Linking the International Wheat Genome Sequencing Consortium Bread Wheat Reference Genome Sequence to Wheat Genetic and Phenomic Data

Michael Alaux

Michael Alaux. Linking the International Wheat Genome Sequencing Consortium Bread Wheat Reference Genome Sequence to Wheat Genetic and Phenomic Data. PAG XXVII - Plant and Animal Genome Conference, Jan 2019, San Diego, United States. hal-04536576

HAL Id: hal-04536576
https://hal.inrae.fr/hal-04536576v1

Submitted on 8 Apr 2024

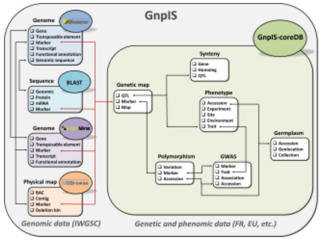
Linking the IWGSC bread wheat reference genome sequence to wheat genetic and phenomic data

Michael Alaux

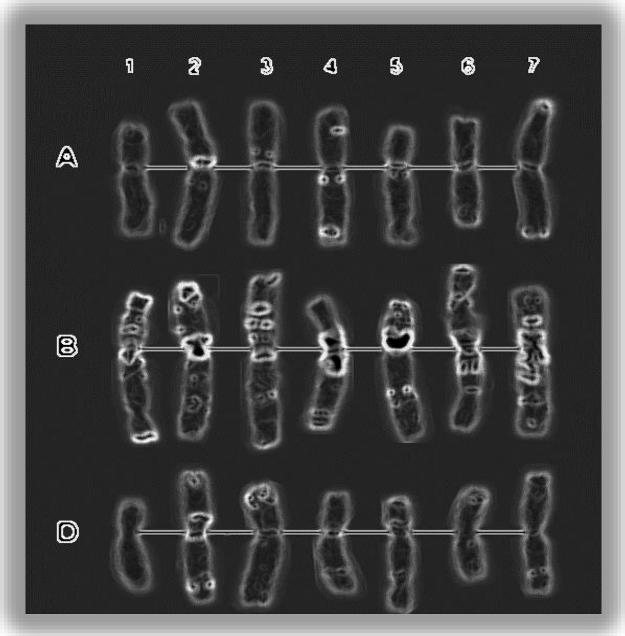
Triticeae Genetics and Genomics, Session 1: Progress in structural and functional genomics

Xwmc430

Examples: Xwmc430, QTL, Tae_1272327



Outline



A. The wheat reference sequence

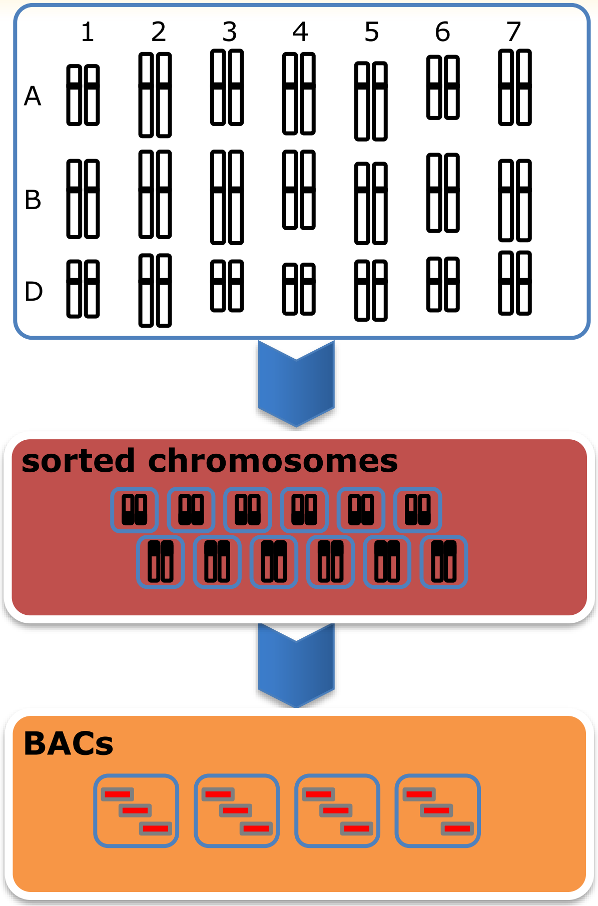
B. A wealth of data available

D. Data integration

➡ **Wheat@URGI portal**

The wheat reference sequence

IWGSC sequencing strategy



NRGene Illumina:
IWGSC NRGene WGS

Illumina PE: CSS v2 (2014)
+MP libs: CSS v3 (2016)

Physical maps: 100%
BAC-seq: 3B + 8 chr.

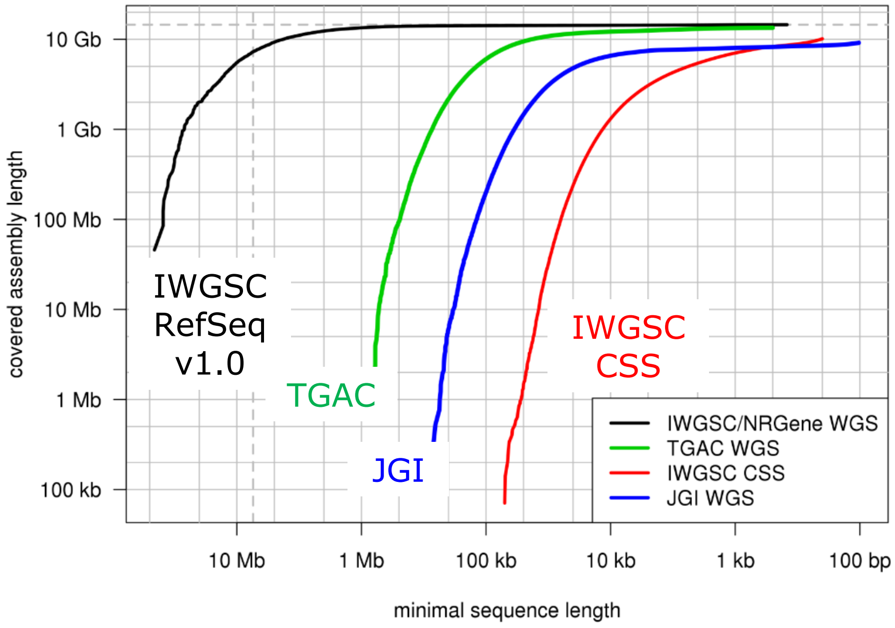
IWGSC
WGA
v0.4
(2016)
+
HiC
BACs
WGP tags
Genetic maps
Bionano maps
Rad. hyb maps

IWGSC
RefSeq
v1.0
(2017)

Courtesy of Fred Choulet

IWGSC RefSeq assembly metrics

total size	14.5 Gb (21 pseudomolecules = 14.1 Gb)
completeness	97-99%
superscaffold N50	23 Mb

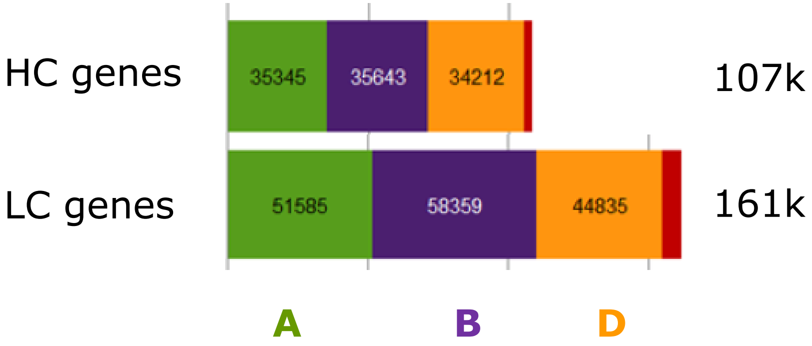


Courtesy of Fred Choulet

IWGSC RefSeq annotation

Genes

v1.0: automated
v1.1: incorporate 3685 manually annotated genes



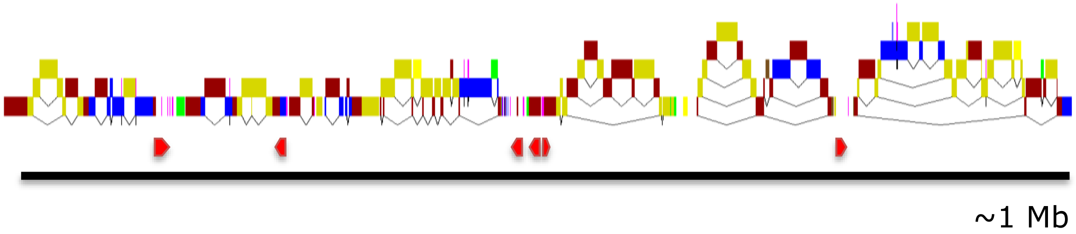
IWGSC, *Science* 2018
https://doi.org/10.1126/science.aar7191

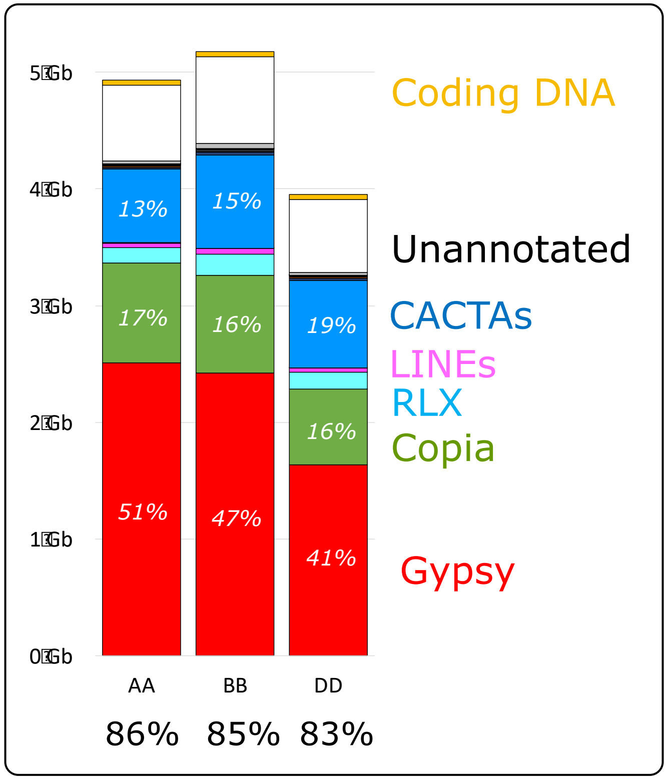
Courtesy of Fred Choulet

IWGSC RefSeq annotation

Transposable Elements

4 million TEs





Wicker et al., *Genome Biology* 2018
https://doi.org/10.1186/s13059-018-1479-0

+ ncRNA, markers, functional annotation, manually curated genes, etc.

Courtesy of Fred Choulet

IWGSC RefSeq data availability

IWGSC WGA v0.4

Toronto agreement

- June 2016:
  - Download
  - BLAST

IWGSC RefSeq v1.0 assembly

Toronto agreement

- January 2017:
  - Download
  - BLAST

IWGSC RefSeq v1.0 annotation

Toronto agreement

- May 2017:
  - Download
- July 2017:
  - Browser

IWGSC RefSeq v1.0

Open access

- July 2018:
  - Assembly and annotations (v1.0 and v1.1) download, BLAST and browser
- August 2018:
  - Publications

IWGSC Data Repository hosted in the **Wheat@URGI portal**

https://wheat-urgi.versailles.inra.fr/Seq-Repository

Michael Alaux

Publications


The tragic ripples of an epic fraud p. 636
Insect pest profits from maize defenses pp. 642 & 694
Photoredox activation of methane pp. 647 & 668

Science

$15
17 AUGUST 2018
sciencemag.org
AAAS

ROAD MAP FOR
WHEAT
Ordered sequence will speed research pp. 635, 661 & 662


RESEARCH ARTICLE

Shifting the limits in wheat research and breeding using a fully annotated reference genome

The International Wheat Genome Sequencing Consortium (IWGSC), IWGSC RefSeq principal investigators:, Rudi Appels[1,36,*]...

+ See all authors and affiliations

Science 17 Aug 2018:
Vol. 361, Issue 6403, eaar7191
DOI: 10.1126/science.aar7191


RESEARCH ARTICLE

The transcriptional landscape of polyploid wheat

R. H. Ramírez-González[1,*], P. Borrill[1,*,†], D. Lang[2], S. A. Harrington[1], J. Brinton[1], L. Venturini[3], M. Davey[4], J. Jacobs[4], F. van...

+ See all authors and affiliations

Science 17 Aug 2018:
Vol. 361, Issue 6403, eaar6089
DOI: 10.1126/science.aar6089


RESEARCH ARTICLE | PLANT SCIENCES

Genome mapping of seed-borne allergens and immunoresponsive proteins in wheat

Angéla Juhász[1,2,*], Tatiana Belova[3,*], Chris G. Florides[1], Csaba Maulis[1], Iris Fischer[4], Gyöngyvér Gell[2], Zsófia Birinyi[2], Jami...

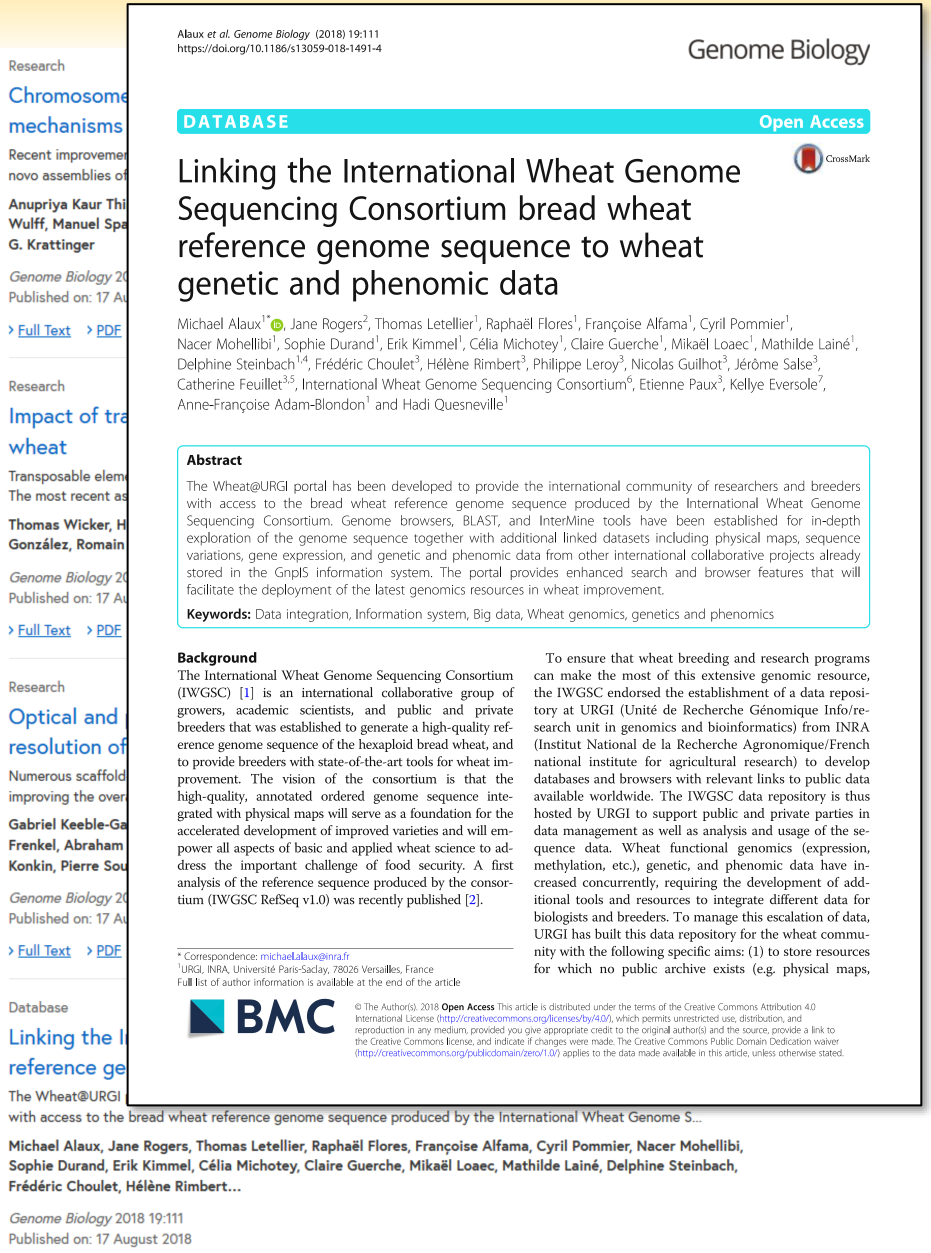
+ See all authors and affiliations

Science Advances 17 Aug 2018:
Vol. 4, no. 8, eaar8602
DOI: 10.1126/sciadv.aar8602

Science Advances
AAAS

URGI

Publications



Alaux *et al. Genome Biology* (2018) 19:111
https://doi.org/10.1186/s13059-018-1491-4

Genome Biology

DATABASE **Open Access**

Linking the International Wheat Genome Sequencing Consortium bread wheat reference genome sequence to wheat genetic and phenomic data

Michael Alaux[1*], Jane Rogers[2], Thomas Letellier[1], Raphaël Flores[1], Françoise Alfama[1], Cyril Pommier[1], Nacer Mohellibi[1], Sophie Durand[1], Erik Kimmel[1], Célia Michotey[1], Claire Guerche[1], Mikaël Loaec[1], Mathilde Lainé[1], Delphine Steinbach[1,4], Frédéric Choulet[3], Hélène Rimbert[3], Philippe Leroy[3], Nicolas Guilhot[3], Jérôme Salse[3], Catherine Feuillet[3,5], International Wheat Genome Sequencing Consortium[6], Etienne Paux[3], Kellye Eversole[7], Anne-Françoise Adam-Blondon[1] and Hadi Quesneville[1]

Abstract

The Wheat@URGI portal has been developed to provide the international community of researchers and breeders with access to the bread wheat reference genome sequence produced by the International Wheat Genome Sequencing Consortium. Genome browsers, BLAST, and InterMine tools have been established for in-depth exploration of the genome sequence together with additional linked datasets including physical maps, sequence variations, gene expression, and genetic and phenomic data from other international collaborative projects already stored in the GnpIS information system. The portal provides enhanced search and browser features that will facilitate the deployment of the latest genomics resources in wheat improvement.

Keywords: Data integration, Information system, Big data, Wheat genomics, genetics and phenomics

Background

The International Wheat Genome Sequencing Consortium (IWGSC) [1] is an international collaborative group of growers, academic scientists, and public and private breeders that was established to generate a high-quality reference genome sequence of the hexaploid bread wheat, and to provide breeders with state-of-the-art tools for wheat improvement. The vision of the consortium is that the high-quality, annotated ordered genome sequence integrated with physical maps will serve as a foundation for the accelerated development of improved varieties and will empower all aspects of basic and applied wheat science to address the important challenge of food security. A first analysis of the reference sequence produced by the consortium (IWGSC RefSeq v1.0) was recently published [2].

To ensure that wheat breeding and research programs can make the most of this extensive genomic resource, the IWGSC endorsed the establishment of a data repository at URGI (Unité de Recherche Génomique Info/research unit in genomics and bioinformatics) from INRA (Institut National de la Recherche Agronomique/French national institute for agricultural research) to develop databases and browsers with relevant links to public data available worldwide. The IWGSC data repository is thus hosted by URGI to support public and private parties in data management as well as analysis and usage of the sequence data. Wheat functional genomics (expression, methylation, etc.), genetic, and phenomic data have increased concurrently, requiring the development of additional tools and resources to integrate different data for biologists and breeders. To manage this escalation of data, URGI has built this data repository for the wheat community with the following specific aims: (1) to store resources for which no public archive exists (e.g. physical maps,

---

* Correspondence: michael.alaux@inra.fr
[1] URGI, INRA, Université Paris-Saclay, 78026 Versailles, France
Full list of author information is available at the end of the article

Research

Chromosome
mechanisms

Recent improveme
novo assemblies of

Anupriya Kaur Thi
Wulff, Manuel Spa
G. Krattinger

Genome Biology 2
Published on: 17 Au

› Full Text › PDF

Research

Impact of tra
wheat

Transposable elem
The most recent as

Thomas Wicker, H
González, Romain

Genome Biology 2
Published on: 17 Au

› Full Text › PDF

Research

Optical and
resolution of

Numerous scaffold
improving the over

Gabriel Keeble-Ga
Frenkel, Abraham
Konkin, Pierre Sou

Genome Biology 2
Published on: 17 Au

› Full Text › PDF

Database

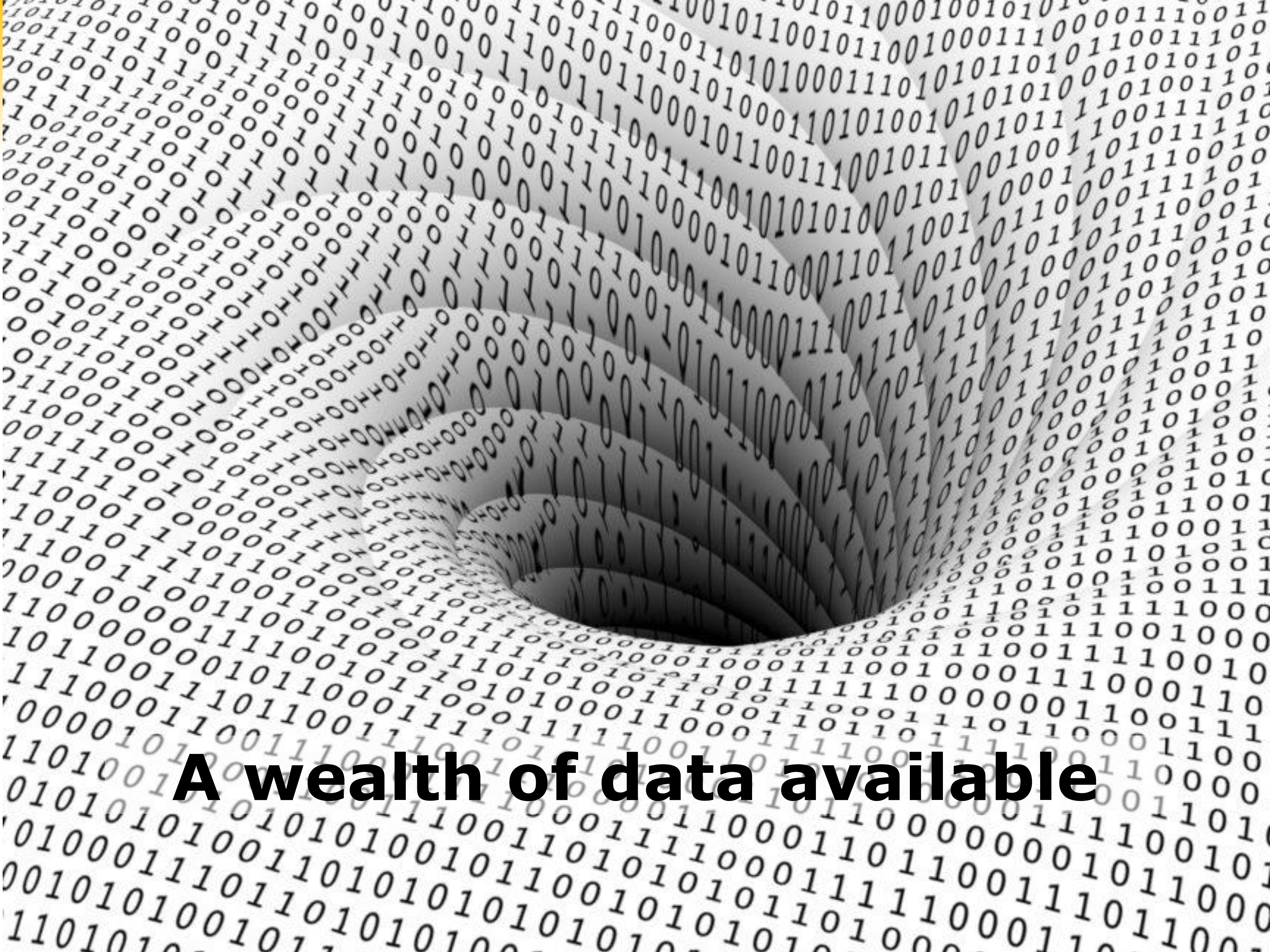
Linking the I
reference ge

The Wheat@URGI
with access to the bread wheat reference genome sequence produced by the International Wheat Genome S...

Michael Alaux, Jane Rogers, Thomas Letellier, Raphaël Flores, Françoise Alfama, Cyril Pommier, Nacer Mohellibi, Sophie Durand, Erik Kimmel, Célia Michotey, Claire Guerche, Mikaël Loaec, Mathilde Lainé, Delphine Steinbach, Frédéric Choulet, Hélène Rimbert...

Genome Biology 2018 19:111
Published on: 17 August 2018

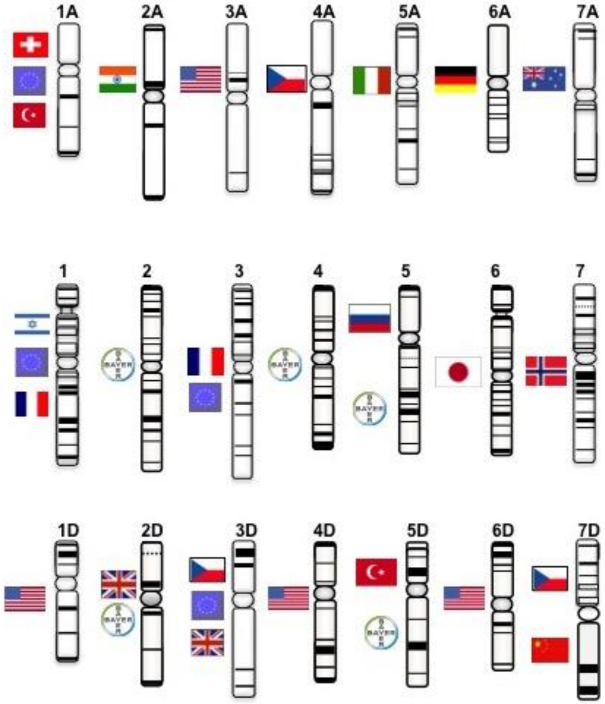
Alaux et al., *Genome Biology* 2018
https://doi.org/10.1186/s13059-018-1491-4



A wealth of data available

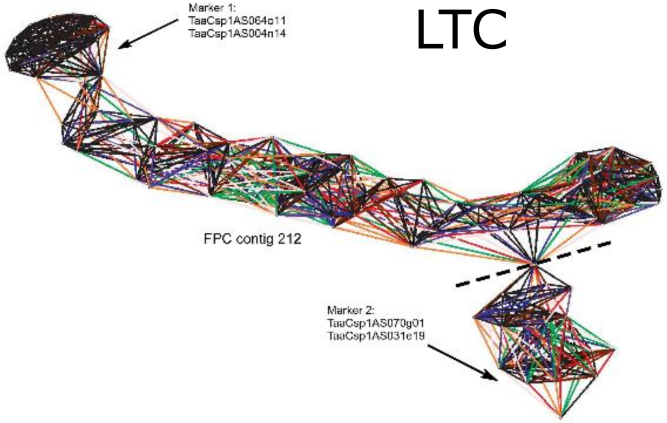
Other IWGSC data hosted

Physical maps



T. aestivum

cv Chinese Spring



LTC

Frenkel et al., *BMC Bioinformatics* 2010
https://doi.org/10.1186/1471-2105-11-584

Other IWGSC data hosted

Expression data

Transcriptome atlas

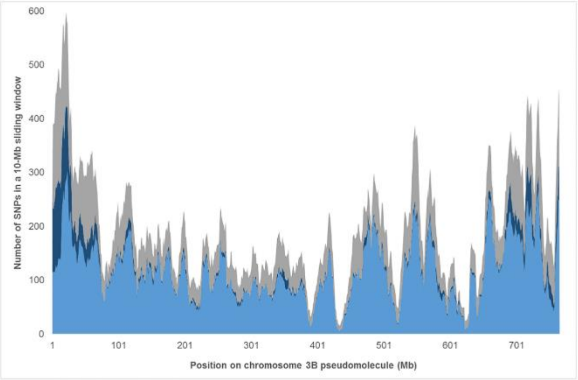
RNA-seq mapping of 850 samples – 32 tissues

- *85% HC genes expressed*
- *49% LC genes expressed*

Ramírez-González et al., *Science* 2018
https://doi.org/10.1126/science.aar6089

Variation data

SNPs
GBS
Exome capture
Varietal SNPs



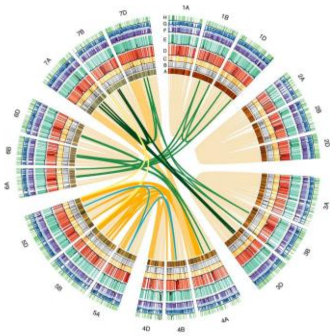
Rimbert et al., *PLOS ONE* 2018
https://doi.org/10.1371/journal.pone.0186329

Dubcovsky, Akhunov

Wheat gene pool

Other CS assembly

TGACv1



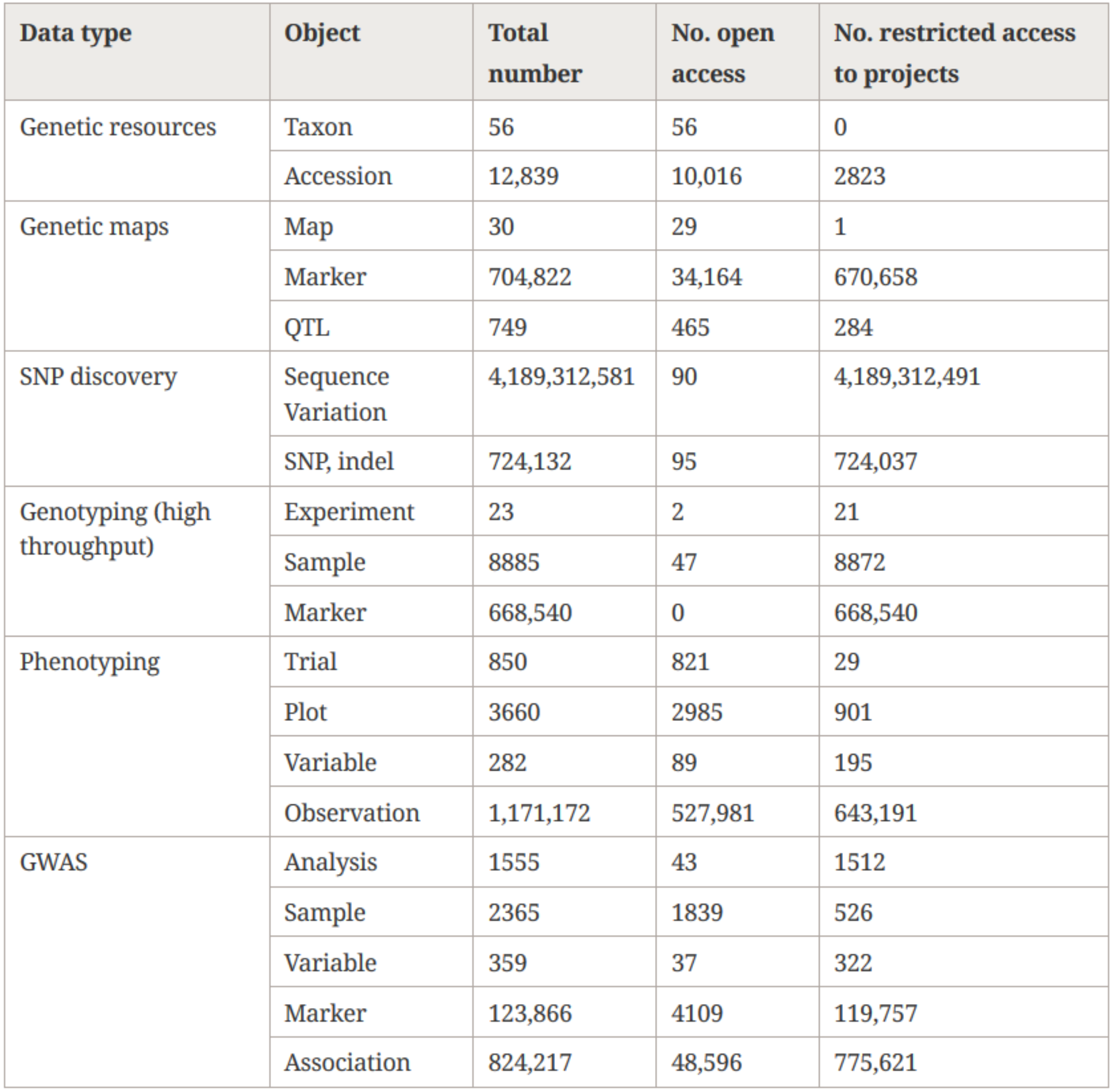
Clavijo et al., *Genome Research* 2017
https://doi.org/10.1101/gr.217117.116

Other wheat species

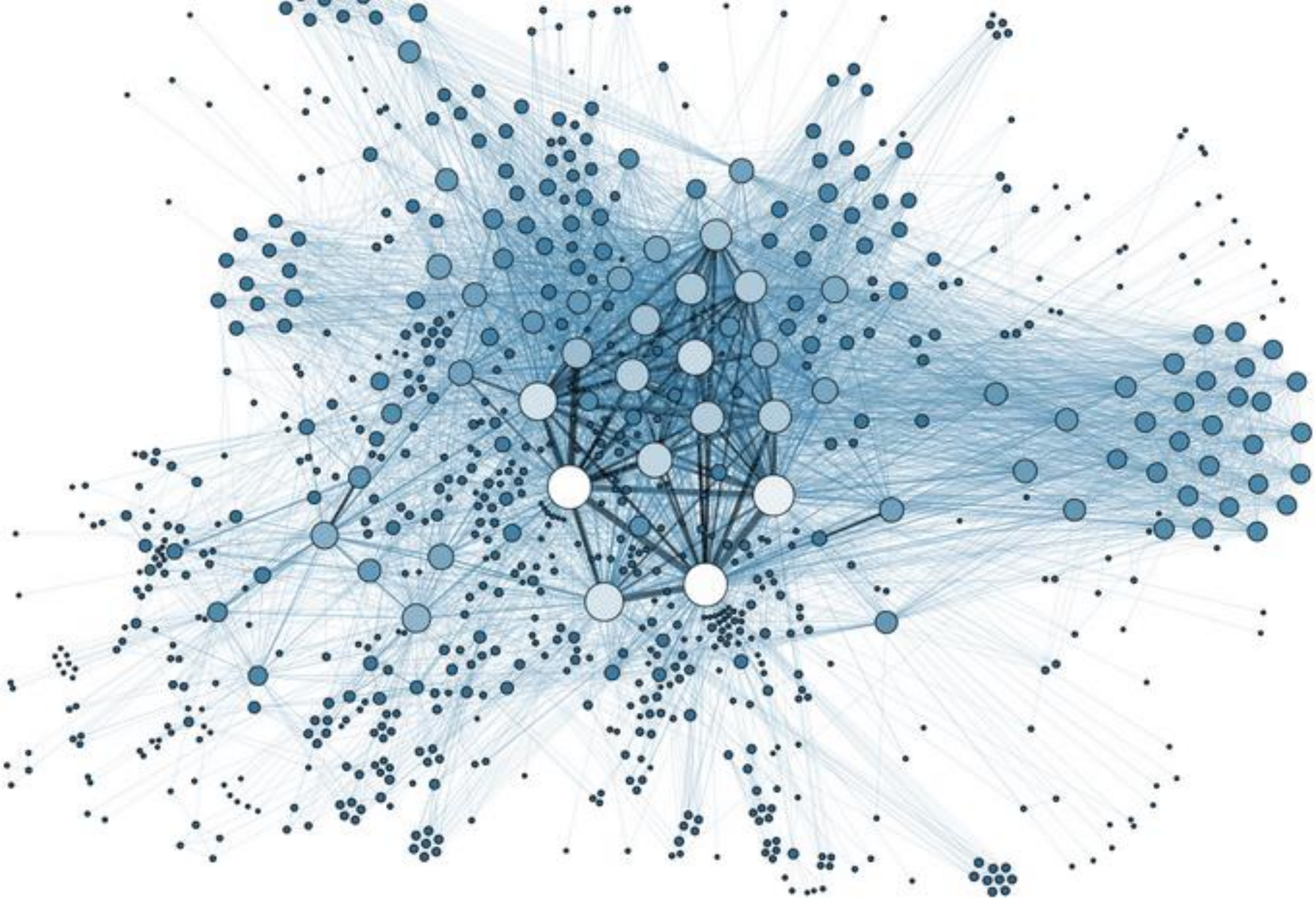
- Diploid progenitor of Aegilops taushii
- Triticum durum
- etc.

Wheat genetic and phenomic data

from French and European projects since 2000



Data type	Object	Total number	No. open access	No. restricted access to projects
Genetic resources	Taxon	56	56	0
	Accession	12,839	10,016	2823
Genetic maps	Map	30	29	1
	Marker	704,822	34,164	670,658
	QTL	749	465	284
SNP discovery	Sequence Variation	4,189,312,581	90	4,189,312,491
	SNP, indel	724,132	95	724,037
Genotyping (high throughput)	Experiment	23	2	21
	Sample	8885	47	8872
	Marker	668,540	0	668,540
Phenotyping	Trial	850	821	29
	Plot	3660	2985	901
	Variable	282	89	195
	Observation	1,171,172	527,981	643,191
GWAS	Analysis	1555	43	1512
	Sample	2365	1839	526
	Variable	359	37	322
	Marker	123,866	4109	119,757
	Association	824,217	48,596	775,621



Data integration

Make your data great

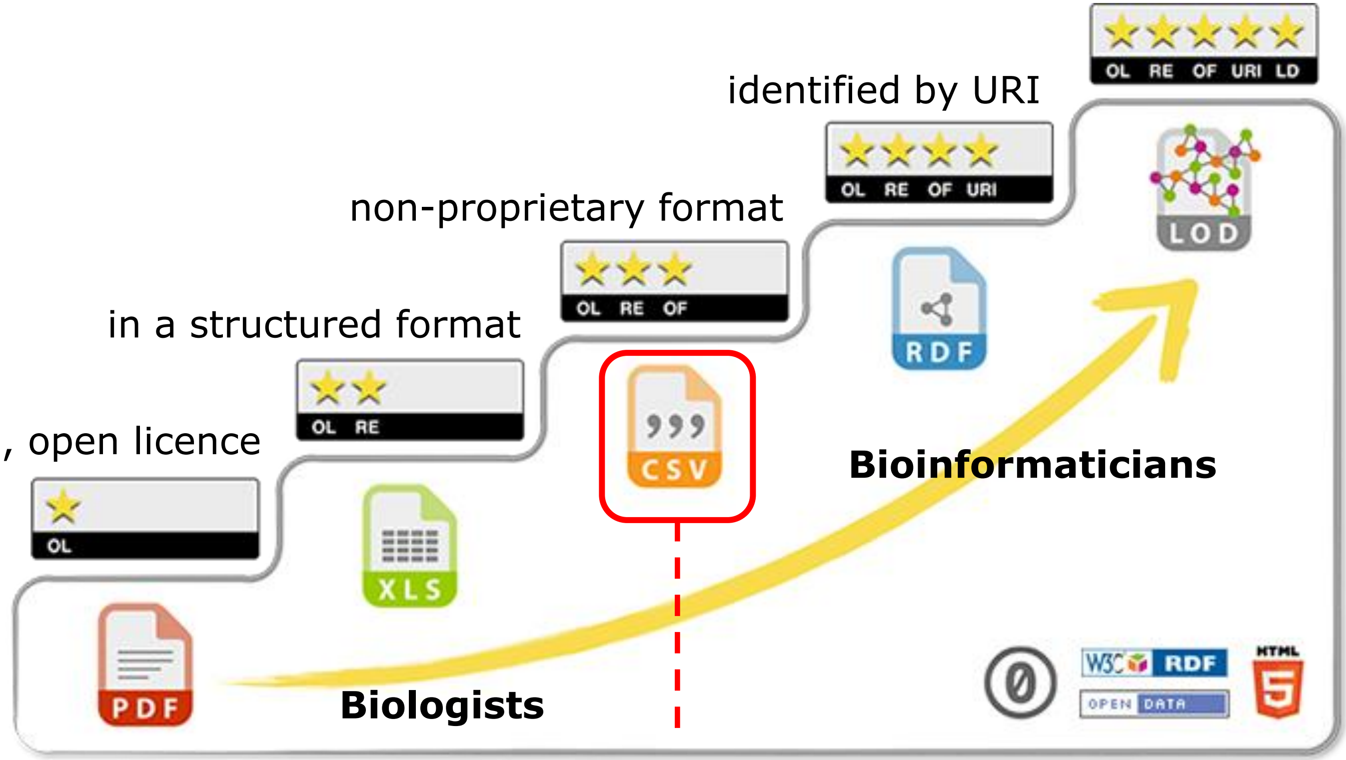
Persistent identifier (e.g. URI, DOI)
Metadata
Indexed

Data portal
Open data

Web API
Metadata
DMP
Ontologies

Standards
Licence
Documentation
Publication

5 ★Open Data



related to other data

identified by URI

non-proprietary format

in a structured format

Data on the web, open licence

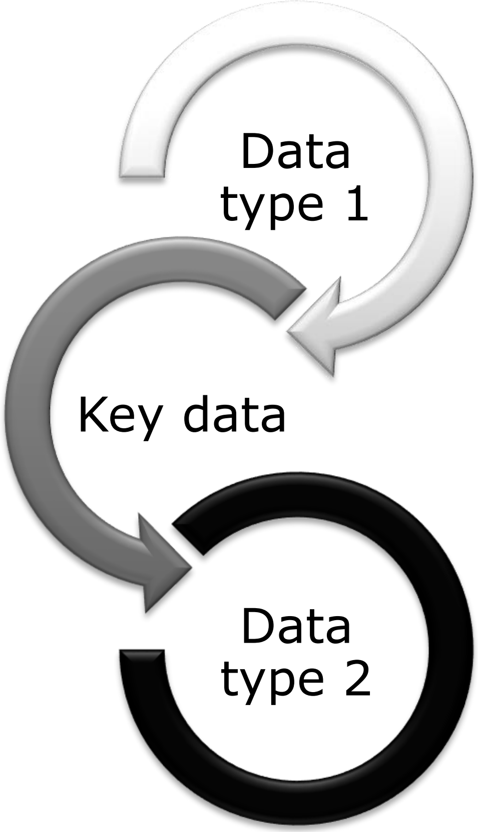
Bioinformaticians

Biologists

Courtesy of Daniel Jacob

Integration using key data

Key data are interoperability « pivots »



Data type 1

Key data

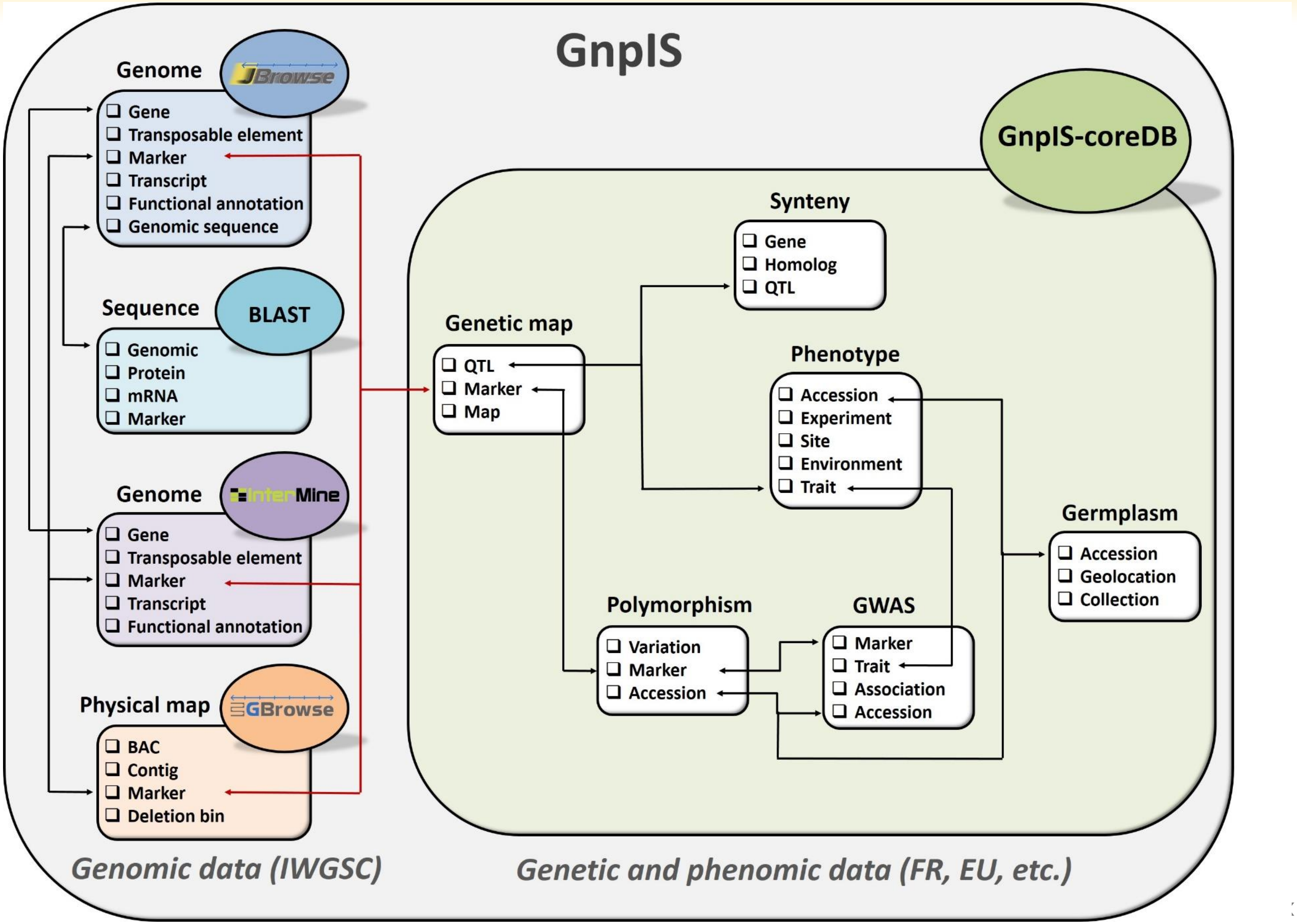
Data type 2

Phenotype

Trait

GWAS

Data integration in GnpIS



GnpIS

Genome JBrowse

☐ Gene
☐ Transposable element
☐ Marker
☐ Transcript
☐ Functional annotation
☐ Genomic sequence

Sequence BLAST

☐ Genomic
☐ Protein
☐ mRNA
☐ Marker

Genome InterMine

☐ Gene
☐ Transposable element
☐ Marker
☐ Transcript
☐ Functional annotation

Physical map GBrowse

☐ BAC
☐ Contig
☐ Marker
☐ Deletion bin

Genomic data (IWGSC)

GnpIS-coreDB

Genetic map

☐ QTL
☐ Marker
☐ Map

Synteny

☐ Gene
☐ Homolog
☐ QTL

Phenotype

☐ Accession
☐ Experiment
☐ Site
☐ Environment
☐ Trait

Germplasm

☐ Accession
☐ Geolocation
☐ Collection

Polymorphism

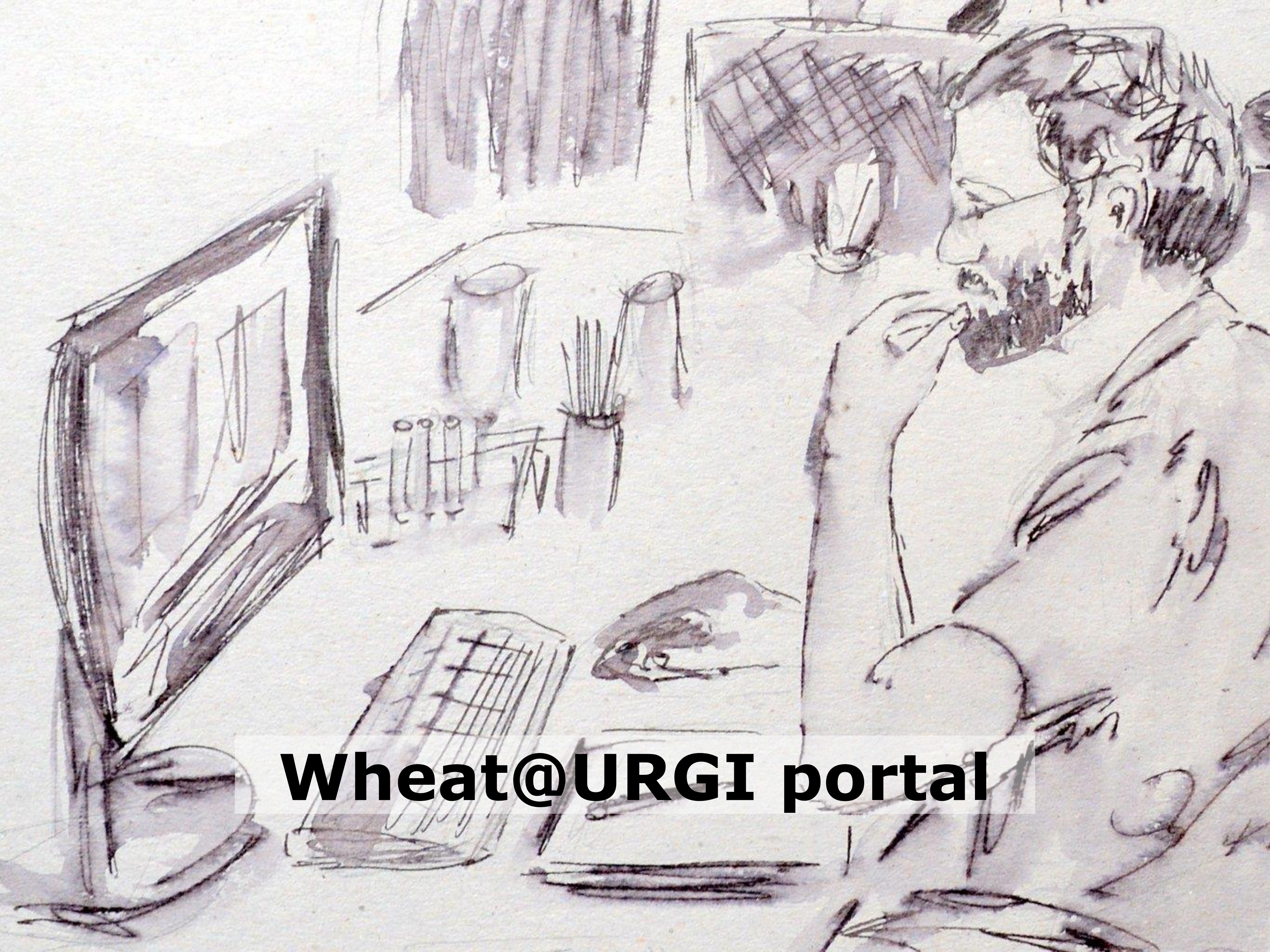
☐ Variation
☐ Marker
☐ Accession

GWAS

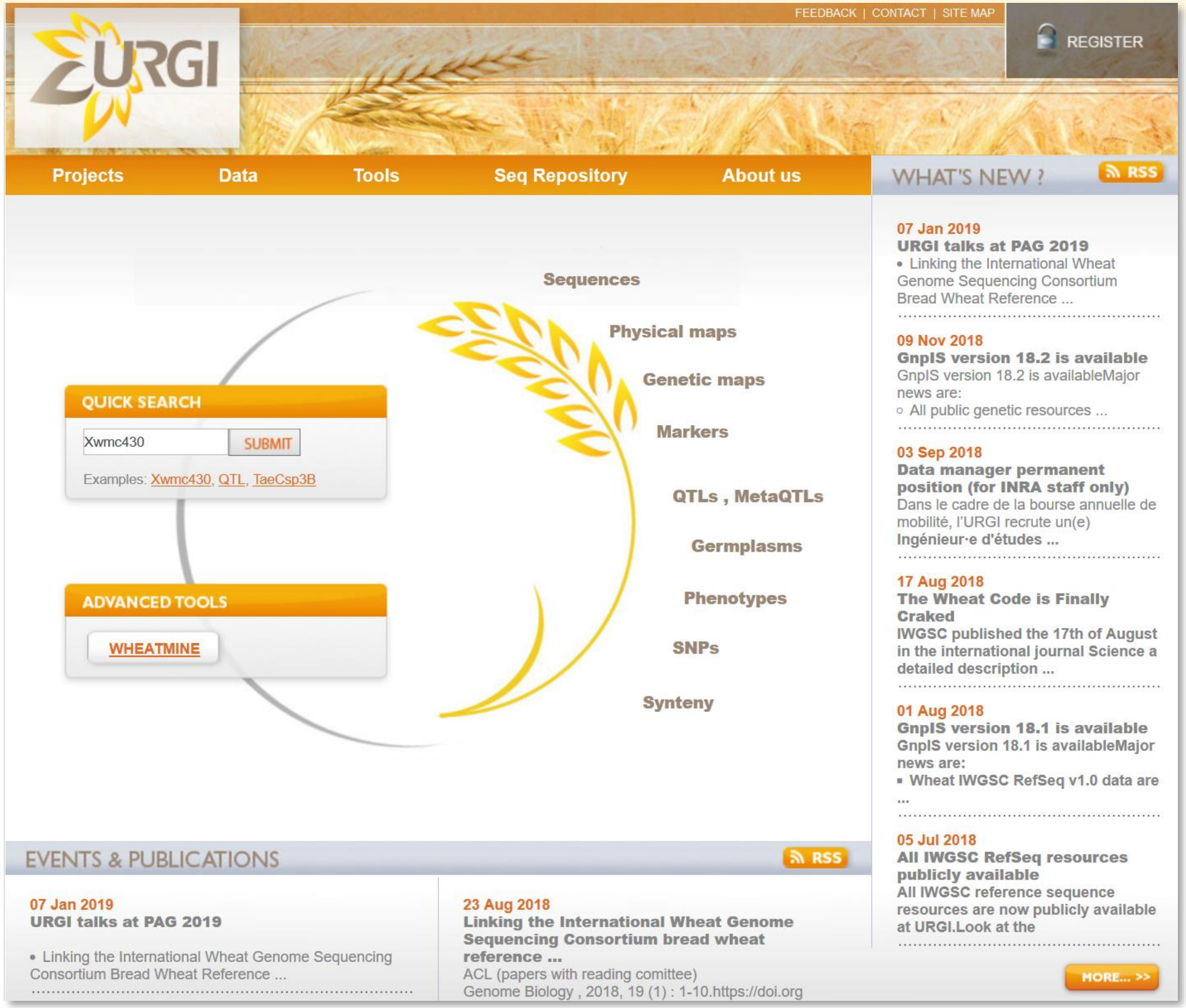
☐ Marker
☐ Trait
☐ Association
☐ Accession

Genetic and phenomic data (FR, EU, etc.)

Wheat@URGI portal



Website



Examples: Xwmc430, QTL, TaeCsp3B

WHAT'S NEW ?

07 Jan 2019
URGI talks at PAG 2019
- Linking the International Wheat Genome Sequencing Consortium Bread Wheat Reference ...

09 Nov 2018
GnpIS version 18.2 is available
GnpIS version 18.2 is availableMajor news are:
- All public genetic resources ...

03 Sep 2018
Data manager permanent position (for INRA staff only)
Dans le cadre de la bourse annuelle de mobilité, l'URGI recrute un(e) Ingénieur·e d'études ...

17 Aug 2018
The Wheat Code is Finally Craked
IWGSC published the 17th of August in the international journal Science a detailed description ...

01 Aug 2018
GnpIS version 18.1 is available
GnpIS version 18.1 is availableMajor news are:
- Wheat IWGSC RefSeq v1.0 data are ...

05 Jul 2018
All IWGSC RefSeq resources publicly available
All IWGSC reference sequence resources are now publicly available at URGI.Look at the

EVENTS & PUBLICATIONS

07 Jan 2019
URGI talks at PAG 2019
- Linking the International Wheat Genome Sequencing Consortium Bread Wheat Reference ...

23 Aug 2018
Linking the International Wheat Genome Sequencing Consortium bread wheat reference ...
ACL (papers with reading comittee)
Genome Biology , 2018, 19 (1) : 1-10.https://doi.org

https://wheat-urgi.versailles.inra.fr/

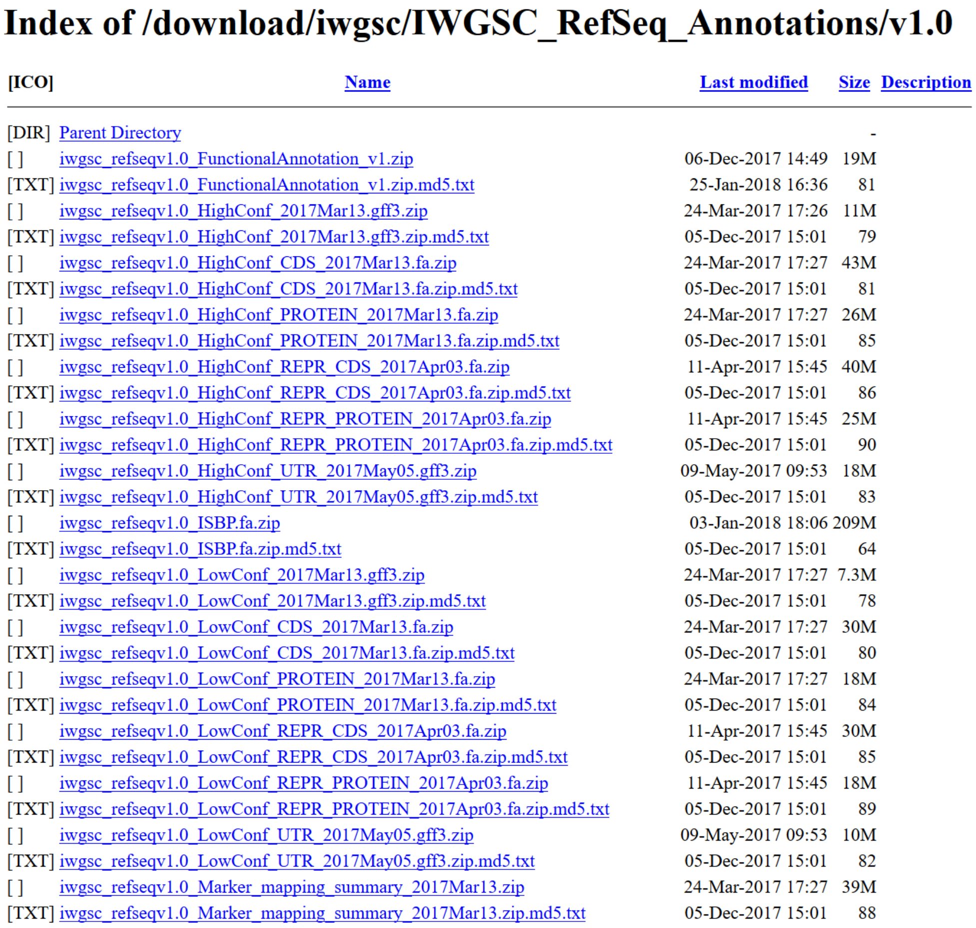
Michael Alaux

Tools - developed by URGI

IWGSC Data Repository

Webpages + File Download including README, MD5

27k downloads in 2018



Index of /download/iwgsc/IWGSC_RefSeq_Annotations/v1.0

[ICO]	Name	Last modified	Size	Description
[DIR]	Parent Directory		-	
[ ]	iwgsc_refseqv1.0_FunctionalAnnotation_v1.zip	06-Dec-2017 14:49	19M	
[TXT]	iwgsc_refseqv1.0_FunctionalAnnotation_v1.zip.md5.txt	25-Jan-2018 16:36	81	
[ ]	iwgsc_refseqv1.0_HighConf_2017Mar13.gff3.zip	24-Mar-2017 17:26	11M	
[TXT]	iwgsc_refseqv1.0_HighConf_2017Mar13.gff3.zip.md5.txt	05-Dec-2017 15:01	79	
[ ]	iwgsc_refseqv1.0_HighConf_CDS_2017Mar13.fa.zip	24-Mar-2017 17:27	43M	
[TXT]	iwgsc_refseqv1.0_HighConf_CDS_2017Mar13.fa.zip.md5.txt	05-Dec-2017 15:01	81	
[ ]	iwgsc_refseqv1.0_HighConf_PROTEIN_2017Mar13.fa.zip	24-Mar-2017 17:27	26M	
[TXT]	iwgsc_refseqv1.0_HighConf_PROTEIN_2017Mar13.fa.zip.md5.txt	05-Dec-2017 15:01	85	
[ ]	iwgsc_refseqv1.0_HighConf_REPR_CDS_2017Apr03.fa.zip	11-Apr-2017 15:45	40M	
[TXT]	iwgsc_refseqv1.0_HighConf_REPR_CDS_2017Apr03.fa.zip.md5.txt	05-Dec-2017 15:01	86	
[ ]	iwgsc_refseqv1.0_HighConf_REPR_PROTEIN_2017Apr03.fa.zip	11-Apr-2017 15:45	25M	
[TXT]	iwgsc_refseqv1.0_HighConf_REPR_PROTEIN_2017Apr03.fa.zip.md5.txt	05-Dec-2017 15:01	90	
[ ]	iwgsc_refseqv1.0_HighConf_UTR_2017May05.gff3.zip	09-May-2017 09:53	18M	
[TXT]	iwgsc_refseqv1.0_HighConf_UTR_2017May05.gff3.zip.md5.txt	05-Dec-2017 15:01	83	
[ ]	iwgsc_refseqv1.0_ISBP.fa.zip	03-Jan-2018 18:06	209M	
[TXT]	iwgsc_refseqv1.0_ISBP.fa.zip.md5.txt	05-Dec-2017 15:01	64	
[ ]	iwgsc_refseqv1.0_LowConf_2017Mar13.gff3.zip	24-Mar-2017 17:27	7.3M	
[TXT]	iwgsc_refseqv1.0_LowConf_2017Mar13.gff3.zip.md5.txt	05-Dec-2017 15:01	78	
[ ]	iwgsc_refseqv1.0_LowConf_CDS_2017Mar13.fa.zip	24-Mar-2017 17:27	30M	
[TXT]	iwgsc_refseqv1.0_LowConf_CDS_2017Mar13.fa.zip.md5.txt	05-Dec-2017 15:01	80	
[ ]	iwgsc_refseqv1.0_LowConf_PROTEIN_2017Mar13.fa.zip	24-Mar-2017 17:27	18M	
[TXT]	iwgsc_refseqv1.0_LowConf_PROTEIN_2017Mar13.fa.zip.md5.txt	05-Dec-2017 15:01	84	
[ ]	iwgsc_refseqv1.0_LowConf_REPR_CDS_2017Apr03.fa.zip	11-Apr-2017 15:45	30M	
[TXT]	iwgsc_refseqv1.0_LowConf_REPR_CDS_2017Apr03.fa.zip.md5.txt	05-Dec-2017 15:01	85	
[ ]	iwgsc_refseqv1.0_LowConf_REPR_PROTEIN_2017Apr03.fa.zip	11-Apr-2017 15:45	18M	
[TXT]	iwgsc_refseqv1.0_LowConf_REPR_PROTEIN_2017Apr03.fa.zip.md5.txt	05-Dec-2017 15:01	89	
[ ]	iwgsc_refseqv1.0_LowConf_UTR_2017May05.gff3.zip	09-May-2017 09:53	10M	
[TXT]	iwgsc_refseqv1.0_LowConf_UTR_2017May05.gff3.zip.md5.txt	05-Dec-2017 15:01	82	
[ ]	iwgsc_refseqv1.0_Marker_mapping_summary_2017Mar13.zip	24-Mar-2017 17:27	39M	
[TXT]	iwgsc_refseqv1.0_Marker_mapping_summary_2017Mar13.zip.md5.txt	05-Dec-2017 15:01	88	

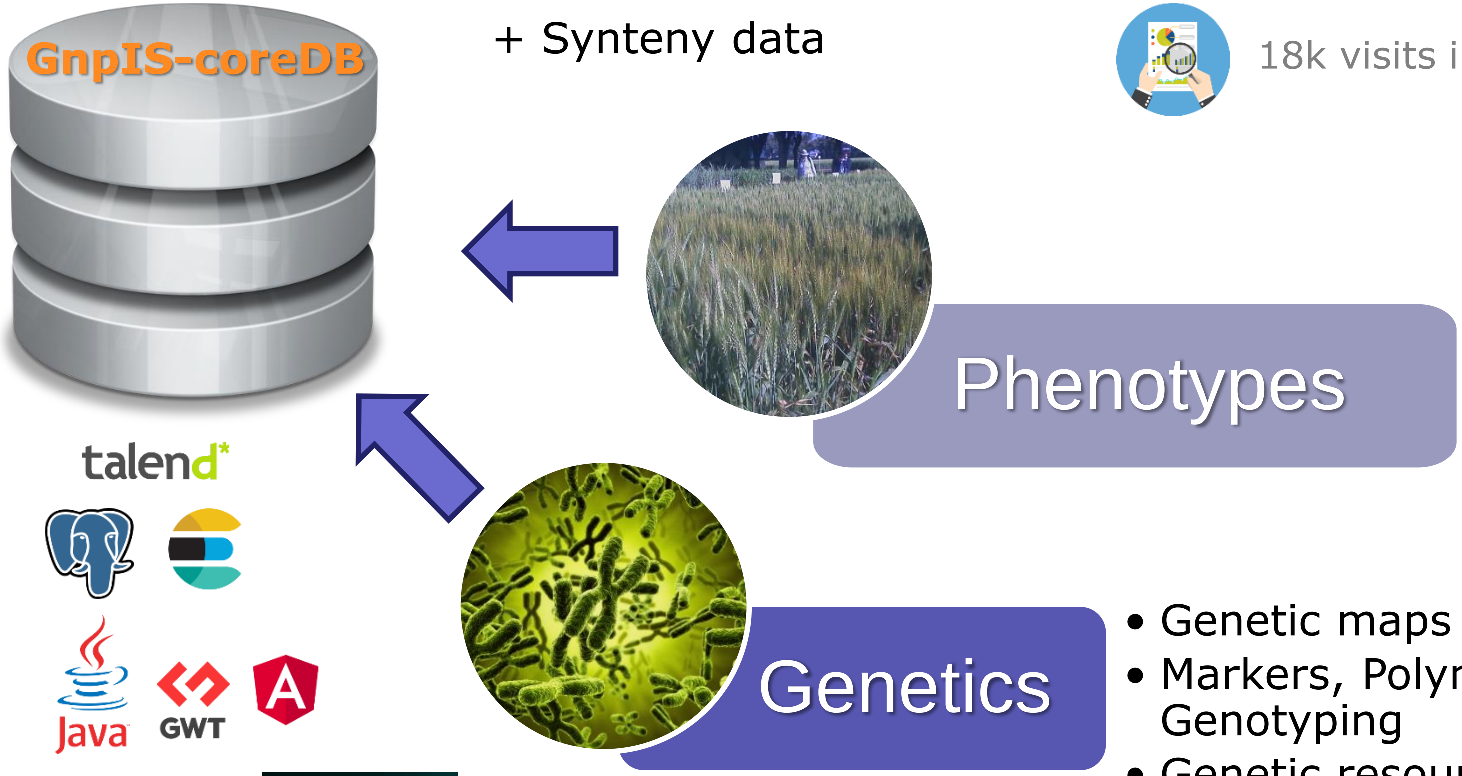
Tools - developed by URGI

GnpIS-coreDB

Information System to manage genetic and phenomic data

+ Synteny data

18k visits in 2018



- GxE
- QTLs
- GWAS

- Genetic maps
- Markers, Polymorphisms, Genotyping
- Genetic resources

talend

Java GWT

miappe

BrAPI

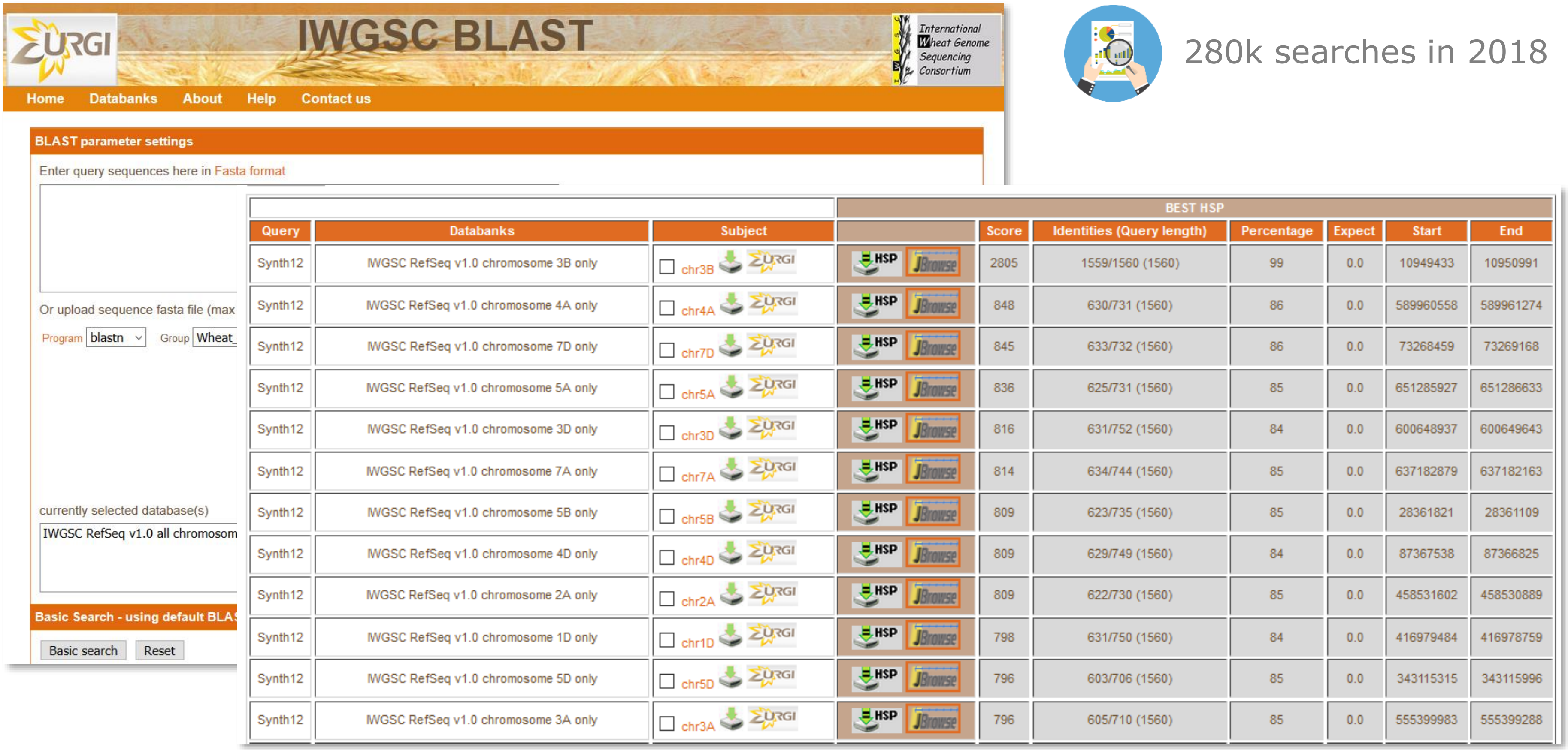
Findable Accessible Interoperable Reusable

URGI

Michael Alaux

Tools – instantiated by URGI

BLAST



IWGSC BLAST

International Wheat Genome Sequencing Consortium

BLAST parameter settings

Enter query sequences here in Fasta format

Or upload sequence fasta file (max

Program: blastn Group: Wheat_

currently selected database(s)

IWGSC RefSeq v1.0 all chromosom

Basic Search - using default BLAS

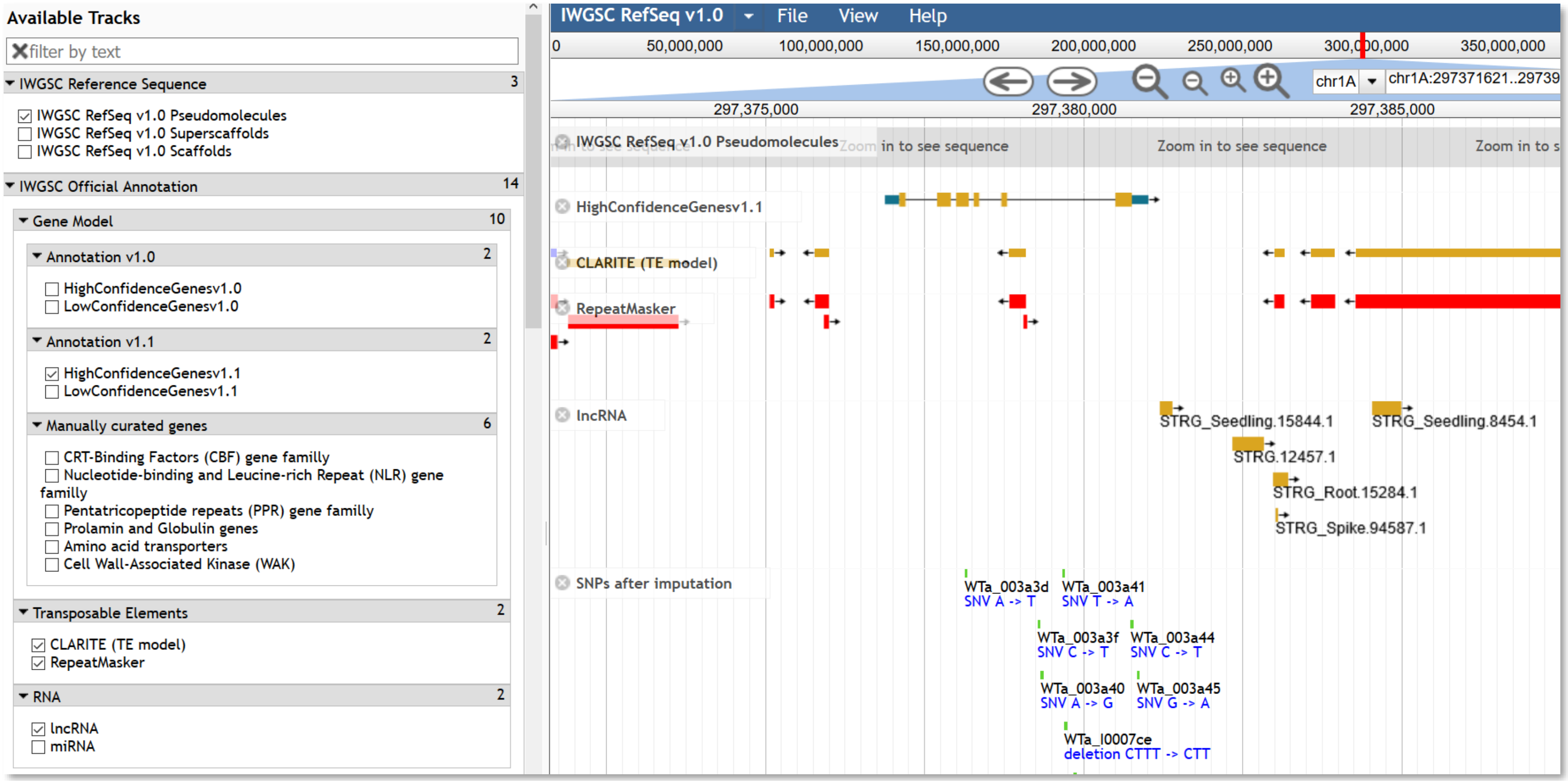
			BEST HSP						
Query	Databanks	Subject		Score	Identities (Query length)	Percentage	Expect	Start	End
Synth12	IWGSC RefSeq v1.0 chromosome 3B only	☐ chr3B	HSP JBrowse	2805	1559/1560 (1560)	99	0.0	10949433	10950991
Synth12	IWGSC RefSeq v1.0 chromosome 4A only	☐ chr4A	HSP JBrowse	848	630/731 (1560)	86	0.0	589960558	589961274
Synth12	IWGSC RefSeq v1.0 chromosome 7D only	☐ chr7D	HSP JBrowse	845	633/732 (1560)	86	0.0	73268459	73269168
Synth12	IWGSC RefSeq v1.0 chromosome 5A only	☐ chr5A	HSP JBrowse	836	625/731 (1560)	85	0.0	651285927	651286633
Synth12	IWGSC RefSeq v1.0 chromosome 3D only	☐ chr3D	HSP JBrowse	816	631/752 (1560)	84	0.0	600648937	600649643
Synth12	IWGSC RefSeq v1.0 chromosome 7A only	☐ chr7A	HSP JBrowse	814	634/744 (1560)	85	0.0	637182879	637182163
Synth12	IWGSC RefSeq v1.0 chromosome 5B only	☐ chr5B	HSP JBrowse	809	623/735 (1560)	85	0.0	28361821	28361109
Synth12	IWGSC RefSeq v1.0 chromosome 4D only	☐ chr4D	HSP JBrowse	809	629/749 (1560)	84	0.0	87367538	87366825
Synth12	IWGSC RefSeq v1.0 chromosome 2A only	☐ chr2A	HSP JBrowse	809	622/730 (1560)	85	0.0	458531602	458530889
Synth12	IWGSC RefSeq v1.0 chromosome 1D only	☐ chr1D	HSP JBrowse	798	631/750 (1560)	84	0.0	416979484	416978759
Synth12	IWGSC RefSeq v1.0 chromosome 5D only	☐ chr5D	HSP JBrowse	796	603/706 (1560)	85	0.0	343115315	343115996
Synth12	IWGSC RefSeq v1.0 chromosome 3A only	☐ chr3A	HSP JBrowse	796	605/710 (1560)	85	0.0	555399983	555399288

280k searches in 2018

Tools – instantiated by URGI

JBrowse

40k visits in 2018



Available Tracks

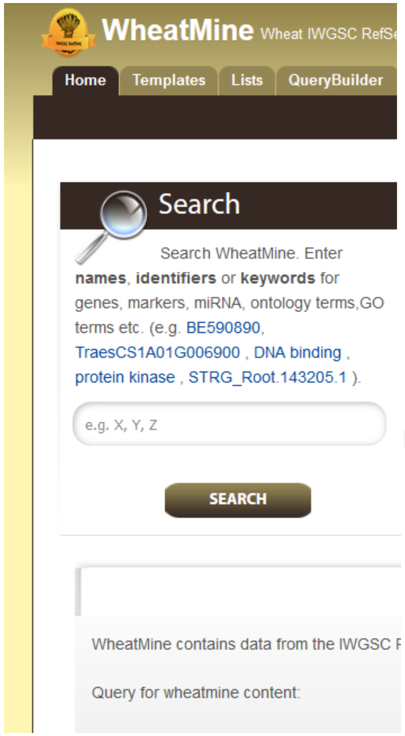
filter by text

- IWGSC Reference Sequence 3
  - ☑ IWGSC RefSeq v1.0 Pseudomolecules
  - ☐ IWGSC RefSeq v1.0 Superscaffolds
  - ☐ IWGSC RefSeq v1.0 Scaffolds
- IWGSC Official Annotation 14
  - Gene Model 10
    - Annotation v1.0 2
      - ☐ HighConfidenceGenesv1.0
      - ☐ LowConfidenceGenesv1.0
    - Annotation v1.1 2
      - ☑ HighConfidenceGenesv1.1
      - ☐ LowConfidenceGenesv1.1
    - Manually curated genes 6
      - ☐ CRT-Binding Factors (CBF) gene familly
      - ☐ Nucleotide-binding and Leucine-rich Repeat (NLR) gene familly
      - ☐ Pentatricopeptide repeats (PPR) gene familly
      - ☐ Prolamin and Globulin genes
      - ☐ Amino acid transporters
      - ☐ Cell Wall-Associated Kinase (WAK)
  - Transposable Elements 2
    - ☑ CLARITE (TE model)
    - ☑ RepeatMasker
  - RNA 2
    - ☑ lncRNA
    - ☐ miRNA

Michael Alaux

Tools – instantiated by URGI

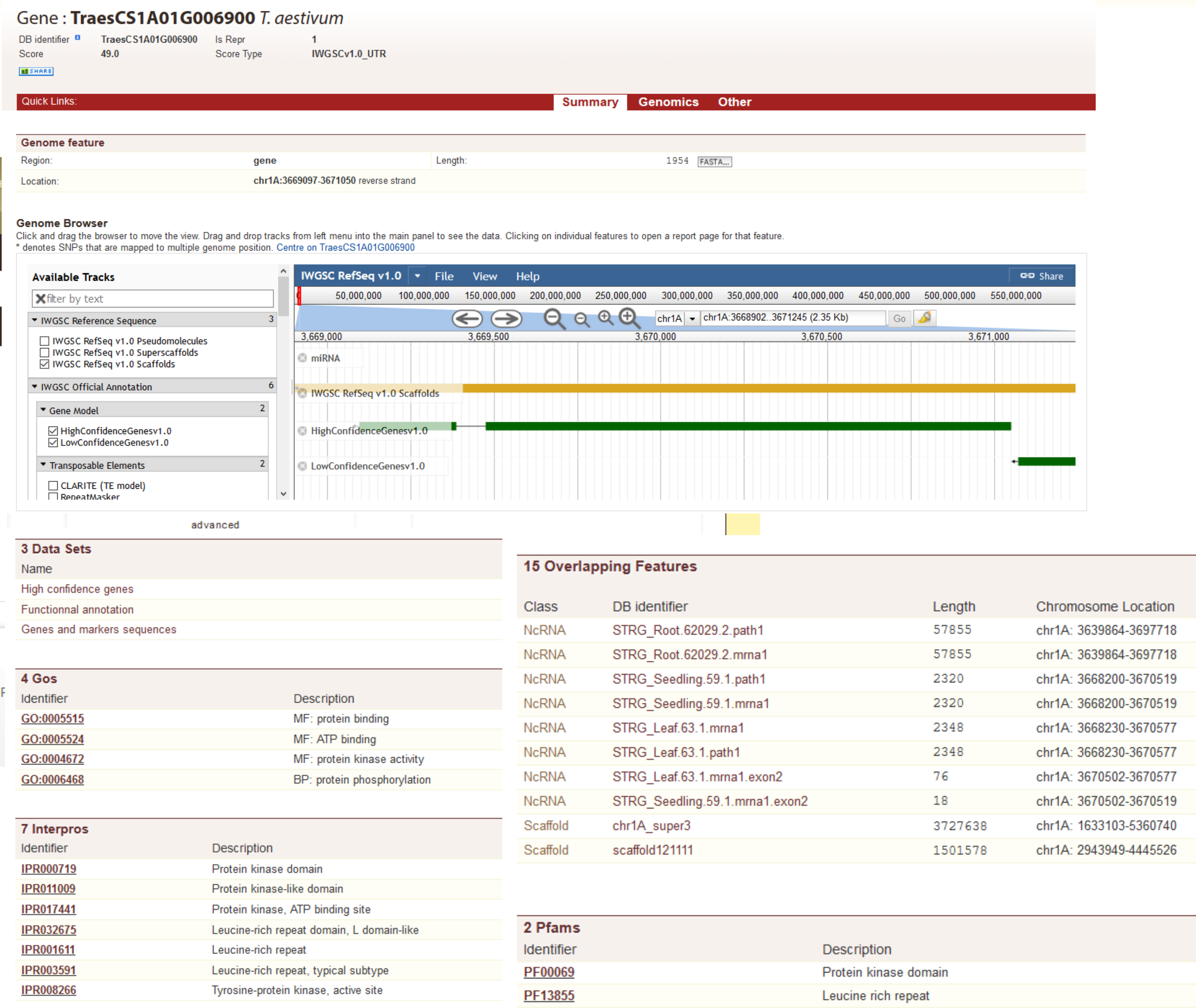
InterMine



Search WheatMine. Enter **names**, **identifiers** or **keywords** for genes, markers, miRNA, ontology terms,GO terms etc. (e.g. BE590890, TraesCS1A01G006900 , DNA binding , protein kinase , STRG_Root.143205.1).

WheatMine contains data from the IWGSC R

Query for wheatmine content:



Gene : **TraesCS1A01G006900** *T. aestivum*

DB identifier	TraesCS1A01G006900	Is Repr	1
Score	49.0	Score Type	IWGSCv1.0_UTR

Genome feature

Region:	gene	Length:	1954
Location:	chr1A:3669097-3671050 reverse strand		

Genome Browser

Click and drag the browser to move the view. Drag and drop tracks from left menu into the main panel to see the data. Clicking on individual features to open a report page for that feature.
* denotes SNPs that are mapped to multiple genome position. Centre on TraesCS1A01G006900

Available Tracks

filter by text

- IWGSC Reference Sequence 3
  - ☐ IWGSC RefSeq v1.0 Pseudomolecules
  - ☐ IWGSC RefSeq v1.0 Superscaffolds
  - ☑ IWGSC RefSeq v1.0 Scaffolds
- IWGSC Official Annotation 6
  - Gene Model 2
    - ☑ HighConfidenceGenesv1.0
    - ☑ LowConfidenceGenesv1.0
  - Transposable Elements 2
    - ☐ CLARITE (TE model)
    - ☐ RepeatMasker

3 Data Sets

Name
High confidence genes
Functional annotation
Genes and markers sequences

4 Gos

Identifier	Description
GO:0005515	MF: protein binding
GO:0005524	MF: ATP binding
GO:0004672	MF: protein kinase activity
GO:0006468	BP: protein phosphorylation

7 Interpros

Identifier	Description
IPR000719	Protein kinase domain
IPR011009	Protein kinase-like domain
IPR017441	Protein kinase, ATP binding site
IPR032675	Leucine-rich repeat domain, L domain-like
IPR001611	Leucine-rich repeat
IPR003591	Leucine-rich repeat, typical subtype
IPR008266	Tyrosine-protein kinase, active site

15 Overlapping Features

Class	DB identifier	Length	Chromosome Location
NcRNA	STRG_Root.62029.2.path1	57855	chr1A: 3639864-3697718
NcRNA	STRG_Root.62029.2.mrna1	57855	chr1A: 3639864-3697718
NcRNA	STRG_Seedling.59.1.path1	2320	chr1A: 3668200-3670519
NcRNA	STRG_Seedling.59.1.mrna1	2320	chr1A: 3668200-3670519
NcRNA	STRG_Leaf.63.1.mrna1	2348	chr1A: 3668230-3670577
NcRNA	STRG_Leaf.63.1.path1	2348	chr1A: 3668230-3670577
NcRNA	STRG_Leaf.63.1.mrna1.exon2	76	chr1A: 3670502-3670577
NcRNA	STRG_Seedling.59.1.mrna1.exon2	18	chr1A: 3670502-3670519
Scaffold	chr1A_super3	3727638	chr1A: 1633103-5360740
Scaffold	scaffold121111	1501578	chr1A: 2943949-4445526

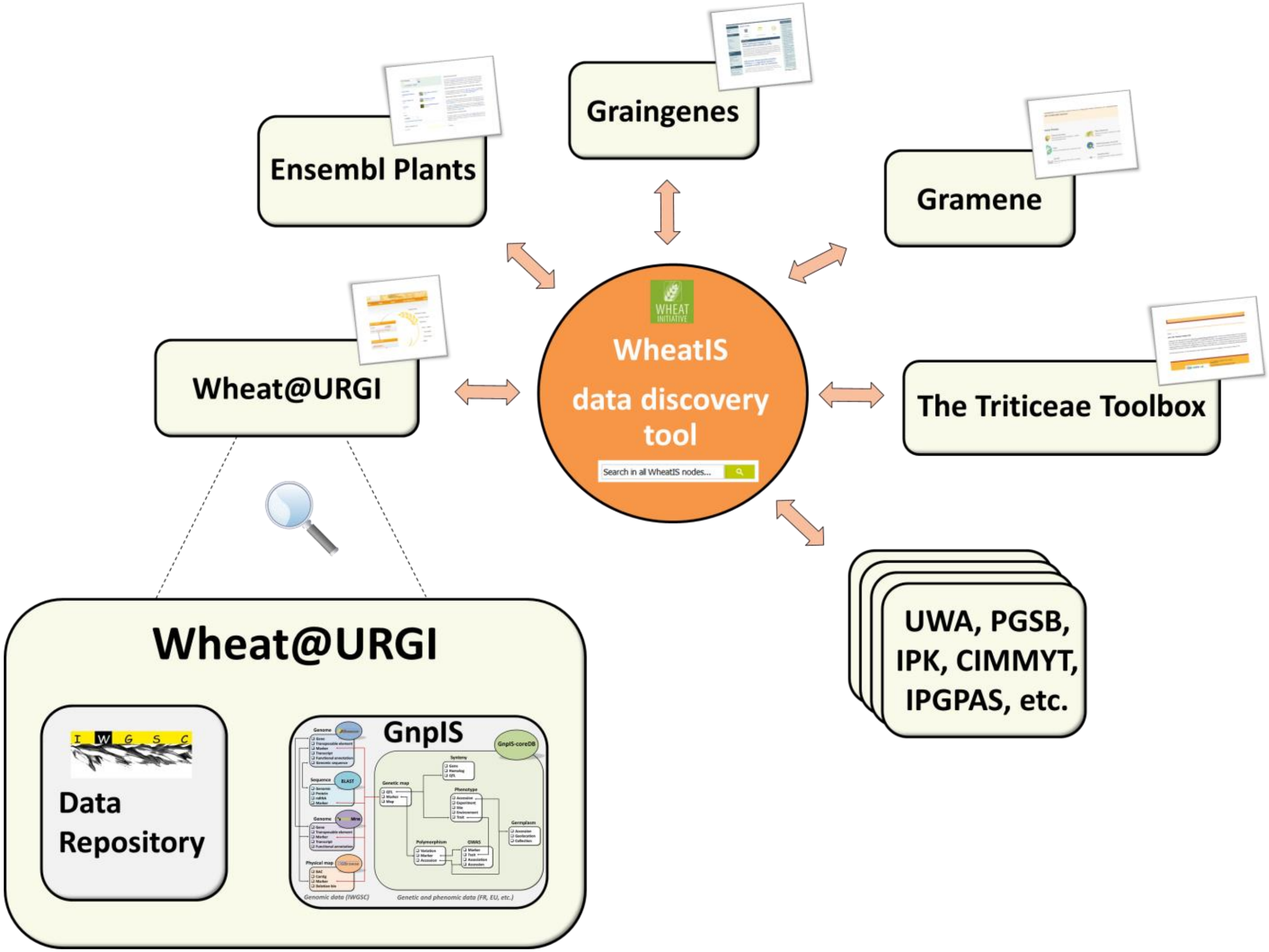
2 Pfams

Identifier	Description
PF00069	Protein kinase domain
PF13855	Leucine rich repeat

4k visits in 2018

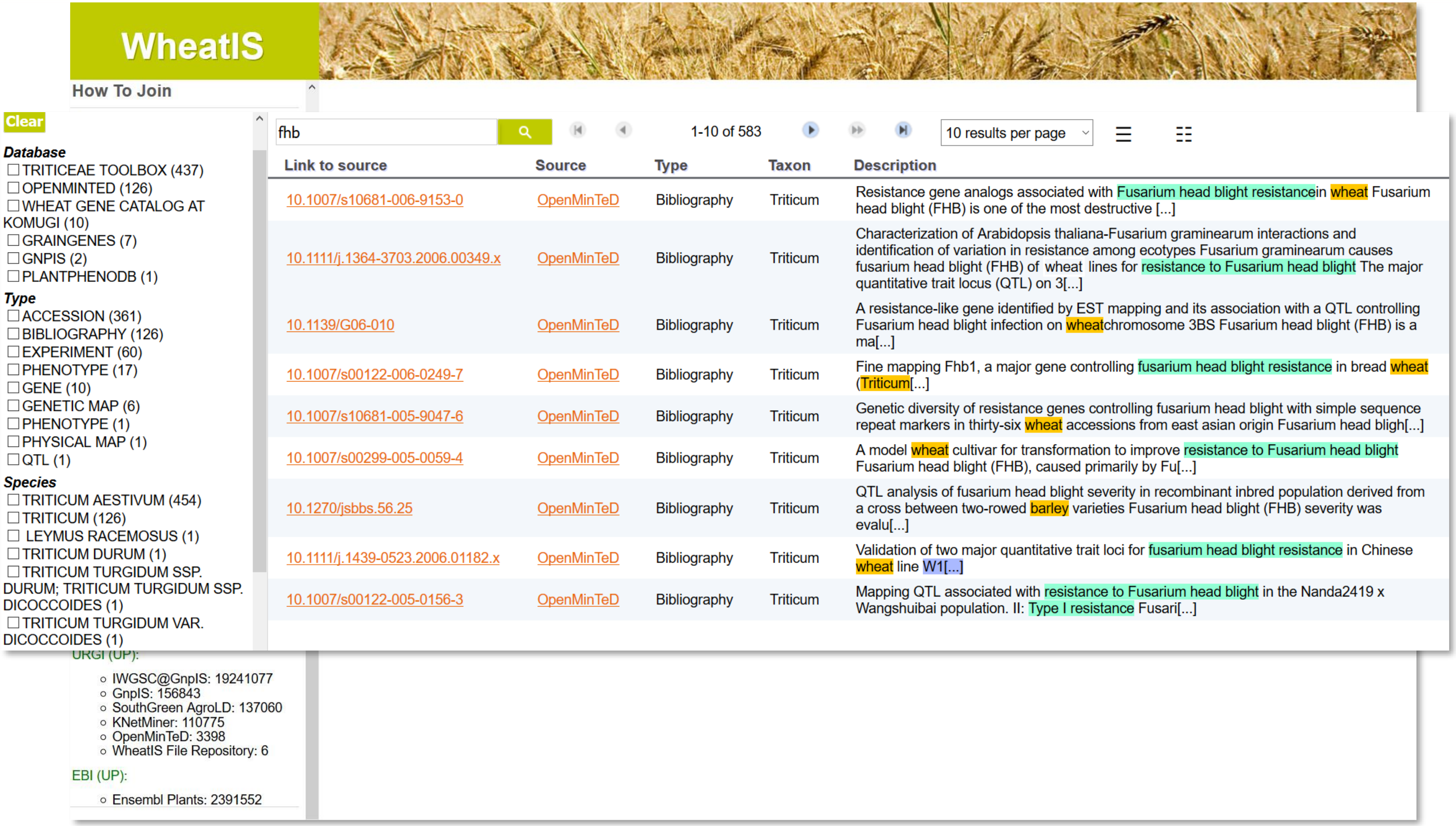
Tools – collaborative development

WheatIS data discovery



Michael Alaux

Tools – collaborative development



Database

☐ TRITICEAE TOOLBOX (437)
☐ OPENMINTED (126)
☐ WHEAT GENE CATALOG AT KOMUGI (10)
☐ GRAINGENES (7)
☐ GNPIS (2)
☐ PLANTPHENODB (1)

Type

☐ ACCESSION (361)
☐ BIBLIOGRAPHY (126)
☐ EXPERIMENT (60)
☐ PHENOTYPE (17)
☐ GENE (10)
☐ GENETIC MAP (6)
☐ PHENOTYPE (1)
☐ PHYSICAL MAP (1)
☐ QTL (1)

Species

☐ TRITICUM AESTIVUM (454)
☐ TRITICUM (126)
☐ LEYMUS RACEMOSUS (1)
☐ TRITICUM DURUM (1)
☐ TRITICUM TURGIDUM SSP. DURUM; TRITICUM TURGIDUM SSP. DICOCCOIDES (1)
☐ TRITICUM TURGIDUM VAR. DICOCCOIDES (1)

fhb

1-10 of 583

10 results per page

Link to source	Source	Type	Taxon	Description
10.1007/s10681-006-9153-0	OpenMinTeD	Bibliography	Triticum	Resistance gene analogs associated with Fusarium head blight resistancein wheat Fusarium head blight (FHB) is one of the most destructive [...]
10.1111/j.1364-3703.2006.00349.x	OpenMinTeD	Bibliography	Triticum	Characterization of Arabidopsis thaliana-Fusarium graminearum interactions and identification of variation in resistance among ecotypes Fusarium graminearum causes fusarium head blight (FHB) of wheat lines for resistance to Fusarium head blight The major quantitative trait locus (QTL) on 3[...]
10.1139/G06-010	OpenMinTeD	Bibliography	Triticum	A resistance-like gene identified by EST mapping and its association with a QTL controlling Fusarium head blight infection on wheatchromosome 3BS Fusarium head blight (FHB) is a ma[...]
10.1007/s00122-006-0249-7	OpenMinTeD	Bibliography	Triticum	Fine mapping Fhb1, a major gene controlling fusarium head blight resistance in bread wheat (Triticum[...]
10.1007/s10681-005-9047-6	OpenMinTeD	Bibliography	Triticum	Genetic diversity of resistance genes controlling fusarium head blight with simple sequence repeat markers in thirty-six wheat accessions from east asian origin Fusarium head bligh[...]
10.1007/s00299-005-0059-4	OpenMinTeD	Bibliography	Triticum	A model wheat cultivar for transformation to improve resistance to Fusarium head blight Fusarium head blight (FHB), caused primarily by Fu[...]
10.1270/jsbbs.56.25	OpenMinTeD	Bibliography	Triticum	QTL analysis of fusarium head blight severity in recombinant inbred population derived from a cross between two-rowed barley varieties Fusarium head blight (FHB) severity was evalu[...]
10.1111/j.1439-0523.2006.01182.x	OpenMinTeD	Bibliography	Triticum	Validation of two major quantitative trait loci for fusarium head blight resistance in Chinese wheat line W1[...]
10.1007/s00122-005-0156-3	OpenMinTeD	Bibliography	Triticum	Mapping QTL associated with resistance to Fusarium head blight in the Nanda2419 x Wangshuibai population. II: Type I resistance Fusari[...]

URGI (UP):

- IWGSC@GnpIS: 19241077
- GnpIS: 156843
- SouthGreen AgroLD: 137060
- KNetMiner: 110775
- OpenMinTeD: 3398
- WheatIS File Repository: 6

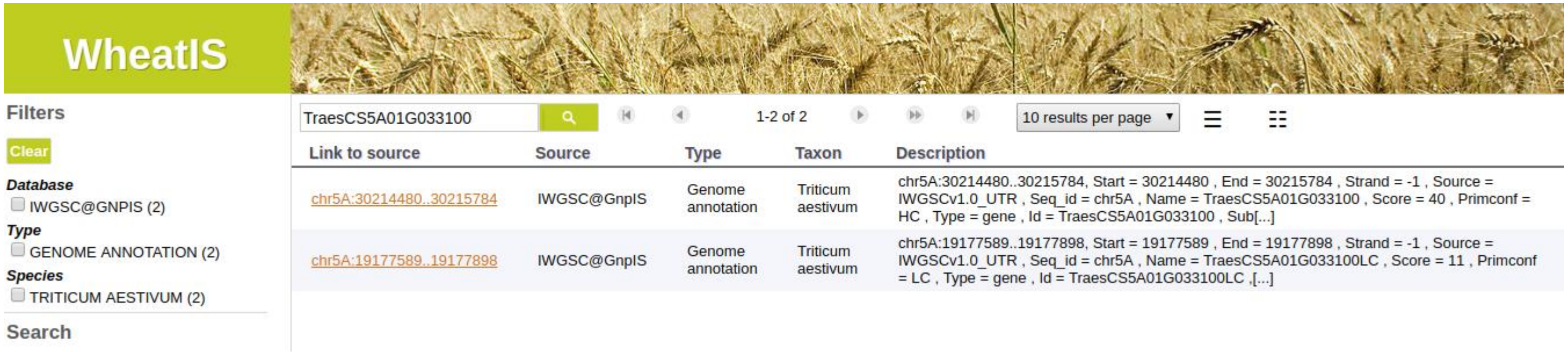
EBI (UP):

- Ensembl Plants: 2391552

27k searches in 2018

Use case:
From a gene to genetic studies

https://wheat-urgi.versailles.inra.fr/

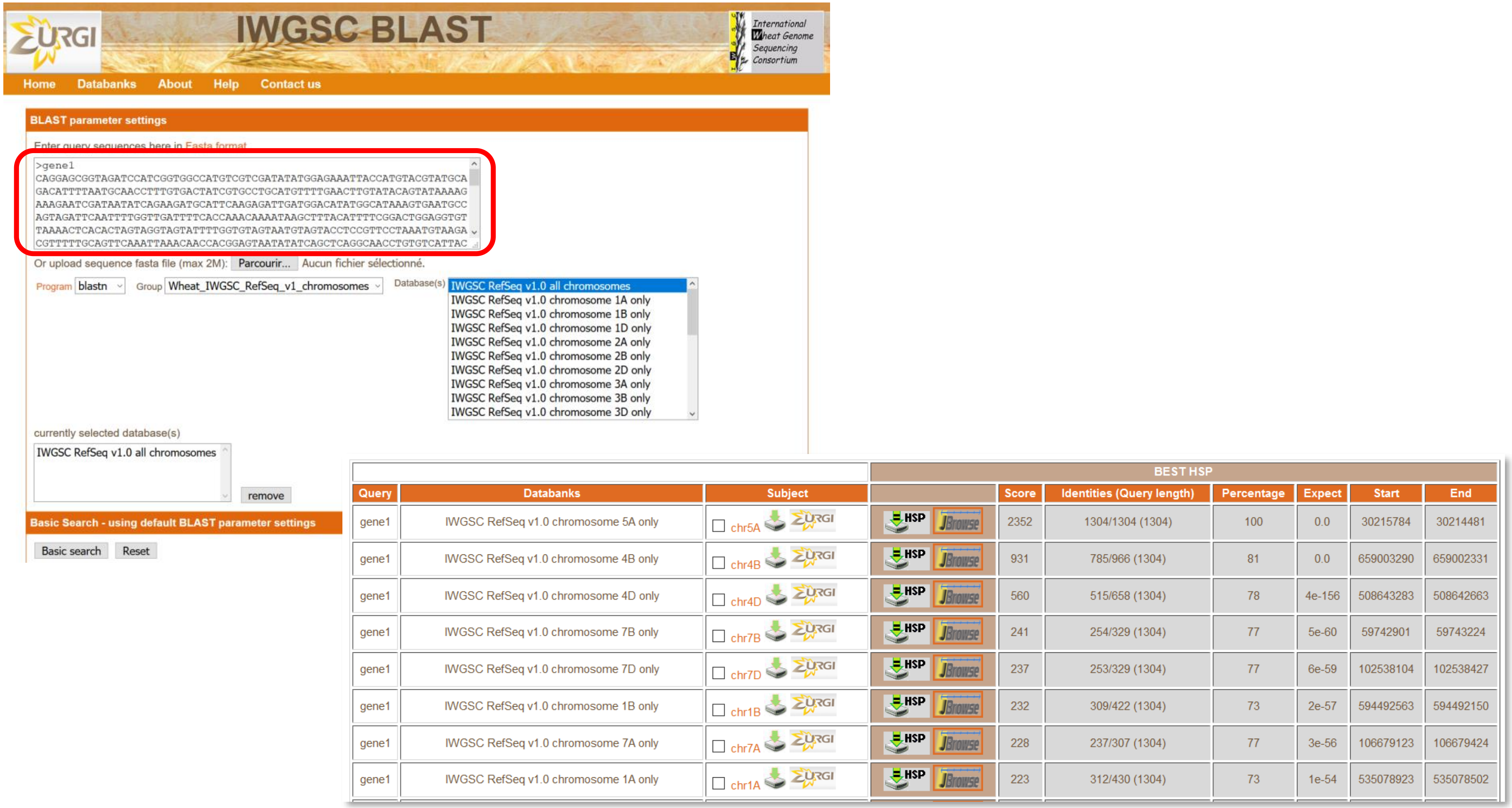


Link to source	Source	Type	Taxon	Description
chr5A:30214480..30215784	IWGSC@GnpIS	Genome annotation	Triticum aestivum	chr5A:30214480..30215784, Start = 30214480 , End = 30215784 , Strand = -1 , Source = IWGSCv1.0_UTR , Seq_id = chr5A , Name = TraesCS5A01G033100 , Score = 40 , Primconf = HC , Type = gene , Id = TraesCS5A01G033100 , Sub[...]
chr5A:19177589..19177898	IWGSC@GnpIS	Genome annotation	Triticum aestivum	chr5A:19177589..19177898, Start = 19177589 , End = 19177898 , Strand = -1 , Source = IWGSCv1.0_UTR , Seq_id = chr5A , Name = TraesCS5A01G033100LC , Score = 11 , Primconf = LC , Type = gene , Id = TraesCS5A01G033100LC ,[...]

Michael Alaux

Use case:
From a gene to genetic studies

https://urgi.versailles.inra.fr/blast_iwgsc



BLAST parameter settings

Enter query sequences here in Fasta format

```
>gene1
CAGGAGCGGTAGATCCATCGGTGGCCATGTCGTCGATATATGGAGAAATTACCATGTACGTATGCA
GACATTTTAATGCAACCTTTGTGACTATCGTGCCTGCATGTTTTGAACTTGTATACAGTATAAAAG
AAAGAATCGATAATATCAGAAGATGCATTCAAGAGATTGATGGACATATGGCATAAAGTGAATGCC
AGTAGATTCAATTTTGGTTGATTTTCACCAAACAAAATAAGCTTTACATTTTCGGACTGGAGGTGT
TAAAACTCACACTAGTAGGTAGTATTTTGGTGTAGTAATGTAGTACCTCCGTTCCTAAATGTAAGA
CGTTTTTGCAGTTCAAATTAAACAACCACGGAGTAATATATCAGCTCAGGCAACCTGTGTCATTAC
```

Or upload sequence fasta file (max 2M): Aucun fichier sélectionné.

Program: blastn

Group: Wheat_IWGSC_RefSeq_v1_chromosomes

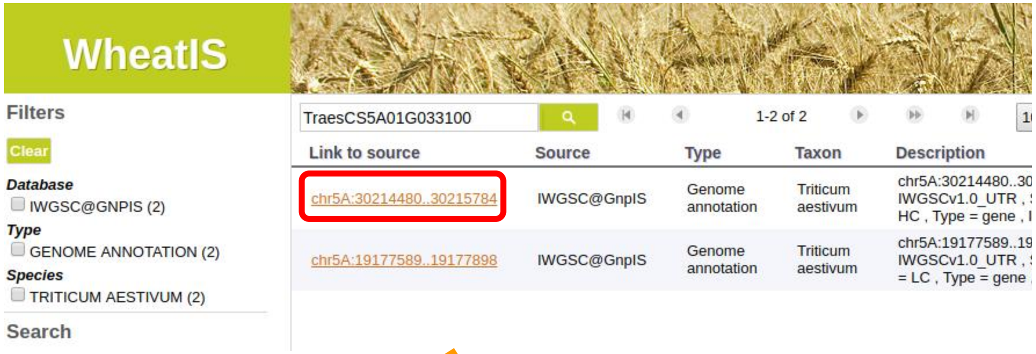
Database(s):
- IWGSC RefSeq v1.0 all chromosomes
- IWGSC RefSeq v1.0 chromosome 1A only
- IWGSC RefSeq v1.0 chromosome 1B only
- IWGSC RefSeq v1.0 chromosome 1D only
- IWGSC RefSeq v1.0 chromosome 2A only
- IWGSC RefSeq v1.0 chromosome 2B only
- IWGSC RefSeq v1.0 chromosome 2D only
- IWGSC RefSeq v1.0 chromosome 3A only
- IWGSC RefSeq v1.0 chromosome 3B only
- IWGSC RefSeq v1.0 chromosome 3D only

currently selected database(s): IWGSC RefSeq v1.0 all chromosomes

Basic Search - using default BLAST parameter settings

			BEST HSP						
Query	Databanks	Subject		Score	Identities (Query length)	Percentage	Expect	Start	End
gene1	IWGSC RefSeq v1.0 chromosome 5A only	☐ chr5A	HSP JBrowse	2352	1304/1304 (1304)	100	0.0	30215784	30214481
gene1	IWGSC RefSeq v1.0 chromosome 4B only	☐ chr4B	HSP JBrowse	931	785/966 (1304)	81	0.0	659003290	659002331
gene1	IWGSC RefSeq v1.0 chromosome 4D only	☐ chr4D	HSP JBrowse	560	515/658 (1304)	78	4e-156	508643283	508642663
gene1	IWGSC RefSeq v1.0 chromosome 7B only	☐ chr7B	HSP JBrowse	241	254/329 (1304)	77	5e-60	59742901	59743224
gene1	IWGSC RefSeq v1.0 chromosome 7D only	☐ chr7D	HSP JBrowse	237	253/329 (1304)	77	6e-59	102538104	102538427
gene1	IWGSC RefSeq v1.0 chromosome 1B only	☐ chr1B	HSP JBrowse	232	309/422 (1304)	73	2e-57	594492563	594492150
gene1	IWGSC RefSeq v1.0 chromosome 7A only	☐ chr7A	HSP JBrowse	228	237/307 (1304)	77	3e-56	106679123	106679424
gene1	IWGSC RefSeq v1.0 chromosome 1A only	☐ chr1A	HSP JBrowse	223	312/430 (1304)	73	1e-54	535078923	535078502

Use case:
From a gene to genetic studies



Filters

Clear

Database

☐ IWGSC@GNPIS (2)

Type

☐ GENOME ANNOTATION (2)

Species

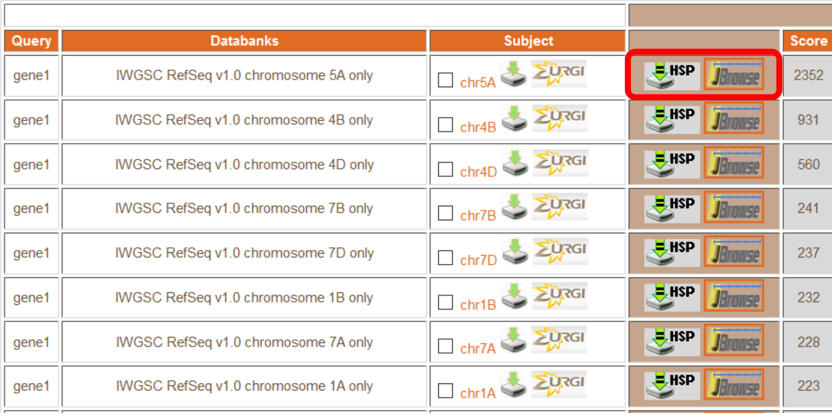
☐ TRITICUM AESTIVUM (2)

Search

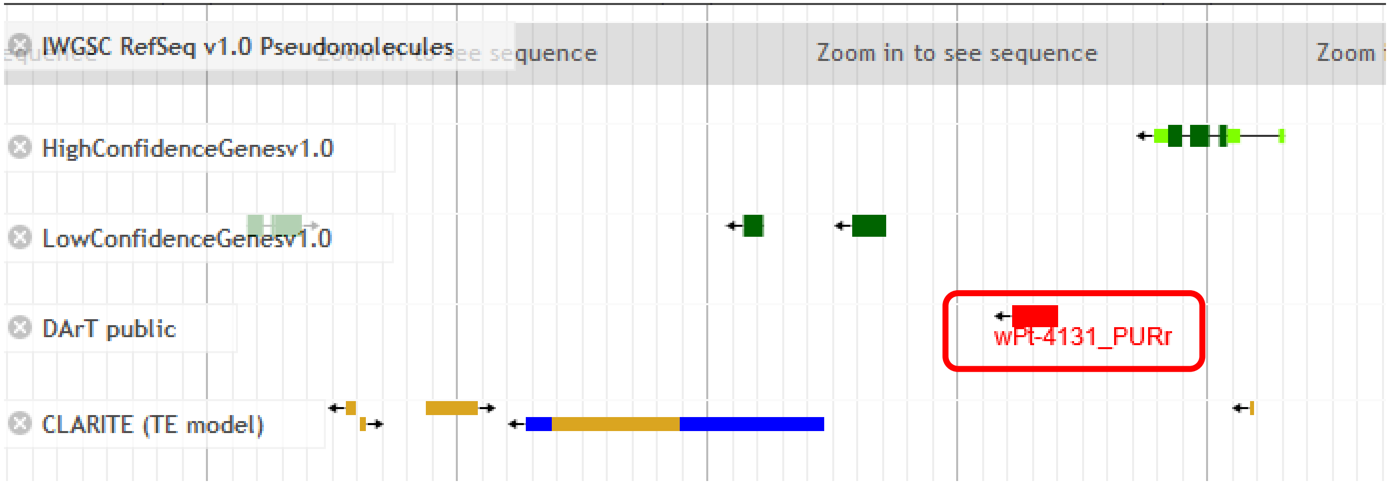
TraesCS5A01G033100

1-2 of 2

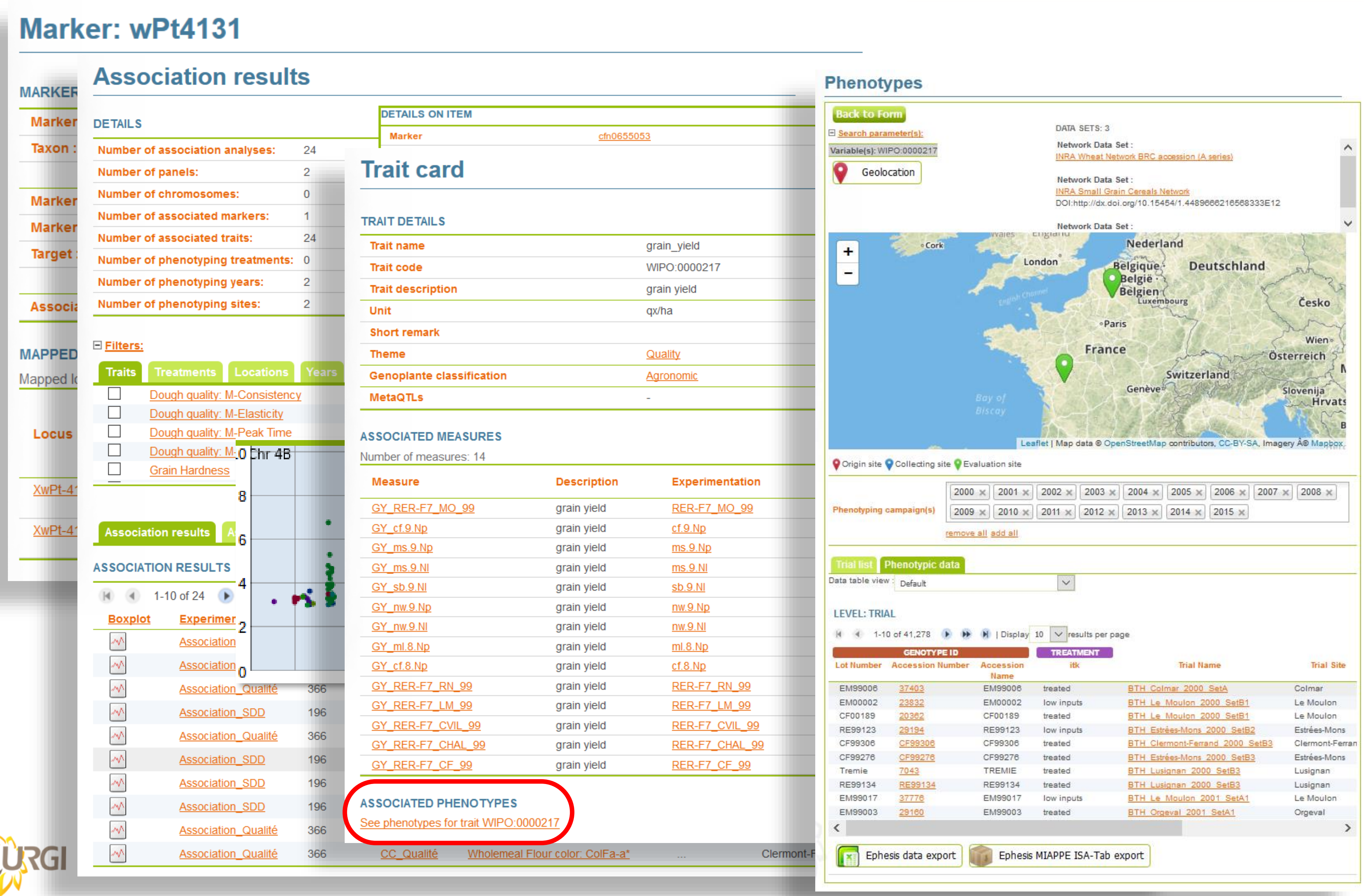
Link to source	Source	Type	Taxon	Description
chr5A:30214480..30215784	IWGSC@GnpIS	Genome annotation	Triticum aestivum	chr5A:30214480..30 IWGSCv1.0_UTR , HC , Type = gene ,
chr5A:19177589..19177898	IWGSC@GnpIS	Genome annotation	Triticum aestivum	chr5A:19177589..19 IWGSCv1.0_UTR , = LC , Type = gene



Query	Databanks	Subject		Score
gene1	IWGSC RefSeq v1.0 chromosome 5A only	☐ chr5A	HSP JBrowse	2352
gene1	IWGSC RefSeq v1.0 chromosome 4B only	☐ chr4B	HSP JBrowse	931
gene1	IWGSC RefSeq v1.0 chromosome 4D only	☐ chr4D	HSP JBrowse	560
gene1	IWGSC RefSeq v1.0 chromosome 7B only	☐ chr7B	HSP JBrowse	241
gene1	IWGSC RefSeq v1.0 chromosome 7D only	☐ chr7D	HSP JBrowse	237
gene1	IWGSC RefSeq v1.0 chromosome 1B only	☐ chr1B	HSP JBrowse	232
gene1	IWGSC RefSeq v1.0 chromosome 7A only	☐ chr7A	HSP JBrowse	228
gene1	IWGSC RefSeq v1.0 chromosome 1A only	☐ chr1A	HSP JBrowse	223



Use case: From a gene to genetic studies



Marker: wPt4131

MARKER

Marker

Taxon :

Marker

Marker

Target :

Associa

MAPPED

Mapped lo

Locus

XwPt-41

XwPt-41

Association results

DETAILS

Number of association analyses:	24
Number of panels:	2
Number of chromosomes:	0
Number of associated markers:	1
Number of associated traits:	24
Number of phenotyping treatments:	0
Number of phenotyping years:	2
Number of phenotyping sites:	2

⊟ Filters:

- ☐ Dough quality: M-Consistency
- ☐ Dough quality: M-Elasticity
- ☐ Dough quality: M-Peak Time
- ☐ Dough quality: M-
- ☐ Grain Hardness

ASSOCIATION RESULTS

1-10 of 24

Boxplot	Experimen	
	Association	
	Association	
	Association_Qualité	366
	Association_SDD	196
	Association_Qualité	366
	Association_SDD	196
	Association_SDD	196
	Association_SDD	196
	Association_Qualité	366
	Association_Qualité	366

CC_Qualité Wholemeal Flour color: ColFa-a* ... Clermont-F

DETAILS ON ITEM

Marker cfn0655053

Trait card

TRAIT DETAILS

Trait name	grain_yield
Trait code	WIPO:0000217
Trait description	grain yield
Unit	qx/ha
Short remark	
Theme	Quality
Genoplante classification	Agronomic
MetaQTLs	-

ASSOCIATED MEASURES

Number of measures: 14

Measure	Description	Experimentation
GY_RER-F7_MO_99	grain yield	RER-F7_MO_99
GY_cf.9.Np	grain yield	cf.9.Np
GY_ms.9.Np	grain yield	ms.9.Np
GY_ms.9.Nl	grain yield	ms.9.Nl
GY_sb.9.Nl	grain yield	sb.9.Nl
GY_nw.9.Np	grain yield	nw.9.Np
GY_nw.9.Nl	grain yield	nw.9.Nl
GY_ml.8.Np	grain yield	ml.8.Np
GY_cf.8.Np	grain yield	cf.8.Np
GY_RER-F7_RN_99	grain yield	RER-F7_RN_99
GY_RER-F7_LM_99	grain yield	RER-F7_LM_99
GY_RER-F7_CVIL_99	grain yield	RER-F7_CVIL_99
GY_RER-F7_CHAL_99	grain yield	RER-F7_CHAL_99
GY_RER-F7_CF_99	grain yield	RER-F7_CF_99

ASSOCIATED PHENOTYPES

See phenotypes for trait WIPO:0000217

Phenotypes

⊟ Search parameter(s):

Variable(s): WIPO:0000217

DATA SETS: 3

Network Data Set :
INRA Wheat Network BRC accession (A series)

Network Data Set :
INRA Small Grain Cereals Network
DOI:http://dx.doi.org/10.15454/1.4489666216568333E12

Network Data Set :

Leaflet | Map data © OpenStreetMap contributors, CC-BY-SA, Imagery Â© Mapbox

Origin site Collecting site Evaluation site

Phenotyping campaign(s): 2000 × 2001 × 2002 × 2003 × 2004 × 2005 × 2006 × 2007 × 2008 × 2009 × 2010 × 2011 × 2012 × 2013 × 2014 × 2015 ×

remove all add all

Data table view : Default

LEVEL: TRIAL

1-10 of 41,278 | Display 10 results per page

GENOTYPE ID			TREATMENT		
Lot Number	Accession Number	Accession Name	itk	Trial Name	Trial Site
EM99006	37403	EM99006	treated	BTH_Colmar_2000_SetA	Colmar
EM00002	23832	EM00002	low inputs	BTH_Le_Moulon_2000_SetB1	Le Moulon
CF00189	20362	CF00189	treated	BTH_Le_Moulon_2000_SetB1	Le Moulon
RE99123	29194	RE99123	low inputs	BTH_Estrées-Mons_2000_SetB2	Estrées-Mons
CF99306	CF99306	CF99306	treated	BTH_Clermont-Ferrand_2000_SetB3	Clermont-Ferran
CF99276	CF99276	CF99276	treated	BTH_Estrées-Mons_2000_SetB3	Estrées-Mons
Tremie	7043	TREMIE	treated	BTH_Lusignan_2000_SetB3	Lusignan
RE99134	RE99134	RE99134	treated	BTH_Lusignan_2000_SetB3	Lusignan
EM99017	37776	EM99017	low inputs	BTH_Le_Moulon_2001_SetA1	Le Moulon
EM99003	29160	EM99003	treated	BTH_Orgeval_2001_SetA1	Orgeval

Acknowledgements

URGI

Alaux M.
Letellier T.
Flores R.
Alfama F.
Pommier C.
Mohellibi N.
Durand S.
Kimmel E.
Michotey C.
Guerche C.
Loaec M.
Lainé M.
Steinbach D.
Adam-Blondon A.F.
Quesneville H.

Jamilloux V., Amselem J., Charruaud D., Cornut G., Burlot L., Philippe F., Francillonne N., Couderc L., Verdelet D., Brault B., Chennen K.

IWGSC
International Wheat Genome Sequencing Consortium

Rogers J.
Eversole K.

Caugant I.

GDEC
Génétique Diversité Ecophysiologie des Céréales

Choulet F.
Rimbert H.
Leroy P.
Guilhot N.
Salse J.
Feuillet C.
Paux E.

Le Gouis J.
Charmet G.
Balfourier F.
Sourdille P.
Ravel C.
Oury F.X.
Didier A.

ist@inra

Dzale E.
Aubin S.
Hologne O.

Acknowledgements

All data submitters

IWGSC Coordinating Committee
IWGSC Sequencing team
IWGSC Analysis team

WheatIS Expert Working Group

RDA Wheat Data Interoperability Group

Questions

Alaux et al., *Genome Biology* 2018
https://doi.org/10.1186/s13059-018-1491-4


Genome Biology

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

Wheat@URGI portal
https://wheat-urgi.versailles.inra.fr/

Contact me at michael.alaux@inra.fr

@michaelalaux