**Title**
What are we learning from the cancer genome?



**Journal**
Nature Reviews Clinical Oncology, 9(11)

**ISSN**
1759-4774

**Authors**
Collisson, Eric A
Cho, Raymond J
Gray, Joe W

**Publication Date**
2012-11-01

**DOI**
10.1038/nrclinonc.2012.159

Peer reviewed

Published in final edited form as:
*Nat Rev Clin Oncol*. 2012 November ; 9(11): 621–630. doi:10.1038/nrclinonc.2012.159.

# What are we learning from the cancer genome?

**Eric A. Collisson**,
Department of Medicine/Hematology and Oncology, University of California San Francisco, 505 Parnassus Avenue, San Francisco, CA 94143, USA

**Raymond J. Cho**, and
Department of Dermatology, University of California San Francisco, 1701 Divisadero Street,4th Floor, San Francisco, CA 94115, USA

**Joe W. Gray**
Oregon Health & Science University, Knight Cancer Institute, CH13B 3303 South West Bond Avenue, Room #13000, Portland, OR 97239, USA

## Abstract

Massively parallel approaches to nucleic acid sequencing have matured from proof-of-concept to commercial products during the past 5 years. These technologies are now widely accessible, increasingly affordable, and have already exerted a transformative influence on the study of human cancer. Here, we review new features of cancer genomes that are being revealed by large-scale applications of these technologies. We focus on those insights most likely to affect future clinical practice. Foremost among these lessons, we summarize the formidable genetic heterogeneity within given cancer types that is appreciable with higher resolution profiling and larger sample sets. We discuss the inherent challenges of defining driving genomic events in a given cancer genome amidst thousands of other somatic events. Finally, we explore the organizational, regulatory and societal challenges impeding precision cancer medicine based on genomic profiling from assuming its place as standard-of-care.

## Introduction

Our understanding of the human genome and the myriad ways its deregulation contributes to human cancer is expanding rapidly as a result of remarkable advances in genome analysis technologies. The most-established methods analyse genome copy number,[1,2] single nucleotide polymorphism content,[3] or gene expression,[4] by examining quantitative hybridization of labelled nucleic acid fragments from a cancer and comparing the results with representations of the normal genome on a bead or chip. Hybridization intensity at each array element gives a relative measure of nucleic acid abundance. Hybridization conditions can be adjusted to distinguish even single-base mismatches between tumour DNA and target oligo-nucleotide, providing some information about sequence (for example, the presence of a specific DNA sequence polymorphism or mutation). However, hybridization-based

Correspondence to: J. W. Gray grayjo@ohsu.edu.

Author contributions: All authors researched data for and wrote the article, made a substantial contribution to discussion of the content, and edited the manuscript prior to submission.

approaches are complemented and may soon be replaced by massively parallel sequencing strategies,[5] where tumour and normal reference DNA or RNA samples are converted into libraries of billions of small nucleic acid fragments, the ends of which are sequenced using these new technologies.[6] The genomic locations of the ends of each clone are either computationally mapped onto either the normal genome[7] or reassembled *de novo*.[8] Through comparison to matched normal samples, these data identify mutations,[9] genome copy number changes,[10] and structural rearrangements[11] in the genome. Massively parallel technologies can also be used with RNA as the input (RNAseq) and can not only quantify gene expression,[12,13] but also detect genomically encoded variants,[14] alternative splicing,[15] and transcribed fusion genes.[16]

These genome analysis techniques are now being used in international cancer genome analysis efforts to catalogue aberrations driving the pathophysiology of nearly all major cancer types. The Cancer Genome Atlas (TCGA) project[17] and the International Cancer Genome Consortium (ICGC)[18,19] represent the largest of such efforts. The TCGA project is assessing aberrations in 500 to 1,000 tumours from each of approximately 20 major human cancer types. The TCGA and ICGC together are currently committed to analysis of approximately 47 separate tumour lineages. Importantly, the results from these projects are being made widely available for community interpretation in near real time. Prototypical results for glioblastoma,[20] serous ovarian cancer,[7] colorectal cancer,[21] squamous-cell lung cancer,[22] and breast cancer are already available from the TCGA project, and other technology centres around the world have contributed studies of cancers of the pancreas,[23] prostate,[24] lung,[25] kidney,[26] melanoma,[27] myeloma,[28] and acute myeloid leukaemia (AML),[29] to name a few. The breast cancer genomic landscape has been particularly well defined.[30–38] As the cost decreases and the reliability of genome-scale analysis technologies increases, such studies become possible in an increasing number of research and clinical settings. In this Review article, we discuss the current challenges, which include functional interpretation of the results; management and interpretation of the resulting large-scale data sets; and development of genome analysis-based studies that result in clinically validated strategies for precision cancer treatment.

## Pace of discovery and translation

The first human genome was sequenced to about 85% completion at a cost of hundreds of millions of dollars and required more than a decade of cross-institutional work.[39,40] Today, we can realistically expect a single user to sequence the coding portion of the genome (hereafter referred to as the exome) for about US$1,000 in a few hours using a tabletop machine.[41]Figure 1 summarizes some of these technologies. As a consequence of these and other advances, the rate at which genome discoveries are being translated to clinical practice is accelerating (Table 1). We required nearly three decades to leverage the discovery of the Philadelphia chromosome encoding *BCR–ABL*[41] into the approval of imatinib for the treatment of chronic myeloid leukaemia.[43] By comparison, crizotinib[43] was approved for the treatment of non-small-cell lung cancer (NSCLC) harbouring *EML4–ALK* translocations less than 4 years from the discovery of *EML4–ALK*[44] as a recurrent driver aberration. However, the success rate for the development of approved drugs for cancer is still disappointingly low.[46] New strategies are needed to efficiently identify molecular signatures associated with

therapeutic response. Here, we describe international efforts underway to catalogue omic aberrations in major cancer types, summarize their incipient impact on cancer management, and discuss the challenges that impede translation of genomic discoveries into improved cancer care. We focus on genomic aberrations, while recognizing that epigenomic aberrations[46] and microenvironment changes[47] likely have equally important roles in the genesis and progression[48] of many malignancies.

## Lessons learned

International cancer genomics efforts have demonstrated several general and noteworthy features of cancer genomes.

## Driver aberrations

A typical solid tumour can contain tens to thousands of genomic aberrations that include: somatic and germline changes in copy number (gain, loss, homozygous deletion and high-level amplification) that increase or decrease the levels of important coding and noncoding RNA transcripts; somatic mutations and naturally occurring polymorphisms that alter protein structure, protein stability and/or change the way transcripts are spliced; and structural changes (such as translocations, deletions or inversions) that change transcript levels by altering gene-promoter associations or creating new fusion genes. These changes include ‘driver’ aberrations selected for during tumour progression that enable oncogenesis, progression or other cancer hallmarks.[50] Driver aberrations exist alongside ‘passenger’ aberrations that do not contribute to cancer pathophysiology but arise by chance in an unstable tumour. Cancer hallmarks that may be deregulated by driver aberrations include aspects of cell proliferation, death, motility, angio-genesis, differentiation, DNA replication and repair, and senescence. The Sanger Institute now maintains a searchable, up-to-date catalogue of somatic mutations in cancer.[51] Driver aberrations are thought to modulate the behaviour of three classes of cancer genes: oncogenes, tumour-suppressor genes, and DNA-integrity genes.[52]

Oncogenes contribute to the cancer phenotype when their activity is enhanced, usually by constitutive activation by missense mutation or overexpression caused by gene amplification. Examples of oncogenic aberrations include recurrent mutations affecting *KRAS, BRAF, EGFR,* and *PIK3CA*, amplifications in *HER2* and *MYC*, and fusion genes involving *MYC*, *ABL*, ETS family transcription factors and *ALK*.[52] The products of oncogenes are attractive therapeutic targets because inhibition of these aberrant or overabundant proteins is likely to either reverse the oncogenic effect of the driver gene or result in the death of cancer cells addicted to these aberrations.[53] Driver mutations now being discovered tend to occur at a low frequency (often <5%),[54,55] which indicates the efficiency of past, biologically motivated studies in the identification of genes capable of cellular transformation.[56] However, important mutations in *BRAF* in hairy-cell leukaemia,[57] *GRM3*[57] and *TRAPP*[58] in melanoma, and *IDH1* in glioblastoma[59] and AML[60] were discovered in large-scale sequencing studies and are present in meaningfully large fractions of these tumour types.

Tumour-suppressor genes typically act as negative regulators of cancer hallmarks such as cell-cycle progression and intracellular signalling, so that their loss or inactivity contributes to cancer progression. Classic examples of tumour-suppressor genes that are inactivated by genomic aberrations include *RB1*, *TP53*, *NF1*, *APC*, and *CDKN2A*. Inactivation of a single allele contributes to cancer pathophysiology via haploinsufficiency in some cases, but pronounced phenotypes usually result from inactivation of both tumour-suppressor gene alleles.[62] Aberrations in apparent tumour-suppressor genes recently discovered in large-scale cancer genome studies involve the chromatin modification genes *SMARCA4*, *UTX1*, *SETD2*, *KDM5C*,[26] *PBRM1*[62] and *ARID1A*.[64] It is generally more difficult to develop therapeutic strategies to attack cancers associated with mutations in tumour-suppressor genes since this often requires restoration of function of the wild-type gene, an inherently more difficult task than inhibiting an activated kinase, for example.

DNA integrity genes are involved in aspects of DNA repair or replication. Deregulation of these genes can result in chromosome instability and/or inability to repair some classes of DNA damage. The end result is either a rearranged cancer genome with copy number changes and translocations though the chromosomal break-fusion-break cycle or many point mutations in an other wise diploid genome, or both. Prototypical examples include mutations of *BRCA1*, *BRCA2*, *ATM*, *MSH2*, *MLH1*, and *XPA* that lead to cancer hyper-mutation phenotypes. A mutation in *PALB2* in hereditar y and spontaneous pancreatic cancers is another such example.[65] Aberrations involving these genes are often germline and contribute to cancer susceptibility in kindred, but also occur somatically outside of cancer-prone families. Some therapeutic agents attempt to exploit the inability of cancer cells to repair some classes of DNA damage, a so-called synthetic lethal relationship.[66] For example, small-molecule inhibitors of poly(ADP-ribose) polymerase (PARP) sensitize cancers carrying *BRCA1* and/or *BRCA2* mutations to chemotherapy,[67] taking advantage of the fact that the cells with *BRCA1* and/or *BRCA2* mutations are not able to repair DNA damage induced by chemotherapy by homologous recombination.

## Intertumour heterogeneity

The view of the cancer genome that is now emerging from the TCGA and other studies is that the number of genes involved in the genesis of an individual tumour is the sum of the few common drivers (about which much is known), plus tens to thousands of plausible driver aberrations that occur rarely or are idiosyncratic to an individual tumour and about which little is known.[7,20,22] It is likely that the ensemble of aberrations in each individual tumour collaborates in myriad ways to deregulate common signalling pathways, thereby enabling established cancer ‘hallmarks’.

Intertumour heterogeneity is likely caused by the stochastic nature of genome damage that occurs during passage through telomere crisis,[68] differences in tumour-initiating insults, immune surveillance and other aspects of biology that influence cancer progression, and differences in the normal progenitor cell in which each cancer originates. The impact of the outside world is particularly clear; cancers associated with carcinogen exposure (for example, melanoma,[9] lung cancer,[22] and skin cancer[69]) typically have much higher mutation rates than childhood and haematological malignancies.[61] In what might be the

strongest anti-smoking message yet, an unlucky 40 pack-year smoker can expect about 10 mutations in their squamous-cell lung cancer genome per cigarette smoked.[22] Likewise, the age of the cell from which a cancer originates also likely impacts the number of mutations present.[70] Recent DNA sequence-based genome surveys demonstrate that information on cancer aetiology can be gleaned from the nature of the associated mutations. For example, mutations caused by alkylating agents typically demonstrate a greater proportion of base transversions than those induced by exposure to UV radiation or smoking.[9,69]

Importantly, driver abnormalities tend to be specific to a type of cancer. Translocations, for example, are common in haematological tumours and lymphomas, and mutations and copy-number changes may have a more limited role.[70] Somatic mutations are more frequent in epithelial cancers than in haematological malignancies and are often specific to a given carcinoma or subtype of cancer. *KRAS* mutations occur frequently in pancreatic and colorectal cancer,[71,72] whereas *NRAS* and *HRAS* are commonly mutated in melanoma[73] and bladder cancer,[74] respectively. *TP53* mutations occur in many cancers, but are nearly universal in serous ovarian cancer [7] and most squamous-cell carcinomas.[69,75]*EGFR* mutations occur in subsets of lung adenocarcinomas[76,77] and glioblastoma.[78]*PIK3CA* mutations tend to occur preferentially in oestrogen receptor-positive breast cancer, while *TP53* mutations occur more frequently in oestrogen receptor-negative breast cancer.[32,36,79] The spectrum of copy number abnormalities also varies between tumour types. Glioblastoma[20] and luminal A breast cancer [80] typically show relatively few genome copy number changes, but serous ovarian cancers[7] and basal-like breast cancers show many.[32] The prevalence of specific aberration type is likely determined at least in part by the epigenetically determined regulatory structures that must be deregulated to enable cancer tumour formation.

## Intratumour heterogeneity

Substantial genomic heterogeneity exists even within individual tumours. This heterogeneity is revealed using massively parallel sequencing-based approaches that detect even very-low-abundance aberrations in individual tumours and is thought to arise as a result of clonal divergence during tumour progression.[81,82] In the most parsimonious hypothesis, aberrations that occur early in tumour progression appear in most cells of the tumour, while later-occurring aberrations appear only in a subpopulation of the tumour.[29,79] Ding *et al.*[81] showed individual, somatically mutated alleles to be present in anywhere from 5% to 90% of recovered sequencing reads from a primar y tumour, implying the presence of complex subclonal populations in the tumour. Navin *et al.*[83] profiled many single nuclei from two distinct breast carcinomas; they showed that while one carcinoma was comprised of a nearly uniform dominant population, the other carcinoma had several related but independent subclones, all apparently coexisting at a single anatomic site.[83] In the most compelling example of the importance of tumour heterogeneity within a single patient, Gerlinger *et al.*[84] showed that subclones of an individual kidney cancer co-evolve to inactivate tumour-suppressive pathways in different ways.[84] This intramalignancy heterogeneity is likely the result of instability during tumour progression, generating the critical population diversity needed to evolve resistance to therapy. It also signals a formidable adaptive capacity of cancer genomes that must be incorporated into any molecular treatment strategy.[85]

These findings have important therapeutic ramifications, since the opportunity for the emergence of resistant subclones is almost certainly related to the diversity of the treated population.[86] In treating metastatic cancer, oncologists often base treatment decisions on the genomic features of the primary tumour. How tumours evolve over time and under the selective pressure of effective therapy is an active area of research.[87] One concept to organize thinking about heterogeneity in cancer has been introduced, modelled on central precepts in organismal evolution. In this model, individual cells or anatomical areas of a given patient's cancer can be categorized as either ‘palm trees’ or ‘oak trees’ based on the shape of the phylogenetic trees that describe their shared and derived genomic traits (Figure 2).[88] The palm tree model describes relatively late divergence enforcing similarity between primary and metastatic cancers within an individual. By contrast, the oak tree model suggests early genomic divergence, such that the primary tumour and dispersed metastatic lesions can differ substantially both in genome content and, by extension, cellular response to therapy.

In tumours with early genomic divergence (that is, oak trees), it may be critical to identify and therapeutically target genomic aberrations arising early in the evolutionary process (that is: founder mutations) otherwise transient clinical responses will give way to inevitable drug resistance. The order of occurrence of aberrations can be inferred from longitudinal studies of tumour progression[89] or by analysis of the interplay between somatic mutations and genome copy number changes in individual tumour samples.[69,83] The spectrum of aberrations also may differ between metastatic lesions within an individual patient,[90] indicating that different lesions within one patient may evolve independently once they escape from the primary tumour, as illustrated in Figure 2.

The origin of heterogeneity between a given patient's unique tumour genomes is likely multifaceted. Ongoing selective pressures are applied to tumour cells at both the primary[79] and metastatic[81] sites. Omic events that are essential for tumour initiation may be dispensable, or even deleterious, in the progression or metastatic process. Further longitudinal studies of the cancer genome will educate both our treatment as well as our screening and post-operative surveillance approaches.

## Therapeutic targeting

Precision cancer treatment is based on the concept that tumour cells with specific aberrations will respond homogeneously to aberration-specific treatments. The identification of recurrent genome aberrations that define cancer subsets has led to the approval of therapies that target the aberrant genes or pathways deregulated by the aberrations (Table 1). Recurrent aberrations now being discovered by international cancer genome efforts present a growing number of new therapeutic targets. Some abnormalities occur frequently, especially within a given tumour type.[91] However, many more occur at relatively low frequency in individual tumour types.[55] The latter observation suggests an omic aberration-centric rather than a histology-centric approach to therapeutic development will be successful, based on the logic that a druggable aberration present at low frequency in many tumour types is still an attractive therapeutic target when aggregated across diseases. The success of crizotinib in *ALK* rearranged NSCLC and lymphoma supports this notion,[92] but the relative inefficacy of

vemurafenib in *BRAF* V600E-mutant colon cancer [93] reinforces the message that not every mutant kinase detected is an indication for a drug in all tumour types. Finally, some genes have apparently opposing roles in oncogenesis. The Notch family of receptors is one such example, being mutated with high frequency in both haematological cancers[94] and squamous-cell carcinomas.[75,95] The mutations activate the receptor in the former scenario and inactivate it in the latter (Figure 3), reminding us that context matters, and functional characterization must follow omic discovery for therapeutic benefit to emerge.

A major challenge in assessing individual genome aberrations as potential therapeutic targets is our limited understanding of their roles in cancer pathophysiology. Cancer depends on some aberrations so strongly that therapeutic intervention results in a strong and clinically meaningful apoptotic response. Other driver aberrations influence cancer hallmarks such as motility and proliferation so that therapeutic interventions may not translate into even short-term improvements in outcome. Conversely, others may lead to terminal differentiation that may not result in immediate tumour regression, but still profoundly increase long-term survival. In addition to these protein-driven changes, other genetic aberrations likely have effects on the epigenome[63] or RNA splicing,[96] and lack obvious phenotypical criteria to functionally assay at present. Our clinical trials need flexibility beyond the standard assessment tool of Response Evaluation Criteria In Solid Tumors (RECIST) to encompass these possibilities. Early examples to guide us include metabolic imaging[97] or assessment of circulating cells[98] as embedded correlates in clinical trials.

## Pathway-based genomics

The concept that many diverse events in many tumours converge on common pathways to achieve discrete hallmarks suggests the possibility of managing heterogeneity between and within cancers by organizing the spectrum of aberrations within an individual tumour into deregulated pathways. A growing number of pathway analysis tools are now becoming available to organize diverse candidate driver aberrations into functional networks.[99–105] Several recent publications document this approach.[7,106,107] They typically either identify pathways curated from the literature that harbour more genomic aberrations than would be expected by chance, or use genomic information to calculate pathway activity scores that take into account the regulatory ‘directionality’ that is known from the curated literature. Figure 4 shows a recent example of the use of genomic information to calculate pathway activities for more than 1,400 pathways. This information was combined into one superpathway identifying subnodes that differ in activity between two comparison populations (for example, tumour versus normal or drug responsive versus drug resistant). Although promising, the experimental validity of postulated networks and the evidence linking them to specific cancer hallmarks—especially hallmarks other than those involving survival and proliferation—so far is modest. Substantial work will be needed to fully validate this approach, but it is worth the effort since it provides a framework on which to organize an otherwise unmanageably large number of aberrations. Organization of aberrations into networks facilitates drug development because it guides development of therapies that attack common networks, often in different ways. The number of important networks that are deregulated by genome aberrations may well be manageably small, but the

number of mechanisms by which genome aberrations can deregulate these pathways is enormous. Several large-scale systems biology efforts are now underway to identify and validate regulatory networks that are deregulated by genomic aberrations, and to use this information to guide the development of pathway-targeted therapies.[99,106,108]

## Genome diagnostics

Identification and validation of a druggable genomic aberration or pathway it deregulates sets the stage for a clinical trial designed to attack the aberration. This development requires molecular assays to identify the tumours harbouring the target aberration or deregulated pathway. This task is relatively straightforward if the target is a single genomic aberration, such as those listed in Table 1. In these cases, well-established single gene or protein assays, such as immunohistochemistry, PCR and/or fluorescence *in situ* hybridization (FISH), can be applied and interpretation is straightforward. These assays typically are amenable to tissue prepared for standard histopathological analysis (that is, formalin-fixed and paraffin-embedded samples). Many single-gene assays are now routine (for example, FISH-based tests[109] for translocations and amplifications) and an increasing number of laboratories are providing services that assay for actionable mutations.[110,111]

Assay development is somewhat more difficult if the target is a deregulated pathway defined by a multigene signature requiring an interpretative algorithm that converts the measurements into a metric of pathway activity, or score. A few multigene assays are now in routine use.[112] In all cases, assays require careful standardization and validation—especially if they are to be used in ways that impact patient management. Unfortunately, we have not done well coordinating co-development of diagnostics alongside therapeutics, especially for early phase trials where the targeting concept may come entirely from preclinical studies.[113] Strategies for simultaneous development of markers and therapies will continually be refined in coming years. In the meantime, it is important that the information and algorithms used for marker development be available for independent scrutiny.

## The genomic ‘call’

It seems feasible to soon analyse transcriptional changes, copy-number abnormalities and mutations in a few genes or the entire exome in individual tumours overnight at reasonable cost. We can do so using samples consisting of only a few thousand cells and in formalin-preserved materials. The cost of sequencing is dropping so quickly that storing and interpreting data will soon exceed the cost of producing it. Thus, fairly complete genomic analysis of individual tumours seems likely to become a part of the standard-of-care. The utility of such data is only realized when a medical professional changes a medical decision as a result of the genomic ‘call’. Therefore, it is important that researchers in and around the clinical arena identify the real challenges facing implementation of personalized medicine in the coming years. These challenges are many but, in our opinion, the ones with the most impact are: secure and confidential data storage, development of widely available strategies to interpret large-scale genomic data in actionable format, appropriate regulation of genomic testing so that progress continues without harm to patients, and, finally, understanding the impact of treatment on underlying genomic and biological heterogeneity.

Most clinicians (as well as most basic scientists) lack the training to interpret the huge amounts of raw data produced by new sequencing technologies and many institutions are ill-equipped to manage it safely and confidentially: it is a daunting prospect. The amount of data generated by full genome sequencing could exceed 1,000 petabytes per year if full genome sequencing becomes part of the standard-of-care. By comparison, Google is estimated to process only about 24 petabytes of user-generated data per day. These data must be managed securely, properly compressed, linked to metadata, clinically interpreted and distributed to the medical community under increasingly strict privacy requirements.[114] Gaining full benefit from these data will require a substantial infrastructure that does not currently exist. Acquisition, storage, disseminated analysis and convergent communication of large datasets are commonplace in large physical sciences projects and such efforts provide useful templates for the biomedical community.[115]

Once data is interpreted, it is essential that user interfaces succinctly and accurately communicate results from omic testing to the treating physician and patient without oversimplifying the underlying message. Centralizing and standardizing key aspects of interpretation (for example, mutation calling) will simultaneously allow for quality control as well as foster the building of large disease databases for cross-centre and cross-disease comparisons of the interplay between genotype, treatment and clinical outcome. Although daunting, early examples of feasibility already exist. For example, Jones *et al*.[116] demonstrated that medical interpretation of RNA and DNA sequencing can lead to meaningful treatment outcomes, in this case for a single patient with a rare adenocarcinoma.[116]

Genomic cancer data comes from nucleic acids, which (usually) come from biopsies. Practicing clinicians are intimately involved in the crucial biopsy acquisition decision. In a patient considering therapy for metastatic cancer, often recurrent after an initial resection, this decision is often summarized as whether to rebiopsy a metastatic site or to rely on the primary (resected) tissue for omic analysis. The heterogeneity of cancer within an individual supports the rebiopsy approach for many reasons. Interestingly, the known heterogeneity of oestrogen receptor, progesterone receptor, and *HER2* status in breast cancer has made rebiopsy of metastatic sites common,[117,118] whereas most clinicians rely on a resection sample for *KRAS* genotyping in colorectal cancer. Since so many parties (surgeons, radiologists, pathologists, and so on) are involved in the treatment decisions that increasingly rely on genomic results, we should agree that the rebiobsy approach is the right one and obtain tissue as proximate to the current treatment arena as possible. This approach will increase the accuracy and relevance of omic findings in a particular patient's tumour and hopefully make the increased costs and inconvenience of a repeat biopsy worthwhile.

Finally, our understanding of the clinical consequences of genomic heterogeneity is limited. Modern single molecule-based genomic analysis tools provide the ability to detect heterogeneity in aberration content within and between lesions in a single patient. What is lacking is a strategy for when an actionable aberration or aberrant pathway is detected in a small fraction of tumour cells in a patient. Answering this question will require that large-scale genome analyses that are sensitive to low-frequency aberrations be applied to pathway-targeted trials, in order to impact on outcome. These trials will require a substantial

commitment to genome analysis by the international clinical trials and genomic science communities alike.

## Conclusions

In summary, the advent of technologies to read and enumerate nucleic acids is seeing the cancer genome yield some long-held secrets. After over a century of preamble, cancer biology is only now compiling its first complete set of ‘mission statements’, as landmark genomes of each cancer type are systematically deciphered. The charge going forward is no longer to read the genome, but to understand the differences between causal and benign mutations in a given cancer, between mutations that occur in most tumours and those that occur in most cells of a single tumour, and finally to develop treatments to counteract these abnormalities, at the levels of the gene, cell, tissue and patient.

## Acknowledgments

The authors thank Sam Ng, Josh Stuart and Barry Taylor for assistance, comments and figures. This work was supported by the National Institutes of Health, National Cancer Institute grant P50 CA 58207 and the U54 CA 112970, K08CA137153.

## References

1. Pinkel D, et al. High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. Nat Genet. 1998; 20:207–211. [PubMed: 9771718]
2. Pollack JR, et al. Genome-wide analysis of DNA copy-number changes using cDNA microarrays. Nat Genet. 1999; 23:41–46. [PubMed: 10471496]
3. Lindblad-Toh K, et al. Loss-of-heterozygosity analysis of small-cell lung carcinomas using single-nucleotide polymorphism arrays. Nat Biotechnol. 2000; 18:1001–1005. [PubMed: 10973224]
4. Perou CM, et al. Distinctive gene expression patterns in human mammary epithelial cells and breast cancers. Proc Natl Acad Sci USA. 1999; 96:9212–9217. [PubMed: 10430922]
5. Morozova O, Marra MA. Applications of next-generation sequencing technologies in functional genomics. Genomics. 2008; 92:255–264. [PubMed: 18703132]
6. Chin L, Hahn WC, Getz G, Meyerson M. Making sense of cancer genomic data. Genes Dev. 2011; 25:534–555. [PubMed: 21406553]
7. Cancer Genome Atlas Research Network. Integrated genomic analyses of ovarian carcinoma. Nature. 2011; 474:609–615. [PubMed: 21720365]
8. Li Y, et al. Structural variation in two human genomes mapped at single-nucleotide resolution by whole genome *de novo* assembly. Nat Biotechnol. 2011:29, 723–730.
9. Pleasance ED, et al. A comprehensive catalogue of somatic mutations from a human cancer genome. Nature. 2010; 463:191–196. [PubMed: 20016485]
10. Bignell GR, et al. Architectures of somatic genomic rearrangement in human cancer amplicons at sequence-level resolution. Genome Res. 2007; 17:1296–1303. [PubMed: 17675364]
11. Campbell PJ, et al. Identification of somatically acquired rearrangements in cancer using genome-wide massively parallel paired-end sequencing. Nat Genet. 2008; 40:722–729. [PubMed: 18438408]
12. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods. 2008; 5:621–628. [PubMed: 18516045]
13. Morin R, et al. Profiling the HeLa S3 transcriptome using randomly primed cDNA and massively parallel short-read sequencing. Biotechniques. 2008; 45:81–94. [PubMed: 18611170]
14. Morin RD, et al. Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin. Nat Genet. 2010; 42:181–185. [PubMed: 20081860]

15. Griffith M, et al. Alternative expression analysis by RNA sequencing. Nat Methods. 2010; 7:843–847. [PubMed: 20835245]
16. Pflueger D, et al. Discovery of non-ETS gene fusions in human prostate cancer using next-generation RNA sequencing. Genome Res. 2011; 21:56–67. [PubMed: 21036922]
17. National Cancer Institute. The Cancer Genome Atlas. 2012. [online]. http://cancergenome.nih.gov
18. International Cancer Genome Consortium. ICGC Cancer Genome Projects. 2012. [online]. http://www.icgc.org
19. Bhat KP, et al. The transcriptional coactivator TAZ regulates mesenchymal differentiation in malignant glioma. Genes Dev. 2011; 25:2594–2609. [PubMed: 22190458]
20. Cancer Genome Atlas Research Network. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. Nature. 2008; 455:1061–1068. [PubMed: 18772890]
21. Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. Nature. 2012; 487:330–337. [PubMed: 22810696]
22. The TCGA Investigators. A comprehensive genomic characterization of squamous cell lung cancers. Nature. http://dx.doi.org/10.1038/nature11404
23. Jones S, et al. Core signaling pathways in human pancreatic cancers revealed by global genomic analyses. Science. 2008; 321:1801–1806. [PubMed: 18772397]
24. Prensner JR, et al. Transcriptome sequencing across a prostate cancer cohort identifies PCAT-1, an unannotated lincRNA implicated in disease progression. Nat Biotechnol. 2011; 29:742–749. [PubMed: 21804560]
25. Hammerman PS, et al. Mutations in the DDR2 kinase gene identify a novel therapeutic target in squamous cell lung cancer. Cancer Discov. 2011; 1:78–89. [PubMed: 22328973]
26. Dalgliesh GL, et al. Systematic sequencing of renal carcinoma reveals inactivation of histone modifying genes. Nature. 2010; 463:360–363. [PubMed: 20054297]
27. Berger MF, et al. Melanoma genome sequencing reveals frequent PREX2 mutations. Nature. 2012; 485:502–506. [PubMed: 22622578]
28. Chapman MA, et al. Initial genome sequencing and analysis of multiple myeloma. Nature. 2011; 471:467–472. [PubMed: 21430775]
29. Ding L, et al. Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. Nature. 2012; 481:506–510. [PubMed: 22237025]
30. Gray J, Druker B. Genomics: the breast cancer landscape. Nature. 2012; 486:328–329. [PubMed: 22722187]
31. Wood LD, et al. The genomic landscapes of human breast and colorectal cancers. Science. 2007; 318:1108–1113. [PubMed: 17932254]
32. Shah SP, et al. The clonal and mutational evolution spectrum of primary triple-negative breast cancers. Nature. 2012; 486:395–399. [PubMed: 22495314]
33. Russnes HG, et al. Genomic architecture characterizes tumor progression paths and fate in breast cancer patients. Sci Transl Med. 2010; 2:38ra–47.
34. Stephens PJ, et al. Complex landscapes of somatic rearrangement in human breast cancer genomes. Nature. 2009; 462:1005–1010. [PubMed: 20033038]
35. Ellis MJ, et al. Whole-genome analysis informs breast cancer response to aromatase inhibition. Nature. 2012; 486:353–360. [PubMed: 22722193]
36. Banerji S, et al. Sequence analysis of mutations and translocations across breast cancer subtypes. Nature. 2012; 486:405–409. [PubMed: 22722202]
37. Stephens PJ, et al. The landscape of cancer genes and mutational processes in breast cancer. Nature. 2012; 486:400–404. [PubMed: 22722201]
38. Curtis C, et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature. 2012; 486:346–352. [PubMed: 22522925]
39. Venter JC, et al. The sequence of the human genome. Science. 2001; 291:1304–1351. [PubMed: 11181995]
40. Lander ES, et al. Initial sequencing and analysis of the human genome. Nature. 2001; 409:860–921. [PubMed: 11237011]

41. Kedes L, Campany G. The new date, new format, new goals and new sponsor of the Archon Genomics X PRIZE Competition. Nat Genet. 2011; 43:1055–1058. [PubMed: 22030612]
42. Rowley JD. Letter: A new consistent chromosomal abnormality in chronic myelogenous leukaemia identified by quinacrine fluorescence and Giemsa staining. Nature. 1973; 243:290–293. [PubMed: 4126434]
43. Druker BJ, et al. Efficacy and safety of a specific inhibitor of the BCR-ABL tyrosine kinase in chronic myeloid leukemia. N Engl J Med. 2001; 344:1031–1037. [PubMed: 11287972]
44. Kwak EL, et al. Anaplastic lymphoma kinase inhibition in non-small-cell lung cancer. N Engl J Med. 2010; 363:1693–1703. [PubMed: 20979469]
45. Soda M, et al. Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer. Nature. 2007; 448:561–566. [PubMed: 17625570]
46. Kola I, Landis J. Can the pharmaceutical industry reduce attrition rates? Nat Rev Drug Discov. 2004; 3:711–715. [PubMed: 15286737]
47. Baylin SB, Jones PA. A decade of exploring the cancer epigenome - biological and translational implications. Nat Rev Cancer. 2011; 11:726–734. [PubMed: 21941284]
48. Bissell MJ, Hines WC. Why don't we get more cancer? A proposed role of the microenvironment in restraining cancer progression Nat Med. 2011; 17:320–329.
49. Sandoval J, Esteller M. Cancer epigenomics: beyond genomics. Curr Opin Genet Dev. 2012; 22:50–55. [PubMed: 22402447]
50. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011; 144:646–674. [PubMed: 21376230]
51. Forbes S, et al. Cosmic 2005. Br J Cancer. 2006; 94:318–322. [PubMed: 16421597]
52. Vogelstein B, Kinzler KW. Cancer genes and the pathways they control. Nat Med. 2004; 10:789–799. [PubMed: 15286780]
53. Sharma SV, et al. A common signaling cascade may underlie “addiction” to the Src, BCR-ABL, and EGF receptor oncogenes. Cancer Cell. 2006; 10:425–435. [PubMed: 17097564]
54. Barbieri CE, et al. Exome sequencing identifies recurrent SPOP, FOXA1 and MED12 mutations in prostate cancer. Nat Genet. 2012; 44:685–689. [PubMed: 22610119]
55. Takeuchi K, et al. RET, ROS1 and ALK fusions in lung cancer. Nat Med. 2012; 18:378–381. [PubMed: 22327623]
56. Shih C, Shilo BZ, Goldfarb MP, Dannenberg A, Weinberg RA. Passage of phenotypes of chemically transformed cells via transfection of DNA and chromatin. Proc Natl Acad Sci USA. 1979; 76:5714–5718. [PubMed: 230490]
57. Takahashi Y, Mori J, Kami M. BRAF mutations in hairy-cell leukemia. N Engl J Med. 2011; 365:960–961. [PubMed: 21899463]
58. Prickett TD, et al. Exon capture analysis of G protein-coupled receptors identifies activating mutations in GRM3 in melanoma. Nat Genet. 2011; 43:1119–1126. [PubMed: 21946352]
59. Wei X, et al. Exome sequencing identifies GRIN2A as frequently mutated in melanoma. Nat Genet. 2011; 43:442–446. [PubMed: 21499247]
60. Parsons DW, et al. An integrated genomic analysis of human glioblastoma multiforme. Science. 2008; 321:1807–1812. [PubMed: 18772396]
61. Mardis ER, et al. Recurring mutations found by sequencing an acute myeloid leukemia genome. N Engl J Med. 2009; 361:1058–1066. [PubMed: 19657110]
62. Knudson AG Jr. Genetics of human cancer Genetics. 1975; 79(Suppl):305–316.
63. Varela I, et al. Exome sequencing identifies frequent mutation of the SWI/SNF complex gene PBRM1 in renal carcinoma. Nature. 2011; 469:539–542. [PubMed: 21248752]
64. Wiegand KC, et al. ARID1A mutations in endometriosis-associated ovarian carcinomas. N Engl J Med. 2010; 363:1532–1543. [PubMed: 20942669]
65. Jones S, et al. Exomic sequencing identifies PALB2 as a pancreatic cancer susceptibility gene. Science. 2009; 324:217. [PubMed: 19264984]
66. Rehman FL, Lord CJ, Ashworth A. Synthetic lethal approaches to breast cancer therapy. Nat Rev Clin Oncol. 2010; 7:718–724. [PubMed: 20956981]

67. Dedes KJ, et al. Synthetic lethality of PARP inhibition in cancers lacking BRCA1 and BRCA2 mutations. Cell Cycle. 2011; 10:1192–1199. [PubMed: 21487248]
68. Argilla D, et al. Absence of telomerase and shortened telomeres have minimal effects on skin and pancreatic carcinogenesis elicited by viral oncogenes. Cancer Cell. 2004; 6:373–385. [PubMed: 15488760]
69. Durinck S, et al. Temporal dissection of tumorigenesis in primary cancers. Cancer Discov. 2011; 1:137–143. [PubMed: 21984974]
70. Welch JS, et al. The origin and evolution of mutations in acute myeloid leukemia. Cell. 2012; 150:264–278. [PubMed: 22817890]
71. Almoguera C, et al. Most human carcinomas of the exocrine pancreas contain mutant c-K-ras genes. Cell. 1988; 53:549–554. [PubMed: 2453289]
72. Smit VT, et al. KRAS codon 12 mutations occur very frequently in pancreatic adenocarcinomas. Nucleic Acids Res. 1988; 16:7773–7782. [PubMed: 3047672]
73. van 't Veer LJ, et al. N-ras mutations in human cutaneous melanoma from sun-exposed body sites. Mol Cell Biol. 1989; 9:3114–3116. [PubMed: 2674680]
74. Fujita J, et al. Ha-ras oncogenes are activated by somatic alterations in human urinary tract tumours. Nature. 1984; 309:464–466. [PubMed: 6328318]
75. Agrawal N, et al. Exome sequencing of head and neck squamous cell carcinoma reveals inactivating mutations in NOTCH1. Science. 2011; 333:1154–1157. [PubMed: 21798897]
76. Lynch TJ, et al. Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. N Engl J Med. 2004; 350:2129–2139. [PubMed: 15118073]
77. Paez JG, et al. EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. Science. 2004; 304:1497–1500. [PubMed: 15118125]
78. Ekstrand AJ, Sugawa N, James CD, Collins VP. Amplified and rearranged epidermal growth factor receptor genes in human glioblastomas reveal deletions of sequences encoding portions of the N- and/or C-terminal tails. Proc Natl Acad Sci USA. 1992; 89:4309–4313. [PubMed: 1584765]
79. Nik-Zainal S, et al. The life history of 21 breast cancers. Cell. 2012; 149:994–1007. [PubMed: 22608083]
80. Chin K, et al. Genomic and transcriptional aberrations linked to breast cancer pathophysiologies. Cancer Cell. 2006; 10:529–541. [PubMed: 17157792]
81. Ding L, et al. Genome remodelling in a basal-like breast cancer metastasis and xenograft. Nature. 2010; 464:999–1005. [PubMed: 20393555]
82. Shah SP, et al. Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution. Nature. 2009; 461:809–813. [PubMed: 19812674]
83. Navin N, et al. Tumour evolution inferred by single-cell sequencing. Nature. 2011; 472:90–94. [PubMed: 21399628]
84. Gerlinger M, et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. N Engl J Med. 2012; 366:883–892. [PubMed: 22397650]
85. Salk JJ, Fox EJ, Loeb LA. Mutational heterogeneity in human cancers: origin and consequences. Annu Rev Pathol. 2010; 5:51–75. [PubMed: 19743960]
86. Diaz LA Jr, et al. The molecular evolution of acquired resistance to targeted EGFR blockade in colorectal cancers. Nature. 2012; 486:537–540. [PubMed: 22722843]
87. Sequist LV, et al. Genotypic and histological evolution of lung cancers acquiring resistance to EGFR inhibitors. Sci Transl Med. 2011; 3:75ra–26.
88. Yap TA, Gerlinger M, Futreal PA, Pusztai L, Swanton C. Intratumor heterogeneity: seeing the wood for the trees. Sci Transl Med. 2012; 4:127ps–10.
89. Vogelstein B, et al. Genetic alterations during colorectal-tumor development. N Engl J Med. 1988; 319:525–532. [PubMed: 2841597]
90. Liu W, et al. Copy number analysis indicates monoclonal origin of lethal metastatic prostate cancer. Nat Med. 2009; 15:559–565. [PubMed: 19363497]
91. Tomlins SA, et al. Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. Science. 2005; 310:644–648. [PubMed: 16254181]

92. McDermott U, Settleman J. Personalized cancer therapy with selective kinase inhibitors: an emerging paradigm in medical oncology. J Clin Oncol. 2009; 27:5650–5659. [PubMed: 19858389]
93. Prahallad A, et al. Unresponsiveness of colon cancer to BRAF(V600E) inhibition through feedback activation of EGFR. Nature. 2012; 483:100–103. [PubMed: 22281684]
94. Weng AP, et al. Activating mutations of NOTCH in human T cell acute lymphoblastic leukemia. Science. 2004; 306:269–271. [PubMed: 15472075]
95. Wang NJ, et al. Loss-of-function mutations in Notch receptors in cutaneous and lung squamous cell carcinoma. Proc Natl Acad Sci USA. 2011; 108:17761–17766. [PubMed: 22006338]
96. Graubert TA, et al. Recurrent mutations in the U2AF1 splicing factor in myelodysplastic syndromes. Nat Genet. 2012; 44:53–57. [PubMed: 22158538]
97. Mileshkin L, et al. Changes in $^{18}$F-fluorodeoxyglucose and 18F-fluorodeoxythymidine positron emission tomography imaging in patients with non-small cell lung cancer treated with erlotinib. Clin Cancer Res. 2011; 17:3304–3315. [PubMed: 21364032]
98. Punnoose EA, et al. Evaluation of circulating tumor cells and circulating tumor DNA in non-small cell lung cancer: association with clinical endpoints in a phase II clinical trial of pertuzumab and erlotinib. Clin Cancer Res. 2012; 18:2391–2401. [PubMed: 22492982]
99. Dinov ID, et al. iTools: a framework for classification, categorization and integration of computational biology resources. PLoS ONE. 2008; 3:e2265. [PubMed: 18509477]
100. Floratos A, Smith K, Ji Z, Watkinson J, Califano A. geWorkbench: an open source platform for integrative genomics Bioinformatics. 2010; 26:1779–1780.
101. Carter SL, et al. Absolute quantification of somatic DNA alterations in human cancer. Nat Biotechnol. 2012; 30:413–421. [PubMed: 22544022]
102. Shen R, et al. Integrative subtype discovery in glioblastoma using iCluster. PLoS ONE. 2012; 7:e35236. [PubMed: 22539962]
103. Broad Institute. Cancer Program Resource Gateway [online]. 2012. http://www.broadinstitute.org/software/cprg/?q=tool-resource
104. UC Santa Cruz. UCSC Cancer Genomics Browser [online]. 2012. https://genome-cancer.ucsc.edu
105. Memorial Sloan-Kettering Cancer Center. cBio Cancer Genomics Portal [online]. 2012. http://www.cbioportal.org/public-portal
106. Heiser LM, et al. Subtype and pathway specific responses to anticancer compounds in breast cancer. Proc Natl Acad Sci USA. 2012; 109:2724–2729. [PubMed: 22003129]
107. Vaske CJ, et al. Inference of patient-specific pathway activities from multi-dimensional cancer genomics data using PARADIGM. Bioinformatics. 2010; 26:i237–i245. [PubMed: 20529912]
108. Barretina J, et al. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature. 2012; 483:603–607. [PubMed: 22460905]
109. Pinkel D, Straume T, Gray JW. Cytogenetic analysis using quantitative, high-sensitivity, fluorescence hybridization. Proc Natl Acad Sci USA. 1986; 83:2934–2938. [PubMed: 3458254]
110. Sequist LV, et al. Implementing multiplexed genotyping of non-small-cell lung cancers into routine clinical practice. Ann Oncol. 2011; 22:2616–2624. [PubMed: 22071650]
111. Corless CL. Medicine. Personalized cancer diagnostics. Science. 2011; 334:1217–1218. [PubMed: 22144608]
112. Paik S, et al. Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor-positive breast cancer. J Clin Oncol. 2006; 24:3726–3734. [PubMed: 16720680]
113. Chung KY, et al. Cetuximab shows activity in colorectal cancer patients with tumors that do not express the epidermal growth factor receptor by immunohistochemistry. J Clin Oncol. 2005; 23:1803–1810. [PubMed: 15677699]
114. Zerbino DR, Paten B, Haussler D. Integrating genomes. Science. 2012; 336:179–182. [PubMed: 22499938]
115. Nielsen, MA. Reinventing discovery: the new era of networked science. Princeton University Press; Princeton, USA: 2012.
116. Jones SJ, et al. Evolution of an adenocarcinoma in response to selection by targeted kinase inhibitors. Genome Biol. 2010; 11:R82. [PubMed: 20696054]

117. Amir E, et al. Prospective study evaluating the impact of tissue confirmation of metastatic disease in patients with breast cancer. J Clin Oncol. 2012; 30:587–592. [PubMed: 22124102]
118. Lindstrom LS, et al. Clinically used breast cancer markers such as estrogen receptor, progesterone receptor, and human epidermal growth factor receptor 2 are unstable throughout tumor progression. J Clin Oncol. 2012; 30:2601–2608. [PubMed: 22711854]
119. Shah NP, et al. Multiple BCR-ABL kinase domain mutations confer polyclonal resistance to the tyrosine kinase inhibitor imatinib (STI571) in chronic phase and blast crisis chronic myeloid leukemia. Cancer Cell. 2002; 2:117–125. [PubMed: 12204532]
120. Talpaz M, et al. Dasatinib in imatinib-resistant Philadelphia chromosome-positive leukemias. N Engl J Med. 2006; 354:2531–2541. [PubMed: 16775234]
121. de Thé H, Chomienne C, Lanotte M, Degos L, Dejean A. The t(15;17) translocation of acute promyelocytic leukaemia fuses the retinoic acid receptor alpha gene to a novel transcribed locus. Nature. 1990; 347:558–561. [PubMed: 2170850]
122. Huang ME, et al. All-trans retinoic acid with or without low dose cytosine arabinoside in acute promyelocytic leukemia. Report of 6 cases. Chin Med J(Engl). 1987; 100:949–953. [PubMed: 3133168]
123. Joensuu H, et al. Effect of the tyrosine kinase inhibitor STI571 in a patient with a metastatic gastrointestinal stromal tumor. N Engl J Med. 2001; 344:1052–1026. [PubMed: 11287975]
124. King CR, Kraus MH, Aaronson SA. Amplification of a novel v-erbB-related gene in a human mammary carcinoma. Science. 1985; 229:974–976. [PubMed: 2992089]
125. 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]
126. Beatson GT. On the treatment of inoperable cases of carcinoma of the mamma: suggestions for a new method of treatment, with illustrative cases. Lancet. 1896; 148:162–165.
127. Kris MG, et al. Efficacy of gefitinib, an inhibitor of the epidermal growth factor receptor tyrosine kinase, in symptomatic patients with non-small cell lung cancer: a randomized trial. JAMA. 2003; 290:2149–2158. [PubMed: 14570950]
128. Davies H, et al. Mutations of the BRAF gene in human cancer. Nature. 2002; 417:949–954. [PubMed: 12068308]
129. Chapman PB, et al. Improved survival with vemurafenib in melanoma with BRAF V600E mutation. N Engl J Med. 2011; 364:2507–2516. [PubMed: 21639808]
130. Bos JL, et al. Prevalence of ras gene mutations in human colorectal cancers. Nature. 1987; 327:293–297. [PubMed: 3587348]
131. Lièvre A, et al. KRAS mutation status is predictive of response to cetuximab therapy in colorectal cancer. Cancer Res. 2006; 66:3992–3995. [PubMed: 16618717]

### Key points

- The past decade of studying the cancer genome has resulted in major breakthroughs in the treatment of cancers, such as leukaemias, breast and lung cancers and melanoma
- Advances in sequencing now allow interrogation of cancer genomes both broadly, by assessing many genes in many samples, and deeply, by assessing variations within a specific gene in a single sample
- Atypical tumour may carry a few well-understood driver aberrations and tens to thousands of poorly understood aberrations that may collaborate in the development of each individual cancer phenotype
- Organizing the myriad genomic aberrations in an individual tumour onto a limited number of affected pathways may guide development of pathway-targeted therapies
- By expanding our definition of driver aberrations to include those that influence genome stability, chromatin structure, differentiation, RNA processing and beyond, we may increase successes in developing novel therapeutic strategies
- Future challenges include data storage, distribution, analysis and the necessity to update regulatory models; developers, regulators, clinicians, patient advocates and others must approach this exciting time with flexibility and creativity

### Review criteria

Review of the biomedical literature was conducted in English via PubMed covering the years 1980 to 2012. Search terms used were “cancer genome”, “next-generation sequencing”, “predictive biomarkers”, and “clinical genomics”.

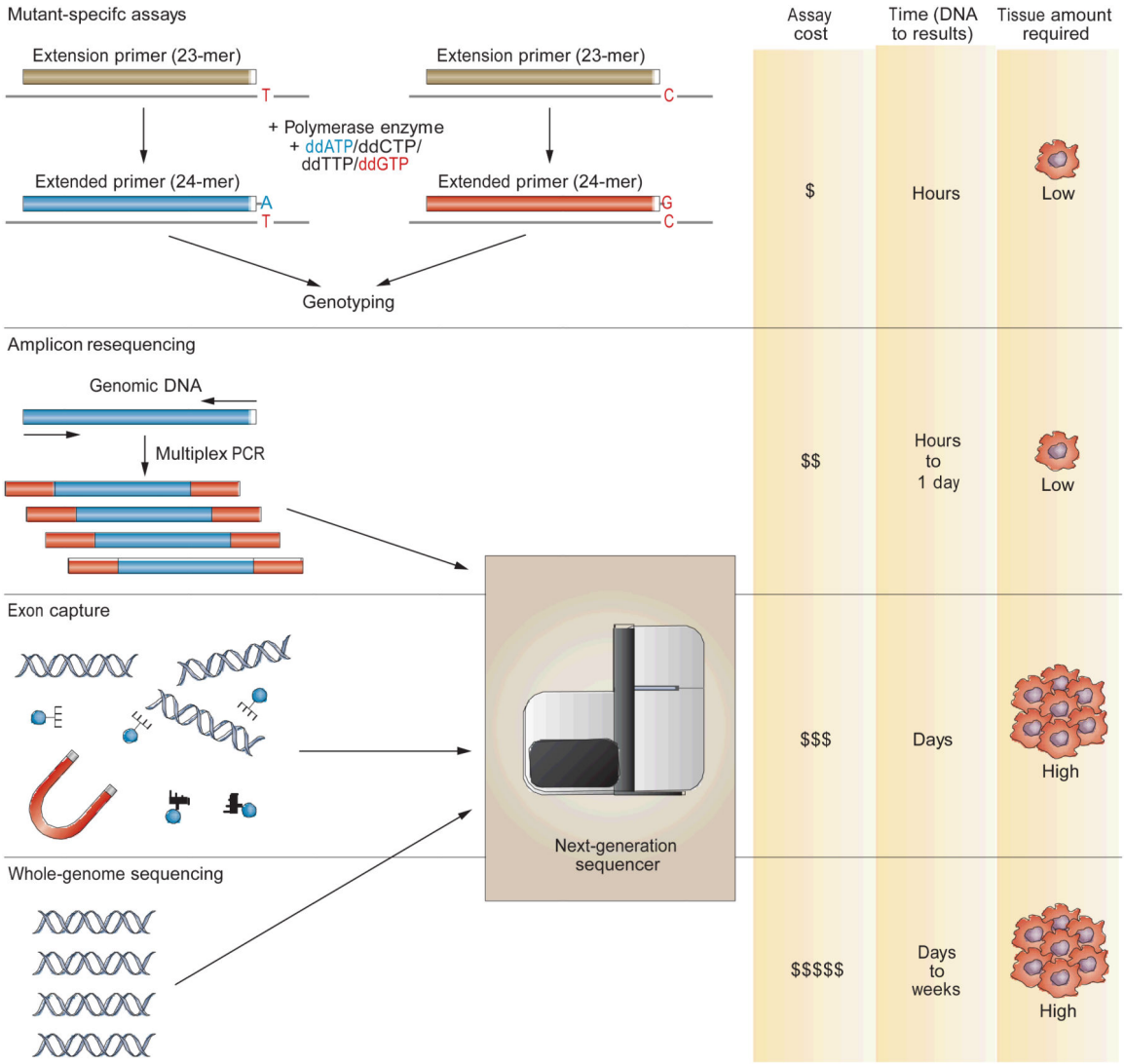


**Figure 1.**

Heterogeniety in Cancer. A breast cancer patient has a primary breast tumor, consisting of three sections (P1, P2, P3), and metastases in the brain (MB), liver (ML) and hip bone (MO). While all the tumors from all sites are related, and differ from the germline genome, their degree of relatedness can be interrogated with genome evaluation. Conceptual patterns of relatedness emerging are those malignancies that continue to evolve as they metastasize or in response to treatment (the Oak tree model), or those that are more uniform(the Palm tree model).

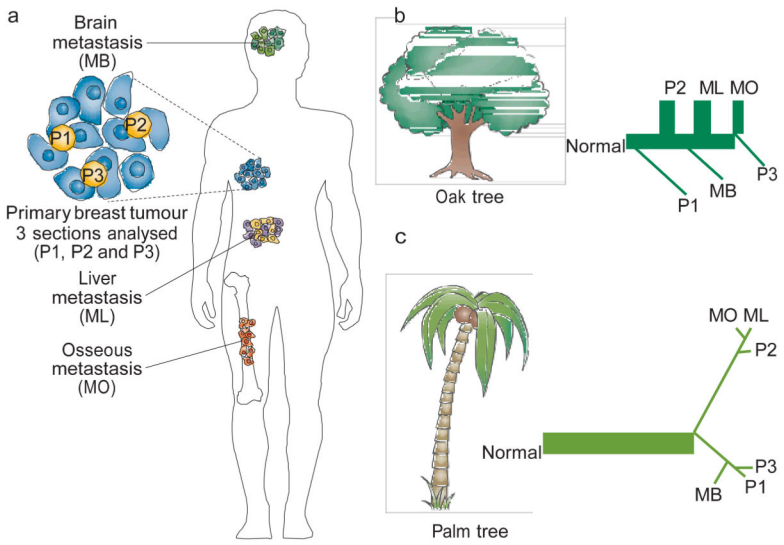


**Figure 2.**

Gain of function mutations in Notch genes are recurrent in T-ALL whereas loss of function mutations predominate in squamous cell cancers of the skin, lung and head and neck.

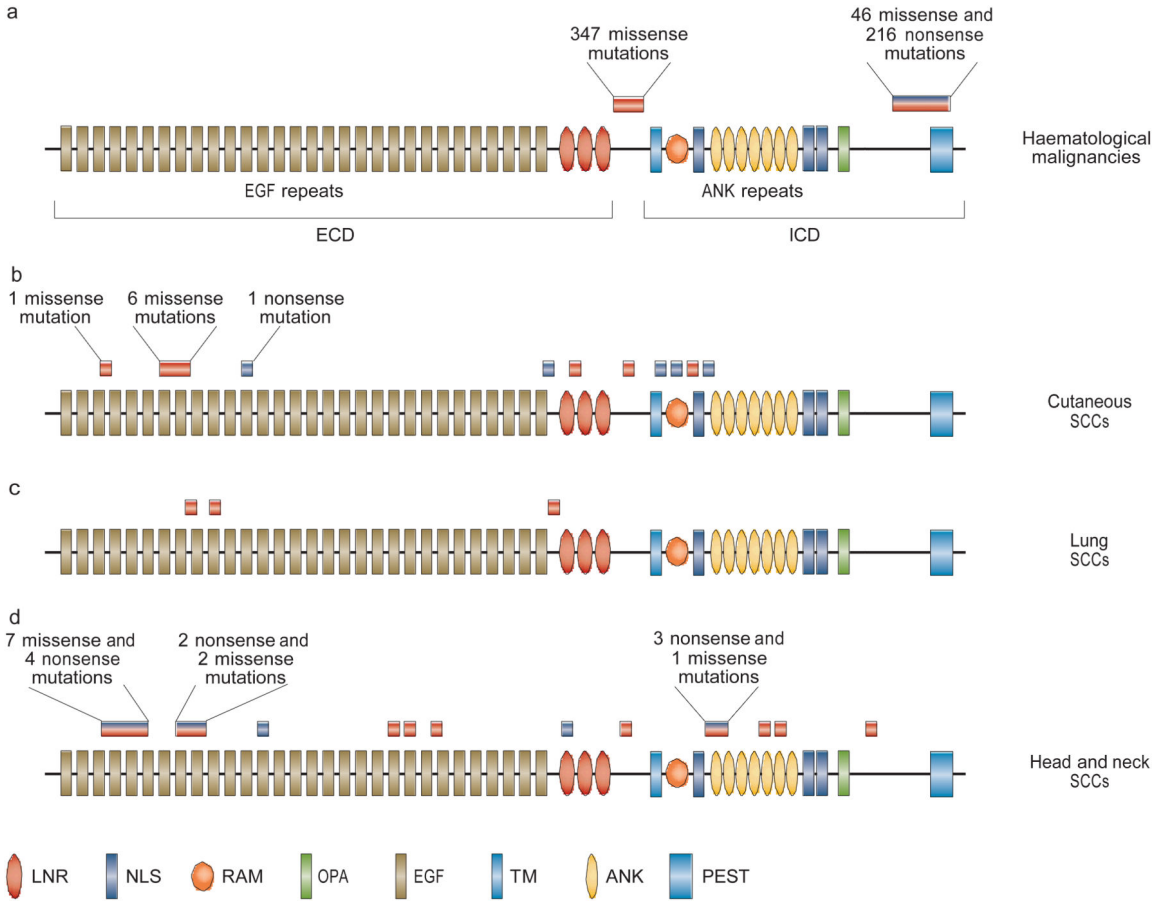


**Figure 3.**

PARADIGM approach to pathway analysis of a drug sensitive vs. drug resistant population of patients.

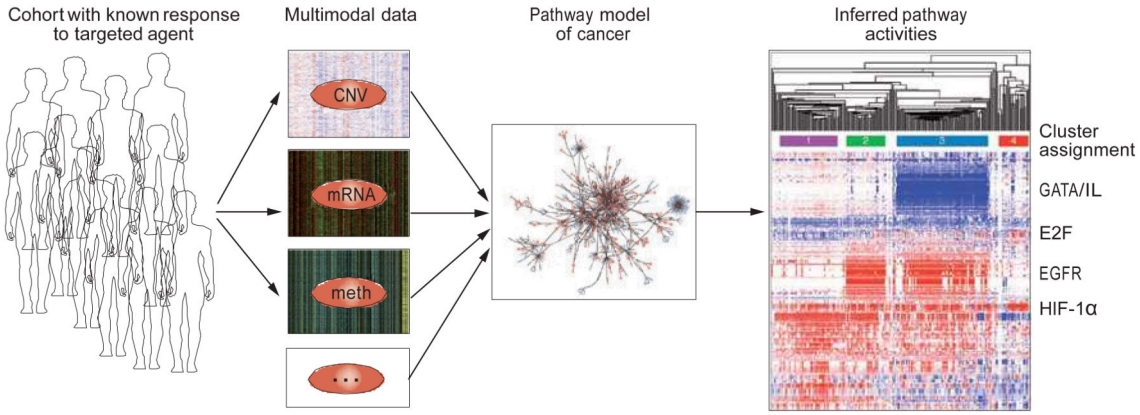


**Figure 4.**

An example of an integrative approach to pathway analysis of a drug-sensitive versus drug-resistant population of patients.

**Table 1**

**Genomic assays and the clinical decisions they support**

| Omic feature | Time of aberration discovery | Clinical setting | Analyte | Analysis | Treatment implication | Time of implementation of first targeted therapy |
|---|---|---|---|---|---|---|
| *BCR–ABL* translocation | 1973[41] | Chronic myeloid leukaemia, acute lymphoblastic leukaemia | DNA, RNA | PCR (translocation) | Imatinib, dasatinib, nilotinib | 2001[42] |
| *BCR–ABL* imatinib resistance mutation | 2002[119] | Chronic myeloid leukaemia | DNA | Sequencing (point mutation) | Dasatinib, nilotinib | 2006[120] |
| *PML–RAR* translocation | 1990[121] | Acute promyelocytic leukaemia | DNA | PCR (translocation) | All *trans* retinoic acid | 1987[122] |
| *KIT* mutation | 2001[123] | Gastrointestinal stromal tumour | DNA | Sequencing | Imatinib | 2001[123] |
| *HER2* amplification | 1985[124] | Breast cancer, gastric cancer | Tissue | Fluorescence *in situ* hybridization, immunohistochemistry | Trastuzumab, lapatinib, pertuzumab | 2001[125] |
| Oestrogen or progesterone receptors | 1896[126] | Breast cancer | Tissue | Immunohistochemistry | Multiple hormonal-based therapies (for example, aromatase inhibitors) | 1896 (oophorectomy)[126] |
| *EGFR* mutation | 2004[76] | Non-small-cell lung cancer | DNA | Sequencing or PCR (point mutations) | Erlotinib, gefitinib | 2003[127] |
| *ALK* fusion gene | 2007[44] | Non-small-cell lung cancer | DNA | Fluorescence *in situ* hybridization (break apart) | Crizotinib | 2010[43] |
| *BRAF* mutation | 2001[128] | Melanoma | DNA | Sequencing or PCR (point mutations) | Vemurafenib | 2011[129] |
| *KRAS* mutation | 1987[130] | Colorectal cancer | DNA | Sequencing or PCR (point mutations) | Withhold cetuximab or panitumumab | 2006[131] |

*Nat Rev Clin Oncol*. Author manuscript; available in PMC 2014 September 19.