Published online 24 November 2021 *Nucleic Acids Research, 2022, Vol. 50, Database issue* ***D1483–D1490***
https://doi.org/10.1093/nar/gkab1087

PRGdb 4.0: an updated database dedicated to genes involved in plant disease resistance process

Joan Calle García[1], Anna Guadagno[2], Andreu Paytuvi-Gallart[1], Alfonso Saera-Vila[1], Ciro Gianmaria Amoroso[2], Daniela D'Esposito[2], Giuseppe Andolfo[2], Riccardo Aiese Cigliano[1], Walter Sanseverino[1,*] and Maria Raffaella Ercolano[2,*]

[1]Sequentia Biotech SL, Calle Comte D'Urgell 240, 08036 Barcelona, Spain and [2]Dipartimento di Agraria, Università di Napoli 'Federico II', Via Università 100, 80055 Portici, Italy

Received September 14, 2021; Revised October 11, 2021; Editorial Decision October 14, 2021; Accepted November 19, 2021

ABSTRACT

The Plant Resistance Genes database (PRGdb; http://prgdb.org/prgdb4/) has been greatly expanded, keeping pace with the increasing amount of available knowledge and data (sequenced proteomes, cloned genes, public analysis data, etc.). The easy-to-use style of the database website has been maintained, while an updated prediction tool, more data and a new section have been added. This new section will contain plant resistance transcriptomic experiments, providing additional easy-to-access experimental information. DRAGO3, the tool for automatic annotation and prediction of plant resistance genes behind PRGdb, has been improved in both accuracy and sensitivity, leading to more reliable predictions. PRGdb offers 199 reference resistance genes and 586.652 putative resistance genes from 182 sequenced proteomes. Compared to the previous release, PRGdb 4.0 has increased the number of reference resistance genes from 153 to 199, the number of putative resistance genes from 177K from 76 proteomes to 586K from 182 sequenced proteomes. A new section has been created that collects plant-pathogen transcriptomic data for five species of agricultural interest. Thereby, with these improvements and data expansions, PRGdb 4.0 aims to serve as a reference to the plant scientific community and breeders worldwide, helping to further study plant resistance mechanisms that contribute to fighting pathogens.

INTRODUCTION

Green plants are ubiquitous in almost every ecosystem in the world, being the fundamental source of food and of a wide range of products (1). Nowadays, near 40% of worldwide crop production is lost due to pests and diseases (2). For this reason, plant breeders and researchers put great efforts in searching for genes involved in plant disease resistance mechanisms. Plants have developed through evolution the ability to recognize potential pathogens and predators and activate defense mechanisms to fight them (3). The activation of these mechanisms is based on specific receptors encoded by the so-called pathogen recognition genes (PRGs).

Even though different architectures can be found among PRGs, the leucine-rich repeat (LRR) domain must be highlighted for its ubiquity. This domain is present in the pattern recognition receptors (PRRs) that, being transmembrane proteins, recognize external signals and lead to the first layer of inducible defenses (also called PAMP-triggered immunity or PTI) (4). These can be divided in two main groups: RLP, that only contain LRR and, of course, transmembrane domain (TM) and RLKs, which in addition contain kinase (KIN) domain (5). Furthermore, the LRR domain is present in NLR proteins, which also contain a nucleotide-binding site (NBS) domain. These receptors are intracellular and lead to a more robust immunity (effector-triggered immunity or ETI) (6), and also can be further divided in two main groups: TNL, which in addition contain a Toll-interleukin 1 receptor (TIR) domain, and CNL, carrying an additional coiled-coil (CC) domain (7). However, there are receptors with other domains instead of LRR. For instance, there are many important PRRs for fungal pathogens that contain Lysin motif (LYSM) instead of LRR, like LYK and LYP proteins, which have a similar architecture to RLK and RLP receptors respectively, but with LYSM instead of LRR domains. Some others contain lectin-like motifs (LECM) instead of LRR, like LECRK proteins, which have similar architecture to RLK and LYK proteins (6).

Important, knowledge about plant-pathogen interactions was gained by working on individual genes. However, disease response is modulated by highly connected gene

*To whom correspondence should be addressed. Tel: +39 812539431; Email: ercolano@unina.it
Correspondence may also be addressed to Walter Sanseverino. Tel: +34 93 010 73 68; Email: wsanseverino@sequentiabiotech.com

© The Author(s) 2021. Published by Oxford University Press on behalf of Nucleic Acids Research.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Downloaded from https://academic.oup.com/nar/article/50/D1/D1483/6438004 by guest on 30 January 2025

networks as well as by a cross talk between various processes and pathways. Until this date, numerous transcriptome studies have been carried out in the area of plant-pathogen interactions, establishing transcriptomics as a suitable platform to elucidate the complexity of the molecular mechanism of such interactions (8–12). Despite the advances in the omics and bioinformatics fields, data exploratory analysis is still a tedious task and the use of bioinformatics tools to study PRGs remains challenging for a significant part of the scientific community. PRGdb (13) was developed to fill this gap and, with the updated version presented here, we believe it could serve as reference to study genes involved in plant disease resistance process PRGs.

In this work, we present version 4.0 of PRGdb. PRGs classes have been expanded to 7, including LYK, LYP and LECRK receptors. The number of references PRGs has increased to 199 and the number of analyzed proteomes to 182. A new section was created that integrates public transcriptome data from studies focused on plant-pathogen interaction for five species of agricultural interest. Our annotation tool, Disease Resistance Analysis and Gene Orthology (DRAGO), has also been improved to render provide more accurate and sensitive annotations of any given DNA or amino acid (AA) sequence. All this information and other features can be browsed in our user-friendly website: http://prgdb.org/prgdb4/.

MATERIALS AND METHODS

New reference PRGs, classes and proteomes

An extensive bibliographic search was conducted to retrieve new cloned resistance genes from 1 August 2016 (last update of PRGdb 3.0) (14) to 1 June 2021. The aim of the search was expanded beyond LRR-containing receptors, encountering cloned genes for LYK, LYP and LECRK proteins. To test whether they were underrepresented in PRGdb 3.0, they were run through the old DRAGO2 tool. As expected, it was not able to identify these new classes. New reference PRGs were established after confirmation of their domain composition using public prediction tools: InterProScan (15), Pfam (16), CDD (17), Smart (18) and Prosite (19). A total of 51 proteins (4 RLP, 6 RLK, 6 LYP, 14 LYK and 21 LECRK) have been included as reference PRGs in PRGdb 4.0 (Supplementary Table S1).

Phytozome V13 (18,20) and Ensembl Plants release 51 (21) public databases were explored to upload any newly released proteomes since the last PRGdb 3.0 update in August 2016.

HMM construction

The AA sequences of reference genes were separated in the seven classes of resistance genes included in PRGdb 4.0 (CNL, TNL, RLK, RLP, LYK, LYP and LECRK). Afterwards, a multiple sequence alignment (MSA) was constructed for each class using MEGA X (MUSCLE algorithm, default parameters) (22,23). Next, the MSAs were used to build Hidden Markov Models (HMMs) as described for PRGdb 3.0 (14). HMMs of each domain were built from the MSA of each class separately, with the exception of the domains LYSM (LYK and LYP proteins were also grouped together to produce additional and more robust HMMs for LYSM) and LECM (additional HMMs were also built for two subgroups within LECRK: one containing domains legume-LECM domains and another bulb-type LECM domains). Resistance domains were located within the MSA using InterProScan (SMART, Pfam, CDD and Prosite tools were activated) and MEGA software was used to visualize the MSAs and determine HMMs domain of origin.

HMMs were further filtered with *hmmsearch* (default parameters; *hmmer* tool; http://hmmer.org/) against the initial FASTA files to test whether they were indeed useful for resistance domain prediction.

DRAGO3 pipeline

The core of DRAGO3 remained the same as the previous version DRAGO2, computing the alignment score of the different hits based on a BLOSUM62 matrix. However, HMMs were updated (except for CNL and TNL classes) and three new protein classes were added: LYK, LYP and LECRK, with all the other non-canonical domain combinations. Minimum score thresholds of DRAGO3 were defined like in the previous version PRGdb 3.0 (14), and CC domains and TM domains were equally predicted by using COILS version 2.2 (24) and TMHMM 2.0c (25) softwares, respectively.

Validation analyses

To evaluate the performance of the new DRAGO3 tool, Araport 11 annotation of the *Arabidopsis thaliana* proteome was analyzed with both DRAGO2 and DRAGO3. The obtained putative resistance genes were then analyzed with command line InterProScan version 5.52-86.0, using Pfam 33.1 and COILS version 2.2 with the other parameters set to default, as a ground truth and compared to DRAGO results according to two criteria: (i) criterion 1 comprised proteins which have been identically predicted by DRAGO and InterProScan, (ii) criterion 2 considered proteins that have been predicted equally or better by DRAGO.

Two different validation sets, one coming from the bibliographic search (12 proteins, Supplementary Table S2) and another one from the unpublished reference database RefPlantNLR (415 proteins) (preprint available here: https://doi.org/10.1101/2020.07.08.193961), were used to calculate various performance metrics based on the Confusion matrix elements: true positive (TP), false positive (FP), true negative (TN), false negative (FN).

RNA-seq data

An extensive literature search was performed to find publicly available RNA-seq experiments in plant-pathogen interaction in five species of agricultural interest: *Solanum lycopersicum*, *Oryza sativa*, *Triticum aestivum*, *Vitis vinifera* and *Arabidopsis thaliana*. In these studies, the different plant species were challenged with various pathogens such as bacteria, fungi, insects and viruses. For rice, wheat, grapevine and Arabidopsis, the lists of DEGs were retrieved from the corresponding works. DEGs lists from the corresponding same study were grouped for each comparative

Downloaded from https://academic.oup.com/nar/article/50/D1/D1483/6438004 by guest on 30 January 2025

condition. For tomato, the raw sequencing data (fastq file) were downloaded from SRA repository of NCBI (https://www.ncbi.nlm.nih.gov/sra) (26) using SRA-toolkit (http://ncbi.github.io/sra-tools/) and analyzed on the same bioinformatic pipeline using web-based A.I.R. RNAseq analysis package (platform) (Sequentia Biotech, Barcelona, Spain) (https://transcriptomics.sequentiabiotech.com/) to explore and compare the different studies. Starting from raw data the AIR platform (27) performs the automatic analysis and comparison of RNA samples including the following steps: (i) quality check, (ii) adapter removal (iii) mapping, (iii) raw count calculation, (iv) DEG analysis. Tomato genome version SL3.0 ITAG 3.0 was used for the mapping.

Relational database

The PRG data was imported into a MySQL (5.5) based relational database hosted in an Ubuntu server (14.04). The web application was developed using NodeJS technology with the ExpressJS web development framework in the back-end and HTML5, CSS3, JavaScript technologies in the front-end, besides importing libraries and frameworks such as Bootstrap or JQuery. PRGdb 4.0 is freely accessible through a web interface at the following address: http://www.prgdb.org.

RESULTS

DRAGO3 improvement

Initially, 209 HMMs were built from the best conserved regions extracted from the MSA of each resistance class. These HMMs were further filtered with the following criteria: (i) HMMs belonging to non-relevant regions (i.e. TM domains, which are analyzed with TMHMM tool, not with DRAGO3 HMMs); (ii) HMMs incapable of recognizing the proteins used to construct them and (iii) new HMMs were compared against DRAGO2 HMMs and those with higher performance were retained. A total of 109 HMMs were finally retained, almost doubling the previous version, DRAGO2.

DRAGO3 validation

The analyses were carried out for all resistance domains in DRAGO3 (TIR, NBS, LRR, LYSM and LECM) except for CC and TM, since they were identified using external softwares. For the shared classes with DRAGO2 (TNL, CNL, RLK and RLP), the percentage of proteins that fulfill criterion 1 was higher in DRAGO3 for all classes (Supplementary Figure S1). For criterion 2, DRAGO3 results also were equal or higher than DRAGO2 for all classes, even though these differences were minimal (Supplementary Figure S1).

Since InterProScan is also a prediction tool, DRAGO3 was validated over two sets of reference resistance genes: (i) a set from the bibliographic search and (ii) a set from RefPlantNLR unpublished database. They were also used to extend the performance comparison with DRAGO2. Again, only the resistance domains depending on DRAGO HMMs were considered. In both validation sets, all the metrics showed overall higher performance for DRAGO3. For our validation set, DRAGO3 showed a slightly better performance. From RefPlantNLR validation, a trade-off was observed between accuracy and sensitivity (increased from DRAGO2) and precision and specificity, which was found to be generally positive for DRAGO3 as reflected by the increase in the accuracy and *F*-score (Supplementary Figure S2), which indicates the balance between precision and sensitivity. Thus, the addition of new reference proteins for almost all classes, allowed the construction of more robust HMMs. For CNL and TNL classes though, HMMs remain the same from PRGdb 3.0.

DRAGO3 outperforms the previous version both in accuracy and sensitivity with both validation sets. It must be noted that RefPlantNLR only contains NLR proteins. Therefore, even though they carry most of the resistance domains in PRGdb 4.0, performance on the prediction of some important domains like LYSM or LECM could not be tested against any gold standard due to the reduced number of cloned genes of LYK, LYP and LECRK classes.

New data and annotations

With the addition of three new resistance classes LYK, LECRK and LYP, PRGdb 4.0 now includes seven canonical classes of resistance proteins in plants and it is able to predict domain combinations that go beyond these well-established classes. Fifty one new proteins (4 RLP, 6 RLK, 6 LYP, 14 LYK and 21 LECRK) have been added as reference PRGs (Supplementary Table S1), leading to a total of 199 reference resistance genes now included in PRGdb 4.0. Also 182 plant proteomes were analyzed with DRAGO3 (Supplementary Table S3) and their putative resistance genes incorporated in PRGdb 4.0. Following NAR guidelines, the Sequentia in-house proteomes present in PRGdb 3.0 were removed and only leave publicly available proteomes were kept in PRGdb 4.0. Of the 182 proteomes, 39 remain unaltered from PRGdb 3.0, 33 have been updated and 110 new proteomes have been added. A total of 586 652 putative resistance genes have been predicted from these 182 proteomes. As in the previous PRGdb 3.0 release, RLK and RLP continue to be the most abundant classes with 192 544 and 150 222 total putative proteins each. Resistance classes LECRK, LYP and LYK are the least abundant with 24 959, 33 293 and 47 035 total putative proteins, respectively. A summary for all the improvements in PRGdb4.0 in comparison to PRGdb3.0 can be found in Supplementary Table S4.

RNA-seq data integration

The extensive literature search regarding plant-pathogen transcriptome analyses yielded a total of 35 RNA-seq studies, nine for *A. thaliana*, nine for *V. vinifera*, eight for *S. lycopersicum*, five for *O. sativa* and four for *T. aestivum*. Table 1 summarizes RNA-seq studies collected, the plant species and the pathogen they interact. The list of differentially expressed genes (DEGs) under given conditions were obtained for all of them and this data has been incorporated into a new section of the PRGdb 4.0, called 'Scientific literature with expression data'. This section is easy

Downloaded from https://academic.oup.com/nar/article/50/D1/D1483/6438004 by guest on 30 January 2025

Table 1. Plant-pathogen transcriptomic data added to PRGdb RNAseq section

Plant species	Pathogen	Pathogen name	Reference
Arabidopsis			
Arabidopsis thaliana	Bacteria	*Pseudomonas syringae*	(28)
Arabidopsis thaliana	Fungus	*Verticillium dahliae*	(29)
Arabidopsis thaliana		*Sclerotinia sclerotiorum*	(30)
Arabidopsis thaliana		*Macrophomina phaseolina*	(31)
Arabidopsis thaliana		*Fusarium oxysporum*	(32)
Arabidopsis thaliana		*Fusarium oxysporum*	(33)
Arabidopsis thaliana	Fungus; insect	*Botrytis cinerea, Pieris rapae*	(34)
Arabidopsis thaliana	Virus	*Turnip crinkle virus*	(35)
Arabidopsis thaliana	Aphid	*Myzus persicae (GPA)*	(36)
Grape wine			
V. vinifera spp. *Sativa, V. vinifera* spp. *Sylvestris*	Fungus	*Erysiphe necator*	(37)
Vitis pseudoreticulata		*Erysiphe necator*	(38)
Vitis. vinifera		*Botrytis cinerea*	(39)
Vitis vinifera		*Botrytis cinerea*	(40)
Vitis vinifera		*Neofusicoccum parvum*	(41)
Vitis vinifera		*Lasiodiplodia theobromae*	(42)
Vitis vinifera		*Erysiphe necator*	(43)
Vitis vinifera		*Plasmopara viticola*	(44)
Vitis vinifera	Phytoplasma; insect	*Flavescence dorèe; Scaphoideus titanus*	(45)
Tomato			
Solanum lycopersicum	Bacteria	*Pseudomonas syringae* pv. tomato DC3000	(46)
Solanum lycopersicum and *Solanum Pimpinellifolium*		*Ralstonia solanacearum*	(47)
Solanum lycopersicum	Fungus	*Phytophthora infestans*	(48)
Solanum lycopersicum		*Cladosporium fulvum*	(49)
Solanum lycopersicum		Sclerotinia sclerotium	(50)
Solanum lycopersicum	Virus	*Southen Tomato Virus*	(51)
Solanum lycopersicum		*Tomato spotted wilt tospovirus*	(52)
Solanum lycopersicum	Pest	*Tuta absoluta*	(53)
Rice			
Oryza sativa L. cv. *Nipponbare*	Bacteria	*Xanthomonas oryzae* pv. *oryzae*	(54)
Oryza sativa L. and *Oryza meyeriana L.*		*Xanthomonas oryzae* pv. *oryzae*	(55)
Oryza sativa L. spp. *Japonica*	Fungus	*Magnaporthe oryzae*	(56)
Oryza sativa		*Fusarium fujikuroi*	(57)
Oryza sativa L. spp. *Japonica*		*Fusarium fujikuroi*	(58)
Wheat			
Triticum aestivum L.	Fungus	*Fusarium pseudograminearum*	(59)
Triticum aestivum L.		*Puccinia striiformis* f. sp. *tritici*	(60)
Triticum aestivum L.		*Tilletia controversa Kühn*	(61)
Triticum aestivum L.	Bacteria	*Azospirillum brasilense*	(62)

to access through the page dedicated to each of the five species with expression analyses (Figure 1A). In the main page available experiments for a particular species are displayed (Figure 1B). Selecting one of them will take the user to the DEG matrix (Figure 1C). The user will be able explore the results of differential expressed analysis of the various studies. This information is displayed as a DEG matrices of up- and down-regulated genes, reporting gene ID, $\log_2$ fold change for the comparison of interest of each study and the gene functional annotation. The user can download the data in csv format, can sort the data by gene ID, log FC and functional description. This addition will provide an experimental perspective on potential genes of interest. Such special section can be used to cross-check activated/repressed genes during different plant stresses. Here we illustrated three usage cases to identify relevant genes for genetic engineering projects, functional characterization studies and/or breeding programs in *A. thaliana* and *S. lycopersicum* datasets:

a) Identification of most relevant plant species stress-related DEGs in response to a given pathogen;
b) Intraspecific approach (a single gene, DEG expressed in response to more stress within the same species);
c) Interspecific approach (orthologous genes, DEGs expressed in different species challenged by the same stress);

a) The list of genes resulting from *A. thaliana* infected with *Verticillium* shows 11.319 DEGs at different time points (29). To identify genes expressed during early plant response against the fungus, we choose to select four moments: 0hpi, 0.25hpi, 0.5hpi and 1hpi. As suggested by Scholz (63) the fungus invades the plant tissue but does not reach the vascular system within the first 24 h of co-culture, so this timeframe is perfect to identify genes differentially regulated during early infection phases. Genes were sorted based on their Log FC (from highest to lowest to know the most activated, or viceversa). Arabidopsis AT1G01260 resulted down regulated at 0hpi (Log FC: −1.62) while its expression profile drastically changes after 0.25, 0.5 and 1hpi (Log FC: 13.04, 11.49 and 11.27, respectively). AT1G01260 (bHLH13) is a transcriptional factor that negatively regulate jasmonate responses, interacting with JAZ proteins. As reported by several studies, Jasmonates regulate plant growth and defence responses (64,65). In Arabidopsis, bHLH13 overexpression affects resistance against the necrotrophic fungi *Botrytis cinerea*

Downloaded from https://academic.oup.com/nar/article/50/D1/D1483/6438004 by guest on 30 January 2025

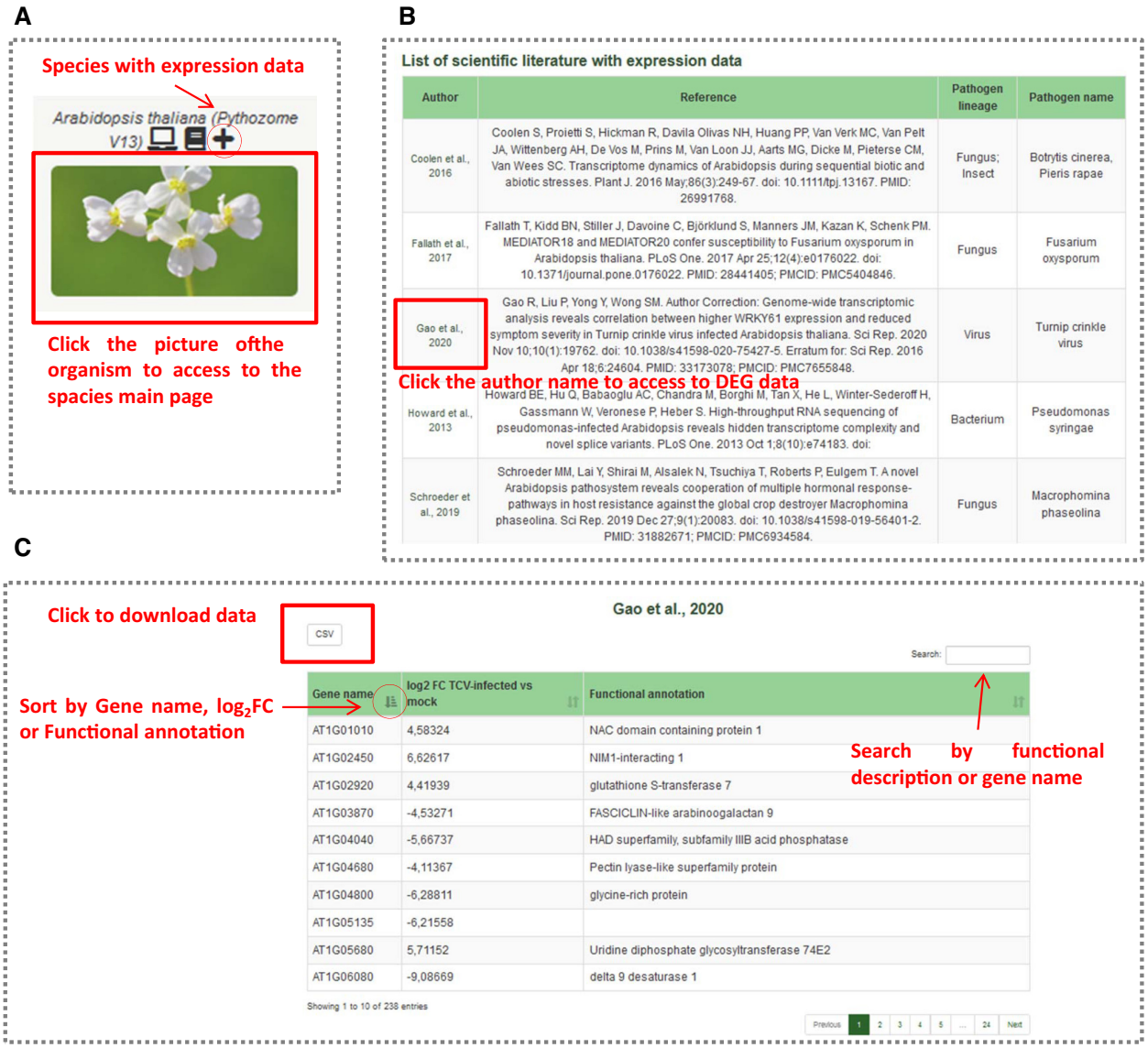

Figure 1. Screenshot of the new section for plant−pathogen differential expression analyses.

and the insect *Spodoptera exigua* (66). *Verticillium dahile* is a chemibiotrophic fungi and AT1G01260 could be selected as target to test its role in plant response to *Verticillium* infection.

b) Comparing two or more list of DEGs within the same species could help to have an overview on the expression profile of a particular gene under different stresses. For example, AT1G01260 (bHLH13) appears up regulated during *Botritis cinerea* and *Pieris rapae* attacks in Coolen (34) dataset and also during the fungus *Sclerotinia sclerotiorum* infections in Sucher (30). These insights could be useful to plane further genetic engineering experiments to better understand gene function, which could affect resistance or other related traits.

c) Interspecific approach could be conducted comparing two different species affected by the same pathogen. In our example, we compare Arabidopsis and tomato infected with *Sclerotinia sclerotiorum*. A short list among the up and down regulated genes in both experiments was analyzed. We found that Arabidopsis AT2G41820 (PXC3) orthologous of tomato Solyc04g015600 (LRR family protein), resulted down regulated in both dataset (Log FC −4.78 and −1.79, respectively).

We consider that data cross-comparison can represent a powerful tool to discover new important genes involved in plant−pathogen interaction. The above case are examples of potential analyses that can be carried out exploring the new expression data sections of our database.

CONCLUSION AND PERSPECTIVES

As sequencing becomes more affordable and efficient, plant genome data availability is rapidly increasing. The extrac-

Downloaded from https://academic.oup.com/nar/article/50/D1/D1483/6438004 by guest on 30 January 2025

tion of meaningful results and conclusions from this data depends on data quality, and exploratory analyses are fundamental to assess this quality. Furthermore, crop research has diversified to several useful species and new resistance strategies are being studied. PRGdb4.0 is an updated database with increased utility to the plant scientific and breeding communities. It can be used to consult plant resistance genes for many plants and algae, to analyze sequences for the prediction of resistance genes and to investigate gene expression under specific plant-pathogen conditions.

The new incorporation of transcriptomic data to PRGdb 4.0 will offer researchers and breeders a more comprehensive perspective of the plant response to pathogens for most relevant plant species. Furthermore, the update of our prediction and annotation tool DRAGO3 will improve the quality of exploratory analyses. The addition of new proteomes will surely allow the study of species that was not possible to date.

PRGdb will continue to grow by incorporating new reference PRGs and genomes as they become publicly available. We will also continue to work on multi-omics data integration in PRGdb, with the aim of becoming a reference database for plant researchers to address critical challenges in plant-pathogen interactions.

DATA AVAILABILITY

PRGdb 4.0 is available at: http://prgdb.org/prgdb4/.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

FUNDING

European Union's Horizon 2020 research and innovation programme through the HARNESSTOM project [101000716]. Funding for open access charge: EU H2020: HARNESSTOM project [101000716].
Conflict of interest statement. None declared.

REFERENCES

1. Kochhar,S.L. (2016) In: *Economic Botany: A Comprehensive Study*. 5th edn, Cambridge University Press.
2. Oerke,E.-C. (2006) Crop losses to pests. *J. Agric. Sci.*, **144**, 31–43.
3. Andolfo,G. and Ercolano,M.R. (2015) Plant innate immunity multicomponent model. *Front. Plant. Sci.*, **6**, 987.
4. Abdul Malik,N.A., Kumar,I.S. and Nadarajah,K. (2020) Elicitor and receptor molecules: orchestrators of plant defense and immunity. *Int. J. Mol. Sci.*, **21**, 963.
5. Torii,K.U. (2004) Leucine-rich repeat receptor kinases in plants: structure, function, and signal transduction pathways. *Int. Rev. Cytol.*, **234**, 1–46.
6. Saijo,Y., Loo,E.P.-I. and Yasuda,S. (2018) Pattern recognition receptors and signaling in plant-microbe interactions. *Plant J.*, **93**, 592–613.
7. Lolle,S., Stevens,D. and Coaker,G. (2020) Plant NLR-triggered immunity: from receptor activation to downstream signaling. *Curr. Opin. Immunol.*, **62**, 99–105.
8. Chakraborty,N. and Basak,J. (2017). Transcriptomics: a successful approach to unravel the molecular mechanism of plant-pathogen interaction in post-genomic era. *Res. J. Biotech.*, **12**, 78–88.
9. Sekhwal,MK., Li,P., Lam,I., Wang,X., Cloutier,S. and You,F.M. (2015) Disease resistance gene analogs (RGAs) in plants. *Int. J. Mol. Sci.*, **16**, 19248–19290.
10. Li,N.Y., Zhou,L., Zhang,D.D., Klosterman,S.J., Li,T.G., Gui,Y.J., Kong,Z.Q., Ma,X.F., Short,D.P.G., Zhang,W.Q. *et al.* (2018) Heterologous expression of the cotton NBS-LRR gene GbaNA1 enhances *Verticillium wilt* resistance in Arabidopsis. *Front. Plant Sci.*, **9**, 119.
11. Xing,L., Hu,P., Liu,J., Witek,K., Zhou,S., Xu,J., Zhou,W., Gao,L., Huang,Z., Zhang,R. *et al.* (2018) Pm21 from *Haynaldia villosa* encodes a CC-NBS-LRR protein conferring powdery mildew resistance in wheat. *Mol. Plant.*, **11**, 874–878.
12. Qi,H., Jiang,Z., Zhang,K., Yang,S., He,F. and Zhang,Z. (2018) PlaD: a transcriptomics database for plant defense responses to pathogens, providing new insights into plant immune system. *Genomics Proteomics Bioinformatics*, **16**, 283–293.
13. Sanseverino,W., Roma,G., De Simone,M., Faino,L., Melito,S., Stupka,E., Frusciante,L. and Ercolano,M.R. (2010) PRGdb: a bioinformatics platform for plant resistance gene analysis. *Nucleic Acids Res.*, **38**, D814–D821.
14. Osuna- Cruz,C.M., Paytuvi- Gallart,A., Di Donato,A., Sundesha,V., Andolfo,G., Aiese Cigliano,R., Sanseverino,W. and Ercolano,M.R. (2018) PRGdb 3.0: a comprehensive platform for prediction and analysis of plant disease resistance genes. *Nucleic Acids Res.*, **46**, D1197–D1201.
15. Jones,P., Binns,D., Chang,H.-Y., Fraser,M., Li,W., McAnulla,C., McWilliam,H., Maslen,J., Mitchell,A., Nuka,G. *et al.* (2014) InterProScan 5: genome-scale protein function classification. *Bioinformatics*, **30**, 1236–1240.
16. Mistry,J., Chuguransky,S., Williams,L., Qureshi,M., Salazar,G.A., Sonnhammer,E.L.L., Tosatto,S.C.E., Paladin,L., Raj,S., Richardson,L.J. *et al.* (2021) Pfam: the protein families database in 2021. *Nucleic Acids Res.*, **49**, D412–D419.
17. Lu,S., Wang,J., Chitsaz,F., Derbyshire,M.K., Geer,R.C., Gonzales,N.R., Gwadz,M., Hurwitz,D.I., Marchler,G.H., Song,J.S. *et al.* (2020) CDD/SPARCLE: the conserved domain database in 2020. *Nucleic Acids Res.*, **48**, D265–D268.
18. Letunic,I., Khedkar,S. and Bork,P. (2021) SMART: recent updates, new developments and status in 2020. *Nucleic Acids Res.*, **49**, D458–D460.
19. de Castro,E., Sigrist,C.J.A., Gattiker,A., Bulliard,V., Langendijk-Genevaux,P.S., Gasteiger,E., Bairoch,A. and Hulo,N. (2006) ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. *Nucleic Acids Res.*, **34**, W362–W365.
20. Goodstein,D.M., Shu,S., Howson,R., Neupane,R., Hayes,R.D., Fazo,J., Mitros,T., Dirks,W., Hellsten,U., Putnam,N. *et al.* (2012) Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res.*, **40**, D1178–D1186.
21. Howe,K.L., Achuthan,P., Allen,J., Allen,J., Alvarez- Jarreta,J., Amode,M.R., Armean,I.M., Azov,A.G., Bennett,R., Bhai,J. *et al.* (2021) Ensembl 2021. *Nucleic Acids Res.*, **49**, D884–D891.
22. Kumar,S., Stecher,G., Li,M., Knyaz,C. and Tamura,K. (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.*, **35**, 1547–1549.
23. Edgar,R.C. (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics*, **5**, 113.
24. Lupas,A., Van Dyke,M. and Stock,J. (1991) Predicting coiled coils from protein sequences. *Science*, **252**, 1162–1164.
25. Sonnhammer,E.L., von Heijne,G. and Krogh,A. (1998) A hidden Markov model for predicting transmembrane helices in protein sequences. *Proc. Int. Conf. Intell. Syst. Mol. Biol.*, **6**, 175–182.
26. Leinonen,R., Sugawara,H. and Shumway,M. (2011) International Nucleotide Sequence Database Collaboration. The sequence read archive. *Nucleic Acids Res.*, **39**, D19–D21.
27. Vara,C., Paytuví- Gallart,A., Cuartero,Y., Le Dily,F., Garcia,F., Salvà- Castro,J., Gómez-H,L., Julià,E., Moutinho,C., Aiese Cigliano,R. *et al.* (2019) Three-dimensional genomic structure and cohesin occupancy correlate with transcriptional activity during spermatogenesis. *Cell Rep.*, **28**, 352–367.
28. Howard,B.E., Hu,Q., Babaoglu,A.C., Chandra,M., Borghi,M., Tan,X., He,L., Winter- Sederoff,H., Gassmann,W., Veronese,P. *et al.* (2013) High-throughput RNA sequencing of pseudomonas-infected Arabidopsis reveals hidden transcriptome complexity and novel splice variants. *PLoS One*, **8**, e74183.

Downloaded from https://academic.oup.com/nar/article/50/D1/D1483/6438004 by guest on 30 January 2025

29. Su,X., Lu,G., Guo,H., Zhang,K., Li,X. and Cheng,H. (2018) The dynamic transcriptome and metabolomics profiling in *Verticillium dahliae* inoculated *Arabidopsis thaliana*. *Sci. Rep.*, **8**, 15404.
30. Sucher,J., Mbengue,M., Dresen,A., Barascud,M., Didelon,M., Barbacci,A. and Raffaele,S. (2020) Phylotranscriptomics of the pentapetalae reveals frequent regulatory variation in plant local responses to the fungal pathogen *Sclerotinia sclerotiorum*. *Plant Cell*, **32**, 1820–1844.
31. Schroeder,M.M., Lai,Y., Shirai,M., Alsalek,N., Tsuchiya,T., Roberts,P. and Eulgem,T. (2019) A novel Arabidopsis pathosystem reveals cooperation of multiple hormonal response-pathways in host resistance against the global crop destroyer *Macrophomina phaseolina*. *Sci Rep.*, **9**, 20083.
32. Zhu,Q.H., Stephen,S., Kazan,K., Jin,G., Fan,L., Taylor,J., Dennis,E.S., Helliwell,C.A. and Wang,M.B. (2013) Characterization of the defense transcriptome responsive to *Fusarium oxysporum*-infection in Arabidopsis using RNA-seq. *Gene*, **512**, 259–266.
33. Fallath,T., Kidd,B.N., Stiller,J., Davoine,C., Björklund,S., Manners,J.M., Kazan,K. and Schenk,P.M. (2017) MEDIATOR18 and MEDIATOR20 confer susceptibility to *Fusarium oxysporum* in *Arabidopsis thaliana*. *PLoS One*, **12**, e0176022.
34. Coolen,S., Proietti,S., Hickman,R., Davila Olivas,N.H., Huang,P.P., Van Verk,M.C., Van Pelt,J.A., Wittenberg,A.H., De Vos,M., Prins,M. *et al.* (2016) Transcriptome dynamics of Arabidopsis during sequential biotic and abiotic stresses. *Plant J.*, **86**, 249–267.
35. Gao,R., Liu,P., Yong,Y. and Wong,S.M. (2020) Author Correction: Genome-wide transcriptomic analysis reveals correlation between higher WRKY61 expression and reduced symptom severity in Turnip crinkle virus infected *Arabidopsis thaliana*. *Sci Rep.*, **10**, 19762.
36. Zhu,L., Guo,J., Ma,Z., Wang,J. and Zhou,C. (2018) Arabidopsis transcription factor MYB102 increases plant susceptibility to aphids by substantial activation of ethylene biosynthesis. *Biomolecules*, **8**, 39.
37. Amrine,K.C., Blanco- Ulate,B., Riaz,S., Pap,D., Jones,L., Figueroa-Balderas,R., Walker,M.A. and Cantu,D. (2015) Comparative transcriptomics of Central Asian *Vitis vinifera* accessions reveals distinct defense strategies against powdery mildew. *Hortic Res.*, **2**, 15037.
38. Weng,K., Li,Z.Q., Liu,R.Q., Wang,L., Wang,Y.J. and Xu,Y. (2014) Transcriptome of *Erysiphe necator*-infected *Vitis pseudoreticulata* leaves provides insight into grapevine resistance to powdery mildew. *Hortic. Res.*, **24**, 14049.
39. Haile,Z.M., Malacarne,G., Pilati,S., Sonego,P., Moretto,M., Masuero,D., Vrhovsek,U., Engelen,K., Baraldi,E. and Moser,C. (2020) Dual transcriptome and metabolic analysis of *Vitis vinifera* cv. Pinot Noir berry and *Botrytis cinerea* during quiescence and egressed infection. *Front. Plant Sci.*, **10**, 1704.
40. Lovato,A., Zenoni,S., Tornielli,G.B., Colombo,T., Vandelle,E. and Polverari,A. (2019) Plant and fungus transcriptomic data from grapevine berries undergoing artificially-induced noble rot caused by *Botrytis cinerea*. *Data Brief.*, **25**, 104150.
41. Czemmel,S., Galarneau,E.R., Travadon,R., McElrone,A.J., Cramer,G.R. and Baumgartner,K. (2015) Genes expressed in grapevine leaves reveal latent wood infection by the fungal pathogen *Neofusicoccum parvum*. *PLoS One*, **10**, e0121828.
42. Zhang,W., Yan,J., Li,X., Xing,Q., Chethana,K.W.T. and Zhao,W. (2019) Transcriptional response of grapevine to infection with the fungal pathogen *Lasiodiplodia theobromae*. *Sci Rep.*, **9**, 5387.
43. Pagliarani,C., Moine,A., Chitarra,W., Meloni,G.R., Abbà,S., Nerva,L., Pugliese,M., Gullino,M.L. and Gambino,G. (2020) The molecular priming of defense responses is differently regulated in grapevine genotypes following elicitor application against powdery mildew. *Int. J. Mol. Sci.*, **21**, 6776.
44. Eisenmann,B., Czemmel,S., Ziegler,T., Buchholz,G., Kortekamp,A., Trapp,O., Rausch,T., Dry,I. and Bogs,J. (2019) Rpv3-1 mediated resistance to grapevine downy mildew is associated with specific host transcriptional responses and the accumulation of stilbenes. *BMC Plant Biol.*, **19**, 343.
45. Bertazzon,N., Bagnaresi,P., Forte,V., Mazzucotelli,E., Filippin,L., Guerra,D., Zechini,A., Cattivelli,L. and Angelini,E. (2019) Grapevine comparative early transcriptomic profiling suggests that *Flavescence dorée* phytoplasma represses plant responses induced by vector feeding in susceptible varieties. *BMC Genomics*, **20**, 526.
46. Yang,Y.X., Wang,M.M., Yin,Y.L., Onac,E., Zhou,G.F., Peng,S., Xia,X.J., Shi,K., Yu,J.Q. and Zhou,Y.H. (2015) RNA-seq analysis reveals the role of red light in resistance against *Pseudomonas syringae* pv. tomato DC3000 in tomato plants. *BMC Genomics*, **16**, 120.
47. French,E., Kim,B.S., Rivera- Zuluaga,K. and Iyer- Pascuzzi,A.S. (2018) Whole root transcriptomic analysis suggests a role for auxin pathways in resistance to *Ralstonia solanacearum* in tomato. *Mol. Plant Microbe Interact.*, **31**, 432–444.
48. Canto- Pastor,A., Santos,B., Valli,A., Summers,W., Schornack,S. and Baulcombe,D. (2019) Enhanced resistance to bacterial and oomycete pathogens by short tandem target mimic RNAs in tomato. *Proc Natl. Acad. Sci. U.S.A.*, **116**, 2755–2760.
49. Zhao,T., Liu,W., Zhao,Z., Yang,H., Bao,Y., Zhang,D., Wang,Z., Jiang,J., Xu,Y., Zhang,H. *et al.* (2019) Transcriptome profiling reveals the response process of tomato carrying Cf-19 and *Cladosporium fulvum* interaction. *BMC Plant Biol.*, **19**, 572.
50. Badet,T., Voisin,D., Mbengue,M., Barascud,M., Sucher,J., Sadon,P., Balagué,C., Roby,D. and Raffaele,S. (2017a) Parallel evolution of the POQR prolyl oligo peptidase gene conferring plant quantitative disease resistance. *PLoS Genet.*, **13**, e1007143.
51. Fukuhara,T., Tabara,M., Koiwa,H. and Takahashi,H. (2020) Effect of asymptomatic infection with southern tomato virus on tomato plants. *Arch. Virol.*, **165**, 11–20.
52. Padmanabhan,C., Ma,Q., Shekasteband,R., Stewart,K.S., Hutton,S.F., Scott,J.W., Fei,Z. and Ling,K.S. (2019) Comprehensive transcriptome analysis and functional characterization of PR-5 for its involvement in tomato Sw-7 resistance to tomato spotted wilt tospovirus. *Sci. Rep.*, **9**, 7673.
53. D' Esposito,D., Manzo,D., Ricciardi,A., Garonna,A.P., De Natale,A., Frusciante,L., Pennacchio,F. and Ercolano,M.R. (2021) Tomato transcriptomic response to *Tuta absoluta* infestation. *BMC Plant. Biol.*, **21**, 358.
54. Yu,C., Chen,H., Tian,F., Leach,J.E. and He,C. (2014) Differentially-expressed genes in rice infected by *Xanthomonas oryzae* pv. oryzae relative to a flagellin-deficient mutant reveal potential functions of flagellin in host-pathogen interactions. *Rice (N Y)*, **7**, 20.
55. Xu,R., Zhou,J., Zheng,E., Yang,Y., Li,D., Chen,Y., Yan,C., Chen,J. and Wang,X. (2021) Systemic acquired resistance plays a major role in bacterial blight resistance in a progeny of somatic hybrids of cultivated rice (*Oryza sativa L.*) and wild rice (*Oryza meyeriana L.*). *J. Plant Dis. Prot.*, **128**, 1023–1040.
56. Bagnaresi,P., Biselli,C., Orrù,L., Urso,S., Crispino,L., Abbruscato,P., Piffanelli,P., Lupotto,E., Cattivelli,L. and Valè,G. (2012) Comparative transcriptome profiling of the early response to *Magnaporthe oryzae* in durable resistant vs susceptible rice (*Oryza sativa L.*) genotypes. *PLoS One*, **7**, e51609.
57. Ji,Z., Zeng,Y., Liang,Y., Qian,Q. and Yang,C. (2016) Transcriptomic dissection of the rice–*Fusarium fujikuroi* interaction by RNA-Seq. *Euphytica*, **211**, 123–137.
58. Matić,S., Bagnaresi,P., Biselli,C., Orrù,L., Amaral Carneiro,G., Siciliano,I., Valé,G., Gullino,M.L. and Spadaro,D. (2016) Comparative transcriptome profiling of resistant and susceptible rice genotypes in response to the seedborne pathogen *Fusarium fujikuroi*. *BMC Genomics*, **17**, 608.
59. Powell,J.J., Carere,J., Fitzgerald,T.L., Stiller,J., Covarelli,L., Xu,Q., Gubler,F., Colgrave,M.L., Gardiner,D.M., Manners,J.M. *et al.* (2017) The Fusarium crown rot pathogen *Fusarium pseudograminearum* triggers a suite of transcriptional and metabolic changes in bread wheat (*Triticum aestivum* L.). *Ann. Bot.*, **119**, 853–867.
60. Dobon,A., Bunting,D.C., Cabrera- Quio,L.E., Uauy,C. and Saunders,D.G. (2016) The host-pathogen interaction between wheat and yellow rust induces temporally coordinated waves of gene expression. *BMC Genomics*, **17**, 380.
61. Ren,Z., Liu,J., Din,G.M.U., Zhang,H., Du,Z., Chen,W., Liu,T., Zhang,J., Zhao,S. and Gao,L. (2020) Transcriptome analysis of wheat spikes in response to *Tilletia controversa* Kühn which cause wheat dwarf bunt. *Sci. Rep.*, **10**, 21567.
62. Camilios- Neto,D., Bonato,P., Wassem,R., Tadra- Sfeir,M.Z., Brusamarello- Santos,L.C., Valdameri,G., Donatti,L., Faoro,H., Weiss,V.A., Chubatsu,L.S. *et al.* (2014) Dual RNA-seq transcriptional analysis of wheat roots colonized by *Azospirillum brasilense* reveals up-regulation of nutrient acquisition and cell cycle genes. *BMC Genomics*, **15**, 378.

Downloaded from https://academic.oup.com/nar/article/50/D1/D1483/6438004 by guest on 30 January 2025

63. Scholz,S.S., Schmidt- Heck,W., Guthke,R., Furch,A.C.U., Reichelt,M., Gershenzon,J. and Oelmüller,R. (2018) *Verticillium dahliae*-Arabidopsis interaction causes changes in gene expression profiles and jasmonate levels on different time scales. *Front. Microbiol.*, **9**, 217.
64. Wasternack,C. (2014) Action of jasmonates in plant stress responses and development–applied aspects. *Biotechnol. Adv.*, **32**, 31–39.
65. Huang,H., Liu,B., Liu,L. and Song,S. (2017) Jasmonate action in plant growth and development. *J Exp. Bot.*, **68**, 1349–1359.
66. Huang,H., Gao,H., Liu,B., Fan,M., Wang,J., Wang,C., Tian,H., Wang,L., Xie,C., Wu,D. *et al.* (2018) bHLH13 regulates jasmonate-mediated defense responses and growth. *Evol. Bioinform. Online*, **14**, 1176934318790265.

Downloaded from https://academic.oup.com/nar/article/50/D1/D1483/6438004 by guest on 30 January 2025