

# CRISPR/Cas9 systems targeting $\beta$ -globin and *CCR5* genes have substantial off-target activity

Thomas J. Cradick, Eli J. Fine, Christopher J. Antico and Gang Bao\*

Department of Biomedical Engineering, Georgia Institute of Technology and Emory University, Atlanta, GA 30332, USA

Received June 15, 2013; Revised July 17, 2013; Accepted July 19, 2013

## ABSTRACT

The ability to precisely modify endogenous genes can significantly facilitate biological studies and disease treatment, and the clustered regularly interspaced short palindromic repeats (CRISPR) systems have the potential to be powerful tools for genome engineering. However, the target specificity of CRISPR systems is largely unknown. Here we demonstrate that CRISPR/Cas9 systems targeting the human hemoglobin  $\beta$  and C-C chemokine receptor type 5 genes have substantial off-target cleavage, especially within the hemoglobin  $\delta$  and C-C chemokine receptor type 2 genes, respectively, causing gross chromosomal deletions. The guide strands of the CRISPR/Cas9 systems were designed to have a range of mismatches with the sequences of potential off-target sites. Off-target analysis was performed using the T7 endonuclease I mutation detection assay and Sanger sequencing. We found that the repair of the on- and off-target cleavage resulted in a wide variety of insertions, deletions and point mutations. Therefore, CRISPR/Cas9 systems need to be carefully designed to avoid potential off-target cleavage sites, including those with mismatches to the 12-bases proximal to the guide strand protospacer-adjacent motif.

## INTRODUCTION

The ability to precisely edit endogenous DNA sequences has greatly facilitated the creation of cell lines and animal models for biological and disease studies, and led to unprecedented opportunities in therapeutics. For example, engineered zinc finger nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs) have generated hundreds of animal models for disease studies (1,2), and nuclease-based treatment strategies are currently undergoing clinical trials. The discovery of a

bacterial defense system that uses RNA-guided DNA cleaving enzymes and clustered, regularly interspaced, short palindromic repeats (CRISPR) (3–7) may provide an exciting alternative to ZFNs and TALENs, as the CRISPR-associated (Cas) protein remains the same for different gene targets; only the short sequence of the guide RNA needs to be changed to redirect the site-specific cleavage (8).

Potential off-target cleavage by engineered nucleases poses concerns both for adverse events in therapeutic applications and confounding variables in biological studies. ZFNs (9,10) have been shown to lack exquisite specificity and may cleave sequences in addition to their intended targets, which often induces unwanted mutations and/or toxicity (11,12). Although recent reports suggest that TALENs have better specificity than ZFNs, off-target activities have been found for TALENs as well (13–15). Previous *in vitro* studies suggested that CRISPR/Cas9 systems have a high potential for off-target activity, as they have more promiscuous binding abilities at positions distal from the protospacer-adjacent motif (PAM) region (8,16–18). Further, because the guide RNA strands typically target a DNA sequence of ~20 bp, relatively short compared with the  $\geq 36$  bp targeted by TALENs, many potential off-target sites may exist in large genomes, such as in mammals. Additionally, because non-Watson–Crick base pairing is known to occur (18), it is possible that CRISPR/Cas9 systems have more off-target activities compared with corresponding ZFNs and TALENs.

To determine the off-target effects of CRISPR/Cas9 systems in the context of the human genome, we constructed several CRISPR/Cas9 systems with guide RNA strands targeting the human hemoglobin  $\beta$  (*HBB*) and C-C chemokine receptor type 5 (*CCR5*) genes, expressed them in human embryonic kidney 293T (HEK-293T) cells, and quantified their on- and off-target activities using the T7 endonuclease I (T7E1) mutation detection assay and Sanger sequencing, with special attention placed on the mismatches between the guide strand and the cleaved sequences. We found that the CRISPR/Cas9 systems targeting the human *HBB* and *CCR5* genes have substantial off-target cleavage and resulted in a wide

\*To whom correspondence should be addressed. Tel: +1 404 385 0373; Fax: +1 404 385 3856; Email: gang.bao@bme.gatech.edu

variety of insertions and deletions (indels), as well as point mutations. Our results have significant implications to the design and optimization of the CRISPR/Cas9 systems for genome editing applications.

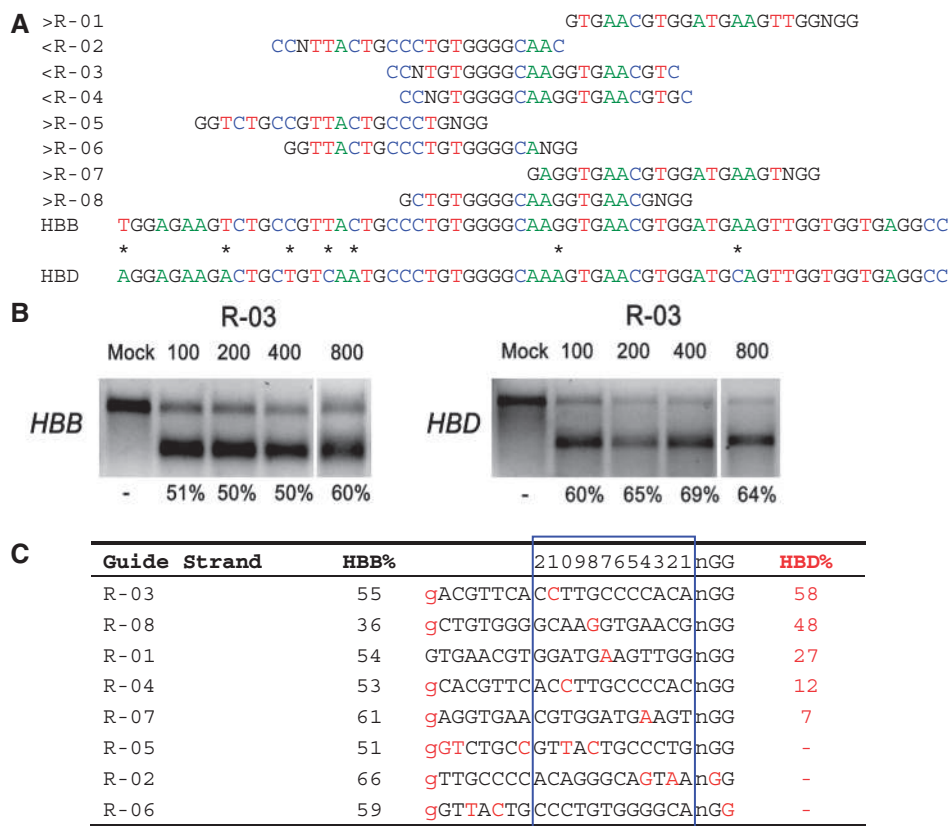
## MATERIALS AND METHODS

### CRISPR design and testing

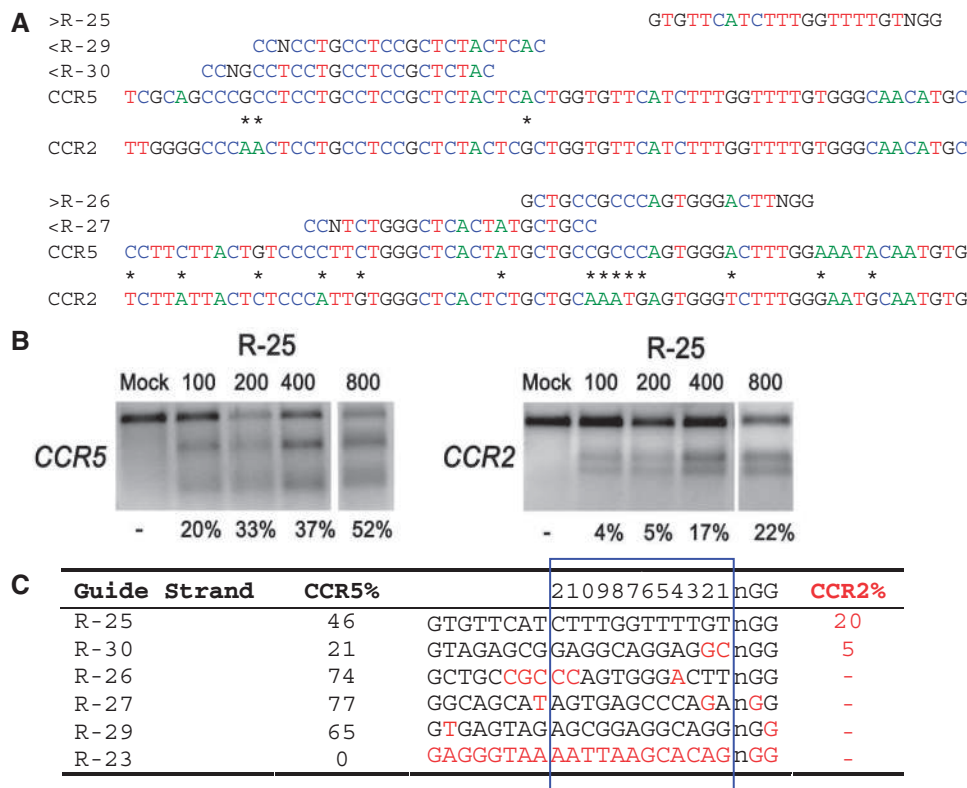
There were no CRISPR target sites in the human *HBB* gene sequence with their proximal 12 bases unique in the human genome (8); therefore, we chose CRISPR/Cas9 guide strands targeting *HBB* by comparing the similar regions in the human hemoglobin  $\delta$  (*HBD*) gene. We designed eight 20-base guide strands to target sites near the sickle mutation in the *HBB* gene (Figure 1a), each adjacent to a PAM sequence that contains the canonical trinucleotide NGG. We also designed five guide strands to target two segments in the human *CCR5* gene (Figure 2a), and tested the corresponding CRISPR/Cas9 systems to determine their on-target cleavage and potential off-target activity at the human C-C chemokine receptor type 2 (*CCR2*) gene. Herein we use the name of the

guide strand (such as R-03) to represent the CRISPR/Cas9 system with the specified guide strand.

CRISPR plasmids were generated by kinasing and annealing oligonucleotides containing a G followed by 19 additional bases of the guide strand plus sticky ends, ligating into the pX330 plasmid that contains a U6 promoter-driven chimeric +85-bp guide strand and a CHb promoter-driven Cas9 expression cassette, and expressed together from the 8.5-kb *Cas9* gene expression plasmid, pX330 (provided by Dr. Feng Zhang, and also available through Addgene 42230) (19). In a 24-well plate, 80 000 HEK-293T cells/well were seeded and cultured in Dulbecco's modified Eagle medium supplemented with 10% fetal bovine serum (FBS) and 2 mM fresh L-glutamine, 24 h prior to transfection. Cells were transfected with 100, 200, 400 or 800 ng of CRISPR plasmids (normalized to 800 ng with pUC18) using FuGENE HD (Promega). The genomic DNA was harvested after 3 days using QuickExtract (EpiCentre). Targeted cleavage was measured at the endogenous loci by the rate of mutations through mis-repair, detected using amplification of these sites using bar-coded or traditional primers (Supplementary Table S1) and the T7E1 assay. The



**Figure 1.** On- and off-target cleavage by CRISPR/Cas9 systems targeting the *HBB* gene. (a) Guide strands are aligned to their target sites and corresponding region in *HBB* and *HBD*. Forward direction guide strands (marked 'greater than') are shown adjacent to NGG, representing the PAM sequence. Guide strands complementary to the reverse strand (marked 'less than') are listed to the right of CCN. Asterisks between *HBB* and *HBD* indicate nucleotides that differentiate the two genes. The first base shown in *HBB* is the sickle cell anemia mutation site. For clarity, the A, C, T and G nucleotides are shown in green, blue, red and black, respectively. (b) R-03-induced cleavage at *HBB* (on-target) and *HBD* (off-target) measured by the T7E1 assay. Cells were transfected with 100, 200, 400 or 800 ng of the R-03 CRISPR plasmid. (c) Guide strands ranked in order of the off-target mutation rates at *HBD*. Differences between the guide sequence and *HBD* are in red. A lowercase g indicates that the first base in *HBB* does not match the guide strands' initial G (for all but R-01). The 12 bases closest to the PAM are boxed in blue and numbered on top.



**Figure 2.** On- and off-target cleavage by CRISPR/Cas9 systems targeting the *CCR5* gene. (a) Guide strands are aligned to their target sites in *CCR5* and corresponding region in *CCR2*. Forward direction guide strands (marked 'greater than') are shown adjacent to NGG, representing the PAM sequence. Guide strands complementary to the reverse strand (marked 'less than') are listed to the right of CCN. Asterisks between *CCR5* and *CCR2* indicate nucleotides that differentiate the two genes. For clarity, the A, C, T and G nucleotides are shown in green, blue, red and black, respectively. (b) Cells were transfected with 100, 200, 400 or 800 ng of the R-25 CRISPR plasmid. Results of the T7E1 mutation detection assay with the on- and off-target mutation rates at *CCR5* and *CCR2*, respectively. R-23 targeting *CFTR* was used as a negative control. (c) Guide strands ranked in order of the off-target mutation rates at *CCR2*. Differences between the guide strand sequence and complementary sequence in *CCR2* are in red. The 12 bases closest to the PAM are boxed in blue and numbered on top.

fragments were separated on agarose gels and quantitated using ImageJ; the mutation frequencies were calculated and averaged. To better determine the mutation rate, amplification bands were cloned using the TOPO<sup>®</sup> TA kit [Invitrogen], Sanger sequenced and aligned to observe the individual mutations and determine the mutational spectra. Sanger sequencing was chosen to ensure the detection of large insertions and deletions, as well as effectively detect single base indels, both of which can be problematic with the next-generation sequencing methods.

### Off-target analysis

Off-target analysis was performed using a bioinformatics-based search tool to select potential off-target sites, which were evaluated using the T7E1 mutation detection assay (data not shown). Sanger sequencing was used to confirm the gene modification frequencies for the CRISPR/Cas9 systems, including guide strand R-02 at *GRIN3A* (Supplementary Figure S1b, Table 1).

### Chromosomal deletion analysis

To assay for gross chromosomal deletions, genomic DNA from cells transfected with R-03 was amplified using the *HBD* forward primer and the reverse primer downstream

of the *HBB* site. Genomic DNA from cells transfected with R-25 or R-30 were similarly amplified using the *CCR2* forward and the *CCR5* reverse primers. Agarose gels were used to confirm that the polymerase chain reaction (PCR) product sizes were consistent with chromosomal deletions between these sites (not shown). The R-03, R-25 and R-30 PCR products were cloned and the individual colonies Sanger sequenced and aligned.

## RESULTS

### On- and off-target cleavage of CRISPR/Cas9 systems targeting *HBB* and *CCR5*

We quantified and compared the on- and off-target cleavage by the CRISPR/Cas9 systems targeting *HBB* and *CCR5*, with special attention placed on the effects of mismatches between the guide strands and the complementary target sequences. This allowed a direct evaluation of the impact of the location and number of mismatches within the 12 bases nearest the PAM region, as well as those in the PAM region (that usually match the canonical NGG motif) (Table 1) on potential off-target activities (8,20). We found that the CRISPR/Cas9 systems targeting

**Table 1.** CRISPR on- and off-target cleavage rates

Guide strand	Mismatches <sup>a</sup>	Indel % by sequencing (%)	Indel % by T7E1 (%)	<sup>b</sup> 210987654321nGG	Region	Gene
R-01	0	67	54	GTGAACGTGGATGAAGTTGGtGG	Exon	<i>HBB</i>
	1	30	27	GTGAACGTGGATGcAGTTGGtGG	Exon	<i>HBD</i>
R-02	1	75	66	cTTGCCCCACAGGGCAGTAAcGG	Exon	<i>HBB</i>
	3	77	33 <sup>c</sup>	tcaGCCCCACAGGGCAGTAAcGG	Intergenic	<i>GRIN3A</i>
R-03	1	70	55	cACGTTACCTTGCCCCACaGG	Exon	<i>HBB</i>
	2	62	58	cACGTTACCTTTGCCCCACaGG	Exon	<i>HBD</i>
R-04	1	47	53	cCACGTTACCTTGCCCCACaGG	Exon	<i>HBB</i>
	2	10	12	cCACGTTACCTTTGCCCCACaGG	Exon	<i>HBD</i>
R-05	1		51	aGTCTGCCGTTACTGCCCTGnGG	Exon	<i>HBB</i>
R-06	1		59	cGTTACTGCCCTTGGGGCAnGG	Exon	<i>HBB</i>
R-07	1	68	61	aAGGTGAACGTGGATGAAGTtGG	Exon	<i>HBB</i>
	2		7	aAGGTGAACGTGGATGcAGTtGG	Exon	<i>HBD</i>
R-08	1	38	36	cCTGTGGGGCAAGTGAACGtGG	Exon	<i>HBB</i>
	2		48	cCTGTGGGGCAaAGTGAACGtGG	Exon	<i>HBD</i>
R-25	0	50	46	GTGTTCACTTTGGTTTTGTgGG	Exon	<i>CCR5</i>
	0	32	20	GTGTTCACTTTGGTTTTGTgGG	Exon	<i>CCR2</i>
R-26	0	76	74	GCTGCCGCCAGTGGGACTTtGG	Exon	<i>CCR5</i>
R-27	0	78	77	GGCAGCATAGTGAGCCAGAGGG	Exon	<i>CCR5</i>
R-29	0		65	GTGAGTAGAGCGGAGGCAGGnGG	Exon	<i>CCR5</i>
R-30	0		21	GTAGAGCGGAGGCAGGAGGcGG	Exon	<i>CCR5</i>
	2		5	GTAGAGCGGAGGCAGGAGt_tGGG	Exon	<i>CCR2</i>

<sup>a</sup>Number of base differences between the guide strand and complementary sequence, including the 5' nucleotide.

<sup>b</sup>Base pair positions from the PAM are numbered above the loci. The differences between the guide strand and complementary sequences are indicated in lowercase underlined nucleotides. The first of the three nucleotides in the PAM sequence is also indicated in lowercase.

<sup>c</sup>T7E1 was performed in duplicate for this off-target site, not triplicate as with all other cases.

the human *HBB* and *CCR5* genes had significant off-target cleavage activities, especially at the *HBD* and *CCR2* genes, which have high sequence homology with *HBB* and *CCR5*, respectively.

Table 1 summarizes the on- and off-target cleavage rates in which, for each CRISPR/Cas9 system, the complementary sequence of the guide strand, the number of mismatches within the guide strand and the name and genetic region of the on- and off-target activities are provided. Specifically, in Table 1, the third and fourth columns list, respectively, the indel percentages determined by Sanger sequencing and T7E1.

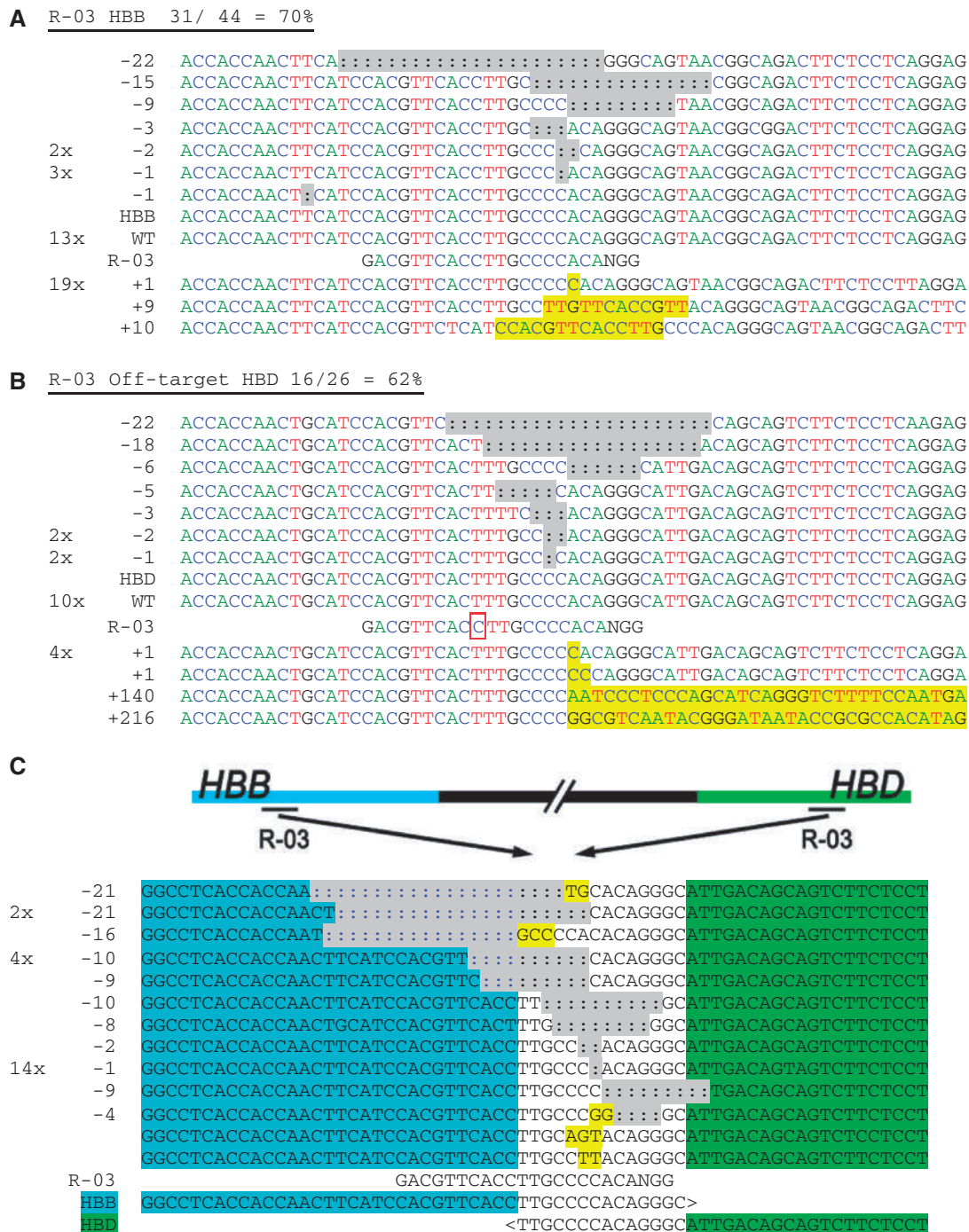
Guide strands directed toward *HBB* resulted in high rates of on-target activity, with an average mutation frequency of 54% measured by the T7E1 assay (Figure 1c, Supplementary Figure S2). Because the T7E1 assay may not cleave the PCR product completely and assumptions must be made about the indel diversity to calculate the mutation percentages (21), we verified the mutation frequencies using Sanger sequencing. We found that for some guide strands and loci, Sanger sequencing gave much higher mutation frequencies than the T7E1 measurements. For example, Sanger sequencing of the *HBB* loci indicated that R-02 and R-03 resulted, respectively, in 60 of 80 (75%) and 31 of 44 (70%) sequences with insertions or deletions (indels) indicative of the error-prone nonhomologous end-joining (NHEJ) DNA repair pathway (Supplementary Figure S1a, Figure 3a). Similarly, HEK-293T cells transfected with CRISPR constructs containing guide strands targeting *CCR5* resulted in high rates of on-target activity, with an average of 57% mutation frequency measured by the T7E1 assay (Figure 2c, Supplementary Figure S2).

Some CRISPR/Cas9 systems with guide strands targeting *HBB* also cleaved *HBD* (some at high rates), even though there are mismatches between the guide strands and the complementary *HBD* sequences. For example, guide strands having just one-base mismatch with the complementary *HBD* sequences, located at positions 4 (R-07), 7 (R-01), 8 (R-08), 10 (R-04) and 11 (R-03) bases from the PAM sequence, resulted in off-target mutation rates ranging from 7 to 58%, roughly corresponding to the distance between the mismatch location and the PAM sequence, with R-04 as an exception (Figure 1c). Note that two off-target sites at *HBD* had mutation rates even higher than the on-target rates at *HBB*, especially R-08, which induced a mutation rate of 48% at *HBD*, much higher than that at *HBB* (36%).

To allow RNA transcription by the U6 polymerase, the guide strand is typically preceded by a guanine (8). We (and others) found that it is not necessary for the guanine base to match the target site for efficient cleavage, as seven guide strands without a guanine at this position induced mutations in *HBB* (R-02 to R-08) and four guide strands (R-03, R-04, R-07, R-08) induced mutations in *HBD* (Figure 1c).

To a lesser extent, *CCR5*-targeting CRISPR/Cas9 systems also induced off-target cleavage on *CCR2*, with mutation rates of 5% and 20% (Figure 2c, Supplementary Figure S2). Specifically, guide strand R-25 was designed with two identical genomic targets in *CCR5* and *CCR2* genes to identify the influence of factors beyond sequence homology, such as genomic context. The CRISPR/Cas9 system with R-25 showed a >2-fold difference in mutation rate at these two sites (46% versus 20% mutation rate, Figure 2c). These results suggest that other





**Figure 3.** Chromosomal deletions in *HBB* and *HBD* induced by CRISPR/Cas9 systems. HEK-293T cells were transfected with each CRISPR construct, and their genomic DNA harvested after 3 days in culture. The (a) on- and (b) off-target loci for guide strands R-03 were amplified with flanking PCR primers, cloned and Sanger sequenced. Sequencing reads are given for each guide strand and aligned to the wild-type sequence. The number of times each read occurred is indicated to the left of the alignment. Unmodified reads are indicated by 'WT'. In (b) the guide strand mismatch is boxed. In (a) and (b), the A, C, T and G nucleotides are shown in green, blue, red and black, respectively, for clarity. (c) Genomic DNA from cells treated with R-03 was amplified using an *HBD* forward primer and reverse primer downstream of the *HBB* site. The PCR products were sequenced and aligned to '*HBB-HBD*' with the bases unique to *HBB* or *HBD* indicated in blue or green, respectively, surrounding an identical area found in both genes. Sequencing detected that each product contained indels and mutations consistent with NHEJ, near the target sites for R-03. Insertions and point mutations are marked in yellow and deletions (:) are highlighted in gray.

features such as genomic context may play an important role in cleavage activity. Surprisingly, although guide strand R-30 had two mismatches with *CCR2* at the two bases proximal to the PAM region, it induced mutations

in *CCR2* at a rate of 5% as measured by T7E1 with 800 ng of plasmid in transfection (Figures 2c, Supplementary Figure S3). R-30 transfections with 1100 ng of plasmid induced mutations of 21% quantified by sequencing, but

only 6% by T7E1 (Supplementary Figures S3, S1h); part of the difference is likely because of the incomplete cleavage of PCR products by T7E1.

A distinct feature of CRISPR off-target activity as related to mismatches in the guide strand is that mismatches in the PAM region can prevent off-target cleavage (19). For example, R-06, which has a one-base mismatch in the PAM, did not induce detectable mutations at *HBD*, although it has a perfect match of the 14 bases proximal to the PAM (Figure 1c, Supplementary Figure S2). Further, R-02 did not induce cleavage at *HBD* because of the one-base mismatch in the PAM and two mismatches at positions 2 and 4 from the PAM (Figure 1c). Similarly, there was no off-site mutagenesis detected at *CCR2* by the *CCR5*-targeting CRISPR/Cas9 systems with guide strands R-27 and R-29 that had NTG and NGT PAM substitutions, respectively. In particular, although R-29 had a perfect match with the 18-bp sequence proximal to the PAM, a one-base mismatch in the PAM region prevented cleavage of *CCR2* (Figure 2c, Supplementary Figure S2). Clearly, off-target cleavage could also be prevented without any mismatch in the PAM, by having multiple mismatches between the guide strand and the complementary target sequence proximal to the PAM, as demonstrated by R-05 (Figure 1c) and R-26 (Figure 2c).

To quantify the change in CRISPR/Cas9 cleavage activity with transfection conditions, CRISPR plasmids were transfected at doses from 100 to 800 ng, and corresponding on- and off-target activities measured by T7E1 (Figures 1b and 2b, Supplementary Figure S4). With the dose decreases, we found that R-04 and R-25 gave lower on- and off-target activities, whereas R-30 resulted in increased on-target activity and decreased off-target activity; the on- and off-target activities of R-03 and R-08 remained roughly the same. In general, transfection with the lowest dose (100 ng) increased the ratio of on-target to off-target activities for R-04, R-25 and R-30, although not for R-03 and R-08. These findings expand the results of a recent study where no appreciable changes in on- and off-target rates were found with two CRISPR guide strands at two doses (22).

### Target loci sequence and chromosomal deletion analysis

As revealed by Sanger sequencing, CRISPR-targeted loci showed a wide variety of insertions, deletions and point mutations. Because *HBD* is located ~7 kb upstream of *HBB* on chromosome 11, cleavage at both sites raises the possibility of chromosomal rearrangements, including a deletion of the intervening segment (23–26). These gross chromosomal deletions are seen with guide strand R-03, which cleaves both *HBB* and *HBD* at high rates, even though it has a mismatch to *HBD* (Figure 3a and b). PCR amplification and sequence analysis revealed gross chromosomal deletions resulting from rejoining the DNA double-strand break ends induced by two cleavage events in (or near) the conserved region of the *HBB* and *HBD* (Figure 3c). Each of these joined *HBD*–*HBB* clones amplified from cells transfected with R-03 had an indel consistent with NHEJ. Quantitative PCR was used to

estimate that 12.6% of *HBB* alleles contained the chromosomal deletion with *HBD* (Supplementary Figure S5).

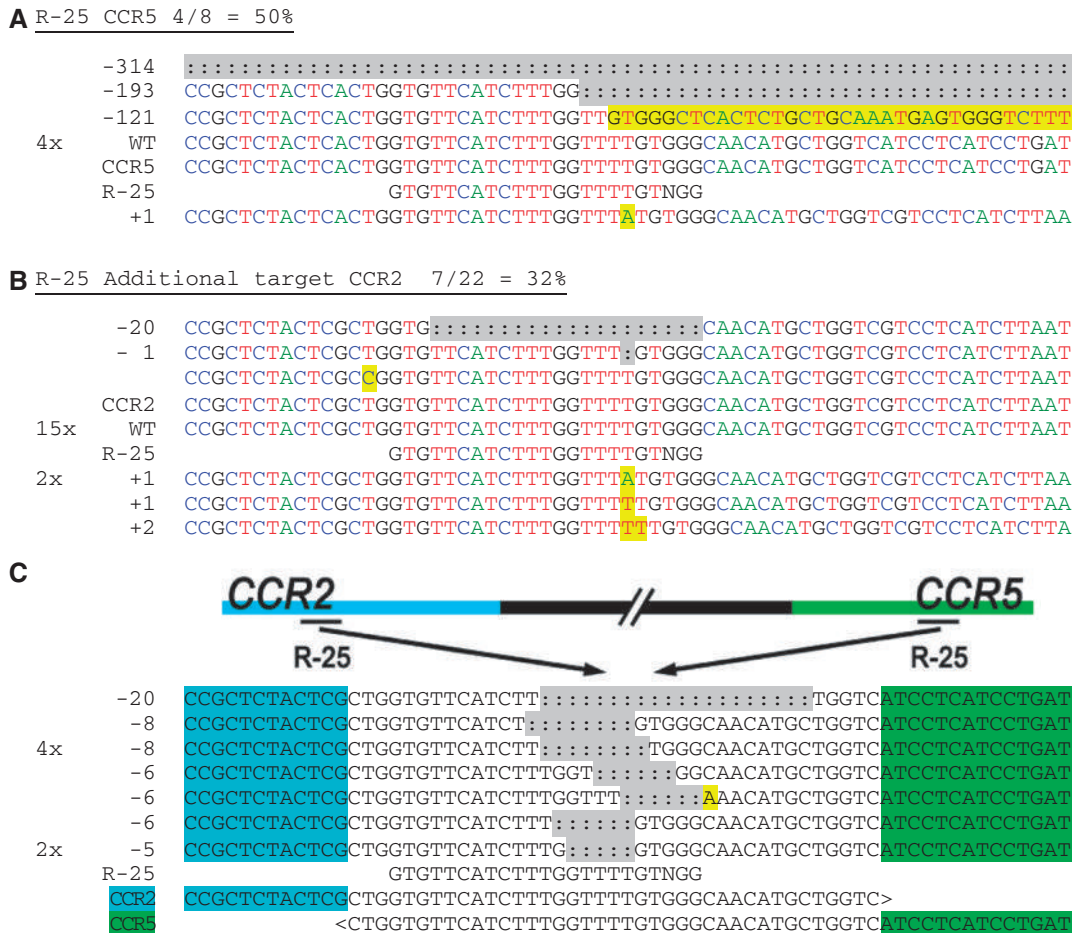
Similarly, *CCR5* is located ~8 kb upstream of *CCR2* on chromosome 3; thus, chromosomal rearrangements may occur with cleavages at both *CCR5* and *CCR2*. These gross chromosomal deletions were detected with the R-25 CRISPR/Cas9 system, which cleaved both genes at high rates (Figure 4a and b). Here again, PCR amplification and sequence analysis revealed two cleavage events in (or near) a conserved region of the *CCR5* and *CCR2* genes, as indicated by indels consistent with NHEJ (Figure 4c). Cells transfected with the R-30 CRISPR/Cas9 system also had chromosomal deletions between *CCR5* and *CCR2* (Supplementary Figure S1h).

Sequencing the on- and off-target loci revealed a range of different indels as a result of CRISPR/Cas9-induced DNA cleavage, including three large insertions (140, 216 and 448 bp). Specifically, our results indicated that one-base insertions and deletions occurred frequently, usually several bases from the PAM sequence, consistent with the reported cleavage between the third and fourth bases from the PAM (27). As shown in Figure 5, the frequency of cleavage-induced gene modifications varied significantly with indels of different sizes, though 21% were one-base insertions and 12% one-base deletions. Interestingly, a common indel size was a 9-bp deletion that occurred in 14% of the clones. Because the range of indels is influenced by sequence differences, microhomologies and/or palindromes in the area being cleaved (28), and our results were primarily from a limited number of overlapping target sites, further sequence analysis is needed to ensure a more general distribution.

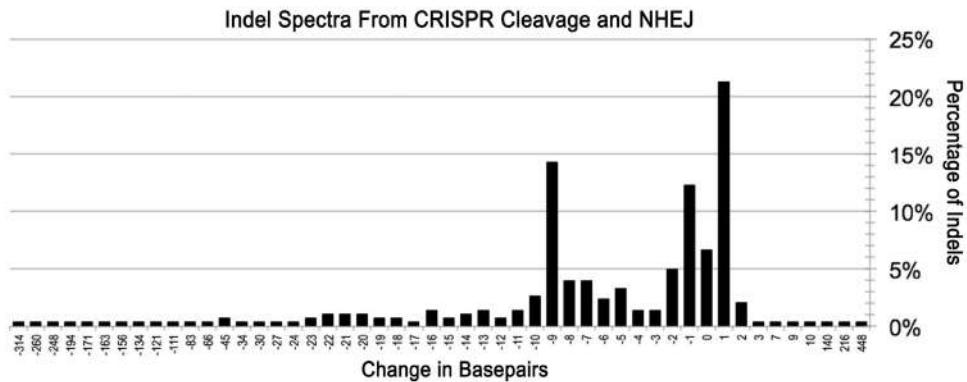
### DISCUSSION

Although CRISPR/Cas9 systems can induce high rates of gene modification in mammalian cells, they do not have perfect specificity, similar to previous observations with ZFNs and TALENs. Our results demonstrate that CRISPR/Cas9 systems can have significant off-target activities even if 10 or 11 of the 12 bases proximal to the PAM sequence match. Therefore, it is likely that there are many more potential off-target sites in the human genome than previously thought (8,29), if cleavage occurs when any permutation of 10 of the 12 bases in the guide strand matches a genomic sequence. Our results suggest that mismatches in, or proximal to, the PAM sequence could block cleavage, as seen by others (19,22,29). However, there are contrary examples, such as R-30 that cleaves *CCR2* with mismatches in the two PAM-proximal bases (Figure 2c, Supplementary Figure S1h). Additional studies are required to deduce the key design rules concerning these mismatches.

The importance of the PAM sequence (30) was corroborated by the lack of cleavage at some complementary sequences similar to the guide strand, but with PAM sequences differing from NGG (Figures 1c and 2c). An example is guide strand R-06 that cleaved *HBB* at 59%, but had no detectable cleavage at *HBD*, presumably due to the NGA in the PAM sequence. Similarly, R-29 cleaves



**Figure 4.** Chromosomal deletions in *CCR5* and *CCR2* induced by CRISPR/Cas9 systems. HEK-293T cells were transfected with each CRISPR construct, and their genomic DNA harvested after 3 days in culture. The (a) on- and (b) off-target loci for guide strands R-25 were amplified with flanking PCR primers, cloned and Sanger sequenced. Sequencing reads are given for each guide strand and aligned to the wild-type sequence. The number of times each read occurred is indicated to the left of the alignment. Unmodified reads are indicated by 'WT'. In (b) the guide strand mismatch is boxed. In (a) and (b), the A, C, T and G nucleotides are shown in green, blue, red and black, respectively, for clarity. (c) Genomic DNA from cells treated with R-25 was amplified using a *CCR2* forward primer and reverse primer downstream of the *CCR5* site. The PCR products were sequenced and aligned to '*CCR2-CCR5*' with the bases unique to *CCR2* or *CCR5* indicated in blue or green, respectively, surrounding an identical area found in both genes. Sequencing detected that each product contained indels and mutations consistent with NHEJ, near the target sites for R-25. Insertions and point mutations are marked in yellow and deletions (:) are highlighted in gray.



**Figure 5.** Indel spectra from CRISPR/Cas9 cleavage and NHEJ repair. The change in number of base pairs resulting from each indel was calculated and compiled. The y-axis represents the percentage of insertion or deletion. The most common indels for the CRISPR/Cas9 systems studied: 1 bp additions, 9 bp deletions and 1 bp deletions occurred in 64, 43 and 37 of 302 clones, respectively.

*CCR5* at 65% efficiency. R-29 failed to cleave at *CCR2* possibly due to the less tolerated, adjacent NGT PAM sequence, although the R-29 guide strand matches the 18 bases closest to the PAM sequence at *CCR2*.

Although Cas9 is thought to generate blunt ends (16,27), our results indicate that CRISPR-directed on- and off-target cleavage can induce a wide range of indels, with a large number of one-base insertions and a few large deletions. The high rate of off-target cleavage may result in large indels, causing a significant potential of mutagenesis and chromosomal rearrangements. For example, if two or more cleavage sites are on the same chromosome, it may lead to gross chromosomal deletions, as seen with R-03 (Figure 3c), R-25 (Figure 4c) and R-30 (Supplementary Figure S1h). These chromosomal deletions and the high levels of on- and off-target cleavage suggest that there might be other chromosomal rearrangements, translocations and inversions. Although the ability of engineered CRISPR/Cas9 systems to target multiple sites/genes with different guide strands is an exciting feature (8,29,31), each system may lead to off-target cleavage. The effect of having multiple guide strands on off-target cleavage and its effect on rates of chromosomal rearrangement have yet to be thoroughly studied (31). A CRISPR/Cas9 system may cause chromosomal rearrangements with one guide strand inducing cleavage at two defined locations, or with a pair of guide strands inducing deletion between the target sites (25); in both cases the off-target effects of each guide strand must be assayed. Therefore, multiplexed gene editing using CRISPR/Cas9-based approaches might have significant limitations unless optimal design of the guide strands can be performed to reduce or even eliminate the potential for gross chromosomal rearrangements.

As demonstrated in this work and elsewhere (19,22), CRISPR/Cas9 systems may have high rates of off-target cleavage; therefore, care must be taken when choosing and evaluating target sites. Even with diligent choice of target sites, in most genome editing applications, quantifying the off-target activities is necessary to identify unintended cleavage and mutagenesis. Transfection conditions, including plasmid dosage, may be optimized to decrease off-target cleavage, although the effects may vary with guide strands (Supplementary Figure S4). The variety of on- and off-target cleavage rates induced by CRISPR/Cas9 systems raises hope that better selection of target sites, possibly through rational design and/or screening in cells, can result in gene editing with improved specificity. Advanced genome searches may be needed in choosing optimal target sites by minimizing the number of potential off-target sites corresponding to different mismatches. More extensive off-target analysis of the CRISPR/Cas9 systems, with a combination of bioinformatics and experimental approaches, may reveal patterns and design guidelines that better predict the target sites that can be effectively cleaved with high specificity.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

## ACKNOWLEDGEMENTS

We thank Dr. Feng Zhang for helpful discussions and providing the Cas9 expression plasmid.

## FUNDING

The National Institutes of Health as an NIH Nanomedicine Development Center Award [PN2EY018244 to G.B.], NSF Graduate Research Fellowship [DGE-1148903 to E.J.F.]. Funding for open access charge: NIH.

*Conflict of interest statement.* None declared.

## REFERENCES

- Perez, E.E., Wang, J., Miller, J.C., Jouvenot, Y., Kim, K.A., Liu, O., Wang, N., Lee, G., Bartsevich, V.V., Lee, Y.L. *et al.* (2008) Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. *Nat. Biotechnol.*, **26**, 808–816.
- Geurts, A.M., Cost, G.J., Freyvert, Y., Zeitler, B., Miller, J.C., Choi, V.M., Jenkins, S.S., Wood, A., Cui, X., Meng, X. *et al.* (2009) Knockout rats via embryo microinjection of zinc-finger nucleases. *Science*, **325**, 433.
- Bolotin, A., Quinquis, B., Sorokin, A. and Ehrlich, S.D. (2005) Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology*, **151**, 2551–2561.
- Horvath, P. and Barrangou, R. (2010) CRISPR/Cas, the immune system of bacteria and archaea. *Science*, **327**, 167–170.
- Marraffini, L.A. and Sontheimer, E.J. (2010) CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. *Nat. Rev. Genet.*, **11**, 181–190.
- Garneau, J.E., Dupuis, M., Villion, M., Romero, D.A., Barrangou, R., Boyaval, P., Fremaux, C., Horvath, P., Magadán, A.H. and Moineau, S. (2010) The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature*, **468**, 67–71.
- Hale, C.R., Zhao, P., Olson, S., Duff, M.O., Graveley, B.R., Wells, L., Terns, R.M. and Terns, M.P. (2009) RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. *Cell*, **139**, 945–956.
- Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A. *et al.* (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science*, **339**, 819–823.
- Pattanayak, V., Ramirez, C.L., Joung, J.K. and Liu, D.R. (2011) Revealing off-target cleavage specificities of zinc-finger nucleases by in vitro selection. *Nat. Methods*, **8**, 765–770.
- Gabriel, R., Lombardo, A., Arens, A., Miller, J.C., Genovese, P., Kaeppl, C., Nowrouzi, A., Bartholomae, C.C., Wang, J., Friedman, G. *et al.* (2011) An unbiased genome-wide analysis of zinc-finger nuclease specificity. *Nat. Biotechnol.*, **29**, 816–823.
- Cornu, T.I. and Cathomen, T. (2010) Quantification of zinc finger nuclease-associated toxicity. *Methods Mol. Biol.*, **649**, 237–245.
- Ramirez, C.L., Certo, M.T., Mussolino, C., Goodwin, M.J., Cradick, T.J., McCaffrey, A.P., Cathomen, T., Scharenberg, A.M. and Joung, J.K. (2012) Engineered zinc finger nickases induce homology-directed repair with reduced mutagenic effects. *Nucleic Acids Res.*, **40**, 5560–5568.
- Tesson, L., Usal, C., Ménoret, S., Leung, E., Niles, B.J., Remy, S., Santiago, Y., Vincent, A.I., Meng, X., Zhang, L. *et al.* (2011) Knockout rats generated by embryo microinjection of TALENs. *Nat. Biotechnol.*, **29**, 695–696.
- Hockemeyer, D., Wang, H., Kiani, S., Lai, C.S., Gao, Q., Cassady, J.P., Cost, G.J., Zhang, L., Santiago, Y., Miller, J.C. *et al.* (2011) Genetic engineering of human pluripotent cells using TALE nucleases. *Nat. Biotechnol.*, **29**, 731–734.
- Mussolino, C., Morbitzer, R., Lutge, F., Dannemann, N., Lahaye, T. and Cathomen, T. (2011) A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. *Nucleic Acids Res.*, **39**, 9283–9293.



16. Gasiunas,G., Barrangou,R., Horvath,P. and Siksnys,V. (2012) Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc. Natl Acad. Sci. USA*, **109**, E2579–E2586.
17. Jinek,M., East,A., Cheng,A., Lin,S., Ma,E. and Doudna,J. (2013) RNA-programmed genome editing in human cells. *Elife*, **2**, e00471.
18. Jiang,W., Bikard,D., Cox,D., Zhang,F. and Marraffini,L.A. (2013) RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat. Biotechnol.*, **31**, 233–239.
19. Hsu,P.D., Scott,D.A., Weinstein,J.A., Ran,F.A., Konermann,S., Agarwala,V., Li,Y., Fine,E.J., Wu,X. *et al.* (2013) DNA targeting specificity of the RNA-guided Cas9 nuclease. *Nat. Biotechnol.*, **31**, 827–832.
20. Sapranaukas,R., Gasiunas,G., Fremaux,C., Barrangou,R., Horvath,P. and Siksnys,V. (2011) The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli*. *Nucleic Acids Res.*, **39**, 9275–9282.
21. Guschin,D.Y., Waite,A.J., Katibah,G.E., Miller,J.C., Holmes,M.C. and Rebar,E.J. (2010) A rapid and general assay for monitoring endogenous gene modification. *Methods Mol. Biol.*, **649**, 247–256.
22. Fu,Y., Foden,J.A., Khayter,C., Maeder,M.L., Reyon,D., Joung,J.K. and Sander,J.D. (2013) High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat. Biotechnol.*, **31**, 822–826.
23. 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.
24. Gupta,A., Hall,V.L., Kok,F.O., Shin,M., McNulty,J.C., Lawson,N.D. and Wolfe,S.A. (2013) Targeted chromosomal deletions and inversions in zebrafish. *Genome Res.*, **23**, 1008–1017.
25. Xiao,A., Wang,Z., Hu,Y., Wu,Y., Luo,Z., Yang,Z., Zu,Y., Li,W., Huang,P., Tong,X. *et al.* (2013) Chromosomal deletions and inversions mediated by TALENs and CRISPR/Cas in zebrafish. *Nucleic Acids Res.*, **41**, e141.
26. Gratz,S.J., Cummings,A.M., Nguyen,J.N., Hamm,D.C., Donohue,L.K., Harrison,M.M., Wildonger,J. and O'Connor-Giles,K.M. (2013) Genome engineering of *Drosophila* with the CRISPR RNA-guided Cas9 nuclease. *Genetics*, **194**, 1029–1035.
27. Jinek,M., Chylinski,K., Fonfara,I., Hauer,M., Doudna,J.A. and Charpentier,E. (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*, **337**, 816–821.
28. Yu,A.M. and McVey,M. (2010) Synthesis-dependent microhomology-mediated end joining accounts for multiple types of repair junctions. *Nucleic Acids Res.*, **38**, 5706–5717.
29. Mali,P., Yang,L., Esvelt,K.M., Aach,J., Guell,M., DiCarlo,J.E., Norville,J.E. and Church,G.M. (2013) RNA-guided human genome engineering via Cas9. *Science*, **339**, 823–826.
30. Mojica,F.J., Díez-Villaseñor,C., García-Martínez,J. and Almendros,C. (2009) Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology*, **155**, 733–740.
31. Wang,H., Yang,H., Shivalila,C.S., Dawlaty,M.M., Cheng,A.W., Zhang,F. and Jaenisch,R. (2013) One-Step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell*, **153**, 910–918.