

Review

Toward precise CRISPR DNA fragment editing and predictable 3D genome engineering

Qiang Wu* and Jia Shou

Center for Comparative Biomedicine, MOE Key Lab of Systems Biomedicine, State Key Laboratory of Oncogenes and Related Genes, Institute of Systems Biomedicine, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai 200240, China

* Correspondence to: Qiang Wu, E-mail: qwu123@gmail.com

Edited by Jinsong Li

Ever since gene targeting or specific modification of genome sequences in mice was achieved in the early 1980s, the reverse genetic approach of precise editing of any genomic locus has greatly accelerated biomedical research and biotechnology development. In particular, the recent development of the CRISPR/Cas9 system has greatly expedited genetic dissection of 3D genomes. CRISPR gene-editing outcomes result from targeted genome cleavage by ectopic bacterial Cas9 nuclease followed by presumed random ligations via the host double-strand break repair machineries. Recent studies revealed, however, that the CRISPR genome-editing system is precise and predictable because of cohesive Cas9 cleavage of targeting DNA. Here, we synthesize the current understanding of CRISPR DNA fragment-editing mechanisms and recent progress in predictable outcomes from precise genetic engineering of 3D genomes. Specifically, we first briefly describe historical genetic studies leading to CRISPR and 3D genome engineering. We then summarize different types of chromosomal rearrangements by DNA fragment editing. Finally, we review significant progress from precise 1D gene editing toward predictable 3D genome engineering and synthetic biology. The exciting and rapid advances in this emerging field provide new opportunities and challenges to understand or digest 3D genomes.

Keywords: CRISPR, DNA fragment editing, 3D genome engineering, repair mechanisms, chromatin loops, precise modifications, predictable indels

Introduction

The successful finishing of the Human Genome Project ushers in a new era to understand and engineer genomes by reverse genetics. However, the folding of 3-billion-bp 1D mammalian genomes, which are ~2 m long, into 3D structures within cell nuclei of ~5 μm in diameter adds another layer of complexity. The secret of 3D genome coding likely resides in the non-coding regions—the 97.5% of mammalian genomes—that were once assumed to be ‘junk DNA’ but are now regarded as ‘crown jewels’. Specifically, high-throughput mapping of functional genomic sequences has revealed numerous non-coding DNA elements, up to 8.4 million in number (Neph et al., 2012). In addition, junk DNA transcribes so-called ‘junk RNA’—numerous long non-coding RNA—whose functions are difficult to study (Cech and Steitz, 2014). The organizational and structural roles

of these non-coding DNA elements in 3D genome regulation and function necessitate functional genetic experiments.

Trekking across time: the long journey of reverse genetics leading to CRISPR and 3D genome editing

Genetic research focuses on heredity and ‘mutants’ (Castle and Little, 1909; Muller, 1930). Some mutants arise spontaneously but specific mutants are usually generated through tedious forward genetic screening experiments (Acevedo-Arozena et al., 2008). Forward genetic screening in mice was performed before the mouse genome sequencing was finished and greatly contributed to our understanding of human physiology (Kile and Hilton, 2005). However, reverse genetics that would generate specific alterations of mammalian genomic sequences or so-called gene targeting was a dream in the early days.

Transgenic: random integration in animal and plant genomes

Transgenes were originally derived from viruses and transposons or so-called jumping genes in animals and plants (McClintock, 1950; Jaenisch and Mintz, 1974; Bevan et al., 1983). A transgene can be integrated randomly into one or very few

Received July 10, 2020. Revised September 23, 2020. Accepted September 23, 2020.

© The Author(s) (2020). Published by Oxford University Press on behalf of *Journal of Molecular Cell Biology*, IBCB, SIBS, CAS.

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

sites of the mouse genome and exhibits expression patterns with position-effect variegations (Figure 1A; Jaenisch and Mintz, 1974; Gordon et al., 1980; Brinster et al., 1981; Costantini and Lacy, 1981). Multiple copies of transgenes are typically integrated at a random genomic site in tandem arrays as a head-to-tail concatemer (Figure 1A; Brinster et al., 1981; Folger et al., 1982). Homologous recombination (HR) was demonstrated convincingly to be the predominant mechanism of head-to-tail transgene integration (Folger et al., 1982). In fact, it is with this conviction that eventually led to the development of gene targeting in mice (Capecchi, 2005).

Gene targeting or knockout mice

Gene targeting is different from transgenic technologies and has greatly accelerated biological researches. Even before the completion of the mouse genome sequencing, the dream of specific modification of any mouse locus had been realized by so-called gene targeting (Figure 1A; Smithies et al., 1985; Thomas et al., 1986). The technique is achieved by constructing a targeting vector with designed modification in a specific locus, which is flanked by two homologous arms. This donor template is then introduced into mouse embryonic stem cells (ESCs) (Evans and Kaufman, 1981; Martin, 1981) and replaces the endogenous sequences through HR (Figure 1A). Finally, the ESC clones carrying the designed specific modification are then injected into the mouse blastocoel cavity to generate chimeric mice. Heterozygous or homozygous mice could then be obtained simply by breeding. The remarkable technique and general protocol for generating knockout mice with any gene targeted were quickly developed (Mansour et al., 1988).

Gene editing with zinc-finger nucleases, transcription activator-like effector nucleases, and CRISPR

Targeted gene replacement through HR has also been achieved for other model organisms such as yeast and flies (Scherer and Davis, 1979; Rong and Golic, 2000). Since free double-strand break (DSB) ends greatly stimulate HR (Figure 1B; Orr-Weaver et al., 1981; Jasin and Berg, 1988), intense efforts were devoted to creating targeted DSBs. A series of programmable endonucleases, including zinc-finger nucleases (ZFNs) (Bibikova et al., 2003), transcription activator-like effector nucleases (TALENs) (Miller et al., 2011), and clustered regularly interspaced short palindromic repeat/CRISPR-associated nuclease 9 (CRISPR/Cas9) (Gasiunas et al., 2012; Jinek et al., 2012; Cong et al., 2013; Mali et al., 2013), were found to be able to introduce not only targeted modifications across genomes but also targeted head-to-tail insertions (Figure 1B–D; Folger et al., 1982; Skryabin et al., 2020). CRISPR, in particular, has revolutionized targeted genome modification because of its simplicity and practicality.

CRISPR: clustered regularly interspaced short palindromic repeats

CRISPR/Cas9 is an RNA-guided adaptive immune system of bacteria and archaea, which defends against phage or virus infection and plasmid conjugation. The type II CRISPR/Cas9 system has been widely used for genome editing. The

programmable CRISPR/Cas9 system consists of a synthetic single-guide RNA (sgRNA; derived from crRNA and tracrRNA) and RNA-guided Cas9 nuclease (Figure 1D; Jinek et al., 2012). Upon recognition of a protospacer adjacent motif (PAM, NGG for *SpCas9* from *Streptococcus pyogenes*) downstream of the targeting sequence, Cas9 cleaves the complementary and non-complementary strands of the target DNA duplex by the HNH and RuvC nuclease domains, respectively (Garneau et al., 2010; Gasiunas et al., 2012; Jinek et al., 2012), resulting in presumed blunt-ended DSBs which are then ligated by cellular endogenous DNA repair machineries (Figure 1D).

Gene-editing outcomes from single DSBs

There are numerous gene-editing applications of single DSBs from CRISPR. The simplest application is the generation of frameshift mutations in the coding region of a protein-encoding gene. Cas9 can be reprogrammed by single sgRNAs to target a coding exon, generating one DSB that often leads to nucleotide insertions and/or deletions (indels). Two-thirds of these indels can cause a shift in the open reading frame of a protein-coding gene, resulting in truncated protein translation or null mutation through the nonsense-mediated mRNA decay. Recent studies demonstrated, however, that single DSBs also lead to large deletions from extended long resections (Li et al., 2015a; Shin et al., 2017; Kosicki et al., 2018, 2020; Jia et al., 2020). In addition, Cas9 with single sgRNAs causes frequent loss-of-heterozygosity or gene conversion as well as allele-specific chromosomal removal in human embryos (Alanis-Lobato et al., 2020; Liang et al., 2020; Zuccaro et al., 2020). Finally, if a donor DNA template is provided, single DSBs often lead to targeted precise gene insertions through HR (Figure 1B).

3D genome primer

Although genetic information is encoded in the finished linear 1D genomic sequences, the extremely long and thin DNA molecules do actually exist in Euclidean 3D space and are physically folded into a cell nucleus. Each interphase chromosome occupies a distinct territory and compartmentalizes further into multiple topologically associated domains (TADs). The recognition sites of architectural protein CCCTC-binding factor (CTCF) are enriched at boundaries of chromatin domains; however, there are also numerous CTCF sites located within topological domains or TADs. Exactly how 3D genomes are folded and regulated remains unknown; however, novel technological developments have enabled tremendous progress in 3D genomics (Banigan and Mirny, 2020; Li et al., 2020a; Zhang and Li, 2020). In particular, DNA fragment editing or CRISPR-induced chromosomal rearrangements have shed significant insights into the mechanisms of 3D genome folding (Liu and Wu, 2020).

There are numerous excellent reviews on CRISPR or 3D genomics (Doudna and Charpentier, 2014; Huang and Wu, 2016; Jiao and Gao, 2016; Yan and Li, 2019; Yang and Huang, 2019; Zhang, 2019; Anzalone et al., 2020; Li et al., 2020a; Yang and Chen, 2020; Zhang and Li, 2020; Zhang et al., 2020). Here, we

focus on chromosomal rearrangements and 3D genome engineering by DNA fragment editing using Cas9 with dual sgRNAs.

Chromosomal rearrangements by CRISPR with dual sgRNAs

Structural chromosomal abnormalities or chromosomal rearrangements include DNA fragment deletions, inversions, duplications, translocations, and insertions (Figure 2; Shaffer and Lupski, 2000; Huang and Wu, 2016). Chromosomal rearrangements are estimated to occur at 0.6% of human newborns (Jacobs et al., 1992). In addition, recurrent chromosomal rearrangements are quite frequent in human neurological diseases (Weckselblatt and Rudd, 2015) and tumors (Rabbitts, 1994; Mitelman et al., 1997). Early studies to model human diseases generated large chromosomal rearrangements of up to tens of millions bp in mice through the combined technologies of gene targeting and Cre/*LoxP* recombination (Ramirez-Solis et al., 1995; Herault et al., 1998; Wu et al., 2007; and reviewed in Mills and Bradley, 2001; Yu and Bradley, 2001). ZFNs and TALENs have also been used to generate chromosomal rearrangements in human cells (Lee et al., 2010; Gupta et al., 2013; Nyquist et al., 2013; Xiao et al., 2013). In this section, we outline 3D genome engineering by modeling chromosomal rearrangements using the CRISPR/Cas9 system with dual sgRNAs (Figure 2; Li et al., 2015b).

Chromosomal rearrangements by DNA fragment editing

Disruption of a specific gene of interest could be easily achieved by Cas9 reprogrammed with single sgRNAs because two-thirds of random indels at a DSB site within a protein-coding region result in frameshifts. For non-coding elements, however, random indels induced by Cas9 with single sgRNAs are usually not enough. A practical way to characterize non-coding regions, of which there are estimated millions in mammalian genomes, is to generate very large deletions containing defined regions with multiple non-coding elements (Wu et al., 2007). Engineering a large DNA fragment could be achieved by Cas9 reprogrammed with dual sgRNAs, which would generate two concurrent DSBs in a genome (Figure 2). Specifically, with the participation of cellular DNA repair proteins, the four DSB ends generated by the two Cas9 cleavages are randomly ligated, resulting in DNA fragment deletion or inversion when concurrent DSBs occur on single chromosomes (Figure 2A and B) and DNA fragment duplication or translocation when the DSBs occur on different chromatids or chromosomes (Figure 2C and D).

DNA fragment deletion by CRISPR

It is well established that Cas9 with dual sgRNAs can easily generate DNA fragment deletions (Figure 2A; Huang and Wu, 2016). However, initial utilization of the CRISPR system with dual sgRNAs has been to mitigate off-target activity. The D10A Cas9 nickase guided by paired sgRNAs in proper configurations and optimized offsets generates double nicking and 5' overhangs (Ran et al., 2013; Shen et al., 2014). Subsequent

targeting of two separate intrachromosomal sites by wildtype Cas9 with dual sgRNAs results in the interstitial deletion of large DNA fragments in zebrafish (Gupta et al., 2013; Xiao et al., 2013), mammalian cells (Cong et al., 2013; Mali et al., 2013; Canver et al., 2014; Guo et al., 2015, 2018; He et al., 2015; Li et al., 2015c; Kim et al., 2017; Schmieder et al., 2018; Shou et al., 2018; Shi et al., 2019; Jia et al., 2020), mice (Zhou et al., 2014; Li et al., 2015a; Jia et al., 2020), rabbits (Song et al., 2016), worms (Chen et al., 2014), and plants (Pauwels et al., 2018; Schmidt et al., 2019) (Table 1).

DNA fragment inversion by CRISPR

In addition to DNA fragment deletions, DNA fragment inversion events also occur through double cutting, which is different from double nicking, within single chromosomes (Figure 2B). Different from DNA fragment deletion, in which there is only one junction after deleting the intervening sequences, DNA fragment inversion has an upstream junction and a downstream junction after inverting the intervening DNA fragment (Huang and Wu, 2016).

DNA fragment inversions using Cas9 guided with dual sgRNAs can be easily achieved in cultured cells (Canver et al., 2014; Choi and Meyerson, 2014; Guo et al., 2015; Kraft et al., 2015; Li et al., 2015a; Park et al., 2015), mice (Blasco et al., 2014; Maddalo et al., 2014; Kraft et al., 2015; Li et al., 2015a; Seruggia et al., 2015; Boroviak et al., 2016; Birling et al., 2017; Lu et al., 2019; Jia et al., 2020), rats (Birling et al., 2017), and plants (Schmidt et al., 2019). In particular, DNA fragment inversion results in the generation of an oncogenic gene from fusion of two genes at an inversion junction in mouse somatic tissues that faithfully models human tumors (Blasco et al., 2014; Maddalo et al., 2014). Finally, Cas9 guided by dual sgRNAs has been used to study the role of the orientation of non-coding regulatory elements such as enhancers and insulators (Guo et al., 2015; Li et al., 2015a).

DNA fragment duplication by CRISPR

Chromosomal duplications can be generated by *trans*-allelic ligations of DSB ends in two homologous chromosomes or chromatids (Figure 2C; Golic and Golic, 1996; Wu et al., 2007; Li et al., 2015a). Specifically, DNA fragment duplications can be generated by complementary *trans*-chromatid ligations of paracentric DSB ends resulting from cleavages by Cas9 guided with dual sgRNAs after DNA replication during both mitosis and meiosis. Thus, Cas9 guided with dual sgRNAs induces DNA fragment duplications in cultured cells (Kraft et al., 2015; Li et al., 2015a). In addition, DNA fragment duplications in mice *in vivo* can be induced by Cas9 with dual sgRNAs through pronuclear microinjection (Li et al., 2015a; Korablev et al., 2017). In particular, a tandem duplication of a 1211-bp DNA fragment was confirmed by Sanger sequencing of the entire duplicated segment (Li et al., 2015a). Finally, quantitative analyses revealed frequent segmental duplications by Cas9 with dual sgRNAs, though with lower efficiency compared with that of DNA fragment deletions and inversions (Li et al., 2015a).

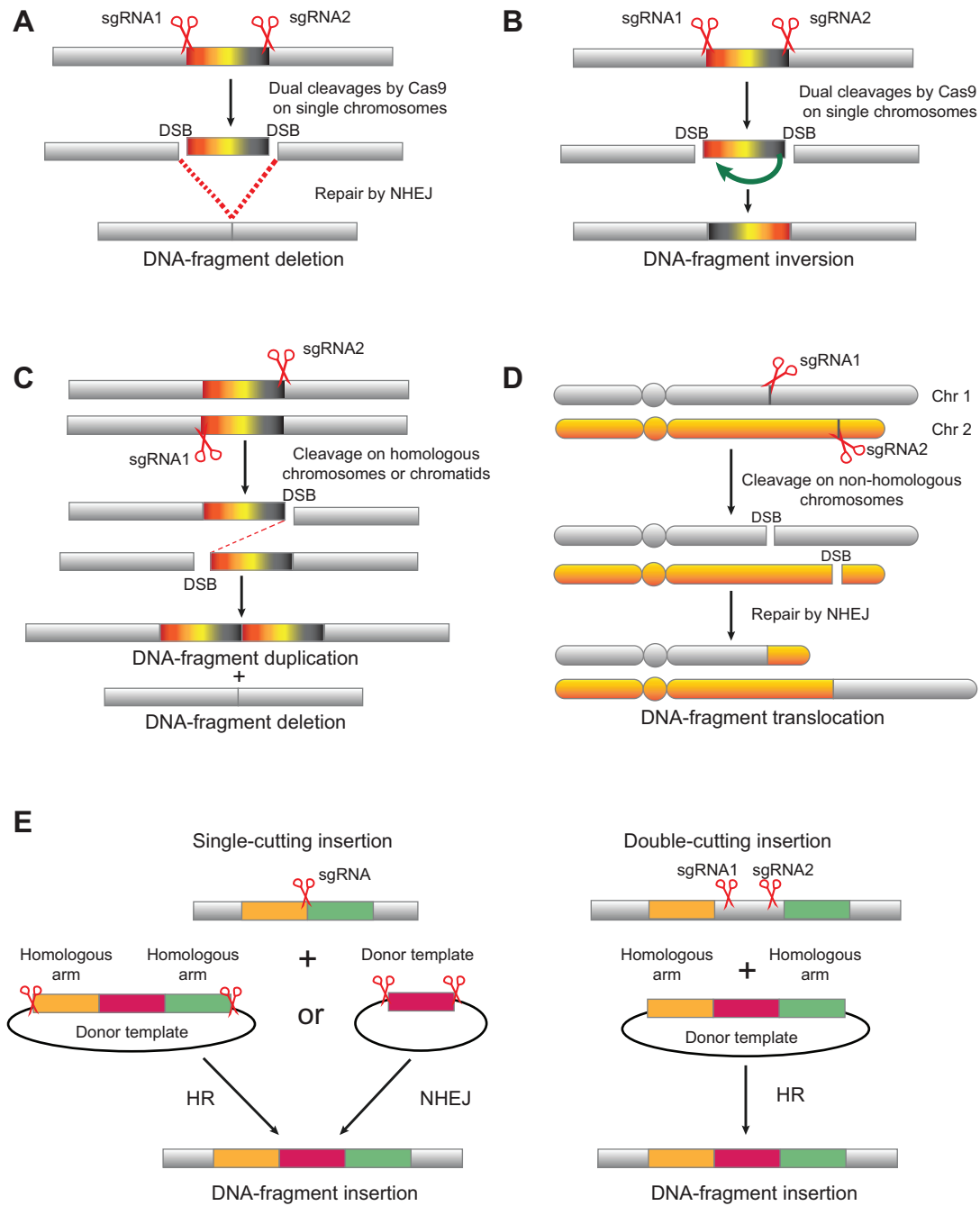


Figure 2 DNA fragment editing induces chromosomal rearrangements including large DNA fragment deletion (A), inversion (B), duplication (C), translocation (D), as well as insertion (E).

Chromosomal translocation by CRISPR

Chromosomal translocations result from joining DSB ends in two distinct chromosomes (Figure 2D). Recurrent chromosomal translocations are frequent in many types of tumors especially in leukemias (Lieber, 2016; Vanoli and Jasin, 2017; Brunet and Jasin, 2018; Cheong et al., 2018). Cas9

reprogrammed with dual sgRNAs that target specific loci in non-homologous chromosomes has been used to induce chromosomal translocations to model human Ewing's sarcoma, desmoplastic small round cell tumors, and acute myeloid leukemia (AML) (Torres et al., 2014; Vanoli et al., 2017).

Table 1 Chromosomal rearrangements by CRISPR with dual sgRNAs.

Chromosomal rearrangement	Cell type or organism	Gene or region of interest	Targeting size (kb)	Targeting efficiency (%)	Efficiency measuring method	References
DNA fragment deletion	Mice	Hypoxanthine phosphoribosyltransferase locus (<i>HPRT</i>)	10	9/27 (33.3%)	Mutant mice by zygote injection	Fujii et al. (2013)
	Murine erythroleukemia (MEL) cells	ND	1.3	18/48 (37.5%)	Screening single cell clones	Canver et al. (2014)
			2.0	60/234 (25.6%)		
			2.8	29/78 (37.2%)		
			4.5	14/122 (11.5%)		
			4.5	10/164 (6.1%)		
			7.3	59/332 (17.8%)		
			8.0	190/800 (23.8%)		
			13.5	20/160 (12.5%)		
			15.0	74/316 (23.4%)		
			19.0	2/68 (2.9%)		
			19.0	21/240 (8.8%)		
			20.3	34/140 (24.3%)		
			23.0	20/142 (14.1%)		
			23.0	5/54 (9.3%)		
			70.5	1/364 (0.3%)		
	1025.3	1/266 (0.4%)				
	1025.7	3/420 (0.7%)				
	HAP1 cells	Chr 15: 61,105,000 to ~89,890,000	~28000	5/400 (1.3%)	Screening single cell clones	Essletzbichler et al. (2014)
	Mouse ESC	<i>Dip2a</i>	65	11/93 (11.8%)	Screening single cell clones	Zhang et al. (2015)
	Mice	<i>Dip2a</i>	65	3/14 (21.4%)	Mutant mice by zygote injection	Zhang et al. (2015)
	Mice	<i>Rab38</i>	3.2	10/27 (37%)	Mutant mice by zygote injection	Brandl et al. (2015)
	HEK293FT cells	<i>HPRT</i>	1.79	3.3%	Digital PCR analysis	He et al. (2015)
			2.14	3.3%		
			13.33	10%		
			0.35	10%		
			11.54	10%		
11.19			1%			
63.07			10%			
112.93			10%			
513.60			10%			
1017.84			1%			
HEK293FT cells	Hypoxanthine phosphoribosyltransferase locus (<i>HPRT</i>)	513.60	8/63 (12.7%)	Screening single cell clones	He et al. (2015)	
Mouse ESC	<i>H2afy</i>	1.189	11/288 (3.8%)	Screening single cell clones	Kraft et al. (2015)	
		<i>Bmp2</i>	3.7			12/192 (6.3%)
		<i>lhh</i>	12.6			121/288 (42%)
		<i>Pitx1</i>	32			9/288 (3.1%)
		<i>Laf4</i>	353			38/288 (13.2%)
HEK293T	<i>Epha4</i>	1672	4/192 (2.1%)	Quantitative PCR	Li et al. (2015a)	
		<i>β-globin RE1</i>	0.709			(28.33 ± 6.19)%
		<i>Pcdh RE1</i>	1.272			(17.51 ± 1.04)%
		<i>β-globin RE2</i>	6.277			(34.49 ± 3.57)%
		<i>HoxD</i>	18.142			(9.15 ± 0.11)%
		<i>β-globin</i>	80.732			(13.39 ± 0.80)%
		<i>Pcdhα cluster</i>	256.744			(8.46 ± 0.24)%
		<i>Pcdh α, β, and γ clusters</i>	807.480			(0.47 ± 0.08)%
Mice	<i>Pcdh</i> locus 1	1.241	26/120 (21.7%)	Mutant mice by zygote injection	Li et al. (2015a)	
		<i>Pcdh</i> locus 2	0.96			6/8 (75%)
		<i>Pcdh</i> locus 3	29.401			5/26 (19.2%)
Mice	Tyrosinase (<i>Tyr</i>) non-coding regulatory DNA elements	1.2	19/64 (29.7%)	Mutant mice by zygote injection	Seruggia et al. (2015)	

continued

Table 1 (continued)

Chromosomal rearrangement	Cell type or organism	Gene or region of interest	Targeting size (kb)	Targeting efficiency (%)	Efficiency measuring method	References
	Human Pluripotent Stem Cells (hESC)	<i>MALAT1</i>	0.5 1 3 8	7/12 (58.3%) 6/8 (75%) 18/32 (56.3%) 18/39 (46.2%)	Screening single cell clones	Liu et al. (2016)
	Mice	<i>Tyr</i> <i>Nox4</i> <i>Grm5</i> <i>Nox4 to Grm5</i>	9.5 65 155 545 1150	3/30 (10%) 13/81 (16%) 11/46 (23.9%) 12/68 (17.6%) 14/48 (29.2%)	Mutant mice by zygote injection	Boroviak et al. (2016)
	Rats	<i>Cbs</i> <i>Dyrk1a</i> <i>Umodl1-Prmt2</i> <i>Lipi-Zfp295</i>	37.2 121.7 3513 24499	12/24 (50%) 4/28 (14.3%) 2/40 (5%) 1/9 (11.1%)	Mutant rat by zygote injection	Birling et al. (2017)
	Mice	<i>Hmgn1</i> <i>Tiam1</i> <i>Runx1-Cbr1</i>	16.8 226 1100	4/8 (50%) 8/41 (19.5%) 1/34 (2.9%)	Mutant mice by zygote injection	Birling et al. (2017)
	CHO cells (Chinese Hamster Ovary cells)	α -1,6-Fucosyltransferase 8 (<i>FUT8</i>)	2.1 12.5 52.6 96.8 150.7	34% 30% 29% 35% 21%	Quantitative PCR	Schmieder et al. (2018)
	Rabbits	Tyrosinase (<i>Tyr</i>)	105	3/17 (17.6%)	Mutant rabbits by zygote injection	Song et al. (2016)
	Rabbits	<i>GJA8</i>	0.054	11/11 (100%)	Mutant rabbits by zygote injection	Yuan et al. (2016)
	Pigs	<i>PDX1</i>	0.204	3/9 (33.3%)	Mutant pigs by zygote injection	Wu et al. (2017)
	Rhesus monkeys	<i>PINK1</i>	7.237	3/11 (27.3%)	Mutant monkeys by zygote injection	Yang et al. (2019)
DNA fragment inversion	HEK293T	<i>KIF5B-RET</i> <i>EML4-ALK</i>	11000 12000	1.6% 8%	Flow cytometry Flow cytometry	Choi and Meyerson (2014)
	Mice	<i>EML4-ALK</i>	11000	1.5×10^{-6}	PCR	Blasco et al. (2014); Maddalo et al. (2014) Park et al. (2015)
	Patient iPSCs	<i>F8 gene</i>	140 563	8/120 (6.7%) 5/135 (3.7%)	Screening single cell clones	Canver et al. (2014)
	Murine erythroleukemia (MEL) cells	ND	2 8 15 20.3 1025.3 1025.7	20/156 (12.8%) 9/96 (9.4%) 17/164 (10.4%) 26/140 (18.6%) 2/266 (0.8%) 2/418 (0.5%)	Screening single cell clones	Kraft et al. (2015)
	Mouse ESC	<i>H2afy</i> <i>Bmp2</i> <i>Ihh</i> <i>Pitx1</i> <i>Laf4</i> <i>Epha4</i>	1.189 3.7 12.6 32 353 1672	2/288 (0.7%) 3/192 (1.6%) 7/288 (2.4%) 3/288 (1%) 12/288 (4.2%) 3/192 (1.6%)	Mutant mice by zygote injection	Li et al. (2015a)
	HEK293T	β -globin RE1 <i>Pcdh RE1</i> β -globin RE2 <i>HoxD</i> β -globin <i>Pcdhα cluster</i> <i>Pcdh α, β, and γ clusters</i>	0.709 1.272 6.277 18.142 80.732 256.744 807.480	(21.12 \pm 4.99)% (23.28 \pm 2.42)% (23.13 \pm 1.13)% (7.28 \pm 1.60)% (5.96 \pm 0.28)% (5.48 \pm 0.37)% (0.71 \pm 0.12)%	Quantitative PCR	Li et al. (2015a)
	Mice	<i>Pcdh</i> locus 1 <i>Pcdh</i> locus 2 <i>Pcdh</i> locus 3	1.241 0.96 29.401	6/120 (5%) 8/8 (100%) 2/26 (7.7%)	Mutant mice by zygote injection	Li et al. (2015a)
	Mice	Tyrosinase (<i>Tyr</i>) non-coding regulatory DNA elements	1.2	7/64 (10.9%)	Mutant mice by zygote injection	Seruggia et al. (2015)
	Mice	<i>Nox4</i> <i>Grm5</i> <i>Nox4 to Grm5</i>	155 545 1150	14/46 (30.4%) 12/68 (17.6%) 10/48 (20.8%)	Mutant mice by zygote injection	Boroviak et al. (2016)

continued

Table 1 (continued)

Chromosomal rearrangement	Cell type or organism	Gene or region of interest	Targeting size (kb)	Targeting efficiency (%)	Efficiency measuring method	References	
	Rat	<i>Cbs</i>	37.2	7/24 (29.2%)	Mutant rat by zygote injection	Birling et al. (2017)	
		<i>Dyrk1a</i>	121.7	3/28 (10.7%)			
	Mice	<i>Runx1-Cbr1</i>	1100	1/34 (2.9%)	Mutant mice by zygote injection	Birling et al. (2017)	
DNA fragment duplication	Mouse ESC	<i>Pitx1</i>	32	2/288 (0.7%)	Screening single cell clones	Kraft et al. (2015)	
	HEK293T	<i>Laf4</i>	353	81/288 (28.1%)	Quantitative PCR	Li et al. (2015a)	
		<i>Pcdh RE1</i>	1.272	(0.23 ± 0.12)%			
		<i>β-globin RE2</i>	6.277	(5.30 ± 1.19)%			
		<i>β-globin</i>	80.732	(5.97 ± 0.33)%			
		<i>Pcdhα cluster</i>	256.744	(0.61 ± 0.02)%			
		<i>Pcdh α, β, and γ clusters</i>	807.480	(0.17 ± 0.03)%			
	Mice	<i>Pcdh locus 1</i>	1.241	1/26 (3.8%)	Mutant mice by zygote injection	Li et al. (2015a)	
	Mice	<i>Nox4</i>	155	1/46 (2.2%)	Mutant mice by zygote injection	Boroviak et al. (2016)	
			<i>Grm5</i>	545	1/68 (1.5%)		
	Rat	<i>Cbs</i>	37.2	1/24 (4.2%)	Mutant rat by zygote injection	Birling et al. (2017)	
			<i>Dyrk1a</i>	121.7	2/28 (7.1%)		
		<i>Lip1-Zfp295</i>	24499	1/9 (11.1%)			
Mice	<i>Tiam1</i>	226	1/41 (2.4%)	Mutant mice by zygote injection	Birling et al. (2017)		

Relationship between DNA fragment size and editing frequency

Deletion frequencies at some loci are inversely correlated with the sizes of the intervening sequences between the two cleavage sites (Canver et al., 2014). However, at other loci, there is no inverse correlation between DNA-fragment-deletion frequency and the fragment size (Table 1; He et al., 2015; Kraft et al., 2015; Li et al., 2015a; Schmieder et al., 2018). In addition, the frequencies of DNA-fragment inversion and DNA-fragment duplication have no relationship with fragment sizes (Table 1). The DNA fragment-editing frequency may be related to the locus-specific 3D chromatin structure as well as the spatial distance between the two cutting sites, which is an unresolved problem requiring further studies.

DNA fragment insertion by CRISPR

DNA fragment insertion can be efficiently achieved through the CRISPR system using Cas9 with either dual sgRNAs or single sgRNAs (Figure 2E). Mechanistically, DNA fragment insertions can be achieved by either HR or non-homologous end-joining (Suzuki et al., 2016). It is known that single cuts by Cas9 stimulate DNA fragment insertion through HR with a donor template harboring flanking homologous arms. One study carefully investigated the DNA fragment insertion efficiencies of HR by Cas9 with dual sgRNAs (Byrne et al., 2015). Moreover, Cas9 with dual sgRNAs targeting both the genome and donor template may be more efficient through homology-mediated end joining (HMEJ) (Yao et al., 2017). However, insertion needs careful screening for single-copy insertional clones or mice because any donor template could result in random head-to-tail tandem insertions just as transgenes (Figure 1B; Folger et al., 1982; Skryabin et al., 2020). Thus, the DNA fragment insertion clones or mice are best screened by Southern blot analyses rather than by PCR only.

Many ways to cut and heal

The mutated sequences obtained from CRISPR/Cas9-editing result from eventual consequences of the opposite forces of Cas9 cleavage and cellular repair. Specifically, the observed random indels by Cas9 with single sgRNAs are eventual repaired outcomes after cycles of repeated ligation and cleavage of precisely ligated DNA ends. In addition to blunt-end cleavage, Cas9 can also cohesively cleave the DNA duplex generating staggered ends with 5' overhangs. Thus, the cohesive cleavage of Cas9 actually generates diverse profiles of DSB ends with distinct 5' overhangs. Finally, rapid progress in the field has made it possible to predict editing outcomes by manipulating DNA repair pathways (Long, 2019; Yeh et al., 2019).

Double cutting vs. single cutting

The plain difference between cleavages of double and single cutting is that double cutting generates four DSB ends. The combinatorial ligations of two of these four DSB ends result in a variety of chromosomal rearrangements (Figure 2). The fundamental difference between double and single cutting is that in single cutting, after precise ligation of the two DSB ends, the repaired sequences still match the targeting sgRNA and thus can be recut. In contrast, the ligations of combinatorial two DSB ends out of the four ends from double cutting cannot be recut since the rearranged junctional sequences no longer match either of the two targeting sgRNAs (Huang and Wu, 2016; Shou et al., 2018; Shi et al., 2019). Therefore, dual-sgRNA-mediated chromosomal rearrangements maintain the integrity of Cas9-cleavage ends and make them less vulnerable to end-processing by repair enzymes (Figure 2). Hence, precise ligations upon direct rejoining of Cas9 blunt-cleavage ends after double cutting are much more frequent than after single

cutting (Li et al., 2015a; Zhu et al., 2016b; Guo et al., 2018; Shou et al., 2018).

Cohesive Cas9 cleavage in vitro and in silico

Since the advent of Cas9-mediated genome editing, it has long been assumed that Cas9 cleaves the targeting DNA duplex at the -3 position upstream of the PAM site, generating blunted DSB ends with no overhang (Figure 1D; Gasiunas et al., 2012; Jinek et al., 2012). In contrast to the earlier finding that Cas9 has potential exonuclease activity, *in silico* molecular dynamics modeling and *in vitro* high-throughput sequencing suggest that Cas9 cleaves the non-complementary strand at the -4 position upstream of the PAM site (Kim et al., 2016; Palermo et al., 2016; Zuo and Liu, 2016). In addition, *in vitro* cleavage of dsDNA, whose non-complementary strand is labeled at the 3' ends, reveals both blunted and cohesive Cas9 cleavages (Shou et al., 2018; Stephenson et al., 2018). Specifically, *in vitro* cleavage of dsDNA duplex with the 3'-biotin-labeled non-complementary strand reveals flexible cleavages at the -4 and -3 positions upstream of the PAM site (Shou et al., 2018). Finally, deep sequencing of *in vitro* Cas9-cleaved products reveals flexible cleavages of the non-complementary strand at the -6 , -5 , -4 , and -3 positions upstream of the PAM site but the exact cleavage of the complementary strand at the -3 position (Shi et al., 2019). Collectively, these studies clearly show that Cas9 endonucleolytically cleaves the non-complementary strand at the -6 , -5 , -4 , and -3 positions *in vitro*, generating cohesive DSB ends with 1–3-nt 5' overhangs as well as blunted ends (Figure 3A).

Cohesive Cas9 cleavage in vivo

Overwhelming evidence suggests cohesive Cas9 cleavage *in vivo*. First, the predicted metal coordination distance to the -3 phosphate is much larger than expected for the typical RuvC catalysis (Figure 3B; Chen and Doudna, 2017). Second, Cas9-mediated nucleotide insertions at junctions of DNA fragment editing are strongly biased toward nucleotides at the -6 , -5 , and -4 positions upstream of the PAM site *in vivo* (Figure 3A; Shou et al., 2018; Shi et al., 2019). Finally, by engineering the Cas9 hinge regions located between the HNH and RuvC nuclease domains, rationally designed Cas9 variants display R-loop-dependent alterations of the scissile profile of the non-complementary strand *in vivo* (Figure 3A; Shou et al., 2018). Taken together, these studies suggest that Cas9 cleaves targeting DNA duplex with flexibility on the non-complementary strand, resulting in DSB ends with 5' overhangs.

Mechanism of cohesive Cas9 cleavage

Cas9 RuvC and HNH nuclease domains cleave non-complementary and complementary strands via putative two-metal-ion and one-metal-ion mechanisms, respectively (Jinek et al., 2014; Nishimasu et al., 2014; Chen and Doudna, 2017). In both the two-metal-ion and one-metal-ion mechanisms,

nucleophilic attack is always in-line from the 5' site of the phosphodiester bond, resulting in 5' phosphate and 3' hydroxyl groups (Figure 3B; Yang, 2010). Whereas one magnesium ion coordinates Cas9 HNH active sites to the scissile phosphate at exactly the -3 position upstream of NGG PAM after a large conformational change, two magnesium ions coordinate Cas9 RuvC active sites to the scissile phosphate at positions further upstream of PAM, resulting in flexible Cas9 cleavages with variable staggered 5' overhangs.

After cutting—DSB repair pathways

DNA damage response pathways are activated after Cas9 cleavage to repair the resulting DSBs. The repair of mammalian DSBs involves three possible pathways: HR, canonical non-homologous end-joining (cNHEJ), and alternative non-homologous end-joining (aNHEJ) that includes microhomology-mediated end joining (MMEJ) (Figure 3C; Chang et al., 2017). In mammalian cells, when a template donor is available, the HR repair pathway is used to achieve precise genome editing, including insertion or replacement of specific sequences. However, the low efficiency of HR limits its usage (Ceccaldi et al., 2016a). When no donor is provided, both cNHEJ and aNHEJ (Figure 3C) are predominant pathways for repairing DSBs introduced by Cas9.

In the cNHEJ repair pathway, the Ku70–Ku80 heterodimer recognizes DSB ends to protect them from being processed by resection nucleases (Figure 3C; Deriano and Roth, 2013). The DNA-dependent protein kinase catalytic subunit (DNA-PKcs) and the endonuclease Artemis are then recruited to the Ku-DNA ends. They form an Artemis–PK–Ku complex at the DSB ends. Finally, precise ligations of the two DSB ends are catalyzed by the ligase IV–XRCC4–XLF complex (Deriano and Roth, 2013). Thus, cNHEJ is an accurate and precise DSB repair pathway (Shou et al., 2018).

The aNHEJ pathway was originally thought to be a backup repair mechanism for cNHEJ and it usually introduces small indels (Figure 3C). If the cNHEJ repair pathway is not available or is disrupted, the DSB ends will be repaired by the aNHEJ pathway, resulting in error-prone large indels or chromosomal rearrangements. Indeed, in species with no cNHEJ pathway, the genomes are prone to chromosomal rearrangements via aNHEJ (Deng et al., 2018).

In the aNHEJ pathway, extensive resections of DSB ends are catalyzed by several resection nucleases including the MRE11–RAD50–NBS1 (MRN) complex (Nijmegen breakage syndrome protein 1 or nibrin). These resections are facilitated by CtBP-interacting protein (CtIP or RBBP8) and FANCD2 (Ceccaldi et al., 2016b; Chang et al., 2017; Shou et al., 2018). The resection exposes single-stranded DNA (ssDNA) overhangs that could be annealed by complementary base pairing. The annealed DSB ends are then ligated by XRCC1 and DNA ligase III of the aNHEJ pathway, generating indels (Chang et al., 2017). Thus, cNHEJ- and aNHEJ-mediated DNA repairs either join the DSB ends directly or modify them slightly, resulting in precise ligation or small indels, respectively (Figure 3C).

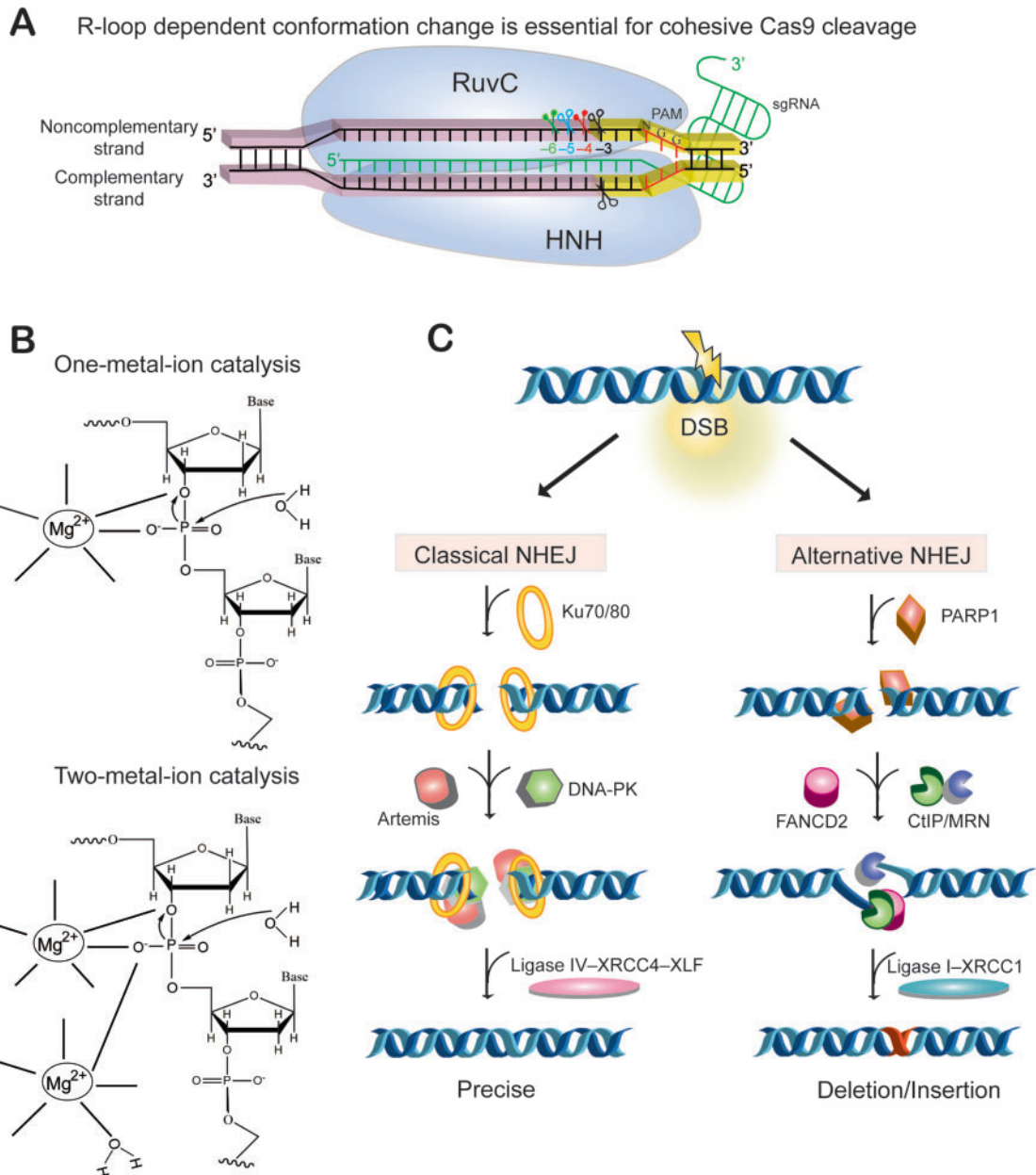


Figure 3 Mechanisms of cohesive Cas9 cleavage and repair. **(A)** Cas9 endonuclease reprogrammed by a synthetic guide RNA (sgRNA) can target any specific site in a genome through forming a structure composed of three strands of nuclear acid chains, known as R-loop. Specifically, the first 20 nucleotides of the sgRNA form a DNA–RNA hybrid with 20 nucleotides of the targeting DNA sequences through base-pairing with the complementary strand, displacing the non-complementary strand (the original protospacer sequences) and resulting in a structure called R-loop. SgRNA guides Cas9 to the targeting site and Cas9 cleaves dsDNA at locations upstream of the PAM site. While the HNH domain of Cas9 cleaves the complementary strand at the exact -3 position upstream of the PAM site, the RuvC domain of Cas9 flexibly cleaves the non-complementary strand at the -6 , -5 , and -4 positions as well as the -3 position upstream of the PAM site, generating a diverse cohesive DSB ends with 1-, 2-, and 3-nt 5' overhangs in addition to blunt ends. **(B)** Diagram of one-metal-ion cleavage mechanism for HNH and two-metal-ion cleavage mechanism for RuvC domain of Cas9 protein. **(C)** Schematic of NHEJ repair pathways for repairing of a targeted DSB. NHEJ includes two competing pathways known as classic or canonical NHEJ (cNHEJ) and alternative NHEJ (aNHEJ). The cNHEJ pathway requires XRCC4 and DNA ligase IV. The aNHEJ pathway includes MMEJ. The cleaved DSB ends are ligated by cellular DNA repairing machineries using either the precise pathway of cNHEJ or the mutagenic pathway of MMEJ.

Random vs. non-random indels

Initial gene editing by CRISPR indicates that prevalent random indels are induced by Cas9 cleavage programmed with single sgRNAs in heterologous systems (Cho et al., 2013; Cong et al., 2013; Jinek et al., 2013; Mali et al., 2013). Similarly, random small indels at the junctions of chromosomal rearrangements—or at the Cas9 cleavage site for the so-called scarring—are also introduced by DNA fragment editing with Cas9 reprogrammed with dual sgRNAs (Canver et al., 2014; Kraft et al., 2015; Li et al., 2015a). These random indels likely result from the NHEJ repair pathway (Figure 3C; Jiang and Marraffini, 2015; Huang and Wu, 2016).

Subsequent studies by Cas9 reprogrammed with dual sgRNAs show that, in addition to random indels or scarring at individual cleavage sites and rearranged junctions (Cong et al., 2013; Mali et al., 2013; Wang et al., 2013; Xiao et al., 2013; Canver et al., 2014; Guo et al., 2015; He et al., 2015; Kraft et al., 2015; Li et al., 2015a; Schmieder et al., 2018; Shou et al., 2018; Shi et al., 2019), there are predominant ligations at exactly the -3 positions and precise chromosomal rearrangements (Figure 4A; Canver et al., 2014; Guo et al., 2015; Li et al., 2015a; Huang and Wu, 2016; Zhu et al., 2016b). Moreover, profiling of DNA repair outcomes demonstrates that indels induced by Cas9 programmed with single sgRNAs are non-random and are related to sequences of the protospacer (van Overbeek et al., 2016). Finally, recent studies revealed that editing outcomes by the CRISPR/Cas9 system are precise (Figure 4A) and predictable (Figure 4B; Allen et al., 2018; Chakrabarti et al., 2018; Shen et al., 2018; Shou et al., 2018; Taheri-Ghahfarokhi et al., 2018; Chen et al., 2019; Iyer et al., 2019; Leenay et al., 2019; Long, 2019; Molla and Yang, 2020).

Predictable deletions

When homologous sequences near the DSB ends generated by Cas9 with single sgRNAs are direct repeats, small deletions could be generated via the MMEJ pathway (Figure 4B; McVey and Lee, 2008; Shou et al., 2018). Specifically, if resections expose short complementary sequences within 3' overhangs, they will form a DNA duplex and the 3' flap will be cleaved by flap endonuclease 1 (FEN1), resulting in predictable deletions (Figure 4B; Iyer et al., 2019). Similarly, when direct repeats flank the two cleavage sites of Cas9 targeted by dual sgRNAs, the intervening sequences could be deleted via the MMEJ pathway (Figure 4C; McVey and Lee, 2008; Shou et al., 2018).

Predictable nucleotide insertions at editing junctions

CRISPR-editing technologies are moving forward at lightning speed. It used to be thought of as uncontrollable or unpredictable but now is considered predictable through machine learning approaches. For example, base-editing outcomes have recently been shown to be predictable (Arbab et al., 2020). In this section, we focus on predictable nucleotide insertions based on the mechanistic understanding of cohesive or staggered Cas9 cleavage. In particular, the cohesive Cas9 cleavage

mechanism has a profound impact on gene-editing outcomes of the CRISPR system in a wide variety of scenarios and species. If Cas9 cleavage ends with single-nucleotide 5' overhangs are filled in and ligated, it will result in duplications of the -4 nucleotide (Table 2). Similarly, if Cas9 cleavage ends with 2-nt overhangs are filled in and ligated, it will lead to repetition of the dinucleotide of the -5 and -4 positions (Table 2). Finally, if Cas9 cleavage ends with 3-nt overhangs are filled in and ligated, it will produce repetition of the trinucleotide of the -6 , -5 , and -4 positions (Table 2).

Predictable single-nucleotide insertions at single cutting sites

Extensive studies have shown that Cas9-mediated single-nucleotide insertions at repair junctions in budding yeast, mouse ESCs, mammalian cell lines, and mice are predictable (Figure 4D; Chakrabarti et al., 2018; Kalhor et al., 2018; Lemos et al., 2018; Shen et al., 2018; Shou et al., 2018; Taheri-Ghahfarokhi et al., 2018; Chen et al., 2019; Gisler et al., 2019; Leenay et al., 2019). When Cas9 reprogrammed with single sgRNAs cleaves the non-complementary strand at the -4 position, it will generate two cohesive ends with 1-nt 5' overhangs, which could be filled-in by an unknown polymerase (Figure 4D). The two filled-in DSB ends are then ligated directly, generating single-nucleotide insertion which is the duplication of the -4 nucleotide upstream of PAM (Figure 4D).

This ligation mechanism is via the cNHEJ pathway since blocking XRCC4 results in a significant decrease of precise ligation in DNA fragment editing (Shou et al., 2018). In addition, knocking down of DNA ligase IV leads to a significant decrease of precise DNA-fragment-deletion efficiency, suggesting that cNHEJ is an error-free DNA repair pathway (Shou et al., 2018). Therefore, numerous cases of 1-bp insertions, which were reported as random insertions, actually result from Cas9 cohesive cleavage at the -4 position (Table 2). For example, the Nana '+1' allele of *CCR5* of the unethically edited baby (Ryder, 2018) is probably generated by cohesive Cas9 cleavage at the -4 position, resulting in two DSB ends with 1-nt 5' overhang, which are then filled in and ligated precisely (Figure 4E). All in all, gene editing via Cas9 cohesive cleavage at the -4 position generates predictable 1-bp insertions (Table 2).

Dinucleotide and trinucleotide insertions at single cutting sites

If Cas9 RuvC domain cleaves the non-complementary strand at the -5 or -6 position upstream of PAM, it will generate two cohesive DSB ends each with a dinucleotide or trinucleotide 5' overhang. After both of them get filled-in, these filled-in ends could be blunt-end ligated via the cNHEJ pathway. This will generate a dinucleotide or trinucleotide insertion, which is the tandem duplication of the dinucleotide or trinucleotide further upstream of the -3 position of PAM (Table 2; Figure 4F).

Prominent predictable nucleotide insertions at rearranged junctions of double cutting

Systematic analyses of the inserted nucleotides reveal predictable nucleotide insertions at the junctions of chromosomal

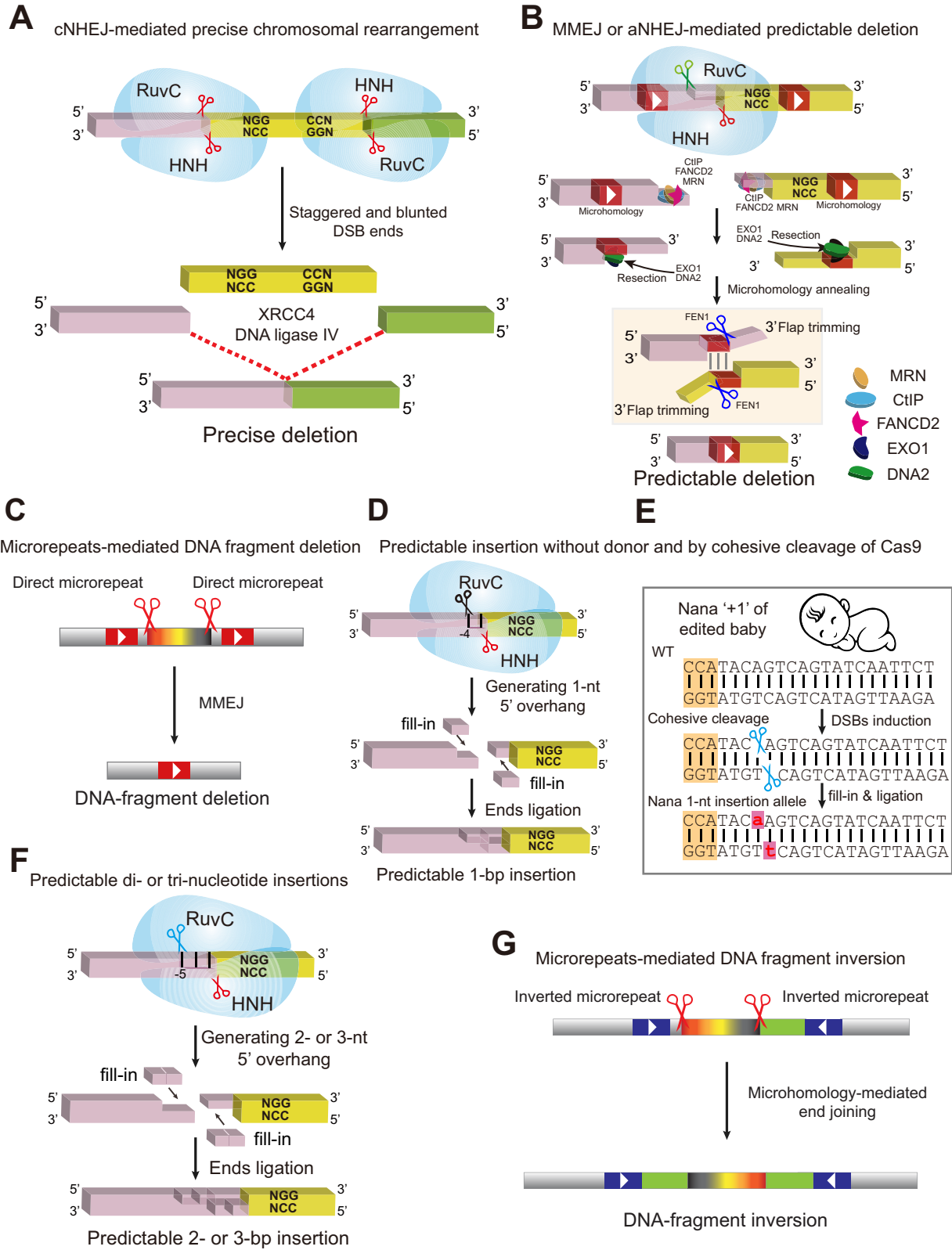


Figure 4 Mechanisms of precise and predictable CRISPR/Cas9 genome editing. **(A)** Precise chromosomal rearrangements by DNA fragment editing. cNHEJ-mediated precise DNA fragment deletion could be generated through direct ligation by XRCC4–DNA ligase IV of the two staggered or blunted DSB ends from Cas9 cleavage with NGG–CCN PAM configuration. In particular, perturbations of CtIP or FANCD2, two proteins involved in the aNHEJ pathway, enhance the cNHEJ-mediated precise DNA fragment deletion. **(B)** Predictable deletions. The cohesive

rearrangements by Cas9 with dual sgRNAs (Table 3; Shou et al., 2018). Interestingly, the frequency of nucleotide insertions (1, 2, or 3 nt) is much higher at junctions of chromosomal rearrangements by double cutting than that by single cutting (Shi et al., 2019). The reason for the increased insertion frequency at rearranged junctions is that the ligated junctions of chromosomal rearrangement after Cas9 double cleavages cannot be recut. For single Cas9 cleavages, the two cohesive DSB ends are always complementary to each other (Figure 3A). After annealing of the cohesive ends and ligation by cellular repair machineries, it will be recut by Cas9 programmed with the same sgRNA. By contrast, any two DSB ends from chromosomal rearrangements, which have distinct 5' overhangs, are rarely complementary to each other, and thus cannot be annealed and recut by Cas9 programmed with either of the two original sgRNAs.

There are barely any 2- or 3-bp insertions with Cas9 reprogrammed with single sgRNAs (Figure 4F; Allen et al., 2018; Shen et al., 2018; Chen et al., 2019; Leenay et al., 2019). In addition, Cas9 reprogrammed with single sgRNA shows significantly higher frequency of 1-bp insertions than 2- or 3-bp insertions (Chen et al., 2019; Shi et al., 2019). The reason that 2- or 3-bp insertions with Cas9 guided by single sgRNAs are much less observable (Allen et al., 2018; Shen et al., 2018; Leenay et al., 2019; Shi et al., 2019) than by dual sgRNAs (Shou et al., 2018; Shi et al., 2019; Figure 4F) is that the annealing efficiencies of 2- or 3-bp overhangs after Cas9 single cleavages are much higher than that of 1-bp overhangs, and thus the repaired 2- or 3-bp cohesive overhangs are more frequent to be recut. Overall, predictable nucleotide insertions are easily observed at junctions of chromosome rearrangements by Cas9 with dual sgRNAs (Figure 5; Shou et al., 2018; Shi et al., 2019).

Toward precise and predictable genome editing

In order to achieve precise and predictable genome editing, the Cas9 endonuclease effector needs first to be located

precisely to a targeting site. Once targeted to a genome site, the Cas9 effector can make a predictable modification on the sequences of the targeting site. Novel derivative gene-editing systems such as base editing and prime editing are developed rapidly (Anzalone et al., 2020; Yang and Chen, 2020). The base-editing system is achieved by fusing dCas9 with a nucleobase deaminase such as cytidine deaminases of the APOBEC/AID family or adenosine deaminase (Komor et al., 2016; Gaudelli et al., 2017). The prime-editing system is achieved by fusing H840A Cas9 with a reverse transcriptase and also fusing sgRNA with designed sequences functioning as a priming RNA template for reverse transcription, so-called prime-editing guide RNA or pegRNA (Anzalone et al., 2019). Both of these new gene-editing systems have advantages of precise editing without the requirement of DNA donor templates and DSBs. In this section, we focus only on precise and predictable genome editing derived from the mechanistic understanding of the Cas9 catalysis.

Factors influencing CRISPR genome editing

Various factors influence the complexity of DNA repair outcomes, including the type of DNA repair pathways chosen by host cells, the diversity of DSB ends from Cas9 cleavage, and the 3D genome sequence context surrounding the DSBs. In particular, inhibiting the aNHEJ pathway by knocking down its component proteins of CtIP or FANCD2 enhances precise DNA fragment deletion since cNHEJ and aNHEJ compete with each other for repair substrates (Figure 3C; Shou et al., 2018). Conversely, overexpression of CtIP protein facilitates usage of the MMEJ pathway and results in predictable deletions (Figure 4B; Nakade et al., 2018). In addition, interplays between structures of DSB ends and cellular repair protein machineries (resection nucleases, polymerases, and ligases) likely determine end-joining patterns. Indeed, DSB polarity influences repair outcomes at the editing junctions of Cas9-induced artificial class switching and translocations in human B cells (So and Martin, 2019).

and blunted DSB ends could be resected by the MRN complex, resulting in 3' overhangs. This resection process could be facilitated by CtIP and FANCD2 proteins. Further resection by EXO1 and DNA2 nucleases exposes micro-homologous sequences in the vicinity of the cleavage site. Base-pairing between the microhomologous sequences and removal of the two 3' overhanging flaps by FEN1 generate predictable deletions. (C) Large DNA fragment deletion could also be achieved by MMEJ. When there are direct repeats flanking the two cleavage sites by Cas9 with dual sgRNAs, MMEJ-mediated repair could induce deletion of the intervening sequences between the two direct repeats (rather than between the two cleavage sites through cNHEJ repair pathway). (D) Predictable single-nucleotide insertions. Cleavage at the -4 position by Cas9 generates cohesive DSB ends with 1-nt 5' overhangs. Fill-in and ends ligation by cellular repair machineries result in predictable 1-bp insertions, which are the duplication of the -4 nucleotide. (E) The Nana '+1' allele of the human *CCR5* gene in the CRISPR-edited baby probably results from cohesive Cas9 cleavage at the -4 position of the non-complementary strand. (F) Predictable di- or tri-nucleotide insertions. Cleavage at the -5 (or -6) position by Cas9 generates cohesive DSB ends with 2-nt (or 3-nt) 5' overhangs. Fill-in and ends ligation by cellular repair machineries result in predictable 2-bp insertions, which are the duplication of dinucleotide from the -5 and -4 positions. Thus, nucleotide insertions mediated by Cas9 reprogrammed with single sgRNAs manifest as tandem repeats. Finally, nucleotide insertions mediated by Cas9 reprogrammed with dual sgRNAs at various junctions of chromosomal rearrangements are generated by filled-in of cohesive DSB ends. (G) Predictable DNA fragment inversion. Large DNA fragment inversion could also be achieved by MMEJ. When there are microhomologous inverted repeats flanking the cleavage sites by Cas9 with dual sgRNAs, MMEJ-mediated repair can induce predictable inversion of the intervening sequence between the inverted repeats (rather than between two cleavage sites through cNHEJ repair pathway).

Table 2 Predictable nucleotide insertions by cohesive Cas9 cleavage with single sgRNAs.

Cell line/organism	Locus	Inserted nt	Reference sequence 5'–3', mutant sequence 5'–3'	Cohesive cleavage	Reference
Humans	<i>EMX1</i>	WT	GAGTCCGAGCAGAAGAAGAGGG	aGAAGGG CTTCCC	Cong et al. (2013)
		(+1)	GAGTCCGAGCAGAAGAAaGAAGGG		
Rats	<i>Tet1</i>	WT	ATGAAGACATTGCTGGAGACTGTCTG	atTGCTGG ACGACC	Li et al. (2013b)
		(+2)	ATGAAGACATa tTGCTGGAGACTGTCTG		
Mice	<i>Tet2</i>	WT	GGCTGCTGTCTCAGGGAGCTCATGG	cTCATGG AGTACC	Wang et al. (2013)
		(+1)	GGCTGCTGTCTCAGGGAGCC cTCATGG		
K562 cells	<i>CCR5</i>	WT	TGACATCAATTATTATACATCGG	aCATCGG GTAGCC	Cho et al. (2013)
		(+1)	TGACATCAATTATTATA aCATCGG		
	<i>C4BPB</i>	WT	AATGACCACTACATCCTCAAGGG	tCAAGGG GTTCCC	
		(+1)	AATGACCACTACATCCT tCAAGGG		
		(+2)	AATGACCACTACATCCT ctCAAGGG	ctCAAGGG GTTCCC	
HEK293T cells	<i>HBB</i>	WT	CCACGTTACCTTGCCCCACAGGG	cACAGGG TGTCCC	Cradick et al. (2013)
		(+1)	CCACGTTACCTTGCCCC cACAGGG		
	<i>CCR2</i>	WT	GTGTTTCATCTTTGGTTTGTGGG	tTGTGGG ACACCC	
		(+1)	GTGTTTCATCTTTGGTTT tTGTGGG		
(+2)	GTGTTTCATCTTTGGTTT ttTGTGGG	ttTGTGGG ACACCC			
Yeast	<i>CAN1</i>	WT	GATACGTTCTCTATGGAGGATGG	aGGATGG CCTACC	DiCarlo et al. (2013)
		(+1)	GATACGTTCTCTATGGAA gGATGG		
Zebrafish	<i>fh</i>	WT	CCCCGGTCGCCATGTACCGCTCC	CCCCGG GGGGCCa	Hwang et al. (2013)
		(+1)	CCCCGG tCGCCATGTACCGCTCC		
<i>Arabidopsis</i>	<i>AtPDS3</i>	WT	GGACTTTTGCCAGCCATGGTCCGG	tGGTCCG CCAGCC	Li et al. (2013a)
		(+1)	GGACTTTTGCCAGCCAT tGGTCCGG		
<i>Nicotiana benthamiana</i>	<i>NbPDS3</i>	WT	GCCGTTAATTGAGAGTCCAAGG	tCCAAGG GGTTCC	Li et al. (2013a)
		(+1)	GCCGTTAATTGAGAGT tCCAAGG		
Rice	<i>OsPDS</i>	WT	GTTGGTCTTTGCTCCTGCAGAGG	gCAGAGG GTCTCC	Shan et al. (2013)
		(+1)	GTTGGTCTTTGCTCCTG gCAGAGG		
Rice	<i>CAO1</i>	WT	CCAAGCTCTTGAGGTGGTCCGGT	CCAAGC GGTTCGa	Miao et al. (2013)
		(+1)	CCAAGC tCTTGAGGTGGTCCGGT		
Mice intestinal stem cells	<i>APC locus</i>	WT	CCCTCAAAAGCGTTTGTAGTGCC	CCCTCA GGGAGTt	Schwank et al. (2013)
Mice	<i>EGFP</i>	WT	GGAGCGCACCATCTTCTCAAGG	tTCAAGG AGTTCC	Shen et al. (2013)
		(+1)	GGAGCGCACCATCTTCT tTCAAGG		
Mice neuron	<i>GRIN1</i>	WT	AACCAGGCCAATAAGCGACACGG	gACACGG TGTGCC	Incontro et al. (2014)
		(+1)	AACCAGGCCAATAAGCG gACACGG		
K562 cells	<i>C4BPB</i>	WT	AATGACCACTACATCCTCAAGGG	tCAAGGG GTTCCC	Kim et al. (2014)
		(+1)	AATGACCACTACATCCT tCAAGGG		
		(+3)	AATGACCACTACATCCT ctCAAGGG	ctCAAGGG GTTCCC	
Mice	<i>NeuN</i>	WT	CCTTCCGGTTCAGGGACCCCGAC	CCTTCC GGAAGGc	Platt et al. (2014)
		(+1)	CCTTCC gGGTTCAGGGACCCCGAC		
		(+2)	CCTTCC ggGGTTCAGGGACCCCGAC	CCTTCC GGAAGGcc	

continued

Table 2 (continued)

Cell line/organism	Locus	Inserted nt	Reference sequence 5'–3', mutant sequence 5'–3'	Cohesive cleavage	Reference
Murine liver	<i>Pten</i>	WT	AGATCGTTAGCAGAAACAAAAGG	cAAAAGG TTTTCC	Xue et al. (2014)
		(+1)	AGATCGTTAGCAGAAACcAAAAGG		
	<i>P53</i>	WT	GCCTCGAGCTCCCTCTGAGCCAGG	aGCCAGG CGGTCC	
		(+1)	GCCTCGAGCTCCCTCTGAaGCCAGG		
Mice	<i>Fgf10</i>	WT	CCACCAACTGCTCTTCTTCTCTCC	CCACCA GGTGGTt	Yasue et al. (2014)
		(+1)	CCACCAaACTGCTCTTCTTCTCTCC		
Mice	<i>Tyr</i>	WT	GGGTGGATGACCGTGAGTCCTGG	gTCCTGG AGGACC	Fujii et al. (2014)
		(+1)	GGGTGGATGACCGTGAGgTCCTGG		
Mice	<i>Tet1</i>	WT	GGCTGCTGTGTCAGGAGCTCATGG	cTCATGG AGTACC	Horii et al. (2014)
		(+1)	GGCTGCTGTGTCAGGAGCtCATGG		
<i>Drosophila</i>	<i>singed (sn)</i>	WT	GCCAGCACAAAGTACATGACCGCGG	gaCCGCGG GGCGCC	Lee et al. (2014b)
		(+2)	GCCAGCACAAAGTACATGagaCCGCGG		
<i>Bombyx mori</i>	<i>Bmku70</i>	WT	GCCATTGGCGCCACCTAACATGG	aACATGG TGTACC	Ma et al. (2014)
		(+1)	GCCATTGGCGCCACCTAaACATGG		
Goat fibroblast	<i>Prp</i>	WT	AACCGCTATCCACCTCAGGGAGG	aGGGAGG CCCTCC	Ni et al. (2014)
		(+1)	AACCGCTATCCACCTCaGGGAGG		
Monkey	<i>Ppar-g</i>	WT	CCCTTCACTACTGTTGACTTCTC	CCCTTC GGGAAGt	Niu et al. (2014)
		(+1)	CCCTTCaACTACTGTTGACTTCTC		
HEK293T cells	<i>CCR5</i>	WT	TGACATCAATTATTATACATCGG	aCATCGG GTAGCC	Ramakrishna et al. (2014)
		(+1)	TGACATCAATTATTATAaCATCGG		
Mice	<i>Tyr</i>	WT	CCTATCGGCCATAACAGAGACTC	CCTATC GGATAGc	Yen et al. (2014)
		(+1)	CCTATCgGGCCATAACAGAGACTC		
Rats	<i>Tyr</i>	WT	TTTCCAGGATTATGTAATAGTGG	aTAGTGG ATCACC	Yoshimi et al. (2014)
		(+1)	TTTCCAGGATTATGTAAaTAGTGG		
		(+2)	TTTCCAGGATTATGTAAaaTAGTGG	aaTAGTGG ATCACC	
Mice	<i>Them2</i>	WT	CCTTAGTGGACAGCATCTCGACC	CCTTAG GGAATCa	Zhu et al. (2014)
		(+1)	CCTTAGtTGGACAGCATCTCGACC		
Mice	<i>Pitx1</i>	WT	CCTCACTAGAGTACAGGTGTGAA	CCTCAC GGAGTGa	Kraft et al. (2015)
		(+1)	CCTCACtTAGAGTACAGGTGTGAA		
HCT116 cells	<i>HPRT gene</i>	WT	CCAGACTGTAAGTGAATTACTTT	CCAGAC GGTCTGa	Liao et al. (2015b)
		(+1)	CCAGACtGTAAGTGAATTACTTT		
HCT116 cells	<i>Trex1</i>	WT	CCGTGTGCGAGTCTGGAGGGGAC	CCGTGT GGCACAc	
		(+1)	CCGTGTgGCGAGTCTGGAGGGGAC		
Zebrafish	<i>urod</i>	WT	AGTTCAGGGAATCACGGCAGGG	gGCAGGG CGTCCC	Ablain et al. (2015)
		(+1)	AGTTCAGGGAATCACGGgGCAGGG		
<i>Nicotiana benthamiana</i>	<i>Tomato yellow leaf curl virus</i>	WT	GGCCATCCGTATAATATTACCGG	tTACCGG ATGGCC	Ali et al. (2015)
		(+1)	GGCCATCCGTATAATAtTACCGG		
Murine myeloid progenitor cells	<i>Bim</i>	WT	GACAATTGCAGCCTGCTGAGAGG	tGAGAGG CTCTCC	Aubrey et al. (2015)
		(+1)	GACAATTGCAGCCTGCTtGAGAGG		
		(+2)	GACAATTGCAGCCTGCTctGAGAGG	ctGAGAGG CTCTCC	
Soybean	<i>GmFEI2</i>	WT	GTTGGACCTATACCTGCTGATGG	cTGATGG ACTACC	Cai et al. (2015)
		(+1)	GTTGGACCTATACCTGCtTGATGG		

continued

Table 2 (continued)

Cell line/organism	Locus	Inserted nt	Reference sequence 5'–3', mutant sequence 5'–3'	Cohesive cleavage	Reference
Tobacco	<i>NtPDS</i>	WT	GAGGCAAGAGATGTCCTAGGTGG	tAGGTGG TCCACC	Gao et al. (2015)
		(+1)	GAGGCAAGAGATGTCCTtAGGTGG		
Ghost cells	<i>CXCR4</i>	WT	GAAGAAACTGAGAAGCATGACGG	aTGACGG ACTGCC	Hou et al. (2015)
		(+1)	GAAGAAACTGAGAAGCAaTGACGG		
Jurkat T cells	<i>CXCR4</i>	WT	GTTCCAGTTTCAGCACATCATGG	aTCATGG AGTACC	
		(+1)	GTTCCAGTTTCAGCACaTCATGG		
Barley (<i>Hordeum vulgare</i>)	<i>HvPM19</i>	WT	GCTCTCCACTCTGGGCTCTTCGG	tCTTCGG GAAGCC	Lawrenson et al. (2015)
		(+1)	GCTCTCCACTCTGGGCTtCTTCGG		
HEK293 cells	<i>GFP</i>	WT	GTCGCCACCATGGTGAGCAAGGG	gCAAGGG GTTCCC	Liao et al. (2015a)
		(+1)	GTCGCCACCATGGTGAGgCAAGGG		
	<i>LTR</i>	WT	GGGAGCTCTCTGGCTAACTAGGG	aCTAGGG GATCCC	
		(+1)	GGGAGCTCTCTGGCTAAaCTAGGG		
Human intestinal organoids	<i>SMAD4</i>	WT	CCACCAAAAACGGCCATCTTCAGC	CCACCA GGTGGTt	Matano et al. (2015)
		(+1)	CCACCAaAAAACGGCCATCTTCAGC		
Soybean	<i>Glyma06g14180</i>	WT	GTGAAATTAACCAGCTGCAGTGG	gCAGTGG GTCACC	Sun et al. (2015)
		(+1)	GTGAAATTAACCAGCTGgCAGTGG		
Mice	<i>Pten</i>	WT	CCATCATCAAAGAGATCGTTAGCA	CCATCA GGTAGTt	Weber et al. (2015)
		(+1)	CCATCATtCAAAGAGATCGTTAGCA		
<i>Nicotiana attenuata</i>	<i>AOC</i>	WT	CAAAAGACTGTCAATTCCCTTGG	cCCTTGG GGAACC	Woo et al. (2015)
		(+1)	CAAAAGACTGTCAATTcCCTTGG		
<i>Arabidopsis</i>	<i>BRI1</i>	WT	TTGGGTCATAACGATATCTCTGG	tCTCTGG GAGACC	Yan et al. (2015)
		(+1)	TTGGGTCATAACGATATtCTCTGG		
<i>Nicotiana benthamiana</i>	<i>NblspH</i>	WT	GAATGGATATGAGTACACTTGGG	aCTTGGG GAACCC	Yin et al. (2015)
		(+1)	GAATGGATATGAGTACAaCTTGGG		
Mice	<i>Kcnj13</i>	WT	CCTGCGATGGACAGCAGTAATTG	CCTGCG GGACGct	Zhong et al. (2015)
		(+1)	CCTGCGaATGGACAGCAGTAATTG		
Mice	<i>Nf1</i>	WT	AGTCAGCACCGAGCACAAACAAGG	aACAAGG TGTTC	Zuckermann et al. (2015)
		(+1)	AGTCAGCACCGAGCACAAaACAAGG		
	<i>Pten</i>	WT	AAAGACTTGAAGGTGTATACAGG	aTACAGG ATGTCC	
		(+1)	AAAGACTTGAAGGTGTaTACAGG		
	<i>Trp53</i>	WT	ACAGCCATCACCTCACTGCATGG	tGCATGG CGTACC	
		(+1)	ACAGCCATCACCTCACTtGCATGG		
HEK293T, K562, HCT116	<i>Non-coding region</i>	WT	GGCAGTGCAGATGAAAACCTGGG	aACTGGG TGACCC	van Overbeek et al. (2016)
		(+1)	GGCAGTGCAGATGAAAAaACTGGG		
HEK293T, K562	<i>Chr1:65349091</i>	WT	GAGGAGCTCCAAGAAGACTGAGG	aCTGAGG GACTCC	
		(+1)	GAGGAGCTCCAAGAAGaACTGAGG		
<i>Yarrowia lipolytica</i>	<i>PEX10</i>	WT	GCCCAGCCCGAAAACATGGAAAGG	tGGAAGG CCTTCC	Gao et al. (2016)
		(+1)	GCCCAGCCCGAAAACATtGGAAGG		
		(+2)	GCCCAGCCCGAAAACATaTGGAAAGG	atGGAAGG CCTTCC	
Murine HSPCs	<i>Eed</i>	WT	TGCTTGCAATGGGCAATCAGG	aATCAGG TAGTCC	Gundry et al. (2016)
		(+1)	TGCTTGCAATGGGCAaATCAGG		

continued

Table 2 (continued)

Cell line/organism	Locus	Inserted nt	Reference sequence 5'–3', mutant sequence 5'–3'	Cohesive cleavage	Reference
<i>Taraxacum</i>	<i>Fructan 1-fructosyltransferase</i>	WT	ACAACCCGTACGCACCAATT TGG	aATTTGG TAAACC	Iaffaldano et al. (2016)
		(+1)	ACAACCCGTACGCACCA aa ATT TGG		
Apple	<i>PDS</i>	WT	ATGGCTTGAGCGTAAAAGACT TGG	aGACTGG CTGACC	Nishitani et al. (2016)
		(+1)	ATGGCTTGAGCGTAAAA a GACT TGG		
<i>Phaeodactylum tricornutum</i> cells	<i>CpSRP54</i>	WT	CCGCCCTTCGTGAAGTACGT CGG	aCGTCGG GCAGCC	Nymark et al. (2016)
		(+1)	CCGCCCTTCGTGAAGT a CGT CGG		
Chardonnay	<i>IdnDH</i>	WT	GGGAAAGGAGGCAACTCTG AGG	tCTGAGG GACTCC	Ren et al. (2016)
		(+1)	GGGAAAGGAGGCAACT t CTG AGG		
Maize immature embryo cells	<i>liguleless1 (LIG)</i>	WT	ATACGCGTACGCGTACGTGT AGG	tGTGAGG CACTCC	Svitashev et al. (2016)
		(+1)	ATACGCGTACGCGTACGT t GT AGG		
SNU719 cells	<i>EBV genomic locus of BART5</i>	WT	CCT CAAGGTGAATATAGCTGCC	CCTCAA GGAGTtc	van Diemen et al. (2016)
		(+1)	CCT CAAgGGTGAATATAGCTGCC		
HEK293 cells	<i>GFP</i>	WT	GGGCGAGGAGCTGTTCCCG GGG	aCCGGG GGCCCC	Yin et al. (2016)
		(+1)	GGGCGAGGAGCTGTTCA a CCG GGG		
Wheat	<i>TaGW2</i>	WT	CCT CTAGAAATGCCCATCCTG	CCTCTA GGAGATc	Zhang et al. (2016)
		(+1)	CCT CTAgGAAATGCCCATCCTG		
Maize	<i>PSY1</i>	WT	GAGACTTGAGGATCTGTTC ACGG	tTCACGG AGTGCC	Zhu et al. (2016a)
		(+1)	GAGACTTGAGGATCTGT t TC ACGG		
Gal4EED HEK293	<i>firefly luciferase</i>	WT	AAGAGATACGCCCTGGTTC TGG	gtTCCTGG AGGACC	Daer et al. (2017)
		(+2)	AAGAGATACGCCCTGGT gt TC TGG		
Chicken DF-1 fibroblasts	<i>Pax7</i>	WT	CCA TGGCTGATGACCAAGATCTG	CCATGG GGTACCg	Gandhi et al. (2017)
		(+1)	CCA TGGcCTGATGACCAAGATCTG		
Cotton	<i>GhPDS</i>	WT	GAAGCGAGAGATGTTCTAG TGG	tAGGTGG TCCACC	Gao et al. (2017)
		(+1)	GAAGCGAGAGATGTTCT t AG TGG		
Mice liver	<i>Ldlr</i>	WT	TGCTGCTGGCCAAGGACAT CGG	cATGCGG TACGCC	Jarrett et al. (2017)
		(+1)	TGCTGCTGGCCAAGGAC c AT CGG		
Bread wheat	<i>TaGW2</i>	WT	CCT CTAGAAATACCCCATCCTG	CCTCTA GGAGATc	Liang et al. (2017)
		(+1)	CCT CTAgGAAATACCCCATCCTG		
TZM-bl cells	<i>CXCR4</i>	WT	GCTTCTACCCCAATGACTT TGG	cTTGTGG AACACC	Liu et al. (2017b)
		(+1)	GCTTCTACCCCAATGAC c TT TGG		
Mice	<i>Kcnk13</i>	WT	CCT GAAACGAGGACAACGCGGCT	CCTGAA GGACTTg	Mianne et al. (2017)
		(+1)	CCT GAAcCGAGGACAACGCGGCT		
Hexaploid Camelina sativa	<i>FAD2</i>	WT	TCAAGGCTGTGTCCTA ACCGG	tAACCGG TTGGCC	Morineau et al. (2017)
		(+1)	TCAAGGCTGTGTCCT t A ACCGG		
T cells	<i>TCR a</i>	WT	TGTGCTAGACATGAGGTCTA TGG	tCTATGG GATACC	Ren et al. (2017)
		(+1)	TGTGCTAGACATGAGGT t CTA TGG		
Watermelon	<i>CIPDS</i>	WT	ATGCCGCTAGAGTGGT C CGG	tGCCCGG CGGGCC	Tian et al. (2017)
		(+1)	ATGCCGCTAGAGTGGT t G CCCGG		
MCF-7 cells	<i>HER2</i>	WT	GGGCATGGAGCACTTGC GAGAGG	cGAGAGG CTCTCC	Wang and Sun (2017)
		(+1)	GGGCATGGAGCACTTGC c G AGAGG		
Reef-building coral	<i>RFP</i>	WT	GTCTTCACTGAATAT CCTCAAGG	cTCAAGG AGTTCC	Cleves et al. (2018)
		(+1)	GTCTTCACTGAATAT c T CAAGG		

continued

Table 2 (continued)

Cell line/organism	Locus	Inserted nt	Reference sequence 5'–3', mutant sequence 5'–3'	Cohesive cleavage	Reference
Solanaceae crop <i>Physalis pruinosa</i>	<i>Ppr-SP</i>	WT	CCTTCCTTAGTCACCTCTAAACC	CCTTCC GGAAGGa	Lemmon et al. (2018)
		(+1)	CCTTCCTtTTAGTCACCTCTAAACC		
K562 cells	<i>ND</i>	WT	GCATCGGCCTGAAAGCAGTGAGG	aGTGAGG CACTCC	Allen et al. (2018)
		(+1)	GCATCGGCCTGAAAGCAaGTGAGG		
HPS1 B-LCL cells	<i>HPS1</i>	WT	CAGCAGGGGAGGCCCCcAGCAGG	cAGCAGG TCGTCC	Iyer et al. (2019)
		(+1)	CAGCAGGGGAGGCCCCcAGCAGG		

Mechanism for computer programs of machine learning

Precise and predictable Cas9-mediated genome editing could be achievable through machine learning. For example, computer programs with machine learning algorithms have been recently developed to predict repair outcomes and to achieve predictable genome editing (Allen et al., 2018; Shen et al., 2018; Chen et al., 2019; Leenay et al., 2019). Specifically, with editing using *SpCas9* with the PAM site of NGG, the presence of a nucleotide of 'T' or 'A' at the –4 position tends to result in more predictable 1-bp insertions. In contrast, the presence of a nucleotide of 'G' at the –4 position tends to generate more predictable deletions. The reason for this deletion preference is related to microhomology between the 'G' at the –4 position and the N'GG' of the PAM site (Shi et al., 2019).

Predictable MMEJ-mediated DNA fragment inversion

Short inverted repeats flanking the two cleavage sites induce microhomology-mediated inversion of the intervening sequences. Namely, when homology sequences near the DSB ends are inverted repeats, the intervening sequences can be inverted via the MMEJ pathway (Figure 4G; McVey and Lee, 2008; Li et al., 2015a). Therefore, MMEJ-mediated precise DNA fragment editing may be predicted from microhomologous sequences around the two cleavage sites.

Toward predictable chromosomal rearrangements

Cas9 programmed with dual sgRNAs induces predictable junctional insertions of DNA fragment editing since specific PAM configurations can generate distinct combinations of DSB ends from cohesive Cas9 cleavages (Figure 5; Shou et al., 2018). For example, in the NGG–NGG PAM configuration, the flexible cleavage profile of Cas9 with sgRNA2 can be obtained by sequencing rearranged junctions of DNA fragment deletion. Similarly, the flexible cleavage profile of Cas9 with sgRNA1 can be obtained by sequencing rearranged junctions of DNA fragment duplication. The nucleotide insertions at the downstream junctions of DNA fragment inversion can be easily predicted by the combined cleavage profiles of both sgRNAs (Figure 5A). Note that the upstream junctions of DNA fragment inversion for the NGG–NGG PAM configuration are always precise (Figure 5A). Similarly, the rearranged junctions of DNA fragment deletion (Figure 5B), the downstream junctions of DNA fragment inversion (Figure 5C),

and the rearranged junctions of DNA fragment duplication (Figure 5D) are always precise for the NGG–CCN, CCN–CCN, and CCN–NGG PAM configurations, respectively. In addition, the nucleotide insertions at rearranged junctions of DNA fragment duplication, the upstream junctions of DNA fragment inversion, and the rearranged junctions of DNA fragment deletion are predictable for the NGG–CCN, CCN–CCN, and CCN–NGG PAM configurations, respectively (Figure 5B–D). Understanding the mechanisms of chromosomal rearrangements will facilitate precise and predictable CRISPR DNA fragment editing.

Chromosomal rearrangement mechanisms in the context of 3D genome

After Cas9 cleavage, the histone H2AX within nucleosomes located in the regions flanking the DSB ends is phosphorylated by the ATM kinase, generating γ H2AX (Iacovoni et al., 2010; Lee et al., 2014a). Interestingly, a recent study showed that Cas9 is a genome mutator and induces γ H2AX accumulation (Xu et al., 2020). In addition, long-distance chromatin interactions are increased within the γ H2AX chromatin domains (Aymard et al., 2017). However, whether these increased chromatin interactions influence the form of the so-called 'DNA repair foci' needs further exploration (Marnef and Legube, 2017).

Several recent studies have shown that CTCF participates in DSB repair through its interaction with the repair proteins of BRCA2, RAD51, Mre11, and CtIP (Han et al., 2017; Hilmi et al., 2017; Lang et al., 2017; Hwang et al., 2019). In addition, cohesin inhibits distal DSB end joining (Gelot et al., 2016). Because CTCF and cohesin are known prominent 3D genome architecture proteins (Merkschlager and Nora, 2016), the recruitment of CTCF and its associated cohesin complex to the regions around DSB ends suggests that 3D genome architecture is closely related to DNA DSB repair.

3D motility of DSB ends in the nuclear space

In order to repair and ligate Cas9-induced DSB ends, they need to be brought into close spatial contact in the 3D nuclear space. Nuclear actin may play an important role in DSB motility required for both HR and NHEJ repairs (Caridi et al., 2018). Clustering of DSB ends and formation of a macro-repair center may be a prerequisite for proper chromosomal rearrangements by DNA fragment editing (Jasin and Rothstein, 2013; Aymard et al., 2017).

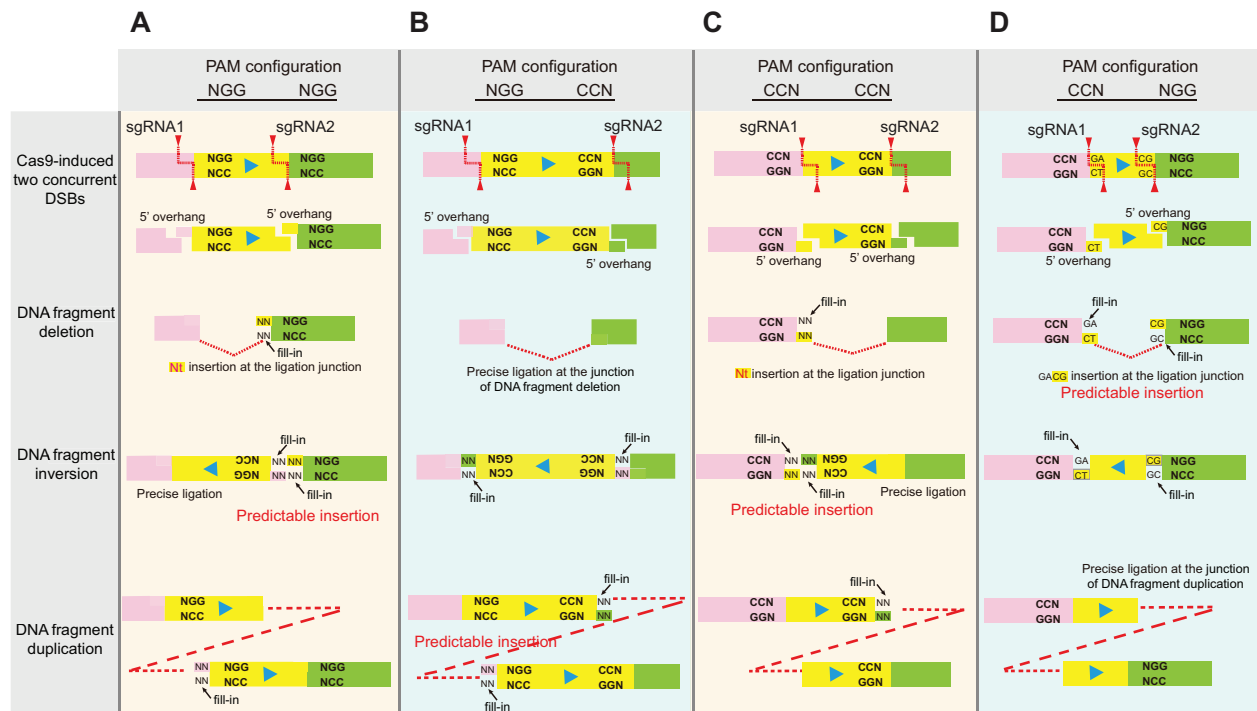


Figure 5 Precise and predictable Cas9-mediated nucleotide insertions at the junctions of chromosome arrangements for the four PAM configurations by Cas9 with dual sgRNAs. **(A)** In the NGG–NGG PAM configuration, the nucleotide insertions at the downstream junctions of DNA fragment inversion could be predicted based on combined flexible cleavage profiles of Cas9 with sgRNA1 and sgRNA2. However, the upstream junctions of DNA fragment inversion in the NGG–NGG PAM configuration are precise. **(B–D)** Similarly, the nucleotide insertions at the junctions of DNA fragment duplication **(B)**, at the upstream junctions of DNA fragment inversion **(C)**, and at the junctions of DNA fragment deletion **(D)** are predictable in the NGG–CCN **(B)**, CCN–CCN **(C)**, and CCN–NGG **(D)** PAM configurations, respectively. In addition, the ligations at the junctions of DNA fragment deletion **(B)**, at the downstream junctions of DNA fragment inversion **(C)**, and at the junctions of DNA fragment duplication **(D)** are precise in the NGG–CCN **(B)**, CCN–CCN **(C)**, and CCN–NGG **(D)** PAM configurations, respectively.

Toward precise and predictable 3D genome editing: from 1D to 3D

The higher order chromatin structure is highly dynamic and is regulated by epigenetic processes of DNA methylation, histone modification, and chromatin remodeling, ensuring proper cellular processes such as DNA replication, RNA transcription, and DNA damage repair in response to developmental or physiological signals (Dekker and Mirny, 2016; Hansen et al., 2018; Bickmore, 2019). Structural variations or chromosomal rearrangements affect 3D genome organization and gene expression. Editing of higher order chromatin structures or engineering chromosomal rearrangements to model genome structural variations not only sheds light on the fundamental mechanisms of 3D genome folding but also contributes to our understanding of aberrant 3D genome folding in human diseases (Wang et al., 2019b). Specifically, 3D genome engineering may pave the way to understanding vast GWAS data and CRISPR correction of aberrant alleles may lead to human disease therapy in the future (Qian et al., 2019).

Proximity ligation-based chromosome conformation capture (3C) technologies, in conjunction with high-throughput next-generation sequencing, have led to tremendous progress in understanding 3D genome architecture (Dekker et al., 2002; Rao

et al., 2014; Liu et al., 2017a; Tan et al., 2019; reviewed in Denker and de Laat, 2016; Zheng and Xie, 2019). In addition, fluorescence-labeled single-molecule imaging with super-resolution microscopy has shed significant light on the mechanisms of genome folding (Hansen et al., 2018; Sigal et al., 2018). Although genetic methods have long been used to investigate the position-effects variegations of chromatin organization (Lewis, 1950; McClintock, 1950), they have not been widely used to probe 3D genome organization compared to various chromosome conformation capture (3C, 4C, 5C, 6C, 7C, Hi-C, capture-C, etc.) ‘C’ technologies and imaging methods.

General principles of 3D genome organization

The 3D genomes in the nuclear space are thought to be assembled in a hierarchical manner composed of successive chromosomal territories, compartments or clustering regions, TADs or topological domains, and chromatin loops (Dekker and Mirny, 2016; Dixon et al., 2016; Bickmore, 2019). Briefly, each interphase chromosome occupies a distinct territory. Within a chromosome territory, chromatin fibers are segregated into active and inactive compartments with distinct histone modifications. Chromatin compartments are further divided into TADs or topological domains which are thought to be enriched in long-

distance chromatin contacts or loops (Bonev and Cavalli, 2016). Emerging evidence suggests, however, that chromosome compartments are smaller than previously thought and could be the consequences of gene activity (Rowley and Corces, 2018). Nevertheless, chromatin loops are fundamental units of the higher order chromatin structures.

CRISPR DNA fragment inversion reveals that the locations and relative orientations of CTCF sites determine the directionality of chromatin looping

Inversion of CTCF sites in the protocadherin alpha (*Pcdhα*) and *β-globin* clusters switches the directionality of chromatin looping (Guo et al., 2015; Shou et al., 2018; Jia et al., 2020). Specifically, the causality between orientation of mammalian insulators known as CTCF sites and directionality of long-distance chromatin looping is demonstrated by inverting CTCF sites using CRISPR DNA fragment-editing methods (Figure 6A; Guo et al., 2015; Shou et al., 2018; Lu et al., 2019; Jia et al., 2020). In addition, haplotype variants that alter chromatin looping topology are linked to human disease risks (Tang et al., 2015). In the *Sox2* and *Fbn2* loci, however, reinserting an inverted CTCF site in the original location does not form new chromatin loops (de Wit et al., 2015). Nevertheless, alterations of native chromatin loops have functional consequence on gene expression (de Wit et al., 2015; Guo et al., 2015). Moreover, genome-wide distributions of forward and reverse CTCF sites tend to be located in close 3D spaces (Rao et al., 2014; Guo et al., 2015). Thus, the relative orientations of CTCF sites determine the directionality of chromatin looping across mammalian genomes (Figure 6A). Specifically, there are strong long-distance chromatin interactions between forward and reverse convergent CTCF sites. However, there are weak long-distance chromatin interactions between two tandem CTCF sites in the same orientation. Finally, the configuration of reverse and forward CTCF sites constrains long-distance chromatin interactions between remote elements (Figure 6A). In summary, 3D genome structures could be predicted from 1D nucleotide sequences based on this CTCF-coding mechanism.

Mechanism of 3D genome folding by cohesin ‘loop extrusion’

The CTCF coding for the 3D genome could be explained by CTCF blocking of cohesin ‘loop extrusion’ along chromatin fibers (Guo et al., 2015; Nichols and Corces, 2015; Sanborn et al., 2015; Fudenberg et al., 2016; Merkschlager and Nora, 2016; Li et al., 2020b). The current model for the formation of TADs or topological domains is the cohesin sliding-mediated ‘loop extrusion’ (Banigan and Mirny, 2020). Specifically, CTCF helps to establish TADs boundaries by stalling the sliding of cohesin on DNA fibers and thus facilitates chromatin loop formations by ‘two-headed’ cohesin complex (Jia et al., 2020). Therefore, the cohesin complex can bring distant DNA elements into close spatial contact by the so-called active ‘loop extrusion’, which requires ATP as an energy source (Davidson et al., 2019; Kim et al., 2019). The genome-wide colocalization of CTCF and cohesin as well as a strong tendency of long-distance chromatin interactions between forward–reverse

convergent CTCF sites provide strong evidence for CTCF stalling of cohesin ‘loop extrusion’ (Parelho et al., 2008; Wendt et al., 2008; Rao et al., 2014; Guo et al., 2015). In addition, consistent with the model of cohesin ‘loop extrusion’, deletion of WAPL, a cohesin releasing factor, thus increasing cohesin enrichments on chromatin, results in a significant increase of TAD size (Gassler et al., 2017; Haerhuis et al., 2017; Wutz et al., 2017). Conversely, deletion of NIPBL, a cohesin loading factor, or deletion of cohesin directly, causes weakening or loss of chromatin loops (Rao et al., 2017; Schwarzer et al., 2017).

Asymmetric reeling of chromatin fibers by cohesin ‘loop extrusion’

In the *Pcdh* gene clusters, a large array of tandem forward CTCF sites in the variable region is followed by tandem reverse CTCF sites in the downstream super-enhancer (Guo et al., 2012; Zhai et al., 2016). CTCF/cohesin-dependent long-distance chromatin interactions bridge the distal enhancer to its target promoters and activate transcription. The reverse CTCF sites in the downstream super-enhancer act as a strong anchor to stall ‘one-head’ of cohesin complex. The other cohesin head still slides along the variable region and thus reels in chromatin fibers (Figure 6B). By inverting or deleting single or arrays of CTCF sites in the variable-promoter or super-enhancer regions of the clustered *Pcdh* genes and assaying the resulting architectural and functional consequences, asymmetric topological effects of long-distance chromatin contacts and disruption of *Pcdh* gene expression can be detected (Lu et al., 2019; Jia et al., 2020).

Topological selections of enhancer–promoter pairing

Genome-editing technologies have facilitated our understanding of 3D chromatin architecture in specific enhancer–promoter contacts (reviewed in Schoenfelder and Fraser, 2019). CTCF/cohesin-mediated chromatin looping regulates the promoter selection of the *Pcdh* gene clusters and their neuron-specific expression patterns (Guo et al., 2012; Jiang et al., 2017; Allahyar et al., 2018; Wu et al., 2020). Specifically, the chromatin conformation capture 3C assay revealed that the enhancer element is spatially close to the promoter of the variable exon in the *Pcdh* gene cluster. In addition, the CTCF protein recognizes its conserved DNA-binding sites with directionality (Guo et al., 2015; Yin et al., 2017; Xu et al., 2018). Finally, single CTCF sites function as traditional insulators to ensure proper activation of target promoters by cognate enhancers; while tandem CTCF sites function as topological insulators to balance spatial chromatin contacts and to allocate enhancer resources for promoter choice (Zhai et al., 2016; Jia et al., 2020; Wu et al., 2020).

Synthetic single-chromosome yeast

Double cutting by Cas9 guided by two sgRNAs, each targeting to a site close to the telomeres of two separate yeast chromosomes, leads to the fusion of the two chromosomes (Shao et al., 2018). Remarkably, a functional single-chromosome

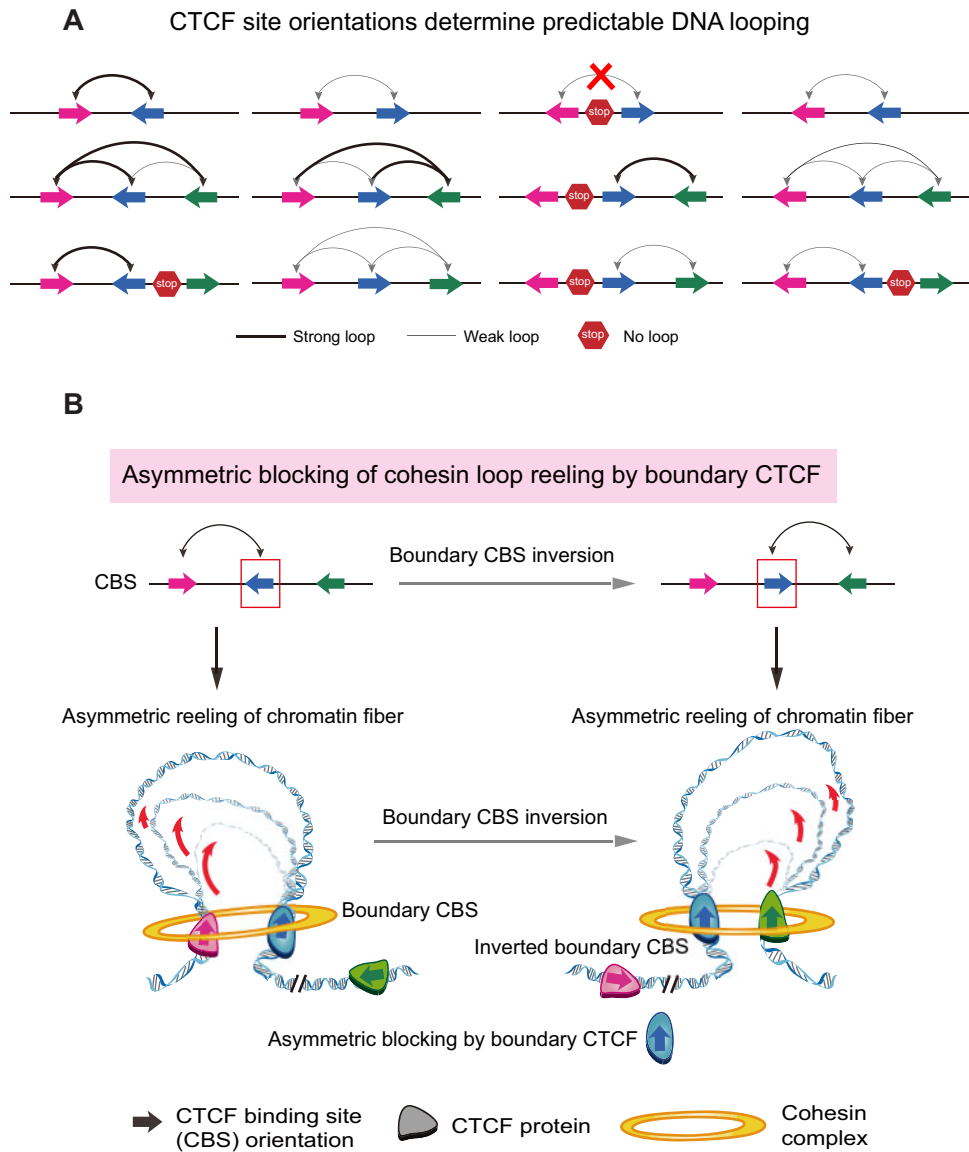


Figure 6 Predictable 3D genome engineering. **(A)** CTCF coding from 1D genomic sequences to 3D genome organization. The topology and strength of chromatin loops can be predicted based on the locations and relative orientations of CTCF sites. **(B)** Schematic of asymmetric ‘loop extrusion’ model revealed by CRISPR inversion of boundary CTCF sites. Genetic manipulation of CTCF sites demonstrates asymmetric blocking of cohesin loop extrusion by directional CTCF binding to oriented CBS elements. Chromatin fibers are compacted by active cohesin ‘loop extrusion’ with ‘two heads’. Cohesin complex reels in chromatin fibers until anchored by oriented CTCF sites. If ‘one head’ of cohesin is anchored by CTCF sites, cohesin can continue to reel in chromatin fibers through the ‘other head’, resulting in so-called asymmetric ‘loop extrusion’.

yeast was created by successive repeated fusions of all 16 yeast chromosomes into one giant chromosome by this CRISPR double cutting method (Shao et al., 2018). The two ends of the single linear chromosome could be further fused to generate a single circular chromosome (Shao et al., 2019). Apparently, both linear and circular single-chromosome yeasts have not been found in nature and thus are artificially synthesized yeast strains. This interesting observation indicates the power of targeted 3D genome engineering in synthetic biology by CRISPR with dual sgRNAs (Sadhu and Kruglyak, 2018).

3D genome synthetic biology

Programmed chromosomal fission and fusion by multiplexed CRISPR have generated synthetic genomes with nucleotide precision in bacteria (Wang et al., 2019a). In prokaryotic *Escherichia coli*, artificial chromosomes in single cells can be fused into a single genome with precise translocation and scarless inversion (Wang et al., 2019a). In eukaryotic yeast, Hi-C experiments revealed that the large-scale 3D organization of the synthetic genome is unaffected by the removal of numerous repeated sequences (Mercy et al., 2017). Interestingly, Hi-C

experiments demonstrated that the single linear-chromosome and circular-chromosome yeasts have similar globular 3D genome conformation (Shao et al., 2019). These studies suggest that global 3D genome structures have significant plasticity and can tolerate local genetic perturbations.

Perspective

We have sampled flavored highlights of some recent advances of genetic engineering of 3D genomes by CRISPR/Cas9 systems with various precise chromosomal rearrangements. Significant progress has been made recently in understanding the cleavage mechanisms of the CRISPR/Cas9 genome-editing system (Chen and Doudna, 2017). In addition, rapid technological advances in predictable DSB repair outcomes of precise CRISPR DNA fragment editing may accelerate its applications in agriculture and biomedicine (Tang and Fu, 2018). Furthermore, recent multiplexing CRISPR epigenetic technologies inform and promise cross-disciplinary revolutions (McCarty et al., 2020). Finally, CRISPR off-targets remain a big challenge but detecting methods are improving rapidly (Wienert et al., 2019).

Genetic engineering of 3D genomes and predictable chromosomal rearrangements by DNA fragment editing require interdisciplinary research. Obviously, fully predictable 3D genome engineering has not been achieved despite rapid progress in precise CRISPR DNA fragment editing in the last few years. Because very little is known in this area, it is a typical genre of desert-wandering night science that is full of darkness but also may stumble into a gold mine if lucky. 3D genomics integrates live biology with physical geometry. Renaissance of understanding and designing 3D genomes in the future may turn this night science into hypothesis-driven day science. Understanding the mechanisms of 3D genome folding will facilitate future precise and predictable CRISPR DNA fragment editing.

Acknowledgements

We apologize to our colleagues whose important contributions could not be cited due to space limitations.

Funding

This work was supported by grants from the National Natural Science Foundation of China (31630039 and 32000425), the Ministry of Science and Technology of China (2017YFA0504203 and 2018YFC1004504), and the Science and Technology Commission of Shanghai Municipality (19JC1412500).

Conflict of interest: none declared.

References

Ablain, J., Durand, E.M., Yang, S., et al. (2015). A CRISPR/Cas9 vector system for tissue-specific gene disruption in zebrafish. *Dev. Cell* 32, 756–764.

Acevedo-Arozena, A., Wells, S., Potter, P., et al. (2008). ENU mutagenesis, a way forward to understand gene function. *Annu. Rev. Genomics Hum. Genet.* 9, 49–69.

Alanis-Lobato, G., Zohren, J., McCarthy, A., et al. (2020). Frequent loss-of-heterozygosity in CRISPR–Cas9-edited early human embryos. *bioRxiv*, <http://doi.org/10.1101/2020.06.05.135913>

Ali, Z., Abulfaraj, A., Idris, A., et al. (2015). CRISPR/Cas9-mediated viral interference in plants. *Genome Biol.* 16, 238.

Allahyar, A., Vermeulen, C., Bouwman, B.A.M., et al. (2018). Enhancer hubs and loop collisions identified from single-allele topologies. *Nat. Genet.* 50, 1151–1160.

Allen, F., Crepaldi, L., Alsinet, C., et al. (2018). Predicting the mutations generated by repair of Cas9-induced double-strand breaks. *Nat. Biotechnol.* 37, 64–72.

Anzalone, A.V., Koblan, L.W., and Liu, D.R. (2020). Genome editing with CRISPR–Cas nucleases, base editors, transposases and prime editors. *Nat. Biotechnol.* 38, 824–844.

Anzalone, A.V., Randolph, P.B., Davis, J.R., et al. (2019). Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* 576, 149–157.

Arbab, M., Shen, M.W., Mok, B., et al. (2020). Determinants of base editing outcomes from target library analysis and machine learning. *Cell* 182, 463–480.

Aubrey, B.J., Kelly, G.L., Kueh, A.J., et al. (2015). An inducible lentiviral guide RNA platform enables the identification of tumor-essential genes and tumor-promoting mutations in vivo. *Cell Rep.* 10, 1422–1432.

Aymard, F., Aguirrebengoa, M., Guillou, E., et al. (2017). Genome-wide mapping of long-range contacts unveils clustering of DNA double-strand breaks at damaged active genes. *Nat. Struct. Mol. Biol.* 24, 353–361.

Banigan, E.J., and Mirny, L.A. (2020). Loop extrusion: theory meets single-molecule experiments. *Curr. Opin. Cell Biol.* 64, 124–138.

Bevan, M.W., Flavell, R.B., and Chilton, M.D. (1983). A chimaeric antibiotic resistance gene as a selectable marker for plant cell transformation. *Nature* 304, 184–187.

Bibikova, M., Beumer, K., Trautman, J.K., et al. (2003). Enhancing gene targeting with designed zinc finger nucleases. *Science* 300, 764.

Bickmore, W.A. (2019). Patterns in the genome. *Heredity* 123, 50–57.

Birling, M.C., Schaeffer, L., Andre, P., et al. (2017). Efficient and rapid generation of large genomic variants in rats and mice using CRISMERE. *Sci. Rep.* 7, 43331.

Blasco, Rafael B., Karaca, E., Ambrogio, C., et al. (2014). Simple and rapid in vivo generation of chromosomal rearrangements using CRISPR/Cas9 technology. *Cell Rep.* 9, 1219–1227.

Bonev, B., and Cavalli, G. (2016). Organization and function of the 3D genome. *Nat. Rev. Genet.* 17, 661–678.

Boroviak, K., Doe, B., Banerjee, R., et al. (2016). Chromosome engineering in zygotes with CRISPR/Cas9. *Genesis* 54, 78–85.

Brandl, C., Ortiz, O., Röttig, B., et al. (2015). Creation of targeted genomic deletions using TALEN or CRISPR/Cas nuclease pairs in one-cell mouse embryos. *FEBS Open Bio.* 5, 26–35.

Brinster, R.L., Chen, H.Y., Trumbauer, M., et al. (1981). Somatic expression of herpes thymidine kinase in mice following injection of a fusion gene into eggs. *Cell* 27, 223–231.

Brunet, E., and Jasin, M. (2018). Induction of chromosomal translocations with CRISPR–Cas9 and other nucleases: understanding the repair mechanisms that give rise to translocations. *Adv. Exp. Med. Biol.* 1044, 15–25.

Byrne, S.M., Ortiz, L., Mali, P., et al. (2015). Multi-kilobase homozygous targeted gene replacement in human induced pluripotent stem cells. *Nucleic Acids Res.* 43, e21.

Cai, Y., Chen, L., Liu, X., et al. (2015). CRISPR/Cas9-mediated genome editing in soybean hairy roots. *PLoS One* 10, e0136064.

Canver, M.C., Bauer, D.E., Dass, A., et al. (2014). Characterization of genomic deletion efficiency mediated by clustered regularly interspaced palindromic repeats (CRISPR)/Cas9 nuclease system in mammalian cells. *J. Biol. Chem.* 289, 21312–21324.

Capecchi, M.R. (2005). Gene targeting in mice: functional analysis of the mammalian genome for the twenty-first century. *Nat. Rev. Genet.* 6, 507–512.

- Caridi, C.P., D'Agostino, C., Ryu, T., et al. (2018). Nuclear F-actin and myosins drive relocalization of heterochromatic breaks. *Nature* 559, 54–60.
- Castle, W.E., and Little, C.C. (1909). The peculiar inheritance of pink eyes among colored mice. *Science* 30, 313–314.
- Ceccaldi, R., Rondinelli, B., and D'Andrea, A.D. (2016a). Repair pathway choices and consequences at the double-strand break. *Trends Cell Biol.* 26, 52–64.
- Ceccaldi, R., Sarangi, P., and D'Andrea, A.D. (2016b). The Fanconi anaemia pathway: new players and new functions. *Nat. Rev. Mol. Cell Biol.* 17, 337–349.
- Cech, T.R., and Steitz, J.A. (2014). The noncoding RNA revolution-trashing old rules to forge new ones. *Cell* 157, 77–94.
- Chakrabarti, A.M., Henser-Brownhill, T., Monserrat, J., et al. (2018). Target-specific precision of CRISPR-mediated genome editing. *Mol. Cell* 73, 699–713.
- Chang, H.H.Y., Pannunzio, N.R., Adachi, N., et al. (2017). Non-homologous DNA end joining and alternative pathways to double-strand break repair. *Nat. Rev. Mol. Cell Biol.* 18, 495–506.
- Chen, J.S., and Doudna, J.A. (2017). The chemistry of Cas9 and its CRISPR colleagues. *Nat. Rev. Chem.* 1, 0078.
- Chen, W., McKenna, A., Schreiber, J., et al. (2019). Massively parallel profiling and predictive modeling of the outcomes of CRISPR/Cas9-mediated double-strand break repair. *Nucleic Acids Res.* 47, 7989–8003.
- Chen, X., Xu, F., Zhu, C., et al. (2014). Dual sgRNA-directed gene knockout using CRISPR/Cas9 technology in *Caenorhabditis elegans*. *Caenorhabditis elegans*. *Sci. Rep.* 4, 7581.
- Cheong, T.C., Blasco, R.B., and Chiarle, R. (2018). The CRISPR/Cas9 system as a tool to engineer chromosomal translocation in vivo. *Adv. Exp. Med. Biol.* 1044, 39–48.
- Cho, S.W., Kim, S., Kim, J.M., et al. (2013). Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat. Biotechnol.* 31, 230–232.
- Cho, S.W., Kim, S., Kim, Y., et al. (2014). Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome Res.* 24, 132–141.
- Choi, P.S., and Meyerson, M. (2014). Targeted genomic rearrangements using CRISPR/Cas technology. *Nat. Commun.* 5, 3728.
- Cleves, P.A., Strader, M.E., Bay, L.K., et al. (2018). CRISPR/Cas9-mediated genome editing in a reef-building coral. *Proc. Natl Acad. Sci. USA* 115, 5235–5240.
- Cong, L., Ran, F.A., Cox, D., et al. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science* 339, 819–823.
- Costantini, F., and Lacy, E. (1981). Introduction of a rabbit β -globin gene into the mouse germ line. *Nature* 294, 92–94.
- Cradick, T.J., Fine, E.J., Antico, C.J., et al. (2013). CRISPR/Cas9 systems targeting β -globin and CCR5 genes have substantial off-target activity. *Nucleic Acids Res.* 41, 9584–9592.
- Daer, R.M., Cutts, J.P., Brafman, D.A., et al. (2017). The impact of chromatin dynamics on Cas9-mediated genome editing in human cells. *ACS Synth. Biol.* 6, 428–438.
- Davidson, I.F., Bauer, B., Goetz, D., et al. (2019). DNA loop extrusion by human cohesin. *Science* 366, 1338–1345.
- de Wit, E., Vos, E.S., Holwerda, S.J., et al. (2015). CTCF binding polarity determines chromatin looping. *Mol. Cell* 60, 676–684.
- Dekker, J., and Mirny, L. (2016). The 3D genome as moderator of chromosomal communication. *Cell* 164, 1110–1121.
- Dekker, J., Rippe, K., Dekker, M., et al. (2002). Capturing chromosome conformation. *Science* 295, 1306–1311.
- Deng, W., Henriët, S., and Chourrout, D. (2018). Prevalence of mutation-prone microhomology-mediated end joining in a chordate lacking the c-NHEJ DNA repair pathway. *Curr. Biol.* 28, 3337–3341.e4.
- Denker, A., and de Laat, W. (2016). The second decade of 3C technologies: detailed insights into nuclear organization. *Genes Dev.* 30, 1357–1382.
- Deriano, L., and Roth, D.B. (2013). Modernizing the nonhomologous end-joining repertoire: alternative and classical NHEJ share the stage. *Annu. Rev. Genet.* 47, 433–455.
- DiCarlo, J.E., Norville, J.E., Mali, P., et al. (2013). Genome engineering in *Saccharomyces cerevisiae* using CRISPR–Cas systems. *Nucleic Acids Res.* 41, 4336–4343.
- Dixon, J.R., Gorkin, D.U., and Ren, B. (2016). Chromatin domains: the unit of chromosome organization. *Mol. Cell* 62, 668–680.
- Doudna, J.A., and Charpentier, E. (2014). Genome editing. The new frontier of genome engineering with CRISPR–Cas9. *Science* 346, 1258096.
- Essletzbichler, P., Konopka, T., Santoro, F., et al. (2014). Megabase-scale deletion using CRISPR/Cas9 to generate a fully haploid human cell line. *Genome Res.* 24, 2059–2065.
- Evans, M.J., and Kaufman, M.H. (1981). Establishment in culture of pluripotential cells from mouse embryos. *Nature* 292, 154–156.
- Folger, K.R., Wong, E.A., Wahl, G., et al. (1982). Patterns of integration of DNA microinjected into cultured mammalian cells: evidence for homologous recombination between injected plasmid DNA molecules. *Mol. Cell Biol.* 2, 1372–1387.
- Fudenberg, G., Imakaev, M., Lu, C., et al. (2016). Formation of chromosomal domains by loop extrusion. *Cell Rep.* 15, 2038–2049.
- Fujii, W., Kawasaki, K., Sugiura, K., et al. (2013). Efficient generation of large-scale genome-modified mice using gRNA and CAS9 endonuclease. *Nucleic Acids Res.* 41, e187.
- Fujii, W., Onuma, A., Sugiura, K., et al. (2014). One-step generation of phenotype-expressing triple-knockout mice with heritable mutated alleles by the CRISPR/Cas9 system. *J. Reprod. Dev.* 60, 324–327.
- Gandhi, S., Piacentino, M.L., Vieceli, F.M., et al. (2017). Optimization of CRISPR/Cas9 genome editing for loss-of-function in the early chick embryo. *Dev. Biol.* 432, 86–97.
- Gao, J., Wang, G., Ma, S., et al. (2015). CRISPR/Cas9-mediated targeted mutagenesis in *Nicotiana tabacum*. *Plant Mol. Biol.* 87, 99–110.
- Gao, S., Tong, Y., Wen, Z., et al. (2016). Multiplex gene editing of the *Yarrowia lipolytica* genome using the CRISPR–Cas9 system. *J. Ind. Microbiol. Biotechnol.* 43, 1085–1093.
- Gao, W., Long, L., Tian, X., et al. (2017). Genome editing in cotton with the CRISPR/Cas9 system. *Front. Plant Sci.* 8, 1364.
- Garneau, J.E., Dupuis, M.E., Villion, M., et al. (2010). The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* 468, 67–71.
- Gasiunas, G., Barrangou, R., Horvath, P., et al. (2012). Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc. Natl Acad. Sci. USA* 109, E2579–E2586.
- Gassler, J., Brandao, H.B., Imakaev, M., et al. (2017). A mechanism of cohesin-dependent loop extrusion organizes zygotic genome architecture. *EMBO J.* 36, 3600–3618.
- Gaudelli, N.M., Komor, A.C., Rees, H.A., et al. (2017). Programmable base editing of AT to GC in genomic DNA without DNA cleavage. *Nature* 551, 464–471.
- Gelot, C., Guirouilh-Barbat, J., Le Guen, T., et al. (2016). The cohesin complex prevents the end joining of distant DNA double-strand ends. *Mol. Cell* 61, 15–26.
- Gisler, S., Goncalves, J.P., Akhtar, W., et al. (2019). Multiplexed Cas9 targeting reveals genomic location effects and gRNA-based staggered breaks influencing mutation efficiency. *Nat. Commun.* 10, 1598.
- Golic, K.G., and Golic, M.M. (1996). Engineering the *Drosophila* genome: chromosome rearrangements by design. *Genetics* 144, 1693–1711.
- Gordon, J.W., Scangos, G.A., Plotkin, D.J., et al. (1980). Genetic transformation of mouse embryos by microinjection of purified DNA. *Proc. Natl Acad. Sci. USA* 77, 7380–7384.
- Gratz, S.J., Cummings, A.M., Nguyen, J.N., et al. (2013). Genome engineering of *Drosophila* with the CRISPR RNA-guided Cas9 nuclease. *Genetics* 194, 1029–1035.

- Gundry, M.C., Brunetti, L., Lin, A., et al. (2016). Highly efficient genome editing of murine and human hematopoietic progenitor cells by CRISPR/Cas9. *Cell Rep.* 17, 1453–1461.
- Guo, T., Feng, Y.L., Xiao, J.J., et al. (2018). Harnessing accurate non-homologous end joining for efficient precise deletion in CRISPR/Cas9-mediated genome editing. *Genome Biol.* 19, 170.
- Guo, Y., Monahan, K., Wu, H., et al. (2012). CTCF/cohesin-mediated DNA looping is required for protocadherin alpha promoter choice. *Proc. Natl Acad. Sci. USA* 109, 21081–21086.
- Guo, Y., Xu, Q., Canzio, D., et al. (2015). CRISPR inversion of CTCF sites alters genome topology and enhancer/promoter function. *Cell* 162, 900–910.
- Gupta, A., Hall, V.L., Kok, F.O., et al. (2013). Targeted chromosomal deletions and inversions in zebrafish. *Genome Res.* 23, 1008–1017.
- Haarhuis, J.H.I., van der Weide, R.H., Blomen, V.A., et al. (2017). The cohesin release factor WAPL restricts chromatin loop extension. *Cell* 169, 693–707.e14.
- Han, D., Chen, Q., Shi, J., et al. (2017). CTCF participates in DNA damage response via poly(ADP-ribosylation). *Sci. Rep.* 7, 43530.
- Hansen, A.S., Cattoglio, C., Darzacq, X., et al. (2018). Recent evidence that TADs and chromatin loops are dynamic structures. *Nucleus* 9, 20–32.
- He, Z., Proudfoot, C., Mileham, A.J., et al. (2015). Highly efficient targeted chromosome deletions using CRISPR/Cas9. *Biotechnol. Bioeng.* 112, 1060–1064.
- Herault, Y., Rassoulzadegan, M., Cuzin, F., et al. (1998). Engineering chromosomes in mice through targeted meiotic recombination (TAMERE). *Nat. Genet.* 20, 381–384.
- Hilmi, K., Jangal, M., Marques, M., et al. (2017). CTCF facilitates DNA double-strand break repair by enhancing homologous recombination repair. *Sci. Adv.* 3, e1601898.
- Horii, T., Arai, Y., Yamazaki, M., et al. (2014). Validation of microinjection methods for generating knockout mice by CRISPR/Cas-mediated genome engineering. *Sci. Rep.* 4, 4513.
- Hou, P., Chen, S., Wang, S., et al. (2015). Genome editing of CXCR4 by CRISPR/cas9 confers cells resistant to HIV-1 infection. *Sci. Rep.* 5, 15577.
- Huang, H., and Wu, Q. (2016). CRISPR double cutting through the labyrinthine architecture of 3D genomes. *J. Genet. Genomics* 43, 273–288.
- Hwang, S.Y., Kang, M.A., Baik, C.J., et al. (2019). CTCF cooperates with CtIP to drive homologous recombination repair of double-strand breaks. *Nucleic Acids Res.* 47, 9160–9179.
- Hwang, W.Y., Fu, Y.F., Reyon, D., et al. (2013). Heritable and precise zebrafish genome editing using a CRISPR–Cas system. *PLoS One* 8, e68708.
- Iacovoni, J.S., Caron, P., Lassadi, I., et al. (2010). High-resolution profiling of γ H2AX around DNA double strand breaks in the mammalian genome. *EMBO J.* 29, 1446–1457.
- Iaffaldano, B., Zhang, Y., and Cornish, K. (2016). CRISPR/Cas9 genome editing of rubber producing dandelion *Taraxacum kok-saghyz* using *Agrobacterium rhizogenes* without selection. *Ind. Crops Prod.* 89, 356–362.
- Incontro, S., Asensio, C.S., Edwards, R.H., et al. (2014). Efficient, complete deletion of synaptic proteins using CRISPR. *Neuron* 83, 1051–1057.
- Iyer, S., Suresh, S., Guo, D., et al. (2019). Precise therapeutic gene correction by a simple nuclease-induced double-stranded break. *Nature* 568, 561–565.
- Jacobs, P.A., Browne, C., Gregson, N., et al. (1992). Estimates of the frequency of chromosome abnormalities detectable in unselected newborns using moderate levels of banding. *J. Med. Genet.* 29, 103–108.
- Jaenisch, R., and Mintz, B. (1974). Simian virus 40 DNA sequences in DNA of healthy adult mice derived from preimplantation blastocysts injected with viral DNA. *Proc. Natl Acad. Sci. USA* 71, 1250–1254.
- Jarrett, K.E., Lee, C.M., Yeh, Y.H., et al. (2017). Somatic genome editing with CRISPR/Cas9 generates and corrects a metabolic disease. *Sci. Rep.* 7, 44624.
- Jasin, M., and Berg, P. (1988). Homologous integration in mammalian cells without target gene selection. *Genes Dev.* 2, 1353–1363.
- Jasin, M., and Rothstein, R. (2013). Repair of strand breaks by homologous recombination. *Cold Spring Harb. Pers. Biol.* 5, a012740.
- Jia, Z., Li, J., Ge, X., et al. (2020). Tandem CTCF sites function as insulators to balance spatial chromatin contacts and topological enhancer-promoter selection. *Genome Biol.* 21, 75.
- Jiang, W., and Marraffini, L.A. (2015). CRISPR–Cas: new tools for genetic manipulations from bacterial immunity systems. *Annu. Rev. Microbiol.* 69, 209–228.
- Jiang, Y., Loh, Y.E., Rajarajan, P., et al. (2017). The methyltransferase SETDB1 regulates a large neuron-specific topological chromatin domain. *Nat. Genet.* 49, 1239–1250.
- Jiao, R., and Gao, C. (2016). The CRISPR/Cas9 genome editing revolution. *J. Genet. Genomics* 43, 227–228.
- Jinek, M., Chylinski, K., Fonfara, I., et al. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816–821.
- Jinek, M., East, A., Cheng, A., et al. (2013). RNA-programmed genome editing in human cells. *eLife* 2, e00471.
- Jinek, M., Jiang, F., Taylor, D.W., et al. (2014). Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. *Science* 343, 1247997.
- Kalhor, R., Kalhor, K., Mejia, L., et al. (2018). Developmental barcoding of whole mouse via homing CRISPR. *Science* 361, eaat9804.
- Kile, B.T., and Hilton, D.J. (2005). The art and design of genetic screens: mouse. *Nat. Rev. Genet.* 6, 557–567.
- Kim, D., Kim, J., Hur, J.K., et al. (2016). Genome-wide analysis reveals specificities of Cpf1 endonucleases in human cells. *Nat. Biotechnol.* 34, 863–868.
- Kim, H., Li, F., He, Q., et al. (2017). Systematic analysis of human telomeric dysfunction using inducible telosome/shelterin CRISPR/Cas9 knockout cells. *Cell Discov.* 3, 17034.
- Kim, J.M., Kim, D., Kim, S., et al. (2014). Genotyping with CRISPR–Cas-derived RNA-guided endonucleases. *Nat. Commun.* 5, 3157.
- Kim, Y., Shi, Z., Zhang, H., et al. (2019). Human cohesin compacts DNA by loop extrusion. *Science* 366, 1345–1349.
- Komor, A.C., Kim, Y.B., Packer, M.S., et al. (2016). Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* 533, 420–424.
- Korablev, A.N., Serova, I.A., and Serov, O.L. (2017). Generation of megabase-scale deletions, inversions and duplications involving the *Contactin-6* gene in mice by CRISPR/Cas9 technology. *BMC Genet.* 18, 112.
- Kosicki, M., Allen, F., and Bradley, A. (2020). Cas9-induced large deletions and small indels are controlled in a convergent fashion. *bioRxiv*, <http://doi.org/10.1101/2020.08.05.216739>
- Kosicki, M., Tomberg, K., and Bradley, A. (2018). Repair of double-strand breaks induced by CRISPR–Cas9 leads to large deletions and complex rearrangements. *Nat. Biotechnol.* 36, 765–771.
- Kraft, K., Geuer, S., Will, A.J., et al. (2015). Deletions, inversions, duplications: engineering of structural variants using CRISPR/Cas in mice. *Cell Rep.* 10, 833–839.
- Lang, F., Li, X., Zheng, W., et al. (2017). CTCF prevents genomic instability by promoting homologous recombination-directed DNA double-strand break repair. *Proc. Natl Acad. Sci. USA* 114, 10912–10917.
- Lawrenson, T., Shorinola, O., Stacey, N., et al. (2015). Induction of targeted, heritable mutations in barley and *Brassica oleracea* using RNA-guided Cas9 nuclease. *Genome Biol.* 16, 258.
- Lee, C.S., Lee, K., Legube, G., et al. (2014a). Dynamics of yeast histone H2A and H2B phosphorylation in response to a double-strand break. *Nat. Struct. Mol. Biol.* 21, 103–109.
- Lee, H.J., Kim, E., and Kim, J.S. (2010). Targeted chromosomal deletions in human cells using zinc finger nucleases. *Genome Res.* 20, 81–89.
- Lee, J.S., Kwak, S.J., Kim, J., et al. (2014b). RNA-guided genome editing in *Drosophila* with the purified Cas9 protein. *G3* 4, 1291–1295.

- Leenay, R.T., Aghazadeh, A., Hiatt, J., et al. (2019). Large dataset enables prediction of repair after CRISPR-Cas9 editing in primary T cells. *Nat. Biotechnol.* **37**, 1034–1037.
- Lemmon, Z.H., Reem, N.T., Dalrymple, J., et al. (2018). Rapid improvement of domestication traits in an orphan crop by genome editing. *Nat. Plants* **4**, 766–770.
- Lemos, B.R., Kaplan, A.C., Bae, J.E., et al. (2018). CRISPR/Cas9 cleavages in budding yeast reveal templated insertions and strand-specific insertion/deletion profiles. *Proc. Natl Acad. Sci. USA* **115**, E2040–E2047.
- Lewis, E.B. (1950). The phenomenon of position effect. *Adv. Genet.* **3**, 73–115.
- Li, J., Shou, J., Guo, Y., et al. (2015a). Efficient inversions and duplications of mammalian regulatory DNA elements and gene clusters by CRISPR/Cas9. *J. Mol. Cell Biol.* **7**, 284–298.
- Li, J., Shou, J., and Wu, Q. (2015b). DNA fragment editing of genomes by CRISPR/Cas9. *Yi Chuan* **37**, 992–1002.
- Li, J.F., Norville, J.E., Aach, J., et al. (2013a). Multiplex and homologous recombination-mediated genome editing in *Arabidopsis* and *Nicotiana benthamiana* using guide RNA and Cas9. *Nat. Biotechnol.* **31**, 688–691.
- Li, M., Gan, J., Sun, Y., et al. (2020a). Architectural proteins for the formation and maintenance of the 3D genome. *Sci. China Life Sci.* **63**, 795–810.
- Li, W., Teng, F., Li, T.D., et al. (2013b). Simultaneous generation and germline transmission of multiple gene mutations in rat using CRISPR-Cas systems. *Nat. Biotechnol.* **31**, 684–686.
- Li, Y., Haarhuis, J.H.I., Cacciatore, Á.S., et al. (2020b). The structural basis for cohesin-CTCF-anchored loops. *Nature* **578**, 472–476.
- Li, Y., Park, A.I., Mou, H., et al. (2015c). A versatile reporter system for CRISPR-mediated chromosomal rearrangements. *Genome Biol.* **16**, 111.
- Liang, D., Gutierrez, N.M., Chen, T., et al. (2020). Frequent gene conversion in human embryos induced by double strand breaks. *bioRxiv*, <http://doi.org/10.1101/2020.06.19.162214>
- Liang, Z., Chen, K., Li, T., et al. (2017). Efficient DNA-free genome editing of bread wheat using CRISPR/Cas9 ribonucleoprotein complexes. *Nat. Commun.* **8**, 14261.
- Liao, H.K., Gu, Y., Diaz, A., et al. (2015a). Use of the CRISPR/Cas9 system as an intracellular defense against HIV-1 infection in human cells. *Nat. Commun.* **6**, 6413.
- Liao, S., Tamaro, M., and Yan, H. (2015b). Enriching CRISPR-Cas9 targeted cells by co-targeting the HPRT gene. *Nucleic Acids Res.* **43**, e134.
- Lieber, M.R. (2016). Mechanisms of human lymphoid chromosomal translocations. *Nat. Rev. Cancer* **16**, 387–398.
- Liu, P.F., and Wu, Q. (2020). Probing 3D genome by CRISPR/Cas9. *Yi Chuan* **42**, 18–31.
- Liu, X., Zhang, Y., Chen, Y., et al. (2017a). In situ capture of chromatin interactions by biotinylated dCas9. *Cell* **170**, 1028–1043.e19.
- Liu, Z., Chen, S., Jin, X., et al. (2017b). Genome editing of the HIV co-receptors CCR5 and CXCR4 by CRISPR-Cas9 protects CD4⁺ T cells from HIV-1 infection. *Cell Biosci.* **7**, 47.
- Liu, Z., Hui, Y., Shi, L., et al. (2016). Efficient CRISPR/Cas9-mediated versatile, predictable, and donor-free gene knockout in human pluripotent stem cells. *Stem Cell Rep.* **7**, 496–507.
- Long, C. (2019). God does not play dice, and neither does CRISPR/Cas9. *Natl Sci. Rev.* **6**, 393.
- Lowder, L.G., Zhang, D., Baltus, N.J., et al. (2015). A CRISPR/Cas9 toolbox for multiplexed plant genome editing and transcriptional regulation. *Plant Physiol.* **169**, 971–985.
- Lu, Y., Shou, J., Jia, Z., et al. (2019). Genetic evidence for asymmetric blocking of higher-order chromatin structure by CTCF/cohesin. *Protein Cell* **10**, 914–920.
- Ma, S., Chang, J., Wang, X., et al. (2014). CRISPR/Cas9 mediated multiplex genome editing and heritable mutagenesis of *BmKu70* in *Bombyx mori*. *Sci. Rep.* **4**, 4489.
- Maddalo, D., Machado, E., Concepcion, C.P., et al. (2014). In vivo engineering of oncogenic chromosomal rearrangements with the CRISPR/Cas9 system. *Nature* **516**, 423–427.
- Mali, P., Yang, L., Esvelt, K.M., et al. (2013). RNA-guided human genome engineering via Cas9. *Science* **339**, 823–826.
- Mansour, S.L., Thomas, K.R., and Capecchi, M.R. (1988). Disruption of the proto-oncogene int-2 in mouse embryo-derived stem cells: a general strategy for targeting mutations to non-selectable genes. *Nature* **336**, 348–352.
- Marnef, A., and Legube, G. (2017). Organizing DNA repair in the nucleus: DSBs hit the road. *Curr. Opin. Cell Biol.* **46**, 1–8.
- Martin, G.R. (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc. Natl Acad. Sci. USA* **78**, 7634–7638.
- Matano, M., Date, S., Shimokawa, M., et al. (2015). Modeling colorectal cancer using CRISPR-Cas9-mediated engineering of human intestinal organoids. *Nat. Med.* **21**, 256–262.
- McCarty, N.S., Graham, A.E., Studená, L., et al. (2020). Multiplexed CRISPR technologies for gene editing and transcriptional regulation. *Nat. Commun.* **11**, 1281.
- McClintock, B. (1950). The origin and behavior of mutable loci in maize. *Proc. Natl Acad. Sci. USA* **36**, 344–355.
- McVey, M., and Lee, S.E. (2008). MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings. *Trends Genet.* **24**, 529–538.
- Mercy, G., Mozziconacci, J., Scolari, V.F., et al. (2017). 3D organization of synthetic and scrambled chromosomes. *Science* **355**, eaaf4597.
- Merkenschlager, M., and Nora, E.P. (2016). CTCF and cohesin in genome folding and transcriptional gene regulation. *Annu. Rev. Genomics Hum. Genet.* **17**, 17–43.
- Mianne, J., Codner, G.F., Caulder, A., et al. (2017). Analysing the outcome of CRISPR-aided genome editing in embryos: screening, genotyping and quality control. *Methods* **121–122**, 68–76.
- Miao, J., Guo, D., Zhang, J., et al. (2013). Targeted mutagenesis in rice using CRISPR-Cas system. *Cell Res.* **23**, 1233–1236.
- Miller, J.C., Tan, S., Qiao, G., et al. (2011). A TALE nuclease architecture for efficient genome editing. *Nat. Biotechnol.* **29**, 143–148.
- Mills, A.A., and Bradley, A. (2001). From mouse to man: generating megabase chromosome rearrangements. *Trends Genet.* **17**, 331–339.
- Mitelman, F., Mertens, F., and Johansson, B. (1997). A breakpoint map of recurrent chromosomal rearrangements in human neoplasia. *Nat. Genet.* **15**, 417–474.
- Molla, K.A., and Yang, Y. (2020). Predicting CRISPR/Cas9-induced mutations for precise genome editing. *Trends Biotechnol.* **38**, 136–141.
- Morineau, C., Bellec, Y., Tellier, F., et al. (2017). Selective gene dosage by CRISPR-Cas9 genome editing in hexaploid *Camelina sativa*. *Plant Biotechnol. J.* **15**, 729–739.
- Muller, H.J. (1930). Types of visible variations induced by X-rays in *Drosophila*. *J. Genet.* **22**, 299–334.
- Nakade, S., Mochida, K., Kunii, A., et al. (2018). Biased genome editing using the local accumulation of DSB repair molecules system. *Nat. Commun.* **9**, 3270.
- Neph, S., Vierstra, J., Stergachis, A.B., et al. (2012). An expansive human regulatory lexicon encoded in transcription factor footprints. *Nature* **489**, 83–90.
- Ni, W., Qiao, J., Hu, S.W., et al. (2014). Efficient gene knockout in goats using CRISPR/Cas9 system. *PLoS One* **9**, e106718.
- Nichols, M.H., and Corces, V.G. (2015). A CTCF code for 3D genome architecture. *Cell* **162**, 703–705.
- Nishimasu, H., Ran, F.A., Hsu, P.D., et al. (2014). Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell* **156**, 935–949.
- Nishitani, C., Hirai, N., Komori, S., et al. (2016). Efficient genome editing in apple using a CRISPR/Cas9 system. *Sci. Rep.* **6**, 31481.
- Niu, Y., Shen, B., Cui, Y., et al. (2014). Generation of gene-modified cynomolgus monkey via Cas9/RNA-mediated gene targeting in one-cell embryos. *Cell* **156**, 836–843.
- Nymark, M., Sharma, A.K., Sparstad, T., et al. (2016). A CRISPR/Cas9 system adapted for gene editing in marine algae. *Sci. Rep.* **6**, 24951.

- Nyquist, M.D., Li, Y., Hwang, T.H., et al. (2013). TALEN-engineered AR gene rearrangements reveal endocrine uncoupling of androgen receptor in prostate cancer. *Proc. Natl Acad. Sci. USA* *110*, 17492–17497.
- Orr-Weaver, T.L., Szostak, J.W., and Rothstein, R.J. (1981). Yeast transformation: a model system for the study of recombination. *Proc. Natl Acad. Sci. USA* *78*, 6354–6358.
- Palermo, G., Miao, Y., Walker, R.C., et al. (2016). Striking plasticity of CRISPR–Cas9 and key role of non-target DNA, as revealed by molecular simulations. *ACS Cent. Sci.* *2*, 756–763.
- Parelho, V., Hadjur, S., Spivakov, M., et al. (2008). Cohesins functionally associate with CTCF on mammalian chromosome arms. *Cell* *132*, 422–433.
- Park, C.-Y., Kim, D.H., Son, J.S., et al. (2015). Functional correction of large factor viii gene chromosomal inversions in hemophilia a patient-derived iPSCs using CRISPR–Cas9. *Cell Stem Cell* *17*, 213–220.
- Pauwels, L., De Clercq, R., Goossens, J., et al. (2018). A dual sgRNA approach for functional genomics in *Arabidopsis thaliana*. *G3* *8*, 2603–2615.
- Platt, Randall J., Chen, S., Zhou, Y., et al. (2014). CRISPR–Cas9 knockin mice for genome editing and cancer modeling. *Cell* *159*, 440–455.
- Qian, Y., Zhang, L., Cai, M., et al. (2019). The prostate cancer risk variant rs55958994 regulates multiple gene expression through extreme long-range chromatin interaction to control tumor progression. *Sci. Adv.* *5*, eaaw6710.
- Rabbitts, T.H. (1994). Chromosomal translocations in human cancer. *Nature* *372*, 143–149.
- Ramakrishna, S., Kwaku Dad, A.B., Beloor, J., et al. (2014). Gene disruption by cell-penetrating peptide-mediated delivery of Cas9 protein and guide RNA. *Genome Res.* *24*, 1020–1027.
- Ramirez-Solis, R., Liu, P., and Bradley, A. (1995). Chromosome engineering in mice. *Nature* *378*, 720–724.
- Ran, F.A., Hsu, P.D., Lin, C.Y., et al. (2013). Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* *154*, 1380–1389.
- Rao, S.S., Huntley, M.H., Durand, N.C., et al. (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* *159*, 1665–1680.
- Rao, S.S.P., Huang, S.C., Glenn St Hilaire, B., et al. (2017). Cohesin loss eliminates all loop domains. *Cell* *171*, 305–320.
- Ren, C., Liu, X., Zhang, Z., et al. (2016). CRISPR/Cas9-mediated efficient targeted mutagenesis in Chardonnay (*Vitis vinifera* L.). *Sci. Rep.* *6*, 32289.
- Ren, J., Liu, X., Fang, C., et al. (2017). Multiplex genome editing to generate universal CAR T cells resistant to pd1 inhibition. *Clin. Cancer Res.* *23*, 2255–2266.
- Rong, Y.S., and Golic, K.G. (2000). Gene targeting by homologous recombination in *Drosophila*. *Science* *288*, 2013–2018.
- Rowley, M.J., and Corces, V.G. (2018). Organizational principles of 3D genome architecture. *Nat. Rev. Genet.* *19*, 789–800.
- Ryder, S.P. (2018). CRISPRbabies: notes on a scandal. *CRISPR J.* *1*, 355–357.
- Sadhu, M.J., and Kruglyak, L. (2018). How low can you go? *CRISPR J.* *1*, 312–313.
- Sakuma, T., Nishikawa, A., Kume, S., et al. (2014). Multiplex genome engineering in human cells using all-in-one CRISPR/Cas9 vector system. *Sci. Rep.* *4*, 5400.
- Sanborn, A.L., Rao, S.S., Huang, S.C., et al. (2015). Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proc. Natl Acad. Sci. USA* *112*, E6456–E6465.
- Scherer, S., and Davis, R.W. (1979). Replacement of chromosome segments with altered DNA sequences constructed in vitro. *Proc. Natl Acad. Sci. USA* *76*, 4951–4955.
- Schmidt, C., Pacher, M., and Puchta, H. (2019). Efficient induction of heritable inversions in plant genomes using the CRISPR/Cas system. *Plant J.* *98*, 577–589.
- Schmieder, V., Bydlinski, N., Strasser, R., et al. (2018). Enhanced genome editing tools for multi-gene deletion knock-out approaches using paired CRISPR sgRNAs in CHO cells. *Biotechnol. J.* *13*, e1700211.
- Schoenfelder, S., and Fraser, P. (2019). Long-range enhancer-promoter contacts in gene expression control. *Nat. Rev. Genet.* *20*, 437–455.
- Schwank, G., Koo, B.K., Sasselli, V., et al. (2013). Functional repair of CFTR by CRISPR/Cas9 in intestinal stem cell organoids of cystic fibrosis patients. *Cell Stem Cell* *13*, 653–658.
- Schwarzer, W., Abdennur, N., Goloborodko, A., et al. (2017). Two independent modes of chromatin organization revealed by cohesin removal. *Nature* *551*, 51–56.
- Seruggia, D., Fernández, A., Cantero, M., et al. (2015). Functional validation of mouse tyrosinase non-coding regulatory DNA elements by CRISPR–Cas9-mediated mutagenesis. *Nucleic Acids Res.* *43*, 4855–4867.
- Shaffer, L.G., and Lupski, J.R. (2000). Molecular mechanisms for constitutional chromosomal rearrangements in humans. *Annu. Rev. Genet.* *34*, 297–329.
- Shan, Q.W., Wang, Y.P., Li, J., et al. (2013). Targeted genome modification of crop plants using a CRISPR–Cas system. *Nat. Biotechnol.* *31*, 686–688.
- Shao, Y., Lu, N., Cai, C., et al. (2019). A single circular chromosome yeast. *Cell Res.* *29*, 87–89.
- Shao, Y., Lu, N., Wu, Z., et al. (2018). Creating a functional single-chromosome yeast. *Nature* *560*, 331–335.
- Shen, B., Zhang, J., Wu, H., et al. (2013). Generation of gene-modified mice via Cas9/RNA-mediated gene targeting. *Cell Res.* *23*, 720–723.
- Shen, B., Zhang, W., Zhang, J., et al. (2014). Efficient genome modification by CRISPR–Cas9 nickase with minimal off-target effects. *Nat. Methods* *11*, 399–402.
- Shen, M.W., Arbab, M., Hsu, J.Y., et al. (2018). Predictable and precise template-free CRISPR editing of pathogenic variants. *Nature* *563*, 646–651.
- Shi, X., Shou, J., Mehryar, M.M., et al. (2019). Cas9 has no exonuclease activity resulting in staggered cleavage with overhangs and predictable di- and tri-nucleotide CRISPR insertions without template donor. *Cell Discov.* *5*, 53.
- Shin, H.Y., Wang, C., Lee, H.K., et al. (2017). CRISPR/Cas9 targeting events cause complex deletions and insertions at 17 sites in the mouse genome. *Nat. Commun.* *8*, 15464.
- Shou, J., Li, J., Liu, Y., et al. (2018). Precise and predictable CRISPR chromosomal rearrangements reveal principles of Cas9-mediated nucleotide insertion. *Mol. Cell* *71*, 498–509.
- Sigal, Y.M., Zhou, R., and Zhuang, X. (2018). Visualizing and discovering cellular structures with super-resolution microscopy. *Science* *361*, 880–887.
- Skryabin, B.V., Kummerfeld, D.M., Gubar, L., et al. (2020). Pervasive head-to-tail insertions of DNA templates mask desired CRISPR–Cas9-mediated genome editing events. *Sci. Adv.* *6*, eaax2941.
- Smithies, O., Gregg, R.G., Boggs, S.S., et al. (1985). Insertion of DNA sequences into the human chromosomal β -globin locus by homologous recombination. *Nature* *317*, 230–234.
- So, C.C., and Martin, A. (2019). DSB structure impacts DNA recombination leading to class switching and chromosomal translocations in human B cells. *PLoS Genet.* *15*, e1008101.
- Song, Y., Yuan, L., Wang, Y., et al. (2016). Efficient dual sgRNA-directed large gene deletion in rabbit with CRISPR/Cas9 system. *Cell. Mol. Life Sci.* *73*, 2959–2968.
- Stephenson, A.A., Raper, A.T., and Suo, Z. (2018). Bidirectional degradation of DNA cleavage products catalyzed by CRISPR/Cas9. *J. Am. Chem. Soc.* *140*, 3743–3750.
- Su, S., Hu, B., Shao, J., et al. (2016). CRISPR–Cas9 mediated efficient PD-1 disruption on human primary T cells from cancer patients. *Sci. Rep.* *6*, 20070.
- Sun, X., Hu, Z., Chen, R., et al. (2015). Targeted mutagenesis in soybean using the CRISPR–Cas9 system. *Sci. Rep.* *5*, 10342.

- Suzuki, K., Tsunekawa, Y., Hernandez-Benitez, R., et al. (2016). In vivo genome editing via CRISPR/Cas9 mediated homology-independent targeted integration. *Nature* *540*, 144–149.
- Svitashev, S., Schwartz, C., Lenderts, B., et al. (2016). Genome editing in maize directed by CRISPR–Cas9 ribonucleoprotein complexes. *Nat. Commun.* *7*, 13274.
- Taheri-Ghahfarokhi, A., Taylor, B.J.M., Nitsch, R., et al. (2018). Decoding non-random mutational signatures at Cas9 targeted sites. *Nucleic Acids Res.* *46*, 8417–8434.
- Tan, L., Xing, D., Daley, N., et al. (2019). Three-dimensional genome structures of single sensory neurons in mouse visual and olfactory systems. *Nat. Struct. Mol. Biol.* *26*, 297–307.
- Tang, Y., and Fu, Y. (2018). Class 2 CRISPR/Cas: an expanding biotechnology toolbox for and beyond genome editing. *Cell Biosci.* *8*, 59.
- Tang, Z., Luo, O.J., Li, X., et al. (2015). CTCF-mediated human 3D genome architecture reveals chromatin topology for transcription. *Cell* *163*, 1611–1627.
- Thomas, K.R., Folger, K.R., and Capecchi, M.R. (1986). High frequency targeting of genes to specific sites in the mammalian genome. *Cell* *44*, 419–428.
- Tian, S., Jiang, L., Gao, Q., et al. (2017). Efficient CRISPR/Cas9-based gene knockout in watermelon. *Plant Cell Rep.* *36*, 399–406.
- Torres, R., Martin, M.C., Garcia, A., et al. (2014). Engineering human tumour-associated chromosomal translocations with the RNA-guided CRISPR–Cas9 system. *Nat. Commun.* *5*, 3964.
- van Diemen, F.R., Kruse, E.M., Hooykaas, M.J., et al. (2016). CRISPR/Cas9-mediated genome editing of herpesviruses limits productive and latent infections. *PLoS Pathog.* *12*, e1005701.
- van Overbeek, M., Capurso, D., Carter, M.M., et al. (2016). DNA repair profiling reveals nonrandom outcomes at Cas9-mediated breaks. *Mol. Cell* *63*, 633–646.
- Vanoli, F., and Jasin, M. (2017). Generation of chromosomal translocations that lead to conditional fusion protein expression using CRISPR–Cas9 and homology-directed repair. *Methods* *121–122*, 138–145.
- Vanoli, F., Tomishima, M., Feng, W., et al. (2017). CRISPR–Cas9-guided oncogenic chromosomal translocations with conditional fusion protein expression in human mesenchymal cells. *Proc. Natl Acad. Sci. USA* *114*, 3696–3701.
- Wang, H., and Sun, W. (2017). CRISPR-mediated targeting of HER2 inhibits cell proliferation through a dominant negative mutation. *Cancer Lett.* *385*, 137–143.
- Wang, H., Yang, H., Shivalila, C.S., et al. (2013). One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell* *153*, 910–918.
- Wang, K., de la Torre, D., Robertson, W.E., et al. (2019a). Programmed chromosome fission and fusion enable precise large-scale genome rearrangement and assembly. *Science* *365*, 922–926.
- Wang, P., Zhang, J., Sun, L., et al. (2018). High efficient multisites genome editing in allotetraploid cotton (*Gossypium hirsutum*/*Gossypium hirsutum*) using CRISPR/Cas9 system. *Plant Biotechnol. J.* *16*, 137–150.
- Wang, W., Zhang, L., Wang, X., et al. (2019b). The advances in CRISPR technology and 3D genome. *Semin. Cell Dev. Biol.* *90*, 54–61.
- Weber, J., Ollinger, R., Friedrich, M., et al. (2015). CRISPR/Cas9 somatic multiplex-mutagenesis for high-throughput functional cancer genomics in mice. *Proc. Natl Acad. Sci. USA* *112*, 13982–13987.
- Weckselblatt, B., and Rudd, M.K. (2015). Human structural variation: mechanisms of chromosome rearrangements. *Trends Genet.* *31*, 587–599.
- Wendt, K.S., Yoshida, K., Itoh, T., et al. (2008). Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature* *451*, 796–801.
- Wienert, B., Wyman, S.K., Richardson, C.D., et al. (2019). Unbiased detection of CRISPR off-targets in vivo using DISCOVER-Seq. *Science* *364*, 286–289.
- Woo, J.W., Kim, J., Kwon, S.I., et al. (2015). DNA-free genome editing in plants with preassembled CRISPR–Cas9 ribonucleoproteins. *Nat. Biotechnol.* *33*, 1162–1164.
- Wu, J., Vilarino, M., Suzuki, K., et al. (2017). CRISPR–Cas9 mediated one-step disabling of pancreatogenesis in pigs. *Sci. Rep.* *7*, 10487.
- Wu, S., Ying, G., Wu, Q., et al. (2007). Toward simpler and faster genome-wide mutagenesis in mice. *Nat. Genet.* *39*, 922–930.
- Wu, Y., Jia, Z., Ge, X., et al. (2020). Three-dimensional genome architectural CCCTC-binding factor makes choice in duplicated enhancers at Pcdhalpha locus. *Sci. China Life Sci.* *63*, 835–844.
- Wutz, G., Varnai, C., Nagasaka, K., et al. (2017). Topologically associating domains and chromatin loops depend on cohesin and are regulated by CTCF, WAPL, and PDS5 proteins. *EMBO J.* *36*, 3573–3599.
- Xiao, A., Wang, Z., Hu, Y., et al. (2013). Chromosomal deletions and inversions mediated by TALENs and CRISPR/Cas in zebrafish. *Nucleic Acids Res.* *41*, e141.
- Xie, K., Minkenberg, B., and Yang, Y. (2015). Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system. *Proc. Natl Acad. Sci. USA* *112*, 3570–3575.
- Xu, D., Ma, R., Zhang, J., et al. (2018). Dynamic nature of CTCF tandem 11 zinc fingers in multivalent recognition of DNA as revealed by NMR spectroscopy. *J. Phys. Chem. Lett.* *9*, 4020–4028.
- Xu, S., Kim, J., Tang, Q., et al. (2020). Cas9 is a genome mutator by directly disrupting DNA-PK dependent DNA repair pathway. *Protein Cell* *11*, 352–365.
- Xue, W., Chen, S., Yin, H., et al. (2014). CRISPR-mediated direct mutation of cancer genes in the mouse liver. *Nature* *514*, 380–384.
- Yan, L., Wei, S., Wu, Y., et al. (2015). High-efficiency genome editing in arabidopsis using YAO promoter-driven CRISPR/Cas9 system. *Mol. Plant* *8*, 1820–1823.
- Yan, M., and Li, J. (2019). The evolving CRISPR technology. *Protein Cell* *10*, 783–786.
- Yang, G., and Huang, X. (2019). Methods and applications of CRISPR/Cas system for genome editing in stem cells. *Cell Regen.* *8*, 33–41.
- Yang, L., and Chen, J. (2020). A tale of two moieties: rapidly evolving CRISPR/Cas-based genome editing. *Trends Biochem. Sci.* *45*, 874–888.
- Yang, W. (2010). Nucleases: diversity of structure, function and mechanism. *Q. Rev. Biophys.* *44*, 1–93.
- Yang, W., Liu, Y., Tu, Z., et al. (2019). CRISPR/Cas9-mediated PINK1 deletion leads to neurodegeneration in rhesus monkeys. *Cell Res.* *29*, 334–336.
- Yao, X., Wang, X., Hu, X., et al. (2017). Homology-mediated end joining-based targeted integration using CRISPR/Cas9. *Cell Res.* *27*, 801–814.
- Yasue, A., Mitsui, S.N., Watanabe, T., et al. (2014). Highly efficient targeted mutagenesis in one-cell mouse embryos mediated by the TALEN and CRISPR/Cas systems. *Sci. Rep.* *4*, 5705.
- Yeh, C.D., Richardson, C.D., and Corn, J.E. (2019). Advances in genome editing through control of DNA repair pathways. *Nat. Cell Biol.* *21*, 1468–1478.
- Yen, S.T., Zhang, M., Deng, J.M., et al. (2014). Somatic mosaicism and allele complexity induced by CRISPR/Cas9 RNA injections in mouse zygotes. *Dev. Biol.* *393*, 3–9.
- Yin, H., Song, C.Q., Dorkin, J.R., et al. (2016). Therapeutic genome editing by combined viral and non-viral delivery of CRISPR system components in vivo. *Nat. Biotechnol.* *34*, 328–333.
- Yin, K., Han, T., Liu, G., et al. (2015). A geminivirus-based guide RNA delivery system for CRISPR/Cas9 mediated plant genome editing. *Sci. Rep.* *5*, 14926.
- Yin, M., Wang, J., Wang, M., et al. (2017). Molecular mechanism of directional CTCF recognition of a diverse range of genomic sites. *Cell Res.* *27*, 1365–1377.
- Yoshimi, K., Kaneko, T., Voigt, B., et al. (2014). Allele-specific genome editing and correction of disease-associated phenotypes in rats using the CRISPR–Cas platform. *Nat. Commun.* *5*, 4240.
- Yu, Y., and Bradley, A. (2001). Engineering chromosomal rearrangements in mice. *Nat. Rev. Genet.* *2*, 780–790.
- Yuan, L., Sui, T., Chen, M., et al. (2016). CRISPR/Cas9-mediated GJA8 knockout in rabbits recapitulates human congenital cataracts. *Sci. Rep.* *6*, 22024.

- Zhai, Y., Xu, Q., Guo, Y., et al. (2016). Characterization of a cluster of CTCF-binding sites in a protocadherin regulatory region. *Yi Chuan* 38, 323–336.
- Zhang, F. (2019). Development of CRISPR–Cas systems for genome editing and beyond. *Q. Rev. Biophys.* 52, e6.
- Zhang, L., Jia, R., Palange, N.J., et al. (2015). Large genomic fragment deletions and insertions in mouse using CRISPR/Cas9. *PLoS One* 10, e0120396.
- Zhang, Y., and Li, G. (2020). Advances in technologies for 3D genomics research. *Sci. China Life Sci.* 63, 811–824.
- Zhang, Y., Liang, Z., Zong, Y., et al. (2016). Efficient and transgene-free genome editing in wheat through transient expression of CRISPR/Cas9 DNA or RNA. *Nat. Commun.* 7, 12617.
- Zhang, Y., Pribil, M., Palmgren, M., et al. (2020). A CRISPR way for accelerating improvement of food crops. *Nat. Food* 1, 200–205.
- Zheng, H., and Xie, W. (2019). The role of 3D genome organization in development and cell differentiation. *Nat. Rev. Mol. Cell Biol.* 20, 535–550.
- Zhong, H., Chen, Y., Li, Y., et al. (2015). CRISPR-engineered mosaicism rapidly reveals that loss of *Kcnj13* function in mice mimics human disease phenotypes. *Sci. Rep.* 5, 8366.
- Zhou, J., Wang, J., Shen, B., et al. (2014). Dual sgRNAs facilitate CRISPR/Cas9-mediated mouse genome targeting. *FEBS J.* 281, 1717–1725.
- Zhu, J., Song, N., Sun, S., et al. (2016a). Efficiency and inheritance of targeted mutagenesis in maize using CRISPR–Cas9. *J. Genet. Genomics* 43, 25–36.
- Zhu, S., Li, W., Liu, J., et al. (2016b). Genome-scale deletion screening of human long non-coding RNAs using a paired-guide RNA CRISPR–Cas9 library. *Nat. Biotechnol.* 34, 1279–1286.
- Zhu, X., Xu, Y., Yu, S., et al. (2014). An efficient genotyping method for genome-modified animals and human cells generated with CRISPR/Cas9 system. *Sci. Rep.* 4, 6420.
- Zuccaro, M.V., Xu, J., Mitchell, C., et al. (2020). Reading frame restoration at the *EYS* locus, and allele-specific chromosome removal after Cas9 cleavage in human embryos. *Cell*, doi: 10.1016/j.cell.2020.10.025.
- Zuckermann, M., Hovestadt, V., Knobbe-Thomsen, C.B., et al. (2015). Somatic CRISPR/Cas9-mediated tumour suppressor disruption enables versatile brain tumour modelling. *Nat. Commun.* 6, 7391.
- Zuo, Z., and Liu, J. (2016). Cas9-catalyzed DNA cleavage generates staggered ends: evidence from molecular dynamics simulations. *Sci. Rep.* 5, 37584.