



Article

# TritiKBdb: A Functional Annotation Resource for Deciphering the Complete Interaction Networks in Wheat-Karnal Bunt Pathosystem

Naveen Duhan<sup>1,2,†</sup> , Raghav Kataria<sup>1,2,†</sup>  and Rakesh Kaundal<sup>1,2,3,\*</sup> 

<sup>1</sup> Department of Plants, Soils, and Climate, College of Agriculture and Applied Sciences, Utah State University, Logan, UT 84322, USA; naveen.duhan@usu.edu (N.D.); raghav.kataria@usu.edu (R.K.)

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

<sup>3</sup> Department of Computer Science, College of Science, Utah State University, Logan, UT 84322, USA

\* Correspondence: rkaundal@usu.edu; Tel.: +1-(435)-797-4117; Fax: +1-(435)-797-2766

† These authors contributed equally to this work.

**Abstract:** The study of molecular interactions, especially the inter-species protein-protein interactions, is crucial for understanding the disease infection mechanism in plants. These interactions play an important role in disease infection and host immune responses against pathogen attack. Among various critical fungal diseases, the incidences of Karnal bunt (*Tilletia indica*) around the world have hindered the export of the crops such as wheat from infected regions, thus causing substantial economic losses. Due to sparse information on *T. indica*, limited insight is available with regard to gaining in-depth knowledge of the interaction mechanisms between the host and pathogen proteins during the disease infection process. Here, we report the development of a comprehensive database and webserver, *TritiKBdb*, that implements various tools to study the protein-protein interactions in the *Triticum species-Tilletia indica* pathosystem. The novel ‘interactomics’ tool allows the user to visualize/compare the networks of the predicted interactions in an enriched manner. *TritiKBdb* is a user-friendly database that provides functional annotations such as subcellular localization, available domains, KEGG pathways, and GO terms of the host and pathogen proteins. Additionally, the information about the host and pathogen proteins that serve as transcription factors and effectors, respectively, is also made available. We believe that *TritiKBdb* will serve as a beneficial resource for the research community, and aid the community in better understanding the infection mechanisms of Karnal bunt and its interactions with wheat. The database is freely available for public use at <http://bioinfo.usu.edu/tritikbdb/>.

**Keywords:** effector proteins; domain; interolog; Karnal bunt; wheat; protein-protein interactions; *Tilletia*; transcription factors; *Triticum*



**Citation:** Duhan, N.; Kataria, R.; Kaundal, R. TritiKBdb: A Functional Annotation Resource for Deciphering the Complete Interaction Networks in Wheat-Karnal Bunt Pathosystem.

*Int. J. Mol. Sci.* **2022**, *23*, 7455.

<https://doi.org/10.3390/ijms23137455>

ijms23137455

Academic Editor: Fucheng Lin

Received: 31 May 2022

Accepted: 30 June 2022

Published: 5 July 2022

**Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Wheat (*Triticum* spp.) is the most important staple crop of the world, and a significant source of various nutrients, including proteins, vitamins, carbohydrates, lipids, and other minor nutrients, and host to a wide range of fungal pathogens [1,2]. Recently, infectious diseases and pests have become a considerable concern in wheat production, owing to about 21.5% of yield loss, depending on environmental conditions and the pathogen’s virulence. Fungal pathogens belonging to *Puccinia*, *Ustilago*, and *Tilletia* species are considered the primary cause of biotic stress in wheat. They are responsible for considerable losses in the quality and quantity of the crop [3]. The quarantine disease in wheat, Karnal bunt (KB), is caused by *Tilletia indica* and has emerged in various countries worldwide, including India, Mexico, South Africa, and the USA [4]. The disease is primarily seed-borne, but the teliospores can also spread through soil or air, thus affecting the grain quality. Following

previous reports, the annual yield loss of the disease ranges from 0.5% to 1%. Still, KB is recognized as having a severe economic impact due to the loss of grain weight (about 0.25%), followed by the imposition of various international regulatory/quarantine restrictions on the crop grown in KB-infected regions [5]. Most countries have a zero tolerance for Karnal bunt in import shipments.

The morphological and physiological variability in *T. indica* isolates enhance their ability to infect a wide range of hosts; thus, they are considered genetically variable for developing resistant crop varieties [6]. Several methods, such as disease diagnosis in the field, immunodiagnostic assays, and molecular techniques, have been implemented to detect the disease, but these have different constraints. For example, disease detection in the field is challenging due to the production of a low number of bunted kernels in each spike. At the same time, the molecular techniques are expensive and time-consuming [7–9].

Molecular interactions form the basis of pathogenicity. The protein-protein interactions (PPIs) between host and pathogen are highly dynamic throughout the disease infection mechanism [10–12]. The pathogens secrete effector proteins for successful invasion of host cellular defense mechanisms. In contrast, the host activates a suite of immune responses such as PAMP-triggered immunity (PTI), effector-triggered immunity (ETI), hypersensitive response (HR), and other metabolic pathways against the pathogen attack [13,14]. Studying these interactions between host and pathogen proteins is vital to researchers. Also, due to the limited information on pathogen pathogenicity-related biomarkers, no successful *Triticum* cultivar resistant to KB has been developed [15]. We developed a database, *TritiKBdb*, using two computational approaches, interolog-based and domain-based, to predict the protein-protein interactions between two widely grown wheat species (*T. aestivum* and *T. turgidum*), and *T. indica*. This database is the first open-source platform which allows the research community to predict protein-protein interactions and obtain the functional annotations of the host and pathogen proteins to better understand the disease infection mechanism of Karnal bunt in *Triticum* species. *TritiKBdb* aims to provide a platform for comparative study of the *Triticum* species proteins on infection with Karnal bunt. One could compare and visualize these interaction networks in an efficient and enriched visualization environment. *TritiKBdb* is freely accessible at <http://bioinfo.usu.edu/tritikbdb/>.

## 2. Results and Discussion

### 2.1. The Compendium

In the past decade, there has been a substantial increase in infectious diseases in crops, leading to reduced yield and crop quality. Understanding the infection mechanism requires in-depth knowledge of the functioning of the host and pathogen proteins and the cellular environment where the interaction between the host and pathogen proteins occurs. The molecular techniques to study these interactions are efficient but time-consuming and expensive. Integrating computational approaches with comprehensive systems biology can enhance the study of protein-protein interactions [16–18]. The primary goal of developing *TritiKBdb* is to provide the research community with a comprehensive platform containing the respective functional annotations of the host and pathogen involved in the protein-protein interactions during the Karnal bunt disease infection in wheat. The database implements different tools to predict the inter- and intra-species PPIs between *Triticum* species and *Tilletia indica* and further study the interactions using the protein annotations available in various database modules.

### 2.2. Interactome Prediction Tool

We implemented a comprehensive ‘Interactomics’ tool to predict the experimentally validated protein-protein interactions between *Triticum* species and *T. indica* (Figure 1). This tool allows the user to predict the host-pathogen interactions based on various gold standard databases (mentioned in Materials and Methods). The user can select any interaction category (Interolog/Domain/Consensus), choose the available interaction databases, and specify the alignment filtering values such as sequence identity (%), sequence coverage (%),

and e-value to obtain the interactions. The resulting interactions include the host-pathogen interaction pair from the source interaction databases, experimental method used for interaction validation, interaction type (physical or direct), confidence score, and PMID of the publication for the validated interactions linked to PubMed (Figure 2). All the proteins are linked to their respective external resources (Ensembl, UniProt, NCBI, etc.) to provide more information to the user.

**Figure 1.** The Interactome search and Comparison tool of *TriticBdb*.

Users can visualize the resulting interactions with the network module, where the size of the nodes represents the degree of the protein. The user can also visualize the network of the interactions of interest. The user can view a specific interaction in the network by selecting a distinct edge between two proteins. The user can also download the network in 'JSON' format for advanced visualization, which is the go-to format for various network visualization tools. *TriticBdb* provides the user a distinct functionality to predict the interactions in two ways: (i) PPI prediction for one host against one pathogen and (ii) comparison of PPIs for two hosts against one pathogen. In the 'one host against one pathogen' option, the user can select a host (*T. aestivum* or *T. turgidum*) and pathogen (*T. indica*) to predict the interactions using the available parameters, while in the 'two hosts against one pathogen' option, the user can obtain the interactions in two ways. The first method is through unique interactions with only one host. These interactions are host-specific and include the *Triticum* proteins that are not the orthologs of each other. The second method is used when the interactions are common to both hosts. These interactions are predicted based on the *Triticum* species that are orthologs of each other. This functionality of *TriticBdb* enables the user to study the species-specific interactions and the interactions that are common from both the host species, thus making it highly efficient for the users to compare two *Triticum* species simultaneously (Figure 3).

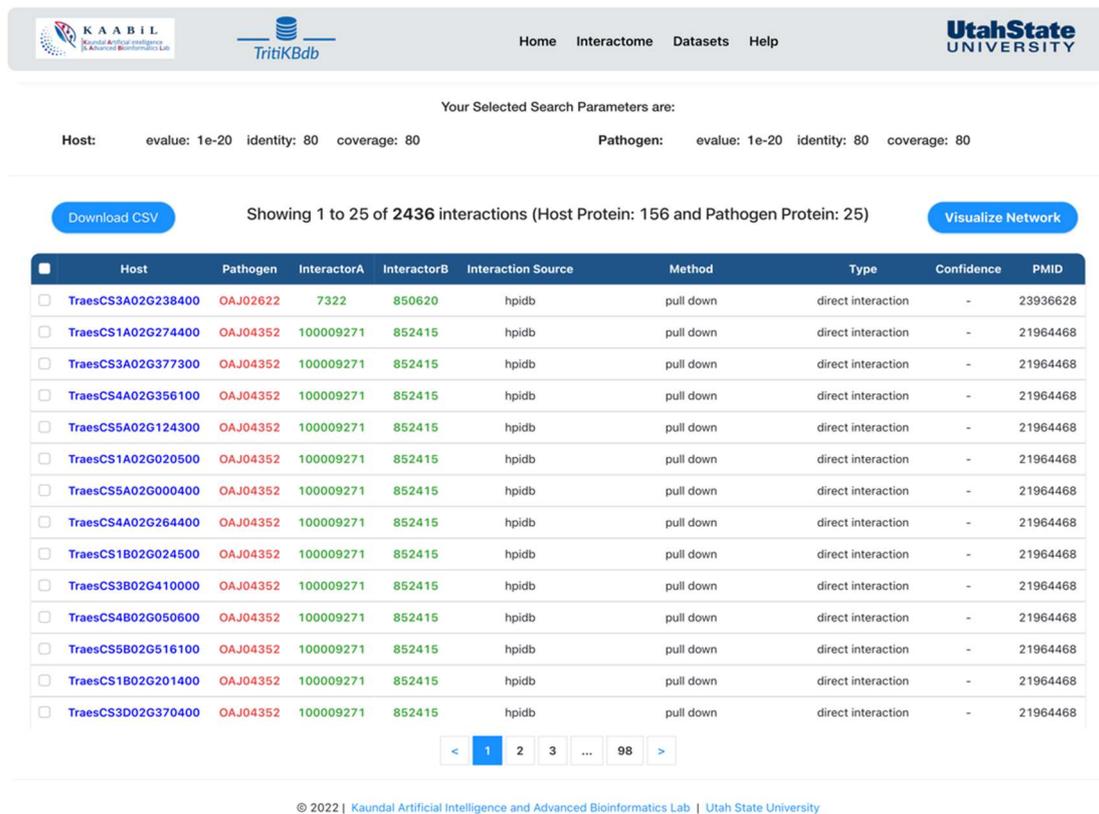


Figure 2. The Interaction Results page of *TriticBdb*.

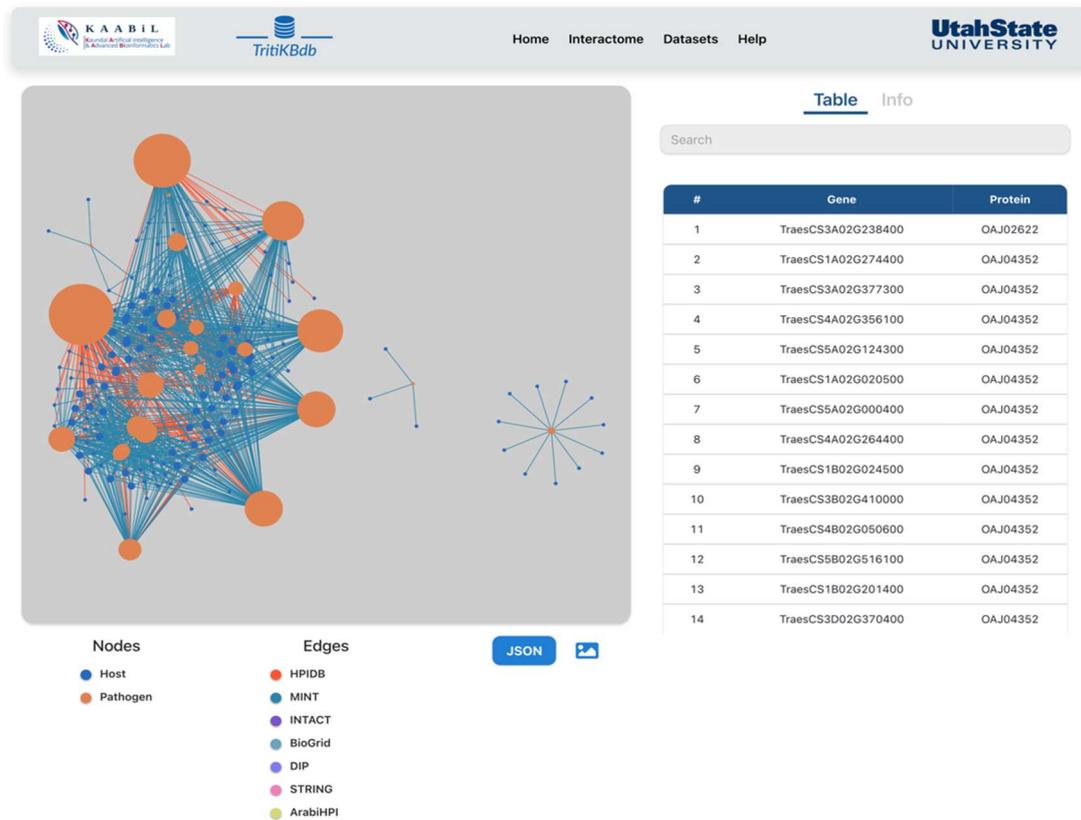


Figure 3. Network of resulting interactions from the host-pathogen interactome module. Node size represents the degree.

### 2.3. BLAST Search

To enable the users to query the protein of interest for similarity against the *Triticum* species and *T. indica* proteomes, we implemented the BLAST (Basic Local Alignment Search Tool) search module on *TritiKBdb*. BLAST module uses in-house Python and JavaScript to perform the alignment. Users can visualize BLAST module results against selected proteomes and export the resulting alignment in a tab-delimited file.

### 2.4. Advanced Search

The advanced search module is a comprehensive module in *TritiKBdb* that allows the user to search the host and pathogen proteins based on specific keywords (Figure 4). These keywords include protein annotation terms, biological terms, protein length, subcellular localization of the protein, and genomic parameters (gene coordinates). The user can provide a keyword (mandatory) and other optional parameters, following which the tool will filter and provide the related proteins for the specific species selected. The tool will search for the provided keyword and other parameters in that same annotation with the chosen annotation.

KAABIL  
Kaundal Artificial Intelligence  
& Advanced Bioinformatics Lab

TritiKBdb

Home Interactome Datasets Help

UtahState  
UNIVERSITY

Advance Search Module

This search module enables search of protein functional annotations like (Gene Ontology, Subcellular Location, Pathways, Transcription Factors, Virulence Proteins)

Select Species

*Triticum aestivum*  *Triticum turgidum*  *Tilletia indica*

Select Annotation

Gene Ontology  KEGG Pathways  Subcellular Localization  Transcription Factors  Functional Domains  Virulence Proteins

Enter a Keyword (e.g: Ubiquitin) OR Enter Protein Accession (e.g: OAJ27264)

keyword Accession

Protein Length

Min - Max

Min Length Max Length

Note: 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.

Advance Search

© 2022 | Kaundal Artificial Intelligence and Advanced Bioinformatics Lab | Utah State University

**Figure 4.** A snapshot of the ‘Advanced search’ module of *TritiKBdb*.

This tool also provides the InterPro accessions associated with the proteins and an external link to the InterProScan website. This module is interconnected with other functional annotations of the proteins available on the database, thus making it extensive and time-efficient.

### 2.5. Functional Annotations

Functional annotation involves integrating the biological information into the proteins, which is crucial to understanding the molecular mechanisms of the disease infection. In

addition to the host-pathogen PPI prediction, we aimed to provide the user with various functional annotations of the host and pathogen proteins involved in the predicted interactions. These annotations include protein subcellular localization, KEGG (Kyoto Encyclopedia of Genes and Genomes) pathways, available domains, gene ontology, host transcription factors, etc. These annotations were generated with the different tools mentioned above in the ‘Materials & Methods’ section.

The ‘functional annotation’ feature for all the species is available on the homepage of *TritiKBdb*. Some of the functional annotations, including subcellular localization, functional domains, KEGG pathways, and gene ontology terms, are available for both the host and pathogen proteins. The determination of the subcellular localization of the protein helps characterize its function in different cellular environments [19]. Since the experimental approaches are time-consuming [20], we implemented diverse computational tools to predict the subcellular localization of the host and pathogen proteins and provide them to users via the web interface. Being structurally and functionally conserved, the domains help the proteins perform various cell functions [21]. The functional domains for the host and pathogen proteins available on *TritiKBdb* represent the specific domain family and the description associated with the domain, along with the source domain database (Pfam, SMART, ProSiteProfiles, PANTHER, PRINTS, Gene3D, TIGRFAM, CDD, Superfamily). The KEGG pathways and Gene Ontology terms (along with their respective description) for the host and pathogen proteins are available on the database.

Further, the database also provides organism-specific annotations of the proteins. The proteins that serve as transcription factors (TFs) are available for *T. aestivum* and *T. turgidum*. Being the primary defense factors in plants, TFs play a crucial role in biotic and abiotic stresses by activating various defense-related molecular responses against pathogen attack [22]. The information about the transcription factor family of the proteins is also made available to the user. TFs available on *TritiKBdb* are 5833 and 3461 for *T. aestivum* and *T. turgidum*, respectively. As both of the *Triticum* species are highly similar concerning their genomes, we implemented the orthologs of *T. aestivum* and *T. turgidum*. We implemented the modules that contain the information about the pathogen proteins that serve as effector and secretory proteins for the pathogen proteins. A new feature module is also implemented whereby the pathogen proteins that serve as both effector and secretory proteins are available. These proteins subvert the host cell’s immune responses during the attack. Apart from this, the effector and secretory proteins also play a role in self-defense, obtaining nutrients from the host, and helping in niche colonization [23]. We believe that the functional annotations of the host and pathogen proteins provided in different modules will enhance the understanding of the disease infection.

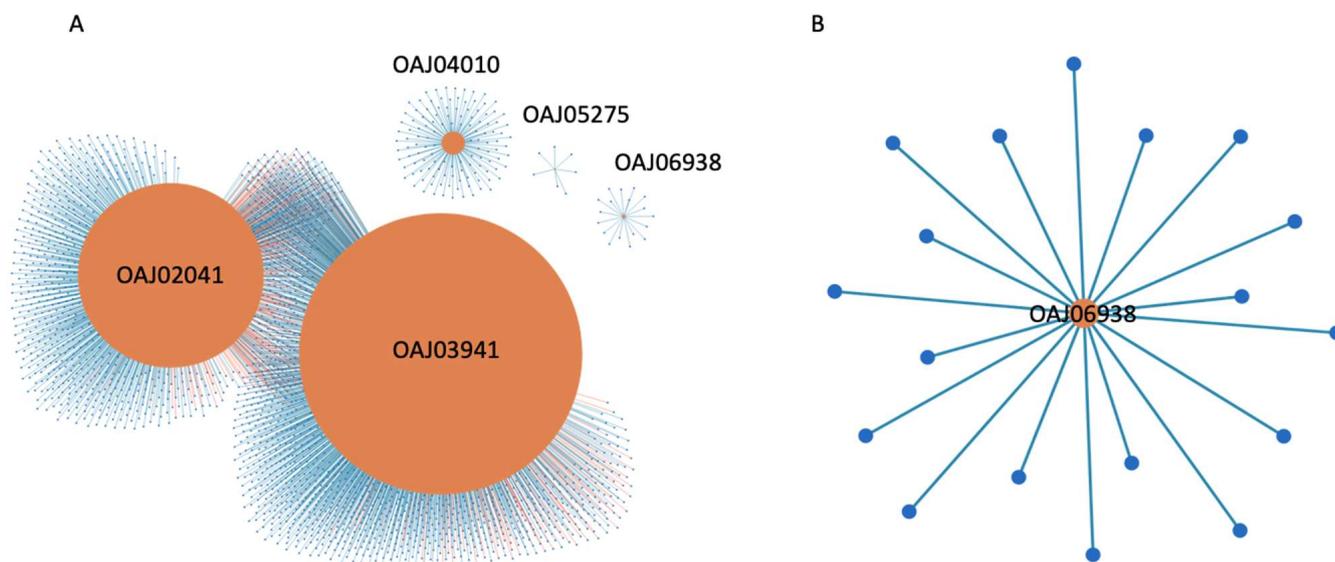
## 2.6. Validation and Use Cases of *TritiKBdb*

We further validated the functional protein annotations in the *TritiKBdb* by integrating it with the disease resistance QTLs/SNPs reported in the literature. Researchers have identified numeral QTLs and SNPs on different chromosomes (1DL, 2BL, 2DL, 3BS, 4AL, 4B, 4D, 5AL, 5AS, 5BL, 6BL, 6BS, 7BS, 7DL) of wheat using genome-wide association studies (GWAS) [24–27]. We found 47,583 wheat proteins in the predicted interactions associated with the bunt-resistance-related chromosomes (Supplementary Material File S1). For database validation, a random host-pathogen interaction pair “TraesCS2D02G583100.1-OAJ04754” was selected from the predicted PPIs, and the proteins were annotated using the functional annotations available on *TritiKBdb*. The host protein in this pair was significantly involved in plant defense-related mechanisms such as biosynthesis of secondary metabolites and response to oxidative stress (GO:0006979) and belonged to the peroxidase domain family. In the previous studies, these mechanisms have been shown to play a crucial role in plant immune responses during biotic stress [28,29]. This protein’s subcellular localization was predicted as ‘plastid,’ which is a major site for the production of defense-related signals and hormones such as ABA, reactive oxygen species (ROS), and others [30]. In comparison, the interacting pathogen protein in the selected PPI pair was

involved in betalain biosynthesis and the small molecule metabolic process (GO:0044281) responsible for fungal pathogenicity and development [31]. In correspondence with the literature, these annotations confirm the viability of the data and makes it a good candidate pair for further experimental validation.

#### HPIs of Experimentally Validated Karnal Bunt Virulence Proteins

To further validate the applicability of *TritiKDb*, we selected some experimentally proven Karnal bunt virulence proteins from the literature [32] and searched those in our database to find their corresponding wheat interacting partners (receptors). About 10 *T. indica* experimentally identified virulence/pathogenicity factors were selected (OAJ02041, OAJ06104, OAJ04365, OAJ03869, OAJ04747, OAJ03941, OAJ05275, OAJ06938, OAJ04010, OAJ06388). The corresponding interactions of these proteins were identified using the ‘Search Interactome’ module on *TritiKDb* in two databases (HPIDB and MINT; users can select as many databases as they want). We found five *T. indica* proteins (at identity > 30, coverage > 40 and e-value <  $1 \times 10^{-4}$ ) interacting with 2101 *T. aestivum* proteins resulting in 2548 unique HPIs (Supplementary Material File S2). The network of these resulting HPIs is shown in Figure 5A. Among these interactions, OAJ03941, OAJ02041, OAJ04010, OAJ06938, and OAJ05275 interact with 1379, 1022, 122, 18, and 7 *T. aestivum* proteins, respectively. These types of analysis with *TritiKDb* can be extremely useful to experimental biologists, plant pathologists, and others when they are inferring functional HPIs within a system which can be directly implemented to develop disease resistant cultivars. Further, one could visualize the individual protein interaction networks on a zoomed scale; for example, OAJ06938 is interacting with 18 *T. aestivum* proteins and the resulting network for this protein is depicted in Figure 5B. The detailed information about the individual interacting pairs, their corresponding NCBI or Uniprot IDs, experimental evidence, confidence score, etc. information can be found in Supplementary Material File S2.



**Figure 5.** (A) Network visualization of 5 karnal bunt virulence proteins interacting with *T. aestivum* proteins. Node size (orange) is based on the degree of interactions, (B) Individual network visualization; shown here OAJ06938 interacting with 18 common wheat (*T. aestivum*) proteins.

Similarly, users can also search the interactions against the durum wheat (*T. turgidum*) and compare the HPIs of Karnal bunt proteins with common wheat as well as durum wheat. Thus, *TritiKDb* can play a significant role in identifying the interacting partners of *T. indica* virulence proteins with the corresponding *T. aestivum* and *T. turgidum* receptor proteins.

### 3. Materials and Methods

#### 3.1. TriticKDb Interface

We developed a user-friendly database, *TriticKDb*, which is freely available at <http://bioinfo.usu.edu/tritikbdb/> (accessed on 9 June 2022) (Figure 6). The database was created using the MERN stack technology and served via NodeJS. MongoDB, a non-structured query language database, was used to store the interaction files, blast files, and functional annotations. The backend APIs were developed using ExpressJS v4.17 and NodeJS v16, and the frontend was developed using REACT. The “Help page” has also been made available for efficient browsing of the database for novice users. Additionally, an information icon on the tools/features page gives a piece of brief information about the functions and datasets available.

**Host Species**

*Triticum aestivum*

Proteins: 104,701  
Gene Ontology: 4904  
KEGG Pathway: 402  
Transcription Factors: 5,833

*Triticum turgidum*

Proteins: 65,409  
Gene Ontology: 3639  
KEGG Pathway: 401  
Transcription Factors: 3,461

**Pathogen Species**

*Tilletia indica*

Proteins: 9,533  
Gene Ontology: 816  
KEGG Pathway: 375  
Effector Proteins: 2651  
Secretory Proteins: 705  
Effector and Secretory: 230

**TriticKDb: a Host-Pathogen Interaction Database for understanding Karnal Bunt disease infection mechanisms in Wheat**

Wheat, the most widely cultivated crop in the world, ranks third and accounts for around 35% of the total grains produced around the globe. Among some critical fungal diseases, the occurrence of Karnal Bunt (*Tilletia indica*, also known as partial bunt) in two of the most widely cultivated species of Wheat (*Triticum aestivum*, *Triticum turgidum*) has affected the grain yield and quality of the crop to a great extent, thus being a major threat to the agriculture industry worldwide. Karnal bunt is recognized to have a severe economic impact due to the loss of grain weight (about 0.25%), followed by the imposition of various international regulatory/quarantine restrictions on the crop grown in Karnal bunt infected regions. Most countries have a zero tolerance for Karnal bunt in import shipments. The morphological and physiological variability in *T. indica* isolates enhance their ability to infect a wide range of hosts; thus, they are considered genetically variable for developing resistant crop varieties. The study of molecular interactions, especially protein-protein interactions, is crucial for understanding the disease infection mechanism in plants. These interactions play an important role in disease infection and host immune responses against the pathogen attack. Here, we report the development of a comprehensive database and webserver, TriticKDb, that implements various tools to study the protein-protein interactions in the Triticum species-Tilletia indica pathosystem. The novel interactomics tool allows the user to visualize / compare the networks of the predicted interactions in an enriched manner. TriticKDb is a user-friendly database that provides functional annotations such as subcellular localization, available domains, KEGG pathways, and GO terms of the host and pathogen proteins. Additionally, the information about the host and pathogen proteins that serve as transcription factors and effectors, respectively, is also made available. We believe that TriticKDb will serve as a beneficial resource to the research community, particularly to the plant breeders and plant pathologists to better understand the Karnal bunt disease mechanisms and interactions with Wheat, and help in developing more efficient and durable disease resistant cultivars.

Search Interactome   Advance Search Module

**Functional Annotations**

*Triticum aestivum*

KEGG Pathways  
Gene Ontology  
Functional Domains  
Subcellular Localization  
Transcription Factors

*Triticum turgidum*

KEGG Pathways  
Gene Ontology  
Functional Domains  
Subcellular Localization  
Transcription Factors

*Tilletia indica*

KEGG Pathways  
Gene Ontology  
Functional Domains  
Subcellular Localization  
Effector Proteins  
Secretory Proteins  
Effector and Secretory Proteins

**Data Sources**

HPIDB<sub>3.0</sub>   IntAct   MINT   BioGRID   STRING   DIP   DOMINE   3did   UniProt   Ensembl

© 2022 | Kaundal Artificial Intelligence and Advanced Bioinformatics Lab | Utah State University

**Figure 6.** The ‘Home’ page of *TriticKDb* database.

#### 3.2. Dataset Collection and Processing

For the protein-protein interaction prediction, the proteomes of *Triticum* species (common wheat and durum wheat) were obtained from Ensembl Plants, while those of *Tilletia indica* were downloaded from Ensembl Fungi (Table 1). The complete proteome of each species was analyzed with CD-HIT (Cluster Database at High Identity with Tolerance) [33] at 100% identity to remove the identical sequences. In the case of *Triticum* species, the

longest sequence of the proteins was considered, followed by analysis with CD-HIT. The information on the datasets used is available on the “Datasets” page of *TritiKDb*.

**Table 1.** Dataset sources of host and pathogen proteomes.

Species	Source	Number of Proteins	
		Downloaded	CD-HIT
<i>Triticum aestivum</i>	Ensembl Plants ( <a href="https://plants.ensembl.org/index.html">https://plants.ensembl.org/index.html</a> , accessed on 8 April 2022)	133,346	104,701
<i>Triticum turgidum</i>	Ensembl Plants ( <a href="https://plants.ensembl.org/index.html">https://plants.ensembl.org/index.html</a> , accessed on 8 April 2022)	196,105	65,409
<i>Tilletia indica</i>	Ensembl Fungi ( <a href="https://fungi.ensembl.org/index.html">https://fungi.ensembl.org/index.html</a> , accessed on 8 April 2022)	9548	9533

Various functional annotations such as subcellular localization, effector proteins, secretory proteins, transcription factors, and functional domains were obtained using different tools (Table 2). Plant-mSubP and DeepLoc employ machine learning-based algorithms to predict the subcellular localization of proteins. EffectorP and SignalP were used to analyze whether a particular pathogen protein serves as an effector and secretory protein. To identify the proteins of *Triticum* species that are transcription factors, we employed PlantTFDB, which determines the transcription factor binding motifs in the promoter region. We also used InterProScan to obtain the functional domains of the host and pathogen proteins.

**Table 2.** Tools employed to obtain functional annotation of the host and pathogen proteins.

Functional Annotation	Tool	Link	Reference
Subcellular localization	Plant-mSubP (Host)	<a href="http://bioinfo.usu.edu/Plant-mSubP/">http://bioinfo.usu.edu/Plant-mSubP/</a> (Accessed on 14 April 2022)	[34]
	DeepLoc 1.0 (Pathogen)	<a href="https://services.healthtech.dtu.dk/service.php?DeepLoc-1.0">https://services.healthtech.dtu.dk/service.php?DeepLoc-1.0</a> (Accessed on 15 April 2022)	[35]
Effector proteins	EffectorP 3.0	<a href="https://effectorp.csiro.au/">https://effectorp.csiro.au/</a> (Accessed on 11 April 2022)	[36]
Secretory proteins	SignalP 6.0	<a href="https://services.healthtech.dtu.dk/service.php?SignalP">https://services.healthtech.dtu.dk/service.php?SignalP</a> (Accessed on 11 April 2022)	[37]
Transcription factors	PlantTFDB v5.0	<a href="http://plantfdb.gao-lab.org/">http://plantfdb.gao-lab.org/</a> (Accessed on 11 April 2022)	[38]
Functional domains	InterProScan	<a href="https://www.ebi.ac.uk/interpro/">https://www.ebi.ac.uk/interpro/</a> (Accessed on 15 April 2022)	[39]

### 3.3. Working of the Interactomics Tool

The interolog and domain-based PPI predictions between *Triticum* species and *T. indica* were accomplished by implementing the gold standard databases (Table 3). These databases include several protein-protein interaction databases: BioGRID [40], DIP [41], HPIDB [42], IntAct [43], MINT [44], STRING [45], and domain-domain interaction databases: 3did [46], DOMINE [47], and IDDI [48]. We filtered these interaction databases for the experimental interactions based on the evidence such as literature support, interaction type/experimental method, and confidence score. For interolog-based prediction, the proteomes were analyzed with BLAST and searched against the PPI databases to obtain the alignments. At the same time, HMMER was used to extract the domains of the proteins of all the species for the domain-based approach. The resulting alignments and domains were then subjected to our in-house python scripts to predict the host-pathogen protein-protein interactions. This tool

also provides the user to visualize the networks of the predicted interactions, for which we implemented Cytoscape [49]. The users can also download the predicted interactions in a tab-delimited format.

**Table 3.** Information of the databases employed inside *TritiKBdb*.

Database	Total Interactions	Link to Database
BioGRID	2,381,484	<a href="https://thebiogrid.org/">https://thebiogrid.org/</a> (accessed on 27 April 2022)
DIP	76,882	<a href="http://dip.doe-mbi.ucla.edu/">http://dip.doe-mbi.ucla.edu/</a> (accessed on 27 April 2022)
HPIDB	69,365	<a href="https://hpidb.igbb.msstate.edu/">https://hpidb.igbb.msstate.edu/</a> (accessed on 27 April 2022)
IntAct	1,184,057	<a href="http://www.ebi.ac.uk/intact/">http://www.ebi.ac.uk/intact/</a> (accessed on 27 April 2022)
MINT	132,249	<a href="http://mint.bio.uniroma2.it/">http://mint.bio.uniroma2.it/</a> (accessed on 27 April 2022)
ArabiHPI	983	Manually curated
STRING	4,313,229	<a href="https://string-db.org/">https://string-db.org/</a> (accessed on 27 April 2022)
3did	11,200	<a href="https://3did.irbbarcelona.org/">https://3did.irbbarcelona.org/</a> (accessed on 27 April 2022)
DOMINE	26,219	<a href="https://manticore.niehs.nih.gov/cgi-bin/Domine">https://manticore.niehs.nih.gov/cgi-bin/Domine</a> (accessed on 27 April 2022)
IDDI	204,716	<a href="http://pcode.kaist.ac.kr/iddi/">http://pcode.kaist.ac.kr/iddi/</a> (accessed on 27 April 2022)

#### 4. Conclusions

The host-pathogen protein-protein interaction database, *TritiKBdb*, was developed to provide the research community with a platform that helps gain in-depth knowledge about the Karnal bunt infection mechanisms in *Triticum* species. The database is built using an extensive framework and hosts various functional annotations of the proteins involved in the host-pathogen interactions. We implemented a customizable ‘interactomics’ tool, which predicts the PPIs based on user-provided parameters. The network visualization of the predicted interactome is displayed in an information-rich and interactive visualization pattern, and the layouts can be customized further based on user-defined schema. Furthermore, the extensive tool ‘advanced search’ is a hub of functional annotation information for the users. We believe *TritiKBdb* will serve as a potential resource to the scientific community, plant breeders, and plant pathologists, to aid them in better understanding the Karnal bunt disease using the protein-protein interaction information with wheat, and aid in the development of more efficient and disease resistant cultivars. In the future, we plan to add more species of wheat (and its pathogens) to the database as their whole genome sequencing datasets become available. Karnal bunt also infects Triticale, a hybrid of wheat and rye, but the whole genome sequences are not available yet for Triticale; when available, they could be added to the *TritiKBdb* resource too.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms23137455/s1>.

**Author Contributions:** R.K. (Rakesh Kaundal) formulated and designed the research. N.D. and R.K. (Raghav Kataria) analyzed the data. N.D. designed and developed the database. Writing—original draft preparation, R.K. (Raghav Kataria); writing—review and editing, R.K. (Rakesh Kaundal) and N.D.; visualization R.K. (Raghav Kataria), N.D. and R.K. (Rakesh Kaundal); supervision, R.K. (Rakesh Kaundal); project administration, R.K. (Rakesh Kaundal); funding acquisition, R.K. (Rakesh Kaundal). All authors have read and agreed to the published version of the manuscript.

**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 9592. 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.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** The authors thank all the members of KAABIL (Kaundal Artificial intelligence & Advanced Bioinformatics Lab) for their valuable suggestions during the development of the database, and for rigorously testing the resource.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Kataria, R.; Kaundal, R. Deciphering the Host-Pathogen Interactome of the Wheat-Common Bunt System: A Step towards Enhanced Resilience in Next Generation Wheat. *Int. J. Mol. Sci.* **2022**, *23*, 2589. [\[CrossRef\]](#)
2. Shewry, P.R.; Hey, S.J. The Contribution of Wheat to Human Diet and Health. *Food Energy Secur.* **2015**, *4*, 178–202. [\[CrossRef\]](#)
3. Savary, S.; Willocquet, L.; Pethybridge, S.J.; Esker, P.; McRoberts, N.; Nelson, A. The Global Burden of Pathogens and Pests on Major Food Crops. *Nat. Ecol. Evol.* **2019**, *3*, 430–439. [\[CrossRef\]](#)
4. Jones, D.R. Arguments for a Low Risk of Establishment of Karnal Bunt Disease of Wheat in Europe. *Eur. J. Plant Pathol.* **2007**, *118*, 93–104. [\[CrossRef\]](#)
5. Bishnoi, S.K.; He, X.; Phuke, R.M.; Kashyap, P.L.; Alakonya, A.; Chhokar, V.; Singh, R.P.; Singh, P.K. Karnal Bunt: A Re-Emerging Old Foe of Wheat. *Front. Plant Sci.* **2020**, *11*, 569057. [\[CrossRef\]](#)
6. Kumar, S.; Singroha, G.; Singh, G.P.; Sharma, P. Karnal Bunt of Wheat: Etiology, Breeding and Integrated Management. *Crop Prot.* **2021**, *139*, 105376. [\[CrossRef\]](#)
7. Tan, M.K.; Brennan, J.P.; Wright, D.; Murray, G.M. A Review of the Methodology to Detect and Identify Karnal Bunt—a Serious Biosecurity Threat. *Australas. Plant Pathol.* **2013**, *42*, 95–102. [\[CrossRef\]](#)
8. McNeil, M.; Roberts, A.M.I.; Cockerell, V.; Mulholland, V. Real-Time PCR Assay for Quantification of *Tilletia Caries* Contamination of UK Wheat Seed. *Plant Pathol.* **2004**, *53*, 741–750. [\[CrossRef\]](#)
9. Kataria, R.; Kaundal, R. Deciphering the Crosstalk Mechanisms of Wheat-Stem Rust Pathosystem: Genome-Scale Prediction Unravels Novel Host Targets. *Front. Plant Sci.* **2022**, *13*. [\[CrossRef\]](#)
10. Kataria, R.; Kaundal, R. AlfaNET: A Database of Alfalfa-Bacterial Stem Blight Protein-Protein Interactions Revealing the Molecular Features of the Disease-Causing Bacteria. *Int. J. Mol. Sci.* **2021**, *22*, 8342. [\[CrossRef\]](#)
11. Jo, E.K. Interplay between Host and Pathogen: Immune Defense and Beyond. *Exp. Mol. Med.* **2019**, *51*, 19–21. [\[CrossRef\]](#)
12. Loaiza, C.D.; Duhan, N.; Lister, M.; Kaundal, R. In silico prediction of host-pathogen protein interactions in melioidosis pathogen *Burkholderia pseudomallei* and human reveals novel virulence factors and their targets. *Brief. Bioinform.* **2020**. [\[CrossRef\]](#)
13. Kataria, R.; Duhan, N.; Kaundal, R. Computational Systems Biology of Alfalfa-Bacterial Blight Host-Pathogen Interactions: Uncovering the Complex Molecular Networks for Developing Durable Disease Resistant Crop. *Front. Plant Sci.* **2022**, *12*, 807354. [\[CrossRef\]](#)
14. Iqbal, Z.; Iqbal, M.S.; Hashem, A.; Abd\_Allah, E.F.; Ansari, M.I. Plant Defense Responses to Biotic Stress and Its Interplay With Fluctuating Dark/Light Conditions. *Front. Plant Sci.* **2021**, *12*, 631810. [\[CrossRef\]](#)
15. Pandey, V.; Singh, M.; Pandey, D.; Kumar, A. Integrated Proteomics, Genomics, Metabolomics Approaches Reveal Oxalic Acid as Pathogenicity Factor in *Tilletia Indica* Inciting Karnal Bunt Disease of Wheat. *Sci. Rep.* **2018**, *8*, 7826. [\[CrossRef\]](#)
16. Kumar, V.; Baweja, M.; Singh, P.K.; Shukla, P. Recent Developments in Systems Biology and Metabolic Engineering of Plant-Microbe Interactions. *Front. Plant Sci.* **2016**, *7*, 1421. [\[CrossRef\]](#)
17. Rodriguez, P.A.; Rothballer, M.; Chowdhury, S.P.; Nussbaumer, T.; Gutjahr, C.; Falter-Braun, P. Systems Biology of Plant-Microbiome Interactions. *Mol. Plant* **2019**, *12*, 804–821. [\[CrossRef\]](#)
18. Kataria, R.; Kaundal, R. WeCoNET: A host-pathogen interactome database for deciphering crucial molecular networks of wheat-common bunt cross-talk mechanisms. *Plant Methods* **2022**, *18*, 1–11. [\[CrossRef\]](#)
19. Scott, M.S.; Calafell, S.J.; Thomas, D.Y.; Hallett, M.T. Refining Protein Subcellular Localization. *PLoS Comput. Biol.* **2005**, *1*, e66. [\[CrossRef\]](#)
20. Mirzaei Mehrabad, E.; Hassanzadeh, R.; Eslahchi, C. PMLPR: A Novel Method for Predicting Subcellular Localization Based on Recommender Systems. *Sci. Rep.* **2018**, *8*, 12006. [\[CrossRef\]](#)
21. Basu, M.K.; Poliakov, E.; Rogozin, I.B. Domain Mobility in Proteins: Functional and Evolutionary Implications. *Brief. Bioinform.* **2009**, *10*, 205–216. [\[CrossRef\]](#)

22. Shahzad, R.; Jamil, S.; Ahmad, S.; Nisar, A.; Amina, Z.; Saleem, S.; Zaffar Iqbal, M.; Muhammad Atif, R.; Wang, X. Harnessing the Potential of Plant Transcription Factors in Developing Climate Resilient Crops to Improve Global Food Security: Current and Future Perspectives. *Saudi J. Biol. Sci.* **2021**, *28*, 2323–2341. [[CrossRef](#)]
23. Snelders, N.C.; Kettles, G.J.; Rudd, J.J.; Thomma, B.P.H.J. Plant Pathogen Effector Proteins as Manipulators of Host Microbiomes? *Mol. Plant Pathol.* **2018**, *19*, 257–259. [[CrossRef](#)]
24. Singh, S.; Sehgal, D.; Kumar, S.; Arif, M.A.R.; Vikram, P.; Sansaloni, C.P.; Fuentes-Dávila, G.; Ortiz, C. GWAS Revealed a Novel Resistance Locus on Chromosome 4D for the Quarantine Disease Karnal Bunt in Diverse Wheat Pre-Breeding Germplasm. *Sci. Rep.* **2020**, *10*, 5999. [[CrossRef](#)]
25. Nelson, J.C.; Autrique, J.E.; Fuentes-Dávila, G.; Sorrells, M.E. Chromosomal Location of Genes for Resistance to Karnal Bunt in Wheat. *Crop Sci.* **1998**, *38*, 231–236. [[CrossRef](#)]
26. Emebiri, L.; Singh, S.; Tan, M.K.; Singh, P.K.; Fuentes-Dávila, G.; Ogonnaya, F. Unravelling the Complex Genetics of Karnal Bunt (*Tilletia indica*) Resistance in Common Wheat (*Triticum Aestivum*) by Genetic Linkage and Genome-Wide Association Analyses. *G3 Genes Genomes Genet.* **2019**, *9*, 1437–1447. [[CrossRef](#)]
27. Gupta, V.; He, X.; Kumar, N.; Fuentes-Dávila, G.; Sharma, R.K.; Dreisigacker, S.; Juliana, P.; Ataei, N.; Singh, P.K. Genome Wide Association Study of Karnal Bunt Resistance in a Wheat Germplasm Collection from Afghanistan. *Int. J. Mol. Sci.* **2019**, *20*, 3124. [[CrossRef](#)]
28. Jan, R.; Asaf, S.; Numan, M.; Lubna, N.M.; Kim, K.M. Plant Secondary Metabolite Biosynthesis and Transcriptional Regulation in Response to Biotic and Abiotic Stress Conditions. *Agronomy* **2021**, *11*, 968–975. [[CrossRef](#)]
29. Huang, C.; Verrillo, F.; Renzone, G.; Arena, S.; Rocco, M.; Scaloni, A.; Marra, M. Response to Biotic and Oxidative Stress in *Arabidopsis thaliana*: Analysis of Variably Phosphorylated Proteins. *J. Proteom.* **2011**, *74*, 1934–1949. [[CrossRef](#)]
30. Park, E.; Nedo, A.; Caplan, J.L.; Dinesh-Kumar, S.P. Plant–Microbe Interactions: Organelles and the Cytoskeleton in Action. *New Phytol.* **2018**, *217*, 1012–1028. [[CrossRef](#)]
31. Fox, E.M.; Howlett, B.J. Secondary Metabolism: Regulation and Role in Fungal Biology. *Curr. Opin. Microbiol.* **2008**, *11*, 481–487. [[CrossRef](#)]
32. Pandey, V.; Gupta, A.K.; Singh, M. Complementary Proteomics, Genomics approaches identifies potential pathogenicity/virulence factors in *Tilletia indica* induced under the influence of host factor. *Sci. Rep.* **2019**, *9*, 553. [[CrossRef](#)]
33. Fu, L.; Niu, B.; Zhu, Z.; Wu, S.; Li, W. CD-HIT: Accelerated for Clustering the next-Generation Sequencing Data. *Bioinformatics* **2012**, *28*, 3150–3152. [[CrossRef](#)]
34. Sahu, S.S.; Loaiza, C.D.; Kaundal, R. Plant-MSubP: A Computational Framework for the Prediction of Single- And Multi-Target Protein Subcellular Localization Using Integrated Machine-Learning Approaches. *AoB PLANTS* **2021**, *12*, plz068. [[CrossRef](#)]
35. Almagro Armenteros, J.J.; Sønderby, C.K.; Sønderby, S.K.; Nielsen, H.; Winther, O. DeepLoc: Prediction of Protein Subcellular Localization Using Deep Learning. *Bioinformatics* **2017**, *33*, 3387–3395. [[CrossRef](#)]
36. Sperschneider, J.; Dodds, P.N. EffectorP 3.0: Prediction of Apoplastic and Cytoplasmic Effectors in Fungi and Oomycetes. *Mol. Plant-Microbe Interact. MPMI* **2022**, *35*, 146–156. [[CrossRef](#)]
37. Teufel, F.; Almagro Armenteros, J.J.; Johansen, A.R.; Gíslason, M.H.; Pihl, S.I.; Tsigirgos, K.D.; Winther, O.; Brunak, S.; von Heijne, G.; Nielsen, H. SignalP 6.0 Predicts All Five Types of Signal Peptides Using Protein Language Models. *Nat. Biotechnol.* **2022**. [[CrossRef](#)]
38. Tian, F.; Yang, D.C.; Meng, Y.Q.; Jin, J.; Gao, G. PlantRegMap: Charting Functional Regulatory Maps in Plants. *Nucleic Acids Res.* **2020**, *48*, D1104–D1113. [[CrossRef](#)]
39. Jones, P.; Binns, D.; Chang, H.Y.; Fraser, M.; Li, W.; McAnulla, C.; McWilliam, H.; Maslen, J.; Mitchell, A.; Nuka, G.; et al. InterProScan 5: Genome-Scale Protein Function Classification. *Bioinformatics* **2014**, *30*, 1236–1240. [[CrossRef](#)]
40. Chatr-Aryamontri, A.; Oughtred, R.; Boucher, L.; Rust, J.; Chang, C.; Kolas, N.K.; O'Donnell, L.; Oster, S.; Theesfeld, C.; Sellam, A.; et al. The BioGRID Interaction Database: 2017 Update. *Nucleic Acids Res.* **2017**, *45*, D369–D379. [[CrossRef](#)]
41. Salwinski, L.; Miller, C.S.; Smith, A.J.; Pettit, F.K.; Bowie, J.U.; Eisenberg, D. The Database of Interacting Proteins: 2004 Update. *Nucleic Acids Res.* **2004**, *32*, 449–451. [[CrossRef](#)]
42. Ammari, M.G.; Gresham, C.R.; McCarthy, F.M.; Nanduri, B. HPIDB 2.0: A Curated Database for Host-Pathogen Interactions. *Database J. Biol. Databases Curation* **2016**, *2016*, baw103. [[CrossRef](#)]
43. Kerrien, S.; Aranda, B.; Breuza, L.; Bridge, A.; Broackes-Carter, F.; Chen, C.; Duesbury, M.; Dumousseau, M.; Feuermann, M.; Hinz, U.; et al. The IntAct Molecular Interaction Database in 2012. *Nucleic Acids Res.* **2012**, *40*, 841–846. [[CrossRef](#)]
44. Licata, L.; Briganti, L.; Peluso, D.; Perfetto, L.; Iannuccelli, M.; Galeota, E.; Sacco, F.; Palma, A.; Nardozza, A.P.; Santonico, E.; et al. MINT, the Molecular Interaction Database: 2012 Update. *Nucleic Acids Res.* **2012**, *40*, 857–861. [[CrossRef](#)]
45. Szklarczyk, D.; Gable, A.L.; Lyon, D.; Jung, A.; Wyder, S.; Huerta-Cepas, J.; Simonovic, M.; Doncheva, N.T.; Morris, J.H.; Bork, P.; et al. STRING V11: Protein-Protein Association Networks with Increased Coverage, Supporting Functional Discovery in Genome-Wide Experimental Datasets. *Nucleic Acids Res.* **2019**, *47*, D607–D613. [[CrossRef](#)]
46. Mosca, R.; Céol, A.; Stein, A.; Olivella, R.; Aloy, P. 3did: A Catalog of Domain-Based Interactions of Known Three-Dimensional Structure. *Nucleic Acids Res.* **2014**, *42*, 374–379. [[CrossRef](#)]

47. Raghavachari, B.; Tasneem, A.; Przytycka, T.M.; Jothi, R. DOMINE: A Database of Protein Domain Interactions. *Nucleic Acids Res.* **2008**, *36*, 656–661. [[CrossRef](#)]
48. Kim, Y.; Min, B.; Yi, G.S. IDDI: Integrated Domain-Domain Interaction and Protein Interaction Analysis System. *Proteome Sci.* **2012**, *10*, S9. [[CrossRef](#)]
49. Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N.S.; Wang, J.T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: A Software Environment for Integrated Models. *Genome Res.* **1971**, *13*, 426. [[CrossRef](#)]