Published in final edited form as:

Curr Opin Genet Dev. 2010 February ; 20(1): 4–14. doi:10.1016/j.gde.2009.11.005.

Breast cancer genomes — form and function

James Korkola and Joe W Gray

Life Sciences Division, Lawrence Berkeley National Laboratory, One Cyclotron Road, MS977-250, Berkeley, CA 94127, United States

Abstract

This review summarizes advances in our understanding of the genomic and epigenomic abnormalities in breast cancers that are being revealed by the increasingly powerful suite of genomic analysis technologies. It summarizes the remarkable genomic heterogeneity that characterizes the disease, describes mechanisms that shape cancer genomes as they evolve toward metastasis, summarizes important recurrent aberrations that exist in spite of the genomic chaos and that contribute to breast cancer pathophysiology, and describes the use of preclinical models to identify drugs that will be effective against subsets of breast cancers carrying specific genomic and epigenomic abnormalities.

Introduction

Cancer is now recognized as disease in which abnormalities in the genome and epigenome accumulate as a result of exposure to endogenous and exogenous damaging agents thereby enabling cells to escape normal regulatory controls. The set of accumulated abnormalities in the tumor and in the microenvironment in which the tumors reside determine the course of the disease including propensity to invade and metastasize and to respond to treatment. A central goal of modern genomic projects like The Cancer Genome Atlas (TCGA; <http://cancergenome.nih.gov/>) Project [1] and the International Cancer Genome Consortium (ICGC; <http://www.icgc.org/>) is to determine the spectrum of genomic aberrations that make up important human cancer types and to understand how these aberrations contribute to the pathophysiology of these cancers. The technologies being employed by these and related projects are providing an increasingly detailed view of breast cancer genomes. Today's comparative genomic hybridization (CGH) platforms provide genome-wide maps of genome copy number abnormalities with subgene resolution [2]. In some cases the measurements are allele specific. Fluorescence *in situ* hybridization (FISH) provides information about copy number and structural heterogeneity including variation between cells and in histological context [3]. Moreover, probes are available to interrogate essentially any part of the genome (<http://bacpac.chori.org/>). Genome sequencing is now providing detailed information about mutations and structural aberrations. Massively parallel sequencing platforms [4•] have now made practical, analysis of hundreds to thousands of genes in individual tumor genomes and comprehensive analyses of entire breast cancer genomes are beginning to emerge [5•]. Analysis of promoter methylation and of DNA sequences immunoprecipitated with antibodies against specific chromatin marks provides information on epigenomic events that regulate gene expression [6,7].

Genome chaos

The picture of breast cancer genomes that emerges from large-scale studies is one of remarkable complexity and variability. Individual tumors often carry aberrations that deregulate hundreds

or even thousands of genes. Genome copy number changes and epigenomic modifications seem to deregulate the largest number of genes. Figure 1 shows genome copy number abnormalities for three primary breast tumors. These show ‘strongly selected’ events such as high-level gene amplification and homozygous deletion as well as abnormalities that change genome copy number by only one or a few copies. In general, the strongly selected regions are narrow in genomic extent and involved only a few genes while low level genome copy number abnormalities (CNAs) that span larger regions of the genome up to and including whole chromosome losses or gains may affect thousands of genes. These data also show the remarkable differences in copy number abnormality structure that occur between clinically similar tumors. Although CGH profiles of breast cancers appear complicated, analyses of cancer genomes using DNA sequence-based methods such as end sequence profiling [8] or paired end sequencing [9•] show that the genome copy number abnormality maps yield an overly simple view of the events that conspire to change copy number. Figure 2, for example, shows results from a DNA sequence analysis of a ~150 kbp sequence from a region of amplification nominally from chromosome 20 [8]. Remarkably, sequences from several normally separate parts of the genome have been ‘stitched together’ in the cancer genome and amplified together as a cassette — perhaps as a result of fusions that occur during a bridge breakage fusion process or during extrachromosomal fusion and amplification. Genome sequencing efforts are revealing large numbers of structural rearrangements in breast cancers. So far, however, important recurrent structural aberrations like the Ets-family fusions found in prostate cancer [10•] have not been discovered in breast cancer. Breast cancer genomes are further deregulated through mutational processes. In general, the number of genes deregulated via mutations is considerably less than through other mechanisms — typically tens of mutations per tumor [11].

Epigenomic events including promoter methylation, histone acetylation, methylation, and phosphorylation and chromatin remodeling further contribute to cancer pathophysiology [7, 12]. Genome-wide analyses of promoter regions show that promoter CpG island hypermethylation is a mechanism by which tumor suppressor genes are inactivated [13,14] and is an especially important early event in breast cancer development [15–17]. Indeed, this appears to be a more common mechanism of tumor suppressor gene inactivation than mutation or homozygous deletion. The importance of epigenomic events as early events in breast cancer is noteworthy considering that the epigenome is subject to environmental-related and treatment-related modification [18•] and may even reflect environmental influences that are passed between generations [19].

Recurrent abnormalities

Although the genomes of clinically similar breast cancers are remarkably different, some regions of the genome are recurrently aberrant. Figure 3, for example, shows the locations of significant genome copy number gains and losses across the genome measured using array CGH [20]. The number of genes mapping to regions of significant abnormality is remarkable. In fact, correlative analyses of genome copy number and gene expression indicate that 10–15% of the entire genome is deregulated by recurrent genome copy number abnormalities [20,21]. Analysis of CGH data suggests the existence of at least three general classes — those with ‘simple’ genomes displaying relatively few genome aberrations, those with ‘complex’ genomes in displaying many aberrations and those displaying high-level amplification [20, 22]. It is likely that these subtypes may be caused by differences in DNA repair defects. For example, complex genomes typically carry p53 mutations while simple genomes do not. CGH profiles for tumors from BRCA1 carriers also display high genome complexity although the spectrum of recurrent abnormalities differs between sporadic and heritable breast cancers. In particular, tumors from BRCA1 mutation carriers typically display recurrent aberrations involving chromosomes 3 and 5 that sporadic tumors do not have [23]. Relatively few recurrent

mutations have been found in breast cancers. High frequency somatic mutations (i.e. present in >3% of reported cases) include PIK3CA, TP53, CDH1, CDKN2A, and AKT1 (<http://www.sanger.ac.uk/genetics/CGP/cosmic>). However, almost two hundred of others have been reported as recurrently mutated but at lower frequency [11]. Prevalent germ line mutations that contribute to breast cancer genesis involve BRCA1, BRCA2, CHEK1, and TP53.

The mechanisms by which low level copy number abnormalities contribute to cancer pathophysiology have not yet been determined definitively. However, we have suggested that these deregulate transcription of large numbers of genes that collectively increase metabolic activity [20]. The contributions of regions of amplification at 8p11–12, 8q24, 11q13, 12q13, 17q11–12, 17q21–24 and 20q13 as somewhat better understood since each encodes a known oncogene (e.g. *FGFR1*, *MYC*, *CCND1*, *MDM2*, *ERBB2*, *PS6K* and *ZNF217*, respectively). However, substantial evidence is now emerging that demonstrates that multiple genes in each region of amplification contribute to the pathophysiology of the disease. Functionally important co-amplified genes demonstrated so far include *LSM1*, *BAG4*, and *C8orf4* at 8p11–12 [24]; *MYC* and *PVT1* at 8q24 [25]; *CCND1* and *EMSY* at 11q13 [26]; and *ERBB2* and *GRB7* at 17q21 [27]. Recent studies also suggest that amplifications of regions that are well separated in the normal genome are not independent events. For example, coamplification of regions at 8p11–8p12 and 11q12–11q14 [28] or 8q24 and 17q21 [29] have been reported to contribute collectively to breast cancer pathophysiology. These interactions may be just the tip of the iceberg since dozens of genes have been identified as over expressed as a result of recurrent, high-level amplification and most have not been functionally assessed [20]. Coamplification of genes that are normally separate along the genome is plausible considering the fact that genes normally separate in the genome may be joined and contiguous in cancer genomes as a result of structural rearrangements [8]. However, this remains to be investigated. Recurrent mutations, although relatively infrequent in breast cancer, provide particularly strong evidence of functional importance. Bringing all of these observations together, recent integrative analyses of homozygous deletions, high-level amplification and recurrent mutations suggest that genomic deregulation of pathways involving *ERBB2*, *EGFR*, and *PI3K* signaling and DNA topology is particularly important in breast cancer genesis and progression [30].

Epigenomic alterations also contribute to breast cancer pathophysiology [19,31]. Targeted and genome wide analyses of promoter region methylation in breast cancer have shown that genes with known or suspected tumor suppressor function including *BRCA1*, *CCND2*, *CDKN2A*, *RAR β* , and *RASSF1A* [32,33] and members of the *WNT* signaling pathway [34] are recurrently down-regulated by hypermethylation. Consistent with this, genes found to be recurrently mutated in large-scale genome analysis studies are frequently targets of hypermethylation and hypermethylation is usually mutually exclusive from genetic changes [13]. At the chromatin level, modification via acetylation and/or phosphorylation directly modifies estrogen receptor α and other steroid hormone receptors superfamily in breast cancer thereby modifying ligand sensitivity and hormone antagonist responses [35]. Chromatin remodeling also has been implicated in growth and metastasis. In particular, *SATB1* has been reported to be a genome organizer that tethers multiple genomic loci and recruits chromatin-remodeling enzymes to regulate chromatin structure and gene expression. Overexpression of this ‘master chromatin regulator’ appears to drive breast cancer progression by upregulating metastasis-associated genes and down-regulating tumor-suppressor genes [36].

Shaping breast cancer genomes

Breast cancers accumulate a remarkable number of genomic aberrations as they develop the cancer hallmarks that eventually enable metastatic spread [37]. The wide variation in number and type of abnormalities that occurs between clinically similar tumors is understandable considering the evolutionary forces that each tumor experiences. It is useful to consider the

shaping of cancer genomes as the result of overcoming checks and balances that have evolved over time to protect against cancer [38]. Several possible barriers are illustrated schematically in Figure 4 and include DNA repair, differentiation, senescence, immune surveillance, limited angiogenesis, apoptotic surveillance, cell cycle regulation, and anticancer treatments. These are depicted separately but of course they interact considerably — for example, DNA repair is tightly linked to cell cycle regulation. Nonetheless, considering them as separate barriers is a convenient way of beginning to parse studies of cancer genome evolution. Since the ‘efficiency’ of each of these barriers is likely influenced by the individual genotype and by the microenvironment of the evolving cancer it is easy to see how individual tumors evolve such different genomic and epigenomic features.

Recent mouse genetic experiments comparing cancer incidence in crosses between cancer prone and cancer resistant mouse strains are revealing some of the important barriers to cancer progression that directly influence individual genotypes. Such experiments already have identified transcriptional networks including mitotic checkpoints, inflammation, apoptosis, and epidermal barrier function whose activity likely determines barrier efficiency [39•]. Genetically engineered mouse models of cancer also clearly demonstrate that the genetic background significantly influences the genome aberration spectrum of the resulting tumors [40]. Combined genetic and genomic studies of mouse cancer models and human tumors show that tumors selectively amplify cancer susceptibility alleles thereby establishing a direct link between tumor genome aberration spectrum and individual genotype [41]. Finally, it is telling that the recurrent genomic aberrations that arise in tumors in BRCA1 carriers are significantly different than those that arise in noncarriers. In short, the individual genetic background plays a major role in determining cancer susceptibility and in shaping cancer genomes [42].

A wealth of evidence now indicates that the extracellular matrix composition and organization strongly influences breast tumorigenesis [43,44] and that the extracellular environment and the tumor genome coevolve during progression. Key microenvironmental factors that contribute to cancer progression include microenvironmental remodeling via matrix metalloproteinases [45–47], macrophage-mediated metastasis [48], senescence-associated cytokine production [49], changes in tissue architecture [50], estrogen-related stromal–epithelial cell crosstalk [51], and TGF- β signaling [52]. Most evidence suggests that the microenvironment coevolves with breast tumors via epigenomic modifications [44] although some evidence supports the possibility that clonal genomic aberrations in breast cancer-associated stromal cells also may play a role [53•,54]. The latter possibility remains controversial [55].

A model of breast cancer progression

Breast cancers progress through multiple genomic and epigenomic steps. However, the evolutionary process is unlikely to be the result of a linear and more or less constant rate of successive genomic and epigenomic aberration accumulation. It seems more likely that cancer progression varies between individuals and over time within an individual. The breast progenitor cell in which the tumor arises likely determines the spectrum of aberrations needed to enable progression. Evidence for this comes from studies showing that the spectrum of genomic aberrations accumulated is strongly influenced by the normal progenitor cell type in which the tumor arises [20,21]. In general, the events associated with early aspects of breast cancer appear to be epigenomic in nature, wherein stepwise DNA methylation changes enable escape of telomerase negative epithelial cells from proliferation barriers [56•]. This leads to proliferation in the absence of telomerase and culminates in a period of high genome instability owing to checkpoint deregulation [57,58] and/or entry into telomere crisis when telomeres become critically short [3,59•]. This barrier is highly effective but not perfectly so. As a consequence, most cells become genomically unstable and die. The increase in genome instability during breast cancer progression is illustrated in Figure 5. Rarely, a single cell

accumulates genomic or epigenomic alterations that reactivate telomerase and confer a proliferative advantage. This cell might be considered the tumor initiation cell and will have multiple characteristics that appear 'stem cell like'. The extent to which this is related to normal stem cells remains unclear. However, it is likely that the genomic characteristics of this cell — both transcriptional and genomic — will be reflected in subsequent progeny. This may explain why tumor genomes appear to evolve relatively slowly after telomere crisis and why metastases that develop years after immortalization usually retain the genomic characteristics of the primary tumor from which they were derived [60]. That said, not all breast tumors progress this way. A recent analysis of tumor progression termed Sector-Ploidy-Profiling (SPP) demonstrates that breast tumor evolution may evolve as a single major clonal subpopulation or as multiple clonal subpopulations [61]. The latter model may explain why a small percentage of metastatic cancers do not resemble the primary tumor from which they were derived.

Figure 5 also shows that the cells that survive telomere crisis remain genomic unstable. Thus, tumors are likely to consist of a large number of cells that are not faithful genomic representations of the tumor-initiating cell and are likely to be biologically compromised. These cells are likely to be much more sensitive to treatment than the tumor initiating cells from which they were derived. This may partially explain why breast tumors initially respond well to treatment but fail to exhibit a durable response.

Clinical associations with breast cancer subtypes

Molecular profiling technology promises to have a profound effect on the classification of breast cancer. Pioneering molecular profiling studies on invasive ductal carcinoma of the breast have recognized the presence of distinct breast cancer classes, based on expression patterns of the 'intrinsic gene set', a large panel of genes that shows heterogeneous expression across entire tumor cohorts [62–64]. The tumors within these distinct clusters are defined by expression patterns that are similar to cells within the normal breast architecture (i.e. basal versus luminal clusters), or have known genomic aberrations (ERBB2/HER2 positive), or most closely resemble normal breast ('normal-like' breast cancer). Some of these clusters can be further subdivided on the basis of additional expression features (e.g. luminal A and luminal B). Patients within the distinct subclasses have significantly different clinical outcomes; basal and ERBB2/HER2 patients have a poorer prognosis than their luminal A and normal subclass counterparts [63,64]. The most common histopathologic types of breast cancer, namely invasive ductal carcinoma and invasive lobular carcinoma, can be classified as any of these five subtypes [65]. In contrast, most rare histologic types of breast cancer belong exclusively to single molecular subclasses, such as medullary and metaplastic (basal) and mucinous and neuroendocrine (lobular) [65]. The classification of tumors into basal, lobular, and HER2 subgroups reflects the molecular characteristics of the disease, an advantage over those studies that simply develop predictive gene sets for outcome prediction. Since these initial studies, larger cohorts of tumors have been profiled on more comprehensive array platforms, leading to the refinement of the classification and the identification of specific genes that are definitive of each of the classes.

In keeping with their relatively good prognosis, the luminal class of tumors tends to be ER positive and express genes that are consistent with this hormone receptor status (*ESR1*, *PGR*, and *GATA3*) [63,66]. At the genomic level, luminal tumors have gain of 1q and 16p and loss of 16q and frequent high-level amplification of 8p11–12, 11q13–14, 12q13–14, 17q11–12, 17q21–24, and 20q13 (luminal A), while luminal B is characterized by gains of chromosomes 1q, 8q, 17q, and 20q and losses involving portions of 1p, 8p, 13q, 16q, 17p, and 22q, with frequent high-level amplification of 8p11–12, two regions of 8q, and 11q13–14 [20,21].

Similarly, the HER2 tumors are defined by the high expression of *ERBB2* and *GRB7*, both of which lie within the 17q amplicon that is almost universally amplified in this type of breast cancer. *ERBB2* tumors are relatively enriched for increased copy number at 1q, 7p, 8q, 16p, and 20q and reduced copy number at 1p, 8p, 13q, and 18q, with high levels of amplification of 17q as expected [20,21].

Basal-like tumors are ER negative, and are defined by the expression of genes associated with both basal epithelium and myoepithelium (e.g. *CK5* and *CK17*) [62–64]. Basal tumors are relatively enriched for gains involving 3q, 8q, and 10p and losses involving 3p, 4p, 4q, 5q, 12q, 13q, 14q, and 15q [20,21]. Interestingly, high-level amplification at any locus is infrequent in these tumors. Basal tumors are aggressive, highly proliferative, and usually present as high-grade tumors, but the basis of this aggressive phenotype is not fully understood. Furthermore, it has become clear that basal tumors share many characteristics with breast tumors derived from patients with *BRCA1* mutations [67,68]. In keeping with this observation, basal tumors often lose expression of *BRCA1* through mechanisms such as promoter hypermethylation. Finally, many but not all of the basal tumors are exemplified by the lack of expression of hormone receptors (ER and PR) and HER2, forming the ‘triple negative’ type of breast cancer [67,69–71]. Although once thought to be exclusively basal in origin, ‘triple negative’ breast tumors are now recognized to include a subset of normal-like breast tumors [67,69]. Triple negative tumors are more common in African-American than Caucasian women, and are associated with shorter duration survival [71]. Recent studies suggest that subsets of aggressive tumors are also defined by stem cell like (CD44 high, CD24 low) and claudin low genotypes as well that may overlap but be distinct from traditional basal subtypes [72].

Subclass specific therapeutic response *in vivo*

Given the heterogeneity of expression and genomic composition within the distinct subclasses, it might also be expected that there will be distinct clinical responses to chemotherapeutic and targeted agents within the different subclasses. If agents can be identified that show enhanced efficacy against specific subclasses, this might represent the first step toward more personalized medicine for the treatment of breast cancer. Such class distinctions have already been identified, in particular for HER2 positive tumors. Targeted therapeutic agents such as the monoclonal antibody trastuzumab or the small molecule tyrosine kinase inhibitor lapatinib have both been approved for use in treating patients with HER2 positive breast cancer, leading to improved survival in these patients [73–75]. Given that HER2-directed therapies were developed before the identification of breast cancer subclasses, it is now of interest to see if additional subclass specific agents can be identified. Recent studies have demonstrated that inhibitors of poly (ADP-ribose) polymerase (PARP), an important transducer of *BRCA*-mediated DNA damage response, selectively target *BRCA1* and *BRCA2* mutant breast cancer [76]. This is a form of synthetic lethality, since only cells with aberrant *BRCA* function will be affected. Indeed, the side effects in patients treated with PARP inhibitors have been much milder than most traditional chemotherapeutic agents. However, when patients with *BRCA* mutant tumors have been treated with PARP inhibitors, the result is prolonged patient survival [76]. Given the similarity between *BRCA1* mutant and the basal subclass of tumors, it is hoped that basal tumors will also demonstrate increased sensitivity to PARP inhibitors. One caveat to this approach has been the observation that *BRCA* may be dispensable for the growth of established breast cancers. This is evidenced by the findings that *BRCA2* mutant tumors develop secondary resistance to DNA damaging agents through mutations that restore normal *BRCA* function and DNA damage response [77,78]. With the restoration of the DNA damage response, these tumors are less sensitive to the actions of DNA damaging agents such as cisplatin and carboplatin. Basal subtype tumors have also been shown to be exquisitely sensitive to dose dense regimens of epirubicin–cyclophosphamide combinations [79]. Interestingly, of the 10 basal type tumors within the cohort, 9 of them showed complete

response. Thus, these results demonstrate that both targeted and conventional chemotherapies are likely to demonstrate subclass specific efficacy.

Linking the genome to treatment response — a quest for predictive markers

The successes of genome marker targeted therapies described above have motivated efforts to identify other response-genome aberration associations that might guide the development of additional breast cancer treatments. This is complicated by the complexity of cancer genomes and the large and growing numbers of candidate anticancer agents that should be considered for use in the treatment. In fact, a 2009 survey by the pharmaceutical industry indicates that more than 800 experimental therapies are now being evaluated clinically including 106 in breast cancer (http://www.phrma.org/files/09-046PhRMACancer09_0331.pdf). Identification of genomic and epigenomic features that predict response to these agents is challenging in the context of clinical trials because of the difficulty of obtaining tumor samples during trials, the cost of comprehensive molecular profiling and the relatively large size of clinical trials needed to establish solid predictive markers. An alternative is to use preclinical models to identify associations between genome aberrations and response. The challenge in this approach is to develop models that reflect the range of genome and epigenome aberrations that influence response. Collections of breast cancer cell lines grown *in vitro* and as xenografts are now being used for this purpose as are collections of genetically engineered mouse models.

Breast cancer cell line models

The use of collections of cell lines to identify molecular features associated with therapeutic response began with an effort by the NCI Developmental Therapeutics Program (<http://dtp.nci.nih.gov/index.html>) to assess responses to potential anticancer agents in a collection of 60 diverse cancer cell lines (the NCI-60). Over the past several decades, the concentrations of drug needed to inhibit growth by a predetermined amount were determined for more than 40 000 agents. Recently, the cell lines have been characterized in substantial molecular depth [80] so that the cell responses can be compared with molecular features to identify response-associated features that might be developed into predictive molecular markers that can be tested clinically. Several studies have illustrated the promise of this approach [80–82].

One problem with the NCI-60 cell line panel for the assessment of specific responses in breast cancer is the lack of representation of several of the major breast cancer subtypes in the cell line collection. To overcome that limitation, we and others have focused on testing drugs intended for breast cancer treatment in large panels of breast cancer cell lines selected to capture the molecular features that determines clinical responses. Analysis of genome copy number and mutations in commonly available breast cancer cell lines showed that the recurrent aberrations in the cell lines were similar to those found in collections of primary breast tumors [83,84]. However, analysis of the transcriptional features showed that the collection represented only three of the transcriptional subtypes seen clinically. These were designated basal B (now redefined as claudin low [72]) and basal A (now redefined as basal) and a single luminal subtype [83,84].

The relative ease of analysis in cell line panels allows efficient identification of candidate molecular predictors of response to new agents and combinations thereof that warrant clinical evaluation. This approach is supported by studies of breast cancer cell line panels that clearly show the association between ERBB2 amplification and response to trastuzumab [85] and lapatinib [86] that has been observed clinically. Other predictive markers suggested by cell line panel studies include:

- p27 relocalization as a predictor of response to trastuzumab [84].

- Sensitivity of basal subtype cells to MEK inhibitors leading to reactivation of PI3K signaling [87]. This study suggests the utility of combinations of MEK and PI3K inhibitors.
- Enhanced sensitivity to PI3K and AKT inhibitors in cells with HER2 amplification, PIK3CA mutation, or PTEN deletion [88–90].
- Association of Ras mutations with resistance to the actions of PI3K inhibitors that combination treatments with PI3K and MEK inhibitors can overcome [89].
- Luminal subtype specificity of a CDK4/6 inhibitor [91,92].

Of course, cell line collections are necessarily imperfect models since they model only aspects of response that are determined by factors that are intrinsic to the tumors from which they were derived. Clearly, they cannot accurately identify predictors of response to agents that inhibit extrinsic cancer hallmarks such as angiogenesis, immune surveillance or metastatic spread. Their utility also is affected by the failure to accurately model interactions with the microenvironment [50]. Culturing cells in different microenvironments can overcome this limitation to some extent but that can never be accomplished perfectly. Indeed, the microenvironments in which tumor cells reside vary within individuals — from *in situ* disease to diverse metastatic sites. Thus, drugs that are effective in a broad range of molecular environments and molecular markers that predict response independent of microenvironment may be most likely to predict durable clinical response. Studies of the role of the microenvironment on drug response may be facilitated by the recent development of microenvironment microarrays that allow the effects of growing cells in many different microenvironments to be assessed [93•].

Mouse models of breast subtypes

Genetically engineered mouse (GEM) tumor models are important preclinical adjuncts to cell line models. Historically, the utility of GEM models in predictive marker development has been limited because they did not model the genetic diversity found in human tumors. However, this is changing thanks to the substantial increase in GEM tumor model diversity because of work in the NCI Mouse Models of Human Cancer Consortium (MMHCC) and throughout the mouse model community (<http://cancermodels.nci.nih.gov/>). A recent comparison of transcriptional characteristics of 13 GEM mouse models with human tumors showed that ‘many of the defining characteristics of human subtypes were conserved among the mouse models’ [94]. Several interesting mammary models have been reported since that study. A first example is a GEM with MMTV driven HER2 overexpression combined with cre-recombinase mediated PTEN deletion [95] to model the luminal type of HER2 positive breast cancer. Tumors that arise in these mice are multifocal and metastatic, with hyperactivation of the PI3K pathway. Given that PTEN deficiency in HER2 positive tumors is associated with trastuzumab resistance, these mice are likely to represent a valuable resource to identify therapeutic combinations that are likely to overcome this resistance. A second example is a GEM with mammary specific expression of mutant Kras to model basal-like breast cancer. These mice are characterized by the overexpression of *Igf1r* [96]. Treatment of tumor bearing mice with an IGF1R inhibitor has been shown to result in shrinkage of the tumor mass. Xenografts of the *KRAS* mutant cell line MDAMB231 into mice has shown that MDAMB231 is also highly sensitive to the IGF1R inhibitor, further reinforcing the basal-like phenotype of this mouse model.

Conclusions and the future

The power of modern genome analysis tools — especially massively parallel sequencing — is increasing at a remarkable rate. It seems clear that the ‘thousand dollar genome’ is nearly at

hand. Thus it seems reasonable to expect that comprehensive tumor scans of genomic and epigenomic aberrations will become an important and routine part of cancer management. The challenges will be to manage the information so that it is generally accessible to the scientific and medical communities and to interpret it in ways that lead to improved cancer management. This will require a substantial investment in functional studies since the roles of most genomic and epigenomic aberrations in cancer pathophysiology are not understood. It will also require full development and deployment of large-scale information management and interpretation systems. Success in these areas will allow the potential of cancer genome and epigenome in treatment personalization to be fully realized.

Acknowledgments

This work was supported by the Director, Office of Science, Office of Biological & Environmental Research, of the U.S. Department of Energy under contract No. DE-AC02-05CH11231, by the National Institutes of Health, National Cancer Institute grants U54 112970, and P50 CA 58207.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
 - • of outstanding interest
1. Collins FS, Barker AD. Mapping the cancer genome. Pinpointing the genes involved in cancer will help chart a new course across the complex landscape of human malignancies. *Sci Am* 2007;296:50–57. [PubMed: 17348159]
 2. Chin L, Gray JW. Translating insights from the cancer genome into clinical practice. *Nature* 2008;452:553–563. [PubMed: 18385729]
 3. Chin K, et al. In situ analyses of genome instability in breast cancer. *Nat Genet* 2004;36:984–988. [PubMed: 15300252]
 4. Schuster SC. Next-generation sequencing transforms today's biology. *Nat Methods* 2008;5:16–18. [PubMed: 18165802] This review summarizes recent developments in massively parallel, single molecule sequencing that are revolutionizing cancer genome analysis.
 5. Shah SP, et al. Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution. *Nature* 2009;461:809–813. [PubMed: 19812674] A DNA sequence level analysis of the mutational changes that occur during progression of a lobular breast cancer.
 6. Collas P. The state-of-the-art of chromatin immunoprecipitation. *Methods Mol Biol* 2009;567:1–25. [PubMed: 19588082]
 7. Jones PA, Baylin SB. The epigenomics of cancer. *Cell* 2007;128:683–692. [PubMed: 17320506]
 8. Volik S, et al. End-sequence profiling: sequence-based analysis of aberrant genomes. *Proc Natl Acad Sci U S A* 2003;100:7696–7701. [PubMed: 12788976]
 9. Fullwood MJ, et al. Next-generation DNA sequencing of paired-end tags (PET) for transcriptome and genome analyses. *Genome Res* 2009;19:521–532. [PubMed: 19339662] A massively parallel nucleic acid sequencing strategy that allows comprehensive analysis of mutations, copy number abnormalities, and structural changes in cancer genomes.
 10. Kumar-Sinha C, Tomlins SA, Chinnaiyan AM. Recurrent gene fusions in prostate cancer. *Nat Rev Cancer* 2008;8:497–511. [PubMed: 18563191] This paper reviews recurrent Ets-family gene fusions that contribute to the genesis of the majority of prostate cancers.
 11. Sjoblom T, et al. The consensus coding sequences of human breast and colorectal cancers. *Science* 2006;314:268–274. [PubMed: 16959974]
 12. Zhang Y, et al. The structure and function of histone deacetylases: the target for anti-cancer therapy. *Curr Med Chem* 2008;15:2840–2849. [PubMed: 18991639]
 13. Chan TA, et al. Convergence of mutation and epigenetic alterations identifies common genes in cancer that predict for poor prognosis. *PLoS Med* 2008;5:e114. [PubMed: 18507500]

14. Brooks J, Cairns P, Zeleniuch-Jacquotte A. Promoter methylation and the detection of breast cancer. *Cancer Causes Control* 2009;20:1539–1550. [PubMed: 19768562]
15. Agrawal A, Murphy RF, Agrawal DK. DNA methylation in breast and colorectal cancers. *Mod Pathol* 2007;20:711–721. [PubMed: 17464311]
16. Bean GR, 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. [PubMed: 18006786]
17. Crawford YG, et al. Histologically normal human mammary epithelia with silenced p16(INK4a) overexpress COX-2, promoting a premalignant program. *Cancer Cell* 2004;5:263–273. [PubMed: 15050918]
18. Dolinoy DC, Jirtle RL. Environmental epigenomics in human health and disease. *Environ Mol Mutagen* 2008;49:4–8. [PubMed: 18172876] This review summarizes epigenomic mechanisms by which cancer progression can be influenced by environmental factors.
19. Das R, Hampton DD, Jirtle RL. Imprinting evolution and human health. *Mamm Genome*. 2009
20. Chin K, et al. Genomic and transcriptional aberrations linked to breast cancer pathophysiologies. *Cancer Cell* 2006;10:529–541. [PubMed: 17157792]
21. Bergamaschi A, et al. Distinct patterns of DNA copy number alteration are associated with different clinicopathological features and gene-expression subtypes of breast cancer. *Genes Chromosomes Cancer* 2006;45:1033–1040. [PubMed: 16897746]
22. Fridlyand J, et al. Breast tumor copy number aberration phenotypes and genomic instability. *BMC Cancer* 2006;6:96. [PubMed: 16620391]
23. Wessels LF, 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. [PubMed: 12460933]
24. Yang ZQ, et al. Multiple interacting oncogenes on the 8p11–p12 amplicon in human breast cancer. *Cancer Res* 2006;66:11632–11643. [PubMed: 17178857]
25. Guan Y, et al. Amplification of PVT1 contributes to the pathophysiology of ovarian and breast cancer. *Clin Cancer Res* 2007;13:5745–5755. [PubMed: 17908964]
26. Brown LA, et al. Co-amplification of CCND1 and EMSY is associated with an adverse outcome in ER-positive tamoxifen-treated breast cancers. *Breast Cancer Res Treat*. 2009
27. Hodgson JG, 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. [PubMed: 16266989]
28. Kwek SS, et al. Co-amplified genes at 8p12 and 11q13 in breast tumors cooperate with two major pathways in oncogenesis. *Oncogene* 2009;28:1892–1903. [PubMed: 19330026]
29. Park K, et al. c-myc amplification is associated with HER2 amplification and closely linked with cell proliferation in tissue microarray of nonselected breast cancers. *Hum Pathol* 2005;36:634–639. [PubMed: 16021569]
30. Leary RJ, et al. Integrated analysis of homozygous deletions, focal amplifications, and sequence alterations in breast and colorectal cancers. *Proc Natl Acad Sci U S A* 2008;105:16224–16229. [PubMed: 18852474]
31. Lo PK, Sukumar S. Epigenomics and breast cancer. *Pharmacogenomics* 2008;9:1879–1902. [PubMed: 19072646]
32. Kenemans P, Verstraeten RA, Verheijen RH. Oncogenic pathways in hereditary and sporadic breast cancer. *Maturitas* 2008;61:141–150. [PubMed: 19434886]
33. Zhu W, et al. Quantitative evaluation of DNA hypermethylation in malignant and benign breast tissue and fluids. *Int J Cancer*. 2009
34. Klarmann GJ, Decker A, Farrar WL. Epigenetic gene silencing in the Wnt pathway in breast cancer. *Epigenetics* 2008;3:59–63. [PubMed: 18398311]
35. Fu M, et al. Acetylation of nuclear receptors in cellular growth and apoptosis. *Biochem Pharmacol* 2004;68:1199–1208. [PubMed: 15313417]
36. Han HJ, et al. SATB1 reprogrammes gene expression to promote breast tumour growth and metastasis. *Nature* 2008;452:187–193. [PubMed: 18337816]

37. Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell* 2000;100:57–70. [PubMed: 10647931]
38. Gatenby RA, Gillies RJ. A microenvironmental model of carcinogenesis. *Nat Rev Cancer* 2008;8:56–61. [PubMed: 18059462]
39. Quigley DA, et al. Genetic architecture of mouse skin inflammation and tumour susceptibility. *Nature* 2009;458:505–508. [PubMed: 19136944] This paper describes a mouse genetics approach to the identification of biological processes that contribute to cancer genesis and progression.
40. Hager JH, 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. [PubMed: 15059892]
41. Ewart-Toland A, et al. Identification of Stk6/STK15 as a candidate low-penetrance tumor-susceptibility gene in mouse and human. *Nat Genet* 2003;34:403–412. [PubMed: 12881723]
42. van Beers EH, 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. [PubMed: 15705879]
43. Kass L, et al. Mammary epithelial cell: influence of extracellular matrix composition and organization during development and tumorigenesis. *Int J Biochem Cell Biol* 2007;39:1987–1994. [PubMed: 17719831]
44. Spencer VA, Xu R, Bissell MJ. Extracellular matrix, nuclear and chromatin structure, and gene expression in normal tissues and malignant tumors: a work in progress. *Adv Cancer Res* 2007;97:275–294. [PubMed: 17419950]
45. Eck SM, et al. Matrix metalloproteinase and G protein coupled receptors: co-conspirators in the pathogenesis of autoimmune disease and cancer. *J Autoimmun.* 2009
46. Ghajar CM, Bissell MJ. Extra cellular matrixcontrol of mammary gland morphogenesis and tumorigenesis: insights from imaging. *Histochem Cell Biol* 2008;130:1105–1118. [PubMed: 19009245]
47. Sternlicht MD, et al. The stromal proteinase MMP3/stromelysin-1 promotes mammary carcinogenesis. *Cell* 1999;98:137–146. [PubMed: 10428026]
48. DeNardo DG, et al. CD4(+) T cells regulate pulmonary metastasis of mammary carcinomas by enhancing protumor properties of macrophages. *Cancer Cell* 2009;16:91–102. [PubMed: 19647220]
49. Coppe JP, et al. Senescence-associated secretory phenotypes reveal cell-nonautonomous functions of oncogenic RAS and the p53 tumor suppressor. *PLoS Biol* 2008;6:2853–2868. [PubMed: 19053174]
50. Nelson CM, Bissell MJ. Of extracellular matrix, scaffolds, and signaling: tissue architecture regulates development, homeostasis, and cancer. *Annu Rev Cell Dev Biol* 2006;22:287–309. [PubMed: 16824016]
51. Yamaguchi Y, Hayashi S. Estrogen-related cancer microenvironment of breast carcinoma. *Endocr J* 2009;56:1–7. [PubMed: 18497452]
52. Barcellos-Hoff MH, Akhurst RJ. Transforming growth factor-beta in breast cancer: too much, too late. *Breast Cancer Res* 2009;11:202. [PubMed: 19291273]
53. Trimboli AJ, et al. Pten in stromal fibroblasts suppresses mammary epithelial tumours. *Nature* 2009;461:1084–1091. [PubMed: 19847259] This paper describes a possible mechanism by which genomic aberrations in the stroma can influence the development of breast cancer.
54. Deng G, et al. Loss of heterozygosity in normal tissue adjacent to breast carcinomas. *Science* 1996;274:2057–2059. [PubMed: 8953032]
55. Weinberg RA. Coevolution in the tumor microenvironment. *Nat Genet* 2008;40:494–495. [PubMed: 18443582]
56. Novak P, et al. Stepwise DNA methylation changes are linked to escape from defined proliferation barriers and mammary epithelial cell immortalization. *Cancer Res* 2009;69:5251–5258. [PubMed: 19509227] This paper elucidates early epigenomic changes that enable escape from regulatory controls and enable progression toward breast cancer.
57. Berman H, 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. [PubMed: 16869768]

58. McDermott KM, et al. p16(INK4a) prevents centrosome dysfunction and genomic instability in primary cells. *PLoS Biol* 2006;4:e51. [PubMed: 16464125]
59. Artandi SE, Depinho RA. Telomeres and telomerase in cancer. *Carcinogenesis*. 2009 This paper summarizes current understanding of the roles of telomeres and telomerase as barriers to cancer progression.
60. Kuukasjarvi T, et al. Genetic heterogeneity and clonal evolution underlying development of asynchronous metastasis in human breast cancer. *Cancer Res* 1997;57:1597–1604. [PubMed: 9108466]
61. Navin N, et al. Inferring tumor progression from genomic heterogeneity. *Genome Res*. 2009
62. Perou CM, et al. Molecular portraits of human breast tumours. *Nature* 2000;406:747–752. [PubMed: 10963602]
63. Sorlie T, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A* 2001;98:10869–10874. [PubMed: 11553815]
64. Sorlie T, et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc Natl Acad Sci U S A* 2003;100:8418–8423. [PubMed: 12829800]
65. Weigelt B, et al. Refinement of breast cancer classification by molecular characterization of histological special types. *J Pathol* 2008;216:141–150. [PubMed: 18720457]
66. Kouros-Mehr H, et al. GATA-3 maintains the differentiation of the luminal cell fate in the mammary gland. *Cell* 2006;127:1041–1055. [PubMed: 17129787]
67. Rakha EA, Reis-Filho JS, Ellis IO. Basal-like breast cancer: a critical review. *J Clin Oncol* 2008;26:2568–2581. [PubMed: 18487574]
68. Turner NC, Reis-Filho JS. Basal-like breast cancer and the BRCA1 phenotype. *Oncogene* 2006;25:5846–5853. [PubMed: 16998499]
69. Reis-Filho JS, Tutt AN. Triple negative tumours: a critical review. *Histopathology* 2008;52:108–118. [PubMed: 18171422]
70. Thike AA, et al. Triple-negative breast cancer: clinicopathological characteristics and relationship with basal-like breast cancer. *Mod Pathol*. 2009
71. Cleator S, Heller W, Coombes RC. Triple-negative breast cancer: therapeutic options. *Lancet Oncol* 2007;8:235–244. [PubMed: 17329194]
72. Hennessy BT, et al. Characterization of a naturally occurring breast cancer subset enriched in epithelial-to-mesenchymal transition and stem cell characteristics. *Cancer Res* 2009;69:4116–4124. [PubMed: 19435916]
73. Cameron DA, Stein S. Drug insight: intracellular inhibitors of HER2 — clinical development of lapatinib in breast cancer. *Nat Clin Pract Oncol* 2008;5:512–520. [PubMed: 18594499]
74. Geyer CE, et al. Lapatinib plus capecitabine for HER2-positive advanced breast cancer. *N Engl J Med* 2006;355:2733–2743. [PubMed: 17192538]
75. Slamon DJ, et al. Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. *N Engl J Med* 2001;344:783–792. [PubMed: 11248153]
76. Fong PC, et al. Inhibition of poly(ADP-ribose) polymerase in tumors from BRCA mutation carriers. *N Engl J Med* 2009;361:123–134. [PubMed: 19553641]
77. Edwards SL, et al. Resistance to therapy caused by intragenic deletion in BRCA2. *Nature* 2008;451:1111–1115. [PubMed: 18264088]
78. Sakai W, et al. Secondary mutations as a mechanism of cisplatin resistance in BRCA2-mutated cancers. *Nature* 2008;451:1116–1120. [PubMed: 18264087]
79. Bertheau P, et al. Exquisite sensitivity of TP53 mutant and basal breast cancers to a dose-dense epirubicin–cyclophosphamide regimen. *PLoS Med* 2007;4:e90. [PubMed: 17388661]
80. Bussey KJ, et al. Integrating data on DNA copy number with gene expression levels and drug sensitivities in the NCI-60 cell line panel. *Mol Cancer Ther* 2006;5:853–867. [PubMed: 16648555]
81. Scherf U, et al. A gene expression database for the molecular pharmacology of cancer. *Nat Genet* 2000;24:236–244. [PubMed: 10700175]
82. Mori S, et al. Utilization of genomic signatures to identify phenotype-specific drugs. *PLoS One* 2009;4:e6772. [PubMed: 19714244]

83. Kao J, et al. Molecular profiling of breast cancer cell lines defines relevant tumor models and provides a resource for cancer gene discovery. *PLoS One* 2009;4:e6146. [PubMed: 19582160]
84. Neve RM, et al. A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. *Cancer Cell* 2006;10:515–527. [PubMed: 17157791]
85. Lewis GD, et al. Growth regulation of human breast and ovarian tumor cells by heregulin: evidence for the requirement of ErbB2 as a critical component in mediating heregulin responsiveness. *Cancer Res* 1996;56:1457–1465. [PubMed: 8640840]
86. Konecny GE, et al. Activity of the dual kinase inhibitor lapatinib (GW572016) against HER-2-overexpressing and trastuzumab-treated breast cancer cells. *Cancer Res* 2006;66:1630–1639. [PubMed: 16452222]
87. Mirzoeva OK, et al. Basal subtype and MAPK/ERK kinase (MEK)-phosphoinositide 3-kinase feedback signaling determine susceptibility of breast cancer cells to MEK inhibition. *Cancer Res* 2009;69:565–572. [PubMed: 19147570]
88. She QB, et al. Breast tumor cells with PI3K mutation or HER2 amplification are selectively addicted to Akt signaling. *PLoS One* 2008;3:e3065. [PubMed: 18725974]
89. Torbett NE, et al. A chemical screen in diverse breast cancer cell lines reveals genetic enhancers and suppressors of sensitivity to PI3K isoform-selective inhibition. *Biochem J* 2008;415:97–110. [PubMed: 18498248]
90. Michelucci A, et al. PIK3CA in breast carcinoma: a mutational analysis of sporadic and hereditary cases. *Diagn Mol Pathol*. 2009
91. Finn RS, et al. PD 0332991, a selective cyclin D kinase 4/6 inhibitor, preferentially inhibits proliferation of luminal estrogen receptor-positive human breast cancer cell lines in vitro. *Breast Cancer Res* 2009;11:R77. [PubMed: 19874578]
92. Finn RS, et al. Dasatinib, an orally active small molecule inhibitor of both the src and abl kinases, selectively inhibits growth of basal-type/“triple-negative” breast cancer cell lines growing in vitro. *Breast Cancer Res Treat* 2007;105:319–326. [PubMed: 17268817]
93. LaBarge MA, et al. Human mammary progenitor cell fate decisions are products of interactions with combinatorial microenvironments. *Integr Biol* 2009;1:70–79. This paper describes microenvironment microarrays that facilitate investigations of the behavior of normal and tumor cells when grown in many different microenvironments.
94. Herschkowitz JI, et al. Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. *Genome Biol* 2007;8:R76. [PubMed: 17493263]
95. Schade B, et al. PTEN deficiency in a luminal ErbB-2 mouse model results in dramatic acceleration of mammary tumorigenesis and metastasis. *J Biol Chem* 2009;284:19018–19026. [PubMed: 19435886]
96. Klinakis A, et al. Igf1r as a therapeutic target in a mouse model of basal-like breast cancer. *Proc Natl Acad Sci U S A* 2009;106:2359–2364. [PubMed: 19174523]
97. Beroukhi R, et al. Assessing the significance of chromosomal aberrations in cancer: methodology and application to glioma. *Proc Natl Acad Sci U S A* 2007;104:20007–20012. [PubMed: 18077431]

Figure 1. Genome copy number profiles measured for three primary breast tumors using comparative genomic hybridization (CGH). Log 10 relative copy number is displayed as a function of distance along the genome with chromosome 1 to the left and chromes 22 and X to the right.

Figure 2.

Analysis of a breast cancer using end sequence profiling (ESP [8]). Panel (a) Genome sequence of a 150 kbp genome segment cloned from a region of amplification at chromosome 20q13. Panel (b) DNA sequences at regions that join normally separate genomic sequences.

Figure 3. Array CGH analysis of 143 primary human breast cancers [20]. Panel (a) Frequencies of gains (positive values) and losses (negative values) displayed as a function of distance along the genome. Panel (b) Genomic identification of significant targets in cancer (GISTIC; log 10 (FDR values)) [97]. This statistic emphasizes regions of high-level amplification or homozygous deletion.

Figure 4. Schematic representation of barriers to cancer progression that shape cancer genome evolution. The influences of the microenvironment and individual genetic composition are illustrated (adapted from [38]).

Figure 5. Genome instability measured during breast cancer progression using FISH [3]. Histological sections depicting stages of evolution are illustrated above. Bivariate measures of genome copy number at the centromere of chromosome 1 and at chromosome 20q13. The onset of instability at the DCIS stage of evolution is apparent.