D452–D459 *Nucleic Acids Research, 2023, Vol. 51, Database issue* *Published online 16 October 2022*
https://doi.org/10.1093/nar/gkac879

PAT: a comprehensive database of prokaryotic antimicrobial toxins

Ya Liu[1], Sheng Liu[1], Zhuo Pan[1], Yu Ren[1], Yiru Jiang[1], Feng Wang[1], Dan-dan Li[1], Yue-zhong Li[1,*] and Zheng Zhang[1,2,*]

[1]State Key Laboratory of Microbial Technology, Institute of Microbial Technology, Shandong University, Qingdao 266237, China and [2]Suzhou Research Institute, Shandong University, Suzhou 215123, China

Received August 12, 2022; Revised September 22, 2022; Editorial Decision September 27, 2022; Accepted September 30, 2022

ABSTRACT

Antimicrobial toxins help prokaryotes win competitive advantages in intraspecific or interspecific conflicts and are also a critical factor affecting the pathogenicity of many pathogens that threaten human health. Although many studies have revealed that antagonism based on antimicrobial toxins plays a central role in prokaryotic life, a database on antimicrobial toxins remains lacking. Here, we present the prokaryotic antimicrobial toxin database (PAT, http://bioinfo.qd.sdu.edu.cn/PAT/), a comprehensive data resource collection on experimentally validated antimicrobial toxins. PAT has organized information, derived from the reported literature, on antimicrobial toxins, as well as the corresponding immunity proteins, delivery mechanisms, toxin activities, structural characteristics, sequences, etc. Moreover, we also predict potential antimicrobial toxins in prokaryotic reference genomes and show the taxonomic information and environmental distribution of typical antimicrobial toxins. These details have been fully incorporated into the PAT database, where users can browse, search, download, analyse and view informative statistics and detailed information. PAT resources have already been used in our prediction and identification of prokaryotic antimicrobial toxins and may contribute to promoting the efficient investigation of antimicrobial toxin functions, the discovery of novel antimicrobial toxins, and an improved understanding of the biological roles and significance of these toxins.

INTRODUCTION

Prokaryotic communities are ubiquitous on Earth. To compete for growth niches and other limited environmental resources, prokaryotic cells employ various means to inhibit the growth of their competitors in communities (1–3). Antimicrobial toxins are proteins that can be transferred from cell to cell to inhibit the growth of other microbes, and are a key weapon in microbial ecological competition (1,4). Many prokaryotes are equipped with multiple antimicrobial toxins; some are limited to intraspecies killing, whereas others are able to act across genera, families, and orders (5). Antimicrobial toxin producers also produce immunity proteins to prevent self-intoxication or intoxication by toxins from kin cells (6). Immunity proteins are usually encoded by genes immediately downstream of antimicrobial toxin genes, thus forming antimicrobial toxin-immunity protein systems (7). These systems are extraordinarily diverse in sequence, and the polymorphism underpins an important mechanism of self/nonself discrimination in prokaryotes (8,9).

The repertoire of experimentally verified antimicrobial toxins and antagonistic systems has substantially increased in recent years (4,10). Antimicrobial toxins have been found in major groups of bacteria and many archaea, and their secretion can be broadly classified in two types: contact-dependent secretion and contact-independent secretion (1,11). Contact-independent secretion refers mainly to bacteriocins, which are the first-characterized antimicrobial toxins and are released into the environment from the producing cells (12). Contact-dependent secretion requires a specialized secretion system (13). The 'contact-dependent growth inhibition' (CDI) phenomenon is performed by a subclass of type V secretion systems (T5SSs) mediated by CdiA proteins (14). The type VI secretion system (T6SS) is a complex nanomolecular machine resembling an inverted phage tail with potent bactericidal activities (15,16). Diverse antimicrobial toxins can be fused to structural components such as VgrG or PAAR as extension domains, or loaded onto T6SS components as cargo effectors (17–20). The type VII secretion system (T7SS) is widespread in Mycobacteria and other gram-positive bacteria and exports substrates with a variety of biological roles, including antimicrobial toxins (21,22). Furthermore, some type IV secretion systems (T4SSs) and extracellular contraction injection sys-

*To whom correspondence should be addressed. Tel: +86 532 58631557; Email: zhangzheng@sdu.edu.cn
Correspondence may also be addressed to Yue-zhong Li. Tel: +86 532 58631539; Email: lilab@sdu.edu.cn

tems (eCISs) deliver toxin effectors directly into target bacteria (23,24). The antimicrobial WapA, MafB and outer membrane exchange (OME) toxins have been characterized from *Bacillus subtilis*, *Neisseria meningitidis* and *Myxococcus xanthus*, respectively (25–28). It is now clear that intercellular antimicrobial toxin transfer is a fundamental and ubiquitous characteristic of prokaryotic biology. Notably, the antimicrobial toxin-immunity protein system is different from the toxin–antitoxin (TA) system, which is also widely distributed in bacteria. Antimicrobial toxins are used as competitive weaponry against competitors, whereas toxins encoded in TA systems are self-poisoning agents and are not secreted (29).

A large number of genes in prokaryotic genomes are expected to encode antimicrobial toxins, immunity proteins and related secretion systems, but experimentally verifying antimicrobial toxin activities and the underlying mechanisms of action remains a significant challenge (6,11). Most of the known antimicrobial toxins have enzymatic activities and can disrupt critical molecular structures of target cells in very small amounts (4). These antimicrobial toxins can be nucleases that target DNA or RNA, phospholipases or pore-forming toxins that target cell membranes, glycoside hydrolases or proteases that degrade cell walls, NADases that disrupt cellular energy balance, ADP-ribosyltransferases that target tubulin-like proteins to prevent cell division, and so on (4,30). Many antimicrobial toxin proteins have been found to have multidomain architectures: the N-terminal domain is involved in secretion, and the C-terminal domain is responsible for bactericidal activity (6). These polymorphic toxins can generate enormous diversity through recombination, even between closely related strains (31).

Although many studies have revealed that the ability to compete for survival based on antimicrobial toxins is a core function of prokaryotes, there is still no comprehensive data resource focusing on antimicrobial toxins. Here, we constructed the PAT database (Prokaryotic Antimicrobial Toxin database, http://bioinfo.qd.sdu.edu.cn/PAT/), which covers bacteriocins (molecular weight > 10 kDa), the toxic effectors of T4SS, CDI, T6SS, T7SS, eCIS and OME systems, with experimental evidence of antimicrobial activity. PAT collates the literature, delivery mechanisms and corresponding immunity proteins of these antimicrobial toxins and provides their sequences and structural information (Figure 1). At the same time, we also predicted potential antimicrobial toxins in the prokaryotic reference genomes and showed the taxonomic sources and environmental distributions of typical antimicrobial toxins. These details have been fully incorporated into PAT, where users can browse, search, download and analyse informative statistics and detailed information to facilitate their investigations into antimicrobial toxins and immunity proteins.

DATA COLLECTION AND DATABASE CONSTRUCTION

Antimicrobial toxins in prokaryotes are extremely diverse, involving multiple delivery mechanisms, the scope of which has only recently been defined (1,4,5). A systematic review of relevant literature is particularly difficult because there is no uniform nomenclature for antimicrobial toxin proteins in the literature. We sorted the literature manually according to different delivery mechanisms and screened out the proteins with experimental evidence of antimicrobial activities. In terms of bacteriocins, we only collected antimicrobial proteins with a molecular weight greater than 10 kDa, including colicins and colicin-like bacteriocins produced by gram-negative bacteria, and class III bacteriocins produced by gram-positive bacteria, excluding antimicrobial peptides and multiprotein complexes (1,32). For the antimicrobial toxins secreted by the T4SS, T6SS and T7SS, we first extracted the effectors from available web resources, and then manually inspected each individual experimental evidence of antimicrobial toxicity, adding the toxic effectors reported recently, as well as those previously overlooked (33–35). The antimicrobial toxins secreted by the CDI, eCIS, OME and T9SS, as well as those secreted by the Sec-dependent system, such as MafB and WapA, were derived entirely from the original literature that reported the experimental evidence. While mining the experimental evidence of antimicrobial toxins, we also collected the corresponding immunity proteins with experimental evidence (if available). PAT is a non-redundant database, and toxins from different studies but with identical sequences are integrated into a unique entry with a commentary of their resources. The PAT database currently contains 441 experimentally validated antimicrobial toxin proteins from 341 research papers.

Information on antimicrobial toxins, corresponding immunity proteins, delivery mechanisms, and toxin activities was all retrieved from the literature and incorporated into the PAT database; their annotations and the PubMed ID from the literature reporting the relevant experimental evidence can be found in the corresponding dedicated detailed information interface (Figure 1). We also comprehensively annotated the features of each protein based on the NCBI (36) and UniProt (37) databases, including gene name, brief description (e.g. protein name), organism, taxonomic lineage, amino acid sequence, and sequence length, as well as the UniProt ID and NCBI ID. Nucleic acid sequences were obtained from the NCBI RefSeq genome database (38). Experimentally determined tertiary structure information was obtained from the PDB (39), and predicted structure information was obtained from the AlphaFold database (40,41). The CDD (42), Pfam (43), SMART (44), COG (45), PRK (46) and TIGRFAM (47) databases were used to collectively annotate and visualize the domain architecture. Signal peptides and corresponding cleavage sites were predicted using SignalP 6.0 software (48). Transmembrane helices were predicted using DeepTMHMM software (49).

Based on the domain architecture of experimentally validated antimicrobial toxin proteins, 63 canonical families of toxicity domains and 38 families of secretion-related markers (such as trafficking domains, repeat domains, protoxins, and conserved motifs) were collected (Supplementary Tables S1 and S2) (6,11). All prokaryotic reference genomes were scanned by RPS-BLAST, which revealed 6064 genes encoding both the antimicrobial toxin domain and the secretion-related marker domain (expected value threshold 0.01) (42). Based on these predicted antimicrobial toxin genes, the taxonomic sources of typical antimi-

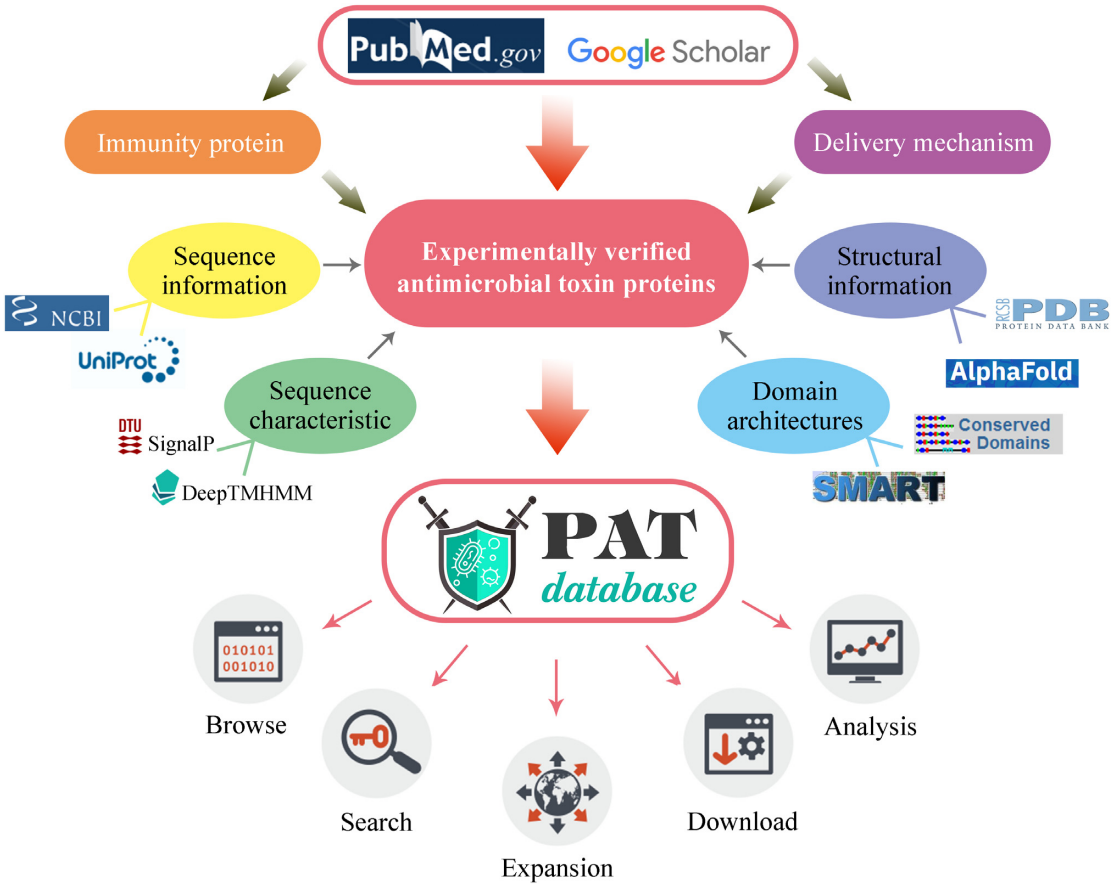

Figure 1. Workflow of PAT database construction.

crobial toxins were counted and visualized. In addition, we investigated the distribution of typical antimicrobial toxins among the global prokaryotic communities through a large-scale alignment between the sequences of 10 000 samples released by the Earth Microbiome Project (EMP) and all RefSeq genomes (38,50). The 16S tags from the EMP data were mapped to those extracted from the genomes with a 97% identity threshold and 95% alignment coverage (51). If multiple genomes were found, those with the highest alignment identity and length were selected. The occurrence frequency of each typical antimicrobial toxin family in EMP samples was counted and visualized at the EMPO_3 level (50).

The PAT database was built based on the Hypertext Preprocessor (PHP) and Bootstrap 4; the latter runs on the Apache web server. All data in the database were stored and managed by the MySQL relational database. The server-backend development was based on the PHP language, and the web-frontend interfaces were implemented in Hyper Text Markup Language (HTML), Cascading Style Sheets (CSS) and JavaScript (JS). PHP language was used for frontend and backend interactions to implement functions such as search and submission, to improve response time, and to speed up browsing. The PAT database is available online without registration, and the Chrome browser is recommended.

DATABASE CONTENT AND USAGE

Experimentally validated antimicrobial toxins

Information on 441 experimentally validated antimicrobial toxin proteins from 70 prokaryotic genera was collected from the literature, of which more than 40% were reported in the past 5 years. Approximately two-thirds of these antimicrobial toxins have corresponding immunity proteins encoded by the genes in the genetic neighbourhood of the toxin encoding genes, and approximately one-sixth have experimentally resolved three-dimensional structures. Currently, the experimentally validated antimicrobial toxins were almost exclusively derived from bacteria, except Halocin, which is found in a few archaea (Supplementary Figure S1). The bacteriocin proteins in the PAT database were experimentally identified in at least six prokaryotic phyla, but 75% are from *Proteobacteria*, mainly *Enterobacterales* and *Pseudomonadales*. The T4SS antimicrobial toxins are presently found only in *Xanthomonadales*. All the CDI antimicrobial toxins are from *Proteobacteria*. The T6SS antimicrobial toxins are mainly from gram-negative *Proteobacteria* and *Bacteroidetes*, while the T7SSs are from gram-positive *Firmicutes* and *Actinobacteria*. Although the number of known eCIS antimicrobial toxins is small, they come from 11 orders in 5 phyla. The OME antimicrobial toxins are only from *Myxococcales*.

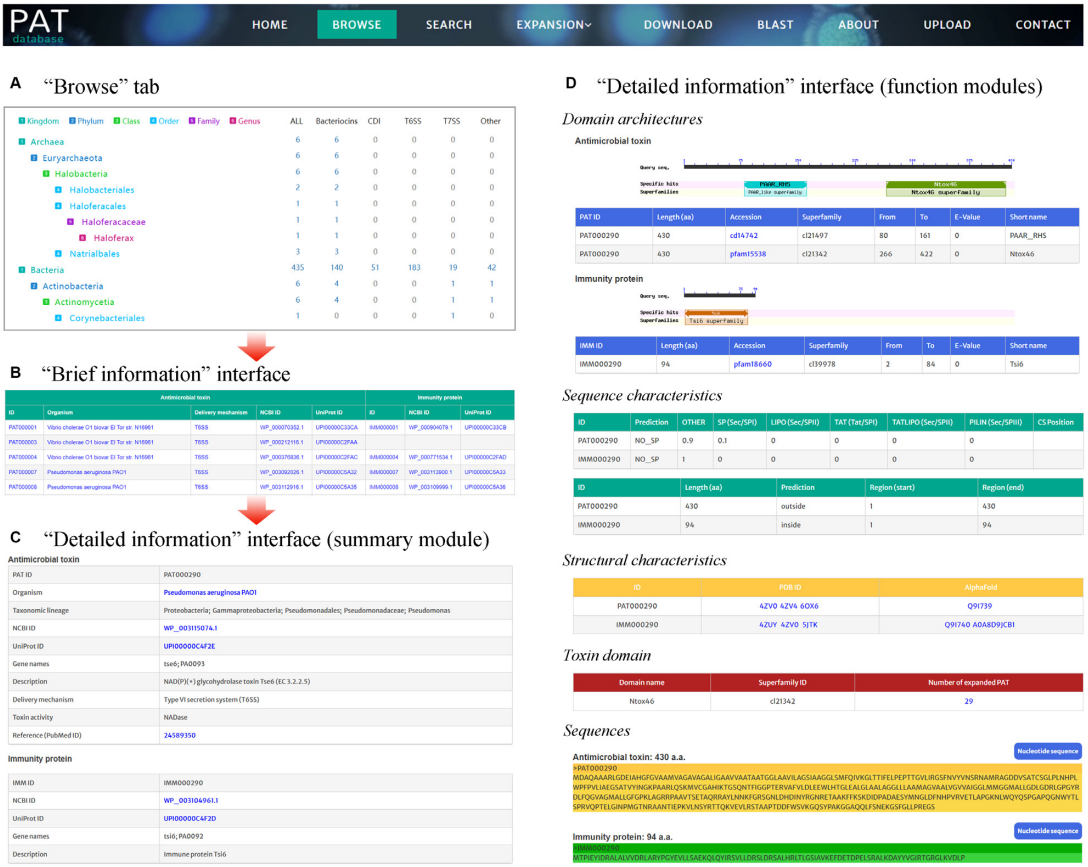

Figure 2. Browse in the database. (**A**) 'Browse' tab. (**B**) 'Brief information' interface. (**C**) 'Detailed information' interface (summary module). (**D**) 'Detailed information' interface (function modules).

Browse in the database

In the 'Browse' tab, antimicrobial toxins are organized by their taxonomic sources and secretion mechanisms to allow rapid identification of proteins of interest (Figure 2A). The number of antimicrobial toxins in each type is also displayed, and upon clicking on the number, users are redirected to the corresponding 'Brief information' interface, where each antimicrobial toxin is shown as an independent entry (Figure 2B). Each entry contains the following information: the source organism, delivery mechanism, NCBI ID and UniProt ID of the antimicrobial toxin and information on the corresponding immunity protein (if available). Users can adjust the number of entries displayed on each interface, and the entries are arranged with an alternately grey or white background. Mouseover provides the corresponding explanation for each table heading. Upon clicking each unique PAT ID, users are redirected to that protein's 'Detailed information' interface to access comprehensive annotations and analyses. Alternatively, users can click on the Organism or NCBI Protein ID to be redirected to external websites.

The 'Detailed information' interface comprises six modules, providing detailed annotations for each experimentally validated antimicrobial toxin (Figure 2C and D). The summary module consists of the source organism (which is clickable to navigate to the corresponding entry in NCBI Taxonomy), taxonomic lineage, NCBI ID (which is clickable to navigate to its entry in the NCBI database), UniProt ID (which is clickable to navigate to its entry in UniProt), gene names, description, delivery mechanism, toxin activity, references (PubMed ID, which is clickable to navigate to the relevant publication associated with the entry), and the NCBI ID, UniProt ID, gene names and description of the corresponding immunity protein. The domain architecture module provides a visualization of the domain composition of antimicrobial toxins and immunity proteins with detailed information on each domain displayed and navigable to the corresponding entry in the CDD database. The sequence characteristic module displays signal peptide and transmembrane region predictions for antimicrobial toxins and immunity proteins. The structural characteristic module provides experimentally determined and AlphaFold-predicted tertiary structure information for antimicrobial toxins and immunity proteins and allows navigation to corresponding entries in the PDB and AlphaFold database. The toxin domain module displays typical families of toxic domains contained in the entry and is navigable to the related entry in the 'Expansion' tab. The protein sequence module shows the amino acid sequences and sequence lengths of antimicrobial toxins and immunity proteins, and the corresponding nucleic acid sequences can be packaged and downloaded.

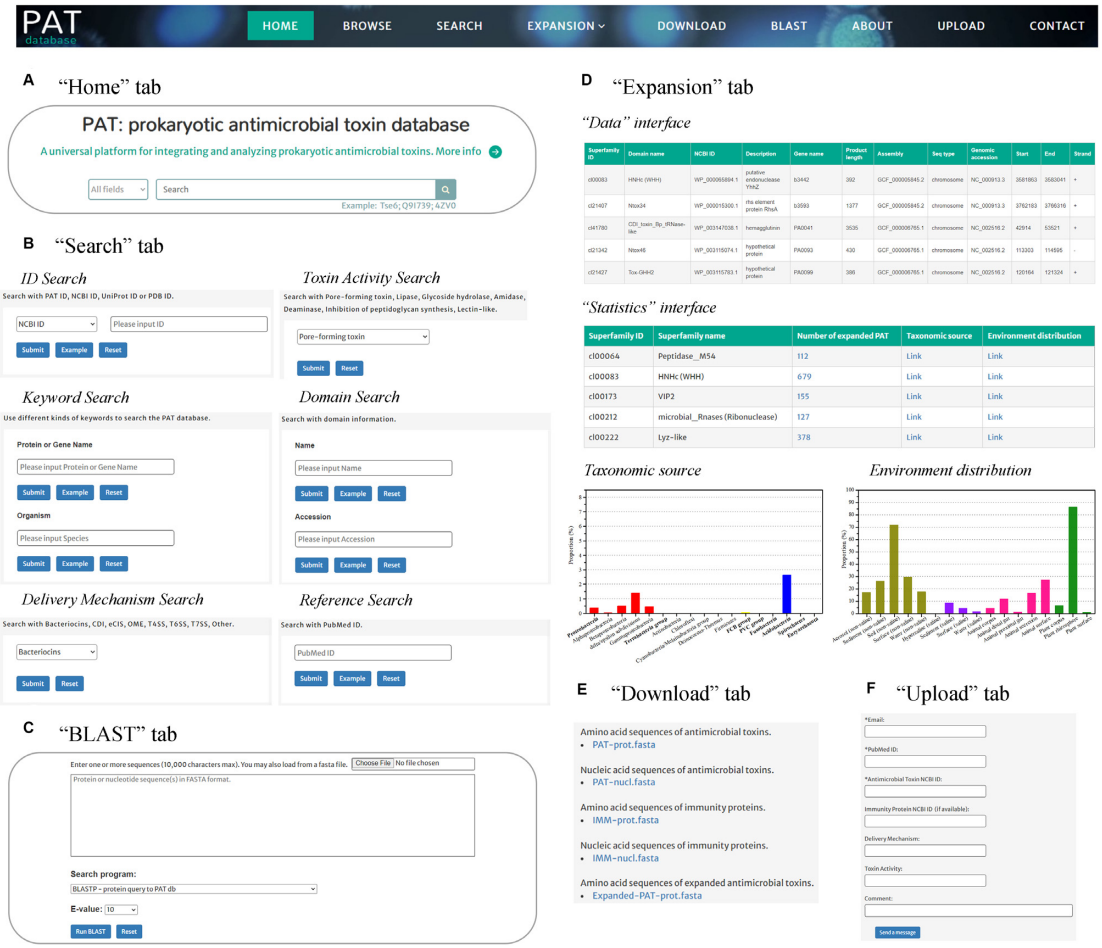

Figure 3. Search and analysis. (**A**) 'Home' tab. (**B**) 'Search' tab. (**C**) 'BLAST' tab. (**D**) 'Expansion' tab. (**E**) 'Download' tab. (**F**) 'Upload' tab.

Search the database

The 'Home' tab allows a quick search of the PAT database (Figure 3A). The 'Search' tab offers a powerful search engine including six searching methods to help users accurately find the entries of interest in the database (Figure 3B). ID search allows exact queries based on the PAT ID, NCBI ID, UniProt ID or PDB ID of antimicrobial toxins or immunity proteins. Keyword search enables fuzzy searches using keywords for protein names, gene names or organisms. Delivery mechanism search and toxin activity search each provide a drop-down filter option to filter for antimicrobial toxin entries. Domain search allows searching for antimicrobial toxins and immunity proteins with corresponding domains by domain name (fuzzy search) or accession (exact search). Reference search uses the PubMed ID to retrieve the antimicrobial toxin entries related to the literature. All the above search results are displayed on the 'Brief Information' interface and further clicked to enter the 'Detailed Information' interface of each antimicrobial toxin entry. In addition, the PAT database also integrates the BLAST program, allowing users to submit nucleic acid or amino acid sequences for sequence alignment of all antimicrobial toxin and immunity protein sequences in the database (Figure 3C).

Expansion

Based on experimentally validated antimicrobial toxin characteristics, potential antimicrobial toxins in prokaryotic reference genomes were predicted. Putative antimicrobial toxins are exhibited in the 'Expansion' tab to distinguish them from experimentally validated antimicrobial toxins (Figure 3D). The 'Methods' interface provides a pipeline for predicting potential antimicrobial toxins in prokaryotic genomes based on combinations of experimentally validated antimicrobial toxin domains and secretion-related marker domains. The 'Statistics' interface displays information on 63 typical antimicrobial toxin families, including predicted antimicrobial toxin numbers (clickable to navigate to data information), visual taxonomic sources and environmental distributions. The 'Data' interface provides detailed information on all predicted antimicrobial toxins.

Data download and upload

Batch download of PAT database entries can be achieved in the 'Download' tab, including the amino acid sequences and nucleic acid sequences of experimentally validated antimicrobial toxins and immunity proteins, as well as the amino

acid sequences of predicted expanded antimicrobial toxins (Figure 3E).

The capture of functional information is extremely time consuming. To improve functional data capture, we have implemented a simple online form in the 'Upload' tab that enables anyone to report one or several functional characterization(s), using a minimal number of fields, namely, a sequence database accession (NCBI ID), PubMed ID (only published peer-reviewed data are taken into account), a textual description of the activity (if available), immunity protein information (if available) and an email address for occasional clarification (Figure 3F).

DISCUSSION

The range of antimicrobial toxin proteins has recently been defined as toxins that are released extracellularly to antagonize other microorganisms, which are functionally distinct from TA system toxins (1,4,5). TA system toxins are used for cellular self-poisoning and play important roles in bacterial physiology, stress response and antimicrobial persistence (29). There are several comprehensively available online resources on TA systems (52–54). Although numerous studies have revealed that antimicrobial toxins have very diverse mechanisms of action, modes of secretion, sequence and structural characteristics, an integrative online resource that summarizes these data is still lacking (4,55). The PAT database provides a unique, readily explorable archive of literature-reported and putative antimicrobial toxins in prokaryotic genomes.

Secretion of proteinaceous antimicrobial toxins is a strategy widely used by prokaryotic organisms to restrict the growth of competitors (2,5). However, the impact of antimicrobial toxins on human health should not be overlooked (56). *In vivo* analyses in host models have shown that some antimicrobial toxins play a role in microbiota-mediated colonization resistance by preventing invasion by pathogens (57). Some pathogens, however, also use toxins to battle with the resident microbiota to invade an ecosystem and cause disease (1). The diversification of antimicrobial toxins is a potential to provide an evolutionary reservoir of toxins, including those used by bacterial pathogens against mammalian hosts and toxins horizontally acquired by eukaryotes to defend against bacteria (5). Hence, we expect that knowledge of the repertoire of antimicrobial toxins involved in bacterial competition will ultimately allow us to better predict the outcome of microbial interactions and will help in the strategic design of methods to manipulate the microbiota for medical purposes.

PAT resources have already been used in the prediction and identification of prokaryotic antimicrobial toxins (8,19,20,58). In the future, the PAT database will continue to integrate the latest knowledge of antimicrobial toxins to satisfy the needs of users worldwide, helping to promote efficient investigation on functions of antimicrobial toxins, the discovery of novel antimicrobial toxins, and an improved understanding of their biological roles and significance.

DATA AVAILABILITY

The PAT database is freely available to the public without registration or login requirements (http://bioinfo.qd.sdu.edu.cn/PAT/).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

FUNDING

National Natural Science Foundation of China [32070030, 32270073]; Special Investigation on Scientific and Technological Basic Resources [2017FY100300]; National Key Research and Development Program [2018YFA0900400, 2018YFA0901704]; Key Research and Development Program of Shandong Province [2018GSF121015]; Natural Science Foundation of Jiangsu Province [BK20190199]; Fundamental Research Funds of Shandong University [2020GN113]; China Postdoctoral Science Foundation [2021M700084, 2021M701987]. Funding for open access charge: National Natural Science Foundation of China [32270073].

Conflict of interest statement. None declared.

REFERENCES

1. Garcia-Bayona,L. and Comstock,L.E. (2018) Bacterial antagonism in host-associated microbial communities. *Science*, **361**, eaat2456.
2. Granato,E.T., Meiller-Legrand,T.A. and Foster,K.R. (2019) The evolution and ecology of bacterial warfare. *Curr. Biol.*, **29**, R521–R537.
3. Raffatellu,M. (2018) Learning from bacterial competition in the host to develop antimicrobials. *Nat. Med.*, **24**, 1097–1103.
4. Ruhe,Z.C., Low,D.A. and Hayes,C.S. (2020) Polymorphic toxins and their immunity proteins: diversity, evolution, and mechanisms of delivery. *Annu. Rev. Microbiol.*, **74**, 497–520.
5. Peterson,S.B., Bertolli,S.K. and Mougous,J.D. (2020) The central role of interbacterial antagonism in bacterial life. *Curr. Biol.*, **30**, R1203–R1214.
6. Zhang,D., de Souza,R.F., Anantharaman,V., Iyer,L.M. and Aravind,L. (2012) Polymorphic toxin systems: comprehensive characterization of trafficking modes, processing, mechanisms of action, immunity and ecology using comparative genomics. *Biol. Direct*, **7**, 18.
7. Benz,J. and Meinhart,A. (2014) Antibacterial effector/immunity systems: it's just the tip of the iceberg. *Curr. Opin. Microbiol.*, **17**, 1–10.
8. Gong,Y., Zhang,Z., Liu,Y., Zhou,X.W., Anwar,M.N., Li,Z.S., Hu,W. and Li,Y.Z. (2018) A nuclease-toxin and immunity system for kin discrimination in *myxococcusxanthus*. *Environ. Microbiol.*, **20**, 2552–2567.
9. Vassallo,C.N. and Wall,D. (2019) Self-identity barcodes encoded by six expansive polymorphic toxin families discriminate kin in myxobacteria. *Proc. Natl. Acad. Sci. U.S.A.*, **116**, 24808–24818.
10. Klein,T.A., Ahmad,S. and Whitney,J.C. (2020) Contact-Dependent interbacterial antagonism mediated by protein secretion machines. *Trends Microbiol.*, **28**, 387–400.
11. Makarova,K.S., Wolf,Y.I., Karamycheva,S., Zhang,D., Aravind,L. and Koonin,E.V. (2019) Antimicrobial peptides, polymorphic toxins, and self-nonself recognition systems in archaea: an untapped armory for intermicrobial conflicts. *Mbio*, **10**, e00715-19.
12. Cascales,E., Buchanan,S.K., Duche,D., Kleanthous,C., Lloubes,R., Postle,K., Riley,M., Slatin,S. and Cavard,D. (2007) Colicin biology. *Microbiol. Mol. Biol. Rev.*, **71**, 158–229.
13. Galan,J.E. and Waksman,G. (2018) Protein-Injection machines in bacteria. *Cell*, **172**, 1306–1318.

14. Aoki,S.K., Pamma,R., Hernday,A.D., Bickham,J.E., Braaten,B.A. and Low,D.A. (2005) Contact-dependent inhibition of growth in *escherichiacoli*. *Science*, **309**, 1245–1248.
15. Hood,R.D., Singh,P., Hsu,F., Guvener,T., Carl,M.A., Trinidad,R.R., Silverman,J.M., Ohlson,B.B., Hicks,K.G., Plemel,R.L. *et al.* (2010) A type VI secretion system of *pseudomonasaeruginosa* targets a toxin to bacteria. *Cell Host Microbe*, **7**, 25–37.
16. MacIntyre,D.L., Miyata,S.T., Kitaoka,M. and Pukatzki,S. (2010) The vibrio cholerae type VI secretion system displays antimicrobial properties. *Proc. Natl. Acad. Sci. U.S.A.*, **107**, 19520–19524.
17. Burkinshaw,B.J., Liang,X., Wong,M., Le,A.N.H., Lam,L. and Dong,T.G. (2018) A type VI secretion system effector delivery mechanism dependent on PAAR and a chaperone-co-chaperone complex. *Nat Microbiol*, **3**, 632–640.
18. Shneider,M.M., Buth,S.A., Ho,B.T., Basler,M., Mekalanos,J.J. and Leiman,P.G. (2013) PAAR-repeat proteins sharpen and diversify the type VI secretion system spike. *Nature*, **500**, 350–353.
19. Liu,Y., Zhang,Z., Wang,F., Li,D.D. and Li,Y.Z. (2020) Identification of type VI secretion system toxic effectors using adaptors as markers. *Comput. Struct. Biotechnol. J.*, **18**, 3723–3733.
20. Zhang,Z., Liu,Y., Zhang,P., Wang,J., Li,D. and Li,Y.Z. (2021) PAAR proteins are versatile clips that enrich the antimicrobial weapon arsenals of prokaryotes. *mSystems*, **6**, e00953-21.
21. Cao,Z., Casabona,M.G., Kneuper,H., Chalmers,J.D. and Palmer,T. (2016) The type VII secretion system of *staphylococcusaureus* secretes a nuclease toxin that targets competitor bacteria. *Nat. Microbiol.*, **2**, 16183.
22. Whitney,J.C., Peterson,S.B., Kim,J., Pazos,M., Verster,A.J., Radey,M.C., Kulasekara,H.D., Ching,M.Q., Bullen,N.P., Bryant,D. *et al.* (2017) A broadly distributed toxin family mediates contact-dependent antagonism between gram-positive bacteria. *Elife*, **6**, e26938.
23. Souza,D.P., Oka,G.U., Alvarez-Martinez,C.E., Bisson-Filho,A.W., Dunger,G., Hobeika,L., Cavalcante,N.S., Alegria,M.C., Barbosa,L.R., Salinas,R.K. *et al.* (2015) Bacterial killing via a type IV secretion system. *Nat. Commun.*, **6**, 6453.
24. Jiang,F., Li,N., Wang,X., Cheng,J., Huang,Y., Yang,Y., Yang,J., Cai,B., Wang,Y.P., Jin,Q. *et al.* (2019) Cryo-EM structure and assembly of an extracellular contractile injection system. *Cell*, **177**, 370–383.
25. Arenas,J., de Maat,V., Caton,L., Krekorian,M., Herrero,J.C., Ferrara,F. and Tommassen,J. (2015) Fratricide activity of MafB protein of *n. meningitidis* strain B16B6. *BMC Microbiol.*, **15**, 156.
26. Jamet,A., Jousset,A.B., Euphrasie,D., Mukorako,P., Boucharlat,A., Ducousso,A., Charbit,A. and Nassif,X. (2015) A new family of secreted toxins in pathogenic *neisseria* species. *PLoS Pathog.*, **11**, e1004592.
27. Koskiniemi,S., Lamoureux,J.G., Nikolakakis,K.C., t'Kint de Roodenbeke,C., Kaplan,M.D., Low,D.A. and Hayes,C.S. (2013) Rhs proteins from diverse bacteria mediate intercellular competition. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 7032–7037.
28. Vassallo,C.N., Cao,P., Conklin,A., Finkelstein,H., Hayes,C.S. and Wall,D. (2017) Infectious polymorphic toxins delivered by outer membrane exchange discriminate kin in myxobacteria. *Elife*, **6**, e29397.
29. Jurenas,D., Fraikin,N., Goormaghtigh,F. and Van Melderen,L. (2022) Biology and evolution of bacterial toxin-antitoxin systems. *Nat. Rev. Microbiol.*, **20**, 335–350.
30. Russell,A.B., Peterson,S.B. and Mougous,J.D. (2014) Type VI secretion system effectors: poisons with a purpose. *Nat. Rev. Microbiol.*, **12**, 137–148.
31. Jamet,A. and Nassif,X. (2015) New players in the toxin field: polymorphic toxin systems in bacteria. *Mbio*, **6**, e00285-15.
32. van Heel,A.J., de Jong,A., Song,C., Viel,J.H., Kok,J. and Kuipers,O.P. (2018) BAGEL4: a user-friendly web server to thoroughly mine ripps and bacteriocins. *Nucleic Acids Res.*, **46**, W278–W281.
33. Li,J., Yao,Y., Xu,H.H., Hao,L., Deng,Z., Rajakumar,K. and Ou,H.Y. (2015) SecReT6: a web-based resource for type VI secretion systems found in bacteria. *Environ. Microbiol.*, **17**, 2196–2202.
34. Dai,W., Li,J., Li,Q., Cai,J., Su,J., Stubenrauch,C. and Wang,J. (2022) PncsHub: a platform for annotating and analyzing non-classically secreted proteins in Gram-positive bacteria. *Nucleic Acids Res.*, **50**, D848–D857.
35. Bi,D., Liu,L., Tai,C., Deng,Z., Rajakumar,K. and Ou,H.Y. (2013) SecReT4: a web-based bacterial type IV secretion system resource. *Nucleic Acids Res.*, **41**, D660–D665.
36. Sayers,E.W., Bolton,E.E., Brister,J.R., Canese,K., Chan,J., Comeau,D.C., Connor,R., Funk,K., Kelly,C., Kim,S. *et al.* (2022) Database resources of the national center for biotechnology information. *Nucleic Acids Res.*, **50**, D20–D26.
37. UniProt, C. (2021) UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res.*, **49**, D480–D489.
38. Li,W., O'Neill,K.R., Haft,D.H., DiCuccio,M., Chetvernin,V., Badretdin,A., Coulouris,G., Chitsaz,F., Derbyshire,M.K., Durkin,A.S. *et al.* (2021) RefSeq: expanding the prokaryotic genome annotation pipeline reach with protein family model curation. *Nucleic Acids Res.*, **49**, D1020–D1028.
39. Burley,S.K., Bhikadiya,C., Bi,C., Bittrich,S., Chen,L., Crichlow,G.V., Christie,C.H., Dalenberg,K., Di Costanzo,L., Duarte,J.M. *et al.* (2021) RCSB protein data bank: powerful new tools for exploring 3D structures of biological macromolecules for basic and applied research and education in fundamental biology, biomedicine, biotechnology, bioengineering and energy sciences. *Nucleic Acids Res.*, **49**, D437–D451.
40. Jumper,J., Evans,R., Pritzel,A., Green,T., Figurnov,M., Ronneberger,O., Tunyasuvunakool,K., Bates,R., Zidek,A., Potapenko,A. *et al.* (2021) Highly accurate protein structure prediction with alphafold. *Nature*, **596**, 583–589.
41. Varadi,M., Anyango,S., Deshpande,M., Nair,S., Natassia,C., Yordanova,G., Yuan,D., Stroe,O., Wood,G., Laydon,A. *et al.* (2022) AlphaFold protein structure database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.*, **50**, D439–D444.
42. Lu,S., Wang,J., Chitsaz,F., Derbyshire,M.K., Geer,R.C., Gonzales,N.R., Gwadz,M., Hurwitz,D.I., Marchler,G.H., Song,J.S. *et al.* (2020) CDD/SPARCLE: the conserved domain database in 2020. *Nucleic Acids Res.*, **48**, D265–D268.
43. Mistry,J., Chuguransky,S., Williams,L., Qureshi,M., Salazar,G.A., Sonnhammer,E.L.L., Tosatto,S.C.E., Paladin,L., Raj,S., Richardson,L.J. *et al.* (2021) Pfam: the protein families database in 2021. *Nucleic Acids Res.*, **49**, D412–D419.
44. Letunic,I., Khedkar,S. and Bork,P. (2021) SMART: recent updates, new developments and status in 2020. *Nucleic Acids Res.*, **49**, D458–D460.
45. Galperin,M.Y., Wolf,Y.I., Makarova,K.S., Vera Alvarez,R., Landsman,D. and Koonin,E.V. (2021) COG database update: focus on microbial diversity, model organisms, and widespread pathogens. *Nucleic Acids Res.*, **49**, D274–D281.
46. Klimke,W., Agarwala,R., Badretdin,A., Chetvernin,S., Ciufo,S., Fedorov,B., Kiryutin,B., O'Neill,K., Resch,W., Resenchuk,S. *et al.* (2009) The national center for biotechnology information's protein clusters database. *Nucleic Acids Res.*, **37**, D216–D23.
47. Haft,D.H., Selengut,J.D., Richter,R.A., Harkins,D., Basu,M.K. and Beck,E. (2013) TIGRFAMs and genome properties in 2013. *Nucleic Acids Res.*, **41**, D387–D95.
48. Teufel,F., Almagro Armenteros,J.J., Johansen,A.R., Gislason,M.H., Pihl,S.I., Tsirigos,K.D., Winther,O., Brunak,S., von Heijne,G. and Nielsen,H. (2022) SignalP 6.0 predicts all five types of signal peptides using protein language models. *Nat. Biotechnol.*, **40**, 1023–1025.
49. Hallgren,J., Tsirigos,K.D., Pedersen,M.D., Almagro Armenteros,J.J., Marcatili,P., Nielsen,H., Krogh,A. and Winther,O. (2022) DeepTMHMM predicts alpha and beta transmembrane proteins using deep neural networks. bioRxiv doi: https://doi.org/10.1101/2022.04.08.487609, 10 April 2022, preprint: not peer reviewed.
50. Thompson,L.R., Sanders,J.G., McDonald,D., Amir,A., Ladau,J., Locey,K.J., Prill,R.J., Tripathi,A., Gibbons,S.M., Ackermann,G. *et al.* (2017) A communal catalogue reveals earth's multiscale microbial diversity. *Nature*, **551**, 457–463.
51. Zhang,Z., Wang,J., Wang,J., Wang,J. and Li,Y. (2020) Estimate of the sequenced proportion of the global prokaryotic genome. *Microbiome*, **8**, 134.
52. Xie,Y., Wei,Y., Shen,Y., Li,X., Zhou,H., Tai,C., Deng,Z. and Ou,H.Y. (2018) TADB 2.0: an updated database of bacterial type II toxin-antitoxin loci. *Nucleic Acids Res.*, **46**, D749–D753.

53. Akarsu,H., Bordes,P., Mansour,M., Bigot,D.J., Genevaux,P. and Falquet,L. (2019) TASmania: a bacterial toxin-antitoxin systems database. *PLoS Comput. Biol.*, **15**, e1006946.
54. Tourasse,N.J. and Darfeuille,F. (2021) T1TAdb: the database of type i toxin-antitoxin systems. *RNA*, **27**, 1471–1481.
55. Jamet,A., Charbit,A. and Nassif,X. (2018) Antibacterial toxins: gram-positive bacteria strike back!*Trends Microbiol.*, **26**, 89–91.
56. Ross,B.D., Verster,A.J., Radey,M.C., Schmidtke,D.T., Pope,C.E., Hoffman,L.R., Hajjar,A.M., Peterson,S.B., Borenstein,E. and Mougous,J.D. (2019) Human gut bacteria contain acquired interbacterial defence systems. *Nature*, **575**, 224–228.
57. Chassaing,B. and Cascales,E. (2018) Antibacterial weapons: targeted destruction in the microbiota. *Trends Microbiol.*, **26**, 329–338.
58. Liu,Y., Wang,J., Zhang,Z., Wang,F., Gong,Y., Sheng,D.H. and Li,Y.Z. (2021) Two PAAR proteins with different C-Terminal extended domains have distinct ecological functions in *Myxococcus xanthus*. *Appl. Environ. Microbiol.*, **87**, e00080-21.