Cancer Research Review

An Overview of Recent Cancer Research Publications Featuring Illumina® Technology

Table of Contents

This document highlights recent publications that demonstrate the use of Illumina technologies in cancer research. To learn more about the platforms and assays cited, visit www.illumina.com.

Introduction

In cancer research each cancer sample presents the researcher with an altered genome that contains a unique and unpredictable number of point mutations, indels, translocations, fusions, and other aberrations. Since many of these alterations might never have been observed before and might not necessarily reside in coding regions of the genome, whole-genome sequencing is increasingly seen as the only rigorous approach that can find all the variants in a cancer genome. Among all these alterations are a select few that drive the progression of the disease. Based on the assumption that changes in gene expression levels impact disease progression, RNA-Seq is increasingly employed as a useful technique to determine if these genetic alterations impact disease progression. Genetic alterations have the potential to impact all cellular processes, including chromatin structure, DNA methylation, RNA splice variants, RNA editing, and microRNA (miRNA) to name but a few. Real progress in cancer research will come through the measurement and integrated analysis of all these interdependent processes.

The key characteristic of next-generation sequencing technologies is that billions of independent sequence reads are generated in parallel, with each read derived from a single molecule of DNA. The resultant data approximate a random sample of DNA molecules which, in turn, represents the genomes of individual cells contained in the tumor sample.[1] This provides us with a powerful toolbox to untangle the causes and mechanisms of cancer. (See **Technical Considerations** for additional information.)
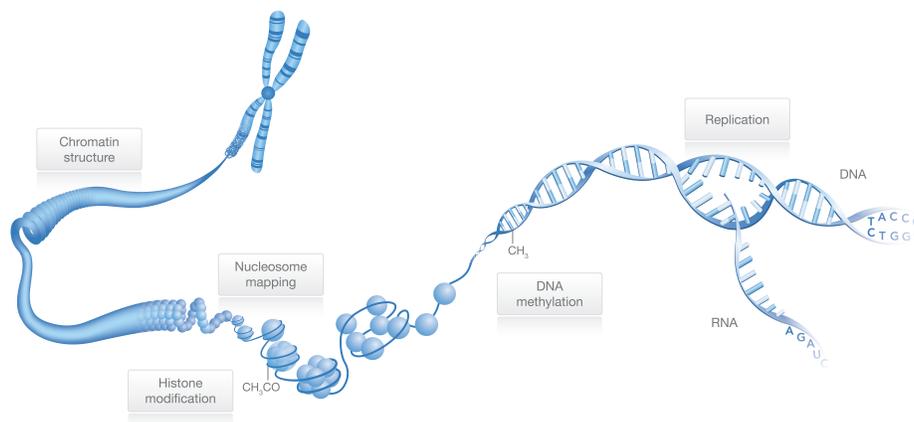
Reviews

Garraway L. A. and Lander E. S. (2013) Lessons from the cancer genome. Cell 153: 17-37

Soon W. W., Hariharan M. and Snyder M. P. (2013) High-throughput sequencing for biology and medicine. Mol Syst Biol 9: 640

Nik-Zainal S., Alexandrov L. B., Wedge D. C., Van Loo P., Greenman C. D., et al. (2012) Mutational processes molding the genomes of 21 breast cancers. Cell 149: 979-993

Shendure J. and Lieberman Aiden E. (2012) The expanding scope of DNA sequencing. Nat Biotechnol 30: 1084-1094

Tuna M. and Amos C. I. (2012) Genomic sequencing in cancer. Cancer Lett

Yates L. R. and Campbell P. J. (2012) Evolution of the cancer genome. Nat Rev Genet 13: 795-806

---

[1] Yates L. R. and Campbell P. J. (2012) Evolution of the cancer genome. Nat Rev Genet 13: 795-806

Cancer Biology

Tumor Heterogeneity

Every individual carries a unique set of inherited germline mutations. As cancer progresses, additional somatic mutations and genomic rearrangements accumulate.[2] These changes can trigger drug resistance and metastasis.[3] Increasing evidence suggests that these processes are deliberate, with a finite number of distinct mechanisms.[4] Longitudinal experiments, where samples are collected over the course of the disease, are useful to elucidate the mechanism of disease progression. These samples are commonly used to understand the causes of relapse[5] and drug resistance.[6]
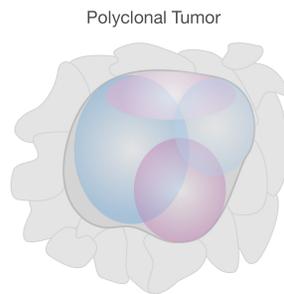
Polyclonal Tumor

A polyclonal tumor in a background of normal tissue. Most tumor samples contain a mixture of tumor and normal cells. The tumor itself may contain several different clonal types, each with a different response to therapy and potential for recurrence.

Tumor samples typically include normal cells, such as stromal cells, blood vessels, and immune cells. Based on conventional pathology estimates, most studies focus on tumors with >60% tumor nuclei present.[7] To determine which mutations are unique to the tumor, a reference normal tissue sample from the same individual is usually included in the analysis.

The tumor itself may be heterogeneous. During cancer progression new mutations may occur in individual cells and these newly mutated cells can go on to proliferate and form clones. As a result late-stage cancers often consist of polyclonal tumors, where each clone has a unique set of mutations, unique pathology, and unique drug responses.[8-9] Deep sequencing has the sensitivity to detect clones comprising as little as 1% of the sample. (See **Technical Considerations** for additional information.)

---

[2] Yates L. R. and Campbell P. J. (2012) Evolution of the cancer genome. Nat Rev Genet 13: 795-806
[3] Caldas C. (2012) Cancer sequencing unravels clonal evolution. Nat Biotechnol 30: 408-410
[4] Ding L., Ley T. J., Larson D. E., Miller C. A., Koboldt D. C., et al. (2012) Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. Nature 481: 506-510
[5] Meyer J. A., Wang J., Hogan L. E., Yang J. J., Dandekar S., et al. (2013) Relapse-specific mutations in NT5C2 in childhood acute lymphoblastic leukemia. Nat Genet 45: 290-294
[6] Tzoneva G., Perez-Garcia A., Carpenter Z., Khiabanian H., Tosello V., et al. (2013) Activating mutations in the NT5C2 nucleotidase gene drive chemotherapy resistance in relapsed ALL. Nat Med 19: 368-371
[7] Mardis E. R. (2012) Genome sequencing and cancer. Curr Opin Genet Dev 22: 245-250
[8] Gerlinger M., Rowan A. J., Horswell S., Larkin J., Endesfelder D., et al. (2012) Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. N Engl J Med 366: 883-892
[9] Landau D. A., Carter S. L., Stojanov P., McKenna A., Stevenson K., et al. (2013) Evolution and impact of subclonal mutations in chronic lymphocytic leukemia. Cell 152: 714-726
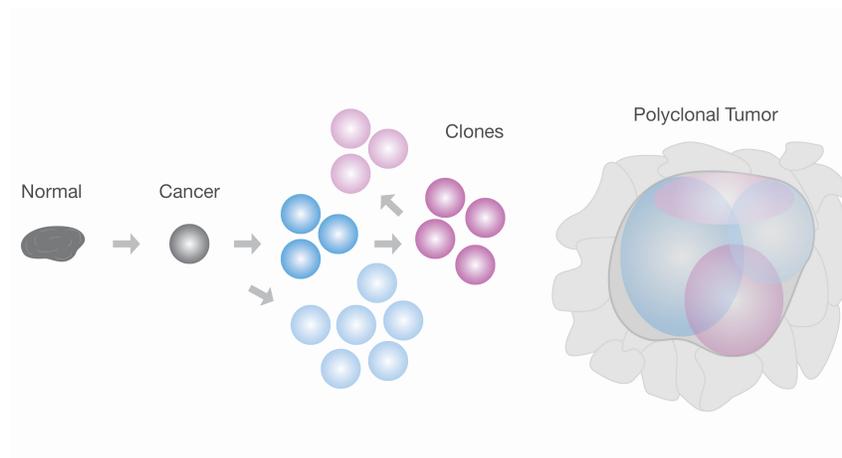
Intratumor heterogeneity. The progressive accumulation of somatic mutations results in a heterogeneous polyclonal tumor in which different clones may respond differently to treatment.

In some genes mutations frequently occur in the same location, which may indicate a specific mechanism at work. However, in the majority of genes mutations can appear apparently randomly throughout the gene, which may reflect the failure of replication and repair mechanisms. Sequencing can detect mutations from both scenarios with equal facility.
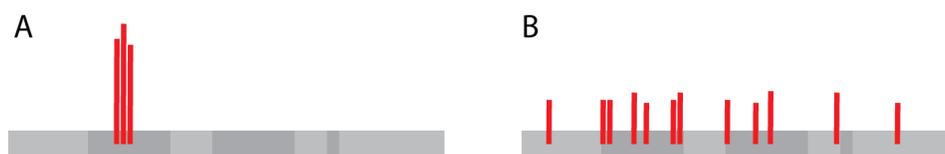
Two hypothetical genes with two different mutation models. The dark boxes indicate exomes and the red bars indicate locations where mutations occur. Panel A: Recurrent mutations in a specific location may indicate the involvement of a biological mechanism to generate the mutations. Panel B: Scattered mutations occurring throughout the gene, such as P53, may be due to the failure of the replication and repair mechanisms. Sequencing can detect mutations generated in both scenarios.

References

Ding L., Ley T. J., Larson D. E., Miller C. A., Koboldt D. C., et al. (2012) Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. Nature 481: 506-510

This study addresses the causes of relapse in acute myeloid leukemia (AML). The authors found two general mechanisms: (1) the founding clone in the primary tumor gained mutations and evolved into the relapse clone; or (2) a subclone of the founding clone survived initial therapy, gained additional mutations, and expanded at relapse. In one case a subclone that made up only 5.1% of the primary tumor became the predominant clone after relapse. In all cases, chemotherapy failed to eradicate the founding clone. This study underscores the importance of detecting and eradicating small cellular populations after diagnosis and also after the initial treatment. The ability of next-generation sequencing to detect *de novo* mutations in very small cell populations makes it uniquely suited to this type of application.

Illumina technology: Genome Analyzer$_{IIx}$ system 100 bp paired-end (PE) reads

Gerlinger M., Rowan A. J., Horswell S., Larkin J., Endesfelder D., et al. (2012) Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. N Engl J Med 366: 883-892

The authors used whole-exome sequencing to investigate multiple samples from spatially separated regions of primary renal carcinomas and associated metastatic sites in two patients. They found extensive heterogeneity within the primary tumor and noted that 63%–69% of all somatic mutations were not detectable across every tumor region. Gene-expression signatures of good and poor prognosis were also detected in different regions of the same tumor. This underscores the importance of early diagnosis before the mutations accumulate, as well as the need for multiple biopsy sites in larger tumors. The use of multiple samples from the same patient allows the authors to reconstruct the progression of the disease. This is a remarkably powerful approach that detected not only the trigger events, but also genes that display parallel evolution. Parallel evolution is usually an indication of genes under evolutionary pressure and it implies that those genes could be effective therapeutic targets.

Illumina technology: Genome Analyzer$_{IIx}$ System and HiSeq® 2000 system

Walter M. J., Shen D., Ding L., Shao J., Koboldt D. C., et al. (2012) Clonal architecture of secondary acute myeloid leukemia. N Engl J Med 366: 1090-1098

Secondary AML develops in approximately one-third of persons with myelodysplastic syndromes. This study is intended to identify mutations in myelodysplastic syndromes that may predict progression to AML. The authors performed whole-genome sequencing of seven paired samples of skin and bone marrow in seven subjects with secondary AML as well as matched bone marrow samples from the antecedent myelodysplastic syndrome. They found that, in all cases, the dominant secondary AML clone was derived from a myelodysplastic syndrome founding clone. This implies that myelodysplastic syndrome samples contain prognostically important mutations. Therapies that target these mutations may improve outcomes.

Illumina Technology: Genome Analyzer$_{IIx}$ System and HiSeq 2000 System with 2 x 75 paired-end reads and 100x coverage

Landau D. A., Carter S. L., Stojanov P., McKenna A., Stevenson K., et al. (2013) Evolution and impact of subclonal mutations in chronic lymphocytic leukemia. Cell 152: 714-726

Streppel M. M., Lata S., Delabastide M., Montgomery E. A., Wang J. S., et al. (2013) Next-generation sequencing of endoscopic biopsies identifies ARID1A as a tumor-suppressor gene in Barrett's esophagus. Oncogene

Shah S. P., Roth A., Goya R., Oloumi A., Ha G., et al. (2012) The clonal and mutational evolution spectrum of primary triple-negative breast cancers. Nature 486: 395-399

Yap T. A., Gerlinger M., Futreal P. A., Pusztai L. and Swanton C. (2012) Intratumor heterogeneity: seeing the wood for the trees. Sci Transl Med 4: 127ps110

Methods

Cibulskis K., Lawrence M. S., Carter S. L., Sivachenko A., Jaffe D., et al. (2013) Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. Nat Biotechnol 31: 213-219

Saunders C. J., Miller N. A., Soden S. E., Dinwiddie D. L., Noll A., et al. (2012) Rapid whole-genome sequencing for genetic disease diagnosis in neonatal intensive care units. Sci Transl Med 4: 154ra135

Metastasis

Metastasis is a complex process in which cancer cells break away from the primary tumor and circulate through the bloodstream or lymphatic system to other sites in the body. At new sites, the cells continue to multiply and eventually form additional tumors comprised of cells that reflect the tissue of origin. The ability of tumors, such as pancreatic cancer and uveal cancers, to metastasize contributes greatly to their lethality. Many fundamental questions remain about the clonal structures of metastatic tumors, phylogenetic relationships among metastases, the scale of ongoing parallel evolution in metastatic and primary sites, how the tumor disseminates, and the role that the tumor microenvironment plays in the determination of the metastatic site.

Review

Caldas C. (2012) Cancer sequencing unravels clonal evolution. Nat Biotechnol 30: 408-410
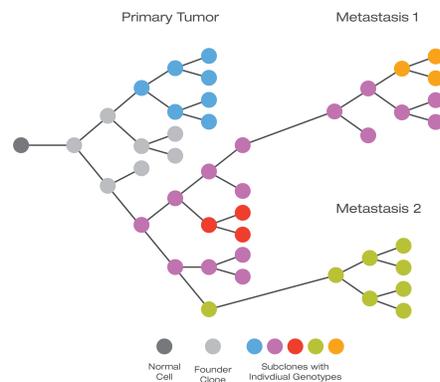
Metastases can originate from either a major clone in the primary tumor (metastasis 1), or from minor clones (metastasis 2). Metastases can also undergo clonal evolution (as shown in metastasis 1).

References

Hsieh A. C., Liu Y., Edlind M. P., Ingolia N. T., Janes M. R., et al. (2012) The translational landscape of mTOR signalling steers cancer initiation and metastasis. Nature 485: 55-61

The authors demonstrated a specialized translation of the prostate cancer genome by oncogenic mTOR signaling, which resulted in a remarkably specific repertoire of genes involved in cell proliferation, metabolism, and invasion. They then functionally characterized a class of translationally controlled proinvasion messenger RNAs that orchestrate prostate cancer invasion and metastasis.

Illumina Technology: Genome Analyzer$_{IIx}$ system for mRNA-Seq and Ribo-Seq

Chen D., Sun Y., Wei Y., Zhang P., Rezaeian A. H., et al. (2012) LIFR is a breast cancer metastasis suppressor upstream of the Hippo-YAP pathway and a prognostic marker. Nat Med 18: 1511-1517

Gerlinger M., Rowan A. J., Horswell S., Larkin J., Endesfelder D., et al. (2012) Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. N Engl J Med 366: 883-892

Turajlic S., Furney S. J., Lambros M. B., Mitsopoulos C., Kozarewa I., et al. (2012) Whole genome sequencing of matched primary and metastatic acral melanomas. Genome Res 22: 196-207

Genomic Mutations

All tumors accumulate somatic mutations during their development. Most common cancers are associated with diverse cancer genes that are mutated at a low frequency. One of the most striking observations from large cancer databases is the genetic heterogeneity among cancers and even within individual cancer types. However, it appears that a limited number of cellular pathways are central to tumor cell biology.[10-11] Comprehensive catalogs of somatic mutations are being compiled for various cancer types to better understand the mechanisms that underlie this disease.[12]

References

Nik-Zainal S., Alexandrov L. B., Wedge D. C., Van Loo P., Greenman C. D., et al. (2012) Mutational processes molding the genomes of 21 breast cancers. Cell 149: 979-993

The authors generated catalogs of somatic mutation from 21 breast cancers. Cancers with BRCA1 or BRCA2 mutations exhibited a characteristic combination of substitution mutation signatures and a distinctive profile of deletions. They also described a localized hypermutation phenomenon, termed "kataegis." Base substitutions in these regions were almost exclusively of cytosine at TpC dinucleotides.

Illumina Technology: Genome Analyzer$_{IIx}$ system or HiSeq 2000 system with no-PCR protocol
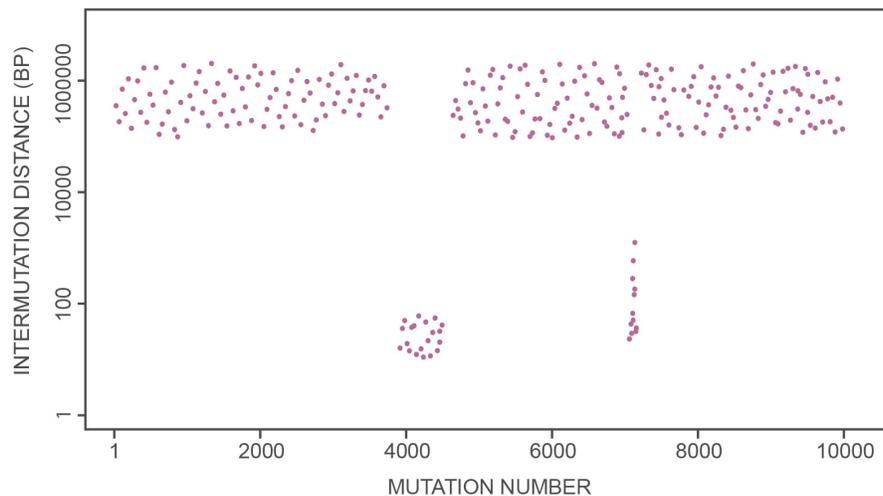
A Kataegis (rainfall) plot. The intermutation distance is plotted on the vertical axis on a log scale. Most mutations in this hypothetical genome have an intermutation distance of ~10^5 bp to ~10^6 bp. Mutations in a region of hypermutation present as a cluster of lower intermutation distances. The Kataegis plot is very useful to illustrate clusters of mutations. Nik-Zainal S., Alexandrov L. B., Wedge D. C., Van Loo P., Greenman C. D., et al. (2012) Mutational processes molding the genomes of 21 breast cancers. Cell 149: 979-993

---

[10] Ellis M. J., Ding L., Shen D., Luo J., Suman V. J., et al. (2012) Whole-genome analysis informs breast cancer response to aromatase inhibition. Nature 486: 353-360
[11] Ellis M. J. and Perou C. M. (2013) The genomic landscape of breast cancer as a therapeutic roadmap. Cancer Discov 3: 27-34
[12] Pleasance E. D., Cheetham R. K., Stephens P. J., McBride D. J., Humphray S. J., et al. (2010) A comprehensive catalogue of somatic mutations from a human cancer genome. Nature 463: 191-196

Govindan R., Ding L., Griffith M., Subramanian J., Dees N. D., et al. (2012) Genomic landscape of non-small cell lung cancer in smokers and never-smokers. Cell 150: 1121-1134

This paper reports the whole-genome and transcriptome sequencing of tumor and adjacent normal tissue samples from 17 patients with non-small cell lung carcinoma (NSCLC). The observed mutation frequency was 10-fold higher in smokers than in never-smokers. Deep sequencing revealed diverse clonal patterns in both of these populations. All validated EFGR and KRAS mutations were present in the founder clones, suggesting possible roles in cancer initiation. Of the perturbed genes, 54 are potentially targetable with currently available drugs.

Illumina Technology: Genome Analyzer$_{IIx}$ system 100 bp paired-end reads, RNA-Seq library, Human OmniExpress BeadChip

Palles C., Cazier J. B., Howarth K. M., Domingo E., Jones A. M., et al. (2013) Germline mutations affecting the proofreading domains of POLE and POLD1 predispose to colorectal adenomas and carcinomas. Nat Genet 45: 136-144

Huang S., Holzel M., Knijnenburg T., Schlicker A., Roepman P., et al. (2012) MED12 controls the response to multiple cancer drugs through regulation of TGF-beta receptor signaling. Cell 151: 937-950

Park C., Han S., Lee K. M., Choi J. Y., Song N., et al. (2012) Association between CASP7 and CASP14 genetic polymorphisms and the risk of childhood leukemia. Hum Immunol 73: 736-739

Schulz E., Valentin A., Ulz P., Beham-Schmid C., Lind K., et al. (2012) Germline mutations in the DNA damage response genes BRCA1, BRCA2, BARD1 and TP53 in patients with therapy related myeloid neoplasms. J Med Genet 49: 422-428

Thompson-Wicking K., Francis R. W., Stirnweiss A., Ferrari E., Welch M. D., et al. (2012) Novel BRD4-NUT fusion isoforms increase the pathogenic complexity in NUT midline carcinoma. Oncogene

Mosaicism

Most of the mutations found in AML genomes are actually random events that occurred in hematopoietic stem/progenitor cells (HSPCs) before they acquired the initiating mutation; the mutational history of that cell is ''captured'' as the clone expands. In many cases, only one or two additional, cooperating mutations are needed to generate the malignant founding clone.[13]
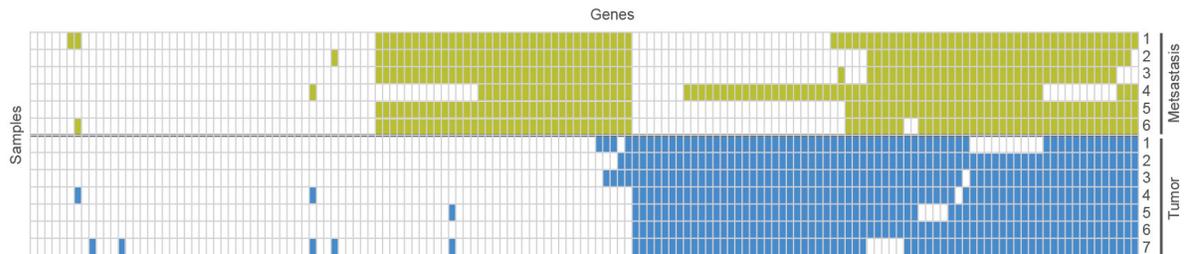
Mosaicism of primary tumors and metastasis. A hypothetical heat map of the regional distribution of nonsynonymous point mutations and indels (green blocks). The rows represent samples from seven primary-tumor regions and six metastasis regions. Gerlinger M., Rowan A. J., Horswell S., Larkin J., Endesfelder D., et al. (2012) Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. N Engl J Med 366: 883-892

References

Abyzov A., Mariani J., Palejev D., Zhang Y., Haney M. S., et al. (2012) Somatic copy number mosaicism in human skin revealed by induced pluripotent stem cells. Nature 492: 438-442

The authors found that, on average, an iPSC line manifests two copy number variants (CNVs) not apparent in the fibroblasts from which the iPSC was derived. They showed that at least 50% of those CNVs are present as low-frequency somatic genomic variants in parental fibroblasts. Based on this observation they estimated that approximately 30% of the fibroblast cells have somatic CNVs in their genomes, suggesting widespread somatic mosaicism in the human body.

Illumina Technology: HiSeq 2000 system whole-genome paired-end sequencing and HumanHT-12 v4 BeadChip

---

[13] Welch J. S., Ley T. J., Link D. C., Miller C. A., Larson D. E., et al. (2012) The origin and evolution of mutations in acute myeloid leukemia. Cell 150: 264-278

Ruark E., Snape K., Humburg P., Loveday C., Bajrami I., et al. (2013) Mosaic PPM1D mutations are associated with predisposition to breast and ovarian cancer. Nature 493: 406-410

Forsberg L. A., Rasi C., Razzaghian H. R., Pakalapati G., Waite L., et al. (2012) Age-related somatic structural changes in the nuclear genome of human blood cells. Am J Hum Genet 90: 217-228

Jasmine F., Rahaman R., Dodsworth C., Roy S., Paul R., et al. (2012) A genome-wide study of cytogenetic changes in colorectal cancer using SNP microarrays: opportunities for future personalized treatment. PLoS ONE 7: e31968

Perez-Mancera P. A., Rust A. G., van der Weyden L., Kristiansen G., Li A., et al. (2012) The deubiquitinase USP9X suppresses pancreatic ductal adenocarcinoma. Nature 486: 266-270

Turajlic S., Furney S. J., Lambros M. B., Mitsopoulos C., Kozarewa I., et al. (2012) Whole genome sequencing of matched primary and metastatic acral melanomas. Genome Res 22: 196-207

Welch J. S., Ley T. J., Link D. C., Miller C. A., Larson D. E., et al. (2012) The origin and evolution of mutations in acute myeloid leukemia. Cell 150: 264-278

Gene Fusions

Gene fusions are widespread and are the hallmarks of some cancer types.[14-15] The combination of a strong promoter with a functional gene (proto-oncogene) downstream is common in some cancers. It is estimated that half of prostate cancers harbor gene fusions between *TMPRSS2* and members of the *ETS* transcription factor family.[14] Gene fusions are formed by the joining of two previously separate genes or loci and may lead to a gene product with a new or different function from the two fusion partners. It may also result in an oncogenic activation, as in the case of Philadelphia chromosome positive-acute lymphoblastic leukemia. This gene fusion results in expression of the *BCR-ABL* tyrosine kinase, which activates cellular proliferation.[15] Gene fusions can be generated by several mechanisms, which can be characteristic for some cancer types. Pancreatic cancer is characterized by frequent breakage–fusion–bridge cycles of chromosomal rearrangement.[16] There are several approaches to studying fusion events through sequencing, such as whole-genome sequencing of the tumor and mRNA-Seq. The combination of RNA-Seq with whole-genome sequencing is particularly effective. mRNA-Seq provides an additional layer of evidence to support the observation of the fusion event and provides evidence if the fused gene is expressed.[17]

Review

Wang Q., Xia J., Jia P., Pao W. and Zhao Z. (2012) Application of next generation sequencing to human gene fusion detection: computational tools, features and perspectives. Brief Bioinform

Chr 19   Fragment   Chr 19
                of chr 6

Fusion events caused by fold-back inversions can capture fragments of distant regions of the genome, such as centromeric repeats or regions involved in somatic rearrangements. In this example, a fragment of chromosome 6 is captured between a duplicated part of chromosome 19. Note that the second copy of chromosome 19 is inverted. This is characteristic of a fold-back inversion. Campbell P. J., Yachida S., Mudie L. J., Stephens P. J., Pleasance E. D., et al. (2010) The patterns and dynamics of genomic instability in metastatic pancreatic cancer. Nature 467: 1109-1113

---

[14] Pflueger D., Terry S., Sboner A., Habegger L., Esgueva R., et al. (2011) Discovery of non-ETS gene fusions in human prostate cancer using next-generation RNA sequencing. Genome Res 21: 56-67
[15] Roberts K. G., Morin R. D., Zhang J., Hirst M., Zhao Y., et al. (2012) Genetic alterations activating kinase and cytokine receptor signaling in high-risk acute lymphoblastic leukemia. Cancer Cell 22: 153-166
[16] Campbell P. J., Yachida S., Mudie L. J., Stephens P. J., Pleasance E. D., et al. (2010) The patterns and dynamics of genomic instability in metastatic pancreatic cancer. Nature 467: 1109-1113
[17] Kangaspeska S., Hultsch S., Edgren H., Nicorici D., Murumagi A., et al. (2012) Reanalysis of RNA-sequencing data reveals several additional fusion genes with multiple isoforms. PLoS ONE 7: e48745
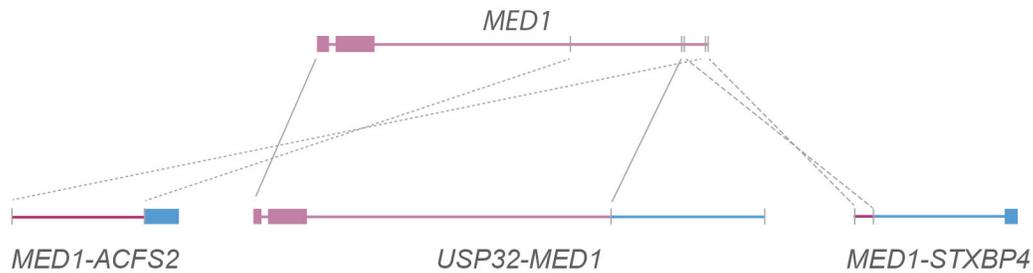
MED1 (red) forms fusions with several partner genes (blue): ACSF2, USP32 and STXBP4. Kangaspeska S., Hultsch S., Edgren H., Nicorici D., Murumagi A., et al. (2012) Reanalysis of RNA-sequencing data reveals several additional fusion genes with multiple isoforms. PLoS ONE 7: e48745

Experimental Considerations

Whole-genome sequencing with paired-end reads is the most accurate and comprehensive tool available to detect all gene fusions, including duplications, inversions, read-throughs, and single-base indels. In particular, the use of paired-end sequencing is a key factor in the success of fusion gene detection.

Whole-genome sequencing excels in the *de novo* discovery of fusion break points. Deep sequencing, in conjunction with longer reads, enables base-pair resolution of the microhomology at the fusion junctions. This ability, unique to sequencing, provides a footprint of the mechanisms involved in generating the fusions.

mRNA-Seq is a highly efficient, cost-effective approach to detect fused genes in large numbers of samples. This approach is based on the hypothesis that highly expressed fusion genes will have the greatest biological impact. mRNA-Seq is particularly effective for detecting highly expressed oncogenes. However, it is limited to expressed genes with poly(A) tails and will not capture information on intergenic regions and UTRs.

References

Robinson D. R., Wu Y. M., Kalyana-Sundaram S., Cao X., Lonigro R. J., et al. (2013) Identification of recurrent NAB2-STAT6 gene fusions in solitary fibrous tumor by integrative sequencing. Nat Genet 45: 180–185

The authors used whole-exome and transcriptome sequencing to find a gene fusion of the transcriptional repressor NAB2 with the transcriptional activator STAT6. Transcriptome sequencing of 27 additional solitary fibrous tumors (SFTs) identified the presence of a NAB2-STAT6 gene fusion in all tumors. Overexpression of the NAB2-STAT6 gene fusion induced proliferation in cultured cells and activated the expression of EGR-responsive genes.

Illumina Technology: HiSeq 2000 system to 100 bp paired-end reads. TruSeq DNA Sample Prep kit and TruSeq RNA protocol

Seshagiri S., Stawiski E. W., Durinck S., Modrusan Z., Storm E. E., et al. (2012) Recurrent R-spondin fusions in colon cancer. Nature 488: 660-664

The authors analyzed the exomes, transcriptomes, and copy-number alterations of 70 primary human colon tumors. Copy-number and RNA-Seq data analysis identified amplifications and corresponding overexpression of IGF2 in a subset of colon tumors. They also used RNA-Seq to find gene fusions involving R-spondin family members (RSPO2 and RSPO3) in 10% of the colon tumors. This study demonstrates the importance of combining multiple technologies to understand the complexity of the cancer genome.

Illumina Technologies: HiSeq 2000 system exome sequencing with 75 bp paired-end reads, TruSeq® RNA Sample Preparation kit with 75 bp paired-end reads, Human Omni 2.5 arrays

---

Thompson-Wicking K., Francis R. W., Stirnweiss A., Ferrari E., Welch M. D., et al. (2012) Novel BRD4-NUT fusion isoforms increase the pathogenic complexity in NUT midline carcinoma. Oncogene

This paper shows that a novel BRD4–NUT fusion in PER-624 encodes a functional protein that is central to the oncogenic mechanism in these cells. The generation of BRD4–NUT fusion transcripts through post-translocation RNA splicing appears to be a common feature of these carcinomas. It is the first report of this mechanism, which facilitates the expression of alternative isoforms of the fusion genes.

Illumina Technology: Genome Analyzer$_{II}$ system for RNA-Seq

---

Wen H., Li Y., Malek S. N., Kim Y. C., Xu J., et al. (2012) New fusion transcripts identified in normal karyotype acute myeloid leukemia. PLoS ONE 7: e51203

In this study the authors used paired-end RNA-Seq to find fusions in karyotypes that had no aberrations detectable by conventional cytogenetic analysis. They found fusion transcripts between adjacent genes and seven fusions that were exclusively present in normal karyotypes.

Illumina Technology: Genome Analyzer$_{IIx}$ system 50 bp paired-end RNA-Seq

Robinson D. R., Wu Y. M., Kalyana-Sundaram S., Cao X., Lonigro R. J., et al. (2013) Identification of recurrent NAB2-STAT6 gene fusions in solitary fibrous tumor by integrative sequencing. Nat Genet 45: 180-185

Banerji S., Cibulskis K., Rangel-Escareno C., Brown K. K., Carter S. L., et al. (2012) Sequence analysis of mutations and translocations across breast cancer subtypes. Nature 486: 405-409

Ju Y. S., Lee W. C., Shin J. Y., Lee S., Bleazard T., et al. (2012) A transforming KIF5B and RET gene fusion in lung adenocarcinoma revealed from whole-genome and transcriptome sequencing. Genome Res 22: 436-445

Kangaspeska S., Hultsch S., Edgren H., Nicorici D., Murumagi A., et al. (2012) Reanalysis of RNA-sequencing data reveals several additional fusion genes with multiple isoforms. PLoS ONE 7: e48745

Roberts K. G., Morin R. D., Zhang J., Hirst M., Zhao Y., et al. (2012) Genetic alterations activating kinase and cytokine receptor signaling in high-risk acute lymphoblastic leukemia. Cancer Cell 22: 153-166

Rudin C. M., Durinck S., Stawiski E. W., Poirier J. T., Modrusan Z., et al. (2012) Comprehensive genomic analysis identifies SOX2 as a frequently amplified gene in small-cell lung cancer. Nat Genet 44: 1111-1116

Wen H., Li Y., Malek S. N., Kim Y. C., Xu J., et al. (2012) New fusion transcripts identified in normal karyotype acute myeloid leukemia. PLoS ONE 7: e51203

Chromothripsis

Chromothripsis is a one-off cellular crisis during which tens to hundreds of genomic rearrangements occur in a single event.[18-19] The consequences of this catastrophic event are complex local rearrangements and copy number variants where a limited range of two (or occasionally three) copy-number states are detectable along the chromosome.[20] This model of a single catastrophic event is different from the typical model of cancer progression through the progressive accumulation of mutations. In a cancer progression model where mutations accumulate, there is no upper limit to the copy numbers, so it is common to see a wide range. It is estimated that chromothripsis occurs in 2%–3% of all cancers, across many subtypes, and in ~25% of bone cancers.[21]
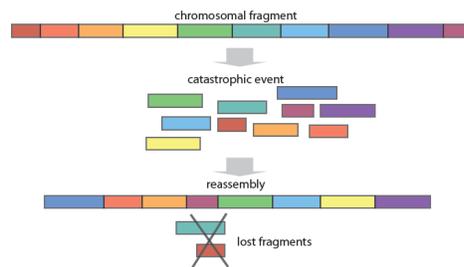
A pictorial representation of chromothripsis. Tubio JM (2011) Cancer: When catastrophe strikes a cell. Nature 470:476-7.

References

Rausch T., Jones D. T., Zapatka M., Stutz A. M., Zichner T., et al. (2012) Genome Sequencing of Pediatric Medulloblastoma Links Catastrophic DNA Rearrangements with TP53 Mutations. Cell 148: 59-71

The authors report massive, complex chromosome rearrangements in a Sonic-Hedgehog medulloblastoma (SHH-MB) brain tumor from a patient with a germline TP53 mutation (Li-Fraumeni syndrome). In a larger screen of 11 Li-Fraumeni syndrome patients, 36% of the tumors showed rearrangements consistent with chromothripsis. This is much higher than the 2% incidence observed in general tumor populations. A germline mutation in P53 is consistent with the hypothesis of aborted apoptosis as the cause for chromothripsis in some tumors.

Illumina Technology: Genome Analyzer$_{IIx}$ system and HiSeq 2000 system with paired-end and mate-pair protocols

Chiang C., Jacobsen J. C., Ernst C., Hanscom C., Heilbut A., et al. (2012) Complex reorganization and predominant non-homologous repair following chromosomal breakage in karyotypically balanced germline rearrangements and transgenic integration. Nat Genet 44: 390-397, S391

Korbel J. O. and Campbell P. J. (2013) Criteria for inference of chromothripsis in cancer genomes. Cell 152: 1226-1236

---

[18] Tubio J. M. and Estivill X. (2011) Cancer: When catastrophe strikes a cell. Nature 470: 476-477
[19] Forment J. V., Kaidi A. and Jackson S. P. (2012) Chromothripsis and cancer: causes and consequences of chromosome shattering. Nat Rev Cancer 12: 663-670
[20] Stephens P. J., Greenman C. D., Fu B., Yang F., Bignell G. R., et al. (2011) Massive genomic rearrangement acquired in a single catastrophic event during cancer development. Cell 144: 27-40
[21] Chiang C., Jacobsen J. C., Ernst C., Hanscom C., Heilbut A., et al. (2012) Complex reorganization and predominant non-homologous repair following chromosomal breakage in karyotypically balanced germline rearrangements and transgenic integration. Nat Genet 44: 390-397, S391

Copy Number Variants (CNV)

Structural variants impact gene dosage, the number of functional copies of the gene that are available for transcription. Tumor progression, drug response, and the onset of drug resistance are commonly driven by underlying gene amplifications and deletions. These genomic alterations can be divided into large aberrations and small aberrations. Large aberrations include the loss or duplication of whole or partial chromosomes, also called aneuploidy. Small alterations can span as little as one base, as in the case of point mutations and indels. Unlike the healthy genome where changes in gene expression are carefully controlled through transcription factors, the cancer genome adapts through the duplication and deletion of genes. The development of drug resistance is an excellent demonstration of the speed and efficiency of this response.

Both genotyping arrays and high-throughput sequencing successfully detect CNVs. However, next-generation sequencing can reveal additional small regions of CNVs and novel structural variants.[22]

References

Holmfeldt L., Wei L., Diaz-Flores E., Walsh M., Zhang J., et al. (2013) The genomic landscape of hypodiploid acute lymphoblastic leukemia. Nat Genet 45: 242-252

Sarhadi V. K., Lahti L., Scheinin I., Tyybakinoja A., Savola S., et al. (2013) Targeted resequencing of 9p in acute lymphoblastic leukemia yields concordant results with array CGH and reveals novel genomic alterations. Genomics

Yang D., Sun Y., Hu L., Zheng H., Ji P., et al. (2013) Integrated analyses identify a master microRNA regulatory network for the mesenchymal subtype in serous ovarian cancer. Cancer Cell 23: 186-199

Ha G., Roth A., Lai D., Bashashati A., Ding J., et al. (2012) Integrative analysis of genome-wide loss of heterozygosity and monoallelic expression at nucleotide resolution reveals disrupted pathways in triple-negative breast cancer. Genome Res 22: 1995-2007

---

[22] Sarhadi V. K., Lahti L., Scheinin I., Tyybakinoja A., Savola S., et al. (2013) Targeted resequencing of 9p in acute lymphoblastic leukemia yields concordant results with array CGH and reveals novel genomic alterations. Genomics

Gene Expression

Gene expression analysis measures the product of gene transcription, RNA processing, and epigenetic control. As a result, gene expression analysis provides an overview of the health of these processes as well as insight into molecular functions within the cell. Microarray-based mRNA analysis has been used extensively to study gene expression in cancer research, but the advent of sequencing-based mRNA analysis (mRNA-Seq) represents a quantum leap forward in the ability to measure and interpret the products of gene expression.[23] The ability of mRNA-Seq to detect modified RNAs[24] and RNAs expressed at very low levels makes it uniquely suited to cancer research. Methods based on mRNA-Seq can also detect very rapid changes in transcription,[25] splice variants, fusion genes, and the use of alternative polyadenylation sites.[26]

Application	Notes	Reference
Detect driver mutations	Expressed genes are more likely to play a role in tumor progression than genes that are not expressed. For some genes, such as TP53 and RB1, loss of function can act as a driver.	[27]
Confirm fusion genes	The mRNAs from fusion genes are relatively easy to detect. It can also confirm the expression of fused genes.	[28, 29]
Impact of CNVs	Changes in gene copy numbers can have a significant impact on expression levels.	[28]
Confirm gene mutations	mRNA sequencing can confirm the presence of a gene mutation.	

References

Feng H., Qin Z. and Zhang X. (2012) Opportunities and methods for studying alternative splicing in cancer with RNA-Seq. Cancer Lett

This review focuses on the use of RNA-Seq to study cancer-related alternative splicing. It includes a table of bioinformatics tools and an extensive discussion on estimating the expression levels of alternative splicing isoforms.

[23] McGettigan P. A. (2013) Transcriptomics in the RNA-seq era. Curr Opin Chem Biol 17: 4-11
[24] Jiang Q., Crews L. A., Barrett C. L., Chun H. J., Court A. C., et al. (2013) ADAR1 promotes malignant progenitor reprogramming in chronic myeloid leukemia. Proc Natl Acad Sci U S A 110: 1041-1046
[25] Ingolia N. T., Ghaemmaghami S., Newman J. R. and Weissman J. S. (2009) Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. Science 324: 218-223
[26] Shepard P. J., Choi E. A., Lu J., Flanagan L. A., Hertel K. J., et al. (2011) Complex and dynamic landscape of RNA polyadenylation revealed by PAS-Seq. RNA 17: 761-772
[27] Peifer M., Fernandez-Cuesta L., Sos M. L., George J., Seidel D., et al. (2012) Integrative genome analyses identify key somatic driver mutations of small-cell lung cancer. Nat Genet 44: 1104-1110
[28] Krauthammer M., Kong Y., Ha B. H., Evans P., Bacchiocchi A., et al. (2012) Exome sequencing identifies recurrent somatic RAC1 mutations in melanoma. Nat Genet 44: 1006-1014
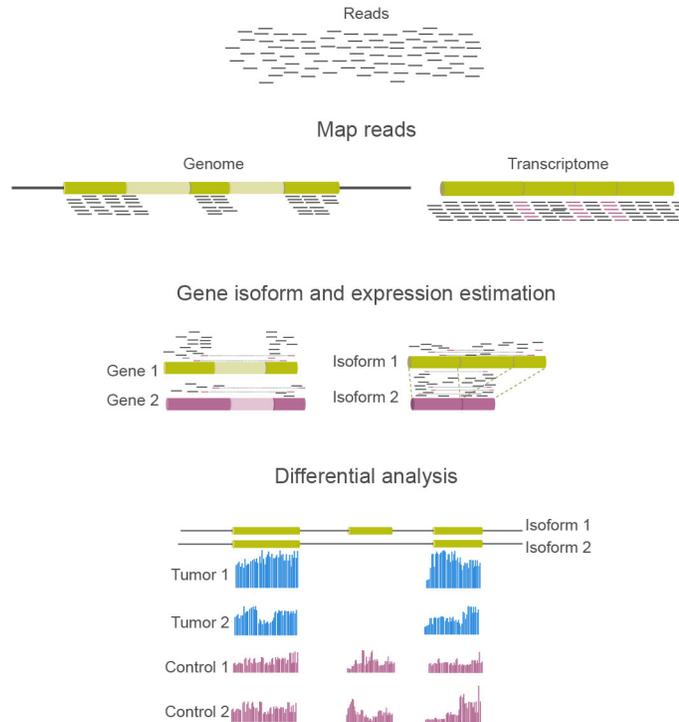
A typical bioinformatics pipeline for studying gene expression and alternative splicing in cancers using RNA-Seq. Short reads are first mapped to a reference genome or transcriptome. After mapping, the expression and splicing of the annotated genes and transcripts can be estimated. Feng H., Qin Z. and Zhang X. (2012) Opportunities and methods for studying alternative splicing in cancer with RNA-Seq. Cancer Lett in press

van Delft J., Gaj S., Lienhard M., Albrecht M. W., Kirpiy A., et al. (2012) RNA-Seq provides new insights in the transcriptome responses induced by the carcinogen benzo[a]pyrene. Toxicol Sci 130: 427-439

The authors showed that RNA-Seq detects approximately 20% more genes than microarray-based technology but almost threefold more significantly differentially expressed genes. As a result, they detected two- to five-fold more affected pathways and biological processes. The authors also detected alternative isoform expression in many genes, including regulators of cell death and DNA repair, such as TP53, BCL2, and XPA, which are relevant for genotoxic responses. They also found potentially novel isoforms with unknown function, such as fragments of known transcripts, transcripts with additional exons, intron retention, or exon-skipping events.

Illumina Technology: Genome Analyzer™ System with 51 bp paired-end reads.

Kaur H., Mao S., Li Q., Sameni M., Krawetz S. A., et al. (2012) RNA-Seq of human breast ductal carcinoma in situ models reveals aldehyde dehydrogenase isoform 5A1 as a novel potential target. PLoS ONE 7: e50249

The authors determined the expression changes that are common among three DCIS models (MCF10.DCIS, SUM102 and SUM225) compared to the MCF10A model of non-tumorigenic mammary epithelial cells in three-dimensional (3D) overlay culture. They found differentially expressed genes encoding for proteins that are associated with a number of signaling pathways.

Illumina Technology: Genome Analyzer$_{IIx}$ System for 76 cycles of single-end sequencing

---

Meyer J. A., Wang J., Hogan L. E., Yang J. J., Dandekar S., et al. (2013) Relapse-specific mutations in NT5C2 in childhood acute lymphoblastic leukemia. Nat Genet 45: 290-294

The authors report the transcriptome profiles of matched diagnosis and relapse bone marrow specimens from ten individuals with pediatric B-lymphoblastic leukemia using RNA sequencing. Transcriptome sequencing identified 20 newly acquired, novel mutations not present at initial diagnosis, with 2 individuals harboring relapse-specific mutations. All individuals who harbored NT5C2 mutations relapsed within 36 months of initial diagnosis.

Illumina Technology: Genome Analyzer$_{IIx}$ system RNA-Seq with 54 bp reads

Feng H., Qin Z. and Zhang X. (2012) Opportunities and methods for studying alternative splicing in cancer with RNA-Seq. Cancer Lett

Huang S., Holzel M., Knijnenburg T., Schlicker A., Roepman P., et al. (2012) MED12 controls the response to multiple cancer drugs through regulation of TGF-beta receptor signaling. Cell 151: 937-950

Ju Y. S., Lee W. C., Shin J. Y., Lee S., Bleazard T., et al. (2012) A transforming KIF5B and RET gene fusion in lung adenocarcinoma revealed from whole-genome and transcriptome sequencing. Genome Res 22: 436-445

Kaur H., Mao S., Li Q., Sameni M., Krawetz S. A., et al. (2012) RNA-Seq of human breast ductal carcinoma in situ models reveals aldehyde dehydrogenase isoform 5A1 as a novel potential target. PLoS ONE 7: e50249

Liang W. S., Craig D. W., Carpten J., Borad M. J., Demeure M. J., et al. (2012) Genome-wide characterization of pancreatic adenocarcinoma patients using next generation sequencing. PLoS ONE 7: e43192

Nyquist K. B., Panagopoulos I., Thorsen J., Haugom L., Gorunova L., et al. (2012) Whole-transcriptome sequencing identifies novel IRF2BP2-CDX1 fusion gene brought about by translocation t(1;5)(q42;q32) in mesenchymal chondrosarcoma. PLoS ONE 7: e49705

Sorrells S., Carbonneau S., Harrington E., Chen A. T., Hast B., et al. (2012) Ccdc94 protects cells from ionizing radiation by inhibiting the expression of p53. PLoS Genet 8: e1002922

Thompson-Wicking K., Francis R. W., Stirnweiss A., Ferrari E., Welch M. D., et al. (2012) Novel BRD4-NUT fusion isoforms increase the pathogenic complexity in NUT midline carcinoma. Oncogene

Notes on experimental design

RNA-Seq had become a routine application in investigating molecular changes in the tumor and most investigators use the manufacturer's protocols. The use of rRNA depletion may improve signal-to-noise ratios and allow the detection of low-expressing transcripts.

Somatic mutations in cancer are essentially *de novo*. Sequencing does not require prior knowledge of the mutations and can map mutations accurately along with the transcript abundance.

Tumors often contain mixtures of cells. The extended dynamic range and accuracy of mRNA-Seq are invaluable to detect small changes in expression. If a tumor transcript contains a unique somatic mutation or splice variant, it can be distinguished from normal cells.

The sensitivity for detecting gene fusions by paired-end, next-generation sequencing depends upon many factors including expression level, transcript length, the sample preparation method used, and cDNA library fragment length.

Most protocols use poly(A)-enriched RNA preparations to measure mRNA levels. However noncoding RNAs, such as miRNA, play an important role in the biology of the cell and often mediate processes critical to tumor growth and survival.[29] Noncoding RNA can easily be analyzed by next-generation sequencing with current poly(A)⁻ (rRNA-depleted) protocols.

RNA expression is tissue- and cell-type specific. This should be considered when controls are selected in tumor-normal controls.

---

[29] Hermeking H. (2012) MicroRNAs in the p53 network: micromanagement of tumour suppression. Nat Rev Cancer 12: 613-626

Alternative Splicing

The biogenesis, development and metastases of cancer are associated with many variations in the transcriptome. Cancer-specific alternative splicing is a widespread phenomenon and a major post-transcriptional regulation mechanism that is involved in many types of cancer.[30]

Review

Feng H., Qin Z. and Zhang X. (2012) Opportunities and methods for studying alternative splicing in cancer with RNA-Seq. Cancer Lett

Frenkel-Morgenstern M., Lacroix V., Ezkurdia I., Levin Y., Gabashvili A., et al. (2012) Chimeras taking shape: potential functions of proteins encoded by chimeric RNA transcripts. Genome Res 22: 1231-1242

Merkin J., Russell C., Chen P. and Burge C. B. (2012) Evolutionary dynamics of gene and isoform regulation in Mammalian tissues. Science 338: 1593-1599

Seo J. S., Ju Y. S., Lee W. C., Shin J. Y., Lee J. K., et al. (2012) The transcriptional landscape and mutational profile of lung adenocarcinoma. Genome Res 22: 2109-2119

The authors analyzed 200 Korean lung adenocarcinomas. They found novel driver mutations in LMTK2, ARID1A, NOTCH2, and SMARCA4. They also found 45 fusion genes, 8 of which were chimeric tyrosine kinases. Among 17 recurrent alternative splicing events, exon 14 skipping in the proto-oncogene MET may be a cancer driver. This study demonstrates the complexity of this cancer and the value of bringing several technologies to bear.

Illumina Technology: HiSeq 2000 system to 100-bp paired-end reads for exon sequencing. RNA-Seq

Liu J., Lee W., Jiang Z., Chen Z., Jhunjhunwala S., et al. (2012) Genome and transcriptome sequencing of lung cancers reveal diverse mutational and splicing events. Genome Res 22: 2315-2327

The authors performed whole-genome sequencing and transcriptome sequencing on 19 lung cancer cell lines and 3 lung tumor/normal pairs. They identified 106 splice-site mutations associated with cancer-specific aberrant splicing, including mutations in several known cancer-related genes. RAC1b, an isoform of the RAC1 GTPase that includes one additional exon, was found to be preferentially upregulated in lung cancer and showed sensitivity to a MAP2K (MEK) inhibitor PD-0325901.

Illumina Technology: Genome Analyzer$_{II}$ System RNA-Seq with Ribo-minus karyotype sample prep (Invitrogen) to 75 bp and Illumina HumanOmni 2.5 arrays

---

[30] Feng H., Qin Z. and Zhang X. (2012) Opportunities and methods for studying alternative splicing in cancer with RNA-Seq. Cancer Lett

Thompson-Wicking K., Francis R. W., Stirnweiss A., Ferrari E., Welch M. D., et al. (2012) Novel BRD4-NUT fusion isoforms increase the pathogenic complexity in NUT midline carcinoma. Oncogene

This paper shows that a novel BRD4–NUT fusion in PER-624 encodes a functional protein that is central to the oncogenic mechanism in these cells. The generation of BRD4–NUT fusion transcripts through post-translocation RNA splicing appears to be a common feature of these carcinomas. This observation, and the mechanism facilitating the expression of alternative isoforms of the fusion, has not been appreciated previously.

Illumina Technology: Genome Analyzer$_{II}$ System for RNA-Seq

Eswaran J., Cyanam D., Mudvari P., Reddy S. D., Pakala S. B., et al. (2012) Transcriptomic landscape of breast cancers through mRNA sequencing. Sci Rep 2: 264

Frenkel-Morgenstern M., Lacroix V., Ezkurdia I., Levin Y., Gabashvili A., et al. (2012) Chimeras taking shape: potential functions of proteins encoded by chimeric RNA transcripts. Genome Res 22: 1231-1242

Ha G., Roth A., Lai D., Bashashati A., Ding J., et al. (2012) Integrative analysis of genome-wide loss of heterozygosity and monoallelic expression at nucleotide resolution reveals disrupted pathways in triple-negative breast cancer. Genome Res 22: 1995-2007

Ren S., Peng Z., Mao J.-H., Yu Y., Yin C., et al. (2012) RNA-seq analysis of prostate cancer in the Chinese population identifies recurrent gene fusions, cancer-associated long noncoding RNAs and aberrant alternative splicings. Cell Res

Sajnani M. R., Patel A. K., Bhatt V. D., Tripathi A. K., Ahir V. B., et al. (2012) Identification of novel transcripts deregulated in buccal cancer by RNA-seq. Gene 507: 152-158

RNA Editing

In humans, differences between the sequences of DNA and RNA—also called RNA editing—are widespread.[31,32 and 33] The most frequent type of RNA editing is conversion of adenosine to inosine by adenosine deaminases acting on RNA (ADARs). The splicing and translational machineries subsequently recognize the inosine as a guanosine.[34] Some tumor genomes have a higher percentage of RNA-DNA differences than their matched normal genomes.[35]

References

Jiang Q., Crews L. A., Barrett C. L., Chun H. J., Court A. C., et al. (2013) ADAR1 promotes malignant progenitor reprogramming in chronic myeloid leukemia. Proc Natl Acad Sci U S A 110: 1041-1046

The authors showed that blast crisis chronic myeloid leukemia (CML) progenitors have an increased IFN-γ pathway gene expression as well as BCR-ABL amplification. During CML progression, they also found enhanced expression of the IFN-responsive ADAR1 p150 isoform and increased adenosine-to-inosine RNA editing.

Illumina Technology: HiSeq 2000 system for RNA-Seq 50 bp paired-end reads

Chen L., Li Y., Lin C. H., Chan T. H., Chow R. K., et al. (2013) Recoding RNA editing of AZIN1 predisposes to hepatocellular carcinoma. Nat Med 19: 209-216

Bahn J. H., Lee J. H., Li G., Greer C., Peng G., et al. (2012) Accurate identification of A-to-I RNA editing in human by transcriptome sequencing. Genome Res 22: 142-150

Li M., Wang I. X. and Cheung V. G. (2012) Response to Comments on "Widespread RNA and DNA Sequence Differences in the Human Transcriptome". Science 335: 1302-1302

Liu J., Lee W., Jiang Z., Chen Z., Jhunjhunwala S., et al. (2012) Genome and transcriptome sequencing of lung cancers reveal diverse mutational and splicing events. Genome Res 22: 2315-2327

Peng Z., Cheng Y., Tan B. C., Kang L., Tian Z., et al. (2012) Comprehensive analysis of RNA-Seq data reveals extensive RNA editing in a human transcriptome. Nat Biotechnol 30: 253-260

---

[31] Li M., Wang I. X., Li Y., Bruzel A., Richards A. L., et al. (2011) Widespread RNA and DNA sequence differences in the human transcriptome. Science 333: 53-58

[32] Ju Y. S., Kim J. I., Kim S., Hong D., Park H., et al. (2011) Extensive genomic and transcriptional diversity identified through massively parallel DNA and RNA sequencing of eighteen Korean individuals. Nat Genet 43: 745-752

[33] Wu D., Lamm A. T. and Fire A. Z. (2011) Competition between ADAR and RNAi pathways for an extensive class of RNA targets. Nat Struct Mol Biol 18: 1094-1101

[34] McGettigan P. A. (2013) Transcriptomics in the RNA-seq era. Curr Opin Chem Biol 17: 4-11

[35] Liu J., Lee W., Jiang Z., Chen Z., Jhunjhunwala S., et al. (2012) Genome and transcriptome sequencing of lung cancers reveal diverse mutational and splicing events. Genome Res 22: 2315-2327

MicroRNA and Noncoding RNAs

MicroRNAs (miRNAs) range in size from 17 to 25 bp and are members of the noncoding RNA (ncRNA) family. They regulate a variety of biological functions, including development, cell proliferation, cell differentiation, signal transduction, apoptosis, metabolism, and life span.[36-37] miRNAs suppress the gene's post-transcriptional expression through the interaction of the RNA-induced silencing complex (RISC) with its target recognition sites in the 3'-untranslated region (3'-UTR) or the coding regions of the transcript.[38-39] Many miRNAs are located in genomic regions that are deleted or amplified in various cancer types, which indicate that they might play a prominent role in cancer progression.[40] Editing sites have also been observed in miRNAs, suggesting a potential link between RNA editing and miRNA-mediated regulation.[41] The ease of measurement, relative stability, and role in the control of large numbers of mRNAs make miRNAs attractive markers for the detection and staging of cancer during diagnosis and treatment.[42] miRNA preparation and detection have become routine and users can expect very high sensitivity and specificity from the manufacturers' protocols.

References

Law P. T., Qin H., Ching A. K., Lai K. P., Co N. N., et al. (2013) Deep sequencing of small RNA transcriptome reveals novel non-coding RNAs in hepatocellular carcinoma. J Hepatol

The paper describes a new PIWI-interacting RNA (piRNA) piR-Hep1 involved in liver tumor progression. PiR-Hep1 was upregulated in 46.6% of hepatocellular carcinoma (HCC) tumors compared to adjacent normal liver. Silencing of piR-Hep1 inhibited cell viability, motility, and invasiveness. The authors also found abundant expression of miR-1323 in HCC and its distinct association in tumors arising from a cirrhotic background.

Illumina Technology: HiSeq 2000 system with small-RNA library preparation

Whittaker S. R., Theurillat J. P., Van Allen E., Wagle N., Hsiao J., et al. (2013) A Genome-Scale RNA Interference Screen Implicates NF1 Loss in Resistance to RAF Inhibition. Cancer Discov 3: 350-362

Lian L., Qu L., Chen Y., Lamont S. J. and Yang N. (2012) A systematic analysis of miRNA transcriptome in Marek's disease virus-induced lymphoma reveals novel and differentially expressed miRNAs. PLoS ONE 7: e51003

[36] Kloosterman W. P. and Plasterk R. H. (2006) The diverse functions of microRNAs in animal development and disease. Dev Cell 11: 441-450

[37] Cheng W. C., Chung I. F., Huang T. S., Chang S. T., Sun H. J., et al. (2013) YM500: a small RNA sequencing (smRNA-seq) database for microRNA research. Nucleic Acids Res 41: D285-294

[38] Guo H., Ingolia N. T., Weissman J. S. and Bartel D. P. (2010) Mammalian microRNAs predominantly act to decrease target mRNA levels. Nature 466: 835-840

[39] Lewis B. P., Burge C. B. and Bartel D. P. (2005) Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell 120: 15-20

[40] Enfield K. S., Pikor L. A., Martinez V. D. and Lam W. L. (2012) Mechanistic Roles of Noncoding RNAs in Lung Cancer Biology and Their Clinical Implications. Genet Res Int 2012: 737416

[41] Peng Z., Cheng Y., Tan B. C., Kang L., Tian Z., et al. (2012) Comprehensive analysis of RNA-Seq data reveals extensive RNA editing in a human transcriptome. Nat Biotechnol 30: 253-260

[42] Frampton A. E., Gall T. M., Castellano L., Stebbing J., Jiao L. R., et al. (2013) Towards a clinical use of miRNAs in pancreatic cancer biopsies. Expert review of molecular diagnostics 13: 31-34

Notes on experimental design:

Sequencing depth has a direct relationship with the sensitivity of detection. In a typical experiment, where one sample is applied to one lane of a flow cell, the sequencing depth is very high and leads to extremely sensitive detection. For this reason, miRNA read depth is rarely a consideration. In screening applications, or studies where such a high level of detection is not required, the samples can be indexed and several samples applied to a single channel in the flow cell. When deciding on the depth of coverage, keep in mind that miRNAs control gene expression and small changes in miRNA levels may impact many protein-coding genes.

A newly discovered miRNA should be confirmed with a functional assay, such as Ago2 binding[43] or knockout experiments.

The experiment should contain enough samples to establish statistical confidence. The presence of a miRNA in the tumors of a few patients is enough to create the hypothesis that the miRNA may play a role in the disease. Usually a large number of patients will be needed to test the hypothesis and establish statistical confidence. At present, there is no universally recognized approach for establishing statistical confidence and multiple test correction in sequencing studies. Sequencing-based miRNA profiling does not provide absolute measurements of miRNA expression, but rather the relative counts of different miRNAs, such as in a tumor-normal pair.

Sample stratification is an issue in cancer samples. A particular cancer phenotype may represent several different etiologies and mechanisms. For a rigorous analysis, there should be enough samples in the study to represent each tumor subtype adequately. Expression of miRNAs may change with the progression of the tumor, so establishing the stage and grade of the tumor should be incorporated in the experimental design. For biomarker discovery, the requirements and experimental designs have been well-established. Any newly discovered marker must be validated in large, independent cohorts.

---

[43] Persson H., Kvist A., Rego N., Staaf J., Vallon-Christersson J., et al. (2011) Identification of new microRNAs in paired normal and tumor breast tissue suggests a dual role for the ERBB2/Her2 gene. Cancer Res 71: 78-86

RNA-Protein Binding (CLIP-Seq)

In human cells most mRNAs (or pre-mRNAs) are associated with heterogeneous nuclear ribonucleoparticle (hnRNP) proteins, forming large hnRNP-RNA complexes.[44] hnRNP proteins play a role in all crucial aspects of RNA processing, including pre-mRNA splicing, and mRNA export, localization, translation, and stability.[45-46] The hnRNP proteins of dozens of other RNA-binding proteins (RBPs) and genes are associated with cancer.[47]

RNA-protein interactions can be measured with cross-linked immunoprecipitation sequencing (CLIP-Seq). In CLIP-Seq, cells are treated with ultraviolet light to covalently cross-link RBP-RNA complexes. The cells are then lysed, the RBP-RNA complexes are immunoprecipitated, and the RNA sequenced.[48]

References

Wilbert M. L., Huelga S. C., Kapeli K., Stark T. J., Liang T. Y., et al. (2012) LIN28 binds messenger RNAs at GGAGA motifs and regulates splicing factor abundance. Mol Cell 48: 195-206

LIN28 is a conserved RNA-binding protein implicated in pluripotency, reprogramming, and oncogenesis. Aberrant upregulation of LIN28 has been found in a range of different cancer cells and primary tumor tissues. In this paper the authors use CLIP-Seq to identify discrete LIN28-binding sites in a quarter of human transcripts. These sites revealed that LIN28 binds to GGAGA sequences enriched within loop structures in mRNAs. They also found that LIN28 expression causes widespread downstream changes in alternative splicing.

Illumina Technology: Illumina Genome Analyzer$_{II}$ System RNA-Seq and small RNA-Seq with TruSeq barcodes

Hoque M., Ji Z., Zheng D., Luo W., Li W., et al. (2013) Analysis of alternative cleavage and polyadenylation by 3' region extraction and deep sequencing. Nat Methods 10: 133-139

Huelga S. C., Vu A. Q., Arnold J. D., Liang T. Y., Liu P. P., et al. (2012) Integrative genome-wide analysis reveals cooperative regulation of alternative splicing by hnRNP proteins. Cell Rep 1: 167-178

---

[44] Dreyfuss G., Matunis M. J., Pinol-Roma S. and Burd C. G. (1993) hnRNP proteins and the biogenesis of mRNA. Annu Rev Biochem 62: 289-321

[45] Han S. P., Tang Y. H. and Smith R. (2010) Functional diversity of the hnRNPs: past, present and perspectives. Biochem J 430: 379-392

[46] Yeo G. W., Coufal N. G., Liang T. Y., Peng G. E., Fu X. D., et al. (2009) An RNA code for the FOX2 splicing regulator revealed by mapping RNA-protein interactions in stem cells. Nat Struct Mol Biol 16: 130-137

[47] Huelga S. C., Vu A. Q., Arnold J. D., Liang T. Y., Liu P. P., et al. (2012) Integrative genome-wide analysis reveals cooperative regulation of alternative splicing by hnRNP proteins. Cell Rep 1: 167-178

[48] Katz Y., Wang E. T., Airoldi E. M. and Burge C. B. (2010) Analysis and design of RNA sequencing experiments for identifying isoform regulation. Nat Methods 7: 1009-1015

Epigenetic changes during cancer progression are associated with aberrant gene expression. Recent evidence indicates that epigenetic changes may play a role in cancer initiation. Epigenetic control is mediated thorough multiple processes, including DNA modification via methylation or acetylation, histone modification, and nucleosome remodeling.[49] Mutations in genes that control the epigenome are surprisingly common in human cancers. Next-generation sequencing provides an extensive set of tools to map mutations and measure their impact on cancer progression.
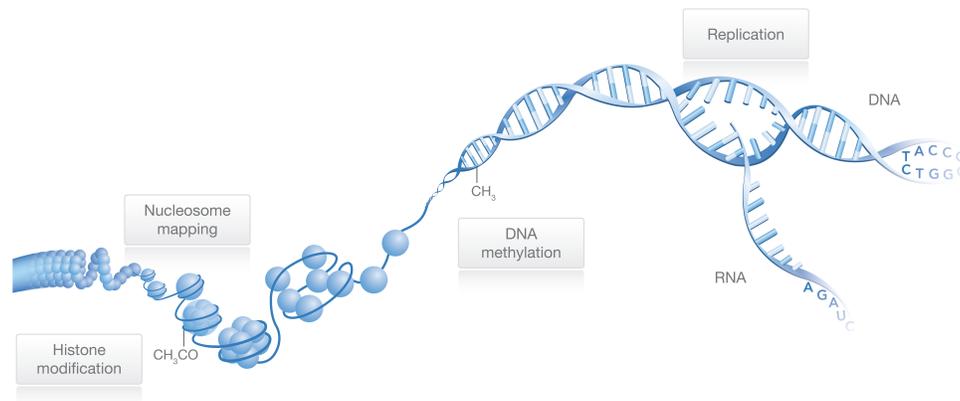
Genetic mutations in epigenetic modifiers in cancer. Mutations in the three classes of epigenetic modifiers are frequently observed in various types of cancers, which highlight the crosstalk between genetics and epigenetics. Mutations of epigenetic modifiers have the potential to cause genome-wide epigenetic alterations in cancer. Understanding the relationship of genetic and the epigenetic changes will offer novel insights for cancer therapies.

Review

Umer M. and Herceg Z. (2013) Deciphering the Epigenetic Code: An Overview of DNA Methylation Analysis Methods. Antioxid Redox Signal

DNA Modifications

DNA modifications can be readily determined via a variety of techniques. The choice of technique depends on the throughput and resolution required.

---

[49] Easwaran H. and Baylin S. B. (2013) Epigenetic abnormalities in cancer find a "home on the range". Cancer Cell 23: 1-3

Technique	Characteristics
Whole-genome bisulfite sequencing (BS-Seq or WGBS)	In bisulfite-treated DNA, unmethylated cytosines are converted into thymidines.[50] Next-generation sequencing provides a complete overview of CpG methylation level at base-pair resolution.
Reduced-representation bisulfite sequencing (RRBS) or restriction enzyme-enriched sequencing (rrBS-Seq)	RRBS involves digesting DNA with a methylation-insensitive enzyme to enrich the sample for CpG islands. The CpG-enriched sample is then bisulfite-treated and sequenced. RRBS is an efficient technique that is suitable for obtaining information from most CpG islands and information about sequences outside CpG-rich regions.[51-52]
Affinity-enrichment-based sequencing techniques (MBD-Seq or MeDIP-Seq)	MBD-Seq[53] and MeDIP-seq[54] combine the advantages of next-generation sequencing and enrichment of methylated regions by immunoprecipitation.
DNA methylation arrays	CpG-specific array technology is an alternative option for determining a genome-wide DNA methylation profile. The Human Methylation 450 beadchip assay (Illumina) covers 99% of all human RefSeq[55] genes and approximately 450,000 CpGs overall.[56]
Locus-specific DNA methylation analysis	In addition to genome-wide technologies, locus-specific identification of the DNA methylation level is a cost-effective strategy, especially if single genes are already established as biomarkers for diagnosis or prognosis.

References

Bert S. A., Robinson M. D., Strbenac D., Statham A. L., Song J. Z., et al. (2013) Regional activation of the cancer genome by long-range epigenetic remodeling. Cancer Cell 23: 9-22

The authors identified a mechanism of domain gene deregulation through coordinated long-range epigenetic activation (LREA). These regions typically span 1 Mb and include key oncogenes, microRNAs, and cancer biomarker genes. Gene promoters within LREA domains are characterized by a gain of active chromatin markers and a loss of repressive markers.

Illumina Technology: Illumina 450K arrays, Genome Analyzer$_{II}$ System for MBDCap-seq

Brastianos P. K., Horowitz P. M., Santagata S., Jones R. T., McKenna A., et al. (2013) Genomic sequencing of meningiomas identifies oncogenic SMO and AKT1 mutations. Nat Genet 45: 285-289

To identify and validate somatic genetic alterations in meningiomas, the authors performed whole-genome or whole-exome sequencing on 17 meningiomas and focused sequencing on an additional 48 tumors. The spectrum of observed mutations is extensive, but they confirmed focal NF2 inactivation in 43% of tumors and found alterations in epigenetic modifiers in an additional 8% of tumors.

Illumina Technology: HiSeq 2000 system Whole-genome sequencing, whole exome sequencing and targeted sequencing

[50] Frommer M., McDonald L. E., Millar D. S., Collis C. M., Watt F., et al. (1992) A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. Proc Natl Acad Sci U S A 89: 1827-1831

[51] Meissner A., Gnirke A., Bell G. W., Ramsahoye B., Lander E. S., et al. (2005) Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. Nucleic Acids Res 33: 5868-5877

[52] Lee E. J., Luo J., Wilson J. M. and Shi H. (2012) Analyzing the cancer methylome through targeted bisulfite sequencing. Cancer Lett

[53] Serre D., Lee B. H. and Ting A. H. (2010) MBD-isolated Genome Sequencing provides a high-throughput and comprehensive survey of DNA methylation in the human genome. Nucleic Acids Res 38: 391-399

[54] Jacinto F. V., Ballestar E. and Esteller M. (2008) Methyl-DNA immunoprecipitation (MeDIP): hunting down the DNA methylome. Biotechniques 44: 35, 37, 39 passim

[55] www.ncbi.nlm.nih.gov/RefSeq/

[56] Dedeurwaerder S., Defrance M., Calonne E., Denis H., Sotiriou C., et al. (2011) Evaluation of the Infinium Methylation 450K technology. Epigenomics 3: 771-784

Duncan C. G., Barwick B. G., Jin G., Rago C., Kapoor-Vazirani P., et al. (2012) A heterozygous IDH1R132H/WT mutation induces genome-wide alterations in DNA methylation. Genome Res 22: 2339-2355

Monoallelic point mutations of the NADP$^+$-dependent isocitrate dehydrogenases IDH1 and IDH2 occur frequently in gliomas, acute myeloid leukemias, and chondromas. The authors show that heterozygous expression of the IDH1^{R132H} allele is sufficient to induce the genome-wide alterations in DNA methylation characteristic of these tumors. This demonstrates a causal role for IDH1$^{R132H/WT}$ mutants in driving epigenetic instability in human cancer cells.

Illumina Technology: Illumina Infinium® Methylation27 and Illumina HumanMethylation450

Zhang J., Benavente C. A., McEvoy J., Flores-Otero J., Ding L., et al. (2012) A novel retinoblastoma therapy from genomic and epigenetic analyses. Nature 481: 329-334

Retinoblastoma is an aggressive childhood cancer of the developing retina. It is initiated by RB1 inactivation, but the underlying mechanism is not known. In a highly aggressive cancer such as this, many genes are involved but RB1 was the only known cancer gene mutated. In contrast to the limited number of somatic mutations present, the tumor showed profound changes in its methylation profile relative to normal retinoblasts. One of the most striking results was the induction of the expression of the proto-oncogene spleen tyrosine kinase (SYK) in human retinoblastoma. SYK is required for tumor cell survival. The researchers went on to show that small-molecule inhibition of SYK caused cell death in retinoblastoma cells in culture and *in vivo.*

Illumina Technology: Genome Analyzer$_{IIx}$ system 101-bp paired-end both targeted and whole genome sequencing; Methylation 27 arrays

Cahill N., Bergh A. C., Kanduri M., Goransson-Kultima H., Mansouri L., et al. (2013) 450K-array analysis of chronic lymphocytic leukemia cells reveals global DNA methylation to be relatively stable over time and similar in resting and proliferative compartments. Leukemia 27: 150-158

Jithesh P. V., Risk J. M., Schache A. G., Dhanda J., Lane B., et al. (2013) The epigenetic landscape of oral squamous cell carcinoma. Br J Cancer 108: 370-379

Boyle P., Clement K., Gu H., Smith Z. D., Ziller M., et al. (2012) Gel-free multiplexed reduced representation bisulfite sequencing for large-scale DNA methylation profiling. Genome Biol 13: R92

Brinkman A. B., Gu H., Bartels S. J., Zhang Y., Matarese F., et al. (2012) Sequential ChIP-bisulfite sequencing enables direct genome-scale investigation of chromatin and DNA methylation cross-talk. Genome Res 22: 1128-1138

Leidenroth A., Sorte H. S., Gilfillan G., Ehrlich M., Lyle R., et al. (2012) Diagnosis by sequencing: correction of misdiagnosis from FSHD2 to LGMD2A by whole-exome analysis. Eur J Hum Genet 20: 999-1003

Selamat S. A., Chung B. S., Girard L., Zhang W., Zhang Y., et al. (2012) Genome-scale analysis of DNA methylation in lung adenocarcinoma and integration with mRNA expression. Genome Res 22: 1197-1211

Sturm D., Witt H., Hovestadt V., Khuong-Quang D. A., Jones D. T., et al. (2012) Hotspot mutations in H3F3A and IDH1 define distinct epigenetic and biological subgroups of glioblastoma. Cancer Cell 22: 425-437

Notes on experimental design:

Each tissue and cell type has a unique methylation pattern; therefore, the tissue of interest must be available for analysis. Tissue-normal adjacent pairs help simplify the analysis.

The extremely large number of CpG markers produced by bisulfite sequencing is challenging to interpret, and robust statistical analysis is still elusive. However, there are practical approaches to simplify the analysis:
- RRBS-Seq simplifies analysis by limiting the coverage.
- Integrated analysis improves the interpretability of the results substantially. For example, combining expression analysis with methylation assays allows the researcher to focus on genes whose expression levels have changed.
- Limit the analysis to a gene or region of interest. This approach is effective in a follow-up for a GWAS study or where there is already experimental evidence for gene regulation or chromatin remodeling in the region of interest. Unlike reduced-representation approaches, this method enables analysis of additional regions as more information becomes available.

Tissue cultures should be used with caution. Over time, with extended-proliferation tissue, cultures may change their level of methylation and become less representative of the original tissue samples.[57]

Histone Modifications (Methylation)

Histone modifications usually refer to methylation and acetylation. Methylation of histones H3K9, H3K27, and H4K20 often correlates with repression of gene transcription, while trimethylation of H3K4 and H3K36 is associated with actively transcribed chromatin. Histone acetylation is almost always associated with chromatin accessibility and increased levels of transcriptional activity. By manipulating chromatin states and DNA accessibility, epigenetic modification plays a critical role in the control of gene expression across diverse developmental stages, tissue types, and diseases.[58]

Review

Caren H., Pollard S. M. and Beck S. (2012) The good, the bad and the ugly: Epigenetic mechanisms in glioblastoma. Mol Aspects Med in press

References

Wilkinson A. C., Ballabio E., Geng H., North P., Tapia M., et al. (2013) RUNX1 is a key target in t(4;11) leukemias that contributes to gene activation through an AF4-MLL complex interaction. Cell Rep 3: 116-127

This paper reports a mechanism of transformation whereby two oncogenic fusion proteins cooperate by activating a target gene and then modulating the function of its downstream product.

Illumina Technology: Genome Analyzer$_{IIx}$ system or HiSeq 2000 system for ChIP-Seq

---

[57] Liu J., Zhang Z., Bando M., Itoh T., Deardorff M. A., et al. (2010) Genome-wide DNA methylation analysis in cohesin mutant human cell lines. Nucleic Acids Res 38: 5657-5671
[58] Bernstein B. E., Meissner A. and Lander E. S. (2007) The mammalian epigenome. Cell 128: 669-681

Paul S., Kuo A., Schalch T., Vogel H., Joshua-Tor L., et al. (2013) Chd5 requires PHD-mediated histone 3 binding for tumor suppression. Cell Rep 3: 92-102

Wilkinson A. C., Ballabio E., Geng H., North P., Tapia M., et al. (2013) RUNX1 is a key target in t(4;11) leukemias that contributes to gene activation through an AF4-MLL complex interaction. Cell Rep 3: 116-127

Guo Y., Monahan K., Wu H., Gertz J., Varley K. E., et al. (2012) CTCF/cohesin-mediated DNA looping is required for protocadherin alpha promoter choice. Proc Natl Acad Sci U S A 109: 21081-21086

Ntziachristos P., Tsirigos A., Van Vlierberghe P., Nedjic J., Trimarchi T., et al. (2012) Genetic inactivation of the polycomb repressive complex 2 in T cell acute lymphoblastic leukemia. Nat Med 18: 298-301

Robinson G., Parker M., Kranenburg T. A., Lu C., Chen X., et al. (2012) Novel mutations target distinct subgroups of medulloblastoma. Nature 488: 43-48

Notes on experimental design:

Each tissue and cell type has a unique methylation pattern; therefore, the tissue of interest must be available for analysis. Tissue-normal adjacent pairs help simplify the analysis.

Histone methylation is measured with a variation of ChIP-Seq where the antibody is specific to the methylated histone of interest. A wide variety of methylated histones are available commercially.

Tissue cultures should be used with caution. Over time, with extended-proliferation tissue, cultures may change their level of methylation and become less representative of the original tissue samples.[59]

---

[59] Liu J., Zhang Z., Bando M., Itoh T., Deardorff M. A., et al. (2010) Genome-wide DNA methylation analysis in cohesin mutant human cell lines. Nucleic Acids Res 38: 5657-5671

Chromatin Structure and Rearrangements

Chromosomal rearrangements require formation and joining of DNA double-strand breaks. These events disrupt the integrity of the genome and are frequently observed in leukemias, lymphomas, and sarcomas.[60] The recurrence of gene fusions between specific genes in multiple individuals indicates that those genes must be physically close at some stage in the cell cycle.[61]
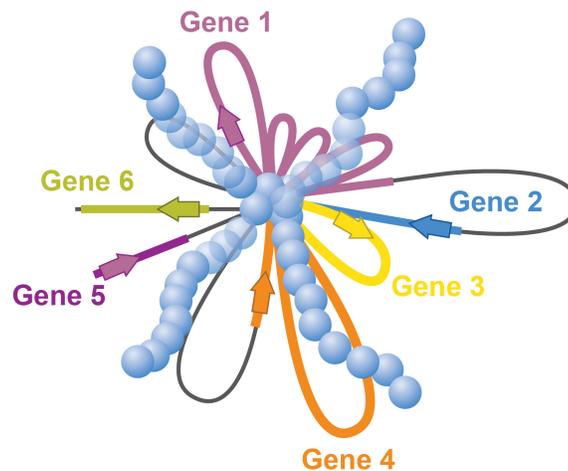
A hypothetical three-dimensional, transcriptionally active complex containing dense looping positions. This schematic diagram is based on the looping events detected, assuming that all looping events can occur in a single cell. In this model, all small loops converge onto a common core base (blue spheres). Loops reduce the physical size of the active transcriptional complex to enhance the accessibility of transcription factors to specific genomic sites. Kohwi-Shigematsu T., Kohwi Y., Takahashi K., Richards H. W., Ayers S. D., et al. (2012) SATB1-mediated functional packaging of chromatin into loops. Methods 58: 243-254

---

[60] Klein I. A., Resch W., Jankovic M., Oliveira T., Yamane A., et al. (2011) Translocation-capture sequencing reveals the extent and nature of chromosomal rearrangements in B lymphocytes. Cell 147: 95-106
[61] Hakim O., Resch W., Yamane A., Klein I., Kieffer-Kwon K. R., et al. (2012) DNA damage defines sites of recurrent chromosomal translocations in B lymphocytes. Nature 484: 69-74

a Chromatin conformation capture

b CHiA-PET

a DNA-binding protein ChiP–seq

Crosslink proteins and DNA

Crosslink proteins and DNA

Crosslink proteins and DNA

Exonuclease

Sample fragmentation
• Restriction enzymes

Sample fragmentation
• Sonication

Sample fragmentation
• Sonication
• Endonuclease (ChiP–exo)

Immunoprecipitate

Immunoprecipitate and
then purify DNA

PCR amplify ligated junctions

Ligation

Restriction enzyme digestion

DNA library creation and
paired-end sequencing

DNA library creation and
paired-end sequencing

Detecting chromatin interactions. In three-dimensional space, distal genomic regions on the same or different chromosomes interact, and this interaction can be mediated by one or more DNA-binding proteins. a) ChIP-Seq uses a chromatin immunoprecipitation step to identify DNA-protein interactions. Various DNA fragmentation approaches and exonucleases can be used to narrow the size distribution of the fragments. b) Chromatin conformation capture experiments use a ligation step to join interacting chromatin fragments. This approach can identify proteins bound to distant sequences. c) Chromatin interaction analysis with paired-end tag sequencing (ChIA-PET) similarly detects chromatin interactions using a ligation step to pair nonadjacent interacting regions. However, ChIA-PET uses a chromatin immunoprecipitation (ChIP) step to only identify only interactions with a particular protein, such as RNA polymerase II. Furey T. S. (2012) ChIP-seq and beyond: new and improved methodologies to detect and characterize protein-DNA interactions. Nat Rev Genet 13: 840-852

References

Papantonis A., Kohro T., Baboo S., Larkin J. D., Deng B., et al. (2012) TNFalpha signals through specialized factories where responsive coding and miRNA genes are transcribed. EMBO J 31: 4404-4414

The authors use sequencing, combined with chromosome conformation capture (3C)[62] and ChIA-PET,[63] to show that TNFα induces responsive genes to congregate in discrete "NF-κB factories." Some factories further specialize in transcribing responsive genes encoding miRNAs that target downregulated mRNAs.

Illumina Technology: Genome Analyzer system to sequence Chromosome Conformation Capture (3C) and ChIA-PET (>10 ng) fragments.

---

[62] Zhao Z., Tavoosidana G., Sjolinder M., Gondor A., Mariano P., et al. (2006) Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. Nat Genet 38: 1341-1347
[63] Li G., Fullwood M. J., Xu H., Mulawadi F. H., Velkov S., et al. (2010) ChIA-PET tool for comprehensive chromatin interaction analysis with paired-end tag sequencing. Genome Biol 11: R22

Rocha P. P., Micsinai M., Kim J. R., Hewitt S. L., Souza P. P., et al. (2012) Close proximity to Igh is a contributing factor to AID-mediated translocations. Mol Cell 47: 873-885

Nuclear organization may determine "off-target" activity and the choice of fusion partners. This study indicates that the vast majority of known activation-induced cytidine deaminase (AID)-mediated Igh translocation partners are found in chromosomal domains that contact this locus during class switching. Further, these interaction domains can be used to identify other genes that are targeted by AID.

Illumina Technology: Genome Analyzer$_{IIx}$ system. The 4C library was sequenced to 72 bases in a single-read

Theodoratou E., Montazeri Z., Hawken S., Allum G. C., Gong J., et al. (2012) Systematic meta-analyses and field synopsis of genetic association studies in colorectal cancer. J Natl Cancer Inst 104: 1433-1457

The authors show that the T-cell-specific transcription factor GATA3 plays an important role in mediating enhancer accessibility at regulatory regions involved in estrogen receptor 1 (ESR1)-mediated transcription. GATA3 silencing resulted in a global redistribution of cofactors and active histone marks prior to estrogen stimulation.

Illumina Technology: Genome Analyzer$_{IIx}$ system for ChIP-Seq fragments and Gene expression analysis was carried out on Illumina Human HT12 V3 arrays

Hakim O., Resch W., Yamane A., Klein I., Kieffer-Kwon K. R., et al. (2012) DNA damage defines sites of recurrent chromosomal translocations in B lymphocytes. Nature 484: 69-74

The authors find that, in the absence of recurrent DNA damage, translocations between Igh or Myc and all other genes are directly related to their contact frequency in cultured mouse B lymphocytes. Conversely, translocations associated with recurrent site-directed DNA damage are proportional to the rate of DNA break formation. They conclude that nontargeted rearrangements reflect nuclear organization, whereas DNA break formation governs the location and frequency of recurrent translocations, including those driving B-cell malignancies.

Illumina Technology: Genome Analyzer$_{IIx}$ system with 36 or 54 bp paired-end reads

Jankovic M., Feldhahn N., Oliveira T. Y., Silva I. T., Kieffer-Kwon K. R., et al. (2013) 53BP1 alters the landscape of DNA rearrangements and suppresses AID-induced B cell lymphoma. Mol Cell 49: 623-631

Cowper-Sal lari R., Zhang X., Wright J. B., Bailey S. D., Cole M. D., et al. (2012) Breast cancer risk-associated SNPs modulate the affinity of chromatin for FOXA1 and alter gene expression. Nat Genet 44: 1191-1198

Greenman C. D., Pleasance E. D., Newman S., Yang F., Fu B., et al. (2012) Estimation of rearrangement phylogeny for cancer genomes. Genome Res 22: 346-361

Guo Y., Monahan K., Wu H., Gertz J., Varley K. E., et al. (2012) CTCF/cohesin-mediated DNA looping is required for protocadherin alpha promoter choice. Proc Natl Acad Sci U S A 109: 21081-21086

Kohwi-Shigematsu T., Kohwi Y., Takahashi K., Richards H. W., Ayers S. D., et al. (2012) SATB1-mediated functional packaging of chromatin into loops. Methods 58: 243-254

McBride D. J., Etemadmoghadam D., Cooke S. L., Alsop K., George J., et al. (2012) Tandem duplication of chromosomal segments is common in ovarian and breast cancer genomes. J Pathol 227: 446-455

Ng C. K., Cooke S. L., Howe K., Newman S., Xian J., et al. (2012) The role of tandem duplicator phenotype in tumour evolution in high-grade serous ovarian cancer. J Pathol 226: 703-712

Zhang X., Cowper-Sal lari R., Bailey S. D., Moore J. H. and Lupien M. (2012) Integrative functional genomics identifies an enhancer looping to the SOX9 gene disrupted by the 17q24.3 prostate cancer risk locus. Genome Res 22: 1437-1446

Methods

Rausch T., Zichner T., Schlattl A., Stutz A. M., Benes V., et al. (2012) DELLY: structural variant discovery by integrated paired-end and split-read analysis. Bioinformatics 28: i333-i339

van de Werken H. J., Landan G., Holwerda S. J., Hoichman M., Klous P., et al. (2012) Robust 4C-seq data analysis to screen for regulatory DNA interactions. Nat Methods 9: 969-972

Zhang J., Poh H. M., Peh S. Q., Sia Y. Y., Li G., et al. (2012) ChIA-PET analysis of transcriptional chromatin interactions. Methods 58: 289-299

Integrative Analysis

All biological processes are interconnected, and every change in one process in the cancer cell impacts all other processes. A mutation can impact the activity of an expressed protein, which in turn can impact the methylation of DNA, which in turn can impact the expression of many other genes and so on.[64] The vast number of mutations that are unique to every individual, when coupled with this chain of events, gives insight into the wide range of disease phenotypes that characterize many cancers. An integrated analysis is a step towards reflecting the true complexity of cancer biology. Researchers now have the ability to measure most of these processes individually, but real progress in the understanding and treatment of cancer will come from an integrated analysis of all these processes.

References

Weischenfeldt J., Simon R., Feuerbach L., Schlangen K., Weichenhan D., et al. (2013) Integrative genomic analyses reveal an androgen-driven somatic alteration landscape in early-onset prostate cancer. Cancer Cell 23: 159-170

The authors show that early-onset prostate cancer formation involves androgen-driven structural rearrangements. By comparison, elderly-onset prostate cancers accumulate nonandrogen-associated structural rearrangements, indicative of a different tumor formation mechanism.

Illumina Technology: HiSeq 2000 system 101 bp paired-end reads, mate-pair libraries, strand-specific mRNA-Seq, and miRNA-Seq

Cowper-Sal lari R., Zhang X., Wright J. B., Bailey S. D., Cole M. D., et al. (2012) Breast cancer risk-associated SNPs modulate the affinity of chromatin for FOXA1 and alter gene expression. Nat Genet 44: 1191-1198

The authors show that breast cancer risk–associated SNPs are enriched in the cistromes of FOXA1 1 and ESR1 1, and the epigenome of histone H3 lysine 4 monomethylation (H3K4me1). The majority of the risk-associated SNPs modulate the affinity of chromatin for FOXA1 at distal regulatory elements, which results in allele-specific gene expression.

Illumina Technology: Genome Analyzer$_{II}$ System for ChIP-Seq and Human OmniExpress

Peifer M., Fernandez-Cuesta L., Sos M. L., George J., Seidel D., et al. (2012) Integrative genome analyses identify key somatic driver mutations of small-cell lung cancer. Nat Genet 44: 1104-1110

The authors found evidence for inactivation of TP53 and RB1, and recurrent mutations in the genes that encode histone modifiers. Furthermore, they observed mutations in PTEN, SLIT2, and EPHA7, as well as focal amplifications of the FGFR1 tyrosine kinase gene. This integrated analysis indicates that histone modifications may be involved in small-cell lung cancer (SCLC).

Illumina Technology: Genome Analyzer$_{II}$ system mRNA-Seq and whole-exome sequencing to 95 bp paired-end reads. HiSeq 2000 System whole-genome sequencing to 100 bp paired-end reads.

---

[64] Peifer M., Fernandez-Cuesta L., Sos M. L., George J., Seidel D., et al. (2012) Integrative genome analyses identify key somatic driver mutations of small-cell lung cancer. Nat Genet 44: 1104-1110

Jankovic M., Feldhahn N., Oliveira T. Y., Silva I. T., Kieffer-Kwon K. R., et al. (2013) 53BP1 alters the landscape of DNA rearrangements and suppresses AID-induced B cell lymphoma. Mol Cell 49: 623-631

Robinson D. R., Wu Y. M., Kalyana-Sundaram S., Cao X., Lonigro R. J., et al. (2013) Identification of recurrent NAB2-STAT6 gene fusions in solitary fibrous tumor by integrative sequencing. Nat Genet 45: 180-185

Sausen M., Leary R. J., Jones S., Wu J., Reynolds C. P., et al. (2013) Integrated genomic analyses identify ARID1A and ARID1B alterations in the childhood cancer neuroblastoma. Nat Genet 45: 12-17

Straessler K. M., Jones K. B., Hu H., Jin H., van de Rijn M., et al. (2013) Modeling clear cell sarcomagenesis in the mouse: cell of origin differentiation state impacts tumor characteristics. Cancer Cell 23: 215-227

Yang D., Sun Y., Hu L., Zheng H., Ji P., et al. (2013) Integrated analyses identify a master microRNA regulatory network for the mesenchymal subtype in serous ovarian cancer. Cancer Cell 23: 186-199

Carter S. L., Cibulskis K., Helman E., McKenna A., Shen H., et al. (2012) Absolute quantification of somatic DNA alterations in human cancer. Nat Biotechnol 30: 413-421

Jones D. T., Jager N., Kool M., Zichner T., Hutter B., et al. (2012) Dissecting the genomic complexity underlying medulloblastoma. Nature 488: 100-105

Kresse S. H., Rydbeck H., Skarn M., Namlos H. M., Barragan-Polania A. H., et al. (2012) Integrative analysis reveals relationships of genetic and epigenetic alterations in osteosarcoma. PLoS ONE 7: e48262

Peifer M., Fernandez-Cuesta L., Sos M. L., George J., Seidel D., et al. (2012) Integrative genome analyses identify key somatic driver mutations of small-cell lung cancer. Nat Genet 44: 1104-1110

Sturm D., Witt H., Hovestadt V., Khuong-Quang D. A., Jones D. T., et al. (2012) Hotspot mutations in H3F3A and IDH1 define distinct epigenetic and biological subgroups of glioblastoma. Cancer Cell 22: 425-437

Technical Considerations

A good experimental design will optimize the performance of the technology to produce the most interpretable and robust results. This section is intended to highlight the unique characteristics of the biology and the technology that researchers should keep in mind when designing their experiments.

Experimental designs in cancer research offer some unique challenges. A typical tumor sample consists of two genomes: the germline inherited from the parents and the somatic mutations that accumulate during progression of the disease.[65] The percentage of tumor cells in the sample can vary between 10% and 100%.[66] The tumor genome is also dynamic and can accumulate *de novo* mutations rapidly. As a result, tumors can consist of several clonal types.[67]

The number of samples involved in most currently published studies is very small and can be regarded as hypothesis-generating. As more sequencing information becomes available, most cancer types can be divided into several subpopulations based on their molecular phenotype.[68] This severely decreases the power of the experiment and increases the number of samples required for a rigorous analysis. A partial solution is to use whole-genome sequencing in the discovery phase to find new mutations. In the second phase, whole-exome or targeted sequencing can be used to confirm the newly discovered mutations and determine their abundance in a large cohort. However, statistically rigorous whole-genome sequencing experiments in the future will likely be very large, requiring in the order of thousands of samples.

Deep sequencing with next-generation sequencing technology refers to the generation of reads that map to the same region multiple times—sometimes hundreds of times or more. Since every read was generated from a single DNA molecule, deep sequencing allows the detection of clones comprising as little as 1% of the original sample. Sequence reads from infiltrating normal tissue can easily be identified by comparing the sequences from the tumor and adjacent normal tissue from the same individual. The optimal read depth will vary depending on the cancer type and the sensitivity required, although a typical current recommendation is a minimum of 40-fold coverage for normal genomes and 80-fold for cancer genomes. When tumors are highly heterogeneous it may take several biopsies from different sites in the tumor to represent all the cell types.[69]

---

[65] Landau D. A., Carter S. L., Stojanov P., McKenna A., Stevenson K., et al. (2013) Evolution and impact of subclonal mutations in chronic lymphocytic leukemia. Cell 152: 714-726
[66] Mardis E. R. (2012) Genome sequencing and cancer. Curr Opin Genet Dev 22: 245-250
[67] Gerlinger M., Rowan A. J., Horswell S., Larkin J., Endesfelder D., et al. (2012) Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. N Engl J Med 366: 883-892
[68] Ellis M. J. and Perou C. M. (2013) The genomic landscape of breast cancer as a therapeutic roadmap. Cancer Discov 3: 27-34
[69] Gerlinger M., Rowan A. J., Horswell S., Larkin J., Endesfelder D., et al. (2012) Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. N Engl J Med 366: 883-892

adjacent normal tissue   A A C A C A T G A A C A C A T G

tumor sample

   A A C A C A T G A A C A C A T G
   A A G A C A T G A A C A C A T G   *normal*
   A A C A C A T T A A C A C A T G
   A A C A C A T T A A C A C C T G
   A A C A C A T T A T C A C A T G
   A A C A C A T T A A C A C A T G   *major clone*
   A A C A C A T T A A C A C A T G
   A T C A G A T T A A C A C A T G
   A T C A G A T T A A C A C A T G
   A T C A G A T T A A C A C A T G   *minor clone*

polyclonal tumor
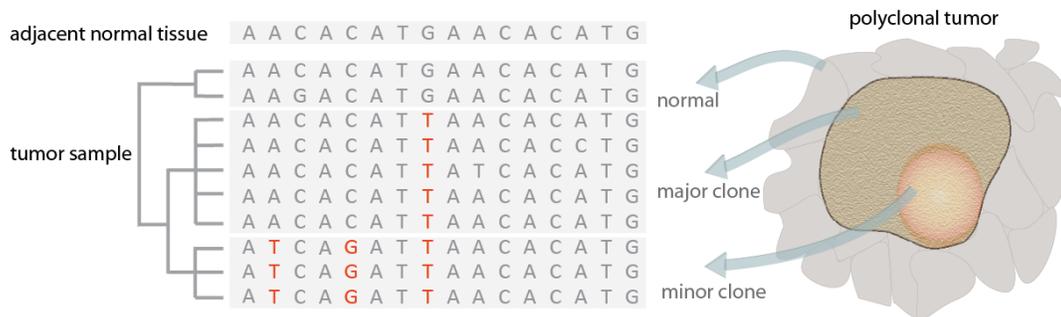
A hypothetical example of a tumor with two cancer clones and contaminating adjacent tissue. The sequences produced by the normal cells in the tumor sample (top two sequences in the tumor alignment) can be identified by comparison to the sequence produced by the adjacent normal tissue. The remaining sequences in the tumor sample can be separated into two groups that represent the major and minor tumor clones. Minor clones, if left untreated, may become major components of the tumor upon relapse. In an actual analysis, the tumor sample will have at least 40-fold coverage and cover targeted sets of genes, whole exomes, or the whole genome.

There are three general approaches to detect somatic mutations in the cancer genome: whole-genome sequencing, whole-exome sequencing, and targeted gene sequencing. The table below contains a brief summary of the advantages and disadvantages of the respective approaches. In a comparison between whole-genome and exome sequencing in a multiple myeloma study, half of all the protein-coding mutations occurred via chromosomal aberrations such as translocations, most of which would not have been discovered by exome sequencing alone.[70] Targeted resequencing is a useful technique to catalog variants of known cancer-related genes in very large cohorts. In the long run, as knowledge of the genome grows and the ability to handle and interpret the large data sets improves, whole-genome sequencing will clearly be the optimal approach for molecular characterization of tumors. In the immediate future, targeted gene sequencing can map drugs already on the market to patients who can derive immediate benefit from them.[71-72]

---

[70] Chapman M. A., Lawrence M. S., Keats J. J., Cibulskis K., Sougnez C., et al. (2011) Initial genome sequencing and analysis of multiple myeloma. Nature 471: 467-472

[71] Holbrook J. D., Parker J. S., Gallagher K. T., Halsey W. S., Hughes A. M., et al. (2011) Deep sequencing of gastric carcinoma reveals somatic mutations relevant to personalized medicine. J Transl Med 9: 119

[72] Lipson D., Capelletti M., Yelensky R., Otto G., Parker A., et al. (2012) Identification of new ALK and RET gene fusions from colorectal and lung cancer biopsies. Nat Med 18: 382-384

Approach	Advantages	Disadvantages
Whole-genome sequencing	- Comprehensive view of whole genome - Can detect all types of mutations including structural variants - Standardized processing and analysis for all patients and all tumor types - Does not require any prior knowledge of the disease	- More expensive - Large dataset presents a challenge for data management, analysis and interpretation - Findings may not be actionable - Risk of incidental findings - Shallow sequencing less sensitive than targeted approaches
Whole-exome sequencing	- About half the cost of whole-genome sequencing - Small data set is easier to manage, analyze and interpret - Standardized processing and analysis for all patients and all tumor types - Will detect indels, SNPs and CNVs - Does not require any prior knowledge of the disease - Provides deep sequencing with good sensitivity for rare clones	- Only 1.5% of the genome is sequenced - May miss fusion genes and oncogenes[73] - Findings may not be actionable - Risk of incidental findings
Targeted gene sequencing	- Cost-effective - Results are easy to interpret - Findings actionable for cancer-relevant genes - Very deep sequencing with very high sensitivity for rare clones	- Will miss many mutations - Requires a prior knowledge of the genes of interest - Delays diagnosis of patients with rare tumors that are not represented on the panel

---

[73] Chapman M. A., Lawrence M. S., Keats J. J., Cibulskis K., Sougnez C., et al. (2011) Initial genome sequencing and analysis of multiple myeloma. Nature 471: 467-472

Whole-Genome Resequencing

Whole-genome sequencing of tumor-normal pair samples provides a comprehensive picture of all the unique mutations present in the tumor. It has become relatively inexpensive and fast to sequence complete genomes and it is an excellent choice for hypothesis-free discovery applications.

References

Sausen M., Leary R. J., Jones S., Wu J., Reynolds C. P., et al. (2013) Integrated genomic analyses identify ARID1A and ARID1B alterations in the childhood cancer neuroblastoma. Nat Genet 45: 12-17

Six samples were analyzed by both exome and high-coverage whole-genome sequencing, permitting independent validation of the somatic alterations as well as a comparison of these approaches for the detection of sequence alterations. Over 91% of the whole-genome and 94% of the whole-exome targeted bases were represented by at least ten reads (Supplementary Tables 2 and 3). A total of 245 somatic alterations in coding regions were detected by either approach, with 219 mutations identified by whole-genome sequencing and 240 alterations identified by whole-exome sequencing. Exomic and genomic sequencing detected 98% and 89% of the mutations, respectively, consistent with similar comparisons made by others.

Illumina Technology: Genome Analyzer$_{IIx}$ System and HiSeq™ System with 31-fold and 94-fold coverage for whole-genome and exome sequencing with 100 bp and 200 bp paired-end reads

Roberts K. G., Morin R. D., Zhang J., Hirst M., Zhao Y., et al. (2012) Genetic alterations activating kinase and cytokine receptor signaling in high-risk acute lymphoblastic leukemia. Cancer Cell 22: 153-166

The authors performed whole-genome sequencing and mRNA-Seq on 15 cases of Ph-like acute lymphoblastic leukemia (ALL). This is a high-risk B-progenitor ALL with a gene expression profile similar to BCR-ABL1-positive ALL. They found activating mutations of IL7R and FLT3, and deletion of SH2B3 (which encodes the JAK2-negative regulator LNK). Several of these alterations were attenuated with tyrosine kinase inhibitors, suggesting that the treatment outcome of these patients may be improved with targeted therapy.

Illumina Technology: Genome Analyzer$_{IIx}$ System and HiSeq 2000 system whole-genome sequencing, mRNA-Seq

Zhang, J., Ding L., Holmfeldt L., Wu G., Heatley S. L., et al. (2012) The genetic basis of early T-Cell precusor acute lymphoblastic leukaemia. Nature 481: 157-163

Early T-cell precursor acute lymphoblastic leukemia (ETP ALL) is a rare and aggressive malignancy of unknown genetic basis. The authors performed whole-genome sequencing of matched normal and leukemic samples from 12 ETP ALL cases and determined the frequency of somatic mutations in a separate cohort of 52 ETP and 42 non-ETP childhood T-ALL cases. The mutational spectrum is similar to that of myeloid tumors. The global transcriptional profile of ETP ALL was also similar to that of normal and myeloid leukemia hematopoietic stem cells. These findings suggest that addition of myeloid-directed therapies might improve the poor outcome of ETP ALL. In a study such as this, where there are few samples and the genetic alterations are unknown, whole-genome sequencing is a good tool to find genetic alterations.

Illumina Technology: Genome Analyzer$_{IIx}$ System for 101 bp paired end reads

Parker S. C., Gartner J., Cardenas-Navia I., Wei X., Ozel Abaan H., et al. (2012) Mutational signatures of de-differentiation in functional non-coding regions of melanoma genomes. PLoS Genet 8: e1002871

Turajlic S., Furney S. J., Lambros M. B., Mitsopoulos C., Kozarewa I., et al. (2012) Whole genome sequencing of matched primary and metastatic acral melanomas. Genome Res 22: 196-207

Welch J. S., Ley T. J., Link D. C., Miller C. A., Larson D. E., et al. (2012) The origin and evolution of mutations in acute myeloid leukemia. Cell 150: 264-278

Exome Sequencing

Exome sequencing focuses only on the 1% to 2% of the genome that codes for proteins and is therefore less expensive to run and simpler to analyze. There have been many notable successes using this approach on Mendelian diseases.[74-75] Although it produces only one-fiftieth of the whole-genome sequence, the cost saving is only half, due to the more expensive and labor-intensive processing of the genetic material.[76] In cancer research, where gross genomic rearrangements are common, exome sequencing may miss key mutations.[77]

References

Tzoneva G., Perez-Garcia A., Carpenter Z., Khiabanian H., Tosello V., et al. (2013) Activating mutations in the NT5C2 nucleotidase gene drive chemotherapy resistance in relapsed ALL. Nat Med 19: 368-371

The authors found mutations in NT5C2 in matched diagnostic remission and relapsed DNA samples from five patients with T-ALL. In an extended panel of 98 relapse T-ALL and 35 relapse B-precursor ALL samples, 17 harbored mutations in NT5C2. This gene encodes a 5'-nucleotidase enzyme that is responsible for the inactivation of nucleoside-analog chemotherapy drugs. *In vitro*, the encoded protein increased in nucleotidase activity in ALL lymphoblasts and conferred resistance to chemotherapy with 6-mercaptopurine and 6-thioguanine.

Illumina Technology: HiSeq 2000 system 100 bp paired-end reads from whole-exome captured DNA

De Keersmaecker K., Atak Z. K., Li N., Vicente C., Patchett S., et al. (2013) Exome sequencing identifies mutation in CNOT3 and ribosomal genes RPL5 and RPL10 in T-cell acute lymphoblastic leukemia. Nat Genet 45: 186-190

The authors detected protein-altering mutations in 508 genes in 67 T-ALL cases. They identified CNOT3 as a tumor suppressor mutated in 7 of 89 (7.9%) adult T-ALLs and mutations affecting the ribosomal proteins RPL5 and RPL10 in 12 of 122 (9.8%) pediatric T-ALLs. The study is a good example of sequencing followed up with biological validation.

Illumina Technology: HiSeq 2000 system at 2 × 100 bp paired-end exome sequencing

Krauthammer M., Kong Y., Ha B. H., Evans P., Bacchiocchi A., et al. (2012) Exome sequencing identifies recurrent somatic RAC1 mutations in melanoma. Nat Genet 44: 1006-1014

In this study the authors used exome sequencing, RNA-Seq, CNV, and loss of heterozygosity (LOH) to examine 147 melanoma samples. They found that sun-exposed melanomas had markedly more ultraviolet (UV)-like C>T somatic mutations compared to sun-shielded melanomas. Apart from mutations in BRAF or NRAS, they identified a recurrent activating mutation in RAC1 in 9.2% of sun-exposed melanomas.

Illumina Technology: Genome Analyzer$_{IIx}$ system and HiSeq 2000 system exome sequencing with 75-bp paired-end reads, RNA-Seq libraries

---

[74] Gilissen C., Hoischen A., Brunner H. G. and Veltman J. A. (2012) Disease gene identification strategies for exome sequencing. Eur J Hum Genet 20: 490-497

[75] Majewski J., Schwartzentruber J., Lalonde E., Montpetit A. and Jabado N. (2011) What can exome sequencing do for you? J Med Genet 48: 580-589

[76] Waters H. (2012) New NIH genetics center focuses its lens on exome, despite doubts. Nat Med 18: 8

[77] Chapman M. A., Lawrence M. S., Keats J. J., Cibulskis K., Sougnez C., et al. (2011) Initial genome sequencing and analysis of multiple myeloma. Nature 471: 467-472

Piazza R., Valletta S., Winkelmann N., Redaelli S., Spinelli R., et al. (2013) Recurrent SETBP1 mutations in atypical chronic myeloid leukemia. Nat Genet 45: 18-24

The authors performed exome sequencing of eight atypical chronic myeloid leukemia (aCML) cases and identified somatic alterations of SETBP1 in two cases. Targeted resequencing of 70 aCMLs, 574 diverse hematological malignancies, and 344 cancer cell lines identified SETBP1 mutations in 24 cases. It appears that SETBP1 mutations are present in approximately one-quarter of aCML cases, where they confer a worse clinical course. Expression of mutant SETBP1 Gly870Ser in the TF1 cell line resulted in higher SETBP1 protein levels, SET protein stabilization, PP2A inhibition, and higher proliferation rates.

Illumina Technology: Genome Analyzer$_{IIx}$ system with 76-bp paired-end reads using Illumina TruSeq SBS kit v5. Exome sequencing with TruSeq Exome Enrichment kit. RNA-Seq with TruSeq RNA Sample Preparation kit

Robinson D. R., Wu Y. M., Kalyana-Sundaram S., Cao X., Lonigro R. J., et al. (2013) Identification of recurrent NAB2-STAT6 gene fusions in solitary fibrous tumor by integrative sequencing. Nat Genet 45: 180-185

Jones D. T., Jager N., Kool M., Zichner T., Hutter B., et al. (2012) Dissecting the genomic complexity underlying medulloblastoma. Nature 488: 100-105

Targeted Resequencing

Targeted resequencing focuses on a restricted set of genes that was compiled based on some prior knowledge. By using only cancer-relevant genes, the results are relatively easy to interpret and potentially actionable. A panel that contains the appropriate genes could be used on different cancer types to streamline laboratory processing and data interpretation. Larger studies in the future may show potential stratification of the patients according to disease progression, genetic profile, environmental exposure, or other factors. The studies to date indicate that this approach may have significant potential as a diagnostic tool.

References

Lipson D., Capelletti M., Yelensky R., Otto G., Parker A., et al. (2012) Identification of new ALK and RET gene fusions from colorectal and lung cancer biopsies. Nat Med 18: 382-384

The authors targeted 145 cancer-relevant genes in 40 colorectal cancer and 24 non–small cell lung cancer formalin-fixed paraffin-embedded (FFPE) samples. Of the samples tested, 59% contained mutations represented in this panel. The remaining patients who did not have mutations in the genes represented on the panel would be good candidates for whole-genome sequencing to potentially expand the cancer-relevant gene panel. A small group of genes represent the majority of mutations, but the diversity of remaining mutations is remarkable. It clearly indicates the benefit of a molecular diagnosis to treat each of the patients appropriately.

Illumina Technology: HiSeq 2000 system with 36 bp paired-end reads to average depth of 229 fold

Ruark E., Snape K., Humburg P., Loveday C., Bajrami I., et al. (2013) Mosaic PPM1D mutations are associated with predisposition to breast and ovarian cancer. Nature 493: 406–410

The authors used targeted sequencing of pooled samples to find rare variants. They sequenced the 507 genes implicated in the repair of DNA in 1,150 primary samples and a 13,642-individual replication experiment. They found protein-truncating variants (PTVs) in the p53-inducible protein phosphatase PPM1D associated with predisposition to breast cancer and ovarian cancer. Interestingly, these truncations have a gain-of-function effect. This approach may be generally useful for rare or mosaic genetic variants.

Illumina Technology: HiSeq 2000 system with a minimum coverage of 4803 per pool of 24 individuals. The sensitivity of base substitution calling was estimated at 99.6%. MiSeq sequencing of PCR amplicons to a median coverage of 3,3873 across the PPM1D mutation to confirm the mutations.

Wagle N., Berger M. F., Davis M. J., Blumenstiel B., Defelice M., et al. (2012) High-throughput detection of actionable genomic alterations in clinical tumor samples by targeted, massively parallel sequencing. Cancer Discov 2: 82-93

FFPE Samples

Tissue samples are commonly stored as formalin-fixed, paraffin-embedded (FFPE) preparations. In some cases these may be the only samples available.[78] FFPE samples yield relatively short DNA fragments, but they can deliver excellent results when the appropriate care is taken during collection, embedding, and storage.[79] In addition, the method of nucleic acid isolation and sample preparation method for next-generation sequencing are also critical factors in the success of analyzing FFPE-derived samples.

The following table provides examples of the successful application of Illumina sequencing technology to FFPE samples for various applications.

Reference	Material	Target
Fanelli et.al[80]	Whole Genome	ChIP-Seq
Gu et al.[81]	Whole Genome	Methylation (Bs-Seq)
Kerick et al.[82]	Targeted Regions	Single Nucleotide Variants (SNV)
Lipson et al.[83]	Targeted Regions	Exome Sequencing
Schweiger et al.[84]	Whole Genome	Copy Number Variants (CNV)
Wagle et al.[85]	Targeted Regions	Exome Sequencing
Weng et al.[86]	RNA	miRNA

References

Hadd A. G., Houghton J., Choudhary A., Sah S., Chen L., et al. (2013) Targeted, high-depth, next-generation sequencing of cancer genes in formalin-fixed, paraffin-embedded and fine-needle aspiration tumor specimens. J Mol Diagn 15: 234-247

Corless C. L. and Spellman P. T. (2012) Tackling formalin-fixed, paraffin-embedded tumor tissue with next-generation sequencing. Cancer Discov 2: 23-24

Morlan J. D., Qu K. and Sinicropi D. V. (2012) Selective depletion of rRNA enables whole transcriptome profiling of archival fixed tissue. PLoS ONE 7: e42882

Sinicropi D., Qu K., Collin F., Crager M., Liu M. L., et al. (2012) Whole transcriptome RNA-Seq analysis of breast cancer recurrence risk using formalin-fixed paraffin-embedded tumor tissue. PLoS ONE 7: e40092

[78] Corless C. L. and Spellman P. T. (2012) Tackling formalin-fixed, paraffin-embedded tumor tissue with next-generation sequencing. Cancer Discov 2: 23-24

[79] Yost S. E., Smith E. N., Schwab R. B., Bao L., Jung H., et al. (2012) Identification of high-confidence somatic mutations in whole genome sequence of formalin-fixed breast cancer specimens. Nucleic Acids Res 40: e107

[80] Fanelli M., Amatori S., Barozzi I. and Minucci S. (2011) Chromatin immunoprecipitation and high-throughput sequencing from paraffin-embedded pathology tissue. Nat Protoc 6: 1905-1919

[81] Gu H., Bock C., Mikkelsen T. S., Jager N., Smith Z. D., et al. (2010) Genome-scale DNA methylation mapping of clinical samples at single-nucleotide resolution. Nat Methods 7: 133-136

[82] Kerick M., Isau M., Timmermann B., Sultmann H., Herwig R., et al. (2011) Targeted high throughput sequencing in clinical cancer settings: formaldehyde fixed-paraffin embedded (FFPE) tumor tissues, input amount and tumor heterogeneity. BMC Med Genomics 4: 68

[83] Lipson D., Capelletti M., Yelensky R., Otto G., Parker A., et al. (2012) Identification of new ALK and RET gene fusions from colorectal and lung cancer biopsies. Nat Med 18: 382-384

[84] Schweiger M. R., Kerick M., Timmermann B., Albrecht M. W., Borodina T., et al. (2009) Genome-wide massively parallel sequencing of formaldehyde fixed-paraffin embedded (FFPE) tumor tissues for copy-number- and mutation-analysis. PLoS ONE 4: e5548

[85] Wagle N., Berger M. F., Davis M. J., Blumenstiel B., DeFelice M., et al. (2011) High-Throughput Detection of Actionable Genomic Alterations in Clinical Tumor Samples by Targeted, Massively Parallel Sequencing. Cancer Discovery 2: 82-93

[86] Weng L., Wu X., Gao H., Mu B., Li X., et al. (2010) MicroRNA profiling of clear cell renal cell carcinoma by whole-genome small RNA deep sequencing of paired frozen and formalin-fixed, paraffin-embedded tissue specimens. J Pathol 222: 41-51

Single-Cell and Low-Abundance Samples

Tumors commonly contain several clonal populations that reflect the ongoing accumulation of mutations.[87-88] Single-cell genomic methods have the capacity to resolve these complex mixtures of cells. Molecular assays of tissues reflect an average signal of the population, or alternatively only the dominant clone, which may not be the most malignant clone present in the tumor.[89]

The clinical value of single-cell genomic methods will be in profiling scarce cancer cells, monitoring, and detecting rare clones that may be resistant to chemotherapy. These applications are likely to improve all three major themes of oncology: detection, progression, and prediction of therapeutic efficacy.[90]

Review

Swanton C. (2013) Plasma-derived tumor DNA analysis at whole-genome resolution. Clin Chem 59: 6-8

References

Hou Y., Song L., Zhu P., Zhang B., Tao Y., et al. (2012) Single-cell exome sequencing and monoclonal evolution of a JAK2-negative myeloproliferative neoplasm. Cell 148: 873-885

This pilot study shows the initial characterization of the cancer genome at the single-cell, single-nucleotide level.

Illumina Technology: HiSeq 2000 system with 100 bp paired-end reads and 1M genotyping arrays

Falconer E., Hills M., Naumann U., Poon S. S., Chavez E. A., et al. (2012) DNA template strand sequencing of single-cells maps genomic rearrangements at high resolution. Nat Methods 9: 1107-1112

The authors show that the sequencing of individual parental DNA template strands can map sister chromatid exchanges (sces) at orders-of-magnitude greater resolution than was previously possible.

Illumina Technology: Genome Analyzer$_{IIx}$ system or HiSeq 2000 system 76 bp paired-end reads

Baslan T., Kendall J., Rodgers L., Cox H., Riggs M., et al. (2012) Genome-wide copy number analysis of single cells. Nat Protoc 7: 1024-1041

Falconer E., Hills M., Naumann U., Poon S. S., Chavez E. A., et al. (2012) DNA template strand sequencing of single-cells maps genomic rearrangements at high resolution. Nat Methods 9: 1107-1112

Hardt O., Wild S., Oerlecke I., Hofmann K., Luo S., et al. (2012) Highly sensitive profiling of CD44+/CD24- breast cancer stem cells by combining global mRNA amplification and next generation sequencing: evidence for a hyperactive PI3K pathway. Cancer Lett 325: 165-174

Xu X., Hou Y., Yin X., Bao L., Tang A., et al. (2012) Single-cell exome sequencing reveals single-nucleotide mutation characteristics of a kidney tumor. Cell 148: 886-895

---

[87] Stephens P. J., Greenman C. D., Fu B., Yang F., Bignell G. R., et al. (2011) Massive genomic rearrangement acquired in a single catastrophic event during cancer development. Cell 144: 27-40
[88] Navin N., Kendall J., Troge J., Andrews P., Rodgers L., et al. (2011) Tumour evolution inferred by single-cell sequencing. Nature 472: 90-94
[89] Navin N. E. and Hicks J. (2010) Tracing the tumor lineage. Mol Oncol 4: 267-283
[90] Navin N. and Hicks J. (2011) Future medical applications of single-cell sequencing in cancer. Genome Med 3: 31

Stem Cells and Cell Cultures

Cultured cell lines are versatile tools to study biological processes. When cell lines are cultured over a long period of time, they can be expected to accumulate mutations. Also, the genetic bottlenecks that result from population reduction during the culturing process can significantly accelerate the accumulation of mutations.[91] In a sample containing multiple clones, culturing can skew the relative abundance of the clones. Recent studies used deep sequencing to show that as many as 50% of the genomic variants that are seen in cell lines derive from low-frequency somatic genomic variants that are present in the parental fibroblasts.[92] Next-generation sequencing provides a highly sensitive and precise tool that can help researchers distinguish between authentic biology and artifacts from the immortalization and propagation process to make these versatile models even more effective.[93]

References

Abyzov A., Mariani J., Palejev D., Zhang Y., Haney M. S., et al. (2012) Somatic copy number mosaicism in human skin revealed by induced pluripotent stem cells. Nature 492: 438-442

The authors find that, on average, an induced pluripotent stem cell (iPSC) line manifests two CNVs not apparent in the fibroblasts from which the iPSC was derived. They show that at least 50% of those CNVs are present as low-frequency somatic genomic variants in parental fibroblasts. They go on to estimate that approximately 30% of the fibroblast cells have somatic CNVs in their genomes, suggesting widespread somatic mosaicism in the human body.

Illumina Technology: HiSeq 2000 system whole-genome paired-end sequencing and HumanHT-12 v4 BeadChip

Kaur H., Mao S., Li Q., Sameni M., Krawetz S. A., et al. (2012) RNA-Seq of human breast ductal carcinoma in situ models reveals aldehyde dehydrogenase isoform 5A1 as a novel potential target. PLoS ONE 7: e50249

The authors determined the expression changes that are common among three ductal carcinoma *in situ* (DCIS) models (MCF10.DCIS, SUM102 and SUM225) compared to the MCF10A model of nontumorigenic mammary epithelial cells in three-dimensional (3D) overlay culture. They found differentially expressed genes encoding for proteins that are associated with a number of signaling pathways.

Illumina Technology: Genome Analyzer$_{IIx}$ system for 76 cycles of single-end sequencing

---

[91] Gisselsson D., Lindgren D., Mengelbier L. H., Ora I. and Yeger H. (2010) Genetic bottlenecks and the hazardous game of population reduction in cell line based research. Exp Cell Res 316: 3379-3386

[92] Abyzov A., Mariani J., Palejev D., Zhang Y., Haney M. S., et al. (2012) Somatic copy number mosaicism in human skin revealed by induced pluripotent stem cells. Nature 492: 438-442

[93] Nishi M., Sakai Y., Akutsu H., Nagashima Y., Quinn G., et al. (2013) Induction of cells with cancer stem cell properties from nontumorigenic human mammary epithelial cells by defined reprogramming factors. Oncogene

Leich E., Weissbach S., Klein H. U., Grieb T., Pischimarov J., et al. (2013) Multiple myeloma is affected by multiple and heterogeneous somatic mutations in adhesion- and receptor tyrosine kinase signaling molecules. Blood Cancer J 3: e102

Nishi M., Sakai Y., Akutsu H., Nagashima Y., Quinn G., et al. (2013) Induction of cells with cancer stem cell properties from nontumorigenic human mammary epithelial cells by defined reprogramming factors. Oncogene

Liu T., Wang Y., Peng X., Zhang L., Cheng J., et al. (2012) Establishment of mouse teratocarcinomas stem cells line and screening genes responsible for malignancy. PLoS ONE 7: e43955

Parker S. C., Gartner J., Cardenas-Navia I., Wei X., Ozel Abaan H., et al. (2012) Mutational signatures of de-differentiation in functional non-coding regions of melanoma genomes. PLoS Genet 8: e1002871

Poell J. B., van Haastert R. J., de Gunst T., Schultz I. J., Gommans W. M., et al. (2012) A functional screen identifies specific microRNAs capable of inhibiting human melanoma cell viability. PLoS ONE 7: e43569

Bibliography

Abyzov A., Mariani J., Palejev D., Zhang Y., Haney M. S., et al. (2012) Somatic copy number mosaicism in human skin revealed by induced pluripotent stem cells. Nature 492: 438-442

Bahn J. H., Lee J. H., Li G., Greer C., Peng G., et al. (2012) Accurate identification of A-to-I RNA editing in human by transcriptome sequencing. Genome Res 22: 142-150

Banerji S., Cibulskis K., Rangel-Escareno C., Brown K. K., Carter S. L., et al. (2012) Sequence analysis of mutations and translocations across breast cancer subtypes. Nature 486: 405-409

Baslan T., Kendall J., Rodgers L., Cox H., Riggs M., et al. (2012) Genome-wide copy number analysis of single cells. Nat Protoc 7: 1024-1041

Bernstein B. E., Meissner A. and Lander E. S. (2007) The mammalian epigenome. Cell 128: 669-681

Bert S. A., Robinson M. D., Strbenac D., Statham A. L., Song J. Z., et al. (2013) Regional activation of the cancer genome by long-range epigenetic remodeling. Cancer Cell 23: 9-22

Boyle P., Clement K., Gu H., Smith Z. D., Ziller M., et al. (2012) Gel-free multiplexed reduced representation bisulfite sequencing for large-scale DNA methylation profiling. Genome Biol 13: R92

Brastianos P. K., Horowitz P. M., Santagata S., Jones R. T., McKenna A., et al. (2013) Genomic sequencing of meningiomas identifies oncogenic SMO and AKT1 mutations. Nat Genet 45: 285-289

Brinkman A. B., Gu H., Bartels S. J., Zhang Y., Matarese F., et al. (2012) Sequential ChIP-bisulfite sequencing enables direct genome-scale investigation of chromatin and DNA methylation cross-talk. Genome Res 22: 1128-1138

Cahill N., Bergh A. C., Kanduri M., Goransson-Kultima H., Mansouri L., et al. (2013) 450K-array analysis of chronic lymphocytic leukemia cells reveals global DNA methylation to be relatively stable over time and similar in resting and proliferative compartments. Leukemia 27: 150-158

Caldas C. (2012) Cancer sequencing unravels clonal evolution. Nat Biotechnol 30: 408-410

Campbell P. J., Yachida S., Mudie L. J., Stephens P. J., Pleasance E. D., et al. (2010) The patterns and dynamics of genomic instability in metastatic pancreatic cancer. Nature 467: 1109-1113

Caren H., Pollard S. M. and Beck S. (2012) The good, the bad and the ugly: Epigenetic mechanisms in glioblastoma. Mol Aspects Med Carter S. L., Cibulskis K., Helman E., McKenna A., Shen H., et al. (2012) Absolute quantification of somatic DNA alterations in human cancer. Nat Biotechnol 30: 413-421

Chapman M. A., Lawrence M. S., Keats J. J., Cibulskis K., Sougnez C., et al. (2011) Initial genome sequencing and analysis of multiple myeloma. Nature 471: 467-472

Chen D., Sun Y., Wei Y., Zhang P., Rezaeian A. H., et al. (2012) LIFR is a breast cancer metastasis suppressor upstream of the Hippo-YAP pathway and a prognostic marker. Nat Med 18: 1511-1517

Chen L., Li Y., Lin C. H., Chan T. H., Chow R. K., et al. (2013) Recoding RNA editing of AZIN1 predisposes to hepatocellular carcinoma. Nat Med 19: 209-216

Cheng W. C., Chung I. F., Huang T. S., Chang S. T., Sun H. J., et al. (2013) YM500: a small RNA sequencing (smRNA-seq) database for microRNA research. Nucleic Acids Res 41: D285-294

Chiang C., Jacobsen J. C., Ernst C., Hanscom C., Heilbut A., et al. (2012) Complex reorganization and predominant non-homologous repair following chromosomal breakage in karyotypically balanced germline rearrangements and transgenic integration. Nat Genet 44: 390-397, S391

Cibulskis K., Lawrence M. S., Carter S. L., Sivachenko A., Jaffe D., et al. (2013) Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. Nat Biotechnol 31: 213-219

Corless C. L. and Spellman P. T. (2012) Tackling formalin-fixed, paraffin-embedded tumor tissue with next-generation sequencing. Cancer Discov 2: 23-24

Cowper-Sal lari R., Zhang X., Wright J. B., Bailey S. D., Cole M. D., et al. (2012) Breast cancer risk-associated SNPs modulate the affinity of chromatin for FOXA1 and alter gene expression. Nat Genet 44: 1191-1198

De Keersmaecker K., Atak Z. K., Li N., Vicente C., Patchett S., et al. (2013) Exome sequencing identifies mutation in CNOT3 and ribosomal genes RPL5 and RPL10 in T-cell acute lymphoblastic leukemia. Nat Genet 45: 186-190

Dedeurwaerder S., Defrance M., Calonne E., Denis H., Sotiriou C., et al. (2011) Evaluation of the Infinium Methylation 450K technology. Epigenomics 3: 771-784

Ding L., Ley T. J., Larson D. E., Miller C. A., Koboldt D. C., et al. (2012) Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. Nature 481: 506-510

Dreyfuss G., Matunis M. J., Pinol-Roma S. and Burd C. G. (1993) hnRNP proteins and the biogenesis of mRNA. Annu Rev Biochem 62: 289-321

Duncan C. G., Barwick B. G., Jin G., Rago C., Kapoor-Vazirani P., et al. (2012) A heterozygous IDH1R132H/WT mutation induces genome-wide alterations in DNA methylation. Genome Res 22: 2339-2355

Easwaran H. and Baylin S. B. (2013) Epigenetic abnormalities in cancer find a "home on the range". Cancer Cell 23: 1-3

Ellis M. J., Ding L., Shen D., Luo J., Suman V. J., et al. (2012) Whole-genome analysis informs breast cancer response to aromatase inhibition. Nature 486: 353-360

Ellis M. J. and Perou C. M. (2013) The genomic landscape of breast cancer as a therapeutic roadmap. Cancer Discov 3: 27-34

Enfield K. S., Pikor L. A., Martinez V. D. and Lam W. L. (2012) Mechanistic Roles of Noncoding RNAs in Lung Cancer Biology and Their Clinical Implications. Genet Res Int 2012: 737416

Eswaran J., Cyanam D., Mudvari P., Reddy S. D., Pakala S. B., et al. (2012) Transcriptomic landscape of breast cancers through mRNA sequencing. Sci Rep 2: 264

Falconer E., Hills M., Naumann U., Poon S. S., Chavez E. A., et al. (2012) DNA template strand sequencing of single-cells maps genomic rearrangements at high resolution. Nat Methods 9: 1107-1112

Fanelli M., Amatori S., Barozzi I. and Minucci S. (2011) Chromatin immunoprecipitation and high-throughput sequencing from paraffin-embedded pathology tissue. Nat Protoc 6: 1905-1919

Feng H., Qin Z. and Zhang X. (2012) Opportunities and methods for studying alternative splicing in cancer with RNA-Seq. Cancer Lett
Forment J. V., Kaidi A. and Jackson S. P. (2012) Chromothripsis and cancer: causes and consequences of chromosome shattering. Nat Rev Cancer 12: 663-670

Forsberg L. A., Rasi C., Razzaghian H. R., Pakalapati G., Waite L., et al. (2012) Age-related somatic structural changes in the nuclear genome of human blood cells. Am J Hum Genet 90: 217-228

Frampton A. E., Gall T. M., Castellano L., Stebbing J., Jiao L. R., et al. (2013) Towards a clinical use of miRNAs in pancreatic cancer biopsies. Expert review of molecular diagnostics 13: 31-34

Frenkel-Morgenstern M., Lacroix V., Ezkurdia I., Levin Y., Gabashvili A., et al. (2012) Chimeras taking shape: potential functions of proteins encoded by chimeric RNA transcripts. Genome Res 22: 1231-1242

Frommer M., McDonald L. E., Millar D. S., Collis C. M., Watt F., et al. (1992) A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. Proc Natl Acad Sci U S A 89: 1827-1831

Furey T. S. (2012) ChIP-seq and beyond: new and improved methodologies to detect and characterize protein-DNA interactions. Nat Rev Genet 13: 840-852

Garraway L. A. and Lander E. S. (2013) Lessons from the cancer genome. Cell 153: 17-37

Gerlinger M., Rowan A. J., Horswell S., Larkin J., Endesfelder D., et al. (2012) Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. N Engl J Med 366: 883-892

Gilissen C., Hoischen A., Brunner H. G. and Veltman J. A. (2012) Disease gene identification strategies for exome sequencing. Eur J Hum Genet 20: 490-497

Gisselsson D., Lindgren D., Mengelbier L. H., Ora I. and Yeger H. (2010) Genetic bottlenecks and the hazardous game of population reduction in cell line based research. Exp Cell Res 316: 3379-3386

Govindan R., Ding L., Griffith M., Subramanian J., Dees N. D., et al. (2012) Genomic landscape of non-small cell lung cancer in smokers and never-smokers. Cell 150: 1121-1134

Greenman C. D., Pleasance E. D., Newman S., Yang F., Fu B., et al. (2012) Estimation of rearrangement phylogeny for cancer genomes. Genome Res 22: 346-361

Gu H., Bock C., Mikkelsen T. S., Jager N., Smith Z. D., et al. (2010) Genome-scale DNA methylation mapping of clinical samples at single-nucleotide resolution. Nat Methods 7: 133-136

Guo H., Ingolia N. T., Weissman J. S. and Bartel D. P. (2010) Mammalian microRNAs predominantly act to decrease target mRNA levels. Nature 466: 835-840

Guo Y., Monahan K., Wu H., Gertz J., Varley K. E., et al. (2012) CTCF/cohesin-mediated DNA looping is required for protocadherin alpha promoter choice. Proc Natl Acad Sci U S A 109: 21081-21086

Ha G., Roth A., Lai D., Bashashati A., Ding J., et al. (2012) Integrative analysis of genome-wide loss of heterozygosity and monoallelic expression at nucleotide resolution reveals disrupted pathways in triple-negative breast cancer. Genome Res 22: 1995-2007

Hadd A. G., Houghton J., Choudhary A., Sah S., Chen L., et al. (2013) Targeted, high-depth, next-generation sequencing of cancer genes in formalin-fixed, paraffin-embedded and fine-needle aspiration tumor specimens. J Mol Diagn 15: 234-247

Hakim O., Resch W., Yamane A., Klein I., Kieffer-Kwon K. R., et al. (2012) DNA damage defines sites of recurrent chromosomal translocations in B lymphocytes. Nature 484: 69-74

Han S. P., Tang Y. H. and Smith R. (2010) Functional diversity of the hnRNPs: past, present and perspectives. Biochem J 430: 379-392

Hardt O., Wild S., Oerlecke I., Hofmann K., Luo S., et al. (2012) Highly sensitive profiling of CD44+/CD24- breast cancer stem cells by combining global mRNA amplification and next generation sequencing: evidence for a hyperactive PI3K pathway. Cancer Lett 325: 165-174

Hermeking H. (2012) MicroRNAs in the p53 network: micromanagement of tumour suppression. Nat Rev Cancer 12: 613-626

Holbrook J. D., Parker J. S., Gallagher K. T., Halsey W. S., Hughes A. M., et al. (2011) Deep sequencing of gastric carcinoma reveals somatic mutations relevant to personalized medicine. J Transl Med 9: 119

Holmfeldt L., Wei L., Diaz-Flores E., Walsh M., Zhang J., et al. (2013) The genomic landscape of hypodiploid acute lymphoblastic leukemia. Nat Genet 45: 242-252

Hoque M., Ji Z., Zheng D., Luo W., Li W., et al. (2013) Analysis of alternative cleavage and polyadenylation by 3' region extraction and deep sequencing. Nat Methods 10: 133-139

Hou Y., Song L., Zhu P., Zhang B., Tao Y., et al. (2012) Single-cell exome sequencing and monoclonal evolution of a JAK2-negative myeloproliferative neoplasm. Cell 148: 873-885

Hsieh A. C., Liu Y., Edlind M. P., Ingolia N. T., Janes M. R., et al. (2012) The translational landscape of mTOR signalling steers cancer initiation and metastasis. Nature 485: 55-61

Huang S., Holzel M., Knijnenburg T., Schlicker A., Roepman P., et al. (2012) MED12 controls the response to multiple cancer drugs through regulation of TGF-beta receptor signaling. Cell 151: 937-950

Huelga S. C., Vu A. Q., Arnold J. D., Liang T. Y., Liu P. P., et al. (2012) Integrative genome-wide analysis reveals cooperative regulation of alternative splicing by hnRNP proteins. Cell Rep 1: 167-178

Ingolia N. T., Ghaemmaghami S., Newman J. R. and Weissman J. S. (2009) Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. Science 324: 218-223

Jacinto F. V., Ballestar E. and Esteller M. (2008) Methyl-DNA immunoprecipitation (MeDIP): hunting down the DNA methylome. Biotechniques 44: 35, 37, 39 passim

Jankovic M., Feldhahn N., Oliveira T. Y., Silva I. T., Kieffer-Kwon K. R., et al. (2013) 53BP1 alters the landscape of DNA rearrangements and suppresses AID-induced B cell lymphoma. Mol Cell 49: 623-631

Jasmine F., Rahaman R., Dodsworth C., Roy S., Paul R., et al. (2012) A genome-wide study of cytogenetic changes in colorectal cancer using SNP microarrays: opportunities for future personalized treatment. PLoS ONE 7: e31968

Jiang Q., Crews L. A., Barrett C. L., Chun H. J., Court A. C., et al. (2013) ADAR1 promotes malignant progenitor reprogramming in chronic myeloid leukemia. Proc Natl Acad Sci U S A 110: 1041-1046

Jithesh P. V., Risk J. M., Schache A. G., Dhanda J., Lane B., et al. (2013) The epigenetic landscape of oral squamous cell carcinoma. Br J Cancer 108: 370-379

Jones D. T., Jager N., Kool M., Zichner T., Hutter B., et al. (2012) Dissecting the genomic complexity underlying medulloblastoma. Nature 488: 100-105

Ju Y. S., Kim J. I., Kim S., Hong D., Park H., et al. (2011) Extensive genomic and transcriptional diversity identified through massively parallel DNA and RNA sequencing of eighteen Korean individuals. Nat Genet 43: 745-752

Ju Y. S., Lee W. C., Shin J. Y., Lee S., Bleazard T., et al. (2012) A transforming KIF5B and RET gene fusion in lung adenocarcinoma revealed from whole-genome and transcriptome sequencing. Genome Res 22: 436-445

Kangaspeska S., Hultsch S., Edgren H., Nicorici D., Murumagi A., et al. (2012) Reanalysis of RNA-sequencing data reveals several additional fusion genes with multiple isoforms. PLoS ONE 7: e48745

Katz Y., Wang E. T., Airoldi E. M. and Burge C. B. (2010) Analysis and design of RNA sequencing experiments for identifying isoform regulation. Nat Methods 7: 1009-1015

Kaur H., Mao S., Li Q., Sameni M., Krawetz S. A., et al. (2012) RNA-Seq of human breast ductal carcinoma in situ models reveals aldehyde dehydrogenase isoform 5A1 as a novel potential target. PLoS ONE 7: e50249

Kerick M., Isau M., Timmermann B., Sultmann H., Herwig R., et al. (2011) Targeted high throughput sequencing in clinical cancer settings: formaldehyde fixed-paraffin embedded (FFPE) tumor tissues, input amount and tumor heterogeneity. BMC Med Genomics 4: 68

Klein I. A., Resch W., Jankovic M., Oliveira T., Yamane A., et al. (2011) Translocation-capture sequencing reveals the extent and nature of chromosomal rearrangements in B lymphocytes. Cell 147: 95-106

Kloosterman W. P. and Plasterk R. H. (2006) The diverse functions of microRNAs in animal development and disease. Dev Cell 11: 441-450

Kohwi-Shigematsu T., Kohwi Y., Takahashi K., Richards H. W., Ayers S. D., et al. (2012) SATB1-mediated functional packaging of chromatin into loops. Methods 58: 243-254

Korbel J. O. and Campbell P. J. (2013) Criteria for inference of chromothripsis in cancer genomes. Cell 152: 1226-1236

Krauthammer M., Kong Y., Ha B. H., Evans P., Bacchiocchi A., et al. (2012) Exome sequencing identifies recurrent somatic RAC1 mutations in melanoma. Nat Genet 44: 1006-1014

Kresse S. H., Rydbeck H., Skarn M., Namlos H. M., Barragan-Polania A. H., et al. (2012) Integrative analysis reveals relationships of genetic and epigenetic alterations in osteosarcoma. PLoS ONE 7: e48262

Landau D. A., Carter S. L., Stojanov P., McKenna A., Stevenson K., et al. (2013) Evolution and impact of subclonal mutations in chronic lymphocytic leukemia. Cell 152: 714-726

Law P. T., Qin H., Ching A. K., Lai K. P., Co N. N., et al. (2013) Deep sequencing of small RNA transcriptome reveals novel non-coding RNAs in hepatocellular carcinoma. J Hepatol

Lee E. J., Luo J., Wilson J. M. and Shi H. (2012) Analyzing the cancer methylome through targeted bisulfite sequencing. Cancer Lett

Leich E., Weissbach S., Klein H. U., Grieb T., Pischimarov J., et al. (2013) Multiple myeloma is affected by multiple and heterogeneous somatic mutations in adhesion- and receptor tyrosine kinase signaling molecules. Blood Cancer J 3: e102

Leidenroth A., Sorte H. S., Gilfillan G., Ehrlich M., Lyle R., et al. (2012) Diagnosis by sequencing: correction of misdiagnosis from FSHD2 to LGMD2A by whole-exome analysis. Eur J Hum Genet 20: 999-1003

Lewis B. P., Burge C. B. and Bartel D. P. (2005) Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell 120: 15-20

Li G., Fullwood M. J., Xu H., Mulawadi F. H., Velkov S., et al. (2010) ChIA-PET tool for comprehensive chromatin interaction analysis with paired-end tag sequencing. Genome Biol 11: R22

Li M., Wang I. X., Li Y., Bruzel A., Richards A. L., et al. (2011) Widespread RNA and DNA sequence differences in the human transcriptome. Science 333: 53-58

Li M., Wang I. X. and Cheung V. G. (2012) Response to Comments on "Widespread RNA and DNA Sequence Differences in the Human Transcriptome". Science 335: 1302-1302

Lian L., Qu L., Chen Y., Lamont S. J. and Yang N. (2012) A systematic analysis of miRNA transcriptome in Marek's disease virus-induced lymphoma reveals novel and differentially expressed miRNAs. PLoS ONE 7: e51003

Liang W. S., Craig D. W., Carpten J., Borad M. J., Demeure M. J., et al. (2012) Genome-wide characterization of pancreatic adenocarcinoma patients using next generation sequencing. PLoS ONE 7: e43192

Lipson D., Capelletti M., Yelensky R., Otto G., Parker A., et al. (2012) Identification of new ALK and RET gene fusions from colorectal and lung cancer biopsies. Nat Med 18: 382-384

Liu J., Zhang Z., Bando M., Itoh T., Deardorff M. A., et al. (2010) Genome-wide DNA methylation analysis in cohesin mutant human cell lines. Nucleic Acids Res 38: 5657-5671

Liu J., Lee W., Jiang Z., Chen Z., Jhunjhunwala S., et al. (2012) Genome and transcriptome sequencing of lung cancers reveal diverse mutational and splicing events. Genome Res 22: 2315-2327

Liu T., Wang Y., Peng X., Zhang L., Cheng J., et al. (2012) Establishment of mouse teratocarcinomas stem cells line and screening genes responsible for malignancy. PLoS ONE 7: e43955

Majewski J., Schwartzentruber J., Lalonde E., Montpetit A. and Jabado N. (2011) What can exome sequencing do for you? J Med Genet 48: 580-589

Mardis E. R. (2012) Genome sequencing and cancer. Curr Opin Genet Dev 22: 245-250

McBride D. J., Etemadmoghadam D., Cooke S. L., Alsop K., George J., et al. (2012) Tandem duplication of chromosomal segments is common in ovarian and breast cancer genomes. J Pathol 227: 446-455

McGettigan P. A. (2013) Transcriptomics in the RNA-seq era. Curr Opin Chem Biol 17: 4-11

Meissner A., Gnirke A., Bell G. W., Ramsahoye B., Lander E. S., et al. (2005) Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. Nucleic Acids Res 33: 5868-5877

Merkin J., Russell C., Chen P. and Burge C. B. (2012) Evolutionary dynamics of gene and isoform regulation in Mammalian tissues. Science 338: 1593-1599

Meyer J. A., Wang J., Hogan L. E., Yang J. J., Dandekar S., et al. (2013) Relapse-specific mutations in NT5C2 in childhood acute lymphoblastic leukemia. Nat Genet 45: 290-294

Morlan J. D., Qu K. and Sinicropi D. V. (2012) Selective depletion of rRNA enables whole transcriptome profiling of archival fixed tissue. PLoS ONE 7: e42882

Navin N. and Hicks J. (2011) Future medical applications of single-cell sequencing in cancer. Genome Med 3: 31

Navin N., Kendall J., Troge J., Andrews P., Rodgers L., et al. (2011) Tumour evolution inferred by single-cell sequencing. Nature 472: 90-94

Navin N. E. and Hicks J. (2010) Tracing the tumor lineage. Mol Oncol 4: 267-283

Ng C. K., Cooke S. L., Howe K., Newman S., Xian J., et al. (2012) The role of tandem duplicator phenotype in tumour evolution in high-grade serous ovarian cancer. J Pathol 226: 703-712

Nik-Zainal S., Alexandrov L. B., Wedge D. C., Van Loo P., Greenman C. D., et al. (2012) Mutational processes molding the genomes of 21 breast cancers. Cell 149: 979-993

Nishi M., Sakai Y., Akutsu H., Nagashima Y., Quinn G., et al. (2013) Induction of cells with cancer stem cell properties from nontumorigenic human mammary epithelial cells by defined reprogramming factors. Oncogene
Ntziachristos P., Tsirigos A., Van Vlierberghe P., Nedjic J., Trimarchi T., et al. (2012) Genetic inactivation of the polycomb repressive complex 2 in T cell acute lymphoblastic leukemia. Nat Med 18: 298-301

Nyquist K. B., Panagopoulos I., Thorsen J., Haugom L., Gorunova L., et al. (2012) Whole-transcriptome sequencing identifies novel IRF2BP2-CDX1 fusion gene brought about by translocation t(1;5)(q42;q32) in mesenchymal chondrosarcoma. PLoS ONE 7: e49705

Palles C., Cazier J. B., Howarth K. M., Domingo E., Jones A. M., et al. (2013) Germline mutations affecting the proofreading domains of POLE and POLD1 predispose to colorectal adenomas and carcinomas. Nat Genet 45: 136-144

Papantonis A., Kohro T., Baboo S., Larkin J. D., Deng B., et al. (2012) TNFalpha signals through specialized factories where responsive coding and miRNA genes are transcribed. EMBO J 31: 4404-4414

Park C., Han S., Lee K. M., Choi J. Y., Song N., et al. (2012) Association between CASP7 and CASP14 genetic polymorphisms and the risk of childhood leukemia. Hum Immunol 73: 736-739

Parker S. C., Gartner J., Cardenas-Navia I., Wei X., Ozel Abaan H., et al. (2012) Mutational signatures of de-differentiation in functional non-coding regions of melanoma genomes. PLoS Genet 8: e1002871

Paul S., Kuo A., Schalch T., Vogel H., Joshua-Tor L., et al. (2013) Chd5 requires PHD-mediated histone 3 binding for tumor suppression. Cell Rep 3: 92-102

Peifer M., Fernandez-Cuesta L., Sos M. L., George J., Seidel D., et al. (2012) Integrative genome analyses identify key somatic driver mutations of small-cell lung cancer. Nat Genet 44: 1104-1110

Peng Z., Cheng Y., Tan B. C., Kang L., Tian Z., et al. (2012) Comprehensive analysis of RNA-Seq data reveals extensive RNA editing in a human transcriptome. Nat Biotechnol 30: 253-260

Perez-Mancera P. A., Rust A. G., van der Weyden L., Kristiansen G., Li A., et al. (2012) The deubiquitinase USP9X suppresses pancreatic ductal adenocarcinoma. Nature 486: 266-270

Persson H., Kvist A., Rego N., Staaf J., Vallon-Christersson J., et al. (2011) Identification of new microRNAs in paired normal and tumor breast tissue suggests a dual role for the ERBB2/Her2 gene. Cancer Res 71: 78-86

Pflueger D., Terry S., Sboner A., Habegger L., Esgueva R., et al. (2011) Discovery of non-ETS gene fusions in human prostate cancer using next-generation RNA sequencing. Genome Res 21: 56-67

Piazza R., Valletta S., Winkelmann N., Redaelli S., Spinelli R., et al. (2013) Recurrent SETBP1 mutations in atypical chronic myeloid leukemia. Nat Genet 45: 18-24

Pleasance E. D., Cheetham R. K., Stephens P. J., McBride D. J., Humphray S. J., et al. (2010) A comprehensive catalogue of somatic mutations from a human cancer genome. Nature 463: 191-196

Poell J. B., van Haastert R. J., de Gunst T., Schultz I. J., Gommans W. M., et al. (2012) A functional screen identifies specific microRNAs capable of inhibiting human melanoma cell viability. PLoS ONE 7: e43569

Rausch T., Jones D. T., Zapatka M., Stutz A. M., Zichner T., et al. (2012) Genome Sequencing of Pediatric Medulloblastoma Links Catastrophic DNA Rearrangements with TP53 Mutations. Cell 148: 59-71

Rausch T., Zichner T., Schlattl A., Stutz A. M., Benes V., et al. (2012) DELLY: structural variant discovery by integrated paired-end and split-read analysis. Bioinformatics 28: i333-i339

Ren S., Peng Z., Mao J.-H., Yu Y., Yin C., et al. (2012) RNA-seq analysis of prostate cancer in the Chinese population identifies recurrent gene fusions, cancer-associated long noncoding RNAs and aberrant alternative splicings. Cell Res
Roberts K. G., Morin R. D., Zhang J., Hirst M., Zhao Y., et al. (2012) Genetic alterations activating kinase and cytokine receptor signaling in high-risk acute lymphoblastic leukemia. Cancer Cell 22: 153-166

Robinson D. R., Wu Y. M., Kalyana-Sundaram S., Cao X., Lonigro R. J., et al. (2013) Identification of recurrent NAB2-STAT6 gene fusions in solitary fibrous tumor by integrative sequencing. Nat Genet 45: 180-185

Robinson G., Parker M., Kranenburg T. A., Lu C., Chen X., et al. (2012) Novel mutations target distinct subgroups of medulloblastoma. Nature 488: 43-48

Rocha P. P., Micsinai M., Kim J. R., Hewitt S. L., Souza P. P., et al. (2012) Close proximity to Igh is a contributing factor to AID-mediated translocations. Mol Cell 47: 873-885

Ruark E., Snape K., Humburg P., Loveday C., Bajrami I., et al. (2013) Mosaic PPM1D mutations are associated with predisposition to breast and ovarian cancer. Nature 493: 406-410

Rudin C. M., Durinck S., Stawiski E. W., Poirier J. T., Modrusan Z., et al. (2012) Comprehensive genomic analysis identifies SOX2 as a frequently amplified gene in small-cell lung cancer. Nat Genet 44: 1111-1116

Sajnani M. R., Patel A. K., Bhatt V. D., Tripathi A. K., Ahir V. B., et al. (2012) Identification of novel transcripts deregulated in buccal cancer by RNA-seq. Gene 507: 152-158

Sarhadi V. K., Lahti L., Scheinin I., Tyybakinoja A., Savola S., et al. (2013) Targeted resequencing of 9p in acute lymphoblastic leukemia yields concordant results with array CGH and reveals novel genomic alterations. Genomics
Saunders C. J., Miller N. A., Soden S. E., Dinwiddie D. L., Noll A., et al. (2012) Rapid whole-genome sequencing for genetic disease diagnosis in neonatal intensive care units. Sci Transl Med 4: 154ra135

Sausen M., Leary R. J., Jones S., Wu J., Reynolds C. P., et al. (2013) Integrated genomic analyses identify ARID1A and ARID1B alterations in the childhood cancer neuroblastoma. Nat Genet 45: 12-17

Schulz E., Valentin A., Ulz P., Beham-Schmid C., Lind K., et al. (2012) Germline mutations in the DNA damage response genes BRCA1, BRCA2, BARD1 and TP53 in patients with therapy related myeloid neoplasms. J Med Genet 49: 422-428

Schweiger M. R., Kerick M., Timmermann B., Albrecht M. W., Borodina T., et al. (2009) Genome-wide massively parallel sequencing of formaldehyde fixed-paraffin embedded (FFPE) tumor tissues for copy-number- and mutation-analysis. PLoS ONE 4: e5548

Selamat S. A., Chung B. S., Girard L., Zhang W., Zhang Y., et al. (2012) Genome-scale analysis of DNA methylation in lung adenocarcinoma and integration with mRNA expression. Genome Res 22: 1197-1211

Seo J. S., Ju Y. S., Lee W. C., Shin J. Y., Lee J. K., et al. (2012) The transcriptional landscape and mutational profile of lung adenocarcinoma. Genome Res 22: 2109-2119

Serre D., Lee B. H. and Ting A. H. (2010) MBD-isolated Genome Sequencing provides a high-throughput and comprehensive survey of DNA methylation in the human genome. Nucleic Acids Res 38: 391-399

Seshagiri S., Stawiski E. W., Durinck S., Modrusan Z., Storm E. E., et al. (2012) Recurrent R-spondin fusions in colon cancer. Nature 488: 660-664

Shah S. P., Roth A., Goya R., Oloumi A., Ha G., et al. (2012) The clonal and mutational evolution spectrum of primary triple-negative breast cancers. Nature 486: 395-399

Shendure J. and Lieberman Aiden E. (2012) The expanding scope of DNA sequencing. Nat Biotechnol 30: 1084-1094

Shepard P. J., Choi E. A., Lu J., Flanagan L. A., Hertel K. J., et al. (2011) Complex and dynamic landscape of RNA polyadenylation revealed by PAS-Seq. RNA 17: 761-772

Sinicropi D., Qu K., Collin F., Crager M., Liu M. L., et al. (2012) Whole transcriptome RNA-Seq analysis of breast cancer recurrence risk using formalin-fixed paraffin-embedded tumor tissue. PLoS ONE 7: e40092

Soon W. W., Hariharan M. and Snyder M. P. (2013) High-throughput sequencing for biology and medicine. Mol Syst Biol 9: 640

Sorrells S., Carbonneau S., Harrington E., Chen A. T., Hast B., et al. (2012) Ccdc94 protects cells from ionizing radiation by inhibiting the expression of p53. PLoS Genet 8: e1002922

Stephens P. J., Greenman C. D., Fu B., Yang F., Bignell G. R., et al. (2011) Massive genomic rearrangement acquired in a single catastrophic event during cancer development. Cell 144: 27-40

Straessler K. M., Jones K. B., Hu H., Jin H., van de Rijn M., et al. (2013) Modeling clear cell sarcomagenesis in the mouse: cell of origin differentiation state impacts tumor characteristics. Cancer Cell 23: 215-227

Streppel M. M., Lata S., Delabastide M., Montgomery E. A., Wang J. S., et al. (2013) Next-generation sequencing of endoscopic biopsies identifies ARID1A as a tumor-suppressor gene in Barrett's esophagus. Oncogene
Sturm D., Witt H., Hovestadt V., Khuong-Quang D. A., Jones D. T., et al. (2012) Hotspot mutations in H3F3A and IDH1 define distinct epigenetic and biological subgroups of glioblastoma. Cancer Cell 22: 425-437

Swanton C. (2013) Plasma-derived tumor DNA analysis at whole-genome resolution. Clin Chem 59: 6-8

Theodoratou E., Montazeri Z., Hawken S., Allum G. C., Gong J., et al. (2012) Systematic meta-analyses and field synopsis of genetic association studies in colorectal cancer. J Natl Cancer Inst 104: 1433-1457

Thompson-Wicking K., Francis R. W., Stirnweiss A., Ferrari E., Welch M. D., et al. (2012) Novel BRD4-NUT fusion isoforms increase the pathogenic complexity in NUT midline carcinoma. Oncogene
Tubio J. M. and Estivill X. (2011) Cancer: When catastrophe strikes a cell. Nature 470: 476-477

Tuna M. and Amos C. I. (2012) Genomic sequencing in cancer. Cancer Lett

Turajlic S., Furney S. J., Lambros M. B., Mitsopoulos C., Kozarewa I., et al. (2012) Whole genome sequencing of matched primary and metastatic acral melanomas. Genome Res 22: 196-207

Tzoneva G., Perez-Garcia A., Carpenter Z., Khiabanian H., Tosello V., et al. (2013) Activating mutations in the NT5C2 nucleotidase gene drive chemotherapy resistance in relapsed ALL. Nat Med 19: 368-371

Umer M. and Herceg Z. (2013) Deciphering the Epigenetic Code: An Overview of DNA Methylation Analysis Methods. Antioxid Redox Signal
van de Werken H. J., Landan G., Holwerda S. J., Hoichman M., Klous P., et al. (2012) Robust 4C-seq data analysis to screen for regulatory DNA interactions. Nat Methods 9: 969-972

van Delft J., Gaj S., Lienhard M., Albrecht M. W., Kirpiy A., et al. (2012) RNA-Seq provides new insights in the transcriptome responses induced by the carcinogen benzo[a]pyrene. Toxicol Sci 130: 427-439

Wagle N., Berger M. F., Davis M. J., Blumenstiel B., DeFelice M., et al. (2011) High-Throughput Detection of Actionable Genomic Alterations in Clinical Tumor Samples by Targeted, Massively Parallel Sequencing. Cancer Discovery 2: 82-93

Wagle N., Berger M. F., Davis M. J., Blumenstiel B., Defelice M., et al. (2012) High-throughput detection of actionable genomic alterations in clinical tumor samples by targeted, massively parallel sequencing. Cancer Discov 2: 82-93

Walter M. J., Shen D., Ding L., Shao J., Koboldt D. C., et al. (2012) Clonal architecture of secondary acute myeloid leukemia. N Engl J Med 366: 1090-1098

Wang Q., Xia J., Jia P., Pao W. and Zhao Z. (2012) Application of next generation sequencing to human gene fusion detection: computational tools, features and perspectives. Brief Bioinform
Waters H. (2012) New NIH genetics center focuses its lens on exome, despite doubts. Nat Med 18: 8

Weischenfeldt J., Simon R., Feuerbach L., Schlangen K., Weichenhan D., et al. (2013) Integrative genomic analyses reveal an androgen-driven somatic alteration landscape in early-onset prostate cancer. Cancer Cell 23: 159-170

Welch J. S., Ley T. J., Link D. C., Miller C. A., Larson D. E., et al. (2012) The origin and evolution of mutations in acute myeloid leukemia. Cell 150: 264-278

Wen H., Li Y., Malek S. N., Kim Y. C., Xu J., et al. (2012) New fusion transcripts identified in normal karyotype acute myeloid leukemia. PLoS ONE 7: e51203

Weng L., Wu X., Gao H., Mu B., Li X., et al. (2010) MicroRNA profiling of clear cell renal cell carcinoma by whole-genome small RNA deep sequencing of paired frozen and formalin-fixed, paraffin-embedded tissue specimens. J Pathol 222: 41-51

Whittaker S. R., Theurillat J. P., Van Allen E., Wagle N., Hsiao J., et al. (2013) A Genome-Scale RNA Interference Screen Implicates NF1 Loss in Resistance to RAF Inhibition. Cancer Discov 3: 350-362

Wilbert M. L., Huelga S. C., Kapeli K., Stark T. J., Liang T. Y., et al. (2012) LIN28 binds messenger RNAs at GGAGA motifs and regulates splicing factor abundance. Mol Cell 48: 195-206

Wilkinson A. C., Ballabio E., Geng H., North P., Tapia M., et al. (2013) RUNX1 is a key target in t(4;11) leukemias that contributes to gene activation through an AF4-MLL complex interaction. Cell Rep 3: 116-127

Wu D., Lamm A. T. and Fire A. Z. (2011) Competition between ADAR and RNAi pathways for an extensive class of RNA targets. Nat Struct Mol Biol 18: 1094-1101

Xu X., Hou Y., Yin X., Bao L., Tang A., et al. (2012) Single-cell exome sequencing reveals single-nucleotide mutation characteristics of a kidney tumor. Cell 148: 886-895

Yang D., Sun Y., Hu L., Zheng H., Ji P., et al. (2013) Integrated analyses identify a master microRNA regulatory network for the mesenchymal subtype in serous ovarian cancer. Cancer Cell 23: 186-199

Yap T. A., Gerlinger M., Futreal P. A., Pusztai L. and Swanton C. (2012) Intratumor heterogeneity: seeing the wood for the trees. Sci Transl Med 4: 127ps110

Yates L. R. and Campbell P. J. (2012) Evolution of the cancer genome. Nat Rev Genet 13: 795-806

Yeo G. W., Coufal N. G., Liang T. Y., Peng G. E., Fu X. D., et al. (2009) An RNA code for the FOX2 splicing regulator revealed by mapping RNA-protein interactions in stem cells. Nat Struct Mol Biol 16: 130-137

Yost S. E., Smith E. N., Schwab R. B., Bao L., Jung H., et al. (2012) Identification of high-confidence somatic mutations in whole genome sequence of formalin-fixed breast cancer specimens. Nucleic Acids Res 40: e107

Zhang J., Benavente C. A., McEvoy J., Flores-Otero J., Ding L., et al. (2012) A novel retinoblastoma therapy from genomic and epigenetic analyses. Nature 481: 329-334

Zhang J., Poh H. M., Peh S. Q., Sia Y. Y., Li G., et al. (2012) ChIA-PET analysis of transcriptional chromatin interactions. Methods 58: 289-299

Zhang X., Cowper-Sal lari R., Bailey S. D., Moore J. H. and Lupien M. (2012) Integrative functional genomics identifies an enhancer looping to the SOX9 gene disrupted by the 17q24.3 prostate cancer risk locus. Genome Res 22: 1437-1446

Zhao Z., Tavoosidana G., Sjolinder M., Gondor A., Mariano P., et al. (2006) Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. Nat Genet 38: 1341-1347