

REVIEW

Engineering chromosome rearrangements in cancer

Salvador Alonso^{1,2,*} and Lukas E. Dow^{1,3,*}

ABSTRACT

The identification of large chromosomal rearrangements in cancers has multiplied exponentially over the last decade. These complex and often rare genomic events have traditionally been challenging to study, in part owing to lack of tools that efficiently engineer disease-associated inversions, deletions and translocations in model systems. The emergence and refinement of genome editing technologies, such as CRISPR, have significantly expanded our ability to generate and interrogate chromosomal aberrations to better understand the networks that govern cancer growth. Here we review how existing technologies are employed to faithfully model cancer-associated chromosome rearrangements in the laboratory, with the ultimate goal of developing more accurate pre-clinical models of and therapeutic strategies for cancers driven by these genomic events.

KEY WORDS: CRISPR, Cancer, Chromosomal rearrangements, Fusion oncogenes

Introduction

Despite significant investments in cancer research, cancer death rates over the past 20 years have only seen modest annual decreases of 1.8% for men and 1.4% for women (Henley et al., 2020). Cancer therapies represent ~25% of all clinical trials in the US but <5% of all investigational drugs are ultimately approved for patient use by the FDA (Kola and Landis, 2004). The use of disease models that do not faithfully recapitulate human cancer has been partly blamed for the high rate of attrition regarding compounds that enter clinical trials and for the failure to translate scientific advances from bench to clinic (Sharpless and Depinho, 2006). Given the finite research resources, advancing technologies that more accurately mimic human disease and – at the same time – are simple, efficient and cheap, is a critical goal for driving more-effective pre-clinical studies. As our understanding of the molecular basis for cancer expands, laboratory-based models will play an even more crucial role in characterizing recurrent genetic mutations and validating targets for precision medicine approaches, ultimately narrowing the gap between preclinical and clinical scientific findings.

Historically, modeling chromosomal rearrangements and other structural variants in the laboratory has relied on transgenic approaches, in which gene products – often fusion proteins – are overexpressed under the control of an exogenous promoter (Shtivelman et al., 1985; Heisterkamp et al., 1990, 1991; Adams et al., 1985). Although these technologies are simple and efficient, and

have enabled several important basic discoveries, they do not faithfully recapitulate the events that occur during tumorigenesis. Other traditional approaches that are based on homologous recombination accurately model endogenous rearrangements but their technical complexity and low efficiency makes them an impractical tool to characterize the hundreds of structural variants that are being identified with increasing speed (Box 1). The emergence of new genome-editing technologies, in particular CRISPR, has drastically increased our ability to mirror the complexity of human disease, offering new opportunities to advance our understanding of cancer biology and, ultimately, develop more-effective treatments.

Biology of chromosomal rearrangements

In 1960, David Hungerford and Peter Nowell first described that cancer cells from patients with chronic myeloid leukemia (CML) had an abnormally short chromosome 22 (Nowell and Hungerford, 1960). This represented the first cytogenetic defect linked to cancer and was named the Philadelphia chromosome after the city in which it was discovered (Nowell, 2007). Analysis of various tumor types in the years that followed revealed that most cancers are associated with chromosomal rearrangements that were more extensive as the disease progressed (Sandberg, 1966). However, it was unknown whether these aberrations are a cause or a consequence of the oncogenic process. As cytogenetic and molecular techniques improved over the next two decades, subsequent studies revealed that the Philadelphia chromosome results from translocation of chromosomes 9 and 22, generating a fusion tyrosine kinase protein between the breakpoint cluster region (BCR) and the tyrosine-protein kinase ABL1 (BCR-ABL). Ultimately, these pivotal studies led to the discovery of imatinib, the first targeted therapy approved for cancer treatment (Druker et al., 2001).

Chromosomal rearrangements are large genomic alterations that result from double-strand DNA breaks (DSBs) at two different loci, which are then aberrantly repaired by non-homologous end joining (Richardson and Jasin, 2000). Chromosomal rearrangements are arbitrarily defined as involving ≥50 base pairs; they are classified as balanced when there is an even exchange of genetic material between two loci, such as reciprocal translocations and inversions, and as unbalanced when parts of a chromosome are lost or gained, such as insertions, duplications and deletions (Table 1). These rearrangements drive tumor growth by disrupting tumor-suppressor genes, altering gene copy number, creating oncogenic fusion proteins or juxtaposing a gene with the regulatory elements of another gene (Li et al., 2020).

In recent years, the exponential increase in cancer genome sequencing has led to the identification of thousands of novel recurrent chromosomal rearrangements (see Mitelman Database of Chromosome Aberrations and Gene Fusions in Cancer; Box 1). These discoveries have led to groundbreaking treatments in select groups of patients. For instance, the use of small-molecule kinase inhibitors substantially improved treatment-response rates in patients with ALK receptor tyrosine kinase [*ALK*- (Kwak et al., 2010; Solomon et al., 2014)], *RET*- (Drilon et al., 2020) and

¹Sandra and Edward Meyer Cancer Center, Weill Cornell Medicine, New York, NY 10021, USA. ²Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY 10021, USA. ³Department of Medicine, Weill Cornell Medicine, New York, NY 10065, USA.

*Authors for correspondence (alonsoms@mskcc.org; lud2005@med.cornell.edu)

 L.E.D., 0000-0001-7048-1418

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution and reproduction in any medium provided that the original work is properly attributed.

Box 1. Cataloging structural variants

The exponential increase of next-generation sequencing (NGS) technologies has enabled the use of high-throughput genomic analysis in patient care and the diagnosis of actionable genomic changes. For instance, established clinical diagnostic tests, such as MSK-IMPACT are designed to detect known fusions of a small subset of genes, i.e. ALK, ROS1 and RET. Identification of other clinically actionable fusions is possible through both DNA- and RNA-based methods using direct amplification or capture-based sequencing (Benayed et al., 2019; Reeser et al., 2017; Heydt et al., 2021). For mutation discovery, research studies often employ transcriptome (RNAseq), whole-exome (WES) or whole-genome sequencing (WGS), which enable the identification of novel cancer-associated chromosome rearrangements (Seshagiri et al., 2012; Ju et al., 2012). RNAseq is particularly effective to identify expressed fusion products and can pinpoint uncharacterized drivers (Benayed et al., 2019). WGS, although more expensive, allows the identification of structural variants in non-coding regions (Rheinbay et al., 2020).

ROS1-rearranged cancers (Shaw et al., 2014). Although such rearrangements – which drive the expression of constitutively active kinases, are of particular interest as they represent ‘druggable targets’, the oncogenic potential of the vast majority of recurrent structural variants remains untested.

For patients whose tumors harbor select oncogenic rearrangements, large randomized clinical trials have demonstrated improved outcomes after treatment with small-molecule inhibitors compared with chemotherapy. In the PROFILE 1014 trial, treatment with crizotinib was associated with longer progression-free survival (PFS) (10.9 months versus 7.0 months $P<0.001$) and improved response rates (74% versus 45%, $P<0.001$) among treatment-naïve patients diagnosed with *ALK*-rearranged non-small cell lung cancer (NSCLC) compared with patients who had received chemotherapy (Solomon et al., 2014). Similarly, the ASCEND-5 trial showed a significant improvement in PFS with ceritinib compared to chemotherapy in patients with *ALK*-rearranged NSCLC who had previously received crizotinib (5.4 months versus 1.6 months, $P<0.001$) in (Shaw et al., 2017). Randomized trials evaluating small-molecule inhibitors in patients diagnosed with *ROS1*-, *NTRK1*- and *RET*-rearranged cancer are ongoing but preliminary phase 1 and phase 2 studies have shown promising results, with improved outcomes compared to historical controls undergoing chemotherapy (Shaw et al., 2014; Drlon et al., 2018; 2020).

Rarely found, rarely studied – the importance of engineering rearrangements

The study of cancer-associated mutations has often relied on patient-derived cell lines that carry a particular genotype. Although this strategy has facilitated the characterization of common oncogenic drivers associated with hematologic malignancies, for which patient samples are more easily accessible, efforts to study rare or diverse large-scale genomic events in carcinomas are often hampered by the lack of clinical specimens. Genetically engineered preclinical models provide a platform to study rare oncogenic drivers and offer some advantages over the traditional patient-derived systems. In particular, the ability to build models with any combination of cooperating events expands the ‘genetic space’ in which fusions can be investigated, and provides a platform to more thoroughly test new therapies. Furthermore, the generation of chromosome rearrangements in murine models enables studying cancer initiation and progression in the context of immunocompetent hosts.

Preclinical models may also provide proof-of-concept to test tailored therapeutic strategies in subgroups of patients for whom standard therapies do not exist or are ineffective. For instance, cancers harboring *ROS1* fusions often respond poorly to standard chemotherapy but are exquisitely sensitive to small-molecule kinase inhibitors (Solomon et al., 2014). Securing a meaningful number of clinical specimens to identify prognostic and predictive biomarkers would be challenging, as *ROS1* rearrangements are present in only 1-2% of all NSCLC cases. The use of preclinical models for these and other relatively rare but clinically significant alterations may accelerate the understanding of rare genomic events, and the development of novel therapeutic strategies (Arai et al., 2013). Although the individual incidence of cancers driven by specific gene fusions is low, collectively, they represent a high number of patients who may gain significant clinical benefit.

Traditional tools to model chromosomal rearrangements

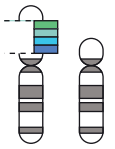

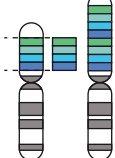

All chromosomal rearrangements – translocations, inversions, deletions and duplications – require the induction of DSBs at two separate loci and joining of otherwise unrelated genomic fragments. Not surprisingly, the efficient induction of specific chromosome rearrangements in the laboratory has proven difficult. Traditional approaches based on ectopic transgene expression, homologous recombination or Cre-loxP (Table 2) either poorly model endogenous rearrangements or are too inefficient to allow rapid characterization of newly identified cancer-associated mutations (Torres et al., 2014; Collins et al., 2000; Piganeau et al., 2013). Since 2014, older technologies have largely been replaced with CRISPR-based strategies that offer efficiency, simplicity and flexibility. In the following sections, we discuss the advantages and limitations of different genome-editing technologies for engineering chromosome rearrangements.

Fusion gene expression

A large proportion of chromosome rearrangements result in the *de novo* generation of gene and protein fusions. As such, many efforts to model gene fusions have involved the expression of fusion protein-encoding cDNAs from heterologous promoters. In these systems, the fusion transgene of interest is introduced into cell lines or embryonic stem cells via lentiviral or transgenic expression cassettes. The approach was first used in the 1980s and 90s to engineer leukemia and lymphoma models driven by the *IGH-MYC* and *BCR-ABL* translocations, and quickly became the most widely used technique to model chromosomal rearrangements (Shtivelman et al., 1985; Heisterkamp et al., 1990, 1991; Adams et al., 1985). A major advantage of this system is its simplicity: virtually any oncogene or gene fusion can be cloned into an expression vector and rapidly employed to engineer transgenic cell lines or mice. For *in vivo* experiments, the ectopic transgenic method offers the added benefits of high penetrance and short tumor latency, cutting research costs by allowing the experimental cohorts to be maintained at a reasonable size (Sharpless and Depinho, 2006). For example, Soda and colleagues engineered mice that ectopically express the *ELM4-ALK* fusion protein in alveolar epithelial cells under the control of the surfactant C promoter (Soda et al., 2008). A few weeks after birth, *ELM4-ALK* transgenic mice developed hundreds of pulmonary nodules, allowing researchers to rapidly and reproducibly test the effect of small-molecule kinase inhibitors in brief and simple experiments (Soda et al., 2008).

Highly penetrant and aggressive phenotypes, by contrast, come with the trade-off of poorly recapitulating the stochastic evolution of

Table 1. Types of chromosomal rearrangement and notable clinical examples

	Cytogenetic abnormality	Fusion product	Function of fusion product [†]	Associated malignancies	
Deletion		del(2)(p21p23)	<i>ELM4-ALK*</i>	RTK	NSCLC
		del(8)(q23q23)	<i>EIF3E-RSPO2*</i>	Ligand	CRC
		del(19)(p13p13)	<i>DNAJB1-PRKACA</i>	Kinase	Fibrolamellar hepatocarcinoma
		del(8)(q13q21)	<i>HEY1-NCOA4</i>	TF	Sarcoma
Inversion		Inv(1)(q23q31)	<i>TPR-NTRK1*</i>	RTK	NSCLC, thyroid, cholangiocarcinoma
		Inv(2)(p21p23)	<i>EML4-ALK*</i>	RTK	NSCLC and thyroid cancer
		Inv(10)(p11q11)	<i>KIF5B-RET*</i>	RTK	NSCLC
		Inv(6)(q22q25)	<i>EZR-ROS1*</i>	RTK	NSCLC
		Inv(6)(q22q22)	<i>PTPRK-RSPO3*</i>	Ligand	CRC
Amplification		dup(4)(p16p16)	<i>FGFR-TACC3*</i>	RTK	Urothelial, glioblastoma
		amp(8q24)	<i>MYC</i>	TF	Lymphoma, ovarian, breast, CRC, lung
		amp(17q12)	<i>Her2*</i>	RTK	Breast, esophageal, gastric, CRC
		amp(7p11)	<i>EGFR*</i>	RTK	NSCLC, colorectal, urothelial
		amp(11q13)	<i>CCND1</i>	Cyclin	Breast, urothelial, NSCLC
Translocation		t(9;22)(q34;q11)	<i>BCR-ABL*</i>	TK	CML, ALL
		t(8;21)(q22;q22)	<i>AML1-ETO</i>	TF	AML
		t(2;5)(p23;q35)	<i>NPM1-ALK</i>	RTK	AML
		t(15;17)(q24;q21)	<i>PML-RARA*</i>	TF	APL
		t(11;22)(q24;q12)	<i>EWS-ETS</i>	TF	Ewing sarcoma
		t(12;15)(p13;p25)	<i>ETV6-NTRK3*</i>	RTK	Thyroid, colon
	t(1;1)(q23;q23)	<i>BCAN-NTRK1*</i>	RTK	Glioblastoma	

Abbreviations: *ABL*, tyrosine-protein kinase ABL; *ALK*, ALK receptor tyrosine kinase; ALL, acute lymphocytic leukemia; AML, acute myeloid leukemia; *AML1*, RUNX family transcription factor 1 (officially known as *RUNX1*); APL, acute promyelocytic leukemia; *BCR*, breakpoint cluster region protein; *BCAN*, brevicin; *CCND1*, cyclin D1; CML, chronic myeloid leukemia; CRC, colorectal cancer; *EIF3E*, eukaryotic translation initiation factor 3 subunit E; *DNAJB1*, DnaJ heat shock protein family (Hsp40) member B1; *EGFR*, epidermal growth factor receptor; *EML4*, EMAP-like 4; *ETO*, RUNX1 partner transcriptional co-repressor 1 (officially known as *RUNX1T1*); *ETS*, ETS proto-oncogene 1; *ETV6*, ETS variant transcription factor 6; *EWS*, EWS RNA-binding protein 1 (officially known as *EWSR1*); *EZR*, ezrin; *FGFR*, fibroblast growth factor receptor; *Her2*, erb-b2 receptor tyrosine kinase 2 (officially known as *ERBB2*); *HEY1*, hairy/enhancer-of-split related with YRPW motif 1; *KIF5B*, kinesin family member 5B; *MYC*, MYC proto-oncogene; *NCOA4*, nuclear receptor coactivator 4; *NPM1*, nucleophosmin 1; NSCL, non-small cell lung cancer; *NTRK1*, neurotrophic receptor tyrosine kinase 1; *NTRK3*, neurotrophic receptor tyrosine kinase 3; *PML*, promyelocytic leukemia protein; *PRKACA*, cAMP-dependent protein kinase catalytic subunit alpha; *PTPRK*, protein tyrosine phosphatase receptor type K; *RARA*, retinoic acid receptor alpha; *RET*, ret proto-oncogene; *ROS1*, proto-oncogene tyrosine-protein kinase ROS; *RSPO2*, R-spondin 2; *RSPO3*, R-spondin 3; RTK, receptor tyrosine kinase *TACC3*, transforming acidic coiled-coil-containing protein 3; *TPR*, translocated promoter region; TF, transcription factor; TK, tyrosine kinase.

*The rearrangement is druggable with standard or investigational therapies.

[†]The N-terminal partner typically drives overexpression of the fusion product, amplifying the oncogenic function of the C-terminal fusion protein.

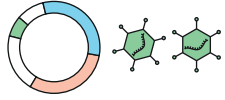
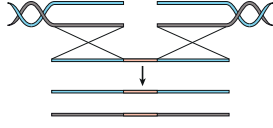
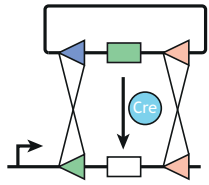
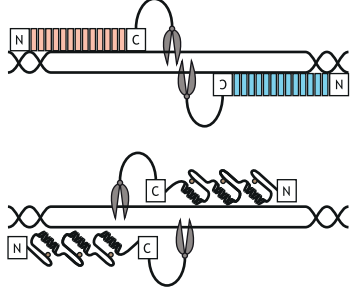
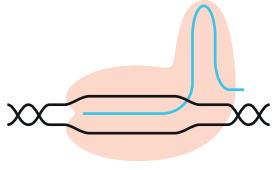
human cancer, as the emergence of synchronous and multifocal large tumors is likely to blunt microenvironmental interactions and the effect of secondary oncogenic events (Sharpless and Depinho, 2006). These features are a consequence of what is, perhaps, the main limitation of transgenic approaches – supraphysiologic oncogene expression that does not mimic the tumor context. For instance, the pioneering mouse models of *BCR-ABL* leukemia were associated with embryonic or early postnatal lethality, a stark contrast to the indolent behavior of CML in humans (Heisterkamp et al., 1990, 1991; Sawyers, 1999). Similarly, mouse models of *MYC*-rearranged B-cell lymphoma and *ALK*-rearranged NSCLC displayed highly penetrant phenotypes, precluding the study of the stochastic events driving cancer growth (Adams et al., 1985; Soda et al., 2008). Selecting a tissue-specific promoter can help limit the effects to the cellular compartment of interest, but even transient or low-level transgene expression carries the risk of transforming off-target tissues (Chiarle et al., 2003).

One approach to limit uniform tissue transformation and unwanted transgene expression in non-target tissues is a ‘mosaic’ or somatic introduction of gene fusions. In these models, somatic

cells may be transduced *ex vivo* and then transplanted into syngeneic recipients (Zuber et al., 2009; Lange et al., 2003), thus allowing temporal control of disease onset and facilitating the development of more-complex genotypes by introducing multiple oncogenes or inactivating mutations at once. Although this approach does overcome several limitations associated with germline transgenic methods, supraphysiologic expression of the transgene remains the main concern, with the potential to transform unintended cellular compartments. For example, two independent research groups detected B-cell malignancies in mice upon transplantation of HSCs that express the *NPM-ALK* fusion protein, which – in humans – is exclusively observed in T-cell anaplastic large-cell lymphoma (Lange et al., 2003; Kuefer et al., 1997). Moreover, the transgenic somatic approach only had success in the hematopoietic system that allows for *ex vivo* transduction of stem cells.

A further limitation of the ectopic transgenic approach is that the endogenous genes remain unmodified and the transgenes, therefore, do not entirely recapitulate the oncogenic insult. For example, fusions driven by chromosomal deletions can involve the heterozygous loss of a large number of genes, whereas inversions and

Table 2. Advantages and limitations of genome-editing technologies

	Overview	Advantages	Limitations
<p>Expression of fusion gene</p> 	<ul style="list-style-type: none"> Fusion transgene of interest is introduced into cell lines or embryonic stem cells via lentiviral or transgenic expression cassettes 	<ul style="list-style-type: none"> High penetrance and short tumor latency Simple and efficient 	<ul style="list-style-type: none"> Does not model the stochastic evolution of cancer Risk of leaky expression and transformation of off-target tissues Endogenous loci and intervening regions are left intact
<p>Homologous recombination (knock-in)</p> 	<ul style="list-style-type: none"> Transgenes are introduced into the genome downstream of endogenous regulatory elements 	<ul style="list-style-type: none"> Less prone to supraphysiologic activation of signaling pathways, and off-target effects 	<ul style="list-style-type: none"> Does not completely recapitulate endogenous rearrangements, as the reciprocal locus of the gene fusion remains intact Inefficient and time-consuming; not an optimal tool for large-scale experiments
<p>Cre-loxP</p> 	<ul style="list-style-type: none"> Insertion of <i>loxP</i> sites at the desired chromosomal breakpoints by two consecutive rounds of homologous recombination. Transient expression of Cre-recombinase leads to recombination of the <i>loxP</i> sites and the desired chromosomal rearrangement 	<ul style="list-style-type: none"> More accurately mimics cancer-associated events Recombination can be targeted to specific tissues by expressing Cre under tissue-specific promoters Highly precise and efficient for small genomic segments 	<ul style="list-style-type: none"> Relatively inefficient for distantly separated loxP sites, and ineffective for engineering translocations between different chromosomes Possible lack of synteny between human genome and model system
<p>TALENs and ZFNs</p> 	<ul style="list-style-type: none"> Links the DNA-binding domain of a ZF protein with the nuclease domain of the FOK1 restriction enzyme, targeting DNA DSBs to specific loci 	<ul style="list-style-type: none"> Accurately models cancer-associated events Exquisite target specificity 	<ul style="list-style-type: none"> Technically challenging; the enzymes are markedly difficult to engineer Possible lack of synteny between human genome and model system
<p>CRISPR-Cas9</p> 	<ul style="list-style-type: none"> Expression of two sgRNAs enables the introduction of simultaneous DNA breaks mediated by the Cas9 endonuclease, which through a Lig4-dependent process, can induce an array of large chromosomal aberrations 	<ul style="list-style-type: none"> Accurately mirrors the expression levels of native tumor conditions Recapitulates the reciprocal loss of the allele associated with the rearrangement Ability to engineer somatic alterations <i>in vivo</i>, more closely modeling the stochastic growth of cancer Decreased need for mouse breeding Low cost, technically simple and efficient 	<ul style="list-style-type: none"> All possible mutation and rearrangement events of the target loci are induced (Fig. 1) Off-target activity of either sgRNA can drive unexpected genomic consequences Possible lack of synteny between human genome and model system

TALE, transcription activator-like effector; ZF protein, zinc-finger protein.

translocations can disrupt ‘partner’ genes or regulatory elements, such as enhancers, within intervening regions. Classic examples of this phenomenon are nucleophosmin 1 (*NPM1*) and *PTPRK*, putative tumor suppressor genes that are recurrently fused to *ALK1* and *RSPO3*, respectively (Grisendi et al., 2006; Chang et al., 2020; Shimozato et al., 2015). Disruption of *PTPRK* was shown to accelerate tumor growth through the phosphorylation of the cancer stem cell marker PROM1 and through activation of AKT signaling (Shimozato et al., 2015). Therefore, overexpression of *RSPO3* and loss of *PTPRK* could confer a double hit in colorectal cancer (CRC) harboring the *RSPO3-PTPRK* fusion. Similarly, loss of *NPM1* may destabilize tumor suppressor proteins, such as p53 and synergize with *ALK* overexpression to promote tumor growth in lymphomas with *NPM1-ALK* fusions (Grisendi et al., 2006).

One final, albeit rare issue with transgenic approaches is that they do not recapitulate the reciprocal product (e.g. ABL-BCR) of the translocation. Although expression of a reciprocal fusion product only occurs in some situations (Grisendi et al., 2006), there are cases in which both gene fusions are expressed and contribute to cancer growth. For example, both products of the reciprocal translocation *t(11;17)*, *PLZF-RAR α* and *RAR α -PLZF*, are necessary to induce acute promyelocytic leukemia in mice (He et al., 2000).

Knock-in and conditional approaches

In contrast to the expression of fusion cDNAs under the control of heterologous promoters, transgenes may be introduced into the genome downstream of endogenous regulatory elements, thus more closely reflecting the expression pattern of cancer-linked gene fusions. One of the first examples was the generation of the E μ -Myc mouse, created by transgenic insertion of the *MYC* protooncogene downstream of the immunoglobulin heavy chain enhancer (Adams et al., 1985), mimicking the translocation of *MYC* observed in almost all Burkitt lymphomas (Taub et al., 1982). In this model, the fusion cassette is not positioned at the endogenous Ig locus, *MYC* expression is still restricted to B cells and these mice form B-cell lymphomas. Corral and colleagues employed a similar strategy to characterize the role of *Mll-AF9* in acute leukemia (Corral et al., 1996). Despite the presence of the gene fusion in the germline, mice exclusively developed acute myeloid leukemia (AML), recapitulating the spectrum of human cancers with the translocation *t(9;11)*. Subsequent studies employed homologous recombination of transgenes to characterize the role of BCR-ABL in CML and acute lymphoblastic leukemia (Castellanos et al., 1997; Foley et al., 2013).

The main advantage of this knock-in approach is that gene expression is controlled through endogenous regulatory elements and, therefore, is less prone to supraphysiologic pathway activation and off-target tissue effects. However, this approach does not completely recapitulate rearrangements, as the reciprocal locus of the gene fusion remains intact. The technique is also time-consuming and not ideal for large-scale experiments or rapid interrogation of novel rearrangements.

The Cre-loxP system more accurately mimics cancer-associated events by creating rearrangements between two separate target loci (Van Deursen et al., 1995). In this approach, *loxP* sites are inserted at the desired chromosomal breakpoints by two consecutive rounds of homologous recombination. Transient expression of Cre-recombinase leads to recombination of the *loxP* sites and the desired chromosomal rearrangement. By expressing Cre-recombinase under the control of a tissue-specific promoter, recombination events can be targeted to specific tissues. Conditional expression of Cre-recombinase, e.g. by using tetracycline- or tamoxifen-regulated alleles, can be used to time events of recombination. Cre-loxP

recombination is remarkably precise and efficient for the deletion of small genomic segments, e.g. floxed KO alleles, but is relatively inefficient for distantly separated *loxP* sites and generally ineffective when modeling translocations between different chromosomes (Yu and Bradley, 2001). Despite this issue, the Cre-loxP system has been employed to model recurrent gene fusions observed in sarcomas and hematologic malignancies (Keller et al., 2004; Collins et al., 2000; Smith et al., 1995; Buchholz et al., 2000; Forster et al., 2003; Drynan et al., 2005). However, with the exception of the *Mll-Enl* fusion leukemia mouse model (Forster et al., 2003; Drynan et al., 2005), the chromosomal rearrangements were not sufficient to induce malignancies in the engineered mice. The lack of malignant transformation is most probably related to low recombination efficiencies in target tissues and decreased expression of oncogenes compared with that in transgenic mouse models (Yu and Bradley, 2001).

Recently, Lowe and colleagues used the Cre-loxP system to engineer mouse models of AML and lymphoma with 17p deletion. Somatic heterozygous deletion of the mouse chromosome 11B3, a region syntenic to human 17p13 and encompassing the *Trp53* locus, resulted in a more-aggressive phenotype compared with homozygous loss of *Trp53* only (Liu et al., 2016). The aggressive phenotypes were the result of simultaneously deleted tumor suppressor genes on mouse chromosome 11B3, which underscores the selective advantage segmental deletions or other chromosomal rearrangements may confer to cancer cells due to the disruption of multiple genes. These findings further highlight the importance of accurately modeling these understudied oncogenic events.

Chromosomal rearrangements through genome editing ZFNs and TALENs

Zinc-finger nucleases (ZFNs) link the DNA-binding domain of a customizable zinc finger protein with the nuclease domain of the FOK1 restriction enzyme (Urnov et al., 2010), targeting DNA DSBs to specific genomic loci. Like ZFNs, transcription activator-like effector (TALE) nucleases (TALENs) contain the nuclease domain of FOK1 but use TALE prokaryotic transcription factors as the DNA-binding domain. By simultaneously targeting two separate loci, TALENs and ZFNs have been employed to engineer the *Ewsr1-Fli1* and *Npm1-Alk* fusions implicated in Ewing sarcoma and anaplastic large-cell lymphoma, respectively (Piganeau et al., 2013). Because both ZFNs and TALENs require binding of two FOK1-linked proteins for each target locus, they have exquisite target specificity. However, these enzymes are also much more difficult to engineer and have quickly been overshadowed by the emergence of CRISPR-Cas9 tools.

CRISPR-Cas9

Clustered regularly interspaced short palindromic repeats (CRISPR) is a programmable, RNA-guided genome-editing system that has completely revolutionized the field of cancer genetics owing to its low cost, ease of use and high efficiency. First identified as a crucial component of bacterial immunity against phage infection, the system was subsequently engineered to target alternate DNA sequences in bacteria (Jinek et al., 2012) and mammalian cells (Cong et al., 2013; Jinek et al., 2013; Mali et al., 2013; Ran et al., 2013). The functional unit is made up of a dual RNA complex or single guide RNA (sgRNA) and a CRISPR-associated endonuclease, usually CRISPR-associated protein 9 (Cas9). These two components form a ribonucleoprotein complex that scans the genome for complementary DNA sequences adjacent to small consensus sequences called protospacer adjacent motifs (PAMs) (Sternberg et al., 2014). Given sufficient DNA-RNA homology, the endonuclease domains of Cas9 mediate a DNA DSB

(Cong et al., 2013; Jinek et al., 2013; Mali et al., 2013). Most importantly, targeting specificity is achieved by simply modifying a 17–20 bp sequence within the sgRNA, thus providing a means to engineer specific tools for multiple loci with relative ease.

Whereas Cas9-mediated DNA cleavage is often repaired by error-prone non-homologous end joining leading to small indels at the break site, expression of two sgRNAs enables the introduction of simultaneous DNA breaks that, through a Lig4-dependent process, can induce an array of large chromosomal aberrations (Li et al., 2015). In most cases, generating precise fusions of the two breakpoints is not crucial because the sgRNAs are usually engineered to target intronic regions. As a result, splicing of the exons will most often create the desired fusion transcript. Using this approach, multiple complex chromosomal rearrangements involved in hematological malignancies as well as lung, liver, brain and intestinal cancer have been engineered (Blasco et al., 2014; Maddalo et al., 2014; Li et al., 2015; Xue et al., 2014; Cook et al., 2017; Han et al., 2017).

CRISPR-Cas9 was first used to model recurrent oncogenic chromosomal rearrangements in cancer cell lines and primary cells, including the translocations *t*(11;22) and *t*(8;21) observed in Ewing sarcoma and AML, and the inversions *inv*(2)(p21p23) and *inv*(10)(p11q11) observed in NSCLC (Torres et al., 2014; Choi and Meyerson, 2014). Shortly after, Maddalo et al. and Blasco et al. used an *in vivo* somatic approach to engineer mouse models carrying *Eml4-Alk* (inversion) fusion-driven lung cancers via intratracheal instillation of recombinant adenoviruses (Maddalo et al., 2014; Blasco et al., 2014). Expression of Cas9 and sgRNAs in the endobronchial epithelium induced the endogenous inversion of chromosome 17, rearrangement of the *Elm4-Alk* loci and tumor growth with 100% penetrance (Maddalo et al., 2014; Blasco et al., 2014). The efficiency of the system in generating the *Elm4-Alk* fusion *in vivo* was estimated to be 1.5 rearrangements per 10⁶ cells (Blasco et al., 2014). As expected, *Alk*-rearranged tumors were sensitive to the small-molecule kinase inhibitor crizotinib (Maddalo et al., 2014). These pioneering studies demonstrated that CRISPR can be readily adapted to model cancer-associated chromosomal rearrangements, opening opportunities to better understand cancer initiation and progression, explore novel therapeutic strategies and investigate drivers of drug resistance *in vivo*. Subsequently, other groups adapted an optimized version of the same approach to engineer rearrangements in the liver and brain through tail-vein or intracranial injection of Cas9 and sgRNAs, respectively (Li et al., 2015; Xue et al., 2014; Cook et al., 2017).

For tissues that are not easily transduced *in vivo*, such as the intestine, inducible transgenic platforms allow temporal regulation of Cas9 expression to induce the desired rearrangements. For instance, Han and colleagues generated transgenic mice carrying a doxycycline (dox)-regulated Cas9 transgene and two sgRNAs targeting introns within *Eif3e* and *Rspo2* or *Ptprk* and *Rspo3* (Han et al., 2017). Treatment with dox induced the expected *Eif3e-Rspo2* deletion and *Rspo3-Ptprk* inversion. After 6 weeks, both models developed hyperproliferative and dysplastic lesions throughout the small intestine, although the phenotype was much less pronounced in *Eif3e-Rspo2* mice. It is worth noticing that this specific fusion is far less common in human CRC (Seshagiri et al., 2012; Sackstein et al., 2021) and is often associated with amplification of the 8q locus, suggesting a requirement for even further elevated expression of *Rspo2* to induce tumor growth. Recently, Kawasaki et al. described the development of both *PTPRK-RSPO3* and *EIF3E-RSPO2* fusions in human colon organoids (Kawasaki et al., 2020). Unlike the murine model where only *Rspo3* fusions enabled organoid growth in RSP0-free medium (Han et al., 2017) both fusions enabled niche

independence in the human organoids (Kawasaki et al., 2020). The precise functional difference between the two models is unclear but could reflect differences in the 5′ untranslated region of mouse and human *Rspo2* fusions that impact gene expression (Han et al., 2017). The polytypes of the *Ptprk-Rspo3 in vivo* model were widespread in the small intestine and harbored at least one copy of the inversion in most of the tumor cells, suggesting a cell-intrinsic advantage of carrying the fusion. This observation contradicts an independent study in which a Cre-dependent *Rspo3* cDNA transgene was induced into LGR5⁺ intestinal stem cells and the resulting epithelial hyperproliferation was reportedly driven by paracrine secretion of the *Rspo3* ligand (Hilkens et al., 2017). These conflicting findings underscore how subtle differences in genetic models can have different effects on disease phenotypes.

The simplicity of using CRISPR to create fusions has provided an incentive to characterize the oncogenic potential of newly identified and often very rare cancer-associated structural variants (Box 1). These new model systems can then be used to prospectively test sensitivity to existing or novel targeted therapies. In one example, Cook and colleagues sought to characterize multiple novel recurrent chromosomal rearrangements observed in high-grade gliomas (Cook et al., 2017). In one of the engineered models, intracranial injection of adenoviruses expressing Cas9 and paired sgRNAs led to the expected deletion on chromosome 3, and growth of high-grade gliomas carrying the *Bcan-Ntrk1* fusion. These tumors were sensitive to entrectinib, a pan-TRK inhibitor that has now been approved by the FDA for patients who have tumors harboring neurotrophic receptor tyrosine kinase (NTRK) gene fusions. By engineering an interstitial deletion in chromosome 3, the study provided proof of concept for using CRISPR somatic editing to model chromosomal rearrangements that do not result in gene fusions (Cook et al., 2017).

Using the same approach, Kasthuber et al. (2017) applied CRISPR editing to develop a mouse model of fibrolamellar hepatocellular carcinoma. This extremely rare type of liver cancer primarily affects adolescents and young adults, and is invariably associated with a segmental deletion on chromosome 19, generating an in-frame fusion of *DNAJB1* and *PRKACA* (Honeyman et al., 2014). The authors used different genome-editing approaches to develop a series of mouse models and showed that tumor growth depends on the kinase domain of the fusion protein, paving the way for clinical trials employing small-molecule inhibitors (Abou-Alfa et al., 2021).

A CRISPR-based approach to model a chromosomal rearrangement overcomes several of the limitations associated with the more traditional strategies of transgene expression or homologous recombination. By targeting the endogenous loci of the rearrangement, the model accurately mirrors the expression levels observed in the native tumor conditions. In addition, it recapitulates the loss of the reciprocal allele associated with the rearrangement, which may also promote tumor growth. The ability to engineer somatic alterations *in vivo* – impossible with traditional approaches due to their poor recombination efficiencies – offers additional advantages. By targeting only a subset of cells, somatic engineering more closely mirrors the natural evolution and stochastic growth of human cancer. For example, somatic CRISPR models of *Alk*-rearranged lung cancer displayed more-indolent growth compared with those of transgenic germline systems (Blasco et al., 2014; Soda et al., 2008). Furthermore, a single Cas9/sgRNA vector can be readily adapted to model rearrangements or other cooperating mutations in different genetic backgrounds, without the need for mouse breeding, thereby, significantly decreasing research costs.

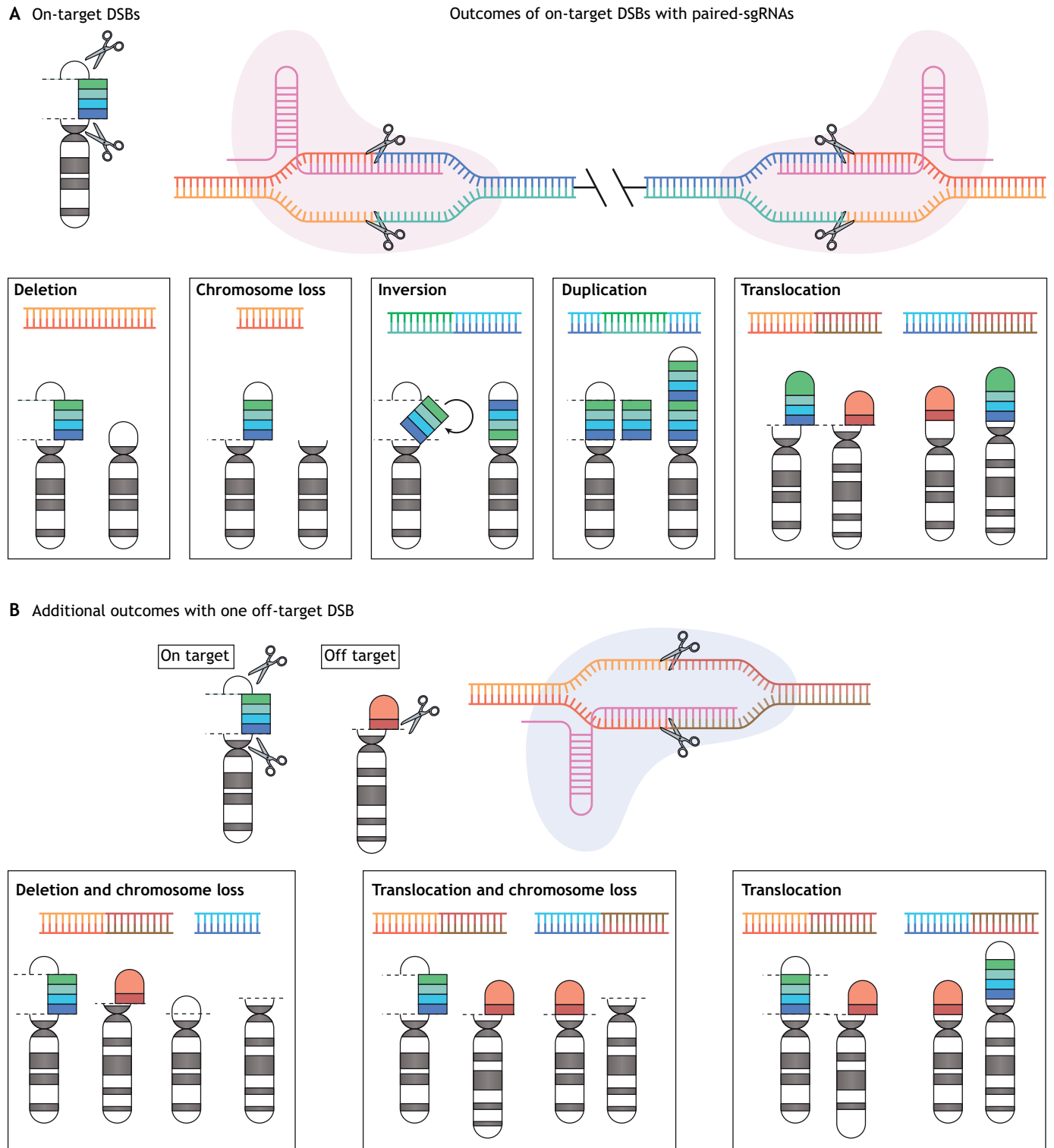


Fig. 1. Unintended rearrangements involving on-target and off-target loci upon CRISPR-based editing. (A) All possible rearrangement events of the target loci – deletions, inversions and duplications – are induced following CRISPR-mediated double-strand DNA breaks (DSBs) with paired sgRNAs. Positive selection within a bulk population of gene-edited cells often drives the enrichment of oncogenic rearrangements. (B) When DSBs are induced in one or more off-target loci, the number of possible unintended rearrangements grows exponentially, including novel fusion events, loss of entire chromosome segments, dicentric and acentric chromosomes (not shown).

Limitations of CRISPR-based genome editing

Widespread use of CRISPR over the past 8 years has led to significant improvements in its potency and specificity. However,

some challenges remain to be addressed (Table 2). One caveat of the dual sgRNA targeting system is that, within a population of cells, all possible mutation and rearrangement events of the target loci are

induced (Fig. 1). These include focal indels at each site, representing the majority of events, as well as inversions, deletions and duplications. Fortunately, in many cancer models, positive selection can drive the enrichment of an oncogenic lesion. For example, Cook and colleagues showed that expression of Cas9 and paired sgRNAs in neural stem cells led to both inversions and deletions on chromosome 3; yet, only clones harboring a deletion expressed the BCAN-NTRK1 fusion protein and drove tumor growth (Cook et al., 2017). However, for events with weak tumor-promoting potential, identifying relatively rare clones or interpreting data obtained from mixed populations can be challenging.

In addition to unintended rearrangements between the on-target loci, off-target activity of either sgRNA can drive unexpected and, possibly, deleterious consequences, including the generation of completely novel fusion events, dicentric chromosomes and loss of entire chromosome segments (Perez et al., 2017). If an sgRNA has multiple off-target loci, the number of possible unanticipated events grows exponentially, especially when dealing with aneuploid or hyper-diploid cancer cells (Fig. 1). Although the development of effective high-fidelity Cas9 nuclease variants can limit off-target activity (Kleinstiver et al., 2016; Zafra et al., 2018; Vakulskas et al., 2018), in many cases – even sgRNAs predicted to have high specificity – can target multiple genomic loci with near-identical sequences (Perez et al., 2017; Fu et al., 2013). The optimization of NGS-based tools may facilitate identification of off-target effects, and increase the specificity and reproducibility of CRISPR-based models (Wienert et al., 2020; Zuo et al., 2020). The final major limitation for the generation of chromosome rearrangement is not unique to CRISPR but a problem for all non-human models. Even with highly specific and active sgRNAs, some rearrangements seen in human disease are simply impossible to recreate in a model due to lack of synteny or lack of conservation of intron-exon boundaries across species.

Future directions

Chromosome rearrangements are a frequent and diverse group of cancer-associated genetic events. Most importantly, rearrangements frequently drive the production of gene fusions that act as oncogenic drivers. Although the overall incidence of cancers harboring individual rearrangements is low, the aggregate represents a high number of patients. Thus, developing fast, flexible and cost-effective methods to characterize the increasing number of recurrent chromosomal rearrangements that are being identified each year is a crucial step toward realizing the goals of precision medicine (Li et al., 2020). As the examples above demonstrate, CRISPR technologies have been instrumental in engineering and characterizing chromosomal rearrangements that directly contribute to cancer growth through expression of fusion proteins or disruption of tumor suppressor genes. Yet, CRISPR tools also offer flexibility and efficiency to functionally characterize putative oncogenic structural variants that involve non-coding regions (Rheinbay et al., 2020; Fujimoto et al., 2016; Quigley et al., 2018). Such rearrangement within non-coding regions may mediate the upregulation of nearby oncogenes and downregulation of tumor suppressor genes by, for example, altering regulatory elements or non-coding RNAs (Rheinbay et al., 2020; Quigley et al., 2018; Fujimoto et al., 2016). Larger datasets and advances in sequencing technologies (Box 1) will continue to reveal new coding and non-coding structural variants, and CRISPR will play a central role in modeling and characterizing the impact of these changes – both in cancer and other genetic disorders that are linked to chromosome rearrangements.

Acknowledgements

We thank Maria Paz Zafra and Alyna Katti for advice and comments on preparation of the manuscript. This work was supported by the T32 Clinical Scholars Grant from Memorial Sloan Kettering (T32-CA009512).

Competing interests

L.E.D. holds equity and is a scientific advisor for Mirimus Inc. and is a consultant for Volastra Therapeutics.

Funding

This work was supported by the T32 Clinical Scholars Grant to S.A. from Memorial Sloan Kettering (grant number: T32-CA009512).

References

- Abou-Alfa, G. K., Meyer, T., Zhang, J., Sherrin, S., Yaquib, A., O'neill, A. C., Xu, F., Eli, L. D., Harding, J. J., O'reilly, E. M. et al. (2021). Evaluation of neratinib (N), pembrolizumab (P), everolimus (E), and nivolumab (V) in patients (pts) with fibrolamellar carcinoma (FLC). *J. Clin. Oncol.* **39**, 310. doi:10.1200/JCO.2021.39.3_suppl.310
- Adams, J. M., Harris, A. W., Pinkert, C. A., Corcoran, L. M., Alexander, W. S., Cory, S., Palminter, R. D. and Brinster, R. L. (1985). The c-myc oncogene driven by immunoglobulin enhancers induces lymphoid malignancy in transgenic mice. *Nature* **318**, 533-538. doi:10.1038/318533a0
- Arai, Y., Totoki, Y., Takahashi, H., Nakamura, H., Hama, N., Kohno, T., Tsuta, K., Yoshida, A., Asamura, H., Mutoh, M. et al. (2013). Mouse model for ROS1-rearranged lung cancer. *PLoS One* **8**, e56010. doi:10.1371/journal.pone.0056010
- Benayed, R., Offin, M., Mullaney, K., Sukhadia, P., Rios, K., Desmeules, P., Ptashkin, R., Won, H., Chang, J., Halpenny, D. et al. (2019). High yield of RNA sequencing for targetable kinase fusions in lung adenocarcinomas with no mitogenic driver alteration detected by DNA sequencing and low tumor mutation burden. *Clin. Cancer Res.* **25**, 4712-4722. doi:10.1158/1078-0432.CCR-19-0225
- Blasco, R. B., Karaca, E., Ambrogio, C., Cheong, T. C., Karayol, E., Minero, V. G., Voena, C. and Chiarle, R. (2014). Simple and rapid in vivo generation of chromosomal rearrangements using CRISPR/Cas9 technology. *Cell Rep* **9**, 1219-1227. doi:10.1016/j.celrep.2014.10.051
- Buchholz, F., Refaeli, Y., Trumpp, A. and Bishop, J. M. (2000). Inducible chromosomal translocation of AML1 and ETO genes through Cre/loxP-mediated recombination in the mouse. *EMBO Rep.* **1**, 133-139. doi:10.1093/embo-reports/kvd027
- Castellanos, A., Pintado, B., Weruaga, E., Arevalo, R., López, A., Orfao, A. and Sanchez-Garcia, I. (1997). A BCR-ABL(p190) fusion gene made by homologous recombination causes B-cell acute lymphoblastic leukemias in chimeric mice with independence of the endogenous bcr product. *Blood* **90**, 2168-2174. doi:10.1182/blood.V90.6.2168
- Chang, L. S., Kim, M., Glinka, A., Reinhard, C. and Niehrs, C. (2020). The tumor suppressor PTPRK promotes ZNRF3 internalization and is required for Wnt inhibition in the Spemann organizer. *Elife* **9**. doi:10.7554/eLife.51248
- Chiarle, R., Gong, J. Z., Guasparri, I., Pesci, A., Cai, J., Liu, J., Simmons, W. J., Dhall, G., Howes, J., Piva, R. et al. (2003). NPM-ALK transgenic mice spontaneously develop T-cell lymphomas and plasma cell tumors. *Blood* **101**, 1919-1927. doi:10.1182/blood-2002-05-1343
- Choi, P. S. and Meyerson, M. (2014). Targeted genomic rearrangements using CRISPR/Cas technology. *Nat. Commun.* **5**, 3728. doi:10.1038/ncomms4728
- Collins, E. C., Pannell, R., Simpson, E. M., Forster, A. and Rabbitts, T. H. (2000). Inter-chromosomal recombination of Mll and Afl9 genes mediated by cre-loxP in mouse development. *EMBO Rep.* **1**, 127-132. doi:10.1093/embo-reports/kvd021
- Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P. D., Wu, X., Jiang, W., Marraffini, L. A. et al. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819-823. doi:10.1126/science.1231143
- Cook, P. J., Thomas, R., Kannan, R., De Leon, E. S., Drilon, A., Rosenblum, M. K., Scaltriti, M., Benezra, R. and Ventura, A. (2017). Somatic chromosomal engineering identifies BCAN-NTRK1 as a potent glioma driver and therapeutic target. *Nat. Commun.* **8**, 15987. doi:10.1038/ncomms15987
- Corral, J., Lavenir, I., Impey, H., Warren, A. J., Forster, A., Larson, T. A., Bell, S., McKenzie, A. N., King, G. and Rabbitts, T. H. (1996). An Mll-AF9 fusion gene made by homologous recombination causes acute leukemia in chimeric mice: a method to create fusion oncogenes. *Cell* **85**, 853-861. doi:10.1016/S0092-8674(00)81269-6
- Drilon, A., Laetsch, T. W., Kummar, S., Dubois, S. G., Lassen, U. N., Demetri, G. D., Nathanson, M., Doebele, R. C., Farago, A. F., Pappo, A. S. et al. (2018). Efficacy of Larotrectinib in TRK fusion-positive cancers in adults and children. *N. Engl. J. Med.* **378**, 731-739. doi:10.1056/NEJMoa1714448
- Drilon, A., Oxnard, G. R., Tan, D. S. W., Loong, H. H. F., Johnson, M., Gainor, J., Mccoach, C. E., Gautschi, O., Besse, B., Cho, B. C. et al. (2020). Efficacy of Selpercatinib in RET fusion-positive non-small-cell lung cancer. *N. Engl. J. Med.* **383**, 813-824. doi:10.1056/NEJMoa2005653
- Druker, B. J., Talpaz, M., Resta, D. J., Peng, B., Buchdunger, E., Ford, J. M., Lydon, N. B., Kantarjian, H., Capdeville, R., Ohno-Jones, S. et al. (2001). Efficacy and safety of a specific inhibitor of the BCR-ABL tyrosine kinase in

- chronic myeloid leukemia. *N. Engl. J. Med.* **344**, 1031-1037. doi:10.1056/NEJM200104053441401
- Drynan, L. F., Pannell, R., Forster, A., Chan, N. M., Cano, F., Daser, A. and Rabbitts, T. H. (2005). *MLL* fusions generated by Cre-loxP-mediated de novo translocations can induce lineage reassignment in tumorigenesis. *EMBO J.* **24**, 3136-3146. doi:10.1038/sj.emboj.7600760
- Foley, S. B., Hildenbrand, Z. L., Soyombo, A. A., Magee, J. A., Wu, Y., Oravec-Wilson, K. I. and Ross, T. S. (2013). Expression of BCR/ABL p210 from a knockin allele enhances bone marrow engraftment without inducing neoplasia. *Cell Rep* **5**, 51-60. doi:10.1016/j.celrep.2013.08.037
- Forster, A., Pannell, R., Drynan, L. F., McCormack, M., Collins, E. C., Daser, A. and Rabbitts, T. H. (2003). Engineering de novo reciprocal chromosomal translocations associated with *MLL* to replicate primary events of human cancer. *Cancer Cell* **3**, 449-458. doi:10.1016/S1535-6108(03)00106-5
- Fu, Y., Foden, J. A., Khayter, C., Maeder, M. L., Reynon, D., Joung, J. K. and Sander, J. D. (2013). High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat. Biotechnol.* **31**, 822-826. doi:10.1038/nbt.2623
- Fujimoto, A., Furuta, M., Totoki, Y., Tsunoda, T., Kato, M., Shiraiishi, Y., Tanaka, H., Taniguchi, H., Kawakami, Y., Ueno, M. et al. (2016). Whole-genome mutational landscape and characterization of noncoding and structural mutations in liver cancer. *Nat. Genet.* **48**, 500-509. doi:10.1038/ng.3547
- Grisendi, S., Mecucci, C., Falini, B. and Pandolfi, P. P. (2006). Nucleophosmin and cancer. *Nat. Rev. Cancer* **6**, 493-505. doi:10.1038/nrc1885
- Han, T., Schatoff, E. M., Murphy, C., Zafra, M. P., Wilkinson, J. E., Elemento, O. and Dow, L. E. (2017). R-Spondin chromosome rearrangements drive Wnt-dependent tumour initiation and maintenance in the intestine. *Nat. Commun.* **8**, 15945. doi:10.1038/ncomms15945
- He, L.-Z., Bhaumik, M., Tribioli, C., Rego, E. M., Ivins, S., Zelent, A. and Pandolfi, P. P. (2000). Two critical hits for promyelocytic leukemia. *Mol. Cell* **6**, 1131-1141. doi:10.1016/S1097-2765(00)00111-8
- Heisterkamp, N., Jenster, G., Ten Hoeve, J., Zovich, D., Pattengale, P. K. and Groffen, J. (1990). Acute leukaemia in *bcr/abl* transgenic mice. *Nature* **344**, 251-253. doi:10.1038/344251a0
- Heisterkamp, N., Jenster, G., Kioussis, D., Pattengale, P. K. and Groffen, J. (1991). Human *bcr-abl* gene has a lethal effect on embryogenesis. *Transgenic Res.* **1**, 45-53. doi:10.1007/BF02512996
- Henley, S. J., Ward, E. M., Scott, S., Ma, J., Anderson, R. N., Firth, A. U., Thomas, C. C., Islami, F., Weir, H. K., Lewis, D. R. et al. (2020). Annual report to the nation on the status of cancer, part I: National cancer statistics. *Cancer* **126**, 2225-2249. doi:10.1002/cncr.32802
- Heydt, C., Wölwer, C. B., Velazquez Camacho, O., Wagener-Rydzek, S., Pappesch, R., Siemanowski, J., Rehker, J., Haller, F., Agaimy, A., Worm, K. et al. (2021). Detection of gene fusions using targeted next-generation sequencing: a comparative evaluation. *BMC Med Genomics* **14**, 62. doi:10.1186/s12920-021-00909-y
- Hilkens, J., Timmer, N. C., Boer, M., Ikink, G. J., Schewe, M., Sacchetti, A., Koppens, M. A. J., Song, J. Y. and Bakker, E. R. M. (2017). RSPQ3 expands intestinal stem cell and niche compartments and drives tumorigenesis. *Gut* **66**, 1095-1105. doi:10.1136/gutjnl-2016-311606
- Honeyman, J. N., Simon, E. P., Robine, N., Chiaroni-Clarke, R., Darcy, D. G., Lim, I., Gleason, C. E., Murphy, J. M., Rosenberg, B. R., Teegan, L. et al. (2014). Detection of a recurrent DNAJB1-PRKACA chimeric transcript in fibrolamellar hepatocellular carcinoma. *Science* **343**, 1010-1014. doi:10.1126/science.1249484
- Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A. and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816-821. doi:10.1126/science.1225829
- Jinek, M., East, A., Cheng, A., Lin, S., Ma, E. and Doudna, J. (2013). RNA-programmed genome editing in human cells. *Elife* **2**, e00471. doi:10.7554/eLife.00471
- Ju, Y. S., Lee, W.-C., Shin, J.-Y., Lee, S., Bleazard, T., Won, J.-K., Kim, Y. T., Kim, J.-I., Kang, J.-H. and Seo, J.-S. (2012). A transforming KIF5B and RET gene fusion in lung adenocarcinoma revealed from whole-genome and transcriptome sequencing. *Genome Res.* **22**, 436-445. doi:10.1101/gr.133645.111
- Kastenhuber, E. R., Lalazar, G., Houlihan, S. L., Tschaharganeh, D. F., Baslan, T., Chen, C.-C., Requena, D., Tian, S., Bosbach, B., Wilkinson, J. E. et al. (2017). DNAJB1-PRKACA fusion kinase drives FL-HCC. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 13076-13084. doi:10.1073/pnas.1716483114
- Kawasaki, K., Fujii, M., Sugimoto, S., Ishikawa, K., Matano, M., Ohta, Y., Toshimitsu, K., Takahashi, S., Hosoe, N., Sekine, S. et al. (2020). Chromosome engineering of human colon-derived organoids to develop a model of traditional serrated adenoma. *Gastroenterology* **158**, 638-651e8. doi:10.1053/j.gastro.2019.10.009
- Keller, C., Hansen, M. S., Coffin, C. M. and Capecchi, M. R. (2004). Pax3:Fkhr interferes with embryonic Pax3 and Pax7 function: implications for alveolar rhabdomyosarcoma cell of origin. *Genes Dev.* **18**, 2608-2613. doi:10.1101/gad.1243904
- Kleinstiver, B. P., Pattanayak, V., Prew, M. S., Tsai, S. Q., Nguyen, N. T., Zheng, Z. and Joung, J. K. (2016). High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* **529**, 490-495. doi:10.1038/nature16526
- Kola, I. and Landis, J. (2004). Can the pharmaceutical industry reduce attrition rates? *Nat. Rev. Drug Discov.* **3**, 711-715. doi:10.1038/nrd1470
- Kuefer, M. U., Look, A. T., Pulford, K., Behm, F. G., Pattengale, P. K., Mason, D. Y. and Morris, S. W. (1997). Retrovirus-mediated gene transfer of NPM-ALK causes lymphoid malignancy in mice. *Blood* **90**, 2901-2910. doi:10.1182/blood.V90.8.2901
- Kwak, E. L., Bang, Y. J., Camidge, D. R., Shaw, A. T., Solomon, B., Maki, R. G., Ou, S. H., Dezube, B. J., Janne, P. A., Costa, D. B. et al. (2010). Anaplastic lymphoma kinase inhibition in non-small-cell lung cancer. *N. Engl. J. Med.* **363**, 1693-1703. doi:10.1056/NEJMoa1006448
- Lange, K., Uckert, W., Blankenstein, T., Nadrowitz, R., Bittner, C., Renaud, J. C., Van Snick, J., Feller, A. C. and Merz, H. (2003). Overexpression of NPM-ALK induces different types of malignant lymphomas in IL-9 transgenic mice. *Oncogene* **22**, 517-527. doi:10.1038/sj.onc.1206076
- Li, Y., Park, A. I., Mou, H., Colpan, C., Bizhanova, A., Akama-Garren, E., Joshi, N., Hendrickson, E. A., Feldser, D., Yin, H. et al. (2015). A versatile reporter system for CRISPR-mediated chromosomal rearrangements. *Genome Biol.* **16**, 111. doi:10.1186/s13059-015-0680-7
- 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. (2020). Patterns of somatic structural variation in human cancer genomes. *Nature* **578**, 112-121. doi:10.1038/s41586-019-1913-9
- Liu, Y., Chen, C., Xu, Z., Scuoppo, C., Rillahan, C. D., Gao, J., Spitzer, B., Bosbach, B., Kastenhuber, E. R., Baslan, T. et al. (2016). Deletions linked to TP53 loss drive cancer through p53-independent mechanisms. *Nature* **531**, 471-475. doi:10.1038/nature17157
- Maddalo, D., Manchado, E., Concepcion, C. P., Bonetti, C., Vidigal, J. A., Han, Y. C., Ogrodowski, P., Crippa, A., Rekhman, N., De Stanchina, E. et al. (2014). In vivo engineering of oncogenic chromosomal rearrangements with the CRISPR/Cas9 system. *Nature* **516**, 423-427. doi:10.1038/nature13902
- Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., Dicarlo, J. E., Norville, J. E. and Church, G. M. (2013). RNA-guided human genome engineering via Cas9. *Science* **339**, 823-826. doi:10.1126/science.1232033
- Nowell, P. C. (2007). Discovery of the Philadelphia chromosome: a personal perspective. *J. Clin. Invest.* **117**, 2033-2035. doi:10.1172/JCI31771
- Nowell, P. C. and Hungerford, D. A. (1960). Chromosome studies on normal and leukemic human leukocytes. *J. Natl. Cancer Inst.* **25**, 85-109.
- Perez, A. R., Pritykin, Y., Vidigal, J. A., Chhangawala, S., Zamparo, L., Leslie, C. S. and Ventura, A. (2017). GuideScan software for improved single and paired CRISPR guide RNA design. *Nat. Biotechnol.* **35**, 347-349. doi:10.1038/nbt.3804
- Piganeau, M., Ghezraoui, H., De Cian, A., Guittat, L., Tomishima, M., Perrouault, L., Rene, O., Katibah, G. E., Zhang, L., Holmes, M. C. et al. (2013). Cancer translocations in human cells induced by zinc finger and TALE nucleases. *Genome Res.* **23**, 1182-1193. doi:10.1101/gr.147314.112
- Quigley, D. A., Dang, H. X., Zhao, S. G., Lloyd, P., Aggarwal, R., Alumkal, J. J., Foye, A., Kothari, V., Perry, M. D., Bailey, A. M. et al. (2018). Genomic hallmarks and structural variation in metastatic prostate cancer. *Cell* **175**, 889. doi:10.1016/j.cell.2018.10.019
- Ran, F. A., Hsu, P. D., Lin, C. Y., Gootenberg, J. S., Konermann, S., Trevino, A. E., Scott, D. A., Inoue, A., Matoba, S., Zhang, Y. et al. (2013). Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* **154**, 1380-1389. doi:10.1016/j.cell.2013.08.021
- Reeser, J. W., Martin, D., Miya, J., Kautto, E. A., Lyon, E., Zhu, E., Wing, M. R., Smith, A., Reeder, M., Samorodnitsky, E. et al. (2017). Validation of a targeted RNA sequencing assay for kinase fusion detection in solid tumors. *J. Mol. Diagn.* **19**, 682-696. doi:10.1016/j.jmoldx.2017.05.006
- Rheinbay, E., Nielsen, M. M., Abascal, F., Wala, J. A., Shapira, O., Tiao, G., Hornshoj, H., Hess, J. M., Juul, R. I., Lin, Z. et al. (2020). Analyses of non-coding somatic drivers in 2,658 cancer whole genomes. *Nature* **578**, 102-111. doi:10.1038/s41586-020-1965-x
- Richardson, C. and Jasin, M. (2000). Frequent chromosomal translocations induced by DNA double-strand breaks. *Nature* **405**, 697-700. doi:10.1038/35015097
- Sackstein, P., Buechner, G. and Weinberg, B. A. (2021). Patient characteristics and survival outcomes in gastrointestinal cancers harboring RNF43 mutations or RSPQ3 fusions: a retrospective analysis. *J. Clin. Oncol.* **39**, 473. doi:10.1200/JCO.2021.39.3_suppl.473
- Sandberg, A. A. (1966). The chromosomes and causation of human cancer and leukemia. *Cancer Res.* **26**, 2064-2081.
- Sawyers, C. L. (1999). Chronic myeloid leukemia. *N. Engl. J. Med.* **340**, 1330-1340. doi:10.1056/NEJM199904293401706
- Seshagiri, S., Stawicki, E. W., Durinck, S., Modrusan, Z., Storm, E. E., Conboy, C. B., Chaudhuri, S., Guan, Y., Janakiraman, V., Jaiswal, B. S. et al. (2012). Recurrent R-spondin fusions in colon cancer. *Nature* **488**, 660-664. doi:10.1038/nature11282

- Sharpless, N. E. and Depinho, R. A.** (2006). The mighty mouse: genetically engineered mouse models in cancer drug development. *Nat. Rev. Drug Discov.* **5**, 741-754. doi:10.1038/nrd2110
- Shaw, A. T., Ou, S.-H., Bang, Y.-J., Camidge, D. R., Solomon, B. J., Salgia, R., Riely, G. J., Varella-Garcia, M., Shapiro, G. I., Costa, D. B. et al.** (2014). Crizotinib in ROS1-rearranged non-small-cell lung cancer. *N. Engl. J. Med.* **371**, 1963-1971. doi:10.1056/NEJMoa1406766
- Shaw, A. T., Kim, T. M., Crino, L., Gridelli, C., Kiura, K., Liu, G., Novello, S., Bearz, A., Gautschi, O., Mok, T. et al.** (2017). Ceritinib versus chemotherapy in patients with ALK-rearranged non-small-cell lung cancer previously given chemotherapy and crizotinib (ASCEND-5): a randomised, controlled, open-label, phase 3 trial. *Lancet Oncol.* **18**, 874-886. doi:10.1016/S1470-2045(17)30339-X
- Shimozato, O., Waraya, M., Nakashima, K., Souda, H., Takiguchi, N., Yamamoto, H., Takenobu, H., Uehara, H., Ikeda, E., Matsushita, S. et al.** (2015). Receptor-type protein tyrosine phosphatase κ directly dephosphorylates CD133 and regulates downstream AKT activation. *Oncogene* **34**, 1949-1960. doi:10.1038/ncr.2014.141
- Shtivelman, E., Lifshitz, B., Gale, R. P. and Canaani, E.** (1985). Fused transcript of *abl* and *bcr* genes in chronic myelogenous leukaemia. *Nature* **315**, 550-554. doi:10.1038/315550a0
- Smith, A. J., De Sousa, M. A., Kwabi-Addo, B., Heppell-Parton, A., Impey, H. and Rabbitts, P.** (1995). A site-directed chromosomal translocation induced in embryonic stem cells by Cre-loxP recombination. *Nat. Genet.* **9**, 376-385. doi:10.1038/ng0495-376
- Soda, M., Takada, S., Takeuchi, K., Choi, Y. L., Enomoto, M., Ueno, T., Haruta, H., Hamada, T., Yamashita, Y., Ishikawa, Y. et al.** (2008). A mouse model for EML4-ALK-positive lung cancer. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 19893-19897. doi:10.1073/pnas.0805381105
- Solomon, B. J., Mok, T., Kim, D.-W., Wu, Y.-L., Nakagawa, K., Mekhail, T., Felip, E., Cappuzzo, F., Paolini, J., Usari, T. et al.** (2014). First-line crizotinib versus chemotherapy in ALK-positive lung cancer. *N. Engl. J. Med.* **371**, 2167-2177. doi:10.1056/NEJMoa1408440
- Sternberg, S. H., Redding, S., Jinek, M., Greene, E. C. and Doudna, J. A.** (2014). DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature* **507**, 62-67. doi:10.1038/nature13011
- Taub, R., Kirsch, I., Morton, C., Lenoir, G., Swan, D., Tronick, S., Aaronson, S. and Leder, P.** (1982). Translocation of the *c-myc* gene into the immunoglobulin heavy chain locus in human Burkitt lymphoma and murine plasmacytoma cells. *Proc. Natl. Acad. Sci. U.S.A.* **79**, 7837-7841. doi:10.1073/pnas.79.24.7837
- Torres, R., Martin, M. C., Garcia, A., Cigudosa, J. C., Ramirez, J. C. and Rodriguez-Perales, S.** (2014). Engineering human tumour-associated chromosomal translocations with the RNA-guided CRISPR-Cas9 system. *Nat. Commun.* **5**, 3964. doi:10.1038/ncomms4964
- Urnov, F. D., Rebar, E. J., Holmes, M. C., Zhang, H. S. and Gregory, P. D.** (2010). Genome editing with engineered zinc finger nucleases. *Nat. Rev. Genet.* **11**, 636-646. doi:10.1038/nrg2842
- Vakulskas, C. A., Dever, D. P., Rettig, G. R., Turk, R., Jacobi, A. M., Collingwood, M. A., Bode, N. M., McNeill, M. S., Yan, S., Camarena, J. et al.** (2018). A high-fidelity Cas9 mutant delivered as a ribonucleoprotein complex enables efficient gene editing in human hematopoietic stem and progenitor cells. *Nat. Med.* **24**, 1216-1224. doi:10.1038/s41591-018-0137-0
- Van Deursen, J., Fornerod, M., Van Rees, B. and Grosveld, G.** (1995). Cre-mediated site-specific translocation between nonhomologous mouse chromosomes. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 7376-7380. doi:10.1073/pnas.92.16.7376
- Wienert, B., Wyman, S. K., Yeh, C. D., Conklin, B. R. and Corn, J. E.** (2020). CRISPR off-target detection with DISCOVER-seq. *Nat. Protoc.* **15**, 1775-1799. doi:10.1038/s41596-020-0309-5
- Xue, W., Chen, S., Yin, H., Tammela, T., Papagiannakopoulos, T., Joshi, N. S., Cai, W., Yang, G., Bronson, R., Crowley, D. G. et al.** (2014). CRISPR-mediated direct mutation of cancer genes in the mouse liver. *Nature* **514**, 380-384. doi:10.1038/nature13589
- Yu, Y. and Bradley, A.** (2001). Engineering chromosomal rearrangements in mice. *Nat. Rev. Genet.* **2**, 780-790. doi:10.1038/35093564
- Zafra, M. P., Schatoff, E. M., Katti, A., Foronda, M., Breinig, M., Schweitzer, A. Y., Simon, A., Han, T., Goswami, S., Montgomery, E. et al.** (2018). Optimized base editors enable efficient editing in cells, organoids and mice. *Nat. Biotechnol.* **36**, 888-893. doi:10.1038/nbt.4194
- Zuber, J., Radtke, I., Pardee, T. S., Zhao, Z., Rappaport, A. R., Luo, W., Mccurrach, M. E., Yang, M. M., Dolan, M. E., Kogan, S. C. et al.** (2009). Mouse models of human AML accurately predict chemotherapy response. *Genes Dev.* **23**, 877-889. doi:10.1101/gad.1771409
- Zuo, E., Sun, Y., Wei, W., Yuan, T., Ying, W., Sun, H., Yuan, L., Steinmetz, L. M., Li, Y. and Yang, H.** (2020). Goti, a method to identify genome-wide off-target effects of genome editing in mouse embryos. *Nat. Protoc.* **15**, 3009-3029. doi:10.1038/s41596-020-0361-1