nature genetics

Acquired *CYP19A1* amplification is an early specific mechanism of aromatase inhibitor resistance in ERα metastatic breast cancer

Luca Magnani[1,13], Gianmaria Frigè[2,13], Raffaella Maria Gadaleta[1], Giacomo Corleone[1], Sonia Fabris[3], Hermannus Kempe[4], Pernette J Verschure[4], Iros Barozzi[5], Valentina Vircillo[6], Sung-Pil Hong[1], Ylenia Perone[1], Massimo Saini[7,8], Andreas Trumpp[7,8], Giuseppe Viale[9], Antonino Neri[3,10], Simak Ali[1], Marco Angelo Colleoni[11], Giancarlo Pruneri[8] & Saverio Minucci[2,12]

© 2017 Nature America, Inc., part of Springer Nature. All rights reserved.

Tumor evolution is shaped by many variables, potentially involving external selective pressures induced by therapies[1]. After surgery, patients with estrogen receptor (ERα)-positive breast cancer are treated with adjuvant endocrine therapy[2], including selective estrogen receptor modulators (SERMs) and/or aromatase inhibitors (AIs)[3]. However, more than 20% of patients relapse within 10 years and eventually progress to incurable metastatic disease[4]. Here we demonstrate that the choice of therapy has a fundamental influence on the genetic landscape of relapsed diseases. We found that 21.5% of AI-treated, relapsed patients had acquired *CYP19A1* (encoding aromatase) amplification (*CYP19A1*^{amp}). Relapsed patients also developed numerous mutations targeting key breast cancer–associated genes, including *ESR1* and *CYP19A1*. Notably, *CYP19A1*^{amp} cells also emerged *in vitro*, but only in AI-resistant models. *CYP19A1* amplification caused increased aromatase activity and estrogen-independent ERα binding to target genes, resulting in *CYP19A1*^{amp} cells showing decreased sensitivity to AI treatment. These data suggest that AI treatment itself selects for acquired *CYP19A1*^{amp} and promotes local autocrine estrogen signaling in AI-resistant metastatic patients.

ERα activation characterizes more than 70% of breast cancers, where it represents the key prognostic factor and therapeutic target[5]. ERα activation is primarily dependent on circulating estrogens and results in genome-wide chromatin binding at thousands of regulatory regions[6]. ERα binding leads to the transcription of hundreds of genes central to breast cancer growth[6]. Endocrine therapies, including SERMs and AIs, were developed to prevent ERα activation and block breast cancer growth[5]. The mechanisms behind drug resistance are only partially understood and often involve transcriptional activation of alternative survival pathways, especially at later stages of the disease[7]. Nonetheless, recent genomic studies highlight how ERα signaling might still have a role in metastatic disease. For example, activating somatic mutations targeting *ESR1* (the gene encoding ERα) are found at higher frequencies after endocrine therapy[8,9]. These mutations have been characterized in metastatic lesions from patients that received several cycles of endocrine therapy and chemotherapy[10,11], suggesting that the selective pressure imposed by endocrine treatments might favor the development of focused genetic aberrations during tumor evolution[11]. However, it is impossible to infer from most studies when genetic aberrations originate and how they are selected, because patients are biopsied after multiple treatments. The SERM tamoxifen (TAM) directly blocks ERα coactivation in the tumor cell, whereas AI targets CYP19A1 (aromatase) in the peripheral tissue, thereby lowering estrogen availability. We recently reported that ERα-positive breast cancer cells activate alternative epigenetic programs in response to TAM or AI[12], suggesting that choice of endocrine therapies might affect tumor evolution. Here we examine, in parallel, a cohort of ERα-positive patients who were treated with single-agent adjuvant endocrine therapies (TAM or nonsteroidal AI) and re-biopsied when they had their first distal relapse (**Fig. 1a** and **Supplementary Figs. 1** and **2**).

We initially assessed copy number alterations (CNAs) of the genes encoding the targets of AI and TAM (*CYP19A1* (15q21) and *ESR1* (6q25), respectively), considering the central role of copy number

[1]Department of Surgery and Cancer, Imperial College London, London, UK. [2]Department of Experimental Oncology, European Institute of Oncology, Milan, Italy. [3]Hematology Unit, Fondazione IRCCS Ca' Granda, Ospedale Maggiore Policlinico, Milan, Italy. [4]Swammerdam Institute for Life Sciences, University of Amsterdam, Amsterdam, the Netherlands. [5]Genomics Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA. [6]Department of Pharmacy, Health and Nutritional Sciences, University of Calabria, Arcavacata di Rende, Italy. [7]Division of Stem Cells and Cancer, Deutsches Krebsforschungszentrum (DKFZ), Heidelberg, Germany. [8]Institute for Stem Cell Technology and Experimental Medicine GmbH, Deutsches Krebsforschungszentrum (DKFZ), Heidelberg, Germany. [9]Division of Pathology, European Institute of Oncology and University of Milan, School of Medicine, Milan, Italy. [10]Department of Oncology and Hemato-oncology, University of Milano, Milan, Italy. [11]Division of Medical Senology, European Institute of Oncology (IEO), Milan, Italy. [12]Department of Biosciences, University of Milano, Milan, Italy. [13]These authors contributed equally to this work. Correspondence should be addressed to S.M. (saverio.minucci@ieo.eu), G.P. (giancarlo.pruneri@ieo.it) or L.M. (l.magnani@imperial.ac.uk).

Received 13 September 2016; accepted 21 December 2016; published online 23 January 2017; corrected online 31 January 2017 (details online); doi:10.1038/ng.3773

© 2017 Nature America, Inc., part of Springer Nature. All rights reserved.

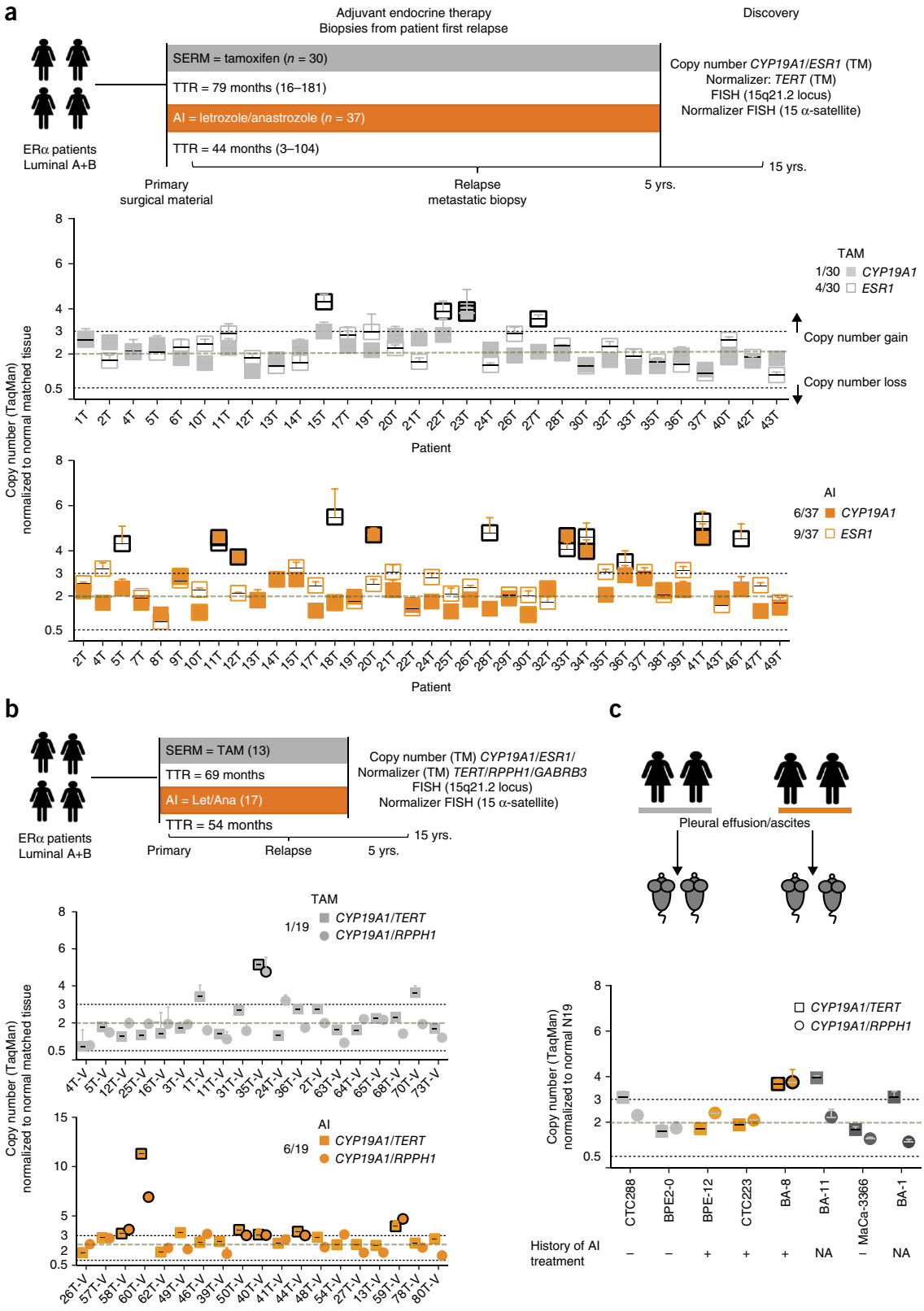

Figure 1 Clinical treatments shape cancer genetic evolution. (**a**) Clinical discovery cohorts and sample design used in the study. CNA profiles for the *CYP19A1* and *ESR1* loci in the first relapse of patients treated with adjuvant tamoxifen (TAM) or AI monotherapy (letrozole or anastrozole (Let/Ana). TM, TaqMan assay. (**b**) Clinical discovery cohorts and sample design used in the study. CNA profiles for the *CYP19A1* and *ESR1* loci in the first relapse of patients treated with adjuvant tamoxifen or AI monotherapy (*ESR1* data are shown in **Supplementary Fig. 4**). Data were normalized to the *TERT* and *RPPH1* loci. (**c**) PDX cohort. CNA profiles for the *CYP19A1* and *ESR1* loci in PDXs from patients treated with tamoxifen or AI. Labels indicate the origin of unique PDXs (Online Methods). (*ESR1* data are shown in **Supplementary Fig. 4**). Boxes and error bars, mean ± s.d. of 3 technical replicates.

changes in breast cancer[13]. Meta-analysis of data from primary, treatment-naive patients using GISTIC-based[14] cBioPortal[15] showed that *CYP19A1* and *ESR1* CNAs are exceedingly rare in ERα-positive primary breast cancers (0.006%, 2/321 for *CYP19A1* and 0.018%, 6/321 for *ESR1* in ERα-positive primary breast cancer, based on The Cancer Genome Atlas (TCGA) CNAs data[16]; threshold = 1.5-fold change). Using an independent database of SNP-array-based studies with an alternative CNA algorithm[17], we confirmed the rarity of *CYP19A1* amplification events (**Supplementary Table 1**). *CYP19A1* and *ESR1* amplification are rare in other primary cancers as well (**Supplementary Fig. 3** and **Supplementary Table 1**). These data demonstrate that *CYP19A1* and *ESR1* loci are not rearrangement hot spots in untreated primary cancers. We then analyzed our discovery cohort consisting of tumor samples collected from the first relapse after single therapy using a TaqMan CNA assay, comparing metastatic with matched normal breast tissue. Notably, we found that the *CYP19A1* locus is amplified (*CYP19A1*^{amp}) in 6/37 (16%) of patients that received AI. Conversely, only one patient (3%) that received TAM showed evidence of *CYP19A1*^{amp} (**Fig. 1a**). The *ESR1* locus was also substantially amplified in material from relapsed patients (24% and 13%, AI- and TAM-treated cohorts, respectively) (**Fig. 1a**). To confirm these data, we investigated an independent validation cohort with similar clinical characteristics. In agreement with the discovery cohort, we found that *CYP19A1* was amplified in 6/19 (32%) of AI-treated patients and only 1/19 (5%) of TAM-treated patients (**Fig. 1b**). *ESR1* was amplified in 4/19 (21%) of AI-treated and 0/19 of the TAM-treated relapse samples (**Supplementary Fig. 4a**). The *CYP19A1* locus showed evidence for both focal and chromosome-wide amplification (**Supplementary Fig. 5a**). *CYP19A1* and *ESR1* CNAs might work cooperatively, considering the rate of coamplification in AI-treated patients (8/12 *CYP19A1*^{amp} patients also carried *ESR1*^{amp}; **Supplementary Fig. 5b**). Notably, we also identified *CYP19A1* and *ESR1* amplification in patient-derived xenografts (PDXs) from patients previously treated with nonsteroidal AI (**Fig. 1c** and **Supplementary Fig. 4b**). Collectively, these data show that treatment with reversible AI significantly increased the frequency of *CYP19A1*^{amp} at first distal relapse (21.5% versus 4%, AI versus TAM, $P = 0.009$, $P = 0.004$ including PDXs, two-tailed Fisher's exact test). Similarly, we observed a trend in AI-treated patients of preferential amplification of the *ESR1* locus (23% versus 8%, AI versus TAM, $P = 0.06$, $P = 0.03$ including PDXs, two-tailed Fisher's exact test). *CYP19A* and *ESR1* amplification in distal relapses from AI-resistant breast cancers is strongly reminiscent of androgen receptor (AR) amplification in patients with castration-resistant prostate cancer[18,19].

We next designed a DNA-FISH assay to validate *CYP19A1* amplification and investigate its degree of heterogeneity. We examined four cases that were found to be amplified by TaqMan; all of them present strong evidence for cluster amplification (**Fig. 2a**,**b**). FISH analysis also confirmed 100% of TaqMan calls in the validation data set (**Supplementary Fig. 6**). More than 90% of nuclei from each of the metastatic samples examined by FISH have *CYP19A1* amplification signals, indicating that *CYP19A1*^{amp} cells represent the dominant clone. Additionally, the ratio of the tandemly repeated centromeric 15 α-satellite to *CYP19A1* strongly suggests that *CYP19A1* amplification is not a consequence of unspecific aneuploidy (**Fig. 2b**). Using DNA-FISH, we did not find convincing evidence of *CYP19A1* amplification in the respective primary samples (**Fig. 2a**). Therefore, these results support the notion that *CYP19A1* amplification occurs under treatment, although we cannot exclude the presence of very small *CYP19A1*^{amp} subclones at diagnosis.

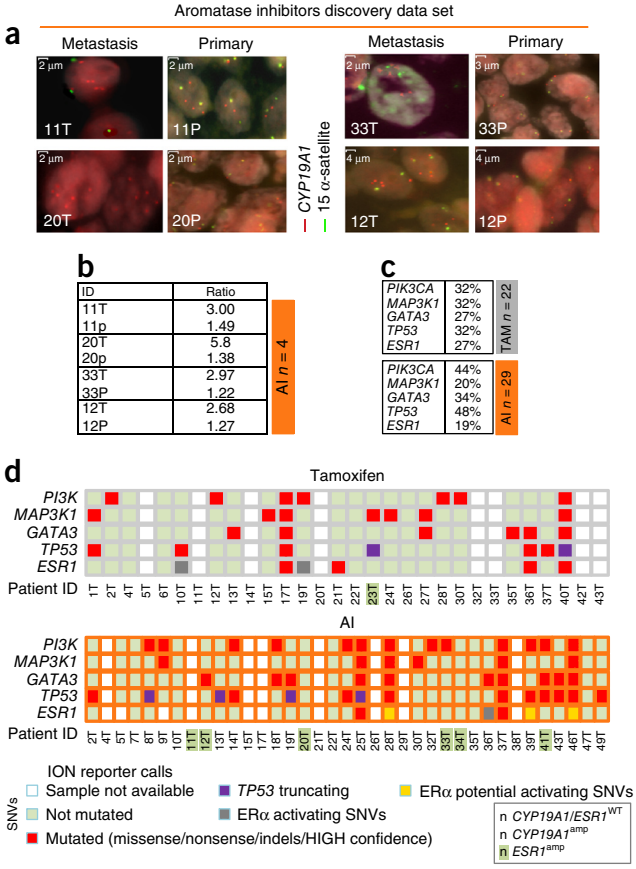

Figure 2 AI-resistant metastases develop cluster amplification of *CYP19A1*. (**a**) Double-color FISH analyses to measure DNA copy number using probes for 15 α-satellite and *CYP19A1* identify cluster amplification of the *CYP19A1* locus. (**b**) Ratio of amplification obtained by computing *CYP19A1*/α-satellite signal in 30 representative individual cancer cells from each validated tumor sample. (**c**) Breast cancer–associated mutations in tamoxifen- and AI-treated metastatic samples. (**d**) Mutations in the two cohorts.

We then investigated the frequency of metastasis-specific *ESR1* activating mutations[11] and other commonly occurring mutations[13,16] (in *PIK3CA*, *MAPK*, *TP53* and *GATA3*) in AI- or TAM-treated distal relapses compared to matched normal germline DNA using an AmpliSeq custom panel for targeted sequencing (**Supplementary Table 2** and Online Methods). Overall, we found similar patterns of mutations between the two cohorts (**Fig. 2c** and **Supplementary Data 1** and **2**). These patterns, however, may be different from those previously characterized in primary breast cancers (**Supplementary Fig. 7**). For example, we identified several novel *ESR1* mutations, including predicted activating mutations encoding p.Leu536His[20] (patient ID TAM 19T), p.Met543Ala[21] (AI 28T and AI 46T) p.Asp538Ala (TAM 10T) and p.Arg503Gln (AI 28T and AI 39T) (**Fig. 2c**,**d** and **Supplementary Data 1** and **2**), in addition to the common p.Tyr537Ser (AI 36T). These mutations occur at relatively high allele frequencies and can be polyclonal (AI 28T) (**Supplementary Fig. 8a**). In addition, we identified several truncating *TP53* mutations (encoding, for example, p.Arg213* (AI 8T), p.Cys242* (AI 13T) and p.Glu294* (TAM 23T); **Fig. 2c**,**d** and **Supplementary Data 1** and **2**). Most of these mutations were confirmed by a second assay (**Supplementary Fig. 8b**). Notably, we designed probes against *CYP19A1* and identified the novel recurrent alteration p.Pro410Ser/Leu (AI 5T and 18T; **Supplementary Fig. 8b**

© 2017 Nature America, Inc., part of Springer Nature. All rights reserved.

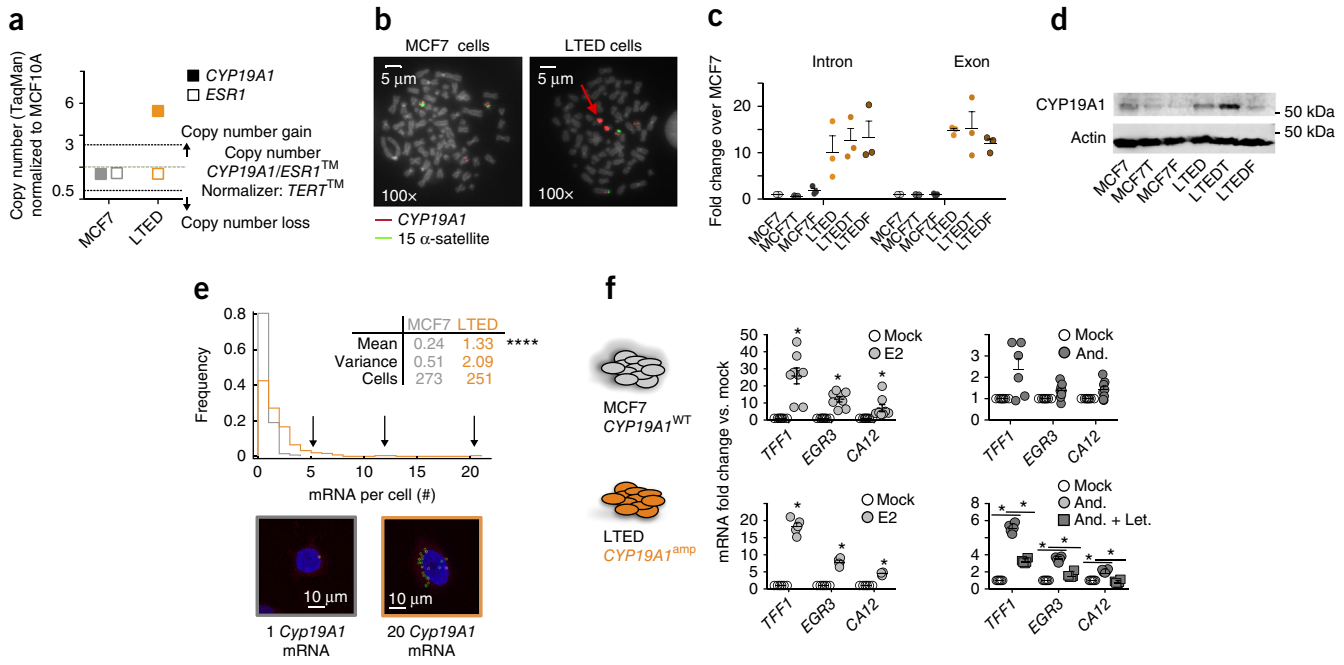

Figure 3 *CYP19A1* amplification leads to increased aromatase activity. (**a**) CNA profiles for the *CYP19A1* and *ESR1* loci in treatment-naive (MCF7) and estrogen-deprived (LTED) cells. Boxes and error bars, mean ± s.d. of 3 technical replicates. Data are normalized to the *TERT* locus. (**b**) DNA-FISH using *CYP19A1*-centered probes identifies widespread cluster amplification in LTED cells. (**c**,**d**) RT-qPCR (**c**) and immunoblot (**d**) analysis of aromatase expression at the CYP19A1 locus in LTED clones. T suffix, tamoxifen-resistant; F, fulvestrant-resistant. Data represent mean ± s.e.m. from 3 independent experiments (full blots shown in **Supplementary Fig. 11**). (**e**) Single-cell RNA-FISH highlights heterogeneity in aromatase expression. $^{****}P < 10^{-20}$, Mann-Whitney test. (**f**) Transcriptional activation of estrogen target genes in *CYP19A1*^{amp} cells treated with ethanol (mock) or aromatizable androstenedione (and.) alone or in combination with letrozole (let.). Data are mean ± s.e.m. from 5 independent experiments. $^{*}P < 0.05$, two-way or one-way ANOVA (bottom right).

and **Supplementary Data 2**). This alteration was mutually exclusive with *CYP19A1* amplification, further suggesting its functional significance (**Supplementary Fig. 8b**). Interestingly, Pro410 might be involved in the formation of a functional channel[22]. Collectively, these data demonstrate that several metastasis-specific mutations might have a role in early tumor progression.

In postmenopausal women, aromatase inhibitors target the peripheral conversion of testosterone or androstenedione to estrogens by inhibiting aromatase. Cholesterol is the common precursor of sex hormones, including testosterone and androstenedione, the two substrates of the aromatase enzyme[23]. *In vitro,* AI-resistant cells develop *de novo* cholesterol biosynthesis via epigenetic events to promote autonomous ERα activation[12]. Current treatment protocols can effectively remove all estradiol from serum used in culture, but they are far less efficient in removing testosterone[24]. Thus, we hypothesized that *CYP19A1* amplification might favor autocrine estrogen production starting from epigenetically driven *de novo* cholesterol biosynthesis or residual male hormones. We investigated the possibility of some CNAs contributing to AI resistance using a shallow whole-genome sequencing approach[25] to profile several cell lines derived from ERα-positive breast cancer MCF7 cells and exposed to distinct endocrine treatments[12,26] (**Supplementary Fig. 9a**). Among those, LTED cells are derived upon chronic estrogen deprivation[26,27], mimicking AI-treated breast cancer. Genomic analysis suggests that all endocrine-resistant cells maintain DNA rearrangements of parental MCF7 cells (3p13–p14 and 20q12–qter gains). However, LTED cells also acquired CNAs around the *CYP19A1* locus (**Supplementary Fig. 9b**). *CYP19A1* CNAs were identified in LTED cells but not in tamoxifen-fulvestrant-resistant models (**Supplementary Fig. 9b**). We next quantified *CYP19A1* CNAs using our TaqMan assay (**Fig. 3a**) and DNA-FISH (**Fig. 3b**). Notably, the degree of amplification was comparable to that observed in AI-treated patients *in vivo* (**Figs. 1c** and **3a**). We could not identify a single *CYP19A1*^{amp} cell among MCF7 cells, suggesting that *CYP19A1* CNAs were acquired during treatment. Compared to parental MCF7 cells, LTED cells undergo a tenfold increase in ERα protein levels[12], although we did not find significant evidence for *ESR1* amplification or ERα activating mutations. Collectively, these data indicate that estrogen deprivation promotes *CYP19A1* CNAs *in vitro* similarly to AI treatment *in vivo*. *CYP19A1*^{amp} LTED cells have higher levels of both *CYP19A1* mRNA and CYP19A1 protein than MCF7 cells carrying wild-type *CYP19A1* (*CYP19A1*^{WT}) (**Fig. 3c–e**). Single-cell RNA-FISH (Online Methods) showed a large degree of heterogeneity in *CYP19A1* expression; however, *CYP19A1*^{amp} cells had a significantly higher chance of accumulating more *CYP19A1* mRNA molecules per cell (**Fig. 3e**). We also found that LTED cells had significantly higher aromatase activity than MCF7 cells, and this could be partially antagonized with AI ($P < 0.05$, one-way ANOVA with Tukey's post-test; **Supplementary Fig. 10a**). *CYP19A1*^{amp} LTED cells, but not *CYP19A1*^{WT} MCF7 cells, showed significant transcriptional activation of three well-characterized estrogen target genes[7] (*TFF1*, *EGR3* and *CA12*) in response to androstenedione (**Fig. 3f**). More notably, transcriptional activation by androstenedione could be partially blocked by AI (letrozole) (**Fig. 3f**). These data indicate that acquired *CYP19A1* amplification drives increased endogenous aromatase activity and promotes estrogen-independent transcription.

When we investigated genome-wide ERα binding to target genes in MCF7 and LTED cells in the absence of estrogens, we found that the chromatin of LTED cells was strongly enriched in ERα binding

© 2017 Nature America, Inc., part of Springer Nature. All rights reserved.

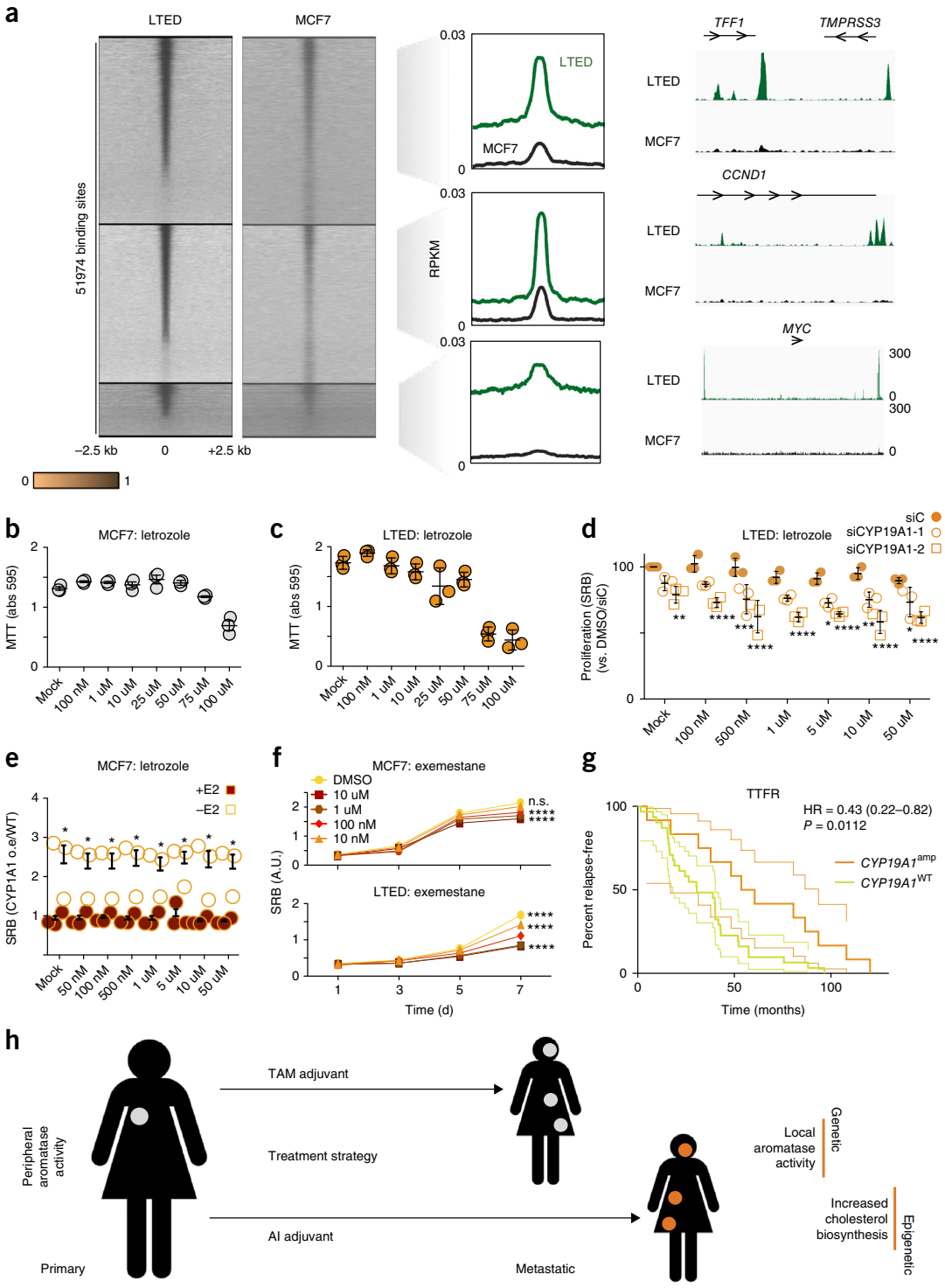

Figure 4 *CYP19A1*^{amp} cells endogenously activate ERα and develop tolerance to AI. (**a**) ChIP-seq heat maps for ERα in MCF7 and LTED cells. Binding sites have been assigned to three clusters. The average profile of each cluster is shown (middle). Examples of ERα enrichment near important estrogen target genes are shown in the insets (right). (**b**) IC_{50} growth curve for MCF7 cell treated with increasing doses of AI in combination with estradiol. (**c**) LTED treatment with AI in the absence of estradiol. (**d**) LTED cells treated with siRNA against *CYP19A1* (siCYP19A1-1 and siCYP19A1-2) have increased sensitivity to AI. SRB, sulforhodamine B assay; siC, control siRNA targeting luciferase. (**e**) CYP19A1-overexpressing cells have a growth advantage over wild-type in the absence of estradiol. Relative increase in growth rate is shown as the ratio of the growth of CYP19A1-overexpressing cells to CYP19A1 wild-type cells under letrozole challenge. (**f**) *CYP19A1*^{amp} LTED cells respond to low levels of irreversible steroidal AI. (**g**) Kaplan–Meier curve showing time to first relapse (TTFR) for AI-treated patients stratified retrospectively for *CYP19A1* amplification. Dotted lines represent 95% CI. HR, hazard ratio. *P* value determined by log–rank test. (**h**) Working hypothesis for therapy-specific breast cancer progression. Genetic and epigenetic changes collaborate to increase tumor fitness by creating an estrogen-independent niche at metastatic sites treated with AI therapy. For **b**–**f**, data are mean ± s.e.m. (**b**,**d**,**e**) or mean ± s.d. (**c**) from 3 independent experiments or mean and 95% CI from 4 independent replicates (**f**). $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$, Student's *t*-test (**c**) or two-way ANOVA and Bonferroni's post-test (**d**–**f**). A.U., arbitrary units.

© 2017 Nature America, Inc., part of Springer Nature. All rights reserved.

compared to parental MCF7 cells (**Fig. 4a**), which is consistent with our transcriptional data. As expected, MCF7 cells grown in the presence of estrogen were not sensitive to AI (**Fig. 4b**), whereas *CYP19A1*^{amp} LTED cells had low sensitivity (half-maximal inhibitory concentration (IC_{50}) = 80 μM; **Fig. 4c**). To test the role of *CYP19A1*^{amp} on AI sensitivity, we then treated LTED cells and parental MCF7 cells with two independent small interfering RNAs (siRNAs) targeting *CYP19A1* and measured cell viability in response to the AI letrozole (**Fig. 4b–d** and **Supplementary Fig. 10b**). siRNAs against *CYP19A1* significantly increased sensitivity to AI treatment in *CYP19A1*^{amp} LTED cells (**Fig. 4d**) but did not affect MCF7 cells grown in estrogen-supplemented conditions (**Fig. 4e** and **Supplementary Fig. 10c**). *CYP19A1* overexpression did not confer a growth advantage to MCF7 cells grown in the presence of estradiol. However, *CYP19A1* overexpression was sufficient to relieve cell cycle arrest in MCF7 cells cultured in absence of estrogens. Notably, this effect was not antagonized by letrozole (**Fig. 4e**). Finally, we confirmed that ERα still has a role in *CYP19A1*^{amp} LTED cell growth, as shown by LTED cell sensitivity to fulvestrant treatment (**Supplementary Fig. 10d**). Moreover, our results in *CYP19A1*^{amp} LTED cells treated with an irreversible AI (exemestane) suggested that increased aromatase activity can be antagonized with a steroidal AI (**Fig. 4f**). Collectively, these data support our initial hypothesis and suggest that *CYP19A1* amplification might induce reduced sensitivity to reversible AI treatment.

CYP19A1 amplification triggers ERα activity by converting male sex hormones obtained through endogenous epigenetic cholesterol biosynthesis[12] or circulating within the tumor microenvironment. *In vivo* and *in vitro* data indicate that *CYP19A1* CNAs are acquired rather than selected. Indeed, LTED cells develop *CYP19A1* CNAs over the course of chronic estrogen deprivation (>1 year). We then analyzed the time to first relapse, merging our AI patient data sets. Relapses characterized by *CYP19A1* CNAs emerged significantly later than those with wild-type *CYP19A1* (median 57 versus 30 months, amplified versus wild-type $P = 0.0112$, log-rank Mantel-Cox test; **Fig. 4g**). These data support the notion that breast cancer cells treated with AI slowly evolve to create a favorable autocrine microenvironment through genetic and epigenetic reprogramming (**Fig. 4h**). One unexplained clinical observation is that patients progressing under reversible AI treatment (letrozole or anastrozole) occasionally respond to irreversible AI (i.e., exemestane). Notably, five *CYP19A1*^{amp} AI-treated patients in our study showed stabilized disease for ~1 year, on average, after switching to exemestane (**Supplementary Figs. 1** and **2**). Thus, it is tempting to speculate that *CYP19A1* amplification might arise in response to reversible inhibitors but could be antagonized by switching to irreversible inhibitors. Alternatively, it should be clinically feasible to directly antagonize the low levels of circulating male hormones commonly found in postmenopausal women. Considering that AI normally targets peripheral tissues, our data also warrant AI pharmacodynamics studies to evaluate the ability of this class of drugs to target tumor cells directly. Taken together, our clinical data demonstrate that the evolution of breast cancer is shaped by clinical intervention and thus advocate the development of treatment- and setting-specific biomarkers.

METHODS

Methods, including statements of data availability and any associated accession codes and references, are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We thank all participants and their families. We thank A. Bardelli for his comments. We thank D. Patten for help with the exemestane study. We thank L. Watson for her help with the manuscript. We thank J.Bean for support. For these studies, S.M. and G.P. were supported by Associazione Italiana Ricerca sul Cancro (AIRC) (5x1000 campaign). L.M. was supported by the Imperial College Junior Research Fellowship. S.-P.H. was supported by Cancer Research UK (CRUK) grant C37/A18784. Y.P. was supported by CRUK PhD studentship P55374. G.C. was supported by the EpiPredict project (European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement 642691).

AUTHOR CONTRIBUTIONS

L.M. conceived the study and wrote the manuscript. L.M., S.M. and G.P. planned and supervised all experiments. L.M., G.F., S.-P.H., Y.P., R.M.G., S.F., H.K., and V.V. performed experiments. G.C. and I.B. performed bioinformatics analyses. P.J.V., G.V., A.N., M.S., A.T., S.A. and M.A.C. provided reagents, samples and intellectual contribution. All authors discussed the results and commented on the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Nowell, P.C. The clonal evolution of tumor cell populations. *Science* **194**, 23–28 (1976).
2. Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Relevance of breast cancer hormone receptors and other factors to the efficacy of adjuvant tamoxifen: patient-level meta-analysis of randomised trials. *Lancet* **378**, 771–784 (2011).
3. Pagani, O. *et al.* Adjuvant exemestane with ovarian suppression in premenopausal breast cancer. *N. Engl. J. Med.* **371**, 107–118 (2014).
4. Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Aromatase inhibitors versus tamoxifen in early breast cancer: patient-level meta-analysis of the randomised trials. *Lancet* **386**, 1341–1352 (2015).
5. Musgrove, E.A. & Sutherland, R.L. Biological determinants of endocrine resistance in breast cancer. *Nat. Rev. Cancer* **9**, 631–643 (2009).
6. Carroll, J.S. *et al.* Chromosome-wide mapping of estrogen receptor binding reveals long-range regulation requiring the forkhead protein FoxA1. *Cell* **122**, 33–43 (2005).
7. Magnani, L. *et al.* Genome-wide reprogramming of the chromatin landscape underlies endocrine therapy resistance in breast cancer. *Proc. Natl. Acad. Sci. USA* **110**, E1490–E1499 (2013).
8. Fuqua, S.A., Chamness, G.C. & McGuire, W.L. Estrogen receptor mutations in breast cancer. *J. Cell. Biochem.* **51**, 135–139 (1993).
9. Fribbens, C. *et al.* Plasma *ESR1* mutations and the treatment of estrogen receptor–positive advanced breast cancer. *J. Clin. Oncol.* **34**, 2961–2968 (2016).
10. Robinson, D.R. *et al.* Activating *ESR1* mutations in hormone-resistant metastatic breast cancer. *Nat. Genet.* **45**, 1446–1451 (2013).
11. Toy, W. *et al. ESR1* ligand-binding domain mutations in hormone-resistant breast cancer. *Nat. Genet.* **45**, 1439–1445 (2013).
12. Nguyen, V.T.M. *et al.* Differential epigenetic reprogramming in response to specific endocrine therapies promotes cholesterol biosynthesis and cellular invasion. *Nat. Commun.* **6**, 10044 (2015).
13. Curtis, C. *et al.* The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature* **486**, 346–352 (2012).
14. Beroukhim, R. *et al.* Assessing the significance of chromosomal aberrations in cancer: methodology and application to glioma. *Proc. Natl. Acad. Sci. USA* **104**, 20007–20012 (2007).
15. Gao, J. *et al.* Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci. Signal.* **6**, pl1 (2013).
16. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature* **490**, 61–70 (2012).
17. Cao, Q. *et al.* CaSNP: a database for interrogating copy number alterations of cancer genome from SNP array data. *Nucleic Acids Res.* **39**, D968–D974 (2011).
18. Koivisto, P. *et al.* Androgen receptor gene amplification: a possible molecular mechanism for androgen deprivation therapy failure in prostate cancer. *Cancer Res.* **57**, 314–319 (1997).
19. Visakorpi, T. *et al. In vivo* amplification of the androgen receptor gene and progression of human prostate cancer. *Nat. Genet.* **9**, 401–406 (1995).
20. Chen, Z., Katzenellenbogen, B.S., Katzenellenbogen, J.A. & Zhao, H. Directed evolution of human estrogen receptor variants with significantly enhanced androgen specificity and affinity. *J. Biol. Chem.* **279**, 33855–33864 (2004).
21. Dunn, C.A., Clark, W., Black, E.J. & Gillespie, D.A.F. Estrogen receptor activation function 2 (AF-2) is essential for hormone-dependent transactivation and cell transformation induced by a v-Jun DNA binding domain–estrogen receptor chimera. *Biochim. Biophys. Acta* **1628**, 147–155 (2003).
22. Park, J., Czapla, L. & Amaro, R.E. Molecular simulations of aromatase reveal new insights into the mechanism of ligand binding. *J. Chem. Inf. Model.* **53**, 2047–2056 (2013).

© 2017 Nature America, Inc., part of Springer Nature. All rights reserved.

23. Ali, S., Buluwela, L. & Coombes, R.C. Antiestrogens and their therapeutic applications in breast cancer and other diseases. *Annu. Rev. Med.* **62**, 217–232 (2011).
24. Cao, Z. *et al*. Effects of resin or charcoal treatment on fetal bovine serum and bovine calf serum. *Endocr. Res.* **34**, 101–108 (2009).
25. Li, Y., Sidore, C., Kang, H.M., Boehnke, M. & Abecasis, G.R. Low-coverage sequencing: implications for design of complex trait association studies. *Genome Res.* **21**, 940–951 (2011).
26. Shaw, L.E., Sadler, A.J., Pugazhendhi, D. & Darbre, P.D. Changes in oestrogen receptor-α and -β during progression to acquired resistance to tamoxifen and fulvestrant (Faslodex, ICI 182,780) in MCF7 human breast cancer cells. *J. Steroid Biochem. Mol. Biol.* **99**, 19–32 (2006).
27. Jeng, M.H. *et al*. Estrogen receptor expression and function in long-term estrogen-deprived human breast cancer cells. *Endocrinology* **139**, 4164–4174 (1998).

© 2017 Nature America, Inc., part of Springer Nature. All rights reserved.

© 2017 Nature America, Inc., part of Springer Nature. All rights reserved.

ONLINE METHODS

Adjuvant setting patient selection and tissue preparation. The IEO Data Quality Control Unit selected from the institutional database a set of consecutive breast cancer patients fulfilling the following criteria: (i) tumors classified as luminal A-like and luminal B-like (HER2 negative), in accordance with St. Gallen 2013 recommendations[28]; (ii) patients receiving exclusively AIs (50 patients) or TAM (50 patients) as systemic adjuvant therapy; (iii) patients with at least 1 year of follow up; (iv) patients who experienced a distant metastasis as first event after surgery and upon adjuvant therapy. Patients who initially presented with bilateral breast tumors, were receiving neoadjuvant treatments or had metastatic breast disease at the time of presentation or within 12 months after surgery were excluded. The initial data set used for this project comprised 26,495 women who had undergone surgery for a first primary breast cancer at the IEO between 1994 and 2014. All the cases prospectively entered the IEO breast cancer database and were discussed at the weekly multidisciplinary meeting. Patients had follow-up physical examinations every 6 months, annual mammography and breast ultrasounds, blood tests every 6–12 months and further evaluations when symptomatic. Data on the patients' medical history, concurrent diseases, surgery, pathological evaluation, results of staging procedures, radiotherapy, adjuvant systemic treatments, events occurring during the follow up and treatments for metastatic disease were available. Use of patients' data was approved by the ethics committee of the IEO and by the Italian Data Protection Authority. All patients provided written informed consent. All the primary tumors were fresh sampled, fixed in 4% buffered formalin and embedded in paraffin. All the metastatic biopsies were fixed in 4% buffered formalin. Detailed information regarding tumor type and grade, ER/PgR and HER2 status and Ki-67 labeling index were available in all the cases of primary and metastatic tumors. ER/PgR and HER2 immunoreactivity was assessed in line with the clinical practice procedures applicable at diagnosis. HER2 immunoreactivity was assessed using the monoclonal antibody CB11 (Novocastra, 1:800) from 1995 until 2005, and the HercepTest (Dako) thereafter. Cases classified as $HER2^+$ by immunohistochemistry were tested by FISH analysis with Vysis probes, in accordance with the ASCO/CAP guidelines. Ki-67 labeling index was assessed with the Mib-1 monoclonal antibody (Dako, 1:200) by counting at least 500 invasive tumor cells independently of their staining intensity and without focusing on hot spots. Only tumors classified as luminal A-like (ER- and PgR-positive, absence of HER2 overexpression and Ki-67 <20%) and luminal B-like (ER-positive, HER2-negative and at least one of Ki-67 ≥20% and PgR <20%) in accordance with St. Gallen recommendations were included in the study[28]. All the samples (primary tumors and paired metastatic deposits) from the patients satisfying the aforementioned criteria were reviewed at the IEO Division of Pathology (University of Milan) for assessing tumor cellularity and for tumor enrichment by macrodissection, if necessary. DNA from all the samples was extracted using commercially available kits (QIAmp DNA FFPE Tissue Kit, Qiagen), and the DNA yield was measured by Qubit Fluorometric Quantitation (Thermo Fisher Scientific). Finally, two 3-µm thick slides were cut from all the samples and put on charged slides for FISH analysis.

DNA extraction and TaqMan assay. For tumor samples, hematoxylin- and eosin-stained sections were prepared assessing the percentage of tumor cells and evaluated by a pathologist. Samples with less than 80% tumor cells were microdissected to increase the percentage of tumor cells. Matched normal DNA was extracted from nonmetastatic axillary lymph nodes or histologically non-neoplastic breast samples obtained from mammary quadrants macroscopically free of disease. Genomic DNA from FFPE tissue sections was extracted using QIAamp FFPE Tissue kit (Qiagen) according to the manufacturer's instructions. TaqMan Copy Number Assay (Applied Biosystems) for *CYP19A1* (Hs00116110_cn) and *ESR1* (Hs02488982_cn) was performed using the 7900HT Fast Real-Time PCR Systems (Applied Biosystems) according to the manufacturer's protocol. *TERT*, RNAse P (*RPPH1*), and *GARBR3* genes were analyzed as endogenous reference genes (Applied Biosystems, 4458373, 4403326 and hs_05365082_cn). Copy number (CN) for each sample was estimated using the Copy Caller Software V1.0 (Applied Biosystems) using the matched normal counterpart as reference. CN range bars indicated the minimum and maximum CN calculated for the sample replicates.

Targeted sequencing and *in silico* analysis. An AmpliSeq custom panel was designed using Ion AmpliSeq Designer 2.2 (http://www.ampliseq.com/) against the exons of *TP53*, *ESR1*, *PIK3CA*, *GATA3*, *MAP3K1* and *CYP19A1*. For the preliminary analysis, we filtered the design taking into consideration previously identified mutations (COSMIC, 141 amplicons, **Supplementary Table 2**). For the validation replicate, we then included the entire set of exons (184 amplicons). Libraries were generated from 10 ng DNA (tumor and normal) using the Ion AmpliSeq Library Kit v2.0 (Life Technologies) according to the manufacturer's instructions. Quantification of the libraries was performed using the Quant-iT dsDNA HS assay kit and a Qubit2.0 fluorimeter (Life Technologies). Templates were prepared from a pool of equimolar amounts of each library using the Ion PGM Template OT2 200 Kit with OneTouch2 system (Life Technologies). Samples were sequenced on the Ion Torrent PGM sequencer using the Ion PGM 200 Sequencing Kit v2.0 on Ion 318 chips. Data were analyzed using IonReporter and MuTect[29] to compare metastatic samples with normal DNA (normal breast samples extracted from the same patient).

Fluorescence *in situ* hybridization (FISH) analysis. FISH co-hybridization using a specific clone covering the altered locus and the specific α-satellite, as control probe for ploidy status, was performed on formalin-fixed paraffin embedded sections. Specifically, RP11-66L23 BAC clone for the *CYP19A1* locus at 15q21.2 (red signal) and 15 α-satellite probe (green signal) were used to identify *CYP19A1* gene amplification. The BAC clone was selected using the University of California Santa Cruz Genome Browser Database (http://genome.ucsc.edu/) and was tested on normal human metaphase cells to verify the absence of cross-hybridization, while the α-satellite probes were kindly provided by M. Rocchi (University of Bari, Italy). FISH experiments were performed as previously described, with minor modifications[30]. An average 30 representative nuclei scored per sample, scanning several areas to account for potential heterogeneity was counted to calculate the amplification ratio.

Single cell RNA-FISH. The protocol for adherent mammalian cell lines was optimized for Stellaris FISH probes. Hybridization was performed overnight, and no anti-fade was used for imaging. The sequence of the CAL Fluor Red 590 tagged probe targeting the *CYP19A1* mRNA is provided in the **Supplementary Note**. Samples were imaged using a Nikon Ti-E scanning laser confocal inverted microscope (A1) with 60× oil objective in tandem with Nikon NIS-Elements imaging software. Excitation was by 561.5-nm diode-pumped solid state. Detection was via 595/50-nm filter. Optical sections were captured at 0.300-µm intervals and a resolution of 256 by 256 pixels and zoom factor of 6.8, resulting in a voxel size of 0.0047 μm^3 (0.1243 µm by 0.1243 µm by 0.3 µm). Four times averaging was used to reduce photon and camera noise. An automated spot-count algorithm determined the number of mRNAs[31]. For the analysis, we included 30 positive-control cells to better define an mRNA spot.

Cell lines and hormone manipulation. Parental MCF7 breast cancer cell lines were maintained in DMEM containing 10% FCS. MCF7 cells and derivatives were authenticated using STR profiling. Cells were routinely tested for Mycoplasma contamination. The chronically estrogen-deprived MCF7-derived LTED breast cancer cell lines were maintained in phenol-red free DMEM containing 10% charcoal-stripped FCS (SFCS). Both media were supplemented with 2 mM L-glutamine and 100 units/mL penicillin 0.1 mg/mL. Estradiol 10^{-8} M (E2758 Sigma-Aldrich) was added routinely to MCF7. Both cell lines were starved for 48 h before further treatment. Subsequently, both cell lines were treated with 10^{-8} M estradiol or androstenedione (25 nM final concentration). LTED cells were also treated in the presence or absence of the aromatase inhibitor letrozole (100 nM final concentration). After 24 h, cells were lysed and RNA extracted with RNAeasy Micro Kit (Qiagen) according to the manufacturer's instructions. The quantity, quality and integrity of isolated mRNA were confirmed by absorption measurement and RNA gel electrophoresis. Then, 1 µg RNA was retrotranscribed by using iSCRIPT (Bio-Rad) containing random hexamers. Afterwards, quantitative real-time PCR was carried out using SYBR select master mix (Life Technologies), and expression of ERα targets (*TFF1*, *EGR3* and *CA12*) assessed. Values were quantified using the comparative threshold cycle method, and target gene mRNA expression was normalized to *GAPDH*. Primers are available upon request. Results are expressed as means ± s.e.m. One-way ANOVA statistical analysis was performed using GraphPad Prism, and with GraphPad Software (GraphPad Software, Inc.). Two-sided $P < 0.05$ was considered statistically significant and are expressed

as $^{*}P < 0.05$. For CYP19A1 protein quantification, we used the Abcam ab71264 antibody. siRNAs for CYP19A1 were obtained from Thermo Fisher (siSilencer Select prevalidated s3875 and s3877). siRNA was transfected at 5 nM final concentration 2 d before SRB analysis (day 0). SRB proliferation measurements were obtained after 3 further days of culture in the presence of increasing amount of letrozole. Experiments were conducted using 5 technical replicates and 3 independent biological replicates. CYP19A1-overexpressing cells were obtained by transfecting MCF7 cells with full-length *CYP19A1* (RC205890, OriGene Technologies) and selection using G418. SRB proliferation experiments were conducted as described above. For the exemestane challenge, MCF7 and LTED cells were plated in identical numbers and then treated with increased dose of exemestane (Tocris BioScience). SRB proliferation experiments were conducted as described above.

Patient-derived xenografts of ER^{+} breast cancer patients. Patient-derived xenografts were established from patients with ER^{+} metastatic breast cancer by injecting circulating and disseminated cancer cells isolated from the peripheral blood (CTC) pleural effusion fluids (BPE) or ascites (BA) into NOD.Cg-*Prkdc*^{scid}*Il2rg*^{tm1Wjl} mice, as described[32,33]. Analysis of DNA was performed from first- or second-passage xenograft tumors established in a mouse mammary fat pad. Animal care and procedures were carried out according to German legal regulations and were previously approved by the governmental review board of the federal state of Baden-Württemberg, Germany (Regierungspräsidium Karlsruhe authorization number G240/11). Human material for xenograft experiments was obtained either from patients admitted to the University Clinic Mannheim Department of Gynecology or from patients recruited at the division of Gynecologic Oncology of the Heidelberg University Hospital. The study was approved by the ethics committee of the University of Heidelberg-Mannheim (case number 2011-380N-MA). All patients gave written consent.

Survival analysis. Kaplan–Meier plots were generated using PRISM (v5). We analyzed the time that separated the date of surgery from the date of first relapse (for all patients with univocal histology numbers). Data were analyzed using a log-rank (Mantel–Cox (ratio between number of estimated DNA content vs. expected)) test. Curves were also significantly different when analyzed using a Gehan–Breslow–Wilcoxon test.

Aromatase activity assay. Aromatase activity was evaluated using a 3H-water release assay using 0.5 µmol/L of [1β-^{3}H(N)]androst-4-ene-3,17-dione as substrate[34]. The incubations were performed at 37 °C for 2 h under an air–CO_2 (5%) atmosphere. The results obtained were expressed as femtomole or picomole per hour and normalized to abundance of protein (pmol/h/mg).

CNV meta-analysis and shallow sequencing. Meta-analysis of previously published data was conducted using cBioPortal (http://www.cbioportal.org/index.do). Amplification was scored as positive for GISTIC values of >2 (Amplified). SNPs array data were interrogated using caSNP (http://cistrome.org/CaSNP/) using a CT (estimated allele number) threshold >3. Shallow-sequencing analysis was conducted using previously published data[12]. Briefly, input SAM files from resistant cell lines were used as ChIP tracks in MACS 1.4 against input tracks generated in MCF7 cells. BED and WIG files were generated using default settings[35].

ChIP-seq. ERα ChIP-seq data were re-analyzed from previously published data[12]. ERα-bound regions were clustered using CHASE and *K*-means clustering ($n = 3$)[36] (http://chase.cs.univie.ac.at/overview). RPKM plots were created for each specific cluster comparing tags from MCF7 and LTED cells. ChIP-seq data can be accessed at GEO (GEO GSE60517).

Statistical methods. Mann–Whitney's, Student's *t*-test (two-sided) and one-way or two-way ANOVA with Bonferroni's post test were used as indicated in the figure legends. Assumptions on normal distribution and equal variance were tested before statistical tests using ANOVA.

Data availability. ChIP-seq data for ERα used in this study can be accessed in GEO (GSE60517). Primary cancer data sets analyzed for CNAs in this manuscript were retrieved from cBioPortal http://www.cbioportal.org/. For breast cancer–specific mutational analyses, we used all available data sets from cBioPortal. Each study was labeled according to the publisher's instructions to indicate which cancer type and which study was analyzed via cBioportal.

28. Goldhirsch, A. *et al.* Personalizing the treatment of women with early breast cancer: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2013. *Ann. Oncol.* **24**, 2206–2223 (2013).
29. Cibulskis, K. *et al.* Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat. Biotechnol.* **31**, 213–219 (2013).
30. Pruneri, G. *et al.* The transactivating isoforms of p63 are overexpressed in high-grade follicular lymphomas independent of the occurrence of p63 gene amplification. *J. Pathol.* **206**, 337–345 (2005).
31. Kempe, H., Schwabe, A., Crémazy, F., Verschure, P.J. & Bruggeman, F.J. The volumes and transcript counts of single cells reveal concentration homeostasis and capture biological noise. *Mol. Biol. Cell* **26**, 797–804 (2015).
32. Baccelli, I. *et al.* Identification of a population of blood circulating tumor cells from breast cancer patients that initiates metastasis in a xenograft assay. *Nat. Biotechnol.* **31**, 539–544 (2013).
33. Al-Hajj, M., Wicha, M.S., Benito-Hernandez, A., Morrison, S.J. & Clarke, M.F. Prospective identification of tumorigenic breast cancer cells. *Proc. Natl. Acad. Sci. USA* **100**, 3983–3988 (2003).
34. Lephart, E.D. & Simpson, E.R. Assay of aromatase activity. *Methods Enzymol.* **206**, 477–483 (1991).
35. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
36. Younesy, H. *et al.* An interactive analysis and exploration tool for epigenomic data. *Comput. Graph. Forum* **32**, 91–100 (2013).

© 2017 Nature America, Inc., part of Springer Nature. All rights reserved.

doi:10.1038/ng.3773

Corrigendum: Acquired *CYP19A1* amplification is an early specific mechanism of aromatase inhibitor resistance in ERα metastatic breast cancer

Luca Magnani, Gianmaria Frigè, Raffaella Maria Gadaleta, Giacomo Corleone, Sonia Fabris, Mannus H Kempe, Pernette J Vershure, Iros Barozzi, Valentina Vircillo, Sung-Pil Hong, Ylenia Perone, Massimo Saini, Andreas Trumpp, Giuseppe Viale, Antonino Neri, Simak Ali, Marco Angelo Colleoni, Giancarlo Pruneri & Saverio Minucci
***Nat. Genet.*; doi:10.1038/ng.3773; corrected online 31 January 2017**

In the version of this article initially published online, the names of authors Hermannus Kempe and Pernette J. Verschure were spelled incorrectly. These errors have been corrected in the print, PDF and HTML versions of this article.

© 2017 Nature America, Inc., part of Springer Nature. All rights reserved.