Comprehensive molecular portraits of human breast tumours

The Cancer Genome Atlas Network*

We analysed primary breast cancers by genomic DNA copy number arrays, DNA methylation, exome sequencing, messenger RNA arrays, microRNA sequencing and reverse-phase protein arrays. Our ability to integrate information across platforms provided key insights into previously defined gene expression subtypes and demonstrated the existence of four main breast cancer classes when combining data from five platforms, each of which shows significant molecular heterogeneity. Somatic mutations in only three genes (TP53, PIK3CA and GATA3) occurred at >10% incidence across all breast cancers; however, there were numerous subtype–associated and novel gene mutations including the enrichment of specific mutations in GATA3, PIK3CA and MAP3K1 with the luminal A subtype. We identified two novel protein–expression–defined subgroups, possibly produced by stromal/microenvironmental elements, and integrated analyses identified specific signalling pathways dominant in each molecular subtype including a HER2/phosphorylated EGFR signature within the HER2–enriched expression subtype. Comparison of basal–like breast tumours with high–grade serous ovarian tumours showed many molecular commonalities, indicating a related aetiology and similar therapeutic opportunities. The biological finding of the four main breast cancer subtypes caused by different subsets of genetic and epigenetic abnormalities raises the hypothesis that much of the clinically observable plasticity and heterogeneity occurs within, and not across, these major biological subtypes of breast cancer.

Breast cancer is one of the most common cancers with greater than 1,300,000 cases and 450,000 deaths each year worldwide. Clinically, this heterogeneous disease is categorized into three basic therapeutic groups. The oestrogen receptor (ER) positive group is the most numerous and diverse, with several genomic tests to assist in predicting outcomes for ER+ patients receiving endocrine therapy1,2. The HER2 (also called ERBB2) amplified group3 is a great clinical success because of effective therapeutic targeting of HER2, which has led to intense efforts to characterize other DNA copy number aberrations4,5. Triple-negative breast cancers (TNBCs), lacking expression of ER, progesterone receptor (PR) and HER2, also known as basal-like breast cancers6, are a group with only chemotherapy options, and HER2/EGFR/phosphorylated EGFR signature within the HER2–enriched expression subtype. Intense efforts to characterize other DNA copy number aberrations4,5, which has led to effective therapeutic targeting of HER2, which has led to intense efforts to characterize other DNA copy number aberrations4,5.

Significantly mutated genes in breast cancer

Overall, 510 tumours from 507 patients were subjected to whole-exome sequencing, identifying 30,626 somatic mutations comprised of 28,319 point mutations, 4 dinucleotide mutations, and 2,302 insertions/deletions (indels) (ranging from 1 to 53 nucleotides). The point mutations included 6,486 silent, 19,045 missense, 1,437 nonsense, 26 read-through, 506 splice-site mutations, and 819 mutations in RNA genes. Comparison to COSMIC and OMIM databases identified 619 mutations across 177 previously reported cancer genes. Of 19,045 missense mutations, 9,484 were predicted to have a high probability of being deleterious by Condel16. The MuSiC package18, which determines the significance of the observed mutation rate of each gene based on the background mutation rate, identified 35 significantly mutated genes (excluding LOC or Ensembl gene IDs) by at least two tests (convolution and likelihood ratio tests) with false discovery rate (FDR) <5%. (Supplementary Table 2).

In addition to identifying nearly all genes previously implicated in breast cancer (PIK3CA, PTEN, AKTI, TP53, GATA3, CDHI, RB1, ML3, MAP3K1 and CDKN1B), a number of novel significantly mutated genes were included identifying TP53, RUNXI, CBFB, AFF2, PIK3R1, PTPN22, PTPRD, NF1, SF3B1 and CCND3. TP53, which is mutated in ulnar-mammary syndrome and involved in mammary gland development19, harboured 13 mutations (8 frame-shift indels,
Mutations and mRNA-expression subtype associations

We analysed the somatic mutation spectrum within the context of the four mRNA-expression subtypes, excluding the normal-like group owing to small numbers (n = 8) (Fig. 1). Several significantly mutated genes showed mRNA-subtype–specific (Supplementary Figs 1–3) and clinical-subtype–specific patterns of mutation (Supplementary Table 2). Significantly mutated genes were considerably more diverse and recurrent within luminal A and luminal B tumours than within basal-like and HER2-enriched (HER2E) subtypes; however, the overall mutation rate was lowest in luminal A subtype and highest in the basal-like and HER2E subtypes. The luminal A subtype harboured the most significantly mutated genes, with the most frequent being PIK3CA (45%), followed by MAP3K1, GATA3, TP53, CDH1 and MAP2K4. Twelve per cent of luminal A tumours contained likely inactivating mutations in MAP3K1, GATA3, TP53 and CDH1, respectively.

N, node status; T, tumour size. The right panel shows significantly mutated genes with frequent copy number amplifications (red) or deletions (blue). The average mutation rate for each gene was calculated per megabase, adjusted for coverage. The average mutation rate for each expression subtype is indicated. Hypermutated: mutation rates >3 s.d. above the mean (>4.688, indicated by grey line).

Figure 1 | Significantly mutated genes and correlations with genomic and clinical features. Tumour samples are grouped by mRNA subtype: luminal A (n = 225), luminal B (n = 126), HER2E (n = 57) and basal-like (n = 93). The left panel shows non-silent somatic mutation patterns and frequencies for significantly mutated genes. The middle panel shows clinical features: dark grey, positive or T2–4; white, negative or T1; light grey, N/A or equivocal. The far-right panel shows non-silent mutation rate per tumour (mutations per megabase, adjusted for coverage). The average mutation rate for each expression subtype is indicated. Hypermutated: mutation rates >3 s.d. above the mean (>4.688, indicated by grey line).
These data confirmed the association between the presence of germline BRCA1 mutations and basal-like breast cancers8.

**Gene expression analyses (mRNA and miRNA)**

Several approaches were used to look for structure in the mRNA expression data. We performed an unsupervised hierarchical clustering analysis of 525 tumours and 22 tumour-adjacent normal tissues using the top 3,662 variably expressed genes (Supplementary Fig. 5); SigClust analysis identified 12 classes (5 classes with >9 samples per class). We performed a semi-supervised hierarchical cluster analysis using a previously published ‘intrinsic gene list’14, which identified 13 classes (9 classes with >9 samples per class) (Supplementary Fig. 6). We also classified each sample using the 50-gene PAM50 model14 (Supplementary Fig. 5). High concordance was observed between all three analyses; therefore, we used the PAM50-defined subtype predictor as a common classification metric. There were only eight normal-like and eight claudin-low tumours27, thus we did not perform focussed analyses on these two subtypes.

MicroRNA expression levels were assayed via Illumina sequencing, using 1,222 miRBase28 v16 mature and star strands as the reference database of miRNA transcripts/genes. Seven subtypes were identified by consensus non-negative matrix factorization (NMF) clustering using an abundance matrix containing the 25% most variable miRNAs (306 transcripts/genes or MIMATs (miRNA IDs)). These subtypes correlated with mRNA subtypes, ER, PR and HER2 clinical status (Supplementary Fig. 7). Of note, miRNA groups 4 and 5 showed high overlap with the basal-like mRNA subtype and contained many TP53 mutations. The remaining miRNA groups (1–3, 6 and 7) were composed of a mixture of luminal A, luminal B and HER2E with little correlation with the PAM50 defined subtypes. With the exception of TP53—which showed a strong positive correlation—and PIK3CA and GATA3—which showed negative associations with groups 4 and 5, respectively—there was little correlation with mutation status and mRNA subtype.

**DNA methylation**

Illumina Infinium DNA methylation arrays were used to assay 802 breast tumours. Data from HumanMethylation27 (HM27) and HumanMethylation450 (HM450) arrays were combined and filtered to yield a common set of 574 probes used in an unsupervised clustering analysis, which identified five distinct DNA methylation groups (Supplementary Fig. 8). Group 3 showed a hypermethylated phenotype and was significantly enriched for luminal B mRNA subtype and under-represented for PIK3CA, MAP3K1 and MAP2K4 mutations. Group 5 showed the lowest levels of DNA methylation, overlapped with the basal-like mRNA subtype, and showed a high frequency of TP53 mutations. HER2-positive (HER2+2) clinical status, or the HER2E mRNA subtype, had only a modest association with the methylation subtypes.

A supervised analysis of the DNA methylation and mRNA expression data was performed to compare DNA methylation group 3 (N = 49) versus all tumours in groups 1, 2 and 4 (excluding group 5, which consisted predominantly of basal-like tumours). This analysis identified 4,283 genes differentially methylated (3,735 higher in group 3 tumours) and 1,899 genes differentially expressed (1,232 downregulated); 490 genes were both methylated and showed lower expression in group 3 tumours (Supplementary Table 4). A DAVID (database for annotation, visualization and integrated discovery) functional annotation analysis identified ‘extracellular region part’ and ‘Wnt signalling pathway’ to be associated with this 490-gene set; the group 3 hypermethylated samples showed fewer PIK3CA and MAP3K1 mutations, and lower expression of Wnt-pathway genes.

**DNA copy number**

A total of 773 breast tumours were assayed using Affymetrix 6.0 SNP arrays. Segmentation analysis and GISTIC were used to identify focal amplifications/deletions and arm-level gains and losses (Supplementary Table 5). These analyses confirmed all previously reported copy number variations and highlighted a number of significantly mutated genes including focal amplification of regions containing PIK3CA, EGFR, FOXA1 and HER2, as well as focal deletions of regions containing MLL3, PTEN, RB1 and MAP2K4 (Supplementary Fig. 9); in all cases, multiple genes were included within each altered region. Importantly, many of these copy number changes correlated with mRNA subtype including characteristic loss of 5q and gain of 10p in basal-like cancers6,24 and gain of 1q and/or 16q loss in luminal tumours4. NMF clustering of GISTIC segments identified five copy number clusters/groups that correlated with mRNA subtypes, ER, PR and HER2 clinical status, and TP53 mutation status (Supplementary Fig. 10). In addition, this aCGH subtype classification was highly correlated with the aCGH subtypes recently defined by ref. 30 (Supplementary Fig. 11).

**Reverse phase protein arrays**

Quantified expression of 171 cancer-related proteins and phospho-proteins by RPPA was performed on 403 breast tumours41. Unsupervised hierarchical clustering analyses identified seven subtypes; one class contained too few cases for further analysis (Supplementary Fig. 12). These protein subtypes were highly concordant with the mRNA subtypes, particularly with basal-like and HER2E mRNA subtypes. Closer examination of the HER2-containing RPPA-defined subgroup showed coordinated overexpression of HER2 and EGFR with a strong concordance with phosphorylated HER2 (pY1248) and EGFR (pY992), probably from heterodimerization and cross-phosphorylation. Although there is a potential for modest cross reactivity of antibodies against these related total and phospho-proteins, the concordance of phosphorylation of HER2 and EGFR was confirmed using multiple independent antibodies.

In RPPA-defined luminal tumours, there was high protein expression of ER, PR, AR, BCL2, GATA3 and INPP4B, defining mostly luminal A cancers and a second more heterogeneous protein subgroup composed of both luminal A and luminal B cancers. Two potentially novel protein-defined subgroups were identified: reactive I consisted primarily of a subset of luminal A tumours, whereas reactive II consisted of a mixture of mRNA subtypes. These groups are termed ‘reactive’ because many of the characteristic proteins are probably produced by the microenvironment and/or cancer-activated fibroblasts including fibronectin, caveolin 1 and collagen VI. These two RPPA groups did not have a marked difference in the percentage tumour cell content when compared to each other, or the other protein subtypes, as assessed by SNP array analysis or pathological examination. In addition, supervised analyses of reactive I versus I12 groups using mRNA expression, DNA methylation, mutation, or DNA copy number data identified no significant differences between these groups, whereas similar supervised analyses using protein and mRNA expression identified many differences.

**Multipatform subtype discovery**

To reveal higher-order structure in breast tumours based on multiple data types, significant clusters/subtypes from each of five platforms were analysed using a multipatform data matrix subjected to unsupervised consensus clustering (Fig. 2). This ‘cluster of clusters’ (C-of-C) approach illustrated that basal-like cancers had the most distinct multipatform signature as all the different platforms for the basal-like groups clustered together. To a great extent, the four major C-of-C subtypes correlated well with the previously published mRNA subtypes (driven, in part, by the fact that the four intrinsic subtypes were one of the inputs). Therefore, we also performed C-of-C analysis with no mRNA data present (Supplementary Fig. 13) or with the 12 unsupervised mRNA subtypes (Supplementary Fig. 14), and in each case 4–6 groups were identified. Recent work identified ten copy-number-based subgroups in a 997 breast cancer set30. We evaluated
A confirmatory finding here was the high mutation frequency of PIK3CA in luminal/ER+ breast cancers34,35. Through multiple technology platforms, we examined possible relationships between PIK3CA mutation, PTEN loss, INPP4B loss and multiple gene and protein expression signatures of pathway activity. RPPA data demonstrated that pAKT, pS6 and p4EBP1, typical markers of phosphatidylinositol-3-OH kinase (PI(3)K) pathway activation, were not elevated in PIK3CA-mutated luminal A cancers; instead, they were highly expressed in basal-like and HER2E mRNA subtypes (the latter having frequent PIK3CA mutations) and correlated strongly with INPP4B and PTEN loss, and to a degree with PIK3CA amplification. Similarly, protein36 and three mRNA signatures37–39 of PI(3)K pathway activation were enriched in basal-like over luminal A cancers (Fig. 3a). This apparent disconnect between the presence of PIK3CA mutations and biomarkers of pathway activation has been previously noted36.

Another striking luminal/ER+ subtype finding was the frequent mutation of MAP3K1 and MAP2K4, which represent two contiguous steps within the p38–JNK1 pathway24,40. These mutations are predicted to be inactivating, with MAP2K4 also a target of focal DNA loss in luminal tumours (Supplementary Fig. 9). To explore the possible interplay between PIK3CA, MAP3K and MAP2K4 signalling, MEMo analysis41 was performed to identify mutually exclusive alterations targeting frequently altered genes likely to belong to the same pathway (Fig. 4). Across all breast cancers, MEMo identified a set of modules that highlight the differential activation events within the receptor tyrosine kinase (RTK)–PI(3)K pathway (Fig. 4a); mutations of PIK3CA were very common in luminal/ER+ cancers whereas PTEN loss was more common in basal-like tumours. Almost all MAP3K1 and MAP2K4 mutations were in luminal tumours, yet MAP3K1 and MAP2K4 appeared almost mutually exclusive relative to one another.

The TP53 pathway was differentially inactivated in luminal/ER+ breast cancers, with a low TP53 mutation frequency in luminal A (12%) and a higher frequency in luminal B (29%) cancers (Fig. 1). In addition to TP53 itself, a number of other pathway-inactivating events occurred including ATM loss and MDM2 amplification (Figs 3b and 4b), both of which occurred more frequently within luminal B cancers. Gene expression analysis demonstrated that individual markers of functional TP53 (GADD45A and CDKN1A), and TP53 activity42,43 signatures, were highest in luminal A cancers (Fig. 3b). These data indicate that the TP53 pathway remains largely intact in luminal A cancers but is often inactivated in the more aggressive luminal B cancers44. Other PARADIGM-based pathway differences driving luminal B versus luminal A included hyperactivation of transcriptional activity associated with MYC and FOXM1 proliferation.

The critical retinoblastoma/RB1 pathway also showed mRNA-subtype-specific alterations (Fig. 3c). RB1 itself, by mRNA and protein expression, was detectable in most luminal cancers, with highest levels within luminal A. A common oncogenic event was cyclin D1 amplification and high expression, which preferentially occurred within luminal tumours, and more specifically within luminal B. In contrast, the presumed tumour suppressor CDKN2C (also called p18) was at its lowest levels in luminal A cancers, consistent with observations in mouse models45. Finally, RB1 activity signatures were also high in luminal cancers46–48. Luminal A tumours, which have the best prognosis, are the most likely to retain activity of the major tumour suppressors RB1 and TP53.

These genomic characterizations also provided clues for drugable targets. We compiled a drug target table in which we defined a target as a gene/protein for which there is an approved or investigational drug in human clinical trials targeting the molecule or canonical pathway (Supplementary Table 6). In luminal/ER+ cancers, the high frequency of PIK3CA mutations suggests that inhibitors of this activated kinase or its signalling pathway may be beneficial. Other potential significantly mutated gene drug candidates include AKT1 inhibitors (11 out of 12 AKT1 variants were luminal) and PARP inhibitors for BRCA1/BRCA2 mutations. Although still unapproved as biomarkers, many potential copy-number-based drug targets
HER2-amplicon-associated genes that in part define the HER2E mRNA subtype (Supplementary Fig. 5). However, not all clinically HER2+ tumours are of the HER2E mRNA subtype, and not all tumours in the HER2E mRNA subtype are clinically HER2+.

Integrated analysis of the RPPA and mRNA data clearly identified a HER2+ group (Supplementary Fig. 12). When the HER2+ protein and HER2E mRNA subtypes overlapped, a strong signal of EGFR, pEGFR, HER2, and pHER2 was observed. However, only ~50% of clinically HER2+ tumours fall into this HER2E-mRNA-subtype/HER2-protein group, the rest of the clinically HER2+ tumours were observed predominantly in the luminal mRNA subtypes.

These data indicate that there exist at least two types of clinically defined HER2+ tumours. To identify differences between these groups, a supervised gene expression analysis comparing 36 HER2E-mRNA-subtype/HER2+ versus 31 luminal-mRNA-subtype/HER2- tumours was performed and identified 302 differentially expressed genes (q-value = 0%) (Supplementary Fig. 18 and Supplementary Table 7). These genes largely track with ER status but also indicated that HER2E-mRNA-subtype/HER2+ tumours showed significantly higher expression of a number of RTKs including FGFR4, EGFR, HER2 itself, as well as genes within the HER2 amplicon (including GRB7). Conversely, the luminal-mRNA-subtype/HER2- tumours showed higher expression of the luminal cluster of genes including GATA3, BCL2 and ESR1. Further support for two types of clinically defined HER2+ disease was evident in the somatic mutation data supervised by either mRNA subtype or ER status; TP53 mutations were significantly enriched in HER2E or ER-negative tumours whereas GATA3 mutations were only observed in luminal subtypes or ER+ tumours.

Analysis of the RPPA data according to mRNA subtype identified 36 differentially expressed proteins (q-value <5%) (Supplementary Fig. 18G and Supplementary Table 8). The EGFR/pEGFR/HER2/phER2 signal was again observed and present within the HER2E-mRNA-subtype/HER2+ tumours, as high pS6 and pS6; conversely, many protein markers of luminal cancers again distinguished the luminal-mRNA-subtype/HER2- tumours. Given the importance of clinical HER2 status, a more focused analysis was performed based on the RPPA-defined protein expression of HER2 (Supplementary Fig. 19)—the results strongly recapitulated findings from the RPPA and mRNA subtypes including a high correlation between HER2 clinical status, HER2 protein by RPPA, pHER2, EGFR and pEGFR. These multiple signatures, namely HER2E mRNA subtype, HER2 amplicon genes by mRNA expression, and RPPA EGFR/pEGFR/HER2/phER2 signature, ultimately identify at least two groups/subtypes within clinically HER2+ tumours (Table 1). These signatures represent breast cancer biomarker(s) that could potentially predict response to anti-HER2 targeted therapies.

Many therapeutic advances have been made for clinically HER2+ disease. This study has identified additional somatic mutations that represent potential therapeutic targets within this group, including a high frequency of PIK3CA mutations (39%), a lower frequency of PTEN and PIK3R1 mutations (Supplementary Table 6), and genomic losses of PTEN and INPP4B. Other possible druggable mutations included variants within HER family members including two somatic mutations in HER2, two within EGFR, and five within HER3. Pertuzumab, in combination with trastuzumab, targets the HER2–HER3 heterodimer43; however, these data suggest that targeting EGFR with HER2 could also be beneficial. Finally, the HER2E mRNA subtype typically showed high aneuploidy, the highest somatic mutation rate (Table 1), and DNA amplification of other potential therapeutic targets including FGFRs, EGFR, CDK4 and cyclin D1.

**Basal-like summary analysis**

The basal-like subtype was discovered more than a decade ago by first-generation cDNA microarrays43. These tumours are often referred to as triple-negative breast cancers (TNBCs) because most basal-like tumours are typically negative for ER, PR and HER2.
Cancers. Even though chromosome 8q24 is amplified across all subtypes (Supplementary Fig. 5), basal-like tumours alone. The top-scoring module included ATM mutations, defects at BRCA1 and BRCA2, and deregulation of the RB1 pathway. A gene expression heat map is below the fingerprint to show expression levels.

Expression features of basal-like tumours include a characteristic signature containing keratins 5, 6 and 17 and high expression of genes associated with cell proliferation (Supplementary Fig. 5). A PARADIGM analysis of basal-like versus luminal tumours emphasized the importance of hyperactivated FOXM1 as a transcriptional driver of this enhanced proliferation signature (Supplementary Fig. 17). PARADIGM also identified hyperactivated MYC and HIF1-α/ARNT network hubs as key regulatory features of basal-like cancers. Even though chromosome 8q24 is amplified across all subtypes (Supplementary Fig. 9), high MYC activation seems to be a basal-like characteristic.

Given the striking contrasts between basal-like and luminal/HER2E subtypes, we performed a MEMo analysis on basal-like tumours alone. The top-scoring module included ATM mutations, BRCA1 and BRCA2 inactivation, RB1 loss and cyclin E1 amplification (Fig. 4c). Notably, these same modules were identified previously for serous ovarian cancers. Furthermore, the basal-like (and TNBC) mutation spectrum was reminiscent of the spectrum seen in serous ovarian cancers with only one gene (that is, TP53) at >10% mutation frequency. To explore possible similarities between serous ovarian and the breast basal-like cancers, we performed a number of analyses comparing ovarian versus breast luminal, ovarian versus breast basal-like, and breast basal-like versus breast luminal cancers.

However, ~75% of TNBCs are basal-like with the other 25% comprised of all other mRNA subtypes. In this data set, there was a high degree of overlap between these two distinction with 76 TNBCs, 81 basal-like, and 65 that were both TNBCs and basal-like. Given the degree of overlap between these two distinctions with 76 TNBCs, 81 basal-like, and 65 that were both TNBCs and basal-like. Given the striking contrasts between basal-like and luminal/HER2E subtypes, we performed a MEMo analysis on basal-like tumours alone. The top-scoring module included ATM mutations, BRCA1 and BRCA2 inactivation, RB1 loss and cyclin E1 amplification (Fig. 4c). Notably, these same modules were identified previously for serous ovarian cancers. Furthermore, the basal-like (and TNBC) mutation spectrum was reminiscent of the spectrum seen in serous ovarian cancers with only one gene (that is, TP53) at >10% mutation frequency. To explore possible similarities between serous ovarian and the breast basal-like cancers, we performed a number of analyses comparing ovarian versus breast luminal, ovarian versus breast basal-like, and breast basal-like versus breast luminal cancers.

Expression features of basal-like tumours include a characteristic signature containing keratins 5, 6 and 17 and high expression of genes associated with cell proliferation (Supplementary Fig. 5). A PARADIGM analysis of basal-like versus luminal tumours emphasized the importance of hyperactivated FOXM1 as a transcriptional driver of this enhanced proliferation signature (Supplementary Fig. 17). PARADIGM also identified hyperactivated MYC and HIF1-α/ARNT network hubs as key regulatory features of basal-like cancers. Even though chromosome 8q24 is amplified across all subtypes (Supplementary Fig. 9), high MYC activation seems to be a basal-like characteristic.

Given the striking contrasts between basal-like and luminal/HER2E subtypes, we performed a MEMo analysis on basal-like tumours alone. The top-scoring module included ATM mutations, BRCA1 and BRCA2 inactivation, RB1 loss and cyclin E1 amplification (Fig. 4c). Notably, these same modules were identified previously for serous ovarian cancers. Furthermore, the basal-like (and TNBC) mutation spectrum was reminiscent of the spectrum seen in serous ovarian cancers with only one gene (that is, TP53) at >10% mutation frequency. To explore possible similarities between serous ovarian and the breast basal-like cancers, we performed a number of analyses comparing ovarian versus breast luminal, ovarian versus breast basal-like, and breast basal-like versus breast luminal cancers.
(percentages are based on 466 tumor overlap list. Amp, amplification; mut, mutation.)

## Table 1: Highlights of genomic, clinical and proteomic features of subtypes

| Subtype               | Luminal A                      | Luminal B                      | Basal-like                      | HER2   |
|-----------------------|--------------------------------|--------------------------------|--------------------------------|--------|
| ER” / HER2” (%)       | 87                             | 82                             | 10                             | 20     |
| HER2” (%)             | 7                              | 15                             | 2                              | 68     |
| TNBCs (%)             | 2                              | 1                              | 80                             | 9      |
| TP53 pathway          | TP53 mut (12%); gain of MDM2 (14%) | TP53 mut (32%); gain of MDM2 (31%) | TP53 mut (84%); gain of MDM2 (14%) | TP53 mut (75%); gain of MDM2 (30%) |
| PIK3CA/PTEN pathway   | PIK3CA mut (49%); PTEN mut/loss (13%); INPP4B loss (9%) | PIK3CA mut (32%); PTEN mut/loss (24%); INPP4B loss (16%) | PIK3CA mut (7%); PTEN mut/loss (35%); INPP4B loss (30%) | PIK3CA mut (42%); PTEN mut/loss (19%); INPP4B loss (30%) |
| RB1 pathway           | Cyclin D1 amp (29%); CDK4 gain (14%); low expression of CDKN2A; high expression of RB1 | Cyclin D1 amp (58%); CDK4 gain (25%) | RB1 mut/loss (20%); cyclin E1 amp (9%); high expression of CDKN2A; low expression of RB1 | Cyclin D1 amp (38%); CDK4 gain (24%) |
| mRNA expression       | High ER cluster; high proliferation | Lower ER cluster; high proliferation | Basal signature; high proliferation | HER2 ampiclon signature; high proliferation |
| Copy number           | Most diploid; many with quiet genomes; 1q, 8q, 8p11 gain; 8p; 16q loss; 11q13.3 amp (24%) | Most aneuploid; many with focal amp; 1q, 8q, 8p11 gain; 8p; 16q loss; 11q13.3 amp (51%); 8p11.23 amp (28%) | Most aneuploid; high genomic instability; 1q, 10p gain; 8p, 5q loss; MCF7 gain (40%) | Most aneuploid; high genomic instability; 1q, 8q gain; 8p loss; 17q12 focal ERRB2 amp (71%) |
| DNA mutations         | PIK3CA (49%); TP53 (12%); GATA3 (14%); MAP3K1 (14%) | TP53 (32%); PIK3CA (32%); MAP3K1 (5%) | TP53 (84%); PIK3CA (7%) | TP53 (75%); PIK3CA (42%); PIK3R1 (8%) |
| DNA methylation       | –                              | Hypermethylated phenotype for subset | –                              | –      |
| Protein expression    | High oestrogen signalling; high MYB; RPPA reactive subtypes | Less oestrogen signalling; high FOXM1 and MYC; RPPA reactive subtypes | High expression of DNA repair proteins; PTEN and INPP4B loss signature (pAKT) | High protein and phospho-protein expression of EGFR and HER2 |

## Figure 5 | Comparison of breast and serum ovarian carcinomas.

(a) Significantly enriched genomic alterations identified by comparing basal-like or serous ovarian tumours to luminal cancers. b, Inter-sample correlations (yellow, positive) between gene transcription profiles of breast tumours (columns; TCGA data, arranged by subtype) and profiles of cancers from various tissues of origin (rows; external ‘TGEN expO’ data set, GSE2109) including ovarian cancers.

Supplementary Fig. 20A, B; in both cases, basal-like tumours were the most similar to the serous ovarian carcinomas.

We systematically looked for other common features between serous ovarian and basal-like tumours when each was compared to luminal. We identified: (1) BRCA1 inactivation; (2) RB1 loss and cyclin E1 amplification; (3) high expression of AKT3; (4) MYC amplification and high expression; and (5) a high frequency of TP53 mutations (Fig. 5a). An additional supervised analysis of large, external multitumour type transcriptomic data set (Gene Expression Omnibus accession GSE2109) was performed where each TCGA (The Cancer Genome Atlas) breast tumour expression profile was compared via a correlation analysis to that of each tumour in the multitumour set. Basal-like breast cancers clearly showed high mRNA expression correlations with serous ovarian cancers, as well as with lung squamous carcinomas (Fig. 5b). A PARADIGM analysis that calculates whether a gene or pathway feature is both differentially activated in basal-like versus luminal cancers and has higher overall activity across the TCGA ovarian samples was performed; this identified comparably high pathway activity of the HIF1-α/ARNT, MYC and FOXM1 regulatory hubs in both ovarian and basal-like cancers (Supplementary Fig. 20C). The common findings of TP53, RB1 and BRCA1 loss, with MYC amplification, strongly suggest that these are shared driving events for basal-like and serous ovarian carcinogenesis. This suggests that common therapeutic approaches should be considered, which is supported by the activity of platinum analogues and taxanes in breast basal-like and serous ovarian cancers.

Given that most basal-like cancers are TNBCs, finding new drug targets for this group is critical. Unfortunately, the somatic mutation repertoire for basal-like breast cancers has not provided a common target aside from BRCA1 and BRCA2. Here we note that ~20% of basal-like tumours had a germline (n = 12) and/or somatic (n = 8) BRCA1 or BRCA2 variant, which suggests that one in five basal-like patients might benefit from PARP inhibitors and/or platinum compound.

The copy number landscape of basal-like cancers showed widespread genomic instability and common gains of 1q, 8q, 8p11 gain; 8p, 16q loss; 11q13.3 amp (24%) (Fig. 5). Comparing copy number landscapes, we observed several shared driving events for basal-like and serous ovarian carcinogenesis. Inactivation of RB1 and the shared RB1 pathway (Supplementary Fig. 20A, B) in both cases, basal-like tumours were the most similar to the serous ovarian carcinomas.

Given that most basal-like cancers are TNBCs, finding new drug targets for this group is critical. Unfortunately, the somatic mutation repertoire for basal-like breast cancers has not provided a common target aside from BRCA1 and BRCA2. Here we note that ~20% of basal-like tumours had a germline (n = 12) and/or somatic (n = 8) BRCA1 or BRCA2 variant, which suggests that one in five basal-like patients might benefit from PARP inhibitors and/or platinum compound. The copy number landscape of basal-like cancers showed widespread genomic instability and common gains of 1q, 8q, 8p11 gain; 8p, 16q loss; 11q13.3 amp (24%) (Fig. 5). Comparing copy number landscapes, we observed several shared driving events for basal-like and serous ovarian carcinogenesis. Inactivation of RB1 and the shared RB1 pathway (Supplementary Fig. 20A, B) in both cases, basal-like tumours were the most similar to the serous ovarian carcinomas.

Given that most basal-like cancers are TNBCs, finding new drug targets for this group is critical. Unfortunately, the somatic mutation repertoire for basal-like breast cancers has not provided a common target aside from BRCA1 and BRCA2. Here we note that ~20% of basal-like tumours had a germline (n = 12) and/or somatic (n = 8) BRCA1 or BRCA2 variant, which suggests that one in five basal-like patients might benefit from PARP inhibitors and/or platinum compound. The copy number landscape of basal-like cancers showed widespread genomic instability and common gains of 1q, 8q, 8p11 gain; 8p, 16q loss; 11q13.3 amp (24%) (Fig. 5). Comparing copy number landscapes, we observed several shared driving events for basal-like and serous ovarian carcinogenesis. Inactivation of RB1 and the shared RB1 pathway (Supplementary Fig. 20A, B) in both cases, basal-like tumours were the most similar to the serous ovarian carcinomas.

Given that most basal-like cancers are TNBCs, finding new drug targets for this group is critical. Unfortunately, the somatic mutation repertoire for basal-like breast cancers has not provided a common target aside from BRCA1 and BRCA2. Here we note that ~20% of basal-like tumours had a germline (n = 12) and/or somatic (n = 8) BRCA1 or BRCA2 variant, which suggests that one in five basal-like patients might benefit from PARP inhibitors and/or platinum compound. The copy number landscape of basal-like cancers showed widespread genomic instability and common gains of 1q, 8q, 8p11 gain; 8p, 16q loss; 11q13.3 amp (24%) (Fig. 5). Comparing copy number landscapes, we observed several shared driving events for basal-like and serous ovarian carcinogenesis. Inactivation of RB1 and the shared RB1 pathway (Supplementary Fig. 20A, B) in both cases, basal-like tumours were the most similar to the serous ovarian carcinomas.
some basal-like cancers include FGFIR1, FGFIR2, IGF1R, KIT, MET and PDGFRα. Finally, the PARADIGM identification of high HIF1-α/ARNT pathway activity suggests that these malignancies might be susceptible to angiogenesis inhibitors and/or bioreductive drugs that become activated under hypoxic conditions.

Concluding remarks

The integrated molecular analyses of breast carcinomas that we report here significantly extends our knowledge base to produce a comprehensive catalogue of likely genomic drivers of the most common breast cancer subtypes (Table 1). Our novel observation that diverse genetic and epigenetic alterations converge phenotypically into four main breast cancer classes is not only consistent with convergent evolution of gene circuits, as seen across multiple organisms, but also with models of breast cancer clonal expansion and in vivo cell selection proposed to explain the phenotypic heterogeneity observed within defined breast cancer subtypes.

METHODS SUMMARY

Specimens were obtained from patients with appropriate consent from institutional review boards. Using a co-isolation protocol, DNA and RNA were purified. In total, 800 patients were assayed on at least one platform. Different numbers of patients were used for each platform using the largest number of patients available at the time of data freeze; 406 samples (463 patients) were in common across 5 out of 6 platforms. Technologies used include: (1) gene expression DNA microarrays; (2) reverse phase protein arrays; (3) exome sequencing; and (6) reverse phase protein arrays. Each platform, except for the exome sequencing, was used in a complementary manner.

Received 22 March; accepted 11 July 2012.

Published online 23 September; corrected online 3 October 2012.

1. Paik, S. et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. N. Engl. J. Med. 351, 2817–2826 (2004).

2. van’t Veer, L. J. et al. Gene expression profiling predicts clinical outcome of breast cancer. Nature 415, 530–536 (2002).

3. Slamon, D. J. et al. Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. Science 235, 177–182 (1987).

4. Chin, K. et al. Genomic and transcriptional architecture of 2,000 breast tumors reveals novel subgroups. Nature 466, 445–452 (2010).

5. Haney, B. T. et al. A technical assessment of the utility of reverse phase protein arrays for the study of the functional proteome in non-small-cell lung cancers. Clin. Proteomics 6, 129–151 (2010).

6. Daub, H. et al. Kinase-selective enrichment enables quantitative phosphoproteomics of the kinome across the cell cycle. Mol. Cell 31, 438–448 (2008).

7. Vaske, C. J. et al. Inference of patient-specific pathway activities from multidimensional cancer genomics data using PARADIGM. Bioinformatics 26, 237–245 (2010).

8. Campbell, I. G. et al. Mutation of the PIK3CA gene in ovarian and breast cancer. Cancer Res. 64, 7678–7681 (2004).

9. Bachman, K. E. et al. The PIK3CA gene is mutated with high frequency in inflammatory breast cancers. Cancer Biol. Ther. 3, 772–775 (2004).

10. Stempke-Hale, K. et al. An integrative genomic and proteomic analysis of PIK3CA, PTEN, and AKT mutations in breast cancer. Cancer Res. 68, 6084–6091 (2008).

11. Creighton, C. J. et al. Proteomic and transcriptomic profiling reveals a link between the PI3K pathway and lower estrogen-receptor (ER) levels and activity in ER‘ breast cancer. Breast Cancer Res. 12, R40 (2010).

12. Majumder, P. K. et al. mTOR inhibition reverses Akt-dependent prostate intraepithelial neoplasia through regulation of apoptotic and HIF-1-dependent pathways. Nature Med. 10, 594–601 (2004).

13. Saal, L. H. et al. Recurrent gross mutations of the PTEN tumor suppressor gene in breast cancers with deficient DSB repair. Nature Genet. 40, 102–107 (2008).

14. Wagner, E. F. & Nebreda, A. R. Signal integration by JNK and p38 MAPK pathways. Annu. Rev. Genomics Hum. Genet. 9, 398–406 (2008).

15. Kannan, K. et al. DNA microarray identification of primary and secondary target genes regulated by p53. Oncogene 20, 2225–2234 (2001).

16. Troester, M. A. et al. Gene expression patterns associated with p53 status in breast cancer. BMC Cancer 6, 276 (2006).

17. Deinonen, C. et al. DKK inhibitor p18Kix is a downstream target of GATA3 and restrains mammary luminal progenitor cell proliferation and tumorigenesis. Cancer Cell 15, 389–401 (2009).

18. Chicas, A. et al. Dissecting the unique role of the retinoblastoma tumor suppressor during cellular senescence. Cancer Cell 17, 376–387 (2010).

19. Happle, R., Meisinger, C. & Stoltenberg, J. The functional loss of the retinoblastoma tumour suppressor is a common event in basal-like and luminal B breast carcinomas. Breast Cancer Res. 10, R75 (2008).

20. Liberg, A. et al. PTEN, and AKT mutations in breast cancer. Clin. Cancer Res. 6, 1080–1087 (2000).

21. Lucas, D. P. et al. PTEN amplification is a novel mechanism of PTEN gene inactivation. Cancer Res. 59, 529–541 (1999).

22. Bergamaschi, A. et al. Distinct patterns of DNA copy number alteration are associated with different clinopathological features and gene-expression subtypes of breast cancer. Genes Chromosom. Cancer 45, 1033–1040 (2006).

23. Perou, C. M. Molecular stratification of triple-negative breast cancers. Oncologist 16 (suppl. 1), 61–70 (2011).

24. Sorlie, T. et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. Proc. Natl Acad. Sci. USA 100, 8418–8423 (2003).

25. Wei, M. et al. Germline BRCA1 mutations and a basal epithelial phenotype in breast cancer. J. Natl Cancer Inst. 95, 1482–1485 (2003).

26. Carey, L. A. et al. Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study. J. Am. Med. Assoc. 295, 2492–2502 (2006).

27. Li, Q. Y. et al. The Brachyury (T) gene family. Hum. Genet. 128, 529–541 (2010).

28. Nelson, P. W. et al. Gene expression patterns associated with p53 status in breast cancer. BMC Cancer 6, 276 (2006).

29. Baserga, J. et al. Pertuzumab plus trastuzumab plus docetaxel for metastatic breast cancer. N. Engl. J. Med. 366, 109–119 (2012).

30. Jiang, Z. et al. RB deletion in mouse mammary progenitors induces luminal-B or basal-like/EMT tumor subtypes depending on p53 status. J. Clin. Invest. 120, 3296–3309 (2010).

31. Chandriani, S. et al. A core MYC gene expression signature is prominent in basal-like breast cancer but only partially overlaps the core response. PLoS ONE 4, e6693 (2009).

32. The Cancer Genome Atlas Research Network. Integrated genomic analyses of ovarian carcinoma. Nature 474, 609–615 (2011).

33. Auduch, M. W. et al. Oral poly(ADP-ribose) polymerase inhibitor olaparib in patients with BRCA1 or BRCA2 mutations and recurrent ovarian cancer: a proof-of-concept trial. Lancet 376, 245–251 (2010).

34. Fong, P. C. et al. Inhibition of poly(ADP-ribose) polymerase in tumors from BRCA mutation carriers. N. Engl. J. Med. 361, 123–134 (2009).

35. Cheung, L. W. et al. High frequency of PIK3R1 and PIK3R2 mutations in endometrial cancer elucidates a novel mechanism for regulation of PTEN protein stability. Cancer Discov. 1, 170–185 (2011).

36. Stemke-Hale, K. et al. Mutations of SF3B1 and other novel cancer genes in myelodysplastic syndromes and myelodysplastic/myeloproliferative neoplasms. Blood 118, 6239–6246 (2011).

37. Wang, L. et al. SF3B1 and other novel cancer genes in chronic lymphocytic leukemia. N. Engl. J. Med. 356, 2497–2506 (2011).

38. Ding, L. et al. Somatic mutations affect key pathways in lung adenocarcinoma. Nature 455, 1069–1075 (2008).

39. Johnson, G. L. & Lapadat, R. Mitogen-activated protein kinase pathway mediated by ERK, JNK, and p38 protein kinases. Science 298, 1911–1912 (2002).

40. Wagner, E. F. & Nebreda, A. R. Signal integration by JNK and p38 MAPK pathways. Annu. Rev. Genomics Hum. Genet. 9, 398–406 (2008).

41. Walsh, T. et al. Detection of inherited mutations for breast and ovarian cancer using genomic capture and massively parallel sequencing. Proc. Natl Acad. Sci. USA 107, 12529–12533 (2010).

42. Prat, A. et al. Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. Breast Cancer Res. 12, R68 (2010).
55. Fedele, C. G. et al. Inositol polyphosphate 4-phosphatase II regulates PI3K/Akt signaling and is lost in human basal-like breast cancers. Proc. Natl Acad. Sci. USA 107, 22231–22236 (2010).

56. Gewinner, C. et al. Evidence that inositol polyphosphate 4-phosphatase type II is a tumor suppressor that inhibits PI3K signaling. Cancer Cell 16, 115–125 (2009).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank M. Sheth and S. Lucas for administrative coordination of TCGA activities, and C. Gunter for critical reading of the manuscript. This work was supported by the following grants from the USA National Institutes of Health: U24CA143883, U24CA143858, U24CA143840, U24CA147999, U24CA143835, U24CA143845, U24CA143882, U24CA143867, U24CA143866, U24CA143848, U24CA144025, U54HG03079, P50CA16201 and P50CA58223. Additional support was provided by the Susan G. Komen for the Cure, the US Department of Defense through the Henry M. Jackson Foundation for the Advancement of Military Medicine, and the Breast Cancer Research Foundation. The views expressed in this paper are those of the authors and do not reflect the official policy of the Department of Defense, or US Government.

Author Contributions TCGA research network contributed collectively to this study. Biospecimens were provided by tissue source sites and processed by a Biospecimens Core Resource. Data generation and analyses were performed by genome sequencing centres, cancer genome characterization centres, and genome data analysis centres. RPPA analysis was performed at the MD Anderson Cancer Center in association with the genome data analysis centre. All data were released through the Data Coordinating Center. Project activities were coordinated by NCI and NCI/HRG project teams. We also acknowledge the following TCGA investigators of the Breast Analysis Working Group who contributed substantially to the analysis and writing of this manuscript: Project leaders, C.M.P., M.J.E.; manuscript coordinator, C.M.P., K.A.H.; data coordinator, K.A.H.; authors, C.M.P., M.J.E.; manuscript leaders, C.M.P., M.J.E.; manuscript coordinator, C.M.P., K.A.H.; analysis coordinator, C.M.P., K.A.H.; DNA sequence analysis, D.C.K., L.D.; microRNA analysis, K.C.H.; mRNA sequence analysis, A.D.C., M.M.; histone analysis, G.B.; microRNA sequence analysis, A.G.R.; DNA methylation analysis, S. Malik, M. Sarker, P.W.L.; copy number analysis; A.D.C., M.M.; manuscript leaders, C.M.P., M.J.E.; manuscript coordinator, C.M.P., K.A.H.; data coordinator, K.A.H.; authors, C.M.P., M.J.E.; manuscript leaders, C.M.P., M.J.E.; manuscript coordinator, C.M.P., K.A.H.; analysis coordinator, C.M.P., K.A.H.; DNA sequence analysis, D.C.K., L.D.; microRNA analysis, K.C.H.; mRNA sequence analysis, A.D.C., M.M.; histone analysis, G.B.; microRNA sequence analysis, A.G.R.; DNA methylation analysis, S. Malik, M. Sarker, P.W.L.; copy number analysis; A.D.C., M.M.; protein analysis, W.L., R.G.W., G.B.M.; pathway/integrated analysis, C.J.C., C.Y., J.M.S., C.C.B., G.C., S.S., R.S.; biospecimen core resource, T.L.B., J.M.G.; pathology and clinical expertise, T.A.K., H.H., R.J.M., J.N.I., T.S., F.W.

Author Information All of the primary sequence files are deposited in CCHub (https://cchub.unc.edu/); all other data including mutation annotation file are deposited at the Data Coordinating Center (http://cancergenomics.ucsf.edu/); Sample list version of the paper and supporting data can be found at https://tcga-data.nci.nih.gov/docs/publications/brc2a_2012/. The data can be explored via the ISB Regulome Explorer (http://explorer.cancerregulome.org/) and the cBio Cancer Genomics Portal (http://cbioportal.org). Data donations can be found at https://wiki.nci.nih.gov/display/brca_2012/. The data can be explored via the ISB Regulome Explorer (http://explorer.cancerregulome.org/) and the cBio Cancer Genomics Portal (http://cbioportal.org). Data donations can be found at http://tcga-data.nci.nih.gov/docs/publications/brc2a_2012/. The data can be explored via the ISB Regulome Explorer (http://explorer.cancerregulome.org/) and the cBio Cancer Genomics Portal (http://cbioportal.org). Data donations can be found at https://wiki.nci.nih.gov/display/brca_2012/. The data can be explored via the ISB Regulome Explorer (http://explorer.cancerregulome.org/) and the cBio Cancer Genomics Portal (http://cbioportal.org). Data donations can be found at http://tcga-data.nci.nih.gov/docs/publications/brc2a_2012/.

Acknowledgements

Author Contributions

TCGA research network contributed collectively to this study. Biospecimens were provided by tissue source sites and processed by a Biospecimens Core Resource. Data generation and analyses were performed by genome sequencing centres, cancer genome characterization centres, and genome data analysis centres. RPPA analysis was performed at the MD Anderson Cancer Center in association with the genome data analysis centre. All data were released through the Data Coordinating Center. Project activities were coordinated by NCI and NCI/HRG project teams. We also acknowledge the following TCGA investigators of the Breast Analysis Working Group who contributed substantially to the analysis and writing of this manuscript: Project leaders, C.M.P., M.J.E.; manuscript coordinator, C.M.P., K.A.H.; analysis coordinator, C.M.P., K.A.H.; DNA sequence analysis, D.C.K., L.D.; microRNA analysis, K.C.H.; mRNA sequence analysis, A.D.C., M.M.; histone analysis, G.B.; microRNA sequence analysis, A.G.R.; DNA methylation analysis, S. Malik, M. Sarker, P.W.L.; copy number analysis; A.D.C., M.M.; protein analysis, W.L., R.G.W., G.B.M.; pathway/integrated analysis, C.J.C., C.Y., J.M.S., C.C.B., G.C., S.S., R.S.; biospecimen core resource, T.L.B., J.M.G.; pathology and clinical expertise, T.A.K., H.H., R.J.M., J.N.I., T.S., F.W.

Author Information

All of the primary sequence files are deposited in CCHub (https://cchub.unc.edu/); all other data including mutation annotation file are deposited at the Data Coordinating Center (http://cancergenomics.ucsf.edu/); Sample list version of the paper and supporting data can be found at http://tcga-data.nci.nih.gov/docs/publications/brc2a_2012/. The data can be explored via the ISB Regulome Explorer (http://explorer.cancerregulome.org/) and the cBio Cancer Genomics Portal (http://cbioportal.org). Data donations can be found at https://wiki.nci.nih.gov/display/brca_2012/. The data can be explored via the ISB Regulome Explorer (http://explorer.cancerregulome.org/) and the cBio Cancer Genomics Portal (http://cbioportal.org). Data donations can be found at https://wiki.nci.nih.gov/display/brca_2012/. The data can be explored via the ISB Regulome Explorer (http://explorer.cancerregulome.org/) and the cBio Cancer Genomics Portal (http://cbioportal.org). Data donations can be found at https://wiki.nci.nih.gov/display/brca_2012/.

Acknowledgements

Author Contributions

TCGA research network contributed collectively to this study. Biospecimens were provided by tissue source sites and processed by a Biospecimens Core Resource. Data generation and analyses were performed by genome sequencing centres, cancer genome characterization centres, and genome data analysis centres. RPPA analysis was performed at the MD Anderson Cancer Center in association with the genome data analysis centre. All data were released through the Data Coordinating Center. Project activities were coordinated by NCI and NCI/HRG project teams. We also acknowledge the following TCGA investigators of the Breast Analysis Working Group who contributed substantially to the analysis and writing of this manuscript: Project leaders, C.M.P., M.J.E.; manuscript coordinator, C.M.P., K.A.H.; analysis coordinator, C.M.P., K.A.H.; DNA sequence analysis, D.C.K., L.D.; microRNA analysis, K.C.H.; mRNA sequence analysis, A.D.C., M.M.; histone analysis, G.B.; microRNA sequence analysis, A.G.R.; DNA methylation analysis, S. Malik, M. Sarker, P.W.L.; copy number analysis; A.D.C., M.M.; protein analysis, W.L., R.G.W., G.B.M.; pathway/integrated analysis, C.J.C., C.Y., J.M.S., C.C.B., G.C., S.S., R.S.; biospecimen core resource, T.L.B., J.M.G.; pathology and clinical expertise, T.A.K., H.H., R.J.M., J.N.I., T.S., F.W.

Author Information

All of the primary sequence files are deposited in CCHub (https://cchub.unc.edu/); all other data including mutation annotation file are deposited at the Data Coordinating Center (http://cancergenomics.ucsf.edu/); Sample list version of the paper and supporting data can be found at http://tcga-data.nci.nih.gov/docs/publications/brc2a_2012/. The data can be explored via the ISB Regulome Explorer (http://explorer.cancerregulome.org/) and the cBio Cancer Genomics Portal (http://cbioportal.org). Data donations can be found at https://wiki.nci.nih.gov/display/brca_2012/. The data can be explored via the ISB Regulome Explorer (http://explorer.cancerregulome.org/) and the cBio Cancer Genomics Portal (http://cbioportal.org). Data donations can be found at https://wiki.nci.nih.gov/display/brca_2012/. The data can be explored via the ISB Regulome Explorer (http://explorer.cancerregulome.org/) and the cBio Cancer Genomics Portal (http://cbioportal.org). Data donations can be found at https://wiki.nci.nih.gov/display/brca_2012/.
