Our changing view of the genomic landscape of cancer

Daphne W Bell[*]
National Human Genome Research Institute, National Institutes of Health, Bethesda, MD, USA

Abstract

Sporadic tumours, which account for the majority of all human cancers, arise from the acquisition of somatic, genetic and epigenetic alterations leading to changes in gene sequence, structure, copy number and expression. Within the last decade, the availability of a complete sequence-based map of the human genome, coupled with significant technological advances, has revolutionized the search for somatic alterations in tumour genomes. Recent landmark studies, which resequenced all coding exons within breast, colorectal, brain and pancreatic cancers, have shed new light on the genomic landscape of cancer. Within a given tumour type there are many infrequently mutated genes and a few frequently mutated genes, resulting in incredible genetic heterogeneity. However, when the altered genes are placed into biological processes and biochemical pathways, this complexity is significantly reduced and shared pathways that are affected in significant numbers of tumours can be discerned. The advent of next-generation sequencing technologies has opened up the potential to resequence entire tumour genomes to interrogate protein-encoding genes, non-coding RNA genes, non-genic regions and the mitochondrial genome. During the next decade it is anticipated that the most common forms of human cancer will be systematically surveyed to identify the underlying somatic changes in gene copy number, sequence and expression. The resulting catalogues of somatic alterations will point to candidate cancer genes requiring further validation to determine whether they have a causal role in tumourigenesis. The hope is that this knowledge will fuel improvements in cancer diagnosis, prognosis and therapy, based on the specific molecular alterations that drive individual tumours. In this review, I will provide a historical perspective on the identification of somatic alterations in the pre- and post-genomic eras, with a particular emphasis on recent pioneering studies that have provided unprecedented insights into the genomic landscape of human cancer.

Keywords

Introduction

In 1914, Theodor Boveri hypothesized that numerical alterations in the chromosome content of a cell might contribute to tumourigenesis [1]. However, it would take almost another 50 years before the first specific chromosomal abnormality, a marker chromosome termed the Philadelphia chromosome, was identified within cancer genomes of patients with chronic

myelogenous leukaemia (CML) [2]. We now know that sporadic human tumours, which account for the majority of all human cancers, result from the accumulation of numerous genetic and epigenetic alterations, leading to the dysregulation of protein-encoding genes as well as non-coding RNAs (ncRNAs) [3–6]. Such alterations underlie the acquired attributes shared by most cancer cells, namely their capacity to generate their own mitogenic signals, evade apoptosis, resist exogenous growth-inhibitory signals, proliferate without limits, and acquire angiogenic, invasive and metastatic properties [7].

Early attempts to identify protein-encoding cancer genes were hampered by both the limited resolution of technologies available to detect genomic alterations and the enormous complexity of tumour genomes. This complexity was first recognized from cytogenetic studies cataloguing numerical and structural abnormalities within solid tumours [8]. Unlike haematological malignancies and certain sarcomas, which were often characterized by a signature chromosome abnormality, most solid tumours of epithelial origin exhibited numerous chromosomal aberrations, both clonal and nonclonal [8–10]. The karyotypic heterogeneity of tumour genomes led to the proposal by Peter Nowell, in the mid-1970s, that tumourigenesis occurs by a stepwise evolutionary process [11]. Nowell suggested that, following an initiating event that converts a normal cell to a neoplastic cell, cancer progression results from the acquisition of genetic instability, leading to the accumulation of genetic alterations and the continual selective outgrowth of variant subpopulations of tumour cells with a proliferative advantage. This model not only accounted for the genetic heterogeneity of solid tumours, but also led Nowell to speculate that each cancer patient might require individual specific therapy [11].

In the ensuing years, cancer gene identification was a laborious task, significantly impeded by the lack of a sequence-based map of the normal human genome. This changed in 2003 with the completion of the Human Genome Project [12]. In the so-called post-genomic era, the availability of a complete map of the human genome and significant technological advances together powered systematic interrogations of cancer genomes, at unprecedented resolution and throughput, to catalogue somatic alterations and thus pinpoint candidate cancer genes. Within the last decade the concept of individualized therapy, based on the presence of specific molecular alterations within a patient's tumour, has become a reality, spurring the research community to systematically catalogue the somatic alterations that underlie all human cancers [13–26]. Here I review progress that has been made in the search for the genetic basis of sporadic cancers, with a particular emphasis on recent efforts to comprehensively and systematically catalogue somatic alterations using integrated genomic approaches. For comprehensive reviews of other factors that shape the cancer landscape, including the role of ncRNAs, dysregulation of protein synthesis, altered splicing, changes in chromatin structure, epi-genetic alterations and germline cancer susceptibility, the reader is referred to additional articles within this issue and elsewhere [27–38].

Searching for cancer genes in the pre-genomics era

To appreciate the significance of recent developments in the search for cancer genes, it is useful to understand the pace of cancer gene discovery between 1960 and 2003, the years marking the identification of the first cancer-specific chromosomal abnormality and the completion of the Human Genome Project (Figure 1) [2,12]. In the late 1970s it was discovered that the genes responsible for the oncogenicity of certain retroviruses were actually altered versions (oncogenes) of normal cellular genes (proto-oncogenes). The first proto-oncogene, termed *c-src*, was identified in 1976, based on homology to the transforming gene (v-src) of the Rous sarcoma virus, a retrovirus that induces sarcomas in chickens [39,40].

The significance of proto-oncogenes in relation to human cancer causation became clear in 1982, when a transforming gene isolated from a human bladder carcinoma cell line was discovered to be a mutated version of a normal cellular proto-oncogene [41–44]. This marked the first description of a sequence change, a point mutation resulting in the oncogenic activation of *H-RAS*, in a human cancer gene.

The development of high-resolution chromosome banding techniques in the early 1970s allowed for detailed karyotypic analysis of cancer genomes and the identification of recurrent chromosomal abnormalities in haematological malignancies, cancer-derived cell lines and, to a lesser extent, solid tumours (reviewed in [8]). The molecular characterization of genes positioned at the breakpoints of chromosomal translocations, or those contained within homogeneously staining regions and double minutes, revealed that gene rearrangement, amplification and increased expression also resulted in oncogene activation in human cancers [45–54] (reviewed in [55]). In the pre-genomics era, oncogene identification largely relied on cloning genes at the site of proviral integrations, functional assays to isolate cellular genes capable of inducing foci formation in NIH3T3 cells, and positional cloning following the identification of structural and numerical chromosomal abnormalities, using both conventional and, later, molecular cytogenetic approaches (reviewed in [56,57]). By 2002, more than 100 oncogenes had been described [58]. Many of these encoded protein tyrosine kinases that modulate mitogenic signal transduction pathways [59]. Their oncogenic activation was associated with increased kinase activity and cellular transformation, properties that would have important therapeutic implications in later years (reviewed in [60]).

In 1971, by studying the inheritance pattern of retinoblastoma, a childhood cancer of the eye with a familial component, Alfred Knudson hypothesized the existence of a distinct class of cancer gene, acting recessively and requiring 'two hits' to contribute to tumourigenesis [61]. The idea that neoplasia was sometimes a genetically recessive phenotype was supported by the cell fusion experiments of Henry Harris. The retinoblastoma gene (*RB1*), the first of the co-called tumour suppressor genes, was cloned 16 years later [62–68]. As Knudson had predicted, familial forms of retinoblastoma are caused by the inheritance of a germline alteration within one allele of *RB1* and the acquisition of a somatic alteration in the other allele (the second hit) during the patient's lifetime. In contrast, sporadic forms of retinoblastoma arise later in life after a single cell acquires two independent somatic alterations, leading to a complete loss of *RB1* function. A variety of genetic alterations lead to tumour suppressor gene inactivation, including point mutations, gene deletions and epigenetic alterations leading to gene silencing [64,69]. The earliest tumour suppressor genes identified were isolated by positional cloning, following the delineation of regions of linkage to disease susceptibility in cancer kindreds [62–68,70–88]. By 1993, 13 tumour suppressor genes had been positionally cloned or mapped to specific chromosomal locations [89].

Later, candidate gene approaches were also successful in uncovering cancer susceptibility genes [90–92]. With the development of molecular methods to search for somatic loss-of-heterozygosity and homozygous deletions in sporadic tumours, additional recessively acting cancer genes were described [70,71,93–95]. In 1986, Renato Dulbecco contended that the research community was faced with a choice — either to continue searching for cancer genes by a piecemeal approach or to decode the human genome and lay a foundation for future systematic searches for cancer genes [96]. Four years later saw the launch of the Human Genome Project, a publicly funded, international effort to decode the human genome, led by the International Human Genome Sequencing Consortium (IHGSC).

Searching for cancer genes in the post-genomics era

In 2001, the IHGSC and Celera Genomics reported draft versions of the human genome sequence [97,98]. Shortly thereafter, the complete human genome sequence was reported by the IHGSC [12]. The availability of a sequence-based map of the human genome and the development of high-throughput Sanger sequencing, a by-product of the human genome project, provided new opportunities to systematically catalogue somatic mutations within both causal and candidate cancer genes on a larger scale, and at a faster pace, than previously possible. Given the clinical successes of genotype-directed therapies in treating *HER2*-positive breast cancer patients with trastuzumab, and treating *BCR-ABL*-positive CML patients with imatinib, many of the earliest high-throughput resequencing efforts centred on protein kinase-encoding genes [13,99,100]. As a result, a number of novel therapeutic targets were revealed for genotypically-defined subgroups of cancer patients [101–118] (reviewed in [119]). This paved the way for other high-throughput resequencing efforts targeting specific biochemical pathways, large gene families and large numbers of candidate cancer genes [120–130]. Not all cancer genes encode proteins that are amenable to direct therapeutic intervention. However, identifying synthetic lethal partners of these so-called 'undruggable targets' holds promise in leveraging such alterations for targeted cancer therapy [23–26,131] (reviewed in [132]).

In parallel with the human genome project, the development of array-based comparative genomic hybridization and progressive refinements of this technology significantly increased the resolution and throughput at which genome-wide somatic alterations in DNA copy number could be ascertained (reviewed in [133]). The completion of the human genome project also facilitated the construction of the HapMap, a map of naturally occurring human genomic variation, in the form of single nucleotide polymorphisms (SNPs) and its underlying genomic structure [134–136]. This allowed for the development of high-density SNP genotyping arrays that could be used to screen not only for copy number alterations but also for copy-neutral loss-of-heterozygosity throughout the genome, at unparalleled resolution [137–143] (reviewed in [133]). The implementation of genome-wide expression profiling revealed novel transcriptional signatures that can be utilized to molecularly classify cancer subtypes or predict clinical phenotypes, and was also instrumental in facilitating the identification of previously unrecognized gene fusions in common epithelial tumours (reviewed in [144,145]). Furthermore, integrated studies that combined genome-wide searches for copy number alterations with global gene expression profiling significantly increased the power to hone in on candidate cancer genes [146,147]. Nonetheless, the capacity to understand the full compendium of genomic alterations that drive human tumourigenesis was limited by the inability to rapidly and systematically sequence entire tumour genomes.

Our first glimpse into the true genetic complexity of human cancers has come from recent pioneering studies that sequenced the exomes, the coding exons of more than 18 000 protein-encoding genes, from a series of breast, colorectal, pancreatic and brain cancers [148–151]. Remarkably, these explorations were achieved using high-throughput Sanger sequencing, and were integrated with genomic analyses to interrogate gene expression and copy number [149,150,152]. These studies revealed that cancer genomes are highly complex, with an average of 48–101 somatic alterations in each tumour, depending on the cancer type [149,150]. Within a given cancer type there is considerable inter-tumour heterogeneity, resulting in large numbers of altered genes. However, this complexity is reduced significantly by considering the biological pathways and processes on which altered genes converge, rather than the altered genes themselves [151]. For example, 12 core biological processes or pathways appear to be deregulated in the majority of pancreatic tumours, although precisely how this is achieved varies from tumour to tumour (Figure 2)

[149]. As has been noted, this may have practical implications for the development of targeted therapeutics, in that it may be more prudent to consider targeting functional pathways or processes rather than individual proteins encoded by mutated genes [151]. Prior to these investigations, most cancer genes had been identified because they were frequently altered in tumourigenesis. However, the resequencing of tumour exomes revealed that, for a given type of cancer, the majority of somatically mutated genes are altered in just a fraction of tumours. This new view of the genomic landscape of human cancer suggested that the acquisition of numerous somatic mutations, each with a small fitness advantage, may also drive tumourigenesis [151,153].

Somatic alterations can be so-called driver mutations that confer a selective growth advantage to the tumour cell, or passenger mutations that have no effect on tumourigenesis [101,154]. Thus, the identification of a somatically altered gene indicates a candidate cancer gene rather than a causal cancer gene. What necessarily follows are detailed biochemical and cellular studies comparing the functional properties of the wild-type and mutant proteins. To guide such studies, statistical calculations, based on the frequency and nature of the observed somatic mutations, can be applied to prioritize or rank candidate cancer genes based on the likelihood that they represent driver genes [155]. The statistical assumptions that are most appropriate to use in this type of predictive modelling have been the subject of some debate, because of the inherent difficulty in setting a background mutation rate for each tumour type [155–158]. Other computational approaches predict driver mutations rather than driver genes [159–161]. One such method has estimated that approximately 8% of missense mutations identified by exomic sequencing of glioblastomas are likely to be functionally significant, with the majority of these affecting infrequently mutated genes [161]. Although synonymous somatic mutations are generally not considered in statistical predictions because they do not result in amino acid changes, it is worth noting that they can, on occasion, encode proteins with altered functional activity [162]. It is important to note that, in addition to statistical predictions, functional genetic screens in mice and large-scale RNA interference screens can also guide the identification of causal cancer genes [131,163–172] (reviewed in [173–175]).

Although exomic resequencing of cancer genomes captures the spectrum of mutations within protein-encoding genes, it does not assess the sequence integrity of non-coding regions of the genome. These regions contain functionally relevant elements, including ultra-conserved elements and ncRNAs, which are being systematically mapped by the Encyclopedia of DNA Elements (ENCODE) project [176–180]. Non-coding RNAs have been implicated in a variety of processes, including the regulation of transcription and chromosome structure, RNA processing and modification, mRNA stability and translation and protein stability and transport (reviewed in [181]). Within the past few years, our vision of the cancer landscape has been reshaped with the realization that the dysregulation of micro-RNAs (miRNAs), a subset of ncRNAs, contributes to tumourigenesis (reviewed in [32,35]). MiRNAs are small ncRNAs that negatively regulate gene expression, including that of protein-encoding cancer genes. Dysregulated miRNAs have been described in human cancers and in some instances are associated with oncogenic properties, tumour-suppressive properties or both, depending on the cellular context (reviewed in [32]). Furthermore, miRNA expression profiling of a mouse model of pancreatic cancer revealed distinct miRNA expression signatures at each step in the progression of tumourigenesis, correlating with the acquired attributes shared by most cancer cells [6,7]. The full extent to which ncRNAs contribute to cancer has yet to be revealed. In addition to miRNAs, another class of ncRNAs, represented by transcribed ultraconserved regions of the genome, has also been implicated in tumourigenesis [179,180]. Moreover, inherited mutations within the gene encoding DICER1, an endonuclease that regulates the processing of ncRNAs, have been linked to familial pleuropulmonary blastoma [182].

Next-generation sequencing

The impetus for the human genome project was to generate a sequence-based map that would allow researchers to identify the germline and somatic variation that underlies human disease, including cancer [96]. One implication is that of personalized medicine, the ability to predict individual disease risk and drug response based on personal genomics. In the context of sporadic tumourigenesis, decoding individual tumour genomes holds the promise of more accurate diagnosis and prognosis and of guiding personalized treatment strategies for cancer patients.

Sanger sequencing has been the gold standard of DNA sequencing since it was first described over 30 years ago [183]. However, the expense and throughput of this approach prohibits its application in sequencing large numbers of genomes. Therefore, in order to make personalized medicine incorporating whole-genome sequencing a reality, several new and revolutionary sequencing technologies have been developed in the past few years. These so-called next-generation sequencing approaches have been reviewed in detail elsewhere [184–188]. Suffice it to say that, by sequencing DNA in a massively parallel fashion, next-generation sequencing methodologies have significantly lowered the cost and time required to decode an entire human genome [189–193]. Nonetheless, the short reads obtained, coupled with the relatively high error rate, require deep coverage of each genome, keeping current costs relatively high. In addition, data analysis and confirmation of sequence variants are non-trivial. To bring whole-genome sequencing in line with other clinical diagnostic tests, the goal of a $1000 genome has been set (reviewed in [194,195]). Although this has not yet been attained, it is believed to be within reach with further improvements in technology. In addition to their higher throughput and lowered costs, next-generation sequencing methodologies have several other advantages compared with Sanger sequencing that are key to comprehensively deciphering tumour genomes. They are quantitative in nature and can be used to simultaneously determine both nucleotide sequence and copy number. They are more sensitive than Sanger sequencing, and therefore can detect somatic mutations present in just a subset of tumour cells [196,197]. In addition, these methods can be coupled with a so-called paired-end read strategy, in which both ends of individual clones are sequenced and mapped back to the genome [198]. This strategy can be used to identify structural alterations, including insertions, deletions, duplications and rearrangements [199–201]. It is important to note that sequencing entire genomes encompasses not only the cellular genome but also the mitochondrial genome and, potentially, any virally-associated genomes. Prior to the development of next-generation sequencing, searching for changes in sequence, copy number and structure required the integration of data from multiple platforms. As next-generation sequencing eventually becomes routine, it will transform the analyses of cancer genomes, using a single platform.

Next-generation sequencing of whole tumour genomes and transcriptomes

Decoding entire genomes by next-generation sequencing is feasible, but is not yet commonplace [189–193]. One consideration that pertains to whole-genome sequencing of cancer genomes is the current need to survey both the tumour and the constitutional genomes from the same individual, to accurately discriminate polymorphic variants from potential somatic mutations. At the time of this writing, the genomes of two cytogenetically normal cases of acute myeloid leukaemia (AML), as well as DNA from the normal skin cells of these individuals, have been resequenced and partially analysed. The extraordinary number of potential somatic mutations identified (20 000–30 000 in each tumour), together with the high error rate, necessitated the prioritization of mutations for confirmatory sequencing by other methods. To date, only those mutations localizing within protein-encoding genes of the first AML genome have been validated [202]. For the second AML

genome, mutations within the coding regions of annotated genes and RNA genes, as well as those within highly conserved non-genic regions and regulatory regions, have been validated [203]. Eight to twelve true somatic alterations were identified within the coding regions of each tumour genome [202,203]. Most of the altered genes were not previously implicated in the pathogenesis of AML and would not have been obvious choices for resequencing using a candidate gene approach. The coding regions of the genome comprise only 1–2% of the entire human genome. It has been estimated that 500–1000 additional somatic mutations exist within the non-coding regions of each AML genome. Most are anticipated to be passenger mutations with no contribution to tumourigenesis, but this awaits experimental confirmation [203].

Other studies have applied next-generation sequencing to search for somatic alterations within only the transcribed genes, or transcriptome, of tumours and tumour-derived cell lines, by a process known as RNA sequencing [204–213]. Because transcriptomes are significantly smaller than whole genomes, RNA sequencing represents a more cost-effective alternative than whole-genome sequencing to search for mutations within coding genes. In addition, transcriptome analysis can provide insights into the nature and abundance of alternative splice forms. However, one disadvantage of transcriptome sequencing versus exome or whole-genome sequencing is that unstable or low-abundance transcripts might escape detection [214]. Notably, this approach has thus far led to the identification of novel candidate cancer genes within malignant mesotheliomas, as well as a novel genetic signature within granulosa-cell tumours of the ovary [205,206]. A recurrent somatic mutation within the *FOXL2* gene was identified within almost all adult cases of ovarian granulosa cell tumours, opening up the possibility for improved diagnostic classification of this rare malignancy [205]. The relevance of massively parallel transcriptome analysis to histopathology is reviewed elsewhere in this issue [215]. Paired end-read transcriptome sequencing has also been applied to search for gene fusions resulting from chromosomal rearrangements within tumours [216,217]. Likewise, massively parallel sequencing, using a paired end-read strategy, has been used to catalogue structural rearrangements and copy number alterations present within two lung cancer cell lines [199]. Approximately one-third of somatic rearrangements identified among these cell lines had escaped detection by other methods, testifying to the sensitivity of this approach to catalogue genomic rearrangements.

Large-scale initiatives to search comprehensively for somatic alterations in human cancers

The past few years have seen the formation of increasingly more organized efforts to comprehensively and systematically search for the somatic alterations that underlie human cancer. The Cancer Genome Project of the Wellcome Trust Sanger Institute (UK) spearheaded the systematic search for somatic alterations in human tumours and tumour-derived cell lines, and continues this mission. Ongoing projects include the resequencing of 4000 candidate cancer genes from a series of cancer cell lines derived from a variety of solid tumours and from a collection of clear cell renal carcinomas (http://www.sanger.ac.uk/genetics/CGP/Studies/). Additionally, 800 tumour-derived cell lines are being interrogated for copy number alterations.

In 2005, the National Human Genome Research Institute and the National Cancer Institute (USA) launched The Cancer Genome Atlas (TCGA), a publicly-funded initiative to systematically catalogue genomic alterations present within the major forms of human cancer (http://cancergenome.nih.gov/about/mission.asp). This began with a 3 year pilot study interrogating glioblastoma multiforme, ovarian cancer and lung cancer, using integrated genomic approaches to search for changes in gene copy number, sequence and expression. In the pilot phase, ~600 candidate cancer genes were selected for Sanger

sequencing [218]. The analysis of glioblastomas by the TCGA network revealed novel somatic alterations in this disease, including mutations within the *PIK3R1* gene, the regulatory subunit of PI3K, and indicated that disruption of the TP53, RB1 and receptor tyrosine kinase-mediated signalling pathways is a core element of gliomagenesis. It also pointed to a possible mechanistic link between the methylation of the *MGMT* gene, which identifies subgroups of glioblastoma patients sensitive to temozolomide treatment, and the appearance of *MSH6* mutations, which are associated with temozolomide resistance [218, 219]. TCGA is moving out of its pilot phase with the objective of comprehensively identifying somatic alterations in 20–25 major tumour types by 2014 (http://www.cancer.gov/recovery). The data are publicly available through both open and controlled access via the Cancer Genome Workbench (CGWB), a feature of the Cancer Biomedical Informatics Grid (caBIG) [220,221].

One limitation experienced in the pilot phase of TCGA was the difficulty in obtaining sufficient amounts of tumour tissue for integrated genomic analyses using multiple platforms [218]. This is likely to become less of a limitation as next-generation sequencing, a single method that requires significantly less starting material for analysis, is incorporated into this endeavour. Building on the experience gained in the pilot phase of TCGA, the NCI and the Foundation for the National Institutes of Health launched the Childhood Cancer Therapeutically Applicable Research to Generate Effective Treatments (TARGET) Initiative (http://target.cancer.gov/). TARGET aims to identify therapeutic targets for childhood cancers, including acute lymphoblastic leukaemia (ALL) and neuroblastoma, by characterizing cancer genomes and transcriptomes. Already studies within this initiative have revealed that somatic alterations within the *IKZF1* and janus kinase (*JAK*) genes occur in a subset of paediatric patients with B cell progenitor ALL and are predictive of poor outcome in such cases [222]. This is also significant because it suggests the potential for therapeutic intervention of molecularly defined subsets of paediatric ALL patients with JAK inhibitors [223].

In 2008, the International Cancer Genome Consortium (ICGC) was formed in an effort to standardize the approaches by which genomic alterations are identified in human cancers (http://www.icgc.org/). The consortium aims to produce comprehensive catalogues of the genomic alterations of up to 50 clinically and societally significant cancer types and subtypes over the course of a decade. Tumours of the pancreas, ovary, stomach, liver, breast and oral cavity, as well as chronic lymphocytic leukaemia, have thus far been prioritized for analysis.

Catalogues of genetic alterations in cancer

With the emergence of the post-genomic era, and the potential to interrogate human tumour genomes for somatic alterations at unprecedented resolution, came the need to organize and integrate mutation data into searchable catalogues. In 2004, the Wellcome Trust Sanger Institute established the Cancer Gene Census, a catalogue of genes mutated in human cancer at a higher frequency than expected by chance alone [224]. This catalogue includes genes that have undergone somatic mutations in sporadic tumours as well as those that are mutated in the germline of cancer families. From 2004 to 2009 the number of consensus cancer genes increased by 40%, from 291 to 410 genes, coinciding with concerted efforts to systematically interrogate cancer genomes for somatic alterations in gene sequence and copy number. A complementary resource, the Catalogue of Somatic Mutations in Cancer (COSMIC), catalogues somatic mutation frequencies in benign and malignant tumours as well as tumour-derived cell lines [225,226]. The genotypic information available through COSMIC is not limited to just the consensus cancer genes but also includes mutated genes that do not meet the requirement for a consensus cancer gene, as well as genes that have

been found not to harbour cancer-associated mutations. Currently, almost 5000 genes and ~340 000 tumours have been curated, resulting in a catalogue of approximately 13 000 unique mutations (Figure 3). Yet other catalogues of mutations serve more specific purposes. The Mitelman Database of Chromo-some Aberrations in Cancer systematically catalogues structural and numerical chromosomal alterations and their clinical associations, as reported in the literature [227,228]. The Genetic Alterations in Cancer database curates mutations associated with exposures to specific chemical, physical and biological agents implicated in tumourigenesis [229]. Numerous other databases catalogue the frequency and spectrum of mutations within individual cancer genes with the goal of facilitating the clinical interpretation of variants (reviewed in [230]).

As the number of somatic mutations identified in cancer genomes continues to grow, it seems likely that additional, gene-specific catalogues will emerge. In this regard, guidelines to standardize the reporting of both somatically mutated and non-mutated tumour samples have recently been proposed, in order to establish a foundation for the accurate interpretation of their clinical and biological significance [230]. Setting and adhering to such guidelines will be critical as we enter a new era in cancer genomics, in which next-generation sequencing of tumour genomes will undoubtedly lead to an explosion in the wealth of mutation data.

Concluding remarks

Over the course of the next decade we will gain an unparalleled appreciation of the somatic alterations present within human cancers, within protein-coding genes, non-coding RNA genes and non-coding regions of the genome, as well as the mitochondrial genome. However, this is just a first step towards identifying bone fide cancer genes. What necessarily follows will be the detailed functional characterization of individual candidate cancer genes, to determine whether and how they contribute to a tumourigenic phenotype. In parallel, functional genetic screens in mice and large-scale RNA interference screens may also guide the confirmation of causal cancer genes [131,163–172] (reviewed in [173–175]).

Once a causal cancer gene has been identified, it is important to recognize that not all somatic mutations within the gene will be functionally equivalent; some mutations might be consequential and others functionally insignificant [154]. For example, certain *EGFR* mutations predict clinical responses to targeted therapies, whereas other mutations within this gene predict clinical resistance [115,231–236]. Likewise, the larger genomic context of a mutated gene can also affect its interpretation. For example, in glioblastomas it is the co-expression of a mutated form of *EGFR* (*EGFRvIII*) and *PTEN* that correlates with clinical sensitivity to small molecule inhibitors of *EGFR* [237]. Conversely, the absence of a somatic mutation within tumour cells can also be used to guide molecularly directed therapies. For example, in both the USA and Europe the drugs cetuximab and panitumumab (monoclonal antibodies against EGFR) are recommended only for advanced colorectal patients whose tumours contain a non-mutated form of the *KRAS* gene (reviewed in [238]). Therefore, before new tests are introduced into the clinical setting it will be imperative to establish the clinicopathological associations of individual mutations, or lack thereof.

How the systematic cataloguing of somatic alterations in cancer genomes will transform diagnostic and clinical practices remains to be seen. One could envision a new battery of specific tests for each tumour type and subtype, based on altered genes or altered pathways, or eventually routine sequencing of individual tumour genomes. Such approaches might ultimately be coupled with methods that noninvasively isolate circulating tumour cells from peripheral blood samples, permitting the serial analyses of cancer genomes to dynamically monitor tumour evolution and adjust clinical management accordingly (reviewed in [239]).

Whatever the final outcome, the systematic analyses of cancer genomes using new technologies [242] that will take place over the next decade has the potential to create a paradigm shift in the diagnosis, prognosis and treatment of individual cancer patients, as Nowell speculated over 30 years ago [11].

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. Boveri, T. Zur Frage der Entstehung Maligner Tumouren. Gustav Fischer; Jena: 1914.

2. Nowell P, Hungerford D. A minute chromosome in human chronic granulocytic leukemia. Science. 1960; 132:1497.

3. Vogelstein B, Fearon ER, Hamilton SR, Kern SE, Preisinger AC, Leppert M, et al. Genetic alterations during colorectal tumour development. N Engl J Med. 1988; 319:525–532. [PubMed: 2841597]

4. Feinberg AP, Vogelstein B. Hypomethylation distinguishes genes of some human cancers from their normal counterparts. Nature. 1983; 301:89–92. [PubMed: 6185846]

5. Calin GA, Dumitru CD, Shimizu M, Bichi R, Zupo S, Noch E, et al. Frequent deletions and down-regulation of micro-RNA genes *miR15* and *miR16* at 13q14 in chronic lymphocytic leukemia. Proc Natl Acad Sci USA. 2002; 99:15524–15529. [PubMed: 12434020]

6. Olson P, Lu J, Zhang H, Shai A, Chun MG, Wang Y, et al. MicroRNA dynamics in the stages of tumourigenesis correlate with hallmark capabilities of cancer. Genes Dev. 2009; 23:2152–2165. [PubMed: 19759263]

7. Hanahan D, Weinberg RA. The hallmarks of cancer. Cell. 2000; 100:57–70. [PubMed: 10647931]

8. Wolman. Karyotypic progression in human tumours. Cancer Metast Rev. 1983; 2:257–293.

9. Sandberg AA. Cytogenetics and molecular genetics of bone and soft-tissue tumours. Am J Med Genet. 2002; 115:189–193. [PubMed: 12407700]

10. Atkin NB. Solid tumour cytogenetics. Progress since 1979. Cancer Genet Cytogenet. 1989; 40:3–12. [PubMed: 2667752]

11. Nowell PC. The clonal evolution of tumour cell populations. Science. 1976; 194:23–28. [PubMed: 959840]

12. International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. Nature. 2004; 431:931–945. [PubMed: 15496913]

13. Druker BJ, Talpaz M, Resta DJ, Peng B, Buchdunger E, Ford JM, 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]

14. Slamon DJ, Leyland-Jones B, Shak S, Fuchs H, Paton V, Bajamonde A, 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]

15. Vogel CL, Cobleigh MA, Tripathy D, Gutheil JC, Harris LN, Fehrenbacher L, et al. Efficacy and safety of trastuzumab as a single agent in first-line treatment of HER2-overexpressing metastatic breast cancer. J Clin Oncol. 2002; 20:719–726. [PubMed: 11821453]

16. Romond EH, Perez EA, Bryant J, Suman VJ, Geyer CE Jr, Davidson NE, et al. Trastuzumab plus adjuvant chemotherapy for operable HER2-positive breast cancer. N Engl J Med. 2005; 353:1673–1684. [PubMed: 16236738]

17. Piccart-Gebhart MJ, Procter M, Leyland-Jones B, Goldhirsch A, Untch M, Smith I, et al. Trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer. N Engl J Med. 2005; 353:1659–1672. [PubMed: 16236737]

18. Guilhot F, Apperley J, Kim DW, Bullorsky EO, Baccarani M, Roboz GJ, et al. Dasatinib induces significant hematologic and cytogenetic responses in patients with imatinib-resistant or - intolerant chronic myeloid leukemia in accelerated phase. Blood. 2007; 109:4143–4150. [PubMed: 17264298]

19. Kantarjian H, Pasquini R, Hamerschlak N, Rousselot P, Holowiecki J, Jootar S, et al. Dasatinib or high-dose imatinib for chronic-phase chronic myeloid leukemia after failure of first-line imatinib: a randomized phase 2 trial. Blood. 2007; 109:5143–5150. [PubMed: 17317857]

20. Kantarjian HM, Giles F, Gattermann N, Bhalla K, Alimena G, Palandri F, et al. Nilotinib (formerly AMN107), a highly selective BCR-ABL tyrosine kinase inhibitor, is effective in patients with Philadelphia chromosome-positive chronic myelogenous leukemia in chronic phase following imatinib resistance and intolerance. Blood. 2007; 110:3540–3546. [PubMed: 17715389]

21. Kantarjian H, Giles F, Wunderle L, Bhalla K, O'Brien S, Wassmann B, et al. Nilotinib in imatinib-resistant CML and Philadelphia chromosome-positive ALL. N Engl J Med. 2006; 354:2542–2551. [PubMed: 16775235]

22. Talpaz M, Shah NP, Kantarjian H, Donato N, Nicoll J, Paquette R, et al. Dasatinib in imatinib-resistant Philadelphia chromosome-positive leukemias. N Engl J Med. 2006; 354:2531–2541. [PubMed: 16775234]

23. Hartwell LH, Szankasi P, Roberts CJ, Murray AW, Friend SH. Integrating genetic approaches into the discovery of anticancer drugs. Science. 1997; 278:1064–1068. [PubMed: 9353181]

24. Farmer H, McCabe N, Lord CJ, Tutt AN, Johnson DA, Richardson TB, et al. Targeting the DNA repair defect in *BRCA* mutant cells as a therapeutic strategy. Nature. 2005; 434:917–921. [PubMed: 15829967]

25. Jones C, Plummer ER. PARP inhibitors and cancer therapy — early results and potential applications. Br J Radiol. 2008; 81(spec no. 1):S2–5. [PubMed: 18819994]

26. Ratnam K, Low JA. Current development of clinical inhibitors of poly(ADP-ribose) polymerase in oncology. Clin Cancer Res. 2007; 13:1383–1388. [PubMed: 17332279]

27. Davidson, E. The Regulatory Genome. Gene Regulatory Networks in Development and Evolution. Academic Press; Amsterdam: 2006.

28. Lever E, Sheer D. The role of nuclear organization in cancer. J Pathol. 2009 DOI:10.1002/path. 2651.

29. Cazier J- B, Tomlinson I. General lessons from large-scale studies to identify human cancer predisposition genes. J Pathol. 2009 DOI:10.1002/path.2650.

30. Le Quesne JP, Spriggs KA, Bushell M, Willis AE. Dysregulation of protein synthesis and disease. J Pathol. 2009 DOI:10.1002/path.2627.

31. Hall PA, Russell SH. New perspectives on neoplasia and the RNA world. Hematol Oncol. 2005; 23:49–53. [PubMed: 16216032]

32. Croce CM. Causes and consequences of microRNA dysregulation in cancer. Nat Rev Genet. 2009; 10:704–714. [PubMed: 19763153]

33. Kampranis SC, Tsichlis PN. Histone demethylases and cancer. Adv Cancer Res. 2009; 102:103–169. [PubMed: 19595308]

34. Venables JP. Unbalanced alternative splicing and its significance in cancer. Bioessays. 2006; 28:378–386. [PubMed: 16547952]

35. Sotiropoulou G, Pampalakis G, Lianidou E, Mourelatos Z. Emerging roles of microRNAs as molecular switches in the integrated circuit of the cancer cell. RNA. 2009; 15:1443–1461. [PubMed: 19561119]

36. Jones PA, Baylin SB. The epigenomics of cancer. Cell. 2007; 128:683–692. [PubMed: 17320506]

37. Bernstein BE, Meissner A, Lander ES. The mammalian epigenome. Cell. 2007; 128:669–681. [PubMed: 17320505]

38. Sonenberg N, Hinnebusch AG. Regulation of translation initiation in eukaryotes: mechanisms and biological targets. Cell. 2009; 136:731–745. [PubMed: 19239892]

39. Stehelin D, Varmus HE, Bishop JM, Vogt PK. DNA related to the transforming gene(s) of avian sarcoma viruses is present in normal avian DNA. Nature. 1976; 260:170–173. [PubMed: 176594]

40. Rous P. A sarcoma of the fowl transmissible by an agent seperable from the tumour cells. J Exp Med. 1911; 13:397–410. [PubMed: 19867421]

41. Reddy EP, Reynolds RK, Santos E, Barbacid M. A point mutation is responsible for the acquisition of transforming properties by the T24 human bladder carcinoma oncogene. Nature. 1982; 300:149–152. [PubMed: 7133135]

42. Parada LF, Tabin CJ, Shih C, Weinberg RA. Human EJ bladder carcinoma oncogene is homologue of Harvey sarcoma virus *ras* gene. Nature. 1982; 297:474–478. [PubMed: 6283357]

43. Tabin CJ, Bradley SM, Bargmann CI, Weinberg RA, Papageorge AG, Scolnick EM, et al. Mechanism of activation of a human oncogene. Nature. 1982; 300:143–149. [PubMed: 6290897]

44. Taparowsky E, Suard Y, Fasano O, Shimizu K, Goldfarb M, Wigler M. Activation of the T24 bladder carcinoma transforming gene is linked to a single amino acid change. Nature. 1982; 300:762–765. [PubMed: 7177195]

45. 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]

46. de Klein A, van Kessel AG, Grosveld G, Bartram CR, Hagemeijer A, Bootsma D, et al. A cellular oncogene is translocated to the Philadelphia chromosome in chronic myelocytic leukaemia. Nature. 1982; 300:765–767. [PubMed: 6960256]

47. Groffen J, Stephenson JR, Heisterkamp N, de Klein A, Bartram CR, Grosveld G. Philadelphia chromosomal breakpoints are clustered within a limited region, bcr, on chromosome 22. Cell. 1984; 36:93–99. [PubMed: 6319012]

48. Heisterkamp N, Stam K, Groffen J, de Klein A, Grosveld G. Structural organization of the *bcr* gene and its role in the Ph′ translocation. Nature. 1985; 315:758–761. [PubMed: 2989703]

49. Konopka JB, Watanabe SM, Singer JW, Collins SJ, Witte ON. Cell lines and clinical isolates derived from Ph1-positive chronic myelogenous leukemia patients express c-abl proteins with a common structural alteration. Proc Natl Acad Sci USA. 1985; 82:1810–1814. [PubMed: 3856862]

50. Ben-Neriah Y, Daley GQ, Mes-Masson AM, Witte ON, Baltimore D. The chronic myelogenous leukemia-specific P210 protein is the product of the *bcr/abl* hybrid gene. Science. 1986; 233:212–214. [PubMed: 3460176]

51. Alitalo K, Schwab M, Lin CC, Varmus HE, Bishop JM. Homogeneously staining chromosomal regions contain amplified copies of an abundantly expressed cellular oncogene (*c-myc*) in malignant neuroendocrine cells from a human colon carcinoma. Proc Natl Acad Sci USA. 1983; 80:1707–1711. [PubMed: 6300869]

52. Dalla-Favera R, Bregni M, Erikson J, Patterson D, Gallo RC, Croce CM. Human *cmyc* oncogene is located on the region of chromosome 8 that is translocated in Burkitt lymphoma cells. Proc Natl Acad Sci USA. 1982; 79:7824–7827. [PubMed: 6961453]

53. Erikson J, Nishikura K, ar-Rushdi A, Finan J, Emanuel B, Lenoir G, et al. Translocation of an immunoglobulin *κ* locus to a region 3′ of an unrearranged *cmyc* oncogene enhances c-myc transcription. Proc Natl Acad Sci USA. 1983; 80:7581–7585. [PubMed: 6424112]

54. Nishikura K, ar-Rushdi A, Erikson J, Watt R, Rovera G, Croce CM. Differential expression of the normal and of the translocated human *c-myc* oncogenes in B cells. Proc Natl Acad Sci USA. 1983; 80:4822–4826. [PubMed: 6308654]

55. Edwards PA. Fusion genes and chromosome translocations in the common epithelial cancers. J Pathol. 2009 DOI:10.1002/ path.2632.

56. Weinberg RA. Fewer and fewer oncogenes. Cell. 1982; 30:3–4. [PubMed: 6751559]

57. Sozzi G, Testi MA, Croce CM. Advances in cancer cytogenetics. J Cell Biochem. 1999; 32–33(suppl):173–182.

58. Weinstein IB. Cancer. Addiction to oncogenes — the Achilles heel of cancer. Science. 2002; 297:63–64. [PubMed: 12098689]

59. Hunter T, Cooper JA. Protein–tyrosine kinases. Annu Rev Biochem. 1985; 54:897–930. [PubMed: 2992362]

60. Blume-Jensen P, Hunter T. Oncogenic kinase signalling. Nature. 2001; 411:355–365. [PubMed: 11357143]

61. Knudson AG Jr. Mutation and cancer: statistical study of retinoblastoma. Proc Natl Acad Sci USA. 1971; 68:820–823. [PubMed: 5279523]

62. Knudson AG Jr, Meadows AT, Nichols WW, Hill R. Chromosomal deletion and retinoblastoma. N Engl J Med. 1976; 295:1120–1123. [PubMed: 980006]

63. Francke U, Kung F. Sporadic bilateral retinoblastoma and 13q-chromosomal deletion. Med Pediatr Oncol. 1976; 2:379–385. [PubMed: 1004381]

64. Cavenee WK, Dryja TP, Phillips RA, Benedict WF, Godbout R, Gallie BL, et al. Expression of recessive alleles by chromosomal mechanisms in retinoblastoma. Nature. 1983; 305:779–784. [PubMed: 6633649]

65. Friend SH, Bernards R, Rogelj S, Weinberg RA, Rapaport JM, Albert DM, et al. A human DNA segment with properties of the gene that predisposes to retinoblastoma and osteosarcoma. Nature. 1986; 323:643–646. [PubMed: 2877398]

66. Lee WH, Shew JY, Hong FD, Sery TW, Donoso LA, Young LJ, et al. The retinoblastoma susceptibility gene encodes a nuclear phosphoprotein associated with DNA binding activity. Nature. 1987; 329:642–645. [PubMed: 3657987]

67. Fung YK, Murphree AL, T'Ang A, Qian J, Hinrichs SH, Benedict WF. Structural evidence for the authenticity of the human retinoblastoma gene. Science. 1987; 236:1657–1661. [PubMed: 2885916]

68. Huang HJ, Yee JK, Shew JY, Chen PL, Bookstein R, Friedmann T, et al. Suppression of the neoplastic phenotype by replacement of the *RB* gene in human cancer cells. Science. 1988; 242:1563–1566. [PubMed: 3201247]

69. Ferguson-Smith AC, Reik W, Surani MA. Genomic imprinting and cancer. Cancer Surv. 1990; 9:487–503. [PubMed: 2101722]

70. Vogelstein B, Fearon ER, Kern SE, Hamilton SR, Preisinger AC, Nakamura Y, et al. Allelotype of colorectal carcinomas. Science. 1989; 244:207–211. [PubMed: 2565047]

71. Baker SJ, Fearon ER, Nigro JM, Hamilton SR, Preisinger AC, Jessup JM, et al. Chromosome 17 deletions and *p53* gene mutations in colorectal carcinomas. Science. 1989; 244:217–221. [PubMed: 2649981]

72. Baker SJ, Markowitz S, Fearon ER, Willson JK, Vogelstein B. Suppression of human colorectal carcinoma cell growth by wild-type p53. Science. 1990; 249:912–915. [PubMed: 2144057]

73. Chen PL, Chen YM, Bookstein R, Lee WH. Genetic mechanisms of tumour suppression by the human p53 gene. Science. 1990; 250:1576–1580. [PubMed: 2274789]

74. Goldgar DE, Green P, Parry DM, Mulvihill JJ. Multipoint linkage analysis in neurofibromatosis type I: an international collaboration. Am J Hum Genet. 1989; 44:6–12. [PubMed: 2491784]

75. Cawthon RM, Weiss R, Xu GF, Viskochil D, Culver M, Stevens J, et al. A major segment of the neurofibromatosis type 1 gene: cDNA sequence, genomic structure, and point mutations. Cell. 1990; 62:193–201. [PubMed: 2114220]

76. Wallace MR, Marchuk DA, Andersen LB, Letcher R, Odeh HM, Saulino AM, et al. Type 1 neurofibromatosis gene: identification of a large transcript disrupted in three NF1 patients. Science. 1990; 249:181–186. [PubMed: 2134734]

77. European Chromosome 16 Tuberous Sclerosis Consortium. Identification and characterization of the tuberous sclerosis gene on chromosome 16. Cell. 1993; 75:1305–1315. [PubMed: 8269512]

78. van Slegtenhorst M, de Hoogt R, Hermans C, Nellist M, Janssen B, Verhoef S, et al. Identification of the tuberous sclerosis gene TSC1 on chromosome 9q34. Science. 1997; 277:805–808. [PubMed: 9242607]

79. Miki Y, Swensen J, Shattuck-Eidens D, Futreal PA, Harsh-man K, Tavtigian S, et al. A strong candidate for the breast and ovarian cancer susceptibility gene *BRCA1*. Science. 1994; 266:66–71. [PubMed: 7545954]

80. Wooster R, Bignell G, Lancaster J, Swift S, Seal S, Mangion J, et al. Identification of the breast cancer susceptibility gene *BRCA2*. Nature. 1995; 378:789–792. [PubMed: 8524414]

81. Latif F, Tory K, Gnarra J, Yao M, Duh FM, Orcutt ML, et al. Identification of the von Hippel–Lindau disease tumour suppressor gene. Science. 1993; 260:1317–1320. [PubMed: 8493574]

82. Guilford P, Hopkins J, Harraway J, McLeod M, McLeod N, Harawira P, et al. E-cadherin germline mutations in familial gastric cancer. Nature. 1998; 392:402–405. [PubMed: 9537325]

83. Haber DA, Buckler AJ, Glaser T, Call KM, Pelletier J, Sohn RL, et al. An internal deletion within an 11p13 zinc finger gene contributes to the development of Wilms' tumour. Cell. 1990; 61:1257–1269. [PubMed: 2163761]

84. Kinzler KW, Nilbert MC, Su LK, Vogelstein B, Bryan TM, Levy DB, et al. Identification of FAP locus genes from chromosome 5q21. Science. 1991; 253:661–665. [PubMed: 1651562]

85. Nishisho I, Nakamura Y, Miyoshi Y, Miki Y, Ando H, Horii A, et al. Mutations of chromosome 5q21 genes in FAP and colorectal cancer patients. Science. 1991; 253:665–669. [PubMed: 1651563]

86. Groden J, Thliveris A, Samowitz W, Carlson M, Gelbert L, Albertsen H, et al. Identification and characterization of the familial adenomatous polyposis coli gene. Cell. 1991; 66:589–600. [PubMed: 1651174]

87. Joslyn G, Carlson M, Thliveris A, Albertsen H, Gelbert L, Samowitz W, et al. Identification of deletion mutations and three new genes at the familial polyposis locus. Cell. 1991; 66:601–613. [PubMed: 1678319]

88. Trofatter JA, MacCollin MM, Rutter JL, Murrell JR, Duyao MP, Parry DM, et al. A novel moesin-, ezrin-, radixin-like gene is a candidate for the neurofibromatosis 2 tumour suppressor. Cell. 1993; 72:791–800. [PubMed: 8453669]

89. Knudson AG. Anti-oncogenes and human cancer. Proc Natl Acad Sci USA. 1993; 90:10914–10921. [PubMed: 7902574]

90. Papadopoulos N, Nicolaides NC, Liu B, Parsons R, Lengauer C, Palombo F, et al. Mutations of GTBP in genetically unstable cells. Science. 1995; 268:1915–1917. [PubMed: 7604266]

91. Papadopoulos N, Nicolaides NC, Wei YF, Ruben SM, Carter KC, Rosen CA, et al. Mutation of a *mutL* homolog in hereditary colon cancer. Science. 1994; 263:1625–1629. [PubMed: 8128251]

92. Nicolaides NC, Papadopoulos N, Liu B, Wei YF, Carter KC, Ruben SM, et al. Mutations of two PMS homologues in hereditary nonpolyposis colon cancer. Nature. 1994; 371:75–80. [PubMed: 8072530]

93. Kamb A, Gruis NA, Weaver-Feldhaus J, Liu Q, Harshman K, Tavtigian SV, et al. A cell cycle regulator potentially involved in genesis of many tumour types. Science. 1994; 264:436–440. [PubMed: 8153634]

94. Li J, Yen C, Liaw D, Podsypanina K, Bose S, Wang SI, et al. PTEN, a putative protein tyrosine phosphatase gene mutated in human brain, breast, and prostate cancer. Science. 1997; 275:1943–1947. [PubMed: 9072974]

95. Narla G, Heath KE, Reeves HL, Li D, Giono LE, Kimmelman AC, et al. KLF6, a candidate tumour suppressor gene mutated in prostate cancer. Science. 2001; 294:2563–2566. [PubMed: 11752579]

96. Dulbecco R. A turning point in cancer research: sequencing the human genome. Science. 1986; 231:1055–1056. [PubMed: 3945817]

97. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, et al. The sequence of the human genome. Science. 2001; 291:1304–1351. [PubMed: 11181995]

98. McPherson JD, Marra M, Hillier L, Waterston RH, Chinwalla A, Wallis J, et al. A physical map of the human genome. Nature. 2001; 409:934–941. [PubMed: 11237014]

99. Slamon DJ, Clark GM, Wong SG, Levin WJ, Ullrich A, McGuire WL. Human breast cancer: correlation of relapse and survival with amplification of the *HER-2/neu* oncogene. Science. 1987; 235:177–182. [PubMed: 3798106]

100. Cobleigh MA, Vogel CL, Tripathy D, Robert NJ, Scholl S, Fehrenbacher L, et al. 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. 1999; 17:2639–2648. [PubMed: 10561337]

101. Davies H, Bignell GR, Cox C, Stephens P, Edkins S, Clegg S, et al. Mutations of the *BRAF* gene in human cancer. Nature. 2002; 417:949–954. [PubMed: 12068308]

102. Parsons DW, Wang TL, Samuels Y, Bardelli A, Cummins JM, DeLong L, et al. Colorectal cancer: mutations in a signalling pathway. Nature. 2005; 436:792. [PubMed: 16094359]

103. Loriaux MM, Levine RL, Tyner JW, Frohling S, Scholl C, Stoffregen EP, et al. High-throughput sequence analysis of the tyrosine kinome in acute myeloid leukemia. Blood. 2008; 111:4788–4796. [PubMed: 18252861]

104. Tomasson MH, Xiang Z, Walgren R, Zhao Y, Kasai Y, Miner T, et al. Somatic mutations and germline sequence variants in the expressed tyrosine kinase genes of patients with *de novo* acute myeloid leukemia. Blood. 2008; 111:4797–4808. [PubMed: 18270328]

105. Bardelli A, Parsons DW, Silliman N, Ptak J, Szabo S, Saha S, et al. Mutational analysis of the tyrosine kinome in colorectal cancers. Science. 2003; 300:949. [PubMed: 12738854]

106. Rand V, Huang J, Stockwell T, Ferriera S, Buzko O, Levy S, et al. Sequence survey of receptor tyrosine kinases reveals mutations in glioblastomas. Proc Natl Acad Sci USA. 2005; 102:14344–14349. [PubMed: 16186508]

107. Dutt A, Salvesen HB, Chen TH, Ramos AH, Onofrio RC, Hatton C, et al. Drug-sensitive *FGFR2* mutations in endometrial carcinoma. Proc Natl Acad Sci USA. 2008; 105:8713–8717. [PubMed: 18552176]

108. Brown JR, Levine RL, Thompson C, Basile G, Gilliland DG, Freedman AS. Systematic genomic screen for tyrosine kinase mutations in CLL. Leukemia. 2008; 22:1966–1969. [PubMed: 18754031]

109. Samuels Y, Wang Z, Bardelli A, Silliman N, Ptak J, Szabo S, et al. High frequency of mutations of the *PIK3CA* gene in human cancers. Science. 2004; 304:554. [PubMed: 15016963]

110. Wang Z, Shen D, Parsons DW, Bardelli A, Sager J, Szabo S, et al. Mutational analysis of the tyrosine phosphatome in colorectal cancers. Science. 2004; 304:1164–1166. [PubMed: 15155950]

111. Hunter C, Smith R, Cahill DP, Stephens P, Stevens C, Teague J, et al. A hypermutation phenotype and somatic *MSH6* mutations in recurrent human malignant gliomas after alkylator chemotherapy. Cancer Res. 2006; 66:3987–3991. [PubMed: 16618716]

112. Bignell G, Smith R, Hunter C, Stephens P, Davies H, Greenman C, et al. Sequence analysis of the protein kinase gene family in human testicular germ-cell tumours of adolescents and adults. Genes Chromosomes Cancer. 2006; 45:42–46. [PubMed: 16175573]

113. Davies H, Hunter C, Smith R, Stephens P, Greenman C, Bignell G, et al. Somatic mutations of the protein kinase gene family in human lung cancer. Cancer Res. 2005; 65:7591–7595. [PubMed: 16140923]

114. Stephens P, Edkins S, Davies H, Greenman C, Cox C, Hunter C, et al. A screen of the complete protein kinase gene family identifies diverse patterns of somatic mutations in human breast cancer. Nat Genet. 2005; 37:590–592. [PubMed: 15908952]

115. Paez JG, Janne PA, Lee JC, Tracy S, Greulich H, Gabriel S, et al. *EGFR* mutations in lung cancer: correlation with clinical response to gefitinib therapy. Science. 2004; 304:1497–1500. [PubMed: 15118125]

116. Jiang J, Paez JG, Lee JC, Bo R, Stone RM, DeAngelo DJ, et al. Identifying and characterizing a novel activating mutation of the FLT3 tyrosine kinase in AML. Blood. 2004; 104:1855–1858. [PubMed: 15178581]

117. Levine RL, Wadleigh M, Cools J, Ebert BL, Wernig G, Huntly BJ, et al. Activating mutation in the tyrosine kinase *JAK2* in polycythemia vera, essential thrombocythemia, and myeloid metaplasia with myelofibrosis. Cancer Cell. 2005; 7:387–397. [PubMed: 15837627]

118. Greenman C, Stephens P, Smith R, Dalgliesh GL, Hunter C, Bignell G, et al. Patterns of somatic mutation in human cancer genomes. Nature. 2007; 446:153–158. [PubMed: 17344846]

119. Engelman JA. Targeting PI3K signalling in cancer: opportunities, challenges and limitations. Nat Rev Cancer. 2009; 9:550–562. [PubMed: 19629070]

120. Thomas RK, Baker AC, Debiasi RM, Winckler W, Laframboise T, Lin WM, et al. High-throughput oncogene mutation profiling in human cancer. Nat Genet. 2007; 39:347–351. [PubMed: 17293865]

121. Tyner JW, Erickson H, Deininger MW, Willis SG, Eide CA, Levine RL, et al. High-throughput sequencing screen reveals novel, transforming RAS mutations in myeloid leukemia patients. Blood. 2009; 113:1749–1755. [PubMed: 19075190]

122. Teng DH, Chen Y, Lian L, Ha PC, Tavtigian SV, Wong AK. Mutation analyses of 268 candidate genes in human tumour cell lines. Genomics. 2001; 74:352–364. [PubMed: 11414763]

123. Ding L, Getz G, Wheeler DA, Mardis ER, McLellan MD, Cibulskis K, et al. Somatic mutations affect key pathways in lung adenocarcinoma. Nature. 2008; 455:1069–1075. [PubMed: 18948947]

124. Link DC, Kunter G, Kasai Y, Zhao Y, Miner T, McLellan MD, et al. Distinct patterns of mutations occurring in *de novo* AML versus AML arising in the setting of severe congenital neutropenia. Blood. 2007; 110:1648–1655. [PubMed: 17494858]

125. Wang TL, Rago C, Silliman N, Ptak J, Markowitz S, Will-son JK, et al. Prevalence of somatic alterations in the colorectal cancer cell genome. Proc Natl Acad Sci USA. 2002; 99:3076–3080. [PubMed: 11867767]

126. Barber TD, McManus K, Yuen KW, Reis M, Parmigiani G, Shen D, et al. Chromatid cohesion defects may underlie chromosome instability in human colorectal cancers. Proc Natl Acad Sci USA. 2008; 105:3443–3448. [PubMed: 18299561]

127. Wang Z, Cummins JM, Shen D, Cahill DP, Jallepalli PV, Wang TL, et al. Three classes of genes mutated in colorectal cancers with chromosomal instability. Cancer Res. 2004; 64:2998–3001. [PubMed: 15126332]

128. van Haaften G, Dalgliesh GL, Davies H, Chen L, Bignell G, Greenman C, et al. Somatic mutations of the histone H3K27 demethylase gene *UTX* in human cancer. Nat Genet. 2009; 41:521–523. [PubMed: 19330029]

129. Palavalli LH, Prickett TD, Wunderlich JR, Wei X, Burrell AS, Porter-Gill P, et al. Analysis of the matrix metalloproteinase family reveals that MMP8 is often mutated in melanoma. Nat Genet. 2009; 41:518–520. [PubMed: 19330028]

130. Ikediobi ON, Davies H, Bignell G, Edkins S, Stevens C, O'Meara S, et al. Mutation analysis of 24 known cancer genes in the NCI-60 cell line set. Mol Cancer Ther. 2006; 5:2606–2612. [PubMed: 17088437]

131. Scholl C, Frohling S, Dunn IF, Schinzel AC, Barbie DA, Kim SY, et al. Synthetic lethal interaction between oncogenic KRAS dependency and STK33 suppression in human cancer cells. Cell. 2009; 137:821–834. [PubMed: 19490892]

132. Martin SA, Hewish M, Lord CJ, Ashworth A. Genomic instability and the selection of treatments for cancer. J Pathol. 2009 DOI:10.1002/path.2631.

133. Carter NP. Methods and strategies for analyzing copy number variation using DNA microarrays. Nat Genet. 2007; 39:S16–21. [PubMed: 17597776]

134. International HapMap Consortium. The International HapMap Project. Nature. 2003; 426:789–796. [PubMed: 14685227]

135. International HapMap Consortium. A haplotype map of the human genome. Nature. 2005; 437:1299–1320. [PubMed: 16255080]

136. Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL, Gibbs RA, et al. A second generation human haplotype map of over 3.1 million SNPs. Nature. 2007; 449:851–861. [PubMed: 17943122]

137. Greshock J, Feng B, Nogueira C, Ivanova E, Perna I, Nathanson K, et al. A comparison of DNA copy number profiling platforms. Cancer Res. 2007; 67:10173–10180. [PubMed: 17968032]

138. Cox C, Bignell G, Greenman C, Stabenau A, Warren W, Stephens P, et al. A survey of homozygous deletions in human cancer genomes. Proc Natl Acad Sci USA. 2005; 102:4542–4547. [PubMed: 15761058]

139. Zhao X, Weir BA, LaFramboise T, Lin M, Beroukhim R, Garraway L, et al. Homozygous deletions and chromosome amplifications in human lung carcinomas revealed by single nucleotide polymorphism array analysis. Cancer Res. 2005; 65:5561–5570. [PubMed: 15994928]

140. Weir BA, Woo MS, Getz G, Perner S, Ding L, Beroukhim R, et al. Characterizing the cancer genome in lung adenocarcinoma. Nature. 2007; 450:893–898. [PubMed: 17982442]

141. George RE, Attiyeh EF, Li S, Moreau LA, Neuberg D, Li C, et al. Genome-wide analysis of neuroblastomas using high-density single nucleotide polymorphism arrays. PLoS One. 2007; 2:e255. [PubMed: 17327916]

142. Yamamoto G, Nannya Y, Kato M, Sanada M, Levine RL, Kawamata N, et al. Highly sensitive method for genome-wide detection of allelic composition in nonpaired, primary tumour specimens by use of affymetrix single-nucleotide-polymorphism genotyping microarrays. Am J Hum Genet. 2007; 81:114–126. [PubMed: 17564968]

143. Peiffer DA, Le JM, Steemers FJ, Chang W, Jenniges T, Garcia F, et al. High-resolution genomic profiling of chromosomal aberrations using Infinium whole-genome genotyping. Genome Res. 2006; 16:1136–1148. [PubMed: 16899659]

144. Nevins JR, Potti A. Mining gene expression profiles: expression signatures as cancer phenotypes. Nat Rev Genet. 2007; 8:601–609. [PubMed: 17607306]

145. Prensner JR, Chinnaiyan AM. Oncogenic gene fusions in epithelial carcinomas. Curr Opin Genet Dev. 2009; 19:82–91. [PubMed: 19233641]

146. Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, et al. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science. 1999; 286:531–537. [PubMed: 10521349]

147. Garraway LA, Widlund HR, Rubin MA, Getz G, Berger AJ, Ramaswamy S, et al. Integrative genomic analyses identify MITF as a lineage survival oncogene amplified in malignant melanoma. Nature. 2005; 436:117–122. [PubMed: 16001072]

148. Sjoblom T, Jones S, Wood LD, Parsons DW, Lin J, Barber TD, et al. The consensus coding sequences of human breast and colorectal cancers. Science. 2006; 314:268–274. [PubMed: 16959974]

149. Jones S, Zhang X, Parsons DW, Lin JC, Leary RJ, Angenendt P, et al. Core signaling pathways in human pancreatic cancers revealed by global genomic analyses. Science. 2008; 321:1801–1806. [PubMed: 18772397]

150. Parsons DW, Jones S, Zhang X, Lin JC, Leary RJ, Angenendt P, et al. An integrated genomic analysis of human glioblastoma multiforme. Science. 2008; 321:1807–1812. [PubMed: 18772396]

151. Wood LD, Parsons DW, Jones S, Lin J, Sjoblom T, Leary RJ, et al. The genomic landscapes of human breast and colorectal cancers. Science. 2007; 318:1108–1113. [PubMed: 17932254]

152. Leary RJ, Lin JC, Cummins J, Boca S, Wood LD, Parsons DW, et al. Integrated analysis of homozygous deletions, focal amplifications, and sequence alterations in breast and colorectal cancers. Proc Natl Acad Sci USA. 2008; 105:16224–16229. [PubMed: 18852474]

153. Beerenwinkel N, Antal T, Dingli D, Traulsen A, Kinzler KW, Velculescu VE, et al. Genetic progression and the waiting time to cancer. PLoS Comput Biol. 2007; 3:e225. [PubMed: 17997597]

154. Frohling S, Scholl C, Levine RL, Loriaux M, Boggon TJ, Bernard OA, et al. Identification of driver and passenger mutations of FLT3 by high-throughput DNA sequence analysis and functional assessment of candidate alleles. Cancer Cell. 2007; 12:501–513. [PubMed: 18068628]

155. Parmigiani G, Lin J, Boca S. Response to comments on 'The consensus coding sequences of human breast and colorectal cancers'. Science. 2007; 317:1500d.

156. Forrest WF, Cavet G. Comment on 'The consensus coding sequences of human breast and colorectal cancers'. Science. 2007; 317:1500. [PubMed: 17872427]

157. Getz G, Hofling H, Mesirov JP, Golub TR, Meyerson M, Tibshirani R, et al. Comment on 'The consensus coding sequences of human breast and colorectal cancers'. Science. 2007; 317:1500. [PubMed: 17872428]

158. Rubin AF, Green P. Comment on 'The consensus coding sequences of human breast and colorectal cancers'. Science. 2007; 317:1500. [PubMed: 17872429]

159. Ng PC, Henikoff S. Predicting deleterious amino acid substitutions. Genome Res. 2001; 11:863–874. [PubMed: 11337480]

160. Ramensky V, Bork P, Sunyaev S. Human non-synonymous SNPs: server and survey. Nucleic Acids Res. 2002; 30:3894–3900. [PubMed: 12202775]

161. Carter H, Chen S, Isik L, Tyekucheva S, Velculescu VE, Kinzler KW, et al. Cancer-specific high-throughput annotation of somatic mutations: computational prediction of driver missense mutations. Cancer Res. 2009; 69:6660–6667. [PubMed: 19654296]

162. Kimchi-Sarfaty C, Oh JM, Kim IW, Sauna ZE, Calcagno AM, Ambudkar SV, et al. A "silent" polymorphism in the MDR1 gene changes substrate specificity. Science. 2007; 315:525–528. [PubMed: 17185560]

163. Starr TK, Allaei R, Silverstein KA, Staggs RA, Sarver AL, Bergemann TL, et al. A transposon-based genetic screen in mice identifies genes altered in colorectal cancer. Science. 2009; 323:1747–1750. [PubMed: 19251594]

164. Schlabach MR, Luo J, Solimini NL, Hu G, Xu Q, Li MZ, et al. Cancer proliferation gene discovery through functional genomics. Science. 2008; 319:620–624. [PubMed: 18239126]

165. Silva JM, Marran K, Parker JS, Silva J, Golding M, Schlabach MR, et al. Profiling essential genes in human mammary cells by multiplex RNAi screening. Science. 2008; 319:617–620. [PubMed: 18239125]

166. Brummelkamp TR, Fabius AW, Mullenders J, Madiredjo M, Velds A, Kerkhoven RM, et al. An shRNA barcode screen provides insight into cancer cell vulnerability to MDM2 inhibitors. Nat Chem Biol. 2006; 2:202–206. [PubMed: 16474381]

167. Berns K, Hijmans EM, Mullenders J, Brummelkamp TR, Velds A, Heimerikx M, et al. A large-scale RNAi screen in human cells identifies new components of the p53 pathway. Nature. 2004; 428:431–437. [PubMed: 15042092]

168. Westbrook TF, Martin ES, Schlabach MR, Leng Y, Liang AC, Feng B, et al. A genetic screen for candidate tumour suppressors identifies REST. Cell. 2005; 121:837–848. [PubMed: 15960972]

169. Berns K, Horlings HM, Hennessy BT, Madiredjo M, Hij-mans EM, Beelen K, et al. A functional genetic approach identifies the PI3K pathway as a major determinant of trastuzumab resistance in breast cancer. Cancer Cell. 2007; 12:395–402. [PubMed: 17936563]

170. Kolfschoten IG, van Leeuwen B, Berns K, Mullenders J, Beijersbergen RL, Bernards R, et al. A genetic screen identifies PITX1 as a suppressor of RAS activity and tumourigenicity. Cell. 2005; 121:849–858. [PubMed: 15960973]

171. Ngo VN, Davis RE, Lamy L, Yu X, Zhao H, Lenz G, et al. A loss-of-function RNA interference screen for molecular targets in cancer. Nature. 2006; 441:106–110. [PubMed: 16572121]

172. Luo B, Cheung HW, Subramanian A, Sharifnia T, Okamoto M, Yang X, et al. Highly parallel identification of essential genes in cancer cells. Proc Natl Acad Sci USA. 2008; 105:20380–20385. [PubMed: 19091943]

173. Kool J, Berns A. High-throughput insertional mutagenesis screens in mice to identify oncogenic networks. Nat Rev Cancer. 2009; 9:389–399. [PubMed: 19461666]

174. Quigley D, Balmain A. Systems genetics analysis of cancer susceptibility: from mouse models to humans. Nat Rev Genet. 2009; 10:651–657. [PubMed: 19636343]

175. Zender L, Lowe SW. Integrative oncogenomic approaches for accelerated cancer-gene discovery. Curr Opin Oncol. 2008; 20:72–76. [PubMed: 18043259]

176. Birney E, Stamatoyannopoulos JA, Dutta A, Guigo R, Gingeras TR, Margulies EH, et al. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature. 2007; 447:799–816. [PubMed: 17571346]

177. Ambros V. The functions of animal microRNAs. Nature. 2004; 431:350–355. [PubMed: 15372042]

178. Wilusz JE, Sunwoo H, Spector DL. Long noncoding RNAs: functional surprises from the RNA world. Genes Dev. 2009; 23:1494–1504. [PubMed: 19571179]

179. Calin GA, Liu CG, Ferracin M, Hyslop T, Spizzo R, Sevignani C, et al. Ultraconserved regions encoding ncRNAs are altered in human leukemias and carcinomas. Cancer Cell. 2007; 12:215–229. [PubMed: 17785203]

180. Bejerano G, Pheasant M, Makunin I, Stephen S, Kent WJ, Mattick JS, et al. Ultraconserved elements in the human genome. Science. 2004; 304:1321–1325. [PubMed: 15131266]

181. Storz G. An expanding universe of noncoding RNAs. Science. 2002; 296:1260–1263. [PubMed: 12016301]

182. Hill DA, Ivanovich J, Priest JR, Gurnett CA, Dehner LP, Desruisseau D, et al. DICER1 mutations in familial pleuropulmonary blastoma. Science. 2009; 325:965. [PubMed: 19556464]

183. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA. 1977; 74:5463–5467. [PubMed: 271968]

184. Shendure JA, Porreca GJ, Church GM. Overview of DNA sequencing strategies. Curr Protoc Mol Biol. 2008; 7 Unit 7.1.

185. Mardis ER. Next-generation DNA sequencing methods. Annu Rev Genom Hum Genet. 2008; 9:387–402.

186. Ansorge WJ. Next-generation DNA sequencing techniques. Nat Biotechnol. 2009; 25:195–203.

187. Shendure J, Ji H. Next-generation DNA sequencing. Nat Biotechnol. 2008; 26:1135–1145. [PubMed: 18846087]

188. Voelkerding KV, Dames SA, Durtschi JD. Next-generation sequencing: from basic research to diagnostics. Clin Chem. 2009; 55:641–658. [PubMed: 19246620]

189. Kim JI, Ju YS, Park H, Kim S, Lee S, Yi JH, et al. A highly annotated whole-genome sequence of a Korean individual. Nature. 2009; 460:1011–1015. [PubMed: 19587683]

190. Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J, Brown CG, et al. Accurate whole human genome sequencing using reversible terminator chemistry. Nature. 2008; 456:53–59. [PubMed: 18987734]

191. Wang J, Wang W, Li R, Li Y, Tian G, Goodman L, et al. The diploid genome sequence of an Asian individual. Nature. 2008; 456:60–65. [PubMed: 18987735]

192. Wheeler DA, Srinivasan M, Egholm M, Shen Y, Chen L, McGuire A, et al. The complete genome of an individual by massively parallel DNA sequencing. Nature. 2008; 452:872–876. [PubMed: 18421352]

193. Pushkarev D, Neff NF, Quake. Single-molecule sequencing of an individual human genome. Nat Biotechnol. 2009; 27:847–852. [PubMed: 19668243]

194. Mardis ER. Anticipating the 1000 dollar genome. Genome Biol. 2006; 7:112. [PubMed: 17224040]

195. Service RF. Gene sequencing. The race for the $1000 genome. Science. 2006; 311:1544–1546. [PubMed: 16543431]

196. Thomas RK, Nickerson E, Simons JF, Janne PA, Tengs T, Yuza Y, et al. Sensitive mutation detection in heterogeneous cancer specimens by massively parallel picoliter reactor sequencing. Nat Med. 2006; 12:852–855. [PubMed: 16799556]

197. Campbell PJ, Pleasance ED, Stephens PJ, Dicks E, Rance R, Goodhead I, et al. Subclonal phylogenetic structures in cancer revealed by ultra-deep sequencing. Proc Natl Acad Sci USA. 2008; 105:13081–13086. [PubMed: 18723673]

198. Kidd JM, Cooper GM, Donahue WF, Hayden HS, Sampas N, Graves T, et al. Mapping and sequencing of structural variation from eight human genomes. Nature. 2008; 453:56–64. [PubMed: 18451855]

199. Campbell PJ, Stephens PJ, Pleasance ED, O'Meara S, Li H, Santarius T, 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]

200. Korbel JO, Urban AE, Affourtit JP, Godwin B, Grubert F, Simons JF, et al. Paired-end mapping reveals extensive structural variation in the human genome. Science. 2007; 318:420–426. [PubMed: 17901297]

201. Ng P, Tan JJ, Ooi HS, Lee YL, Chiu KP, Fullwood MJ, et al. Multiplex sequencing of paired-end ditags (MS-PET): a strategy for the ultra-high-throughput analysis of transcriptomes and genomes. Nucleic Acids Res. 2006; 34:e84. [PubMed: 16840528]

202. Ley TJ, Mardis ER, Ding L, Fulton B, McLellan MD, Chen K, et al. DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome. Nature. 2008; 456:66–72. [PubMed: 18987736]

203. Mardis ER, Ding L, Dooling DJ, Larson DE, McLellan MD, Chen K, et al. Recurring mutations found by sequencing an acute myeloid leukemia genome. N Engl J Med. 2009; 361:1058–1066. [PubMed: 19657110]

204. Zhao Q, Caballero OL, Levy S, Stevenson BJ, Iseli C, de Souza SJ, et al. Transcriptome-guided characterization of genomic rearrangements in a breast cancer cell line. Proc Natl Acad Sci USA. 2009; 106:1886–1891. [PubMed: 19181860]

205. Shah SP, Kobel M, Senz J, Morin RD, Clarke BA, Wiegand KC, et al. Mutation of FOXL2 in granulosa-cell tumours of the ovary. N Engl J Med. 2009; 360:2719–2729. [PubMed: 19516027]

206. Sugarbaker DJ, Richards WG, Gordon GJ, Dong L, De Rienzo A, Maulik G, et al. Transcriptome sequencing of malignant pleural mesothelioma tumours. Proc Natl Acad Sci USA. 2008; 105:3521–3526. [PubMed: 18303113]

207. Morin R, Bainbridge M, Fejes A, Hirst M, Krzywinski M, Pugh T, et al. Profiling the HeLa S3 transcriptome using randomly primed cDNA and massively parallel short-read sequencing. Biotechniques. 2008; 45:81–94. [PubMed: 18611170]

208. Bainbridge MN, Warren RL, Hirst M, Romanuik T, Zeng T, Go A, et al. Analysis of the prostate cancer cell line LNCaP transcriptome using a sequencing-by-synthesis approach. BMC Genom. 2006; 7:246.

209. Guffanti A, Iacono M, Pelucchi P, Kim N, Solda G, Croft LJ, et al. A transcriptional sketch of a primary human breast cancer by 454 deep sequencing. BMC Genomics. 2009; 10:163. [PubMed: 19379481]

210. Tsuchihara K, Suzuki Y, Wakaguri H, Irie T, Tanimoto K, Hashimoto S, et al. Massive transcriptional start site analysis of human genes in hypoxia cells. Nucleic Acids Res. 2009; 37:2249–2263. [PubMed: 19237398]

211. Hashimoto S, Qu W, Ahsan B, Ogoshi K, Sasaki A, Nakatani Y, et al. High-resolution analysis of the 5′ -end transcriptome using a next generation DNA sequencer. PLoS One. 2009; 4:e4108. [PubMed: 19119315]

212. Li H, Lovci MT, Kwon YS, Rosenfeld MG, Fu XD, Yeo GW. Determination of tag density required for digital transcriptome analysis: application to an androgen-sensitive prostate cancer model. Proc Natl Acad Sci USA. 2008; 105:20179–20184. [PubMed: 19088194]

213. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet. 2009; 10:57–63. [PubMed: 19015660]

214. Stalder L, Muhlemann O. The meaning of nonsense. Trends Cell Biol. 2008; 18:315–321. [PubMed: 18524595]

215. Aparicio SA, Huntsman DG. Does massively parallel DNA resequencing signify the end of histopathology as we know it? J Pathol. 2009 DOI:10.1002/path.2636.

216. Maher CA, Palanisamy N, Brenner JC, Cao X, Kalyana-Sundaram S, Luo S, et al. Chimeric transcript discovery by paired-end transcriptome sequencing. Proc Natl Acad Sci USA. 2009; 106:12353–12358. [PubMed: 19592507]

217. Maher CA, Kumar-Sinha C, Cao X, Kalyana-Sundaram S, Han B, Jing X, et al. Transcriptome sequencing to detect gene fusions in cancer. Nature. 2009; 458:97–101. [PubMed: 19136943]

218. Cancer Genome Atlas Research Network. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. Nature. 2008; 455:1061–1068. [PubMed: 18772890]

219. Cahill DP, Levine KK, Betensky RA, Codd PJ, Romany CA, Reavie LB, et al. Loss of the mismatch repair protein MSH6 in human glioblastomas is associated with tumour progression during temozolomide treatment. Clin Cancer Res. 2007; 13:2038–2045. [PubMed: 17404084]

220. Zhang J, Finney RP, Rowe W, Edmonson M, Yang SH, Dracheva T, et al. Systematic analysis of genetic alterations in tumours using Cancer GenomeWorkBench (CGWB). Genome Res. 2007; 17:1111–1117. [PubMed: 17525135]

221. von Eschenbach AC, Buetow K. Cancer Informatics Vision: caBIG. Cancer Inform. 2007; 2:22–24. [PubMed: 19458755]

222. Mullighan CG, Su X, Zhang J, Radtke I, Phillips LA, Miller CB, et al. Deletion of IKZF1 and prognosis in acute lymphoblastic leukemia. N Engl J Med. 2009; 360:470–480. [PubMed: 19129520]

223. Mullighan CG, Zhang J, Harvey RC, Collins-Underwood JR, Schulman BA, Phillips LA, et al. JAK mutations in high-risk childhood acute lymphoblastic leukemia. Proc Natl Acad Sci USA. 2009; 106:9414–9418. [PubMed: 19470474]

224. Futreal PA, Coin L, Marshall M, Down T, Hubbard T, Wooster R, et al. A census of human cancer genes. Nat Rev Cancer. 2004; 4:177–183. [PubMed: 14993899]

225. Bamford S, Dawson E, Forbes S, Clements J, Pettett R, Dogan A, et al. The COSMIC (Catalogue of Somatic Mutations in Cancer) database and website. Br J Cancer. 2004; 91:355–358. [PubMed: 15188009]

226. Forbes SA, Bhamra G, Bamford S, Dawson E, Kok C, Clements J, et al. The Catalogue of Somatic Mutations in Cancer (COSMIC). Curr Protoc Hum Genet. 2008; 10 Unit 10.1.

227. Mitelman, F.; Johansson, B.; Mertens, F. Mitelman Database of Chromosome Aberrations in Cancer. 2009. http://cgapncinihgov/Chromosomes/Mitelman

228. Mitelman F. Catalogue of chromosome aberrations in cancer. Cytogenet Cell Genet. 1983; 36:1–515. [PubMed: 6627995]

229. Jackson MA, Lea I, Rashid A, Peddada SD, Dunnick JK. Genetic alterations in cancer knowledge system: analysis of gene mutations in mouse and human liver and lung tumours. Toxicol Sci. 2006; 90:400–418. [PubMed: 16410370]

230. Olivier M, Petitjean A, Teague J, Forbes S, Dunnick JK, den Dunnen JT, et al. Somatic mutation databases as tools for molecular epidemiology and molecular pathology of cancer: proposed guidelines for improving data collection, distribution, and integration. Hum Mutat. 2009; 30:275–282. [PubMed: 19006239]

231. Lynch TJ, Bell DW, Sordella R, Gurubhagavatula S, Okimoto RA, Brannigan BW, 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]

232. Kobayashi S, Boggon TJ, Dayaram T, Janne PA, Kocher O, Meyerson M, et al. EGFR mutation and resistance of non-small-cell lung cancer to gefitinib. N Engl J Med. 2005; 352:786–792. [PubMed: 15728811]

233. Bean J, Riely GJ, Balak M, Marks JL, Ladanyi M, Miller VA, et al. Acquired resistance to epidermal growth factor receptor kinase inhibitors associated with a novel T854A mutation in a patient with *EGFR*-mutant lung adenocarcinoma. Clin Cancer Res. 2008; 14:7519–7525. [PubMed: 19010870]

234. Balak MN, Gong Y, Riely GJ, Somwar R, Li AR, Zakowski MF, et al. Novel D761Y and common secondary T790M mutations in epidermal growth factor receptor-mutant lung adenocarcinomas with acquired resistance to kinase inhibitors. Clin Cancer Res. 2006; 12:6494–6501. [PubMed: 17085664]

235. Costa DB, Halmos B, Kumar A, Schumer ST, Huberman MS, Boggon TJ, et al. BIM mediates EGFR tyrosine kinase inhibitor-induced apoptosis in lung cancers with oncogenic *EGFR* mutations. PLoS Med. 2007; 4:1669–1679. [PubMed: 17973572]

236. Greulich H, Chen TH, Feng W, Janne PA, Alvarez JV, Zappaterra M, et al. Oncogenic transformation by inhibitor-sensitive and -resistant *EGFR* mutants. PLoS Med. 2005; 2:e313. [PubMed: 16187797]

237. Mellinghoff IK, Wang MY, Vivanco I, Haas-Kogan DA, Zhu S, Dia EQ, et al. Molecular determinants of the response of glioblastomas to EGFR kinase inhibitors. N Engl J Med. 2005; 353:2012–2024. [PubMed: 16282176]

238. Normanno N, Tejpar S, Morgillo F, De Luca A, Van Cutsem E, Ciardiello F. Implications for KRAS status and EGFR-targeted therapies in metastatic CRC. Nat Rev Clin Oncol. 2009; 6:519–527. [PubMed: 19636327]

239. Alix-Panabieres C, Riethdorf S, Pantel K. Circulating tumour cells and bone marrow micrometastasis. Clin Cancer Res. 2008; 14:5013–5021. [PubMed: 18698019]

240. Sandberg AA. Application of cytogenetics in neoplastic diseases. Crit Rev Clin Lab Sci. 1985; 22:219–274. [PubMed: 3899509]

241. Forbes S, Clements J, Dawson E, Bamford S, Webb T, Dogan A, et al. Cosmic 2005. Br J Cancer. 2006; 94:318–322. [PubMed: 16421597]

242. McCaughan F, Dear PH. Single-molecule genomics. J Pathol. 2009 DOI:10.1002/path.2647.

Figure 1.
Timeline of seminal hypotheses, research discoveries and research initiatives that have led to an improved understanding of the genetic aetiology of human tumourigenesis within the last century [1,2,4,5,11,12,39,41–44,58,61,65,66,135,148,240]. The consensus cancer gene data were obtained from the Wellcome Trust Sanger Institute Cancer Genome Project website (http://www.sanger.ac.uk/genetics/CGP)

Figure 2.
Signalling pathways and processes. (A) The 12 pathways and processes whose component genes were genetically altered in most pancreatic cancers. (B, C) Two pancreatic cancers (Pa14C and Pa10X) and the specific genes that are mutated in them. The positions around the circles in (B, C) correspond to the pathways and processes in (A). Several pathway components overlapped, as illustrated by the *BMPR2* mutation that presumably disrupted both the SMAD4 and Hedgehog signalling pathways in Pa10X. Additionally, not all 12 processes and pathways were altered in every pancreatic cancer, as exemplified by the fact that no mutations known to affect DNA damage control were observed in Pa10X. N.O., not observed. From Jones *et al. Science* 2008; **321**: 1801–1806 [149]. Reprinted with permission from AAAS
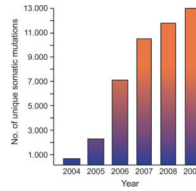
Figure 3.
Number of unique somatic mutations catalogued in human cancers. The mutation data were obtained from the Wellcome Trust Sanger Institute Catalogue Of Somatic Mutations In Cancer (COSMIC) website (http://www.sanger.ac.uk/cosmic) [225,226,241]. The mutation data displayed for 2004–2008 are derived from the year-end COSMIC data. The mutation data for 2009 is derived from the COSMIC v42 release. The COSMIC v42 release does not incorporate data from large-scale resequencing projects [102,148,150,151]