



# feature

## Exploiting genetic complexity in cancer to improve therapeutic strategies

Mathew J. Garnett\*, mj12@sanger.ac.uk and Ultan McDermott

Advances in genome sequencing technologies are enabling researchers to make rapid progress in defining the entire repertoire of causal genetic changes in cancer. The response of patients with cancer to therapy is often highly variable and there is an increasing number of examples where mutations in cancer genomes have been shown to have a profound effect on the clinical effectiveness of drugs. An urgent challenge for the research and clinical communities is how to translate these genomic data sets into new and improved therapeutic strategies for the treatment of patients. The use of large-scale cell line-based drug screens to identify genomic 'biomarkers' of drug response for the stratification of patients has the potential to transform how patients with cancer are treated.

All cancers arise as a result of the acquisition of somatic mutations in their genomes that fundamentally alter the function of the protein product of key cancer genes [1]. In many cases, cancers harbour several hundred mutations, a small number of which (commonly 5–10 'driver' mutations) are thought to be necessary for acquisition of the malignant phenotype and, ultimately, carcinogenesis (Fig. 1). Such mutations are responsible not only for the development of the cancer in the first instance, but also for maintaining the proliferation and evasion of cell death that are the hallmarks of cancer [2]. A detailed knowledge of such mutational events is crucial if one is to understand how cancers develop as well as to design rational therapies to target the appropriate dysregulated pathways.

Over the past decade, several meticulous studies involving gene resequencing have begun to characterise the genetic changes that occur in cancer and have revealed the presence of substantial genomic heterogeneity across

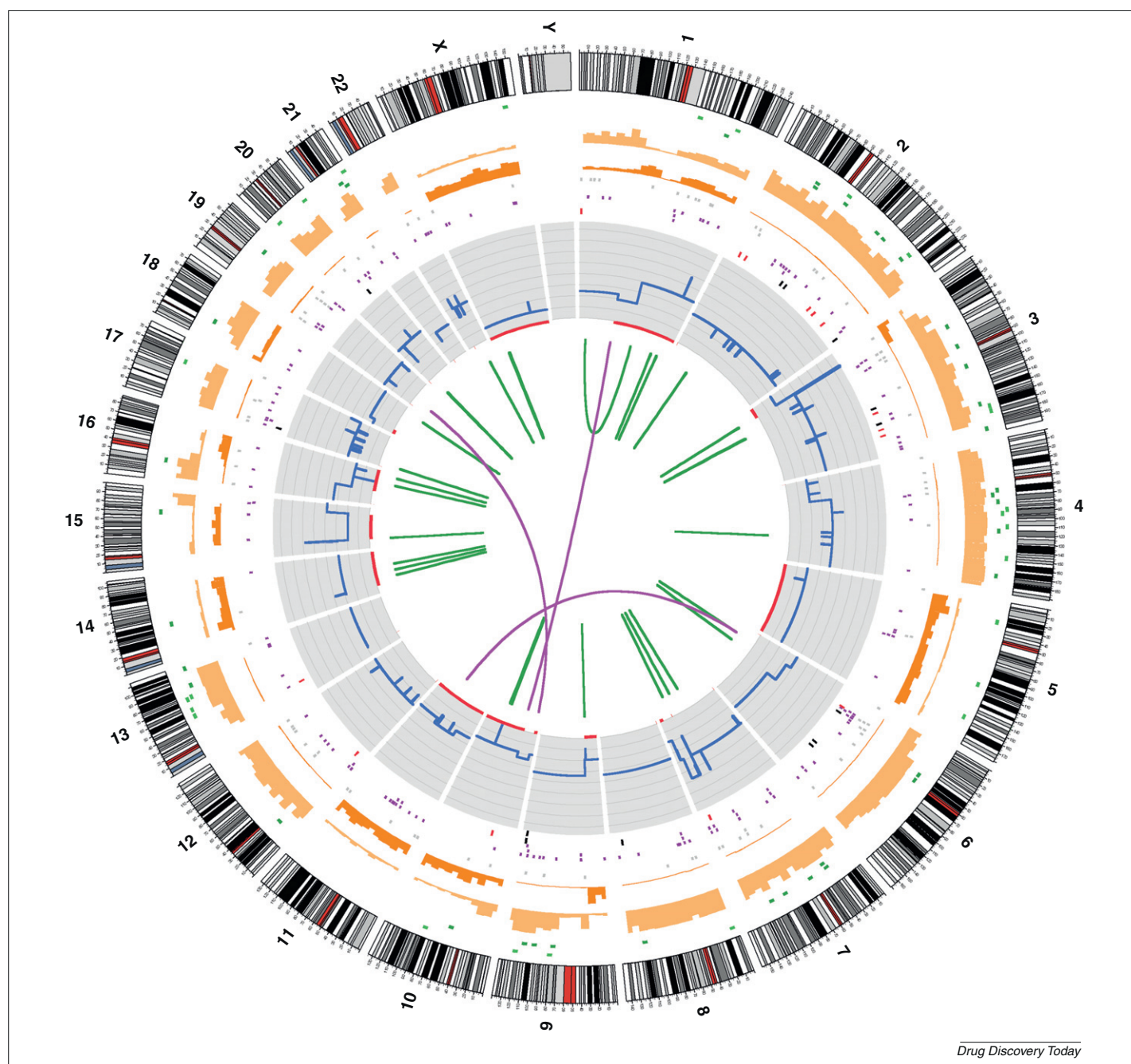
cancer genomes. To date, >400 genes have been identified for which mutations (including somatic coding changes, amplification, deletions and fusion genes) have been causally implicated in cancer (so-called 'cancer genes'; <http://www.sanger.ac.uk/genetics/CGP/Census/>) [3]. In many cases, these mutations occur only in the context of a specific tissue or cancer subtype, and even within these many mutations occur at a low frequency. Moreover, as discussed below, improvements in DNA sequencing technologies are enabling scientists and clinicians to expand rapidly upon the work of these studies and a complete genomic landscape of human cancers is beginning to emerge.

One of the hopes held following completion of the human genome sequence more than 10 years ago was that it would hasten a move towards personalised medicine for many human diseases based on a detailed knowledge of the alterations in germline and cancer genomes. The fundamental principles that underlie per-

sonalised cancer medicine are that: (i) significant genomic heterogeneity exists among tumours, even those derived from the same tissue and (ii) these differences can have a profound impact on the likelihood of a clinical response to treatment with particular therapeutic agents. Thus, a detailed analysis of the genomic landscape of human cancer, coupled with detailed information on how cancer mutations impact the clinical response to a drug, should in principle enable the improvement of clinical effectiveness by accurately stratifying patients who are most likely to respond to a specific treatment based on genomic biomarkers. Moreover, the identification of predictive biomarkers during the early-phase development of new drugs would have an impact on the cost, design and success of new cancer drug development.

### Genomic biomarkers of drug response

Most of the current treatment regimens for cancer are based around the tissue of origin and

**FIGURE 1**

The complexity of somatic alterations in human cancer genomes. A Circos plot showing the whole-genome catalogue of somatic mutations from the malignant melanoma cell line COLO-829. This genome has approximately 30,000 somatic base substitutions and 1000 somatic insertions and/or deletions. In coding exons, 272 somatic substitutions are present, including 155 missense changes, 16 nonsense changes and 101 silent changes. The number and types of mutation are highly variable across different cancer genomes. Chromosome number and karyotype are indicated on the exterior of the plot. Key: blue lines, copy number across each chromosome; red lines, sites of loss of heterozygosity (LOH); green lines, intrachromosomal rearrangements; purple lines, interchromosomal rearrangements; red spots, nonsense mutations; green spots, missense mutations; black spots, silent mutations; brown spots, intronic and intergenic mutations (merged).

the clinical response of patients with cancer to treatment with a particular drug is often highly variable. However, there is a compelling body of evidence, both clinical and experimental, that for an increasing number of drugs used in the clinic, the likelihood of a patient's cancer responding to treatment is strongly influenced by alterations in the cancer genome (Table 1). Arguably the most

celebrated example of this has been the use of imatinib, a small molecule inhibitor of the c-ABL oncogene 1, nonreceptor tyrosine kinase (ABL1) to target the fusion protein product of the breakpoint cluster region (*BCR*)–*ABL* translocation seen in chronic myeloid leukaemia (CML). The five-year survival rate for patients newly diagnosed with chronic-phase CML who are

treated with imatinib is 89% [4]. On the heels of this discovery came the finding that therapeutic targeting of the ERBB family member, ERBB2 [human epidermal growth factor receptor 2 (HER2)], resulted in response rates of 15–26% in HER2-overexpressing breast cancers [5,6]. *HER2* is amplified in 15–30% of breast cancers and carries an adverse prognosis [7].

TABLE 1

**US Food and Drug Administration-approved targeted cancer therapeutics in clinical use against drug-sensitising mutations**

<i>Tumour</i>	<i>Gene (mutation)</i>	<i>Prevalence of gene alteration (%)</i>	<i>FDA-approved drug</i>	<i>Year approved</i>	<i>Therapeutic target</i>	<i>Response rate in mutant tumours (%)</i>	<i>Refs</i>
<b>Chronic myeloid leukaemia</b>	<i>BCR-ABL</i> (translocation)	>95	Imatinib	2001	ABL1	>95	[17]
<b>Gastrointestinal stromal tumour</b>	<i>KIT</i> (mutation), <i>PDGFRA</i> (mutation)	85 (KIT), 5–8 (PDGFRA)	Imatinib	2002	KIT, PDGFRA	>80	[18]
<b>Non-small cell lung cancer (adenocarcinoma)</b>	<i>EGFR</i> (mutation)	10	Gefitinib, erlotinib	2003, 2004	EGFR	70	[8]
<b>Chronic myeloid leukaemia (imatinib-resistant)</b>	<i>BCR-ABL</i> (translocation)	>95	Dasatinib	2006	ABL1	>90	[19]
<b>Melanoma</b>	<i>BRAF</i> (mutation)	40–70	Vemurafenib	2011	BRAF	>50	[20]
<b>Non-small cell lung cancer</b>	<i>EML4-ALK</i> (translocation)	2–7	Crizotinib	2011	ALK	57	[9]

Abbreviations: KIT; v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homologue; PDGFRA; platelet-derived growth factor receptor; alpha polypeptide.

More recently, the use of epidermal growth factor receptor (EGFR) and anaplastic lymphoma receptor tyrosine kinase (ALK) inhibitors in patients with lung cancer whose tumours harbour *EGFR* mutations and echinoderm microtubule associated protein like 4 (*EML4*)–*ALK* rearrangements, respectively, has resulted in significantly improved response rates compared with conventional therapies in those subsets of patients [8,9]. Importantly, these genomic alterations typically account for <10% of the patient population and exemplify the degree of diversity faced when considering personalised cancer medicine targeting distinct subpopulations. Although recent studies have demonstrated that targeting the 40–70% of cutaneous melanomas that harbour activating mutations in v-raf murine sarcoma viral oncogene homologue B1 (*BRAF*) with the *BRAF* inhibitor vemurafenib resulted in an approximately 50% response rate and improved survival of patients with this aggressive cancer [10], it is probable that, for most solid tumours, the prevalence of specific mutational events that are sensitised to drugs will be much lower. These observations highlight not only the therapeutic potential of incorporating genomic biomarkers to define patient populations, but also the challenge of identifying relatively low-frequency subpopulations of patients who are drug responsive, even within a single tissue type. This degree of heterogeneity challenges the traditional pharmaceutical industry model of the 'blockbuster' cancer drug and raises important questions as to how the industry will adapt its business model to the potential need for a multitude of drugs targeting distinct patient subpopulations even within the same tissue class.

### Cancer genomics and next-generation sequencing

In recent years, remarkable advances in DNA sequencing technologies, known as next-generation sequencing, have enabled the analysis of genes and genomes at a scale unimaginable a decade ago [11]. So rapid are the advances it is now possible to imagine realistically a time in the near future where it will be technically possible, and affordable, to sequence all cancers and define their mutational burden. These advances are transforming understanding of cancer by making possible for the first time the complete characterisation of the cancer genome across multiple tissue and cancer subtypes, providing new insights into the origins, evolution and progression of cancer.

A major impetus for these technological advances in sequencing has been the potential for improved cancer therapies, based on a detailed understanding of the genomic alterations present. The drug-sensitising genotypes described previously have been used to argue that the identification of additional cancer genes will help in identifying new drug targets as well as defining drug response in subsets of patients with cancer. The International Cancer Genome Consortium (ICGC) was launched with the goal of generating comprehensive catalogues of genomic abnormalities (i.e. somatic mutations, abnormal expression of genes and epigenetic modifications) in tumours from 50 different cancer types and to make the data available to the entire research community to accelerate research into the causes and treatment of cancer (<http://www.icgc.org/>) [12]. Currently, it has received commitments from funding organisations in Asia, Australia, Europe and North America for 39 projects to study >18,000 cancer

genomes. Catalogues of somatic mutations for many of the most prevalent cancer types are already becoming available and, in the near future, it will be possible for any researcher to know the prevalence of almost every somatic mutation in a given cancer.

The identification of the entire repertoire of human cancer genes has the potential to transform cancer therapeutics. First, it could help identify new therapeutic targets for the drug development pipelines of academia and the pharmaceutical industry. Many of the most effective molecularly targeted drugs currently available (e.g. vemurafenib) inhibit the enzymatic activity of the protein products of cancer genes with gain-of-function mutations (i.e. oncogenes). Second, this information could empower researchers to determine experimentally how catalogues of cancer gene mutations affect specific biological outcomes. To exploit this wealth of knowledge fully towards improved cancer therapies, there is a significant need to develop model systems for functional studies to determine experimentally whether specific mutations have a functional role in drug response. We argue here that established cancer cell lines, if screened at a sufficient scale to capture much of the tissue-type and genetic diversity of human cancers, can in many instances faithfully model the effect of cancer mutations on drug response and are a powerful model system to identify new biomarkers of drug sensitivity.

### Drug sensitivity profiling in cancer cell lines

The first human cancer cell line, HeLa, was established over 50 years ago and, since then, cell lines have been generated from almost every

cancer type and have become a standard research tool in molecular biology. Although certain elements of cancer biology, including invasion and metastasis, can only properly be studied in the context of more complex experimental systems, such as mice models, cell lines have been shown to be robust models for certain cancer cell-intrinsic studies. Significantly, they have been found to recapitulate many of the important drug-sensitising genotypes observed in clinical practice [13]. For example, the acute sensitivity of patients with CML to imatinib can be readily modelled in cancer cell lines bearing the *BCR-ABL* translocation. Indeed, much of the understanding of cancer cell biology, including aspects of signalling and gene regulation, has come from studies of cancer cells in culture.

With an increased understanding of the extent of genomic heterogeneity that exists in cancer has come a need for experimental systems that are both capable of recapitulating this variation and amenable to perturbations of biological function to understand the impact of cancer gene mutations. To this end, many researchers have turned to established human cancer cell lines. The first systematic approach to using cell lines to address heterogeneity in drug response was the 60-cell line panel (NCI60) of the National Cancer Institute [14]. This screen of 60 cell lines, encompassing a range of tissue types across a large panel of chemical compounds, was the forerunner of all high-throughput cell-based drug profiling. Although it introduced the concept of screening compounds across cell lines to identify sensitive and resistant populations [15], it is now known that the number of cell lines screened was insufficient to capture the genomic diversity observed in cancer, which was likely to impact on drug response. This is because it is now known that cancer genomes are remarkably heterogeneous and that many cancer genes are present in only a small fraction of any tumour type. Therefore, large numbers of cell lines are required in any screen to identify rare mutant subsets that have altered sensitivity to cancer therapeutics. Indeed, it is probable that in excess of 1000 cell lines would be required to have sufficient power to detect drug-sensitising mutations that are present in <10% of any given tumour type.

Cell-based drug screens of chemical libraries have traditionally been the preserve of the pharmaceutical industry, given the expense of operating large-scale automated screening platforms. In most instances, these screens were performed using a small panel of cell lines, or even a single cell line, against a very large panel

of chemical compounds. However, more recently, academic research groups have begun to perform large-scale screens of cancer drugs across hundreds of different cancer cell lines. The aim of these efforts is to screen a smaller number of compounds across a large number of cell lines to capture as much of the genetic variation that occurs in cancer and that impacts on drug sensitivity as possible.

The Genomics of Drug Sensitivity in Cancer (GDSC) project, a joint UK–USA initiative funded by the Wellcome Trust, was established in 2009 to screen >1000 genetically characterised human cancer cell lines against 400 cancer drugs, with the goal of identifying genomic alterations that affect drug response (Fig. 2). These data are being made publically available, with sensitivity data for >8000 drug–cell line combinations for 23 drugs currently available through the web portal (<http://www.cancerRx-gene.org>). A similar initiative is being pursued in a collaboration between the Broad Institute and the pharmaceutical company Novartis, namely the Cancer Cell line Encyclopedia (<http://www.broadinstitute.org/ccle>).

A key element of these efforts is combining large drug-sensitivity data sets across a broad collection of cell lines together with detailed genomic information, including somatic point mutation, gene amplifications and deletions, and gene expression data. The integration of these data can be tremendously effective for identifying genomic features associated with drug response. It is probable that the immediate focus of these analyses will be the 464 genes that are currently known to be mutated and causally implicated in cancer development [3]. However, over the next few years, the power of these studies to make connections between mutations and drug response is likely to increase significantly as the efforts of the ICGC, and other large-scale cancer-sequencing projects, expand and refine the list of human cancer genes. Additionally, ongoing efforts to sequence the coding exons of all approximately 22,000 human genes across all 1000 cancer cell lines used for drug screening will make them an increasingly powerful model to link drug sensitivity with the genomic changes known to occur in cancer.

### Moving towards clinical applications

The central role of mutated cancer genes in mediating the proliferation and survival of cancer cells render them potential ‘Achilles’ heels’ to be exploited for biomarker discovery. Many thousands of patients each year are recruited into clinical trials of new cancer drugs and almost every drug used in the clinic today has been part

of this process. These clinical trials offer both an opportunity and a challenge. On the one hand, there are tumour banks linked to meticulously collected clinical data, including information identifying which patients responded to the therapy. On the other hand, this tissue is invariably in the form of formalin-fixed paraffin-embedded (FFPE) samples, with issues such as DNA degradation and contamination with necrotic or apoptotic cells that preclude their analysis by conventional next-generation sequencing protocols. However, it appears increasingly possible to develop assays that specifically target only the exons of mutated cancer genes for next-generation sequencing and to do this using FFPE tissue samples. This area is currently under investigation by several research groups and it appears probable that, in the future, many hundreds of genes can be profiled from each tumour in a cost-efficient manner [16]. In the context of clinical trials, the application of next-generation sequencing technology to archived FFPE samples with linked response and survival data opens up the possibility of improving the use of current drugs and potentially repositioning compounds that were initially unsuccessful during early-phase patient trials. Additionally, we argue that next-generation sequencing of tumour samples, where possible, should be considered as standard practice to inform the outcome of all new clinical trials.

The need to integrate genomic biomarkers effectively into routine clinical practice for the treatment of cancer represents an equally large challenge. As mentioned above, a limited repertoire of genetic biomarkers is already used for patient stratification during treatment, such as the use of imatinib for the treatment of patients with CML bearing the *BCR-ABL* gene fusion; however, genetic biomarkers are not widely used in general (Table 1). This appears likely to change in the years to come; for example, since 2009, Massachusetts General Hospital in Boston has already genetically tested >1000 cancers for >100 recurrent mutations across 19 cancer-related genes. In several instances, these mutations are linked to clinical end-points and can act as biomarkers to help identify the best course of treatment. Although the impact of many mutations on clinical response to treatment is currently unknown, the generation of large data sets linking patient response with genetic information might ultimately prove invaluable in identifying new genomic biomarkers. In a similar effort, the National Health Service (NHS) of the UK has launched a stratified medicine programme to use genetic information to group patients based on their probable response to treatment. A phase I



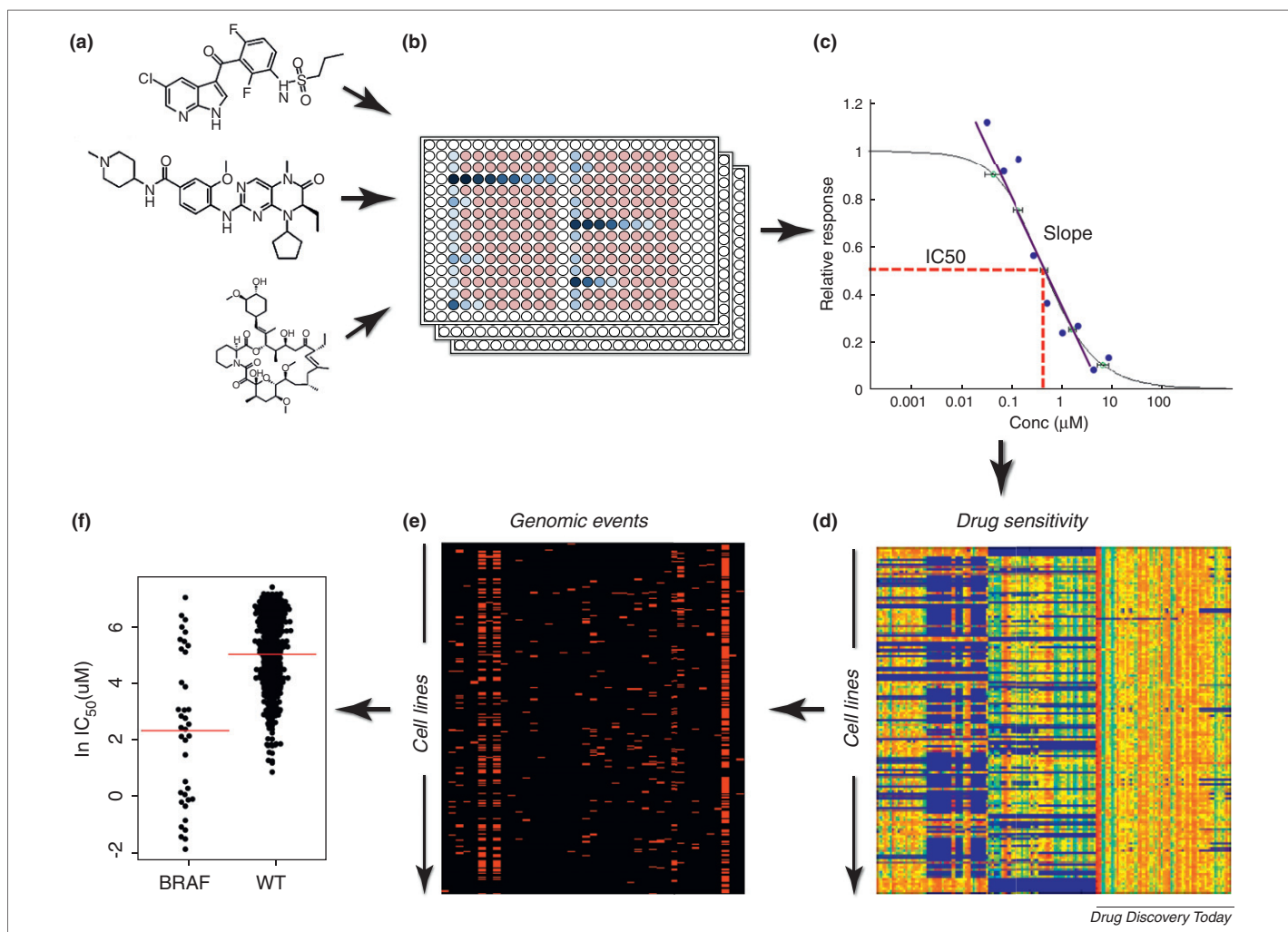


FIGURE 2

A schematic workflow for high-throughput cell-line screening to identify genomic features associated with drug response. **(a)** A library of clinical and preclinical cancer drugs is screened **(b)** against a panel of cancer cell lines in 384-well plate format. **(c)** Drug sensitivity is measured over a nine-point titration of drug concentration and a curve-fitting algorithm is used to derive a signature of response, including the half-maximal inhibitory concentration (IC<sub>50</sub>) and slope of the dose response curve. **(d and e)** Drug-sensitivity data are correlated with genomic features, including point mutations, gene amplification and deletion, as well as with gene-expression data to identify genomic features associated with drug sensitivity. **(f)** In this example, the association between v-ras murine sarcoma viral oncogene homologue B1 (*BRAF*) mutations and sensitivity to a *BRAF* inhibitor is identified. Each circle represents the IC<sub>50</sub> (natural log) of a single cell line and the red line indicates the mean IC<sub>50</sub> for *BRAF*-mutated or wild-type (WT) cell lines.

feasibility study will involve testing for approximately 20 genetic markers already linked to treatment response in >9000 colorectal, breast, prostate, lung, ovarian and skin cancers, with the ultimate goal of rolling out the service nationally.

To be most effective in the clinical setting, these types of service will need to identify accurately specific genetic changes in tumours and provide clear information in a timely manner to assist oncologists in determining the best course of treatment for their patients. Indeed, the emergence of web-based tools with the aim of matching cancer mutations to specific therapies are another indication of the fundamental shift that is occurring in the approach to the diagnosis and treatment of patients with cancer by embracing the use of genetic

biomarkers (e.g. <http://cancercommons.org/> and <http://www.mycancergenome.org/>).

### Concluding remarks

The transformation in understanding of cancer genetics that is currently underway will have the greatest utility if combined with biological assays where one can begin to interpret the functional significance of cancer gene mutations. This is a crucial unmet need at present if one is to begin to determine the extent and contexts in which alterations in cancer genomes are able to subvert normal cellular processes. One approach already discussed has been to screen a large panel of cancer cell lines that are being subjected to the same level of comprehensive genomic characterisation as human tumour samples, against a

range of cancer drugs and to correlate drug response with mutational status. These, and other systematic efforts, will provide a rich source of data-led hypotheses for clinicians and scientists alike and are the inevitable next step in the journey from an understanding of the cancer genome to improving the survival of patients with cancer.

### References

- 1 Stratton, M.R. *et al.* (2009) The cancer genome. *Nature* 458, 719–724
- 2 Hanahan, D. and Weinberg, R.A. (2000) The hallmarks of cancer. *Cell* 100, 57–70
- 3 Futreal, P.A. *et al.* (2004) A census of human cancer genes. *Nat. Rev. Cancer* 4, 177–183
- 4 Druker, B.J. *et al.* (2006) Five-year follow-up of patients receiving imatinib for chronic myeloid leukemia. *N. Engl. J. Med.* 355, 2408–2417

- 5 Cobleigh, M.A. *et al.* (1999) Multinational study of the efficacy and safety of humanized anti-HER2 monoclonal antibody in women who have HER2-overexpressing metastatic breast cancer that has progressed after chemotherapy for metastatic disease. *J. Clin. Oncol.* 17, 2639–2648
- 6 Vogel, C.L. *et al.* (2002) Efficacy and safety of trastuzumab as a single agent in first-line treatment of HER2-overexpressing metastatic breast cancer. *J. Clin. Oncol.* 20, 719–726
- 7 Slamon, D.J. *et al.* (1989) Studies of the HER-2/neu proto-oncogene in human breast and ovarian cancer. *Science* 244, 707–712
- 8 Mok, T.S. *et al.* (2009) Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. *N. Engl. J. Med.* 361, 947–957
- 9 Kwak, E.L. *et al.* (2010) Anaplastic lymphoma kinase inhibition in non-small-cell lung cancer. *N. Engl. J. Med.* 363, 1693–1703
- 10 Chapman, P.B. *et al.* (2011) Improved survival with vemurafenib in melanoma with BRAF V600E mutation. *N. Engl. J. Med.* 364, 2507–2516
- 11 Stratton, M.R. (2011) Exploring the genomes of cancer cells: progress and promise. *Science* 331, 1553–1558
- 12 Hudson, T.J. *et al.* (2010) International network of cancer genome projects. *Nature* 464, 993–998
- 13 McDermott, U. *et al.* (2007) Identification of genotype-correlated sensitivity to selective kinase inhibitors by using high-throughput tumor cell line profiling. *Proc. Natl. Acad. Sci. U.S.A.* 104, 19936–19941
- 14 Shoemaker, R.H. *et al.* (1988) Development of human tumor cell line panels for use in disease-oriented drug screening. *Prog. Clin. Biol. Res.* 276, 265–286
- 15 Weinstein, J.N. *et al.* (1997) An information-intensive approach to the molecular pharmacology of cancer. *Science* 275, 343–349
- 16 Ross, J.S. *et al.* (2011) Comprehensive next-generation sequencing for clinically actionable mutations from formalin-fixed cancer tissues. *J. Clin. Oncol.* 29 (Suppl.), abstr 10564
- 17 Druker, B.J. *et al.* (2001) Efficacy and safety of a specific inhibitor of the BCR-ABL tyrosine kinase in chronic myeloid leukemia. *N. Engl. J. Med.* 344, 1031–1037
- 18 Verweij, J. *et al.* (2004) Progression-free survival in gastrointestinal stromal tumours with high-dose imatinib: randomised trial. *Lancet* 364, 1127–1134
- 19 Talpaz, M. *et al.* (2006) Dasatinib in imatinib-resistant Philadelphia chromosome-positive leukemias. *N. Engl. J. Med.* 354, 2531–2541
- 20 Flaherty, K.T. *et al.* (2010) Inhibition of mutated, activated BRAF in metastatic melanoma. *N. Engl. J. Med.* 363, 809–819

Mathew J. Garnett\*,  
Ultan McDermott

Cancer Genome Project, Wellcome Trust Sanger  
Institute, Hinxton CB10 1SA, UK