Diagnostic Pathology: Open Access

Liu, et al., Diagn Pathol Open 2016, 1:1
DOI: 10.4172/2476-2024.1000108

Review Open Access

Fusion Genes and Their Detection through Next Generation Sequencing in Malignant Hematological Diseases and Solid Tumors

Jiaqi Liu[1, 2, 3#], **Linqian Weng**[4#], **Yue Ming**[5#], **Sen Liu**[1, 2], **Nan Wu**[1, 2*], **Zhihong Wu**[1, 2*]

[1]*Department of Orthopedic Surgery, Peking Union Medical College Hospital, Peking Union Medical College and Chinese Academy of Medical Sciences, Beijing, PR China*

[2]*Beijing Key Laboratory for Genetic Research of Bone and Joint Disease, Beijing, PR China*

[3]*Department of Breast Surgical Oncology, Cancer Hospital of Chinese Academy of Medical Sciences, Beijing, PR China*

[4]*Peking Union Medical College, Peking Union Medical College and Chinese Academy of Medical Sciences, Beijing, PR China*

[5]*The PET-CT Center, Cancer Hospital of Chinese Academy of Medical Sciences, Beijing, PR China*

[#]*These authors have contributed equally to this work*

[*]***Corresponding author:*** *Zhihong Wu, Beijing Key Laboratory for Genetic Research of Skeletal Deformity; Department of Central Laboratory, Peking Union Medical College Hospital, Peking Union Medical College and Chinese Academy of Medical Sciences, No.1 Shuaifuyuan, Dongcheng District, Beijing 100730, PR China; Tel: +8601069156081; Fax: +8601069156081, E-mail: orthoscience@126.com*

Nan Wu, Department of Orthopaedic Surgery, Peking Union Medical College Hospital, Peking Union Medical College and Chinese Academy of Medical Sciences; Beijing Key Laboratory for Genetic Research of Skeletal Deformity, No.1 Shuaifuyuan, Dongcheng District, Beijing 100730, PR China; Tel: +8601069156081; E-mail: dr.wunan@pumch.cn

Received date: January 02, 2016; **Accepted date:** February 20, 2016; **Published date:** February 26, 2016

Abstract

Fusion genes are neoplasia-associated mutations, which play a particularly significant role in tumorgenesis and exhibit great importance for clinical applications in malignant hematological diseases and solid tumors. Simultaneously with copy number variants (CNVs), gene fusions are resulting from balanced and unbalanced chromosomal rearrangements. Thus, understanding the mutagenesis and instability of CNV, as well as the underlying molecular mechanisms of chromosomal rearrangements will improve our comprehension of gene fusions. Recently, next generation sequencing (NGS), especially transcriptome sequencing or RNA-Sequencing (RNA-seq), has become a very useful tool to identify gene alterations in cancer and a powerful approach for investigating the tumorgenesis. However, we are still facing with the challenge of minimizing false positives in results of RNA-seq. Whole-genome sequencing (WGS) is also used for the fusion gene detection, which provides us a more comprehensive and integrative way to detect structural variants. WGS may correct the false-negative results from RNA-seq. Additionally; many computational tools with more sensitivity and specificity have been developed for the detection of fusion transcripts from NGS datas. In the future, multi-omics analysis, third-generation sequencing and liquid-biopsy technique all provide opportunities to comprehensively interpret gene fusions and understand the biology of cancer genomes.

Keywords: Fusion gene; Next generation sequencing (NGS); Transcriptome sequencing or RNA-seq (RNA-Seqencing); Whole genome sequencing (WGS)

Introduction

Fusion genes, also called chimeric genes or hybrid genes, are neoplasia-associated mutations arising from structural chromosome rearrangements, such as chromosomal insertion, deletion, translocation or inversion that juxtaposes two separated genes [1,2]. They have been reported to be important genomic events in human cancer because their fusion gene products can drive the development of cancer, and thus are potential prognostic markers or therapeutic targets in cancer treatment. On the basis of transposons studies, human cancers could also be the result of the translocations and chromosome rearrangements which lead to the abnormal expression of genes located at breakpoints [3]. Up to now, the current next generation sequencing(NGS)-based approaches for detection, such as transcriptome sequencing or RNA-Sequencing (RNA-seq) and whole genome sequencing (WGS), have become a very useful tool to identify new tumor-associated gene fusions and investigate their impact on tumorgenesis [4]. In this study, we then comprehensively reviewed NGS studies to detect gene fusions in malignant hematological diseases and solid tumors, to update our knowledge about the advances and challenges in the gene fusion detection through NGS, especially RNA-seq.

Methods

The literatures searching was conducted on PubMed, ScienceDirect and Google Scholar comprehensively, using keywords included "gene fusion", "RNA-seq"/"whole-genome sequencing"/"next generation sequencing" AND "cancer"/ "tumor"/ "leukemia"/ "lymphoma". After the relevant literatures were carefully read and analyzed, we found 71 publications directly related to our study purpose by this searching method.

Fusion genes: tumorgenesis, biomarker and therapeutic target

Fusion genes play a particularly significant role in tumorgenesis, which has been identified with great importance for clinical

applications [5]. Gene fusion events can be observed in cancer samples more frequently than benign samples. They are present in approximately 20% of all human neoplasms. Although the functional outcomes of many gene fusions are still under exploration, it is well established that most of them will lead to tumorigenesis. Since a strong correlation can be found between recurrent gene fusions and tumor types, gene fusion detection has been suggested to be used for screening of common tumor types. Subtypes identification provides a roadmap for targeted therapies. Although recent studies have thus far defined a large quantity of gene fusions that involve different cancer related genes, which constitute an important diagnostic and prognostic parameter especially malignant hematological diseases and sarcomas, gene fusions in solid tumors have rather limited clinical and biological impact [5].

The *BCR-ABL1* fusion gene in the well-known Philadelphia (Ph) chromosome is the prototypic fusion oncogene, which is associated with chronic myeloid leukemia (CML). It is now used as a biomarker during diagnosis and monitoring patient response to treatment. As some morphologically homogeneous malignancies are heterogeneous because of gene fusion status, they play an important role in treatment stratification, such as different *MLL* fusions in AML or fusion-positive versus fusion-negative Alveolar rhabdomyosarcoma (ARMS) [6,7]. Nowadays, many technologies have already been used in detecting gene fusions and other genetic aberrations, such as chromosome banding analysis, reverse transcriptase-polymerase chain reaction (RT-PCR) and Sanger sequencing etc [8]. Except for the hematological malignancies, large amount of data emerges from the studies of malignant solid tumors, including most sarcomas and a few carcinomas.

Ewing's sarcoma is defined by a recurrent chromosomal translocation between the *EWSR1* gene and various *ETS* genes, and *EWS-FLI1* is the most common gene fusion in Ewing's sarcoma, which present in 85% of cases [9]. In the study carried out by Saravana et al. [10], genes such as *CLK1*, *CASP3*, *PPFIBP1* and *TERT*, which potentially participate in oncogenesis, are alternatively spliced by *EWS-FLI1*. Thus *EWS-FLI1* can be used as a diagnostic biomarker for Ewing's sarcoma. While, there are still many important questions to be solved to understand the molecular mechanism of *EWS-FLI1* and its potential value for cancer therapy.

As is known, the oncogenic potential of *ETS*-related gene (*ERG*) is involved in Ewing's sarcoma and leukemia. However, in the past decades, *ERG* is found to be highly associated with prostate cancer [11]. It is showed by Tomlins et al. [12] that *ERG* is overexpressed in most prostate carcinomas because of a gene fusion with the androgen-driven promoter of *TMPRSS2* gene. Many other studies have also shown *TMPRSS2-ERG* gene rearrangements to be the most commonly found *TMPRSS2:ETS* family pairing in prostate cancer, demonstrating the specificity of *TMPRSS2-ERG* for prostate cancer and a role for *TMPRSS2-ERG* in the development and progression of prostate cancer [13]. The *TMPRSS2-ERG* is showed to trigger carcinogenesis by inhibiting apoptosis of prostate gland cells and at the same time, increasing cell proliferation [14]. The proto-oncogenes *ETV1*, were also found to be highly expressed in a subset of prostate cancers [12]. Recently, it is recommended that the classification of prostate cancer can be divided by distinct molecular subtypes, which includes mutually exclusive *ETS* fusions (*ETS*-positive), *SPINK1*-overexpressing, and *CHD1*-loss cancers etc [15]. In this way, a simple molecular barcode (includes *ETS/SPINK1/SPOP/CHD1/RAS-RAF/PTEN/TP53* status) can be used in molecular prostate cancer subtypes, and thus may allow stratification of patients for different management strategies in the future.

Approximately 40%-70% of men with castration resistant prostate cancers have *ERG* rearrangements, which may respond better to anti-hormonal therapy than *ERG*-negative ones [13]. Currently, there are many studies targeting at the *TMPRSS2-ERG* fusion and its downstream signaling. It was shown that knockdown of the *TMPRSS2-ERG* fusion in a cancer cell line can lead to primary tumour growth inhibition, which made *TMPRSS2-ERG* a potential therapeutic target [16]. There is also study showing that targeting the most common and clinically significant alternatively spliced isoforms of the *TMPRSS2-ERG* mRNA with specific *siRNAs* via liposomal nanovectors can be promising therapy for men with prostate cancer [17]. For example, siRNA has been used to target the *BCR-ABL* fusion successfully in CML and against the *AML1-ETO* in *AML-M2* [18]. Because specific *ETS* factors could be found in many other solid tumor types, their downstream effectors are very likely to be in common, therefore providing more possible novel drug targets for treatment of these malignancies.

Relatively few recurrent gene fusion events have been associated with breast cancer. In a study of whole-transcriptome sequencing of 120 fresh-frozen primary breast cancer samples, six newly validated gene fusions were recurrent, including three in-frame and three out-frame ones [19]. A recurrent gene fusion, *RPS6KB1* kinase, and *EGFR*, which is a therapeutically important receptor kinase and involving in the rapamycin signaling, was discovered in the analysis of 14 breast cancer cell lines [1]. Not only in common tumors, recent study also indicated a novel *FN1-FGFR1* fusion gene might participate the tumorgenesis of phosphaturic mesenchymal tumors (PMTs), which typically cause hypophosphataemia and tumor-induced osteomalacia (TIO) [20].

As people getting more into the clinical importance of gene fusions and other types of genetic rearrangement, greater emphasis has been putting on genetic features in the classification of neoplasms. In the latest World Health Organization (WHO) classifications, translocation and/or gene fusion status is mandatory for the diagnosis of some types of tumors, such as "AML with t(8;21)(q22;q22), *RUNX1-RUNX1T1*" and "B lympho-blastic leukemia/lymphoma with t(5;14)(q31;q32), *IL3-IGH*" [4]. And for other malignancies, the Xp11 translocation renal cell carcinomas (RCC) harbor gene fusions involving *TFE3*, which is among the MiT subfamily of transcription factors. Thus it was first officially recognized in the 2004 WHO renal tumor classification [21].

Therapeutic approaches based on oncogene addiction can offer significant anticancer benefit, among which the identification of anaplastic lymphoma kinase *(ALK)* rearrangements is a key aspect. For all the lung cancer patients, 4-8% can be detected with the *EML4-ALK* gene fusion, especially in light smokers and nonsmokers [22]. Crizotinib was the first approved medication for *ALK*-positive patients. In the phase III PROFILE 1014 study, crizotinib is associated with a median progression-free survival of 10.9 months when used as the first-line treatment [23]. Imatinib, the tyrosine kinase inhibitor, which was the first drug that was specifically designed to target a fusion gene, *BCR-ABL1* in CML. There are various other common malignancies that have been shown to display various fusions involving kinase-encoding regions, e.g. *BRAF*, *FGFR3*, *NTRK1, RET* and *ROS1* etc [4]. With more and more novel drugs under approval of FDA regarding to these gene fusions, stratification of diagnosis and treatment could be of great importance in clinical practice.

Chromosome rearrangement: the origin of gene fusions

Chromosomal rearrangements are very pervasive in cancer, while their impacts are hard to characterize and interpret [24]. Gene fusions are resulting from balanced and unbalanced chromosomal rearrangements. Balanced changes are the prototypical mechanism behind gene fusions, including translocations, insertions and inversions. While gene fusions can also arise through unbalanced chromosomal rearrangements, such as interstitial deletions, as to a deletion of an interstitial chromosomal segment. Both balanced and unbalanced aberrations may lead to create a chimeric gene through the fusion of parts of the two genes from each side of breakpoint, or juxtapose the coding sequences in one gene with the regulatory sequences of another gene from the other breakpoint. Even there should be two derivative chromosomes and each of which may harbour the pathogenetic gene fusion through a balanced chromosomal rearrangement, usually only one of these genes will produce an in-frame fusion transcript [4]. However, genes in one of the breakpoints may also become truncated and lose their function as haploinsufficiency. As the gene fusions can upregulate or deregulate genes depending on the breakpoints, it may lead to tumorgenesis through activation of oncogene or inactivation of tumor suppressor gene.

Interestingly, gene fusion always occurred simultaneously with CNVs, which also has a significant role in tumorigenesis in many cancers, such as gastric cancer [25], ovarian cancer [26], hepatocellular carcinoma [27], colorectal cancer [28], bladder cancer [29] and so on. CNV involves deletions, duplications and insertions of DNA segments larger than 1 kb, which is variable among individuals [30]. Many seemingly balanced translocations that result in gene fusions are accompanied by extensive deletions, duplications or amplifications among the breakpoints [31,32]. In most cases, CNV generates more than one breakpoint. When a breakpoint located between the functional elememts of the two genes, a fusion gene may occur. Fusion partner genes can be found to contribute promoters (5' UTR), coding sequences and 3' UTRs. Consequently, genes affected by CNV are potential candidates for fusion events [4]. Thus, understanding the mutagenesis and instability of CNV, as well as the underlying molecular mechanisms of chromosomal rearrangements will improve our understanding of gene fusions.

In addition, transcript fusions may also originate from non-adjacent genes without a corresponding fusion at the DNA level, resulting in so-called transcription-induced gene fusions (TIGFs), including cis-TIGFs (neighbouring genes located on the same DNA strand) and trans-TIGFs (genes located far apart or on different chromosomes). Some cis-TIGFs have been identified associated with particular tumor types, which indicates TIGFs may play important roles in tumor development [33]. Although trans-TIGFs have been identified in human cells [34], no trans-TIGFs have yet been verified in any independent studies [4].

Next generation sequencing (NGS): a high-performing strategy for fusion gene discovery

Although cytogenetics and fluorescence in situ hybridization (FISH) approach will continue to be indispensable tools for fusion gene diagnostics in hematological diseases and solid tumors, the modern high-throughput NGS have showed their great impact to identify new tumor-associated gene fusions [35]. Recently, NGS has become a very useful tool to identify gene alterations in cancer and a powerful approach for investigating the tumorgenesis [36]. Chromosomal rearrangements, such as deletion, duplication, translocation, insertion and inversion, can be detected by paired-end information and apparent fragment length and orientation of NGS [37]. Additionally, chimera read analysis can detect gene fusions and also reveal their breakpoints directly [37,38], and the de novo assembly approach can be used for some complex fusions [39]. Over the past few years, advances of NGS and affordable price provide an opportunity for detection of cancer transcriptomes, including the expressed fusion genes. The first NGS study to detect gene fusions in cancer were carried out on cell lines [37], and quickly extended to numerous investigations in different cancer types. As another landscape, Maher et al. [40] successfully re-discovered the *BCR-ABL1* gene fusion in a CML cell line and the *TMPRSS2-ERG* gene fusion in a prostate cancer cell line and tissues through RNA-Seq. Yoshihara et al. [2] queried transcriptome data from 4,366 neoplasms from 13 different cancer types, which had been studied within the Cancer Genome Atlas (TCGA) network, and detected more than 8,600 different fusion transcripts. During only the past 3 years, more than 9,000 novel gene fusions have been identified mostly through NGS technologies [4], while most of them have now been described as probably passenger mutations which show little or no effect on tumorigenesis [2]. RNA-Seq is a useful tool for the discovery of gene fusions in cancer transcriptomes and has already become the primary technology for discovering gene fusions. Some open databases of gene fusions in cancer from RNA-seq data have been set up, including Fusion Cancer [41], while we are still faced with the challenge of minimizing false positives in RNA-seq result [19,40,42]. In addition, there are lower proportions (about 3%) of recurrent fusion genes detected by RNA-seq [4].

WGS is also pervasively used for the fusion gene detection [43]. And it provides us a more comprehensive and integrative way to detect structural variants than RNA-seq, especially for de novo gene fusions. WGS would correct the false-negative results from RNA-seq [4,42,44]. As an example, WGS revealed a distinct phenomenon named “chromothripsis” [39], which means chromosomes in a tumor cell produce hundreds of clustered rearrangements [45]. This complicated rearrangement phenomenon was generated as distinct chromosomes or genomic regions shatter into many segments, which are then pieced together by DNA repair mechanisms inaccurately [46]. Recent WGS study suggested this genomic instability phenomenon in cancers co-segregated with inactivation of DNA maintenance genes, like *BRCA1/2* [47], and increasing from patients with germline p53 mutations [48]. Some structural variants without producing fusion genes can also change the expression of nearby genes by changing the functional elements. Although RNA-Seq data can detect most of the transcriptional fusions of these genomic alterations [13], there are still much potential transcriptional consequences of structural variants to be further explored. Integrating data from RNA-Seq and WGS would disclose more genetic variants, as TIGF. However, up to now, there are only few studies comprehensively evaluate the transcriptional fusions from WGS and RNA-Seq [49].

Due to widespread applications of high-throughput NGS technologies, major advancements have been made in computational strategies for fusion gene discovery in recent years [50]. Several computational tools have also been developed for the detection of fusion transcripts using RNA-Seq data, such as MapSplice[51], ShortFuse[52], FusionHunter [44], FusionMap [53], SnowShoes-FTD [54],defuse [55], chimerascan [56], FusionCatcher [57], TopHat-Fusion [44], BreakFusion [58], EricScript [59], SOAPfuse [60],

FusionQ [61] , PRADA [62] and JAFFA [63]. Liu et al.[64] performed a large-scale comparative study by applying these above 15 fusion transcript detection pipelines to 3 synthetic data sets and 3 real paired-end RNA-seq studies and developed a meta-caller algorithm to combine three top-performing methods (FusionCatcher, SOAPfusea and JAFFA). If possible, it is recommended to apply all three above pipelines and combine the results in applications. FusionMatcher (FuMa) is a recently designed fusion genes identical program which can automatically compare and summarize all combinations of two or more datasets in a single run and use one gene annotation, to avoid mismatches caused by tool specific gene annotations [65]. It's believed that both WGS and RNA-seq have their limitations when used independently, and orthogonal validating both data could generate a more sensitive and specific gene fusion detection. To integrate both RNA-seq and WGS data, INTEGRATE was developed to analysis both data to reconstruct gene fusion junctions and genomic breakpoints by split-read mapping. As a result, it was confirmed to be a highly sensitive and accurate approach for detecting high-confidence gene fusion predictions [66]. However, developing the new generation of fusion genes identifying tools from RNA-seq or other NGS data with both sensitivity and specificity remains an important and open question.

Perspective

Gene fusions have strong association with CNVs and whole genomic instability in cancer, which makes it impossible that revealing the complete genomic consequence through only one strategy up to now. In the future, multi-omics analysis of molecular data, such as DNA sequence mutations, CNVs, RNA profiling, DNA methylation, protein expression and chromatin structure may be required to comprehensively interpret gene fusions in order to understand the biology of cancer genomes. Another integrated approach should be done to interpret gene fusions and identify their impact. It is better to combine the NGS result with high-throughput functional cellular assays and more functional data in cancer genomics. In addition, third-generation sequencing which can produce long read sequences is now attempted to clarify complicated genomic structures, including gene fusions, in cancer genome [67].

Nowadays, as the circulating tumor DNA (ctDNA) and circulating tumor cells (CTCs) are more frequently utilized in research and clinical medicine, the 'liquid biopsies' can provide the opportunity to promptly track cancer genome evolution of all cancerous lesions [68]. With the rapid development of highly sensitive and accurate technologies of NGS, it can not only predict the response to treatment, but also monitor minimal residual disease [69,70]. As an example, *FGFR2* fusion in ctDNA was readily detectable by quantitative real-time reverse transcription-polymerase chain reaction and corroborated to be more sensitive and specific than previous biomarkers, such as *CA125* [71]. It is promising that fusion genes can be detected by NGS in liquid biopsies, in the near future.

Competing Interests

The authors declare that they have no competing interests.

Authors' Contributions

All authors participated in searching and reviewing literatures. JL and LW drafted the manuscript.

References

1. Kalyana-Sundaram S, Shankar S, Deroo S, Iyer MK, Palanisamy N, et al. (2012) Gene fusions associated with recurrent amplicons represent a class of passenger aberrations in breast cancer. Neoplasia 14: 702-708.
2. Yoshihara K, Wang Q, Torres-Garcia W, Zheng S, et al. (2015) The landscape and therapeutic relevance of cancer-associated transcript fusions. Oncogene 34: 4845-4854.
3. Klein G (1981) The role of gene dosage and genetic transpositions in carcinogenesis. Nature 294: 313-318.
4. Mertens F, Johansson B, Fioretos T, Mitelman F (2015) The emerging complexity of gene fusions in cancer. Nat Rev Cancer 15: 371-381.
5. Mitelman F, Johansson B, Mertens F (2007) The impact of translocations and gene fusions on cancer causation. Nat Rev Cancer 7: 233-245.
6. Balgobind BV, Raimondi SC, Harbott J, Zimmermann M, Alonzo TA, et al. (2009) Novel prognostic subgroups in childhood 11q23/MLL-rearranged acute myeloid leukemia: results of an international retrospective study. Blood 114: 2489-2496.
7. Williamson D, Missiaglia E, de Reynies A, Pierron G, Thuille B, et al. (2010) Fusion gene-negative alveolar rhabdomyosarcoma is clinically and molecularly indistinguishable from embryonal rhabdomyosarcoma. J Clin Oncol 28: 2151-2158.
8. Mertens F, Tayebwa J (2014) Evolving techniques for gene fusion detection in soft tissue tumours. Histopathology 64: 151-162.
9. Gordon DJ, Motwani M, Pellman D (2015) Modeling the initiation of Ewing sarcoma tumorigenesis in differentiating human embryonic stem cells. Oncogene .
10. Selvanathan SP, Graham GT, Erkizan HV, Dirksen U, Natarajan TG, et al. (2015) Oncogenic fusion protein EWS-FLI1 is a network hub that regulates alternative splicing. Proc Natl Acad Sci U S A 112: E1307-1316.
11. Adamo P, Ladomery MR (2016) The oncogene ERG: a key factor in prostate cancer. Oncogene 35: 403-414.
12. Pellegrini KL, Sanda MG, Moreno CS (2015) RNA biomarkers to facilitate the identification of aggressive prostate cancer. Mol Aspects Med 45: 37-46.
13. Tomlins SA, Rhodes DR, Perner S, Dhanasekaran SM, Mehra R, et al. (2005) Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. Science 310: 644-648.
14. Attard G, Parker C, Eeles RA, Schröder F, Tomlins SA, et al. (2016) Prostate cancer. Lancet 387: 70-82.
15. Lotan TL, Toubaji A, Albadine R, Latour M, Herawi M, et al. (2009) TMPRSS2-ERG gene fusions are infrequent in prostatic ductal adenocarcinomas. Mod Pathol 22: 359-365.
16. Wang J, Cai Y, Yu W, Ren C, Spencer DM, et al. (2008) Pleiotropic biological activities of alternatively spliced TMPRSS2/ERG fusion gene transcripts. Cancer Res 68: 8516-8524.
17. Shao L, Tekedereli I, Wang J, Yuca E, Tsang S, et al. (2012) Highly specific targeting of the TMPRSS2/ERG fusion gene using liposomal nanovectors. Clin Cancer Res 18: 6648-6657.
18. Scherr M, Battmer K, Winkler T, Heidenreich O, Ganser A, et al. (2003) Specific inhibition of bcr-abl gene expression by small interfering RNA. Blood 101: 1566-1569.
19. Kim J, Kim S, Ko S, In YH, Moon HG, et al. (2015) Recurrent fusion transcripts detected by whole-transcriptome sequencing of 120 primary breast cancer samples. Genes Chromosomes Cancer 54: 681-691.
20. Lee JC, Jeng YM, Su SY, Wu CT, Tsai KS, et al. (2015) Identification of a novel FN1-FGFR1 genetic fusion as a frequent event in phosphaturic mesenchymal tumour. J Pathol 235: 539-545.
21. Argani P (2015) MiT family translocation renal cell carcinoma. Semin Diagn Pathol 32: 103-113.
22. Duchemann B, Friboulet L, Besse B (2015) Therapeutic management of ALK+ nonsmall cell lung cancer patients. Eur Respir J 46: 230-242.
23. Garraway LA, Lander ES (2013) Lessons from the cancer genome. Cell 153: 17-37.

24. Liang L, Fang JY, Xu J (2015) Gastric cancer and gene copy number variation: emerging cancer drivers for targeted therapy. Oncogene.
25. Despierre E, Moisse M, Yesilyurt B, Sehouli J, Braicu I, et al. (2014) Somatic copy number alterations predict response to platinum therapy in epithelial ovarian cancer. Gynecol Oncol 135: 415-422.
26. Xu H, Zhu X, Xu Z, Hu Y, Bo S, et al. (2015) Non-invasive Analysis of Genomic Copy Number Variation in Patients with Hepatocellular Carcinoma by Next Generation DNA Sequencing. J Cancer 6: 247-253.
27. Wang H, Liang L, Fang JY, Xu J (2015) Somatic gene copy number alterations in colorectal cancer: new quest for cancer drivers and biomarkers. Oncogene .
28. Xie J, Zhang L, Li M, Du J, Zhou L, et al. (2014) Functional analysis of the involvement of apurinic/apyrimidinic endonuclease 1 in the resistance to melphalan in multiple myeloma. BMC Cancer 14: 11.
29. Feuk L, Carson AR, Scherer SW (2006) Structural variation in the human genome. Nat Rev Genet 7: 85-97.
30. Möller E, Hornick JL, Magnusson L, Veerla S, Domanski HA, et al. (2011) FUS-CREB3L2/L1-positive sarcomas show a specific gene expression profile with upregulation of CD24 and FOXL1. Clin Cancer Res 17: 2646-2656.
31. Sinclair PB, Nacheva EP, Leversha M, Telford N, Chang J, et al. (2000) Large deletions at the t(9;22) breakpoint are common and may identify a poor-prognosis subgroup of patients with chronic myeloid leukemia. Blood 95: 738-743.
32. Nacu S, Yuan W, Kan Z, Bhatt D, Rivers CS, et al. (2011) Deep RNA sequencing analysis of readthrough gene fusions in human prostate adenocarcinoma and reference samples. BMC Med Genomics 4: 11.
33. Zaphiropoulos PG1 (2011) Trans-splicing in Higher Eukaryotes: Implications for Cancer Development? Front Genet 2: 92.
34. Liehr T, Othman MA, Rittscher K, Alhourani E (2015) The current state of molecular cytogenetics in cancer diagnosis. Expert Rev Mol Diagn 15: 517-526.
35. Wang E, Zaman N, Mcgee S, Milanese JS, Masoudi-Nejad A, et al. (2015) Predictive genomics: a cancer hallmark network framework for predicting tumor clinical phenotypes using genome sequencing data. Semin Cancer Biol 30: 4-12.
36. Campbell PJ, Stephens PJ, Pleasance ED, O'Meara S, Li H, Santarius T, Stebbings LA, Leroy C, Edkins S, Hardy C and others. Identification of somatically acquired rearrangements in cancer using genome-wide massively parallel paired-end sequencing. Nat Genet 2008;40: 722-729.
37. Yang L, Luquette LJ, Gehlenborg N, Xi R, Haseley PS, et al. (2013) Diverse mechanisms of somatic structural variations in human cancer genomes. Cell 153: 919-929.
38. Nagarajan N, Pop M (2013) Sequence assembly demystified. Nat Rev Genet 14: 157-167.
39. Campbell PJ, Stephens PJ, Pleasance ED, O'Meara S, Li H, Santarius T, et al. (2008) Identification of somatically acquired rearrangements in cancer using genome-wide massively parallel paired-end sequencing. Nat Genet 40: 722-729.
40. Maher CA, Kumar-Sinha C, Cao X, Kalyana-Sundaram S, Han B, et al. (2009) Transcriptome sequencing to detect gene fusions in cancer. Nature 458: 97-101.
41. Wang Y, Wu N, Liu J, Wu Z, et al. (2015) FusionCancer: a database of cancer fusion genes derived from RNA-seq data. Diagn Pathol 10: 131.
42. Ozsolak F1, Milos PM (2011) RNA sequencing: advances, challenges and opportunities. Nat Rev Genet 12: 87-98.
43. Brohl AS, Solomon DA, Chang W, Wang J, Song Y, et al. (2014) The genomic landscape of the Ewing Sarcoma family of tumors reveals recurrent STAG2 mutation. PLoS Genet 10: e1004475.
44. Kim D, Salzberg SL (2011) TopHat-Fusion: an algorithm for discovery of novel fusion transcripts. Genome Biol 12: R72.
45. Stephens PJ, Greenman CD, Fu B, Yang F, Bignell GR, et al. (2011) Massive genomic rearrangement acquired in a single catastrophic event during cancer development. Cell 144: 27-40.
46. Korbel JO, Campbell PJ (2013) Criteria for inference of chromothripsis in cancer genomes. Cell 152: 1226-1236.
47. Waddell N, Pajic M, Patch AM, Chang DK, Kassahn KS, et al. (2015) Whole genomes redefine the mutational landscape of pancreatic cancer. Nature 518: 495-501.
48. Rausch T, Jones DT, Zapatka M, Stutz AM, Zichner T, et al. (2011) Genome sequencing of pediatric medulloblastoma links catastrophic DNA rearrangements with TP53 mutations. Cell 148: 59-71.
49. Nakagawa H, Wardell CP, Furuta M, Taniguchi H, Fujimoto A (2015) Cancer whole-genome sequencing: present and future. Oncogene 34: 5943-5950.
50. Wang Q, Xia J, Jia P, Pao W, Zhao Z (2013) Application of next generation sequencing to human gene fusion detection: computational tools, features and perspectives. Brief Bioinform 14: 506-519.
51. Wang K, Singh D, Zeng Z, Coleman SJ, Huang Y, et al. (2010) MapSplice: accurate mapping of RNA-seq reads for splice junction discovery. Nucleic Acids Res 38: e178.
52. Kinsella M, Harismendy O, Nakano M, Frazer KA, Bafna V (2011) Sensitive gene fusion detection using ambiguously mapping RNA-Seq read pairs. Bioinformatics 27: 1068-1075.
53. Ge H1, Liu K, Juan T, Fang F, Newman M, et al. (2011) FusionMap: detecting fusion genes from next-generation sequencing data at base-pair resolution. Bioinformatics 27: 1922-1928.
54. Asmann YW, Hossain A, Necela BM, Middha S, Kalari KR, et al. (2011) A novel bioinformatics pipeline for identification and characterization of fusion transcripts in breast cancer and normal cell lines. Nucleic Acids Res 39:e100.
55. McPherson A, Hormozdiari F, Zayed A, Giuliany R, Ha G, et al. (2011) deFuse: an algorithm for gene fusion discovery in tumor RNA-Seq data. PLoS Comput Biol 7: e1001138.
56. Iyer MK, Chinnaiyan AM, Maher CA (2011) ChimeraScan: a tool for identifying chimeric transcription in sequencing data. Bioinformatics 27: 2903-2904.
57. Edgren H, Murumagi A, Kangaspeska S, Nicorici D, Hongisto V, et al. (2011) Identification of fusion genes in breast cancer by paired-end RNA-sequencing. Genome Biol 12: R6.
58. Chen K, Wallis JW, Kandoth C, Kalicki-Veizer JM, Mungall KL, et al. (2012) BreakFusion: targeted assembly-based identification of gene fusions in whole transcriptome paired-end sequencing data. Bioinformatics 28(14):1923-1924.
59. Benelli M, Pescucci C, Marseglia G, Severgnini M, Torricelli F, et al. (2012) Discovering chimeric transcripts in paired-end RNA-seq data by using EricScript. Bioinformatics 28: 3232-3239.
60. Jia W, Qiu K, He M, Song P, Zhou Q, et al. (2013) SOAPfuse: an algorithm for identifying fusion transcripts from paired-end RNA-Seq data. Genome Biol 14: R12.
61. Liu C, Ma J, Chang CJ, Zhou X (2013) FusionQ: a novel approach for gene fusion detection and quantification from paired-end RNA-Seq. BMC Bioinformatics 14: 193.
62. Torres-García W, Zheng S, Sivachenko A, Vegesna R, Wang Q, et al. (2014) PRADA: pipeline for RNA sequencing data analysis. Bioinformatics 30: 2224-2226.
63. Davidson NM, Majewski IJ, Oshlack A (2015) JAFFA: High sensitivity transcriptome-focused fusion gene detection. Genome Med 7: 43.
64. Liu S, Tsai WH, Ding Y, Chen R, Fang Z, et al. (2015) Comprehensive evaluation of fusion transcript detection algorithms and a meta-caller to combine top performing methods in paired-end RNA-seq data. Nucl Acids Res.
65. Hoogstrate Y, Böttcher R, Hiltemann S, van der Spek PJ, Jenster G, et al. (2015) FuMa: reporting overlap in RNA-seq detected fusion genes. Bioinformatics .
66. Zhang J, White NM, Schmidt HK, Fulton RS, Tomlinson C, et al. (2016) INTEGRATE: gene fusion discovery using whole genome and transcriptome data. Genome Res 26: 108-118.

67. Koren S, Schatz MC, Walenz BP, Martin J, Howard JT, et al. (2012) Hybrid error correction and de novo assembly of single-molecule sequencing reads. Nat Biotechnol 30: 693-700.

68. Hokland P, Ommen HB (2011) Towards individualized follow-up in adult acute myeloid leukemia in remission. Blood 117: 2577-2584.

69. Crowley E, Di Nicolantonio F, Loupakis F, Bardelli A (2013) Liquid biopsy: monitoring cancer-genetics in the blood. Nat Rev Clin Oncol 10: 472-484.

70. Cai X, Janku F, Zhan Q, Fan JB (2015) Accessing Genetic Information with Liquid Biopsies. Trends Genet 31: 564-575.

71. Martignetti JA, Camacho-Vanegas O, Priedigkeit N, Camacho C, Pereira E, et al. (2014) Personalized ovarian cancer disease surveillance and detection of candidate therapeutic drug target in circulating tumor DNA. Neoplasia 16: 97-103.