



Systems and Network-Based Approaches for Personalized Medicine

Sarath Chandra Janga^{1,2,3*} and Mohan Mallikarjuna Rao Edupuganti⁴¹Department of Biohealth Informatics, School of Informatics and Computing, Indiana University Purdue University, 719 Indiana Ave Ste 319, Walker Plaza Building, Indianapolis, Indiana 46202, USA²Center for Computational Biology and Bioinformatics, Indiana University School of Medicine, 5021 Health Information and Translational Sciences (HITS), 410 West 10th Street, Indianapolis, Indiana, 46202, USA³Department of Medical and Molecular Genetics, Indiana University School of Medicine, Medical Research and Library Building, 975 West Walnut Street, Indianapolis, Indiana, 46202, USA⁴Division of Cardiovascular Medicine, Department of Internal Medicine, University of Arkansas for Medical Sciences, 4301 West Markham Street, Little Rock, Arkansas, 72205, USA***Corresponding author:** Sarath Chandra Janga, Department of Biohealth Informatics, Center for Computational Biology and Bioinformatics, Department of Medical and Molecular Genetics, Indianapolis, Indiana, USA, Tel: +1-317-278-4147; Fax: +1-317-278-9201; E-mail: scjanga@iupui.edu**Rec Date:** October 30, 2014, **Acc date:** October 31, 2014, **Pub date:** November 2, 2014**Copyright:** © 2014 Janga SC, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract

Most biological outcomes in a cell arise from a complex interplay between different cellular entities such as proteins, DNA, RNA and metabolites. Therefore, a key challenge for biology in the twenty-first century is to understand the structure and dynamics of the complex web of interactions in a cell that contribute to its proper functioning. Recent years have seen a surge in the amount of "omics" data and an integration of several disciplines which has influenced all areas of life sciences, from molecular biology to medicine. With the emergence of a number of sophisticated tools and technologies as a result of genomics revolution, we are now in a position to view the molecular aspects of diseases at a systems level by incorporating various cellular entities into a network framework. Such systems/network-based approaches are not only enabling us to develop models of disease and wellness in a population but also contributing to our efforts to reverse engineer the molecular networks corresponding to disease states by perturbing using drug cocktails. These multi-scale personalized medicine approaches are likely to significantly re-shape the health care industry in the coming decades and decrease the division that we currently see between medicine and other biotechnology disciplines.

Keywords: Systems biology; Network pharmacology; Drug discovery; Personal genomics; Patient profiling; Health care

Editorial

Reductionism, which has been the paradigm in biological research for more than a century, has provided us with a wealth of knowledge about the individual cellular components, their functions and mechanisms. Despite its huge success in the last century, post-genomic biology has increasingly made it clear that discrete biological function can only rarely be attributed to an individual molecule. Instead, most biological outcomes in a cell arise from a complex interplay between different cellular entities such as proteins, DNA, RNA and metabolites. Therefore, a key challenge for biology in the twenty-first century is to understand the structure and dynamics of the complex web of interactions in a cell that contribute to its proper functioning. Although, we cannot answer this question in full in the very near future, the analyses, concepts and frameworks that have resulted in the recent years have provided the scientific community with tools to interpret and better understanding the logic behind the several layers of complex web of interactions happening in the cell.

In the last few years there has been a rapid development in various high-throughput technologies which has led to the accumulation of a large amount of data from different areas of molecular and cellular biology. These developments together with increasing interest in the community for gaining a systems-wide understanding of the cellular machinery have provided us unprecedented insights into the structure,

organization and dynamics of various major cellular processes such as transcription, translation, degradation, replication, metabolism etc. Likewise, efforts to understand the interaction of the cell with external environment have generated global phenotypic maps such as those due to small-molecule perturbations [1] and human microbiomes, which provide us with unparalleled information on the wide variety of microbes which interact with the host's tissues and play an important role in health and disease of an individual [2-4]. Despite the growing amount of data representing each of these processes it should be admitted that none of these cellular processes work in isolation but rather form an integrated network of different wiring diagrams which is responsible for the observed behavior of the cell with in the context of its environment. While there is mounting evidence from several recent studies that each of these networks of associations associated with a particular cellular process can be studied in detail to provide meaningful insights into how they contribute to the functioning of the cell, as well as to identify the factors that constrain their structure and how they influence the genomes on which they are encoded, it is clear that an open challenge of the contemporary biology is to integrate these diverse cellular programs to first understand and model in quantitative terms the topological and dynamic properties of such a unified cellular network and then to exploit it for the therapeutic benefit of mankind.

Conventional medicines that we are aware off today had a start as empirical therapies based primarily in the context of physiology. Even as mechanism-based therapies started to evolve, lack of efficacy and adverse effects were noted and have been accepted to a certain extent.

Most of the conventional medicines were developed as universal drugs to all members of a population suffering from a particular disease. For diseases with multiple pharmacotherapies, the choice was usually left to the prescribing physician's experience and preferences. In other words, traditional molecular pathology has been performed by analyzing well-characterized individual genes, proteins or other molecules identified by targeted studies. However with recent advances in genetics and our understanding of pharmacological behavior of different drugs, it is evident that genetic variations among individuals lead to differences in response to drugs and most complex multi-gene associated diseases could be treated with different drugs depending on the exact genetic variant responsible for the observed phenotype seen in the population. In addition, conventional medicines which were geared towards avoiding side-effects often result in adverse effects despite extensive engineering and often lack efficacy [5]. All these factors together with the observation that conventional clinical trials usually fail for several newly developed drugs because they take in statistical information about the general population of patients and apply it to the individual, has resulted in the notion of personalized medicine. In contrast to trial and error approach of conventional therapies, personalized medicine aims to achieve a better match of drugs to patients so that the right treatments are given to the right patients at the right time. Personalized medicine has become a reality with the sequencing of the human genome, advances in medical genetics, and several technologies including medical diagnostics, single nucleotide polymorphism (SNP) genotyping, transcriptomics and proteomics. In case of complex disorders, the conventional 'one-drug-fits-all' approach involves trial and error before an appropriate treatment is found. In fact, even if appropriate treatment is found using this approach, clinical trial data for a new drug merely show the average response of a study group thereby limiting its success to a small percentage of the population. This is because there is considerable individual variation; some patients show no response whereas others show a dramatic response. This has been increasingly attributed to the fact that although approximately 99.9% of our DNA sequence is identical, the 0.1% difference between any two individuals is medically significant. Buried within this small percentage of difference, lie the clues to hereditary susceptibility to virtually all diseases. At the DNA level, this 0.1% difference translates into 3 million sites of genomic variation. In fact, studies of structural variations (SV) in the human genome, indicate that differences between individuals are far higher than 0.1% as a result of recent sequencing of several human genomes from the HapMap and 1000 genomes projects [6,7]. Therefore, it is obvious that the concept of one drug for all patients with the same disease does not hold and a more individualized approach is needed. One critical aspect of most complex diseases, for example breast cancer, is that the term encompasses many different types of the disease – as defined by differing combinations of disease-perturbed networks i.e, ensemble of network variations can all be the cause of the same disease phenotype, so the same drug might not work for these sub-types. Therefore, the aim of the personalized medicine is to match the right drug to the right patient and in some cases, even to design the treatment for a patient according to genotype and other individual characteristics causing the particular disease phenotype. Key to the use of these data, and also a prime challenge, will be to mine and integrate these data in the context of the dynamics of cellular networks and molecular machines and to construct models of wellness and disease that are predictive of possible disease outcomes so that appropriate action can be taken before the onset of the disease. Recent technologies have made this approach to drug discovery increasingly feasible. Genetic

variation between individual cells or populations can be studied using SNP genotyping arrays or clone-based comparative genomic hybridization (CGH) techniques. Changes in a cell's entire transcriptome or proteome can be profiled in detail. Integration of networks encompassing interactions between different cellular entities can then be studied within the framework of genetic and/or transcriptomic variations responsible for these variations by comparing at different levels; from cells, organs, tissues to individuals to identify the nodes which are responsible for these changes. Detailed changes in cellular phenotype can be quantitatively measured using high content phenotypic screens and hence libraries of small molecules, peptides or poly-nucleotides such as siRNA can be screened to target specific cellular entities that need to be modulated in order to revert the cellular state from a diseased to a healthy one (Figure 1) [5].

One clear advantage of this approach is that molecular entities specific to the disease of interest can be perturbed while still affecting connected pathways and networks [8]. The power of these technologies lies in their ability to generate complex biological data at massive scales. Integration and analysis of this multi-parametric data is vital to personalized medicine. Patterns and relationships within these data sets can be revealed using modelling and comparative network analysis. These patterns can point to pathways that are relevant to specific biological processes making the ultimate goal of understanding the biology of a cell for therapeutic benefit, a reality at the systems level.

Most current cancer therapies that target key molecules have not fulfilled expected promises for common malignancies. Major challenges include the incomplete understanding and validation of these targets in patients, multiplicity and complexity of genetic and epigenetic changes in majority of cancers, and the redundancies and cross-talk found in key signaling pathways. Collectively, the uses of single-pathway targeted approaches are not effective therapies for human malignancies. To overcome these barriers, it is important to understand the molecular cross-talk among key signaling pathways and how they may be combinatorially altered by targeted agents. Such strategies will aid the design of novel therapeutics and their combinations against multifaceted diseases, in which efficacious combination therapies will focus on altering multiple pathways rather than single proteins- a goal of systems medicine, an important step towards personalized approaches to medical treatments. In this context, integrated network modeling and systems biology have emerged as powerful tools benefiting our understanding of drug mechanisms of action in real time.

Environmental factors can also play an important role in the cause of a disease in an individual, who might otherwise be genetically predisposed to it. Most differences in responses to drugs in human are multifactorial, caused by genetic plus environmental factors and this is yet another argument for the broader approach of personalized medicine rather than limiting it to account for the genetic causes. For instance, some adverse drug reactions are caused by interaction of the drugs with environmental toxins, infectious organisms, or dietary constituents. Therefore, prescription of drugs based on genotype tests to individuals considered safe to receive the drugs, may not completely eliminate the possibility of such a reaction. A patient matched to a drug on the basis of a genotyping test may not necessarily respond to it. Although there is considerable improvement in safety and efficacy of a limited number of drugs available now in combination with diagnostics, investigation of environmental factors must continue in order to identify other factors, which will vary from one patient to

another and would still come under the scope of personalized medicine. One such environmental factor is the human intestinal microflora composed of 10¹³ to 10¹⁴ microorganisms whose

collective genome (microbiome) contains at least 100 times as many genes as the human genome [9].

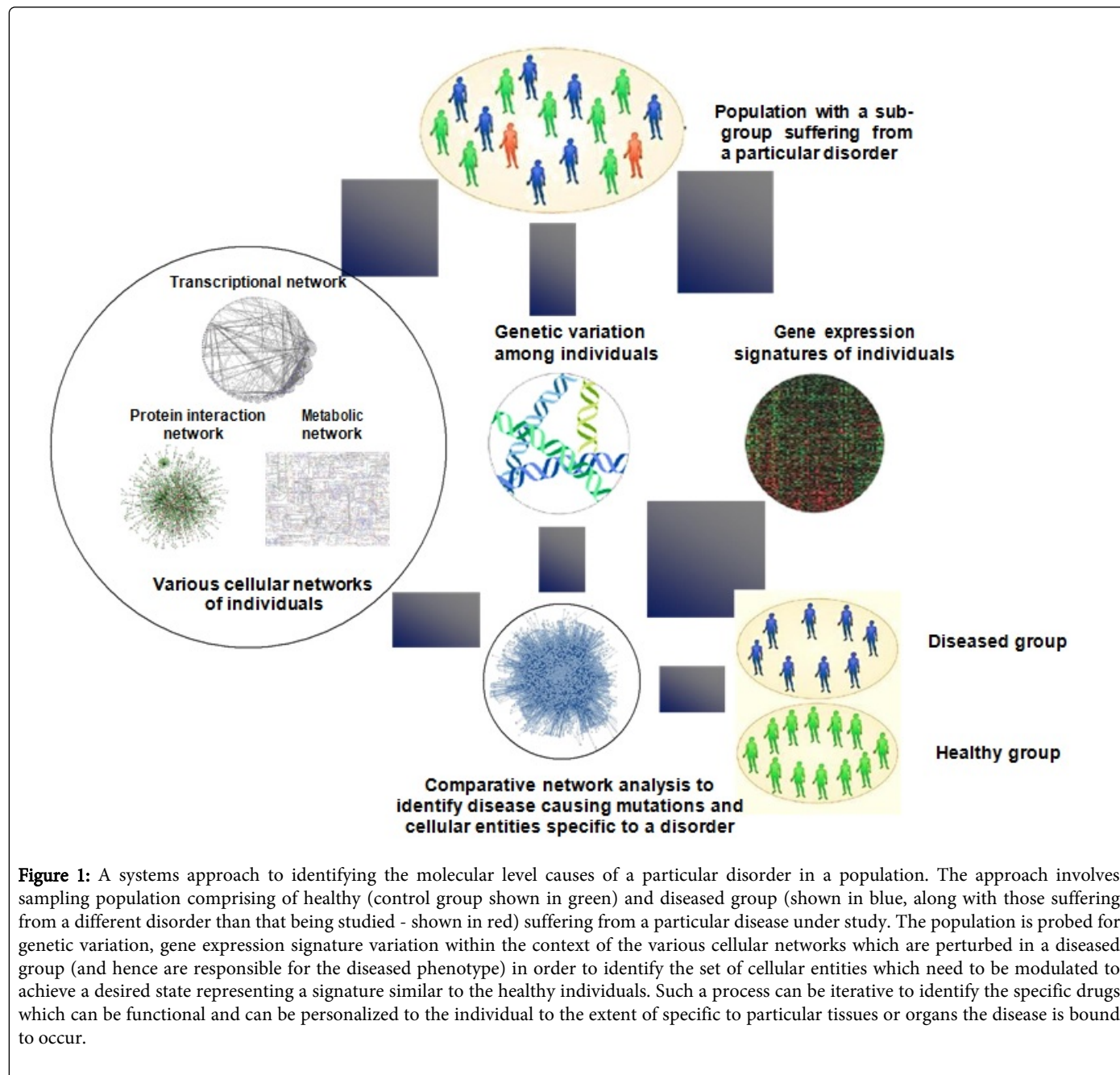


Figure 1: A systems approach to identifying the molecular level causes of a particular disorder in a population. The approach involves sampling population comprising of healthy (control group shown in green) and diseased group (shown in blue, along with those suffering from a different disorder than that being studied - shown in red) suffering from a particular disease under study. The population is probed for genetic variation, gene expression signature variation within the context of the various cellular networks which are perturbed in a diseased group (and hence are responsible for the diseased phenotype) in order to identify the set of cellular entities which need to be modulated to achieve a desired state representing a signature similar to the healthy individuals. Such a process can be iterative to identify the specific drugs which can be functional and can be personalized to the individual to the extent of specific to particular tissues or organs the disease is bound to occur.

The general consensus of the recent studies analyzing human microbiomes is that humans are super organisms whose metabolism represents an amalgamation of microbial and human attributes with microbiomes between tissues and organs vastly varying with time, although the genomic component of the humans remains the same across the cell types further supporting the role of microbiomes in shaping the state of an individual [3,4,9-11]. Without understanding the interactions between human and microbial genomes, it is impossible to obtain a complete picture of human biology. Therefore, the next frontier in this field is the metagenomic analysis of human microbiomes corresponding to different body parts in health and

disease as this has implications for clinical diagnosis and the treatment of many human diseases associated with environmental microbial influences. With the knowledge gained in this area, one can use biomarkers to identify the bacterial population of the individual which needs perturbation. Physicians can then manipulate the population bacteria to be consistent with the optimal health of an individual. Such an analysis would also identify bacteria that are resistant to certain antibiotics, and enable the selection of the appropriate antibiotic for a patient. For instance, in the future, healthy individuals could undergo a metagenomic analysis of their gut to determine their immune status and susceptibility to certain diseases. Such an analysis may enable the

assessment of the effects of age, diet and diseases such as inflammatory bowel disease, cancer, obesity and bacterial vaginosis on the microbial flora of the body parts in persons living in different environments with different dietary habits. In fact, consortiums such as the Human Microbiome Project (HMP) [12] aim to generate resources to describe these microbial communities and to analyze their role in health and disease. For instance, HMP has already proven to be very effective by sequencing and making it public the microbiomes of more than 15 body habitats for 300 healthy human volunteers, providing a wealth of information for understanding the differences between different body parts across individuals. Such datasets not only allow the identification of bacterial species and their interplay in an ecosystem context, responsible for maintaining a homeostasis in the respective body parts but also permit dissecting their role in non-genetic diseases in unhealthy patients, providing a wealth of opportunities for network analysis and integration in the context of human-microbiome associations and more generally their relevance in the context of personalized medicine as more data becomes available.

Personalized medicine is a challenge because each of our genomes differ by approximately 6 million bases, this large difference in individuals ultimately requires the use of the patient's own baseline healthy data as a 'control' that can then be used for subtractive analyses of longitudinal accumulations of molecular and cellular information that signal the transitions from wellness to disease. Such a medicine can become a reality as the details of disease-perturbed molecular networks open the possibility of using drugs targeted at key nodes to stop disease progression. In other words, network-based view of drug discovery and development seeks to define disease states as molecular networks which integrate the following major aspects 1) Gene expression profiles 2) Genome-wide association data reflecting the genomic variations among populations 3) integrated interaction networks which amalgamate the associations between genes, proteins, RNA, cellular metabolites etc. The complex cellular network of interactions are then compared between diseased and healthy populations at the level of cells, tissues or organs, informed by the DNA variation and gene expression patterns to identify one or more of the cellular entities responsible for the rewiring patterns seen in the disease phenotype (Figure 1). Such an approach takes into account the multi-factorial nature of most complex disorders and identifies the part of the diseased state (network) which needs to be perturbed (perhaps combinatorially) with possible combinations of drugs to get back to the desired healthy state. For instance, in complex diseases such as cancer which are known to be caused by a wide variety of phenomenon, the future of personalized medicine is to have a large number of drugs to choose from, each associated with one or more biomarkers that predict responsiveness to the particular sub-type under study in a patient. For example, ~15-20% of breast cancers have HER2 amplification, which corresponds to high levels of responsiveness to trastuzumab [13,14]. Without knowing the relevant population of breast cancer patients to treat with this inhibitor, the drug would not have had significant improvement in the overall population. However, even when we use available tools to predict patient sensitivity, such as FISH for HER2 amplification, we still are not able to perfectly predict who will respond. As an example, approximately one third breast cancer patients with amplified HER2 do not respond to trastuzumab, in part due to deregulation of downstream or parallel pathway components such as PTEN or P13K [15,16]. It is in this context the integrated framework of genetic variation, gene expression profiling and network analysis can provide a huge advantage to uncover the patient population which is likely to

respond to appropriate drug cocktails. Personalized medicine is catalyzing a transformation of medicine that promises to deal with the current impenetrable barriers of incredible complexities of disease through systems approaches, emerging technologies and powerful analytical tools. The ultimate promise is that the focus of medicine will be shifted from disease, which is currently based on physiology and phenotypes, to that of wellness, where signatures of population with healthy states corresponding to billions of data points for each individual will define with exquisite specificity the nature of their wellness – and any transitions into disease can be identified as early as possible to prevent the onset of the disease. Central to this view is the idea that the molecular, cellular and phenotypic data of eventually hundreds of millions of patients will be available for systems analyses which will allow the integration, mining and development of predictive models of each disease. Such a concept of shared patient data and disease models not only help to continuously refine existing disease maps built in an open-access manner that includes researchers from around the world but also can fundamentally change the speed of the discovery and development of therapies for new diseases.

Acknowledgments

SCJ acknowledges support from the School of Informatics and Computing at IUPUI and thank a number of colleagues for providing helpful comments on a previous version of this manuscript. The author reports no financial or other conflict of interest relevant to the subject of this article.

References

1. Lamb, J. et al. (2006). The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *Science* 313, 1929-35.
2. Parfrey LW, Walters WA, Knight R (2011) Microbial eukaryotes in the human microbiome: ecology, evolution, and future directions. *Front Microbiol* 2: 153.
3. White, B.A., Creedon, D.J., Nelson, K.E. and Wilson, B.A. (2011). The vaginal microbiome in health and disease. *Trends Endocrinol Metab*
4. Kinross JM, Darzi AW, Nicholson JK (2011) Gut microbiome-host interactions in health and disease. *Genome Med* 3: 14.
5. Janga SC, Tzakos A (2009) Structure and organization of drug-target networks: insights from genomic approaches for drug discovery. *Mol Biosyst* 5: 1536-1548.
6. Korbel JO, Urban AE, Affourtit JP, Godwin B, Grubert F, et al. (2007) Paired-end mapping reveals extensive structural variation in the human genome. *Science* 318: 420-426.
7. Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, et al. (2006) Global variation in copy number in the human genome. *Nature* 444: 444-454.
8. Schadt EE, Friend SH, Shaywitz DA (2009) A network view of disease and compound screening. *Nat Rev Drug Discov* 8: 286-295.
9. Weinstock GM (2011) The volatile microbiome. *Genome Biol* 12: 114.
10. Caporaso JG, Lauber CL, Costello EK, Berg-Lyons D, Gonzalez A, et al. (2011) Moving pictures of the human microbiome. *Genome Biol* 12: R50.
11. Hamady M, Knight R (2009) Microbial community profiling for human microbiome projects: Tools, techniques, and challenges. *Genome Res* 19: 1141-1152.
12. Turnbaugh PJ, Ley RE, Hamady M, Fraser-Liggett CM, Knight R, et al. (2007) The human microbiome project. *Nature* 449: 804-810.
13. Barron JJ, Cziraky MJ, Weisman T, Hicks DG (2009) HER2 testing and subsequent trastuzumab treatment for breast cancer in a managed care environment. *Oncologist* 14: 760-768.
14. Coulson SG, Kumar VS, Manifold IM, Hatton MQ, Ramakrishnan S, et al. (2010) Review of testing and use of adjuvant trastuzumab across a

-
- cancer network--are we treating the right patients? *Clin Oncol (R Coll Radiol)* 22: 289-293.
15. Marty B, Maire V, Gravier E, Rigai G, Vincent-Salomon A, et al. (2008) Frequent PTEN genomic alterations and activated phosphatidylinositol 3-kinase pathway in basal-like breast cancer cells. *Breast Cancer Res* 10: R101.
16. Stemke-Hale K, Gonzalez-Angulo AM, Lluch A, Neve RM, Kuo WL, et al. (2008) An integrative genomic and proteomic analysis of PIK3CA, PTEN, and AKT mutations in breast cancer. *Cancer Res* 68: 6084-6091.