Genomic approaches in breast cancer research

Henry J. Donahue and Damian C. Genetos

Advance Access publication date 20 June 2013

Abstract

Microarray technologies provide high-throughput analysis of genes that are differentially expressed in humans and other species, and thereby provide a means to measure how biological systems are altered during development or disease states. Within, we review how high-throughput genomic technologies have increased our understanding about the molecular complexity of breast cancer, identified distinct molecular phenotypes and how they can be used to increase the accuracy of predicted clinical outcome.

Keywords: *breast cancer; microarray; genomics; tumor; histology*

INTRODUCTION

Breast cancer (BCa) is the most frequently diagnosed cancer and leading cause of cancer death in women and accounted for an estimated 1 380 000 new diagnoses and 450 000 deaths in 2008. It is currently recognized as a clinically heterogeneous and complex disease, and that histological classifications do not fully correlate to clinical course and outcome [1–3].

BCa is well-documented throughout world history, with fascinating theories and various resulting clinical approaches over the course of 5 millennia. The earliest record comes from the Egyptian Imhotep *circa* 2650 BCE who envisioned no therapy for BCa patients [4,5]. Two millennia later, Hippocrates, Herodotus and Galen suggested that black bile (*melan cholos*) [6], one of the four humors, accumulated during hematosis to cause BCa [7]; surgical approaches were perceived of limited value, owing to poor results and belief in the persistent systemic nature of BCa, which would simply return due to remaining imbalances in black bile [6]. Scientific advances in the 17th and 18th centuries, such as development of the microscope and the cell theory, saw a paradigm shift in BCa from a systemic disease to a localized disease which warranted surgical intervention, and hypotheses as to the cause of BCa ranged from frequency or vigor of sexual activity [8] to drinking curdled milk [9]. In the 1900s, Halstead's radical mastectomy approach—involving removal of the breast, axillary nodes and both chest muscles—persisted as the surgical approach *de rigueur* until the 1930s and 1940s as the introduction of radio- and chemotherapy, as well as re-examining BCa as a systemic disease, suggested that alternate approaches could provide better clinical outcomes [9].

Prior to the advent of genomic analysis, BCa was classified based on histopathological appearance, tumor grade and tumor staging. Multiple strategies for grading BCa have been adopted over the last century. Currently, the Nottingham (Elston–Ellis) modification of Scarff–Bloom–Richardson grading system (NGS) [1–5,10,11] is used by the World Health Organization and other organizations to assign tumor stage and is based on the classification of primary tumor, regional nodes and metastasis (TMN). Histologic grading in general has demonstrated prognostic relevance [11,12], but problems exist with histologic grading—interobserver

Corresponding author. D.C. Genetos, Department of Anatomy, Physiology, and Cell Biology, School of Veterinary Medicine, University of California-Davis, Davis, CA, USA. Tel.: +1 530 754 0146; Fax: +1 530 753 7690; E-mail: dgenetos@ucdavis.edu

Henry J. Donahue, PhD, is a director of the Division of Musculoskeletal Sciences and Vice Chair for Research in the Department of Orthopaedics and Rehabilitation at Penn State Hershey College of Medicine in Hershey, PA, USA. His research interests include mechanosensing and mechanotransduction in the skeleton and mechanisms of BCa metastasis to bone.

Damian C. Genetos, PhD, is an assistant professor in the School of Veterinary Medicine at the University of California at Davis. His research focuses on components of and mechanisms whereby the pericellular environment influences bone cell physiology.

variation and the degree of heterogeneity from tumor to tumor and even within a tumor sample [13]—indicating that classification using these parameters is insufficient for determining individual therapy [11,14].

The sequencing of the human and mouse genome, the high throughput of genomic approaches and decreasing cost of these technologies have enhanced BCa research, from basic research to translational studies. As an example of how genomic approaches have been successfully applied to BCa research, we will review how genomic approaches have defined molecular subtypes, enhanced the predictive power of histologic grading and identified cellular signaling pathways within tumor types that could be targeted pharmacologically.

GENOMIC CLASSIFICATION OF BCA

Despite correlations between BCa classification and clinical prognosis, BCa is a heterogeneous disease that cannot be fully classified for diagnostic or therapeutic purposes based on grade, stage and histopathology alone [13]. Immunohistochemistry allowed for the separation of BCa tumors into estrogen receptor (ER) positive or negative subgroups, which enabled decision making upon whether endocrine therapy would be of benefit. Discovery of *in situ* hybridization enabled discrimination between HER2/ERBB2 amplified and non-amplified disease [15], and later, the relative risk of recurrence. Linkage analysis identified the first genes involved in BCa susceptibility, *BRCA1* [16] and *BRCA2* [17]; mutations in these genes (which encode for proteins involved in DNA damage repair [18]) are rare in the general population, but greatly increase the susceptibility to both breast and ovarian cancer [19]. Other genes implicated in susceptibility, such as TP53 or PTEN, were individually implicated in risk for BCa. Yet, such techniques allowed examination of generally one gene at a time. The introduction of genomic analysis, and the high-throughput they provide, enabled analysis of entire cohorts of BCa tumor samples.

Perou *et al.* [20] hypothesized that the phenotypic diversity observed in BCa resulted from multiple genomic differences from tumor to tumor. To test this, they examined 65 BCa tissues from 42 individuals and 17 BCa cell lines for expression of 8102 genes. Hierarchical clustering algorithms demonstrated that greatest differences in gene expression were observed between tumor samples: multiple samples from the same individual, acquired before or after chemotherapy or surgical intervention, were most closely related to each other. Those genes that showed little variance between repeated sampling of the same tumor, and which showed high variance from tumor to tumor, were termed 'intrinsic' genes and revealed four distinct molecular phenotypes of BCa—luminal-like, basal-like, normal-like and *ERBB2-/HER2*-enriched (ERB2/HER2$^+$). Later studies [21,22] using additional data sets and tumor samples validated these molecular subtypes and also indicated further subsets of the luminal-like phenotype, luminal A and luminal B subtypes. Luminal A are generally ER- or progesterone receptor (PR)-positive, HER2-negative and are associated with better disease-free and overall survival, whereas luminal B subtypes express low levels of ER/PR and have HER2 amplification, although there remains some skepticism whether luminal tumors can be reliably subclassified [15].

This early approach enabled identification of a series of genes that could be used to classify BCa into various molecular subtypes, but does so without incorporating the clinical outcome of patients who provided the tumor samples. Nonetheless, Sørlie *et al.* [21] compared subtypes with overall survival and relapse-free survival. They found significant differences in overall survival among the subtypes, with shortest overall and relapse-free survival time for those classified into basal-like and ERBB2$^+$/HER2$^+$ subtypes. Further studies demonstrated different responsiveness to chemotherapy, as basal-like and ERBB2 show greater chemosensitivity than luminal types [21,23,24].

The approaches discussed above by Perou *et al.*, Sorlie *et al.* and others demonstrated that patterns of gene expression could be used to identify distinct molecular subtypes of BCa. These approaches used unsupervised molecular class discovery methods, wherein gene expression patterns alone defined molecular subtypes and then stratified samples into molecular subtypes. Other criteria, such as relapse-free survival and clinical outcome, were not included in generating the subtypes [1,11,14]. An alternate strategy incorporates such endpoints as overall or relapse-free survival; this strategy, termed supervised classification or class prediction, assigns tumor samples from known clinical outcome (i.e. recurrence or not) to identify which gene differences or signatures (predictors) are present in one outcome (recurrence) versus another (no recurrence). As a result, class

prediction is able to more accurately predict, not correlate (as in unsupervised class discovery), clinical outcome [13,15]. This has led to the generation of prognostic gene signatures that are used in clinical practice [3,15].

GENE EXPRESSION PROFILING AND HISTOLOGIC GRADING

Genomic profiling identified distinct molecular subtypes of BCa and suggested different therapeutic approaches or clinical outcomes [22], yet if or how these subtypes related to previously defined histopathologic methods of grading tumors, resultant treatment strategy and predicted outcomes, remained inconclusive.

Histologic grading of BCa incorporates morphologic and cytologic features of tumor samples, examining the extent of tubule formation, nuclear pleomorphism and mitotic index to ultimately assign a histologic grade of 1, 2 or 3 [25]. Analysis of large patient cohorts has demonstrated that histologic grade is a highly relevant indicator of disease recurrence and patient death: G1-scored patients have a 95% survival rate at 5 years which decreases to 75% and 50% in those with scores of G2 and G3, respectively. Yet, concerns regarding inconsistent grading between multiple pathologists or institutions suggest that histologic grade does not sufficiently assist clinicians in deciding upon a therapeutic approach. Additionally, a high frequency (30–60%) of BCa are defined as histologic grade 2, which provides little prognostic information in terms of risk of recurrence [13,25]. Genomic profiling has the potential to address these concerns and discrepancies.

The potential of genomic profiling was first revealed through genetic analyses of BCa development. BCa development is thought to involve the progression from premalignant atypical ductal hyperplasia (ADH) to preinvasive ductal stage *in situ* (DCIS) to potentially lethal invasive ductal carcinoma (IDC) [26], yet the molecular mechanisms in BCa tumorigenesis remain elusive. In 2003, Ma *et al.* used laser capture microdissection, RNA amplification and cDNA microarrays to identify changes in gene expression in BCa progression. Of note, they did not observe a pattern for differential gene expression across the three pathologic stages (ADH, DCIS and IDC) as initially hypothesized, but instead found discrete differences in gene expression among the three histologic grades, with clear reciprocal demarcation between G1 gene signatures and G3 gene signatures, and G2 samples demonstrating a hybrid of histologic G1 and G3 gene signatures. These data provided the first evidence that different tumor histologic grades had distinct gene expression profiles [26].

Utilizing multiple large patient cohorts, Sotiriou *et al.* [13] undertook gene expression profiling to determine whether it could improve stratification of Elston–Ellis histologic G2 tumors into lower or higher risk tumors. They identified 97 genes, principally involved in cell cycle regulation, proliferation and differentiation, which were differentially expressed between ER-positive histologic G1 and G3 tumors. From these histologic G1 and G3 tumors, gene expression grade was defined as gene expression grade of 1 (from histologic G1 tumors) or 3 (histologic G3 tumors); this enabled generation of a gene expression grade index (GGI), which summarizes the similarity between gene expression and histologic grade. GGI strongly associated with histologic G1 and G3 tumors, but GGI from histologic G2 tumors spanned GGI from histologic G1 and G3 tumors, similar to the finding by Ma *et al.* [26], suggesting that either histologic G2 gene expression is intermediate between G1 and G3 or is a heterogeneous mixture of G1 and G3 tumors.

GGI was able to identify differences in relapse-free survival among histologic G2 tumors. Histologic G2 tumors with a gene expression profile more similar to gene expression grade 1 (i.e. gene expression profile more like that from histologic G1 tumors) had significantly longer relapse-free survival compared with histologic G2 tumors with gene expression grade 3. Ultimately, these data indicate that GGI is a better predictor of clinical outcome than histologic grading in systemically untreated patients [13]. Additionally, they were also able to identify two subtypes of ER-positive tumors, comparable with luminal A and B [20], with distinct clinical outcomes in both untreated and tamoxifen-treated patients [27].

Similar to the GGI generated by Sotiriou *et al.* [13], Ivshina *et al.* [25] sought to determine whether histologic grade 2 diagnosis could be improved by genomic analysis and better define prognosis and responsiveness to clinical treatment. Genomic class prediction used both prediction analysis of microarray and statistically weighted syndromes. Similar to Sotiriou *et al.* they observed distinct sets of gene

expression from G1 (histologic grade 1) compared with G3 samples, indicating that both prediction techniques could successfully distinguish between histologic grade on the basis of differential gene expression. There were high misclassification error rates when comparing histologic G1 to G2, and G2 to G3, suggesting, as noted by Sotiriou *et al.* that G2 tumors are not genomically distinct from G1 or G3 tumors. Instead, they were able to genetically separate histologic G2 tumors into low- (G2a) or high-grade-like (G2b) classes. Survival curve analysis of G1 and G2a gene grade demonstrated no statistical difference between the two grades and no difference between G2b and G3. Survival curves were statistically different between G2a and G2b gene grades in the absence of any therapy, including endocrine therapy, chemotherapy or endocrine and chemotherapy. Similar results were observed when examining disease recurrence, wherein gene expression grade was significantly associated with disease recurrence and also performed at predicted survival better than other grading schemes, such as lymph node status or tumor size. These data demonstrate that gene expression grade is a better prognosticator than histologic grading for disease-free and overall survival, and also better at identifying which patients should be excluded from adjuvant therapy because it would be ineffective.

GGI has also been shown to predict relapses in response to tamoxifen or aromatase inhibitors [28]. Tamoxifen is effective at prolonging disease-free and overall survival in postmenopausal women with early stage BCa [29], yet there remains a significant risk of relapse, and it is associated with adverse effects such as endometrial cancer and thromboembolism [29,30]. Aromatase inhibitors, such as letrozole, are an alternate adjuvant to tamoxifen in women and are effective as second-line treatment of metastatic BCa [31]. Desmedt *et al.* [28] examined whether GGI could predict relapses in post-menopausal women with hormone-positive BCa treated with tamoxifen or letrozole from the BIG 1-98 study, in which patients were followed for 5 years of treatment with tamoxifen, letrozole, tamoxifen followed by letrozole or letrozole followed by tamoxifen [28]. They observed no statistical difference in risk of relapse between samples with high or low GGI, and 4.8-fold greater incidence of relapse in patients who received letrozole and had a high GGI compared with those patients who received letrozole and had a low GGI. Thus, CGI has the potential to responsiveness to at least some chemotherapies.

PATHWAY-BASED PROFILING OF BCa

Genomic profiling, such as GGI, of BCa tumors used both unsupervised (i.e. classifying solely on changes in gene expression) and supervised analysis (i.e. incorporating factors such as clinical outcome) to identify gene signatures that classify the molecular subtype, associated or predicted clinical outcome and suggested therapeutic outcomes. While clearly important in understanding the molecular nature of BCa and how that may translate into a clinical outcome, these studies were limited insofar as genomic analysis does not accurately relate to activation or deactivation of signaling pathways implicated in oncogenesis. To begin to address this, Bild *et al.* used adenoviral techniques to individually activate oncogenic pathways in otherwise quiescent cells (human primary mammary epithelial cell cultures), isolated RNA and used hierarchical clustering to create gene expression signatures associated with activation of a given cell signaling pathway. Doing so enabled generation of microarray-based oncogenic pathway signatures, which were then used to identify patterns of pathway dysregulation in tumor samples as well as clinically relevant associations with disease outcome [32].

Applying this strategy, Gatza *et al.* [33] identified oncogenic and tumor suppressor pathway dysregulation in 1143 BCa tumor specimens. They identified 18 cancer subgroups (signatures; Table 1) with distinct pathway activation as well as Pearson correlation to identify clusters of pathways that were statistically found to be co-activated. A subset of these pathways recapitulated co-activation as found in previous work (ER/PR/p53 and IFNα/IFNγ [34]), but also revealed novel pathways that were co-activated such as PI3K and E2F1 with β-catenin, Akt/p63/Src and EGFR/TGF-β.

Table I: Pathways activated in BCa subgroups by Gatza *et al.* [33]

ER	β-Catenin	Myc	IFNγ	Src	TGFβ
PR	E2FI	Ras	Akt	Her2	STAT3
p53	PI3K	IFNα	p63	EGFR	TNFα

Previous genomic analysis work defined intrinsic subtypes; within this work, the authors compared their pathway-defined subgroups with intrinsic subtypes and found multiple relationships between pathway-defined subgroups and intrinsic subtypes originally identified by Perou *et al.* Intrinsic subtypes also showed distinct patterns of pathway activation: basal subgroups had low ER and PR activity but elevated Myc and Ras, whereas luminal subgroups had an inverse pattern for these pathways. Using this approach, the authors were able to further stratify intrinsic subtypes by Kaplan–Meier survival analysis. Basal-like tumors in subgroups 2 and 5 had low EGFR activity, whereas EGFR activity was higher in subgroup 8, and this correlated to differences in overall survival: the median survival in subgroup 8 was >130 months, whereas median survival in subgroup 5 was 80.6 months. Similar results observed in Luminal A subgroups 15 versus 11. The benefit of predicted pathway-based subgroupings is that it matches a rational, pharmaceutically targetable pathway with patients whose tumors have distinct biological properties, and thereby enables rationale design for a therapeutic regimen.

CONCLUDING REMARKS

Genomic approaches to studying BCa have enabled tremendous advances in the past decade since first applied to understand the complex heterogeneity of BCa, and the studies described above are but an example of how genomic approaches have been used. The studies described within have highlighted how gene expression profiling has contributed to the identification of molecular subtypes with correlation to clinical outcomes and responsiveness to adjuvant therapy, as well as expanded our understanding of the molecular basis for BCa development and progression. Numerous other approaches exist for using genomics in BCa research, including identifying: (1) how structural variation (such as copy number variation, chromosomal translocations, insertions and deletions) contributes to development and disease, gene amplification during evolution and generation of cellular diversity [35]; (2) how many genes are involved in BCa tumorigenesis [36,37] and (3) to what extent BCa evolves within a patient, in the absence or presence of treatment [38–40]. These, and future studies, have positioned genomic analysis as a requisite component of BCa research.

Key Points

- Transcriptome-wide analyses have defined distinct molecular subtypes of BCa with varying clinical outcome.
- Genomic approaches have enhanced the predictive power of histologic grading.
- Genomic analysis of BCa has used both unsupervised and supervised approaches.
- Gene expression-based predictors are used clinically to provide prognostic information and identify those who could be spared chemotherapy.

FUNDING

NIH NCI R01 CA090991 and Michael and Myrtle Baker Endowment (to H.J.D.) and NIH NIAMS R03 AR057547 (to D.C.G.).

References

1. Pusztai L, Mazouni C, Anderson K, *et al*. Molecular classification of breast cancer: limitations and potential. *Oncologist* 2006;**11**:868–77.
2. Jemal A, Bray F, Center MM, *et al*. Global cancer statistics. *CA Cancer J Clin* 2011;**61**:69–90.
3. Cavallaro S, Paratore S, de Snoo F, *et al*. Genomic analysis: toward a new approach in breast cancer management. *Crit Rev Oncol Hematol* 2012;**81**:207–23.
4. Breasted JH. New York HS *The Edwin Smith Surgical Papyrus*. Chicago, IL: The University of Chicago press, 1930.
5. Mukherjee S. *Emperor of Maladies: A Biography of Cancer*. New York: Simon & Schuster, 2010.
6. Ackerknecht EH. Historical notes on cancer. *Med Hist* 1958;**2**:114.
7. Papavramidou N, Papavramidis T, Demetriou T. Ancient Greek and Greco–Roman methods in modern surgical treatment of cancer. *Ann Surg Oncol* 2010;**17**:665.
8. MUSTACCHI P. Ramazzini and Rigoni-Stern on parity and breast cancer. Clinical impression and statistical corroboration. *Arch Intern Med* 1961;**108**:639–42.
9. Olson JS. *Bathsheba's Breast*. Baltimore, MD: Johns Hopkins University Press, 2005.
10. Elston CW, Ellis IO. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. *Histopathology* 1991;**19**:403–10.
11. Rakha E, Reis-Filho J, Baehner F, *et al*. Breast cancer prognostic classification in the molecular era: the role of histological grade. *Breast Cancer Res* 2010;**12**:207.
12. Elston CW, Gresham GA, Rao GS, *et al*. The cancer research campaign (King's/Cambridge trial for early breast cancer: clinico-pathological aspects. *Br J Cancer* 1982;**45**:655.
13. Sotiriou C, Wirapati P, Loi S, *et al*. Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis. *J Natl Cancer Inst* 2006;**98**:262–72.
14. Stevens E, Kimler BF, Davis MK, *et al*. A newly proposed semi-automated method of grading invasive lobular carcinoma: a unifying concept and correlation with prognostic

markers and patient survival. *Ann Clin Lab Sci* 2009;**39**: 25–31.

15. Reis-Filho JS, Pusztai L. Gene expression profiling in breast cancer: classification, prognostication, and prediction. *Lancet* 2011;**378**:1812–23.
16. Miki Y, Swensen J, Shattuck-Eidens D, *et al*. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science* 1994;**266**:66–71.
17. Wooster R, Bignell G, Lancaster J, *et al*. Identification of the breast cancer susceptibility gene BRCA2. *Nature* 1995;**378**: 789–92.
18. Zhang H, Tombline G, Weber BL. BRCA1, BRCA2, and DNA damage response: collision or collusion? *Cell* 1998;**92**: 433–6.
19. Antoniou A, Pharoah PDP, Narod S, *et al*. Average risks of breast and ovarian cancer associated with BRCA1 or BRCA2 mutations detected in case Series unselected for family history: a combined analysis of 22 studies. *Am J Hum Genet* 2003;**72**:1117–30.
20. Perou CM, Sørlie T, Eisen MB, *et al*. Molecular portraits of human breast tumours. *Nature* 2000;**406**:747–52.
21. Sørlie T, Perou CM, Tibshirani R, *et al*. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci USA* 2001;**98**: 10869–74.
22. Sørlie T, Tibshirani R, Parker J, *et al*. Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc Natl Acad Sci USA* 2003;**100**:8418–23.
23. Rouzier R, Perou CM, Symmans WF, *et al*. Breast cancer molecular subtypes respond differently to preoperative chemotherapy. *Clin Cancer Res* 2005;**11**:5678–85.
24. Carey LA, Perou CM, Livasy CA, *et al*. Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study. *JAMA* 2006;**295**:2492–502.
25. Ivshina AV, George J, Senko O, *et al*. Genetic reclassification of histologic grade delineates new clinical subtypes of breast cancer. *Cancer Res* 2006;**66**:10292–301.
26. Ma X-J, Salunga R, Tuggle JT, *et al*. Gene expression profiles of human breast cancer progression. *Proc Natl Acad Sci USA* 2003;**100**:5974–9.
27. Loi S, Haibe-Kains B, Desmedt C, *et al*. Definition of clinically distinct molecular subtypes in estrogen receptor-positive breast carcinomas through genomic grade. *J Clin Oncol* 2007;**25**:1239–46.
28. Desmedt C, Giobbie-Hurder A, Neven P, *et al*. The Gene expression Grade Index: a potential predictor of relapse for endocrine-treated breast cancer patients in the BIG 1–98 trial. *BMC Med Genomics* 2009;**2**:40.
29. Breast International Group (BIG) 1-98 Collaborative Group, Thürlimann B, Keshaviah A, Coates AS, *et al*. A comparison of letrozole and tamoxifen in postmenopausal women with early breast cancer. *N Engl J Med* 2005;353:2747–57.
30. Group EBCTC. Tamoxifen for early breast cancer: an overview of the randomised trials. *Lancet* 1998;**351**:1451–67.
31. Barnadas A, Estévez LG, Lluch-Hernández A, *et al*. An overview of letrozole in postmenopausal women with hormone-responsive breast cancer. *Adv Therapy* 2011;**28**: 1045–8.
32. Bild AH, Yao G, Chang JT, *et al*. Oncogenic pathway signatures in human cancers as a guide to targeted therapies. *Nature* 2006;**439**:353–7.
33. Gatza ML, Lucas JE, Barry WT, *et al*. A pathway-based classification of human breast cancer. *Proc Natl Acad Sci USA* 2010;**107**:6994–9.
34. Nielsen TO, Hsu FD, Jensen K, *et al*. Immunohistochemical and clinical characterization of the basal-like subtype of invasive breast carcinoma. *Clin Cancer Res* 2004;**10**:5367–74.
35. Gresham D, Dunham MJ, Botstein D. Comparing whole genomes using DNA microarrays. *Nat Rev Genet* 2008;**9**: 291–302.
36. Akavia UD, Litvin O, Kim J, *et al*. An integrated approach to uncover drivers of cancer. *Cell* 2010;**143**:1005–17.
37. Alderton GK. Genomics: driving cancer biology. *Nat Rev Cancer* 2011;**11**:79.
38. Ding L, Ellis MJ, Li S, *et al*. Genome remodelling in a basal-like breast cancer metastasis and xenograft. *Nature* 2010;**464**: 999–1005.
39. Nik-Zainal S, Alexandrov LB, Wedge DC, *et al*. Mutational processes molding the genomes of 21 breast cancers. *Cell* 2012;**149**:979–93.
40. Ellis MJ, Ding L, Shen D, *et al*. Whole-genome analysis informs breast cancer response to aromatase inhibition. *Nature* 2012;**486**:353–60.