



Published in final edited form as:

Breast. 2013 August 1; 22 Suppl 2: S19–S21. doi:10.1016/j.breast.2013.07.003.

Mutational analysis of breast cancer: Guiding personalized treatments

Matthew J. Ellis

Division of Medical Oncology Section of Breast Oncology Washington University School of Medicine Siteman Cancer Center 660 South Euclid Ave, CB 8069 St Louis, MO, 63110 Ph: 314-362-8903 Fax: 314-747-9320 mellis@dom.wustl.edu

Summary

The application of high throughput techniques to profile DNA, RNA and protein in breast cancer samples from hundreds of patients has profoundly increased our knowledge of the disease. However there remain many knowledge gaps that will require a long process of extended clinical correlation studies, deeper integrated 'omic analysis and functional annotation to address. This article reviews conclusions from recent breast cancer 'omics profiling' papers and considers pathways forward for extracting medically valuable information from large dimension data sets [1–6].

Cancer genome sequencing – broad principles

Prior to large scale cancer sequencing analysis extensive data on mRNA expression, gene copy number aberration (CNA) and small scale Sanger-based sequencing has already been extensively mined for prognostic biomarkers and therapeutic targets. Massively parallel sequencing has brought a new level of genomic resolution and when conducted at the whole genome level, provides an unbiased catalog of somatic mutations of all classes [7]. The somatic changes (mutations) are classified as signal nucleotide variants (SNV), indels (small *insertions* or *deletions*) and structural variants (SV), which include translocations, large deletions and intra-chromosomal inversions. These events can occur in coding sequence (Tier 1), regulatory sequence (Tier 2), other mappable sequence (Tier 3) and repetitive sequences where mapping is ambiguous (Tier 4). Naturally the current focus is on Tier 1, but it is likely that cancer relevant events are occurring in additional areas of the genome since recent ENCODE project results has disproved the concept that much of the genome is inactive [8]. Since whole genome analysis remains a relatively expensive proposition, exome sequencing is the basis for many recent papers, but this technique will miss aberrations that begin or end in non-coding sequence, or occur in RNA genes, or occur in regulatory sequence. For example recurrent mutations in the *lncRNA* MALAT1 in luminal type breast cancer was discovered by whole genome sequencing [4]. Extraordinary degrees of complexity have been revealed by cancer genome sequencing studies. This heterogeneity comes in two forms. *Intratumoral* heterogeneity is the concept that tumors are multiclonal

© 2013 Elsevier Ltd. All rights reserved.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Conflicts of Interest:

ME has received funding from University Genomics, Bioclassifier LLC. He is a consultant for Novartis and Astra Zeneca. PAM50 Patent and licensing to Nanostring.

and not all mutations are in all cells in the tumor. *Intertumoral* heterogeneity is the concept that essentially every tumor has taken its own unique pathway to becoming malignant. Both represent serious challenges for clinical interpretation and the clinical utility of sequencing based diagnostics.

Significantly mutated genes

The initial step in the analysis of cancer genome sequenced cohort is to conduct a Significantly Mutated Gene (SMG) test. SMG are genes that accumulate missense, nonsense, and small deletions or insertions at a rate that is above what would be expected by chance, and therefore are likely to be mutational events that drive the disease process. The significance tests assumes a uniform background mutation rate across the genome (which is unlikely to be true) and adjusts for gene size, as clearly larger genes have a greater chance of accumulating random cancer-irrelevant mutations than smaller ones. There is also a correction for multiple testing. The published SMG patterns in breast cancer consist mostly of scattered mutations in tumor suppressor genes, many of which were known, but a few were novel such as MAP3K1, a stress kinase [4]. Gain-of-function mutations, with the exception of PIK3CA, are rare. PIK3CA mutations are commonest in luminal type breast cancer as previously described in Sanger-based targeted sequencing analyses [9], and are the target of a number of trials that combine endocrine therapy with a PIK3CA pathway inhibitor and an endocrine agent [10]. In contrast, for poor prognosis ER- breast cancers, TP53 mutation is the commonest event. SMG analysis is an insensitive tool and does not identify relatively uncommon, but druggable mutations. This was dramatically illustrated by the finding that HER2 is mutated in about 1.6% of HER2 normal breast cancer [11]. Whilst this is considered rare, and as a result, not identified by SMG testing, breast cancer is so common that a 2% incidence still produces an estimated 4000 cases a year, comparable to other diseases routinely treated with tyrosine kinase inhibitors such as chronic myeloid leukemia. Clustered HER2 mutations in both the extracellular domain and the kinase domain were shown to be activating by the Washington University Breast Cancer Program. Thus functional annotation is a critical orthogonal step that can efficiently identify druggable mutations. A multiple institution clinical trial targeting the HER2 mutant metastatic population with neratinib is underway as some of the observed mutations, including the commonest (L755S) are resistant to lapatinib [11]. Screening for patients whose tumors harbor HER2 mutations remains a challenge as upwards of 50 patients must be screened to find one case for the trial. However, patients are increasingly gaining access to clinical cancer sequencing and therefore may present to study sites with a HER2 mutation already diagnosed. Ideally a publically funded central sequencing faculty should be made available to the clinical research community so trials of this type can be efficiently executed.

Gene Copy Number and Structural Variation

Massively parallel sequencing also generates data on copy number variation (CNV) because the technology is digital. There are more “read counts” from regions of gene copy gain and fewer from regions of loss. Copy number aberrations (CNV) deliver additional therapeutic targets, but the patterns are highly complex and sometimes hard to interpret from a functional perspective because many genes maybe deregulated in regions of gain or loss that can span millions of base pairs [5]. The TCGA is currently combining DNA and RNA sequencing to analyze breast cancer amplicons in greater detail. Finally whole genome analysis has revealed that many breast cancers harbor complex chromosomal aberrations; some involve complex criss-cross translocation patterns that span who chromosomal locations. These have been referred to as “chromotrypsis” events and are thought to result from a cell recovering from a catastrophic cell division event where multiple chromosomal breaks occurred in a single cell cycle[12]. How this process transforms cells is far from clear

and the clinical associations of tumors that harbor large numbers of rearrangements of this type is only just begun to be explored.

Relating newly discovered somatic mutations with breast cancer outcomes

The wide variation in outcome experienced by breast cancer patients has been a fertile question for biomarker research for many years. Mutational analysis has at least three contributions to make in this area. First, since mutations imply causality there are greater direct mechanistic implications when a mutation is found to be prognostic or predictive, rather than, for example, a multigene expression profile that tracks a biological parameter like cell cycle activity. Second, it is possible that mutation status may add orthogonal data to gene expression profiling and improve prognostic models, perhaps particularly for intermediate risk groups where mRNA profiling has not provided an answer. Third, it may be possible to identify rare clones by massively parallel sequencing that carry a higher risk of progression and death than the majority clone which is providing a misleading benign prognosis [13]. The study of these questions is of course statistically challenging and will take large numbers of patients and carefully annotated data sets. A particular problem is that many SMG occur at relatively low frequencies – 10% or less and so many cases have to be screened to provide an adequate mutation positive population to compare with the majority mutation negative population. This issue is nicely illustrated by the sequencing analysis of the Z1031 neoadjuvant aromatase inhibitor trial. Unlike other recent cancer genome sequencing reports, this project focused on a clinical trial, with deep sample annotation relevant to aromatase inhibitor response, including Ki67 changes induced by treatment, PAM50-based definitions of intrinsic subtype, grade, information on a prognostic model called the preoperative endocrine prognostic index (PEPI), which integrates surgical stage (T and N) with Ki67 values and ER status of the surgical specimen that is obtained after 4 to 5 months of treatment [14]. To discover the mutational landscape of luminal type breast cancer seventy seven pretreatment samples were subjected to whole or partial genome sequencing. All SMG (and some potential therapeutic targets) were taken forward into targeted sequencing assays. Significant interactions were observed between trial endpoints and binary mutation status (mutant versus wildtype) in three instructive cases that were consistent in both a discovery set and a test set. TP53 was a driver for high proliferation, high grade and luminal B status and MAP3K1 loss of function mutations had the opposite association. While the TP53 result was predictable, a role for the MAP3K1 stress kinase in breast cancer was not previously known, but was consistent with reports of LOH in its substrate in the JNK pathway MAP2K4 [5], which was also mutant in the Z1031 series. GATA3 mutation was associated with a greater decrease in Ki67 than GATA3 wildtype cases suggesting that GATA3 status might be useful for prediction of endocrine therapy response [4].

A Sequencing based Therapeutic Road Map

Despite the complexity of breast cancer genomes a therapeutic and biological map of luminal-type breast cancer is emerging based on the functions of the genes that have been found to be mutant [15]. Clearly plasma membrane tyrosine kinases beyond HER2 are emerging as drivers, particularly FGFR1 amplification and perhaps other FGFR family members [16]. Phosphoinositol-3-kinase signaling is currently our best example of a druggable pathway in breast cancer, because of the efficacy of everolimus in endocrine therapy refractory ER positive disease [17]. However, the somatic genomics in the PI3 kinase pathway are complex, with multiple mutations occurring in both positive regulators (PIK3CA and AKT1) as well as negative regulators (PTEN and PIK3R1) [18]. This complicates both sequencing-based diagnostics and molecular pharmacology. It seems inherently likely that targeting a PTEN null tumor will require a broad spectrum PI3 kinase

inhibitor as multiple PI3K catalytic subunits are activated [19–21]. In contrast, a PIK3CA mutant tumor might well respond to a PIK3CA restricted inhibitor since the activation event is selective [22]. Treating PIK3R1 mutant and AKT1 mutant tumors may be subject to similar principles. Other important areas of therapeutics include CDK inhibition in Cyclin D1 amplified/CDK4 activated disease [23, 24] and inhibitors of histone modifications in the setting of a growing list of mutations in histone trimethylases such as MLL3 [15]. Another promising area for clinical investigations is MDM2 inhibition in luminal-type TP53 wild-type breast cancer. MDM2 gene amplification occurs in up to one third of Luminal B tumors and acts to accelerate TP53 degradation, creating a TP53 null state. Pharmacological inhibition of MDM2 restores TP53 activity, triggering cell cycle arrest and apoptosis [25].

Integrated ‘Omics Analysis the new cutting edge

One of the more promising approaches to deciphering cancer genomes is deeper integration of cancer proteomics in our spectrum of analytical approaches. Advanced pathway based analysis, such as PARADIGM, proved to be a useful informatics approach that successfully interprets genomics data at a functional level [26]. For example, this analysis suggested that MALAT1 mutation is associated with poor outcome clinical features in ER+ disease [4]. However, the active signaling networks (activomes) are inferred from a priori knowledge of protein function and signaling pathways and not from directly observed biochemistry. The technical step currently being addressed is how to achieve deep analysis of proteins and post translational modifications to complement data generated by cancer genome and transcriptome sequencing. This will produce a more complete picture of the activomes present in each individual tumor.

Conclusions

In conclusion, the combined knowledge base provided by recent next generation sequencing studies are unprecedented but it will take many years before all the therapeutic hypotheses raised by this vast data repository will be addressed. Nonetheless new therapeutic road maps are emerging and the opportunities in luminal-type breast cancer are particularly compelling. A new treatment paradigm is therefore evolving whereby deep genomic analysis will drive treatment decisions based on a pharmacopeia of cell-type and pathway-matched therapies. Deeper annotation using next generation proteomic approaches is likely to contribute to this conversation as the technology improves and informatics approaches are in place to deal with multiple-tiered integrated omics analysis.

Acknowledgments

Dr Ellis is supported by a Susan G. Komen for the Cure - Promise Grant, BCRP-Idea Award-BC112014, NIH/NCI U24 CA160035, NIH/NCI R01 CA095614, NIH/NCI P30 CA91842, NIH/NCI U10 CA077440 and the Breast Cancer Research Foundation

References

1. Shah SP, Roth A, Goya R, et al. The clonal and mutational evolution spectrum of primary triple-negative breast cancers. *Nature*. 2012; 486:395–399. [PubMed: 22495314]
2. Stephens PJ, Tarpey PS, Davies H, et al. The landscape of cancer genes and mutational processes in breast cancer. *Nature*. 2012; 486:400–404. [PubMed: 22722201]
3. Banerji S, Cibulskis K, Rangel-Escareno C, et al. Sequence analysis of mutations and translocations across breast cancer subtypes. *Nature*. 2012; 486:405–409. [PubMed: 22722202]
4. Ellis MJ, Ding L, Shen D, et al. Whole-genome analysis informs breast cancer response to aromatase inhibition. *Nature*. 2012; 486:353–360. [PubMed: 22722193]

5. Curtis C, Shah SP, Chin SF, et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature*. 2012; 486:346–352. [PubMed: 22522925]
6. Koboldt DC, Fulton RS, McLellan MD, et al. Comprehensive molecular portraits of human breast tumours. *Nature*. 2012
7. Walter MJ, Graubert TA, Dipersio JF, et al. Next-generation sequencing of cancer genomes: back to the future. *Personalized medicine*. 2009; 6:653. [PubMed: 20161678]
8. Dunham I, Kundaje A, Aldred SF, et al. An integrated encyclopedia of DNA elements in the human genome. *Nature*. 2012; 489:57–74. [PubMed: 22955616]
9. Samuels Y, Wang Z, Bardelli A, et al. High frequency of mutations of the PIK3CA gene in human cancers. *Science*. 2004; 304:554. [PubMed: 15016963]
10. Sanchez CG, Ma CX, Crowder RJ, et al. Preclinical modeling of combined phosphatidylinositol-3-kinase inhibition with endocrine therapy for estrogen receptor-positive breast cancer. *Breast cancer research : BCR*. 2011; 13:R21. [PubMed: 21362200]
11. Bose R, Kavuri SM, Searleman AC, et al. Activating HER2 Mutations in HER2 Gene Amplification Negative Breast Cancer. *Cancer discovery*. 2012
12. Stephens PJ, Greenman CD, Fu B, et al. Massive genomic rearrangement acquired in a single catastrophic event during cancer development. *Cell*. 2011; 144:27–40. [PubMed: 21215367]
13. Ding L, Ellis MJ, Li S, et al. Genome remodelling in a basal-like breast cancer metastasis and xenograft. *Nature*. 2010; 464:999–1005. [PubMed: 20393555]
14. Ellis MJ, Suman VJ, Hoog J, et al. Randomized phase II neoadjuvant comparison between letrozole, anastrozole, and exemestane for postmenopausal women with estrogen receptor-rich stage 2 to 3 breast cancer: clinical and biomarker outcomes and predictive value of the baseline PAM50-based intrinsic subtype--ACOSOG Z1031. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2011; 29:2342–2349. [PubMed: 21555689]
15. Ellis MJ, Perou CM. The genomic landscape of breast cancer as a therapeutic roadmap. *Cancer discovery*. 2013; 3:27–34. [PubMed: 23319768]
16. Turner N, Pearson A, Sharpe R, et al. FGFR1 amplification drives endocrine therapy resistance and is a therapeutic target in breast cancer. *Cancer research*. 2010; 70:2085–2094. [PubMed: 20179196]
17. Baselga J, Campone M, Piccart M, et al. Everolimus in postmenopausal hormonereceptor-positive advanced breast cancer. *The New England journal of medicine*. 2012; 366:520–529. [PubMed: 22149876]
18. Comprehensive molecular portraits of human breast tumours. *Nature*. 2012; 490:61–70. [PubMed: 23000897]
19. Crowder RJ, Phommaly C, Tao Y, et al. PIK3CA and PIK3CB inhibition produce synthetic lethality when combined with estrogen deprivation in estrogen receptor-positive breast cancer. *Cancer research*. 2009; 69:3955–3962. [PubMed: 19366795]
20. Carvalho S, Milanezi F, Costa JL, et al. PIKING the right isoform: the emergent role of the p110beta subunit in breast cancer. *Virchows Archiv : an international journal of pathology*. 2010; 456:235–243. [PubMed: 20130907]
21. Ni J, Liu Q, Xie S, et al. Functional characterization of an isoform-selective inhibitor of PI3K-p110beta as a potential anticancer agent. *Cancer discovery*. 2012; 2:425–433. [PubMed: 22588880]
22. Jamieson S, Flanagan JU, Kolekar S, et al. A drug targeting only p110alpha can block phosphoinositide 3-kinase signalling and tumour growth in certain cell types. *The Biochemical journal*. 2011; 438:53–62. [PubMed: 21668414]
23. Miller TW, Balko JM, Fox EM, et al. ERalpha-dependent E2F transcription can mediate resistance to estrogen deprivation in human breast cancer. *Cancer discovery*. 2011; 1:338–351. [PubMed: 22049316]
24. Roberts PJ, Bisi JE, Strum JC, et al. Multiple roles of cyclin-dependent kinase 4/6 inhibitors in cancer therapy. *Journal of the National Cancer Institute*. 2012; 104:476–487. [PubMed: 22302033]
25. Zhuang C, Miao Z, Zhu L, et al. Discovery, synthesis, and biological evaluation of orally active pyrrolidone derivatives as novel inhibitors of p53-MDM2 protein-protein interaction. *Journal of medicinal chemistry*. 2012; 55:9630–9642. [PubMed: 23046248]

26. Ng S, Collisson EA, Sokolov A, et al. PARADIGM-SHIFT predicts the function of mutations in multiple cancers using pathway impact analysis. *Bioinformatics*. 2012; 28:i640–i646. [PubMed: 22962493]