Section IV

Pathology and Biological Markers of Invasive Breast Cancer

Chapter 27

Breast Cancer Genomics

Paul T. Spellman, Laura Heiser, and Joe W. Gray

Affiliation(s): Lawrence Berkeley National Laboratory, Life Sciences Division

Breast cancer is predominantly a disease of the genome with cancers arising and progressing through accumulation of aberrations that alter the genome—by changing DNA sequence, copy number, and structure in ways that that contribute to diverse aspects of cancer pathophysiology. Classic examples of genomic events that contribute to breast cancer pathophysiology include inherited mutations in BRCA1, BRCA2, TP53, and CHK2 that contribute to the initiation of breast cancer, amplification of ERBB2 (formerly HER2) and mutations of elements of the PI3-kinase pathway that activate aspects of epidermal growth factor receptor (EGFR) signaling and deletion of CDKN2A/B that contributes to cell cycle deregulation and genome instability. It is now apparent that accumulation of these aberrations is a time-dependent process that accelerates with age (1). Although American women living to an age of 85 have a 1 in 8 chance of developing breast cancer, the incidence of cancer in women younger than 30 years is uncommon. This is consistent with a multistep cancer progression model whereby mutation and selection drive the tumor's development, analogous to traditional Darwinian evolution (2, 3). In the case of cancer, the driving events are changes in sequence, copy number, and structure of DNA and alterations in chromatin structure or other epigenetic marks.

Our understanding of the genetic, genomic, and epigenomic events that influence the development and progression of breast cancer is increasing at a remarkable rate through application of powerful analysis tools that enable genome-wide analysis of DNA sequence and structure, copy number, allelic loss, and epigenomic modification. Application of these techniques to elucidation of the nature and timing of these events is enriching our understanding of mechanisms that increase breast cancer susceptibility, enable tumor initiation and progression to metastatic disease, and determine therapeutic response or resistance. These studies also reveal the molecular differences between cancer and normal that may be exploited to therapeutic benefit or that provide targets for molecular assays that may enable early cancer detection, and predict individual disease progression or response to treatment. This chapter reviews current and future directions in genome analysis and summarizes studies that provide insights into breast cancer pathophysiology or that suggest strategies to improve breast cancer management.

CANCER GENOME SEQUENCE

Mutations and Polymorphisms

The discovery of germ-line and somatic mutations has long been a critical component of cancer research. Driven by increasingly powerful normal and tumor DNA sequencing capabilities, a substantial

number of germ-line and somatic mutations and polymorphisms have been associated with an increased risk of breast cancer and with aspects of breast cancer pathogenesis.

To date, DNA sequence-based studies have established associations between germ-line mutations in TP53, BRCA1, BRCA2, and PTEN with high breast cancer risk and mutations in CHEK2, ATM, NBS1, RAD50, BRIP1, and PALB2 with approximately twofold increased breast cancer risk (4). In addition, genome-wide analyses of single nucleotide polymorphisms have established associations between polymorphisms in FGFR2, TNRC9, MAP3K1, and LSP1 with significant but modestly increased breast cancer risk (5).

Early DNA sequence analyses of primary tumors and cell lines revealed somatic mutations that contribute strongly to breast cancer pathophysiology. These include mutations such as TP53, CDH1, and PIK3CA that are commonly mutated in breast cancer (from COSMIC http://www.sanger.ac.uk/genetics/CGP/cosmic/). The human genome project proved the viability of sequencing entire genomes and since that time, it has been an obvious goal to search systematically for the mutations that drive cancer. Recent work from groups at the Sanger Center (6), Johns Hopkins University (7), and The Cancer Genome Atlas (8) highlight just how achievable this goal is, and what we can expect to learn when it is completed.

New Discovery Approaches

The critical issue is to determine which genes, when mutated, are responsible for cancer pathophysiology. Two approaches have been proposed to identify statistically significant recurrent mutations. One approach is to target specific classes of genes in a large survey of samples (100s), which provides the ability to detect rare but still recurrent mutational events. The other approach is to sequence all genes in a limited number of samples (10) to identify the common mutational events that drive tumorigenesis. The result of these studies has been to show that (a) a relatively large number of mutated genes (20) are likely to be driving each individual cancer and (b) hundreds of genes can act as drivers of breast cancer in total. Examining the functions of genes that can drive cancer when mutated has shown that they may not always be in genes predicted to harbor oncogenic or tumor suppressor functions.

Sequencing tumor DNA to identify somatic mutations is technically and logistically complicated (Fig. 27.1). Work published to date has used industry standard dideoxy chain termination based sequencing (9). The standard approach is to use polymerase chain reaction (PCR) to amplify each exon of the gene to be sequenced and then sequence each exon individually. For example, sequencing 1,000 typical genes with 10 exons each, in 100 samples entails 1,000,000 DNA sequencing reactions covering approximately 150 million basepairs (bp) of total DNA sequence, which is a substantial amount of work. These sequences are compared with the reference human genome and if they do not match are flagged as possible mutations. It is critical to note that there will be thousands of sequences in the tumor different than the reference genome caused by sequencing errors and normal sequence polymorphisms. Each potential mutation must be resequenced to ensure it is not an error and then the tumor and normal sequence must be compared to ensure that the variation from the reference genome is not a

polymorphism. Technically, Sanger sequencing is challenging because each reaction produces a single electropherogram that is the average sequence of all DNA molecules that are PCR amplified at that region. This is critical because DNA from tumors is not uniform, being contaminated by normal tissue and by heterogeneity within the tumor itself. In some rare cases a tumor may be homozygous for a mutation (caused by loss of heterozygosity see below), but in most cases tumors are heterozygous for a mutation, meaning that the electrophoretic trace will have two peaks at a given base pair. As the ratio of mutated to normal sequence decreases, the ability to detect a mutation correspondingly decreases. Typically, Sanger sequencing cannot reliably detect mutant alleles that are less than 30% of the total (8).

The Sanger Center (10) performed a screen to identify mutations in kinases predicated on two assumptions (a) kinases are key regulators of cell signaling and (b) targeted inhibition of kinases is a well-understood process, which is amenable to therapeutic intervention making genes that are mutated in the survey useful drug targets. In a survey of all 518 kinase genes across 210 diverse human cancers (including 16 breast cancers), they observed more than 900 nonsynonomous somatic mutations that altered the amino acid sequence of a kinase. Not all of the mutations are actually responsible for cancer, however, and the identification of those mutations that drive cancer (versus those that are random passengers) is a critical question. Through the use of selection pressure analysis (based on comparing the rate of synonymous mutations that did not change amino acid sequence with nonsynonymous, those that change the amino acid sequence), the Sanger Center group estimated that most (80%) of the observed mutations were passenger mutations that did not influence the development of the cancer. This meant that in the 210 cancers they sequenced they identified approximately 150 mutations in protein kinase genes that were likely to drive oncogenesis. The overall power of the effort was not sufficient to estimate which genes are oncogenic drivers (most genes showed only single mutations), but it was possible to rank the genes as shown in Table 27.1. Many of the genes with the highest selection pressures (and overall ranking) are either known to be involved in cancer through genetics or previous mutation studies (e.g., ATM, FGFR2, and STK11) (11–13), or have homologs that are known to be involved in oncogenesis (e.g., MAP2K4).

The Johns Hopkins group took the complementary approach. They sequenced every gene in 11 breast cancer cell lines samples as a discovery phase (7, 14) followed by a validation phase of 24 breast tumor samples in which they sequenced only those genes that showed mutations in the discovery phase. Part of the rationale for the limited number of samples was to be genomically complete but another part was a resource limitation; the Johns Hopkins group used cell lines of which only a few matched the normal DNA is available. Sequencing cell lines for the discovery phase makes the sequencing traces more easily interpreted because there is neither normal DNA contamination nor tumor heterogeneity. The results of these studies were revolutionary, suggesting that hundreds, and possibly thousands, of genes have the potential to promote cancer when mutated. Further, individual breast tumors are likely to harbor 20 or more somatic mutations that are responsible for that specific cancer. These observations caused considerable controversy, but have now been generally accepted and new evidence from sequencing other tumor types has strengthened this argument (15, 16).

Recent Results

Nearly 10% of all human genes in these discovery efforts harbored mutations in the discovery screen. This seems extraordinarily high but the background mutation rate of the tumors is roughly one mutation per megabase of genomic DNA (which corresponds to about 3,000 random mutations in the genome) that it is not wholly unexpected. The validation phase, however proved truly remarkable; 167 genes harbored mutations in both the discovery and validation phases, which is approximately 1% of all human genes. Beyond the large numbers of genes observed to be mutated, two key observations emerge from this analysis: first, many recurrent mutations occur in genes that are not obviously related to cancer (i.e., the glycosylase GALNT5 and the transglutaminase TGM3); second and more importantly, many of the mutations appear to be clustered in pathways. For example, at least seven biological pathways, including ATM signaling and apoptosis induction, show significant levels of mutation (17) (Table 27.2). The observation that particular pathways are significantly mutated provides a framework in which to understand their functional significance, and further suggests that a few key pathways may be especially critical for the development of oncogenesis.

The next step, which is still in its infancy, is to relate the patterns of mutation to clinical outcome and treatment. For genes previously known to drive breast cancer when mutated, such as TP53, PIK3CA, and PTEN, the relationship between mutation and outcome for breast cancer has been examined and shown to be significant (18–20). In fact, conditional analysis of the activating mutation of PIK3CA (the catalytic subunit of PI3-kinase) or inactivating events of its negative regulator PTEN, provided better discrimination of outcome than mutation of either gene alone (Fig. 27.2).

STRUCTURAL ANALYSIS OF THE CANCER GENOME

Metaphase Chromosome Analysis

One of the most methodologically challenging questions in breast carcinomas is to understand the structural organization of a tumor genome. This is a critical area of research, which can easily be evidenced by the large number of chromosomal fusion events that drive malignancies in the leukemias and lymphomas. Recurrent mutations have become even more interesting since the recent identification of recurrent gene fusion events including TMPRSS-ERG in prostate cancer (21). Analyses of metaphase chromosome spreads from cultures of human tumors using classic banding techniques provided the first views of the extent of structural rearrangements that exist in human breast cancers. The catalogue by Mitelman et al. (22) provides a comprehensive assessment of breast cancer chromosome changes discovered using this approach. The general structural and numerical chaos is clear from these studies, but the approach is difficult because of the difficulty of preparing metaphase spreads of sufficiently high-quality metaphase chromosome preparations to allow identification of rearrangements with confidence.

The introduction of fluorescence in situ hybridization (FISH) with whole chromosome probes (23, 24) and the subsequent development of combinatorial multicolor labeling and analysis (25–27) substantially simplified the identification interpretation of these complex karyotypic rearrangements. Molecular cytogenetic analyses using whole chromosome analysis techniques are, however, limited in resolution to a few million base pairs by the complex organization of DNA along chromosomes. FISH, with multiple,

region-specific probes, enables high resolution mapping of the structures of numerical and structural chromosome abnormalities. However, this approach is not well suited to discovery or high-resolution analysis of complex structural aberrations.

The development of end sequence profiling (ESP) and paired end deep sequencing more generally have revolutionized analysis of structural aberrations in human breast cancers. Genome sequencing methodologies have used the paired end read technology since its proof of principle in sequencing de novo genomes (28, 29). Pair end reads, where DNA sequence from both ends of a longer piece of DNA (with known sequence length) allow the information of the sequences to be organized in useful ways, primarily to provide larger clone coverage of the genome or to span intermediary sequences that are common repeat elements (nearly one-half of the human genome). A diagram of paired end read methods is provided in Figure 27.3. Paired end reads have historically been performed on highly size-restricted pieces of DNA (either DNA clones of 1–10 kb), fosmids (of 50 kb), or bacterial artificial chromosome (BAC) of (150 kb). The longer the insert, the greater the clone coverage. A total of 2,000 BAC each 150 kb in length would provide average one-fold (1X) clone coverage of the entire genome (3 gigabases), whereas it would require 3 million, 1-kb clones to perform the same. If the goal is to understand which large regions of the genome are hooked together, larger inserts are more effective. As a note, insert lengths above 3 kb are sufficient to bridge most common repetitive sequences.

Initial work in the area of cancer genome structure discovery published in 2003 (30) involved reconstruction of the genomes of cancers by aligning paired end BAC reads against the human genome and searching for cases in which the ends do not map to consistent locations. Those clones that violate genome assembly indicate the presence of a rearrangement in the cancer genome (or an artifact in the BAC DNA library construction). These possibilities can be separated if multiple BACs with different end sequences support the same rearrangement

New work is ongoing both to identify rearrangement breakpoints and to reconstruct the likely genome of the cancer (31,111). These reconstructions may aid in understanding tumor evolution and disease progression, and may identify new therapeutic targets or interventions. New DNA sequences created in the tumor genome might even be targets for patient-directed personalized therapy if gene fusions create novel proteins or chimeric transcripts that can be targeted with emerging small interfering RNA (siRNA)-based therapeutics (32).

New sequencing technologies have obviated the BAC-based approaches developed in the previous decade. The new methods still perform paired end reads but, compared with conventional

Sanger-based sequencing, the costs have dropped more than 1,000 fold, meaning that it is cheaper to sequence 1,000,000 3-kb DNA paired-end fragments than 2,000 BACs. As prices drop further, to the estimated $1,000 genome, evaluating the structures of genomes will become even easier. The first example of this exciting new work from the Sanger Center group in lung cancer (33) has shown what these observations are likely to allow (Fig. 27.4). In just two cell lines, 103 somatic rearrangements were observed, including chromosomal fusions, tandem duplications, and inverted duplications. The deep

sequencing approach provides both structural data and very high resolution copy number data, because the number of sequencable reads is proportional to the DNA copy number.

GENOME COPY NUMBER

Comparative genomic hybridization (CGH) is a hybridization-based analysis strategy that maps changes in genome copy number onto a normal representation of the genome. In CGH, DNA from a test tumor is labeled and hybridized to a normal genome representation and the amount of bound, labeled tumor DNA relative to that for a normal genome is measured along the genome representation as an indication of relative genome copy number. CGH is particularly informative because it provides a direct link between a genome copy number abnormality and its gene content.

The first CGH analyses mapped changes of recurrent aberrations onto normal metaphase chromosomes (34). This approach, however, was quickly supplanted by array CGH in which the genome representation was replaced by arrays of nucleic acid probes (35, 36). Initially these arrays were composed of cloned probes, such as yeast artificial chromosome (YAC), BAC (35), and complementary DNA (cDNA) (37). Eventually, however, it was demonstrated that CGH could be accomplished by hybridization to arrays of synthetic oligonucleotides (38). This enabled use of commercial arrays so that today, commercially available oligonucleotide arrays that carry millions of probes are in common use. Current platforms provide kilobase pair resolution (39). In addition, some platforms have been developed to allow allele-specific discrimination so that the analysis yields allele-specific copy number information. Prominent commercial platforms are listed in Table 27.3.

Individual Tumor Genome Landscapes

One of the remarkable features of breast cancer genomes is the extent of abnormality within individual tumors. Figure 27.5 shows a typical breast cancer CGH profile with copy number abnormalities caused by gains or losses of single copies of portions of the genome, homozygous deletions, and high-level amplification. It is not unusual to find as much as 30% of a breast cancer genome present at abnormal copy number. The extent of the abnormalities may encompass entire chromosomes or as little as a few hundred base pairs. In general, the low-level copy number abnormalities tend to extend over significant parts of the genome, whereas homozygous losses and high-level amplifications involve relatively narrow parts of the genome. These differences in genomic extent of abnormality may reflect the mechanisms by which the abnormalities arise. Low-level copy number gains and losses involving whole chromosomes or chromosome arms may be caused by errors in chromosomal segregation—for example, owing to centrosome dysfunction (40). On the other hand, homozygous deletions and high-level amplifications are evidence of strong and active selection for or against regions of the genome that are particularly important in the pathophysiology of the disease—for example, by bridge-breakage-fusion (41) or through production and amplification of extrachromosomal elements (42).

Another remarkable feature of breast cancer genomes is the extent of variation between tumors. Figure 27.6, for example, shows CGH profiles for two breast tumors with similar clinical characteristics. One genome shows almost no aberrations whereas the other is shattered and displays numerous regions of high-level amplification and homozygous deletion. Some of these aberrations—both structural (43) and

numerical (39)—are extremely complex. Complex aberrations (39) of closely spaced aberrations have been referred to as “firestorms.” In most tumors, the aberrations accumulate during a relatively restricted portion of tumorigenesis and change slowly thereafter. In fact, the CGH profiles of metastatic tumors typically are similar to the profiles of the primary tumors from which they arise—even though the time between primary tumor and metastatic tumor development may be decades (44,45). 368 Section VI Pathology and Biological Markers of Invasive Breast Cancer

Recurrent Aberrations

The ability to map genome copy number abnormalities onto a normal representation of the genome facilitates identification of recurrent aberrations because many tumor profiles can be integrated onto the same representations. Figure 27.7, for example, shows the frequencies of recurrent copy number gains, losses, and amplifications in 145 primary breast tumors measured using BAC array CGH (46) as well as the locations of 9 regions of recurrent high-level amplification involving regions of chromosomes 8, 11, 12, 17, and 20. Numerous CGH studies in the last 5 years support the general locations of recurrent copy number increases involving chromosomes depicted in Figure 27.7 as well as the observation that relatively few parts of the genome are not abnormal in at least 15% of breast tumors (37, 39, 46–56). Recent studies using high-resolution oligonucleotide array CGH have defined the extents of these regions of recurrent abnormality with subgene resolution and demonstrate the existence of an increasing number of aberrations that involve very small regions of the genome that were missed with lower resolution technologies (39, 57, 58).

Integrative analyses of gene expression and genome copy number data indicate that expression levels of greater than 10% of the entire genome are deregulated by these recurrent abnormalities in breast cancers (46, 53). Interestingly, a recent meta-analysis of 5,918 malignant epithelial tumors showed that the copy number gains involving 1q, 3q, 5p, 7q, 8q, 17q, and 20q, and losses at, 4q, 13q, 17p, and 18q found in breast cancer were also common in many other epithelial neoplasias (59). This suggests that these abnormalities may be generally important in carcinogenesis and it is consistent with the observation that the ensemble of genes deregulated by low-level genome copy number abnormalities preferentially affects genes that may contribute to increased metabolic fitness (46).

Several studies have compared recurrent genome copy number changes between clinicopathologic subtypes. These studies showed significant differences between estrogen receptor-positive (ER +) and ER-negative (ER -) tumors (60), between subtypes defined by transcriptional profiling (e.g., basal, ERBB2, luminal A and luminal B) (46, 56) and tumors defined according to histologic features (e.g., ductal vs. lobular) (61). For example, higher numbers of gains or losses are associated with the basal-like, ER - tumor subtype, whereas high-level DNA amplification is more frequent in luminal-B subtype tumors (46, 56). Interestingly, aging does not seem to influence the recurrent abnormality content (62).

A growing number of publications in breast cancer and other tumor types suggest that the pattern of recurrent abnormalities is also influenced by the presence of germ-line mutations, polymorphisms, or both that influence aspects of the DNA repair machinery. For example, breast tumors in which p53 (52) or BRCA1 (63, 64) are aberrant tend to accumulate more abnormalities than do tumors with normal p53

and BRCA1 function. Moreover, the pattern of recurrent aberrations is different in tumors that arise in individuals with BRCA1 mutations than in sporadic tumors (63, 64). Studies in mouse models also support the concept that the underlying individual genotype influences the spectrum of aberrations that arise. For example, CGH studies of five mouse models of breast cancer induced by wild-type and mutated forms of oncogenic ERBB2 or the polyomavirus middle T antigen (PyMT) showed that the pattern of genome copy number abnormalities was strongly influenced by the driving oncogene (65).

GENOME INSTABILITY AND EVOLUTION

Genome instability in solid tumors is thought to enable accumulation of the spectrum of genomic abnormalities needed for tumor progression (67–68). Tools, such as CGH, demonstrate the result of instability and show that the number of aberrations increases during tumor progression. The cartoon in Figure 27.8, for example, suggests that aberrations measured using CGH in breast cancer tend to increase dramatically during progression to ductal carcinoma in situ (DCIS) (69). CGH analyses do not, however, provide a direct measure of instability because they only show aberrations that are present in most of the cells in the tumor.

Studies of the rate of instability in breast cancers assessed in thick tissue sections using FISH with chromosome-specific probes (69) showed that variability was low in normal ductal epithelium and usual ductal hyperplasia (UDH) but remarkably high in DCIS. Variability remained high in invasive cancer (IC). DCIS and IC also showed regions of increased overall ploidy. The variations in genome copy number in DCIS and IC were dramatic between adjacent cells, suggesting that the copy number changes were not the result of clonal evolution but rather were the result of continuing high instability. The instability observed in DCIS was similar to that observed in cultures of breast epithelial cells (70) and in epithelial cells in mice lacking protective telomere function (71–73). Quantitative FISH analyses of relative telomere length during breast cancer progression show decreasing telomere length during the period leading up to the time of increased instability, as detected using FISH and the rapid accumulation of chromosome aberrations as measured by CGH (69).

MECHANISMS OF ABERRATION FORMATION

The remarkable genomic complexity and diversity in the genome sequence, structure, and copy number observed in breast cancers raises the question of how these aberrations form. It is likely that several mechanisms are involved.

Telomere Crisis

One important issue to understand is the remarkable increase in genome instability and number of numerical and structural aberrations that occur during transition to DCIS. A likely explanation is that breast cancers begin with hyperplastic growth—likely initiated by epigenomic events (40, 74, 75)—in cells lacking functional telomerase. This causes progressive telomere loss and culminates in loss of protective telomere function and dramatically increased genome instability. In most cases, this is the end of the story because the dysfunctional cells die. Rarely, however, the genome instability may produce a genomic composition that reactivates telomerase and confers a proliferative advantage; DCIS

or IC is the result. The extremely low probability of progressing through telomere crisis and reactivating telomerase would explain why hyperplasia, although it has many of the proliferative hallmarks of cancer, is associated with only modest cancer risk (76). The stochastic nature of passage through telomere crisis may explain, in large part, the remarkable variation in genome composition between individuals.

The modest changes in the spectrum of genome copy number abnormalities observed as breast cancers progress from DCIS to metastatic cancer (77, 78) is in seeming contradiction to the continued high rate of instability observed in DCIS and IC. One possible explanation is that most of the genome copy number abnormalities observed in this study leave the cells at a proliferative disadvantage relative to a genomically unstable tumor initiating population that maintains the overall genotype during progression. This possibility is supported by analyses of breast cancer cell lines grown in vitro and as xenografts in which the average genome copy number profiles for such cultures or tumors measured using CGH evolve very slowly although FISH analyses demonstrate dramatic genome instability (69).

Although the telomere crisis model explains the large variation in genome composition between individual tumors, it does not explain recurrent aberrations. It seems likely that recurrent aberrations are the result of positive and negative selection during the progression process. This is clearly established for strong oncogenes and tumor suppressor genes that are associated with high level amplification (e.g., ERBB2, MYC, CCND1) and homozygous deletions (e.g., CDKN2A and PTEN). However, it is likely also the case for low level aberrations. Evidence of this is the observation that these aberrations preferentially deregulate genes associated with increased metabolic activity (69).

Of course the section process also is influenced by the environment and the genotype of the individual in which the cancer arises. For example, breast tumors in which p53 or BRCA1/2 are mutationally inactivated accumulate more abnormalities than do tumors with wild-type p53 and BRCA1. Studies in other tumor types implicate other DNA repair genes as modulators of genome copy number abnormality accumulation. Studies in mouse models also support the concept that the underlying individual genotype influences the spectrum of aberrations that arise. For example, CGH studies of five mouse models of breast cancer induced by wild-type and mutated forms of oncogenic ERBB2 or the PyMT showed distinctive, oncogene-associated patterns of genome copy number abnormality. Likewise, CGH analyses of oncogene-induced mouse pancreatic islet cell carcinomas showed that the individual genetic background strongly influenced genome copy number formation (79).

Microenvironment-related factors, such as expression of matrix metalloproteinases (MMP) in the stromal compartment (80, 81), also strongly influence tumorigenesis and the onset of genome instability.

GENOME TARGETED THERAPIES AND MARKERS

Clinical Markers

Several recurrent genome aberrations have been associated with poor outcome. Examples of prognostic markers reported to date for breast cancer include (a) association of the total number of genome copy

number aberrations with reduced survival duration (82); (b) gain of 3q as a stronger predictor of recurrence in lymph node-negative invasive cancer (83); (c) simultaneous chromosome 1q gain and 16q loss as a predictor of slow proliferation (84); (d) gain of chromosome 3q, 9p, 11p, and 11q and loss of 17p associated with short-term survival (85); (e) gain of 11q13, 12q24, 17, and 18p associated with metastasis-free survival (82); (f) gain at 8q24 associated with mutational status of p53 and reduced survival duration (52); (g) increased DNA copy number of RAB25 associated with markedly decreased disease-free survival or overall survival (86); and (h) amplification and overexpression of 66 genes in regions of amplification at 8p11, 8q24, 11q13, 17q21, and 20q13 with reduced survival duration (56) and the presence of aberration hotspots or fire storms (39).

Therapeutic Targets

Several recurrently aberrant genes associated with reduced survival duration and other aspects of breast cancer pathophysiology have been suggested as therapeutic targets in breast cancer. Amplified genes implicated as therapeutic targets include ERBB2 (87), TOP2A, (88), CCND1 and EMS1 (89), MYC (90), ZNF217 (91), RAB25 (86), MDM2 (92), TBX2 (93), RPS6KB1, and the microRNA mir-21 (94). More recently, correlative analyses of gene expression and high-level amplification have identified 66 genes in regions of amplification that are associated with reduced survival duration that are candidate therapeutic targets, 9 of which (FGFR1, IKBKB, ERBB2, PROCC, ADAM9, FNTA, ACACA, PNMT, and NR1D1) were predicted to be drugable (46). Recurrent mutations or deletions in breast cancers include TP53 (95), PIK3CA (96), PTEN (97), BRCA1 (98), and BRCA2 (99).

Recurrent genomic aberrations in breast cancer are attractive targets because they are events for which strong evidence indicates positive selection so the tumors may be addicted to the aberration. In addition, aberrations are not present in normal tissues so that therapies against them are likely to be relatively nontoxic. ERBB2 is the prototypic genome-based therapeutic target (100,101). This receptor tyrosine kinase is highly amplified in about 30% of human breast cancers and the antibody, trastuzumab (102), and the small molecule inhibitor, lapatinib (103), have proved to be clinically effective against tumors in which ERBB2 is amplified. One of the advantages of aberration-targeted therapies is that markers can be readily developed to identify tumors carrying the aberration that are most likely to respond to therapy. In the case of ERBB2-targeted therapies, FISH with probes to ERBB2 readily identify tumor to be treated (104). Following this lead, therapies directed against tumors with aberrations involving TP53 (105), MDM2 (106), TOP2A (107), PI3-kinase mutations (PTEN and PIK3CA) (108), and BRCA1 (109) are now being developed or tested. Clearly, the development of aberration-targeted therapy is only beginning.

FUTURE: WHAT DO WE NEED TO KNOW?

Application genome-wide analysis tolls, such as CGH and high throughput sequencing, are revealing the recurrent aberrations that contribute to breast cancer genesis and progression.

They also are demonstrating the existence of hundreds of low-frequency aberrations. We also are beginning to understand how these aberrations influence the expression of coding and noncoding transcripts to influence cancer pathophysiology. We do not know, however, how these aberrations

cooperate with each other (or substitute for each other) and with epigenomic aberrations to result in the overall cancer phenotype. We do not know how the aberrations arise during progression. And we do not know how the aberrations contribute to the development of drug resistance—especially in metastatic disease.

Progression Map

Recent genome-wide, high resolution breast cancer analyses so far have focused mostly on assessment of aberrations in invasive breast cancers and breast cancer cell lines. This provides a working aberration parts list and identifies aberrations that may be useful as prognostic or predictive markers or as therapeutic targets. They have not, however, provided substantial information about how these aberrations arise and evolve during progression. This information will come from longitudinal integrated "omic" analyses of genome aberration appearance during tumor progression. Studies in mouse models can provide some information, but studies of changes during the evolution of individual human tumors will be invaluable. This will require a long-term, sustained effort from the breast cancer research community to collect the necessary samples and continued refinement of large-scale omic analysis technologies so they are capable of analyzing the small amounts of neoplastic tissue that may be available at early stages of evolution.

Drug Resistance

Less than 30% of women with metastatic breast cancer will survive 5 years. This is in contrast to treatment of early disease where outcomes have improved greatly over the last decade. Dozens of next generation therapies are being designed to target these aberrations. The best developed of these in breast cancer are trastuzumab and lapatinib, which target ERBB2-positive tumors. The responses to these agents, even when combined with conventional chemotherapeutic agents, are not, however, durable in patients with metastatic disease and long-term survival is rare. It seems likely that current treatment strategies fail because these strategies do not take into account the genomic and epigenomic aberrations that contribute to resistance in metastatic breast cancer, resistance-related homeostatic, or feedback loops induced by pathway-targeted therapies and factors unique to the metastatic microenvironments in the bone marrow, lung, liver, and brain that contribute to therapeutic resistance. Development of a detailed "omic" understanding of drug-resistant, metastatic cancer will greatly facilitate treatment of this important aspect of the disease. This will require development of clinical trials in which samples of metastatic breast tissue suitable for large-scale omic analysis are acquired before and after development of drug resistance as well as the development of experimental model systems that mirror the aberrations that contribute to resistance.

Model Systems

Development of well-characterized human and murine model systems in which these aberrations function singly and together will be essential to sort out aberration function. Manipulations of gene function in the compendium of mouse breast cancer models being developed by the National Cancer Institute (NCI) Mouse Models of Human Cancer Consortium (http://emice.nci.nih.gov/) and the

collections of well-characterized human breast cancer cell lines mirroring the aberrations found in human tumors (110) will facilitate these efforts.

ACKNOWLEDGEMENTS

This work was supported by the National Institutes of Health (NIH) (CA58207, CA112970, CA126477), by the Director, Office of Science, Office of Biological and Environmental Research, of the U.S. Department of Energy under Contract No. DE–AC02–05CH11231, and the Avon foundation. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does the mention of trade names, commercial products, or organization imply endorsement by the U.S. Government.

References

1. Wun LM, Merrill RM, Feuer EJ. Estimating lifetime and age-conditional probabilities of developing cancer. *Lifetime Data Anal* 1998;4:169–186.
2. Foulds L. The experimental study of tumor progression: a review. *Cancer Res* 1954;14:327–339.
3. Nowell PC. The clonal evolution of tumor cell populations. *Science* 1976;194: 23–28.
4. Walsh T, King MC. Ten genes for inherited breast cancer. *Cancer Cell* 2007;11: 103–105.
5. Easton DF, Pooley KA, Dunning AM, et al. Genome-wide association study identifies novel breast cancer susceptibility loci. *Nature* 2007;447:1087–1093.
6. Stephens P, Edkins S, Davies H, et al. A screen of the complete protein kinase gene family identifies diverse patterns of somatic mutations in human breast cancer. *Nat Genet* 2005;37:590–592.
7. Wood LD, Parsons DW, Jones S, et al. The genomic landscapes of human breast and colorectal cancers. *Science* 2007;318:1108–1113.
8. Cancer Genome Atlas Network. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature* 2008;455:1061–1068.
9. Sanger F, Coulson AR. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *J Mol Biol* 1975;94:441–448.
10. Greenman C, Stephens P, Smith R, et al. Patterns of somatic mutation in human cancer genomes. *Nature* 2007;446:153–158.
11. Katoh M. Cancer genomics and genetics of FGFR2 [Review]. *Int J Oncol* 2008;33: 233–237.
12. Ahmed M, Rahman N. ATM and breast cancer susceptibility. *Oncogene* 2006;25: 5906–5911.
13. Campeau PM, Foulkes WD, Tischkowitz MD. Hereditary breast cancer: new genetic developments, new therapeutic avenues. *Hum Genet* 2008;124:31–42.
14. Wood LD, Calhoun ES, Silliman N, et al. Somatic mutations of GUCY2F, EPHA3, and NTRK3 in human cancers. *Hum Mutat* 2006;27:1060–1061.
15. Jones AM, Mitter R, Springall R, et al. A comprehensive genetic profile of phyllodes tumours of the breast detects important mutations, intra-tumoral genetic heterogeneity and new genetic changes on recurrence. *J Pathol* 2008;214: 533–544.
16. Parsons DW, Jones S, Zhang X, et al. An integrated genomic analysis of human glioblastoma multiforme. *Science* 2008;321:1807–1812.

17. Lin J, Gan CM, Zhang X, et al. A multidimensional analysis of genes mutated in breast and colorectal cancers. *Genome Res* 2007;17:1304–1318.
18. Langerod A, Zhao H, Borgan O, et al. TP53 mutation status and gene expression profiles are powerful prognostic markers of breast cancer. *Breast Cancer Res* 2007;9:R30.
19. Berns K, Horlings HM, Hennessy BT, et al. A functional genetic approach identifies the PI3K pathway as a major determinant of trastuzumab resistance in breast cancer. *Cancer Cell* 2007;12:395–402.
20. Bertheau P, Turpin E, Rickman DS, et al. Exquisite sensitivity of TP53 mutant and basal breast cancers to a dose-dense epirubicin-cyclophosphamide regimen. *PLoS Med* 2007;4:e90.
21. Mehra R, Tomlins SA, Shen R, et al. Comprehensive assessment of TMPRSS2 and ETS family gene aberrations in clinically localized prostate cancer. *Mod Pathol* 2007;20:538–544.
22. Mitelman F, Johansson B, Mertens F. *Mitelman database of chromosome aberrations in cancer.* 2002. http://CGAP.NCI.NIH.GOV/Chromosomes/Mitelman.
23. Pinkel D, Landegent J, Collins C, et al. Fluorescence in situ hybridization with human chromosome-specific libraries: detection of trisomy 21 and translocations of chromosome 4. *Proc Natl Acad Sci U S A* 1988;85:9138–9142.
24. Lichter P, Cremer T, Tang CJ, et al. Rapid detection of human chromosome 21 aberrations by *in situ* hybridization. *Proc Natl Acad Sci U S A* 1988;85:9664–9668.
25. Speicher MR, Gwyn Ballard S, Ward DC. Karyotyping human chromosomes by combinatorial multi-fluor FISH. *Nat Genet* 1996;12:368–375.
26. Liyanage M, Coleman A, du Manoir S, et al. Multicolour spectral karyotyping of mouse chromosomes. *Nat Genet* 1996;14(3):312–315.
27. Schrock E, du Manoir S, Veldman T, et al. Multicolor spectral karyotyping of human chromosomes. *Science* 1996;273:494–497.
28. Fleischmann RD, Adams MD, White O, et al. Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. *Science* 1995;269:496–512.
29. Myers EW, Sutton GG, Delcher AL, et al. A whole-genome assembly of Drosophila. *Science* 2000;287:2196–2204.
30. Volik S, Zhao S, Chin K, et al. End-sequence profiling: sequence-based analysis of aberrant genomes. *Proc Natl Acad Sci U S A* 2003;100:7696–7701.
31. Bashir A, Volik S, Collins C, et al. Evaluation of paired-end sequencing strategies for detection of genome rearrangements in cancer. *PLoS Comput Biol* 2008;4: e1000051.
32. Shen Y. Advances in the development of siRNA-based therapeutics for cancer. *IDrugs* 2008;11:572–578.
33. Campbell PJ, Stephens PJ, Pleasance ED, et al. Identification of somatically acquired rearrangements in cancer using genome-wide massively parallel paired-end sequencing. *Nat Genet* 2008;40:722–729.
34. Kallioniemi A, Kallioniemi OP, Sudar D, et al. Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. *Science* 1992;258:818–821.
35. Pinkel D, Seagraves R, Sudar D, et al. High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. *Nat Genet* 1998;20:207–211.
36. Solinas-Toldo S, Lampel S, Stilgenbauer S, et al. Matrix-based comparative genomic hybridization: biochips to screen for genomic imbalances. *Genes Chromosomes Cancer* 1997;20:399–407.

37. Pollack JR, Perou CM, Alizadeh AA, et al. Genome-wide analysis of DNA copynumber changes using cDNA microarrays. *Nat Genet* 1999;23:41–46.
38. Lucito R, Healy J, Alexander J, et al. Representational oligonucleotide microarray analysis: a high-resolution method to detect genome copy number variation. *Genome Res* 2003;13:2291–2305.
39. Hicks J, Krasnitz A, Lakshmi B, et al. Novel patterns of genome rearrangement and their association with survival in breast cancer. *Genome Res* 2006;16: 1465–1479.
40. Berman H, Zhang J, Crawford YG, et al. Genetic and epigenetic changes in mammary epithelial cells identify a subpopulation of cells involved in early carcinogenesis. *Cold Spring Harb Symp Quant Biol* 2005;70:317–327.
41. McClintock B. The production of homozygous deficient tissues with mutant characteristics by means of the aberrant mitotic behavior of ring-shaped chromosomes. *Genetics* 1938;23:315–376.
42. Kanda T and Wahl G. The dynamics of acentric chromosomes in cancer cells revealed by GFP-based chromosome labeling strategies. *J Cell Biochem Suppl* 2000;35:107–114.
43. Volik S, Raphael BJ, Huang G, et al. Decoding the fine-scale structure of a breast cancer genome and transcriptome. *Genome Res* 2006;16:394–404.
44. Kuukasjarvi T, Karhu R, Tanner M, et al. Genetic heterogeneity and clonal evolution underlying development of asynchronous metastasis in human breast cancer. *Cancer Res* 1997;57:1597–1604.
45. Nishizaki T, DeVries S, Chew K, et al. Genetic alterations in primary breast cancers and their metastases: direct comparison using modified comparative genomic hybridization. *Genes Chromosomes Cancer* 1997;19:267–272.
46. Chin K, DeVries S, Fridlyand J, et al. Genomic and transcriptional aberrations linked to breast cancer pathophysiologies. *Cancer Cell* 2006;10:529–541.
47. Adelaide J, Finetti P, Bekhouche I, et al. Integrated profiling of basal and luminal breast cancers. *Cancer Res* 2007;67:11565–11575.
48. Hermsen MA, Baak JP, Meijer GA, et al. Genetic analysis of 53 lymph node-negative breast carcinomas by CGH and relation to clinical, pathological, morphometric, and DNA cytometric prognostic factors. *J Pathol* 1998;186: 356–362.
49. Tirkkonen M, Tanner M, Karhu R, et al. Molecular cytogenetics of primary breast cancer by CGH. *Genes Chromosomes Cancer* 1998;21:177–184.
50. Moore E, Magee H, Coyne J, et al. Widespread chromosomal abnormalities in high-grade ductal carcinoma *in situ* of the breast. Comparative genomic hybridization study of pure high-grade DCIS. *J Pathol* 1999;187:403–409.
51. Richard F, Pacyna-Gengelbach M, Schluns K, et al. Patterns of chromosomal imbalances in invasive breast cancer. *Int J Cancer* 2000;89:305–310.
52. Jain AN, Chin K, Borresen-Dale AL, et al. Quantitative analysis of chromosomal CGH in human breast tumors associates copy number abnormalities with p53 status and patient survival. *Proc Natl Acad Sci U S A* 2001;98:7952–7957.
53. Pollack JR, Sorlie T, Perou CM, et al. Microarray analysis reveals a major direct role of DNA copy number alteration in the transcriptional program of human breast tumors. *Proc Natl Acad Sci U S A* 2002;99:12963–12968.
54. Naylor TL, Greshock J, Wang Y, et al. High resolution genomic analysis of sporadic breast cancer using array-based comparative genomic hybridization. *Breast Cancer Res* 2005;7:R1186–R1198.

55. Fridlyand J, Snijders AM, Ylstra B, et al. Breast tumor copy number aberration phenotypes and genomic instability. *BMC Cancer* 2006;in press.
56. Bergamaschi A, Kim YH, Wang P, et al. Distinct patterns of DNA copy number alteration are associated with different clinicopathological features and geneexpression subtypes of breast cancer. *Genes Chromosomes Cancer* 2006;45: 1033–1040.
57. Staaf J, Lindgren D, Vallon-Christersson J, et al. Segmentation-based detection of allelic imbalance and loss-of-heterozygosity in cancer cells using whole genome SNP arrays. *Genome Biol* 2008;9:R136.
58. Komatsu A, Nagasaki K, Fujimori M, et al. Identification of novel deletion polymorphisms in breast cancer. *Int J Oncol* 2008;33:261–270.
59. Baudis M. Genomic imbalances in 5,918 malignant epithelial tumors: an explorative meta-analysis of chromosomal CGH data. *BMC Cancer* 2007;7:226.
60. Loo LW, Grove DI, Williams EM, et al. Array comparative genomic hybridization analysis of genomic alterations in breast cancer subtypes. *Cancer Res* 2004;64: 8541–8549.
61. Nishizaki T, Chew K, Chu L, et al. Genetic alterations in lobular breast cancer by comparative genomic hybridization. *Int J Cancer* 1997;74:513–517.
62. Yau C, Fedele V, Roydasgupta R, et al. Aging impacts transcriptomes but not genomes of hormone-dependent breast cancers. *Breast Cancer Res* 2007;9: R59.
63. Wessels LF, van Welsem T, Hart AA, et al. Molecular classification of breast carcinomas by comparative genomic hybridization: a specific somatic genetic profile for BRCA1 tumors. *Cancer Res* 2002;62:7110–7117.
64. van Beers EH, van Welsem T, Wessels LF, et al. Comparative genomic hybridization profiles in human *BRCA1* and *BRCA2* breast tumors highlight differential sets of genomic aberrations. *Cancer Res* 2005;65:822–827.
65. Hodgson JG, Malek T, Bornstein S, et al. Copy number aberrations in mouse breast tumors reveal loci and genes important in tumorigenic receptor tyrosine kinase signaling. *Cancer Res* 2005;65:9695–9704.
66. Lengauer C, Kinzler KW, Vogelstein B. Genetic instabilities in human cancers. *Nature* 1998;396:643–649.
67. Loeb KR, Loeb LA. Significance of multiple mutations in cancer. *Carcinogenesis* 2000;21:379–385.
68. Myung K, Chen C, Kolodner RD. Multiple pathways cooperate in the suppression of genome instability in Saccharomyces cerevisiae. *Nature* 2001;411:1073–1076.
69. Chin K, de Solorzano CO, Knowles D, et al. *In situ* analyses of genome instability in breast cancer. *Nat Genet* 2004;36:984–988.
70. Romanov S, Kozakliewicz B, Holst C, et al. Normal human mammary epithelial cells spontaneously escape senescence and acquire genomic changes. *Nature* 2001;409:633–637.
71. Artandi SE, Chang S, Lee SL, et al. Telomere dysfunction promotes non-reciprocal translocations and epithelial cancers in mice. *Nature* 2000;406:641–645.
72. Artandi SE, DePinho RA. A critical role for telomeres in suppressing and facilitating carcinogenesis. *Curr Opin Genet Dev* 2000;10:39–46.
73. Chang S, Khoo C, DePinho RA. Modeling chromosomal instability and epithelial carcinogenesis in the telomerase-deficient mouse. *Semin Cancer Biol* 2001;11: 227–239.

74. Bean GR, Bryson AD, Pilie PG, et al. Morphologically normal-appearing mammary epithelial cells obtained from high-risk women exhibit methylation silencing of INK4a/ARF. *Clin Cancer Res* 2007;13:6834–6841.
75. Crawford YG, Gauthier ML, Joubel A, et al. Histologically normal human mammary epithelia with silenced p16(INK4a) overexpress COX-2, promoting a premalignant program. *Cancer Cell* 2004;5:263–273.
76. Dupont WD, Page DL. Risk factors for breast cancer in women with proliferative breast disease. *N Engl J Med* 1985;312:146–151.
77. Kuukasjarvi T, Tanner M, Pennanen S, et al. Genetic changes in intraductal breast cancer detected by comparative genomic hybridization. *Am J Pathol* 1997;150: 1465–1471.
78. Waldman FM, DeVries S, Chew KL, et al. Chromosomal alterations in ductal carcinomas *in situ* and their in situ recurrences. *J Natl Cancer Inst* 2000;92:313–320.
79. Hager JH, Hodgson JG, Fridlyand J, et al. Oncogene expression and genetic background influence the frequency of DNA copy number abnormalities in mouse pancreatic islet cell carcinomas. *Cancer Res* 2004;64:2406–2410.
80. Sternlicht MD, Lochter A, Sympson CJ, et al. The stromal proteinase MMP3/ stromelysin-1 promotes mammary carcinogenesis. *Cell* 1999;98:137–146.
81. Rizki A, Weaver VM, Lee SY, et al. A human breast cell model of preinvasive to invasive transition. *Cancer Res* 2008;68:1378–1387.
82. Aubele M, Auer G, Braselmann H, et al. Chromosomal imbalances are associated with metastasis-free survival in breast cancer patients. *Anal Cell Pathol* 2002;24: 77–87.
83. Janssen EA, Baak JP, Guervos MA, et al. In lymph node-negative invasive breast carcinomas, specific chromosomal aberrations are strongly associated with high mitotic activity and predict outcome more accurately than grade, tumour diameter, and oestrogen receptor. *J Pathol* 2003;201:555–561.
84. Farabegoli F, Hermsen MA, Ceccarelli C, et al. Simultaneous chromosome 1q gain and 16q loss is associated with steroid receptor presence and low proliferation in breast carcinoma. *Mod Pathol* 2004;17:449–455.
85. Blegen H, Will JS, Ghadimi BM, et al. DNA amplifications and aneuploidy, high proliferative activity and impaired cell cycle control characterize breast carcinomas with poor prognosis. *Anal Cell Pathol* 2003;25:103–114.
86. Cheng KW, Lahad JP, Kuo WL, et al. The RAB25 small GTPase determines aggressiveness of ovarian and breast cancers. *Nat Med* 2004;10:1251–1256.
87. Slamon D, Clark G, Wong S, et al. Human breast cancer: Correlation of relapse and survival with amplification of the *HER-2/neu* oncogene. *Science* 1987;235: 177–182.
88. Al-Kuraya K, Novotny H, Bavi P, et al. HER2, TOP2A, CCND1, EGFR and C-MYC oncogene amplification in colorectal cancer. *J Clin Pathol* 2007;60:768–772.
89. Fantl V, Smith R, Brookes S, et al. Chromosome 11q13 abnormalities in human breast cancer. *Cancer Surv* 1993;18:77–94.
90. Escot C, Theillet C, Lidereau R, et al. Genetic alteration of the c-myc protooncogene (MYC) in human primary breast carcinomas. *Proc Natl Acad Sci U S A* 1986;83:4834–4838.
91. Collins C, Rommens JM, Kowbel D, et al. Positional cloning of ZNF217 and NABC1: genes amplified at 20q13.2 and overexpressed in breast carcinoma. *Proc Natl Acad Sci U S A* 1998;95:8703–8708.

92. Courjal F, Cuny M, Rodriguez C, et al. DNA amplifications at 20q13 and MDM2 define distinct subsets of evolved breast and ovarian tumours. *Br J Cancer* 1996; 74:1984–1989.
93. Sinclair CS, Adem C, Naderi A, et al. TBX2 is preferentially amplified in *BRCA1*- and *BRCA2*-related breast tumors. *Cancer Res* 2002;62:3587–3591.
94. Haverty PM, Fridlyand J, Li L, et al. High-resolution genomic and expression analyses of copy number alterations in breast tumors. *Genes Chromosomes Cancer* 2008;47:530–542.
95. Tennis M, Krishnan S, Bonner M, et al. p53 Mutation analysis in breast tumors by a DNA microarray method. *Cancer Epidemiol Biomarkers Prev* 2006;15:80–85.
96. Samuels Y, Wang Z, Bardelli A, et al. High frequency of mutations of the *PIK3CA* gene in human cancers. *Science* 2004;304:554.
97. Li J, Yen C, Liaw D, et al. PTEN, a putative protein tyrosine phosphatase gene mutated in human brain, breast, and prostate cancer. *Science* 1997;275: 1943–1947.
98. Futreal PA, Liu Q, Shattuck-Eidens D, et al. *BRCA1* mutations in primary breast and ovarian carcinomas. *Science* 1994;266:120–122.
99. Wooster R, Bignell G, Lancaster J, et al. Identification of the breast cancer susceptibility gene *BRCA2. Nature* 1995;378:789–792.
100. Ross JS, Fletcher JA. The *HER-2/neu* oncogene in breast cancer: prognostic fFactor, predictive factor, and target for therapy. *Oncologist* 1998;3:237–252.
101. Slamon D, Pegram M. Rationale for trastuzumab (Herceptin) in adjuvant breast cancer trials. *Semin Oncol* 2001;28:13–19.
102. Vogel CL, Cobleigh MA, Tripathy D, et al. First-line Herceptin monotherapy in metastatic breast cancer. *Oncology* 2001;61:37–42.
103. Moy B, Goss PE. Lapatinib: current status and future directions in breast cancer. *Oncologist* 2006;11:1047–1057.
104. Kallioniemi OP, Kallioniemi A, Kurisu W, et al. *ERBB2* amplification in breast cancer analyzed by fluorescence *in situ* hybridization. *Proc Natl Acad Sci U S A* 1992;89:5321–5325.
105. Heise C, Sampson-Johannes A, Williams A, et al. ONYX-015, an E1B gene-attenuated adenovirus, causes tumor-specific cytolysis and antitumoral efficacy that can be augmented by standard chemotherapeutic agents. *Nat Med* 1997;3: 639–645.
106. Wang H, Nan L, Yu D, et al. Antisense anti-MDM2 oligonucleotides as a novel therapeutic approach to human breast cancer: in vitro and in vivo activities and mechanisms. *Clin Cancer Res* 2001;7:3613–3624.
107. Burgess DJ, Doles J, Zender L, et al. Topoisomerase levels determine chemotherapy response *in vitro* and *in vivo. Proc Natl Acad Sci U S A* 2008;105:9053–9058.
108. Hennessy BT, Smith DL, Ram PT, et al. Exploiting the PI3K/AKT pathway for cancer drug discovery. *Nat Rev Drug Discov* 2005;4:988–1004.
109. Rottenberg S, Jaspers JE, Kersbergen A, et al. High sensitivity of BRCA1-deficient mammary tumors to the PARP inhibitor AZD2281 alone and in combination with platinum drugs. *Proc Natl Acad Sci U S A* 2008;105:17079–17084.
110. Neve RM, Chin K, Fridlyand J, et al. A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. *Cancer Cell* 2006;10:515–527.

111. Raphael BJ, Volik S, Yu P, et al. A sequence-based survey of the complex structural organization of tumor genomes. *Genome Biol* 2008;9:R59.

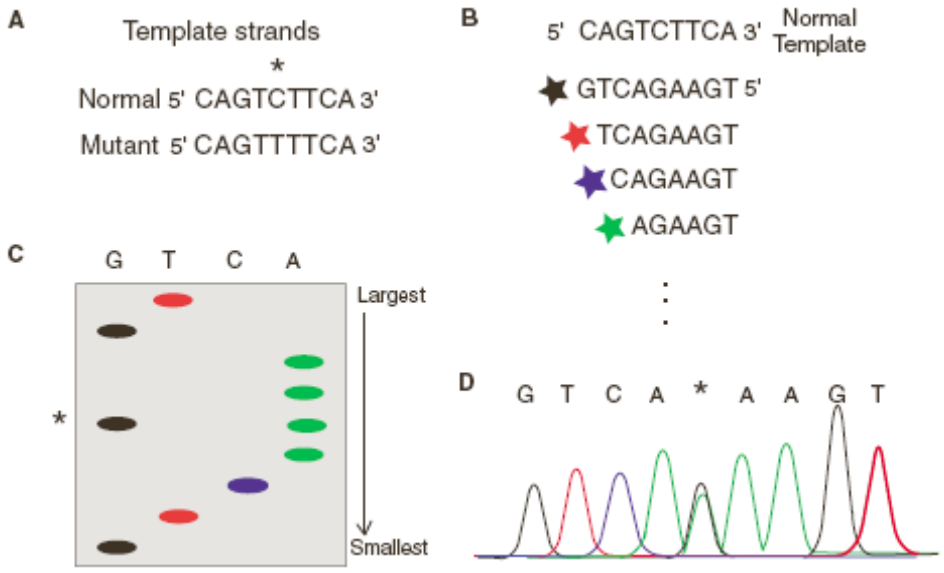


FIGURE 27.1. Sanger DNA sequencing using florescent chain terminators. A: Template DNA from tumor or normal cells is used to synthesize complementary molecules (B) that are terminated with differentially fluorescently labeled deoxyribose nucleotide triphosphates (dNTPs). Chain terminating nucleotides are present at a much lower concentration than normal nucleotides. If a chain terminator is incorporated no additional bases can be added and a DNA chain of a fixed length is created, tagging the molecule with a particular florescent dye that indicates a certain nucleotide. C: Chain terminated molecules are put in a gel-filled capillary, which allows smaller molecules to move through it more easily than larger molecules. In the presence of an electric field, DNA, which is negatively charged, migrates toward the positive electrode which is placed on the opposite end of the capillary from the side that is loaded. Before reaching the end of the capillary the molecules pass in front of a detector, which records the color of the fluorescence. D: Fluorescent spectra are analyzed so the DNA sequence can be interpreted; in the trace shown, the fifth base from the left is a mixture of G (*black*) and A (*green*) rather than a single peak of one color.

Table 27.1 MUTATION PREVALENCE IN THE KINASE SCREEN. EACH GENE IS RANKED BY ITS MUTATION FREQUENCY AND SELECTIVE PRESSURE. THE TOP 20 GENES ARE SHOWN

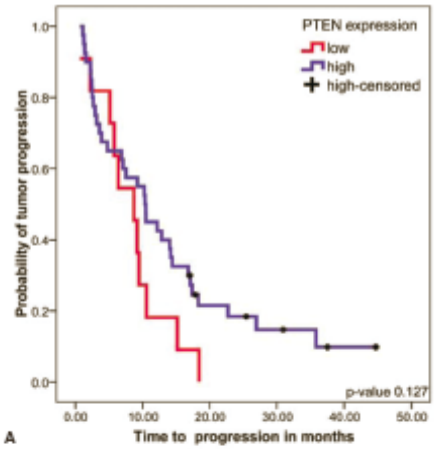
Gene	Ranking	Selection Pressure	Number of Nonsynonymous Mutations
TTN	1	2.0	63
BRAF	2	8.4	8
ATM	3	2.9	10
TAF1L	4	3.6	8
ERN1	5	4.5	6
MAP2K4	6	8.7	4
CHUK	7	5.4	5
FGFR2	8	5.1	5
NTRK3	9	4.8	5
MGC42105	10	7.1	4
TGFBR2	11	5.9	4
EPHA6	12	3.9	5
FLJ23074	13	5.4	4
ITK	14	4.9	4
DCAMKL3	15	4.7	4
STK11	16	7.2	3
PAK7	17	4.2	4
STK6	18	6.0	3
BRD2	19	3.8	4
RPS6KA2	20	3.7	4

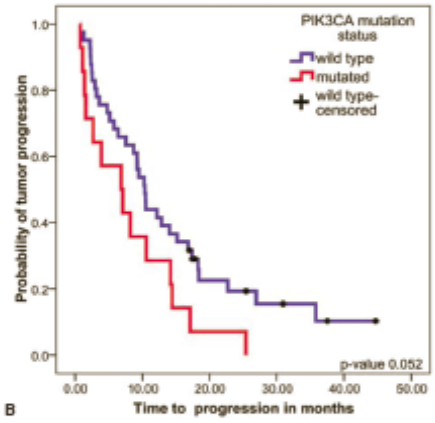
Adapted from Greenman C, Stephens P, Smith R, et al. Patterns of somatic mutation in human cancer genomes. *Nature* 2007;446:153–158, with permission.

Table 27.2 PATHWAYS WITH SIGNIFICANT MUTATIONS IN BREAST CANCER

Pathway	Observed Mutations (N)	Expected Mutations (N)	Significance Score
dsRNA induced gene expression	20	0.3	23.87
Interferon signaling	22	0.5	22.70
ATM signaling	25	0.8	22.37
BRCA1/2	24	1.4	15.68
Apoptosis signaling	20	0.8	16.13
Apoptosis induction	8	1.1	1.76
Fas and TNF	11	1.9	2.13

RNA, double-stranded RNA; TNF, tumor necrosis factor.
Adapted from Lin J, Gan CM, Zhang X, et al. A multidimensional analysis of genes mutated in breast and colorectal cancers. Genome Res 2007;17:1304–1318, with permission.





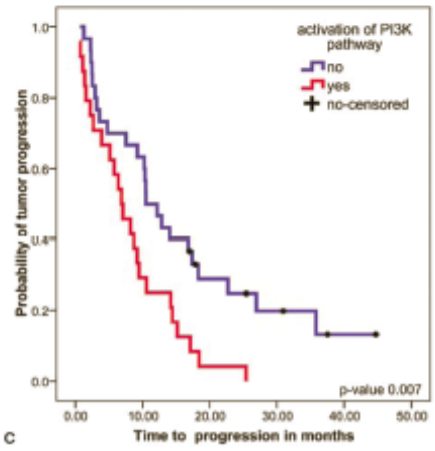


FIGURE 27.2. Mutations affecting breast cancer outcome. Mutations in either of two components of the PI3K pathway are more predictive than outcomes of individual analysis of mutations as assayed by Kaplan-Meier plots. A: Disruptions in *PTEN* activity or (B) mutations in *PIK3CA* are significantly less strongly associated with poor outcome than (C) the unified status of both *PIK3CA* and *PTEN*. (Adapted from Berns K, Horlings HM, Hennessy BT, et al. A functional genetic approach identifies the PI3K pathway as a major determinant of trastuzumab resistance in breast cancer. *Cancer Cell* 2007;12:395–402, with permission.)

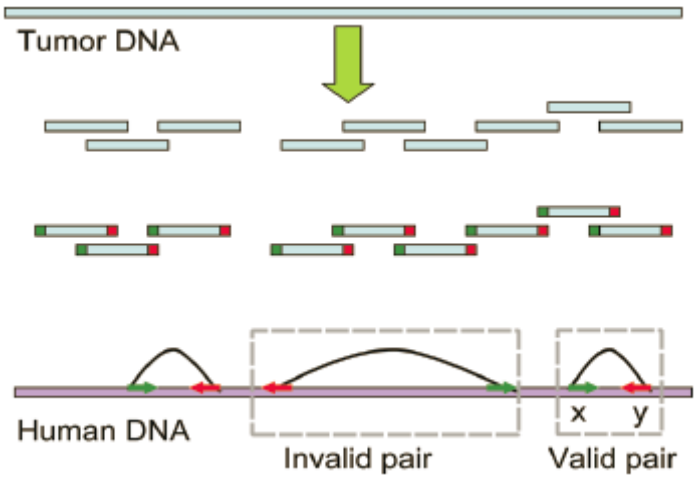


FIGURE 27.3. End sequence profiling (ESP). The principle of ESP is that given a piece of genomic DNA from within a known size distribution, the sequence of the two ends should map to the reference human genome in a particular orientation and within a certain distance. An invalid pair of sequences indicates the potential for a rearrangement. Practically, if more than one pair of sequences indicates the same aberration it has a high probability of being true. (Adapted from Raphael BJ, Volik S, Yu P, et al. A sequence-based survey of the complex structural organization of tumor genomes. *Genome Biol* 2008;9:R59, with permission.)

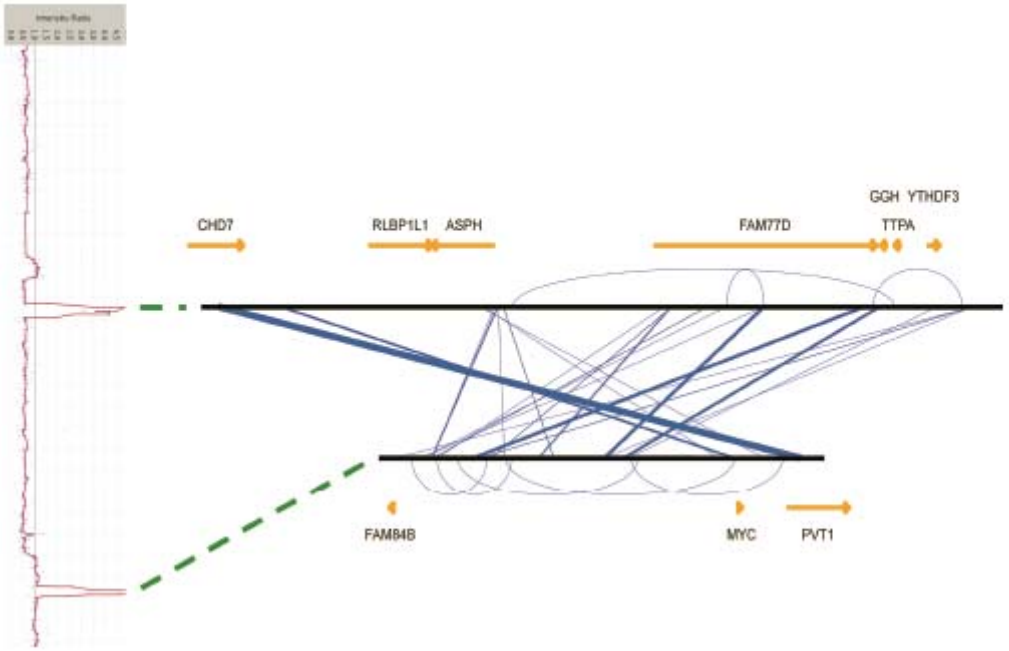


FIGURE 27.4. End sequence profiling (ESP) using high throughput, short read sequencing. Short reads aligned to the human genome estimate copy number and illustrate structural aberrations at the inter- and intra-chromosomal level. A plot of copy number reads that align to chromosome 8 from two regions of chromosome 8 NCI-H2171 are shown is graphed on the left. Intensity ratio corresponds to copy number estimates with two regions that appear at approximately 20 copies. These regions of the genome are shown in linear chromosomal order (x-axis) with the position of known genes drawn and with blue lines showing observed intra-chromosomal fusion events. Numerous events are observed to join various portions of these amplified regions. (Adapted from Campbell PJ, Stephens PJ, Pleasance ED, et al. Identification of somatically acquired rearrangements in cancer using genome-wide massively parallel paired-end sequencing. *Nat Genet* 2008;40:722–729, with permission.)

Table 27.3 COMMONLY USED COMMERCIAL ARRAY CGH PLATFORMS

Manufacturer	Array Element Size	Allele Specific?
Affymetrix http://www.affymetrix.com	24 mer	Yes
Illumina http://www.illumina.com	50 mer	Yes
Agilent http://www.chem.agilent.com	60 mer	No
Roche Nimblegen http://www.nimblegen.com	Flexible	Possible

CGH, comparative genomic hybridization.

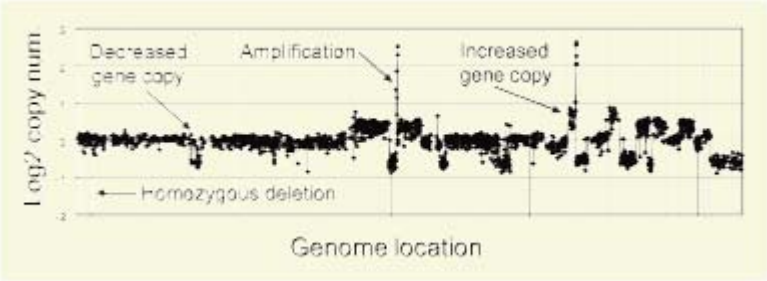


Figure 27.5. Genome copy number abnormalities measured using comparative genomic hybridization (CGH). Log 2 relative copy number is displayed along the genome with chromosome 1 to the left and chromosomes 22 and X to the right. *Vertical lines* show the chromosome boundaries. Relative copy number gains show as significant excursions above zero and losses show as significant excursions below zero.

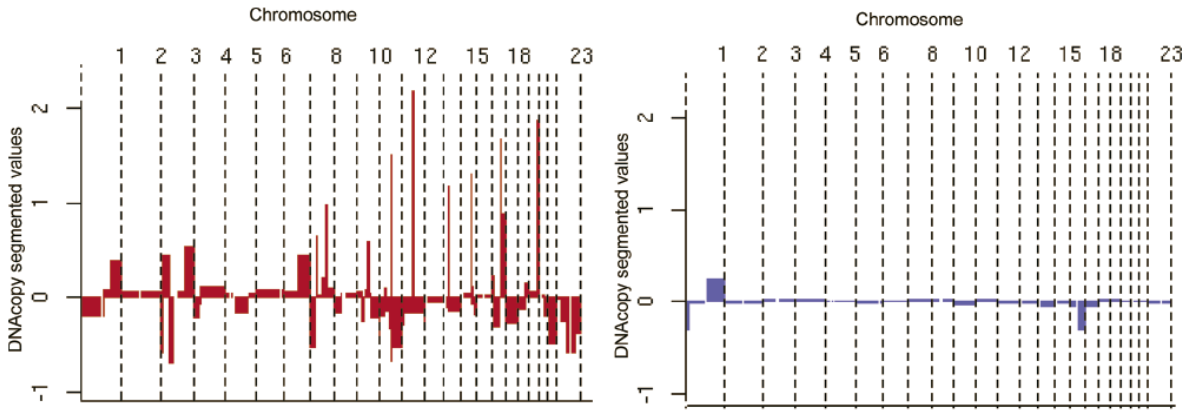


Figure 27.6. Comparative genomic hybridization (CGH) profiles measured for two clinically similar breast tumors after analysis using circular binary segmentation (113). Data are displayed as described in Figure 27.5. Relative copy number values for one tumor are displayed in *red* and the other in *blue.*

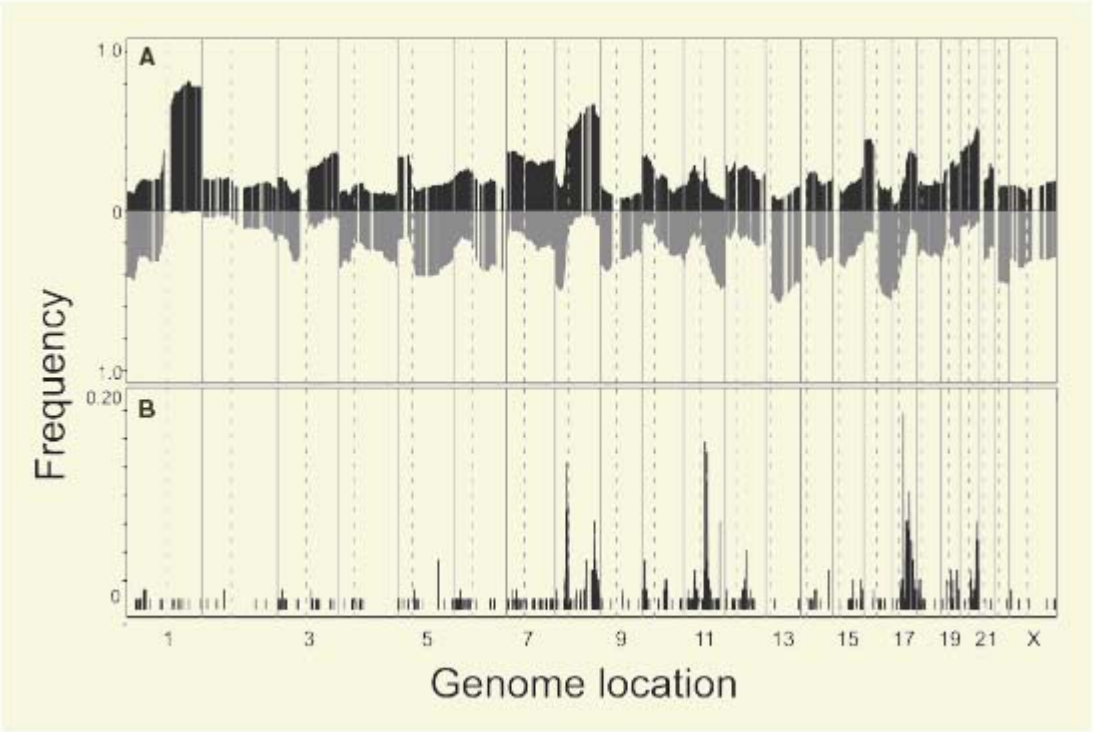


FIGURE 27.7. Recurrent genome copy number aberrations in breast cancer. Panel A. Frequencies of recurrent genome copy number changes in 145 breast cancers. Frequencies of copy number gains are plotted as positive values and frequencies of copy number losses are plotted as negative values. Panel B. Frequencies of amplification. Data from the same 145 tumors is plotted showing only the positions and frequencies of amplification events in the tumors (note the frequent amplifications at 8, 11, and 17). Data in both panels are plotted as a function of genome location as described in Figure 27.5. (Adapted from 46).

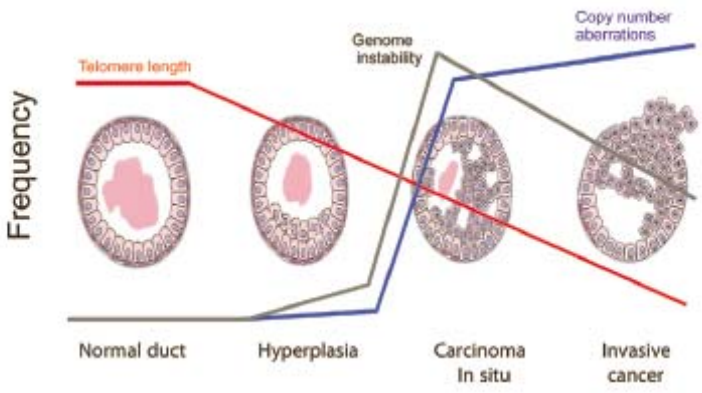


FIGURE 27.8. Schematic diagram of changes in number of genome copy number abnormalities measured using comparative genomic hybridization (CGH) (*blue*), genome instability measured using fluorescence *in situ* hybridization (FISH) (*black*) and relative telomere length (*red*). (Adapted from Chin K, de Solorzano CO, Knowles D, et al. In situ analyses of genome instability in breast cancer. *Nat Genet* 2004;36:984–988, with permission.)

LEGAL DISCLAIMER