Interaction Between SNP Genotype and Efficacy of Anastrozole and Exemestane in Early-Stage Breast Cancer

Junmei Cairns[1], Krishna R. Kalari[2], James N. Ingle[3], Lois E. Shepherd[4], Matthew J. Ellis[5], Paul E. Goss[6], Poulami Barman[2], Erin E. Carlson[2], Barbara Goodnature[7], Matthew P. Goetz[3], Richard M. Weinshilboum[1], Huanyao Gao[1] and Liewei Wang[1,*]

Aromatase inhibitors (AIs) are the treatment of choice for hormone receptor–positive early breast cancer in postmenopausal women. None of the third-generation AIs are superior to the others in terms of efficacy. We attempted to identify genetic factors that could differentiate between the effectiveness of adjuvant anastrozole and exemestane by examining single-nucleotide polymorphism (SNP)-treatment interaction in 4,465 patients. A group of SNPs were found to be differentially associated between anastrozole and exemestane regarding outcomes. However, they showed no association with outcome in the combined analysis. We followed up common SNPs near LY75 and GPR160 that could differentiate anastrozole from exemestane efficacy. LY75 and GPR160 participate in epithelial-to-mesenchymal transition and growth pathways, in both cases with SNP-dependent variation in regulation. Collectively, these studies identified SNPs that differentiate the efficacy of anastrozole and exemestane and they suggest additional genetic biomarkers for possible use in selecting an AI for a given patient.

Study Highlights

WHAT IS THE CURRENT KNOWLEDGE ON THE TOPIC?

☑ None of the third-generation aromatase inhibitors (AIs) are superior to the others in terms of efficacy. As a result, there is currently no basis upon which to choose one AI over the others as adjuvant therapy for early-stage breast cancer.

WHAT QUESTION DID THIS STUDY ADDRESS?

☑ To identify genetic factors that could differentiate between the effectiveness of adjuvant anastrozole and exemestane.

WHAT DOES THIS STUDY ADD TO OUR KNOWLEDGE?

☑ We have obtained evidence that anastrozole and exemestane have differing efficacy depending on genetic background.

HOW MIGHT THIS CHANGE CLINICAL PHARMACOLOGY OR TRANSLATIONAL SCIENCE?

☑ We identified a single-nucleotide polymorphism biomarker panel for future individualized therapy of selected patients who might benefit from anastrozole vs. exemestane.

Breast cancer is the most frequent cancer diagnosed among women, impacting more than two million women each year worldwide and causing over 600,000 deaths in 2018.[1] The majority of those cases were in postmenopausal women with early-stage estrogen receptor α (ERα)-positive tumors for which adjuvant endocrine treatment with aromatase inhibitors (AIs), inhibitors of estrogen synthesis,[2] is the first-line treatment. Aromatase is the enzyme responsible for the conversion of testosterone to estrone (E1) and of androstenedione to estradiol (E2). Inhibitors of aromatase were developed to block the synthesis of estrogen in peripheral tissues and, thus, as antiestrogen therapy for the treatment of ERα positive breast cancer in postmenopausal women.[3] The third-generation AIs that are widely used in the clinic include anastrozole and letrozole, two nonsteroidal compounds, and exemestane, a steroidal aromatase inhibitor. All three of the third-generation AIs are superior to tamoxifen in the adjuvant setting for the treatment of hormone-dependent early breast cancer in postmenopausal women.[4–7] Pharmacodynamic studies have shown that letrozole more effectively suppresses plasma estradiol levels than does anastrozole in patients with advanced breast cancer,[8] but the Femara Versus Anastrozole Clinical Evaluation (FACE) trial in 4,136 patients comparing the

[1]Division of Clinical Pharmacology, Department of Molecular Pharmacology and Experimental Therapeutics, Mayo Clinic, Rochester, Minnesota, USA; [2]Division of Biomedical Statistics and Informatics, Department of Health Sciences Research, Mayo Clinic, Rochester, Minnesota, USA; [3]Division of Medical Oncology, Department of Oncology, Mayo Clinic, Rochester, Minnesota, USA; [4]Canadian Cancer Trials Group, Kingston, Ontario, Canada; [5]Department of Medicine, Baylor University College of Medicine, Houston, Texas, USA; [6]Massachusetts General Hospital Cancer Center, Harvard University, Boston, Massachusetts, USA; [7]Patient Advocate, Mayo Clinic Breast Cancer Specialized Program of Research Excellence, Rochester, Minnesota, USA. *Correspondence: Liewei Wang (wang.liewei@mayo.edu)

Received March 24, 2021; accepted May 8, 2021. doi:10.1002/cpt.2311

efficacy of adjuvant anastrozole vs. letrozole head-to-head showed no statistically significant difference in disease-free survival or overall survival between the treatment arms.[9] Another head-to-head comparison of anastrozole and exemestane in the MA.27 trial in 7,576 patients also indicated that treatment with exemestane for 5 years was not better than anastrozole for 5 years.[10] The more recent FATA-GIM3 trial directly compared anastrozole, exemestane, and letrozole for the first time as adjuvant treatment for hormone receptor-positive early breast cancer, and found no significant difference in disease-free survival 5 years after the initiation of treatments among the 3 groups.[11] As a result, there is currently no basis upon which to choose one AI over the others as adjuvant therapy for early-stage breast cancer.

The "dogma" that the three AIs are essentially the same has been revisited based on our recent data showing that anastrozole, but not exemestane or letrozole, displays an association between the degree of estrogen suppression and outcomes in early breast cancer, and that anastrozole has a mechanism of action in addition to aromatase inhibition in that it is a ligand for ERα.[12,13] These findings have potential implications for the individualization of AIs therapy.

Genomewide association studies (GWAS) have identified several genomic regions that affect risks for breast cancer recurrence in women treated with AIs.[13–15] However, prior studies did not systematically test whether there might be an interaction between individual AI treatment and genetic variability. In the present study, we have investigated germline genetic variation as a predictive marker for differential efficacy of the nonsteroidal AI anastrozole and the steroidal exemestane and as a prognostic marker in the adjuvant Canadian Cancer Trials Group (CCTG) MA.27 trial[10] in patients with hormone receptor-positive early breast cancer. We found that a group of single-nucleotide polymorphisms (SNPs) were differentially associated between anastrozole and exemestane regarding outcomes.

METHODS

Source of patients, study design, and study population

The design of the open-label, randomized phase III trial CCTG MA.27 (ClinicalTrials.gov identifier: NCT00066573), including eligibility criteria, randomization, treatment plan, follow-up, and monitoring and primary statistical analyses have been described elsewhere.[10] In brief, 7,576 postmenopausal patients with hormone receptor-positive breast cancer were randomized to anastrozole (1 mg/day) or exemestane (25 mg/day) as adjuvant therapy with planned treatment duration of 5 years. Only North American patients (6,827 of 7,576 of MA.27 accrual) were offered participation in the collection of blood specimens, and 5,221 of the North American patients contributed blood and gave consent for genetic testing. This research was approved by health regulatory authorities and the institutional review boards of participating centers.

Genotyping was performed in three cohorts as detailed in Supplementary Material.[14,16,17] For the current analysis, we restricted the studies to the White only population that included 4,465 patients with 245 breast events: 2,226 patients on the anastrozole arm with 121 breast events, and 2,239 patients on the exemestane arm with 124 breast events. African American and Han Chinese cohorts had nine breast events in the combined exemestane and anastrozole cohorts. In addition, we did not have breast events in Han Chinese for the exemestane cohort; therefore, the covariate models with all races were unstable and did not converge. Hence, we chose to analyze only the White cohort.

Primary end point

The primary outcome was the STEEP[18] end point of Breast Cancer Free Interval (BCFI), defined as the time from randomization to the first local-regional breast cancer recurrence (including ipsilateral ductal carcinoma in situ (DCIS)), distant breast cancer recurrence, contralateral breast cancer (invasive or DCIS), or death with or from breast cancer without a prior recurrence date. Follow-up was censored at non-breast cancer death or the longest follow-up without recurrence.

Statistical design

The pharmacogenomic analysis focused on identifying SNPs effects whose associations with BCFI were significantly different based on the treatment (anastrozole or exemestane). Thus, our analysis used a stratified genomewide Cox-proportional hazards model using an SNP/treatment interaction term and was controlled for significant stratification factors (chemotherapy, nodal status, and trastuzumab), and additional clinical covariates, including treatment arm, cohort, estrogen receptor (ER)/progesterone receptor status, T stage, Eastern Cooperative Oncology Group (ECOG) performance score, and bisphosphonate use. To avoid bias that might arise from differences in genetic ancestry (i.e., population stratification), the EigenStrat software was used to determine eigenvalues for the SNP correlation matrix that statistically differed from zero based on Tracy–Widom *P* values.[19,20] All the analyses were run using the R statistical computing package, PLINK,[21] and SAS (SAS Institute, Cary, NC, USA).

Reporter gene assay

Luciferase reporter gene constructs containing rs1877193 wild type (WT) and variant genotypes were generated by gene synthesis. Specifically, a 1,001 bp segment of the *LY75* promoter containing estrogen response element (ERE) and rs1877193 SNP WT or variant genotype was synthesized and cloned into the Kpnl and Xhol sites upstream of the luc2 gene in the pGL4-Basic vector (Promega, Madison, WI, USA). Reporter gene constructs containing WT or variant SNP genotypes were transfected into cells using Lipofectamine 2000 (Invitrogen, Waltham, MA, USA). A pRL-CMV vector (Promega) that expresses renilla luciferase was co-transfected as an internal control. After AI treatment, the cells were lysed, and luciferase activity was determined using the Dual-Luciferase Reporter assay system (Promega). Renilla luciferase activity was used to correct for transfection efficiency.

Statistical analysis

For cell survival, proliferation, gene expression, CHIP, reporter gene assays, and quantifications, data are represented as the mean ± SEM of three independent experiments. Unless otherwise described, two-way analysis of variance (ANOVA) was performed to test group difference. Then post hoc analysis was carried out to check if specific groups are significantly different or similar. Tukey Honest Significant Differences (HSD; R function: TukeyHSD), which is essentially a modified *t*-test corrected for multiple comparisons, was applied in this analysis. Statistical significance level is 0.05.

RESULTS

Genetic variation interaction with aromatase inhibitor therapy

The Participant Flow Diagram for the MA.27 trial (**Figure S1**) shows the patients included in the SNP-treatment interaction studies. The analysis cohort was restricted to 4,465 White patients (2,226 patients treated with anastrozole and 2,239 patients treated with exemestane, with 121 and 124 breast events in each arm). To determine the possible SNP-treatment interaction effect using the MA.27 cohort, we performed a stratified Cox-proportional hazard analysis utilizing stratification factors and other covariates as

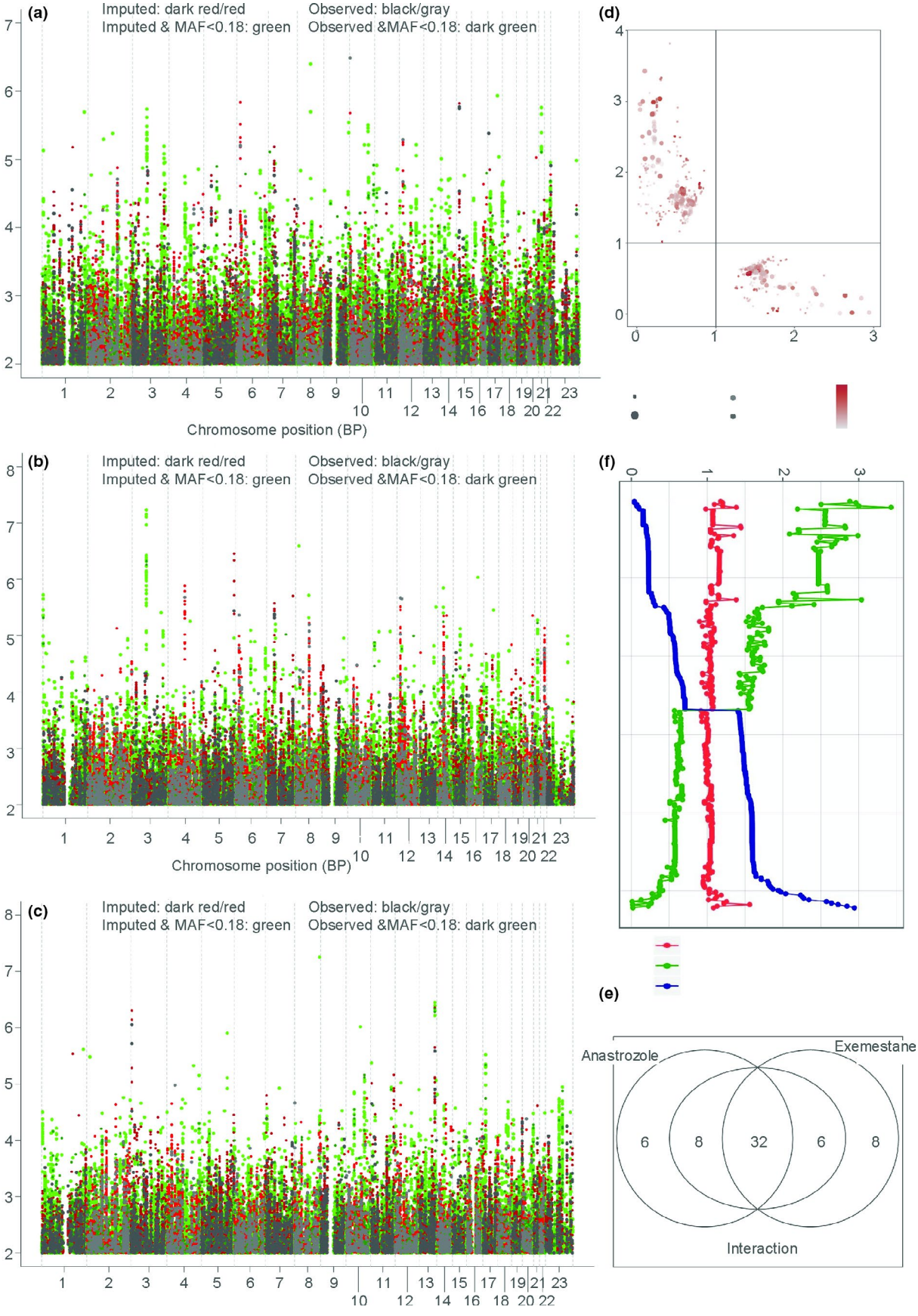
(a)
Imputed: dark red/red
Imputed & MAF<0.18: green
Observed: black/gray
Observed &MAF<0.18: dark green
Chromosome position (BP)
(b)
Imputed: dark red/red
Imputed & MAF<0.18: green
Observed: black/gray
Observed &MAF<0.18: dark green
Chromosome position (BP)
(c)
Imputed: dark red/red
Imputed & MAF<0.18: green
Observed: black/gray
Observed &MAF<0.18: dark green
(d)
(e)
Anastrozole
Exemestane
6
8
32
6
8
Interaction
(f)

Figure 1 SNP treatment interaction analysis. (**a**) SNP-treatment interaction GWAS Manhattan Plot. (**b**) Anastrozole arm GWAS Manhattan Plot. (**c**) Exemestane arm GWAS Manhattan Plot. (**d**) SNP-treatment interaction GWAS signals show opposite effects on BCFI in the anastrozole arm GWAS vs. the exemestane arm GWAS. Hazard ration from two GWASs. X axis: Hazard ration from the exemestane arm GWAS (BCFI_EXE_HR). Y axis: Hazard ration from the anastrozole arm GWAS (BCFI_ANA_HR). Significance indicated by size, and averaged log (p value) by color. (**e**) 261 SNP-treatment interaction GWAS signals showed opposite effects on BCFI in the anastrozole arm GWAS (green) vs. the exemestane arm GWAS (blue), but not associated with BCFI in the combined cohort GWAS (red). Three sets of ordered HRs of 261 SNPs from three GWASs; the whole cohort (red), patients who received only anastrozole (green) and patients who received only exemestane (blue). (**f**) Venn diagram of the SNPs that were selected after performing clumping using three different criteria (see Methods). BCFI, Breast Cancer Free Interval; GWAS, genomewide association studies; HR, hazard ratio; MAF, Minor Allele Frequency; SNP, single-nucleotide polymorphism.

detailed in the Methods section. The Manhattan plot (**Figure 1a**) shows the results of SNP-treatment interaction analyses for BCFI, and the quantile-quantile plot (**Figure S2A**) revealed a lambda of 1.025. Although none of these SNPs met the criteria for genome-wide significance, we had used the largest sample set available to perform the GWAS, so we chose to pursue the results functionally and mechanistically. The results of those studies supported the novel insights that are described subsequently regarding genetic mechanisms for variation in AI effect.

Table S1 presents the 887 significant SNP-by-drug interactions (P < 1.0E-04) on BCFI when comparing exemestane to anastrozole. Signals with an Exemestane/Anastrozole hazard ratio (HR) of > 1 indicate that patients carrying a variant SNP genotype have reduced efficacy when treated with exemestane compared with anastrozole. On the contrary, signals with an Exemestane/Anastrozole HR of < 1, indicate that patients carrying a variant SNP genotype have a better outcome when treated with exemestane compared with anastrozole. We next investigated these 887 SNPs for their associations with anastrozole (**Figure 1b**, **Figure S2B**) and exemestane (**Figure 1c**, **Figure S2C**) efficacy (BCFI) separately based on GWAS conducted within each treatment arm in the MA.27 study. One hundred thirty-four of 887 SNPs (**Figure 1d**, **Table S2**) were associated with better outcomes (longer BCFI) in exemestane-treated patients (HR = 0.04~0.71, P = 2.21E-04~0.009), but with shorter BCFI in anastrozole-treated patients (HR = 1.42~3.43, P = 6.36E-05~0.0097). One hundred twenty-seven of 887 SNPs (**Figure 1d**, **Table S2**) were associated with shorter BCFI in exemestane-treated patients (HR=1.41~2.94, P = 2.21E-04~0.0099), but with longer BCFI in anastrozole-treated patients (HR = 0.01~0.68, P = 1.99E-05~0.0095). However, these two groups of SNPs (134 + 127 = 261 SNPs; **Figure 1e**, **Figure S3A**) showed no association with BCFI in the combined analysis of anastrozole and exemestane (HR = 0.90~1.56, P = 0.08~0.99) based on our previous published GWAS results for the MA.27 trial.[14] **Figure 1e** plots 3 sets of ordered HRs of 261 SNPs from 3 GWASs; the whole cohort (red), patients who received anastrozole (green), and patients who received exemestane (blue). In the left half of the plot, SNPs whose HRs range from 0 to 1 in exemestane cohort have an HR ranging from 1 to 3 in the anastrozole cohort implying the presence of minor alleles in these SNPs will be protective if treated with exemestane as compared to anastrozole. Similarly, in the right half of the plot, we see a set of SNPs whose HRs suggest being protective alleles for anastrozole and risk alleles for exemestane. HRs obtained from the whole cohort remained constant around one, suggesting low power to detect any SNP effect. In summary, the current analyses revealed a difference in anastrozole and exemestane efficacy in the context of the host genome and identified germline genetic variants that might be able to identify which AI might be superior in preventing breast cancer recurrence in that specific patient. Importantly, over 98.7% of the signals (876 of 887 SNPs) were lost (P > 0.01) in the GWAS analysis when anastrozole and exemestane were combined. Moreover, 261 SNPs were observed to be risk alleles for one AI and protective for the other (**Figure 1e**).

Genetic model for guiding aromatase inhibitor selection

We next set out to build a genetic model for guiding AI selection. Of the 261 SNPs, 105 SNPs marginally affected BCFI based on any of the clinical covariates and were removed from further analysis. After filtering the 105 SNPs, the remaining 156 SNPs underwent linkage disequilibrium (LD) –clumping analyses.[22] In brief, "clumping" reported the SNPs that were within 500 kb of the significant SNP, were in high LD (r^2 > 0.8), and had a low P value. For anastrozole and exemestane cohorts, we chose a P value cutoff < 0. 001, and for the interaction GWAS analysis, we chose a P value cutoff < 9.99E-05. From the clumping analysis for each study, we obtained a set of SNPs that were common and unique to each study, as shown in **Figure 1f** and **Table 1**. Coincidentally, the number of SNPs remaining after clumping was 46 for all 3 studies. As shown in the Venn diagram, we overlapped the SNP IDs from the three studies after clumping to identify a reduced list of SNPs that were significantly associated with anastrozole and with exemestane. Thirty-two SNPs were retained after clumping in all the 3 studies, 14 of those SNPs were present in 2 or more studies (interaction and anastrozole or interaction and exemestane LD analysis), 8 SNPs were unique to exemestane only, and 6 SNPs were retained in the anastrozole only clumping analysis, as shown in **Table S3**. We ultimately obtained a list of 60 unique SNPs (32 + 14 + 8 + 6 = 60) as our SNP panel. Note that these SNPs had been obtained from a pure discovery activity and had not been tested or validated in an independent set. **Figure S3b,c** show the HR for SNPs selected based on each of the drug-specific association results.

Table 1 SNPs that remained significant after clumping in each of the three GWAS

GWAS used	P value cutoff	No. of SNPs from clumping
BCFI Interaction	9.99E-05	46
BCFI with anastrozole cohort	0.001	46
BCFI with exemestane cohort	0.001	46

BCFI, Breast Cancer Free Interval; GWAS, genomewide association studies; SNP, single-nucleotide polymorphism.

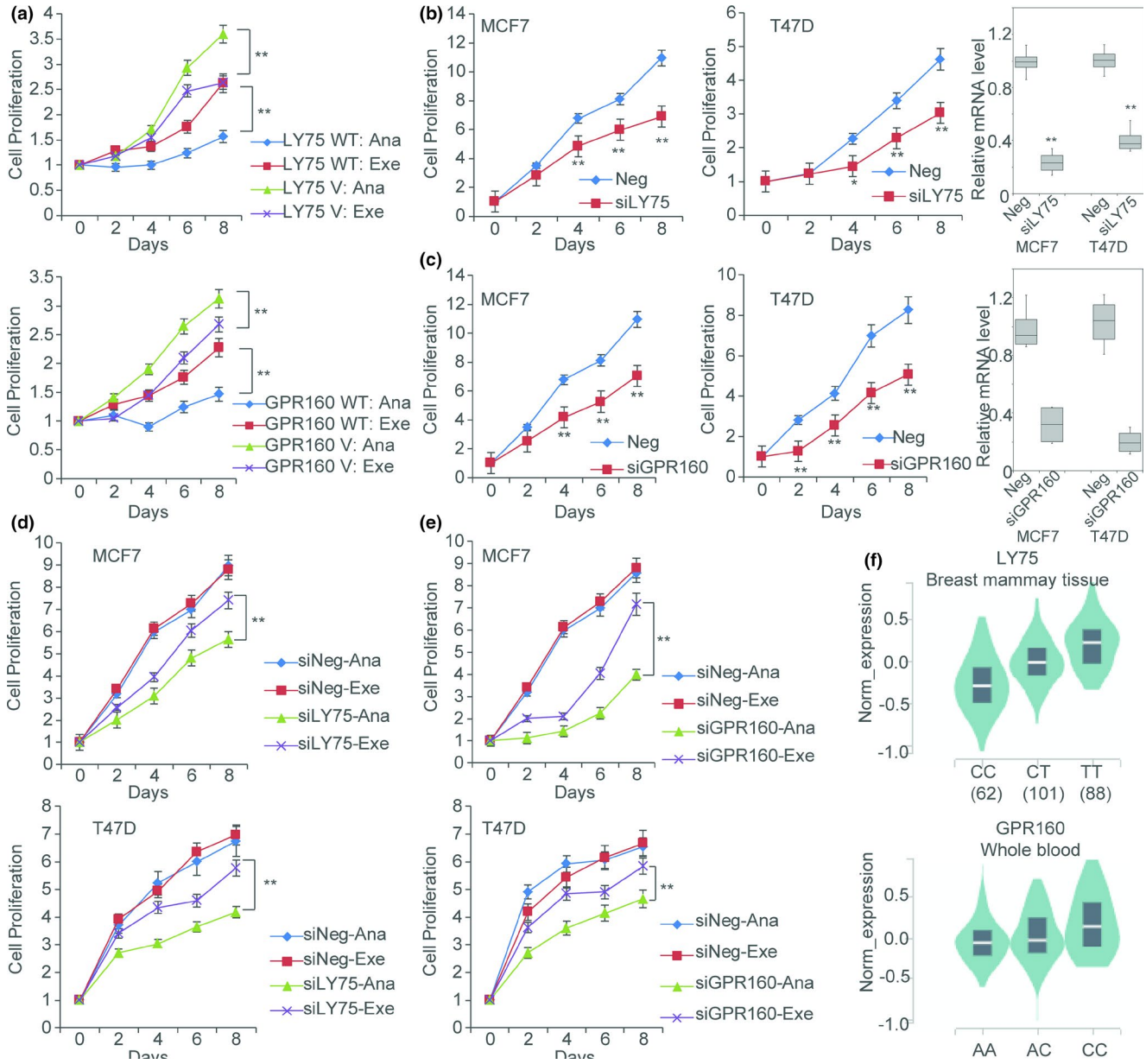

Figure 2 *LY75* and *GPR160* SNPs dependent effect on cell proliferation. (**a**) Cell growth of LCLs with known genotypes for variant (V) and wild type (WT) *LY75* SNPs and *GPR160* SNP were incubated with 100 nM anastrozole (Ana) or exemestane (Exe) for the indicated times. WT: homozygous WT, V: heterozygous and v/v: homozygous variant LCLs (n = 3, means ± SEM). (**b**) Cell growth of MCF-7 and T47D cells after knocking down *LY75*. (**c**) Cell growth of MCF-7 and T47D cells after knocking down *GPR160* (n = 3, means ± SEM). $^{**}P < 0.01$. Statistical test: two-way ANOVA. (**d**) MCF-7 and T47D cells were transfected with *LY75* siRNA, 24 hours later cells were treated with Ana or Exe. Cell growth of MCF-7 and T47D cells were then measured (n = 3, means ± SEM). $^{**}P < 0.01$. (**e**) MCF-7 and T47D cells were transfected with *GPR160* siRNA, 24 hours later cells were treated with Ana or Exe. Cell growth of MCF-7 and T47D cells were then measured (n = 3, means ± SEM). $^{**}P < 0.01$. Statistical test: two-way ANOVA. (**f**) The rs1877193 SNP and the rs62293499 SNP cis-eQTLs for *LY75* and *GPR160*, respectively in GTEX. ANOVA, analysis of variance; eQTL, expression quantitative trait loci; LCL, lymphoblastoid cell line; SNP, single-nucleotide polymorphism.

Functional genomics of the top treatment interacting GWAS SNPs

The next series of experiments were performed to determine the function of the 261 SNPs from our SNP-treatment interaction analysis. Expression quantitative trait loci (eQTL) are particularly useful for annotating variants associated with complex traits.[23] We, therefore, determined whether any of the 261 SNPs were eQTL SNPs using the Genotype-Tissue Expression (GTEx) website.[24] Cis-eQTL analysis identified 46 SNPs that were eQTLs with 2 genes in breast tissue, and an additional 4 genes in other tissues (**Table S4**). Collectively, the eQTL analysis identified six "eGenes" (genes exhibiting association with genotypes for at least one SNP). The SNPs associated with any given gene were in high LD with each other, with pairwise correlation $r^2 = 1$.

We then investigated which SNP-eGenes were differentially associated with efficacy for anastrozole and exemestane. We took

advantage of cell lines selected on the basis of each of the 6 eQTL SNP genotypes from a genomic data-rich panel of 300 human lymphoblastoid cell lines (LCLs) that has proven to be a powerful tool for testing pharmacogenomic hypotheses.[15,25–29] We selected four LCLs with homozygous WT genotypes for each SNP and four LCLs with the variant genotypes with detectable aromatase expression (**Figure S4a**) to test the impact of anastrozole and exemestane on cell growth. We ensured that the aromatase activity was inhibited by anastrozole and exemestane (**Figure S4b**). The eQTL SNPs for the following genes, *LY75*, *GPR160*, *PLA2R1*, *NPLOC4*, and *NT5C3A*, showed that the variant genotypes all had a differential effect on cell growth between exemestane and anastrozole (**Figure 2a**, **Figure S4c**). Among these five eQTL relationships, three were not consistent with our analysis results. The directional effect of the eQTL SNPs for *NT5C3A* on individual drug efficacy identified through SNP-drug interaction GWAS was not consistent with experimental results (**Figure S4**, **Table S1**). As for the *NPLOC4* SNPs, anastrozole, in comparison to exemestane, had greater growth inhibition on cells with variant SNP genotypes, which was consistent with the SNP-drug interaction GWAS results (Minor Allele Frequency (MAF) = 0.037, HR: Anas vs. Exem = 0). However, the anastrozole to exemestane HR was 0 due to the low MAF, resulting in low statistical power,[30,31] raising the possibility that the association might be a false positive. Indeed, regardless of the *NPLOC4* SNP genotype, cells grew faster in the presence of exemestane compared with anastrozole (**Figure S4**). The eQTL SNPs for *LY75* (lymphocyte antigen 75), *GPR160* (G Protein-Coupled Receptor 160), and *PLA2R1* (Phospholipase A2 Receptor 1) showed a consistent effect on drug response between GWAS result and the experimental validation (**Figure S4**, **Figure 2a**, **Table S1**), suggesting that these SNPs may differentially affect anastrozole and exemestane response through regulation of the *LY75*, *GPR160*, and *PLA2R1* genes. We sought to study the functional impact of *LY75*, *GPR160*, and *PLA2R1* genes, which are expressed in breast tissue (**Figure S5**). Knockdown of *LY75* and *GPR160*, but not *PLA2R1*, decreased breast cancer cell proliferation (**Figure 2b,c**, **Figure S6a**). We then validated the interaction effects of these three genes with AIs. *LY75* and *GPR160*, but not *PLA2R1*, knockdown cells grew more slowly in the presence of anastrozole compared to exemestane (**Figure 2d,e**, **Figure S6b**). Therefore, we chose to focus on the functional implications of the *LY75* and *GPR160* SNPs because those SNPs may represent regulatory variants.

Associations with breast cancer free interval in treatment-stratified analysis

The two common SNPs (MAF = 0.455), rs1877193 and rs6735923, were eQTLs for *LY75* gene expression in breast tissue (**Figure 2f**), and they were also in perfect LD (r^2 = 1). These 2 SNPs mapped upstream of *LY75* on chromosome 2 (**Figure 3a,b**) and had significant interaction with treatments (SNP rs1877193: HR: Exe/Ana = 0.45; SNP rs6735923: HR: Exe/Ana = 0.46; **Table S5**), indicating ~ 55% reduction in the hazard for a breast cancer event for patients carrying the variant genotype treated with exemestane compared with anastrozole. In treatment-specific cohort GWAS analysis, the two SNPs were associated with opposite effects on risk for a breast event between the two treatment cohorts: increased risk in the anastrozole cohort (HR = 1.57, 1.58; *P* = 5.3E-04, 6.9E-04), and decreased risk in the exemestane cohort (HR = 0.70, 0.71; *P* = 6.0E-03, 8.0E-03). However, the two SNPs showed no association with breast events if the two treatment arms were combined[14] (HR = 1.07, 1.08; *P* = 0.37, 0.40; **Table S5**).

The *GPR160* SNP rs62293499 mapped downstream of *GPR160* on chromosome 3p7, and the variant allele was associated with increased *GPR160* expression (**Figure 2f**). The minor allele (MAF = 0.338) was associated with decreased risk for breast events when treated with exemestane compared with anastrozole (HR: Exe/Ana = 0.43; **Table S5**). This effect was also confirmed by independent GWAS analysis in each treatment arm, in that a trend for increased risk of a breast event within the anastrozole arm (HR = 1.39, *P* = 0.014), and for significantly decreased risk within the exemestane arm (HR = 0.58, *P* = 2.0E-04) were observed (**Table S5**). Similar to the *LY75* SNPs, rs62293499 was not associated with breast events when the two treatment arms were combined (HR = 0.92, *P* = 0.39; **Table S5**).

LY75 SNP effects differ in the presence of anastrozole vs. exemestane

We next set out to determine whether and how the *LY75* SNPs might have a differential impact on response to the two AIs through *LY75*. We selected LCLs with WT and variant genotypes for *LY75* SNPs, and then exposed them to an AI in the presence of androstenedione, a precursor of estrogen. The variant cell lines showed a reduction of *LY75* expression in response to androstenedione exposure, whereas WT LCLs did not (**Figure 3c,d**). Addition of anastrozole showed a striking reversal of the *LY75* expression pattern in LCLs with variant SNP genotypes: *LY75* expression increased almost threefold with the addition of anastrozole. In contrast, WT expression did not change (**Figure 3c**). Of importance, exemestane did not change the expression patterns of *LY75* in either WT or variant genotype LCLs (**Figure 3d**).

Anastrozole, but not exemestane, regulates LY75 SNP-dependent ERα/ERE binding

To determine whether the SNP- and AI-dependent *LY75* gene regulation was mediated by ERα, we first identified two EREs; 2 bp overlap with each other) located ~ 200 bp distant from the *LY75* SNP rs1877193 based on ENCODE data (**Figure 3e**). We then performed ChIP assays with ERα antibody and found that the rs1877193 SNP (**Figure 3f**) showed striking differential binding to the ERE that was located closest to the SNP, with greater binding for the variant SNP genotype in the presence of androstenedione, but less binding when additional anastrozole was present. However, this differential binding pattern was not detected after exemestane treatment (**Figure 3f**). *LY75* expression was reduced whereas ERα/ERE binding increased in cells with the variant genotypes in response to androstenedione exposure, suggesting that ERα functions as a transcription repressor of *LY75*. Indeed, *LY75* expression increased when ESR1 was knocked down (**Figure 3g**).

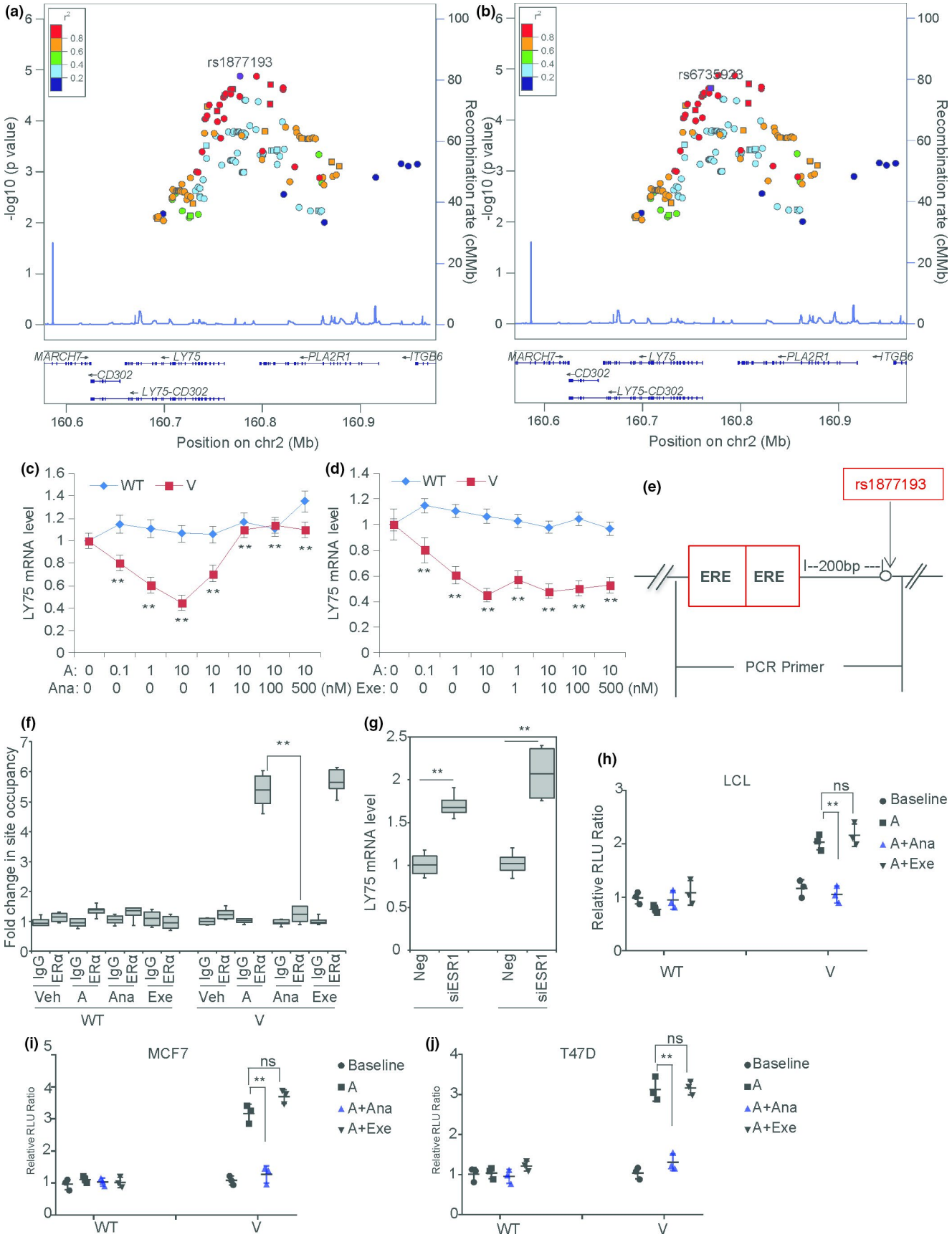
(a)
(b)
rs1877193
rs6735923
-log10 (p value)
Recombination rate (cM/Mb)
MARCH7
LY75
PLA2R1
ITGB6
CD302
LY75-CD302
Position on chr2 (Mb)
(c)
(d)
WT
V
LY75 mRNA level
A: 0 0.1 1 10 10 10 10 10
Ana: 0 0 0 0 1 10 100 500 (nM)
Exe: 0 0 0 0 1 10 100 500 (nM)
(e)
rs1877193
ERE
ERE
I--200bp ---I
PCR Primer
(f)
Fold change in site occupancy
IgG
ERα
Veh
A
Ana
Exe
WT
V
(g)
LY75 mRNA level
Neg
siESR1
(h)
LCL
Relative RLU Ratio
ns
Baseline
A
A+Ana
A+Exe
(i)
MCF7
(j)
T47D

Figure 3 *LY75* SNP- and anastrozole-dependent gene regulation. (**a, b**) Locus zoom plot of the chromosome 2 region surrounding the *LY75* gene. (**c, d**) SNP-androstenedione- and anastrozole (**c**, Ana) or exemestane (**d**, Exe) -dependent regulation of LY75 gene expression in LCLs selected based on *LY75* SNP genotypes. WT: homozygous WT, V: heterozygous variant LCLs. (n = 3, means ± SEM). $^{**}P < 0.01$. (**e**) Schematic figure showing the estrogen response elements (ERE) surrounding the rs1877193 SNP. (**f**) ERα ChIP assay showing SNP dependent ERα binding to the ERE region that is 200 bp upstream from the rs1877193 SNP in LCLs with different genotypes treated with the indicated drugs: androstenedione (A), anastrozole (Ana), and exemestane (Exe) (n = 3, means ± SEM). $^{**}P < 0.01$. Statistical test: two-way ANOVA. (**g**) MCF-7 and T47D cells were transfected with ESR1 siRNA. Forty-eight hours later, mRNA was harvested for qRT-PCR (n = 3, means ± SEM). $^{**}P < 0.01$. (**h–j**) Reporter gene assays for the *LY75* rs1877193 SNP. The rs1877193 SNP was included in the *LY75* promotor sequence that was subcloned upstream of the luciferase reporter gene, Luc2. The plasmids contained either WT or variant sequences for the SNP and were transfected into **h** LCLs, **i** MCF-7, and **j** T47D cells. Values shown for **h–j** are ratios of relative light units (RLUs) compared with the internal reference after vehicle (baseline), A, Ana, or Exe treatment (n = 3, means ± SEM). $^{**}P < 0.01$, ns, not significant. Statistical test: two-way ANOVA. ANOVA, analysis of variance; LCL, lymphoblastoid cell line; PCR, polymerase chain reaction; qRT-PCR, quantitative real-time polymerase chain reaction; SNP, single-nucleotide polymorphism.

These results were compatible with our SNP-treatment interaction analysis results in that variant *LY75* SNP genotypes could differentiate anastrozole from exemestane in the risk for a breast event during 5 years of AI therapy for those enrolled in MA.27.

Finally, to determine whether the differences in binding shown in **Figure 3f** might have functional consequences; we created reporter gene constructs that included ~ 1.1 kb of the *LY75* promoter with WT and variant genotypes for the rs1877193 SNP. These reporter gene constructs were then transfected into different cell lines to compare *LY75* transcriptional activities for WT and variant SNP genotypes in those cells. within the presence of androstenedione, cells transfected with the variant construct showed higher *LY75* transcriptional activity than did cells transfected with the WT construct. However, when the cells were incubated with the addition of 100 nM anastrozole, the transcriptional activity in cells transfected with the variant SNP construct decreased (**Figure 3h–j**). We did not observe any change when we repeated the same experiments with exemestane (**Figure 3h–j**).

LY75 suppression directs breast cancer epithelial-to-mesenchymal transition

To identify molecular mechanisms relevant to the *LY75*—treatment interaction, we interrogated The Cancer Genome Atlas (TCGA)[32] for genes co-expressed with *LY75* in ER+ breast cancer sample set[33] using the cBioPortal online platform. After filtering the co-expression gene prediction results, 2,352 predicted genes (**Figure 4a**) with Spearman's correlations ≥ 0.3 or ≤ −0.3 were selected (**Table S6**). Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis of these 2,352 genes identified PI3K-AKT, MAPK, cell adhesion, and JAK-STAT pathways (**Figure 4b**). We evaluated the top pathways, and our results revealed that JAK/STAT signaling was dysregulated by *LY75* depletion (**Figure 4c**). Activation of the JAK/STAT3 signaling via IL-6/IL-6R leads to increased metastasis resulting from enhanced epithelial-to-mesenchymal (EMT).[34–37] Remarkably, LY75 suppression induced EMT as shown by the expression of specific EMT markers: a significant decrease of E-cadherin and EPCAM; an increase of N-cadherin and SNAIL1 (**Figure 4c**). *LY75* knockdown cells also displayed decreased proliferation rates compared with the control (**Figure 2b**), which was associated with decreased AKT and MAPK phosphorylation (**Figure 4c**). Moreover, *LY75* suppression significantly promoted cell migration (**Figure 4d**). Thus, the *in vitro* invasiveness and motility of *LY75*-knockdown cells was inversely correlated with their proliferative potential which might be due to the acquiring of the mesenchymal phenotype.[38] To provide potential clues to the observed genetic correlations, we examined the SNP-treatment interaction effect on the JAK/STAT and MAPK activities in LCLs with WT and variant genotypes for the *LY75* SNPs. The variant cells showed increased expression of pSTAT3 and decreased pERK in response to androstenedione treatment. In contrast, in WT LCLs, pSTAT3, and pERK, expressions were virtually unchanged with androstenedione. Importantly, the addition of anastrozole, but not exemestane, significantly reduced the pSTAT3 and pERK levels in variant cells (**Figure 4e**). The next question that we addressed was the possible mechanism by which the *GPR160* SNP might result in differential effects in response to anastrozole and exemestane treatment.

Anastrozole differs from exemestane regarding GPR160 SNP-dependent gene regulation

The *GPR160* SNP rs62293499 (**Figure 5a**) variant allele was associated with longer BCFI after treatment with exemestane as compared to anastrozole (**Table S5**). We began our functional genomic studies of the *GPR160* SNP starting with the question of whether *GPR160* expression might be regulated in an SNP-treatment-dependent fashion. *GPR160* mRNA expression was induced by exposure to increasing concentrations of androstenedione in cells with WT genotype, but not in those with the variant SNP genotype (**Figure 5b**). We observed a striking reversal of expression with anastrozole, but not exemestane, that is, increased expression in the presence of anastrozole in cells with variant *GPR160* SNP genotype, but decreased expression in cells with WT genotype (**Figure 5b**). Several EREs were located within 500 bp of the *GPR160* SNP (**Figure 5c**), and the rs62293499 SNP genotype was associated with striking differential binding to the EREs, with more binding for the WT SNP genotype in the presence of androstenedione, but a greater binding for the variant SNP when anastrozole was present. Exemestane, however, did not change the binding activity (**Figure 5d**).

We identified 223 genes that were correlated with GPR160 in TCGA ER+ breast cancer data set (**Table S6**, **Figure S7a**), enriched in metabolic, PI3K-AKT, and MAPK signaling pathways. GPR160 knockdown resulted in the downregulation of cell proliferation pathway, such as PI3K-AKT (**Figure S7b**).

DISCUSSION

In the present investigation, we evaluated the interaction between genetic variation and 2 AIs, exemestane and anastrozole,

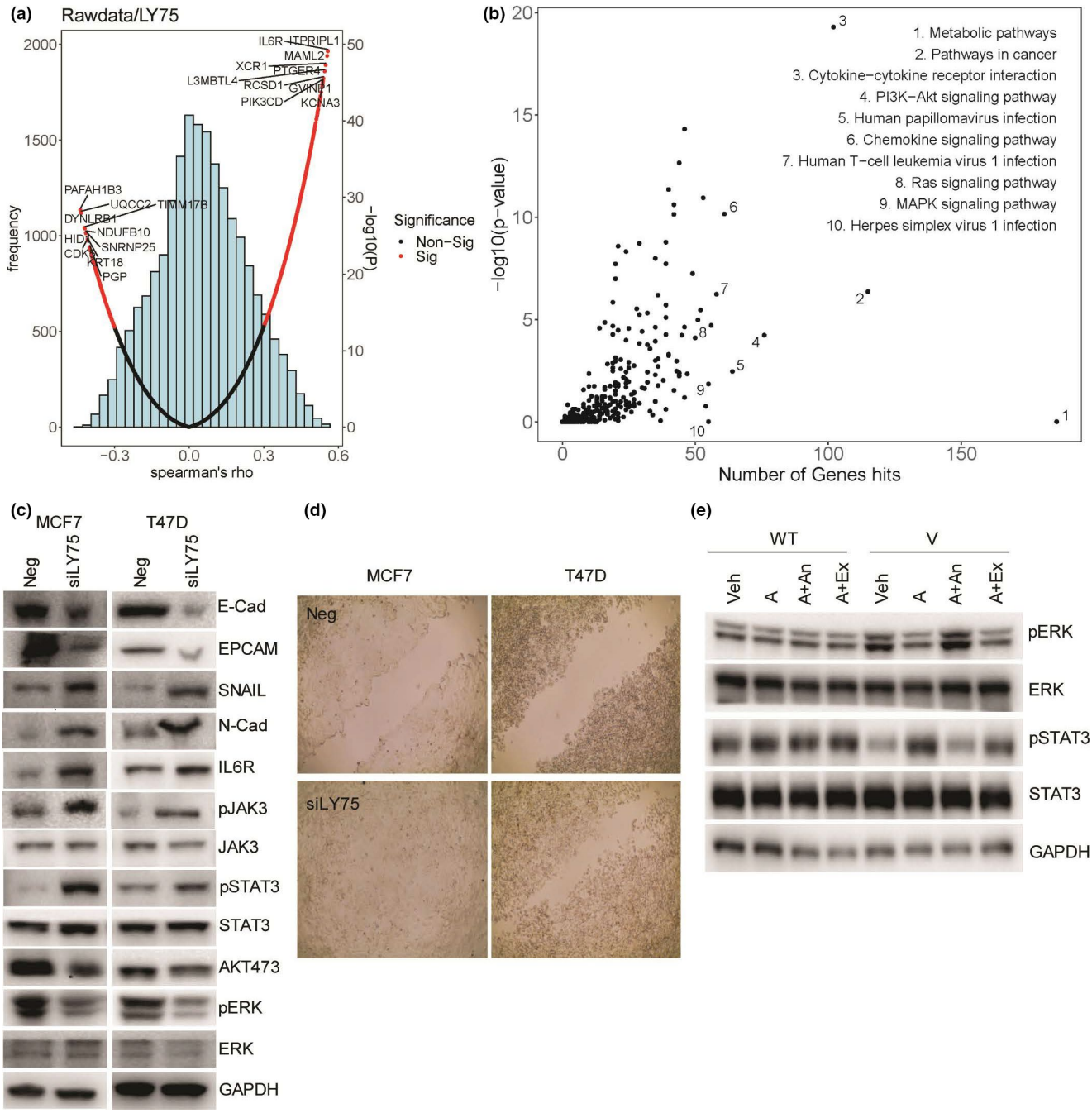

Figure 4 LY75 suppression directs epithelial-to-mesenchymal transition (EMT). (**a**) *LY75* correlated genes in 593 TCGA ER+ patients. (**b**) Pathway analysis of *LY75* correlated gene in TCGA ER+ population. (**c**) Western blot analysis of the expression of different EMT (epithelial and mesenchymal) markers in the control and the *LY75* knockdown breast cancer cells. (**d**) Cell migration of *LY75* knockdown cells was compared to the control cells (Neg). Equal numbers of cells from Neg and siLY75 groups were seeded into 12-well plates for wound healing assay. Original magnification: ×10. (**e**) *LY75* SNP effect on pSTAT3 and pERK protein levels in LCLs with either WT or variant sequences for the SNP treated with androstenedione (A), anastrozole (Ana), and exemestane (Exe). ER, estrogen receptor; LCL, lymphoblastoid cell line; SNP, single-nucleotide polymorphism; TCGA, The Cancer Genome Atlas; V, variant; WT, wild type. [Colour figure can be viewed at wileyonlinelibrary.com]

on BCFI in the MA.27 clinical trial in which postmenopausal women with resected early-stage breast cancer were treated with anastrozole or exemestane adjuvant therapy (**Figure 1a**). The SNPs identified in the interaction analysis were analyzed for their association with BCFI in the treatment-specific cohort (**Figure 1b,c**), as well as in the combined cohort using our pervious GWAS results.[14] This approach allowed the identification of genetic variants that have a differential effect on BCFI between the two AIs (**Figure 1d**). The SNPs that interacted with treatments were otherwise lost in the BCFI analysis when the GWAS was adjusted for the treatment arms (**Figure 1e**, **Table S2**). Our study has provided additional evidence of potential differences among

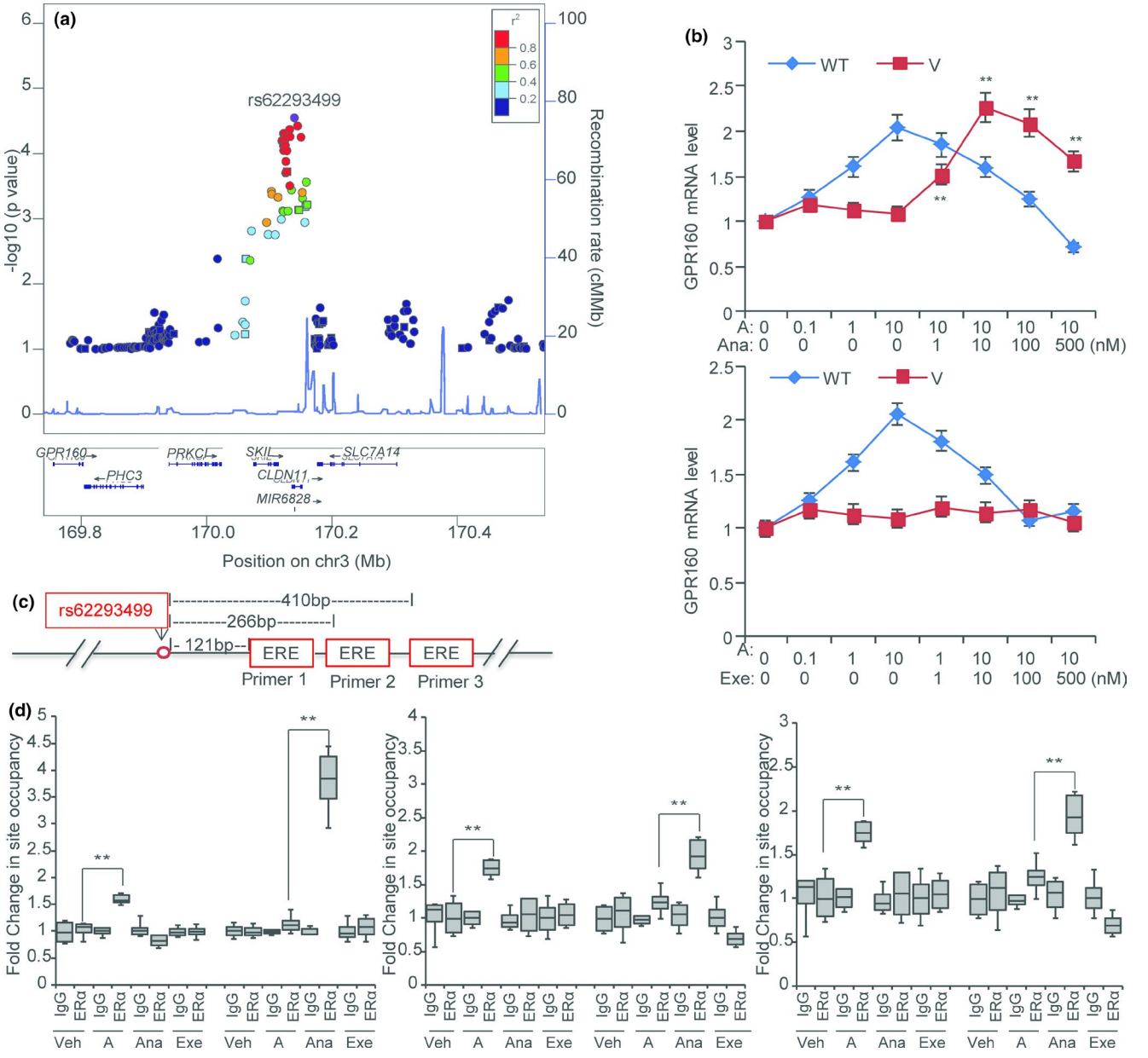

Figure 5 *GPR160* SNP- and anastrozole-dependent gene regulation. (**a**) Locus zoom plot of the chromosome 3 region surrounding the *GPR160* gene. (**b**) SNP-androstenedione- and anastrozole-dependent regulation of *GPR160* gene expression in LCLs selected based on the *GPR160* SNP genotypes. WT: homozygous WT, V: homozygous variant LCLs (n = 3, means ± SEM).**P < 0.01. (**c**) Schematic figure showing the EREs surrounding the rs62293499 SNP. (**d**) ERα ChIP assay that show SNP-dependent ERα binding to the EREs region that are upstream from the rs62293499 SNP in LCLs with different genotypes treated with the indicated drugs: androstenedione (A), anastrozole (Ana), and exemestane (Exe) (n = 3, means ± SEM). **P < 0.01. Statistical test: two-way ANOVA. ANOVA, analysis of variance; ERE, estrogen response element; LCL, lymphoblastoid cell line; SNP, single-nucleotide polymorphism.

the third-generation AIs given differences in individual genetic background. We identified an SNP-based model (**Table S3**, **Figure 1f**) as a biomarker signature panel for predicting superiority in efficacy between anastrozole and exemestane in a given patient. Our current study only focused on White patients because we only had nine patients from Asian or African decedents who have developed a breast event. In the future, it needs to extend these genetic biomarkers to other ethnicities.

Support for the position that anastrozole differs from exemestane is provided by our functional validation studies for the SNPs near the *LY75*[39] and *GPR160*[40] genes (Figures 3a,b and 5a). Of note, there had been no prior reports of a relationship of either of these genes with AI efficacy or breast cancer risk. The expression of *LY75* in cells with variant *LY75* SNP sequences was dysregulated by exposure to anastrozole, but not exemestane (**Figure 3c,d**), which was compatible with the association of variant *LY75* SNP with protection from a breast cancer recurrence while subjects were being treated with exemestane, but not anastrozole. We also showed that the SNP regulated *LY75* expression through ERα (**Figure 3e–j**). We report for the first time a role for the *LY75* gene in controlling

breast cancer cellular phenotypes and corresponding metastatic potential (Figures 4 and 2b). These findings also raise the possibility of targeting downstream EMT in combination with a specific AI to increase its sensitivity for future testing. The *GPR160* SNP mediated *GPR160* expression in response to anastrozole, but not exemestane, by regulating ERα binding to EREs (**Figure 5**). *GPR160* played a role in breast cancer cell growth (**Figure S7b**, **Figure 2c**).

Although all three AIs are interchangeable clinically based on large clinical studies, those studies did not consider patient genetic variability for recruitment and stratification. Our study raises the possibility of salvaging specific AI-treated patients identified to be at increased risk of breast cancer recurrence with a given AI in that they could be switched to an alternative AI. Thus, the genetic markers identified in this study could help select the appropriate AI for a given patient. A limitation of our study is that no other large adjuvant trial of monotherapy with anastrozole and exemestane collected genotyping data is available. However, we have obtained evidence that anastrozole and exemestane have differing efficacy depending on genetic background and have identified an SNP biomarker panel that could potentially be used to select optimal AI for a given patient.

SUPPORTING INFORMATION

Supplementary information accompanies this paper on the *Clinical Pharmacology & Therapeutics* website (www.cpt-journal.com).

ACKNOWLEDGMENTS

The authors acknowledge the women who participated in the MA.27 clinical trial and provided blood samples used in these analyses.

FUNDING

This research was supported in part by NIH grants P50CA116201 (Mayo Clinic Breast Cancer Specialized Program of Research Excellence), U1961388 (The Pharmacogenomics Research Network), U54 GM114838, BCRF-17-075 (the Breast Cancer Research Foundation), U10CA77202, R01CA196648, CCS 015469 from the Canadian Cancer Society, the George M. Eisenberg Foundation for Charities, and the Nan Sawyer Breast Cancer Fund.

CONFLICT OF INTEREST

M.J.E. reports employment with Bioclassifier, LLC; royalty income from Prosigma/Nanostring; honoraria from Nanostring, Novartis, AstraZeneca, Pfizer, Sermonix, Abbvie; patent interest from Bioclassifier, LLC, Prosigma/PAM50; P.E.G. reports research funding from Amgen. M.P.G. reports consulting roles with Lilly, Bovica, Novartis, Sermonix, Context Pharmaceuticals, and Genomic Health; R.M.W. and L.W. are co-founders of, and stockholders in OneOme, LLC. All other authors declared no competing interests for this work.

AUTHOR CONTRIBUTIONS

J.C., K.R.K., J.N.I., R.M.W., and L.W. wrote the manuscript. J.C., J.N.I., and L.W. designed the research. J.C., J.N.I., K.R.K., E.E.C., and P.B. performed the research. J.C., J.N.I., K.R.K., E.E.C., P.B., M.J.E., P.E.G., L.E.S., M.P.G., B.G., and H.G. analyzed the data. K.R.K., E.E.C., and P.B. contributed new reagents/analytical tools.

1. Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R.L., Torre, L.A. & Jemal, A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **68**, 394–424 (2018).
2. Burstein, H.J. *et al*. Adjuvant endocrine therapy for women with hormone receptor-positive breast cancer: American society of clinical oncology clinical practice guideline focused update. *J. Clin. Oncol.* **32**, 2255–2269 (2014).
3. Smith, I.E. & Dowsett, M. Aromatase inhibitors in breast cancer. *N. Engl. J. Med.* **348**, 2431–2442 (2003).
4. Baum, M. *et al*. Anastrozole alone or in combination with tamoxifen versus tamoxifen alone for adjuvant treatment of postmenopausal women with early-stage breast cancer: results of the ATAC (arimidex, tamoxifen alone or in combination) trial efficacy and safety update analyses. *Cancer* **98**, 1802–1810 (2003).
5. Breast International Group (BIG) 1-98 Collaborative group *et al*. A comparison of letrozole and tamoxifen in postmenopausal women with early breast cancer. *N. Engl. J. Med.* **353**, 2747–2757 (2005).
6. Coombes, R.C. *et al*. A randomized trial of exemestane after two to three years of tamoxifen therapy in postmenopausal women with primary breast cancer. *N. Engl. J. Med.* **350**, 1081–1092 (2004).
7. Early Breast Cancer Trialists' Collaborative Group. Aromatase inhibitors versus tamoxifen in early breast cancer: patient-level meta-analysis of the randomised trials. *Lancet* **386**, 1341–1352 (2015).
8. Geisler, J. *et al*. Letrozole is superior to anastrozole in suppressing breast cancer tissue and plasma estrogen levels. *Clin. Cancer Res.* **14**, 6330–6335 (2008).
9. Smith, I. *et al*. Comparative efficacy and safety of adjuvant letrozole versus anastrozole in postmenopausal patients with hormone receptor-positive, node-positive early breast cancer: final results of the randomized phase III femara versus anastrozole clinical evaluation (FACE) trial. *J. Clin. Oncol.* **35**, 1041–1048 (2017).
10. Goss, P.E. *et al*. Exemestane versus anastrozole in postmenopausal women with early breast cancer: NCIC CTG MA.27–a randomized controlled phase III trial. *J. Clin. Oncol.* **31**, 1398–1404 (2013).
11. De Placido, S. *et al*. Adjuvant anastrozole versus exemestane versus letrozole, upfront or after 2 years of tamoxifen, in endocrine-sensitive breast cancer (FATA-GIM3): a randomised, phase 3 trial. *Lancet Oncol.* **19**, 474–485 (2018).
12. Ingle, J.N. *et al*. Anastrozole has an association between degree of estrogen suppression and outcomes in early breast cancer and is a ligand for estrogen receptor alpha. *Clin. Cancer Res.* **26**, 2986–2996 (2020).
13. Cairns, J. *et al*. Pharmacogenomics of aromatase inhibitors in postmenopausal breast cancer and additional mechanisms of anastrozole action. *Science* **5(16)**, e137571 (2020). https://doi.org/10.1172/jci.insight.137571
14. Ingle, J.N. *et al*. Genetic polymorphisms in the long noncoding RNA MIR2052HG offer a pharmacogenomic basis for the response of breast cancer patients to aromatase inhibitor therapy. *Cancer Res.* **76**, 7012–7023 (2016).
15. Cairns, J. *et al*. The lncRNA MIR2052HG regulates ERalpha levels and aromatase inhibitor resistance through LMTK3 by recruiting EGR1. *Breast Cancer Res.* **21**, 47 (2019).
16. Ingle, J.N. *et al*. Genome-wide associations and functional genomic studies of musculoskeletal adverse events in women receiving aromatase inhibitors. *J. Clin. Oncol.* **28**, 4674–4682 (2010).
17. Liu, M. *et al*. Aromatase inhibitor-associated bone fractures: a case-cohort GWAS and functional genomics. *Mol. Endocrinol.* **28**, 1740–1751 (2014).
18. Hudis, C.A. *et al*. Proposal for standardized definitions for efficacy end points in adjuvant breast cancer trials: the STEEP system. *J. Clin. Oncol.* **25**, 2127–2132 (2007).
19. Price, A.L., Patterson, N.J., Plenge, R.M., Weinblatt, M.E., Shadick, N.A. & Reich, D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* **38**, 904–909 (2006).

20. Patterson, N., Price, A.L. & Reich, D. Population structure and Eigen analysis. *PLoS Genet.* **2**, e190 (2006).
21. Purcell, S. *et al*. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
22. Price, A.L. *et al*. Long-range LD can confound genome scans in admixed populations. *Am. J. Hum. Genet.* **83**, 132–135 (2008), author reply 135–139.
23. Zhu, Z. *et al*. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat. Genet.* **48**, 481–487 (2016).
24. GTEx Consortium. The Genotype-Tissue Expression (GTEx) pilot analysis. *Science* **348**, 648–660 (2015).
25. Cairns, J., Ingle, J.N., Wickerham, L.D., Weinshilboum, R., Liu, M. & Wang, L. SNPs near the cysteine proteinase cathepsin O gene (CTSO) determine tamoxifen sensitivity in ERalpha-positive breast cancer through regulation of BRCA1. *PLoS Genet.* **13**, e1007031 (2017).
26. Hanson, C., Cairns, J., Wang, L. & Sinha, S. Principled multi-omic analysis reveals gene regulatory mechanisms of phenotype variation. *Genome Res.* **28**, 1207–1216 (2018).
27. Cairns, J. *et al*. A network-based phenotype mapping approach to identify genes that modulate drug response phenotypes. *Sci. Rep.* **6**, 37003 (2016).
28. Niu, N. *et al*. Radiation pharmacogenomics: a genome-wide association approach to identify radiation response biomarkers using human lymphoblastoid cell lines. *Genome Res.* **20**, 1482–1492 (2010).
29. Ingle, J.N. *et al*. Selective estrogen receptor modulators and pharmacogenomic variation in ZNF423 regulation of BRCA1 expression: individualized breast cancer prevention. *Cancer Discov.* **3**, 812–825 (2013).
30. Wang, W.Y., Barratt, B.J., Clayton, D.G. & Todd, J.A. Genome-wide association studies: theoretical and practical concerns. *Nat. Rev. Genet.* **6**, 109–118 (2005).
31. Lachance, J. Disease-associated alleles in genome-wide association studies are enriched for derived low frequency alleles relative to HapMap and neutral expectations. *BMC Med. Genomics* **3**, 57 (2010).
32. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature* **490**, 61–70 (2012).
33. Ciriello, G. *et al*. Comprehensive molecular portraits of invasive lobular breast cancer. *Cell* **163**, 506–519 (2015).
34. Yu, H., Pardoll, D. & Jove, R. STATs in cancer inflammation and immunity: a leading role for STAT3. *Nat. Rev. Cancer* **9**, 798–809 (2009).
35. Sullivan, N.J. *et al*. Interleukin-6 induces an epithelial-mesenchymal transition phenotype in human breast cancer cells. *Oncogene* **28**, 2940–2947 (2009).
36. Yadav, A., Kumar, B., Datta, J., Teknos, T.N. & Kumar, P. IL-6 promotes head and neck tumor metastasis by inducing epithelial-mesenchymal transition via the JAK-STAT3-SNAIL signaling pathway. *Mol. Cancer Res.* **9**, 1658–1667 (2011).
37. Gong, C. *et al*. Abnormally expressed JunB transactivated by IL-6/STAT3 signaling promotes uveal melanoma aggressiveness via epithelial-mesenchymal transition. *Biosci. Rep.* **38**(4), BSR20180532 (2018).
38. Kalluri, R. & Weinberg, R.A. The basics of epithelial-mesenchymal transition. *J. Clin. Invest.* **119**, 1420–1428 (2009).
39. East, L. & Isacke, C.M. The mannose receptor family. *Biochim. Biophys. Acta* **1572**, 364–386 (2002).
40. Regard, J.B., Sato, I.T. & Coughlin, S.R. Anatomical profiling of G protein-coupled receptor expression. *Cell* **135**, 561–571 (2008).