



OPEN ACCESS

EDITED BY

Ragini Bodade,
Ministry of Science and Technology, India

REVIEWED BY

Yang Liu,
Shenzhen University, China
Yogendra Singh Rajpurohit,
Bhabha Atomic Research Center (BARC),
India

*CORRESPONDENCE

Long Jin
✉ isackkim@alumni.kaist.ac.kr
Kee-Sun Shin
✉ ksshin@kribb.re.kr

RECEIVED 01 April 2025

ACCEPTED 02 September 2025

PUBLISHED 15 September 2025

CITATION

Yang Z-Q, Li M-J, Ahmad F, Jin C-Z, Li T,
Jin F-J, Shin K-S and Jin L (2025) Application
of the transposon-associated TnpB system
of CRISPR-Cas in bacteria: *Deinococcus*.
Front. Microbiol. 16:1604032.
doi: 10.3389/fmicb.2025.1604032

COPYRIGHT

© 2025 Yang, Li, Ahmad, Jin, Li, Jin, Shin and
Jin. This is an open-access article distributed
under the terms of the [Creative Commons
Attribution License \(CC BY\)](#). The use,
distribution or reproduction in other forums
is permitted, provided the original author(s)
and the copyright owner(s) are credited and
that the original publication in this journal is
cited, in accordance with accepted academic
practice. No use, distribution or reproduction
is permitted which does not comply with
these terms.

Application of the transposon-associated TnpB system of CRISPR-Cas in bacteria: *Deinococcus*

Zi-Qi Yang¹, Mei-Ju Li², Faizan Ahmad¹, Chun-Zhi Jin³,
Taihua Li¹, Feng-Jie Jin¹, Kee-Sun Shin^{3*} and Long Jin^{1*}

¹Co-Innovation Center for Sustainable Forestry in Southern China, College of Ecology
and Environment, Nanjing Forestry University, Nanjing, China, ²College of Forestry and Grassland,
Nanjing Forestry University, Nanjing, China, ³Microbiome Convergence Research Center, Korea
Research Institute of Bioscience and Biotechnology (KRIBB), Daejeon, Republic of Korea

Deinococcus radiodurans is one of the most radioresistant organisms found on Earth to date, showing extreme resistance to damage factors such as UV, drought, and mutagens, and is of great interest to scientists around the world. It was determined that the TnpB protein from *D. radiodurans* ISDra2 functions as an RNA-guided endonuclease, serving as a functional ancestor for the widely used CRISPR-Cas endonucleases. The CRISPR-Cas system is an “acquired immune system” found in most Bacteria and Archaea, and used in a wide range of biological and medical research fields. Cas12f is the smallest RNA-directed nuclease that is currently known and possesses unique characteristics. There has been extensive research conducted on the origin, classification, and mechanism of action of CRISPR-Cas12f, as well as its application in the field of gene editing. TnpB, as the protein closest to Cas12f in the evolutionary tree, has the potential to be used as a new micro-editing tool. Systematic studies have been conducted on it to develop smaller volumes of precision gene editing and treatment tools. In this review, the research progress, mechanism, and application of TnpB protein in *D. radiodurans* were reviewed. In addition, the classification of CRISPR-Cas system and the application and function of CRISPR-Cas12f in gene editing are also introduced and summarized.

KEYWORDS

Deinococcus, *Deinococcus radiodurans*, CRISPR/Cas, TnpB, transposon

1 Introduction

Deinococcus radiodurans was first isolated from radiation-sterilized canned meat in 1956 by the American scientist Anderson et al. (Anderson et al., 1956; Jung et al., 2017). It is one of the most radiation-resistant organisms found on Earth to date, and therefore is known as the world's most resistant organism (Anderson et al., 1956). The most remarkable feature of *D. radiodurans* is that it can tolerate a radiation dose of 15 kGy during its stable growth period, which is more than 250 times the radiation resistance of *E. coli* and 3,000 times that of humans (Cox and Battista, 2005). Because of its extreme resistance characteristics, *D. radiodurans* is of great importance in the study of radiation resistance mechanism, environmental remediation and tumorigenesis, and is therefore of great interest to scientists worldwide as a promising microbial resource for exploitation (Desai et al., 2011). *Deinococcus* is capable of withstanding extreme radiation and stress by

means of its survival mechanisms, which are supported by DNA repair systems, genomic features, and unique cell wall structures. The condensed genomes of *Deinococcus* contain repetitive sequences that are involved in the degradation of damaged DNA. The cell envelope of certain *Deinococcus* strains is composed of six layers, each of which is composed of distinct components such as lipids, proteins, and carbohydrates (Jin et al., 2019; Liu et al., 2023). *D. radiodurans* has a very special cell wall structure. According to the structure and composition of its cell wall should belong to Gram-negative bacteria, but because of the thick peptidoglycan layer in the cell wall, it is difficult to decolorize after crystal violet staining, and the Gram stain is positive. Meanwhile, the cell wall has an outer membrane but lacks the teichoic acid contained in ordinary Gram-positive bacteria. This multi-layer cell wall structure has a certain blocking effect on ionizing rays and ultraviolet rays, which can slow down the damage caused by radiation to the bacteria, and play a role in supporting and protecting the cells (Sharma et al., 2024).

Deinococcus radiodurans not only exhibits extraordinary radiation resistance but also possesses the remarkable ability to fully restore its genome from hundreds of double-strand breaks (DSBs) and thousands of single-strand breaks (SSBs). Therefore, studying the DNA damage repair mechanisms in this bacterium is of significant importance—not only for elucidating the molecular basis of DNA repair but also for applications in environmental protection, bioremediation, human health, and even the exploration and utilization of extraterrestrial environments. This has made it a major research focus in recent years (Mishra et al., 2021; Sharma et al., 2025). The chromosome structure is also special. During the stationary growth phase, the nucleoid of *D. radiodurans* presents a very compact circular structure. Researchers believe that this dense ring structure can prevent the fragments formed after double-strand breaks from dispersing during the repair process, thus facilitating the repair of DNA damage and providing passive defense against radiation (Zhang et al., 2009; Mishra et al., 2021). However, this hypothesis has also been disputed based on studies with *E. coli* (Zimmerman and Battista, 2005). Nevertheless, a highly dense genomic structure that prevents the diffusion of damaged DNA and thus facilitates the search for repair templates (Kota and Misra, 2006; Rajpurohit et al., 2013; Kota et al., 2014) is still a reasonable explanation. *D. radiodurans* has an efficient DNA damage repair mechanism. The most obvious feature of *D. radiodurans* is that it can withstand gamma rays of more than 15 kGy and repair up to thousands of DNA double-strand breaks (Redox Signals) produced by radiation within a dozen hours, reconstructing the genome without mutation and without affecting its viability (Minton and Daly, 1995; Bihani et al., 2018). The known DNA repair methods of *D. radiodurans* include five types: base excision repair, direct damage repair, nucleotide excision repair, base mismatch repair and recombination repair (Timmins and Moe, 2016; Sharma et al., 2025), among which DNA recombination repair includes homologous recombination (HR), non-homologous end joining (NHEJ), single-strand end annealing (SSA), and synthesis-dependent single-stranded end joining (SDSA). SSA and SDSA, as well as a new DNA recombination repair pathway, extended synthesis-dependent strand annealing (ESDSA), were proposed by Zahradka et al. (2006). Before DNA damage, the DNA template has some DNA duplexes as patches, and after DNA damage, the newly

synthesized DNA duplexes link the patched duplexes to complete the repair (Bentchikou et al., 2010). Its diverse DNA damage repair mechanism is expected to be further exploited and developed as an effective tool for gene editing. The CRISPR-Cas system is a defense mechanism widely present in bacteria and archaea. Currently, genome editing tools developed based on this system have achieved targeted editing in numerous species. However, their protein sizes all exceed 1,000 amino acids, which hinders delivery efficiency. The TnpB protein from transposon families in *D. radiodurans* has been identified as the ancestor of Cas12. These proteins are collectively referred to as Obligate Mobile Element-Guided Activity (OMEGA) proteins, with their guide RNA called ω RNA. Consequently, the OMEGA system has emerged as a research hotspot in the field of gene editing (Sasnauskas et al., 2023). Further exploration of *D. radiodurans*' efficient DNA repair mechanisms and natural transformation capabilities may lead to the development of precision gene editing tools suitable for extreme environments, thereby promoting innovative applications of this bacterial strain in broader biotechnological scenarios.

In the past, radioresistance of *D. radiodurans* was mainly attributed to its resistance and repair mechanisms against DNA damage, and a lot of research has been conducted in this area. However, as research progressed, it was discovered that *D. radiodurans* also has a superb oxidative defense system, which is another important reason for its extreme radioresistance (Kota et al., 2014; Maurya et al., 2018; Misra and Rajpurohit, 2024). It was found that along with the direct damage to DNA and proteins caused by ionizing radiation, a large number of highly cytotoxic reactive oxygen species (ROS) are produced in the cells, which can inhibit the activity of intracellular proteins, break DNA chains and damage cell membranes containing large amounts of unsaturated fatty acids, causing various ROS can inhibit intracellular protein activity, break DNA strands and damage cell membranes containing large amounts of unsaturated fatty acids, thus causing various metabolic defects, aging, mutations and even cell death (Bauche and Laval, 1999; Lim et al., 2019).

Studies have shown that only 20% of the direct damage to DNA is caused by ionizing radiation, while the remaining 80% is indirectly caused by the attack of ROS generated by radiation. There are three main types of radiation-generated ROS: hydroxyl radicals ($\cdot\text{OH}$), superoxide anion radicals ($\text{O}_2^{\cdot-}$), and hydrogen peroxide (H_2O_2), which constitute strong oxidative damage to biomolecules such as proteins, lipids, nucleic acids, and sugars in cells (Bauche and Laval, 1999; Misra et al., 2012). Therefore, in addition to the efficient DNA damage repair system, the strong scavenging ability and extreme resistance to ROS by the antioxidant defense system in *D. radiodurans* is another important factor for its extreme radioresistance. The antioxidant protection system, overseen by *D. radiodurans*, primarily encompasses two categories of substances: the antioxidant enzyme system and non-enzymatic ROS scavengers. Within the antioxidant enzyme system, notable components are superoxide dismutase (SOD), catalase (CAT), and peroxidase (POD). The non-enzyme category comprises substances such as carotenoids, Mn²⁺ + complexes, pyrroloquinoline quinone, and DNA protection during starvation (Dps) protein (Sadowska-Bartosz and Bartosz, 2023). Two Dps proteins, Dps1 (DR2263) and Dps2 (DRB0092), have been identified in *D. radiodurans*. The dimerization of the Dps1 protein protects DNA from hydroxyl radical damage (Grove and Wilkinson, 2005).

Deinococcus radiodurans, as a tough organism that can withstand extreme radiation and various adversities, represents the limit of survival of life against radiation and other adversities. Therefore, the study of *D. radiodurans* extreme radiation resistance mechanism has become a source of new ideas and methods in the fields of environmental protection and bioremediation, human health, food/drug research and development, plant and animal breeding and cosmetics, and even extraterrestrial space development and utilization, which has a very broad application prospect and important significance. *D. radiodurans* is an attractive candidate for genetic engineering applications due to its powerful DNA repair mechanism and efficient DNA uptake system (Rajpurohit and Misra, 2013).

The engineered strain *D. radiodurans* DRG300 was produced in a study in which the enhanced green fluorescent protein (eGFP) was fused with the promoter of the key DNA damage-inducible *recA* gene derived from *D. radiodurans* and transformed into the *D. radiodurans* R1 (Kota and Misra, 2006; Mishra et al., 2019; Rajpurohit et al., 2021). This strain serves as a potential whole-cell biosensor, enabling the construction of a detection system for real-time monitoring of radioactive and toxic pollutant hazards in the environment (Guanjun et al., 2008). A γ -radiation-inducible small RNA (sRNA), *DrsS*, was discovered in *D. radiodurans*, which is induced by oxidative (e.g., H_2O_2) and genotoxic stress. The research indicates that *DrsS* activates catalase and detoxifies intracellular reactive oxygen species (ROS) under oxidative stress (Kamble et al., 2010; Rajpurohit and Misra, 2010; Kota et al., 2014; Rai and Dutta, 2024). Recently, a research team successfully developed the recombinant strain *Deino-Ure* by genetically engineering *D. radiodurans* to express the urease gene (*Ure*) from *Sporosarcina pasteurii* DSM33. The urease in this strain hydrolyzes urea to produce CO_3^{2-} , which, in the presence of calcium ions, facilitates calcium carbonate precipitation, thereby enhancing *D. radiodurans*' biomineralization and uranium immobilization capabilities (Liang et al., 2024). Furthermore, exosomes from *D. radiodurans* have been shown to substantially increase the survival rate of mice exposed to 8 Gy radiation (70%) while simultaneously safeguarding the hematopoietic and gastrointestinal systems. These exosomes are more effective than current single-organ radioprotectants and do not cause any toxicity (Han et al., 2025). Despite extensive research on the highly efficient DNA repair mechanisms, extreme antioxidant system and unique physicochemical properties of *D. radiodurans*, many genes, proteins and other biomolecules with unknown functions still exist (Bauche and Laval, 1999). TnpB from *D. radiodurans* ISDra2 is an RNA-directed nuclease that cleaves the 5'-TTGAT transposon-associated motifs next to DNA. TnpB was also identified as a progenitor of the CRISPR-Cas (Clustered Regularly Interspaced Short Palindromic Repeats) nuclease and used as a prototype for a new genome editing system (Karvelis et al., 2021).

2 Overview of the CRISPR/Cas system

In about half of the fine bacteria and almost all archaea, CRISPR and CRISPR-Cas associated proteins constitute an adaptive

immune system against phage and exogenous plasmid invasion, called the CRISPR-Cas system (Jansen et al., 2002). This system was discovered in *E. coli* by Japanese scientists Ishino et al. (1987). The CRISPR/Cas system can recognize exogenous DNA and cut them off to silence the expression of exogenous genes. This is similar to the principle of RNA interference (RNAi) in eukaryotes. It is due to this precise targeting function that the CRISPR/Cas system has been developed as an efficient genome editing tool. Unlike genome editing that requires the assistance of Cre and I-SceI nucleases, CRISPR/Cas-based technologies do not require the pre-integration of unique enzyme recognition sequences into the genome, but instead use synthetic single-guide RNA-sgRNA (Jinek et al., 2012) (a chimera of crRNA and tracrRNA), or just crRNA, to direct the nucleic acid endonuclease Cas protein to target any site on the genome for genomic cleavage (Knott and Doudna, 2018). In recent years, transposon families and highly abundant OMEGA systems (including TnpB and IscB) have attracted widespread attention (Altae-Tran et al., 2021). These proteins are considered ancestors to Cas9 and Cas12, characterized by compact molecular sizes (approximately 400 amino acids) and diverse types, and have been demonstrated to possess RNA-guided nuclease activity (Altae-Tran et al., 2021; Karvelis et al., 2021). Notably, the Fanzor protein (a TnpB homologous protein widely present in eukaryotes) has also been shown to exhibit RNA-guided nuclease activity, indicating that RNA-guided nuclease systems are extensively distributed across trilobate organisms (Saito et al., 2023). In nature, CRISPR/Cas systems possess several classes, among which the CRISPR/Cas9 system is the most intensively studied and the most maturely used one.

2.1 CRISPR/Cas system classification

Depending on the structure and characteristics of the complex formed by the Cas effector protein during its function, CRISPR-Cas systems can be generally classified into two major classes (Haft et al., 2005; Makarova et al., 2011, 2020), including six types and 48 isoforms (Figure 1). These systems include three types: type I, type III, and type IV, represented by Cas3, Cas10, and Csf1, respectively. The type II, CRISPR-Cas, requires only a single Cas protein and performs immune functions by binding to crRNA to degrade exogenous invasive DNA. Approximately 90% of all CRISPR-Cas systems known to date fall into the first category. Class II CRISPR/Cas systems are mainly found in bacteria, with a relatively small distribution in archaea. Class II CRISPR-Cas systems also include three types: type II, type V and type VI, represented by Cas9, Cas12a (Cpf1), Cas13a (C2c2) and Cas12f (Cas14), respectively. Since this system relies on the formation of a complex between a single Cas nuclease and crRNA to function (Chylinski et al., 2014; O'Connell, 2019; Makarova et al., 2020), it is simple and efficient, so the understanding of its gene composition, mechanism of action and molecular structure of effectors is more thorough and clearer than that of type I system.

Current genome editing methods generally use Type II and Type V systems, using Cas9 and Cas12 as effector proteins. TnpB and IscB transposon proteins may be the ancestor proteins of Cas12 and Cas9, respectively (Kapitonov et al., 2015; Shmakov et al., 2015, 2017; Makarova et al., 2020). Notably, while CRISPR-like systems

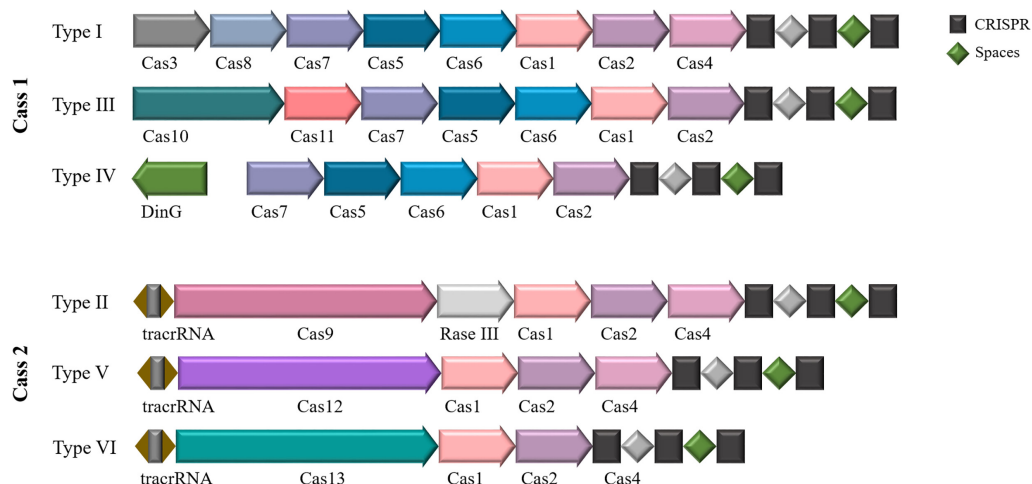


FIGURE 1

CRISPR-Cas system classification. Two classes and six types based on the structure and function of the Cas protein. Class I includes types I, III, and IV, and class II includes types II, V, and VI.

were previously thought to exist exclusively in prokaryotes, recent studies have identified homologous proteins of TnpB and IscB in eukaryotes, demonstrating their RNA-guided nuclease activity (Altae-Tran et al., 2021; Karvelis et al., 2021; Saito et al., 2023).

The CRISPR-Cas system is the most popular gene editing tool in the world today, but the size of conventional Cas nucleases is generally too large, limiting its application to mammals. In recent years, scientists have been seeking to develop smaller gene editing tools. Three research teams from China, Korea, and the United States have published research papers that identify the same class of ultra-small CRISPR nucleases, Cas12f (Wu et al., 2021; Xu et al., 2021; Kim et al., 2022). Cas12f is the smallest RNA-guided nuclease known to play a role in human cell editing and has been less studied compared to Cas9, Cas12 and Cas13. Bioinformatic analysis suggests that TnpB may be the predecessor of CRISPR-Cas9/Cas12 nucleases (Kapitonov et al., 2015; Shmakov et al., 2017; Makarova et al., 2020), and the nearest TnpB neighbor in the evolutionary tree is the miniature Cas12f nuclease. Therefore, a more detailed description of Cas12f nucleases is given below.

3 Overview of micro-CRISPR nuclease Cas12f and its application in the field of gene editing

3.1 CRISPR - Cas12f system structural features

In, Harrington et al. (2018) identified a CRISPR-Cas system containing Cas14 in uncultured archaea and found that it has similar structural domains to Cas9 and Cas12, but with functional differences. Sequence databases identified Cas12f as belonging to the class II CRISPR V-F subtype of the R-Cas system, and a recent classification classified the family into Cas12f1 (Cas14a and V-U3 types), Cas12f2 (Cas14b) and Cas12f3 (Cas14c, V-U2 and U4 types)

(Shmakov et al., 2017; Harrington et al., 2018; Yan et al., 2019; Makarova et al., 2020).

Cas12f nucleases share a common set of features (Jinek et al., 2012; Zetsche et al., 2015; Shmakov et al., 2017; Liu et al., 2019; Yan et al., 2019; Pausch et al., 2020): the inclusion of a RuvC nuclease structural domain similar to Cas12; preferential binding and cleavage of double-stranded DNA (dsDNA) at higher temperatures (45 °C–55 °C); and non-specific single-stranded DNA (ssDNA) degradation after target binding. However, compared to other typical class II nucleases, Cas12f is again distinctive in several properties: (i) Cas12f nuclease is relatively small, with a size of 400–700 amino acid residues, which is only half the size of other class II CRISPR effector nucleases (950–1400 aa), and is the smallest RNA-guided nuclease proven to date (Harrington et al., 2018; Savage, 2019). One of them, Cas12f1 (also known as Cas14a1), consists of 529 residues and lacks detectable sequence identity with other known proteins except for the presence of the RuvC structural domain. (ii) Cas12f1 is associated with a dual crRNA: tracrRNA guide (which can be combined into single guide RNA, sgRNA) and uses TTTR (where R is A or G) PAM (protospacer adjacent motifs) to cleave single- and double-stranded DNA targets (Jansen et al., 2002; Makarova et al., 2011; Harrington et al., 2018). The sequences of DNA cleavage products showed that Cas12f cleaves TS and NTS at 24 and 22 nucleotides upstream of PAM, respectively (Takeda et al., 2021). (iii) Dimerization after binding single-copy gRNA. (iv) CRISPR-Cas12f nuclease, in a PAM-dependent manner recognizes and cleaves dsDNA, whereas cutting ssDNA does not require the target sequence PAM. Cas12f recognizes T-rich or C-rich PAM (Bigelyte et al., 2021), Cas12a (Cpf1) recognizes T-rich PAM and Cas9 recognizes G-rich PAM for cleavage (Jinek et al., 2012). (v) Asymmetric double-stranded cleavage pattern. It was found that the mechanism of DNA cleavage by As-Cas12f1 differs greatly from previous Cas proteins (Wu et al., 2021). Previously reported Cas proteins usually cut two times inside the DNA targeting sequence, severing two strands of DNA and leaving a cut that is usually a neat flat end or a sticky end with only a few bases. In contrast, the core cut site of AsCas12f1 is outside of the target sequence,

making one cut on the target strand and two cuts on the non-target strand, respectively, for a total of three cuts. This creates a set of asymmetric cuts that leave a longer stretch of sticky ends on the target strand. This asymmetric double-stranded cut pattern is currently rare. (vi) Cas14a shows high fidelity in the recognition of ssDNA substrates with targeting cleavage activity against ssDNA, as this recognition is mediated by the seed sequence near the middle of the ssDNA target interaction mediated, which is different from the seed region required to target the CRISPR-Cas12 system of dsDNA and PAM distal sequences, which are located in the middle of Cas14a (Wiedenheft et al., 2011; Chen et al., 2018; Aquino-Jarquín, 2019). (vii) The tracrRNA is very long. Despite the fact that Cas12f1 proteins are the most compact class II CRISPR-Cas nucleases (<500 aa) characterized so far, their gRNA length clearly exceeds the length of other class II effectors. These include crRNAs that are 40–50 nucleotides, component repeat sequences and spacer regions, and a long tracrRNA (153 and 169 nt for SpCas12f1 and AsCas12f1, respectively) (Wu et al., 2021). (viii) The guide RNA of Cas12f1 lacks sequence similarity to other Cas12 enzymes.

3.2 Mechanism of action of the CRISPR - Cas12f system

The defense process of the CRISPR-Cas12f system is similar to other CRISPR-Cas systems and consists of 3 phases: In the first phase, the Cas protein recognizes foreign nucleic acids, targets the pro-spacer sequences of the invading genome for cleavage, acquires new spacer sequences, and integrates them into the host genomic CRISPR array. In the second stage, these incorporated pro-spacer sequences are transcribed into precursor CRISPR ribonucleic acid (pre-crRNA). In the final stage, the pre-crRNA is cleaved into short crRNAs and paired with tracrRNAs to form guide RNA (sgRNA) by the action of Cas proteins. Cas12f requires dimerization of guide RNA, a catalytic effector complex that recognizes sgRNA and its target DNA to form the Cas12f-sgRNA-targetDNA complex, which targets the target nucleic acid region by crRNA and enables cleavage of exogenous nucleic acids to counteract reinvasion by homophages or exogenous DNA re-invasion (Mohanraju et al., 2016; Hille et al., 2018; Takeda et al., 2021; Figure 2). The asymmetric cleavage pattern of AsCas12f1 originates from its unique structural dynamics mechanism. AsCas12f1 (only 422 amino acids in size) first unwinds the DNA within the protospacer and creates a single-strand nick on the non-target strand (NTS) upon binding to target DNA. Subsequently, its bidirectional exonuclease activity extends the unwinding region to a 5-bp area beyond the protospacer, ultimately cleaving the target strand (TS) at a position 3 nucleotides downstream of the protospacer, generating three break sites and resulting in asymmetric cleavage (Song et al., 2024). This pattern differs from other Cas proteins. For instance, Cas9 typically makes two cuts within the DNA target sequence, cleaving both strands and leaving either blunt ends or sticky ends with only a few base overhangs. From evolutionary and functional perspectives, AsCas12f1's asymmetric cleavage mode holds significant implications. While Cas12a exhibits more regular cleavage patterns that facilitate precise control over gene knockout and insertion in conventional gene editing, the unique termini generated by AsCas12f1's asymmetric cleavage

may be more conducive to ligation with specially designed DNA sequences, offering novel strategies for gene editing. Furthermore, studies have shown that Cas12f nucleases (including AsCas12f1) produce relatively fewer harmful byproducts such as chromosomal translocations and large fragment deletions during editing, making them particularly promising for applications requiring high genomic stability, such as gene therapy (Xin et al., 2022). Additionally, the small molecular size of AsCas12f1, combined with its cleavage pattern, enables easier packaging into adeno-associated virus (AAV) vectors for *in vivo* delivery, providing a superior tool for *in vivo* gene therapy and related applications.

3.3 CRISPR - Cas12f in the field of gene editing

To correct mammalian genetic defects, a set of gene editing tools must be delivered precisely to a specific tissue or cell, a process that typically uses adeno-associated viruses (AAVs), which do not make people sick and do not replicate themselves and are currently the most reliable vectors. The disadvantage of AAVs, however, is their limited loading capacity, and the current mainstream gene editing tools, dominated by Cas9 and Cas12a, generally exceed the loading limit of a single AAV. Cas12f nucleases are relatively small, half the size of other class II CRISPR effector nucleases, and hold promise for development as tools for mammalian cell editing.

Although natural Cas12f and its sgRNA are inactive in mammalian cells, Kim et al. (2022) have modified the system to greatly enhance the gene editing activity of the Un1Cas12f1 system in mammals. Based on the structural information of the Un1Cas12f1/sgRNA/DNA complex, truncating the RNA portion of the structure defined as disordered, while adding uridylic acid-rich sequences at the end of the crRNA, the engineered Un1Cas12f1 system showed efficiency comparable to that of SpCas9 and specificity similar to that of AsCas12a. At the same time, the system's volume of only 529 aa allows Un1Cas12f1 to be efficiently delivered via AAV, which meets the volume requirements of AAV viral delivery systems (Kim et al., 2022). The high gene editing efficiency and specific genome editing ability of Cas12f in mammalian cells have revealed its great potential for clinical therapeutic applications.

Wu et al. (2021) characterized the DNA recognition and cleavage mechanism of AsCas12f1, a very small CRISPR nuclease. Cas12f1 can effectively execute gene editing in *E. coli* and *Klebsiella pneumoniae*, as well as a variety of human cells, through plasmids, ribonucleoprotein (RNP) and AAV. This discovery offers a new idea for the development of miniature precision gene editing and therapeutic tools. Wang et al. (2021) successfully demonstrated genome editing in *Bacillus anthracis* based on the AsCas12f1 nuclease from *Acidibacillus sulfuroxidans*. Additionally, they constructed a dual plasmid CRISPR-AsCas12f1 system. The results indicated that the CRISPR-Cas12f system is a highly effective tool for genetic manipulation in *Bacillus anthracis* and may be broadly applicable to other *Bacillus* strains. The CRISPR-Cas12f system has a clear comparative advantage over the CRISPR-Cas9 system, as it has smaller genome editing plasmid size and higher transformation efficiency in *B. anthracis* related experiments using this effector. Bigelyte et al. (2021) demonstrated that two

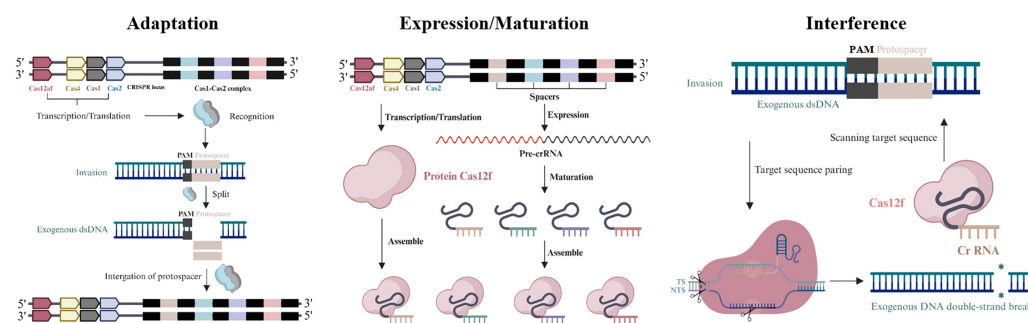


FIGURE 2

Stages of CRISPR-Cas12f system immunity: adaptation, expression/maturation, interference. Adaptation: the Cas protein complex like Cas1-Cas2 complex integrates fragments of invader DNA as spacers in the CRISPR array. Expression/Maturation: precursor CRISPR ribonucleic acid (pre-crRNA) is synthesized by transcription of pro-spacer sequences. Interference: the pre-crRNA is cleaved into short crRNAs and paired with tracrRNAs to form sgRNA by Cas proteins.

V-FCRISPR-Cas nucleases, SpCas12f1 (497 aa) and AsCas12f1 (422 aa), are capable of targeted DNA modification in both plant and human cells. This discovery paves the way for the development of miniature Cas12f1-based genome editing tools. At the same time, the development of diagnostics based on the CRISPR-Cas12f system, as well as the design and modification of these systems, has enabled the development of a variety of rapid assays. Cas14a binds more specifically to its targets than Cas12a, making it better suited for the development of assays that require single-base resolution. Harrington et al. (2018) used phosphorothioate amplification to protect a strand from exonuclease degradation, leaving behind ssDNA substrates that Cas14a can detect. They developed the Cas14a-DETECTR assay, which has high fidelity in distinguishing between ssDNA substrates and can be used for DNA single nucleotide polymorphism (SNP) genotyping, ssDNA pathogen detection, and diagnosis. The team successfully used Cas14a - DETECTR to detect SNP typing of the human E3 ubiquitin ligase (HERC2) gene, which contains SNP typing that leads to iris color variation. The team designed the Cas14-DETECTR assay, which can successfully differentiate between blue and brown eye genotypes with a single base of specificity (only one base difference in the sequence of the genes controlling the two eye colors) (Harrington et al., 2018). Meanwhile, the Cas14- DETECTR platform can also be used as a viral screening method. The CRISPR-Cas14 system is used in conjunction with the HUDSON method (Gootenberg et al., 2017), which does not involve complex sample extraction and transportation, to achieve rapid detection of viruses, and has been successfully applied to the fine virus detection of human bocavirus (HBoV1) (Aquino-Jarquin, 2019).

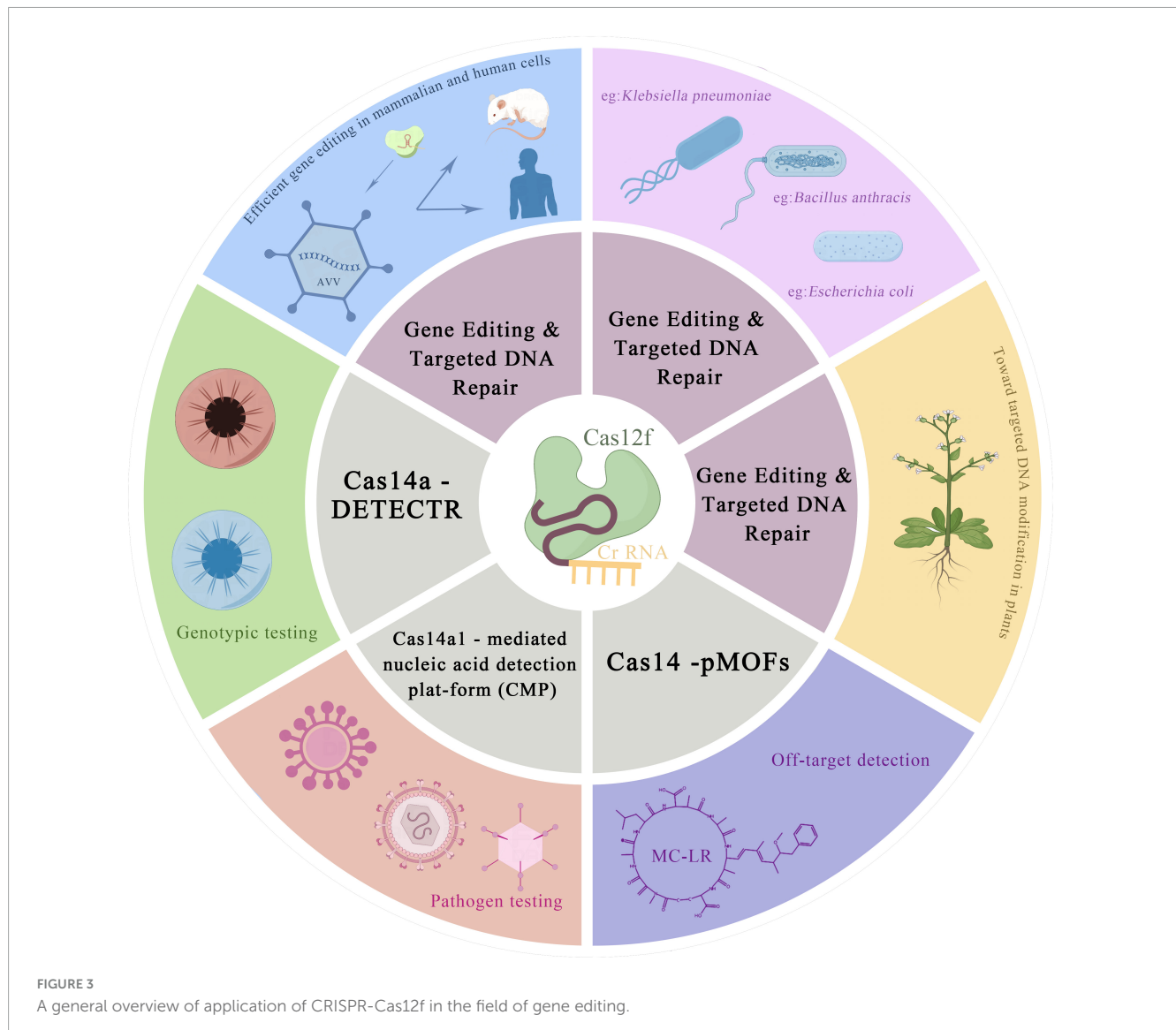
Wei et al. (2021) developed a highly sensitive and specific fluorescent tool for bacterial detection using Cas14a1-mediated nucleic acid detection platform (CMP). Combining CMP with molecular amplification to construct a CRISPR-Cas-based bioanalytical technology allows for rapid nucleic acid detection with high sensitivity and specificity (Wei et al., 2021). The results show that CMP is a reliable and accurate pathogen detection method. The technique can accurately identify different types of pathogens in milk samples. Wu et al. (2022) developed a novel universal Cas14 -pMOFs fluorescent sensor by combining the CRISPR- Cas14a system with two-dimensional porphyrin metal-organic skeleton nanosheets. Using an aptamer to introduce

a homogeneous competitive reaction between Microcystin-LR (MC-LR) and complementary DNA, this novel sensor was applied to sensitively determine MC-LR with a limit of detection of 19 pg/mL (Ma et al., 2014; Yang et al., 2020; Wu et al., 2022). The Cas14- pMOFs fluorescence sensor was shown to be feasible and applicable as a favorable MC-LR assay, and the application of CRISPR-Cas14a-based assays for non-nucleic acid target detection was extended.

However, it is also important to recognize that while CRISPR-Cas12f is an effective and cost-effective method for high-throughput pathogenic mutation screening, sgRNA-mediated Cas12f targets the cleavage of ssDNA, and additional steps are required to generate ssDNA from dsDNA targets prior to detection, and this particular property of CRISPR-Cas12f may increase the complexity of its application (Figure 3).

4 The evolutionary relationship between CRISPR-Cas and TnpB

Over the past decade, scientists have applied the CRISPR-Cas9 system, a precisely programmable system for modifying DNA, to gene editing technology. In addition to generating DNA double-strand breaks, deactivated Cas (dCas) or Cas nickase (nCas) versions have been designed as reprogrammable DNA binding or cleavage systems. Different domains have been fused to dCas or nCas to establish epigenome editing, base editing and primitive editing technologies. The origin of CRISPR-Cas enzymes has also been a hot topic of research in recent years, and researchers have made new discoveries on the subject, as they have identified a new programmable DNA modification system. The TnpB system consists of nuclease and ω RNA components. The ω RNA comprises a scaffold region forming complex three-dimensional structures and a guide region that complements the target DNA. While the OMEGA system contains small nuclease molecules, the ω RNA in TnpB is typically longer and more structurally complex compared to crRNA (Kato et al., 2022; Schuler et al., 2022; Nakagawa et al., 2023; Saito et al., 2023; Sasnauskas et al., 2023). The TnpB ω RNA measures approximately 150 nt, consisting of a 116 nt scaffold region and a 16 nt guide region. It features four stems and one



pseudoknot (a circular sequence in the stem-loop structure that pairs with base sequences elsewhere, forming a “knot-like” shape), which is also preserved in all Cas12 systems (Nakagawa et al., 2023; Sasnauskas et al., 2023). However, TnpB differs significantly from Cas12 systems. Structural comparisons reveal that certain regions of the TnpB ω RNA function similarly to specific domains of the Cas12 protein (Nakagawa et al., 2023; Sasnauskas et al., 2023). Furthermore, a single ISDra2 TnpB protein is guided by a single ω RNA, while certain Cas12 proteins form asymmetric dimers (e.g., Cas12f) that require dual RNA guidance (tracrRNA and crRNA) (Takeda et al., 2021; Xiao et al., 2021). Structural analysis suggests that during evolution, Cas12 likely identified PAM-terminal guide sequences through either forming asymmetric dimers or inserting diversified REC2 domains, thereby establishing the structural basis for CRISPR-Cas adaptation (Nakagawa et al., 2023). The dual-RNA guidance sequence may have evolved from ω RNA. Cas12f1 and dual crRNA: tracrRNA-guided (capable of forming sgRNA) using TTTR (where R is A or G) PAM to cleave single-stranded and double-stranded DNA target (Harrington et al., 2018; Karvelis et al., 2020; Takeda et al., 2021). Cas12f requires dimerization of

the guiding RNA. This catalytic effector complex recognizes sgRNA and target DNA, forming Cas12f-sgRNA-targetDNA complexes (Takeda et al., 2021). Through crRNA targeting of nucleic acid regions, it achieves exogenous DNA cleavage to combat re-invasion by phage or foreign DNA (Mohanraju et al., 2016; Hille et al., 2018; Takeda et al., 2021).

4.1 The widespread IS200/605 transposon family encodes TnpA and TnpB

Transposition plays a key role in remodeling the genomes of all organisms. Insertion sequences of the IS200/IS605 and IS607 families are among the simplest mobile genetic elements, containing only the genes required for their transposition and their regulation (Siguier et al., 2014; Koonin, 2016). These elements encode tnpA transposases, which are essential for mobilization, and usually carry an accessory tnpB gene, which is dispensable for transposition. Although the role of TnpA in IS200/IS605

transposon mobilization is well documented, the function of TnpB remains largely unknown. Three independent experiments on IS607, ISHp608 and ISDra2 elements (the latter two belonging to the IS200/IS605 family) have shown that TnpB is essential for transposition in *E. coli* and *D. radiodurans* (Ishino et al., 1987; Kersulyte et al., 2000, 2002). Interestingly, many IS elements (e.g., IS1341, IS809 and IS1136) encode TnpB as the sole protein (putative transposase), but supporting evidence for TnpB-mediated transposition is still lacking. As with other elements of the IS200/IS605 and IS607 families, these TnpB-only transposons lack terminal inverted repeat sequences (TIR) and target site replication (TSD). Studies have shown that prokaryotic TnpB proteins are encoded by bacterial transposons of the IS200/605 or IS607 families. TnpB-like proteins (Fanzor1 and Fanzor2) are found in a wide range of eukaryotic transposable elements (TES) and large dsDNA viruses that infect eukaryotic cells (Bao and Jurka, 2013). The simplest and oldest mobile genetic elements (MGEs) are insertion sequences from the IS200/IS605 and IS607 families. Typically, encode TnpA and TnpB genes, or isolated TnpA or TnpB genes, in various configurations and contain subterminal left-end (LE) and right-end (RE) palindromic elements at the end of the MGE (ISfinder database).

4.2 TnpB protein is the ancestor of Cas12 nuclease

The fully characterized *D. radiodurans* ISDra2 of the IS200/IS605 family consists of partially overlapping TnpA and TnpB genes flanked by LE and RE elements (May and Craig, 1996; Bernales et al., 1999; Islam et al., 2003; Hickman et al., 2010; Pasternak et al., 2010; Koonin, 2016). For a given IS605 transposon element, the TnpA protein binds to subterminal palindromic sequences located at the left and right ends (LE/RE) of the element, and catalyzes the cleavage and rejoining of DNA substrates through a “Peel and Paste” mechanism (He et al., 2015). During this process, TnpA recognizes a cleavage sequence positioned upstream of the LE (referred to as CL, spanning 4–5 nucleotides), which also serves as the target site (TS) for transposon insertion. In studies examining the editing activity of TnpB derived from ISDra2 (a member of the IS605 family), the aforementioned research proposed a “Peel, Paste and Copy” model (Karvelis et al., 2021). This model involves three key steps: (1) TnpA facilitates the excision of the transposon from its original genomic location and its subsequent insertion into a new site; (2) guided by right-end element RNA (reRNA), TnpB identifies the transposon-associated motif (TAM)—which appears to match the CL sequence—and induces a double-strand break at the transposon’s original position; (3) the resulting break is repaired via recombination with the allele containing the transposon. This function of TnpB as a homing endonuclease allows for the direct prediction of TnpB protein, reRNA, and TAM sequences from genomic data. Such predictability would pave the way for large-scale screening of TnpB as a compact genome-editing tool (Xiang et al., 2024). *D. radiodurans* ISDra2 TnpA is a very small Y1 transposase (140 aa) of the HUH family that excises a specific DNA strand near the 5'-TTGAT sequence to form a single-stranded transposon loop, which is then integrated 3' into a new position in the TTGAT target

to complete the transposition cycle without repeating the target site (Siguier et al., 2006; Pasternak et al., 2010).

Recent research has elucidated the evolutionary pathways of the CRISPR-Cas system, indicating that the prokaryotic transposon-encoded IscB and TnpB proteins may serve as ancestors of the Cas9 and Cas12 nucleases, respectively (Altae-Tran et al., 2021; Karvelis et al., 2021). These ancestors exhibited a small size, yet there was insufficient evidence for their nuclease activity was lacking; it was not until 2021 that IscB and TnpB were found to cleave dsDNA under the guidance of non-coding RNAs (ω RNA or reRNA), thereby confirming a similar working mechanism to the CRISPR-Cas system (Altae-Tran et al., 2021; Karvelis et al., 2021). TnpB is encoded by the IS200/IS605 and other prokaryotic transposon family and is presumed to be involved in transposon expansion. TnpB is widely distributed, with over a million copies present in the currently known genome. Nevertheless, previous studies have identified only one TnpB nuclease (ISDra2) that is active in human cells and inefficient. Additionally, TnpB targets key DNA-dependent cleavage components (e.g., reRNA) that may be associated with transposons and hence can be predicted based on transposon information, which would facilitate the development of tools.

OMEGA (obligate mobile element-guided activity) was discovered by researchers within the IscB protein family, which is believed to be the ancestors of CRISPR-Cas9 enzymes (Altae-Tran et al., 2021). Both IscB and TnpB belong to the IS200/IS605 superfamily transposon-encoding proteins. Among them, IscB is considered as the ancestor of Cas9, while TnpB is the ancestor of Cas12. They are collectively referred to as OMEGA proteins, which guide RNA called ω RNA, revealing their deep evolutionary association with transposons (Shmakov et al., 2017; Altae-Tran et al., 2021). Structurally, IscB contains RuvC and HNH nuclease domains homologous to Cas9, while TnpB possesses the Cas12 family-conserved RuvC domain. These core structures form the basis of their RNA-guided nuclease activity. Their ω RNA contains complex structures such as pseudoknots and stem loops, which are considered evolutionary precursors to crRNA and tracrRNA in CRISPR-Cas systems (Nakagawa et al., 2023; Sasnauskas et al., 2023). Additionally, OMEGA systems are widely distributed among trilobate organisms, not only abundant in prokaryotes but also present in eukaryotic proteins like Fanzor (homolog of TnpB) that exhibit RNA-guided nuclease activity (Altae-Tran et al., 2021; Sutandy et al., 2023; Burstein et al., 2017). This indicates that RNA-guided nuclease activity is an ancient and conserved feature initially serving selfish purposes like transposon retention and dissemination. Subsequently inherited by CRISPR-Cas systems, it has been adapted into immune tools against exogenous DNA. This evolution clearly demonstrates the evolutionary origin pathway of IscB and TnpB as RNA-guided nuclease systems. Systematic phylogenetic analysis reveals that all Cas9 proteins share a common evolutionary origin within the IscB lineage. Altae-Tran et al. (2021) proposed that IscB likely evolved from a compact RuvC nuclease, later acquiring the HNH domain (possibly through insertions of IshB-like proteins) to form the IscB family. Structural characterization further supports these conclusion (Kato et al., 2022; Schuler et al., 2022). Further research is needed to clarify TnpB’s evolutionary relationship with IscB and its origin from RuvC nuclease.

In addition the several *IscB* and *IsrB* families, most IS200/IS605 transposons encode RuvC-like endonucleases from another family, TnpB, which is thought to be the ancestor of the V-type CRISPR effector Cas12s (Shmakov et al., 2017; Altae-Tran et al., 2021). TnpB may also be the ancestor of the bigger protein Fanzors, which are encoded in different eukaryotic transposons (Pasternak et al., 2010). The TnpB family, including Fanzor, is significantly more diverse than the *IscB* family; an HMMER search in publicly available prokaryotic genomes identified more than one million TnpB motifs. Conserved non-coding regions were found downstream of the CDS of many TnpBs, indicating the presence of associated ncRNAs that could serve as RNA guides (Altae-Tran et al., 2021). These TnpB families are considered to be the ancestor of CRISPR-Cas12 enzymes.

However, there are important differences between TnpB and Cas12f nucleases. First, Cas12f nuclease uses gRNA from the CRISPR array, whereas TnpB uses right transposon element-derived reRNA as a guide. Second, TnpB is a monomer and requires a single reRNA molecule, whereas Cas12f nuclease is a dimer and binds to the crRNA (CRISPR RNA)-tracrRNA (transactivating crRNA) duplex (Bao et al., 2009; Cock et al., 2010). Target Adjacent Motif (TAM) and Protospacer Adjacent Motif (PAM) are critical components in genome editing tools that enable nuclease recognition of target DNA. Their core function is to ensure nuclease specificity in identifying and cleaving target sequences, directly impacting editing efficiency, accuracy, and target coverage. These elements form the foundation for effective genome editing. Finally, although the TAM (transposon-associated motif) sequence required for TnpB cleavage appears to be equivalent to the PAM sequence that licenses Cas12f cleavage, Cas12f proteins show a different PAM requirement (Bernales et al., 1999). However, the evolutionary relationship between Cas12 nucleases and the IS200/IS605/IS607 family of MGES reflected in the extent to which the PAM diversity of Cas12f nucleases correlates with the different TAM sequence requirements of TnpB proteins. The main characteristics of TnpB, Cas12f and Cas9 are summarized in Table 1.

The TAM of TnpB is closely associated with the insertion sequence of transposons, resulting in its target selection being limited to specific TAM sequences. For example, TnpB from radiation-resistant *Pseudomonas aeruginosa* only recognizes the 5'-TTGAT-3' TAM, while the archaeal *SisTnpB1* specifically targets TTAA. Although this high specificity reduces off-target risks, it significantly narrows the range of editable targets, particularly when genomic regions lacking corresponding TAM are unavailable. In contrast, Cas12f's PAM-dependent mechanism is more flexible and can be engineered for expansion. For instance, while natural Cas12f primarily recognizes T-rich PAMs, the engineered CnCas12f1 can recognize C-rich PAMs, and Cas12Pro (modified from Cas12n) even identifies 5'-ANN PAMs, achieving 80% insertion/inducement efficiency in human cells (Chen et al., 2023). This modifiability allows broader genomic coverage and wider target selection. Reported TAM recognition sequences of TnpB from different microbial sources vary by approximately 5–6 bases, making existing TnpB protein recognition windows smaller compared to Cas9's PAM recognition motifs (Sasnauskas et al., 2023; Xiang et al., 2024). Both *IscB* and TnpB utilize single non-coding RNAs for RNA-guided double-strand DNA cleavage, demonstrating genome-editing activity in human cells similar to

the CRISPR/Cas genome editing system. However, the *IscB* system operates by targeting cleavage at 3 nt upstream of the TAM sequence, mirroring the CRISPR/Cas9 mechanism. In contrast, the non-target strand cleavage occurs at 8 or 12 nt upstream of the TAM, ultimately producing 5' overhangs of 5 nt or 9 nt length. This characteristic differs from Cas9's blunt-end cleavage while aligning with the Cas12 protein's cutting pattern (Altae-Tran et al., 2021).

4.3 Action mechanism of TnpB

TnpB, the ancestor of CRISPR-Cas nucleases that is a reprogrammable RNA-directed functional nuclease. It is a TAM-dependent RNA-guided dsDNA nuclease, guided by the similarity of TnpB to the CRISPR-Cas12f effector complex that acts as an RNA-guided dsDNA nuclease (Bernales et al., 1999). It has been proposed that the approximately 16-nt 3' ends of reRNA may act as guide sequences to orient TnpB to its target and activate DNA cleavage. These ends are derived from DNA in the vicinity of the transposon, which is itself variable. The TTGAT sequence of the TnpB cleavage plasmid corresponds to the target region required for TnpA-mediated ISDra2 transposon excision and insertion. This region, known as TAM, is equivalent to the PAM sequence required by Cas9 or Cas12 nucleases to trigger DNA excision (Tavakoli and Derbyshire, 2001). TnpB cleaves the donor junction generated after transposon excision and proposes that TnpB-mediated DSB triggers homology-directed repair to transposon back to its original position (Rajpurohit et al., 2016). This process would be similar to the homing of group I introns facilitated by intron encoding nucleic acid endonucleases (Yang and Bennetzen, 2009). There, in MGES that contain both *tnpA* and *tnpB*, two types of transposition would be possible: (i) excision of the transposon and its insertion into a new site (catalyzed by TnpA); (ii) transposon "homing," i.e., which is the process by which TnpB cleaves the DNA in the non-transposon allele, triggering the demonstration of the RNA-guided dsDNA cleavage activity of TnpB. This direct experimental confirmation of the evolutionary scenario of class 2 CRISPR-CAS systems identifies MGES as the predecessor of Cas9 and Cas12 effectors (Robertson, 1981; Adams et al., 1997; Smith and Thorpe, 2002). A comparison of the sequences of TnpB and Cas12 family proteins revealed that they share a similar structural domain organization, which includes conserved RuvC endonuclease-like motifs. The miniature Cas12f nuclease is the closest TnpB neighbor on the evolutionary tree (Robertson, 1981; Bernales et al., 1999; Roberts et al., 2015). The typical TAM sequence for ISDra2 TnpB is 5'-TTGAT-3'. However, unlike the strict constraints imposed by CRISPR's PAM requirements, TnpB's TAM sequence demands exhibit greater flexibility in genome editing. The PAM sequences in CRISPR-Cas systems are highly regulated, with the most widely used *Streptococcus pyogenes* Cas9 requiring two G nucleotides as its PAM sequence – a requirement that significantly limits the number of target sites it can accommodate. In contrast, TnpB demonstrates relatively flexible TAM sequence requirements. A study conducted by Professor Qunxin She's research team at Shandong University on April 24, 2024 revealed that thermophilic archaeal TnpB can recognize diverse TAM variant sequences and mediate gene editing, including many non-casogenic TAM variants. The research also demonstrated that TnpB exhibits exceptional targeting specificity

TABLE 1 Comparative characteristics of TnpB, Cas12f and Cas9.

Characteristics	TnpB	Cas12f	Cas9
Structural domain	N-terminal fragment (REC), C-terminal fragment (NUC)	RuvC	HNH, RuvC
Protein size (amino acids, aa)	369–408 aa (ISDra2 TnpB, 408 aa; ISYmu1, 369 aa)	400–700 aa (Cas12f1, 422 aa)	1,000–1,400 aa (SpCas9, 1368 aa)
Guide RNA type	ω RNA, reRNA	tracrRNA, crRNA	tracrRNA, crRNA
PAM/TAM requirement	TAM (ISDra2 recognizes 5'-TTGAT-3', ISAam1 recognizes 5'-TTYAA-3')	PAM (commonly identified as 5'-TTTN-3')	PAM (5'-NGG-3')
Cleavage pattern	5' overhang	5' overhang (three cuts)	Blunt end
Collateral activity	Yes/no	Yes	No
Host application	Prokaryotes, plants, mammalian cells.	Prokaryotes, eukaryotes, mammalian cells.	Prokaryotes, eukaryotes, plants and mammalian cells.

at weak TAM sites. Leveraging this property, scientists successfully achieved efficient single-nucleotide editing through TnpB-based homologous recombination (Feng et al., 2024). In addition, the research of Professor Nan Peng's group at Huazhong Agricultural University also pointed out that SisTnpB1 only needs a non-conserved TAM sequence (5'-WNHNN-3') to induce the cleavage of target DNA (Xu et al., 2023). Current strategies to overcome the limitations of TAM sequences primarily leverage TnpB's recognition of weak TAM sequences. The use of diverse weak TAM sequences not only enables more flexible gene editing and improved cell survival rates, but also significantly expands the targeting scope. This approach may prove applicable to various CRISPR-Cas systems.

4.4 Applications of the TnpB protein

In the past 5 years, with the rapid development of gene editing technology, TnpB has shown significant potential in gene therapy, disease modeling and microbial engineering, and has become a new generation of miniature gene editing tools after the CRISPR-Cas system.

Altae-Tran et al. (2021) demonstrated that three different transposon-encoded proteins, IscB, IsrB, and TnpB, are naturally occurring, reprogrammable RNA-guided DNA nucleases. TnpB was first demonstrated to possess *in vitro* and *in vivo* cleavage activity for application as a genome editing tool in mammalian gene therapy (Altae-Tran et al., 2021). Experiments have demonstrated that TnpB editing in *E. coli* can efficiently cleave plasmids, as well as work in human cells. Karvelis et al. (2021) demonstrated that TnpB of *D. radiodurans* ISDra2 is an RNA-directed nuclease, which is induced by RNA, with a right-terminal element of the transposon that cleaves DNA next to the 5'-TTGAT transpo AAV-mediated son-related motif. TnpB can be reprogrammed to cleave DNA target sites in human cells and improves the efficiency and specificity by predicting RNA sequences using deep learning. Xiang et al. (2024) screened 78 TnpB systems and identified 33 active ones, including 5 that function in human cells. Notably, the ISAam1 and ISYmu1 TnpB outperformed three previously published Cas12f editors (e.g., Un1Cas12f1) in performance, demonstrating SaCas9-like capabilities while optimizing Nme2Cas9's editing efficiency.

However, TnpB transposase systems have not yet been applied to plant gene editing (Xiang et al., 2024). Li Q. et al. (2024) subsequently demonstrated the successful implementation of TnpB-mediated plant gene editing, revealing that the ISDra2 and ISYmu1 nuclease systems exhibit remarkably high efficiency. Notably, as TnpB nuclease specifically recognizes transactylated (TA)-rich target sequences in transacting actin (TAM), the TnpB system should complement the Cas9 system, which targets the adjacent NGG motif in the original spacer region containing guanine (G). This study establishes a foundation for developing alternative plant genome editing tools (Li Q. et al., 2024). The variants of TnpB with distinct activities (e.g., ISDra2, ISAam1, and ISYmu1) are proved in Table 2.

Sasnauskas et al. (2023) and Nakagawa et al. (2023), on the other hand, analyzed and demonstrated the cryo-electron microscopy structures of the TnpB protein-bound ω RNA complex from *D. radiodurans* in DNA-bound and free forms, respectively. They discovered that the ω RNA has a pseudo-knot structure and that the pseudo-knot in the sgRNA of the Cas12 protein is conserved (Nakagawa et al., 2023; Sasnauskas et al., 2023). These profiles of the protein's physical structure suggest that TnpB is the minimal structural and functional core of the Cas12 protein family. This framework enables the development of TnpB-based genome editing tools and enhances our comprehension of the evolution from transposon-encoded TnpB proteins to CRISPR-Cas12 effectors (Sasnauskas et al., 2023; Marquart et al., 2024).

Given previous experience optimizing Cas proteins and gRNAs, TnpB activity could be improved while addressing TAM limitations. The recent discovery of the structure of the TnpB-reRNA-DNA complex will hasten this process. Researchers have modified TnpB to enhance its interaction with DNA, thereby increasing the efficiency of gene editing. A deep learning model was developed to predict the efficiency of TnpB targeting, which increased the editing activity by 4.4-fold, and the editing efficiency of TnpB in the mouse brain reached 65.9% (Marquart et al., 2024). Wang et al. (2024) utilized ISDra2-TnpB for gene editing in mice and validated the method *in vivo* in an *in vivo*-based AAV system for genome editing. Furthermore, the study successfully developed a truncated ultra-compact TnpB editor (only 379 aa in size) with a shortened CTD structural domain, which is smaller in size but retains efficient gene editing capabilities in mammalian cells and

TABLE 2 The variants of TnpB with distinct activities.

TnpB variant	Source	Protein size (aa)	Editing efficiency/characteristics	TAM characteristics	References
ISDra2 TnpB	<i>Deinococcus radiodurans</i>	408	- Low efficiency of dsDNA cutting <i>in vitro</i>. - Requiring long crRNA (about 170 nt). - The first identified TnpB protease, highly efficient in bacteria (>80% target efficiency), and 30%–50% editing efficiency in human cells.	Identifying the TAM sequence 5'-TTGAT 3' which must be present upstream of the target sequence.	Karvelis et al., 2021
ISAam1 TnpB	<i>Anoxybacillus amylolyticus</i>	369	- The editing efficiency is lower than ISDra2. - The cutting window is located 18–23 bp downstream of PAM. - It supports simplified design of sgRNA (about 60 nt).	Identifying the TAM sequence 5'-TTTAA 3'.	Badon et al., 2024
ISYmu1 TnpB	<i>Youngiibacter multivorans</i>	382	- Editable but lower than the ISDra2 variant in human cells. - The smallest natural TnpB. - Plant cell editing efficiency reaches 20%–30% - Off-target rate is significantly lower than Cas9. - Showing DNA cutting activity but low editing efficiency.	The recognition of the 5'-TTGAT 3' TAM sequence was consistent with that of the ISDra2 TAM recognition sequence.	Badon et al., 2024

mice. The extremely compact non-Cas nuclease TnpB and the truncated ultra-mini ISDra2-TnpB can serve as very promising multifunctional tools that show great potential for a variety of genome editing applications (Wang et al., 2024). Li Q. et al. (2024) improved the nuclease activity of TnpB by stepwise modification of non-coding RNAs (ω RNAs or reRNAs) of the ISDra2 TnpB system. Modified TnpB- ω RNA system can be efficiently delivered *in vivo* with individual AAV and correct the disease phenotype in a mouse model of tyrosinemia, demonstrating its applicability for *in vivo* genome editing. In addition, against the Klkb1 gene target of hereditary angioedema (HAE), the TnpB- ω RNA system also showed high specificity for adult editing (Li Q. et al., 2024). It predicts the potential of miniature TnpB editing tools in the treatment of hereditary diseases, such as high cholesterol genetic defects, through gene editing. Chen et al. (2023) found, by evolutionary and biochemical analyses, that the miniature Cas12n nuclease (400–700 aa) was similar to the progenitor of the V-type CRISPR system - the TnpB, the transposition-associated nuclease, which may be the earliest evolutionary intermediate for TnpB to evolve into other V-type Cas nucleases. Cas12n nuclease recognizes a specific 5'-AAN PAM sequence, in which the A nucleotide at the -2 position is also required for TnpB. In addition, a highly effective CRISPR-Cas12n (called Cas12Pro) system with up to 80% insertion or deletion efficiency was developed for efficient base editing in human cells (Chen et al., 2023).

The above findings demonstrate that TnpB proteins with smaller molecular structures, which can be more easily delivered via viral vectors (e.g., AAV), are particularly important for practical applications in gene therapy. And it can be used as a powerful gene editing tool for precise gene editing in bacteria.

While the TnpB system (e.g., ISDra2-TnpB) has demonstrated high targeting efficiency in multiple studies, its off-target risks still require systematic evaluation. For instance, Wang et al.'s (2024) research further confirmed the significant potential of ISDra2-TnpB as an efficient tool for achieving site-specific modifications *in vitro* and *in vivo*. Their groundbreaking study using ISDra2-TnpB to generate mutant mice demonstrated its utility in genome editing based on *in vivo* adeno-associated virus (AAV). Additionally, the research introduced a truncated ultra-small TnpB

editor (<400 aa) generated by shortening the C-terminal domain (CTD) of TnpB. This engineered TnpB editor achieved effective gene editing in mammalian cells and mice, with a 9.5% editing efficiency for the target gene PCSK9 and no detectable off-target effects verified through whole-genome sequencing (Wang et al., 2024). The modified TnpB editor not only maintained high editing efficiency but also enhanced AAV payload flexibility. The University of Zurich team optimized the design of TnpB from *D. radiodurans* (ISDra2), resulting in TnpBmax for mammalian cells that achieved an average 4.4-fold improvement in editing efficiency. They also developed variants with mutations at the K76 site, which can recognize adjacent motifs (TAM) of alternative targets, thereby expanding the targeting range of ISDra2 TnpB. Through AI model predictions on 10,211 target sites, the optimized variant showed 75.3% efficiency in mouse liver, though some natural variants still carry non-specific cleavage risks. In a 2024 study by Xu Chunlong's team, engineered ω RNA significantly enhanced TnpB's targeting in tyrosinemia mouse models, though further RNA design optimization is needed to reduce off-target effects (Li Q. et al., 2024). While TnpB's size advantages have been fully validated, clinical applications still face challenges in specificity, delivery efficiency, and safety. Future optimizations will focus on improving editing accuracy and delivery efficiency. AI-assisted design could develop highly specific and off-target-resistant variants, while engineered ω RNA structures could further enhance targeting recognition. Additionally, developing TnpB-based base editing tools could reduce DNA double-strand break risks, paving the way for safer gene therapies. The coordinated development of these technical approaches will drive practical clinical applications of TnpB systems.

4.5 Current limitations and future challenges of the TnpB system

Although TnpB, as a supercompact RNA-guided nuclease (approximately 400 amino acids), has demonstrated significant potential in gene editing applications such as AAV delivery compatibility and efficient eukaryotic editing, it still faces several

critical limitations in widespread adoption that require further optimization and exploration. While proof-of-concept studies have been conducted in mammalian cells regarding editing efficiency, its overall performance remains inferior to CRISPR-Cas nuclease systems. The varying cleavage efficiencies across species also restrict its application in scenarios demanding high editing precision. Currently, TnpB is primarily used for gene knockout (indels), while base editing and prime editing systems remain underdeveloped (Komor et al., 2016; Tan et al., 2022; Xiang et al., 2025). In contrast, Cas9 and Cas12 have successfully integrated deaminases or reverse transcriptases, whereas similar tools for TnpB are still in their early stages of development. In terms of targeting range, TnpB's DNA cleavage strictly depends on TAM sequences (e.g., 5'-TTGAT-3' or 5'-TTTAA-3'), significantly limiting the flexibility in selecting edit sites. Compared to CRISPR-Cas9 (PAM NGG) or Cas12f (PAM TTN), TnpB's TAM appears less frequently across genomes, restricting its application in complex organisms like humans and plants. Recent protein engineering approaches (such as TnpBmax) have relaxed TAM restrictions by recognizing 5'-WNHNN-3', yet their editing efficiency still lags behind Cas9 variants. This poses a major obstacle for research and applications requiring precise gene editing across extensive regions. The editing efficiency and off-target effects of TnpB remain key limitations hindering its development. Wild-type TnpB typically achieves only 0%–20% editing efficiency in mammalian cells, significantly lower than Cas9's (>70%). Regarding off-target risks: Although TnpB's RuvC domain demonstrates specificity, its potential for unintended cleavage has not been systematically evaluated. This risk is particularly concerning in eukaryotic cells, where non-standard TAM recognition may lead to unexpected gene cuts.

The expression of TnpB in host organisms also shows significant variations. Most TnpB systems (such as ISDra2 and SisTnpB1) were initially validated in prokaryotes (e.g., bacteria and archaea), but their suitability for eukaryotic cells (e.g., mammals and plants) still requires optimization. Additionally, efficient delivery of TnpB systems to target cells within the body remains a major challenge. Although TnpB's compact structure theoretically offers advantages in *in vivo* delivery, practical implementation still faces difficulties in ensuring precise and efficient transport to specific cells while maintaining stable functionality. Current research on TnpB is predominantly confined to laboratory settings, with commercial applications limited to IDT's Cas9 kit, which restricts its broader application.

In summary, the TnpB system faces significant limitations and challenges in development and application, including editing efficiency, targeting range, applicability, and *in vivo* delivery, which require further research and technical optimization. Future improvements to the TnpB system can be pursued through multiple approaches: First, modifying TnpB via directed evolution or computational protein design to enhance its TAM sequence compatibility, thereby expanding the number of gene editing sites. Second, addressing the low editing efficiency in eukaryotic cells by optimizing codon usage, enhancing nuclear localization signals, and applying AI prediction models to improve gRNA design strategies. Third, developing TnpB-based base editors by fusing TnpB with deaminases or reverse transcriptases to achieve more precise single-base editing. Finally, employing high-throughput methods like whole genome sequencing (WGS) or CIRCLE-seq to systematically evaluate TnpB's off-target effects will ensure safety

for clinical applications. Advancing these research directions will significantly enhance the TnpB system's potential in gene editing.

5 Prospects

Deinococcus radiodurans exhibit extreme resistance to damage factors such as UV, drought and hydrogen peroxide, and their robust DNA repair mechanisms and efficient DNA uptake system make them attractive candidates for genetic engineering applications. The TnpB protein of *D. radiodurans* ISDra2 is believed to be the ancestor of the CRISPR nuclease Cas12f. The traditional CRISPR system has been limited in its application due to its large size. Although CRISPR-Cas12f is currently known as the smallest CRISPR system, it should be noted that while CRISPR-Cas12f is an effective and cost-efficient method for high-throughput pathogenic mutation screening, the sgRNA-mediated Cas12f targeted cleavage of ssDNA requires additional steps to generate ssDNA from dsDNA targets before detection. This particular characteristic of CRISPR-Cas12f could potentially increase the complexity of its application. Furthermore, the mechanism of action of TnpB protein and its related studies suggest that TnpB protein is expected to become a more convenient and practical gene editing tool in the future.

TnpB proteins are derived from mobile genetic elements such as transposons and are capable of inserting DNA sequences at specific genomic loci in a targeted manner. This integration process can be used to introduce exogenous DNA sequences, disrupt target genes or engineer specific genomic modifications. The natural diversity of TnpB homologs, including eukaryotic variants yet to be identified³³ and their miniature size suitable for adeno-associated viral delivery, opens up new horizons for human therapeutic applications. Li Z. et al. (2024) achieved disease phenotype correction in tyrosinemia mouse models by modifying the ω RNA of ISDra2 TnpB, with no significant off-target effects detected through whole-genome analysis. This approach can be extended to treat other genetic disorders such as hypercholesterolemia and hemophilia, particularly suitable for scenarios requiring long-term expression editing tools (Li Z. et al., 2024). As with the Cas system, homology searches and phylogenetic trees allowed us to identify the TnpB system and understand its diversity. At the same time the predictability of TAMs is further enhanced, which promises to further expand the editing scope of TnpB. With CRISPR-Cas research as a foundation, TnpB is expected to break through the limitations of TAMs and become a smaller, simpler, and more practical gene editing tool in the future. Li Z. et al. (2024) achieved a significant boost in editing efficiency by progressively modifying the ω RNA of ISDra2 TnpB. This strategy can be further enhanced by integrating deep learning models to predict ω RNA sequences [as demonstrated by Marquart et al. (2024) model that increased activity by 4.4 times], enabling efficient editing of complex genomic sites (such as oncogenes with repetitive sequences). A research team from the Institute of Zoology, Chinese Academy of Sciences and the School of Life Sciences at Southwest University has identified TnpB proteins including ISAam1 (369 amino acids), ISYmu1 (382 amino acids), and ISdGe10 through genome mining of prokaryotic genomes. These proteins demonstrated comparable editing efficiency to

SaCas9 in both human cells and rice, surpassing some Cas12f proteins and their variants (Xiang et al., 2024). Further screening of high-activity TnpB nuclease enzymes and optimization of their editing performance will facilitate the development of compact, efficient, and precise gene-editing tools to meet the demands of gene therapy and plant genetic improvement.

Author contributions

Z-QY: Conceptualization, Validation, Writing – original draft, Investigation, Visualization. M-JL: Conceptualization, Writing – original draft, Investigation. FA: Validation, Writing – original draft. C-ZJ: Validation, Writing – original draft, Writing – review & editing, Conceptualization. TL: Validation, Writing – review & editing. F-JJ: Validation, Writing – review & editing. K-SS: Validation, Writing – review & editing, Funding acquisition. LJ: Conceptualization, Resources, Validation, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This study was supported by IPET through the Agricultural Machinery/Equipment Localization Technology Development Program funded by MAFRA (321056), Republic of Korea.

References

- Adams, C. W., Nanassy, O., Johnson, R. C., and Hughes, K. T. (1997). Role of arginine-43 and arginine-69 of the Hin recombinase catalytic domain in the binding of Hin to the hix DNA recombination sites. *Mol. Microbiol.* 24, 1235–1247. doi: 10.1046/j.1365-2958.1997.4141789.x
- Altae-Tran, H., Kannan, S., Demircioglu, F. E., Oshiro, R., Nety, S. P., McKay, L. J., et al. (2021). The widespread IS200/IS605 transposon family encodes diverse programmable RNA-guided endonucleases. *Science* 374, 57–65. doi: 10.1126/science.abj6856
- Anderson, A. W., Nordan, H. C., Cain, R. F., Parrish, G., and Duggan, D. (1956). Studies on a radio-resistant *Micrococcus*. I. Isolation, morphology, cultural characteristics, and resistance to gamma radiation. *Food Technol.* 10, 575–577.
- Aquino-Jarquín, G. (2019). CRISPR-Cas14 is now part of the artillery for gene editing and molecular diagnostic. *Nanomedicine* 18, 428–431. doi: 10.1016/j.nano.2019.03.006
- Badon, I. W., Oh, Y., Kim, H. J., and Lee, S. H. (2024). Recent application of CRISPR-Cas12 and OMEGA system for genome editing. *Mol. Ther.* 32, 32–43. doi: 10.1016/j.ymthe.2023.11.013
- Bao, W., and Jurka, J. (2013). Homologues of bacterial TnpB_{IS605} are widespread in diverse eukaryotic transposable elements. *Mob. DNA* 4:12. doi: 10.1186/1759-8753-4-12
- Bao, W., Jurka, M. G., Kapitonov, V. V., and Jurka, J. (2009). New superfamilies of eukaryotic DNA transposons and their internal divisions. *Mol. Biol. Evol.* 26, 983–993. doi: 10.1093/molbev/msp013
- Bauche, C., and Laval, J. (1999). Repair of oxidized bases in the extremely radiation-resistant bacterium *Deinococcus radiodurans*. *J. Bacteriol.* 181, 262–269. doi: 10.1128/jb.181.1.262-269.1999
- Bentchikou, E., Servant, P., Coste, G., and Sommer, S. (2010). A major role of the RecFOR pathway in DNA double-strand-break repair through ESDSA in *Deinococcus radiodurans*. *PLoS Genet.* 6:e1000774. doi: 10.1371/journal.pgen.1000774
- Bernales, I., Mendiola, M. V., and de la Cruz, F. (1999). Intramolecular transposition of insertion sequence IS91 results in second-site simple insertions. *Mol. Microbiol.* 33, 223–234. doi: 10.1046/j.1365-2958.1999.01432.x
- Bigelyte, G., Young, J. K., Karvelis, T., Budre, K., Zedaveinyte, R., Djukanovic, V., et al. (2021). Miniature type V-F CRISPR-Cas nucleases enable targeted DNA modification in cells. *Nat. Commun.* 12:6191. doi: 10.1038/s41467-021-26469-4
- Bihani, S. C., Panicker, L., Rajpurohit, Y. S., Misra, H. S., and Kumar, V. (2018). drFrnE represents a hitherto unknown class of eubacterial cytoplasmic disulfide oxidoreductases. *Antioxid. Redox Signal.* 28, 296–310. doi: 10.1089/ars.2016.6960
- Burstein, D., Harrington, L. B., Strutt, S. C., Probst, A. J., Anantharaman, K., Thomas, B. C., et al. (2017). New CRISPR-Cas systems from uncultivated microbes. *Nature* 542, 237–241. doi: 10.1038/nature21059
- Chen, J. S., Ma, E., Harrington, L. B., Da Costa, M., Tian, X., Palefsky, J. M., et al. (2018). CRISPR-Cas12a target binding unleashes indiscriminate single-stranded DNase activity. *Science* 360, 436–439. doi: 10.1126/science.aar6245
- Chen, W., Ma, J., Wu, Z., Wang, Z., Zhang, H., Fu, W., et al. (2023). Cas12n nucleases, early evolutionary intermediates of type V CRISPR, comprise a distinct family of miniature genome editors. *Mol. Cell* 83:2768–2780.e2766. doi: 10.1016/j.molcel.2023.06.014
- Chylinski, K., Makarova, K. S., Charpentier, E., and Koonin, E. V. (2014). Classification and evolution of type II CRISPR-Cas systems. *Nucleic Acids Res.* 42, 6091–6105. doi: 10.1093/nar/gku241
- Cock, J. M., Sterck, L., Rouzé, P., Scornet, D., Allen, A. E., Amoutzias, G., et al. (2010). The Ectocarpus genome and the independent evolution of multicellularity in brown algae. *Nature* 465, 617–621. doi: 10.1038/nature09016
- Cox, M. M., and Battista, J. R. (2005). *Deinococcus radiodurans* - the consummate survivor. *Nat. Rev. Microbiol.* 3, 882–892. doi: 10.1038/nrmicro1264
- Desai, S. S., Rajpurohit, Y. S., Misra, H. S., and Deobagkar, D. N. (2011). Characterization of the role of the RadS/RadR two-component system in the radiation resistance of *Deinococcus radiodurans*. *Microbiology* 157, 2974–2982. doi: 10.1099/mic.0.049361-0
- Feng, X., Xu, R., Liao, J., Zhao, J., Zhang, B., Xu, X., et al. (2024). Flexible TAM requirement of TnpB enables efficient single-nucleotide editing with expanded targeting scope. *Nat. Commun.* 15:3464. doi: 10.1038/s41467-024-47697-4

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Gootenberg, J. S., Abudayyeh, O. O., Lee, J. W., Essletzbichler, P., Dy, A. J., Joung, J., et al. (2017). Nucleic acid detection with CRISPR-Cas13a/C2c2. *Science* 356, 438–442. doi: 10.1126/science.aam9321
- Grove, A., and Wilkinson, S. P. (2005). Differential DNA binding and protection by dimeric and dodecameric forms of the ferritin homolog Dps from *Deinococcus radiodurans*. *J. Mol. Biol.* 347, 495–508. doi: 10.1016/j.jmb.2005.01.055
- GuanJun, G., Lu, F., HuiMing, L., and YueJin, H. (2008). Engineering *Deinococcus radiodurans* into biosensor to monitor radioactivity and genotoxicity in environment. *Chin. Sci. Bull.* 53, 1675–1681. doi: 10.1007/s11434-008-0224-6
- Haft, D. H., Selengut, J., Mongodin, E. F., and Nelson, K. E. (2005). A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. *PLoS Comput. Biol.* 1:e60. doi: 10.1371/journal.pcbi.0010060
- Han, J. M., Mwititi, G., Yeom, S. J., Lim, J., Kim, W. S., Lim, S., et al. (2025). Radiation-Resistant bacteria *Deinococcus radiodurans*-derived extracellular vesicles as potential radioprotectors. *Adv. Health. Mater.* 14:e2403192. doi: 10.1002/adhm.202403192
- Harrington, L. B., Burstein, D., Chen, J. S., Paez-Espino, D., Ma, E., Witte, I. P., et al. (2018). Programmed DNA destruction by miniature CRISPR-Cas14 enzymes. *Science* 362, 839–842. doi: 10.1126/science.aav4294
- He, S., Corneloup, A., Guynet, C., Lavatine, L., Caumont-Sarcos, A., Siguier, P., et al. (2015). The IS200/IS605 family and “peel and paste” single-strand transposition mechanism. *Microbiol. Spectr.* 3:9. doi: 10.1128/microbiolspec.MDNA3-0039-2014
- Hickman, A. B., James, J. A., Barabas, O., Pasternak, C., Ton-Hoang, B., Chandler, M., et al. (2010). DNA recognition and the precleavage state during single-stranded DNA transposition in *D. radiodurans*. *EMBO J.* 29, 3840–3852. doi: 10.1038/emboj.2010.241
- Hille, F., Richter, H., Wong, S. P., Bratovič, M., Ressel, S., and Charpentier, E. (2018). The biology of CRISPR-Cas: Backward and forward. *Cell* 172, 1239–1259. doi: 10.1016/j.cell.2017.11.032
- Ishino, Y., Shinagawa, H., Makino, K., Amemura, M., and Nakata, A. (1987). Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *J. Bacteriol.* 169, 5429–5433. doi: 10.1128/jb.169.12.5429-5433.1987
- Islam, S. M., Hua, Y., Ohba, H., Satoh, K., Kikuchi, M., Yanagisawa, T., et al. (2003). Characterization and distribution of IS8301 in the radioresistant bacterium *Deinococcus radiodurans*. *Genes Genet. Syst.* 78, 319–327. doi: 10.1266/ggs.78.319
- Jansen, R., Embden, J. D., Gaastra, W., and Schouls, L. M. (2002). Identification of genes that are associated with DNA repeats in prokaryotes. *Mol. Microbiol.* 43, 1565–1575. doi: 10.1046/j.1365-2958.2002.02839.x
- Jin, M., Xiao, A., Zhu, L., Zhang, Z., Huang, H., and Jiang, L. (2019). The diversity and commonalities of the radiation-resistance mechanisms of *Deinococcus* and its up-to-date applications. *AMB Exp.* 9:138. doi: 10.1186/s13568-019-0862-x
- 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. doi: 10.1126/science.1225829
- Jung, K.-W., Lim, S., and Bahn, Y.-S. (2017). Microbial radiation-resistance mechanisms. *J. Microbiol.* 55, 499–507. doi: 10.1007/s12275-017-7242-5
- Kamble, V. A., Rajpurohit, Y. S., Srivastava, A. K., and Misra, H. S. (2010). Increased synthesis of signaling molecules coincides with reversible inhibition of nucleolytic activity during postirradiation recovery of *Deinococcus radiodurans*. *FEMS Microbiol. Lett.* 303, 18–25. doi: 10.1111/j.1574-6968.2009.01855.x
- Kapitonov, V. V., Makarova, K. S., and Koonin, E. V. (2015). ISC, a novel group of bacterial and archaeal DNA transposons that encode Cas9 homologs. *J. Bacteriol.* 198, 797–807. doi: 10.1128/jb.00783-15
- Karvelis, T., Bigelyte, G., Young, J. K., Hou, Z., Zedaveinyte, R., Budre, K., et al. (2020). PAM recognition by miniature CRISPR-Cas12f nucleases triggers programmable double-stranded DNA target cleavage. *Nucleic Acids Res.* 48, 5016–5023. doi: 10.1093/nar/gkaa208
- Karvelis, T., Druteika, G., Bigelyte, G., Budre, K., Zedaveinyte, R., Silanskas, A., et al. (2021). Transposon-associated TnpB is a programmable RNA-guided DNA endonuclease. *Nature* 599, 692–696. doi: 10.1038/s41586-021-04058-1
- Kato, K., Okazaki, S., Kannan, S., Altae-Tran, H., Esra Demircioglu, F., Isayama, Y., et al. (2022). Structure of the IscB- ω RNA ribonucleoprotein complex, the likely ancestor of CRISPR-Cas9. *Nat. Commun.* 13:6719. doi: 10.1038/s41467-022-34378-3
- Kersulyte, D., Mukhopadhyay, A. K., Shirai, M., Nakazawa, T., and Berg, D. E. (2000). Functional organization and insertion specificity of IS607, a chimeric element of *Helicobacter pylori*. *J. Bacteriol.* 182, 5300–5308. doi: 10.1128/jb.182.19.5300-5308.2000
- Kersulyte, D., Velapatiño, B., Dailide, G., Mukhopadhyay, A. K., Ito, Y., Cahuayme, L., et al. (2002). Transposable element ISHp608 of *Helicobacter pylori*: Nonrandom geographic distribution, functional organization, and insertion specificity. *J. Bacteriol.* 184, 992–1002. doi: 10.1128/jb.184.4.992-1002.2002
- Kim, D. Y., Lee, J. M., Moon, S. B., Chin, H. J., Park, S., Lim, Y., et al. (2022). Efficient CRISPR editing with a hypercompact Cas12f1 and engineered guide RNAs delivered by adeno-associated virus. *Nat. Biotechnol.* 40, 94–102. doi: 10.1038/s41587-021-01009-z
- Knott, G. J., and Doudna, J. A. (2018). CRISPR-Cas guides the future of genetic engineering. *Science* 361, 866–869. doi: 10.1126/science.aat5011
- Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A., and Liu, D. R. (2016). Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* 533, 420–424. doi: 10.1038/nature17946
- Koonin, E. V. (2016). Viruses and mobile elements as drivers of evolutionary transitions. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 371:42. doi: 10.1098/rstb.2015.0442
- Kota, S., and Misra, H. S. (2006). PprA: A protein implicated in radioresistance of *Deinococcus radiodurans* stimulates catalase activity in *Escherichia coli*. *Appl. Microbiol. Biotechnol.* 72, 790–796. doi: 10.1007/s00253-006-0340-7
- Kota, S., Charaka, V. K., Ringgaard, S., Waldor, M. K., and Misra, H. S. (2014). PprA contributes to *Deinococcus radiodurans* resistance to nalidixic acid, genome maintenance after DNA damage and interacts with deinococcal topoisomerases. *PLoS One* 9:e85288. doi: 10.1371/journal.pone.0085288
- Li, Q., Wang, Y., Hou, Z., Zong, H., Wang, X., Zhang, Y. E., et al. (2024). Genome editing in plants using the TnpB transposase system. *aBIOTECH* 5, 225–230. doi: 10.1007/s42994-024-00172-6
- Li, Z., Guo, R., Sun, X., Li, G., Shao, Z., Huo, X., et al. (2024). Engineering a transposon-associated TnpB- ω RNA system for efficient gene editing and phenotypic correction of a tyrosinaemia mouse model. *Nat. Commun.* 15:831. doi: 10.1038/s41467-024-45197-z
- Liang, Y., Guo, K., Jiang, F., Peng, G., and Xiao, F. (2024). Research on urease-modified *Deinococcus radiodurans* and the control mechanism of the contamination of U(VI). *J. Radioanal. Nucl. Chem.* 333, 4091–4106. doi: 10.1007/s10967-024-09563-x
- Lim, S., Jung, J. H., Blanchard, L., and de Groot, A. (2019). Conservation and diversity of radiation and oxidative stress resistance mechanisms in *Deinococcus* species. *FEMS Microbiol. Rev.* 43, 19–52. doi: 10.1093/femsre/fuy037
- Liu, F., Li, N., and Zhang, Y. (2023). The radioresistant and survival mechanisms of *Deinococcus radiodurans*. *Radiat. Med. Protect.* 4, 70–79. doi: 10.1016/j.radmp.2023.03.001
- Liu, J. J., Orlova, N., Oakes, B. L., Ma, E., Spinner, H. B., Baney, K. L. M., et al. (2019). CasX enzymes comprise a distinct family of RNA-guided genome editors. *Nature* 566, 218–223. doi: 10.1038/s41586-019-0908-x
- Ma, J., Yan, F., Chen, F., Jiang, L., Li, J., and Chen, L. (2014). C18-functionalized magnetic silica nanoparticles for solid phase extraction of microcystin-Lr in reservoir water samples followed by HPLC-DAD determination. *J. Liquid Chromatogr. Relat. Technol.* 38:6611. doi: 10.1080/10826076.2014.936611
- Makarova, K. S., Aravind, L., Wolf, Y. I., and Koonin, E. V. (2011). Unification of Cas protein families and a simple scenario for the origin and evolution of CRISPR-Cas systems. *Biol. Direct.* 6:38. doi: 10.1186/1745-6150-6-38
- Makarova, K. S., Wolf, Y. I., Iranzo, J., Shmakov, S. A., Alkhnbashi, O. S., Brouns, S. J. J., et al. (2020). Evolutionary classification of CRISPR-Cas systems: A burst of class 2 and derived variants. *Nat. Rev. Microbiol.* 18, 67–83. doi: 10.1038/s41579-019-0299-x
- Marquart, K. F., Mathis, N., Mollaysa, A., Müller, S., Kissling, L., Rothgangl, T., et al. (2024). Effective genome editing with an enhanced ISDra2 TnpB system and deep learning-predicted ω RNAs. *Nat. Methods* 21, 2084–2093. doi: 10.1038/s41592-024-02418-z
- Maurya, G. K., Modi, K., Banerjee, M., Chaudhary, R., Rajpurohit, Y. S., and Misra, H. S. (2018). Phosphorylation of FtsZ and FtsA by a DNA damage-responsive Ser/Thr protein kinase affects their functional interactions in *Deinococcus radiodurans*. *mSphere* 3:e00325–18. doi: 10.1128/mSphere.00325-18
- May, E. W., and Craig, N. L. (1996). Switching from cut-and-paste to replicative Tn7 transposition. *Science* 272, 401–404. doi: 10.1126/science.272.5260.401
- Minton, K. W., and Daly, M. J. (1995). A model for repair of radiation-induced DNA double-strand breaks in the extreme radiophile *Deinococcus radiodurans*. *Bioessays* 17, 457–464. doi: 10.1002/bies.950170514
- Mishra, S., Chaudhary, R., Singh, S., Kota, S., and Misra, H. S. (2019). Guanine quadruplex DNA regulates gamma radiation response of genome functions in the radioresistant bacterium *Deinococcus radiodurans*. *J. Bacteriol.* 201:e00154–19. doi: 10.1128/jb.00154-19
- Mishra, S., Kota, S., Chaudhary, R., and Misra, H. S. (2021). Guanine quadruplexes and their roles in molecular processes. *Crit. Rev. Biochem. Mol. Biol.* 56, 482–499. doi: 10.1080/10409238.2021.1926417
- Misra, H. S., and Rajpurohit, Y. S. (2024). DNA damage response and cell cycle regulation in bacteria: A twist around the paradigm. *Front. Microbiol.* 15:1389074. doi: 10.3389/fmicb.2024.1389074
- Misra, H. S., Rajpurohit, Y. S., and Khairnar, N. P. (2012). Pyrroloquinoline-quinone and its versatile roles in biological processes. *J. Biosci.* 37, 313–325. doi: 10.1007/s12038-012-9195-5

- Mohanraju, P., Makarova, K. S., Zetsche, B., Zhang, F., Koonin, E. V., and van der Oost, J. (2016). Diverse evolutionary roots and mechanistic variations of the CRISPR-Cas systems. *Science* 353:aad5147. doi: 10.1126/science.aad5147
- Nakagawa, R., Hirano, H., Omura, S. N., Nety, S., Kannan, S., Altae-Tran, H., et al. (2023). Cryo-EM structure of the transposon-associated TnpB enzyme. *Nature* 616, 390–397. doi: 10.1038/s41586-023-05933-9
- O'Connell, M. R. (2019). Molecular mechanisms of RNA targeting by Cas13-containing Type VI CRISPR-Cas systems. *J. Mol. Biol.* 431, 66–87. doi: 10.1016/j.jmb.2018.06.029
- Pasternak, C., Ton-Hoang, B., Coste, G., Bailone, A., Chandler, M., and Sommer, S. (2010). Irradiation-induced *Deinococcus radiodurans* genome fragmentation triggers transposition of a single resident insertion sequence. *PLoS Genet.* 6:e1000799. doi: 10.1371/journal.pgen.1000799
- Pausch, P., Al-Shayeb, B., Bisom-Rapp, E., Tsuchida, C. A., Li, Z., Cress, B. F., et al. (2020). CRISPR-CasΦ from huge phages is a hypercompact genome editor. *Science* 369, 333–337. doi: 10.1126/science.abb1400
- Rai, S. N., and Dutta, T. (2024). A novel ionizing radiation-induced small RNA, DrsS, promotes the detoxification of reactive oxygen species in *Deinococcus radiodurans*. *Appl. Environ. Microbiol.* 90:e0153823. doi: 10.1128/aem.01538-23
- Rajpurohit, Y. S., and Misra, H. S. (2010). Characterization of a DNA damage-inducible membrane protein kinase from *Deinococcus radiodurans* and its role in bacterial radioresistance and DNA strand break repair. *Mol. Microbiol.* 77, 1470–1482. doi: 10.1111/j.1365-2958.2010.07301.x
- Rajpurohit, Y. S., and Misra, H. S. (2013). DR1769, a protein with N-terminal beta propeller repeats and a low-complexity hydrophilic tail, plays a role in desiccation tolerance of *Deinococcus radiodurans*. *J. Bacteriol.* 195, 3888–3896. doi: 10.1128/jb.00418-13
- Rajpurohit, Y. S., Bihani, S. C., Waldor, M. K., and Misra, H. S. (2016). Phosphorylation of *Deinococcus radiodurans* RecA regulates its activity and may contribute to radioresistance. *J. Biol. Chem.* 291, 16672–16685. doi: 10.1074/jbc.M116.736389
- Rajpurohit, Y. S., Desai, S. S., and Misra, H. S. (2013). Pyrroloquinoline quinone and a quinoprotein kinase support γ -radiation resistance in *Deinococcus radiodurans* and regulate gene expression. *J. Basic Microbiol.* 53, 518–531. doi: 10.1002/jobm.201100650
- Rajpurohit, Y. S., Sharma, D. K., and Misra, H. S. (2021). PprA protein inhibits DNA strand exchange and ATP hydrolysis of *Deinococcus* RecA and regulates the recombination in gamma-irradiated cells. *Front. Cell. Dev. Biol.* 9:636178. doi: 10.3389/fcell.2021.636178
- Roberts, R. J., Vincze, T., Posfai, J., and Macelis, D. (2015). REBASE—a database for DNA restriction and modification: Enzymes, genes and genomes. *Nucleic Acids Res.* 43, D298–D299. doi: 10.1093/nar/gku1046
- Robertson, D. S. (1981). Mutator activity in maize: Timing of its activation in ontogeny. *Science* 213, 1515–1517. doi: 10.1126/science.213.4515.1515
- Sadowska-Bartos, I., and Bartosz, G. (2023). Antioxidant defense of *Deinococcus radiodurans*: How does it contribute to extreme radiation resistance? *Int. J. Radiat. Biol.* 99, 1803–1829. doi: 10.1080/09553002.2023.2241895
- Saito, M., Xu, P., Faure, G., Maguire, S., Kannan, S., Altae-Tran, H., et al. (2023). Fanzor is a eukaryotic programmable RNA-guided endonuclease. *Nature* 620, 660–668. doi: 10.1038/s41586-023-06356-2
- Sasnauskas, G., Tamulaitiene, G., Druteika, G., Carabias, A., Silanskas, A., Kazlauskas, D., et al. (2023). TnpB structure reveals minimal functional core of Cas12 nuclease family. *Nature* 616, 384–389. doi: 10.1038/s41586-023-05826-x
- Savage, D. F. (2019). Cas14: Big Advances from Small CRISPR Proteins. *Biochemistry* 58, 1024–1025. doi: 10.1021/acs.biochem.9b00035
- Schuler, G., Hu, C., and Ke, A. (2022). Structural basis for RNA-guided DNA cleavage by IscB- ω RNA and mechanistic comparison with Cas9. *Science* 376, 1476–1481. doi: 10.1126/science.abq7220
- Sharma, D. K., Soni, I., and Rajpurohit, Y. S. (2025). Surviving the storm: Exploring the role of natural transformation in nutrition and DNA repair of stressed *Deinococcus radiodurans*. *Appl. Environ. Microbiol.* 91:e0137124. doi: 10.1128/aem.01371-24
- Sharma, D. K., Soni, I., Misra, H. S., and Rajpurohit, Y. S. (2024). Natural transformation-specific DprA coordinate DNA double-strand break repair pathways in heavily irradiated *D. radiodurans*. *Appl. Environ. Microbiol.* 90:e0194823. doi: 10.1128/aem.01948-23
- Shmakov, S., Abudayyeh, O. O., Makarova, K. S., Wolf, Y. I., Gootenberg, J. S., Semenova, E., et al. (2015). Discovery and functional characterization of diverse class 2 CRISPR-Cas systems. *Mol. Cell* 60, 385–397. doi: 10.1016/j.molcel.2015.10.008
- Shmakov, S., Smargon, A., Scott, D., Cox, D., Pyzocha, N., Yan, W., et al. (2017). Diversity and evolution of class 2 CRISPR-Cas systems. *Nat. Rev. Microbiol.* 15, 169–182. doi: 10.1038/nrmicro.2016.184
- Siguier, P., Gourbeyre, E., and Chandler, M. (2014). Bacterial insertion sequences: Their genomic impact and diversity. *FEMS Microbiol. Rev.* 38, 865–891. doi: 10.1111/1574-6976.12067
- Siguier, P., Perochon, J., Lestrade, L., Mahillon, J., and Chandler, M. (2006). ISfinder: The reference centre for bacterial insertion sequences. *Nucleic Acids Res.* 34, D32–D36. doi: 10.1093/nar/gkj014
- Smith, M. C., and Thorpe, H. M. (2002). Diversity in the serine recombinases. *Mol. Microbiol.* 44, 299–307. doi: 10.1046/j.1365-2958.2002.02891.x
- Song, X., Chen, Z., Sun, W., Yang, H., Guo, L., Zhao, Y., et al. (2024). CRISPR-AsCas12f1 couples out-of-protospacer DNA unwinding with exonuclease activity in the sequential target cleavage. *Nucleic Acids Res.* 52, 14030–14042. doi: 10.1093/nar/gkac989
- Sutandy, F. X. R., Gößner, I., Tascher, G., and Münch, C. (2023). A cytosolic surveillance mechanism activates the mitochondrial UPR. *Nature* 618, 849–854. doi: 10.1038/s41586-023-06142-0
- Takeda, S. N., Nakagawa, R., Okazaki, S., Hirano, H., Kobayashi, K., Kusakizako, T., et al. (2021). Structure of the miniature type V-F CRISPR-Cas effector enzyme. *Mol. Cell* 81:558–570.e553. doi: 10.1016/j.molcel.2020.11.035
- Tan, J., Zeng, D., Zhao, Y., Wang, Y., Liu, T., Li, S., et al. (2022). PhieABEs: A PAM-less/free high-efficiency adenine base editor toolbox with wide target scope in plants. *Plant Biotechnol. J.* 20, 934–943. doi: 10.1111/pbi.13774
- Tavakoli, N. P., and Derbyshire, K. M. (2001). Tipping the balance between replicative and simple transposition. *EMBO J.* 20, 2923–2930. doi: 10.1093/emboj/20.11.2923
- Timmins, J., and Moe, E. (2016). A decade of biochemical and structural studies of the DNA Repair machinery of *Deinococcus radiodurans*: Major findings, functional and mechanistic insight and challenges. *Comput. Struct. Biotechnol. J.* 14, 168–176. doi: 10.1016/j.csbj.2016.04.001
- Wang, M., Sun, Z., Liu, Y., Yin, P., Liang, C., Tan, L., et al. (2024). Hypercompact TnpB and truncated TnpB systems enable efficient genome editing in vitro and in vivo. *Cell Discov.* 10:31. doi: 10.1038/s41421-023-00645-w
- Wang, Y., Sang, S., Zhang, X., Tao, H., Guan, Q., and Liu, C. (2021). Efficient genome editing by a miniature CRISPR-AsCas12f1 nuclease in *Bacillus anthracis*. *Front. Bioeng. Biotechnol.* 9:825493. doi: 10.3389/fbioe.2021.825493
- Wei, Y., Yang, Z., Zong, C., Wang, B., Ge, X., Tan, X., et al. (2021). trans single-stranded DNA cleavage via CRISPR/Cas14a1 activated by target RNA without destruction. *Angew. Chem. Int. Ed. Engl.* 60, 24241–24247. doi: 10.1002/anie.202110384
- Wiedenheft, B., van Duijn, E., Bultema, J. B., Waghmare, S. P., Zhou, K., Barendregt, A., et al. (2011). RNA-guided complex from a bacterial immune system enhances target recognition through seed sequence interactions. *Proc. Natl. Acad. Sci. U.S.A.* 108, 10092–10097. doi: 10.1073/pnas.1102716108
- Wu, P., Ye, X., Wang, D., Gong, F., Wei, X., Xiang, S., et al. (2022). A novel CRISPR/Cas14a system integrated with 2D porphyrin metal-organic framework for microcystin-LR determination through a homogeneous competitive reaction. *J. Hazard. Mater.* 424, 127690. doi: 10.1016/j.jhazmat.2021.127690
- Wu, Z., Zhang, Y., Yu, H., Pan, D., Wang, Y., Wang, Y., et al. (2021). Programmed genome editing by a miniature CRISPR-Cas12f nuclease. *Nat. Chem. Biol.* 17, 1132–1138. doi: 10.1038/s41589-021-00868-6
- Xiang, G., Li, Y., Sun, J., Huo, Y., Cao, S., Cao, Y., et al. (2024). Evolutionary mining and functional characterization of TnpB nucleases identify efficient miniature genome editors. *Nat. Biotechnol.* 42, 745–757. doi: 10.1038/s41587-023-01857-x
- Xiang, W., Lin, X., Yang, Y., Huang, L., Chen, Y., Chen, J., et al. (2025). Cas12h is a crRNA-guided DNA nickase that can be utilized for precise gene editing. *Cell Rep.* 44:115718. doi: 10.1016/j.celrep.2025.115718
- Xiao, R., Li, Z., Wang, S., Han, R., and Chang, L. (2021). Structural basis for substrate recognition and cleavage by the dimerization-dependent CRISPR-Cas12f nuclease. *Nucleic Acids Res.* 49, 4120–4128. doi: 10.1093/nar/gkab179
- Xin, C., Yin, J., Yuan, S., Ou, L., Liu, M., Zhang, W., et al. (2022). Comprehensive assessment of miniature CRISPR-Cas12f nucleases for gene disruption. *Nat. Commun.* 13:5623. doi: 10.1038/s41467-022-33346-1
- Xu, X., Chemparathy, A., Zeng, L., Kempton, H. R., Shang, S., Nakamura, M., et al. (2021). Engineered miniature CRISPR-Cas system for mammalian genome regulation and editing. *Mol. Cell* 81:4333–4345.e4334. doi: 10.1016/j.molcel.2021.08.008
- Xu, Y., Liu, T., Wang, J., Xiong, B., Liu, L., and Peng, N. (2023). Reprogramming an RNA-guided archaeal TnpB endonuclease for genome editing. *Cell Discov.* 9:112. doi: 10.1038/s41421-023-00615-2
- Yan, W. X., Hunnewell, P., Alfonse, L. E., Carte, J. M., Keston-Smith, E., Sothiselvam, S., et al. (2019). Functionally diverse type V CRISPR-Cas systems. *Science* 363, 88–91. doi: 10.1126/science.aav7271
- Yang, F., Huang, F., Feng, H., Wei, J., Massey, I. Y., Liang, G., et al. (2020). A complete route for biodegradation of potentially carcinogenic cyanotoxin microcystin-LR in a novel indigenous bacterium. *Water Res.* 174:115638. doi: 10.1016/j.watres.2020.115638

- Yang, L., and Bennetzen, J. L. (2009). Structure-based discovery and description of plant and animal Helitrons. *Proc. Natl. Acad. Sci. U.S.A.* 106, 12832–12837. doi: 10.1073/pnas.0905563106
- Zahradka, K., Slade, D., Bailone, A., Sommer, S., Auerbeck, D., Petranovic, M., et al. (2006). Reassembly of shattered chromosomes in *Deinococcus radiodurans*. *Nature* 443, 569–573. doi: 10.1038/nature05160
- Zetsche, B., Gootenberg, J. S., Abudayyeh, O. O., Slaymaker, I. M., Makarova, K. S., Essletzbichler, P., et al. (2015). Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell* 163, 759–771. doi: 10.1016/j.cell.2015.09.038
- Zhang, H. Y., Li, X. J., Gao, N., and Chen, L. L. (2009). Antioxidants used by *Deinococcus radiodurans* and implications for antioxidant drug discovery. *Nat. Rev. Microbiol.* 7:476; author reply 476. doi: 10.1038/nrmicro2073-c1
- Zimmerman, J. M., and Battista, J. R. (2005). A ring-like nucleoid is not necessary for radioresistance in the Deinococcaceae. *BMC Microbiol.* 5:17. doi: 10.1186/1471-2180-5-17