Nucleic Acids Research, 2025, **53**, gkaf357
https://doi.org/10.1093/nar/gkaf357
Genomics

Unveiling the multifaceted domain polymorphism of the Menshen antiphage system

Huan Li[1], Yongjun Tan[1], Dwaipayan Basu[2], Kevin D. Corbett[3,4], Dapeng Zhang[1,5,*]

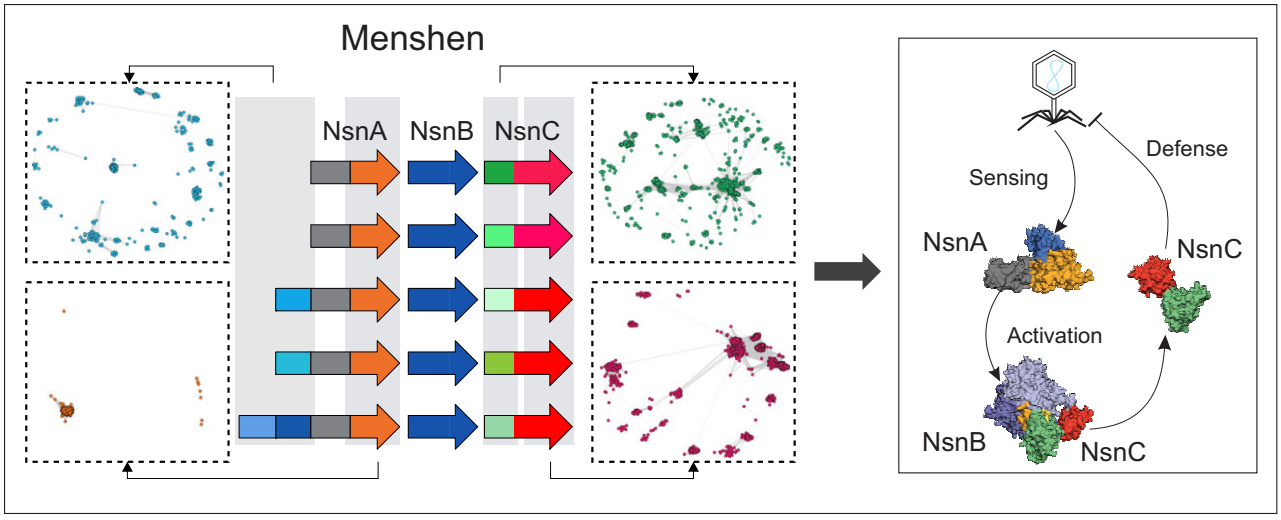
[1]Department of Biology, College of Arts and Sciences, Saint Louis University, Saint Louis, MO 63103, United States
[2]Department of Chemistry and Biochemistry, University of California San Diego, La Jolla, CA 92093, United States
[3]Department of Cellular and Molecular Medicine, University of California San Diego, La Jolla, CA 92093, United States
[4]Department of Molecular Biology, University of California San Diego, La Jolla, CA 92093, United States
[5]Program of Bioinformatics and Computational Biology, School of Science and Engineering, Saint Louis University, Saint Louis, MO 63103, United States

*To whom correspondence should be addressed. Email: dapeng.zhang@slu.edu

Abstract

Recent advances have significantly enriched our understanding of complex bacteria–phage interactions. To date, over one hundred bacterial antiphage systems have been identified, yet the mechanisms of many, including the recently discovered Menshen system, remain elusive. We employed comparative genomics and protein bioinformatics for a systematic investigation of the Menshen system, focusing on its organization, structure, function, and evolution. By delineating six primary domain determinants and predicting their functions, we propose that the three components (NsnA-B-C) of Menshen likely act as sensor, transducer, and effector modules, respectively. Notably, we unveil remarkable polymorphism in domain composition within both NsnA and NsnC. NsnA proteins universally share ParB-DUF262 and DNA-binding ParBDB domains, and often include additional DNA-binding modules at their N-termini. NsnC effectors exhibit diverse inactive PIN (inPIN)-like domains for target recognition in their N-termini, and multiple nuclease domains for toxicity in their C-termini. We demonstrate that this multifaceted polymorphism results from the independent integration of various sensor domains into NsnA, alongside constant shuffling and diversification of the inPIN and effector domains in NsnC. These findings not only elucidate the functional diversity and inter-subunit interactions of the Menshen system, but also underscore its exceptional capacity for adaptability and versatility in the ongoing arms race between bacteria and phages.

Graphical abstract



Introduction

The battle between bacteria and phages has been a driving force for the evolution of bacteria, resulting in a multitude of defense and counter-defense systems [1]. In the early stages of research into prokaryotic immunity, the known arsenal of bacterial antiphage systems was limited to restriction-modification (R-M) systems and CRISPR-Cas systems [2]. However, recent years have witnessed a transformative

Received: January 23, 2024. **Revised:** April 14, 2025. **Editorial Decision:** April 15, 2025. **Accepted:** April 17, 2025

expansion in our knowledge of antiphage systems, largely driven by the genomic "guilt by association" principle [3]: the recognition that multiple antiphage systems often cluster within specific genomic loci in bacteria [4, 5]. By mining genomes, researchers have uncovered over one hundred antiphage systems [6–10], and extensive experimental studies have revealed many systems' molecular mechanisms [1]. Among the identified antiphage systems, several are associated with the DUF262 domain [7], a member of the ParB NTPase superfamily that functions as a sensor module to recognize modified DNA in phage genomes [8–16]. The first example is the GmrSD system, a single-gene Type IV R-M system that contains an N-terminal DUF262 domain followed by a DUF1524 HNH nuclease domain [9], which together enable the recognition and degradation of phage genomes marked by glucosyl-5-hydroxymethyl cytosines [10]. A homolog of GmrSD, the SspE gene, was found in the SspABCD-SspE system, which is a phosphorothioate-based antiphage system [11]. The DUF262 domain in SspE recognizes and binds genomic regions containing phosphorothioate modifications [12]. In addition, there are other DUF262-containing antiphage systems, including PD-T4-2 [13], Dazbog, and Menshen [14], which are not associated with the DUF1524 domain and remain largely uncharacterized.

Menshen is a recently identified antiphage system associated with the DUF262 domain [14], taking its name from a pair of Chinese deities known for safeguarding homes. This nomenclature aligns with other antiphage systems like Thoeris [15, 16], Gabija [17, 18], and Shedu [19, 20], which draw inspiration from protective deities of various world mythologies [21, 14]. Menshen comprises three components known as NsnA, NsnB, and NsnC. Among them, NsnA is characterized by the DUF262 domain, while NsnB possesses a predicted ABC ATPase domain. In a recent study conducted by Millman *et al.*, a Menshen system from *Solibacillus silvestris* effectively defended against *Escherichia coli* phages T2, T4, and T6, and *Bacillus subtilis* phage Fado [14, 22]. However, significant knowledge gaps still exist concerning the molecular mechanisms of the Menshen system. Here, we employ bioinformatic approaches to systematically identify Menshen systems, dissect their structural and genomic organization, and explore potential mechanisms. Our findings reveal Menshen as a highly polymorphic system that leverages a wide range of protein domains likely involved in phage infection sensing, target identification, and effector function. We further provide functional and evolutionary insights into the major components within Menshen systems, demonstrating their varied strategies to counteract phage infections and resistance. Our study lays the groundwork for further investigations into the mechanisms of this system, and serves as an illustrative example of comprehending novel antiphage systems through a bioinformatics lens.

Materials and methods

Protein domain-centric homologous sequence search and analysis

Our computational analyses were based on the protein domains instead of the full length of proteins [23]. To collect homologous sequences for each domain family, PSI-BLAST [24] search was performed against NCBI nr database iteratively until convergence, with an e-value cut-off of 0.001. To build Multiple Sequence Alignment (MSA) for each domain family, highly similar sequences were removed from the collection via the BLASTCLUST program (https://ftp.ncbi.nih.gov/blast/documents/blastclust.html) and the remaining sequences were then aligned using KALIGN [25], MUSCLE [26], or PROMALS3D [27] programs. All the MSAs of novel domains identified in this study are available in Supplementary File 1. To identify conserved residues for each domain family, a custom Perl script was used to analyze the conservation pattern in the MSA based on the classification of amino acids [28]. In addition, MSAs of all newly identified domains were used to generate HMM profiles with the HMMER package [29] for further studies. WebLogo was used to generate sequence logos for families if necessary [30].

Genomic retrieval of the Menshen system using gene neighborhood analysis

Given that Menshen is an antiphage system with multiple components, we employed gene neighborhood analysis [31] to comprehensively collect instances of the Menshen system. Homologous DUF262 proteins, collected through PSI-BLAST, were initially annotated based on their domain architectures. The accession numbers of these proteins were then used as the input to retrieve their five upstream and five downstream genes using a custom script. Subsequently, all neighbor proteins were clustered by the BLASTCLUST program based on their sequence similarities. The conserved neighbors were determined by the criteria: (i) close association with DUF262 genes on their genomic loci, and (ii) presence of these associations across at least two different bacterial phyla. Furthermore, these conserved neighbor proteins were annotated based on their domain architectures by using the HMMSCAN program [29] searching against the Pfam database [32] and our custom HMM profile database. The operons were then classified according to similarities in domain architectures of DUF262 proteins and their conserved neighborhoods. We ultimately identify multiple systems centered around the DUF262 domain, encompassing, but not limited to, the previously identified antiphage systems, such as GmrSD, PD-T4-2, Dazbog, SspABCD-SspE, and Menshen. The Menshen operons were subsequently isolated for further analyses, and their taxonomic information was extracted from NCBI GenPept files.

Sequence polymorphism analysis for the Menshen system

Many antiphage systems were observed to have polymorphisms [20, 33], which is rational given the evolutionary force driven by the positive selection between bacteria and phages. Based on our pilot analysis of the typical NsnABC components, we dissected these proteins into a few primary domains including NsnA-N (only for Type II Menshen), NsnA-C, NsnC-N, and NsnC-C regions, in addition to the DUF262 domain in NsnA and the globular domain in NsnB. To systematically investigate the polymorphism of Menshen, we performed network clustering analysis using the CLANS program [34] on sequences from the NsnA-N, NsnA-C, NsnC-N, and NsnC-C regions, respectively (Supplementary Figs S1 and S2). In this analysis, significant high-scoring segment pairs (HSPs) among sequences were first identified through all-against-all BLASTP searches (*P*-value cutoff: .0001). These HSPs were then used to organize the sequences into clusters using the Fruchterman and Reingold force-directed layout algorithm [35, 36], where

sequences are represented as nodes, and the edges connecting them represent attractive forces proportional to the negative logarithm of the alignment significance (*P*-values). The degree of dispersion in the resulting network reflects the level of polymorphism in each region, while grouped nodes indicate discrete domain families. For the NsnA-N, NsnA-C, and NsnC-C regions, isolating their families is relatively straightforward as their boundaries are clearly defined in the CLANS network. However, in the NsnC-N sequence space, where many distinct families share the same structural topology (inPIN domains), we employed the embedded "find clusters" function of the CLANS program to extract NsnC-N clusters ($n > 2$) based on the average linkage method. Sequences that did not form such clusters were designated as the "ungrouped" cluster.

Structure prediction and analysis of the Menshen system

Protein structures obtained through experimental studies were retrieved from the PDB database [37] while all other structural models presented in this study were predicted using AlphaFold 2 [38, 39] or AlphaFold 3 [40]. The corresponding .pdb or .cif files of all modeled structures and complexes are provided in Supplementary Files 2, 3, and 4. Both the pLDDT scores, which provide residue-level confidence for each structural prediction, and AF3 ipTM scores, which assess the confidence of predicted protein-protein interactions, are available in Supplementary Figs S4, S5, and S7. The determination of domain boundaries for each family was guided by both the structure models and the PAE matrix provided by AlphaFold 2/3. To gain further insights into the functions of novel domain families, the DALI program [41] was employed to search their distant structural homologs. Finally, visualization of structures was conducted via PyMOL [42] and the NGL viewer [43].

Phylogenetic analysis and Sankey diagram construction

To delineate the evolutionary trajectory of the Menshen system, a comprehensive phylogenetic analysis was conducted. We first extracted the DUF262 domain sequences from all 1 997 Menshen instances. These sequences were then used to construct a phylogenetic tree, employing the IQ-TREE software [44] with the model selection parameter (-m TEST). Branch supports were assessed through 1000 iterations of ultrafast bootstrap approximation and 1000 iterations of the SH-aLRT test. The MSA file of DUF262 domains and the corresponding tree file can be found in Supplementary File 5. The resultant phylogenetic tree was visualized using iTOL [45], incorporating annotations of various Menshen types, NsnA-N sensor domains, NsnC-N inPIN domains, NsnC-C effector domains, and taxonomy information. To understand the intricate relationships and functional linkages among the various domains (NsnA-N sensors, NsnC-N inPINs, and NsnC-C effector domains) in the Menshen system, Sankey diagrams were created using the Flourish platform (https://flourish.studio/).

Results and discussion

A first glimpse of the Menshen system via protein structure predictions

The Menshen system from *S. silvestris* StLB046 represents the first example with experimentally validated antiphage activity (Fig. 1A) [14], which we referenced to obtain the sequences of NsnA, NsnB, and NsnC in the NCBI database. Next, we used HMMSCAN profile detection [29], AlphaFold prediction [38], and structural analysis to determine each protein's domain architecture.

NsnA

As depicted in Fig. 1B, we found that NsnA is predicted to contain two distinct globular domains: the N-terminal DUF262 domain and an uncharacterized C-terminal domain. The DUF262 domain is known to have three conserved motifs: a QRxxxW motif (I), a DGxQR motif (II), and a FxxxN motif (III), the second of which is the predicted nucleotide binding center [9]. The AlphaFold 2 model shows that these three motifs are clustered together, suggesting that they work in concert with one another. We also constructed a model of nucleotide-bounded NsnA by overlaying the AlphaFold model of *S. silvestris* NsnA with that of an ATP-binding sulfiredoxin (PDB: 3CYI_A), another member of the ParB superfamily [46]. The modeled structure illustrates the interaction of the DGxQR motif with the β and γ phosphate groups of the bound nucleotide (Fig. 1B). Different variants of the ParB domain are known to function as NTPases in several molecular systems, such as the canonical ParB protein which functions in bacterial chromosome partitioning [8], the fertility inhibition factor Osa [47], and the DndB protein from the DndABCDE antiphage system [48]. Importantly, recent studies have demonstrated that the DUF262 domain acts as a sensor in two antiphage systems, SspABCD-SspE [11] and GmrSD [49]. In these systems, the DUF262 domain detects specific phage phosphorothioate-modified DNA and glucosyl-5-hydroxymethyl cytosine, respectively, and further activates their associated nucleases (DUF1524) through its NTPase activity [49, 12] to initiate the antiphage response. The fact that DUF262 acts as a sensor in other antiphage systems suggests a similar role for NsnA in the Menshen system.

For the C-terminal region of NsnA, our structure predictions revealed similarities to the downstream domains found in other ParB/DndB/DUF262-containing proteins such as DndB, GmrSD, and SspE (Fig. 1C and Supplementary Fig. S3A). These domains share eight conserved helices, with the exception of the DndB C-terminal domain, whose fourth helix is degenerated. These domains were previously considered part of the DUF1524 domain or linkers [49, 12]. However, our findings suggest that they instead form distinct, independent domains with unique conservation patterns (Supplementary Fig. S3B). Notably, multiple conserved charged or polar residues within each of these domain families are located on the inner surface of the groove formed by the dimer, such as those on GmrSD-M (D249) and SspE-M (E334 and K337) domains (Supplementary Fig. S3C). These residues lie adjacent to the QRxxxW motif of DUF262, suggesting they may coordinate their interactions with substrates. Indeed, Gao *et al.* demonstrated that SspE-M domain (a homolog of the NsnA-C domain) is essential for DNA binding [12], suggesting that DNA-binding activity is a shared function among these domains. Further, the combination of ParB-like DUF262 and NsnA-C DNA-binding domains mirrors the arrangement seen in the canonical chromosome partitioning protein ParB, in which the ParB domain precedes an HTH-type DNA-binding domain [50]. Accordingly, we have assigned the DUF262 domain as "ParB-DUF262" and the C-terminal domain as the "ParBDB" (**ParB**-associated **D**NA-**B**inding) domain.

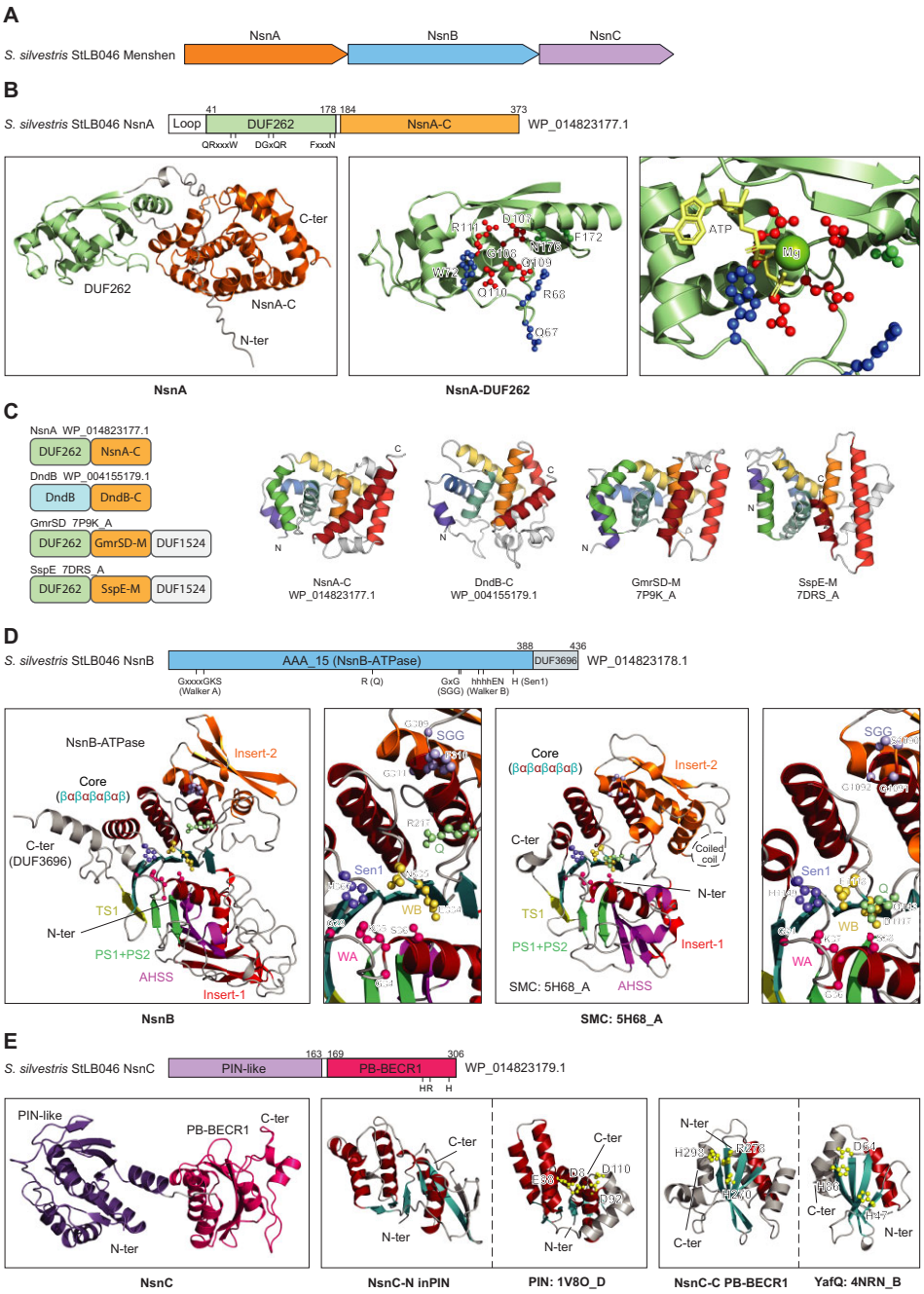


Figure 1. Structure predictions and domain organization of the Menshen system in *S. silvestris* StLB046 strain. (**A**) Operonic structure of the Menshen system in *S. silvestris* StLB046. (**B**) *Upper*: The domain architecture of *S. silvestris* StLB046 NsnA with the N-terminal loop in white, DUF262 domain in green, and C-terminal NsnA-C domain in orange. Domain boundaries are indicated above, conserved motifs of the DUF262 domain below, and the NCBI protein accession ID on the right. *Bottom left*: Cartoon representation of the full-length NsnA protein. *Bottom middle*: Cartoon representation of the DUF262 domain with conserved residues depicted as ball-and-stick models, motif I in blue, motif II in red, and motif III in green. *Bottom right*: Model of the DUF262 domain interacting with ATP, where ATP is shown in pale yellow sticks, and the magnesium atom is represented as a green sphere. (**C**) *Left*: Shared domain architectures of the *S. silvestris* StLB046 NsnA protein and ParB/DUF262-containing proteins from other antiphage systems, including DndB, GmrSD, and SspE. *Right*: Structural similarity among NsnA-C, DndB-C, GmrSD-M, and SspE-M, with corresponding helices highlighted in distinct colors. (**D**) *Upper*: The *S. silvestris* StLB046 NsnB containing an extended version of ABC-ATPase fold (NsnB-ATPase) and a C-terminal helix (DUF3696). *Bottom left two*: (i) Cartoon representation of the NsnB-ATPase domain, highlighting the core regions in firebrick (helices) and teal (strands), major insertions (Insert-1 in red, Insert-2 in orange), two strands N-terminal to the core (PS1 + PS2) in lime green, additional helix-strand-strand at the N-termini of Insert-1 (AHSS) in magenta, and an additional C-terminal strand (TS1) in yellow. Conserved residues are shown as ball and stick, where the Walker A motif (WA) in hot pink, Walker B motif (WB) in yellow-orange, Q motif in pale green, SGG motif in blue white, and Sen1 motif in slate. (ii) Zoomed-in view of conserved residues in NsnB-ATPase with labels. *Bottom right two*: (i) Cartoon representation of the chromosome partition protein SMC (PDB: 5H68_A, Z-score = 22.1). (ii) Zoomed-in view of conserved residues in SMC with labels. (**E**) *Upper*: Domain architecture of *S. silvestris* StLB046 NsnC with PIN-like domain in purple and PB-BECR1 in hot pink. *Bottom left*: Cartoon representation of the full length NsnC protein. *Bottom middle*: Cartoon representation of the PIN-like domain (left) and its distant homologous PIN domain (right, PDB: 1V8O_D, Z-score = 5.8). Structural cores are shown in firebrick (helices) and teal (strands), with conserved residues shown as ball and stick in yellow. *Bottom right*: Cartoon representation of the PB-BECR1 domain (left) and its distant homologous YafQ domain (right, PDB: 4NRN_B, Z-score = 8.5).

NsnB

In the case of *S. silvestris* NsnB, initial Pfam annotation identified two domains, an AAA_15 ATPase domain (or OLD-ABC, given its frequent association with OLD-family nuclease domains) [51] of the ABC ATPase superfamily, and a C-terminal DUF3696 domain (Fig. 1D). By comparing the predicted structure of *S. silvestris* NsnB with its distant homolog, the chromosome partition protein SMC (PDB: 5H68-A with Z-score = 22.1) [52], we found that the NsnB-ATPase domain exhibits typical structural characteristics of the ABC ATPase superfamily [51], including the α/β core (βαβαβαβαβ), the C-terminal strand (TS1), the N-terminal β-hairpin (PS1 + PS2), the main insertion 1 (Insert-1), and the main insertion 2 (Insert-2) (Fig. 1D). To be noted, like the SMC, the N-terminus of NsnB-ATPase Insert-1 has evolved an additional αββ insertion, which is the hallmark of Type ID Insert-1 of the ABC-ATPase superfamily [51]. However, the remaining parts of Insert-1 of NsnB-ATPase exhibit more strands than the corresponding region in SMC and the insert-2 of NsnB-ATPase lacks the long coiled-coil insertion observed in SMC. Further, both NsnB-ATPase and SMC exhibit typical sequence motifs characteristic of ABC ATPase superfamily, which encompass the Walker A motif (GxxxxGKS), the Walker B motif (hhhhDE), the Q motif, the SGG motif, and the Sen1 motif (H) [51]. Nonetheless, in the NsnB-ATPase, the DE dyad in the Walker B motif has been replaced by EN, the Q motif substituted with an arginine, and the SGG motif replaced by a GxG motif.

NsnC

Finally, the predicted structures of *S. silvestris* NsnC identified two hitherto-uncharacterized domains (Fig. 1E). The N-terminal domain exhibits a Rossmann-like fold and DALI searches retrieved many members of the PIN domain superfamily, such as PDB: 1V8O_D with Z-score 5.8 [53]. Typically, the classical PIN domain functions as an RNase, with its catalytic site being orchestrated by four highly conserved residues D-E-D-D [54]. However, these crucial residues are absent in the NsnC-N domain. During evolution, enzymatic domains that lose their catalytic activity often acquire alternative functional roles, such as binding the same substrates of their active enzymatic counterparts [55, 56]. Given that PIN domains are dedicated RNases and NsnC-N represents an inactive variant, we predict that NsnC-N may have evolved to function as an RNA-binding domain. Consequently, we have named it as an "inactive PIN (inPIN)-like domain". On the other hand, our examination of the C-terminal domain of NsnC conclusively places it within the BECR superfamily [57]. The most significant hit identified through DALI for NsnC-C is the YafQ (PDB: 4NRN-B with Z-score 8.5) RNase domain from a toxin–antitoxin (TA) system [58]. YafQ features a H-D-H triad as its active sites while NsnC-C possesses an H-R-H triad in a comparable position. Therefore, we have named the NsnC-C domain as the PB-BECR1 (**P**ar**B** -associated **BECR 1**) domain.

Thus, our sequence analyses and structure predictions have delineated the primary domain components of the archetypal Menshen system from *S. silvestris* StLB046. NsnA possesses ParB-DUF262 and ParBDB domains; NsnB possesses ABC-ATPase and DUF3696 domains; and NsnC possesses an inPIN domain and an active BECR nuclease domain. Notably, both ParB-DUF262 and ParBDB domains in NsnA have been demonstrated to possess phage DNA sensing and DNA-binding abilities in other antiphage systems [11, 12, 49], suggesting that NsnA functions as an infection sensor in Menshen. Meanwhile, BECR nuclease domains are established as effectors in various polymorphic toxins and TA systems [23, 57] and PB-BECR1 is the only potential active toxin domain identified in this Menshen system. Consequently, we propose that NsnC, which contains the PB-BECR1 domain, operates as an RNase effector in the *S. silvestris* Menshen system. Based on the isolated analysis of each protein, the specific role of NsnB remains unclear.

Comprehensive analysis of Menshen homologs reveals it is a highly polymorphic system

We next sought to explore the repertoire of the Menshen system by mining available genomes (Fig. 2A). We first utilized the ParB-DUF262 domain of *S. silvestris* NsnA as the seed for the iterative PSI-BLAST searches. The retrieved proteins were annotated based on their predicted domain architectures, and their gene neighborhoods were further extracted from the NCBI genome database. Next, all neighbor proteins were clustered based on their sequence similarities. Clusters of neighboring proteins that were associated with ParB-DUF262 proteins in the genomic vicinity in at least two different bacterial phyla were annotated and considered as conserved neighbors. We then classified these ParB-DUF262 homologs and their genomic loci based on their domain architectures and conserved neighbors. This process led us to identify multiple systems that incorporate the ParB-DUF262 domain, including GmrSD, PD-T4-2, Dazbog, and SspABCD-SspE, in addition to Menshen. Importantly, for the Menshen system, we found that when a NsnA-like gene (containing ParB-DUF262 and ParBDB domains) is followed by a NsnB-like (NsnB-ATPase) gene, there is almost always a third gene encoding a hypothetical protein located immediately downstream. By domain analysis, some of the third gene products contain known effector domains at their C-termini, such as Gp49 [59], Ntox18 [23], and PB-BECR1, a situation similar to the NsnC in the *S. silvestris*. For the remaining hypothetical proteins, despite no known domains recovered in either Pfam or CDD databases, our further analysis in a later section confirmed the presence of many novel toxin/effector domains at their C-termini. Therefore, based on the conserved association of these three components, we have identified a comprehensive collection ($n = 1997$) of bacterial Menshen systems (Supplementary Table S1).

As many protein domains in our collection of Menshen systems are uncharacterized, we next sought to systematically dissect their domain compositions and explore potential structural and functional diversity. To facilitate this, we used sequence similarity-based network clustering via the CLANS analysis to identify the major sequence families or domains (Fig. 2B). We found that the three major components of the Menshen system can be divided into six regions of globular domains. NsnA proteins possess an optional N-terminal domain not found in *S. silvestris* NsnA (here termed as NsnA-N for N-terminal), plus ParB-DUF262 (NsnA-M for Middle), and ParBDB domains (NsnA-C for C-terminal). NsnB proteins possess the ABC ATPase domain followed by an α-helical tail (DUF3696 domain), and NsnC possesses separate N-terminal and C-terminal domains (NsnC-N and NsnC-C). Notably, we found that NsnA-N, NsnC-N, and NsnC-C domains all exhibited exceptionally high sequence diversity,

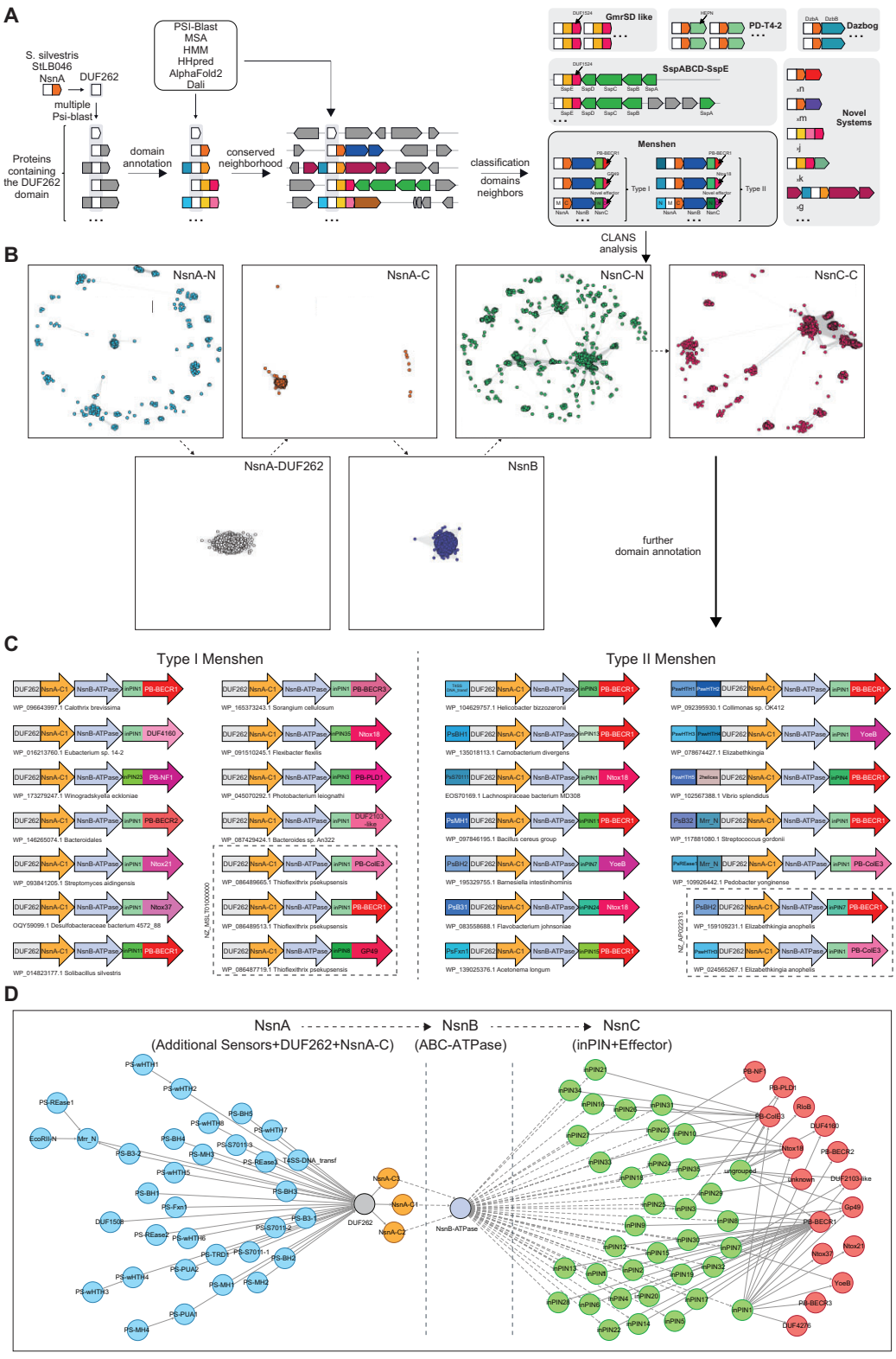


Figure 2. Genomic mining and domain analysis of the Menshen system. (**A**) A domain-centric protein analysis and genome mining pipeline for systematical identification of the Menshen system. For details, please refer to the 'Materials and methods' section or the 'Results and discussion' section. (**B**) CLANS network analysis of the NsnA-N, NsnA-DUF262, NsnA-C, NsnB, NsnC-N, and NsnC-C sequences of the collected Menshen components. Each node corresponds to a sequence. Straight lines indicate significant HSPs detected by all-against-all BLASTP searches with scoring matrix BLOSUM62 and an e-value cutoff of 0.0001. The sequence clusters are organized based on the Fruchterman and Reingold force-directed layout algorithm. For details, please refer to Supplementary Data and the 'Materials and methods' section. (**C**) Representative examples of the Menshen system (Type I, left; Type II, right). Each gene is presented as a block arrow, in which the major domains of the encoded protein are separately shown as rectangular segments. The loci are labeled by the NCBI accession numbers of the NsnA-like proteins followed by their species names. Examples within the same dashed rectangle are from the same bacterial genome. (**D**) A domain architecture and operonic association network of the Menshen system. Solid lines represent direct domain connections within the same protein, while dashed lines indicate operonic linkages between two proteins.

while the NsnA-C region displayed some degree of sequence diversity (Supplementary Figs S1 and S2). The observed diversity within specific protein regions is a characteristic feature of domain polymorphism, a phenomenon evident in polymorphic toxin systems that employ domain shuffling, particularly in the toxin/effector component, to enhance diversity and facilitate organismal inter-specific and intra-specific competitions [23, 57, 60, 61]. In the following sections, we systematically dissected and classified the domain families in Menshen using a series of domain-centric protein analysis strategies, including sequence/profile analysis, structural modeling and comparison, and evolutionary analysis (see the 'Materials and methods' section).

Sequence and structural diversity of NsnA-like proteins

Upon examining NsnA-like proteins, we found that 77% of them, including *S. silvestris* NsnA, either lack any domain or have only the loop region at their N-termini preceding the core ParB-DUF262 and ParBDB domains. However, the remaining 23% of NsnA proteins possess additional N-terminal domains. Consequently, we have designated Menshen systems lacking additional N-terminal domains in NsnA as Type I Menshen, and those systems whose NsnA possesses additional N-terminal domains as Type II Menshen (Fig. 2C). Further exploration through network clustering and structure predictions of NsnA-N sequences within Type II Menshen systems has led to the identification of at least 33 distinct domain families with unique structures (Figs 2D and 3). Substantial evidence suggests that domains in this region serve as extra sensor domains to recognize components of invading phages: (i) Domains with a canonical DNA-binding winged Helix–Turn–Helix (wHTH) fold [62] were the most commonly observed domains in NsnA N-termini, including eight novel families (designated ParB superfamily associated Sensor-wHTH, PS-wHTH families 1 through 8) and the known Mrr_N family (Fig. 3A). Notably, Type IIS restriction enzymes are one example of antiphage systems that utilize wHTH domains to recognize specific DNA sequences, including BpuJI (PDB: 2VLA) [63] and FokI (PDB: 1FOK) [64]. DALI searches further revealed that both BpuJI and FokI contain two wHTH domain copies in their N-terminal regions, which exhibit significant structural similarities with PswHTH1 + PswHTH2 and PswHTH3 + PswHTH4 in the Menshen system, respectively; (ii) Three families belong to the restriction endonuclease (REase) superfamily (designated PS-REase1 to PS-REase3; Fig. 3B), members of which are often nucleases found in R-M systems. However, REase domains in NsnA lack the PD-(D/E)XK active site motif, suggesting that they are enzymatically inactive and solely serve to bind DNA [65]; (iii) Another three families (PS–B3–1, PS–B3–2, and EcoRII-N) have a predicted B3 fold (Fig. 3C), the core of which has seven β-strands (arranged in the order β1–β6–β7–β3-β4–β5–β2) to form a pseudo-β-barrel. B3 family domains are DNA-binding domains often found in plant transcription factors [66]. Nevertheless, some restriction enzymes also use B3 domains as their recognition domains, such as EcoRII [67] and BfiI [68]; (iv) Compared to the above domains that likely recognizing unmodified DNA sequences, two types of NsnA-N domains may bind modified DNA bases. The first one is the PS-Fxn1 (Fig. 3D), which adopts a frataxin-like fold with elements arranged as αββββββα [69]. The second type is PUA domains (PS–PUA1 and PS–PUA2; Fig. 3E), which possess a pseudo-β-barrel that is ordered as β1–β3–β4–β5–β2 [70]. Both members of frataxin-like and PUA superfamilies are presented in Type IV R-M systems, where they sense modified phage DNA [71, 72]; (v) Next, PsTRD1 (Fig. 3F) is another domain with strong evidence to be a sensor domain, since it shows significant structural similarity with the specificity domain of Type I R-M systems [73, 74].

In addition to the above domain families, we also identified many other domains in NsnA-N that are not commonly recognized in antiphage systems. Given that they occupy the same position as the other putative sensor domains in the Menshen system, we propose that these uncharacterized domains also serve as infection sensors. Several NsnA homologs contain two copies of DUF1508 domains (Fig. 3G). We found the two copies are predicted to form an intrinsic dimer that is similar to the topology of the sPC4/Whirly superfamily, many members of which are single-stranded nucleic acid binding domains [75]. In fact, our recent study has found various sPC4/Whirly domains are potential sensor domains located at N-termini of Shedu antiphage nucleases [20]. Further, three NsnA proteins have the T4SS-DNA_transf domain in their N-termini (Fig. 3H), which is a member of the P-loop ATPase family with an α-helical insertion. This domain shares significant structural similarity with the plasmid conjugative coupling protein TrwB (PDB: 1EGR_G), which forms a hexamer and uses a similar α-helical insertion to ensure specificity in recognizing DNA substrates [76]. Finally, the potential sensor domains found in NsnA also include five domains (PS-BH1 to PS-BH5; Fig. 3I) that sharing a topology of a curved β sheet followed by a helix, three domains (PS-S7011-1 to PS-S7011-3; Fig. 3J) that show structural similarities with region 1.1 of sigma-70 factors [77], and four uncharacterized domains made of only α-helices (Fig. 3K).

Sequence and structural diversity of NsnC-like proteins

As the third component of the Menshen system, NsnC-like proteins display remarkable diversity in both their N-terminal and C-terminal regions (Figs 2B and D, and 4). Using CLANS network clustering, we classified NsnC N-terminal regions into 35 major families, and identified an additional 81 sequences that remain ungrouped. Subsequent structural modeling and comparisons among the representative sequences of these 35 families, as well as ungrouped sequences, revealed that the majority (97.9%) adopt a PIN domain-like topology similar to that observed in *S. silvestris* NsnC, despite sharing very limited sequence identity (Fig. 4A and Supplementary Fig. S6). The remaining 2.1% lacked a PIN-like fold and were characterized by disordered structures with notably low pLDDT scores, indicating AlphaFold 2 modeling difficulties due to the absence of homologous sequences. Further sequence conservation analysis revealed that all major PIN-like domains lack the canonical PIN catalytic residues, suggesting that they may be functionally inactive (Supplementary File 1). Accordingly, we named these major domain families inPIN1 to inPIN35, of which the one in *S. silvestris* NsnC is assigned to the inPIN11 family.

For the C-terminal domains of NsnC, we also detected 15 different families (Fig. 4B–E). Detailed sequence analysis and structure predictions allowed us to successfully relate these domains to known protein domain families, with at least 12

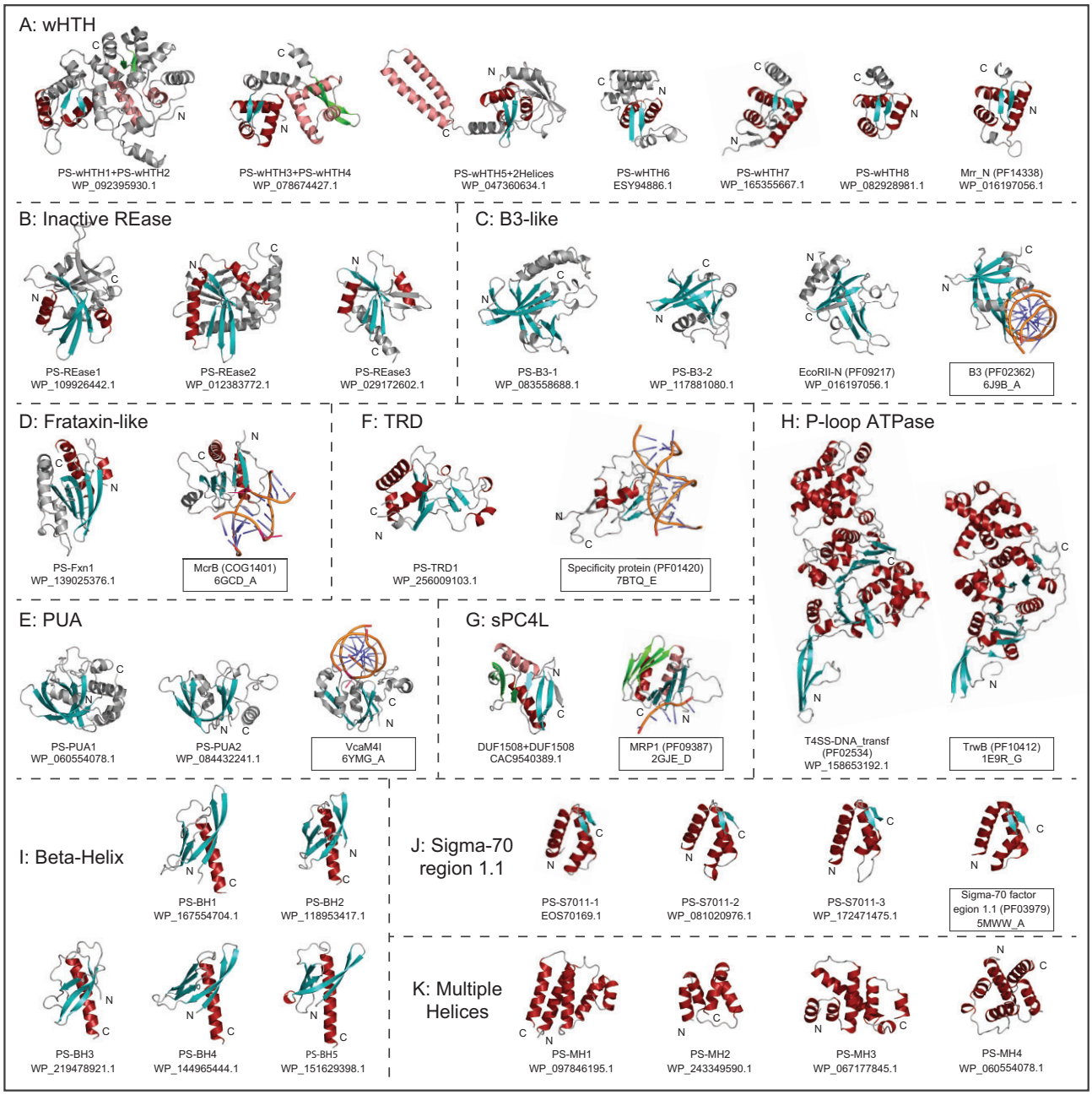


Figure 3. Cartoon representations of diverse sensor domains identified in the NsnA-N region of the Menshen system. (**A**) wHTH domains. (**B**) Inactive Restriction Endonuclease (inREase) domains. (**C**) B3-like domains. (**D**) the Frataxin-like domain. (**E**) PUA-like domains. (**F**) the TRD-like domain. (**G**) sPC4L/Whirly fold domain. (**H**) The P-loop ATPase sensor domain. (**I**) Beta-Helix sensor domains. (**J**) Sigma-70 factor, region 1.1-like sensor domains. (**K**) Multiple Helical (MH) sensor domains. For each domain, a representative structure is displayed alongside its corresponding NCBI protein accession number. In cases where distant structural homologs that interact with nucleic acids are available and considered pertinent, their structures are shown with the name highlighted in box. The structural cores of most domains are colored in firebrick (helices) and teal (strands). However, for the PS-wHTH2 and PS-wHTH4 domains, the second copy of the DUF1508 domain, and the second half of the MRP1 domain, their structural cores are represented in salmon (helices) and tv_green (strands). Nucleic acids are depicted in orange (backbones) and white blue (bases), with modified bases highlighted in hot pink.

of these families being likely nucleases. First, nine NsnC-C domains belong to the BECR (Barnase/EndoU/Colicin/RelE) superfamily, which share a structural core of one α-helix followed by a four-stranded antiparallel β-sheet (Fig. 4B). The RNase activity of several BECR domain families related to those found in Menshen systems has been experimentally validated, including Gp49 [59], YoeB [78], and Ntox21 [79]. Second, we identified DUF4276 and RloB domains as two novel members of the Toprim (Topoisomerase/primase) superfamily (Fig. 4C), which exhibit a three-layered α/β topology with a central sheet arranged in the order of β2–β1–β3–β4 [80]. Typical Toprim domains, such as RNase M5 (PDB: 6Z2B) [81] and those found in OLD nucleases (i.e. PDB:6NK8) [82], are characterized by a conserved glutamic acid on the α1–β1 loop and a conserved DxD motif after the β3 strand; all these catalytic residues are also observed in the Menshen

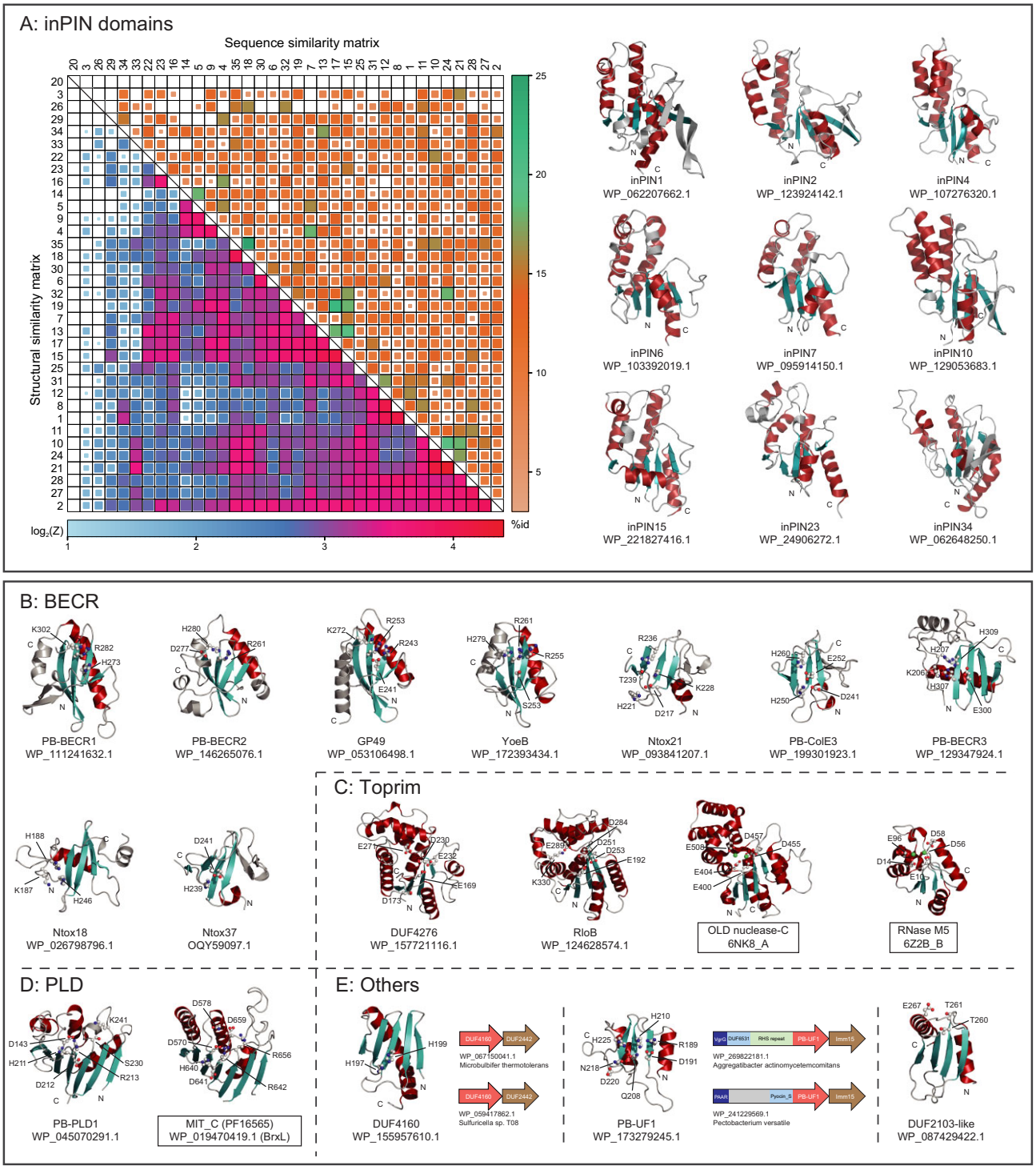


Figure 4. Cartoon representations of various NsnC-N inPIN domains (**A**) and the diverse NsnC-C effector domains (**B**–**E**) of the Menshen system. (**A**) *Left*: Sequence and structural similarity matrices of 35 inPIN domains. The lower-left matrix illustrates the structural similarity between domains, as measured by log-transformed DALI Z-scores, while the upper-right matrix displays sequence identity between these domains. The inPIN20 domain's structural predictions yielded disordered models with low pLDDT scores primarily due to the absence of homologous sequences. *Right*: Predicted structures of nine representative inPIN domains (inPIN1, inPIN2, inPIN3 and inPIN7). (**B**) BECR RNase domains. (**C**) Toprim-fold nuclease domains. (**D**) the PB-PLD1 domain. (**E**) Other novel effector domains, DUF4160, PB-UF1, and DUF2103-like domains. For each domain, a representative structure is displayed alongside its corresponding NCBI protein accession number. In cases where distant structural homologs are available and deemed necessary, their structures are shown with the name highlighted in box. The structural cores of all domains are depicted in firebrick (helices) and teal (strands). Predicted catalytic residues, identified through conservation analysis, are highlighted as ball-and-stick models, with carbon atoms in gray, nitrogen atoms in blue, and oxygen atoms in red. Magnesium atoms are represented as lime green spheres.

DUF4276 and RloB domains (Fig. 4C). Notably, both homologs of DUF4276 (Mad4) and RloB have been shown to be components of other antiphage systems, such as methylation-associated defense system [83] and Type I R-M systems [84], consistent with their predicted effector roles in Menshen. Further, the PB-PLD1 domain (ParB-associated PLD1-like) represents a third class of potential nucleases in NsnC-C domains (Fig. 4D). The predicted structure of PB-PLD1 exhibits strong structural similarity and profile similarity to several nucleases within the phospholipase D superfamily, including MIT_C (PDB: 2YMB; DALI Z-score: 11.4; HHsearch probability: 89%), NucT nuclease (PDB: 6EHI; DALI Z-score: 8.9) [85], and Tyrosyl-DNA phosphodiesterase (PDB: 6DIH; DALI Z-score: 8.7) [86]. Notably, the catalytic site of PB-PLD1, involving a conserved aspartic acid/glutamic acid after the β2 strand and a conserved HDR motif before the β5 strand, is highly similar to that of the MIT_C domain in the MITD1 protein [87], which we further found in the C-termini of many BrxL proteins (e.g. WP_019470419.1) of the BREX antiphage system [88]. This raises an interesting hypothesis that both MITD1 and BrxL are enzymatic proteins, which requires further study.

The remaining three Nsn-C domains—DUF4160 (PF13711), PB-UF1 (Unknown Fold 1), and DUF2103-like—display distinct topologies, and DALI searches using our predicted structures retrieved no confident structural homologs. However, an analysis of their domain architectures and gene neighborhoods provides evidence that they also function as effector components in other toxin systems (Fig. 4E). The majority of DUF4160 domains are found in single-domain proteins, which are typically followed by proteins containing DUF2442 domains within the same genomic loci. A homologous DUF4160/DUF2442 pair has been identified as a novel TA system in *Dickeya dadantii* 3937 strain [89]. The PB-UF1 domain is located at the C-termini of polymorphic toxin proteins secreted via the Type VI secretion system, often followed by Imm15 immunity proteins (Fig. 4E). The DUF2103-like domain is associated with the PDEXK domain (Cas_APE2256) in the RAMP-2 subtype of CRISPR/Cas loci (i.e. WP_193942373.1 from *Sphaerospermopsis aphanizomenoides*). Its distant homolog, DUF2103, has been implicated in Type III CRISPR-Cas antiviral systems [90].

Thus, the highly conserved architecture of all NsnC proteins, comprising N-terminal inPINs and diverse C-terminal effector domains, strongly supports our hypothesis that NsnC functions as the effector in the Menshen system. We propose that once activated, Menshen systems degrade nucleic acids, most likely RNAs, of the invading phage and/or the bacterial host. Consistent with this hypothesis, the N-terminal RNA-binding inPIN domains in NsnC could serve as target recognition modules for phage RNAs. The rampant sequence diversity observed in inPIN domains indicates they have undergone extensive positive selection, reflecting the ongoing arms race between bacteria and phages.

Protein-protein interactions in the Menshen system

To gain insight into the oligomeric state and interactions between Menshen components, we performed AlphaFold 3 structure predictions for homo- and hetero-oligomeric complexes of NsnA, NsnB, and NsnC from *S. silvestris* StLB046, and used the interface predicted template modeling (ipTM) score to judge the likelihood of complex formation. ipTM scores below 0.6 indicate a failed complex prediction, and scores above 0.8 indicate a confidently-predicted complex [91]. NsnA was not predicted to form a complex either with itself, NsnB, or NsnC (Supplementary Table S2). Predictions of NsnB homo-oligomers showed higher confidence, with the highest ipTM score (0.87) obtained with a prediction of an NsnB homotetramer plus ATP and Mg^{2+} (Supplementary Table S2). NsnB was also predicted to interact with NsnC, with an $NsnB_4(ATP{\cdot}Mg^{2+})$–$NsnC_2$ complex showing the highest ipTM score of 0.81. In this model, the NsnB homotetramer is assembled through two symmetric interfaces, one of which (interface 1) is equivalent to the canonical ABC ATPase dimer interface, and the second (interface 2) positioned orthogonal to interface 1 (Fig. 5A). A structural similarity search using the Foldseek server in Multimer mode [92] revealed that the predicted NsnB homotetramer is structurally equivalent to the observed homotetramer assembly of GajA, the ABC ATPase subunit of the Gabija antiphage system (Fig. 5B). In the GajA tetramer, interface 1 involves both the N-terminal ABC ATPase domain and the C-terminal toprim-family nuclease domain, and ATP binding to the GajA ATPase domains is thought to allosterically control nuclease activity of the toprim domain [93, 94]. In the predicted $NsnB_4$–$NsnC_2$ structure, ATPase dimer interface 1 positions two NsnB C-terminal DUF3696 domains close to one another, where they sandwich the NsnC inPIN11 domain through an asymmetric interaction (Fig. 5C). This arrangement—an ABC ATPase homodimer bound asymmetrically to a single effector protein—has been recently observed in two other antiphage systems: Septu and PARIS. In Septu, the PtuA ATPase forms a homohexameric assembly with three PtuA dimers, two of which bind a PtuB nuclease protomer between two PtuA C-terminal domains (Fig. 5D) [95]. In PARIS, the AriA ATPase forms a homohexameric assembly with three AriA dimers, each of which bind a single AriB nuclease subunit via a surface on the dimer interface of the AriA ATPase domain (Fig. 5E) [96, 97]. In both Septu and PARIS, ATP binding and/or hydrolysis by the ABC ATPase subunit (PtuA/AriA) is thought to regulate the activity of the bound nuclease subunit (PtuB/AriB), either by conformational changes (Septu) or by nuclease subunit release (PARIS). Thus, the predicted $NsnB_4$–$NsnC_2$ structure shares key architectural similarities with three other antiphage systems—Gabija, Septu, and PARIS—and suggests a mechanism by which ATP binding and/or hydrolysis by NsnB could regulate NsnC activity, either allosterically via coupled conformational changes, or by triggering binding or release of NsnC.

To determine whether the predicted NsnB–NsnC interaction is conserved across diverse Menshen systems, we performed AlphaFold 3 structure predictions on 24 additional $NsnB_4(ATP{\cdot}Mg^{2+})$–$NsnC_2$ complexes representing 10 different NsnC-N inPIN domain families and all 15 NsnC-C effector domain families. AlphaFold 3 produced high-confidence models (ipTM > 0.8) in eleven cases, and medium-confidence models (ipTM between 0.6 and 0.8) in 10 cases (Supplementary Table S3). All 24 models showed a consistent NsnB tetrameric assembly, and 23 of 24 models showed each pair of NsnB dimers binding an NsnC N-terminal domain asymmetrically through their C-terminal DUF3696 domains (Supplementary Fig. S7). Overall, these predictions suggest that despite the high sequence and structural diversity in NsnC-N domains, these domains likely interact with NsnB in a structurally equivalent manner via the NsnB DUF3696 domain.

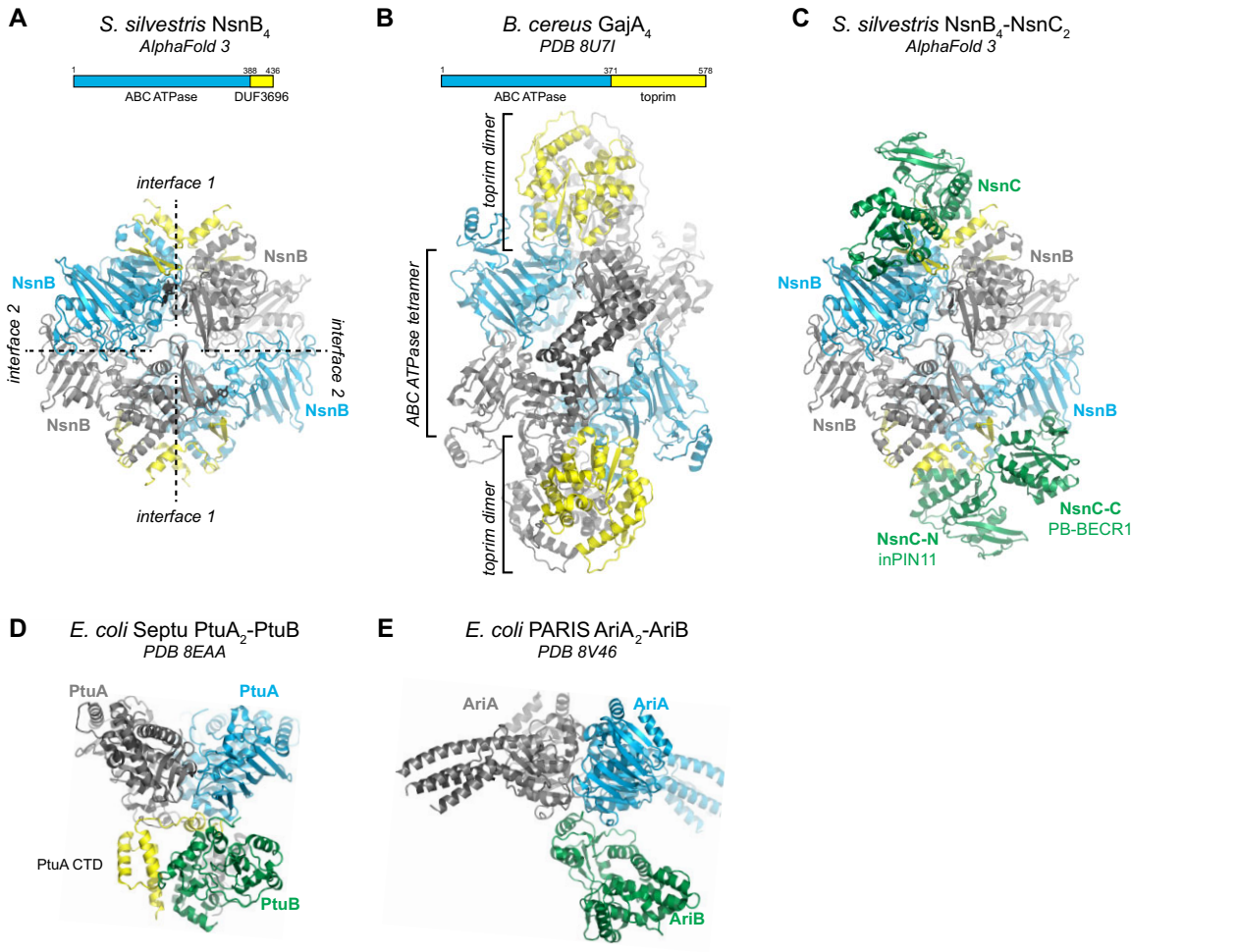


Figure 5. Structure prediction of an NsnB–NsnC complex. (**A**) AlphaFold 3 model of an *S. silvestris* StLB046 NsnB homotetramer. Two NsnB protomers are colored with their ABC ATPase domains blue and their C-terminal DUF3696 domains yellow, and two are colored gray and yellow. Bound ATP and Mg^{2+} ions are shown in black. (**B**) CryoEM structure of *B. cereus* Gabjia GajA homotetramer (PDB ID: 8U7I) [94]. Two GajA protomers are colored with their ABC ATPase domains blue and their C-terminal toprim domains yellow, and two are colored gray. (**C**) AlphaFold 3 model of an *S. silvestris* StLB046 $NsnB_4(ATP{\cdot}Mg^{2+})$–$NsnC_2$ complex. NsnB is colored as in panel (A), and NsnC is colored green. ATP and Mg^{2+} are not shown. See Supplementary Fig. S7 for PAE plot and for models of 24 additional $NsnB_4(ATP{\cdot}Mg^{2+})$–$NsnC_2$ complexes. (**D**) CryoEM structure of *E. coli* Septu, showing two protomers of PtuA (blue/yellow and gray) bound to one protomer of PtuB (green) (PDB ID: 8EEA) [103]. (**E**) CryoEM structure of *E. coli* PARIS, showing two protomers of AriA (blue and gray) bound to one protomer of AriB (green) (PDB ID: 8V46) [96].

Evolution of the Menshen system

In light of the extensive diversity and polymorphism observed in three regions of the Menshen system (Fig. 2D), we sought to unravel its likely evolutionary trajectory. We constructed a phylogenetic tree of the ParB-DUF262 domains collected from all the Menshen loci and examined the distribution patterns of the associated NsnA-N, NsnC-N, and NsnC-C domains, along with the bacterial lineages in which they are found (Fig. 6A). Notably, the majority of the loci across the entire tree belong to Type I Menshen, characterized by the absence of the NsnA-N domain, and possess inPIN1 and PB-BECR1 domains in NsnC. Based on its frequency among Menshen systems, we propose that this domain arrangement in NsnA and NsnC represents the ancestral state of the Menshen system, which was subsequently augmented by addition of NsnA-N domains and diversification of NsnC.

Our analysis revealed a complex diversification process during evolution of Menshen, involving domain acquisition, shuffling, loss, and gene transfers. First, the emergence of the Type II Menshen instances appears to be largely shaped by clade-specific recruitment of NsnA-N sensor domains by Type I Menshen systems (Fig. 6A). Moreover, we observed, beyond the prevalent inPIN1 and PB-BECR1 domains, numerous closely related ParB-DUF262 domains exhibited associations with distinct NsnC-N inPIN domains and NsnC-C effector domains. This suggests that extensive domain shuffling/diversification has occurred to replace ancestral versions with new variations during evolution. Additionally, the acquisition of the same sensor domains in NsnA-N and the substitution of the same effector domains in NsnC-C occurred independently and repeatedly. Examples include the PsPUA1 domain and the Ntox18 domain, which are distributed at multiple positions in the phylogenetic tree (Fig. 6A). We also note certain instances of domain loss following domain acquisition and shuffling for either NsnA-N sensor, NsnC-N inPIN, or NsnC-C effector domains (Fig. 6A and Supplementary Table S1). Finally, the presence of mixed bacterial lineages within the same phylogenetic clade of ParB-DUF262, combined with the similar domain architectures of NsnA and NsnB, suggests extensive gene transfer events of

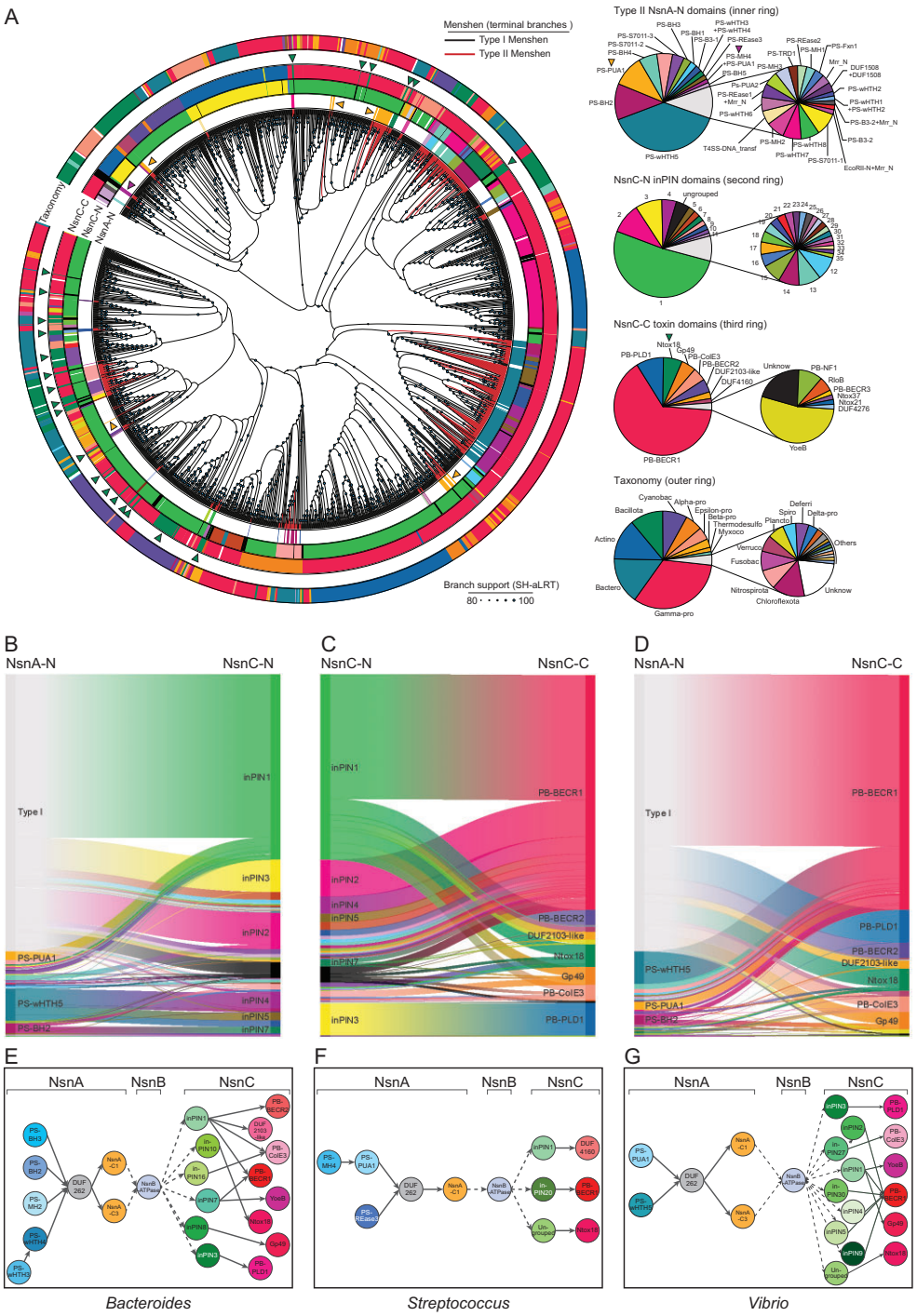


Figure 6. Complex evolutionary trajectory of the Menshen system. (**A**) Phylogeny of the ParB-DUF262 domains in Menshen systems, their associations with diverse NsnA-N sensor domains, NsnC-N inPIN domains, NsnC-C effector domains, and the distribution of these components across bacterial phylogenetic lineages. The terminal branches of the tree are colored based on the types of Menshen instances (Type I: black; Type II: red). Branch supports (SH-aLRT test) above 80 are indicated by blue circles on branches. The tree is further annotated with multiple layers of rings: the inner ring indicates different NsnA-N sensor domains, the second ring indicates different NsnC-N inPIN domains, the third ring indicates various NsnC-C effector domains, and the outermost ring indicates the distribution of bacterial phylogenetic lineages. If no corresponding domain exists in the first three layers, or insufficient taxonomic information is available for the fourth layer, the positions of the rings are left blank. Domain/taxonomy types and their frequencies are summarized as pie charts shown on the right. Specific domains discussed in the main text, including PsPUA1, PsMH4 + PsPUA1, and Ntox18, are marked by triangles positioned along the relevant layers of the rings. The abbreviations in the taxonomy pie chart are: Gamma-pro for Gammaproteobacteria, Bactero for Bacteroidota, Actino for Actinomycetota, Cyanobac for Cyanobacteriota, Alpha-pro for Alphaproteobacteria, Epsilon-pro for Epsilonproteobacteria, Beta-pro for Betaproteobacteria, Thermodesulfo for Thermodesulfobacteriota, Myxoco for Myxococcota, Fusobac for Fusobacteriota, Verruco for Verrucomicrobiota, Plancto for Planctomycetota, Spiro for Spirochaetota, Deferri for Deferribacterota, and Delta-pro Deltaproteobacteria. (**B**) Associations between NsnA-N sensor domains and NsnC-N inPIN domains. (**C**) Associations between NsnC-N inPIN domains and NsnC-C effector domains. (**D**) Associations between NsnA-N sensor domains and NsnC-C effector domains. Each domain is indicated by the same type of color throughout the entire figure, but black indicates ungrouped inPIN sequences or unknown effector domains. (**E**, **F**, **G**) The complexity of the Menshen systems in representative bacterial genus, including *Bacteroides*, *Streptococcus*, and *Vibrio*.

Menshen during evolution (Fig. 6A). These events might have also facilitated the domain acquisition and shuffling that we have observed in the system.

The observed complex diversification patterns (Fig. 6B–D) likely reflect the ongoing arms race between bacteria and phages. First, the independent recruitment of distinct, additional NsnA-N sensor domains to the existing ParB-DUF262 and ParBDB domains likely represent an adaptative strategy of Menshen to expand its ability to sense specific phage DNA elements. These new sensors are likely crucial for Menshen to retain activity as phage inevitably evolve mechanisms to evade the basal recognition provided by the ParB-DUF262 and ParBDB domains. Second, the continuous shuffling and diversification of the NsnC-N inPIN domains and NsnC-C effector domains could also be a strategy to combat evolving phage resistance, especially given that phage molecules targeted by antiphage systems can evolve rapidly [98]. This accounts for Menshen loci that either share the same effector domain but differ their inPIN domains (e.g. PS-wHTH5-PB-BECR1-containing loci) or have the same inPIN domain followed by different effector domains (e.g. loci led by PS-BH2 sensors). Last, the sporadic domain loss in Menshen systems, potentially through pseudogenization of sensor, inPIN, or effector domains, could contribute to the development of phage tolerance, a bacterial adaptation to gain advantage from phages in specific environments.

Taxonomic distribution of Menshen and its association with other antiphage systems

An analysis of the taxonomic distribution of species containing the Menshen operons revealed that Menshen is widely distributed across a diverse range of bacteria, spanning at least 27 different phyla (Fig. 7A and Supplementary Table S1). These include *gammaproteobacteria* (33.1%), *bacteroidota* (15.47%), *actinomycetota* (12.47%), *bacillota* (9.91%), *cyanobacteriota* (8.16%), and *alphaproteobacteria* (5.76%). This broad distribution highlights the prevalence and significance of the Menshen system across the bacterial kingdom. Our analysis also showed that most bacterial genomes contain a single version of the Menshen system. Notably, many genomes from the same genus harbor distinct versions of the system (Fig. 6E–G), suggesting that Menshen is typically unique to each genome. However, we identified 53 genomes spanning a wide range of bacterial lineages that carry two or three distinct Menshen systems (Fig. 2C). This retention of multiple variants within the same genome may enable bacteria to defend against diverse phages, providing a broader and more versatile antiphage defense strategy.

Menshen systems are also frequently associated with other antiphage systems in their gene neighborhood. About 27.9% of Menshen operons are linked with systems such as BREX, Type I and Type III R-M systems, and various methylase genes (Fig. 7B–F). This association supports the role of Menshen as an antiphage system and indicates potential cooperative or complementary interactions with these R-M and R-M-like systems. It is worth noting that for Menshen operons associated with methylase genes, we also often find other nuclease genes in proximity. These nuclease genes could potentially serve as restriction components for the methylases. However, there are instances where nucleases are absent, raising questions about whether the associated methylases are primed for their role in conjunction with the Menshen systems, wherein the ParB-DUF262 domains may target non-modified genomes.

Final remarks

Our detailed investigation has led to a comprehensive dissection of the Menshen antiphage system, revealing its complex architecture involving NsnA, NsnB, and NsnC proteins. Notably, extensive polymorphism is observed in the domain compositions of NsnA and NsnC, underscoring the system's evolutionary adaptability. NsnA is characterized by the core ParB-DUF262 and ParBDB domains, along with multiple N-terminal DNA-binding domains. The shared DNA-binding capabilities of these domains—ParB-DUF262, ParBDB, and various N-terminal DNA-binding domains—strongly suggests that NsnA acts as the sensory module of the system, primarily detecting phage genomic DNA. This detection likely serves as the initial trigger for system activation. NsnC exhibits extensive polymorphism in its N-terminal inPIN domains, which are predicted to possess RNA-binding activity, and in its C-terminal effector domains, most of which are RNases. These features strongly support the hypothesis that NsnC acts as the effector component of the system, with RNA molecules likely being its primary targets. The pronounced polymorphism of NsnC likely reflects an evolutionary arms race between phage and bacteria, suggesting that NsnC may specifically target phage-encoded RNAs. However, our analysis cannot definitively predict whether Menshen targets phage RNAs or instead acts on host RNAs as part of an abortive infection-like defense strategy.

Additionally, our analysis of protein-protein interactions in Menshen suggests that NsnB serves as a transducer for signal transmission from the sensory (NsnA-like) to the effector (NsnC-like) components. We find that NsnB is predicted to form a homotetramer like that of the GajA ATPase subunit of Gabija, and is predicted to bind NsnC in a manner similar to that of Septu and PARIS, two other antiphage systems incorporating ABC ATPases and toxin effectors. In Gabija, the GajA ATPase domain is thought to regulate its C-terminal toprim nuclease domain [93, 94], and in both Septu and PARIS the ATPase subunit regulates the activity of the bound nuclease toxin [95–96]. Relatedly, in antiphage systems involving RloC [99] and PrrC [100], activated ABC ATPases are also known to regulate nuclease domains. We hypothesize that NsnB controls NsnC activity through either coordinated conformational changes or NsnC binding/release that are coupled to ATP binding and/or hydrolysis. Unfortunately, our analysis does not yield a definitive hypothesis regarding how NsnA signals to NsnB. However, in analogous systems like the ParABS and SMC chromosome-partition complexes, the ParB component relies on DNA binding to modulate the activity of an associated ATPase, either the ParA ATPase [101, 104, 105] or the SMC ATPase [101, 102]. By analogy, it is plausible that in Menshen, NsnA, upon binding DNA, similarly stimulates the ATPase activity of NsnB.

Consequently, we can propose a unified model of Menshen's mechanism of action, despite its diversity in domain composition. In both Type I and Type II Menshen, NsnA proteins detect invading phage DNAs via ParB-DUF262, ParBDB, and NsnA-N domains, acting as the initiators of the response. Then NsnAs activate the NsnB-like ABC ATPase, which function as inducers to drive the activation of the NsnC effector proteins. These effectors, armed with inPIN domains, are

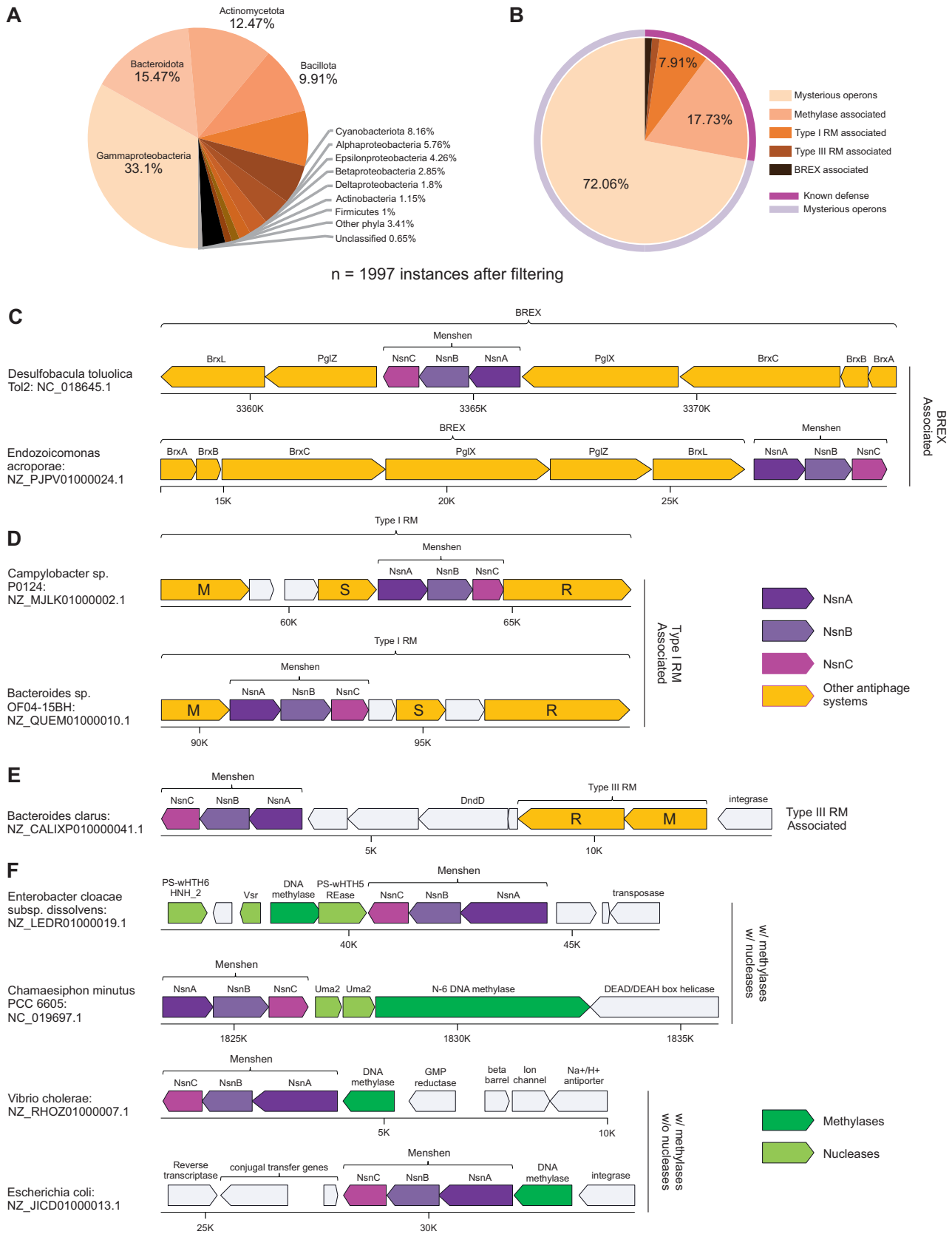


Figure 7. (**A**) A statistical summary of the taxonomic distribution of genomes containing the Menshen system. (**B**) A statistical summary of various types of genomic neighborhoods of the Menshen system. Representative examples of gene neighborhoods of the Menshen system associated with the BREX antiphage system (**C**), Type I R-M system (**D**), Type III R-M system (**E**), and the methylase-nuclease pairs or solitary methylases (**F**).

likely designed to recognize and degrade specific nucleic acids. Our identification of distinct domains, their predicted functions, and the hypothesized interactions within the Menshen system provides a valuable framework for understanding its role in bacterial antiphage defense. We anticipate that these insights will inspire experimental validation and more sophisticated investigations into Menshen's diverse mechanisms and evolutionary dynamics.

Acknowledgements

Author contributions: Huan Li (Conceptualization, Datacuration, Formal analysis, Investigation, Methodology, Software, Visualization, Writing – original draft), Yongjun Tan (Formal analysis, Investigation, Validation, Visualization), Dwaipayan Basu (Formal analysis, Investigation, Methodology, Visualization), Kevin D. Corbett (Conceptualization, Formal analysis, Funding acquisition, Writing – review & editing), Dapeng Zhang (Conceptualization, Formal analysis, Funding acquisition, Project administration, Supervision, Writing – original draft, Writing – review & editing).

Supplementary data

Supplementary data is available at NAR online.

Conflict of interest

The authors declare no competing interests.

Funding

This work was supported by Saint Louis University President's Research Fund (to D.Z.) and National Institutes of Health (R35 GM144121 to K.D.C.). Funding to pay the Open Access publication charges for this article was provided by Saint Louis University.

Data availability

Complete annotations and structural models can be found in the supplementary data.

References

1. Georjon H, Bernheim A. The highly diverse antiphage defence systems of bacteria. *Nat Rev Microbiol* 2023;**21**:686–700. https://doi.org/10.1038/s41579-023-00934-x
2. Tesson F, Herve A, Mordret E *et al.* Systematic and quantitative view of the antiviral arsenal of prokaryotes. *Nat Commun* 2022;**13**:2561. https://doi.org/10.1038/s41467-022-30269-9
3. Aravind L. Guilt by association: contextual information in genome analysis. *Genome Res* 2000;**10**:1074–7. https://doi.org/10.1101/gr.10.8.1074
4. Makarova KS, Wolf YI, Snir S *et al.* Defense islands in bacterial and archaeal genomes and prediction of novel defense systems. *J Bacteriol* 2011;**193**:6039–56. https://doi.org/10.1128/JB.05535-11
5. Makarova KS, Wolf YI, Koonin EV. Comparative genomics of defense systems in archaea and bacteria. *Nucleic Acids Res* 2013;**41**:4360–77. https://doi.org/10.1093/nar/gkt157
6. Burroughs AM, Zhang D, Schaffer DE *et al.* Comparative genomic analyses reveal a vast, novel network of nucleotide-centric systems in biological conflicts, immunity and signaling. *Nucleic Acids Res* 2015;**43**:10633–54. https://doi.org/10.1093/nar/gkv1267
7. Mariano G, Blower TR. Conserved domains can be found across distinct phage defence systems. *Mol Microbiol* 2023;**120**:45–53. https://doi.org/10.1111/mmi.15047
8. Osorio-Valeriano M, Altegoer F, Das CK *et al.* The CTPase activity of ParB determines the size and dynamics of prokaryotic DNA partition complexes. *Mol Cell* 2021;**81**:3992–4007. https://doi.org/10.1016/j.molcel.2021.09.004
9. Machnicka MA, Kaminska KH, Dunin-Horkawicz S *et al.* Phylogenomics and sequence–structure–function relationships in the GmrSD family of type IV restriction enzymes. *BMC Bioinformatics* 2015;**16**:336. https://doi.org/10.1186/s12859-015-0773-z
10. Bair CL, Black LW. A type IV modification dependent restriction nuclease that targets glucosylated hydroxymethyl cytosine modified DNAs. *J Mol Biol* 2007;**366**:768–78. https://doi.org/10.1016/j.jmb.2006.11.051
11. Xiong X, Wu G, Wei Y *et al.* SspABCD-SspE is a phosphorothioation-sensing bacterial defence system with broad anti-phage activities. *Nat Microbiol* 2020;**5**:917–28. https://doi.org/10.1038/s41564-020-0700-6
12. Gao H, Gong X, Zhou J *et al.* Nicking mechanism underlying the DNA phosphorothioate-sensing antiphage defense by SspE. *Nat Commun* 2022;**13**:6773. https://doi.org/10.1038/s41467-022-34505-0
13. Vassallo CN, Doering CR, Littlehale ML *et al.* A functional selection reveals previously undetected anti-phage defence systems in the *E. coli* pangenome. *Nat Microbiol* 2022;**7**:1568–79. https://doi.org/10.1038/s41564-022-01219-4
14. Millman A, Melamed S, Leavitt A *et al.* An expanded arsenal of immune systems that protect bacteria from phages. *Cell Host Microbe* 2022;**30**:1556–69. https://doi.org/10.1016/j.chom.2022.09.017
15. Gao L, Altae-Tran H, Bohning F *et al.* Diverse enzymatic activities mediate antiviral immunity in prokaryotes. *Science* 2020;**369**:1077–84. https://doi.org/10.1126/science.aba0372
16. Ofir G, Herbst E, Baroz M *et al.* Antiviral activity of bacterial TIR domains via immune signalling molecules. *Nature* 2021;**600**:116–20. https://doi.org/10.1038/s41586-021-04098-7
17. Cheng R, Huang F, Wu H *et al.* A nucleotide-sensing endonuclease from the Gabija bacterial defense system. *Nucleic Acids Res* 2021;**49**:5216–29. https://doi.org/10.1093/nar/gkab277
18. Cheng R, Huang F, Lu X *et al.* Prokaryotic Gabija complex senses and executes nucleotide depletion and DNA cleavage for antiviral defense. *Cell Host Microbe* 2023;**31**:1331–44. https://doi.org/10.1016/j.chom.2023.06.014
19. Loeff L, Walter A, Rosalen GT *et al.* DNA end sensing and cleavage by the Shedu anti-phage defense system. *Cell* 2025;**188**:721–33. https://doi.org/10.1016/j.cell.2024.11.030
20. Gu Y, Li H, Deep A *et al.* Bacterial Shedu immune nucleases share a common enzymatic core regulated by diverse sensor domains. *Mol Cell* 2025;**85**:523–36. https://doi.org/10.1016/j.molcel.2024.12.004
21. Doron S, Melamed S, Ofir G *et al.* Systematic discovery of antiphage defense systems in the microbial pangenome. *Science* 2018;**359**:eaar4120. https://doi.org/10.1126/science.aar4120
22. Hochhauser D, Millman A, Sorek R. The defense island repertoire of the *Escherichia coli* pan-genome. *PLoS Genet* 2023;**19**:e1010694. https://doi.org/10.1371/journal.pgen.1010694
23. Li H, Tan Y, Zhang D. Genomic discovery and structural dissection of a novel type of polymorphic toxin system in gram-positive bacteria. *Comput Struct Biotechnol J* 2022;**20**:4517–4531. https://doi.org/10.1016/j.csbj.2022.08.036
24. Altschul SF, Madden TL, Schaffer AA *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search

programs. *Nucleic Acids Res* 1997;**25**:3389–402. https://doi.org/10.1093/nar/25.17.3389
25. Lassmann T, Sonnhammer EL. Kalign–an accurate and fast multiple sequence alignment algorithm. *BMC Bioinformatics* 2005;**6**:298. https://doi.org/10.1186/1471-2105-6-298
26. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 2004;**32**:1792–7. https://doi.org/10.1093/nar/gkh340
27. Pei J, Grishin NV. PROMALS3D: multiple protein sequence alignment enhanced with evolutionary and three-dimensional structural information. *Methods Mol Biol* 2014;**1079**:263–71.
28. Taylor WR. The classification of amino acid conservation. *J Theor Biol* 1986;**119**:205–18. https://doi.org/10.1016/S0022-5193(86)80075-3
29. Eddy SR. Accelerated profile HMM searches. *PLoS Comput Biol* 2011;7:e1002195. https://doi.org/10.1371/journal.pcbi.1002195
30. Crooks GE, Hon G, Chandonia JM *et al.* WebLogo: a sequence logo generator. *Genome Res* 2004;**14**:1188–90. https://doi.org/10.1101/gr.849004
31. Aravind L, Iyer LM, Burroughs AM. Discovering biological conflict systems through genome analysis: evolutionary principles and biochemical novelty. *Annu Rev Biomed Data Sci* 2022;**5**:367–91. https://doi.org/10.1146/annurev-biodatasci-122220-101119
32. Mistry J, Chuguransky S, Williams L *et al.* Pfam: the protein families database in 2021. *Nucleic Acids Res* 2021;**49**:D412–9. https://doi.org/10.1093/nar/gkaa913
33. Kibby EM, Conte AN, Burroughs AM *et al.* Bacterial NLR-related proteins protect against phage. *Cell* 2023;**186**:2410–24. https://doi.org/10.1016/j.cell.2023.04.015
34. Frickey T, Lupas A. CLANS: a Java application for visualizing protein families based on pairwise similarity. *Bioinformatics* 2004;**20**:3702–4. https://doi.org/10.1093/bioinformatics/bth444
35. Fruchterman TMJ, Reingold EM. Graph drawing by force-directed placement. *Softw Pract Exp* 1991;**21**:1129–64. https://doi.org/10.1002/spe.4380211102
36. Enright AJ, Ouzounis CA. BioLayout—an automatic graph layout algorithm for similarity visualization. *Bioinformatics* 2001;**17**:853–4. https://doi.org/10.1093/bioinformatics/17.9.853
37. Berman HM, Westbrook J, Feng Z *et al.* The Protein Data Bank. *Nucleic Acids Res* 2000;**28**:235–42. https://doi.org/10.1093/nar/28.1.235
38. Jumper J, Evans R, Pritzel A *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* 2021;**596**:583–9. https://doi.org/10.1038/s41586-021-03819-2
39. Mirdita M, Schutze K, Moriwaki Y *et al.* ColabFold: making protein folding accessible to all. *Nat Methods* 2022;**19**:679–82. https://doi.org/10.1038/s41592-022-01488-1
40. Abramson J, Adler J, Dunger J *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* 2024;**630**:493–500. https://doi.org/10.1038/s41586-024-07487-w
41. Holm L, Laiho A, Toronen P *et al.* DALI shines a light on remote homologs: one hundred discoveries. *Protein Sci* 2023;**32**:e4519. https://doi.org/10.1002/pro.4519
42. DeLano WL. Pymol: an open-source molecular graphics tool. *CCP4 Newsl Protein Crystallogr* 2002;**40**:82–92.
43. Rose AS, Hildebrand PW. NGL Viewer: a web application for molecular visualization. *Nucleic Acids Res* 2015;**43**:W576–9. https://doi.org/10.1093/nar/gkv402
44. Nguyen LT, Schmidt HA, von Haeseler A *et al.* IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* 2015;**32**:268–74. https://doi.org/10.1093/molbev/msu300
45. Letunic I, Bork P. Interactive Tree of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res* 2021;**49**:W293–6. https://doi.org/10.1093/nar/gkab301
46. Jonsson TJ, Murray MS, Johnson LC *et al.* Reduction of cysteine sulfinic acid in peroxiredoxin by sulfiredoxin proceeds directly through a sulfinic phosphoryl ester intermediate. *J Biol Chem* 2008;**283**:23846–51. https://doi.org/10.1074/jbc.M803244200
47. Maindola P, Raina R, Goyal P *et al.* Multiple enzymatic activities of ParB/Srx superfamily mediate sexual conflict among conjugative plasmids. *Nat Commun* 2014;**5**:5322. https://doi.org/10.1038/ncomms6322
48. Xia S, Chen J, Liu L *et al.* Tight control of genomic phosphorothioate modification by the ATP-modulated autoregulation and reusability of DndB. *Mol Microbiol* 2019;**111**:938–50. https://doi.org/10.1111/mmi.14186
49. Picton DM, Luyten YA, Morgan RD *et al.* The phage defence island of a multidrug resistant plasmid uses both BREX and type IV restriction for complementary protection from viruses. *Nucleic Acids Res* 2021;**49**:11257–73. https://doi.org/10.1093/nar/gkab906
50. Jalal ASB, Tran NT, Stevenson CE *et al.* Diversification of DNA-binding specificity by permissive and specificity-switching mutations in the ParB/Noc protein Family. *Cell Rep* 2020;**32**:107928. https://doi.org/10.1016/j.celrep.2020.107928
51. Krishnan A, Burroughs AM, Iyer LM *et al.* Comprehensive classification of ABC ATPases and their functional radiation in nucleoprotein dynamics and biological conflict systems. *Nucleic Acids Res* 2020;**48**:10045–75. https://doi.org/10.1093/nar/gkaa726
52. Kamada K, Su'etsugu M, Takada H *et al.* Overall shapes of the SMC–ScpAB complex are determined by balance between constraint and relaxation of its structural parts. *Structure* 2017;**25**:603–16. https://doi.org/10.1016/j.str.2017.02.008
53. Arcus VL, Backbro K, Roos A *et al.* Distant structural homology leads to the functional characterization of an archaeal PIN domain as an exonuclease. *J Biol Chem* 2004;**279**:16471–8. https://doi.org/10.1074/jbc.M313833200
54. Senissar M, Manav MC, Brodersen DE. Structural conservation of the PIN domain active site across all domains of life. *Protein Sci* 2017;**26**:1474–92. https://doi.org/10.1002/pro.3193
55. Grishin NV. Mh1 domain of Smad is a degraded homing endonuclease. *J Mol Biol* 2001;**307**:31–7. https://doi.org/10.1006/jmbi.2000.4486
56. Li J, Bonkowski MS, Moniot S *et al.* A conserved NAD(+) binding pocket that regulates protein–protein interactions during aging. *Science* 2017;**355**:1312–7. https://doi.org/10.1126/science.aad8242
57. Zhang D, de Souza RF, Anantharaman V *et al.* Polymorphic toxin systems: comprehensive characterization of trafficking modes, processing, mechanisms of action, immunity and ecology using comparative genomics. *Biol Direct* 2012;**7**:18. https://doi.org/10.1186/1745-6150-7-18
58. Im H, Jang SB, Pathak C *et al.* Crystal structure of toxin HP0892 from *Helicobacter pylori* with two Zn(II) at 1.8 A resolution. *Protein Sci* 2014;**23**:819–32. https://doi.org/10.1002/pro.2465
59. Mansour M, Giudice E, Xu X *et al.* Substrate recognition and cryo-EM structure of the ribosome-bound TAC toxin of *Mycobacterium tuberculosis*. *Nat Commun* 2022;**13**:2641. https://doi.org/10.1038/s41467-022-30373-w
60. Zhang D, Iyer LM, Aravind L. A novel immunity system for bacterial nucleic acid degrading toxins and its recruitment in various eukaryotic and DNA viral systems. *Nucleic Acids Res* 2011;**39**:4532–52. https://doi.org/10.1093/nar/gkr036
61. Zhang D, Burroughs AM, Vidal ND *et al.* Transposons to toxins: the provenance, architecture and diversification of a widespread class of eukaryotic effectors. *Nucleic Acids Res* 2016;**44**:3513–33. https://doi.org/10.1093/nar/gkw221
62. Aravind L, Anantharaman V, Balaji S *et al.* The many faces of the helix-turn-helix domain: transcription regulation and beyond. *FEMS Microbiol Rev* 2005;**29**:231–62. https://doi.org/10.1016/j.femsre.2004.12.008
63. Sukackaite R, Grazulis S, Bochtler M *et al.* The recognition domain of the BpuJI restriction endonuclease in complex with

cognate DNA at 1.3-A resolution. *J Mol Biol* 2008;**378**:1084–93. https://doi.org/10.1016/j.jmb.2008.03.041
64. Wah DA, Hirsch JA, Dorner LF *et al.* Structure of the multimodular endonuclease FokI bound to DNA. *Nature* 1997;**388**:97–100. https://doi.org/10.1038/40446
65. Roberts RJ, Vincze T, Posfai J *et al.* REBASE—a database for DNA restriction and modification: enzymes, genes and genomes. *Nucleic Acids Res* 2015;**43**:D298–9. https://doi.org/10.1093/nar/gku1046
66. Tao Z, Hu H, Luo X *et al.* Embryonic resetting of the parental vernalized state by two B3 domain transcription factors in Arabidopsis. *Nat Plants* 2019;**5**:424–35. https://doi.org/10.1038/s41477-019-0402-3
67. Zhou XE, Wang Y, Reuter M *et al.* Crystal structure of type IIE restriction endonuclease EcoRII reveals an autoinhibition mechanism by a novel effector-binding fold. *J Mol Biol* 2004;**335**:307–19. https://doi.org/10.1016/j.jmb.2003.10.030
68. Golovenko D, Manakova E, Zakrys L *et al.* Structural insight into the specificity of the B3 DNA-binding domains provided by the co-crystal structure of the C-terminal fragment of BfiI restriction enzyme. *Nucleic Acids Res* 2014;**42**:4113–22. https://doi.org/10.1093/nar/gkt1368
69. Andreeva A, Kulesha E, Gough J *et al.* The SCOP database in 2020: expanded classification of representative family and superfamily domains of known protein structures. *Nucleic Acids Res* 2020;**48**:D376–82. https://doi.org/10.1093/nar/gkz1064
70. Iyer LM, Zhang D, Burroughs AM *et al.* Computational identification of novel biochemical systems involved in oxidation, glycosylation and other complex modifications of bases in DNA. *Nucleic Acids Res* 2013;**41**:7635–55. https://doi.org/10.1093/nar/gkt573
71. Pastor M, Czapinska H, Helbrecht I *et al.* Crystal structures of the EVE-HNH endonuclease VcaM4I in the presence and absence of DNA. *Nucleic Acids Res* 2021;**49**:1708–23. https://doi.org/10.1093/nar/gkaa1218
72. Zagorskaite E, Manakova E, Sasnauskas G. Recognition of modified cytosine variants by the DNA-binding domain of methyl-directed endonuclease McrBC. *FEBS Lett* 2018;**592**:3335–45. https://doi.org/10.1002/1873-3468.13244
73. Gao Y, Cao D, Zhu J *et al.* Structural insights into assembly, operation and inhibition of a type I restriction-modification system. *Nat Microbiol* 2020;**5**:1107–18. https://doi.org/10.1038/s41564-020-0731-z
74. Youell J, Firman K. Mechanistic insight into type I restriction endonucleases. *Front Biosci* 2012;**17**:2122–39. https://doi.org/10.2741/4041
75. Graebsch A, Roche S, Niessing D. X-ray structure of Pur-alpha reveals a whirly-like fold and an unusual nucleic-acid binding surface. *Proc Natl Acad Sci USA* 2009;**106**:18521–6. https://doi.org/10.1073/pnas.0907990106
76. Schroder G, Krause S, Zechner EL *et al.* TraG-like proteins of DNA transfer systems and of the *Helicobacter pylori* type IV secretion system: inner membrane gate for exported substrates? *J Bacteriol* 2002;**184**:2767–79. https://doi.org/10.1128/JB.184.10.2767-2779.2002
77. Zachrdla M, Padrta P, Rabatinova A *et al.* Solution structure of domain 1.1 of the sigma(A) factor from *Bacillus subtilis* is preformed for binding to the RNA polymerase core. *J Biol Chem* 2017;**292**:11610–7. https://doi.org/10.1074/jbc.M117.784074
78. Pavelich IJ, Maehigashi T, Hoffer ED *et al.* Monomeric YoeB toxin retains RNase activity but adopts an obligate dimeric form for thermal stability. *Nucleic Acids Res* 2019;**47**:10400–13. https://doi.org/10.1093/nar/gkz760
79. Beck CM, Morse RP, Cunningham DA *et al.* CdiA from *Enterobacter cloacae* delivers a toxic ribosomal RNase into target bacteria. *Structure* 2014;**22**:707–18. https://doi.org/10.1016/j.str.2014.02.012
80. Aravind L, Leipe DD, Koonin EV. Toprim—a conserved catalytic domain in type IA and II topoisomerases, DnaG-type primases, OLD family nucleases and RecR proteins. *Nucleic Acids Res* 1998;**26**:4205–13. https://doi.org/10.1093/nar/26.18.4205
81. Oerum S, Catala M, Bourguet M *et al.* Structural studies of RNase M5 reveal two-metal-ion supported two-step dsRNA cleavage for 5S rRNA maturation. *RNA Biology* 2021;**18**:1996–2006. https://doi.org/10.1080/15476286.2021.1885896
82. Schiltz CJ, Lee A, Partlow EA *et al.* Structural characterization of class 2 OLD family nucleases supports a two-metal catalysis mechanism for cleavage. *Nucleic Acids Res* 2019;**47**:9448–63. https://doi.org/10.1093/nar/gkz703
83. Maestri A, Pons BJ, Pursey E *et al.* The bacterial defense system MADS interacts with CRISPR–Cas to limit phage infection and escape. *Cell Host Microbe* 2024;**32**:1412–26. https://doi.org/10.1016/j.chom.2024.07.005
84. Ernits K, Saha CK, Brodiazhenko T *et al.* The structural basis of hyperpromiscuity in a core combinatorial network of type II toxin–antitoxin and related phage defense systems. *Proc Natl Acad Sci USA* 2023;**120**:e2305393120. https://doi.org/10.1073/pnas.2305393120
85. Celma L, Corbinais C, Vercruyssen J *et al.* Structural basis for the substrate selectivity of *Helicobacter pylori* NucT nuclease activity. *PLoS One* 2017;**12**:e0189049. https://doi.org/10.1371/journal.pone.0189049
86. Lountos GT, Zhao XZ, Kiselev E *et al.* Identification of a ligand binding hot spot and structural motifs replicating aspects of tyrosyl-DNA phosphodiesterase I (TDP1) phosphoryl recognition by crystallographic fragment cocktail screening. *Nucleic Acids Res* 2019;**47**:10134–50. https://doi.org/10.1093/nar/gkz515
87. Hadders MA, Agromayor M, Obita T *et al.* ESCRT-III binding protein MITD1 is involved in cytokinesis and has an unanticipated PLD fold that binds membranes. *Proc Natl Acad Sci USA* 2012;**109**:17424–9. https://doi.org/10.1073/pnas.1206839109
88. Shen BW, Doyle LA, Werther R *et al.* Structure, substrate binding and activity of a unique AAA+ protein: the BrxL phage restriction factor. *Nucleic Acids Res* 2023;**51**:3513–28. https://doi.org/10.1093/nar/gkad083
89. Boss L, Gorniak M, Lewanczyk A *et al.* Identification of three type II toxin–antitoxin systems in model bacterial plant pathogen *Dickeya dadantii* 3937. *Int J Mol Sci* 2021;**22**:5932. https://doi.org/10.3390/ijms22115932
90. Mogila I, Tamulaitiene G, Keda K *et al.* Ribosomal stalk-captured CARF-RelE ribonuclease inhibits translation following CRISPR signaling. *Science* 2023;**382**:1036–41. https://doi.org/10.1126/science.adj2107
91. Evans R, O'Neill M, Pritzel A *et al.* Protein complex prediction with AlphaFold-Multimer. bioRxiv, https://doi.org/10.1101/2021.10.04.463034, 10 March 2022, preprint: not peer reviewed.
92. van Kempen M, Kim SS, Tumescheit C *et al.* Fast and accurate protein structure search with Foldseek. *Nat Biotechnol* 2024;**42**:243–6. https://doi.org/10.1038/s41587-023-01773-0
93. Li J, Cheng R, Wang Z *et al.* Structures and activation mechanism of the Gabija anti-phage system. *Nature* 2024;**629**:467–73. https://doi.org/10.1038/s41586-024-07270-x
94. Antine SP, Johnson AG, Mooney SE *et al.* Structural basis of Gabija anti-phage defence and viral immune evasion. *Nature* 2024;**625**:360–5. https://doi.org/10.1038/s41586-023-06855-2
95. Azam AH, Kondo K, Chihara K *et al.* Evasion of antiviral bacterial immunity by phage tRNAs. *Nat Commun* 2024;**15**:9586. https://doi.org/10.1038/s41467-024-53789-y
96. Deep A, Liang Q, Enustun E *et al.* Architecture and activation mechanism of the bacterial PARIS defence system. *Nature* 2024;**634**:432–9. https://doi.org/10.1038/s41586-024-07772-8
97. Burman N, Belukhina S, Depardieu F *et al.* A virally encoded tRNA neutralizes the PARIS antiviral defence system. *Nature* 2024;**634**:424–31. https://doi.org/10.1038/s41586-024-07874-3

98. Samson JE, Magadan AH, Sabri M *et al.* Revenge of the phages: defeating bacterial defences. *Nat Rev Micro* 2013;**11**:675–87. https://doi.org/10.1038/nrmicro3096
99. Klaiman D, Steinfels-Kohn E, Kaufmann G. A DNA break inducer activates the anticodon nuclease RloC and the adaptive immunity in *Acinetobacter baylyi* ADP1. *Nucleic Acids Res* 2014;**42**:328–39. https://doi.org/10.1093/nar/gkt851
100. Meineke B, Shuman S. Determinants of the cytotoxicity of PrrC anticodon nuclease and its amelioration by tRNA repair. *RNA* 2012;**18**:145–54. https://doi.org/10.1261/rna.030171.111
101. Jalal ASB, Le TBK. Bacterial chromosome segregation by the ParABS system. *Open Biol* 2020;**10**:200097. https://doi.org/10.1098/rsob.200097
102. Gruber S, Errington J. Recruitment of condensin to replication origin regions by ParB/SpoOJ promotes chromosome segregation in *B. subtilis*. *Cell* 2009;**137**:685–96. https://doi.org/10.1016/j.cell.2009.02.035
103. Li Y, Shen Z, Zhang M *et al.* PtuA and PtuB assemble into an inflammasome-like oligomer for anti-phage defense. *Nat Struct Mol Biol* 2024;**31**:413–23. https://doi.org/10.1038/s41594-023-01172-8
104. Gerdes K, Howard M, Szardenings F. Pushing and pulling in prokaryotic DNA segregation. *Cell* 2010;**141**:927–42. https://doi.org/10.1016/j.cell.2010.05.033
105. Davis MA, Martin KA, Austin SJ. Biochemical activities of the parA partition protein of the P1 plasmid. *Mol Microbiol* 1992;**6**:1141–7. https://doi.org/10.1111/j.1365-2958.1992.tb01552.x

Received: January 23, 2024. **Revised:** April 14, 2025. **Editorial Decision:** April 15, 2025. **Accepted:** April 17, 2025