

TRustDB: A comprehensive bioinformatics resource for understanding the complete Wheat–Stem rust host–pathogen interactome

Raghav Kataria ¹ and Rakesh Kaundal ^{1,2,3,*}

¹Department of Plants, Soils, and Climate, College of Agriculture and Applied Sciences, Utah State University, Logan, UT 84322, USA

²Bioinformatics Facility, Center for Integrated BioSystems, College of Agriculture and Applied Sciences, Utah State University, Logan, UT 84322, USA

³Department of Computer Science, College of Science, Utah State University, Logan, UT 84322, USA

*Corresponding author: Tel: +1 (435) 797-4117; Fax: +1 (435) 797-2766; Email: rkaundal@usu.edu

Citation details: Kataria, R. and Kaundal, R. TRustDB: A comprehensive bioinformatics resource for understanding the complete Wheat–Stem rust host–pathogen interactome. *Database* (2022) Vol. 2022: article ID baac068; DOI: <https://doi.org/10.1093/database/baac068>

Abstract

The increasing infectious diseases in wheat immensely reduce crop yield and quality, thus affecting global wheat production. The evolution in phytopathogens hinders the understanding of the disease infection mechanisms. TRustDB is an open-access, comprehensive database that is specifically focused on the disease stem rust (also known as black rust) in *Triticum aestivum*, which is caused by the fungal pathogen *Puccinia graminis* (*Pgt*), strains ‘Ug99’ and ‘21-0’. The database aims at a broader focus of providing the researchers with comprehensive tools to predict the protein–protein interactions and avail the functional annotations of the proteins involved in the interactions that cause the disease. The network of the predicted interactome can also be visualized on the browser. Various modules for the functional annotations of the host and pathogen proteins such as subcellular localization, functional domains, gene ontology annotations, pathogen orthologs and effector proteins have been implemented. The host proteins that serve as transcription factors, along with the respective Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways are also available, which further enhance the understanding of the disease infection mechanisms and the defense responses of the host. The database is also linked with several other databases such as InterPro, KEGG pathways, Ensembl and National Center for Biotechnology Information (NCBI). TRustDB has a user-friendly web interface, which can be accessed through <http://bioinfo.usu.edu/trustdb/>.

Database URL: <http://bioinfo.usu.edu/trustdb/>

Introduction

In the past decade, the impact of infectious diseases in plants has been a primary concern due to the increased incidence of highly virulent pathogens in major crops of the world. Over the course of evolution, various new pathogen strains are emerging that cause highly infectious plant diseases, thus causing huge economic and agricultural losses worldwide (1). Wheat (*Triticum aestivum* L.) is considered as one of the most important staple food grains and is a source of essential nutrients such as carbohydrates, vitamins, proteins, minerals, phytochemicals and others. The future global production of wheat needs to be substantially escalated due to the increased consumption and changing dietary habits of people (2, 3), which can be accomplished through sustainable agriculture strategies. Apart from the high demand for the crop, wheat production is also affected by pests and diseases that accounts for about 21.5% of the crop yield losses (4). The impact of the disease varies annually, depending on the incident rate and pathogenesis of the pathogen, resulting in reduced grain quality and quantity (5).

Biotrophic rust fungi are considered a major constraint in the wheat production and have many strains or races that cause peculiar infections in wheat (6). Recently, wheat stem (or black) rust has been of great significance to researchers, owing to its evolved pathogenic mechanism that overcomes the race-specific host resistance genes. The disease, caused by *Puccinia graminis* f. sp. *tritici* Ericks and Henn (*Pgt*), has been detected in many countries around the globe and is associated with high yield losses, reduced grain size and water loss in the crop (7). Over time, *Pgt* evolved into many strains, of which the African strain ‘Ug99’ (race TTKSK) and the Australian strain ‘21-0’ are of major focus (8, 9). Stem rust infection initiates with the germination of urediniospores, followed by the differentiation of mother cells into haustoria that intakes the nutrients from the host cell (10). The effective fungicides to control the disease are available but the unique mechanism of pathogens develops resistance against the chemicals. Also, the usage of chemicals increases the production cost and is harmful to the environment (11, 12).

The pathogens subvert the host cellular mechanisms during the infection process, while the plant cells involve

Received 7 April 2022; Revised 10 July 2022; Accepted 8 November 2022

© The Author(s) 2022. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License

(<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

gene-for-gene resistance that activates various downstream immune signaling events such as salicylic acid pathway, reactive oxygen species (ROS) production, mitogen-activated protein kinase (MAPK) signaling and others, which are interconnected with the transcription factors (TFs; 13, 14). Furthermore, the TFs are also considered as positive regulators during biotic stresses (15). Thus, the protein-protein interactions (PPIs) play an important role in both pathogen and host species to cause the infection and initiate defense responses against the disease, respectively (16). The complex network of PPIs between the host and pathogen proteins can be analyzed in an efficient manner using the computational systems biology approach, which reveals the biological information at multiple levels (17, 18).

Various studies in the past have mapped stem-rust-resistance quantitative trait locus (QTLs)/genes in wheat using genome-wide approaches (19). Numerous host-pathogen PPI data sources are available for different pathosystems but no open resources are available for the retrieval of the biological information of the proteins involved in PPIs for stem rust disease. In line with this, we developed a user-friendly, open-source web server 'TRustDB' that implements interolog-based PPI prediction and provides functional annotations of the host and pathogen proteins involved in the interactions during stem rust infection in wheat. The database provides a comparative study of the proteins of two *Puccinia* species responsible for causing stem rust disease in wheat. The broad goal of our study includes not only the protein annotation of the predicted PPIs but also providing the research

community with a robust framework that enhances their understanding of the biology underlying the disease infection mechanism. TRustDB is freely available for public use at <http://bioinfo.usu.edu/trustdb/>.

Database overview

TRustDB is an extensive resource, composed of the functional annotations of the proteins involved in stem rust disease in wheat. Three diverse tools: 'interactomics' tool to predict the host-pathogen interactions, 'advanced search' module for the keyword-specific search of the proteins and 'BLAST search' are implemented on the database. The tools provide enhanced visualizations for the retrieval of relevant protein information. Furthermore, various functional annotations of the host and pathogen protein in the form of features have been implemented on the database that gives protein-specific information obtained from different sources. The database provides a user-friendly environment for novice users to obtain the required information for the host and pathogen protein annotations. Additionally, the 'User Guide' page implemented on the database helps the user to browse the database efficiently.

Dataset compilation

For the construction of the database, the proteome of *T. aestivum* was obtained from Ensembl Plants (http://ftp.ebi.ac.uk/ensemblgenomes/pub/release-51/plants/fasta/triticum_aestivum/pep/). On the other hand, the proteomes of *Pgt* Ug99 and *Pgt* 21-0 were obtained from Ensembl Fungi

KAABIL
Kaundal Artificial Intelligence & Advanced Bioinformatics Lab

Utah State University

Home Tools Features Datasets User Guide

Puccinia graminis Ug99 Search protein

TRustDB: Host-Pathogen Interaction Database of Stem Rust disease in *Triticum aestivum*

Database Overview

Wheat (*Triticum aestivum*), the most widely cultivated crop in the world, ranks third and counts to around 35% of the total grains produced around the globe. Recently, the occurrence of stem rust in wheat has affected the grain yield and quality of the crop to around 60-70%, thus being a major threat to the agriculture industry, worldwide. The disease is caused by two races of *Puccinia graminis*: 'Ug99' and '21-0'.

TRustDB, a database for *Triticum aestivum* -*Puccinia* species protein-protein interactions, provides valuable information about the *Puccinia* species, which includes various functional protein annotations of the proteins involved in the interactions.

Host species
Triticum aestivum

Pathogen species
Puccinia graminis isolate Ug99 (*Pgt* Ug99)
Puccinia graminis isolate 21-0 (*Pgt* 21-0)

Download the proteomes of the species

PROTEIN SEQUENCES
162,601 (*T. aestivum*: 104,701 | *Pgt* Ug99: 22,524 | *Pgt* 21-0: 35,376)
BLAST

EFFECTOR PROTEINS
11,500 (*Pgt* Ug99: 4,330 | *Pgt* 21-0: 7,170)
Search

SECRETORY PROTEINS
7,451 (*Pgt* Ug99: 2,193 | *Pgt* 21-0: 5,258)
Search

HOST-PATHOGEN INTERACTOME COMPARISON
Explore the tool

© 2022 | Utah State University | Kaundal Artificial Intelligence & Advanced Bioinformatics Laboratory | Department of Plants, Soils, and Climate | Center for Integrated BioSystems

Figure 1. The 'Homepage' of TRustDB database server.

Databases: hpidb, dipdb, mintdb
Host: *Triticum aestivum*
 Minimum Identity (%): 80 ; Minimum Coverage (%): 50 ; Expected value: 1e-10 ;
Pathogen: *Puccinia graminis Ug99*
 Minimum Identity (%): 50 ; Minimum Coverage (%): 30 ; Expected value: 1e-10 ;
 Email address provided: noemail

Network Visualization

Copy Excel PDF Search:

Host Protein	Pathogen Protein	Database	Host Interactor	Pathogen Interactor
TraesCS1A02G020500.2	GMQ_09878T0	hpidb	P60712	A0A061D645
TraesCS1A02G020500.2	GMQ_09878T0	hpidb	P63258	A0A061D645
TraesCS1A02G020500.2	GMQ_15825T0	hpidb	P60712	A0A2C9C261
TraesCS1A02G020500.2	GMQ_11645T0	hpidb	P60712	A0A2C9C261
TraesCS1A02G020500.2	GMQ_09111T0	hpidb	P63258	A0A3Y1NQI2
TraesCS1A02G020500.2	GMQ_09878T0	hpidb	P63258	A7AX51
TraesCS1A02G020500.2	GMQ_09878T0	hpidb	P60712	A7AX51
TraesCS1A02G020500.2	GMQ_11645T0	hpidb	P60712	Q65810
TraesCS1A02G020500.2	GMQ_11645T0	hpidb	P60712	Q91N49
TraesCS1A02G020500.2	GMQ_15163T0	hpidb	P60712	Q91N49

Showing 1 to 10 of 7,624 entries

Previous 1 2 3 4 5 ... 763 Next

Figure 2. The 'Table' view of resulting host–pathogen interactions. The interologs of the corresponding host and pathogen query proteins are also provided. Results can be downloaded in an Excel or PDF format or can be copied to the clipboard.

(http://ftp.ebi.ac.uk/ensemblgenomes/pub/release-51/fungi/fasta/puccinia_graminisug99/pep/) and NCBI (https://www.ncbi.nlm.nih.gov/assembly/GCA_008522505.1), respectively. All the proteomes were individually analyzed with Cluster Database at High Identity with Tolerance (CD-HIT) at 100% identity to remove redundancy, resulting in 104 701, 22 524 and 35 376 proteins of *T. aestivum*, *Pgt Ug99* and *Pgt 21-0*, respectively. The information of these proteomes has also been made available in 'Datasets' section of the database, which allows the user to download the proteomes from their respective resources. Different tools were used for the prediction of the functional annotations of the host and pathogen proteins, which are described in the respective features section.

Database architecture and web interface

TRustDB is hosted on a CentOS Linux system in a High-Performance Computing cluster facility using a virtual machine. The web server runs using Apache HTTP server v2.4.41 and is designed using multiple scripting languages: MySQL as the data storage engine; PHP scripts for querying the MySQL tables and HTML/CSS/Bootstrap 4/JavaScript stack for the enhanced frontend. Figure 1 represents the homepage of TRustDB.

Prediction tools implemented on TRustDB

Interactomics: predict host–pathogen interactions

PPIs are the most crucial component in causing the disease in plants. Hence, the prediction of the potential PPIs enhances the understanding of the disease infection mechanism and the defense response generated by the host cells (20). TRustDB provides a comprehensive tool, known as 'Interactomics', that predicts the interolog-based interactions between host and pathogen proteins. Interolog approach relies on the sequence homology of the proteins between two species. At the backend, the host and pathogen proteomes were Basic Local Alignment Search Tool (BLAST) searched against the seven standard PPI databases (HPIDB, MINT, DIP, STRING, BioGRID, IntAct and PHI-base), and the resulting alignments were stored. For PPI prediction, the user can select the known database(s) of their choice, input the alignment filtering options (*e*-value, % identity and % coverage) and enter the email address to receive the notification about the job completion. Following this, the stored BLAST alignments are used to predict the interactions using R script and SQL functions, locally. The resulting interactions can be downloaded in Excel or PDF format or can also be copied to the clipboard (Figure 2). Furthermore, the predicted interactions can be visualized on the web interface, which is implemented

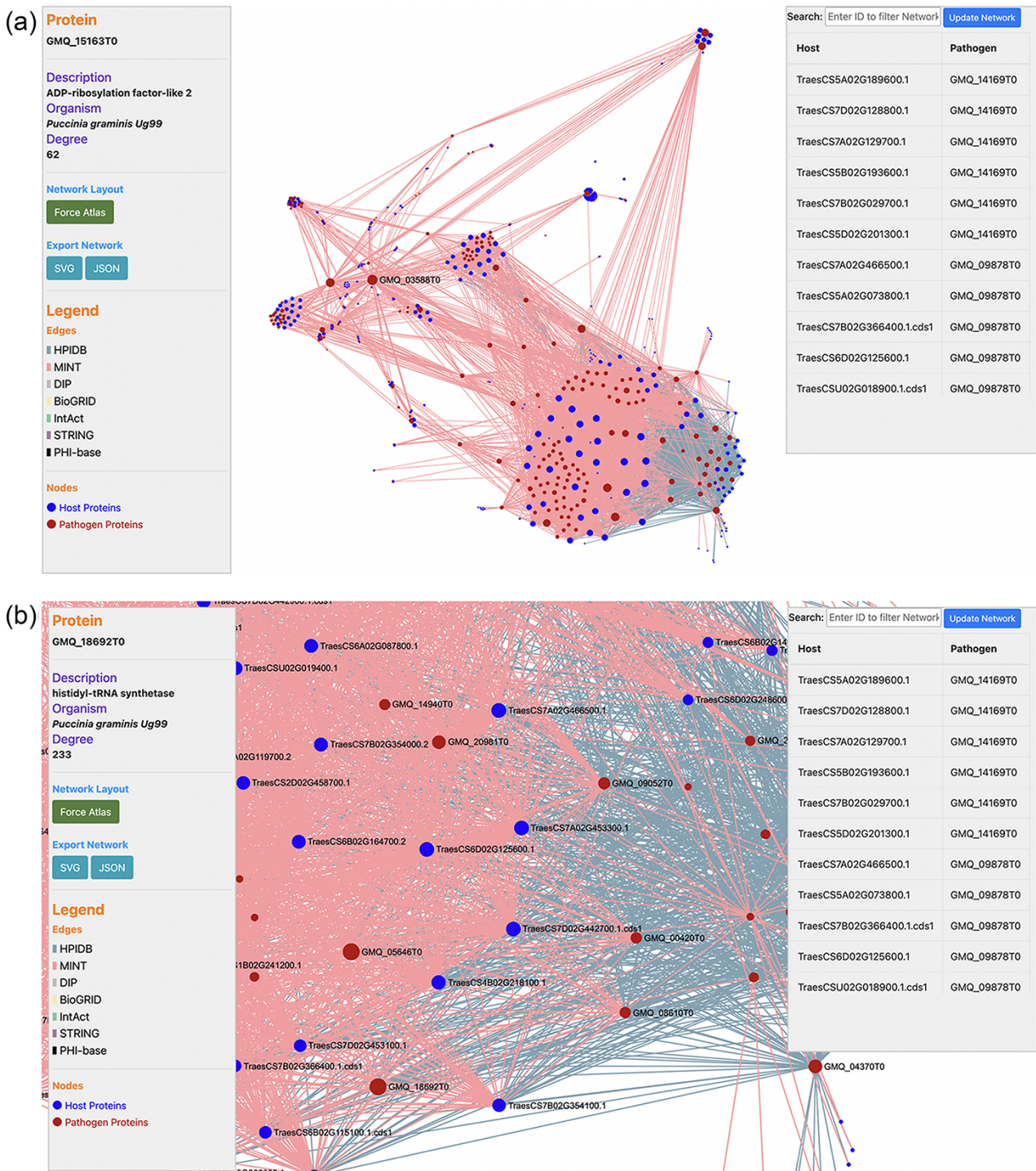


Figure 3. (a) Network visualization of the predicted interactome from the 'Interactomics' tool; (b) A zoomed-in version of the network to view proteins of interest. Users can click (or hover) on any node to view the description of a particular host/pathogen protein in the predicted network.

using the JavaScript-based library, SigmaJS (<https://www.sig-majs.org/>). This library is very flexible and can handle a large number of interactions very efficiently. The resulting network (Figure 3a) is represented by different colors that represent different known PPI database(s) selected initially and can also be downloaded in SVG and JSON format, with JSON being the highly acceptable format by most of the network analyzers for further enhancement of the network. Users can zoom-in on the network visualization to view any proteins of interest

(Figure 3b); the description of a particular host/pathogen protein in the predicted network can be obtained by hovering over the protein of interest. The results from the interactomics tool are also stored on a backup server for 30 days.

Advanced search module

The 'advanced search' interface of TRustDB provides comprehensive functionality to search the database using specific

KAABIL
Kaundi Artificial Intelligence
Advanced Bioinformatics Lab

UtahStateUniversity

Home Tools Features Datasets User guide

Puccinia graminis Ug99 Search e.g: 'GMQ_00017'

Advanced Search Module

You may use this form to search proteins that match selected filtering criteria. Wild cards (partial words/names) are enabled for this search. This module could be use if for example you would like to retrieve certain records of a given length or that are related with some biological process.

This module is comprehensive and will look for protein annotations, Gene Ontology terms, experiment descriptions and many others that would match to the provided keyword.

Keyword * (e.g: GMQ_02857T1, GTPase) Species

Kinase All species

Protein length Genome range Subcellular localization

Min (49) - Max (5035) Start - End All localizations

Protein length and genome range values must be numeric, if those are not numeric the values will be ignored in the search.

*Keyword is the only field that is required to run the search, you may leave the others blank if preferred. Placeholders inside input boxes are just hints for the users to help them to know what to type inside. If the user want to use an additional filter (e.g: Protein length) must type something in those boxes to run a valid search.

Search

Figure 4. A snapshot of the 'Advanced search' module of TRustDB; users can search the database using specific keywords, protein IDs and/or other genomic information of the proteins.

keywords and/or genomic information of the proteins (Figure 4). This module extracts the related proteins by searching multiple SQL tables in accordance with the user query. The users can also use the 'basic' search available under the 'Features' menu on the database. The user can enter biological keyword of choice to obtain all the proteins related to the query. Additionally, the user can also give a particular value for the protein length or genome range in the query, otherwise the program filters out all the proteins for the specific keyword in the results. A specific subcellular localization can also be selected from the dropdown menu, which enables the user to search the database for the proteins in the particular subcellular localization. The search outputs the proteins for the queried keyword along with information such as gene coordinates, protein length, description of the protein and gene ID associated with the proteins. Furthermore, on selecting any protein ID, the user will be taken to a new page that displays additional information of the protein such as NCBI reference ID, locus tag, InterPro annotations and functional domains of the particular protein, if available. From this page, the user can also obtain the amino acid sequence of the protein. The external database IDs such as NCBI, InterPro and others are linked to the respective databases. The interconnection of various features such as subcellular localization, InterPro annotations, functional domain mappings and genomic information makes the 'advanced search' an ideal, time-efficient tool for the users.

BLAST search

BLAST is based on sequence homology and finds the similarity between two sequences. We implemented BLAST on the database to provide a functionality, whereby the users can query the nucleotide/protein of interest against the available proteomes on TRustDB. The user can submit either nucleotide or protein sequences, which are automatically detected by the program, followed by performing BLASTp

or BLASTx alignment, accordingly. A choice of specific weight matrix in BLAST can also be made from the available options. The results of BLAST search are available to be downloaded in PDF or Excel file formats, with an option to download the alignments as a text file. Additionally, a detailed version (Figure 5) of the BLAST alignments is available that includes the alignments in graphical format for enhanced visualization, which is implemented using the JavaScript-based library, BlasterJS (21). These grayscale/colored alignments can be downloaded in JPEG or PNG format.

TRustDB features: Annotations of host and pathogen proteins

The primary goal of TRustDB is to provide the researchers with the functional annotations of the host and pathogen proteins for a better understanding of the protein function during disease infection mechanisms. The proteins were analyzed using various prediction tools to obtain the annotations such as subcellular localization, gene ontology annotations, functional domain mappings, Puccinia orthologs, effector proteins, secretory proteins and host TFs. These annotations are implemented individually in the 'Features' section of the database. Apart from the interactomics tool, the pre-calculated host-pathogen PPIs between *T. aestivum* and Puccinia species are available in the 'Host-pathogen interactions' module under the 'Features' menu, based on default BLAST alignment filters (*e*-value $1e-04$, sequence identity 30% and sequence coverage 40%). These pre-calculated interactions were predicted using our in-house python scripts, and the BLAST parameters combination was selected based on the highest number of effector and secretory proteins obtained in the predicted interactions. More details about the schema followed for pre-calculated interactions can be found under the 'information icon' of the 'Host-pathogen interactions' module



Figure 5. A detailed view of the BLAST result alignments.

of TRustDB. Similarly, most of the feature modules contain the ‘information icon’ that gives brief information about the data available on that particular page. A few of the functional annotations such as subcellular localization, gene ontology (GO) term annotations and functional domain mappings are implemented for both host and pathogen proteins in a separate sub-section for host and pathogen under the features. To obtain the subcellular localization of the host proteins, the proteins were analyzed using the SVM-based tool Plant-mSubP (22), while the pathogen proteins were analyzed in DeepLoc 1.0 (23). Furthermore, the standalone version of InterProScan (24) was used to obtain the GO term annotations and functional domain mappings of the host and pathogen proteins. These modules provide the user with the conserved domains and GO terms associated with the specific protein. The InterPro accessions and domain names have been provided, along with their respective descriptions.

There is a wide range of *P. graminis* strains, which evolved from the African race ‘TTKSK’ (Ug99, named after stem rust

infection in Uganda in 1999) (25). Based on this, we predicted the orthologs of *Pgt* Ug99 and *Pgt* 21-0 using the tool ‘OrthoFinder’ (26). This information is available on the database and contains a unique ‘Ortho Group’, corresponding to the *Pgt* Ug99 and *Pgt* 21-0 proteins that are orthologs of each other. The user can also avail the protein sequence of the orthologs by clicking the identifiers under the ‘Ortho Group’ tab. The pathogens secrete effector and secretory proteins into the host cells for proliferation and successful infection, thus subverting the host cell physiology and suppressing the defense responses (27, 28). Thus, we implemented the modules for the pathogen proteins that serve as effector and secretory proteins separately on the database. The effector and secretory proteins were predicted using the tools EffectorP 2.0 (<https://effectorp.csiro.au/>) (29) and SignalP 5.0 (<https://services.healthtech.dtu.dk/service.php?SignalP-5.0>) (30), respectively. The information about the pathogen protein description and length is also available in these modules. Furthermore, the TFs in the host are known to play a critical role

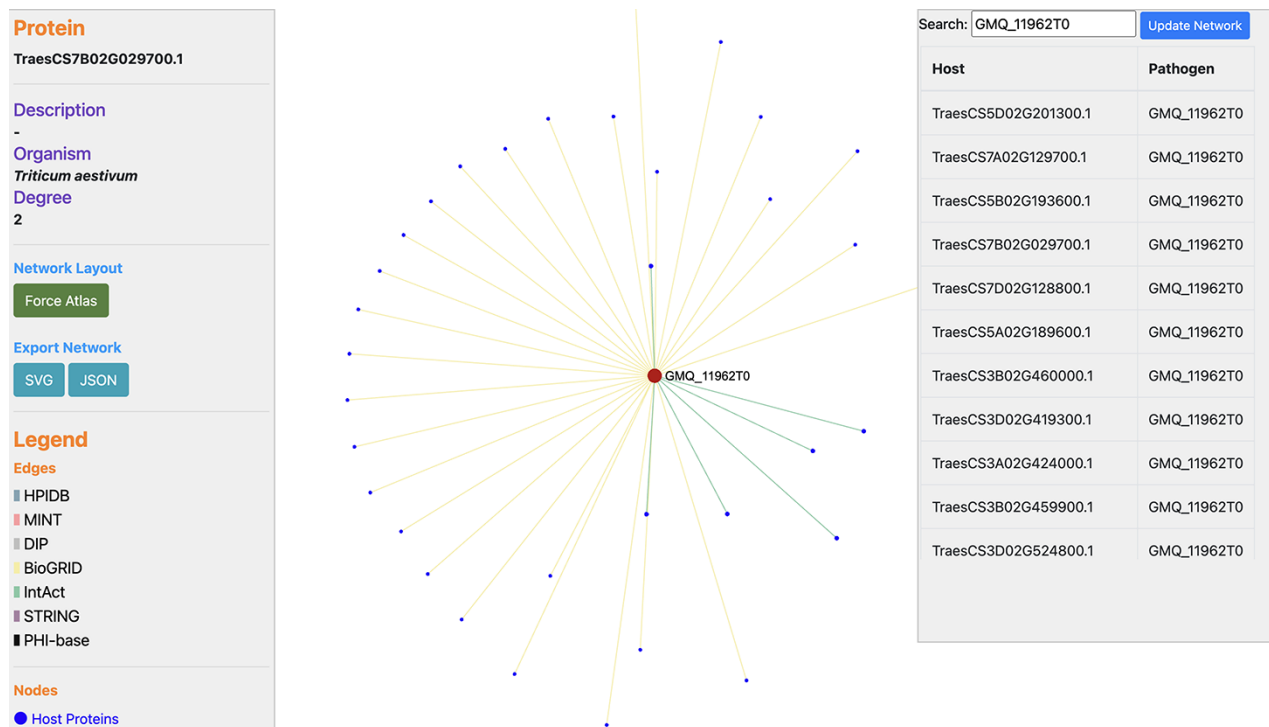


Figure 6. Network visualization of an Ug99 effector protein, GMQ_11962T0, interacting with 34 *T. aestivum* proteins as predicted by TRustDB (Supplementary Material, Sheet 4). GMQ_11962T0 has been shown to be involved in hydrolase activity (GO:0016817) and spliceosome (ko03040), which plays a role in the colonization and pathogenicity of the fungal pathogens.

in defense mechanisms by regulating various biological processes and genes involved in the plant immune responses during various biotic and abiotic stresses (31). The TFs for *T. aestivum* proteins were predicted using PlantTFDB v4.0 (<http://planttfdb.gao-lab.org/>) (32), which is an extensive resource of TFs. Each of the TFs is externally linked to its respective transcription factor family, which gives more information about the TF family. In this module, we have also provided the users with the information on host Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways and their respective KEGG descriptions, which are linked to the KEGG pathway database.

Use case and database validation

In wheat, the disease resistance to stem rust is based on the ‘gene-for-gene’ resistance concept (33). In the past, various QTLs and resistance (R) genes (Sr5, Sr13, Sr22, Sr33, Sr45 and others) against stem rust disease have been identified on wheat chromosomes 1B, 2B, 3B, 5A, 5D, 6A, 6D, 7A and 7B (34, 35). We identified the proteins related to these chromosomes, which resulted in 35 667 wheat proteins, interacting with 1717 pathogen proteins (Supplementary Material, Sheet 1 and Sheet 2, respectively). Among these proteins, 288 host proteins were identified to serve as TFs and 17 pathogen proteins were found to be effectors (Supplementary Material, Sheet 3). For example, Ug99 effector protein ‘GMQ_11962T0’ interacts with 34 *T. aestivum* proteins (Supplementary Material, Sheet 4), as can be visualized in a network (Figure 6). For experimental/literature validation of the PPI pairs, we randomly selected the PPI pair

‘TraesCS1B02G424700.1 - GMQ_11962T0’ from the predicted interactome. The functional annotations of the host protein ‘TraesCS1B02G424700.1’ revealed that the protein is involved in plant-defense-related biological processes such as plant hormone signal transduction pathway (ko04075) and cellular response to auxin stimulus (GO:0071365). This protein was also found to belong to the TF family ‘ARF’ (i.e. Auxin Response Factor) and localized in the nucleus of cell. Researchers in the past have reported the role of signaling pathways such as abscisic acid pathway and jasmonic acid (JA) pathway in the regulation of stomatal closure during biotic/abiotic stress conditions (36). Another research demonstrated that JA and ethylene were required for the activation of the defensin gene (PDF1.2) that was found to be induced on fungal attack (37). These reports show the significant role of the plant hormones in plant defense against the pathogen attack. On the other side, the interacting pathogen protein ‘GMQ_11962T0’ was found to be involved in hydrolase activity (GO:0016817) and spliceosome (ko03040), which plays a role in the colonization and pathogenicity of the fungal pathogens. Several glycoside hydrolase (GH) families have been identified that serve as effectors and play a critical role in the pathogenicity of fungi. These are responsible for initiating various diseases in plants, for example GH3 is responsible for leaf spot disease and GH10 is responsible for vascular disease in tomato. A GH12 protein (PsXEG1) resulted in triggering the cell death in soybean (38). Furthermore, a study on *Sclerotinia sclerotiorum* indicated that the alternative splicing events occur during fungal colonization and are responsible for functional diversity of the proteins (39). To further verify the database, we also queried the host and pathogen protein in the

above-mentioned interaction pair on different modules available in the 'Features' menu in the database. The annotations of the proteins searched on TRustDB were found to be similar as those obtained from the respective annotation resources for GO/KEGG, thus validating the annotations available in the database.

Additionally, we also deciphered some other functional differences between the two pathogen strains used in the study. For example, we found that the *Pgt* Ug99 strain was involved in monoterpenoid biosynthesis pathway, snoRNA binding (GO:0030515) and GTPase complex (GO:1905360). Meanwhile, the *Pgt* 21-0 strain was involved in plant hormone signal transduction pathway, carotenoid biosynthesis, SUMO ligase complex (GO:0106068) and cell wall modifications (GO:0042545). These significantly enriched GO terms and KEGG pathways regulate the fungal growth and development during the infection, as reported earlier (40–43).

Conclusion and future directions

The data available on TRustDB are a beneficial host–pathogen PPI resource for molecular biologists for analyzing and understanding the infection mechanisms of stem rust disease in *T. aestivum* and for the breeders to develop disease-resistant crop lines, thus enhancing the yield and quality of the crop. In the near future, the database is expected to be augmented by incorporating other important rust diseases of wheat.

Supplementary data

Supplementary data are available at Database Online.

Acknowledgements

The authors are thankful to the members of Kaundal Artificial intelligence and Advanced Bioinformatics Lab (KAABiL) for their assistance in the development of the database. We appreciate our colleagues for testing the web server. The authors are also thankful to the referees for their helpful suggestions.

Funding

The authors acknowledge the support to this study from the Utah Agricultural Experiment Station (UAES), Utah State University, grant # A48025 to R Kaundal. This research was also partially supported by UAES office and approved as journal paper number 9599. The funding body did not play any role in the design of this study; the collection, analysis, or interpretation of data; or in the writing of this manuscript.

Data availability

The database is freely available for public use at <http://bioinfo.usu.edu/trustdb/>. The datasets explored in this study to develop the computational models are also available on this website. Supplementary data/results are available in one Excel file containing four different sheets.

Conflict of interest

The authors declare no conflict of interest.

References

- Anderson,P.K., Cunningham,A.A., Patel,N.G. *et al.* (2004) Emerging infectious diseases of plants: Pathogen pollution, climate change and agrotechnology drivers. *Trends Ecol. Evol.*, **19**, 535–544. [10.1016/j.tree.2004.07.021](https://doi.org/10.1016/j.tree.2004.07.021).
- Kataria,R. and Kaundal,R. (2022) Deciphering the host – pathogen interactome of the wheat – common bunt system: a step towards enhanced resilience in next generation wheat. *Int. J. Mol. Sci.*, **23**, 2589. [10.3390/ijms23052589](https://doi.org/10.3390/ijms23052589).
- Shewry,P.R. and Hey,S.J. (2015) The contribution of wheat to human diet and health. *Food Energy Secur.*, **4**, 178–202. [10.1002/fes3.64](https://doi.org/10.1002/fes3.64).
- Simón,M.R., Börner,A. and Struik,P.C. (2021) Editorial: fungal wheat diseases: etiology, breeding, and integrated management. *Front. Plant Sci.*, **12**, 1–5. [10.3389/fpls.2021.671060](https://doi.org/10.3389/fpls.2021.671060).
- Velásquez,A.C., Castroverde,C.D.M. and He,S.Y. (2018) Plant–pathogen warfare under changing climate conditions. *Curr. Biol.*, **28**, R619–R634. [10.1016/j.cub.2018.03.054](https://doi.org/10.1016/j.cub.2018.03.054).
- Singh,R.P., Singh,P.K., Rutkoski,J. *et al.* (2016) Disease impact on wheat yield potential and prospects of genetic control. *Annu. Rev. Phytopathol.*, **54**, 303–322. [10.1146/annurev-phyto-080615-095835](https://doi.org/10.1146/annurev-phyto-080615-095835).
- Figueroa,M., Hammond-Kosack,K.E. and Solomon,P.S. (2018) A review of wheat diseases—a field perspective. *Mol. Plant Pathol.*, **19**, 1523–1536. [10.1111/mpp.12618](https://doi.org/10.1111/mpp.12618).
- Li,F., Upadhyaya,N.M., Sperschneider,J. *et al.* (2019) Emergence of the Ug99 lineage of the wheat stem rust pathogen through somatic hybridisation. *Nat. Commun.*, **10**. [10.1038/s41467-019-12927-7](https://doi.org/10.1038/s41467-019-12927-7).
- Upadhyaya,N.M., Garnica,D.P., Karaoglu,H. *et al.* (2015) Comparative genomics of Australian isolates of the wheat stem rust pathogen *Puccinia graminis* f. Sp. *tritici* reveals extensive polymorphism in candidate effector genes. *Front. Plant Sci.*, **5**, 1–13. [10.3389/fpls.2014.00759](https://doi.org/10.3389/fpls.2014.00759).
- Leonard,K.J. and Szabo,L.J. (2005) Stem rust of small grains and grasses caused by *Puccinia graminis*. *Mol. Plant Pathol.*, **6**, 99–111. [10.1111/j.1364-3703.2005.00273.x](https://doi.org/10.1111/j.1364-3703.2005.00273.x).
- Wang,S., Knox,R.E., Depauw,R.M. *et al.* (2009) Markers to a common bunt resistance gene derived from 'Blizzard' wheat (*Triticum aestivum* L.) and mapped to chromosome arm 1BS. *Theor. Appl. Genet.*, **119**, 541–553. [10.1007/s00122-009-1063-9](https://doi.org/10.1007/s00122-009-1063-9).
- Van de Wouw,A.P., Marcroft,S.J., Ware,A. *et al.* (2014) Breakdown of resistance to the fungal disease, blackleg, is averted in commercial canola (*Brassica napus*) crops in Australia. *F. Crop. Res.*, **166**, 144–151. [10.1016/j.fcr.2014.06.023](https://doi.org/10.1016/j.fcr.2014.06.023).
- Kataria,R. and Kaundal,R. (2021) alfaNET: a database of Alfalfa-bacterial stem blight protein–protein interactions revealing the molecular features of the disease-causing bacteria. *Int. J. Mol. Sci.*, **22**, 8342. [10.3390/ijms22158342](https://doi.org/10.3390/ijms22158342).
- Liu,J.Z. and Lam,H.M. (2019) Signal transduction pathways in plants for resistance against pathogens. *Int. J. Mol. Sci.*, **20**, 2335. [10.3390/ijms20092335](https://doi.org/10.3390/ijms20092335).
- Javed,T., Shabbir,R., Ali,A. *et al.* (2020) Transcription factors in plant stress responses: challenges and potential for sugarcane improvement. *Plants*, **9**, 1–18. [10.3390/plants9040491](https://doi.org/10.3390/plants9040491).
- Kataria,R., Duhan,N. and Kaundal,R. (2022) Computational systems biology of Alfalfa – bacterial blight host-pathogen interactions: uncovering the complex molecular networks for developing durable disease resistant crop. *Front. Plant Sci.*, **12**, 1–17. [10.3389/fpls.2021.807354](https://doi.org/10.3389/fpls.2021.807354).
- Forst,C.V. (2006) Host–pathogen systems biology. *Drug Discov. Today*, **11**, 220–227. [10.1016/S1359-6446\(05\)03735-9](https://doi.org/10.1016/S1359-6446(05)03735-9).
- Kataria,R. and Kaundal,R. (2022) Deciphering the crosstalk mechanisms of wheat-stem rust pathosystem: genome-scale prediction unravels novel host targets. *Front. Plant Sci.*, **13**. [10.3389/fpls.2022.895480](https://doi.org/10.3389/fpls.2022.895480).
- Duplessis,S., Cuomo,C.A., Lin,Y.C. *et al.* (2011) Obligate biotrophy features unraveled by the genomic analysis of rust fungi. *Proc. Natl. Acad. Sci. U. S. A.*, **108**, 9166–9171. [10.1073/pnas.1019315108](https://doi.org/10.1073/pnas.1019315108).

20. Sahu,S.S., Weirick,T. and Kaundal,R. (2014) Predicting genome-scale Arabidopsis-*Pseudomonas syringae* interactome using domain and interolog-based approaches. *BMC Bioinform.*, **15**, S13. [10.1186/1471-2105-15-S11-S13](https://doi.org/10.1186/1471-2105-15-S11-S13).
21. Blanco-Míguez,A., Fdez-Riverola,F., Sánchez,B. *et al.* (2018) BlasterJS: A novel interactive JavaScript visualisation component for BLAST alignment results. *PLoS One*, **13**, 1–10. [10.1371/journal.pone.0205286](https://doi.org/10.1371/journal.pone.0205286).
22. Sahu,S.S., Loaiza,C.D. and Kaundal,R. (2020) Plant-mSubP: a computational framework for the prediction of single- and multi-target protein subcellular localization using integrated machine-learning approaches. *AoB Plants*, **12**, 1–10. [10.1093/aobpla/plz068](https://doi.org/10.1093/aobpla/plz068).
23. Almagro Armenteros,J.J., Sønderby,C.K., Sønderby,S.K. *et al.* (2017) DeepLoc: prediction of protein subcellular localization using deep learning. *Bioinformatics*, **33**, 3387–3395. [10.1093/bioinformatics/btx431](https://doi.org/10.1093/bioinformatics/btx431).
24. Jones,P., Binns,D., Chang,H.Y. *et al.* (2014) InterProScan 5: genome-scale protein function classification. *Bioinformatics*, **30**, 1236–1240. [10.1093/bioinformatics/btu031](https://doi.org/10.1093/bioinformatics/btu031).
25. Olivera,P.D., Badebo,A., Xu,S.S. *et al.* (2012) Resistance to race TTKSK of *Puccinia graminis* f. sp. *tritici* in Emmer wheat. *Crop Sci.*, **52**, 2234–2242. [10.2135/cropsci2011.12.0645](https://doi.org/10.2135/cropsci2011.12.0645).
26. Emms,D.M. and Kelly,S. (2019) OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol.*, **20**, 1–14. [10.1186/s13059-019-1832-y](https://doi.org/10.1186/s13059-019-1832-y).
27. Selin,C., de Kievit,T.R., Belmonte,M.F. *et al.* (2016) Elucidating the role of effectors in plant-fungal interactions: Progress and challenges. *Front. Microbiol.*, **7**, 1–21. [10.3389/fmicb.2016.00600](https://doi.org/10.3389/fmicb.2016.00600).
28. Feldman,D., Yarden,O. and Hadar,Y. (2020) Seeking the roles for fungal small-secreted proteins in affecting saprophytic lifestyles. *Front. Microbiol.*, **11**. [10.3389/fmicb.2020.00455](https://doi.org/10.3389/fmicb.2020.00455).
29. Sperschneider,J., Dodds,P.N., Gardiner,D.M. *et al.* (2018) Improved prediction of fungal effector proteins from secretomes with EffectorP 2.0. *Mol. Plant Pathol.*, **19**, 2094–2110. [10.1111/mpp.12682](https://doi.org/10.1111/mpp.12682).
30. Almagro Armenteros,J.J., Tsirigos,K.D., Sønderby,C.K. *et al.* (2019) SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat. Biotechnol.*, **37**, 420–423. [10.1038/s41587-019-0036-z](https://doi.org/10.1038/s41587-019-0036-z).
31. Falak,N., Imran,Q.M., Hussain,A. *et al.* (2021) Transcription factors as the ‘Blitzkrieg’ of plant defense: A pragmatic view of nitric oxide’s role in gene regulation. *Int. J. Mol. Sci.*, **22**, 1–23. [10.3390/ijms22020522](https://doi.org/10.3390/ijms22020522).
32. Jin,J., Tian,F., Yang,D.C. *et al.* (2017) PlantTFDB 4.0: Toward a central hub for transcription factors and regulatory interactions in plants. *Nucleic Acids Res.*, **45**, D1040–D1045. [10.1093/nar/gkw982](https://doi.org/10.1093/nar/gkw982).
33. Ayliffe,M., Singh,R. and Lagudah,E. (2008) Durable resistance to wheat stem rust needed. *Curr. Opin. Plant Biol.*, **11**, 187–192. [10.1016/j.pbi.2008.02.001](https://doi.org/10.1016/j.pbi.2008.02.001).
34. Bokore,F.E., Cuthbert,R.D., Hiebert,C.W. *et al.* (2021) Mapping stem rust resistance loci effective in Kenya in Canadian spring wheat (*Triticum aestivum* L.) lines ‘AAC Prevail’ and ‘BW961’. *Can. J. Plant Pathol.*, **43**, S263–S274. [10.1080/07060661.2021.1966651](https://doi.org/10.1080/07060661.2021.1966651).
35. Hatta,M.A.M., Arora,S., Ghosh,S. *et al.* (2021) The wheat Sr22, Sr33, Sr35 and Sr45 genes confer resistance against stem rust in barley. *Plant Biotechnol. J.*, **19**, 273–284. [10.1111/pbi.13460](https://doi.org/10.1111/pbi.13460).
36. Lee,S.C. and Luan,S. (2012) ABA signal transduction at the crossroad of biotic and abiotic stress responses. *Plant Cell Environ.*, **35**, 53–60. [10.1111/j.1365-3040.2011.02426.x](https://doi.org/10.1111/j.1365-3040.2011.02426.x).
37. Ku,Y.S., Sintaha,M., Cheung,M.Y. *et al.* (2018) Plant hormone signaling crosstalks between biotic and abiotic stress responses. *Int. J. Mol. Sci.*, **19**, 3206. [10.3390/ijms19103206](https://doi.org/10.3390/ijms19103206).
38. Bradley,E.L., Ökmen,B., Doehlemann,G. *et al.* (2022) Secreted glycoside hydrolase proteins as effectors and invasion patterns of plant-associated fungi and oomycetes. *Front. Plant Sci.*, **13**. [10.3389/fpls.2022.853106](https://doi.org/10.3389/fpls.2022.853106).
39. Ibrahim,H.M.M., Kusch,S., Didelon,M. *et al.* (2021) Genome-wide alternative splicing profiling in the fungal plant pathogen *Sclerotinia sclerotiorum* during the colonization of diverse host families. *Mol. Plant Pathol.*, **22**, 31–47. [10.1111/mpp.13006](https://doi.org/10.1111/mpp.13006).
40. Chiang,Y.M., Szewczyk,E., Davidson,A.D. *et al.* (2009) A gene cluster containing two fungal polyketide synthases encodes the biosynthetic pathway for a polyketide, asperfuranone, in *Aspergillus nidulans*. *J. Am. Chem. Soc.*, **131**, 2965–2970. [10.1021/ja8088185](https://doi.org/10.1021/ja8088185).
41. Noar,R.D., Thomas,E., Xie,D.Y. *et al.* (2019) A polyketide synthase gene cluster associated with the sexual reproductive cycle of the banana pathogen, *Pseudocercospora fijiensis*. *PLoS One*, **14**, 1–26. [10.1371/journal.pone.0220319](https://doi.org/10.1371/journal.pone.0220319).
42. Dallery,J.F., Adelin,É., Le Goff,G. *et al.* (2019) H3K4 trimethylation by CclA regulates pathogenicity and the production of three families of terpenoid secondary metabolites in *Colletotrichum bigginsianum*. *Mol. Plant Pathol.*, **20**, 831–842. [10.1111/mpp.12795](https://doi.org/10.1111/mpp.12795).
43. Avalos,J., Pardo-Medina,J., Parra-Rivero,O. *et al.* (2017) Carotenoid biosynthesis in *Fusarium*. *J. Fungi*, **3**, 39. [10.3390/jof3030039](https://doi.org/10.3390/jof3030039).