

# Fusion Genes and RNAs in Cancer Development

Kenzi Taniue <sup>1,2,\*</sup> and Nobuyoshi Akimitsu <sup>1,\*</sup><sup>1</sup> Isotope Science Center, The University of Tokyo, 2-11-16, Yayoi, Bunkyo-ku, Tokyo 113-0032, Japan<sup>2</sup> Cancer Genomics and Precision Medicine, Division of Gastroenterology and Hematology/Oncology, Department of Medicine, Asahikawa Medical University, 2-1 Midorigaoka Higashi, Asahikawa, Hokkaido 078-8510, Japan

\* Correspondence: kenzui@ric.u-tokyo.ac.jp (K.T.); akimitsu@ric.u-tokyo.ac.jp (N.A.)

**Abstract:** Fusion RNAs are a hallmark of some cancers. They result either from chromosomal rearrangements or from splicing mechanisms that are non-chromosomal rearrangements. Chromosomal rearrangements that result in gene fusions are particularly prevalent in sarcomas and hematopoietic malignancies; they are also common in solid tumors. The splicing process can also give rise to more complex RNA patterns in cells. Gene fusions frequently affect tyrosine kinases, chromatin regulators, or transcription factors, and can cause constitutive activation, enhancement of downstream signaling, and tumor development, as major drivers of oncogenesis. In addition, some fusion RNAs have been shown to function as noncoding RNAs and to affect cancer progression. Fusion genes and RNAs will therefore become increasingly important as diagnostic and therapeutic targets for cancer development. Here, we discuss the function, biogenesis, detection, clinical relevance, and therapeutic implications of oncogenic fusion genes and RNAs in cancer development. Further understanding the molecular mechanisms that regulate how fusion RNAs form in cancers is critical to the development of therapeutic strategies against tumorigenesis.

**Keywords:** fusion RNAs; gene fusions; chromosomal rearrangements; trans-splicing; cis-splicing; FISH; RT-PCR; non-RT-based assay; RNA sequence

**Citation:** Taniue, K.; Akimitsu, N. Fusion Genes and RNAs in Cancer Development. *Non-coding RNA* **2021**, *7*, 10. <https://doi.org/10.3390/ncrna7010010>

Received: 31 December 2020

Accepted: 3 February 2021

Published: 4 February 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Cancer is a disease of the genome [1,2]. Gene fusions or chromosomal rearrangements are an important class of somatic alterations in cancer and can have important roles in the initial steps of tumorigenesis. [3–5]. The first cancer-associated chromosomal rearrangement was identified in 1960 as a translocation of chromosomes 9 and 22 [4,6,7]. The abnormally small resulting chromosome, named the Philadelphia chromosome, was found in over 95% of patients with chronic myelogenous leukemia (CML) and consisted of the breakpoint cluster region (*BCR*) gene fused to the second exon of the Abelson murine leukemia viral oncogene homolog 1 (*ABL1*) gene [8,9]. Additional examples of cancer-associated chromosomal aberrations have been identified in other hematological malignancies and sarcomas; for example, mixed lineage leukemia (MLL) fusions, *RUNX1–RUNX1T1* and *PML–RAR $\alpha$* , *EWSR1–FLI1* and *EVT6–NTRK3* [10,11]. Although originally discovered in hematological malignancies, gene fusions are now known to occur in several solid tumor types [7,12,13]. The first fusion gene found in a solid tumor was *CTNNB1–PLAG1* in salivary gland adenoma, which is usually benign [12,14,15]. Other fusion genes were soon discovered in solid tumors and other malignancies, including glioblastoma, melanoma, and prostate, breast, ovarian, lung, colorectal, and head and neck cancers [13].

Most of the nonprotein-coding region of the human genome was previously considered to be “junk DNA” [16]. With the advent of massive parallel sequencing technology, these regions in the human genome have been clearly shown to transcribe dynamically and differentially into noncoding RNAs (ncRNAs), such as microRNAs (miRNAs), small nucleolar RNAs (snoRNAs), long ncRNAs (lncRNAs), and circular RNAs (circRNAs) [17–

20]. Accumulating evidence indicates that lncRNAs play critical roles in diverse biological processes, including differentiation, stem cell pluripotency, embryogenesis, pathogenic infection, neurogenesis, proliferation, and tumorigenesis [16,20–26]. LncRNAs also function in chromatin and genomic structural remodeling, RNA trafficking, RNA stabilization, transcriptional regulation, translation, signal pathway, and protein degradation [27–31]. Expression of lncRNA, miRNA, and snoRNA have shown close correlations with specific chromosomal rearrangements in cancers [10,32–34]. Moreover, fusion circRNAs (f-circRNAs) that are generated by chromosome rearrangement contribute to oncogenic roles [35]. Furthermore, the fusion RNA *SLC45A3–ELK4*, which regulates cancer cell proliferation, functions as a lncRNA [36]. However, details of mechanisms of the oncogenic roles of these fusion RNAs are unclear.

The ultimate goal of precision medicine in cancer treatment is the development of therapeutic strategies that specifically target cancer cells without affecting normal cells [37]. Targeting oncogenic fusion genes and RNAs specific to cancer tissue for treatment and diagnosis could bring us closer to the approach. Moreover, these fusions are often present at clonal levels within tumors; their generation is frequently the founding genetic abnormality that drives the cancer [37,38]. In this review, we present the function and biogenesis of these fusions and current knowledge regarding their roles in cancer development. We also describe an overview of methodologies for identifying fusion genes and RNAs in cancer development and tumorigenesis. Development of therapeutic strategies that target fusion genes and RNAs, and the study of their mechanisms of production and actions, may provide robust opportunities to eradicate cancers that harbor aberrant genes and RNAs.

## 2. Biological Functions of Fusion Genes and RNAs

Solid tumors and hematopoietic malignancies often have highly complex, unstable genomes. Many gene fusions are random events caused by genetic instability or abnormal splicing machinery [39,40]. These changes at the gene or RNA level are unlikely to result in functional nucleic acids or proteins, as they may occur in regions where there are no known genes [12]. The functions of fusion genes and RNAs are diverse and dependent on the location of the fusion junction. However, the presence of a genomic fusion in a tumor does not necessarily mean that the fusion affects cancer development or tumorigenesis. Fusion RNAs in which the fusion junction is within the protein-coding region may be largely out-of-frame, and such out-of-frame fusions are unlikely to be functional. However, out-of-frame fusion RNAs may function as long noncoding RNAs and may exhibit regulatory functions [41].

An RNA fused between two genes in-frame is translated into a novel fusion protein that may act as a potent oncogenic driver. Kinase genes are often partners in such fusions [7,12]. Kinase fusions often retain kinase activity and result in ligand-independent constitutive activation and enhanced downstream signaling that leads to carcinogenesis [7,12]. Tyrosine kinase fusions that contain kinase-encoding genes, such as *ALK*, *ROS1*, *RET*, *FGFR1/2/3* and *NTRK1*, have been detected in various types of cancer, including glioblastoma, melanoma, and carcinomas of head and neck, breast, lung, prostate, bladder, and thyroid gland [12,40,42,43]. Serine–threonine kinase fusions have also been reported [12,40,43]. These kinase fusions frequently cause activation of signaling pathways that play important roles in cell growth, survival, proliferation, and apoptosis [12,40,42,43]. In addition, kinases are ideal targets for cancer therapy; several inhibitors against kinases, such as *ALK* and *BRAF*, have been used to treat cancers with fused genes [12,44].

Fusion of transcription factors usually produces a fusion protein that leads to constitutive activation or an altered target gene, providing aberrant transcriptional machinery and cell transformation [12]. For example, the *TMPRSS2–ERG* fusion protein, the most common fusion in prostate cancers, mediates overexpression of E26 transformation-specific (ETS) family transcription factors in response to androgen and thus aberrantly acti-

vates downstream oncogenes that play important roles in many biological processes, including cell proliferation, angiogenesis, and invasiveness [45,46]. Moreover, the EWSR1–FLI1 fusion protein can gain the ability to bind to the genome and change the transcriptional mechanism [47,48]. Transcription factor fusions can induce a wide range of phenotypic changes that initiate or promote tumorigenesis. However, they have been generally more difficult to work with as therapeutic targets than kinase fusions [37].

### 3. Biosynthesis Patterns of Fusion Genes and RNAs

Fusion RNAs are known to be generated by three mechanisms [7,49,50]. The best-understood of these is chromosomal rearrangement. Two other mechanisms are grouped together as “splicing”. One is trans-splicing, in which exons from two separate RNA transcripts are spliced together. The other is cis-splicing, which involves adjacent genes on the same strand.

RNA splicing is a form of RNA processing in which a newly made precursor messenger RNA (pre-mRNA) is transformed into a mature messenger RNA (mRNA) [51,52]. It has important functions in regulating the RNA and protein diversity observed in organisms [51,52]. Pre-mRNA splicing involves recognizing and removing noncoding regions (intron excision) and the concomitant joining of coding regions (exon ligation) to produce mature mRNA. For many eukaryotic introns, splicing is performed in a series of reactions that are catalyzed by the spliceosome, a complex of small nuclear ribonucleoproteins that incorporates stepwise assembly and disassembly by several hundred proteins and five small nuclear RNAs [53–55]. Mechanisms for trans- and cis-splicing between neighboring genes are not well understood. In this section, we review the mechanism of fusion genes and RNA generation that are currently known.

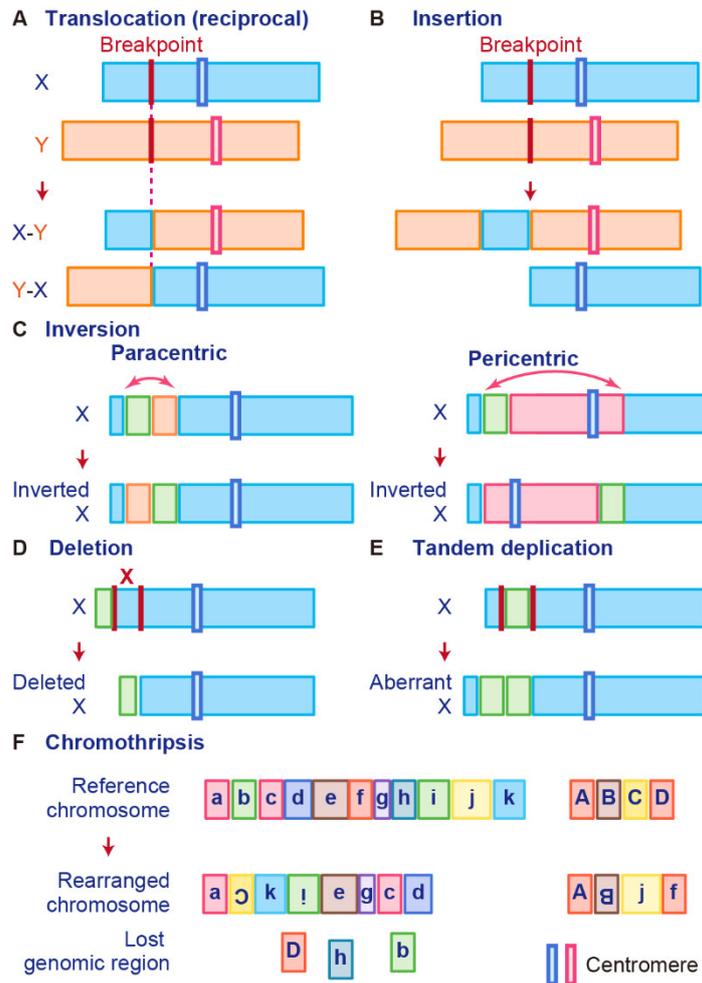
#### 3.1. Chromosomal Rearrangement

Gene fusions are usually caused by alterations in genomic structure resulting from DNA damage and by subsequent erroneous recombination and replication [7]. Genomic rearrangements can occur between one or two independent genes through six different known mechanisms: translocation, insertion, inversion, tandem duplication, deletion, and chromothripsis (Figure 1) [7,50,56]. Inversions, tandem duplications, and chromosomal deletions can occur within one or two adjacent genes, whereas translocations and insertions represent large-scale genomic aberrations that result from interactions between distant regions of the genome (interchromosomal rearrangements) or within the same chromosome (intrachromosomal rearrangements) [7,50,56]. As a result, gene fusions can produce aberrant fusion RNAs and proteins that may activate, reduce, or eliminate their original functions.

A chromosomal translocation is an exchange of parts between two nonhomologous chromosomes, also called a reciprocal translocation. Chromosomal translocation can occur anywhere between any two chromosomes (Figure 1A).

The second type of translocation is an insertion. Insertions are caused by transfer of DNA fragments from one region to another within the same chromosome (intrachromosome) or from one chromosome to another (interchromosome). The latter is also known as a nonreciprocal nonmutual translocation (Figure 1B). The *BCR–ABL1*, the first oncogenic fusion gene ever identified, is formed by a reciprocal chromosomal translocation. *BCR–ABL1*, which is generated from translocation t(9; 22) (q34; q11), is characteristic of CML [9] and also found in acute lymphoid leukemia (ALL) [57] and acute myelogenous leukemia (AML) [58]. The fusion gene *BCR–ABL1* has a constitutive tyrosine kinase activity, which leads to sustained stimulation on proliferation of cancer cells [8]. One chromosomal abnormality discovered after *BCR–ABL1* was a consistent nonhomologous balanced translocation between chromosomes 8 and 21 in leukemia patients [59]. This t(8; 21) translocation is one of the most common genetic defects in AML; it gives rise to the *RUNX1–RUNX1T1* fusion gene (previously called *AML1–ETO*) [60–62]. The *RUNX1–*

*RUNX1T1* fusion protein interacts with other proteins to repress transcription and induce leukemogenesis in myeloid progenitor cells [60,63].



**Figure 1.** Schematic representation of fusion gene formation by structural chromosome rearrangements. (A) Translocation. (B) Insertion. (C) Inversion. (D) Deletion. (E) Tandem duplication. (F) Chromothripsis.

Promyelocytic leukemia (*PML*)–retinoic acid receptor alpha (*RARα*) is a fusion RNA found in almost 95% of acute promyelocytic leukemia (APL). It is generated from the t(15;17) reciprocal translocation in APL [64–68]. *PML* is a key component of *PML* bodies and many proteins have been associated with *PML* in cells [69]. The *PML*–*RARα* protein disrupts *PML* bodies and induces the formation of dispersed microspeckles with the loss of transcriptional activation ability [70,71]. In addition to leukemia, fusion events also occur in Ewing’s sarcoma (EWS). Many EWS cases involve t(11; 22) or t(21; 22) translocation that fuses the 5’ end of the Ewing’s sarcoma breakpoint region-1 (*EWSR1*) gene to the 3’ end of the *FLI1* or *ERG* gene, which generates the fusion genes *EWS–FLI1* or *EWS–ERG*, respectively [47,48,72]. These fusion transcription factors upregulate genes related to the cell cycle, invasion, and proliferation pathways [73–76]. Interestingly, the prion-like domain of *EWS–FLI1*, which is necessary for phase transitions, induces recruitment of BRG1/BRM-associated factors (BAF) complexes to *GGAA* microsatellites that are frequently found in oncogenes and activate the transcription of target genes [77].

A chromosome inversion occurs when a chromosome undergoes a break or rearrangement within a single chromosome (Figure 1C). There are two types of inversions: paracentric and pericentric. Paracentric inversions do not involve centromeres and both

breaks occur in a single chromosome arm, whereas pericentric inversions include a centromere, with one break in each arm (Figure 1C). Many chromosomal rearrangements identified in radiation-induced tumors are known to be paracentric inversions. The most common is the *RET* fusion in papillary thyroid carcinoma, which is present in up to 80% of radiation-related tumors [78]. As another example, in 2007, *EML4-ALK* was identified as a novel fusion oncogenic driver of non-small cell lung cancer (NSCLC) [79]. The fusion of *EML4* with *ALK* is caused by an inversion of chromosome 2 (*inv2*) (p21:p23), by which the kinase domain of the receptor-type tyrosine kinase *ALK* is placed under the control of the constitutive promoter of *EML4* [80]. This fusion allows cancer transformation by activating downstream reactions in the *ALK* signaling pathway [80].

Chromosome deletion, which is the fourth type of rearrangement, is an alteration in which a chromosome fragment is lost during DNA replication (Figure 1D). This chromosome deletion causes the deletion of intergenic regions between two genes that are side by side, and leads to formation of fusion genes by aligning two genes that are transcribed in the same direction. An example is *TMPRSS2-ERG*, generated in prostate cancer via an intron deletion between *TMPRSS2* and *ERG* on chromosome 21q22.2-3 [81,82]. Transmembrane serine protease 2 (*TMPRSS2*) is a prostate-specific androgen-regulated protein, and ETS-related gene (*ERG*) belongs to the ETS family of transcription factors, which can be oncogenic [81–83]. *TMPRSS2-ERG* fusion is reportedly associated with higher tumor stage, increased risk of disease progression, and bone metastasis [46,84].

In tandem duplication, a genomic region is duplicated and fused with a gene from the original region (Figure 1E). *FGFR3-TACC3* in glioblastoma is an example of tandem duplication [85]. A tandem duplication also occurs at 7q34 in pilocytic astrocytoma, resulting in a *KIAA1549-BRAF* fusion gene that exhibits constitutive kinase activity. Moreover, a tandem duplication that leads to *C2orf44-ALK* fusion occurs in-frame on chromosome 2 in colorectal cancer, resulting in overexpression of the *ALK* kinase [86].

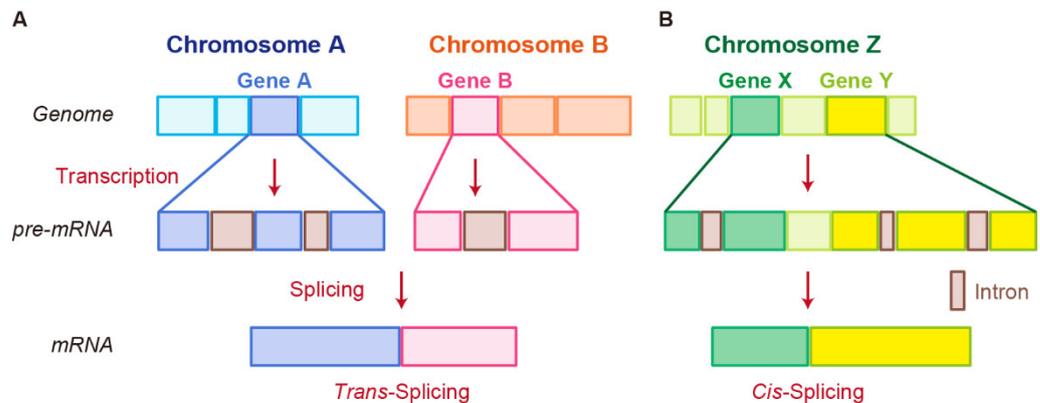
Tumor cells are associated with high genomic instability, and fusions can occur as a result of complex processes involving several rare and/or complex genetic rearrangements. Chromothripsis, which is the sixth type of rearrangement, occurs when a single chromosome, chromosome region, or a small number of chromosomes are shattered into many fragments and the fragments reassemble incorrectly (Figure 1F). Chromothripsis can produce a large number of fusion genes in a single event [87]. The most typical examples of this characteristic event are *PVT1-MYC* and *PVT1-NDRG1* fusions in medulloblastoma. Chromothripsis in medulloblastoma leads to recurrent translocations that eventually fuse a lncRNA *PVT1* to *MYC*, resulting in a continuous oncogenic effect via *MYC* amplification [88].

### 3.2. Trans-splicing

In trans-splicing, exons from different RNA transcripts are spliced and fused together to produce a mature mRNA (Figure 2A) [89–91]. Trans-splicing produces RNAs with exon repetitions or shuffling, as well as RNAs composed of exons transcribed from opposite strands [92–94]. The molecular mechanisms of trans-splicing in vertebrates are largely unexplored; however, several models have been proposed [41]. One model is spliceosome-mediated trans-splicing, which uses a canonical splice site of two different primary RNAs. It shows that the spliceosome mechanism, a fundamental component of splicing, can generate trans-spliced fusion RNAs [95–98].

Examples of trans-spliced fusion RNAs are *JAZF1-JJAZ1* (*SUZ12*) [89] and *PAX3-FOXO1* [99]. In both cases, identical fusions were found as structural chromosomal rearrangements from human tumor tissues and RNA trans-splicing from normal human tissue. *JAZF1-SUZ12*, which is composed of the first 3 exons of *JAZF1* and the last 15 exons of *SUZ12*, canonically resulted from a recurrent translocation t(7;17)(p15;q21) in endometrial stromal tumors. Identical fusion RNA was detected in normal endometrial cells [89]. Whereas *PAX3-FOXO1* with t(2;13) translocation was detected in rhabdomyosarcoma, *PAX3-FOXO1* fusion RNA produced by trans-splicing was transiently present in cells that

underwent differentiation from pluripotent cells into skeletal muscle [99]. In these cases, different mechanisms at genomic or RNA levels may generate identical fusion RNAs, leading to different pathological outcomes.



**Figure 2.** Schematic representation of fusion RNA formation by nonstructural chromosome rearrangements. **(A)** Trans-splicing. **(B)** Cis-splicing.

### 3.3. Cis-splicing

Another splicing mechanism is cis-splicing, in which two neighboring genes are transcribed into a single precursor RNA by transcriptional read-through, followed by RNA splicing between the exons of the two neighboring genes to complete the fusion (Figure 2B) [7,100]. Although cancer-associated fusion RNAs from cis-splicing are uncommon, some have been reported with clear oncogenic roles. *SLC45A3-ELK4* has been discovered by two independent groups and is a potential biomarker in prostate cancer [101,102]. Several *SLC45A3-ELK4* fusions have been reported; the primary form is a fusion of *SLC45A3* exon 1 with the last four exons of *ELK4*. Interestingly, although *SLC45A3-ELK4* functions as a fused lncRNA, its knockdown in cancer cells leads to reduction in cell proliferation, despite its unclear oncogenic roles [36]. *RBM6-RBM5* is found in several types of cancer; its expression is associated with the size of breast tumors [103]. *DUS4L-BCAP29* is found in gastric and prostate cancers and plays a tumor-promoting role in gastric cancer [104,105]. However, *DUS4L-BCAP29* is also present in normal tissues and has a growth-promoting effect in normal as well as cancerous tissues [106].

## 4. Detection of Fusion Genes and RNAs

### 4.1. Guided Approaches to Detect Fusions

The first fusion gene was discovered by chromosome banding techniques in hematologic malignancies. This technique allows each chromosome and chromosome region to be identified on the basis of its unique band pattern, thus subtle rearrangements that were previously undetectable could be found [4,100]. Other techniques that can be used to detect gene fusions include fluorescence in situ hybridization (FISH), polymerase chain reaction (PCR), and comparative genomic hybridization (CGH) array [4,12,100]. FISH can visualize different chromosome structures simultaneously in different colors and greatly improve the depiction of breakpoints in both nondividing cells and on metaphase chromosomes with structural rearrangements [107]. In 1983, PCR technology was developed that would later revolutionize not only biochemistry and molecular biology but diagnostic techniques to detect SARS-CoV-2 [108,109]. However, fusion RNAs that are generated by trans- or cis-splicing cannot be detected by the DNA-based assays described above because they are produced without chromosomal rearrangement. Technologies used to detect these fusion RNAs are RT-PCR, Northern blotting, and RNase protection assays [100].

The cytogenetic approach, which requires indirect and incremental steps, has some important weaknesses [4]. Analysis with chromosome banding techniques requires live cells and rapid transport of tumor samples to the laboratory for cell culture. In addition, highly malignant tumors often have complex genomes, which complicates assigning rearrangements to specific chromosomal bands [4]. The development of array-based platforms for gene expression and copy number profiling in the 1990s provided a new approach to enable the detection of fused genes [110]. Array-based platforms not only enabled a genome-wide view, but provided much higher resolution than chromosome banding techniques, and did not require cell cultures to maintain live cells [110].

#### 4.2. Unbiased Approaches to Detect Fusions

The advent of low-cost, high-throughput, massively parallel sequencing technology has revolutionized sequencing by enabling simultaneous generation of thousands to millions of read sequences [111–113]. Moreover, sequencing technology makes it possible to identify fusions at the DNA and RNA levels in a single experiment. By simultaneously obtaining detailed and comprehensive information about the genome or transcriptome, identification of structural variants and fusion transcripts without prior information about the cytogenetic characteristics of the tumor tissue is now possible [4]. Detection of chromosomal rearrangements and fusion RNAs has become much easier, compared with the guided approaches described above.

Advances in both sequence technology and computational capabilities have enabled discovery of novel fusion gene or RNA candidates through bioinformatics techniques. However, whereas the sequencing technology is highly sensitive and allows for detection of rare events, it is also error-prone [114,115]. Sequencing errors may be introduced before sequencing or arise from use of a sequencing platform [115]. In addition, several bioinformatics pipelines have been developed to facilitate the identification of chimeric transcripts; each of them has its own features, strengths, and weaknesses [116–118]. Algorithms for detecting fusion RNAs differ considerably in their sensitivity and specificity [119], and both false positive and false negative findings from these tools are common [116–118]. Therefore, candidate fusion RNAs obtained through these tools should be experimentally validated using guided approaches. Experimental verification of potential fusions is usually performed using RT-PCR reactions, utilizing primers that span the junction positions of the fusion RNAs. However, template switching and creation of false fusions by stem loops can occur during the RT step of cDNA preparation [120,121]. Therefore, fusions detected by sequencing technology should be validated using non-RT-based assays, such as RNase protection assays and Northern blot analysis [122].

#### 4.3. Databases for Fusion Genes and RNAs

Sequencing technology to detect cancer-associated fusion genes and RNAs has processed established cell lines, followed by a wide range of primary tissue samples from several cancer types, including carcinomas of the breast, colon, lung, prostate, and uterus, as well as leukemias and lymphomas [101,123–129]. The Pan-Cancer Analysis of Whole Genomes (PCAWG) Consortium of the International Cancer Genome Consortium (ICGC) and The Cancer Genome Atlas (TCGA) have also detected various fusions in cancer tissue genomes [5,130–132]. Moreover, thousands of fusion RNAs and genes are now listed in several databases, including Mitelman, ChimerDB, FusionGDB, FusionGDB, and others (Table 1) [49,133]. However, many of these newly identified candidates have not yet been experimentally validated, and some may represent artifacts of the sequencing processes. Moreover, information on fusion transcripts in these databases lacks uniformity [133].

**Table 1.** Available databases hosting fusion transcripts.

Database Name	Tumor	Non-Tumor	First Release	Last Update	Current Version	Total Fusions	URL
Mitleman	Yes	No	1983	2020		32,578	<a href="https://mitelmandatabase.isb-cgc.org">https://mitelmandatabase.isb-cgc.org</a>
ChimerDB	Yes	Yes	2006	2020	4.0	67,610	<a href="https://www.kobic.re.kr/chimerdb_mirror/">https://www.kobic.re.kr/chimerdb_mirror/</a>
FusionGDB	Yes	No	2019	2020		43,895	<a href="https://ccsm.uth.edu/FusionGDB/">https://ccsm.uth.edu/FusionGDB/</a>
COSMIC	Yes	No	2004	2020	92	19,369	<a href="https://cancer.sanger.ac.uk/cosmic">https://cancer.sanger.ac.uk/cosmic</a>
ChiTaRs	Yes	Yes	2012	2019	5.0	23,167	<a href="http://chitars.md.biu.ac.il/">http://chitars.md.biu.ac.il/</a>
Tumor Fusion Gene Data Portal	Yes	Yes	2015	2018		20,731	<a href="https://www.tumorfusions.org/">https://www.tumorfusions.org/</a>
FusionHub	Yes	Yes	2018	2018		150,699	<a href="https://fusionhub.persistent.co.in.">https://fusionhub.persistent.co.in.</a>
TICdb	Yes	No	2007	2013	3.3	1374	<a href="https://genetica.unav.edu/TICdb/">https://genetica.unav.edu/TICdb/</a>
dbCRID	Yes	Yes	2010	2010	0.9	2643	<a href="http://c1.accurascience.com/dbCRID/">http://c1.accurascience.com/dbCRID/</a>

## 5. Clinical Relevance of Fusion Genes

As gene and RNA fusions tend to be tumor-specific, they can be used as biomarkers to identify cancer types [101]. *BCR-ABL1* has been widely used as a biomarker and prognostic factor in patients with ALL [134], and its specific inhibitor, Glivec (imatinib), has been effectively used to target CML and ALL with *BCR-ABL* fusions [135]. *EML4-ALK* is used as a biomarker in NSCLC, which was treated with ALK inhibitors to improve patients' prognoses [136]. In addition, *TMPRSS2-ERG* gene fusions are expressed in early-phase prostate cancer and allow for prognostic evaluation of patients with prostate cancer [137]. Similarly, as the *EVT6-NTRK3* fusion is found in 92% of human secretory breast cancers, it is defined as a diagnostic biomarker [138]. However, although many fusion genes and RNAs have been discovered in cancers and reported as biomarkers, fusions are also present in normal tissues and cells [139–141] and may have important functions in normal physiology.

## 6. Conclusions

Targeting the oncogenic fusion proteins that are present only in cancers is a promising strategy for the development of new cancer therapeutics. This promise needs to be realized because many fusion-driven cancers have particularly poor prognoses. Clinical application of therapies that directly target fusion proteins, such as imatinib in CML, has dramatically altered clinical outcomes. Furthermore, the focus of research has expanded to include fused ncRNAs such as *PVT1-MYC* and *SLC45A3-ELK4*. In the future, therapies will likely be established that target specific fusion RNAs, including antisense oligos (ASOs), siRNAs, or microRNAs. These nucleic acid analogues undergo complementary base-pairing with their targets to promote endogenous RNA degradation and inhibit its translation. Altering the chemical composition of the phosphate backbone and sugar components of oligonucleotides has produced greater binding affinity and in vivo stability, leading to improved cellular uptake and release of these nucleic acid analogues. In addition to developing these technologies, elucidating upstream mechanisms that regulate fusion RNA expression and generation is necessary. Therefore, greater understanding of the mechanisms that underly fusion RNA regulation in cancers, with and without chromosomal rearrangements, is critical.

Advances in sequencing technology have accelerated fusion RNA research, and the implementation of large-scale sequencing projects such as TCGA and PGAWG allow identification of rare oncogenic fusions in a variety of cancer types. Furthermore, future advances in sequencing technology will accelerate the pace of chimeric RNA research. The advent of new sequencing technologies, such as low-cost whole-genome sequencing, long-read sequencing technologies [142], and single-cell full-length total RNA sequencing [143], will enable us to discover more fusion RNAs and elucidate their relationship to cancer heterogeneity. Ongoing research in this area can elucidate new mechanisms of cancer

development and tumorigenesis, which may contribute to significant improvements in therapeutic and diagnostic techniques for cancer.

**Author Contributions:** K.T. and N.A. wrote the manuscript and designed the figures. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the Japan Society for the Promotion of Science (JSPS) KAKENHI (Grant Number: 17KK0163, 18H02570, 18KT0016 and 20H04838). N.A. and K.T. were supported by Takeda Science Foundation. K.T. was supported by Kobayashi Foundation.

**Data Availability Statement:** Data sharing not applicable.

**Conflicts of Interest:** The authors declare no competing interests.

## References

1. Campbell, P.J.; Getz, G.; Korbel, J.O.; Stuart, J.M.; Jennings, J.L.; Stein, L.D.; Perry, M.D.; Nahal-Bose, H.K.; Ouellette, B.F.F.; Li, C.H.; et al. Pan-cancer analysis of whole genomes. *Nature* **2020**, *578*, 82–93, doi:10.1038/s41586-020-1969-6.
2. Vogelstein, B.; Kinzler, K.W. Cancer genes and the pathways they control. *Nat. Med.* **2004**, *10*, 789–799, doi:10.1038/nm1087.
3. Mitelman, F.; Johansson, B.; Mertens, F. The impact of translocations and gene fusions on cancer causation. *Nat. Rev. Cancer* **2007**, *7*, 233–245, doi:10.1038/nrc2091.
4. Mertens, F.; Johansson, B.; Fioretos, T.; Mitelman, F. The emerging complexity of gene fusions in cancer. *Nat. Rev. Cancer* **2015**, *15*, 371–381, doi:10.1038/nrc3947.
5. Gao, Q.; Liang, W.W.; Foltz, S.M.; Mutharasu, G.; Jayasinghe, R.G.; Cao, S.; Liao, W.W.; Reynolds, S.M.; Wyczalkowski, M.A.; Yao, L.; et al. Driver Fusions and Their Implications in the Development and Treatment of Human Cancers. *Cell Rep.* **2018**, *23*, 227–238, doi:10.1016/j.celrep.2018.03.050.
6. Nowell, P.C.; Hungerford, D.A. Chromosome Studies on Normal and Leukemic Human Leukocytes. *J. Natl. Cancer Inst.* **1960**, *25*, 85–109.
7. Dai, X.; Theobard, R.; Cheng, H.; Xing, M.; Zhang, J. Fusion genes: A promising tool combating against cancer. *Biochim. Biophys. Acta-Rev. Cancer* **2018**, *1869*, 149–160, doi:10.1016/j.bbcan.2017.12.003.
8. Shtivelman, E.; Lifshitz, B.; Gale, R.P.; Cananni, E. Fused transcript of *abl* and *bcr* genes in chronic myelogenous leukaemia. *Nature* **1985**, *315*, 550–554.
9. Laemmli, U.K. A new consistent chromosomal abnormality in chronic myelogenous leukaemia identified by quinacrine fluorescence and Giemsa staining. *Nature* **1973**, *243*, 290–293.
10. Han, C.; Sun, L.Y.; Wang, W.T.; Sun, Y.M.; Chen, Y.Q. Non-coding RNAs in cancers with chromosomal rearrangements: The signatures, causes, functions and implications. *J. Mol. Cell Biol.* **2019**, *11*, 886–898, doi:10.1093/jmcb/mjz080.
11. Chen, J.; Odenike, O.; Rowley, J.D. Leukaemogenesis: More than mutant genes. *Nat. Rev. Cancer* **2010**, *10*, 23–36, doi:10.1038/nrc2765.
12. Schram, A.M.; Chang, M.T.; Jonsson, P.; Drilon, A. Fusions in solid tumours: Diagnostic strategies, targeted therapy, and acquired resistance. *Nat. Rev. Clin. Oncol.* **2017**, *14*, 735–748, doi:10.1038/nrclinonc.2017.127.
13. Parker, B.C.; Zhang, W. Fusion genes in solid tumors: An emerging target for cancer diagnosis and treatment. *Chin. J. Cancer* **2013**, *32*, 594–603, doi:10.5732/cjc.013.10178.
14. Kas, K.; Voz, M.L.; Roijer, E.; Astrom, A.; Meyen, E.; Stenman, G.; Ven, W.J.M. Van De Promoter swapping between the genes for a novel zinc finger protein and b-catenin in pleiomorphic adenomas with t(3;8)(p21;q12) translocations. *Nat. Genet.* **1997**, *15*, 170–174.
15. Mark, J.; Dahlenfors, R.; Ekedahl, C.; Stenman, G. The mixed salivary gland tumor—A normally benign human neoplasm frequently showing specific chromosomal abnormalities. *Cancer Genet. Cytogenet.* **1980**, *2*, 231–241, doi:10.1016/0165-4608(80)90030-8.
16. Kopp, F.; Mendell, J.T. Functional Classification and Experimental Dissection of Long Noncoding RNAs. *Cell* **2018**, *172*, 393–407.
17. Palazzo, A.F.; Koonin, E.V. Functional Long Non-coding RNAs Evolve from Junk Transcripts. *Cell* **2020**, *183*, 1151–1161, doi:10.1016/j.cell.2020.09.047.
18. Rinn, J.L.; Chang, H.Y. Genome regulation by long noncoding RNAs. *Annu. Rev. Biochem.* **2012**, *81*, 145–166, doi:10.1146/annurev-biochem-051410-092902.
19. Cech, T.R.; Steitz, J.A. The noncoding RNA revolution—trashing old rules to forge new ones. *Cell* **2014**, *157*, 77–94, doi:10.1016/j.cell.2014.03.008.
20. Anastasiadou, E.; Jacob, L.S.; Slack, F.J. Non-coding RNA networks in cancer. *Nat. Rev. Cancer* **2017**, *18*, 5–18.
21. Wong, C.M.; Tsang, F.H.; Ng, I.O.L. Non-coding RNAs in hepatocellular carcinoma: Molecular functions and pathological implications. *Nat. Rev. Gastroenterol. Hepatol.* **2018**, *15*, 137–151.
22. Batista, P.J.; Chang, H.Y. Long noncoding RNAs: Cellular address codes in development and disease. *Cell* **2013**, *152*, 1298–1307, doi:10.1016/j.cell.2013.02.012.
23. Huarte, M. The emerging role of lncRNAs in cancer. *Nat. Med.* **2015**, *21*, 1253–1261, doi:10.1038/nm.3981.

24. Shirahama, S.; Miki, A.; Kaburaki, T.; Akimitsu, N. Long Non-coding RNAs Involved in Pathogenic Infection. *Front. Genet.* **2020**, *11*, 454, doi:10.3389/fgene.2020.00454.
25. Imamura, K.; Akimitsu, N. Long non-coding RNAs involved in immune responses. *Front. Immunol.* **2014**, *5*, 3–6, doi:10.3389/fimmu.2014.00573.
26. Tano, K.; Akimitsu, N. Long non-coding RNAs in cancer progression. *Front. Genet.* **2012**, *3*, 1–6, doi:10.3389/fgene.2012.00219.
27. Ransohoff, J.D.; Wei, Y.; Khavari, P.A. The functions and unique features of long intergenic non-coding RNA. *Nat. Rev. Mol. Cell Biol.* **2018**, *19*, 143–157.
28. Taniue, K.; Kurimoto, A.; Takeda, Y.; Nagashima, T.; Okada-Hatakeyama, M.; Katou, Y.; Shirahige, K.; Akiyama, T. ASBEL-TCF3 complex is required for the tumorigenicity of colorectal cancer cells. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 12739–12744, doi:10.1073/pnas.1605938113.
29. Taniue, K.; Kurimoto, A.; Sugimasa, H.; Nasu, E.; Takeda, Y.; Iwasaki, K.; Nagashima, T.; Okada-Hatakeyama, M.; Oyama, M.; Kozuka-Hata, H.; et al. Long noncoding RNA *UPAT* promotes colon tumorigenesis by inhibiting degradation of UHRF1. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 1273–1278, doi:10.1073/pnas.1500992113.
30. Yanagida, S.; Taniue, K.; Sugimasa, H.; Nasu, E.; Takeda, Y.; Kobayashi, M.; Yamamoto, T.; Okamoto, A.; Akiyama, T. ASBEL, an ANA/BTG3 antisense transcript required for tumorigenicity of ovarian carcinoma. *Sci. Rep.* **2013**, *3*, 1305, doi:10.1038/srep01305.
31. Slack, F.J.; Chinnaiyan, A.M. The Role of Non-coding RNAs in Oncology. *Cell* **2019**, *179*, 1033–1055, doi:10.1016/j.cell.2019.10.017.
32. Dixon-McIver, A.; East, P.; Mein, C.A.; Cazier, J.B.; Molloy, G.; Chaplin, T.; Lister, T.A.; Young, B.D.; Debernardi, S. Distinctive patterns of microRNA expression associated with karyotype in acute myeloid leukaemia. *PLoS ONE* **2008**, *3*, e2141, doi:10.1371/journal.pone.0002141.
33. Garzon, R.; Volinia, S.; Liu, C.G.; Fernandez-Cymering, C.; Palumbo, T.; Pichiorri, F.; Fabbri, M.; Coombes, K.; Alder, H.; Nakamura, T.; et al. MicroRNA signatures associated with cytogenetics and prognosis in acute myeloid leukemia. *Blood* **2008**, *111*, 3183–3189, doi:10.1182/blood-2007-07-098749.
34. Li, Z.; Lu, J.; Sun, M.; Mi, S.; Zhang, H.; Luo, R.T.; Chen, P.; Wang, Y.; Yan, M.; Qian, Z.; et al. Distinct microRNA expression profiles in acute myeloid leukemia with common translocations. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 15535–15540, doi:10.1073/pnas.0808266105.
35. Guarnerio, J.; Bezzi, M.; Jeong, J.C.; Paffenholz, S.V.; Berry, K.; Naldini, M.M.; Lo-Coco, F.; Tay, Y.; Beck, A.H.; Pandolfi, P.P. Oncogenic Role of Fusion-circRNAs Derived from Cancer-Associated Chromosomal Translocations. *Cell* **2016**, *165*, 289–302, doi:10.1016/j.cell.2016.03.020.
36. Qin, F.; Zhang, Y.; Liu, J.; Li, H. SLC45A3-ELK4 functions as a long non-coding chimeric RNA. *Cancer Lett.* **2017**, *404*, 53–61, doi:10.1016/j.canlet.2017.07.007.
37. Brien, G.L.; Stegmaier, K.; Armstrong, S.A. Targeting chromatin complexes in fusion protein-driven malignancies. *Nat. Rev. Cancer* **2019**, *19*, 255–269, doi:10.1038/s41568-019-0132-x.
38. Crompton, B.D.; Stewart, C.; Taylor-Weiner, A.; Alexe, G.; Kurek, K.C.; Calicchio, M.L.; Kiezun, A.; Carter, S.L.; Shukla, S.A.; Mehta, S.S.; et al. The genomic landscape of pediatric Ewing sarcoma. *Cancer Discov.* **2014**, *4*, 1326–1341, doi:10.1158/2159-8290.CD-13-1037.
39. Vogelstein, B.; Papadopoulos, N.; Velculescu, V.E.; Zhou, S.; Diaz, L.A.; Kinzler, K.W. Cancer genome landscapes. *Science* **2013**, *340*, 1546–1558, doi:10.1126/science.1235122.
40. Yoshihara, K.; Wang, Q.; Torres-Garcia, W.; Zheng, S.; Vegesna, R.; Kim, H.; Verhaak, R.G.W. The landscape and therapeutic relevance of cancer-associated transcript fusions. *Oncogene* **2015**, *34*, 4845–4854, doi:10.1038/onc.2014.406.
41. Chwalenia, K.; Facemire, L.; Li, H. Chimeric RNAs in cancer and normal physiology. *Wiley Interdiscip. Rev. RNA* **2017**, *8*, doi:10.1002/wrna.1427.
42. Kim, H.J.; Kim, N.C.; Wang, Y.D.; Scarborough, E.A.; Moore, J.; Diaz, Z.; MacLea, K.S.; Freibaum, B.; Li, S.; Molliex, A.; et al. Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS. *Nature* **2013**, *495*, 467–473, doi:10.1038/nature11922.
43. Stransky, N.; Cerami, E.; Schalm, S.; Kim, J.L.; Lengauer, C. The landscape of kinase fusions in cancer. *Nat. Commun.* **2014**, *5*, 4846, doi:10.1038/ncomms5846.
44. Shaw, A.T.; Hsu, P.P.; Awad, M.M.; Engelman, J.A. Tyrosine kinase gene rearrangements in epithelial malignancies. *Nat. Rev. Cancer* **2013**, *13*, 772–787, doi:10.1038/nrc3612.
45. Rabbitts, T.H. Commonality but Diversity in Cancer Gene Fusions. *Cell* **2009**, *137*, 391–395, doi:10.1016/j.cell.2009.04.034.
46. Deplus, R.; Delliaux, C.; Marchand, N.; Flourens, A.; Vanpouille, N.; Leroy, X.; de Launoit, Y.; Duterque-Coquillaud, M. TMPRSS2-ERG fusion promotes prostate cancer metastases in bone. *Oncotarget* **2017**, *8*, 11827–11840, doi:10.18632/oncotarget.14399.
47. Delattre, O.; Zucman, J.; Plougastel, B.; Desmaze, C.; Melot, T.; Peter, M.; Kovar, H.; Joubert, I.; De Jong, P.; Rouleau, G.; et al. Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours. *Nature* **1992**, *359*, 162–165, doi:10.1038/359162a0.
48. May, W.A.; Gishizky, M.L.; Lessnick, S.L.; Lunsford, L.B.; Lewis, B.C.; Delattre, O.; Zucman, J.; Thomas, G.; Denny, C.T. Ewing sarcoma 11;22 translocation produces a chimeric transcription factor that requires the DNA-binding domain encoded by FLI1 for transformation. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 5752–5756, doi:10.1073/pnas.90.12.5752.

49. Li, Z.; Qin, F.; Li, H. Chimeric RNAs and their implications in cancer. *Curr. Opin. Genet. Dev.* **2018**, *48*, 36–43, doi:10.1016/j.gde.2017.10.002.
50. Tuna, M.; Amos, C.I.; Mills, G.B. Molecular mechanisms and pathobiology of oncogenic fusion transcripts in epithelial tumors. *Oncotarget* **2019**, *10*, 2095–2111, doi:10.18632/oncotarget.26777.
51. Nilsen, T.W.; Graveley, B.R. Expansion of the eukaryotic proteome by alternative splicing. *Nature* **2010**, *463*, 457–463, doi:10.1038/nature08909.
52. Irimia, M.; Blencowe, B.J. Alternative splicing: Decoding an expansive regulatory layer. *Curr. Opin. Cell Biol.* **2012**, *24*, 323–332.
53. Jurica, M.S.; Moore, M.J. Pre-mRNA splicing: Awash in a sea of proteins. *Mol. Cell* **2003**, *12*, 5–14, doi:10.1016/B978-0-12-374984-0.01205-5.
54. Ule, J.; Blencowe, B.J. Alternative Splicing Regulatory Networks: Functions, Mechanisms, and Evolution. *Mol. Cell* **2019**, *76*, 329–345, doi:10.1016/j.molcel.2019.09.017.
55. Shi, Y. Mechanistic insights into precursor messenger RNA splicing by the spliceosome. *Nat. Rev. Mol. Cell Biol.* **2017**, *18*, 655–670, doi:10.1038/nrm.2017.86.
56. Neckles, C.; Sundara Rajan, S.; Caplen, N.J. Fusion transcripts: Unexploited vulnerabilities in cancer? *Wiley Interdiscip. Rev. RNA* **2020**, *11*, 1–17, doi:10.1002/wrna.1562.
57. Lucas, G.S.; Arden, J.C.; Bartram, C.R. BCR-ABL rearrangements in acute lymphoblastic leukaemia. *Lancet* **1991**, *337*, 1548.
58. Cuenco, G.M.; Ren, R. Cooperation of BCR-ABL and AML1/MDS1/EVI1 in blocking myeloid differentiation and rapid induction of an acute myelogenous leukemia. *Oncogene* **2001**, *20*, 8236–8248, doi:10.1038/sj.onc.1205095.
59. Nucifora, G.; Birn, D.J.; Erickson, P.; Gao, J.; LeBeau, M.M.; Drabkin, H.A.; Rowle, J.D. Detection of DNA rearrangements in the AML1 and ETO loci and of an AML1/ETO fusion mRNA in patients with t(8;21) acute myeloid leukemia. *Blood* **1993**, *81*, 883–888.
60. Higuchi, M.; O'Brien, D.; Kumaravelu, P.; Lenny, N.; Yeoh, E.J.; Downing, J.R. Expression of a conditional AML1-ETO oncogene bypasses embryonic lethality and establishes a murine model of human t(8;21) acute myeloid leukemia. *Cancer Cell* **2002**, *1*, 63–74, doi:10.1016/S1535-6108(02)00016-8.
61. Jurlander, J.; Caligiuri, M.A.; Ruutu, T.; Baer, M.R.; Strout, M.P.; Oberkircher, A.R.; Hoffmann, L.; Ball, E.D.; Frei-Lahr, D.A.; Christiansen, N.P.; et al. Persistence of the AML1/ETO fusion transcript in patients treated with allogeneic bone marrow transplantation for t(8;21) leukemia. *Blood* **1996**, *88*, 2183–2191.
62. Nucifora, G.; Rowley, J.D. AML1 and the 8;21 and 3;21 translocations in acute and chronic myeloid leukemia. *Blood* **1995**, *86*, 1–14, doi:10.1182/blood.v86.1.1.bloodjournal8611.
63. Yuan, Y.; Zhou, L.; Miyamoto, T.; Iwasaki, H.; Harakawa, N.; Hetherington, C.J.; Burel, S.A.; Lagasse, E.; Weissman, I.L.; Akashi, K.; et al. AML1-ETO expression is directly involved in the development of acute myeloid leukemia in the presence of additional mutations. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 10398–10403, doi:10.1073/pnas.171321298.
64. de Thé, H.; Lavau, C.; Marchio, A.; Chomienne, C.; Degos, L.; Dejean, A. The PML-RAR $\alpha$  fusion mRNA generated by the t(15;17) translocation in acute promyelocytic leukemia encodes a functionally altered RAR. *Cell* **1991**, *66*, 675–684, doi:10.1016/0092-8674(91)90113-D.
65. Goddard, A.D.; Borrow, J.; Freemont, P.S.; Solomon, E. Characterization of a zinc finger gene disrupted by the t(15;17) in acute promyelocytic leukemia. *Science* **1991**, *254*, 1371–1374, doi:10.1126/science.1720570.
66. Kakizuka, A.; Miller, W.H.; Umesono, K.; Warrell, R.P.; Frankel, S.R.; Murty, V.V.V.S.; Dmitrovsky, E.; Evans, R.M. Chromosomal translocation t(15;17) in human acute promyelocytic leukemia fuses RAR $\alpha$  with a novel putative transcription factor, PML. *Cell* **1991**, *66*, 663–674, doi:10.1016/0092-8674(91)90112-C.
67. Meroni, G.; Diez-Roux, G. TRIM/RBCC, a novel class of “single protein RING finger” E3 ubiquitin ligases. *BioEssays* **2005**, *27*, 1147–1157, doi:10.1002/bies.20304.
68. Reymond, A.; Meroni, G.; Fantozzi, A.; Merla, G.; Cairo, S.; Luzi, L.; Riganelli, D.; Zanaria, E.; Messali, S.; Cainarca, S.; et al. The tripartite motif family identifies cell compartments. *EMBO J.* **2001**, *20*, 2140–2151.
69. Van Damme, E.; Laukens, K.; Dang, T.H.; Van Ostade, X. A manually curated network of the PML nuclear body interactome reveals an important role for PML-NBs in SUMOylation dynamics. *Int. J. Biol. Sci.* **2010**, *6*, 51–67.
70. Shima, Y.; Shima, T.; Chiba, T.; Irimura, T.; Pandolfi, P.P.; Kitabayashi, I. PML Activates Transcription by Protecting HIPK2 and p300 from SCF<sup>Fbx3</sup>-Mediated Degradation. *Mol. Cell. Biol.* **2008**, *28*, 7126–7138, doi:10.1128/mcb.00897-08.
71. Di Masi, A.; Cilli, D.; Berardinelli, F.; Talarico, A.; Pallavicini, I.; Pennisi, R.; Leone, S.; Antoccia, A.; Noguera, N.I.; Lo-Coco, F.; et al. PML nuclear body disruption impairs DNA double-strand break sensing and repair in APL. *Cell Death Dis.* **2016**, *7*, e2308. doi:10.1038/cddis.2016.115.
72. Sorensen, P.H.B.; Lessnick, S.L.; Lopez-terrada, D.; Liu, X.F.; Triche, T.J.; Denny, C.T.; Cherry, H. A second Ewing’s sarcoma translocation, t(21;22), fuses the EWS gene to another ETS-family transcription factor, ERG. *Nat. Genet.* **1994**, *6*, 146–151.
73. Zinszner, H.; Immanuel, D.; Yin, Y.; Liang, F.X.; Ron, D. A topogenic role for the oncogenic N-terminus of TLS: Nucleolar localization when transcription is inhibited. *Oncogene* **1997**, *14*, 451–461, doi:10.1038/sj.onc.1200854.
74. Yang, L.; Chansky, H.A.; Hickstein, D.D. EWS-Fli-1 fusion protein interacts with hyperphosphorylated RNA polymerase II and interferes with serine-arginine protein-mediated RNA splicing. *J. Biol. Chem.* **2000**, *275*, 37612–37618, doi:10.1074/jbc.M005739200.

75. Bailly, R.; Bosselut, R.; Zucman, J.; Delattre, O.; Roussel, M.; Thomas, G.; Ghysdaell, J. DNA-binding and transcriptional activation properties of the EWS-FLI-1 fusion protein resulting from the t(11;22) translocation in Ewing sarcoma. *Mol. Cell. Biol.* **1994**, *14*, 3230–3241.
76. Braun, B.S.; Frieden, R.; Lessnick, S.L.; May, W.A.; Denny, C.T. Identification of target genes for the Ewing's sarcoma EWS/FLI fusion protein by representational difference analysis. *Mol. Cell. Biol.* **1995**, *15*, 4623–4630, doi:10.1128/mcb.15.8.4623.
77. Boulay, G.; Sandoval, G.J.; Riggi, N.; Iyer, S.; Buisson, R.; Naigles, B.; Awad, M.E.; Rengarajan, S.; Volorio, A.; McBride, M.J.; et al. Cancer-Specific Retargeting of BAF Complexes by a Prion-like Domain. *Cell* **2017**, *171*, 163–178, doi:10.1016/j.cell.2017.07.036.
78. Pierotti, M.A.; Santoro, M.; Jenkins, R.B.; Sozzi, G.; Bongarzone, I.; Grieco, M.; Monzini, N.; Miozzo, M.; Herrmann, M.A.; Fusco, A.; et al. Characterization of an inversion on the long arm of chromosome 10 juxtaposing D10S170 and RET and creating the oncogenic sequence RET/PTC. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 1616–1620, doi:10.1073/pnas.89.5.1616.
79. Soda, M.; Choi, Y.L.; Enomoto, M.; Takada, S.; Yamashita, Y.; Ishikawa, S.; Fujiwara, S.I.; Watanabe, H.; Kurashina, K.; Hatanaka, H.; et al. Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer. *Nature* **2007**, *448*, 561–566, doi:10.1038/nature05945.
80. Rikova, K.; Guo, A.; Zeng, Q.; Possemato, A.; Yu, J.; Haack, H.; Nardone, J.; Lee, K.; Reeves, C.; Li, Y.; et al. Global Survey of Phosphotyrosine Signaling Identifies Oncogenic Kinases in Lung Cancer. *Cell* **2007**, *131*, 1190–1203, doi:10.1016/j.cell.2007.11.025.
81. Pederzoli, F.; Bandini, M.; Marandino, L.; Ali, S.M.; Madison, R.; Chung, J.; Ross, J.S.; Necchi, A. Targetable gene fusions and aberrations in genitourinary oncology. *Nat. Rev. Urol.* **2020**, *17*, 613–625, doi:10.1038/s41585-020-00379-4.
82. Tomlins, S.A.; Laxman, B.; Dhanasekaran, S.M.; Helgeson, B.E.; Cao, X.; Morris, D.S.; Menon, A.; Jing, X.; Cao, Q.; Han, B.; et al. Distinct classes of chromosomal rearrangements create oncogenic ETS gene fusions in prostate cancer. *Nature* **2007**, *448*, 595–599, doi:10.1038/nature06024.
83. Hollenhorst, P.C.; McIntosh, L.P.; Graves, B.J. Genomic and biochemical insights into the specificity of ETS transcription factors. *Annu. Rev. Biochem.* **2011**, *80*, 437–471, doi:10.1146/annurev.biochem.79.081507.103945.
84. Perner, S.; Demichelis, F.; Beroukhi, R.; Schmidt, F.H.; Mosquera, J.M.; Setlur, S.; Tchinda, J.; Tomlins, S.A.; Hofer, M.D.; Pienta, K.G.; et al. TMPRSS2:ERG fusion-associated deletions provide insight into the heterogeneity of prostate cancer. *Cancer Res.* **2006**, *66*, 8337–8341, doi:10.1158/0008-5472.CAN-06-1482.
85. Singh, D.; Chan, J.M.; Zoppoli, P.; Niola, F.; Sullivan, R.; Castano, A.; Liu, E.M.; Reichel, J.; Porrati, P.; Pellegatta, S.; et al. Transforming fusions of FGFR and TACC genes in human glioblastoma. *Science* **2012**, *337*, 1231–1235, doi:10.1126/science.1220834.
86. Lipson, D.; Capelletti, M.; Yelensky, R.; Otto, G.; Parker, A.; Jarosz, M.; Curran, J.A.; Balasubramanian, S.; Bloom, T.; Brennan, K.W.; et al. Identification of new ALK and RET gene fusions from colorectal and lung cancer biopsies. *Nat. Med.* **2012**, *18*, 382–384, doi:10.1038/nm.2673.
87. Holland, A.J.; Cleveland, D.W. Chromoanagenesis and cancer: Mechanisms and consequences of localized, complex chromosomal rearrangements. *Nat. Med.* **2012**, *18*, 1630–1638, doi:10.1038/nm.2988.
88. Northcott, P.A.; Shih, D.J.H.; Peacock, J.; Garzia, L.; Sorana Morrissy, A.; Zichner, T.; Stütz, A.M.; Korshunov, A.; Reimand, J.; Schumacher, S.E.; et al. Subgroup-specific structural variation across 1,000 medulloblastoma genomes. *Nature* **2012**, *487*, 49–56, doi:10.1038/nature11327.
89. Li, H.; Wang, J.; Mor, G.; Sklar, J. A Neoplastic Gene Fusion Mimics Trans-Splicing of RNAs in Normal Human Cells. *Science* **2008**, *321*, 1357–1361.
90. Li, B.L.; Li, X.L.; Duan, Z.J.; Lee, O.; Lin, S.; Ma, Z.M.; Chang, C.C.Y.; Yang, X.Y.; Park, J.P.; Mohandas, T.K.; et al. Human acyl-CoA:cholesterol acyltransferase-1 (ACAT-1) gene organization and evidence that the 4.3-kilobase ACAT-1 mRNA is produced from two different chromosomes. *J. Biol. Chem.* **1999**, *274*, 11060–11071, doi:10.1074/jbc.274.16.11060.
91. Breen, M.A.; Ashcroft, S.J.H. A truncated isoform of Ca<sup>2+</sup>/calmodulin-dependent protein kinase II expressed in human islets of Langerhans may result from trans-splicing. *FEBS Lett.* **1997**, *409*, 375–379, doi:10.1016/S0014-5793(97)00555-3.
92. Horiuchi, T.; Giniger, E.; Aigaki, T. Alternative trans-splicing of constant and variable exons of a Drosophila axon guidance gene, *lola*. *Genes Dev.* **2003**, *17*, 2496–2501, doi:10.1101/gad.1137303.
93. Dorn, R.; Reuter, G.; Loewendorf, A. Transgene analysis proves mRNA trans-splicing at the complex *mod(mdg4)* locus in Drosophila. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 9724–9729, doi:10.1073/pnas.151268698.
94. Lei, Q.; Li, C.; Zuo, Z.; Huang, C.; Cheng, H.; Zhou, R. Evolutionary Insights into RNA trans-Splicing in Vertebrates. *Genome Biol. Evol.* **2016**, *8*, 562–577, doi:10.1093/gbe/evw025.
95. Solnick, D. Trans splicing of mRNA precursors. *Cell* **1985**, *42*, 157–164, doi:10.1016/S0092-8674(85)80111-2.
96. Konarska, M.M.; Padgett, R.A.; Sharp, P.A. Trans splicing of mRNA precursors in vitro. *Cell* **1985**, *42*, 165–171, doi:10.1016/S0092-8674(85)80112-4.
97. Eul, J.; Graessmann, M.; Graessmann, A. Experimental evidence for RNA trans-splicing in mammalian cells. *EMBO J.* **1995**, *14*, 3226–3235, doi:10.1002/j.1460-2075.1995.tb07325.x.
98. Eul, J.; Graessmann, M.; Graessmann, A. Trans-splicing and alternative-tandem-cis-splicing: Two ways by which mammalian cells generate a truncated SV40 T-antigen. *Nucleic Acids Res.* **1996**, *24*, 1653–1661, doi:10.1093/nar/24.9.1653.
99. Yuan, H.; Qin, F.; Movassagh, M.; Park, H.; Golden, W.; Xie, Z.; Zhang, P.; Sklar, J.; Li, H. A chimeric RNA characteristic of rhabdomyosarcoma in normal myogenesis process. *Cancer Discov.* **2013**, *3*, 1394–1403, doi:10.1158/2159-8290.CD-13-0186.
100. Wu, H.; Li, X.; Li, H. Gene fusions and chimeric RNAs, and their implications in cancer. *Genes Dis.* **2019**, *6*, 385–390, doi:10.1016/j.gendis.2019.08.002.

101. Maher, C.A.; Kumar-Sinha, C.; Cao, X.; Kalyana-Sundaram, S.; Han, B.; Jing, X.; Sam, L.; Barrette, T.; Palanisamy, N.; Chinnaiyan, A.M. Transcriptome sequencing to detect gene fusions in cancer. *Nature* **2009**, *458*, 97–101, doi:10.1038/nature07638.
102. Rickman, D.S.; Pflueger, D.; Moss, B.; VanDoren, V.E.; Chen, C.X.; De la Taille, A.; Kuefer, R.; Tewari, A.K.; Setlur, S.R.; Demichelis, F.; et al. SLC45A3-ELK4 is a novel and frequent erythroblast transformation-specific fusion transcript in prostate cancer. *Cancer Res.* **2009**, *69*, 2737–2738, doi:10.1158/0008-5472.CAN-08-4926.
103. Wang, K.; Ubriaco, G.; Sutherland, L.C. RBM6-RBM5 transcription-induced chimeras are differentially expressed in tumours. *BMC Genom.* **2007**, *8*, 1–12, doi:10.1186/1471-2164-8-348.
104. Nacu, S.; Yuan, W.; Kan, Z.; Bhatt, D.; Rivers, C.S.; Stinson, J.; Peters, B.A.; Modrusan, Z.; Jung, K.; Seshagiri, S.; et al. Deep RNA sequencing analysis of readthrough gene fusions in human prostate adenocarcinoma and reference samples. *BMC Med. Genom.* **2011**, *4*, 11, doi:10.1186/1755-8794-4-11.
105. Kim, H.P.; Cho, G.A.; Han, S.W.; Shin, J.Y.; Jeong, E.G.; Song, S.H.; Lee, W.C.; Lee, K.H.; Bang, D.; Seo, J.S.; et al. Novel fusion transcripts in human gastric cancer revealed by transcriptome analysis. *Oncogene* **2014**, *33*, 5434–5441, doi:10.1038/onc.2013.490.
106. Tang, Y.; Qin, F.; Liu, A.; Li, H. Recurrent fusion RNA DUS4L-BCAP29 in non-cancer human tissues and cells. *Oncotarget* **2017**, *8*, 31415–31423, doi:10.18632/oncotarget.16329.
107. Speicher, M.R.; Carter, N.P. The new cytogenetics: Blurring the boundaries with molecular biology. *Nat. Rev. Genet.* **2005**, *6*, 782–792, doi:10.1038/nrg1692.
108. Saiki, R.K.; Scharf, S.; Faloona, F.; Mullis, K.B.; Horn, G.T.; Erlich, H.A.; Arnheim, N. Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. 1985. *Biotechnology* **1992**, *24*, 476–480, doi:10.1007/BF00985904.
109. Vandenberg, O.; Martiny, D.; Rochas, O.; van Belkum, A.; Kozlakidis, Z. Considerations for diagnostic COVID-19 tests. *Nat. Rev. Microbiol.* **2020**, *14*, 1–13, doi:10.1038/s41579-020-00461-z.
110. Pinkel, D.; Albertson, D.G. Array comparative genomic hybridization and its applications in cancer. *Nat. Genet.* **2005**, *37*, S11–S17, doi:10.1038/ng1569.
111. Schuster, S.C. Next-generation sequencing transforms today's biology. *Nat. Methods* **2008**, *5*, 16–18, doi:10.1038/nmeth1156.
112. Wold, B.; Myers, R.M. Sequence census methods for functional genomics. *Nat. Methods* **2008**, *5*, 19–21, doi:10.1038/nmeth1157.
113. Shendure, J.; Ji, H. Next-generation DNA sequencing. *Nat. Biotechnol.* **2008**, *26*, 1135–1145, doi:10.1038/nbt1486.
114. Lou, D.I.; Hussmann, J.A.; McBee, R.M.; Acevedo, A.; Andino, R.; Press, W.H.; Sawyer, S.L. High-throughput DNA sequencing errors are reduced by orders of magnitude using circle sequencing. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 19872–19877, doi:10.1073/pnas.1319590110.
115. Sleep, J.A.; Schreiber, A.W.; Baumann, U. Sequencing error correction without a reference genome. *BMC Bioinformatics* **2013**, *14*, 367, doi:10.1186/1471-2105-14-367.
116. Kumar, S.; Razaq, S.K.; Vo, A.D.; Gautam, M.; Li, H. Identifying fusion transcripts using next generation sequencing. *Wiley Interdiscip. Rev. RNA* **2016**, *7*, 811–823, doi:10.1002/wrna.1382.
117. Kumar, S.; Vo, A.D.; Qin, F.; Li, H. Comparative assessment of methods for the fusion transcripts detection from RNA-Seq data. *Sci. Rep.* **2016**, *6*, 21597, doi:10.1038/srep21597.
118. Liu, S.; Tsai, W.H.; Ding, Y.; Chen, R.; Fang, Z.; Huo, Z.; Kim, S.; Ma, T.; Chang, T.Y.; Priedigkeit, N.M.; et al. Comprehensive evaluation of fusion transcript detection algorithms and a meta-caller to combine top performing methods in paired-end RNA-seq data. *Nucleic Acids Res.* **2015**, *44*, e47, doi:10.1093/nar/gkv1234.
119. Carrara, M.; Beccuti, M.; Lazzarato, F.; Cavallo, F.; Cordero, F.; Donatelli, S.; Calogero, R.A. State-of-the-Art Fusion-Finder Algorithms Sensitivity and Specificity. *BioMed Res. Int.* **2013**, *2013*, 340620.
120. Lu, Z.; Matera, A.G. Vicinal: A method for the determination of ncRNA ends using chimeric reads from RNA-seq experiments. *Nucleic Acids Res.* **2014**, *42*, 1–9, doi:10.1093/nar/gku207.
121. Houseley, J.; Tollervey, D. Apparent non-canonical trans-splicing is generated by reverse transcriptase in vitro. *PLoS ONE* **2010**, *5*, e12271, doi:10.1371/journal.pone.0012271.
122. Yu, C.Y.; Liu, H.J.; Hung, L.Y.; Kuo, H.C.; Chuang, T.J. Is an observed non-co-linear RNA product spliced in trans, in cis or just in vitro? *Nucleic Acids Res.* **2014**, *42*, 9410–9423, doi:10.1093/nar/gku643.
123. Stephens, P.J.; McBride, D.J.; Lin, M.L.; Varela, I.; Pleasance, E.D.; Simpson, J.T.; Stebbings, L.A.; Leroy, C.; Edkins, S.; Mudie, L.J.; et al. Complex landscapes of somatic rearrangement in human breast cancer genomes. *Nature* **2009**, *462*, 1005–1010, doi:10.1038/nature08645.
124. Hammerman, P.S.; Voet, D.; Lawrence, M.S.; Voet, D.; Jing, R.; Cibulskis, K.; Sivachenko, A.; Stojanov, P.; McKenna, A.; Lander, E.S.; et al. Comprehensive genomic characterization of squamous cell lung cancers. *Nature* **2012**, *489*, 519–525, doi:10.1038/nature11404.
125. Maher, C.A.; Palanisamy, N.; Brenner, J.C.; Cao, X.; Kalyana-Sundaram, S.; Luo, S.; Khrebtukova, I.; Barrette, T.R.; Grasso, C.; Yu, J.; et al. Chimeric transcript discovery by paired-end transcriptome sequencing. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 12353–12358, doi:10.1073/pnas.0904720106.
126. Muzny, D.M.; Bainbridge, M.N.; Chang, K.; Dinh, H.H.; Drummond, J.A.; Fowler, G.; Kovar, C.L.; Lewis, L.R.; Morgan, M.B.; Newsham, I.F.; et al. Comprehensive molecular characterization of human colon and rectal cancer. *Nature* **2012**, *487*, 330–337, doi:10.1038/nature11252.
127. Getz, G.; Gabriel, S.B.; Cibulskis, K.; Lander, E.; Sivachenko, A.; Sougnez, C.; Lawrence, M.; Kandoth, C.; Dooling, D.; Fulton, R.; et al. Integrated genomic characterization of endometrial carcinoma. *Nature* **2013**, *497*, 67–73, doi:10.1038/nature12113.

128. Welch, J.S.; Westervelt, P.; Ding, L.; Larson, D.E.; Klco, J.M.; Kulkarni, S.; Wallis, J.; Chen, K.; Payton, J.E.; Fulton, R.S.; et al. Use of whole-genome sequencing to diagnose a cryptic fusion oncogene. *JAMA*. **2011**, *305*, 1577–1584, doi:10.1001/jama.2011.497.
129. Roberts, K.G.; Morin, R.D.; Zhang, J.; Hirst, M.; Zhao, Y.; Su, X.; Chen, S.C.; Payne-Turner, D.; Churchman, M.L.; Harvey, R.C.; et al. Genetic Alterations Activating Kinase and Cytokine Receptor Signaling in High-Risk Acute Lymphoblastic Leukemia. *Cancer Cell* **2012**, *22*, 153–166, doi:10.1016/j.ccr.2012.06.005.
130. Li, Y.; Roberts, N.D.; Wala, J.A.; Shapira, O.; Schumacher, S.E.; Kumar, K.; Khurana, E.; Waszak, S.; Korbel, J.O.; Haber, J.E.; et al. Patterns of somatic structural variation in human cancer genomes. *Nature* **2020**, *578*, 112–121, doi:10.1038/s41586-019-1913-9.
131. Calabrese, C.; Davidson, N.R.; Demircioğlu, D.; Fonseca, N.A.; He, Y.; Kahles, A.; Van Lehmann, K.; Liu, F.; Shiraishi, Y.; Soulette, C.M.; et al. Genomic basis for RNA alterations in cancer. *Nature* **2020**, *578*, 129–136, doi:10.1038/s41586-020-1970-0.
132. Priestley, P.; Baber, J.; Lolkema, M.P.; Steeghs, N.; de Bruijn, E.; Shale, C.; Duyvesteyn, K.; Haidari, S.; van Hoeck, A.; Onstenk, W.; et al. Pan-cancer whole-genome analyses of metastatic solid tumours. *Nature* **2019**, *575*, 210–216, doi:10.1038/s41586-019-1689-y.
133. Barresi, V.; Cosentini, I.; Scuderi, C.; Napoli, S.; Di Bella, V.; Spampinato, G.; Condorelli, D.F. Fusion transcripts of adjacent genes: New insights into the world of human complex transcripts in cancer. *Int. J. Mol. Sci.* **2019**, *20*, 5252, doi:10.3390/ijms20215252.
134. Nashed, A.L.; Rao, K.W.; Gulley, M.L. Clinical applications of BCR-ABL molecular testing in acute leukemia. *J. Mol. Diagnostics* **2003**, *5*, 63–72, doi:10.1016/S1525-1578(10)60454-0.
135. Goldman, J.M.; Druker, B.J. Chronic myeloid leukemia: Current treatment options. *Blood* **2001**, *98*, 2039–2042, doi:10.2146/ajhp060525.
136. Shaw, A.T.; Yeap, B.Y.; Solomon, B.J.; Riely, G.J.; Gainor, J.; Engelman, J.A.; Shapiro, G.I.; Costa, D.B.; Ou, S.H.I.; Butaney, M.; et al. Effect of crizotinib on overall survival in patients with advanced non-small-cell lung cancer harbouring ALK gene rearrangement: A retrospective analysis. *Lancet Oncol.* **2011**, *12*, 1004–1012, doi:10.1016/S1470-2045(11)70232-7.
137. Adamo, P.; Ladomery, M.R. The oncogene ERG: A key factor in prostate cancer. *Oncogene* **2016**, *35*, 403–414, doi:10.1038/onc.2015.109.
138. Tognon, C.; Knezevich, S.R.; Huntsman, D.; Roskelley, C.D.; Melnyk, N.; Mathers, J.A.; Becker, L.; Carneiro, F.; MacPherson, N.; Horsman, D.; et al. Expression of the ETV6-NTRK3 gene fusion as a primary event in human secretory breast carcinoma. *Cancer Cell* **2002**, *2*, 367–376, doi:10.1016/S1535-6108(02)00180-0.
139. Babiceanu, M.; Qin, F.; Xie, Z.; Jia, Y.; Lopez, K.; Janus, N.; Facemire, L.; Kumar, S.; Pang, Y.; Qi, Y.; et al. Recurrent chimeric fusion RNAs in non-cancer tissues and cells. *Nucleic Acids Res.* **2016**, *44*, 2859–2872, doi:10.1093/nar/gkw032.
140. Wu, C.S.; Yu, C.Y.; Chuang, C.Y.; Hsiao, M.; Kao, C.F.; Kuo, H.C.; Chuang, T.J. Integrative transcriptome sequencing identifies trans-splicing events with important roles in human embryonic stem cell pluripotency. *Genome Res.* **2014**, *24*, 25–36, doi:10.1101/gr.159483.113.
141. Chase, A.; Ernst, T.; Fiebig, A.; Collins, A.; Grand, F.; Erben, P.; Reiter, A.; Schreiber, S.; Cross, N.C.P. TFG, a target of chromosome translocations in lymphoma and soft tissue tumors, fuses to GPR128 in healthy individuals. *Haematologica* **2010**, *95*, 20–26, doi:10.3324/haematol.2009.011536.
142. Sakamoto, Y.; Sereewattanawoot, S.; Suzuki, A. A new era of long-read sequencing for cancer genomics. *J. Hum. Genet.* **2020**, *65*, 3–10, doi:10.1038/s10038-019-0658-5.
143. Hayashi, T.; Ozaki, H.; Sasagawa, Y.; Umeda, M.; Danno, H.; Nikaido, I. Single-cell full-length total RNA sequencing uncovers dynamics of recursive splicing and enhancer RNAs. *Nat. Commun.* **2018**, *9*, 619, doi:10.1038/s41467-018-02866-0.