


DATABASE

Open Access



# Rust expression browser: an open source database for simultaneous analysis of host and pathogen gene expression profiles with expVIP

Thomas M. Adams<sup>1</sup>, Tjelvar S. G. Olsson<sup>1</sup>, Ricardo H. Ramírez-González<sup>1</sup>, Ruth Bryant<sup>2</sup>, Rosie Bryson<sup>3</sup>, Pablo Eduardo Campos<sup>4</sup>, Paul Fenwick<sup>5</sup>, David Feuerhelm<sup>6</sup>, Charlotte Hayes<sup>7</sup>, Tina Henriksson<sup>8</sup>, Amelia Hubbard<sup>9</sup>, Radivoje Jevtić<sup>10</sup>, Christopher Judge<sup>9</sup>, Matthew Kerton<sup>11</sup>, Jacob Lage<sup>12</sup>, Clare M. Lewis<sup>1</sup>, Christine Lilly<sup>13</sup>, Udi Meidan<sup>14</sup>, Dario Novoselović<sup>15</sup>, Colin Patrick<sup>16</sup>, Ruth Wanyera<sup>17</sup> and Diane G. O. Saunders<sup>1\*</sup> 

## Abstract

**Background:** Transcriptomics is being increasingly applied to generate new insight into the interactions between plants and their pathogens. For the wheat yellow (stripe) rust pathogen (*Puccinia striiformis* f. sp. *tritici*, *Pst*) RNA-based sequencing (RNA-Seq) has proved particularly valuable, overcoming the barriers associated with its obligate biotrophic nature. This includes the application of RNA-Seq approaches to study *Pst* and wheat gene expression dynamics over time and the *Pst* population composition through the use of a novel RNA-Seq based surveillance approach called “field pathogenomics”. As a dual RNA-Seq approach, the field pathogenomics technique also provides gene expression data from the host, giving new insight into host responses. However, this has created a wealth of data for interrogation.

**Results:** Here, we used the field pathogenomics approach to generate 538 new RNA-Seq datasets from *Pst*-infected field wheat samples, doubling the amount of transcriptomics data available for this important pathosystem. We then analysed these datasets alongside 66 RNA-Seq datasets from four *Pst* infection time-courses and 420 *Pst*-infected plant field and laboratory samples that were publicly available. A database of gene expression values for *Pst* and wheat was generated for each of these 1024 RNA-Seq datasets and incorporated into the development of the rust expression browser (<http://www.rust-expression.com>). This enables for the first time simultaneous ‘point-and-click’ access to gene expression profiles for *Pst* and its wheat host and represents the largest database of processed RNA-Seq datasets available for any of the three *Puccinia* wheat rust pathogens. We also demonstrated the utility of the browser through investigation of expression of putative *Pst* virulence genes over time and examined the host plants response to *Pst* infection.

(Continued on next page)

\* Correspondence: [Diane.Saunders@jic.ac.uk](mailto:Diane.Saunders@jic.ac.uk)

<sup>1</sup>John Innes Centre, Norwich Research Park, Norwich NR4 7UH, UK

Full list of author information is available at the end of the article



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

(Continued from previous page)

**Conclusions:** The rust expression browser offers immense value to the wider community, facilitating data sharing and transparency and the underlying database can be continually expanded as more datasets become publicly available.

**Keywords:** RNA-Seq, expVIP, Gene expression browser, Wheat yellow rust, *Puccinia striiformis* f. sp. *tritici*, Transcriptomics, Open science

## Background

Transcriptomic studies that map fluctuations in the full complement of RNA transcripts, have revolutionized genome-wide gene expression analysis. For plant pathogens, the simultaneous analysis of host and pathogen transcriptomes has enabled many long-standing questions in plant pathology to be addressed particularly regarding how both organisms modulate gene expression at the host-pathogen interface [1]. This has provided new insight into the changes in gene expression profiles of both host and pathogen species. For instance, examination of the rice blast fungus *Magnaporthe oryzae* infecting rice plants identified a set of differentially expressed genes in both the host and the pathogen with more drastic expression changes in incompatible than compatible interactions [2]. Additionally, such analyses have revealed the importance of gene expression polymorphisms. For instance, the gain of virulence for the *Phytophthora infestans* EC-1 lineage on potato carrying *Rpi-vnt1.1* was shown to be due to lack of expression of the corresponding effector *Avrvnt1* [3]. Hence, RNA-based sequencing (RNA-Seq) is being increasingly applied to study the plant-microbe interface, providing an unbiased quantification of expression levels of transcripts that is relatively inexpensive, highly sensitive, and provides high-throughput, high resolution data.

For the wheat yellow (stripe) rust pathogen (*Puccinia striiformis* f. sp. *tritici*, *Pst*) the application of RNA-Seq approaches has proved particularly valuable, overcoming the barriers associated with its obligate biotrophic nature. For instance, evaluating gene expression in wheat plants infected by *Pst* and the powdery mildew pathogen *Blumeria graminis* f. sp. *tritici* (*Bgt*), identified commonalities and differences in the metabolic pathways that were differentially expressed in response to infection through an EST-based approach [4]. Another study, evaluating host responses throughout a time-course of *Pst* infection identified temporally coordinated waves of expression of immune response regulators in wheat that varied in susceptible and resistant interactions [5]. Furthermore, as a pathogen of global concern, an RNA-Seq based surveillance approach was developed for *Pst* called “field pathogenomics” that has been used to study its population dynamics at an unprecedented resolution [6]. The application of this methodology in the UK

uncovered recent changes in the population composition of *Pst*, whilst also revealing varietal and temporal associations of specific *Pst* races (pathotypes) that can help inform disease management [6, 7]. As a dual RNA-Seq approach applied directly to *Pst* infected leaf samples it also provides gene expression data from the host side of the interaction giving new insight into host responses [8]. These approaches generate a wealth of RNA-Seq data that is exceptionally valuable but difficult for those without specialist skills to access, which also inhibits reproducibility of transcriptomic studies.

Currently, the standard for open sharing of RNA-Seq data is to ensure raw reads are deposited in public repositories such as NCBI's Sequence Read Archive (SRA) [9]. However, utilising this data requires specialist bioinformatic expertise and often the use of high-performance computing systems. To overcome this, a series of gene expression browsers have been developed to enable interactive exploration of expression data [10–12]. However, the amount of data included within these databases for *Pst* is limited. The recently released fungi.guru transcriptomic database contains data for *Pst* gene expression from a limited number of samples, however it does not include the large number of field samples currently available or expression profiles for the wheat host [13]. Evaluation of gene expression levels in the wheat host can be undertaken separately using the wheat expression browser; an interactive gene expression browser that uses the RNA-Seq data analysis and visualisation platform expVIP (expression Visualisation and Integration Platform) [14]. However, although this browser hosts a number of RNA-Seq datasets from *Pst*-infected wheat tissue, this data has only been aligned to the wheat host transcriptome, inhibiting the exploration of gene expression profiles on the pathogen side of the interaction. For wheat, the expVIP browser has been extremely useful in providing an open access interface for the visualisation of RNA-Seq datasets. This has been instrumental in improving the understanding of the role of a variety of different wheat genes, such as the iron transporter *TaVIT2* and its potential role in biofortification [15] and the role of *TEOSINTE BRANCHED1* in the regulation of inflorescence architecture and development [16]. As the underlying software is also publicly available [17], an instance was recently developed to support

analysis of fruit development for a wild blackberry species (*Rubus genevieri*) and cultivated red raspberry (*Rubus idaeus* cv. *prestige*) [18]. However, it has yet to be specifically applied to support analysis of plant-microbe interactions.

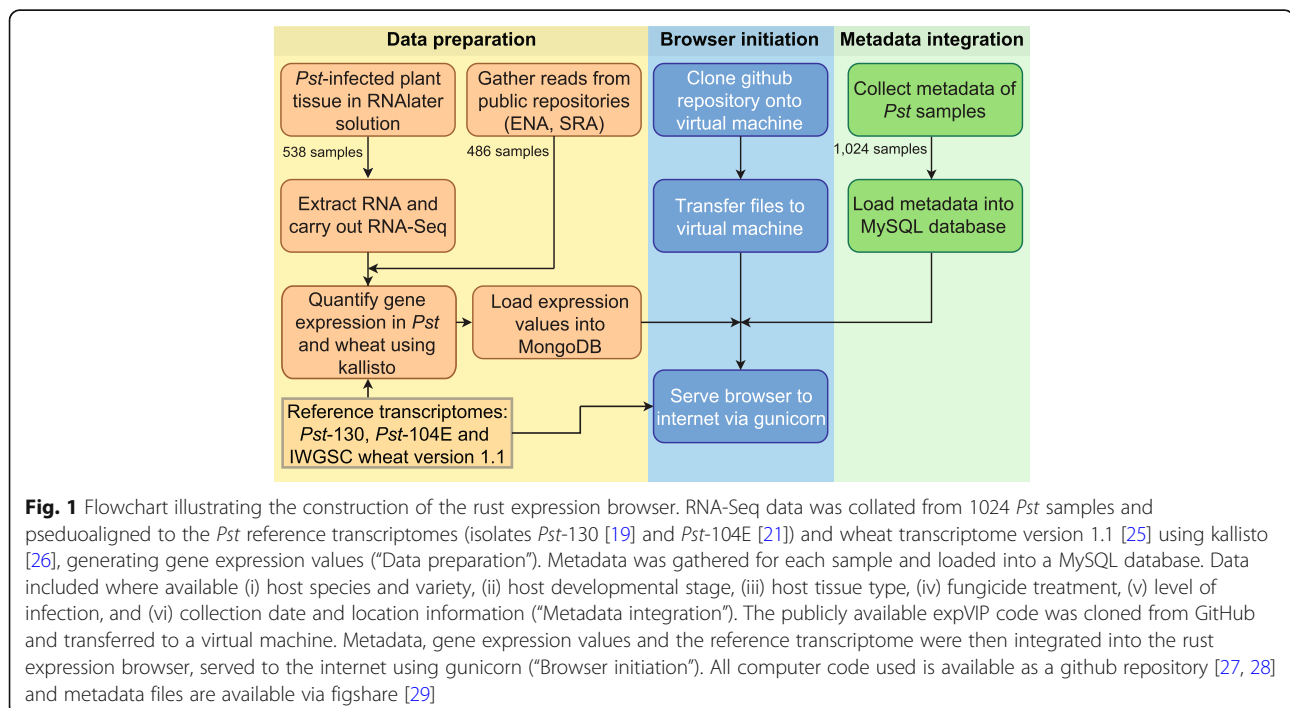
Here we present the first instance of a gene expression browser using the expVIP software that enables simultaneous exploration of both host and pathogen gene expression profiles. Focused on *Pst*, in this initial release we collated and processed 958 RNA-Seq datasets from use of the field pathogenomics methodology and 66 RNA-Seq datasets from *Pst* infection time course experiments for incorporation into the rust expression browser. With 538 of these RNA-Seq datasets generated herein, this has doubled the amount of RNA-Seq data available for this pathosystem and represents the largest collection of processed RNA-Seq datasets available for any of the three wheat rust pathogens. Using our new browser, the underlying database of gene expression values can be easily accessed for both *Pst* and its wheat host under an array of experimental conditions and across developmental stages. We show the utility of the browser for the analysis of putative virulence genes from the pathogen and the response of the host plant to *Pst* infection. This illustrates the immense value of analysing a broad set of RNA-Seq data to provide insight into gene expression regulation during host-pathogen interactions.

## Construction and content

### Generating RNA-Seq data and its incorporation into the rust expression browser

To generate data for incorporation into the *Pst* expression browser we first used a set of 538 *Pst*-infected plant samples that were collected across 30 countries from 2014 to 2018 (Supplementary Table S1). *Pst*-infected wheat leaf samples were collected and initially stored in RNAlater™ solution to preserve nucleic acid integrity (Thermo Fisher Scientific, United Kingdom) as previously described [6]. Total RNA was extracted from each sample, quality checked using an Agilent 2100 Bioanalyzer (Agilent Technologies, United Kingdom) and sequencing libraries prepared using an Illumina TruSeq RNA Sample Preparation Kit (Illumina, United Kingdom). Samples were subjected to RNA-Seq analysis using Illumina short read sequencing either at the Earlham Institute (United Kingdom; until April 2017) or Genewiz (USA; since April 2017) using the Illumina HiSeq 2500.

To further expand this initial dataset, we also identified a total of 486 RNA-Seq datasets from four previously published *Pst* infection time-courses (66 datasets) and *Pst*-infected plant field samples (420 datasets) [5–7, 19–24]. Each of the 1024 transcriptomic datasets were independently pseudoaligned to two *Pst* reference transcriptomes: *Pst* isolate *Pst*-130 [19] and isolate *Pst*-104E [21]. As the vast majority of samples (1004) were from *Pst*-infected wheat tissue, these datasets included both wheat and pathogen-derived reads, thereby samples were

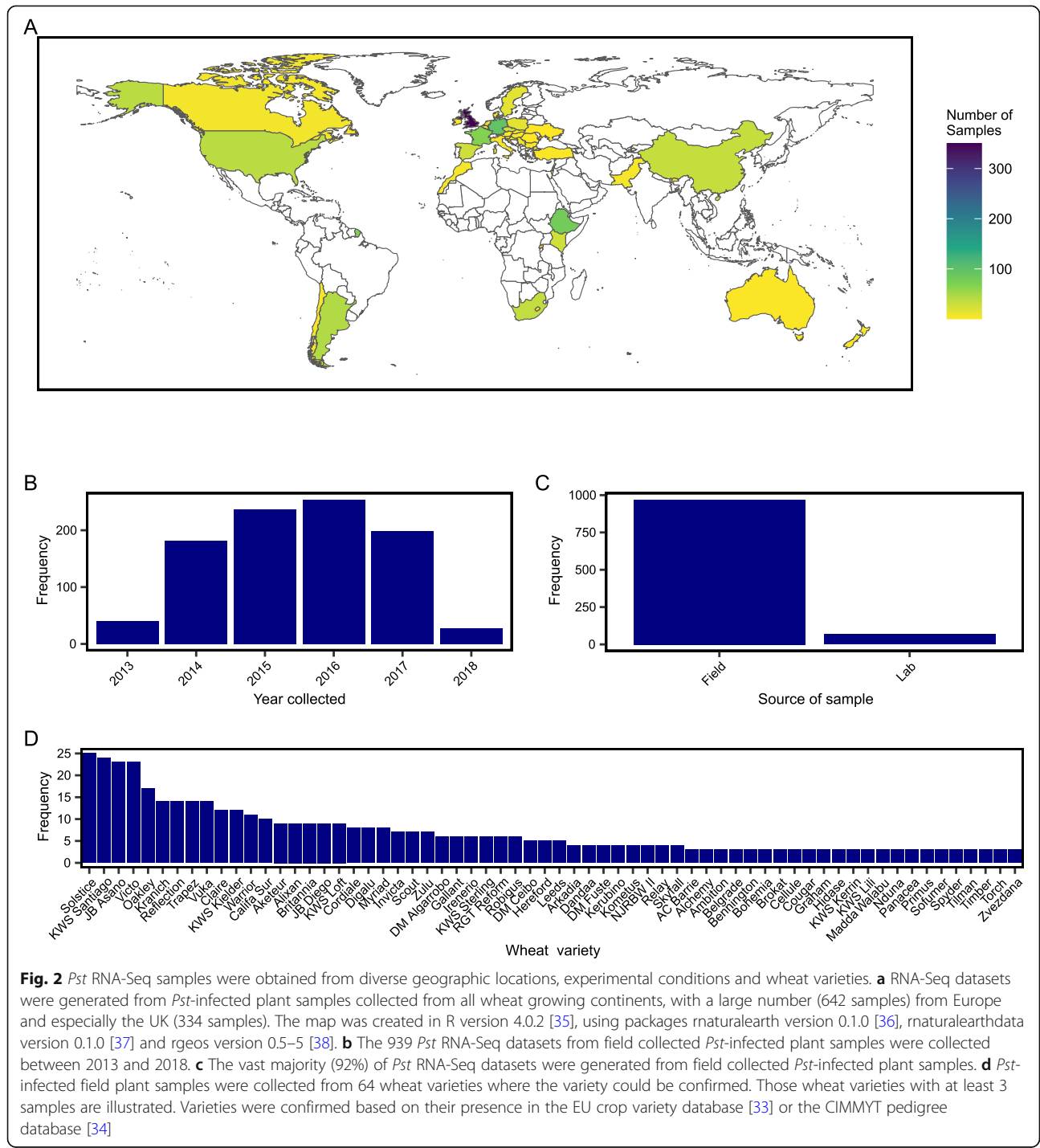


also pseudoaligned to version 1.1 of the wheat transcriptome [25]. To facilitate the processing of large numbers of RNA-Seq datasets, the kallisto aligner version 0.42.3 is used in the expVIP framework as an ultra-fast algorithm that was specifically developed for processing large-scale RNA-Seq datasets of short reads for gene expression quantification [26]. Transcript abundances were determined from the kallisto pseudoalignments and

incorporated into a MongoDB database for integration into the rust expression browser (Fig. 1).

**Construction of the rust expression browser**

The rust expression browser makes use of a modified version of the expVIP code previously used for the wheat expression browser [14] available as a github repository [30]. This repository was cloned onto a virtual machine running



**Fig. 2** *Pst* RNA-Seq samples were obtained from diverse geographic locations, experimental conditions and wheat varieties. **a** RNA-Seq datasets were generated from *Pst*-infected plant samples collected from all wheat growing continents, with a large number (642 samples) from Europe and especially the UK (334 samples). The map was created in R version 4.0.2 [35], using packages *naturalearth* version 0.1.0 [36], *naturalearthdata* version 0.1.0 [37] and *rgeos* version 0.5–5 [38]. **b** The 939 *Pst* RNA-Seq datasets from field collected *Pst*-infected plant samples were collected between 2013 and 2018. **c** The vast majority (92%) of *Pst* RNA-Seq datasets were generated from field collected *Pst*-infected plant samples. **d** *Pst*-infected field plant samples were collected from 64 wheat varieties where the variety could be confirmed. Those wheat varieties with at least 3 samples are illustrated. Varieties were confirmed based on their presence in the EU crop variety database [33] or the CIMMYT pedigree database [34]

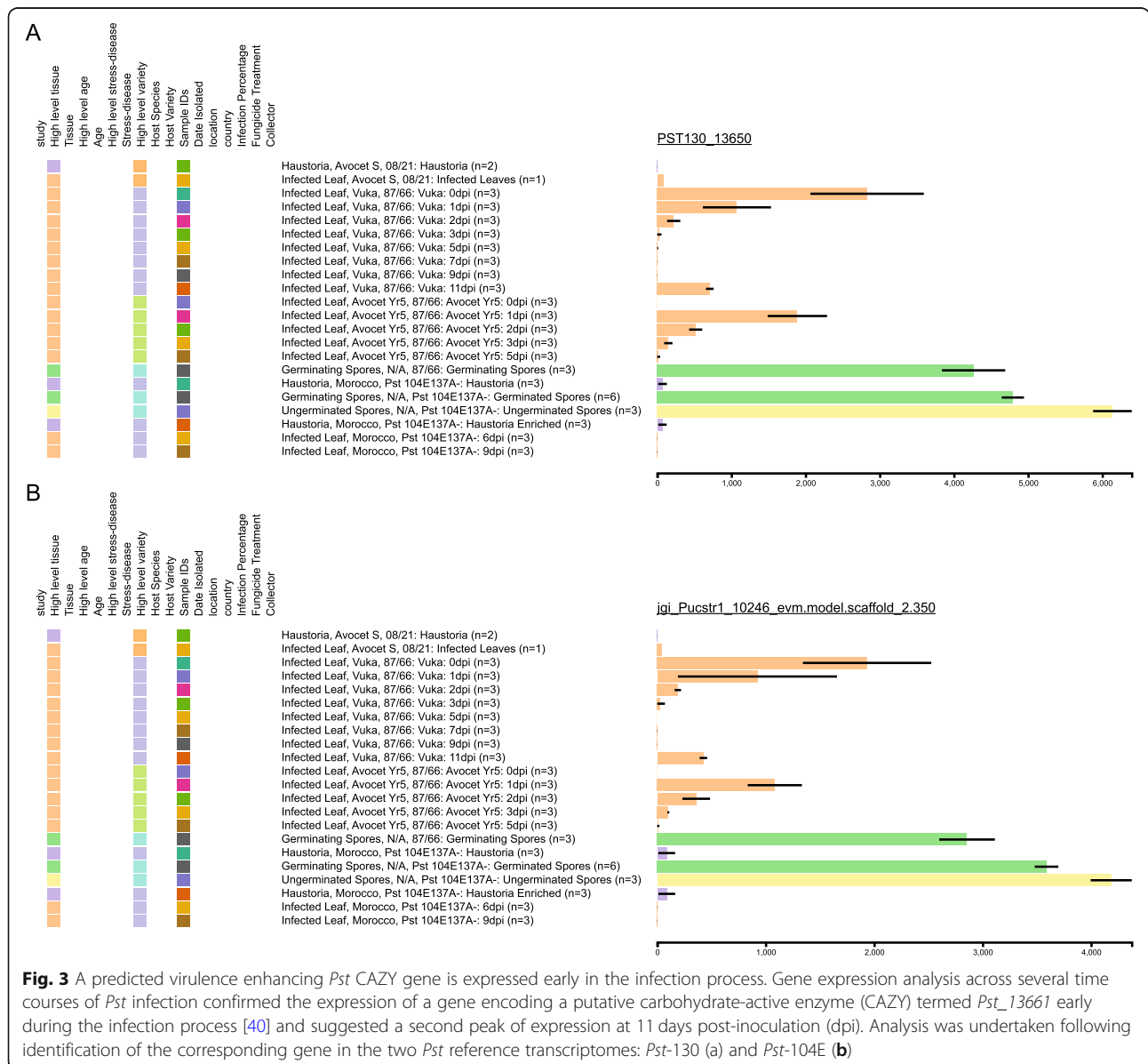
CentOS 7, kernel version 3.10.0-1062.12.1.el7.x86\_64. Metadata information for the samples was loaded into a MySQL database client version 5.5.68-MariaDB and expression values generated using kallisto [26] were loaded into a MongoDB database version 4.0.22 (Fig. 1). Transcript abundances, alongside the metadata and reference transcriptomes, was then integrated into the expVIP database instance for *Pst* [31]. This instance was then made accessible to web browsers through the use of gunicorn v5.5.3.

### Utility and discussion

#### The rust expression browser allows exploration of a broad array of *Pst*-based RNA-Seq datasets

The inclusion of detailed metadata alongside each *Pst* RNA-Seq dataset within the expVIP framework enables

users to easily group data and filter based on categories of interest (Fig. 1; Supplementary Figure S1). To maximise the value of the interface, metadata was gathered for each sample that included where available (i) host species and variety, (ii) host developmental stage, (iii) host tissue type, (iv) fungicide treatment, (v) level of infection, and (vi) collection date and location information. Among the 1024 transcriptomic datasets, 939 represented *Pst*-infected field samples that were collected across all wheat growing continents between 2013 and 2018, with a large number (642 samples) from Europe and especially the UK (334 samples; Fig. 2a). Over 92% of the 939 *Pst*-infected field samples were collected between 2014 and 2017 (Fig. 2b-c), which follows a period of change in the *Pst* population dynamics in Europe and



**Fig. 3** A predicted virulence enhancing *Pst* CAZY gene is expressed early in the infection process. Gene expression analysis across several time courses of *Pst* infection confirmed the expression of a gene encoding a putative carbohydrate-active enzyme (CAZY) termed *Pst\_13661* early during the infection process [40] and suggested a second peak of expression at 11 days post-inoculation (dpi). Analysis was undertaken following identification of the corresponding gene in the two *Pst* reference transcriptomes: *Pst*-130 (a) and *Pst*-104E (b)

hence a flurry of *Pst* surveillance activities and sample collection [32]. For samples where the wheat variety was recorded, this was cross referenced with the EU plant variety database [33] and CIMMYT variety pedigree database [34]. If a variety could be confirmed in either database, it was also included in the browser metadata (Fig. 2d).

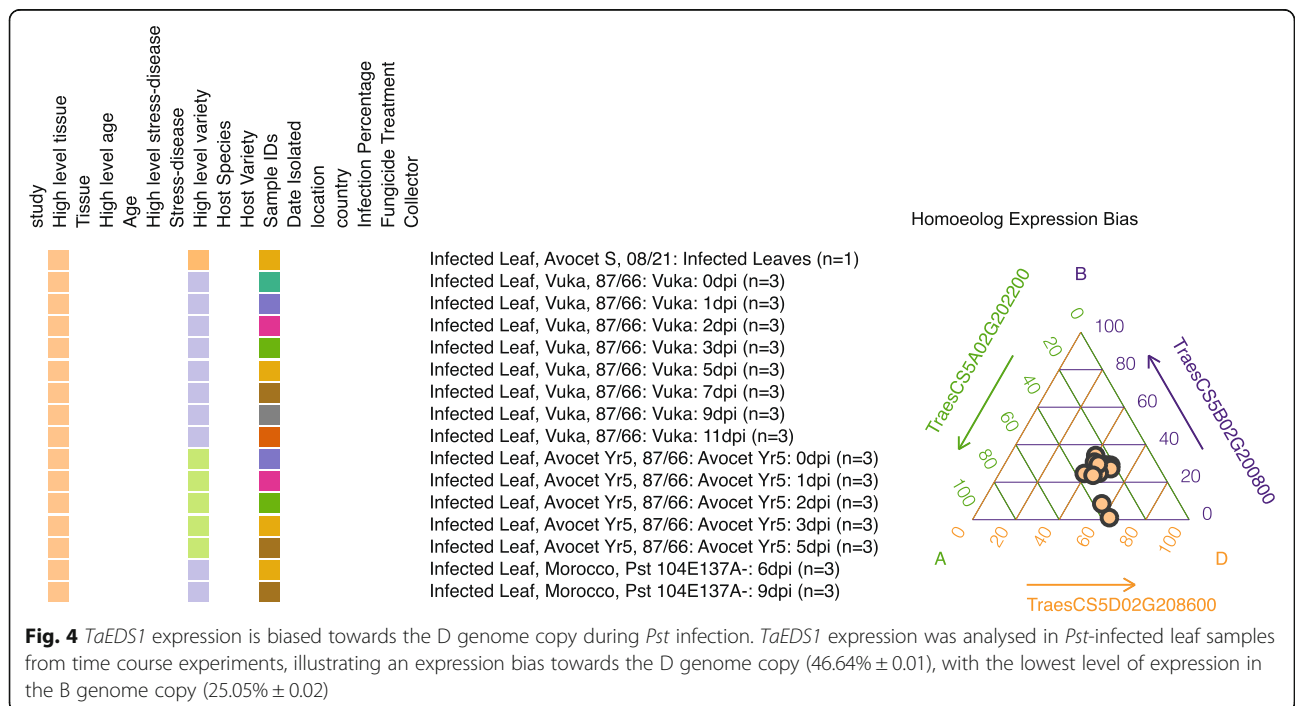
**Simultaneous analysis of multiple RNA-Seq experiments can provide new insight into the expression dynamics of *Pst* virulence factors**

To explore the utility of the rust expression browser, we examined several genes of interest within the browser interface. For *Pst*, we focused on evaluating the expression of a gene (*Pst\_13661*) that was recently reported to encode a putative carbohydrate-active enzyme (CAZY) that are known to be conserved across biotrophic fungi [39]. It was reported that *Pst\_13661* is able to suppress chitin-induced cell death and, through RT-qPCR analysis, to be highly induced early in infection progression, particularly at 12- and 48-h post inoculation (hpi), with a reduction at 72 and 96 hpi [40]. To evaluate *Pst\_13661* expression across all four time-courses of *Pst* infection within the rust expression browser [5, 19–21], we first identified the corresponding gene from the two *Pst* reference genomes using BLASTn [41, 42] conducted via implementation of SequenceServer version 1.0.12 [43] on the main page of the browser (PST130\_13650 and jgi\_Pucstr1\_10246\_evm.model.scaffold\_2.350; Fig. 3). In accordance with the RT-qPCR analysis, high levels of

expression were detected in all cases early in the infection process that was abolished 3 days post-inoculation (dpi). However, within the expression browser we were also able to investigate expression in specific *Pst* developmental stages and across the full infection process in multiple independent experiments. This analysis showed that the gene was highly expressed in ungerminated and germinated urediniospores, had low levels of expression in isolated haustoria, and increased in expression at 11 days post inoculation (dpi) to a level similar to that observed between 1 and 2 dpi. This may suggest a function for this gene later in the infection process or reflect its high level of expression in urediniospores that would begin formation by 11 dpi. The ability to rapidly assess gene expression across an array of time-points, *Pst* developmental stages and experiments provides new insight into the expression of *Pst\_13661* without the need for further lengthy and labour-intensive RT-qPCR analysis.

**Gene expression analysis of wheat responses to *Pst* infection**

As the vast majority of *Pst* RNA-Seq datasets incorporated in the browser were generated from *Pst*-infected wheat tissue, gene expression analysis can also be undertaken on the wheat host during *Pst* infection. To illustrate this, we examined the Enhanced Disease Susceptibility 1 (*EDS1*) gene homologues in wheat. *EDS1* was first defined in *Arabidopsis thaliana* and is essential for *R*-gene mediated and basal defence responses to



**Fig. 4** *TaEDS1* expression is biased towards the D genome copy during *Pst* infection. *TaEDS1* expression was analysed in *Pst*-infected leaf samples from time course experiments, illustrating an expression bias towards the D genome copy (46.64% ± 0.01), with the lowest level of expression in the B genome copy (25.05% ± 0.02)

biotrophic pathogens such as *Hyaloperonospora arabidopsidis* (formerly *Peronospora parasitica*) [44, 45]. Recently, the homologous genes in wheat have been identified as being important in the response of wheat to infection with the powdery mildew pathogen *Bgt* [46]. As a polyploid, bread wheat (*Triticum aestivum*) typically contains three copies of most genes with one each on the A, B and D chromosomes. It has been shown that the expVIP pipeline is able to accurately distinguish the

expression of the three homeologues [14]. Hence, using the expVIP-derived rust expression browser we analysed the expression of the three homeologues of *EDSI* in wheat during *Pst* infection across the samples from four infection time-courses that contained wheat tissue. This analysis revealed that overall expression of the wheat homeologues of *EDSI* tended to be biased towards the D genome copy ( $46.64\% \pm 0.01$ ) with the expression of the B genome copy at the lowest level ( $25.05\% \pm 0.02$ ;



**Fig. 5** The pathogenicity related (*PR*) genes *PR1* and *PR5* were highly expressed during *Pst* infection. A subset of *Pst*-infected wheat field and laboratory samples was examined for expression of *PR1* (TraesCS5A02G183300), *PR2* (TraesCS5A02G017900), *PR3* (TraesCS2B02G125200), *PR5* (TraesCS3A02G517100) and *PR10* (TraesCS4D02G189200). Gene expression is presented as a heatmap and includes only those samples where the wheat variety could be confirmed and at least three entries were present in the browser

Fig. 4). This is in contrast to that reported for *Bgt*-infected wheat plants, where the highest level of expression was observed in the B genome copy and lowest in the D genome copy. This observation could lead to a greater understanding of the response of wheat to biotrophic pathogens through further analysis of the response of the *EDS1* genes to different pathogen species.

We also evaluated the expression of pathogenesis related (*PR*) genes across 939 *Pst*-infected field and 19 *Pst*-infected laboratory wheat samples. In wheat, *PR* gene expression has been reported to be cultivar and pathogen specific, with different *PR* gene expression patterns also associated with resistance to different *Puccinia* species [47]. We examined the expression of *PR1* (TraesCS5A02G183300), *PR2* (TraesCS5A02G017900), *PR3* (TraesCS2B02G125200), *PR5* (TraesCS3A02G517100) and *PR10* (TraesCS4D02G189200) across all *Pst*-infected field samples where the variety had been confirmed and at least 3 entries were present in the browser (Fig. 5). *PR1* and *PR5* were the most highly expressed across all samples, whilst *PR3* showed the lowest expression level. However, we also found a large amount of variation in the expression of each gene across different varieties, potentially reflecting a difference in their response to *Pst* infection.

## Conclusions

Here we report the development of a novel database and tool for simultaneous ‘point and click’ access to gene expression profiles for *Pst* and its wheat host during the infection process. With 1024 *Pst* samples from an array of developmental stages, experimental conditions and wheat varieties, this browser provides rapid access to gene expression values that can be used as an alternative to lengthy RT-qPCR assays where appropriate. With the largest database of processed RNA-Seq datasets available for any of the three wheat rust pathogens, the rust expression browser offers immense value to the wider community. We have shown how this browser can be used to provide new insight into the expression profiles for *Pst* virulence genes over time and the host plants response to *Pst* infection. As new RNA-Seq data becomes available this can easily be incorporated into the browser, continuing to enhance studies into the *Pst*-wheat interaction.

## Abbreviations

*Pst*: *Puccinia striiformis* f. sp. *tritici*; RNA-Seq: RNA-based sequencing; *Bgt*: *Blumeria graminis* f. sp. *tritici*; EST: Expressed sequence tag; NCBI: National Center for Biotechnology Information; SRA: NCBI Sequence Read Archive; expVIP: expression Visualisation and Integration Platform; CIMMYT: International Maize and Wheat Improvement Center; EU: European Union; CAZY: Carbohydrate-active enzyme; RT-qPCR: Quantitative reverse transcription PCR; BLAST: Basic Local Alignment Search Tool; *EDS1*: Enhanced Disease Susceptibility 1 gene; *PR*: Pathogenesis related genes

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-021-07488-3>.

**Additional file 1: Supplementary Table S1.** Metadata for the *Pst* samples used in this study.

**Additional file 2: Supplementary Figure S1.** Flexible filtering can be applied in the wheat expression browser. All expression values presented are for the candidate CAZY effector *Pst\_13661* in the *Pst*-130 genome (termed PST130\_13650) [40]. Illustration of data filtered to display only samples from a single study [22] (A) or data only from germinated and ungerminated urediniospores and purified haustoria (B).

## Acknowledgements

We thank Nicola Cook, Jessica Meades, Phoebe Davey, Laura Reese (JIC) and Hester van Schalkwyk (JIC and the University of Free State, South Africa) for technical assistance, all those that submitted *Pst*-infected plant samples to the UK cereal pathogen virulence survey (UKCPVS) between 2014 and 2018, all members of the BASF Regional Technical Teams (BASF SE), Richard Summers, Jan Panek and Marion Chapus (RAGT 2 N SAS), Richard Oliver, Belinda Cox and Fran Lopez-Ruiz (Curtin University), KWS European breeding teams, Aleksander Vojisavljevic (Institute of Field and Vegetable Crops, Serbia), James Brown (JIC), Jamie Peacher, David Bellamy and Luke Cox (Frontier Agriculture), David Hodson, Ayele Badebo and Bekele Abeyo (CIMMYT) and the team at Kulumsa (EIAR) for fantastic assistance with sample collection, Cristobal Uauy (JIC) for useful discussions regarding the implementation of the rust expression browser and the NBI Computing Infrastructure for Science (CIS) group.

## Authors' contributions

D.G.O.S. conceived the project; T.M.A. performed gene expression quantification, prepared metadata files and constructed the expression browser; T.S.G.O. performed the initial provision of the virtual machine and assisted with constructing the browser; R.H.R.-G. assisted with adaptations to the expVIP code; R.B.1, R.B.2, P.E.C., P.F., D.F., C.H., T.H., A.H., R.J., C.J., M.K., J.L., C.M.L., C.L., U.M., D.N., C.P. and R.W. managed and conducted the collection and preparation of *Pst*-infected plant samples, provided sample metadata and tested the functionality of the browser; T.M.A. and D.G.O.S. drafted the manuscript; all authors reviewed and approved the final manuscript.

## Funding

This project was predominantly supported by grants from the European Union's Horizon 2020 research and innovation programme under grant agreement No 773311 (RustWatch) and an Industrial Partnership Award (BB/M025519/1) from the UK Biological Sciences Research Council (BBSRC). Additional funding was provided to D.G.O.S. by the European Research Council (no. 715638), the BBSRC Institute Strategic Programmes BB/P016855/1 and BB/P012574/1, and the John Innes Foundation. Further financial contributions, assistance with sample and metadata collection, testing of the functionality of the browser and review of the manuscript text was provided by the Agriculture and Horticultural Development Board (AHDB), BASF Plc, Limagrain UK Ltd., RAGT 2 N SAS., KWS UK Ltd., Syngenta UK Ltd. and Masstock Arable (UK) Ltd. (trading as Agrii).

## Availability of data and materials

All raw sequence data used in this study are available at the European Nucleotide Archive (ENA): PRJEB39201, PRJEB33109, PRJEB31334, PRJEB15280 and PRJEB12497; and the NCBI SRA: PRJNA256347, PRJNA181960, PRJNA176472 and PRJNA396589 (Supplementary Table S1). All computer code used in this analysis are available as github repositories [27, 28] and metadata files are available via figshare [29].

The rust expression browser is available at: <http://www.rust-expression.com>. JavaScript should be enabled; we recommend use of Chrome, Firefox, Safari and Edge web browsers for optimal performance.

## Declarations

### Ethics approval and consent to participate

Not applicable.



**Consent for publication**

Not applicable.

**Competing interests**

The authors declare that they have no competing interests.

**Author details**

<sup>1</sup>John Innes Centre, Norwich Research Park, Norwich NR4 7UH, UK. <sup>2</sup>RAGT Seeds Ltd, Grange Road, Ickleton, Essex CB10 1TA, UK. <sup>3</sup>BASF SE, Agricultural Centre, Limburgerhof, Germany. <sup>4</sup>INTA EEA Bordenave, Bordenave, 8187 Buenos Aires, Argentina. <sup>5</sup>Limagrain UK Ltd, Woolpit IP30 9UP, UK. <sup>6</sup>Syngenta Seeds Ltd, Hill Farm Road, Cambridgeshire CB22 4QT, UK. <sup>7</sup>Elsoms Wheat Ltd, Spalding, Lincolnshire PE11 1QG, UK. <sup>8</sup>Lantmännen Lantbruk, Svalöv, Sweden. <sup>9</sup>NIAB, Cambridge CB3 0LE, UK. <sup>10</sup>Institute of Field and Vegetable Crops, Novi Sad, Serbia. <sup>11</sup>DSV United Kingdom Ltd, Banbury, Oxfordshire OX17 1FE, UK. <sup>12</sup>KWS UK Limited, Hertfordshire SG8 7RE, UK. <sup>13</sup>Frontier Agriculture, Witham St Hughs, Lincolnshire LN6 9TN, UK. <sup>14</sup>Hazera Seeds Ltd., Berurim M.P. Shiknim, 7983700 Tel Aviv-Yafo, Israel. <sup>15</sup>Agricultural Institute Osijek, Osijek, Croatia. <sup>16</sup>Masstock Arable (UK) Ltd. (trading as Agrii), Andoversford, Gloucestershire GL54 4LZ, UK. <sup>17</sup>Kenya Agricultural and Livestock Research Organization, Njoro, Nakuru, Kenya.

Received: 20 November 2020 Accepted: 26 February 2021

Published online: 09 March 2021

**References**

- Naidoo S, Visser EA, Zwart L, du Toit Y, Bhaduria V, Shuey LS. Dual RNA-Sequencing to Elucidate the Plant-Pathogen Duel. *Curr Issues Mol Biol.* 2018;27:127–42. <https://doi.org/10.21775/cimb.027.127>.
- Kawahara Y, Oono Y, Kanamori H, Matsumoto T, Itoh T, Minami E. Simultaneous RNA-seq analysis of a mixed transcriptome of rice and blast fungus interaction. *PLoS One.* 2012;7:e49423. <https://doi.org/10.1371/journal.pone.0049423>.
- Pais M, Yoshida K, Giannakopoulou A, Pel MA, Cano LM, Oliva RF, et al. Gene expression polymorphism underpins evasion of host immunity in an asexual lineage of the Irish potato famine pathogen. *BMC Evol Biol.* 2018;18:93. <https://doi.org/10.1186/s12862-018-1201-6>.
- Zhang H, Yang Y, Wang C, Liu M, Li H, Fu Y, et al. Large-scale transcriptome comparison reveals distinct gene activations in wheat responding to stripe rust and powdery mildew. *BMC Genomics.* 2014;15:898. <https://doi.org/10.1186/1471-2164-15-898>.
- Dobon A, Bunting DCE, Cabrera-Quio LE, Uauy C, Saunders DGO. The host-pathogen interaction between wheat and yellow rust induces temporally coordinated waves of gene expression. *BMC Genomics.* 2016;17:380. <https://doi.org/10.1186/s12864-016-2684-4>.
- Hubbard A, Lewis CM, Yoshida K, Ramirez-Gonzalez RH, de Vallavieille-Pope C, Thomas J, et al. Field pathogenomics reveals the emergence of a diverse wheat yellow rust population. *Genome Biol.* 2015;16:23. <https://doi.org/10.1186/s13059-015-0590-8>.
- Bueno-Sancho V, Persoons A, Hubbard A, Cabrera-Quio LE, Lewis CM, Corredor-Moreno P, et al. Pathogenomic analysis of wheat yellow rust lineages detects seasonal variation and host specificity. *Genome Biol Evol.* 2017;9:3282–96. <https://doi.org/10.1093/gbe/evx241>.
- Derevnina L, Michelmore RW. Wheat rusts never sleep but neither do sequencers: will pathogenomics transform the way plant diseases are managed? *Genome Biol.* 2015;16:44. <https://doi.org/10.1186/s13059-015-0615-3>.
- Leinonen R, Sugawara H, Shumway M. International Nucleotide Sequence Database Collaboration. The sequence read archive. *Nucleic Acids Res.* 2011;39(Database issue):D19–21. <https://doi.org/10.1093/nar/gkq1019>.
- Kapushesky M, Emam I, Holloway E, Kurnosov P, Zorin A, Malone J, et al. Gene expression atlas at the European bioinformatics institute. *Nucleic Acids Res.* 2010;38(Database issue):D690–8. <https://doi.org/10.1093/nar/gkp936>.
- Zhang Z, Townsend JP. The filamentous fungal gene expression database (FFGED). *Fungal Genet Biol.* 2010;47:199–204. <https://doi.org/10.1016/j.fgb.2009.12.001>.
- Stajich JE, Harris T, Brunk BP, Brestelli J, Fischer S, Harb OS, et al. FungiDB: an integrated functional genomics database for fungi. *Nucleic Acids Res.* 2012;40(Database issue):D675–81. <https://doi.org/10.1093/nar/gkr918>.
- Lim JJJ, Koh J, Moo JR, Villanueva EMF, Putri DA, Lim YS, et al. Fungi.guru: comparative genomic and transcriptomic database for the Fungi kingdom. *BioRxiv.* 2020. <https://doi.org/10.1101/2020.06.26.174581>.
- Borrill P, Ramirez-Gonzalez R, Uauy C. expVIP: a customizable RNA-seq data analysis and visualization platform. *Plant Physiol.* 2016;170:2172–86. <https://doi.org/10.1104/pp.15.01667>.
- Connorton JM, Jones ER, Rodriguez-Ramiro I, Fairweather-Tait S, Uauy C, Balk J. Wheat vacuolar iron transporter tavt2 transports fe and mn and is effective for biofortification. *Plant Physiol.* 2017;174:2434–44. <https://doi.org/10.1104/pp.17.00672>.
- Dixon LE, Greenwood JR, Bencivenga S, Zhang P, Cockram J, Mellers G, et al. TEOSINTE BRANCHED1 regulates inflorescence architecture and development in bread wheat (*Triticum aestivum*). *Plant Cell.* 2018;30:563–81. <https://doi.org/10.1105/tpc.17.00961>.
- Ramirez-Gonzalez RH. GitHub - Uauy-lab/expvip-web: expVIP. EXPression and Visualisation Platform <https://github.com/Uauy-Lab/expvip-web>. Accessed 7 Sep 2020.
- Thole V, Bassard J-E, Ramirez-Gonzalez R, Trick M, Ghasemi Afshar B, Breitel D, et al. RNA-seq, de novo transcriptome assembly and flavonoid gene analysis in 13 wild and cultivated berry fruit species with high content of phenolics. *BMC Genomics.* 2019;20:995. <https://doi.org/10.1186/s12864-019-6183-2>.
- Cantu D, Govindarajulu M, Kozik A, Wang M, Chen X, Kojima KK, et al. Next generation sequencing provides rapid access to the genome of *Puccinia striiformis* f. sp. *tritici*, the causal agent of wheat stripe rust. *PLoS One.* 2011;6:e24230. <https://doi.org/10.1371/journal.pone.0024230>.
- Garnica DP, Upadhyaya NM, Dodds PN, Rathjen JP. Strategies for wheat stripe rust pathogenicity identified by transcriptome sequencing. *PLoS One.* 2013;8:e67150. <https://doi.org/10.1371/journal.pone.0067150>.
- Schwessinger B, Spersneider J, Cuddy WS, Garnica DP, Miller ME, Taylor JM, et al. A Near-Complete Haplotype-Phased Genome of the Dikaryotic Wheat Stripe Rust Fungus *Puccinia striiformis* f. sp. *tritici* Reveals High Interhaplotype Diversity. *MBio.* 2018;9. <https://doi.org/10.1128/mBio.02275-17>.
- WHP B, Visser B, Lewis CM, Adams TM, DGO S, Terefe T. et al, First Report of *Puccinia striiformis* f. sp. *tritici* , Causing Stripe Rust of Wheat, in Zimbabwe. *Plant Dis.* 2019;PDIS-07-19-1395. <https://doi.org/10.1094/PDIS-07-19-1395-PDN>.
- Radhakrishnan GV, Cook NM, Bueno-Sancho V, Lewis CM, Persoons A, Mitiku AD, et al. MARPLE, a point-of-care, strain-level disease diagnostics and surveillance tool for complex fungal pathogens. *BMC Biol.* 2019;17:65. <https://doi.org/10.1186/s12915-019-0684-y>.
- Corredor-Moreno P, Minter F, Davey PE, Wegel E, Kular B, Brett P, et al. The branched-chain amino acid aminotransferase TaBCAT1 modulates amino acid metabolism and positively regulates wheat rust susceptibility. *Plant Cell.* 2021. <https://doi.org/10.1093/plcell/koab049>.
- International Wheat Genome Sequencing Consortium (IWGSC), IWGSC RefSeq principal investigators, Appels R, Eversole K, Feuillet C, Keller B, et al. Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science.* 2018;361. <https://doi.org/10.1126/science.aar7191>.
- Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol.* 2016;34:525–7. <https://doi.org/10.1038/nbt.3519>.
- Adams T. SaundersLab/Rust\_Expression\_Browser: Version 1. Zenodo. 2020. <https://doi.org/10.5281/zenodo.4252377>.
- Olsson TSG. JIC-CSB/expvip-ansible: Release version 1.0.1. Zenodo. 2020. <https://doi.org/10.5281/zenodo.4252460>.
- Adams TM, Olsson TSG, González RHR, Saunders DGO. Files for building the rust expression browser. Figshare. 2020. <https://doi.org/10.6084/m9.figshare.13177163>.
- Ramirez-Gonzalez RH, Afshar BG, Adams T. TMAAdams/expvip-web: version 1. Zenodo. 2020. <https://doi.org/10.5281/zenodo.4252001>.
- Adams TM, Olsson TSG, Ramirez Gonzalez RH, Saunders DGO. Rust Expression Browser <http://www.rust-expression.com/>. Accessed 27 Oct 2020.
- Hansen JG, Lassen P. Managing global crop disease data. Proceedings of the EFITA-WCCA-CIGR Conference: Sustainable Agriculture Through ICT Innovation; 2013.
- European Commission. EU Plant variety database (v.3.2.1). [https://ec.europa.eu/food/plant/plant\\_propagation\\_material/plant\\_variety\\_catalogues\\_databases/search/public/index.cfm?event=SearchVariety&ctl\\_type=A&species\\_id=](https://ec.europa.eu/food/plant/plant_propagation_material/plant_variety_catalogues_databases/search/public/index.cfm?event=SearchVariety&ctl_type=A&species_id=)

- 258&variety\_name=&listed\_in=0&show\_current=on&show\_deleted=  
Accessed 7 Aug 2020.
34. CIMMYT. CIMMYT pedigree database. 2017. [http://www.wheatpedigree.net/sort/index?search=&\\_action\\_list=Search](http://www.wheatpedigree.net/sort/index?search=&_action_list=Search). Accessed 7 Aug 2020.
  35. R Core Team. R: A language and environment for statistical computing. Vienna: Computer software. R Foundation for Statistical Computing; 2020.
  36. South A. World map data from natural earth [R package rnaturalearth version 0.1.0]. Computer software. CRAN; 2017.
  37. South A. World vector map data from natural earth used in "rnaturalearth" [R package rnaturalearthdata version 0.1.0]. Computer software. CRAN; 2017.
  38. Bivand R, Rundel C. Interface to geometry engine - open source ('GEOS') [R package rgeos version 0.5–5]. Computer software. CRAN; 2020.
  39. Duplessis S, Cuomo CA, Lin Y-C, Aerts A, Tisserant E, Veneault-Fourrey C, et al. Obligate biotrophy features unraveled by the genomic analysis of rust fungi. *Proc Natl Acad Sci U S A*. 2011;108:9166–71. <https://doi.org/10.1073/pnas.1019315108>.
  40. Xu Q, Wang J, Zhao J, Xu J, Sun S, Zhang H, et al. A polysaccharide deacetylase from *Puccinia striiformis* f. sp. *tritici* is an important pathogenicity gene that suppresses plant immunity. *Plant Biotechnol J*. 2020;18:1830–42. <https://doi.org/10.1111/pbi.13345>.
  41. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol*. 1990;215:403–10. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2).
  42. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. *BMC Bioinformatics*. 2009;10:421. <https://doi.org/10.1186/1471-2105-10-421>.
  43. Priyam A, Woodcroft BJ, Rai V, Moghul I, Munagala A, Ter F, et al. Sequenceserver: a modern graphical user interface for custom BLAST databases. *Mol Biol Evol*. 2019;36:2922–4. <https://doi.org/10.1093/molbev/msz185>.
  44. Parker JE, Holub EB, Frost LN, Falk A, Gunn ND, Daniels MJ. Characterization of eds1, a mutation in *Arabidopsis* suppressing resistance to *Peronospora parasitica* specified by several different *RPP* genes. *Plant Cell*. 1996;8:2033–46. <https://doi.org/10.1105/tpc.8.11.2033>.
  45. Falk A, Feys BJ, Frost LN, Jones JD, Daniels MJ, Parker JE. EDS1, an essential component of R gene-mediated disease resistance in *Arabidopsis* has homology to eukaryotic lipases. *Proc Natl Acad Sci U S A*. 1999;96:3292–7. <https://doi.org/10.1073/pnas.96.6.3292>.
  46. Chen H, Shu H, Wang L, Zhang F, Li X, Ochola SO, et al. *Phytophthora* methylomes are modulated by 6mA methyltransferases and associated with adaptive genome regions. *Genome Biol*. 2018;19:181. <https://doi.org/10.1186/s13059-018-1564-4>.
  47. Zhang H, Qiu Y, Yuan C, Chen X, Huang L. Fine-tuning of PR genes in wheat responding to different *Puccinia* rust species. *J Plant Physiol Pathol*. 2018;06. <https://doi.org/10.4172/2329-955X.1000178>.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

