
Cancer Genomics and Precision Medicine: A Way Toward Early Diagnosis and Effective Cancer Treatment

2

Ritika Tiwari and Bushra Ateeq

2.1 Introduction

Since the first draft of the Human Genome Project was completed in April 2003, biomedical researchers have been mining and extrapolating genomic data toward the goal of improving human health and realizing medical benefits. The promise of “personalized oncomedicine,” the matching of therapeutics to appropriate molecular targets in individual cancer patients, lies in the convergence of cancer researchers, computational biologist, and clinicians to identify the driving mutations involved in tumor progression and metastasis and pursue appropriate therapies. The virtual concept of “cancer genome” in the development of uncontrolled cell growth was conceived as early as late nineteenth and early twentieth century by Theodor Boveri (Boveri 2008). Boveri hypothesized that malignant tumors could be the result of a certain abnormal condition of the chromosomes arising from multipolar mitosis. Several decades later the discovery of the Philadelphia chromosome as the genetic driver of chronic myeloid leukemia (CML) provided the experimental evidence for Boveri’s hypothesis (Nowell and Hungerford 1961).

The first description of the translocation between chromosomes 9 and 22 in the Philadelphia chromosome was reported by Janet D. Rowley in 1980 (Rowley 1980); however, it was another 10 years before the genes involved in the rearrangement were identified as breakpoint cluster region (BCR; chromosome 22) and v-abl Abelson murine leukemia viral oncogene homolog (ABL; chromosome 9) (Groffen et al. 1984). BCR-ABL fusion protein was demonstrated to function as a constitutively activated tyrosine kinase that stimulated proliferation of myeloid cells, leading to the development of CML (Lugo et al. 1990). Subsequently, a new therapeutic agent, imatinib mesylate (Gleevec), was developed that targets the kinase domain of

R. Tiwari • B. Ateeq (✉)
Department of Biological Sciences and Bioengineering,
Indian Institute of Technology, Kanpur 208016, UP, India
e-mail: bushra@iitk.ac.in

the fusion protein which essentially reversed the high mortality rate of CML patients (Druker et al. 1996).

The discovery of the Philadelphia chromosome, the identification of BCR-ABL fusion protein, and the development of Gleevec targeting the oncoprotein are a classic early example of personalized medicine. While the identification of BCR-ABL fusion to the development of Gleevec in CML represents a linear path from basic molecular discovery to medical success, most cancers are far more complex. Unlike CML (and most hematological cancers) where there is only a single causative genetic lesion, most solid tumors are highly heterogeneous and many harbor private mutation(s). However, over the last decade or so the field of medical oncology has experienced several remarkable breakthroughs. The amplification of the *HER2/neu* gene that was identified in ~20% of the breast cancer patients (Schechter et al. 1984; Slamon et al. 1987) led to the development of a monoclonal antibody, trastuzumab (Herceptin; Genentech), to treat HER2-positive breast cancer women (Robertson 1998); and lung cancer patients that harbor specific *EGFR* mutations were found to respond to gefitinib (Iressa) and erlotinib (Tarceva) that target these mutations (Lynch et al. 2004; Pao et al. 2004). *PARP1* inhibitor, olaparib, demonstrated antitumor activity in cancers associated with specific *BRCA1* or *BRCA2* mutations that impair the DNA repair pathway (ClinicalTrials.gov number, NCT00516373) (Fong et al. 2009).

The recent advances in high-throughput sequencing technologies that now make clinical sequencing economically feasible, combined with advanced computational approaches and higher-resolution analyses, have allowed us to obtain an unprecedented view of the genomic landscape of mutations/aberrations in individual cancer patients. Integrative sequencing strategies including whole-genome sequencing, targeted whole-exome sequencing, transcriptome sequencing (RNA-Seq), and shallow (5X–15X) paired-end whole-genome sequencing can be applied to uncover clinically significant genetic alterations in tumor specimens of patients and identify targets for existing therapies or direct patients to appropriate clinical trials.

2.2 Integrative Sequencing Strategy

Cancer arises from various genetic/molecular alterations including nucleic acid substitutions, gene fusions and rearrangements, amplifications and deletions, and a host of other aberrations that perturb gene expression (Stratton et al. 2009). Although the cancer genome can harbor multiple mutations, only few are “drivers” that confer clonal growth advantage, are positively selected, and are causally implicated in cancer development in a background of “passenger” mutations. Tumor specimens are often admixtures with varying fractions of normal tissue, or they may contain tumor subclones; therefore high sequencing depth is required for the detection of variants (Fig. 2.1). While whole-genome sequencing could be employed to identify copy number alterations (CNAs) and structural rearrangements at relatively shallow depth (Stephens et al. 2009), accurate identification of

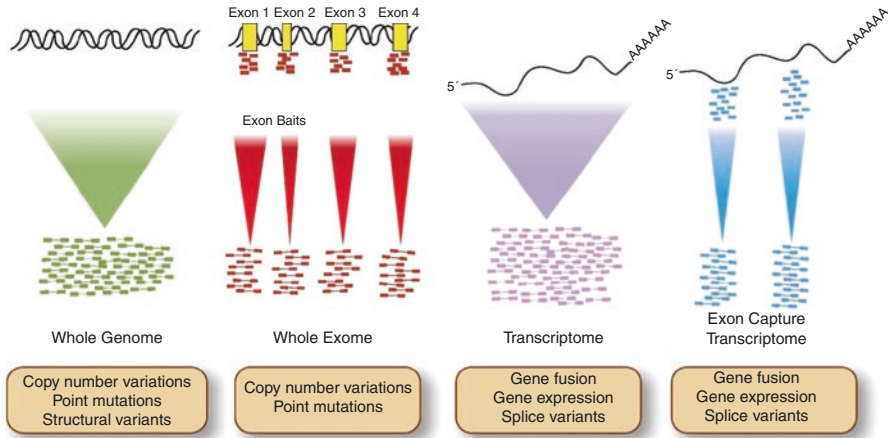


Fig. 2.1 Integrative sequencing of tumors for personalized oncology. Schema representing integration of whole-genome sequencing (green), whole-exome sequencing (red) for 1–2% of the genome, transcriptome or mRNA sequencing (purple), and targeted exome-capture sequencing (blue). Each sequencing strategy can be integrated for detecting genetic aberrations in cancer tissues including structural variants, CNVs, splice variants, point mutations, and gene expression

point mutations requires greater coverage and depth (Meyerson et al. 2010). Somatic mutations distinct from inherited DNA variants are identified by filtering out commonly inherited variants in human populations (>5% allele frequency) that have been registered in databases (Stratton et al. 2009); however, some rare inherited single nucleotide polymorphisms (SNPs) and structural variants may not be registered. Somatic mutations that are highly represented among cancer genes include protein kinases families in various signaling pathways; MAPK/ERK pathway is an example where upstream mutations are found in cell membrane-bound receptor tyrosine kinases such as *EGFR*, *ERBB2*, *FGFR1*, *FGFR2*, *FGFR3*, *PDGFRA*, and *PDGFRB* as well as in the downstream cytoplasmic components *NF1*, *PTPN11*, *HRAS*, *KRAS*, *NRAS*, and *BRAF* (Johnson and Lapadat 2002).

The International Cancer Genome Consortium (ICGC: <http://www.icgc.org/home>) undertook the task of comprehensively characterizing somatically acquired genetic alterations in at least 50 different classes of cancer, including those with the highest global incidence and mortality (Stratton et al. 2009). Currently, the ICGC has received commitments from various funding organizations in Asia, Australia, Europe, North America, and South America for 74 project teams in 17 jurisdictions to study over 25,000 tumor genomes that include cancers of the biliary tract, bladder, blood, bone, brain, breast, cervix, colon, eye, head and neck, kidney, liver, lung, nasopharynx, oral cavity, ovary, pancreas, prostate, rectum, skin, soft tissues, stomach, thyroid, and uterus. The genomic data generated by the participating ICGC members listed in Table 2.1 are made available by the Data Coordination Center through the ICGC website (www.icgc.org).

Table 2.1 List of the participating International Cancer Genome Consortium (ICGC) members for generating genomic data from multiple cancer types

ICGC members	Analyzed tumors
Australia	Ovarian and pancreatic cancer
Canada	Pancreatic, pediatric brain, and prostate cancer
China	Bladder, esophageal, gastric, and renal cancer
European Union/ France	Renal cancer
France	Liver cancer
Germany	Blood, brain, and prostate cancer
India	Oral cancer
Japan	Liver cancer
Saudi Arabia	Thyroid cancer
South Korea	Blood and lung cancer
Spain	Blood cancer
UK	Blood, bone, breast, esophageal, lung, prostate, and skin cancer
USA	Bladder, blood, brain, breast, cervical, colon, gastric, head and neck, liver, lung, ovarian, pancreatic, prostate, rectal, renal, skin, thyroid, and uterine cancer

2.3 Whole-Genome Sequencing

Whole-genome sequencing provides comprehensive characterization of somatic and germline mutations in a specimen. The most commonly used methods to make single nucleotide variant calls are (1) comparison with other sequenced genomes via Single Nucleotide Polymorphism database (dbSNP) and other resources for variant discovery such as the 1000 Genomes Project (www.1000genomes.org) and (2) critical assessment of remaining variant sites by comparison of tumor and matched normal genome. This approach also takes into consideration two primary measures to distinguish high- from low-quality variants (Mardis and Wilson 2009): first, a cumulative base-calling quality value that is summed from the individual quality values of each base identifying the putative variant (assigned by the Illumina's analysis pipeline known as Consensus Assessment of Sequence and Variation—CASAVA software) and second, a mapping quality value assigned by MAQ (Mapping and Assembly with Quality) assessing the genome-wide uniqueness of each aligned read (Li et al. 2008). CASAVA enables genomic builds, SNP calls, insertions/deletions (indels) detection, and count reads from the data generated from one or more runs across a broad range of sequencing applications (Table 2.2). Additionally, MAQ enables both read mapping and genotype calling from simulated and real data by utilizing mate-pair information and estimates the error probability of each read alignment. Recently, whole-genome sequencing of bladder cancers at a median depth of ~80X revealed recurrent protein-inactivating mutations in *CDKN1A* and *FAT1*. Moreover, the Stampy and Platypus programs have been used for mapping and aligning and somatic base substitution/single

Table 2.2 Key terminologies commonly used in next-generation sequencing

Key terminologies	Definition
Number of reads	Total amount of sequence data output by the instrument
Coverage	The average number of reads that align to or “cover” known reference bases
Sequencing depth	The total number of bases sequenced and aligned at a given reference base position
Read length	Number of base pairs of a given read
Error rate	Overall error rates are calculated by dividing the total number of errors by the number of known bases in the reference genome
Paired-end reads	A technology that obtains sequence reads from both ends of a DNA fragment template to generate high-quality, alignable sequence data. Paired-end sequencing facilitates detection of genomic rearrangements and repetitive sequence elements, as well as gene fusions and novel transcripts
Mate-pair reads	Mate-pair sequencing is similar to paired-end sequencing; however, the size of the DNA fragments used as sequencing templates is much longer (1000–10,000 bp). Mate-pair methods are particularly valuable for joining contigs in de novo sequencing and for detecting translocations and large deletions (structural variants)
Multiplexing	Processing of a large number of samples on a high-throughput instrument. Sample multiplexing is a useful technique when targeting specific genomic regions or working with smaller genomes

nucleotide variant (SNV) calls with high confidence ranging from 27,490 to 121,016 (Cazier et al. 2014). A recent study that employed exome sequencing of 50 lethal, heavily treated metastatic castration-resistant prostate cancers (CRPCs) demonstrated recurrent (8.6%) mutations in multiple chromatin- and histone-modifying genes such as *MLL2* and the AR collaborating factor *FOXAI* (3.4%) (Grasso et al. 2012). Nevertheless, evaluating somatic mutations in cancer specimens is often challenging for samples with very low tumor content (percentage of tumor cells in a given specimen) and can limit analysis as the differential variants between tumor and normal sample that could be detected would be very low.

2.4 Transcriptome Sequencing

Transcriptome sequencing (RNA-Seq) provides a comprehensive landscape of the expressed genome that includes all unique RNA transcripts from both coding and noncoding regions. In addition to the sequence and identity of RNA species in a sample, RNA-Seq can identify genomic rearrangements, copy number variations (CNVs), focused indels, and single nucleotide mutations. Importantly, transcriptome sequencing can also provide gene expression data with more sensitivity than microarray experiments (Marioni et al. 2008; Sultan et al. 2008). Utilizing RNA-Seq, Maher et al. discovered novel fusion transcripts by employing single-end reads of various lengths. This approach nominated multiple candidates or chimeras such as *SLC45A3-ELK4* that were independently confirmed as a common “read-through” transcript identified in prostate cancer (Maher et al. 2009). A combination of two

next-generation sequencing platforms was utilized, and the data were integrated to identify fusion transcripts from cancer cell lines; first long reads from RNA-Seq data that partially aligned to the reference genome (Roche 454) were identified as putative fusion transcripts, and then short reads that spanned fusion junctions obtained from a second RNA-Seq dataset (Illumina Genome Analyzer) were integrated with the first dataset to nominate candidate gene fusions. Using this approach, Maher et al. successfully “rediscovered” previously known and novel fusion transcripts in the prostate cancer cell lines LNCaP and VCaP and various prostate tumor samples (Maher et al. 2009). Recently, Kalyana-Sundaram et al. developed a bioinformatics pipeline for explicitly detecting pseudogene transcripts from RNA-Seq data and demonstrated genome-wide expression of pseudogenes, which are ubiquitously expressed in a lineage and/or cancer-specific manner (Kalyana-Sundaram et al. 2012). Briefly, they discovered breast-specific unprocessed pseudogene ATP8A2 Ψ , which possibly arises from the duplication of wild-type ATP8A2, therefore likely harbors similar promoter elements. Similarly, a prostate cancer-specific pseudogene, CXADR- Ψ , was also revealed using the same bioinformatics framework. CXADR- Ψ is a processed pseudogene located on chromosome 15, and parental CXADR protein demonstrates putative tumor suppressor functions, and its loss has been implicated in α -catenin downregulation (Pong et al. 2003).

Chromosomal rearrangements leading to generation of gene fusions represent one of the common mechanisms for the expression of oncogenes in epithelial cancers (Chinnaiyan and Palanisamy 2010). Oncogenic genetic rearrangements were initially thought to be confined to hematological cancers (Mitelman 2000; Mitelman et al. 2007). In 2005, Chinnaiyan and colleagues reported recurrent gene fusions between the transmembrane protease serine 2 (*TMPRSS2*) gene and members of the ETS family of transcription factors, predominantly *ERG* (v-ets erythroblastosis virus E26 oncogene homolog (avian)), in prostate cancer, representing the first discovery of a gene fusion in a solid tumor (Tomlins et al. 2005). Subsequently, various gene fusions in a variety of cancers including breast and lung were discovered (Stephens et al. 2009; Martelli et al. 2009; Natrajan et al. 2014). Gene rearrangements of *SLC45A3-BRAF* (solute carrier family 45, member 3–v-raf murine sarcoma viral oncogene homolog B1) and *ESRP1-RAF1* (epithelial splicing regulatory protein-1–v-raf-1 murine leukemia viral oncogene homolog-1) were discovered in prostate cancer by employing paired-end massively parallel transcriptome (Palanisamy et al. 2010). Importantly, these fusions are potentially “druggable.” Furthermore, identification of these RAF pathway gene rearrangements in a variety of cancer types—prostate and gastric cancers and melanoma—supports the notion that cancers should be stratified by the driving molecular alterations/genetic events rather than by organ site.

Using whole-exome and transcriptome sequencing, a genetic rearrangement between transcriptional repressor *NAB2* and the transcriptional activator *STAT6* was detected in all solitary fibrous tumors (SFT)/hemangiopericytoma cases tested, establishing *NAB2-STAT6* as the causative mutation of SFT (Robinson et al. 2013a). In addition to driving mutations, clinical sequencing can also uncover mechanisms of treatment resistance and disease progression. Recently, ER-positive, metastatic

breast cancer patients underwent sequence analysis that revealed mutations in the ligand-binding domain (LBD) of the estrogen receptor (ESR1) that resulted in constitutive activity and continued responsiveness to anti-estrogen therapies *in vitro*. These results suggest that activating mutations in ESR1 are a key mechanism in acquired endocrine resistance in breast cancer therapy (Robinson et al. 2013b). Moreover, recently identified novel variants of *CDK4*, *LARPI*, *ADD3*, and *PHLPP2* in breast cancer are adding to the repertoire of the cancer transcriptome as well as uncovering novel therapeutic targets (Eswaran et al. 2013). Transcriptome analysis of various cancers also identified recurrent novel fusion involving kinase receptors that can potentially serve as promising drug targets (Stransky et al. 2014).

A distinct advantage of unbiased transcriptome sequencing is the ability to study noncoding RNA species whose role in cellular processes and disease state is becoming increasingly appreciated. Long noncoding RNAs (lncRNAs) play a role in normal cellular processes and are also implicated in cancer progression and metastasis (Crea et al. 2014). Noncoding RNAs (ncRNAs) can be categorized into small (under 200 nucleotides) and large ncRNAs. The small ncRNAs include small nucleolar RNAs (snoRNAs), PIWI-interacting RNAs (piRNAs), small interfering RNAs (siRNAs), and microRNAs (miRNAs) (Amaral et al. 2008). Earlier, the lncRNA known as HOTAIR was found to be aberrantly over-expressed in advanced breast and colorectal cancer, and repressing HOTAIR expression in cancer cells attenuated the invasive potential of the cancer cells (Kogo et al. 2011; Wang and Chang 2011). Thus, RNA-Seq analysis is a powerful tool for understanding the transcriptome landscape of cancers and to molecularly stratify subsets of cancer by mutational classes of genetic aberrations. Further, the resulting datasets from various RNA-Seq methodologies can provide a wide range of information such as gene expression, methylation status, histone modifications, and genomic occupancy of transcription factors and other regulatory protein-binding positions.

2.5 Methylated DNA Immunoprecipitation (MeDIP) Sequencing

The term “epigenetics” was originally coined by Conrad Waddington to describe heritable changes in a cellular phenotype that were independent of alterations in the DNA sequence. All epigenetic changes such as chromatin remodeling, histone modifications, and DNA methylation are highly regulated by a group of chromatin-modifying enzymes. There are at least four known DNA modifications (Baylin and Jones 2011; Wu et al. 2011) and 16 classes of histone modifications (Kouzarides 2007; Tan et al. 2011). Initially, MethylC-seq, a bisulfite conversion method, was developed to analyze the methylome at single-base resolution (Cokus et al. 2008). In this approach, sodium bisulphite converts unmethylated cytosines to uracils leaving 5-methylated cytosines unchanged, and upon amplification by polymerase chain reaction (PCR), unmethylated cytosines appear as thymines and methylated cytosines appear as cytosines (Frommer et al. 1992). A combination of next-generation sequencing (NGS) platforms with established techniques such as chromatin

immunoprecipitation (ChIP-Seq) has yielded an unparalleled view of the epigenome. Importantly, NGS introduced a novel approach to assess genome-wide epigenetic changes in an unbiased manner without the limitations of probe-based microarray platforms. One of the most prevalent epigenetic alterations in cancer is the methylation changes that occur within CpG islands that are present in 70% of all mammalian promoters. CpG island methylation plays a critical role in transcriptional regulation, and it is commonly altered during malignant transformation (Baylin and Jones 2011). Genome-wide mapping of CpG methylation using NGS platforms has confirmed that ~5–10% of normally unmethylated CpG promoter islands become abnormally methylated in various cancer genomes (Ateeq et al. 2008; Szyf 2005). Moreover, CpG hypermethylation of promoters not only affects the expression of protein-coding genes but also the expression of various noncoding RNAs, some of which have a role in oncogenesis (Baylin and Jones 2011).

Kim et al. employed a novel deep-sequencing technique named MethylPlex to enrich for methylated regions of the genome to characterize the DNA methylome map of prostate cancer cells and tissues requiring minimal DNA input. Massively parallel sequencing of the enriched products identified differentially methylated regions (DMRs) and revealed novel insights regarding the genomic loci and functional consequences of DNA methylation in cancer (Kim et al. 2011). This study uncovered 6691 methylated promoters in prostate tissues, 2481 cancer-specific DMRs including several novel DMRs such as WFDC2 promoter that displayed increased levels of methylation in cancer tissues compared to benign tissues and normal prostate epithelial cells.

Whole-genome sequencing in a variety of cancers has identified recurrent somatic mutations in numerous epigenetic regulators as well (Forbes et al. 2011; Stratton et al. 2009). Targeted NGS resequencing of cancer genomes found mutations within EZH2 (enhancer of zeste 2 polycomb repressive complex 2), the catalytic subunit of polycomb repressive complex 2 that is over-expressed in multiple cancers (Cao et al. 2008; Li et al. 2009; Varambally et al. 2008), including lymphoid and myeloid cancers (Khan et al. 2013; Yoshida et al. 2013). Moreover, heterozygous missense mutations resulting in the substitution of tyrosine 641 (Y641) within the SET domain of EZH2 were observed in 22% of patients with diffuse large B-cell lymphoma (Morin et al. 2010). Recurrent mutations in the histone methyltransferase, MLL2, have been discovered in ~90% of follicular lymphoma patients (Morin et al. 2011). Similarly, mutations in UTX, a histone demethylase, were observed in up to 12 histologically distinct cancers (van Haaften et al. 2009). Given these findings, several new drugs against these epigenetic targets are in development (Arrowsmith et al. 2012).

Conclusions

The current genomics era holds tremendous promise for “personalized oncomedicine.” The recent rapid advances in high-throughput sequencing technologies and the field of cancer genomics is catalyzing the discovery of novel “druggable” molecular targets that include oncogenes, protein pathways involved in signaling

cascade, and networks shown to be involved in the pathogenesis of cancer as well as the development of therapies against these targets. In the not too distant future, clinical sequencing of patient tumor specimens to inform therapeutic intervention is likely to be adopted as standard clinical practice.

Acknowledgments We thank Jyoti Athanikar for critically reading the manuscript. B.A. is an Intermediate Fellow of the Wellcome Trust-DBT India Alliance [IA/I(S)/12/2/500635 to BA] and a Young Investigator of the DST-FAST Track scheme.

Competing Interests No competing interests to be disclosed.

About the Author Bushra Ateeq is an Assistant Professor at Department of Biological Sciences and Bioengineering (BSBE), IIT, Kanpur, India. The primary research focus of Dr. Bushra's laboratory is to understand the complex molecular events involved in prostate and breast cancer, identify early diagnostic markers and valuable therapeutic targets. Her lab focuses on experimental evaluation of the functional relevance of the genetic rearrangements, copy number changes, and somatic alterations identified by genomic approaches, using molecular and cellular approaches.

References

- Amaral PP, Dinger ME, Mercer TR, Mattick JS. The eukaryotic genome as an RNA machine. *Science*. 2008;319:1787–9.
- Arrowsmith CH, Bountra C, Fish PV, Lee K, Schapira M. Epigenetic protein families: a new frontier for drug discovery. *Nat Rev Drug Discov*. 2012;11:384–400.
- Ateeq B, Unterberger A, Szyf M, Rabbani SA. Pharmacological inhibition of DNA methylation induces proinvasive and prometastatic genes in vitro and in vivo. *Neoplasia*. 2008;10:266–78.
- Baylin SB, Jones PA. A decade of exploring the cancer epigenome—biological and translational implications. *Nat Rev Cancer*. 2011;11:726–34.
- Boveri T. Concerning the origin of malignant tumours by Theodor Boveri. Translated and annotated by Harris H. *J Cell Sci*. 2008;121 Suppl 1:1–84.
- Cao Q, et al. Repression of E-cadherin by the polycomb group protein EZH2 in cancer. *Oncogene*. 2008;27:7274–84.
- Cazier JB, et al. Whole-genome sequencing of bladder cancers reveals somatic CDKN1A mutations and clinicopathological associations with mutation burden. *Nat Commun*. 2014;5:3756.
- Chinnaiyan AM, Palanisamy N. Chromosomal aberrations in solid tumors. *Prog Mol Biol Transl Sci*. 2010;95:55–94.
- Cokus SJ, et al. Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. *Nature*. 2008;452:215–9.
- Crea F, et al. Identification of a long non-coding RNA as a novel biomarker and potential therapeutic target for metastatic prostate cancer. *Oncotarget*. 2014;5:764–74.
- Druker BJ, et al. Effects of a selective inhibitor of the Abl tyrosine kinase on the growth of Bcr-Abl positive cells. *Nat Med*. 1996;2:561–6.
- Eswaran J, et al. RNA sequencing of cancer reveals novel splicing alterations. *Sci Rep*. 2013;3:1689.
- Fong PC, et al. Inhibition of poly(ADP-ribose) polymerase in tumors from BRCA mutation carriers. *N Engl J Med*. 2009;361:123–34.
- Forbes SA, et al. COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. *Nucleic Acids Res*. 2011;39:D945–50.
- Frommer M, et al. A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proc Natl Acad Sci USA*. 1992;89:1827–31.

- Grasso CS, et al. The mutational landscape of lethal castration-resistant prostate cancer. *Nature*. 2012;487:239–43.
- 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;36:93–9.
- Johnson GL, Lapadat R. Mitogen-activated protein kinase pathways mediated by ERK, JNK, and p38 protein kinases. *Science*. 2002;298:1911–2.
- Kalyana-Sundaram S, et al. Expressed pseudogenes in the transcriptional landscape of human cancers. *Cell*. 2012;149:1622–34.
- Khan SN, et al. Multiple mechanisms deregulate EZH2 and histone H3 lysine 27 epigenetic changes in myeloid malignancies. *Leukemia*. 2013;27:1301–9.
- Kim JH, et al. Deep sequencing reveals distinct patterns of DNA methylation in prostate cancer. *Genome Res*. 2011;21:1028–41.
- Kogo R, et al. Long noncoding RNA HOTAIR regulates polycomb-dependent chromatin modification and is associated with poor prognosis in colorectal cancers. *Cancer Res*. 2011;71:6320–6.
- Kouzarides T. Chromatin modifications and their function. *Cell*. 2007;128:693–705.
- Li X, Gonzalez ME, Toy K, Filzen T, Merajver SD, Kleer CG. Targeted overexpression of EZH2 in the mammary gland disrupts ductal morphogenesis and causes epithelial hyperplasia. *Am J Pathol*. 2009;175:1246–54.
- Li H, Ruan J, Durbin R. Mapping short DNA sequencing reads and calling variants using mapping quality scores. *Genome Res*. 2008;18:1851–8.
- Lugo TG, Pendergast AM, Muller AJ, Witte ON. Tyrosine kinase activity and transformation potency of bcr-abl oncogene products. *Science*. 1990;247:1079–82.
- 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;350:2129–39.
- Maher CA, et al. Transcriptome sequencing to detect gene fusions in cancer. *Nature*. 2009;458:97–101.
- Mardis ER, Wilson RK. Cancer genome sequencing: a review. *Hum Mol Genet*. 2009;18:R163–8.
- Marioni JC, Mason CE, Mane SM, Stephens M, Gilad Y. RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. *Genome Res*. 2008;18:1509–17.
- Martelli MP, et al. EML4-ALK rearrangement in non-small cell lung cancer and non-tumor lung tissues. *Am J Pathol*. 2009;174:661–70.
- Meyerson M, Gabriel S, Getz G. Advances in understanding cancer genomes through second-generation sequencing. *Nat Rev Genet*. 2010;11:685–96.
- Mitelman F. Recurrent chromosome aberrations in cancer. *Mutat Res*. 2000;462:247–53.
- Mitelman F, Johansson B, Mertens F. The impact of translocations and gene fusions on cancer causation. *Nat Rev Cancer*. 2007;7:233–45.
- Morin RD, et al. Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin. *Nat Genet*. 2010;42:181–5.
- Morin RD, et al. Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. *Nature*. 2011;476:298–303.
- Natrajan R, et al. Characterization of the genomic features and expressed fusion genes in micropapillary carcinomas of the breast. *J Pathol*. 2014;232:553–65.
- Nowell PC, Hungerford DA. Chromosome studies in human leukemia. II. Chronic granulocytic leukemia. *J Natl Cancer Inst*. 1961;27:1013–35.
- Palanisamy N, et al. Rearrangements of the RAF kinase pathway in prostate cancer, gastric cancer and melanoma. *Nat Med*. 2010;16:793–8.
- Pao W, Miller V, Zakowski M, Doherty J, Politi K, Sarkaria I, et al. EGF receptor gene mutations are common in lung cancers from “never smokers” and are associated with sensitivity of tumors to gefitinib and erlotinib. *Proc Natl Acad Sci USA*. 2004;101:13306–11.
- Pong RC, et al. Epigenetic regulation of coxsackie and adenovirus receptor (CAR) gene promoter in urogenital cancer cells. *Cancer Res*. 2003;63:8680–6.

- Robertson D. Genentech's anticancer Mab expected by November. *Nat Biotechnol.* 1998;16:615.
- Robinson DR, et al. Identification of recurrent NAB2-STAT6 gene fusions in solitary fibrous tumor by integrative sequencing. *Nat Genet.* 2013a;45:180–5.
- Robinson DR, et al. Activating ESR1 mutations in hormone-resistant metastatic breast cancer. *Nat Genet.* 2013b;45:1446–51.
- Rowley JD. Ph1-positive leukaemia, including chronic myelogenous leukaemia. *Clin Haematol.* 1980;9:55–86.
- Schechter AL, et al. The neu oncogene: an erb-B-related gene encoding a 185,000-Mr tumour antigen. *Nature.* 1984;312:513–6.
- 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;235:177–82.
- Stephens PJ, et al. Complex landscapes of somatic rearrangement in human breast cancer genomes. *Nature.* 2009;462:1005–10.
- Stransky N, Cerami E, Schalm S, Kim JL, Lengauer C. The landscape of kinase fusions in cancer. *Nat Commun.* 2014;5:4846.
- Stratton MR, Campbell PJ, Futreal PA. The cancer genome. *Nature.* 2009;458:719–24.
- Sultan M, et al. A global view of gene activity and alternative splicing by deep sequencing of the human transcriptome. *Science.* 2008;321:956–60.
- Szyf M. Therapeutic implications of DNA methylation. *Future Oncol.* 2005;1:125–35.
- Tan M, et al. Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. *Cell.* 2011;146:1016–28.
- Tomlins SA, et al. Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. *Science.* 2005;310:644–8.
- van Haaften G, et al. Somatic mutations of the histone H3K27 demethylase gene UTX in human cancer. *Nat Genet.* 2009;41:521–3.
- Varambally S, et al. Genomic loss of microRNA-101 leads to overexpression of histone methyltransferase EZH2 in cancer. *Science.* 2008;322:1695–9.
- Wang KC, Chang HY. Molecular mechanisms of long noncoding RNAs. *Mol Cell.* 2011;43:904–14.
- Wu H, et al. Genome-wide analysis of 5-hydroxymethylcytosine distribution reveals its dual function in transcriptional regulation in mouse embryonic stem cells. *Genes Dev.* 2011;25:679–84.
- Yoshida K, et al. The landscape of somatic mutations in down syndrome-related myeloid disorders. *Nat Genet.* 2013;45:1293–9.

Genome Analysis and Human Health

Rawal, L.; Ali, S. (Eds.)

2017, XI, 168 p. 19 illus. in color., Hardcover

ISBN: 978-981-10-4297-3