



**HAL**  
open science

## **Find Genomic, Genetic and Phenomic Data for Wheat Using the WheatIS Data Discovery Tool**

Michael Alaux, Mario Caccamo, Sarah Dyer, David Edwards, Raphaël Flores, Najwa Lakmouri, Cyril Pommier, Taner Z Sen

### ► **To cite this version:**

Michael Alaux, Mario Caccamo, Sarah Dyer, David Edwards, Raphaël Flores, et al.. Find Genomic, Genetic and Phenomic Data for Wheat Using the WheatIS Data Discovery Tool. PAG 32, Jan 2025, San Diego (CA), United States. <hal-05528830>

**HAL Id: hal-05528830**

**<https://hal.inrae.fr/hal-05528830v1>**

Submitted on 26 Feb 2026

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License





# ➤ Find Genomic, Genetic and Phenomic Data for Wheat Using the WheatIS Data Discovery Tool

Michael Alaux *et al.*




## > Intro



**Session**  
“Databases and online tools for displaying the anchored physical maps“  
-  
Wheat mapping and annotation data at URGI

Michaël Alaux

PAG 2010 San Diego  
IWGSC Physical Mapping Standard Protocols Workshop  
organized by Catherine Feuillet



- Genomics data well managed thanks to the coordination of the IWGSC
- Need to find the other types of data (germplasm, phenotyping, genotyping, etc.) hosted in databases around the world  
→ Wheat Information System



INRAE

Find Genomic, Genetic and Phenomic Data for Wheat Using the WheatIS Data Discovery Tool  
PAG32, 14 January 2025



## > Outline

- What is WheatIS and how to use it to discover data ?
  - The Wheat Initiative
  - The WheatIS Expert Working Group
  - Data standards
  - Data discovery
  - Sharing experiences



# COORDINATING GLOBAL WHEAT RESEARCH

THE WHEAT INITIATIVE

TH OF JANUARY 2025 - APPLY NOW TO BECOME A MEMBER - NEW MEMBERS CAN ONLY BE RECIEVED UNTIL THE 6TH OF DECEMBER AND

# > History

Coordinated by JKI (Germany)  
scientific leader: Peter Langridge



Coordinated by INRA (France)  
scientific leader: H el ene Lucas



← chair: Michael Alaux (INRAE-URGI)



← chair: Taner Sen (USDA-ARS)



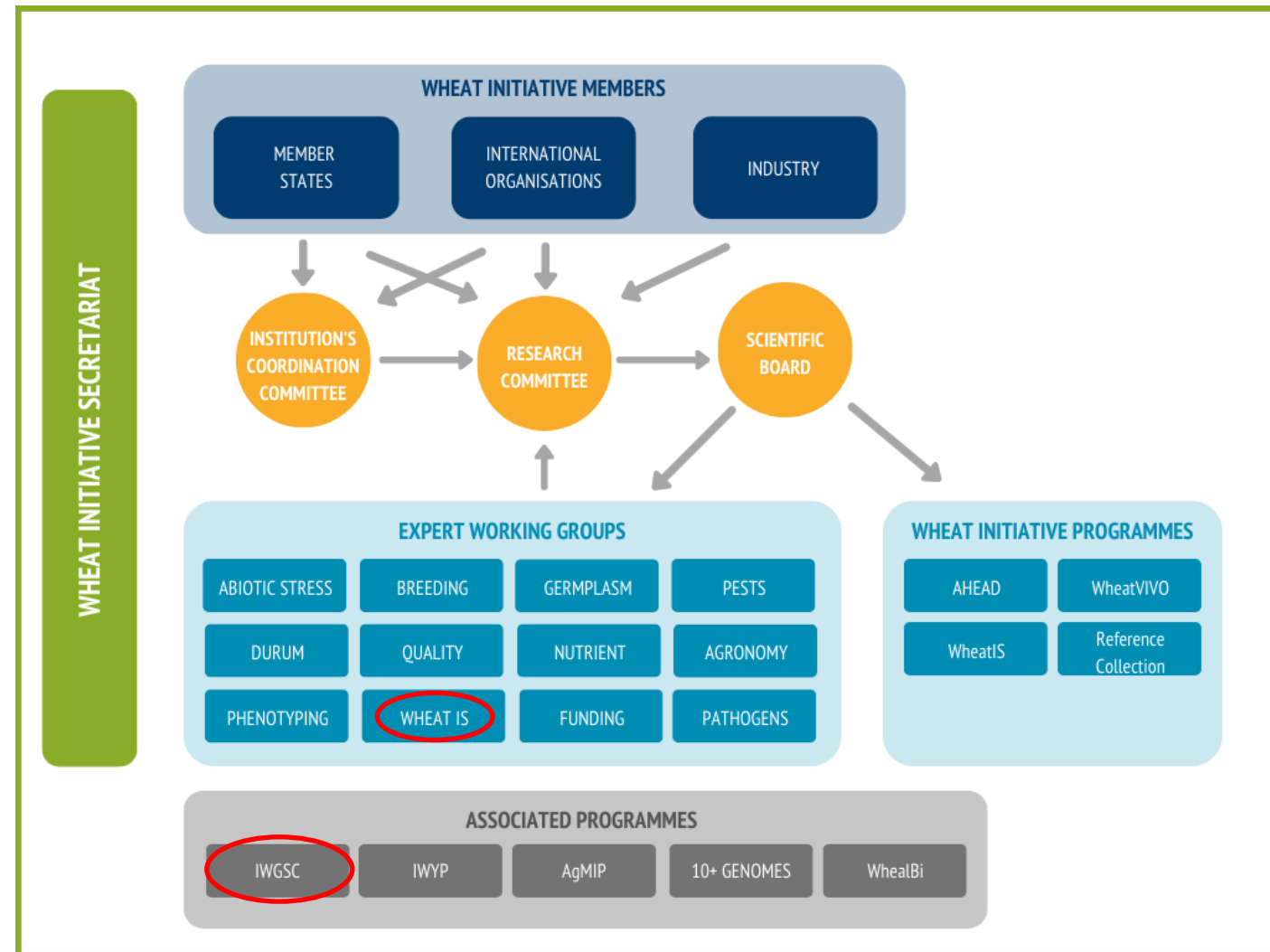
← Launch of the Wheat Information System Expert Working Group  
chair: Hadi Quesneville (INRA-URGI)



INRAE



# Organisation



# ➤ Strategic Research Agenda



Open Access Feature Paper Review

## Meeting the Challenges Facing Wheat Production: The Strategic Research Agenda of the Global Wheat Initiative

by Peter Langridge <sup>1,2,\*</sup> Michael Alaux <sup>3</sup>, Nuno Felipe Almeida <sup>4</sup> Karim Ammar <sup>5</sup>, Michael Baum <sup>6</sup>, Faouzi Bekkaoui <sup>7</sup>, Alison R. Bentley <sup>5</sup> Brian L. Beres <sup>8</sup> Bettina Berger <sup>9</sup> Hans-Joachim Braun <sup>5</sup>, Gina Brown-Guedira <sup>10</sup> Christopher James Burt <sup>11</sup>, Mario Jose Caccamo <sup>12</sup> Luigi Cattivelli <sup>13</sup> Gilles Charmet <sup>14</sup>, Peter Civián <sup>14</sup>, Sylvie Cloutier <sup>15</sup> Jean-Pierre Cohan <sup>16</sup>, Pierre J. Devaux <sup>17</sup> Fiona M. Doohan <sup>18</sup>, + Show full author list

<sup>1</sup> School of Agriculture, Food and Wine, Waite Campus, University of Adelaide, PMB1, Glen Osmond, SA 5064, Australia

<sup>2</sup> Wheat Initiative, JKI (Julius Kühn Institute), Federal Research Centre for Cultivated Plants, 14195 Berlin, Germany

<sup>3</sup> INRAE, URGI, Université Paris-Saclay, 78026 Versailles, France

<sup>4</sup> ASUR Plant Breeding, 60190 Estrées-Saint-Denis, France

<sup>5</sup> CIMMYT (International Maize and Wheat Improvement Center), Texcoco 56237, Mexico

<sup>6</sup> ICARDA (International Center for Agricultural Research in the Dry Areas), Rabat 10106, Morocco

<sup>7</sup> INRA (National Institute for Agricultural Research), Rabat 10090, Morocco

<sup>8</sup> AAFC (Agriculture & Agri-Food Canada), Lethbridge Research and Development Centre, Lethbridge, AB T1J 4B1, Canada

<sup>9</sup> Australian Plant Phenomics Facility, School of Agriculture, Food and Wine, University of Adelaide, Urrbrae, SA 5064, Australia

<sup>10</sup> USDA-ARS (United States Department of Agriculture—Agricultural Research Service), Plant Science Research, Raleigh, NC 27695, USA

+ Show full affiliation list

\* Author to whom correspondence should be addressed.

*Agronomy* **2022**, *12*(11), 2767; <https://doi.org/10.3390/agronomy12112767>

Received: 26 September 2022 / Revised: 28 October 2022 / Accepted: 29 October 2022 /

Published: 7 November 2022



INRAE

Find Genomic, Genetic and Phenomic Data for Wheat Using the WheatIS Data Discovery Tool

PAG32, 14 January 2025



<https://doi.org/10.3390/agronomy12112767>



# wheatIS

**WHEAT INFORMATION SYSTEM**



**INRAE**

Find Genomic, Genetic and Phenomic Data for Wheat Using the WheatIS Data Discovery Tool  
PAG32, 14 January 2025



<http://www.wheatis.org/>


## ➤ Build the community

- Building an expert working group
  - focus on recruiting diverse profiles covering important countries or geographical areas, institutions, interest groups and scientific fields for wheat research
- Seeking help from other communities
  - Wheat Data Interoperability Working Group was created as one of the Research Data Alliance working groups
- Surveys
  - usage of data standards in the wheat research community


OPINION ARTICLE Check for updates

### Building a successful international research community through data sharing: The case of the Wheat Information System (WheatIS)

[version 1; peer review: 2 approved, 1 approved with reservations]

Taner Z. Sen<sup>1</sup>, Mario Caccamo<sup>2</sup>, David Edwards<sup>3</sup>,  Hadi Quesneville<sup>4,5</sup>

[Author details](#)

 This article is included in the [Agriculture, Food and Nutrition gateway](#).

#### Abstract

The International Wheat Information System (WheatIS) Expert Working Group (EWG) was initiated in 2012 under the Wheat Initiative with a broad range of contributing organizations. The mission of the WheatIS EWG was to create an informational infrastructure, establish data standards, and build a single portal that allows search, retrieval, and display of globally distributed wheat data sets that are indexed in standard data formats at servers around the world. The web portal at [WheatIS.org](#) was released publicly in 2015, and by 2020, it expanded to 8 geographically-distributed nodes and around 20 organizations under its umbrella.

In this paper, we present our experience, the challenges we faced, and the answer we brought for establishing an international research community to build an informational infrastructure. Our hope is that our experience with building [wheatis.org](#) will guide current and future research communities to facilitate institutional and international challenges to create global tools and resources to help their respective scientific communities.

**ALL METRICS**

3000  
VIEWES

168  
DOWNLOADS

[Get PDF](#)  
[Get XML](#)  
[Cite](#)  
[Export](#)  
[Track](#)  
[Share](#)

<https://doi.org/10.12688/f1000research.23525.1>



## ➤ Build the community

- Fundings
  - Meetings (e.g. annual meeting at PAG)
  - Workshops
- A successful result
  - WheatIS data discovery tool
- Outreach
  - Trainings
  - Website and social media
  - Joint actions with other EWGs

OPINION ARTICLE

 Check for updates

### Building a successful international research community through data sharing: The case of the Wheat Information System (WheatIS)

[version 1; peer review: 2 approved, 1 approved with reservations]

Taner Z. Sen<sup>1</sup>, Mario Caccamo<sup>2</sup>, David Edwards<sup>3</sup>,  Hadi Quesneville<sup>4,5</sup>

 Author details



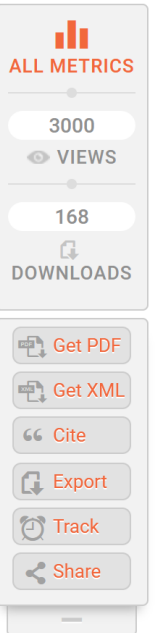
This article is included in the [Agriculture, Food and Nutrition](#) gateway.

#### Abstract

The International Wheat Information System (WheatIS) Expert Working Group (EWG) was initiated in 2012 under the Wheat Initiative with a broad range of contributing organizations. The mission of the WheatIS EWG was to create an informational infrastructure, establish data standards, and build a single portal that allows search, retrieval, and display of globally distributed wheat data sets that are indexed in standard data formats at servers around the world. The web portal at [WheatIS.org](#) was released publicly in 2015, and by 2020, it expanded to 8 geographically-distributed nodes and around 20 organizations under its umbrella.

In this paper, we present our experience, the challenges we faced, and the answer we brought for establishing an international research community to build an informational infrastructure. Our hope is that our experience with building [wheatis.org](#) will guide current and future research communities to facilitate institutional and international challenges to create global tools and resources to help their respective scientific communities.


<https://doi.org/10.12688/f1000research.23525.1>





ALL METRICS


3000  
VIEWES


168  
DOWNLOADS


 Get PDF

 Get XML

 Cite

 Export

 Track

 Share

# ➤ Expert working group

## MAIN RESEARCH AREA

- Information systems are recognized today as essential tools for federating scientific communities. They facilitate the sharing of data

## PURPOSE OF THE EWG

Creation and animation of a network of experts, collaborating to provide the scientific community with wheat genetic and genomic data.



The Wheat Information System EWG federates the diverse information sources (e.g. genetic, genomic, phenotypic, environmental), allowing their findability and integration.

A single entry point for the wheat community to find available data through a full text search engine (WheatIS Data Discovery), allowing searching data in all the wheat databases around the world.

## MAIN CHALLENGES IN THE FIELD

- To maintain effective communication with stakeholders and users to ensure proper tool development and data integration.
- To ensure up-to-date data in the WheatIS Data Discovery.
- To index new databases (especially from Asia) in the WheatIS Data Discovery.
- To improve the FAIRness of the data and metadata.

21 experts from Europe, America, Asia, Oceania including 6 early career researchers

The EWG is open to Early Career Researchers

WheatIS website  
<http://www.wheatis.org/>

WheatIS Data Discovery  
<https://urgi.versailles.inrae.fr/wheatis/>

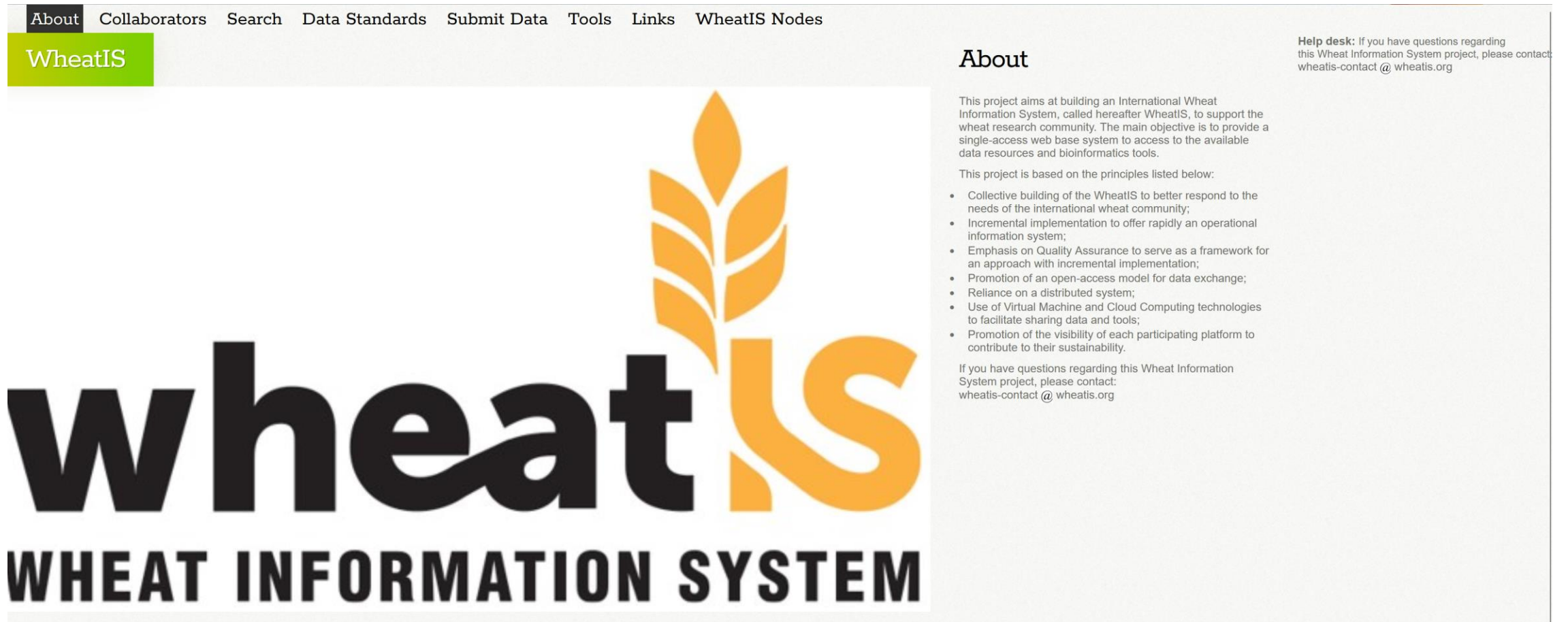
Webinar  
<https://tes.elixir-europe.org/materials/wheatis-datadiscovery-webinar>



INRAE



## > Website



The screenshot shows the WheatIS website homepage. At the top, there is a navigation menu with links: About, Collaborators, Search, Data Standards, Submit Data, Tools, Links, and WheatIS Nodes. The 'WheatIS' link is highlighted in a green box. Below the navigation is a large logo for 'wheatIS' where 'wheat' is in black and 'IS' is in orange, with a stylized wheat stalk above the 'IS'. Below the logo, the text 'WHEAT INFORMATION SYSTEM' is written in bold black capital letters. To the right of the logo, there is an 'About' section with a heading 'About' and a paragraph describing the project's aim to build an International Wheat Information System. Below this is a list of principles and a 'Help desk' section with contact information.

About Collaborators Search Data Standards Submit Data Tools Links WheatIS Nodes

WheatIS

### About

This project aims at building an International Wheat Information System, called hereafter WheatIS, to support the wheat research community. The main objective is to provide a single-access web base system to access to the available data resources and bioinformatics tools.

This project is based on the principles listed below:

- Collective building of the WheatIS to better respond to the needs of the international wheat community;
- Incremental implementation to offer rapidly an operational information system;
- Emphasis on Quality Assurance to serve as a framework for an approach with incremental implementation;
- Promotion of an open-access model for data exchange;
- Reliance on a distributed system;
- Use of Virtual Machine and Cloud Computing technologies to facilitate sharing data and tools;
- Promotion of the visibility of each participating platform to contribute to their sustainability.

If you have questions regarding this Wheat Information System project, please contact: [wheatis-contact@wheatis.org](mailto:wheatis-contact@wheatis.org)

**Help desk:** If you have questions regarding this Wheat Information System project, please contact: [wheatis-contact@wheatis.org](mailto:wheatis-contact@wheatis.org)



INRAE

Find Genomic, Genetic and Phenomic Data for Wheat Using the WheatIS Data Discovery Tool

PAG32, 14 January 2025



<http://www.wheatis.org/>



RESEARCH DATA ALLIANCE



**INRAE**

Find Genomic, Genetic and Phenomic Data for Wheat Using the WheatIS Data Discovery Tool

PAG32, 14 January 2025



# ➤ RDA Wheat Data Interoperability Working Group

## ■ Community driven methodology



OPINION ARTICLE

Check for updates

### REVISED Developing data interoperability using standards: A wheat community use case [version 2; peer review: 2 approved]

Esther Dzale Yeumo<sup>1</sup>, Michael Alaux <sup>2</sup>, Elizabeth Arnaud<sup>3</sup>, Sophie Aubin<sup>1</sup>, Ute Baumann<sup>4</sup>, Patrice Buche<sup>5</sup>, Laurel Cooper <sup>6</sup>, Hanna Ćwiek-Kuczyńska<sup>7</sup>, Robert P. Davey <sup>8</sup>, Richard Allan Fulss<sup>9</sup>, Clement Jonquet <sup>10,11</sup>, Marie-Angélique Laporte<sup>3</sup>, Pierre Larmande <sup>12,13</sup>, Cyril Pommier <sup>2</sup>, Vassilis Protonotarios <sup>14</sup>, Carmen Reverte <sup>15</sup>, Rosemary Shrestha<sup>9</sup>, Imma Subirats<sup>16</sup>, Aravind Venkatesan <sup>12</sup>, Alex Whan<sup>17</sup>, ✉ Hadi Quesneville <sup>2</sup>

Author details



This article is included in the [Agriculture, Food and Nutrition](#) gateway.

### Abstract

In this article, we present a joint effort of the wheat research community, along with data and ontology experts, to develop wheat data interoperability guidelines. Interoperability is the ability of two or more systems and devices to cooperate and exchange data, and interpret that shared information. Interoperability is a growing concern to the wheat scientific community, and agriculture in general, as the need to interpret the deluge of data obtained through high-throughput technologies grows. Agreeing on common data formats, metadata, and vocabulary standards is an important step to obtain the required data interoperability level in order to add value by encouraging data sharing, and subsequently facilitate the extraction of new information from existing and new datasets. During a period of more than 18 months, the RDA Wheat Data Interoperability Working Group (WDI-WG) surveyed the wheat research community about the use of data standards, then discussed and selected a set of recommendations based on consensual criteria. The recommendations promote standards for data types identified by the wheat research community as the most important for the coming years: nucleotide sequence variants, genome annotations, phenotypes, germplasm data, gene expression experiments, and physical maps. For each of these data types, the guidelines recommend best practices in terms of use of data formats, metadata standards and ontologies. In addition to the best practices, the guidelines provide examples of tools and implementations that are likely to facilitate the adoption of the recommendations. To maximize the adoption of the recommendations, the WDI-WG used a community-driven approach that involved the wheat research community from the start, took into account their needs and practices, and provided them with a framework to keep the recommendations up to date. We also report this approach's potential to be generalizable to other (agricultural) domains.

ALL METRICS

2561 VIEWS

455 DOWNLOADS

Get PDF

Get XML

Cite

Export

Track

Share

<https://doi.org/10.12688/f1000research.12234.2>



INRAE

Find Genomic, Genetic and Phenomic Data for Wheat Using the WheatIS Data Discovery Tool  
PAG32, 14 January 2025



# ➤ RDA Wheat Data Interoperability Working Group

- Dissemination
  - Website
  - Agroportal for wheat ontologies
- Adoption

OPINION ARTICLE Check for updates

**REVISED** Developing data interoperability using standards: A wheat community use case [version 2; peer review: 2 approved]

Esther Dzale Yeumo<sup>1</sup>, Michael Alaux <sup>2</sup>, Elizabeth Arnaud<sup>3</sup>, Sophie Aubin<sup>1</sup>, Ute Baumann<sup>4</sup>, Patrice Buche<sup>5</sup>, Laurel Cooper <sup>6</sup>, Hanna Ćwiek-Kupczyńska<sup>7</sup>, Robert P. Davey <sup>8</sup>, Richard Allan Fulss<sup>9</sup>, Clement Jonquet <sup>10,11</sup>, Marie-Angélique Laporte<sup>3</sup>, Pierre Larmande <sup>12,13</sup>, Cyril Pommier <sup>2</sup>, Vassilis Protonotarios <sup>14</sup>, Carmen Reverte <sup>15</sup>, Rosemary Shrestha<sup>9</sup>, Imma Subirats<sup>16</sup>, Aravind Venkatesan <sup>12</sup>, Alex Whan<sup>17</sup>, ✉ Hadi Quesneville <sup>2</sup>

[Author details](#)

 This article is included in the [Agriculture, Food and Nutrition gateway](#).

### Abstract

In this article, we present a joint effort of the wheat research community, along with data and ontology experts, to develop wheat data interoperability guidelines. Interoperability is the ability of two or more systems and devices to cooperate and exchange data, and interpret that shared information. Interoperability is a growing concern to the wheat scientific community, and agriculture in general, as the need to interpret the deluge of data obtained through high-throughput technologies grows. Agreeing on common data formats, metadata, and vocabulary standards is an important step to obtain the required data interoperability level in order to add value by encouraging data sharing, and subsequently facilitate the extraction of new information from existing and new datasets.

During a period of more than 18 months, the RDA Wheat Data Interoperability Working Group (WDI-WG) surveyed the wheat research community about the use of data standards, then discussed and selected a set of recommendations based on consensual criteria. The recommendations promote standards for data types identified by the wheat research community as the most important for the coming years: nucleotide sequence variants, genome annotations, phenotypes, germplasm data, gene expression experiments, and physical maps. For each of these data types, the guidelines recommend best practices in terms of use of data formats, metadata standards and ontologies. In addition to the best practices, the guidelines provide examples of tools and implementations that are likely to facilitate the adoption of the recommendations.

To maximize the adoption of the recommendations, the WDI-WG used a community-driven approach that involved the wheat research community from the start, took into account their needs and practices, and provided them with a framework to keep the recommendations up to date. We also report this approach's potential to be generalizable to other (agricultural) domains.

[Get PDF](#) [Get XML](#) [Cite](#) [Export](#) [Track](#) [Share](#)

2561 VIEWS  
455 DOWNLOADS

<https://doi.org/10.12688/f1000research.12234.2>



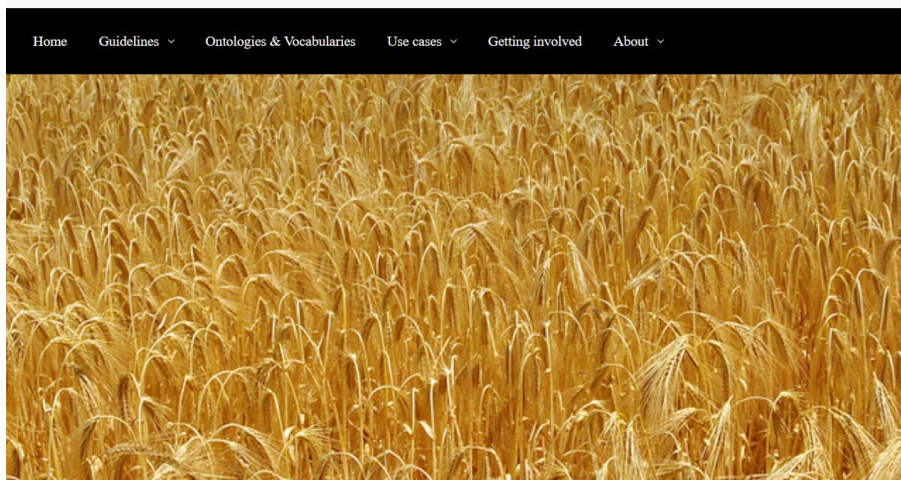
INRAE

Find Genomic, Genetic and Phenomic Data for Wheat Using the WheatIS Data Discovery Tool  
PAG32, 14 January 2025



# ➤ Dedicated website

## Wheat Data Interoperability Guidelines



### Welcome

These recommendations have been prepared by members of the [Wheat Data Interoperability Working Group \(WG\)](#), one of the WGs of the Research Data Alliance and the only WG of the [Agriculture Data Interoperability Interest Group](#). The group is coordinated by members of the [Wheat Initiative](#), a global initiative that aims to reinforce synergies between bread and durum wheat national and international research programmes to increase food security, nutritional value and safety while taking into account societal demands for sustainable and resilient agricultural production systems. All the standards and databases presented in these recommendations are referenced into the [FAIRsharing](#) website.

More specifically, the WG aims to:



**Promote** the adoption of common standards, vocabularies and best practices for Wheat data management



**Facilitate** access, discovery and reuse of wheat data



**Facilitate** wheat data integration

### Guidelines

Under the submenus of this section you will find information on the best practices, tools, recommendations and examples to create, manage and share data related to Wheat. It includes subsections for each of the following data types: What are currently the most used and relevant vocabularies in the context of Wheat Initiative?

**Sequ**  
**Geno**  
**Phen**  
**Gerr**  
**Gene**  
**Physi**

From December 2014 to January 2015 the editorial team conducted a survey "Towards a Comprehensive Overview of Ontologies and Vocabularies for Research on Wheat". The objective was to collect information about the visibility, interoperability, domain, content and other technical aspects of relevant ontologies and vocabularies. As a result, in February 2015 a report ([link](#)) was published, and also a list of vocabularies listed as follows:

1. [AGROVOC](#)
2. [Biorefinery](#)
3. [CAB Thesaurus \(CABT\)](#)
4. [Cell Ontology \(CL\)](#)
5. [Chemical Entities of Biological Interest \(ChEBI\)](#)
6. [Crop Ontology \(CO\)](#)
7. [Crop Research Ontology – part of Crop Ontology \(CO\\_715\)](#)
8. [Environment Ontology \(ENVO\)](#)
9. [Experimental Factor Ontology \(EFO\)](#)
10. [Feature Annotation Location Description Ontology \(FALDO\)](#)
11. [NAL Thesaurus \(NALT\)](#)
12. [Phenotype And Trait Ontology \(PATO\)](#)
13. [Plant Experimental Conditions Ontology \(Plant Environment Ontology, EO, may be changing to PECO\)](#)

### Exempl

### Links

FAIRshari

### Use cases

Click on the titles for details

1. [QUERY ON TRIAL DATA ASSOCIATED WITH VARIETIES](#)
2. [DECIPHER THE EFFECT OF SNPS](#)
3. [IDENTIFICATION OF WHEAT GENES THAT CONTROL ROOT GROWTH](#)
4. [QUERY DATA BY GERMPASM/GENOTYPE](#)
5. [QUERY PHENOTYPE TRIALS TO BUILD INTEGRATIVE DATASET FOR PHENOTYPE OR GENETIC ANALYSIS.](#)
6. [QUERY ON GERMPASM WITH SPECIFIC TRAITS](#)

<https://ist.blogs.inrae.fr/wdi/>



# WheatIS

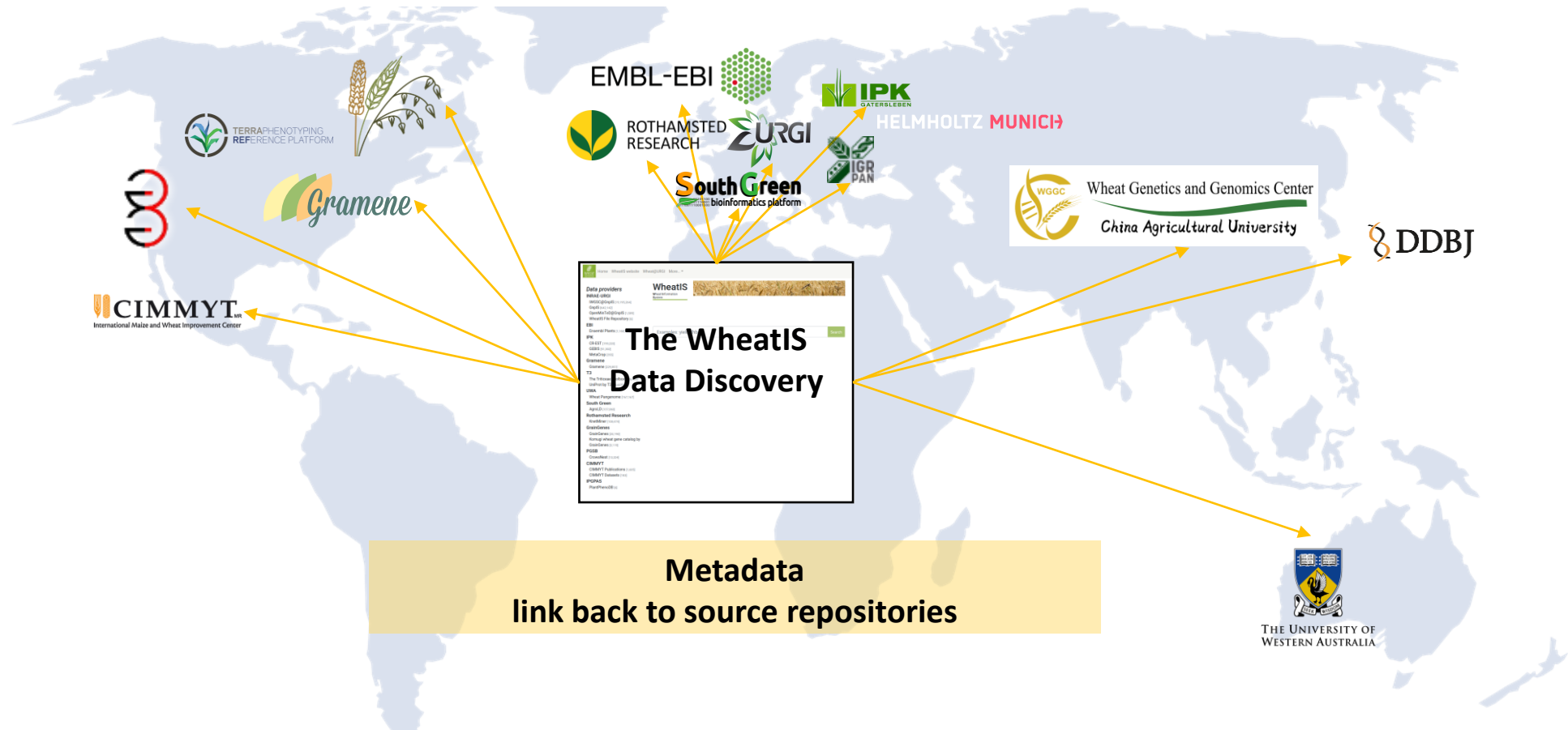
DataDiscovery

Exemples : yield, fhb

Rechercher

<https://urgi.versailles.inrae.fr/wheatis>

# ➤ Federated data portal



## > Repositories

### INRAE-URGI

IWGSC@GnplS [19,195,264]

GnplS [624,827]

brapi@INRAE-URGI [15,949]

OpenMinTeD@GnplS [1,692]

WheatIS File Repository [6]

Research Data Gouv [4]

### CAU

TGT [5,041,169]

wGRN [270,235]

WheatUnion [105,200]

pNOGmap [118]

### EBI

Ensembl Plants [1,954,199]

### Gramene

Gramene Plants [1,954,199]

### IPK

CR-EST [199,220]

GEBIS [51,820]

MetaCrop [355]

e!DAL - PGP Repository [144]

## Data type



Genome

annotation [14,176,618]



Gene annotation [2,861,241]



Physical map

feature [2,157,405]



### CIMMYT

Publications Repository [1,741]

Data Repository [374]

### EVA

brapi@EVA [710]

### TERRA-REF

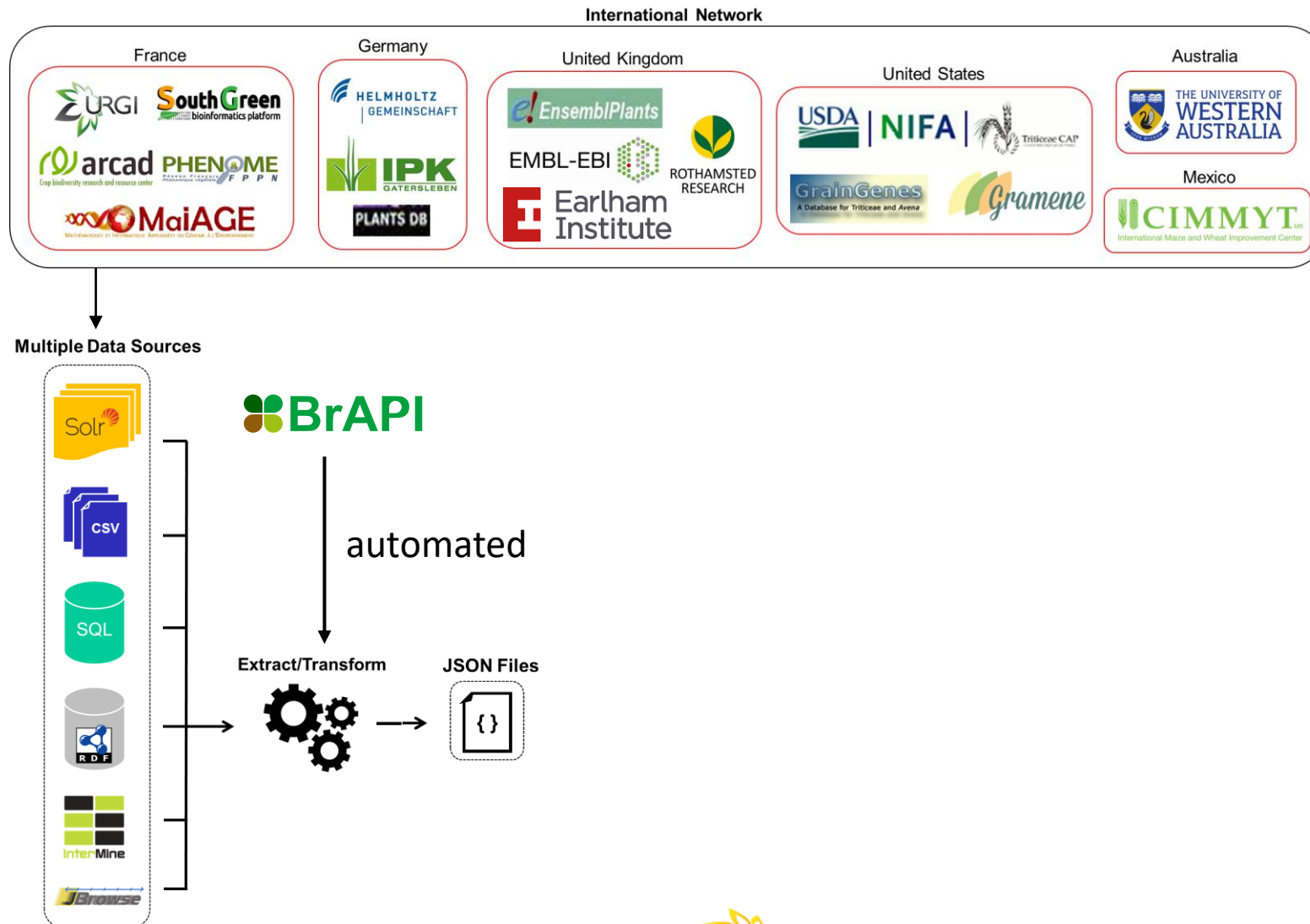
brapi@TERRA-REF [284]

### IPGPAS

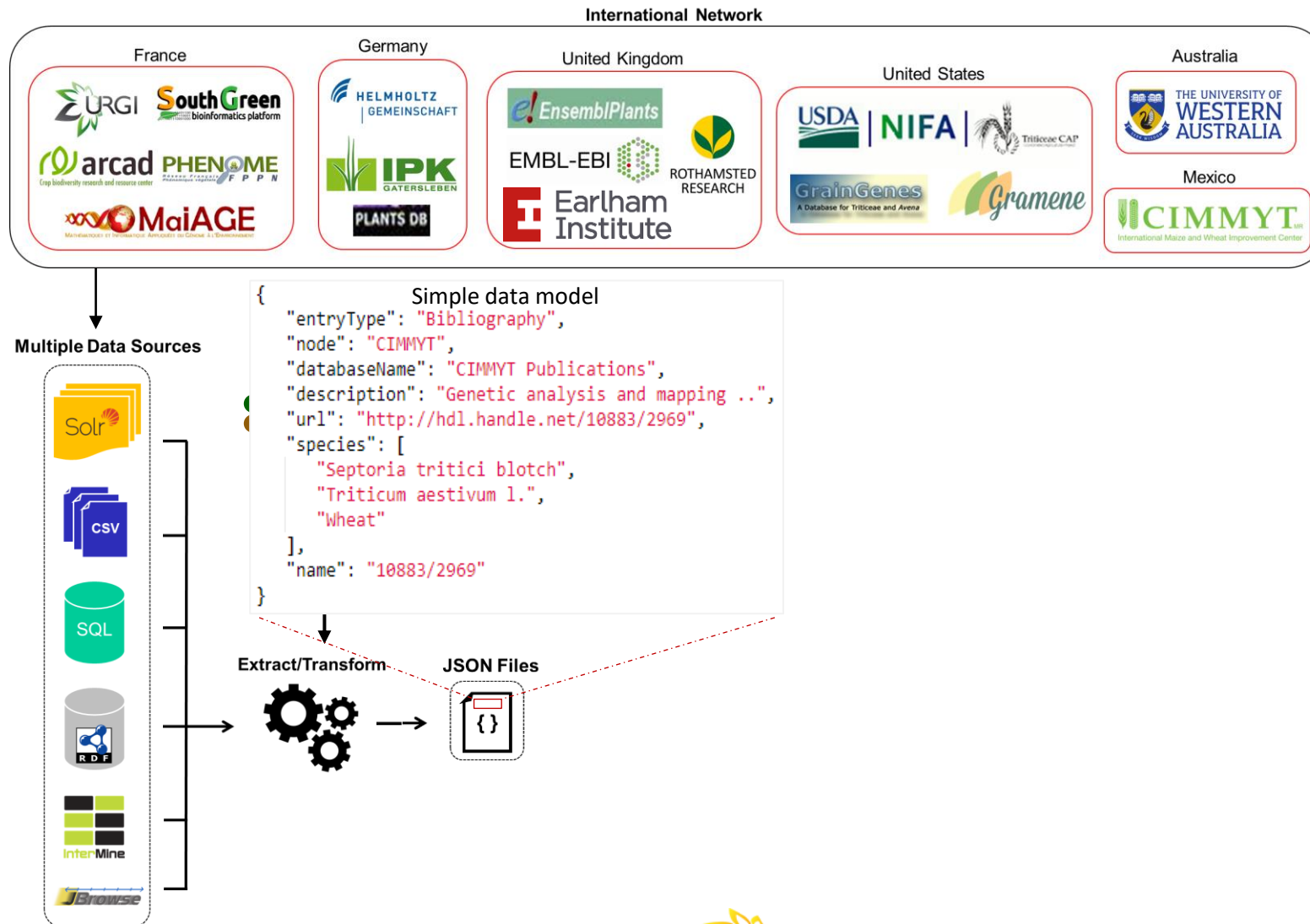
PlantPhenoDB [6]



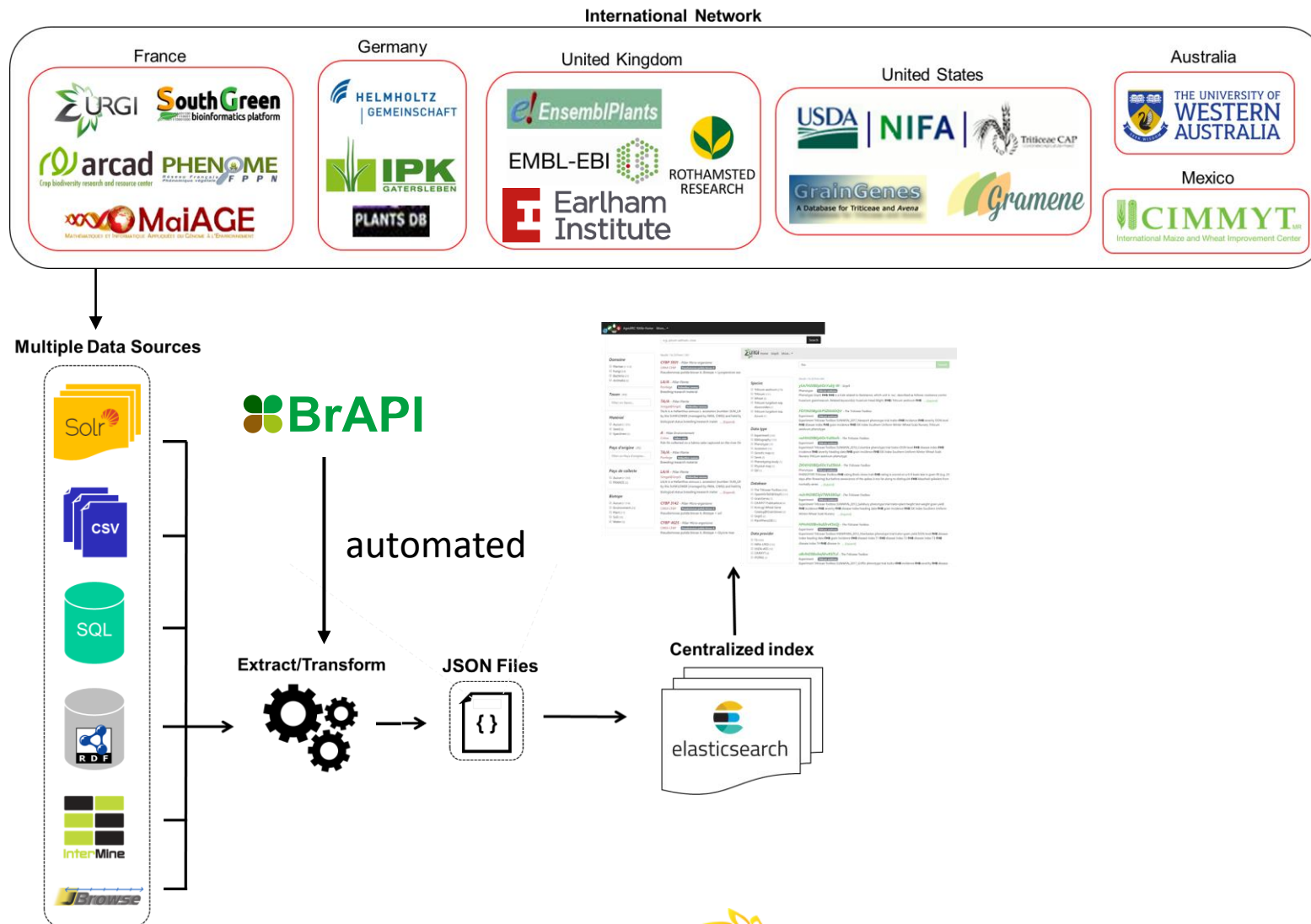
# ➤ Behind the scene



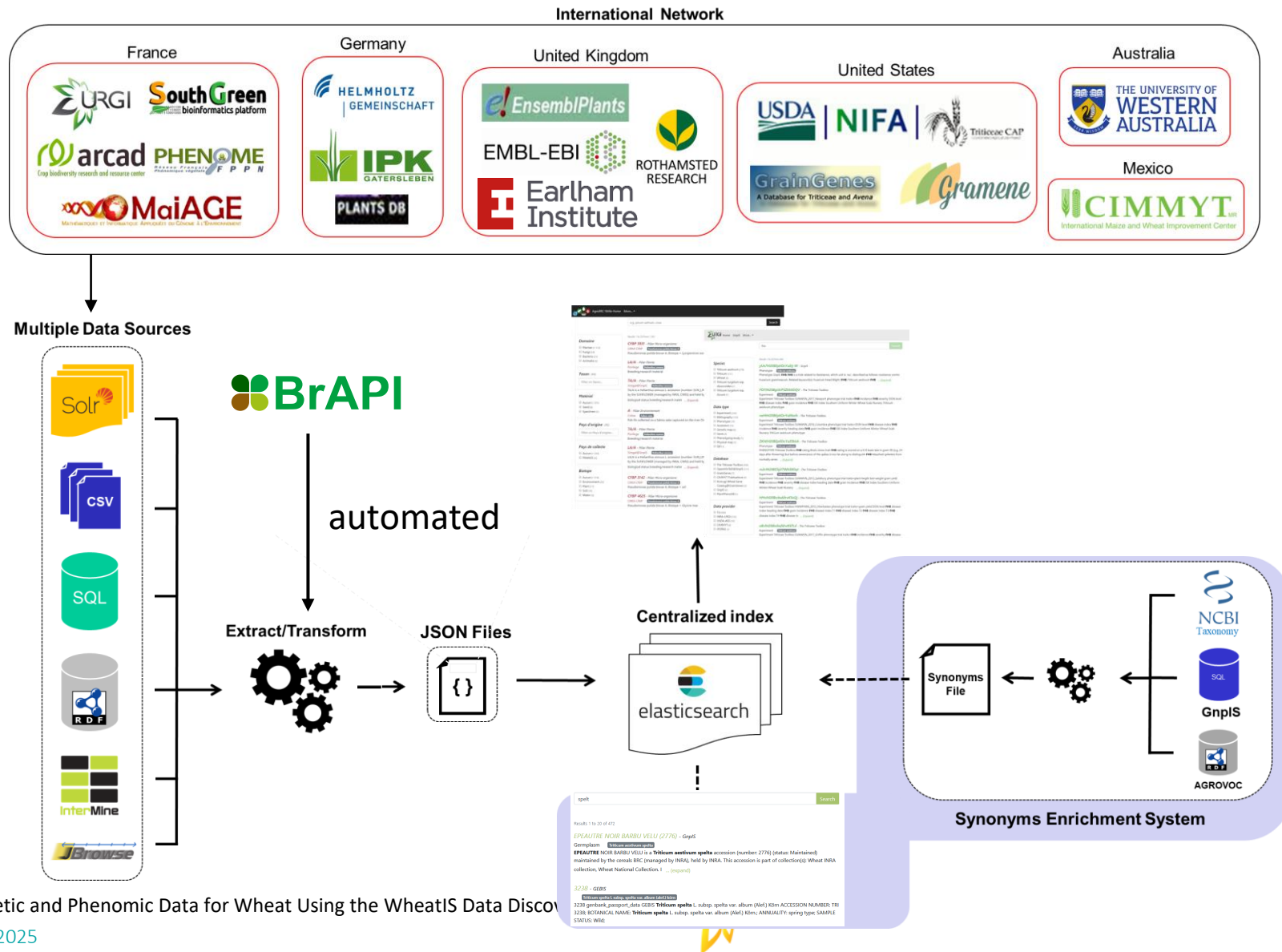
# ➤ Behind the scene



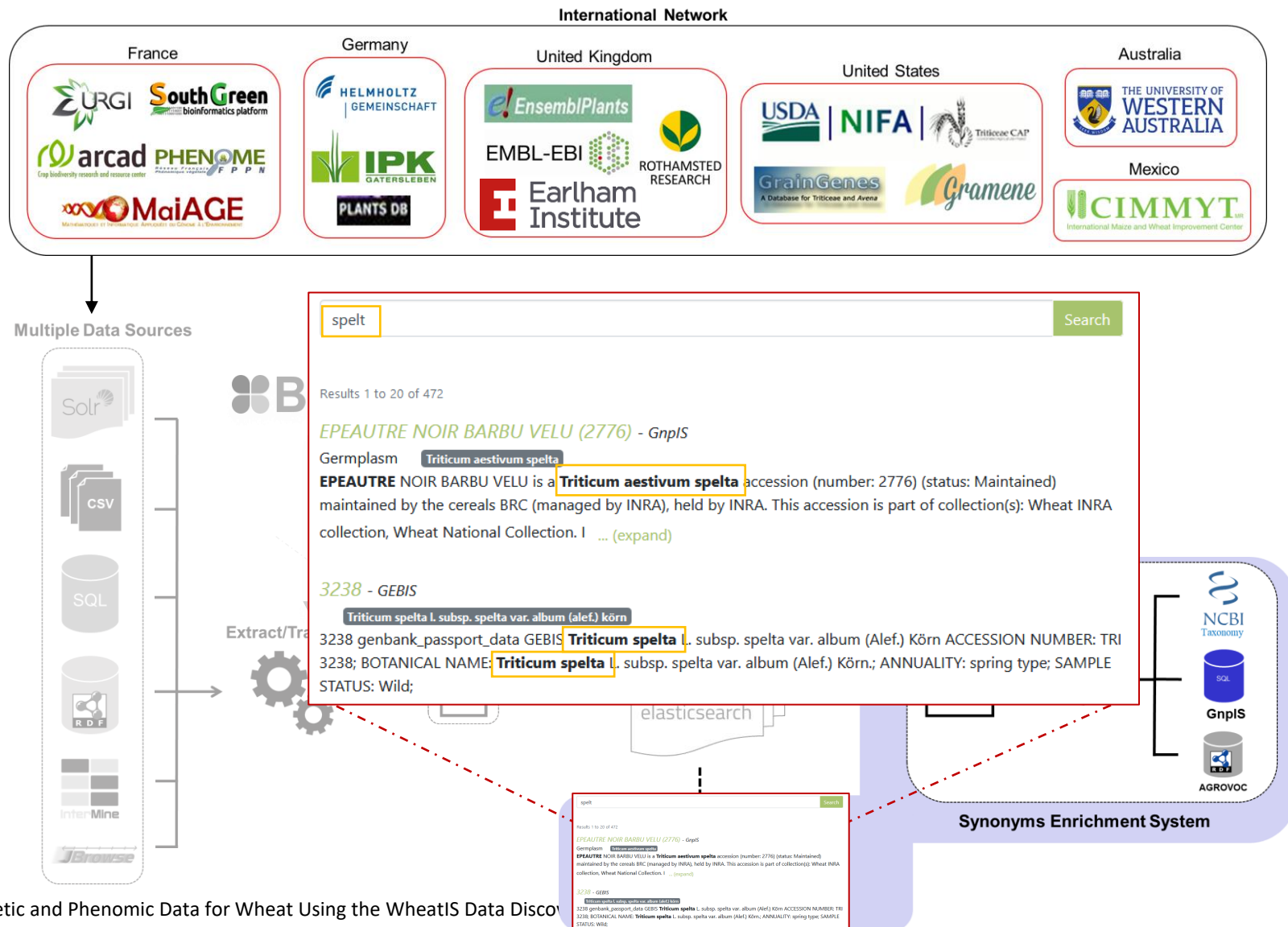
# ➤ Behind the scene



# ➤ Behind the scene



# ➤ Behind the scene



INRAE

Find Genomic, Genetic and Phenomic Data for Wheat Using the WheatIS Data Discovery

PAG32, 14 January 2025

## ➤ Data exchange format



- /brapi/v1/germplasm GET
- /brapi/v1/germplasm/{germplasmDbld} GET
- /brapi/v1/variables GET
- /brapi/v1/locations GET
- /brapi/v1/locations/{locationDbld} GET
- /brapi/v1/studies/{studyDbld}/germplasm GET
- /brapi/v1/studies/{studyDbld}/observationvariables GET
- /brapi/v1/trials GET
- /brapi/v1/trials/{trialDbld} GET

elixir-europe/**plant-brapi-etl-faidare**

Harvest and index meta data from BrAPI endpoints for data access through the plant-faidare data lookup portal (<https://github.com/elixir-europe/plant-faidare>).

11 Contributors 12 Issues 7 Stars 8 Forks



or



Field



- name
- url
- description
- entryType
- species
- node
- databaseName

- WheatIS species list: Aegilops\*, Hordeum\*, Triticum\*, Wheat\*
- Format: JSON or TSV, as you wish!
- Web server: HTTP or FTP, as you wish!

<https://github.com/elixir-europe/plant-faidare>

<https://urgi.versailles.inrae.fr/wheatis/join>

INRAE

Find Genomic, Genetic and Phenomic Data for Wheat Using the WheatIS Data Discovery Tool

PAG32, 14 January 2025



# > Demo



## Data providers

### INRAE-URGI

- IWGSC@GnplS [19,195,264]
- GnplS [642,142]
- OpenMinTeD@GnplS [1,589]
- WheatIS File Repository [6]

### EBI

- Ensembl Plants [1,168,762]

### IPK

- CR-EST [199,220]
- GEBIS [51,302]
- MetaCrop [355]

### Gramene

- Gramene [229,851]

### T3

- The Triticeae Toolbox [206,406]
- UniProt by T3 [16,607]

### UWA

- Wheat Pangenome [167,167]

### South Green

- AgroLD [137,060]

### Rothamsted Research

- KnetMiner [108,474]

### GrainGenes

- GrainGenes [20,190]
- Komugi wheat gene catalog by GrainGenes [3,119]

### PGSB

- CrowsNest [13,324]

### CIMMYT

- CIMMYT Publications [1,605]
- CIMMYT Datasets [183]

### IPGPAS

- PlantPhenoDB [6]

## WheatIS

Wheat Information System



Examples: yield, fhb

Search

More... ▾

- About
- Join us
- Legal mentions
- Help



INRAE

Find Genomic, Gene  
PAG32, 14 January :

# > Demo

WHEAT INITIATIVE Home WheatIS website Wheat@URGI More... ▾

Search:  Search

Results 1 to 1 from 1

**10.1007/s00122-012-1838-2 - OpenMinTeD@GnpIS**

Bibliography **Triticum**

Bibliography OpenMinTeD 10.1007/s00122-012-1838-2 Mapping QTL for <span class='openminted\_phenotype'>resistance to eyespot of <span class='openminted\_taxon' [...] 415</span>, and <span class='openminted\_marker'>**cfd2**</span> are tightly linked to Q.<span class='ope ... (Expand)

**JD10WSE\_0114A\_7333656\_7333940\_CFD\_SSR\_TRIMMED - IWGSC@GnpIS**

Gene annotation **Triticum aestivum**

SEQUENCE FEATURE IWGSC@GnpIS\_chr4A\_7333656\_7333940\_CFD\_SSR\_TRIMMED Start = 7333656, End = 7333940, Strand = 0, Source = ePCR, Seq\_id = chr4A, Size = 283, Marker = **CFD2**, Id = **CFD2**, Type = similarity, Motif = ca(11) Triticum aestivum similarity chr4 ... (Expand)

**10.1007/s00122-012-1838-2 - OpenMinTeD@GnpIS**

Bibliography **Triticum**

Bibliography OpenMinTeD 10.1007/s00122-012-1838-2 Mapping QTL for <span class='openminted\_phenotype'>resistance to eyespot of <span class='openminted\_taxon' [...] 415</span>, and <span class='openminted\_marker'>**cfd2**</span> are tightly linked to Q.<span class='ope ... (Expand)

**Intersection ((AND query))**

- Species
  - Triticum [1]
- Data type
  - Bibliography [1]
- Database
  - IWGSC@GnpIS [4]
  - OpenMinTeD@GnpIS [1]
- Data provider
  - INRAE-URGI [1]



EnsemblPlants ▼ HMMER | BLAST | BioMart | Tools | Downloads | Help & Docs | Blog Login/Register

**Triticum aestivum** (IWGSC) ▼

Location: 2A:100,113,981-100,117,432 Gene: TraesCS2A02G151900 Trans: TraesCS2A02G151900.1

**Gene-based displays**

- Summary
  - Splice variants
  - Transcript comparison
  - Gene alleles
- Sequence
  - Secondary Structure
- Gene families
- Literature
- Plant Compara
  - Genomic alignments
  - Gene tree
  - Gene gain/loss tree
  - Orthologues
  - Paralogues
  - Homoeologues
- Pan-taxonomic Compara
  - Gene Tree
  - Orthologues
- Ontologies
  - GO: Biological process
  - GO: Cellular component
  - GO: Molecular function
- Phenotypes
- Genetic Variation
  - Variant table
  - Variant image
  - Structural variants
- Gene expression
- Pathway
- Regulation
- External references
- Supporting evidence
- ID History
  - Gene history

### Gene: TraesCS2A02G151900

**Description** Uncharacterized protein At3g57150 (Fragment) [Source:Projected from Arabidopsis thaliana (AT3G57150) UniProtKB/TrEMBL;Acc:C0SVF3]

**Location** [Chromosome 2A: 100,113,981-100,117,432](#) forward strand.

**About this gene** This gene has 1 transcript ([splice variant](#)) and [144 orthologues](#).

**Transcripts** [Show transcript table](#)

---

**Summary** ⓘ

**Gene type** Protein coding

**Annotation method** Genes annotated with high confidence by IWGSC

Go to [Region in Detail](#) for more tracks and navigation options (e.g. zooming)

Drag/Select: ↔

**Gene Legend**

Protein Coding

■ IWGSC high confidence protein coding

**Configuring the display**

Tip: use the "Configure this page" link on the left to show additional data in this region.

Ensembl Plants release 48 - August 2020 © [EMBL-EBI](#)

[Ensembl Plants is produced in collaboration with Gramene](#)

- [Configure this page](#)
- [Custom tracks](#)
- [Export data](#)
- [Share this page](#)
- [Bookmark this page](#)

**About Us**

- [About us](#)
- [Contact us](#)
- [Citing Ensembl Genomes](#)

**Get help**

- [Using this website](#)
- [Documentation](#)
- [Adding custom tracks](#)

**Our sister sites**

- [Ensembl](#)
- [Ensembl Bacteria](#)
- [Ensembl Fungi](#)

**Follow us**

- [Blog](#)
- [Twitter](#)



# > Demo

WHEAT INITIATIVE Home WheatIS website Wheat@URGI More... ▾

cfid2 Search

Results 1 to 1 from 1

10.1007/s00122-012-1838-2 - OpenMinTeD@GnpIS

Bibliography **Triticum**

Species  
 Triticum [1]

Springer Nature is making SARS-CoV-2 and COVID-19 research free. [View research](#) | [View latest news](#) | [Sign up for updates](#)

SpringerLink Search Home Log in

Original Paper | Published: 11 March 2012

## Mapping QTL for resistance to eyespot of wheat in *Aegilops longissima*

Hongyan Sheng, Deven R. See & Timothy D. Murray

*Theoretical and Applied Genetics* **125**, 355–366(2012) | [Cite this article](#)

402 Accesses | 14 Citations

### Abstract

Eyespot is an economically important disease of wheat caused by the soilborne fungi *Oculimacula yallundae* and *O. acuformis*. These pathogens infect and colonize the stem base, which results in lodging of diseased plants and reduced grain yield. Disease resistant cultivars are the most desirable control method, but resistance genes are limited in the wheat gene pool. Some accessions of the wheat wild relative *Aegilops longissima* are resistant to eyespot, but nothing is known about the genetic control of resistance. A recombinant inbred line population was developed from the cross PI 542196 (R) × PI 330486 (S) to map the resistance genes and better understand resistance in *Ae. longissima*. A genetic linkage map of the S<sup>1</sup> genome was constructed with 169 wheat microsatellite markers covering 1261.3 cM in 7 groups. F<sub>5</sub> lines (189) were tested for reaction to *O. yallundae* and four QTL were detected in chromosomes 1S<sup>1</sup>, 3S<sup>1</sup>, 5S<sup>1</sup>, and 7S<sup>1</sup>. These QTL explained 44 % of the total phenotypic variation in reaction to eyespot based on GUS scores and 63 % for visual disease ratings. These results demonstrate that genetic control of *O. yallundae* resistance in *Ae. longissima* is polygenic. This is the first report of multiple QTL conferring resistance to eyespot in *Ae. longissima*. Markers *cfid6*, *umc507*, *umc415*, and *cfid2* are tightly linked to *O. yallundae*-1S<sup>1</sup>.

**CFD2** - IWGSC@GnpIS

Genome annotation **Triticum aestivum**

Genome annotation IWGSC@GnpIS **CFD2** **CFD2** is a similarity:ePCR\_cfd of Triticum aestivum located between positions 14028 and 14311 on 4AS\_v2\_5981297 and which properties are Size=283,Motif=gt(9),load\_id=CFD2\_147 Triticum aestivum similarity

Access options

Buy article PDF

34,95 €

Price includes VAT for France

Instant access to the full article PDF.

Buy journal subscription

150,25 €

Price includes VAT for France

Immediate online access to all issues from 2019. Subscription will auto renew annually.

[Rent this article via DeepDyve.](#)

[Learn more about Institutional subscriptions](#)

Sections **Figures** References

Abstract

References



INRAE

Find Genomic, Genetic and P  
PAG32, 14 January 2025

# > Demo

The screenshot displays the Gramene website interface for a gene comparison. At the top, the gene **TRIUR3\_07354** from *Triticum urartu* is identified as a putative deoxyribonuclease TATDN1. A homologous gene, **Zm00001d007881** from *Zea mays*, is highlighted as a "Model Species Homolog" and noted as "TatD related DNase".

The main content area features a "Compara Gene Tree" and a genomic track. The tree shows the evolutionary relationship between the gene and its orthologs across various species, including *Et\_s6749-0.26-1.path1*, *SORBI\_3002G038200*, *Zm00001d007881*, and others. The genomic track displays gene models and orthology relationships across multiple genomes.

Navigation and search options are visible on the left and top. The left sidebar includes sections for "Species" (with checkboxes for various taxa), "Data" (with checkboxes for different data types), and "Data" (with checkboxes for specific datasets). The top navigation includes "Location", "Homology", and "X-refs" tabs. A search bar at the top right contains the text "1 genes in 1 genomes".

At the bottom of the page, there is a footer with "Release Notes (62)", "About", "Cite", "Feedback", and "Privacy" links. A dark green banner contains the text: "Gramene is a collaborative effort among Cold Spring Harbor Laboratory, Oregon State University, and EMBL-EBI. Funding is provided by the NSF and USDA ARS." To the right, an "Application Info" box lists: "Data Release 62", "Gramene Search UI LOCAL DEVELOPMENT BUILD", and "Built on January 14th 2020 at 1:01:57 pm". Below this, an "Actions" section includes a "Reset Gramene Search" button.



**INRAE**

Find Genomic, Genetic and P  
PAG32, 14 January 2025

## ➤ New features: ontology annotation

- Link literature with experimental data (genotyping, phenotyping...) using ontology annotation
- Corpus of ~1600 wheat open access publications
- Text mining ontology: Wheat Trait Ontology
- Experimental data ontology: Crop Ontology

Ontology variable selection

Filter  English ▼

- Wheat Crop Ontology **Ontology**
- Abiotic stress **Trait class**
- Agronomical **Trait class**
  - Grain weight **Trait**
  - Grain yield **Trait**
  - Lodging incidence **Trait**
    - LS\_score: Lodging score **Variable**
  - Plant height **Trait**
  - Precocity **Trait**
- Biotic stress **Trait class**
- Other **Trait class**
- Quality **Trait class**

**LS\_score** **Variable**

Ontology name Wheat Crop Ontology  
Identifier CO\_321:1000099  
Name LS\_score  
Synonyms Lodging score  
Susceptibility to lodging

Institution INRA  
Scientist Jacques Le Gouis  
Date 15/06/2016  
Crop Wheat  
Cross reference WIPO:0000099

**Lodging incidence** **Trait**

Identifier CO\_321:0000167  
Name Lodging incidence  
Description Indicates incidence of lodged plants.  
Entity Plant  
Attribute Lodging incidence  
Class Agronomical  
Main abbreviation Lodg  
Alternative LOD

OK Cancel

# ➤ New features: synonym enrichment

WHEAT INITIATIVE Home WheatIS website Wheat@URGI More... ▾

einkorn wheat Search

Results 1 to 20 from 197 222 (limited to 10 000)

**Species (290)**  
Filter on Species...

**Data type (20)**  
Filter on Data type...

**Database (17)**  
Filter on Database...

**Data provider**

- UWA [167 167]
- INRAE-URGI [12 224]
- GrainGenes [8 008]
- IPK [6 706]
- CIMMYT [1 696]
- T3 [1 353]
- Rothamsted Research [63]
- Gramene [4]
- IPGPAS [1]

**Marquis\*4/2/Stewart\*3/Triticum monococcum - GrainGenes**  
Germplasm **Triticum aestivum**  
Germplasm GrainGenes Marquis\*4/2/Stewart\*3/**Triticum monococcum** Unknown Germplasm Type. Marquis\*4/2 /Stewart\*3/**Triticum monococcum** Triticum aestivum ... (Expand)

**2006380 - GEBIS**  
**Triticum monococcum**  
2006380 genbank\_passport\_data GEBIS **Triticum monococcum L.** ACCESSION NUMBER: TRI 29537; BOTANICAL NAME: **Triticum monococcum L.**; OTHER NAMES: Albini; ANNUALITY: spring type; SAMPLE STATUS: Advanced/improved cultivar; COUNTRY OF ORIGIN: Germany; ... (Expand)

**617 - GEBIS**  
**Triticum monococcum l. var. vulgare korn**  
617 genbank\_passport\_data GEBIS **Triticum monococcum L. var. vulgare Körn** ACCESSION NUMBER: TRI 617; BOTANICAL NAME: **Triticum monococcum L. var. vulgare Körn.**; ANNUALITY: spring type; SAMPLE STATUS: Traditional cultivar/landrace; COUNTRY OF ORIGIN: Albania; ... (Expand)

**2009 - GEBIS**  
**Triticum monococcum l. var. macedonicum papag**  
2009 genbank\_passport\_data GEBIS **Triticum monococcum L. var. macedonicum Papag** ACCESSION NUMBER: TRI 2009; BOTANICAL NAME: **Triticum monococcum L. var. macedonicum Papag.**; ANNUALITY: spring type; SAMPLE STATUS: Traditional cultivar/landrace; COUNTRY OF ORIGIN: ... (Expand)

**644 - GEBIS**  
**Triticum monococcum l. var. vulgare korn**  
644 genbank\_passport\_data GEBIS **Triticum monococcum L. var. vulgare Körn** ACCESSION NUMBER: TRI 644; BOTANICAL NAME: **Triticum monococcum L. var. vulgare Körn.**; ANNUALITY: spring type; SAMPLE STATUS: Traditional



# ➤ New features: BrAPI sources

The screenshot shows the WheatIS website interface. At the top, there are logos for 'wheatIS WHEAT INFORMATION SYSTEM' and 'WHEAT INITIATIVE'. Navigation links include 'Home', 'Wheat@URGI', and 'More...'. The main heading is 'WheatIS DataDiscovery' with a background image of wheat. A search bar contains the text 'Examples: yield, fhb' and a 'Search' button. On the left, a list of data providers is shown, including INRAE-URGI, EBI, IPK, Gramene, T3, UWA, Rothamsted Research, GrainGenes, PGSB, CIMMYT, EVA, TERRA-REF, and IPGPAS. A green arrow points from the BrAPI logo to the 'brapi@INRAE-URGI' entry in the list. Another green arrow points from the BrAPI logo to the 'brapi@EVA' entry. A third green arrow points from the BrAPI logo to the 'brapi@TERRA-REF' entry.

**Data providers**

- INRAE-URGI**
  - IWGSC@GnplS [19,195,264]
  - GnplS [631,498]
  - brapi@INRAE-URGI [15,949]
  - OpenMinTeD@GnplS [1,692]
  - WheatIS File Repository [6]
- EBI**
  - Ensembl Plants [3,071,189]
- IPK**
  - CR-EST [199,220]
  - GEBIS [51,820]
  - MetaCrop [355]
- Gramene**
  - Gramene [229,851]
- T3**
  - The Triticeae Toolbox [206,406]
  - UniProt by T3 [16,607]
- UWA**
  - Wheat Pangenome [167,167]
- Rothamsted Research**
  - KnetMiner [108,474]
- GrainGenes**
  - GrainGenes [20,190]
  - Komugi wheat gene catalog by GrainGenes [3,119]
- PGSB**
  - CrowsNest [13,324]
- CIMMYT**
  - CIMMYT Publications [1,600]
  - CIMMYT Datasets [183]
- EVA**
  - brapi@EVA [710]
- TERRA-REF**
  - brapi@TERRA-REF [284]
- IPGPAS**
  - PlantPhenoDB [6]

**BrAPI**



**INRAE**

Find Genomic, Genetic  
PAG32, 14 January 202

# ➤ How to join ?

## How to join Plant data discovery Federations (FAIDARE, WheatIS)?

---

### Overview

---

The plant data discovery Federations (FAIDARE, wheatIS) provides search data portal that index the metadata from your data resources and then link back to an access page in your system. This indexation can be done using the following approaches:

- Datadiscovery files in a webfolder
- Breeding API (BrAPI) web service endpoint. Provides both datadiscovery and [summary cards](#)
- Breeding API (BrAPI) files in a webfolder. Provides both datadiscovery and [summary cards](#)

Each of those approaches are described below and all assume a minimum information set comprising an URL for link back plus description.

The metadata format must follow the indications below and we invite you to [contact us](#) as soon as possible so that we can provide help and discuss the best way to go ahead.

### Breeding API (BrAPI)

---

This is the richer approach and will bring you all FAIDARE functionalities. The web services building will enable you to plug any [BrAPI](#) client on your database. The BrAPI file generation is simpler and easier to deploy. Only Germplasm and study are indexed from a BreedingAPI endpoint, with their full description. Those metadata will be used to create summary cards [such as](#) The datadiscovery metadata files, following the [specifications](#) below are generated from those summaries. Currently (FEB 2023), FAIDARE indexes BrAPI v1.1+ sources (V1.3 recommended).

### Web services

The breedingAPI full specifications are available on [www.brapi.org](http://www.brapi.org). The resources indexed are germplasm and study only. Information cards are created using the following calls :

- germplasm (mandatory)
- location (recommended)
- ontology (recommended)
- program
- study (mandatory)
- study/{studyDbId}/observationVariable (recommended)
- study/{studyDbId}/germplasm (mandatory)
- study/{studyDbId}/observationUnit (can be resource intensive and therefore not implemented)
- trial (recommended)



INRAE

Find Genomic, Genetic and Phenomic Data for Wheat Using the WheatIS Data Discovery Tool  
PAG32, 14 January 2025



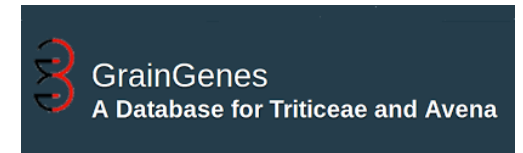
# ➤ Experiences about AI in wheat bioinformatics

**BioCypher**  
A unifying framework for biomedical KGs

- Relies on a collection of **reusable “adapters”** for the different sources of biomedical knowledge
- **Recombination** of those adapters to fit various demands
- Reducing redundant maintenance work through **quasi-standardisation**
- Integrates the controlled vocabularies of ontologies into the process helps to harmonise the data from individual resources and yields a consistent semantic basis for downstream analyses.
- Through unambiguous and simple “low-code” configuration, a reproducible knowledge graph can be created and shared for every specific task

Resources: Tissue, Drug, Gene, ...  
Adapters: .csv, http://, ftp://, GO, DO, etc.  
Configuration: [Settings]  
Optimisation: [Graph]  
Knowledge Graph: SQL, LPG, RDF, etc.  
Actions: Maintain, Analyse, Archive, Share

WheatIS EWG half-yearly meeting AI



INRAE

Find Genomic, Genetic and Phenomic Data for Wheat Using the WheatIS Data Discovery Tool  
PAG32, 14 January 2025



# ➤ Acknowledgements



Former chair:  
Hadi Quesneville

Expert Working Group



Michael Alaux  
Françoise Alfama  
Raphael Flores  
Nicolas Francillonne  
Najwa Lakmouri  
Célia Michotey  
Cyril Pommier  
Anne-Françoise Adam-Blondon



Isabelle Caugant  
Josh Stein  
Kellye Eversole



Teresa Saavedra  
Sylvie Cloutier  
Peter Langridge



**All data and metadata providers !**



Find Genomic, Genetic and Phenomic Data for Wheat Using the WheatIS Data Discovery Tool  
PAG32, 14 January 2025





## WheatIS data discovery

<https://urgi.versailles.inrae.fr/wheatis>

## Related publications:

Alaux et al., *Genome Biology* 2018

<https://doi.org/10.1186/s13059-018-1491-4>



### Linking the International Wheat Genome Sequencin...

The Wheat@URGI portal has been developed to provide the international community of researchers and breeders with access to the bread wheat reference genome se...

[genomebiology.biomedcentral.com](http://genomebiology.biomedcentral.com)

Alaux et al., *The Wheat Genome book* 2023

[https://doi.org/10.1007/978-3-031-38294-9\\_2](https://doi.org/10.1007/978-3-031-38294-9_2)

