Genes & Diseases (2022) 9, 731–740

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: http://ees.elsevier.com/gendis/default.asp

FULL LENGTH ARTICLE

Nucleolus localization of SpyCas9 affects its stability and interferes with host protein translation in mammalian cells

Renke Tan [a], Wenhao Du [b], Yiyang Liu [a], Xiaoji Cong [c], Meirong Bai [d], Chenxiao Jiang [a], Zengxia Li [a], Minjia Tan [c], Dengke K. Ma [d], Qiang Huang [b,], Wei Jiang [a,***], Yongjun Dang [a,*]**

[a] *Key Laboratory of Metabolism and Molecular Medicine, the Ministry of Education, Department of Biochemistry and Molecular Biology, School of Basic Medical Sciences, Shanghai Medical College, Fudan University, Shanghai 200032, PR China*
[b] *State Key Laboratory of Genetic Engineering, School of Life Sciences, Fudan University, Shanghai 200043, PR China*
[c] *Drug Discovery and Design Center, State Key Laboratory of Drug Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai 201203, PR China*
[d] *Cardiovascular Research Institute, University of California, San Francisco, CA 94158, USA*

Received 14 July 2020; received in revised form 14 September 2020; accepted 15 September 2020
Available online 25 September 2020

KEYWORDS
CRISPR/Cas9;
Gene editing;
Global transcription;
Nucleolus detention signal;
Translation

Abstract The CRISPR/Cas9 system, originally derived from the prokaryotic adaptive immune system, has been developed as efficient genome editing tools. It enables precise gene manipulation on chromosomal DNA through the specific binding of programmable sgRNA to target DNA, and the Cas9 protein, which has endonuclease activity, will cut a double strand break at specific locus. However, Cas9 is a foreign protein in mammalian cells, and the potential risks associated with its introduction into mammalian cells are not fully understood. In this study, we performed pull-down and mass spectrometry (MS) analysis of *Streptococcus pyogenes* Cas9 (SpyCas9) interacting proteins in HEK293T cells and showed that the majority of Cas9-associated proteins identified by MS were localized in the nucleolus. Interestingly, we further discovered that the Cas9 protein contains a sequence encoding a nucleolus detention signal (NoDS). Compared with wild-type (WT) Cas9, NoDS-mutated variants of Cas9 (mCas9) are less

* Corresponding author. Key Laboratory of Metabolism and Molecular Medicine, the Ministry of Education, Department of Biochemistry and Molecular Biology, School of Basic Medical Sciences, Shanghai Medical College, Fudan University, Shanghai 200032, PR China
** Corresponding author. State Key Laboratory of Genetic Engineering, School of Life Sciences, Fudan University, Shanghai 200043, PR China
*** Corresponding author. Key Laboratory of Metabolism and Molecular Medicine, the Ministry of Education, Department of Biochemistry and Molecular Biology, School of Basic Medical Sciences, Shanghai Medical College, Fudan University, Shanghai 200032, PR China
E-mail addresses: huangqiang@fudan.edu.cn (Q. Huang), jiangw@fudan.edu.cn (W. Jiang), yongjundang@fudan.edu.cn (Y. Dang).
Peer review under responsibility of Chongqing Medical University.

https://doi.org/10.1016/j.gendis.2020.09.003

stable, although their gene editing activity is minimally affected. Overexpression of WT Cas9, but not mCas9, causes general effects on transcription and protein translation in the host cell. Overall, identification of NoDS in Cas9 will improve the understanding of Cas9's biological function *in vivo*, and the removal of NoDS in Cas9 may enhance its safety for future clinical use.

Introduction

Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) is derived from the prokaryotic adaptive immune system. It protects bacteria against exogenous genetic elements together with CRISPR-associated proteins (Cas).[1] CRISPR effectors with nuclease activity, such as Cas9 and Cas12, have been optimized as genome editing tools for precise DNA and RNA editing.[2,3] This system is widely used for gene sequencing,[4–6] determining compounds' mode of action,[7,8] and functional gene screening.[9] It is also adopted as a highly potent therapeutic approach for some clinical diseases,[10,11] including genetic diseases,[12] viral infection[13] and Car-T cell therapy used for cancer.[14,15] The Cas9 protein and a 20-nt sgRNA form a complex that targets the genome by complementary base pairing, and the Cas9 protein scissors a double-strand break on the host genome. The mammalian cells fix this genomic double-strand break, eventually generating mutations by means of non-homologous end-joining (NHEJ) or homology directed repair (HDR).[2,16,17] The *in vitro* synthesized sgRNA and its precise binding to targeted DNA via base pairing make this new technique a highly efficient and precise genome editing tool that is widely used in a variety of species for different purposes.

Although the CRISPR/Cas technology is a breakthrough in gene editing field, some problems limit its use in clinical and other settings.[18] These include off-target effects, low efficiency of precise genome editing and immune responses triggered by Cas9 expression.[19,20] The off-target effects of CRISPR/Cas may create random indel mutations at unexpected genomic sites, causing unknown risks to the host cell/organism through the unfaithful recognition by Cas9/sgRNA complex of the unintended sites. Various Cas9 mutants, including SpyCas9-HF1 (N497A/R661A/Q695A/Q926A), eSpyCas9 (1.1) (K848A/K1003A/R1060A), and HypaCas9 (N692A/M694A/Q695A/H698A),[21–23] have been shown to improve CRISPR/Cas9 specificity by reducing off-target effects. Moreover, the efficiency of HDR-dependent precise genome editing can be significantly improved through the inhibition of the NHEJ repairing pathway either genetically or with small molecule inhibitors.[24] Furthermore, a few studies have reported strategies for avoiding or eliminating immune responses caused by CRISPR/Cas9 or sgRNAs during genome editing in human cells.[19] Recently, it was shown that phosphatase treatment of *in vitro* transcribed sgRNAs eliminates the sgRNA-triggered immune response.[25] Nevertheless, there have been no report on Cas9 interacting proteins in host cells, and it is unclear whether Cas9 overexpression can affect the host cell's gene expression or protein translation.

To explore the effect of Cas9 overexpression in host cells, we identified the proteins that interact with Apo-*Spy*Cas9 in a mammalian cell line by tandem affinity purification (TAP). We characterized a Nucleolus Detention Signal (NoDS) located at amino acid residues 220—226 in the Cas9 protein. Mutation of the Cas9 NoDS changed the protein's distribution in the nucleolus, minimally influenced the genome editing efficiency, and significantly reduced protein expression level. Moreover, overexpression of WT Cas9 protein, but not mCas9 protein, led to transcriptome changes and translation inhibition in host cells. These findings supply a possible explanation for the side effects of CRISPR/Cas9 and pave the way for future improvements in Cas9 function.

Materials and methods

Cell culture

HEK293T, HeLa, and U2-OS cells were cultured in DMEM (Gibco, Massachusetts, USA) supplemented with 10% fetal bovine serum and maintained in a humidified 5% CO_2 incubator at 37 °C. Mycoplasma was monitored every two weeks by PCR. All cells were maintained and stored in Professor Yongjun Dang's laboratory at the Shanghai Medical College of Fudan University.

Tandem affinity purification

The *Cas9-NLS* gene was cloned into the pcDNA3.1 plasmid, which contained protein A and streptavidin binding protein (SBP) at the N-terminal of the multiple cloning site. HEK293T cells were seeded at a density of 2.2×10^7 cells per 15 cm dish in DMEM (Gibco) with 10% FBS (Gibco) at 24 h before transfection. Transfection mixture was produced by incubating 15—20 μg of plasmid with 37.5—50 μl PEI in DMEM with 10% FBS. One day after transfection, cells were harvested with lysis buffer that contained 50 mM Tris—HCl (pH 7.5), 125 mM NaCl, 0.2% NP-40, 5% glycerol, 1.5 mM $MgCl_2$, protease inhibitors (Sigma, Missouri, USA) and phosphatase inhibitor cocktail (Sigma, Missouri, USA), and lysated by sonication. The supernatant was incubated with 30 μl IgG beads (Sigma, Missouri, USA) for 2 h at 4 °C, and the IgG beads were then incubated with 1.5 μl AcTEV protease (10 U/μl, Invitrogen, Massachusetts, USA) per tube at 4 °C overnight. Eighty μl of streptavidin beads (Invitrogen, Massachusetts, USA) per tube were equilibrated with SBP buffer that contained 10 mM Tris—HCl (pH 7.5), 100 mM NaCl, and 0.2% NP-40. IgG beads treated with AcTEV were

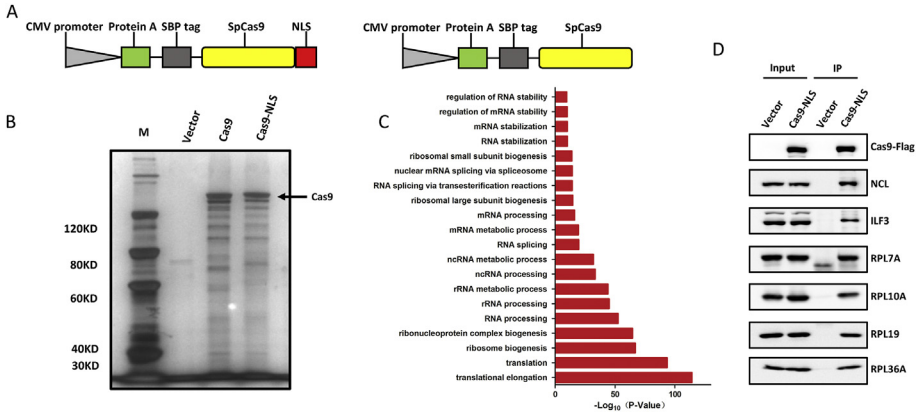

Figure 1 Identification of Cas9 interacting proteins in HEK293T cells. **(A)** A scheme for the interacting protein pull-down assay. A CMV promoter was used to initiate the overexpression of Cas9 protein, and protein A and SBP tags were cloned to the N terminus of the Cas9 protein. A TEV site was located between the protein A tag and the SBP tag for TEV digestion. An additional NLS tag was cloned to the C terminus of the Cas9 protein in one design to promote the protein's nuclear localization. **(B)** Silver staining of the interacting protein pattern. Cas9 and Cas9 NLS indicated the HEK293T cells transfected with ApoCas9 and NLS tagged ApoCas9. Vector showed the pull-down assay background without Cas9 overexpression. Cas9 was indicated by black arrow. **(C)** GO analysis of Cas9 interacting proteins identified by mass spectrometry. **(D)** Western blot verification of the Cas9 interacting proteins. NCL, ILF3, RPL7A, RPL10A, RPL19 and RPL36A were detected in both cell lysate and pull down components; Cas9 protein was detected by Flag.

spun down, the supernatant was added to the streptavidin beads, and the mixture was incubated at 4 °C for 4 h. Streptavidin beads were finally spun down, boiled with 40 μl 2 × SDS sample buffer, and used for mass spectrometry (MS) analysis.

Mass spectrometry

LC—MS/MS analysis of tryptic peptides was performed on an LTQ linear ion trap mass spectrometer (ThermoFisher, Massachusetts, USA) (in the Cas9 interactome experiment) or an Obitrap Fusion mass spectrometer (ThermoFisher, Massachusetts, USA). The LC—MS/MS data were analyzed using Mascot (v2.3, Matrix Science Ltd., London, UK). Peak lists were generated by Proteome Discoverer software (v1.4, ThermoFisher, Massachusetts, USA). Precursor mass tolerance for Mascot analysis was set at ± 10 ppm, and fragment mass tolerance was set at ± 0.5 Da. The Mascot emPAI (exponentially modified protein abundance index) score was used for label-free protein quantification as previously described.[47]

Immunofluorescence

U2-OS cells were seeded on glass slides in a 24-well plate (5 × 103 cells/well) and transfected with GFP-fused Cas9 plasmid using FuGENE (Promega, Wisconsin, USA) after cell adherence. At 72 h after transfection, the slides were gently washed with PBS and fixed in 4% paraformaldehyde, then permeabilized in 0.5% Triton X-100 to perforate nuclear membranes. After blocking with 4% BSA, the samples were incubated with anti-NCL antibody (1:1000) (Proteintech, Illinois, USA) in PBS that contained 4% BSA at 4 °C overnight or at 37 °C for 1 h. They were then stained with Cy3-conjugated anti-rabbit secondary antibody (Jackson Immunoresearch, Pennsylvania, USA) at 37 °C for 30 min. Following extensive washing, the slides were coated with mounting medium (ThermoFisher). Cell fluorescence was visualized under a fluorescence microscope (Olympus Corporation, Tokyo, Japan).

EGFP disruption assay

HeLa EGFP cells harboring a single integrated copy of an LC3—EGFP fusion gene were cultured as described. For transfection, 1.5 × 10^6 cells were cultured in a 6-well plate until cell adherence, then 2 μg of PX330 plasmid and 0.2 μg of RFP expression plasmid were co-transfected with sgRNA targeting EGFP. Cells were analyzed 2 days post-transfection by flow cytometry (Accuri C6, BD, California, USA).

Site-directed mutagenesis

Five amino acids of the NoDS domain were mutated to Ala by four rounds of PCR (KOD, TOYOBO, Osaka, Japan) using the primers listed in Table 1. The PCR products were purified, digested with DpnI restriction enzyme to cleave the methylated template DNA, and then transformed into *E. coli* XL10 strain and plated on LB agar medium that contained 100 μg/ml ampicillin. The mutant sequences was confirmed by Sanger sequencing.

Cas9 protein expression and purification

The protein expression and purification procedures followed those described in a previously published paper.[48] Briefly, the Cas9 encoding sequence was cloned into the pET21a protein expression vector, which was then transferred into the Rosetta (DE3) strain and cultured in LB medium with ampicillin and chloramphenicol. 0.1 mM IPTG was added to induce protein expression, and the protein was expressed at 16 °C for 20 h. Cells were then harvested

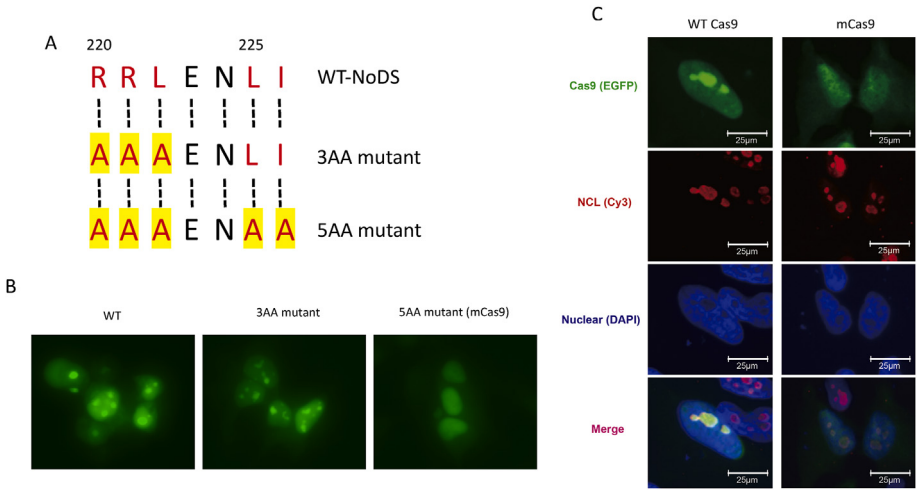

Figure 2 Cas9 contains a nucleolus detention signal and targets nucleolus. **(A)** NoDS in the Cas9 coding sequence. A scheme showing NoDS in Cas9 amino acid sequence, as well as mutations introduced to disrupt NoDS. '3 AA mutant' harbors R to A mutation in R220 and R221, and L to A mutation in L222; '5AA mutant' harbors mutations with R220, R221, L222, L225 and I226 all mutant to A. **(B)** Fluorescent images of cellular distribution of the WT Cas9 protein, 3AA mutant and 5AA mutant in U2OS cells. **(C)** Co-localization of Cas9 (green), NLS (red), and Nucleus (blue) in U2OS cells was detected by immunofluorescence. Scale bar = 50 μm.

and lysed at 10,000 psi in lysis buffer at 4 °C (20 mM HEPES, pH 7.5, 500 mM KCl, 1 mM PMSF). Following centrifugation, the supernatant was incubated with Ni-NTA agarose (Qiagen, Hilden, Germany) for 1.5 h at 4 °C. The Ni-NTA agarose was then washed in a washing buffer that contained 30 mM imidazole and eluted with an elution buffer that contained 100 mM imidazole. The protein was finally concentrated using a 100 kDa filter and stored at −80 °C for further use.

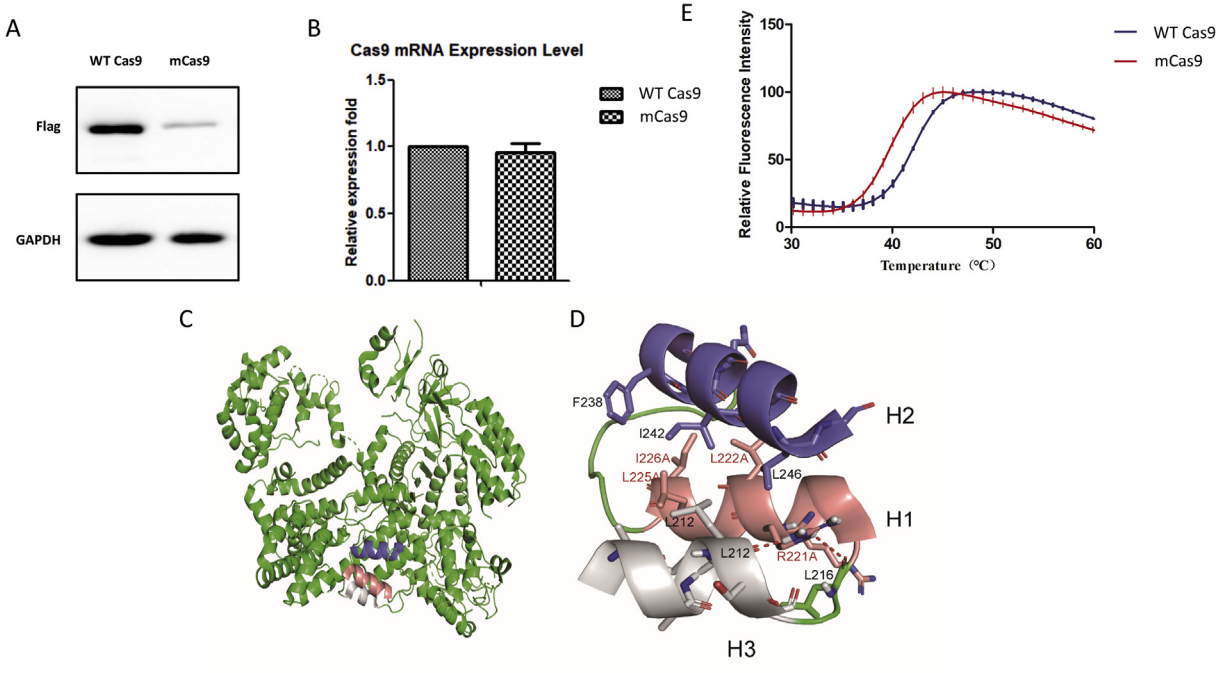

Figure 3 NoDS mutated Cas9 display decreased stability in human cells. **(A)** Western blot detection of Cas9 protein levels in HEK293T cells by Flag tag. GAPDH was used as internal control. **(B)** Transcription level of transfected Cas9 detected by qPCR ($N = 3$). The ampicillin resistance gene was used as internal control. Relative expression of mutant Cas9 is normalized to that of WT Cas9. **(C)** Structure of the Cas9 protein downloaded from the protein data bank (4CMP). The NoDS-containing helix is highlighted in pink, and the adjacent helixes are highlighted in purple and white. **(D)** Close-up of the NoDS-containing helix and the adjacent helix; the amino acid residues involved in hydrophobic interactions are highlighted in red. **(E)** Thermal shift assay of purified WT Cas9 and mCas9. Purified Cas9 and mCas9 protein were mixed with SYPRO Orange in Tris buffer. Changes of SYPRO Orange fluorescence intensity were collected while the protein was heated up. Melting temperature (Tm) was calculated with Boltzmann fitting method. The Tm of WT Cas9 protein was 42 °C and Tm of mCas9 was 39 °C.

Measurement of protein synthesis

HEK293T cells were grown in 6-well plates and transfected with Cas9, mCas9, or vector. Cells were collected 60 hours after transfection. The medium was replaced with methionine-free DMEM with L-AHA and incubated for 30 min, and the cell lysates were collected and subjected to a Click-iT assay to measure the synthesis rate of nascent peptides. Click-iT protein synthesis was carried out following the manufacturer's instructions (Click-iT™ AHA Alexa Fluor™ 488 Protein Synthesis HCS Assay, ThermoFisher, Massachusetts, USA). Nascent peptides were scanned by Typhoon, and total protein was evaluated by Coomassie blue staining.

DNA cleavage assay

The DNA cleavage assay was performed as described in a previous study.[48]

RNAseq and analysis

HEK293T cells were transfected by Cas9 or mCas9 for 48 h and then total RNA was extracted by Trizol. mRNA was purified using polyA magnetic beads. Raw data were aligned to the human genome using a customized pipeline, and annotation and cluster analysis were performed using DAVID online tools (https://david.ncifcrf.gov/).

Western Blot

Cells were washed twice with PBS before harvesting and lysed in lysis buffer (500 mM NaCl, 1% NP40, 1% sodium deoxycholate, 0.1% SDS, 50 mM Tris, pH 7.0). Total protein concentrations were determined with BCA kit (ThermoFisher). The lysate was mixed with sample buffer and boiled at 100 °C for 10 min. Ten microgram of total protein was loaded and separated by SDS-PAGE gel, and then transferred to nitrocellulose membranes (Millipore, Massachusetts, USA). After blocking, the membranes were incubated overnight at 4 °C with antibodies specific to different proteins, and then with secondary antibodies conjugated with horseradish peroxidase. Immunoreactivity was visualized by enhanced chemiluminescence (GE Healthcare, Illinois, USA). Proteins were detected using the following antibodies: Flag (Sigma, F1804, Missouri, USA), NCL (Proteintech, 10556-1-AP, Illinois, USA), ILF3 (Abcam, ab225626 Cambridge, UK), RPL7A (Santa Cruz, sc-79043, Texas, USA), RPL10A (Santa Cruz, sc-100827), RPL19 (Santa Cruz, sc-100830), RPL36A (Santa Cruz, sc-100831), and GAPDH (Sigma, g8795).

Thermal shift

Protein thermal shift assays were performed as described in a previous study.[49] In each well of assay plate, 5 μmol/L Cas9 or mCas9 protein and 5 × SYPRO Orange (Invitrogen,

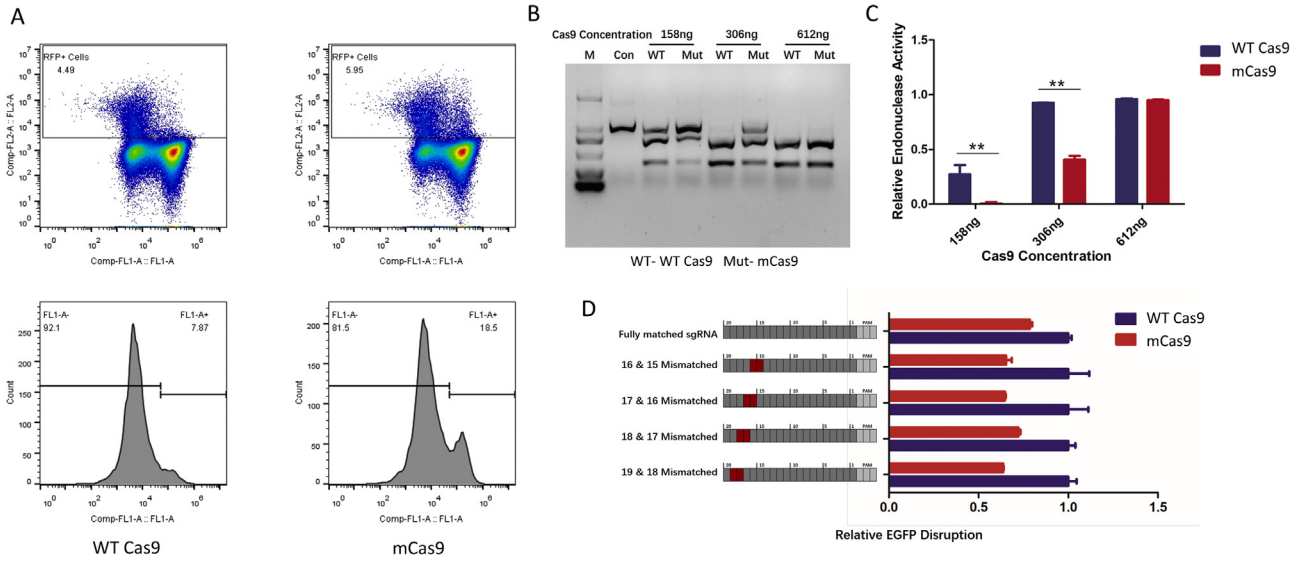

Figure 4 NoDS-mutations minimally affect Cas9's genome editing efficiency. **(A)** A plasmid expressing RFP was co-transfected with lentiCRISPR VII in a 1:10 ratio (WT or NoDS-mutated Cas9). GFP positive cells was analyzed by flow cytometry 60 h after transfection. Representative figure of three independent experiment is shown. **(B)** The endonuclease activity of Cas9 detected by purified Cas9 *in vitro*. Different concentrations of purified Cas9 protein were incubated with DNA fragments and sgRNA at 37 °C for 30 min, and then subjected to agarose gel electrophoresis. **(C)** Quantification of the agarose gel electrophoresis data. The grey values of the bands in (B) were calculated using ImageJ software, and relative endonuclease activity was calculated using the formula: (cut1+cut2)/uncut. $*P < 0.1$, $**P < 0.05$, $***P < 0.01$, Student's *t*-test. **(D)** Characterization of mismatch tolerance of WT Cas9 and mCas9. WT Cas9 and mCas9 were assessed by EGFP disruption assay when programmed with fully matched or partially mismatched sgRNAs. PAM sequence was marked by light grey rectangle; matched base pairs on sgRNAs were marked as dark grey rectangle; mismatched base pairs on the sgRNA was marked as red rectangle. Error bars represent SD. $N = 3$. Relative EGFP disruption was normalized by EGFP disruption level of fully matched sgRNA paired with WT Cas9 or mCas9 separately.

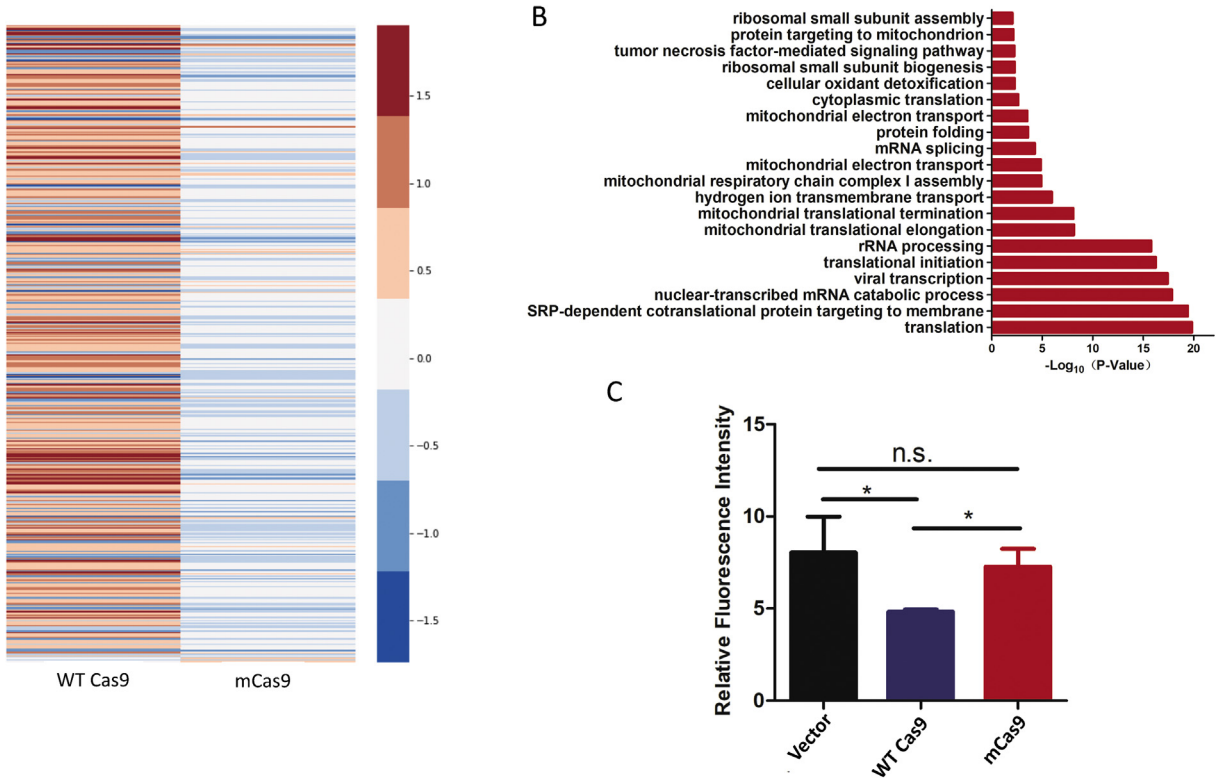

Figure 5 Cas9 but not NoDS-mutated Cas9 affects global protein translation in human cells. **(A)** Heatmap of transcriptional changes triggered by Cas9 or mCas9 based on RNA-seq data of HEK293T cell transfected with WT Cas9 or mCas9. Genes with expression level change greater than 1.5 fold than untreated cell were shown in heatmap. **(B)** Annotation and GO analysis of differentially expressed genes identified by RNA-seq. **(C)** The grey values associated with the Click-iT assay were calculated using ImageJ software, and the relative fluorescence intensities were calculated as the grey value of the labeled lane divided by the grey value of the coomassie-blue-stained lane. ns $P > 0.05$, $^*P < 0.05$, two-sided Student's *t*-test.

Massachusetts, USA) were mixed in a total volume of 20 μL in Tris buffer. The assay plate was heated from 25 °C to 95 °C at a ramp rate of 0.05 °C/s. The changes in SYPRO Orange fluorescence intensity were monitored and collected for further analysis. The melting temperature (Tm) was calculated with the Boltzmann fitting method in Protein Thermal Shift Software v1.2. The experiments were performed three times independently.

qPCR

Total RNA was extracted from cells using Trizol (ThermoFisher). For each sample, 1 μg total RNA was used for reverse transcription with reverse transcriptase (TaKaRa, Kyoto, Japan). qRT-PCR analysis was performed on the LightCycler 480 II PCR platform (Roche, Basel, Switzerland) using the SYBR Green Supermix kit (Takara). The PCR condition was as follows: 1 min of hot start at 95 °C, followed by 45 cycles of 95 °C for 10 s, 60 °C for 10 s, and 72 °C for 10 s. Blank controls with no cDNA templates were used to rule out contamination. The ampicillin resistance gene on the plasmid was used as the internal control. The specificity of the PCR product was confirmed by melting curve analysis and gel electrophoresis. Gene expression levels were normalized to the GAPDH housekeeping gene. Relative expression levels were calculated by the ΔΔCt method. The primers used were as follows:

AmpR-F: TTCTCAGAATGACTTGGTTG.
AmpR-R: GATCAAGGCGAGTTACATGA.
Cas9-F: GACATCGGCACCAACTCTG.
Cas9-R: TCCGGTTCTTCCGTCTGGT.

Results

Identification of Cas9 interacting proteins in HEK293T cells

To investigate whether bacterial-derived Cas9 has interaction partners in human cells, we ectopically overexpressed ApoCas9 from *Streptococcus pyogenes* species, tagged with protein A and streptavidin binding protein in HEK293T cells. To determine whether the nuclear localization of Cas9 protein contributes to differences in its interaction partners, we compared Cas9 with or without a C-terminal tagged nuclear localization sequence (NLS), using the empty vector as control (Fig. 1A). After TAP, the complex of Cas9 and its interaction partners were separated by SDS-PAGE. Multiple interacting proteins were pulled down by Cas9-NLS or Cas9, and the empty vector produced a clean background. All bands were cut out and identified by MS (Fig. 1A, B). GO analysis in the aspect of biological processing of all proteins identified by MS (Table S1) indicated that the majority of the Cas9 interacting proteins were

associated with the biological processes of protein translation, ribosome biogenesis, and RNA processing (Fig. 1C). Similarly, GO enrichment from other aspects agrees with the biological processes that enriched to be associated with Cas9 protein. For example, for which in the aspect of cellular components, ribonucleoprotein complex and ribosome were enriched (Fig. S1A); the enrichment in the aspect of molecular function has enriched functions related to structural constituent of ribosome and RNA binding (Fig. S1B). We did not observe a significant difference in interaction protein lists for Cas9 protein with or without NLS, although Cas9 without NLS clearly displayed stronger bands on the SDS-PAGE gel.

We then performed Western blot to verify the interaction proteins identified by MS. Due to the similarity of the MS results for Cas9 and Cas9-NLS, we focused only on Cas9-NLS in subsequent analyses. As indicated by the GO analysis, many Cas9 interaction partners were ribosome proteins. Interestingly, ribosome biogenesis occurred in the nucleolus, and some nucleolus localized proteins were also detected by MS. We chose ribosome proteins RPL7a, RPL10a, RPL19, RPL36a, the nucleolus localized proteins ILF3 and nucleolin (NCL) for validation. As shown in Figure 1D, all proteins could be detected in the Cas9-NLS associated complex but not in the control. The finding that Cas9 interacts with NCL and ribosomal proteins, which are biomarkers of the nucleolus or assembled in the nucleolus, suggests that Cas9 itself may be localized in the nucleolus.

Cas9 contains a nucleolus detention signal and targets the nucleolus

We thus looked for Cas9 protein localization in previous published papers, and found Cas9 accumulates as some 'bright spots' in nucleus,[2,26] which is highly possible to be nucleolus. When we inspected the protein sequence of SpyCas9, we found amino acids 220—222 (Arg-Arg-Leu) and 225—226 (Leu—Ile) together met the characteristics of a NoDS sequence (Fig. 2A). It has been reported that protein with a NoDS can be trapped in the nucleolus,[27—29] because NoDS consists of highly charged arginine or lysine residues and it can directly bind to non-coding RNAs in the nucleolus.[28] We did not find a classic NoDS sequence in Cas9 proteins from other species, such as SaCas9 or NmeCas9, or in other effector proteins, such as Cpf1, C2c1, or C2c2. But we did find from literature that several type II CRISPR effector proteins are likely to accumulate in the nucleolus, such as SaCas9[30] and Cas13d.[31] To validate this NoDS-like sequence in Cas9 and its effects on Cas9 protein localization, we made two Cas9 mutations, one with amino acids 220—222 mutated to alanine (3 AA mutant) and the other with amino acids 220—222 and 225—226 mutated to alanine (5 AA mutant, mCas9). We fused EGFP to C-terminal of Cas9 or the Cas9 mutants to help track Cas9 localization. As shown in Figure 2B, both WT Cas9 and the 3 AA mutant Cas9 accumulated in nucleus, observed as bright spots, whereas the 5 AA mutant dispersed evenly inside the nucleus. This result indicated that R220, R221, L222, L225, and I226 in SpyCas9 may serve as a NoDS that altered the localization of Cas9 protein in the nucleus. To further understand whether NoDS target Cas9 protein to nucleolus, C terminus EGFP fused WT or 5 AA mutant Cas9 (mCas9) was transfected into U2OS cells. Then the intracellular localization of Cas9-EGFP was monitored by fluorescence microscopy. The position of the nucleolus was determined by immunostaining of NCL, which is a biomarker of the nucleolus. As shown in Figure 2C, the majority of WT Cas9-EGFP protein co-localized with NCL, but the mCas9 protein did not. These results demonstrated that NoDS in Cas9 determines its nucleolus localization.

NoDS mutated Cas9 display decreased stability in human cells

While performing the aforementioned experiments, we accidently found out that the fluorescence of mCas9-EGFP was much weaker than that of WT Cas9-EGFP (Fig. 2B, C). By measuring their transcription and protein levels, we found out that disrupting NoDS in Cas9 caused a significant decrease in mCas9 protein accumulation (Fig. 3A) but not in its mRNA level (Fig. 3B). We speculated that the NoDS mutation may have altered the stability of the Cas9 protein, contributing to the lower protein accumulation. Indeed, as we looked into the structure of the Cas9 protein, we noticed that all the amino acids mutated in the Cas9 protein appeared to be in one α-helix (highlighted in pink in Fig. 3C). The mutation of these highly hydrophobic amino acids affects their hydrophobic interactions with other hydrophobic amino acids in adjacent helices (highlighted in purple and white in Fig. 3C). For example, as shown in Figure 3D, the hydrophobic interactions between L221 and L246, L225 and L212, and I226 and I242 were impaired by the L to A or I to A mutation, which was consistent with the idea that the stability of the mCas9 protein was lower than that of the WT Cas9 protein. We therefore performed a thermal shift assay with purified WT Cas9 and mCas9 protein. The Tm value of the protein shifted from 42 °C to 39 °C, suggesting a decrease in the stability of mCas9 (Fig. 3E). In summary, these results suggested that the

Table 1 Primers for site-directed mutagenesis

R220A-F	GCCAGACTGAGCAAGAGCGCACGGCTGGAAAATCTGATCGCC
R220A-R	GATCAGATTTTCCAGCCGTGCGCTCTTGCTCAGTCTGGCAGA
R221A-F	AGACTGAGCAAGAGCGCAGCGCTGGAAAATCTGATCGCCCAG
R221A-R	GGCGATCAGATTTTCCAGCGCTGCGCTCTTGCTCAGTCTGGC
L222A-F	CTGAGCAAGAGCGCAGCGGCGGAAAATCTGATCGCCCAGCTG
L222A-R	CTGGGCGATCAGATTTTCCGCCGCTGCGCTCTTGCTCAGTCT
L225A, I226A-F	GCGGCGGAAAATGCGGCCGCCCAGCTGCCCGGCGAGAAGAAG
L225A, I226A-R	GGGCAGCTGGGCGGCCGCATTTTCCGCCGCTGCGCTCTTGCT

removal of the NoDS from Cas9 leads to instability of Cas9 protein compared with the WT Cas9 protein and thereby lowers the protein level in the host cell.

NoDS-mutations minimally affect Cas9's genome editing efficiency

In order to test the function of mCas9, especially its genome editing efficiency, we employed EGFP disruption assay.[10] To make comparisons at the same transfection level, we co-delivered the RFP plasmid with the sgEGFP-Cas9 plasmid at a ratio of 1:10 and analyzed EGFP disruption in RFP positive cells. The result showed that the genome editing efficiency of mCas9 (80.86 $\pm$ 1.64%) was only modestly reduced compared to WT Cas9 (90.83 $\pm$ 1.55%) (Fig. 4A). Both decreased protein expression and reduced endonuclease activity of mCas9 may contribute to the lower editing efficiency observed in host cells. To investigate how the mutation causes a decrease in Cas9 endonuclease activity with the same amount of protein, we employed an *in vitro* assay to examine the endonuclease activity of Cas9 and mCas9 (Fig. 4B, C).[32] We tested WT and mCas9 endonuclease activity at three different concentrations. When Cas9 was present at limited amounts in the reaction system (158 ng), mCas9 exhibited little endonuclease activity, and WT Cas9 had about one-quarter of the endonuclease activity as of the Cas9 alone control. As we added more Cas9 to system (306 ng), mCas9 endonuclease activity increased to about 50%, and WT Cas9 endonuclease activity was almost 100%. When we added sufficient Cas9 to system (612 ng), both mCas9 and WT Cas9 exhibited approximately 100% endonuclease activity. These results suggested that mCas9 has lower endonuclease activity than Cas9 and that both activities depend on the concentration of the enzyme (Fig. 4B, C). In addition to genome editing efficiency, we also tested the mismatch sgRNA tolerance of both WT Cas9 and the mCas9 by introducing mismatch mutations to different positions on the sgRNA.[33] There was no significant differences in mismatch tolerance between mCas9 and WT Cas9 (Fig. 4D).

Cas9 but not NoDS-mutated Cas9 affects global protein translation in human cells

To further investigate whether ApoCas9 overexpression may cause global changes in transcription or translation in mammalian cells, we performed RNAseq analysis and measured total protein synthesis. As shown in Figure 5A, transfection of HEK293T cells with WT Cas9 plasmid, in comparison to transfection with empty vector, was associated with the differential expression (>1.5-fold difference) of 498 genes—389 upregulated and 109 downregulated. However, transfection with mCas9 plasmid resulted in differential expression of only seven genes—three upregulated and four downregulated. We performed gene annotation and cluster analysis of the differentially expressed genes in response to WT Cas9 transfection. The result suggested that the differentially expressed genes were involved in biological processes including translation, nuclear-transcribed mRNA catabolic process, viral transcription, rRNA processing, and translational initiation (Fig. 5B). These results were consistent with the MS data, which also detected translation-related biological processes. The RNA-seq and cluster analysis further revealed that the expression of genes associated with mitochondrial-related translation, respiratory chain, and electron transport were also affected by the presence of WT Cas9. Few genes that were differentially expressed in response to mCas9 transfection and were not enriched in any specific biological processes.

To examine the effect of Cas9 expression on general protein translation, we metabolically labeled nascent peptides to compare the protein translation efficiency of HEK293T cells expressing WT Cas9 or mCas9 protein[34]. The result revealed that there was a significant decrease in global protein translation in WT Cas9-transfected cells but not in mCas9-transfected cells (Fig. 5C, Fig. S2). This result suggested that the expression of mCas9 protein has much less impact on general protein translation than that of the WT Cas9 protein in human cells.

Discussion

In this study, we focused on the effects of expression of *Streptococcus pyogenes* Cas9 in mammalian cells. We first demonstrated that Cas9 physically interacts with ribosomal proteins and nucleolus-localized proteins such as ILF3 and NCL (Fig. 1C, D). The nucleolus is involved in critical biological processes, including ribosome biogenesis, cell cycle regulation,[35] stress perception,[36] and viral replication. Recently, the nucleolus was found to have chaperone-like properties and it can promote nuclear protein maintenance under stress.[37] We found out that Cas9 was localized to the nucleolus in a NoDS sequence-dependent fashion (Fig. 2B, C). Mutation of the NoDS in Cas9 not only disrupted its nucleolar location but also reduced its steady-state protein level in human cells (Fig. 2C, 3A, 3B) and the thermal stability *in vitro* (Fig. 3E). Importantly, mCas9 remained as a highly robust gene editor, and it caused much fewer transcriptional and translational changes in comparison to WT Cas9 in human cells (Fig. 4A, 5A).

In addition, Cas9 was recently reported to activate the P53 pathway, resulting in the enrichment of mutant P53.[38] The regulation of P53 and the sequestering of MDM2 by the nucleolus have been widely studied.[39–41] Our findings of Cas9 localization in the nucleolus may play an important role in the enrichment of mutant P53 and therefore warrant further study.

Furthermore, Cas9 is a foreign protein introduced into eukaryotic cells. Understanding its behavior inside human cells is crucial for both CRISPR-based basic research and clinical applications such as the treatment of HIV,[42] cystic fibrosis,[43,44] and hereditary tyrosinemia.[12] The correct dosage and properly control of CRISPR/Cas9 are important for its safety use. Small molecule inhibitors[45] and light[46] have been discussed as means to enable precise control of CRISPR/Cas9 in host cells to avoid potential alteration caused by long term exsitance of Cas9 *in vitro*. Our findings clearly show that long-term and high-dose maintenance of WT Cas9 in human cells is not preferable from the perspective of host cell health. Importantly, our study also

identified that mCas9 variant, which bears 5 amino acids NoDS mutation, is evenly dispersed inside nucleus. This mCas9 variant accumulates less in human cells and loses the ability to interfere with cellular protein translation, but remarkably maintains high genome editing capability. Therefore, mCas9 can serve as a safer version of high-efficiency gene editing tool.

Author contributions

Renke Tan, Zengxia Li, Wei Jiang and Yongjun Dang designed the project; Renke Tan, Yiyang Liu and Chenxiao Jiang performed experiments and collected data; Xiaoji Cong and Minjia Tan performed Mass Spectrum Assay; Wenhao Du and Qiang Huang purified proteins and performed *in vitro* DNA cleavage assay; Meirong Bai and Dengke K. Ma provided technical support; Renke Tan, Wei Jiang and Yongjun Dang wrote the manuscript.

Conflict of interests

All authors declare no conflicts of interest.

Funding

This work was supported by the National Key Research and Development Program of China (No. 2018YFC0310900), the National Natural Science Foundation of China grants (No. 31270830, 21572038, 21877016 to Y.D, No. 31671386 to Q.H, 81972621 to W.J), the Development Fund for Shanghai Talents, Fund of State Key Laboratory of Bioorganic and Natural Products Chemistry, Fund of State Key Laboratory of Drug Research, Chinese Academy of Science (No. SIMM1601KF-08), and Fudan University Graduate Student Research Grant.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.gendis.2020.09.003.

References

1. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*. 2012; 337(6096):816—821.
2. Cong L, Ran FA, Cox D, et al. Multiplex genome engineering using CRISPR/Cas systems. *Science*. 2013;339(6121):819—823.
3. Knott GJ, Doudna JA. CRISPR-Cas guides the future of genetic engineering. *Science*. 2018;361(6405):866—869.
4. Kellner MJ, Koob JG, Gootenberg JS, Abudayyeh OO, Zhang F. SHERLOCK: nucleic acid detection with CRISPR nucleases. *Nat Protoc*. 2019;14(10):2986—3012.
5. Replogle JM, Norman TM, Xu A, et al. Combinatorial single-cell CRISPR screens by direct guide RNA capture and targeted sequencing. *Nat Biotechnol*. 2020;38(8):954—961.
6. Jaitin DA, Weiner A, Yofe I, et al. Dissecting immune circuits by linking CRISPR-pooled screens with single-cell RNA-seq. *Cell*. 2016;167(7):1883—1896. e15.
7. Jost M, Weissman JS. CRISPR approaches to small molecule target identification. *ACS Chem Biol*. 2018;13(2):366—375.
8. Kampmann M. Elucidating drug targets and mechanisms of action by genetic screens in mammalian cells. *Commun (Camb)*. 2017;53(53):7162—7167.
9. Swiech L, Heidenreich M, Banerjee A, et al. In vivo interrogation of gene function in the mammalian brain using CRISPR-Cas9. *Nat Biotechnol*. 2015;33(1):102—106.
10. Lin A, Giuliano CJ, Sayles NM, Sheltzer JM. CRISPR/Cas9 mutagenesis invalidates a putative cancer dependency targeted in on-going clinical trials. *Elife*. 2017;6:e24179.
11. Dominguez AA, Lim WA, Qi LS. Beyond editing: repurposing CRISPR—Cas9 for precision genome regulation and interrogation. *Nat Rev Mol Cell Biol*. 2015;17(1):5—15.
12. Yin H, Xue W, Chen S, et al. Genome editing with Cas9 in adult mice corrects a disease mutation and phenotype. *Nat Biotechnol*. 2014;32(6):551—553.
13. Kennedy EM, Bassit LC, Mueller H, et al. Suppression of hepatitis B virus DNA accumulation in chronically infected cells using a bacterial CRISPR/Cas RNA-guided DNA endonuclease. *Virology*. 2015;476:196—205.
14. Eyquem J, Mansilla-Soto J, Giavridis T, et al. Targeting a CAR to the TRAC locus with CRISPR/Cas9 enhances tumour rejection. *Nature*. 2017;543(7643):113—117.
15. Liu X, Zhang Y, Cheng C, et al. CRISPR-Cas9-mediated multiplex gene editing in CAR-T cells. *Cell Res*. 2017;27(1):154—157.
16. Mali P, Yang L, Esvelt KM, et al. RNA-guided human genome engineering via Cas9. *Science*. 2013;339(6121):823—826.
17. Ran FA, Cong L, Yan WX, et al. In vivo genome editing using Staphylococcus aureus Cas9. *Nature*. 2015;520(7546):186—191.
18. Kang XJ, Caparas CIN, Soh BS, Fan Y. Addressing challenges in the clinical applications associated with CRISPR/Cas9 technology and ethical questions to prevent its misuse. *Protein Cell*. 2017;8(11):791—795.
19. Charlesworth CT, Deshpande PS, Dever DP, et al. Identification of preexisting adaptive immunity to Cas9 proteins in humans. *Nat Med*. 2019;25(2):249—254.
20. Kim S, Koo T, Jee HG, et al. CRISPR RNAs trigger innate immune responses in human cells. *Genome Res*. 2018;28(3):367—373.
21. Kleinstiver BP, Pattanayak V, Prew MS, et al. High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature*. 2016;529(7587):490—495.
22. Slaymaker IM, Gao L, Zetsche B, Scott DA, Yan WX, Zhang F. Rationally engineered Cas9 nucleases with improved specificity. *Science*. 2016;351(6268):84—88.
23. Chen JS, Dagdas YS, Kleinstiver BP, et al. Enhanced proofreading governs CRISPR-Cas9 targeting accuracy. *Nature*. 2017; 550(7676):407—410.
24. Yu C, Liu Y, Ma T, et al. Small molecules enhance CRISPR genome editing in pluripotent stem cells. *Cell Stem Cell*. 2015; 16(2):142—147.
25. Wienert B, Shin J, Zelin E, Pestal K, Corn JE. In vitro-transcribed guide RNAs trigger an innate immune response via the RIG-I pathway. *PLoS Biol*. 2018;16(7):e2005840.
26. Chen B, Gilbert LA, Cimini BA, et al. Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell*. 2013;155(7):1479—1491.
27. Hoellerich E, Dunagan C, Maring D, et al. Nucleolar localization of SmMAK16 protein from Schistosoma mansoni is regulated by three distinct signals that function independent of pH or phosphorylation status. *Mol Biochem Parasitol*. 2014;193(1):9—16.
28. Audas TE, Jacob MD, Lee S. The nucleolar detention pathway. *Cell Cycle*. 2012;11(11):2059—2062.

29. Audas TE, Jacob MD, Lee S. Immobilization of proteins in the nucleolus by ribosomal intergenic spacer noncoding RNA. *Mol Cell*. 2012;45(2):147—157.
30. Sun H, Fu S, Cui S, et al. Development of a CRISPR-SaCas9 system for projection- and function-specific gene editing in the rat brain. *Sci Adv*. 2020;6(12), eaay6687.
31. Konermann S, Lotfy P, Brideau NJ, Oki J, Shokhirev MN, Hsu PD. Transcriptome engineering with RNA-targeting type VI-D CRISPR effectors. *Cell*. 2018;173(3):665—676. e614.
32. Fu Y, Foden JA, Khayter C, et al. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat Biotechnol*. 2013;31(9):822.
33. Kleinstiver BP, Pattanayak V, Prew MS, et al. High-fidelity CRISPR—Cas9 nucleases with no detectable genome-wide off-target effects. *Nature*. 2016;529(7587):490.
34. Xie Q, Weng X, Lu L, Lin Z, Xu X, Fu C. A sensitive fluorescent sensor for quantification of alpha-fetoprotein based on immunosorbent assay and click chemistry. *Biosens Bioelectron*. 2016;77:46—50.
35. Raska I, Shaw PJ, Cmarko D. Structure and function of the nucleolus in the spotlight. *Curr Opin Cell Biol*. 2006;18(3):325—334.
36. Olson MOJ. Sensing cellular stress: another new function for the nucleolus. *Sci STKE*. 2004;2004(224):pe10.
37. Frottin F, Schueder F, Tiwary S, et al. The nucleolus functions as a phase-separated protein quality control compartment. *Science*. 2019;365(6451):342—347.
38. Enache OM, Rendo V, Abdusamad M, et al. Cas9 activates the p53 pathway and selects for p53-inactivating mutations. *Nat Genet*. 2020;52(7):662—668.
39. Bernardi R, Scaglioni PP, Bergmann S, Horn HF, Vousden KH, Pandolfi PP. PML regulates p53 stability by sequestering Mdm2 to the nucleolus. *Nat Cell Biol*. 2004;6(7):665—672.
40. Rubbi CP, Milner J. Disruption of the nucleolus mediates stabilization of p53 in response to DNA damage and other stresses. *EMBO J*. 2003;22(22):6068—6077.
41. Boyd MT, Vlatkovic N, Rubbi CP. The nucleolus directly regulates p53 export and degradation. *J Cell Biol*. 2011;194(5):689—703.
42. Ye L, Wang J, Beyer AI, et al. Seamless modification of wild-type induced pluripotent stem cells to the natural CCR5Δ32 mutation confers resistance to HIV infection. *Proc Natl Acad Sci USA*. 2014;111(26):9591—9596.
43. Schwank G, Koo BK, Sasselli V, et al. Functional repair of CFTR by CRISPR/Cas9 in intestinal stem cell organoids of cystic fibrosis patients. *Cell Stem Cell*. 2013;13(6):653—658.
44. Geurts MH, de Poel E, Amatngalim GD, et al. CRISPR-based adenine editors correct nonsense mutations in a cystic fibrosis organoid biobank. *Cell Stem Cell*. 2020;26(4):503—510.e7.
45. Maji B, Gangopadhyay SA, Lee M, et al. A high-throughput platform to identify small-molecule inhibitors of CRISPR-Cas9. *Cell*. 2019;177(4):1067—1079.e19.
46. Gangopadhyay SA, Cox KJ, Manna D, et al. Precision control of CRISPR-Cas9 using small molecules and light. *Biochemistry*. 2019;58(4):234—244.
47. Li Z, Cheng Z, Raghothama C, et al. USP9X controls translation efficiency via deubiquitination of eukaryotic translation initiation factor 4A1. *Nucleic Acids Res*. 2018;46(2):823—839.
48. Huai C, Li G, Yao R, et al. Structural insights into DNA cleavage activation of CRISPR-Cas9 system. *Nat Commun*. 2017;8(1):1375.
49. Lu T, Hu JC, Lu WC, et al. Identification of small molecule inhibitors targeting the SMARCA2 bromodomain from a high-throughput screening assay. *Acta Pharmacol Sin*. 2018;39(9):1544—1552.