Published in final edited form as:

Clin Pharmacol Ther. 2013 April ; 93(4): 309–311. doi:10.1038/clpt.2012.219.

Functionalizing genomic data for personalization of medicine

Cyril H. Benes

The Center for Molecular Therapeutics, Massachusetts General Hospital Cancer Center, Charlestown Navy Yard, Room 7401, 149, 13th Street, Charlestown, MA 02129, Phone: 617.724.3409, Fax: 617.726.7808

Cyril H. Benes: cbenes@partners.org

Abstract

The wealth of genomic information available today to researcher and clinician has dramatically changed the way some diseases are defined. At the same time leveraging this information is a major challenge in part because previously established concepts and tools used for research and development need are largely inadequate. Using the example of cancer genomics this Translation focuses on systems biology approaches to inform the development of personalized medicine.

Keywords

Genomic; Pharmacogenomic; Systems biology; Personalized medicine

Cancer has been defined as a genetic disease since the identification of the first cancer genes as viral oncogene homologues and the realization that mutations in specific genes can lead to the development of a specific array of tumours. The treatment of cancer has on the other hand been decided based on non-genetic information: primary tumor site and histological classifications. Building on a vast array of studies on cancer genetics and on the molecular underpinning of cancers novel rationally designed therapies are now applied in a small but steadily growing number of cases accompanied by patient selection strategies that lead to marked improvement in treatment outcome. For example, dependency on the activity of the EGFR leads to some dramatic clinical responses in the subset of non-small cell lung cancer patients harboring a set of mutations in the EGFR gene and a panel of mutations found in non-small cell lung cancer now have matched therapeutic strategies (1). This personalization of cancer care based on genetic or more broadly genomic information has the potential to revolutionize the field of cancer therapeutics. The current agents of change are the novel technologies that permit today to interrogate cancer genomes at a rate and depth unimaginable just a few years ago and will tomorrow define the genome of potentially every cancer patient entering the clinic as part of routine diagnostic. However, for the most part, the rational development of cancer therapeutics is driven by our current understanding of tumorigenesis, the genes driving neoplasia and the networks that regulate critical activities necessary for cancer development and in some cases tumor maintenance. This knowledge has been acquired mostly through focused studies of a single gene or gene family, a particular piece of signaling network or other studies which, while leading to tremendous progress represent a relatively narrow scope of investigation in the face of the newly revealed complexity of cancer genomes.

A growing body of work has started to systematically interrogate the genome of cancer cells with the goal of creating a functional map to be used target discovery and biomarker

establishment (see Table). Pioneering efforts have mostly focused on specific genetic variants that are of high disease relevance but are not readily targetable (such as KRAS mutation) as well as mechanisms of therapeutic resistance in otherwise sensitizing genotypes with identified therapeutics (such as BRAF mutation). These functional interrogations of cancer genomes are carried most often using RNA interference and measure effects on cell viability. These experiments generate complex datasets that will become more actionable when integrated with complementary information and results might not necessarily translate across different contexts but they have already delivered clinically relevant insights. The depth of information obtained and to some extent the specificity of the probes in these approaches are currently unparalleled and modeling of RNAi screen datasets obtained across different contexts will almost certainly provide a great deal of valuable information for the development of novel therapeutic strategies (see for example multidisciplinary effort known as the connectivity map http://www.broadinstitute.org/genome_bio/connectivitymap.html). Other systems biology approaches (see for example the CTD2 <http://ctd2.nci.nih.gov/> and LINCS <http://lincsproject.org/projectpages>) use multiple types of probes (RNAi, small molecule or biologics) against a more restricted number of target but sample a larger phenotypic space (not only cell viability) to obtain better mechanistic insight, or combine perturbations (RNAi plus small molecules) to more directly define gene-drug associations. Quantitative measurements of a large number of cellular parameters, such as signaling pathway activation, transcriptional changes or cytokines released are used to feed mathematical models of the cells response to perturbations. Such models can then be used to predict drug sensitivity of a biological system in a given state (here state sometimes equals genotype). A systematic and complete functionalization of cancer genomes will ultimately provide the information to model therapeutic response with much greater accuracy. Clearly, such in depth characterization will take time with the bottleneck in knowledge development likely to move from data acquisition to data analysis very quickly. Thus, partial information needs to be leveraged en route to the goal of global cancer cell networks modeling in order to achieve better clinical responses for today's patients.

Complementary to the approaches described above the profiling of current and upcoming cancer therapeutics across large collection of tumor derived cell lines also functionally interrogates genomic variations and has the potential to inform cancer care rapidly. Here, the size of the cell line collection is used to capture a broad range of the genomic variations encountered in cancer and differences in drug sensitivity across cell lines is correlated with these genomic changes. Pharmacological studies of cancer in the laboratory using tumor derived cell lines are an integral part of therapeutic development and have given insights in many aspects of cancer biology. With the current understanding of the genomic variations observed within each cancer type as classified by organ of origin it becomes obvious that the historical efforts leveraging a relatively small number of cancer cell lines to identify potential anti-cancer therapeutics have to be scaled-up. For example the NCI-60, a panel meant to capture different tumor types to evaluate candidate therapeutics activity could only contain a few cell lines per cancer type. As a straightforward example of the scale issue consider the frequency of mutation that can in some instances be associated with dramatic clinical response. For example, only about 5% of non-small cell lung cancer harbor EML4-ALK translocations associated with extreme sensitivity to ALK inhibitor and remarkable clinical responses, thus screening a cell line collection encompassing several dozens of non-small cell lung cancer cell lines was necessary to capture this dependency in the laboratory (2,3).

Based on these observations, several research centers have assembled larger panels of cancer cell lines, in some instances focusing on a given cancer type (see for example <http://icbp.lbl.gov/index.htm>) with others attempting to capture a larger breadth of genomic variations in cancer by using cell lines representing the majority of known tumor types.

These cell lines have over time been characterized at the genomic level, transcriptome, sequence of genes previously established as cancer genes, copy number changes across the genome. Further characterization is now ongoing using next-gen sequencing technology and soon the sequence of all coding regions (exome) and even the full genome will be collected. High-throughput pharmacological screening is used to define the genomic context most highly correlated with drug sensitivity. This, in some instances at least will be expected to translate into the identification of a subpopulation of patients most likely to present with robust therapeutic response. The genomic correlates identified in these studies thus represent candidate biomarkers that could ultimately be used in patient selection strategies applied to clinical trials. In such an effort, the Center for Molecular Therapeutics at the Massachusetts General Hospital Cancer Center and the team of the Genomic of Drug Sensitivity in Cancer program at the Wellcome Trust Sanger Institute are collaboratively interrogating a collection of ~1,000 cell lines for their sensitivity to current and upcoming cancer therapeutics (<http://www.cancerrxgene.org/>) and similar efforts are ongoing at the Broad Institute (<http://www.broadinstitute.org/ccle/>). While it is clear that drug sensitivity *in vitro* can only represent a fragment of the therapeutic response in patients recently published studies (4, 5) demonstrate that known and novel correlates of drug sensitivity, some of which are already leveraged in the clinic can be readily identified using high-throughput cell line screening. The published data and publicly available updates (see links to websites above) demonstrate the remarkable complexity of drug sensitivity. Statistical analysis and mathematical modeling of these data reveals that a large number of parameters appear to modulate drug sensitivity and the challenge of describing the observed responses with only a handful of strong genomic correlates. While not really surprising, this complexity is not only a hurdle to the validation of the new findings in more physiological contexts but suggests a need for very fine granularity in personalization of care. In particular, while for some drugs a clear major determinant of drug sensitivity can be identified (such as BRAF mutation for BRAF inhibitor sensitivity) in all cases sensitivity among cell lines bearing these genomic events varies broadly, reminiscent of the variation seen in clinical responses. Another important outcome of these studies is the evidence that a given mutation (such as BRAF V600E) might not equivalently sensitize cancer cells to a drug in different contexts, again reminiscent of recent results showing poor performance of BRAF inhibitors in non-melanoma compared to the responses seen in melanoma patients bearing an identical BRAF mutation. Thus, although cell line screens cannot capture the full complexity of the therapeutic response, they are tractable *in vitro* systems that will help in the development of additional biomarkers of sensitivity used to stratify patient populations further.

Nevertheless, the translation of the laboratory findings into a clinical strategy relies on the availability of an equivalent depth of genomic information on patient samples than the drug response models contains. Thus, highly complex models integrating genomic data not currently available for a sufficient number of cases are at this point conceptually informative but largely non-actionable. Similarly, models that rely on post-treatment measurements (for example changes in transcript level) are much more difficult to apply in the clinic. For these reasons, drug sensitivity models that contain variables most widely and robustly acquirable in the clinic are of higher immediate value. This largely means a single or a few genetic events (mutations, amplifications or translocations). Similarly, in terms of rapid clinical translation systems biology approaches are most useful today to identify subgroup of patients most likely to benefit from drugs developed based on prior knowledge of disease mechanism. Systems biology approaches have however the capability of fully functionalizing the cancer genome and other disease genome and will eventually define not only which patients should get which drug but where to focus new drug development efforts.

Acknowledgments

Wellcome Trust Grant 086357. NIH Grant 1U54 HG006097-01.

References

1. Sequist LV, Heist RS, Shaw AT, Fidias P, Rosovsky R, Temel JS, et al. Implementing multiplexed genotyping of non-small-cell lung cancers into routine clinical practice. *Ann Oncol*. 2011 Dec; 22(12):2616–24. [PubMed: 22071650]
2. McDermott U, Iafrate AJ, Gray NS, Shioda T, Classon M, Maheswaran S, et al. Genomic alterations of anaplastic lymphoma kinase may sensitize tumors to anaplastic lymphoma kinase inhibitors. *Cancer Res*. 2008 May 1; 68(9):3389–95. [PubMed: 18451166]
3. Kwak EL, Bang YJ, Camidge DR, Shaw AT, Solomon B, Maki RG, et al. Anaplastic lymphoma kinase inhibition in non-small-cell lung cancer. *N Engl J Med*. 2010 Oct 28; 363(18):1693–703. [PubMed: 20979469]
4. Barretina J, Caponigro G, Stransky N, Venkatesan K, Margolin AA, Kim S, et al. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature*. 2012 Mar 29; 483(7391):603–7. [PubMed: 22460905]
5. Garnett MJ, Edelman EJ, Heidorn SJ, Greenman CD, Dastur A, Lau KW, et al. Systematic identification of genomic markers of drug sensitivity in cancer cells. *Nature*. 2012 Mar 29; 483(7391):570–5. [PubMed: 22460902]

Table
Systems biology approaches for functional interrogation of genomes and major short-term translational impacts.

Approach	Genomic Coverage	Phenotypic Coverage (Type of assays)	Context Coverage	Major Therapeutic Translation	Therapeutic Question
RNAi Screen	Full genome	Low (Viability change)	Limited (Typically 1–2 cell lines)	Novel target identification	Direct targeting difficult, Therapeutic resistance.
Systems pharmacology (Small molecules, biologics, RNAi probes)	Low (<100 targets)	High (Cellular and molecular measurements, signaling events etc.)	Medium (<50 cell lines)	Predictive biomarker, mechanism of action	Biomarker discovery, molecular basis of response.
Large Cell Collection Screen	Moderate to high (disease dependent)	Low (viability change)	High (1,000 Cell lines)	Predictive biomarker, drug application	Patient classification for expected or novel use. Biomarker identification.