Engineering Microbiology 1 (2021) 100004

Contents lists available at ScienceDirect

Engineering Microbiology

journal homepage: www.elsevier.com/locate/engmic

Research Highlights

Genome editing from Cas9 to IscB: Backwards and forwards towards new breakthroughs

Zhenxiao Yu, Qunxin She*

CRISPR and Archaea Biology Research Center, Microbial Technology Institute and State Key Laboratory of Microbial Technology, Shandong University, Qingdao, P. R. China

ARTICLE INFO

Keywords:
CRISPR-Cas
OMEGA systems
RNA-guided nucleases
IscB
IsrB
IshB
Guide RNA
tracrRNA

In a very recent article published in Science, Altae-Tran *et al.* reconstructed the evolution of CRISPR-Cas9 systems and traced their ancestors to unique groups of transposons (Altae-Tran et al., 2021). These transposable elements encode RNA-guided nucleases that show strong potential for developing novel biotechnologies. Structural domains of these nucleases serve as useful building blocks for engineering novel RNA-guided nucleases via synthetic biology to strongly inspire the development of novel and precision genome-editing tools.

Programmable RNA-guided nucleases are encoded by the Clustered Regularly Interspaced Palindromic Repeats (CRISPR) and CRISPR-associated (Cas), which constitutes the adaptive antiviral defense system in archaea and bacteria. This antiviral system produces CRISPR (cr)RNAs and Cas proteins that form ribonucleoprotein (RNP) complexes, and the resulting RNP nucleases exert the immune response in an RNA-guided fashion: The RNP enzymes specifically recognize invading viruses and plasmids by testing the sequence complementarity between their crRNAs and foreign nucleic acids, and target the invaders for destruction (Horvath and Barrangou, 2010). RNPs of CRISPR-Cas systems fall into two main classes. While Class 1 RNPs contain multiple Cas proteins, those of Class 2 only have a single Cas protein (Makarova et al., 2020), including Cas9, the well-known RNA-guided DNA scissor encoded by Type II CRISPR-Cas systems. Cas9 is a multi-domain Cas protein that forms RNPs with guide RNAs, which are composed of crRNA and trans-activating CRISPR (tracr) RNA (Fig. 1). The targeting specificity of Cas9 can be programmed by designed guide RNAs. Due to its simplicity and programmability, the CRISPR-Cas9 system has explosively been exploited for developing novel genome editing tools immediately after its discovery. The resulting Cas9 genome-editing technology (Jinek et al., 2012, Gasiunas et al., 2012) is an innovation that has rapidly transformed scientific researches across the life sciences, including genetic manipulations in plants, animals and fungi as well as in bacteria and archaea, which are otherwise inefficient or even impossible to accomplish. In fact, the development of the system awarded the two pioneer researchers, Emmanuelle M. Charpentier and Jennifer A. Doudna, the Nobel Prize in Chemistry in 2020.

Nevertheless, how the nature could have invented these programmable RNA-guided nucleases has been an intriguing question ever since the discovery of the CRISPR-Cas systems. Very recently, Altae-Tran *et al.* made a major breakthrough and demonstrated that distinctive Insertion Sequence (IS) elements belonging to the IS200/605 family encode novel programmable RNA-guided systems. This has led the authors to conclude these transposons are the ancestors of Class 2 CRISPR-Cas, and to reconstruct the evolutionary history of the Cas9 RNA-guided nucleases (Altae-Tran et al., 2021).

A few years ago, it was reported that IS200/605 family transposons are related to the CRISPR-Cas9 systems (Kapitonov et al., 2015). These transposons possess the genetic organization in which two genes (*tnpA* and *tnpB*) are flanked by left and right ends of subterminal hairpins sequences that can form stem-loop structures. This structure is not observed in the classical IS elements, which instead contain a transposase-encoding gene flanked by inverted repeats. In addition, IS200/605 family transposons only require *tnpA* is essential for the transposition of the genetic element as the gene codes for a transposase that ensures the functionality of the transposon. In contrast, the function of *tnpB*

* Corresponding author.
E-mail address: shequnxin@sdu.edu.cn (Q. She).

https://doi.org/10.1016/j.engmic.2021.100004
Received 7 October 2021; Accepted 11 October 2021
Available online 31 October 2021

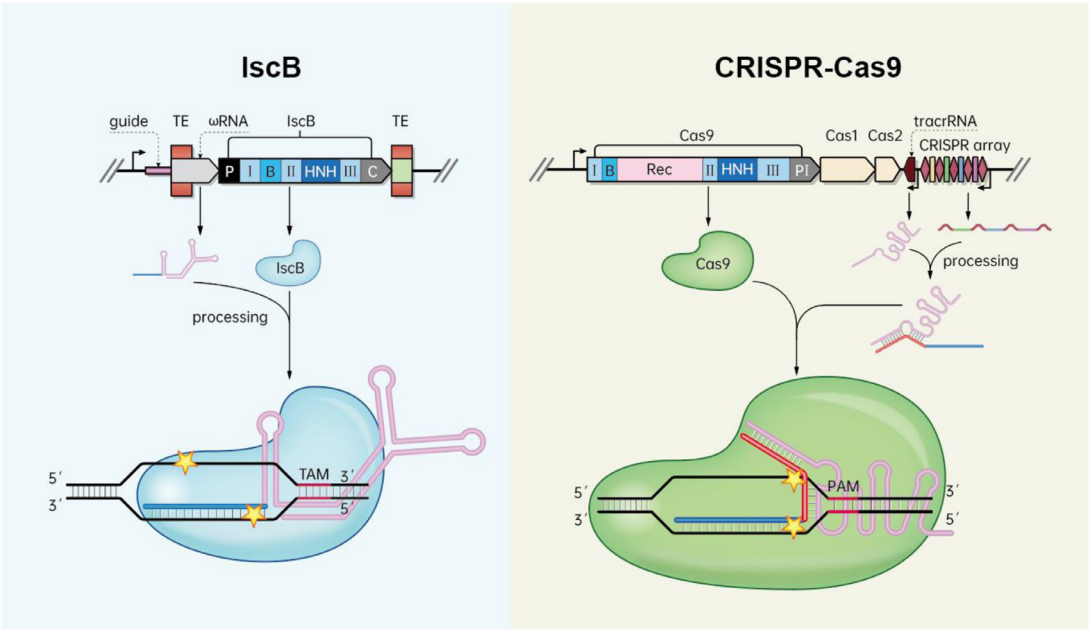

Fig. 1. Comparison of the IscB system and CRISPR-Cas9. Genetic organization is shown for both systems along with the domain organization of IscB and Cas9, the effector protein of each system. The expression of each system and formation of the respective effector complex are illustrated. P: PLMP domain, I~III: split RuvC nuclease, B: bridge helix, HNH: HNH nuclease, C: C terminal domain, TE: transposon end, REC: recognition lobe, PI: PAM interacting domain, tracrRNA: trans-activating CRISPR RNA.

remains unknown, and the gene is not an essential component for all characterized IS200/605 family transposons (He et al., 2015). Interestingly, Kapitonov *et al.* found that some TnpBs are homologous to Cas9 since they share several structural domains and exhibit similar domain organizations. This group of IS200/605 transposons was then classified as the "Insertion sequences encoding Cas9 homologs", and their genes renamed *iscA* and *iscB*, respectively (Kapitonov et al., 2015). In the Science research article, Altae-Tran *et al.* expanded the Cas9 homologs from IscBs to IsrBs (Insertion sequence RuvC-like OrfB) and IshBs (Insertion sequence HNH-like OrfB): The two latter groups lack the HNH domain and the split RuvC domains, respectively, and the deduced proteins have an average of 350 and 180 amino acids, which are smaller than IscBs (about 400 amino acids). These findings have led the authors to propose, for the first time, an evolutionary history for Cas9 proteins: IscBs could be evolved from IsrBs by gaining an HNH domain, and eventually become Cas9 proteins that typically have 1000 or more amino acids (Altae-Tran et al., 2021).

Guide RNA is the other component of the RNA-guided nucleases, and its origin has remained even more mysterious since tracrRNAs of the CRISPR-Cas9 systems do not show any relation to all known non-coding RNAs (Deltcheva et al., 2011). Intriguingly, Altae-Tran et al. showed that IS200/IS605 family transposons code for noncoding (nc)RNAs that form RNPs with the Cas9 homologs. The authors found that a few IscBs systems are associated with CRISPR arrays thanks to the rapid expansion of the metagenomic data in recent years (Altae-Tran et al., 2021). This article has characterized one of the identified IscB systems in detail, the *Ktedonobacter racemifer* IscB-1. The system was genetically expressed in *E. coli*, yielding recombinant IscBs that were functionally characterized. This has led to the discovery that a 329-bp region including a CRISPR array and a noncoding sequence expressing as an ncRNA. Similar to the Cas9 guide RNA, this ncRNA forms RNP complexes with the IscB protein (Fig. 1). Furthermore, when expressed as the recombinant IscB-1 RNP, the effector cleaved target dsDNA in the presence of ATAAA, a 3'-flanking target adjacent motif (TAM). These results led the authors to conclude that the IscB-1 RNP is an RNA-guided nuclease that cleaves dsDNA substrates in a TAM- and ncRNA-dependent fashion, in analogy to all known CRISPR-Cas systems (Altae-Tran et al., 2021). The newly described mechanism is called the Obligate Mobile Element Guided Activity (OMEGA) and the involved ncRNAs show characteristic features of RNA folding and are termed ωRNAs (Fig. 1). In addition, the article reported that ωRNAs are also encoded in a number of OMEGA systems without association of any CRISPR array. This includes Cas9-closely related members IscBs, along with their more distantly related members IsrBs and IshBs. Interestingly, standing alone ωRNAs identified in host chromosomes (Altae-Tran et al., 2021) might functionally interact with the OMEGA systems present in those organisms to modulate the activity of the novel genetic elements. The results presented open an intriguing possibility that tracrRNAs are derived from ωRNAs (Altae-Tran et al., 2021).

These authors also studied the evolution of the CRISPR-Cas12 systems and identified a conserved sequence immediately downstream of *tnpB* that encodes a putative ωRNA in another subset of IS200/IS605 family transposons. Using the *in vitro* plasmid cleavage assay, they showed that a *Ktedonobacter racemifer* TnpB cleaves target dsDNA substrate in a TAM- and ωRNA-dependent fashion, as for the characterized IscB systems. The authors also found that a purified *Alicyclobacillus macrosporangiidus* TnpB not only cleaved the target dsDNA/ssDNA substrates, but also showed a collateral dsDNA/ssDNA cleavage activity (Altae-Tran et al., 2021). This is fascinating since several Cas12 proteins, the predicted descendants of TnpBs, also have target dsDNA/ssDNA-activated collateral ssDNA cleavage activity (Chen et al., 2018, Harrington et al., 2018). These results further demonstrate that these TnpBs and Cas12 proteins are evolutionarily related.

The discovery of OMEGA systems suggests that they could be useful in genome editing. The authors tested several IscB systems, and found that one of them produced appreciable mutations in the human genome at a low efficiency (~4%) (Altae-Tran et al., 2021). Apparently, the OMEGA systems are rather primordial in genome editing, compared to the robust Cas9 and Cas12 systems. Nevertheless, given that the OMEGA systems occur in all three domains of life and their effector proteins (400 amino acids or less) are generally much smaller than Cas9 proteins (usually > 1000 amino acids), there are at least two advantages with these small effectors: (a) it could be safer to utilize them for genome editing in humans since they are widespread in living organisms, and (b) owing

to their large diversity in protein structural domains, it should also be advantageous to use them for building up RNA-guided fusion nucleases for novel applications. For instance, Cas9 has been used for precise gene editing procedures with base editors, which are fusion proteins of Cas9 and cytosine/adenine deaminases (Komor et al., 2016, Gaudelli et al., 2017) as well as prime editing tools utilizing fusion proteins of Cas9 and reverse transcriptases (Anzalone et al., 2019). It would be of great interest to replace individual structural domains of Cas9 with those of IscBs, IsrBs or IshBs for generation of novel engineered RNA-guided nucleases using synthetic biology tools and test the novel DNA scissors for precise genome editing. We envision that such studies would generate RNA-guided nucleases with desired features, and enable the development of new generation tools for precision gene therapy and genetic engineering in the forthcoming years.

Declaration of Competing Interest

We declare no conflict of interest regarding the content presented in this article.

Acknowledgement

Research in our laboratory was supported by grants from the National Key R & D Program of China (2020YFA0906800 to QS), the National Natural Science Foundation of China (Grant No. 31771380 to QS), and State Kay Laboratory of Microbial Technology at Shandong University.

References

Altae-Tran, H, Kannan, S, Demircioglu, FE, Oshiro, R, Nety, SP, McKay, LJ, Dlakic, M, Inskeep, WP, Makarova, KS, Macrae, RK, Koonin, EV, Zhang, F., 2021. The widespread IS200/IS605 transposon family encodes diverse programmable RNA-guided endonucleases. Science 374 (6563), 57–65. doi:10.1126/science.abj6856.

Anzalone, AV, Randolph, PB, Davis, JR, Sousa, AA, Koblan, LW, Levy, JM, Chen, PJ, Wilson, C, Newby, GA, Raguram, A, Liu, DR., 2019. Search-and-replace genome editing without double-strand breaks or donor DNA. Nature 576 (7785), 149–157. doi:10.1038/s41586-019-1711-4.

Chen, JS, Ma, E, Harrington, LB, Da Costa, M, Tian, X, Palefsky, JM, Doudna, JA, 2018. CRISPR-Cas12a target binding unleashes indiscriminate single-stranded DNase activity. Science (New York, N.Y.) 360 (6387), 436–439. doi:10.1126/science.aar6245.

Deltcheva, E, Chylinski, K, Sharma, CM, Gonzales, K, Chao, Y, Pirzada, ZA, Eckert, MR, Vogel, J, Charpentier, E., 2011. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. Nature 471 (7340), 602–607. doi:10.1038/nature09886.

Gasiunas, G, Barrangou, R, Horvath, P, Siksnys, V., 2012. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. PNAS 109 (39), E2579–E2586. doi:10.1073/pnas.1208507109.

Gaudelli, NM, Komor, AC, Rees, HA, Packer, MS, Badran, AH, Bryson, DI, Liu, DR., 2017. Programmable base editing of A*T to G*C in genomic DNA without DNA cleavage. Nature 551 (7681), 464–471. doi:10.1038/nature24644.

Harrington, LB, Burstein, D, Chen, JS, Paez-Espino, D, Ma, E, Witte, IP, Cofsky, JC, Kyrpides, NC, Banfield, JF, Doudna, JA, 2018. Programmed DNA destruction by miniature CRISPR-Cas14 enzymes. Science (New York, N.Y.) 362 (6416), 839–842. doi:10.1126/science.aav4294.

He, S, Corneloup, A, Guynet, C, Lavatine, L, Caumont-Sarcos, A, Siguier, P, Marty, B, Dyda, F, Chandler, M, Ton Hoang, B., 2015. The IS200/IS605 family and "peel and paste" single-strand transposition mechanism. Microbiol. Spectr 3 (4). doi:10.1128/microbiolspec.MDNA3-0039-2014.

Horvath, P, Barrangou, R., 2010. CRISPR/Cas, the immune system of bacteria and archaea. Science 327 (5962), 167–170. doi:10.1126/science.1179555.

Jinek, M, Chylinski, K, Fonfara, I, Hauer, M, Doudna, JA, Charpentier, E., 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 337 (6096), 816–821. doi:10.1126/science.1225829.

Kapitonov, VV, Makarova, KS, Koonin, EV., 2015. ISC, a novel group of bacterial and archaeal DNA transposons that encode Cas9 homologs. J. Bacteriol. 198 (5), 797–807. doi:10.1128/JB.00783-15.

Komor, AC, Kim, YB, Packer, MS, Zuris, JA, Liu, DR., 2016. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. Nature 533 (7603), 420–424. doi:10.1038/nature17946.

Makarova, KS, Wolf, YI, Iranzo, J, Shmakov, SA, Alkhnbashi, OS, Brouns, SJJ, Charpentier, E, Cheng, D, Haft, DH, Horvath, P, Moineau, S, Mojica, FJM, Scott, D, Shah, SA, Siksnys, V, Terns, MP, Venclovas, C, White, MF, Yakunin, AF, Yan, W, Zhang, F, Garrett, RA, Backofen, R, van der Oost, J, Barrangou, R, Koonin, EV., 2020. Evolutionary classification of CRISPR-Cas systems: a burst of class 2 and derived variants. Nat. Rev. Microbiol. 18 (2), 67–83. doi:10.1038/s41579-019-0299-x.