Genetic variation in the androgen estrogen conversion pathway in relation to breast cancer prognosticators

Hatef Darabi, Kamila Czene, Sara Wedrén, Yuqing Li, Jianjun Liu, Per Hall, Keith Humphreys

Genetic Variation in the Androgen Estrogen Conversion Pathway in Relation to Breast Cancer Prognosticators

Hatef Darabi[1*], Kamila Czene[1], Sara Wedrén[2], Yuqing Li[3], Jianjun Liu[3] , Per Hall[1], and Keith Humphreys[1]

[1]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, 171 77 Stockholm, Sweden.

[2] Institute for Environmental Medicine, Karolinska Institutet, 171 77 Stockholm, Sweden.

[3]Human Genetics and Cancer Biology, Genome Institute of Singapore, Singapore 138672.

* Correspondence should be addressed to Hatef Darabi, Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, 171 77 Stockholm, Sweden. (hatef.darabi@ki.se).

Abstract

Purpose:

Genetic variation in the androgen-to-estrogen conversion pathway has been shown to be associated with risk of breast cancer, and in particular with Estrogen Receptor (ER) positive tumours. We aimed at studying how the genetic alterations, which have been identified for risk, are associated with breast cancer prognosticators, with a prior hypothesis that, in general, hormone related breast cancers have a better prognosis than non-hormone related breast cancers.

Methods:

Association between tagging SNPs in genes involved in estrogen metabolism and patient's lymph node status, tumour size and histological grade, were estimated in a sample of 1569 Swedish breast cancer patients.

Results:

Polymorphisms in CYP19A1, which have previously been linked to breast cancer risk, are shown to be associated with breast cancer prognosticators. The strongest association was observed for rs4646, with histological grade. The common allele of rs4646, which has been associated with increased breast cancer risk, was associated with low histological grade and small tumour size ($p=0.001$ and 0.015; 1-sided, respectively). We also found evidence that SNP rs7167936 is associated with histological grade and tumour size ($p=0.010$ and 0.005; 1-sided, respectively). We show that rs4646 and rs7167936 are associated with histological grade even amongst only ER-positive tumours ($p=0.008$ and 0.011; 1-sided, respectively).

Conclusions:

Our results provide new evidence that CYP19A1 is involved in both breast cancer risk and prognosis.

Introduction

Exposure to estrogen is an important determinant of the risk of breast cancer. Breast cancer is believed to be induced by excessive estrogen exposure and genetic variation in genes involved in the estrogen metabolism pathway is therefore likely to be important for the disease. Numerous genetic association studies addressing association with cancer risk and genes in the estrogen metabolism pathway have been reported with conflicting results, probably reflecting the low penetrance of mutations in genes within the estrogen metabolism pathway, suboptimal study designs, underpowered studies or a combination of these factors [1-3]. We have recently been able to show, using a pathway approach, that there is an association between breast cancer risk and some of the genes involved in the androgen-to-estrogen conversion pathway (a sub-pathway of the complete estrogen metabolism pathway), particularly for the risk of estrogen receptor(ER)-positive tumours [4]. Gene-based analysis indicated that CYP19A1 was the major contributor to the observed association.

CYP19A1 is located on chromosome 15q21.1 in humans and encodes the aromatase enzyme, which catalyzes the last step of estrogen biosynthesis from androgens. Mutations in CYP19A1 can result in increased or decreased aromatase activity, and it has been suggested that estrogen functions not only as a sex steroid hormone but also plays a part in cell growth or differentiation [5-6].

In the current study we investigated whether the top risk SNPs within the androgen-to-estrogen conversion sub-pathway, according to [4], are associated with breast cancer prognosticators; tumour grade, size and lymph node status. Prior to analysis we hypothesised that alleles associated with an increased risk of breast cancer are associated with good prognosis (and vice-versa) since hormone related breast cancers, identified in terms of menopausal hormone treatment have also been reported to be associated with favourable prognosis [7]. We assessed association overall and within subgroups of patients stratified by ER-status.

Material and Methods

Study Subjects

A Swedish case-control study was initiated in the early 90's to examine the effect of menopausal hormone use on breast cancer risk. The nation-wide case-control study encompassed all Swedish born women between 50 and 74 years of age and resident in Sweden between October 1993 and March 1995 and diagnosed with breast cancer. The cancer patients were identified through the regional cancer registries in Sweden. Women with previous cancers were excluded in order to minimize the risk of including patients with a metastasis in the study. Controls were randomly chosen from the Swedish Registry of Total Population and frequency matched to cases, on age. Characteristics of the participants in the breast cancer case-control study have been described elsewhere [4,7,8]. For cancer patients, information on tumour characteristics was collected from patient records from surgical and oncological units throughout Sweden. This included estrogen and progesterone receptor expression, histopathological subtype, disease stage according to tumour size, and spread to regional lymph nodes. Tumour samples were not assessed for HER2 status or S-phase. The Bloom-Richardson scale or Nottingham histological grade were used as grading systems for classifying Grade [7]. From the original case-control study, consisting of 3345 cases and 3454 controls, 1569 breast cancer cases and 1730 healthy controls were included in a genetic study. Written informed consent was obtained from all the participating subjects, and the study was approved by the Institutional Review Boards in Sweden and the National University of Singapore [4].

Gene and haplotype tagging SNP selection

The SNPs/genes within the androgen-to-estrogen conversion pathway, which are included in the current study, represent a subset of SNPs/genes from a parent study of the entire estrogen metabolism pathway [4]. Genetic variation in the androgen-to-estrogen conversion (sub-) pathway was shown to be associated with breast cancer risk. In the parent study [4] gene selections were based on the criteria that the gene should code for an enzyme that contributes to estradiol and estrone metabolism. In total, 35 genes were selected and a two-stage genotyping design was used. In the first stage 1007 single nucleotide polymorphisms (SNPs) were selected within the 35 genes and their 30Kb flanking sequences, aiming for a marker density of at least one SNP per 5Kb. Linkage disequilibrium (LD) patterns in these genes were determined using

DNA samples from 92 Swedish control women. Haplotypes were reconstructed using the PLEM algorithm implemented in the tagSNP program [9,10]. TagSNP selection was based on the R^2 coefficient. We chose tagSNPs so that common SNP genotypes (minor allele frequency ≥ 0.03) and common haplotypes (frequency ≥ 0.03) were predicted with a good coverage of R^2 values ≥ 0.8. Overall 302 tagSNPs were selected across 35 genes, and these tagSNPs, in stage 2 of the procedure, were genotyped in all available DNA samples from cases and controls. Of these, 252 tagSNPs were successfully genotyped, from which a further thirteen SNPs were excluded on the basis of failing a Hardy-Weinberg Equilibrium test, or having a call rate < 0.85, or a minor allele frequency < 0.01. This left a final set of 239 SNPs distributed among 34 genes, all of which were genotyped. Of these 34 genes, 15 genes (including 120 tagSNPs) are involved in androgen-to-estrogen conversion. We recently showed that there is an association between breast cancer risk and some of the tagSNPs in these 15 genes [4].

Benjamini and Hochberg [11] introduced an important concept for multiple comparisons that they called the false discovery rate, or FDR. The FDR is the expected proportion of false positives among all tests declared significant. In the current study, the FDR q-value [11-13] was estimated for each of the 120 tagSNPs in [4]. The q-value is the FDR equivalent of the p-value; the q-value of an individual test measures the minimum false discovery rate that is incurred when calling that test significant. Out of the 120 tagSNPs, 13 had a false discovery rate q-value less than 0.20 (with q-values ranging from 0.04 to 0.19); see Table 1. These 13 tagSNPs were included in the present study and examined for association with breast cancer prognosticators. For an overview of the design of the current study see Figure 1. We note that 6 of the 13 tagSNPs are in CYP19A1.

Prognosticators for breast cancer

The most commonly used prognosticators for breast cancer are TNM stage, histological grade, estrogen receptor (ER) status and S-phase [14]. In the current study we decided to include the component variables of TNM stage, lymph node status and tumour size, along with histological grade. We did not include ER status since we have already established, in our study material, that the top risk SNPs, in genes involved in

androgen-to-estrogen conversion, in particular increase the risk of ER-positive breast cancer [4]. S-phase could not be included because it was not measured for patients in our study. The lymph node status is either 1 (if there are no nodes affected), 2 (if up to 3 glands are affected) or 3 (if more than 3 glands are affected). Similarly the tumour grade is scored as either 1 (for a grade I, less aggressive appearance), 2 (for a grade II, intermediate appearance) or 3 (for a grade III, more aggressive appearance).

Statistical Analysis

For assessing association between the prognosticators, lymph node status, tumour size and histological grade, and the (13) tagSNPs selected in the present study, we based analysis on all cases with available data (case only analysis). Ordinal (proportional odds) regression was used to model the relationship between ordinal outcome variables (grade and lymph node status) with SNP covariates. Linear regression was used to model the relationship between continuous variables (tumour size) with covariates of interest. Tumour size (in cm) was logarithm transformed in order to obtain an approximately normally distributed dependent variable. Because of our hypothesis that alleles associated with increased risk of breast cancer are associated with good prognosis (and vice-versa), we performed 1-sided hypothesis tests – for each SNP, the alternative hypothesis being that the high risk allele (from [4]) is associated with a favourable prognosticator value (low grade/low lymph node status/small tumour). We, however, also present p-values for 2-sided tests since these are more conservative/standard. We applied Bonferroni corrections to address multiple testing (39 tests of association).

Overall (global) evidence of association between the (13) genetic variants and each of the prognosticators was evaluated using the Admixture Maximum Likelihood (AML) method, which is described in detail in Tyrer et al [15]. This test assesses the experiment-wise significance by examining the empirical distribution of single marker test statistics, in order to determine whether there exists a cumulative effect from multiple variants. For these analyses we dichotomised the prognosticators (grades 1 & 2 vs. 3, lymph node status 1 vs. 2 & 3, tumour size ≤20mm vs. >20mm), since the AML test is for dichotomised outcomes. The AML

analysis was performed using software obtained from the authors of [15]. All other statistical analyses were performed using the free statistical software R [16].

Results

In Table 2 we present summary statistics for the prognosticators, for the entire case-series, as well as for the subset of cases know to be ER positive. Only just over one-half of the cases have tumour ER status recorded. Of these, approximately three-quarters are ER-postive.

In order to summarise the genetic association with tumour phenotypes in cases overall, we used the AML approach [15] with categorised prognosticators as outcome. We carried out three AML tests of association, in order to assess the cumulative association between the 13 tagSNPs and each of the prognosticators, using all cases. The (2-sided) p-values were 0.038, 0.549 and 0.691, for histological grade, tumour size and lymph node status respectively. Indicating there is a cumulative effect from (some of) the 13 markers included in this study with tumour grade.

To investigate our hypothesis, that the most important genetic variants for breast cancer risk, in genes involved in androgen-to-estrogen conversion, are associated with breast cancer prognosticators, we fitted regression models to grade (ordinal), lymph node status (ordinal) and tumour size (linear). Results are displayed in Table 3. The lowest p-values for association were obtained between histological grade and rs7167936 (p=0.010 and p=0.020 for 1-sided and 2-sided tests, respectively) and rs4646 (p=0.001 and p=0.002 for 1-sided and 2-sided tests, respectively). Markers rs7167936 (p=0.005, 1 sided) and rs4646 (p=0.015, 1 sided) also showed association with tumour size (Table 3). None of the selected markers showed association with lymph node status, coded either as an ordinal variable or as a dichotomized variable. All regression coefficients with a corresponding 2-sided p-value < 0.05, for tagSNPs in CYP19A1, were such that the high risk allele is associated with less aggressive tumours (the 1-sided p-values were less than the 2-sided p-values). The 1-sided p-value for the association between rs4646 and

histological grade remains significant (at α=0.05) after a (conservative) Bonferroni correction for multiple testing (p=0.039), although the 2-sided p-value does not.

We found that the association between the "top" SNPs in CYP19A1 and histological grade remained even when including only cases which are ER-positive (data not shown). For SNPs rs7167936 and rs4646 1-sided p-values of 0.011 and 0.008, respectively, were obtained, with regression coefficients 0.27 and 0.31, corresponding to odds ratios of 1.31 and 1.36. Neither marker remained associated for cases that are ER-negative.

Discussion

We have recently shown an association between breast cancer risk and some of the genes involved in the androgen-to-estrogen conversion pathway [4]. Gene based analysis indicated that the aromatase gene CYP19A1 was a major contributor to the observed genetic association. In the current study we hypothesized that genetic variation in the top risk genes, within the androgen-to-estrogen conversion pathway, is also associated with breast cancer prognosis. In our study material, cancers occurring among women treated with menopausal hormones on average have a more favorable prognosis [7]. Following this line of reasoning we hypothesised that high risk alleles of polymorphisms of genes in the androgen-to-estrogen conversion pathway, are, more specifically, associated with breast cancer prognosticators.

The prognosticator with the strongest association was histological grade. Of the 13 studied markers, in the androgen-to-estrogen conversion pathway, the lowest p-values for association with histological grade were obtained for rs7167936 and rs4646 (1-sided p-values of 0.010 and 0.001), both in CYP19A1.

If genetic variation in the androgen-to-estrogen conversion pathway is primarily associated with ER-positive breast cancer, as we have earlier suggested [4], this will partially explain the association between risk polymorphisms and the prognosticators. This claim, made in [4], was made on the basis that MAFs of "risk associated" SNPs were similar in controls and ER-negative breast cancers, but differed in ER-positive

cancers. Only in a few cases were there statistically significant differences between MAFs in ER-negative and ER-positive cases (e.g. for rs7167936, p=0.041). It appears that the level of association we observed between grade and SNPs in CYP19A1 exceeds that which we would expect due solely to MAF differences in ER-positive and ER-negative cancers.

Multiple studies have shown independent prognostic significance of histological grade in breast cancer. To examine the prognostic value of histological grade Rakha et al. [17] investigated the association of grade in patients with operable breast cancer with different tumour size and lymph node subgroups. Multivariate analysis revealed that tumour grade was a significant independent predictor of breast cancer specific survival, with a hazard ratio of 3.9 for grade 3 relative to grade 1 after a median of 9 years, with significant difference in survival for higher histological grade and poorer patient's outcome. Histological grade together with lymph node status and tumour size are the three strongest prognostic determinants for breast cancer. Assessment of degree of biological aggressiveness of a tumour is crucial as biologically more aggressive tumours are likely to increase in size in a short time, in contrast to less aggressive tumours which exhibit more restricted growth rate. Many breast tumours are estrogen sensitive; estrogen helps them to grow. Aromatase is involved in the production of estrogens, and altered expression of it may be associated with prognosis. The single-nucleotide polymorphism rs4646, of CYP19A1, has been reported to be associated with HER2 status of the tumour [18] and circulating steroid hormones [3], making it biologically possible that polymorphisms in CYP19A1 may be associated with response to aromatase inhibitors. To examine this hypothesis Garcia-Casado et al. [19] analysed polymorphisms in CYP19A1 in a series of postmenopausal woman treated with neoadjuvant letrozole and described their association with response to treatment and with progression free survival. The minor genetic variant of rs4646 was shown to be associated with poor response to letrozole, and presented a lower progression-free survival, concluding that analysis of rs4646 could improve the clinical management of breast cancer patients. Our result is in agreement with the result presented by Garcia-Casado et al. [19], since we found the minor variant of this SNP to be associated with high tumour grade.

Strengths of the present study include a large sample size, extensive coverage of SNPs in genes in the androgen-to-estrogen conversion sub-pathway and a well defined hypothesis – by building on a comprehensive study of breast cancer risk we were able to focus our analysis of breast cancer prognosticators on a small number of polymorphisms (see Figure 1). A limitation of the present work is that several women included in the study had incomplete data on tumour characteristics.

In summary, we found that the polymorphism rs4646 in the 3'-untranslated region of CYP19A1 is associated with histological grade and tumour size in postmenopausal women. This supports the hypothesis that analysis of rs4646 could improve the clinical management of postmenopausal breast cancer patients. Further studies based on larger data series are needed to confirm our findings.

Acknowledgements

This research was funded by The Swedish Research council (grant numbers 521-2008-2335 for HD and 523-2006-972 for KH), Cancerfonden (contract no. 090420 for HD and 5128-B07-01PAF for KC) and Märit and Hans Rausing's Initiative against Breast Cancer. We are grateful to all the women who took the time and effort to participate in this study.

References

[1] Justenhoven C, Hamann U, Schubert F, Zapatka M, Pierl CB, Rabstein S, Selinski S, Mueller T, Ickstadt K, Gilbert M, Ko YD, Baisch C, Pesch B, Harth V, Bolt HM, Vollmert C, Illig T, Eils R, Dippon J, Brauch H. (2008) Breast cancer: a candidate gene approach across the estrogen metabolic pathway. *Breast Cancer Res Treat*. 108:137-49

[2] Torresan C, Oliveira MM, Torrezan GT, de Oliveira SF, Abuázar CS, Losi-Guembarovski R, Lima RS, Urban CA, Cavalli IJ, Ribeiro EM. 2008 Genetic polymorphisms in oestrogen metabolic pathway and breast cancer: a positive association with combined CYP/GST genotypes. *Clin Exp Med.* 8:65-71

[3] Haiman CA, Dossus L, Setiawan VW, Stram DO, Dunning AM, Thomas G, Thun MJ, Albanes D, Altshuler D, Ardanaz E, Boeing H, Buring J, Burtt N, Calle EE, Chanock S, Clavel-Chapelon F, Colditz GA, Cox DG, Feigelson HS, Hankinson SE, Hayes RB, Henderson BE, Hirschhorn JN, Hoover R, Hunter DJ, Kaaks R, Kolonel LN, Le Marchand L, Lenner P, Lund E, Panico S, Peeters PH, Pike MC, Riboli E, Tjonneland A, Travis R, Trichopoulos D, Wacholder S, Ziegler RG. (2007) Genetic variation at the CYP19A1 locus predicts circulating estrogen levels but not breast cancer risk in postmenopausal women. *Cancer Res.* 67: 1893-7

[4] Low YL, Li YQ, Humphreys K, Thalamuthu A, Li Y, Darabi H, Wedren S, Bonnard C, Czene K, Iles M, Heikkinen T, Aittomäki K, Blomqvist C, Nevanlinna H, Hall P, Liu ET, Liu J. (2010) Multi-variant Pathway Association Analysis Reveals the Importance of Genetic Determinants of Estrogen Metabolism in Breast and Endometrial Cancer Susceptibility *PLoS Genet. 6:e1001012.*

[5] Simpson ER, Clyne C, Rubin G, Boon WC, Robertson K, Britt K, Speed C, Jones M. (2002) Aromatase--a brief overview. *Annu Rev Physiol*. 64:93-127.

[6] Entrez Gene Database [http://www.ncbi.nlm.nih.gov/sites/entrez?db=gene]

[7] Rosenberg LU, Granath F, Dickman PW, Einarsdóttir K, Wedrén S, Persson I, Hall P. (2008) Menopausal hormone therapy in relation to breast cancer characteristics and prognosis: a cohort study. *Breast Cancer Res.* 10:R78.

[8] Wedrén S. (2004) Genetic Susceptibility to Breast and Endometrial Cancer. *PhD Thesis 2004*, ISBN 91-7140-053-2.

[9] Qin ZS, Niu T, Liu JS. (2002) Partition-ligation-expectation-maximization algorithm for haplotype inference with single-nucleotide polymorphisms. *Am J Hum Genet.* 71:1242-7.

[10] Stram DO, Haiman CA, Hirschhorn JN, Altshuler D, Kolonel LN, Henderson BE, Pike MC. (2003) Choosing haplotype-tagging SNPS based on unphased genotype data using a preliminary sample of unrelated subjects with an example from the Multiethnic Cohort Study. *Hum Hered* 55:27-36.

[11] Benjamini Y, Hochberg Y. (1995) Controlling the false discovery rate a practical and powerful approach to multiple testing. *J Roy Stat Soc Ser B Meth* 57:289300.

[12] Storey JD. (2002) A direct approach to false discovery rates. *J Roy Stat Soc Ser B Meth.* 64:479498-5

[13] Storey JD. (2003) The positive false discovery rate: A Bayesian interpretation and the q-value. *Annals of Statistics* 31: 2013-2035.

[14] Rakha EA, El-Sayed ME, Menon S, Green AR, Lee AH, Ellis IO. (2008) Histologic grading is an independent prognostic factor in invasive lobular carcinoma of the breast.
Breast Cancer Res Treat. 111:121-127

[15] Tyrer J, Pharoah PD, Easton DF. (2006) The admixture maximum likelihood test: a novel experiment-wise test of association between disease and multiple SNPs. *Genet Epidemiol* 30: 636–643

[16] R Project [http://www-r-project-org/]

[17] Rakha EA, El-Sayed ME, Lee AH, Elston CW, Grainge MJ, Hodi Z, Blamey RW, Ellis IO. (2008) Prognostic significance of Nottingham histologic grade in invasive breast carcinoma. *J Clin Oncol.* 26:3153-8

[18] Fasching PA, Loehberg CR, Strissel PL, Lux MP, Bani MR, Schrauder M, Geiler S, Ringleff K, Oeser S, Weihbrecht S, Schulz-Wendtland R, Hartmann A, Beckmann MW, Strick R. (2008) Single nucleotide

polymorphisms of the aromatase gene (CYP19A1), HER2/neu status, and prognosis in breast cancer patients. *Breast Cancer Res Treat*. 112:89-98.

[19] Garcia-Casado Z, Guerrero-Zotano A, Llombart-Cussac A, Calatrava A, Fernandez-Serra A, Ruiz-Simon A, Gavila J, Climent MA, Almenar S, Cervera-Deval J, Campos J, Albaladejo CV, Llombart-Bosch A, Guillem V, Lopez-Guerrero JA. (2010) A polymorphism at the 3'-UTR region of the aromatase gene defines a subgroup of postmenopausal breast cancer patients with poor response to neoadjuvant letrozole. *BMC Cancer*. 10:36.

Figure Legends

Figure1 : Breast Cancer prognosticators and polymorphisms in genes involved in estrogen metabolism; study design.

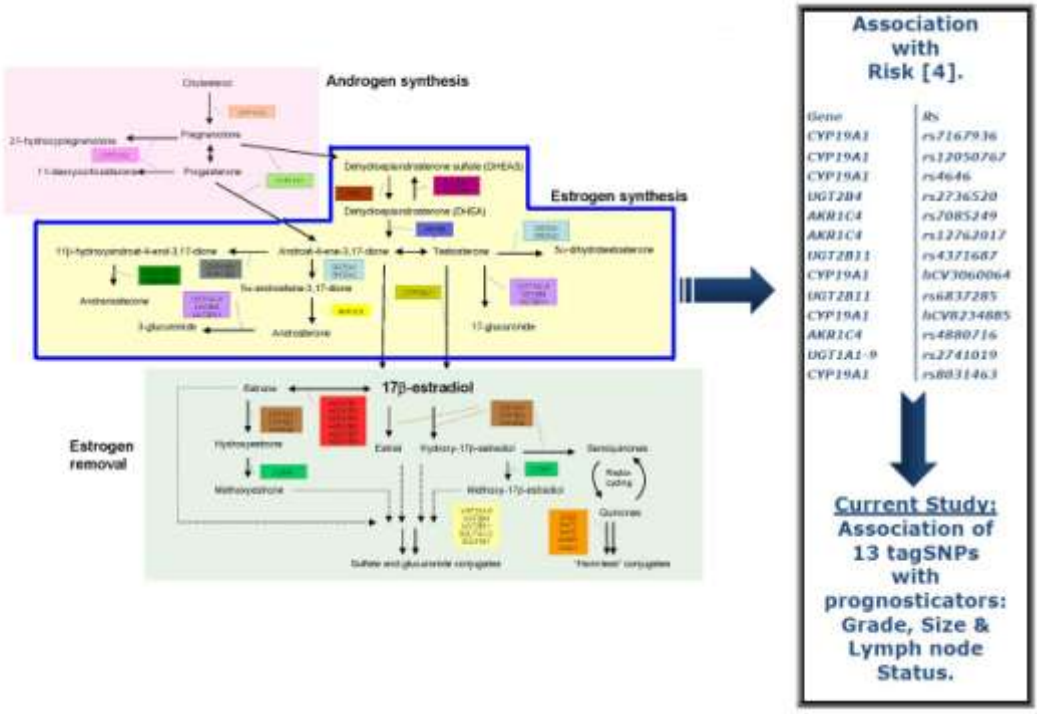

Table 1: Top 13 SNPs in the androgen-to-estrogen conversion pathway for association with breast cancer risk, ordered by p-value.

Rs	Gene	MAF[a]	p-value[b]	OR (95% CI)[b]	q-value[c]	N
rs7167936	CYP19A1	0.49	0.0003	0.83 (0.75,0.92)	0.0405	2997
rs12050767	CYP19A1	0.45	0.0013	1.18 (1.07,1.31)	0.0790	2911
rs4646	CYP19A1	0.30	0.0077	0.86 (0.77,0.96)	0.1915	2994
rs2736520	UGT2B4	0.18	0.0139	0.84 (0.73,0.97)	0.1915	2925
rs7085249	AKR1C4	0.25	0.0144	1.15 (1.03,1.29)	0.1915	2981
rs12762017	AKR1C4	0.15	0.0147	0.83 (0.71,0.96)	0.1915	2791
rs4371687	UGT2B11	0.48	0.0148	1.13 (1.02,1.25)	0.1915	2993
hCV3060064	CYP19A1	0.46	0.0158	0.88 (0.79,0.98)	0.1915	2984

rs6837285	UGT2B11	0.51	0.0176	0.89 (0.80,0.97)	0.1915	2972
hCV8234885	CYP19A1	0.41	0.0180	0.88 (0.79,0.98)	0.1915	2844
rs4880716	AKR1C4	0.25	0.0182	1.15 (1.02,1.28)	0.1915	2982
rs2741019	UGT1A1-9	0.29	0.0201	0.87 (0.78,0.98)	0.1915	3006
rs8031463	CYP19A1	0.05	0.0207	0.75 (0.59,0.96)	0.1915	3016

a: Minor Allele Frequency in Controls.
b: P-value and OR reported in the previous study, [4].
c: Estimated FDR q-values in the current study.

Table 2. Numbers of cases overall and selected subsets of cases selected for available data on histological grade, lymph node status and tumour size.

Data	Total	Grade				Lymph Node Status				Tumour Size		
		1	2	3	NA[a]	1	2	3	NA[a]	≤20mm	>20mm	NA[a]
All	1569	144	445	424	556	1004	313	131	121	1081	389	99
ER+[b]	811	72	291	224	224	525	190	76	20	560	246	5

a: Not Available
b: Out of the 1569 cases, 1031 had ER status recorded, of which 811 were ER-positive.

Table 3 – Single Marker Association with different tumour phenotypes for top 13 risk associated markers.

Rs	Gene	Grade				Lymph node status				Tumour Size		
		β[a]	OR (95%CI)	P-value 1sided	P-value 2sided	β[a]	OR (95%CI)	P-value 1sided	P-value 2sided	β[b]	P-value 1sided	P-value 2sided
rs7167936	CYP19A1	0.20	1.23 (1.03,1.45)	0.010	0.020	-0.04	0.96 (0.82,1.13)	0.698	0.605	0.06	0.005	0.010
rs12050767	CYP19A1	-0.12	0.89 (0.75,1.06)	0.093	0.186	0.05	1.05 (0.89,1.24)	0.726	0.547	-0.04	0.050	0.100
rs4646	CYP19A1	0.31	1.37 (1.12,1.66)	0.001	0.002	0.02	1.02 (0.85,1.22)	0.422	0.844	0.06	0.015	0.030
rs2736520	UGT2B4	-0.03	0.97 (0.77,1.24)	0.587	0.826	-0.01	0.99 (0.79,1.24)	0.542	0.916	0.01	0.332	0.664
rs7085249	AKR1C4	0.12	1.13 (0.94,1.35)	0.899	0.202	-0.05	0.95 (0.80,1.13)	0.268	0.536	-0.01	0.372	0.745
rs12762017	AKR1C4	0.15	1.16 (0.88,1.52)	0.145	0.290	0.03	1.04 (0.80,1.34)	0.398	0.795	0.01	0.345	0.691
rs4371687	UGT2B11	-0.10	0.90 (0.76,1.07)	0.114	0.228	-0.02	0.98 (0.84,1.15)	0.407	0.814	0.00	0.451	0.902

hCV3060064	CYP19A1	0.01	1.01 (0.86,1.20)	0.433	0.866	-0.07	0.93 (0.79,1.09)	0.811	0.377	0.03	0.085	0.171
rs6837285	UGT2B11	0.11	1.12 (0.94,1.33)	0.097	0.194	-0.02	0.98 (0.84,1.15)	0.588	0.824	-0.01	0.680	0.639
hCV8234885	CYP19A1	0.09	1.10 (0.92,1.31)	0.153	0.307	-0.07	0.93 (0.79,1.10)	0.796	0.407	0.02	0.256	0.512
rs4880716	AKR1C4	0.13	1.13 (0.94,1.36)	0.911	0.178	-0.06	0.94 (0.79,1.12)	0.257	0.515	-0.02	0.239	0.479
rs2741019	UGT1A1-9	-0.04	0.96 (0.80,1.16)	0.659	0.682	-0.21	0.81 (0.67,0.97)	0.987	0.026	-0.05	0.978	0.043
rs8031463	CYP19A1	0.27	1.31 (0.86,1.97)	0.103	0.206	0.14	1.15 (0.77,1.70)	0.249	0.499	-0.07	0.878	0.243
Global p-value[c]			0.038				0.549				0.691	

a: Ordinal Regression / Proportional Odds Model : β in $\text{logit}(P(Y \leq j) = \alpha_j - \beta x$.

b: Linear Regression. Tumour size log-transformed: β in $Y = \alpha + \beta x$.

c: Global (2 sided) AML test, based on 5000 permutations.