doi 10.18699/bgrs2024-12.3-25

InTxSEs: a platform for interactions between bacterial secreted proteins and human proteins

Zhu Y.Y., Liu L.Y., Hu Y.M., Chen M.*

Department of Bioinformatics, College of Life Sciences, Zhejiang University, Hangzhou, China

** mchen@zju.edu.cn*

Key words: bacterial secretion system; docking; PPIs; network; 3D structure

Motivation and Aim: Protein secretion plays a central role in modulating the interactions of bacteria with their environments. Bacteria transport numerous substrates across cellular membranes via secretion systems that are essential for virulence and survival [1]. Initial bacterial attachment to host cell surface receptors is commonly followed by induction of signaling events that induce cytoskeletal rearrangements and consequently promote intimate bacteria–host association or bacterial internalization. These events are typically followed by additional manipulations of host cell signaling events such as those involved in attenuation or evasion of the host's innate defenses [2]. In past decades, low-throughput [e. g. co-immunoprecipitation (Co-IP)] and high-throughput [e. g. mass spectrometry (MS) and yeast two-hybrid (Y2H)] techniques [3–5] allowed the determination of human–bacterial PPIs on an unprecedented scale, providing an abundance of PPI data that have been stored in a series of state-of-the-art host-pathogen protein interaction databases, such as IntAct [6]. Currently, there is a lack of specialized databases for the interaction between secreted proteins in bacterial secretion systems and human proteins. Existing databases are incomplete, lacking annotations, including the display of protein 3D structures, as well as the presentation of their interaction interfaces and functional sites. Here we introduce a comprehensive platform for the interaction between bacterial secreted proteins and human proteins. This platform contains data resources for the interaction between secretion system effectors (TxSEs) from ten bacterial secretion systems and human proteins. It provides experimentally validated Protein-Protein Interactions (PPIs), 3D complex structures of protein interactions, as well as displays of the interaction interfaces and functional sites of PPIs. Additionally, we will continuously update predicted PPIs in the future.

Methods and Algorithms:

Collection of experimentally verified PPI data

We collected experimentally verified PPIs from five public databases (i. e. HPIDB [7], PHISTO [8], IntAct [6]) and recently published literature. We obtained a total of 1,853 PPIs, In particular, we mapped protein IDs from different databases to UniProt IDs or NCBI IDs, gene names or protein names as query options in InTxSEs.

Three-dimensional complex structures of PPIs

Retrieve the three-dimensional structures of bacterial secreted proteins and human proteins separately from the PDB database [9] and the AlphaFold Protein Structure Database. For each pair of interacting proteins, InTxSEs will employ ZDOCK 3.0.2 [10], HDOCK [11], and AlphaFold3 [12] to calculate the potential complex structures with the highest docking scores. We used the obtained pdb to analyze the interaction sites using PDBePISA.

Collection of other information on proteins secreted by bacteria

We used TMHMM 2.0 software [13] to predict transmembrane helix sites for bacterial secreted proteins, and SignalP 6.0 [14] to predict signal peptide sites for bacterial secreted proteins.

Construct PPI network

We obtained the secretion protein network for each bacterial species and the corresponding protein network for human proteins from STRING [15]. A JavaScript library Cytoscape [16] is employed to display 2D protein interaction networks for each bacterial species. A JavaScript libraries echarts shows each secreted protein and human protein interaction network.

Results: As a major component, the InTxSEs online resource includes a comprehensive data module and a blast tool. Currently, InTxSEs offers 1,853 experimentally validated PPIs, 193 bacterial-secreted proteins, 102 bacterial species, and the main architecture of HVIDB contains the following key features: (1) For network information, InTxSEs provides a related human-bacterial ppi subnet for each queried PPI and a related human-bacterial PPI network for each bacterial species, which can improve the understanding of the mechanism of viral infection. (2) In terms of structural information, InTxSEs provides 3D visualization of the docking structure of PPI and the corresponding interaction sites, and the interaction sites can be queried. (3) It complements the transmembrane information of bacterial secreted proteins, signal peptide site information and subcellular localization information, which can deepen the understanding of the mechanism of secreted proteins acting on human proteins. (4) InTxSEs has a blast module to analyze homologous proteins.

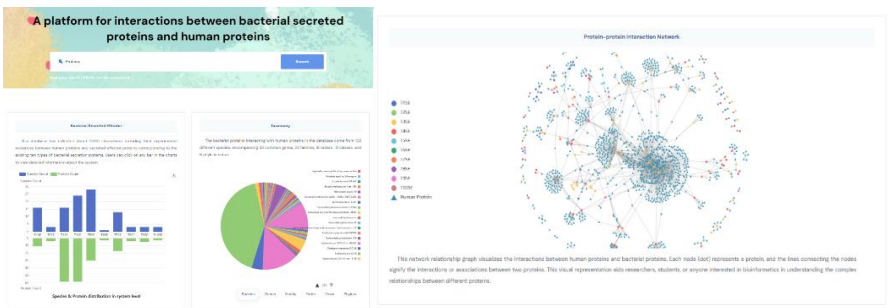

Fig. 1. Main search and browse interfaces of InTxSEs

Conclusion: InTxSEs is a freely accessible resource that provides comprehensive, downloadable PPI data on bacteria-secreted proteins and human proteins for maximum application. In contrast to existing human-bacterial PPI databases, InTxSEs specifically integrates multiple PPI data resources on bacteria-secreted proteins, enabling users to explore the corresponding biological applications of corresponding interaction groups. In terms of future development, we will regularly upgrade InTxSEs by integrating more data resources, developing predictive tools to add to the platform, and designing a simpler and more user-friendly interactive analytics interface. For example, we will add drug targets that act on secreted protein structures to InTxSEs, potentially accelerating drug target discovery and drug retargeting to fight deadly diseases. All in all, InTxSEs can serve as a one-stop knowledge base to understand humans' relationship with

pathogenic bacteria and develop treatments to address the ongoing challenges of bacterial infections.

Funding: The study is supported by the National Natural Sciences Foundation of China (No. 31771477, 32070677).

References

1. Tseng T.T., Tyler B.M., Setubal J.C. Protein secretion systems in bacterial-host associations, and their description in the Gene Ontology. *BMC Microbiol*. 2009;9(Suppl. 1):S2. doi 10.1186/1471-2180-9-S1-S2
2. Mattoo S., Lee Y.M., Dixon J.E. Interactions of bacterial effector proteins with host proteins. *Curr Opin Immunol*. 2007;19(4):392-401. doi 10.1016/j.coi.2007.06.005
3. Walch P., Selkrig J., Knodler L.A. et al. Global mapping of Salmonella enterica-host protein-protein interactions during infection. *Cell Host Microbe*. 2021;29(8):1316-1332.e12. doi 10.1016/j.chom.2021.06.004
4. Paiano A., Margiotta A., De Luca M., Bucci C. Yeast Two-Hybrid Assay to Identify Interacting Proteins. *Curr Protoc Protein Sci*. 2019;95(1):e70. doi 10.1002/cpps.70.
5. Turriziani B., von Kriegsheim A., Pennington S.R. Protein-Protein Interaction Detection Via Mass Spectrometry-Based Proteomics. *Adv Exp Med Biol*. 2016;919:383-396. doi 10.1007/978-3-319-41448-5_18
6. Del Toro N., Shrivastava A., Ragueneau E. et al. The IntAct database: efficient access to fine-grained molecular interaction data. *Nucleic Acids Res*. 2022;50(D1):D648-D653. doi 10.1093/nar/gkab1006
7. Ammari M.G., Gresham C.R., McCarthy F.M., Nanduri B. HPIDB 2.0: a curated database for host-pathogen interactions. *Database (Oxford)*. 2016;2016:baw103
8. Durmuş Tekir S., Çakır T., Ardiç E. et al. PHISTO: pathogen-host interaction search tool. *Bioinformatics*. 2013;29(10):1357-1358. doi 10.1093/bioinformatics/btt137
9. Rose P.W., Prlić A., Altunkaya A. et al. The RCSB protein data bank: integrative view of protein, gene and 3D structural information. *Nucleic Acids Res*. 2017;45(D1):D271-D281. doi 10.1093/nar/gkw1000
10. Pierce B.G., Wiehe K., Hwang H., Kim B.H., Vreven T., Weng Z. ZDOCK server: interactive docking prediction of protein-protein complexes and symmetric multimers. *Bioinformatics*. 2014;30(12):1771-3. doi 10.1093/bioinformatics/btu097
11. Yan Y., Zhang D., Zhou P., Li B., Huang S.Y. HDOCK: a web server for protein-protein and protein-DNA/RNA docking based on a hybrid strategy. *Nucleic Acids Res*. 2017;45(W1):W365-W373. doi 10.1093/nar/gkx407
12. Abramson J., Adler J., Dunger J. et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*. 2024. doi 10.1038/s41586-024-07487-w
13. Krogh A., Larsson B., von Heijne G., Sonnhammer E.L. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol.* 2001;305(3):567-580
14. Teufel F., Almagro Armenteros J.J., Johansen A.R. et al. SignalP 6.0 predicts all five types of signal peptides using protein language models. *Nat Biotechnol*. 2022;40(7):1023-1025. doi 10.1038/s41587-021-01156-3
15. Szklarczyk D., Kirsch R., Koutrouli M. et al. The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res*. 2023;51(D1):D638-D646. doi 10.1093/nar/gkac1000
16. Franz M., Lopes C.T., Huck G. et al. Cytoscape.js: a graph theory library for visualisation and analysis. *Bioinformatics*. 2016;32:309-311