Applications of the CRISPR-Cas9 system in kidney research

Yoshiki Higashijima[1,2], Seiichi Hirano[3], Masaomi Nangaku[1] and Osamu Nureki[3]

[1]Division of Nephrology and Endocrinology, Graduate School of Medicine, The University of Tokyo, Bunkyo-ku, Tokyo, Japan; [2]Isotope Science Center, The University of Tokyo, Bunkyo-ku, Tokyo, Japan; and [3]Department of Biological Sciences, Graduate School of Science, The University of Tokyo, Bunkyo-ku, Tokyo, Japan

The recently discovered clustered regularly interspaced short palindromic repeat (CRISPR)–CRISPR-associated protein 9 (Cas9) is an RNA-guided DNA nuclease, and has been harnessed for the development of simple, efficient, and relatively inexpensive technologies to precisely manipulate the genomic information in virtually all cell types and organisms. The CRIPSR-Cas9 systems have already been effectively used to disrupt multiple genes simultaneously, create conditional alleles, and generate reporter proteins, even *in vivo*. The ability of Cas9 to target a specific genomic region has also been exploited for various applications, such as transcriptional regulation, epigenetic control, and chromosome labeling. Here we first describe the molecular mechanism of the RNA-guided DNA targeting by the CRISPR-Cas9 system and then outline the current applications of this system as a genome-editing tool in mice and other species, to better model and study human diseases. We also discuss the practical and potential uses of the CRISPR-Cas9 system in kidney research and highlight the further applications of this technology beyond genome editing. Undoubtedly, the CRISPR-Cas9 system holds enormous potential for revolutionizing and accelerating kidney research and therapeutic applications in the future.

Kidney International (2017) **92,** 324–335; http://dx.doi.org/10.1016/j.kint.2017.01.037

KEYWORDS: CRISPR-Cas9; gene expression; gene therapy; genome editing

Correspondence: *Osamu Nureki, Department of Biological Sciences, Graduate School of Science, The University of Tokyo, 2-11-16 Yayoi, Bunkyo-ku, Tokyo 113-0032, Japan. E-mail: nureki@bs.s.u-tokyo.ac.jp*

Received 14 September 2016; revised 26 December 2016; accepted 9 January 2017; published online 20 April 2017

The CRISPR-Cas9 system

Clustered regularly interspaced short palindromic repeat (CRISPR)–CRISPR-associated (Cas) is a prokaryotic adaptive immune system against invading genetic elements, such as phages and plasmids.[1–3] The CRISPR loci consist of Cas genes and a CRISPR array, which encodes CRISPR RNA (crRNAs). The CRISPR-Cas systems can be classified into 6 different types (types I–VI), primarily based on the mechanisms of crRNA biogenesis and target degradation.[4] In the type II system, the precursor crRNAs and the *trans*-activating crRNA (tracrRNA) are transcribed from the CRISPR loci, bound to the Cas9 nuclease, and then processed by RNase III, to form a Cas9-crRNA-tracrRNA effector complex. This complex cleaves DNA targets complementary to the crRNA guide sequence, which is originally derived from previously infected mobile genetic elements.[5] In 2012, biochemical studies revealed that Cas9 is a programmable RNA-guided DNA endonuclease, and the Cas9-crRNA-tracrRNA complex cleaves double-stranded DNA targets complementary to the 20-nucleotide guide sequence in the crRNA.[6,7] Cas9 contains 2 endonuclease domains, HNH and RuvC, which cleave the DNA strands that are complementary (target DNA strand) and noncomplementary (nontarget DNA strand) to the crRNA guide, respectively (Figure 1a and b). In addition to the crRNA-target DNA complementarity, target recognition by Cas9 requires a protospacer adjacent motif (PAM), a short nucleotide motif immediately downstream of the target sequence.[8,9] For example, *Streptococcus pyogenes* Cas9 (SpCas9) and *Staphylococcus aureus* (SaCas9) recognize the 5′-NGG-3′ and 5′-NNGRRT-3′ sequences on the nontarget strand as the PAMs, respectively.[6,7,10]

In 2013, several groups demonstrated that SpCas9 can be harnessed to edit genomic DNA sequences in eukaryotic cells.[11–14] Because a single-guide RNA (sgRNA), a synthetic crRNA:tracrRNA fusion, also directs Cas9 to target cleavage[6] (Figure 1a), changing the crRNA guide sequence can allow the simple 2-component Cas9-sgRNA system to edit the DNA target of interest. Thus, SpCas9 is widely used as a cost-effective and convenient genome-editing tool that works in a broad range of cell types and organisms. In addition, the smaller SaCas9 can be efficiently delivered to somatic tissues for genome editing.[10] Importantly, the orthologous CRISPR-Cas9 systems from different microbes enable simultaneous transcriptional activation, repression, and genome editing,[15] because the Cas9 orthologs recognize their cognate guide RNAs in species-specific manners.[10,16]

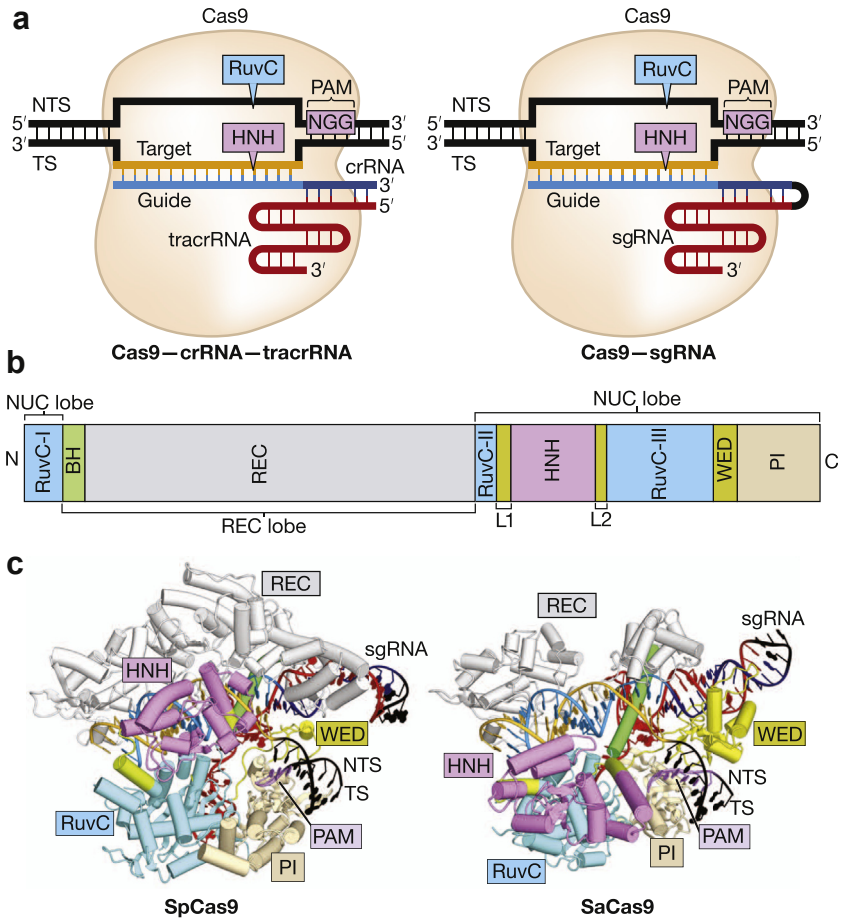

Figure 1 | Function and structures of clustered regularly interspaced short palindromic repeat (CRISPR)–associated 9 (Cas9). (**a**) Schematics of RNA-guided DNA cleavage by Cas9–CRISPR RNA (crRNA)–*trans*-acting crRNA (tracrRNA) and Cas9–single-guide RNA (sgRNA). (**b**) Domain organization of Cas9. (**c**) Crystal structures of *Streptococcus pyogenes* Cas9 (SpCas9) (Protein Data Bank: 4UN3) and *Staphylococcus aureus* Cas9 (SaCas9) (Protein Data Bank: 5CZZ). BH, bridge helix; C, C-Terminus; HNH, a Cas9 endonuclease domain; L1, linker 1; L2, linker 2; N, N-Terminus; NGG, 5′-NGG-′3 PAM sequence of SpCas9; NTS, nontarget DNA strand; NUC, nuclease; PAM, protospacer adjacent motif; PI, PAM-interacting REC, recognition; RuvC, a Cas9 endonuclease domain; TS, target DNA strand; WED, wedge.

Three-dimensional structure of CRISPR-Cas9

The crystal structures of SpCas9 in different functional states advanced our understanding of the mechanism of RNA-guided DNA recognition by Cas9[17–21] (Figure 1c). Cas9 comprises 2 lobes, a recognition lobe and a nuclease lobe. In the absence of the sgRNA, Cas9 adopts an autoinhibited, closed conformation. On sgRNA binding, Cas9 recognizes the sgRNA scaffold and adopts an open conformation, in which the recognition and nuclease lobes form a positively charged, central channel that accommodates the guide RNA-target DNA heteroduplex. The Cas9-sgRNA binary complex initially recognizes the PAM sequence in the target DNA.[22,23] In the crystal structure, the PAM-containing duplex (PAM duplex) is bound between the wedge and PAM-interacting domains, where the PAM nucleotides are recognized from the major-groove side by specific amino acid residues in the PAM-interacting domain.[21] The PAM recognition induces the unwinding of the double-stranded DNA target, thereby initiating the Watson-Crick base-pairing between the crRNA guide and the target sequence.[21,24] The local separation of the DNA target likely causes a conformational change in the HNH domain for target cleavage.[25]

In addition to SpCas9, the crystal structures of SaCas9 and *Francisella novicida* Cas9 revealed the molecular diversity among the orthologous CRISPR-Cas9 systems[26,27] (Figure 1c). The structurally distinct recognition and wedge domains of these Cas9 orthologs recognize their structurally divergent, cognate sgRNA scaffolds in species-specific manners. In the SpCas9 and *Francisella novicida* Cas9 structures, the 5′-NGG-3′ PAMs are recognized by their PAM-interacting domains, in which a distinct set of 2 arginine residues forms direct hydrogen bonds with the GG dinucleotides in the PAM duplex.[21,27] In contrast, in the SaCas9 structure, the 5′-NNGRRT-3′ PAM is recognized by the PAM-interacting domain, in which 2 arginine and 2 asparagine residues form a hydrogen-bonding network with the GRRT nucleotides.[26]

Genome editing in a mouse model

Due to the fundamental genetic similarity between mice and humans, phenotypic analyses of mutant mice are widely used