GrainGenes
A Database for Triticeae and Avena

GrainGenes

a Centralized Nexus for Small Grains Data and Communities

Taner Sen

USDA-ARS / UC Berkeley
Albany, CA

IWGSC
April 4th, 2024

USDA

Berkeley
UNIVERSITY OF CALIFORNIA

Challenges for Sustainable Agriculture

The burden is not shared equally across nations and groups, so we need to be mindful of inequalities

This is not one country's burden, but a global one, requiring global policy solutions and collaborative research across nations

POPULATION INCREASE

8 billion at 1.1%

FOOD SECURITY - YIELD

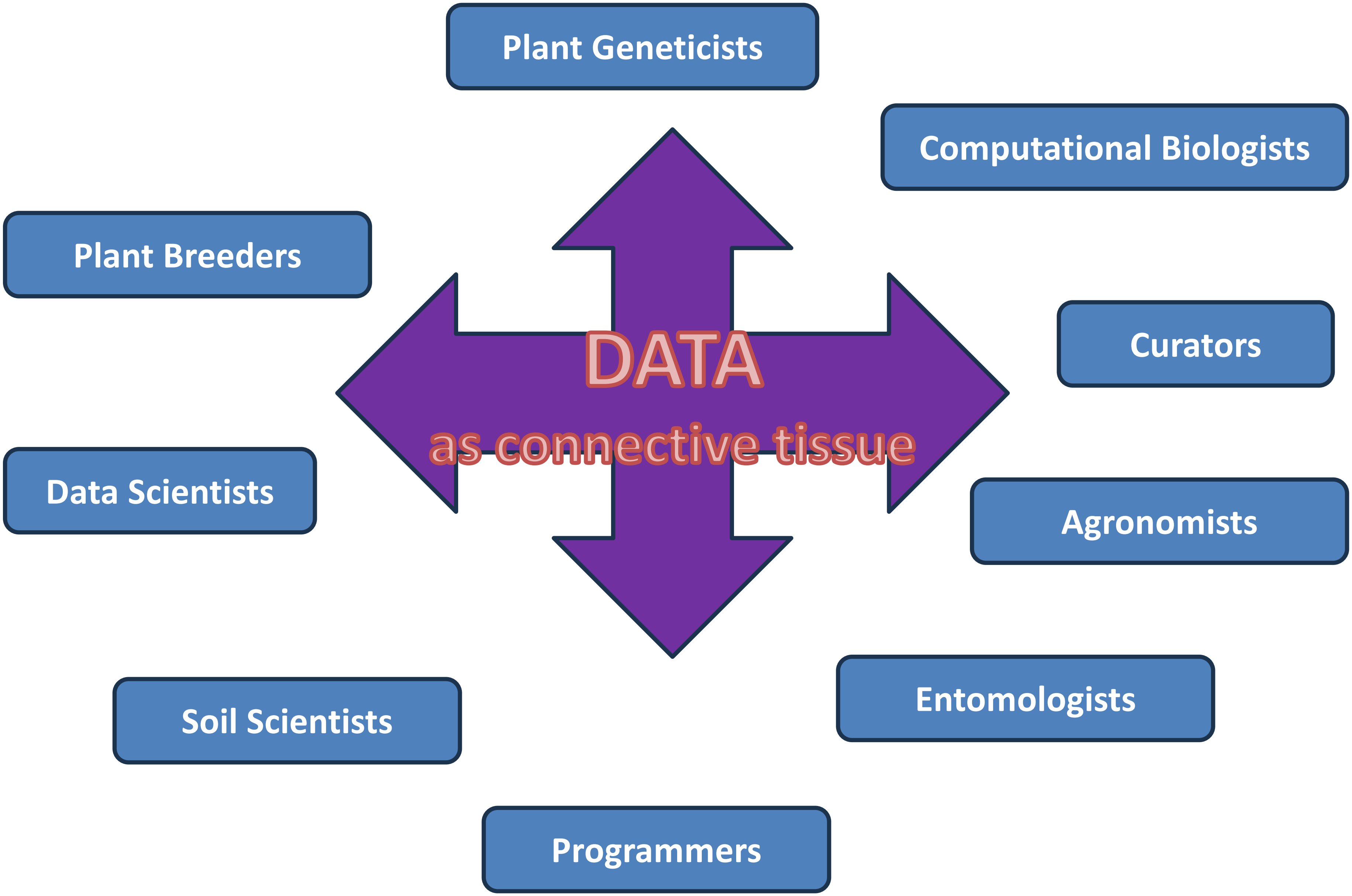
- Insufficient yield
- Limited arable lands

FOOD SECURITY - NUTRITION

Enough food ≠ nutritious food

CLIMATE CHANGE

We need more research to address agricultural challenges



Plant Geneticists

Computational Biologists

Plant Breeders

Curators

DATA

as connective tissue

Data Scientists

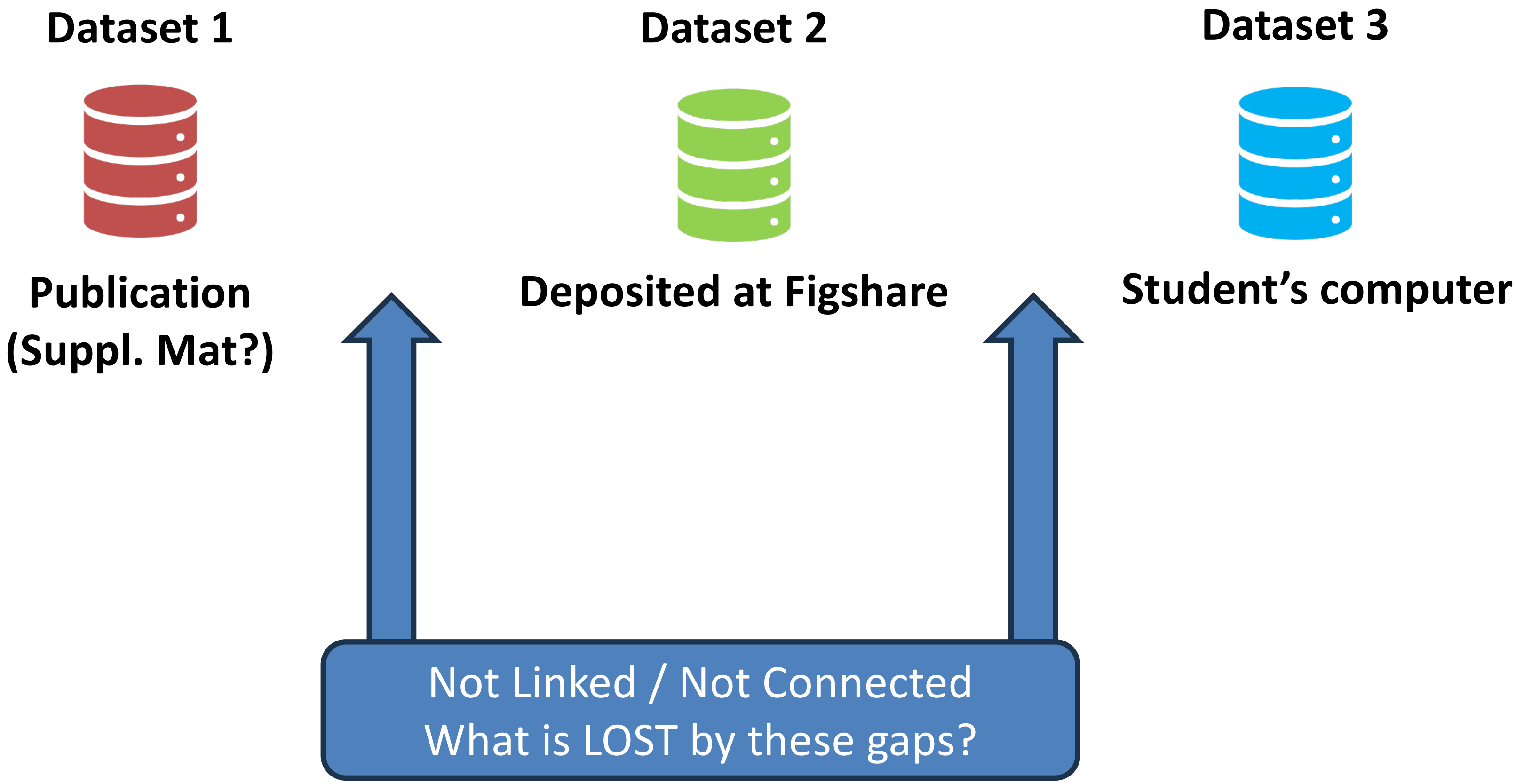
Agronomists

Soil Scientists

Entomologists

Programmers

How can we increase the value of research data even more?



Dataset 1

Publication (Suppl. Mat?)

Dataset 2

Deposited at Figshare

Dataset 3

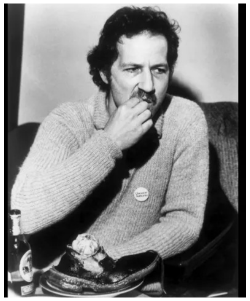
Student's computer

Not Linked / Not Connected
What is LOST by these gaps?

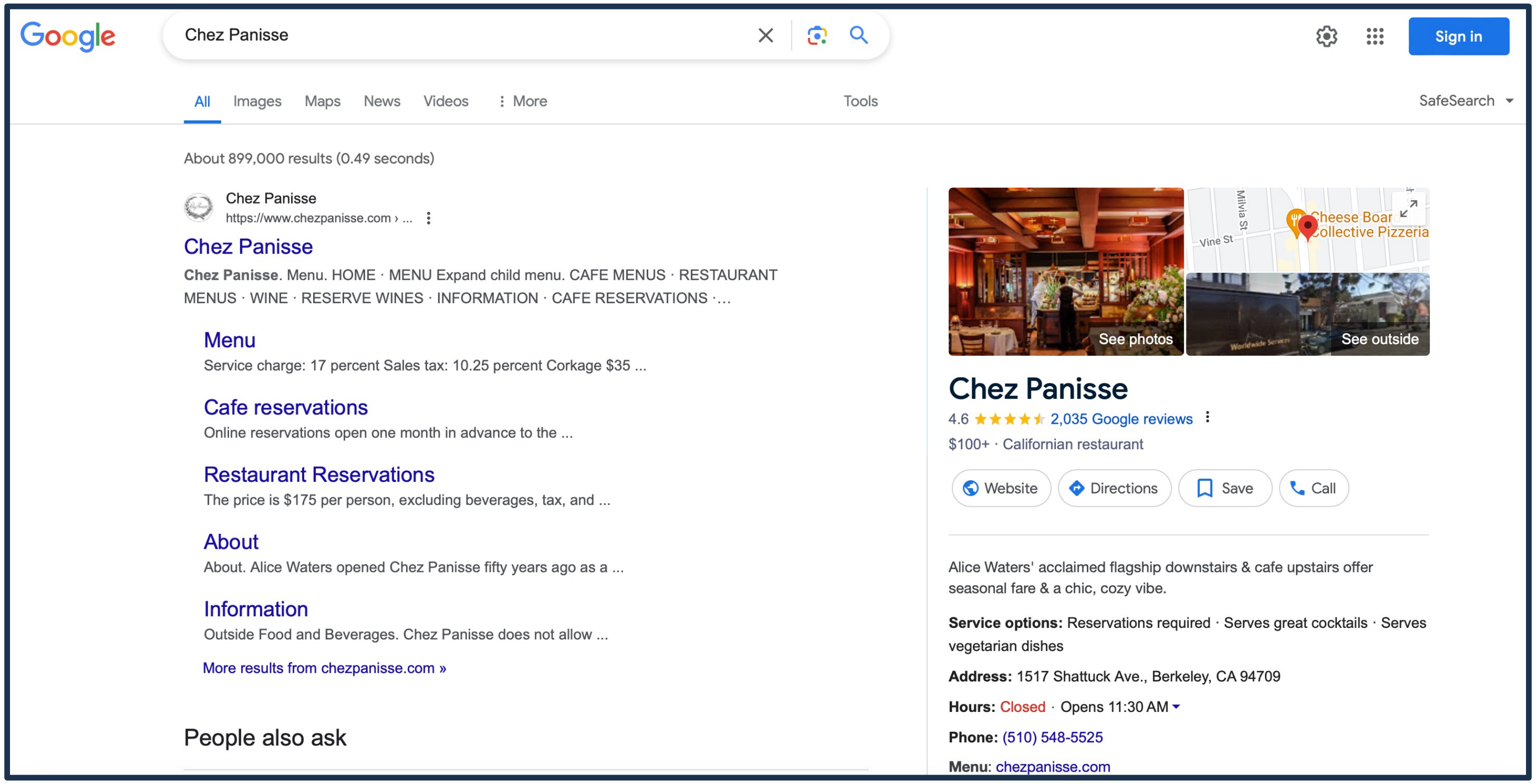
Why linking/associating data important?

Value of Linked Data – "Chez Panisse" Example

Restaurants **Menu** **Reviews** **Addresses**



"Werner Herzog eats his shoe"



About 899,000 results (0.49 seconds)

Chez Panisse
https://www.chezpanisse.com › ...

Chez Panisse

Chez Panisse. Menu. HOME · MENU Expand child menu. CAFE MENUS · RESTAURANT MENUS · WINE · RESERVE WINES · INFORMATION · CAFE RESERVATIONS ·...

Menu
Service charge: 17 percent Sales tax: 10.25 percent Corkage $35 ...

Cafe reservations
Online reservations open one month in advance to the ...

Restaurant Reservations
The price is $175 per person, excluding beverages, tax, and ...

About
About. Alice Waters opened Chez Panisse fifty years ago as a ...

Information
Outside Food and Beverages. Chez Panisse does not allow ...

More results from chezpanisse.com »

People also ask

Chez Panisse
4.6 ★★★★★ 2,035 Google reviews
$100+ · Californian restaurant

Alice Waters' acclaimed flagship downstairs & cafe upstairs offer seasonal fare & a chic, cozy vibe.

Service options: Reservations required · Serves great cocktails · Serves vegetarian dishes

Address: 1517 Shattuck Ave., Berkeley, CA 94709

Hours: Closed · Opens 11:30 AM

Phone: (510) 548-5525

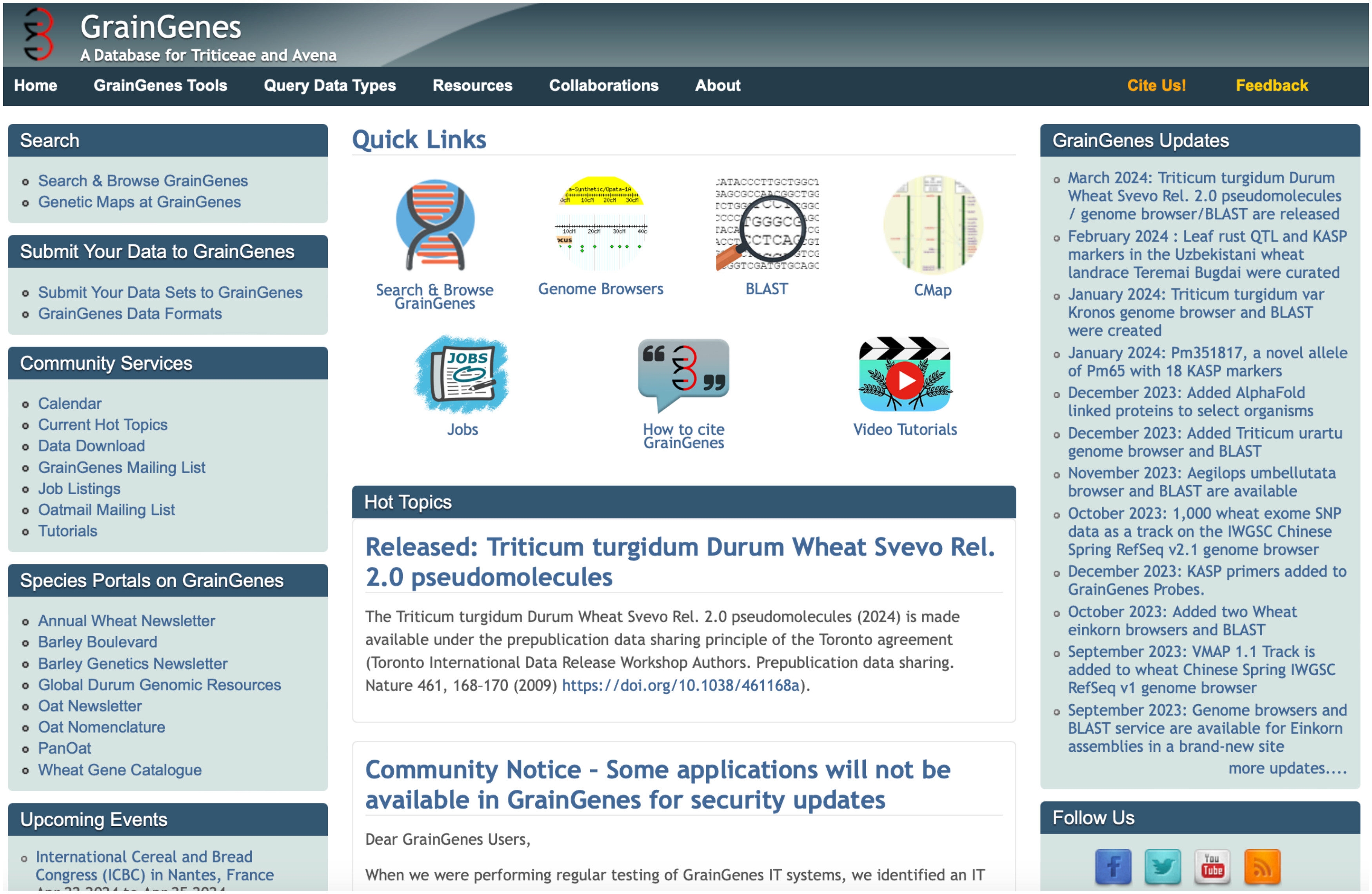
Menu: chezpanisse.com

The value of a centralized location that provides linked datasets and tools to users through the same website/browser

GrainGenes

(Est. 1992)

https://wheat.pw.usda.gov



Search

- Search & Browse GrainGenes
- Genetic Maps at GrainGenes

Submit Your Data to GrainGenes

- Submit Your Data Sets to GrainGenes
- GrainGenes Data Formats

Community Services

- Calendar
- Current Hot Topics
- Data Download
- GrainGenes Mailing List
- Job Listings
- Oatmail Mailing List
- Tutorials

Species Portals on GrainGenes

- Annual Wheat Newsletter
- Barley Boulevard
- Barley Genetics Newsletter
- Global Durum Genomic Resources
- Oat Newsletter
- Oat Nomenclature
- PanOat
- Wheat Gene Catalogue

Upcoming Events

- International Cereal and Bread Congress (ICBC) in Nantes, France

Quick Links

Hot Topics

Released: Triticum turgidum Durum Wheat Svevo Rel. 2.0 pseudomolecules

The Triticum turgidum Durum Wheat Svevo Rel. 2.0 pseudomolecules (2024) is made available under the prepublication data sharing principle of the Toronto agreement (Toronto International Data Release Workshop Authors. Prepublication data sharing. Nature 461, 168-170 (2009) https://doi.org/10.1038/461168a).

Community Notice - Some applications will not be available in GrainGenes for security updates

Dear GrainGenes Users,

When we were performing regular testing of GrainGenes IT systems, we identified an IT

GrainGenes Updates

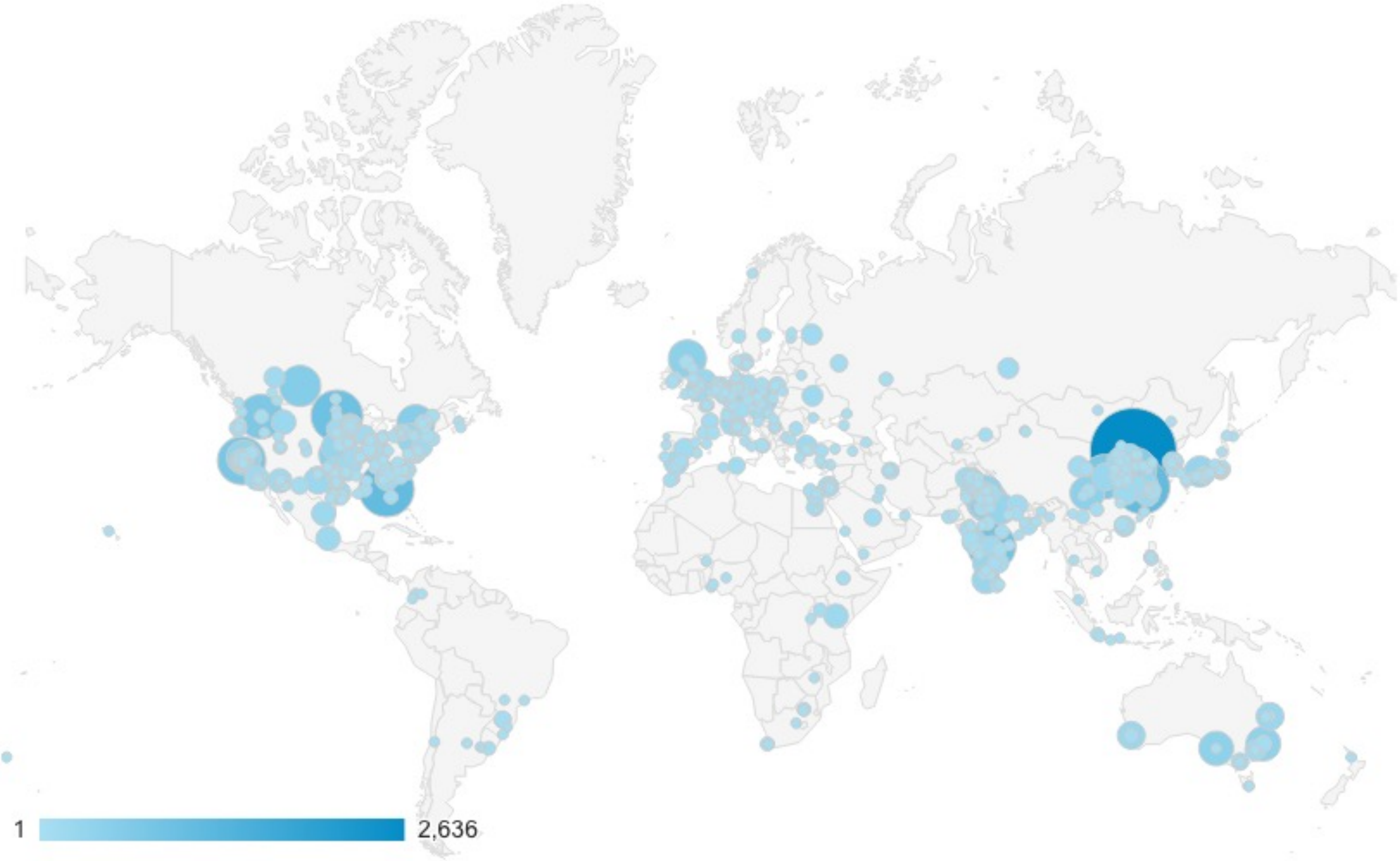
- March 2024: Triticum turgidum Durum Wheat Svevo Rel. 2.0 pseudomolecules / genome browser/BLAST are released
- February 2024 : Leaf rust QTL and KASP markers in the Uzbekistani wheat landrace Teremai Bugdai were curated
- January 2024: Triticum turgidum var Kronos genome browser and BLAST were created
- January 2024: Pm351817, a novel allele of Pm65 with 18 KASP markers
- December 2023: Added AlphaFold linked proteins to select organisms
- December 2023: Added Triticum urartu genome browser and BLAST
- November 2023: Aegilops umbellutata browser and BLAST are available
- October 2023: 1,000 wheat exome SNP data as a track on the IWGSC Chinese Spring RefSeq v2.1 genome browser
- December 2023: KASP primers added to GrainGenes Probes.
- October 2023: Added two Wheat einkorn browsers and BLAST
- September 2023: VMAP 1.1 Track is added to wheat Chinese Spring IWGSC RefSeq v1 genome browser
- September 2023: Genome browsers and BLAST service are available for Einkorn assemblies in a brand-new site

more updates....

Follow Us

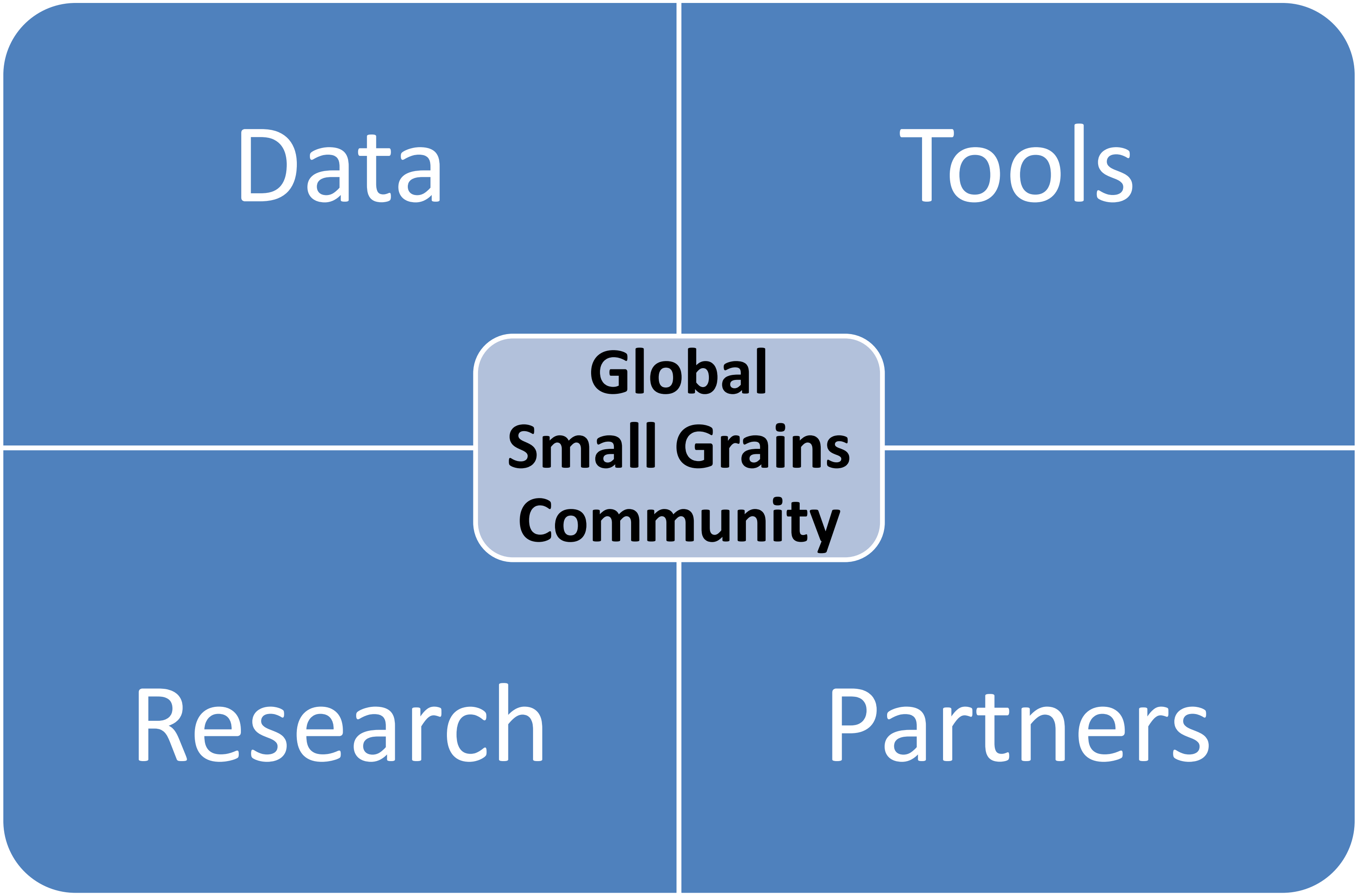
GrainGenes is a Global Resource

We serve small grains researchers



GrainGenes

https://wheat.pw.usda.gov



Data

Tools

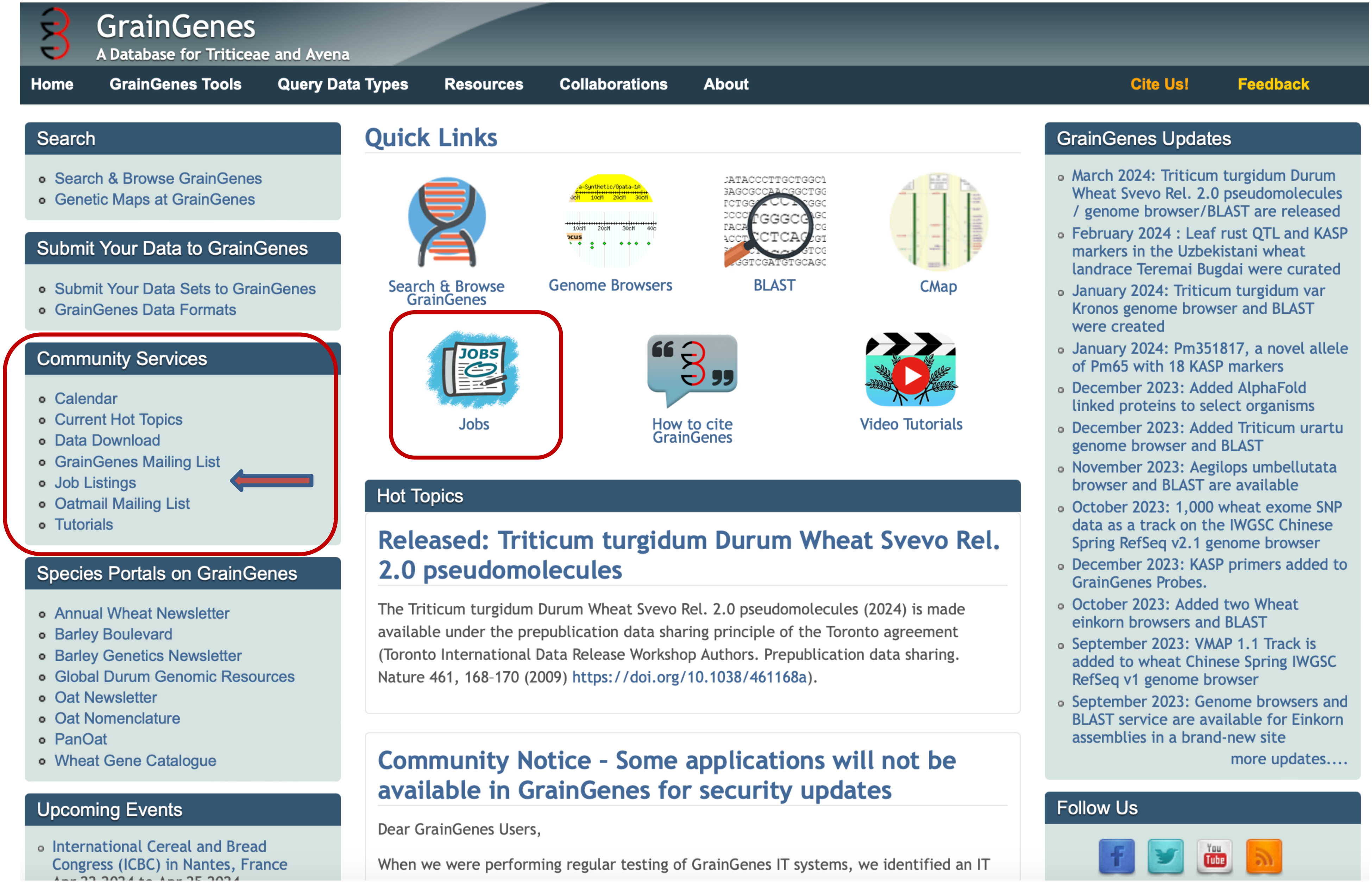
Global
Small Grains
Community

Research

Partners

COMMUNITY TOOLS

GrainGenes - Jobs



Search

- Search & Browse GrainGenes
- Genetic Maps at GrainGenes

Submit Your Data to GrainGenes

- Submit Your Data Sets to GrainGenes
- GrainGenes Data Formats

Community Services

- Calendar
- Current Hot Topics
- Data Download
- GrainGenes Mailing List
- Job Listings
- Oatmail Mailing List
- Tutorials

Species Portals on GrainGenes

- Annual Wheat Newsletter
- Barley Boulevard
- Barley Genetics Newsletter
- Global Durum Genomic Resources
- Oat Newsletter
- Oat Nomenclature
- PanOat
- Wheat Gene Catalogue

Upcoming Events

- International Cereal and Bread Congress (ICBC) in Nantes, France

Quick Links

Hot Topics

Released: Triticum turgidum Durum Wheat Svevo Rel. 2.0 pseudomolecules

The Triticum turgidum Durum Wheat Svevo Rel. 2.0 pseudomolecules (2024) is made available under the prepublication data sharing principle of the Toronto agreement (Toronto International Data Release Workshop Authors. Prepublication data sharing. Nature 461, 168-170 (2009) https://doi.org/10.1038/461168a).

Community Notice - Some applications will not be available in GrainGenes for security updates

Dear GrainGenes Users,

When we were performing regular testing of GrainGenes IT systems, we identified an IT

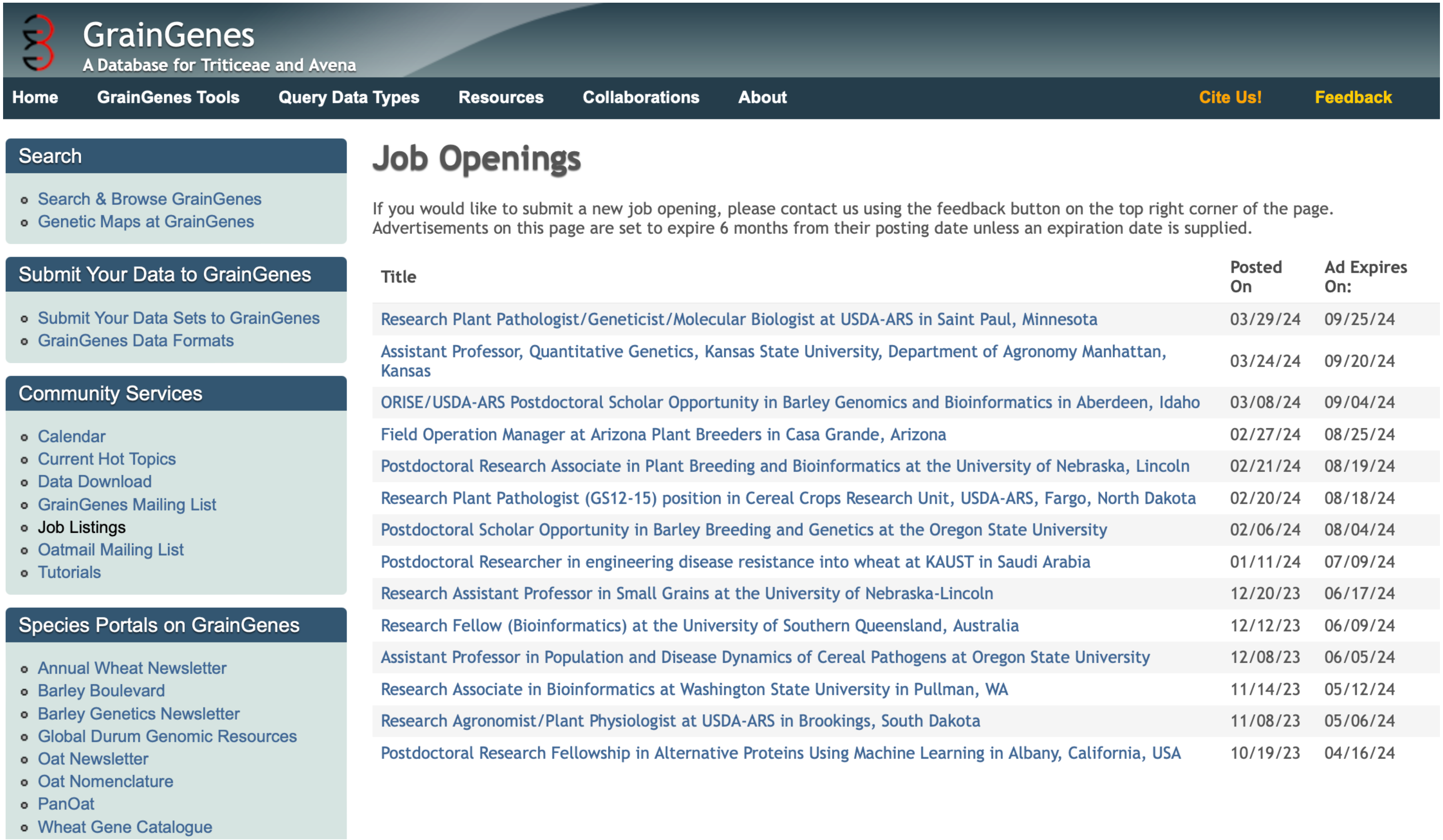
GrainGenes Updates

- March 2024: Triticum turgidum Durum Wheat Svevo Rel. 2.0 pseudomolecules / genome browser/BLAST are released
- February 2024 : Leaf rust QTL and KASP markers in the Uzbekistani wheat landrace Teremai Bugdai were curated
- January 2024: Triticum turgidum var Kronos genome browser and BLAST were created
- January 2024: Pm351817, a novel allele of Pm65 with 18 KASP markers
- December 2023: Added AlphaFold linked proteins to select organisms
- December 2023: Added Triticum urartu genome browser and BLAST
- November 2023: Aegilops umbellutata browser and BLAST are available
- October 2023: 1,000 wheat exome SNP data as a track on the IWGSC Chinese Spring RefSeq v2.1 genome browser
- December 2023: KASP primers added to GrainGenes Probes.
- October 2023: Added two Wheat einkorn browsers and BLAST
- September 2023: VMAP 1.1 Track is added to wheat Chinese Spring IWGSC RefSeq v1 genome browser
- September 2023: Genome browsers and BLAST service are available for Einkorn assemblies in a brand-new site

more updates....

Follow Us

GrainGenes – Job board



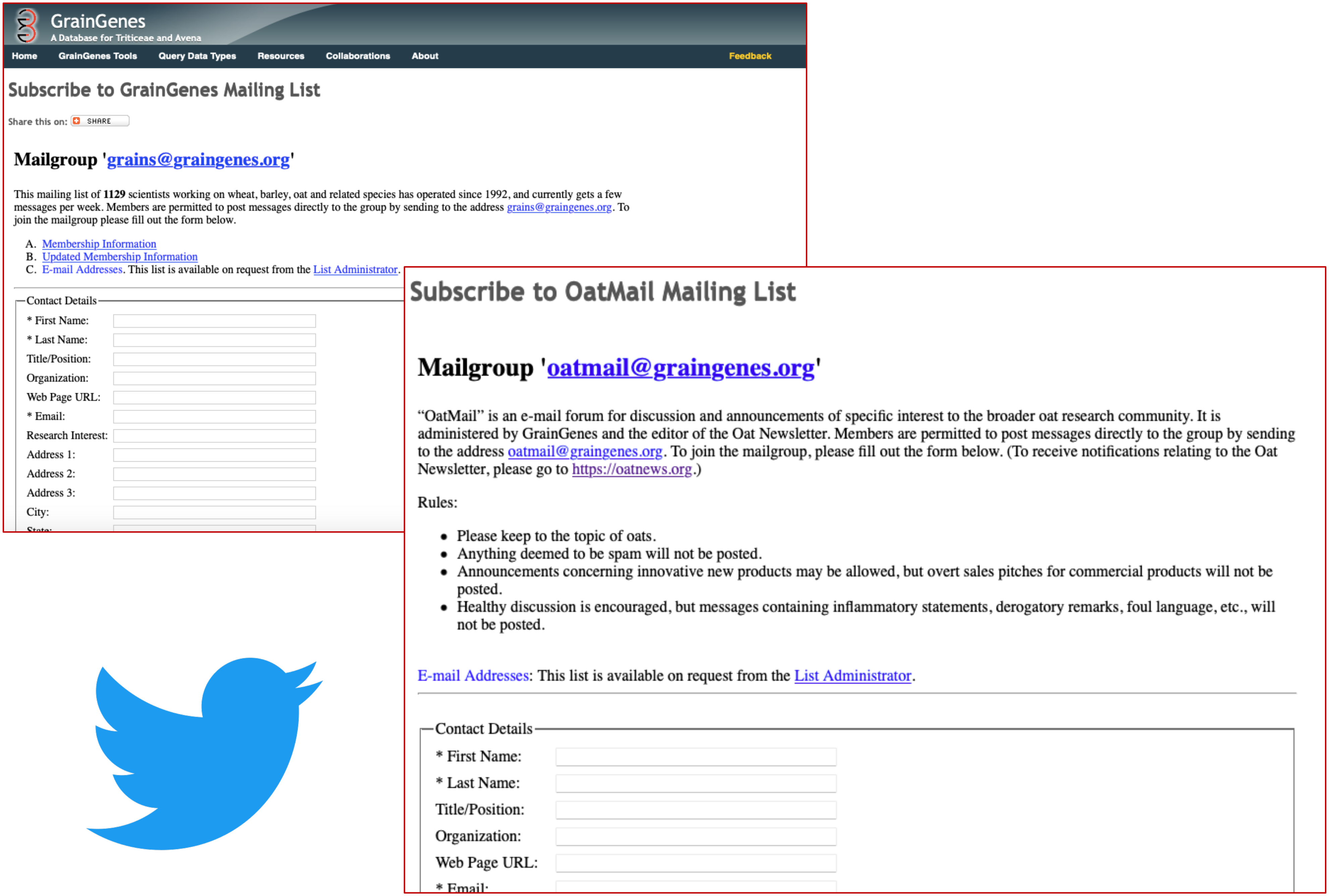
Job Openings

If you would like to submit a new job opening, please contact us using the feedback button on the top right corner of the page. Advertisements on this page are set to expire 6 months from their posting date unless an expiration date is supplied.

Title	Posted On	Ad Expires On:
Research Plant Pathologist/Geneticist/Molecular Biologist at USDA-ARS in Saint Paul, Minnesota	03/29/24	09/25/24
Assistant Professor, Quantitative Genetics, Kansas State University, Department of Agronomy Manhattan, Kansas	03/24/24	09/20/24
ORISE/USDA-ARS Postdoctoral Scholar Opportunity in Barley Genomics and Bioinformatics in Aberdeen, Idaho	03/08/24	09/04/24
Field Operation Manager at Arizona Plant Breeders in Casa Grande, Arizona	02/27/24	08/25/24
Postdoctoral Research Associate in Plant Breeding and Bioinformatics at the University of Nebraska, Lincoln	02/21/24	08/19/24
Research Plant Pathologist (GS12-15) position in Cereal Crops Research Unit, USDA-ARS, Fargo, North Dakota	02/20/24	08/18/24
Postdoctoral Scholar Opportunity in Barley Breeding and Genetics at the Oregon State University	02/06/24	08/04/24
Postdoctoral Researcher in engineering disease resistance into wheat at KAUST in Saudi Arabia	01/11/24	07/09/24
Research Assistant Professor in Small Grains at the University of Nebraska-Lincoln	12/20/23	06/17/24
Research Fellow (Bioinformatics) at the University of Southern Queensland, Australia	12/12/23	06/09/24
Assistant Professor in Population and Disease Dynamics of Cereal Pathogens at Oregon State University	12/08/23	06/05/24
Research Associate in Bioinformatics at Washington State University in Pullman, WA	11/14/23	05/12/24
Research Agronomist/Plant Physiologist at USDA-ARS in Brookings, South Dakota	11/08/23	05/06/24
Postdoctoral Research Fellowship in Alternative Proteins Using Machine Learning in Albany, California, USA	10/19/23	04/16/24

Advertisement cost: $0

GrainGenes email lists



Subscribe to GrainGenes Mailing List

Mailgroup 'grains@graingenes.org'

This mailing list of **1129** scientists working on wheat, barley, oat and related species has operated since 1992, and currently gets a few messages per week. Members are permitted to post messages directly to the group by sending to the address grains@graingenes.org. To join the mailgroup please fill out the form below.

A. Membership Information
B. Updated Membership Information
C. E-mail Addresses. This list is available on request from the List Administrator.

Contact Details

* First Name:
* Last Name:
Title/Position:
Organization:
Web Page URL:
* Email:
Research Interest:
Address 1:
Address 2:
Address 3:
City:
State:

Subscribe to OatMail Mailing List

Mailgroup 'oatmail@graingenes.org'

“OatMail” is an e-mail forum for discussion and announcements of specific interest to the broader oat research community. It is administered by GrainGenes and the editor of the Oat Newsletter. Members are permitted to post messages directly to the group by sending to the address oatmail@graingenes.org. To join the mailgroup, please fill out the form below. (To receive notifications relating to the Oat Newsletter, please go to https://oatnews.org.)

Rules:

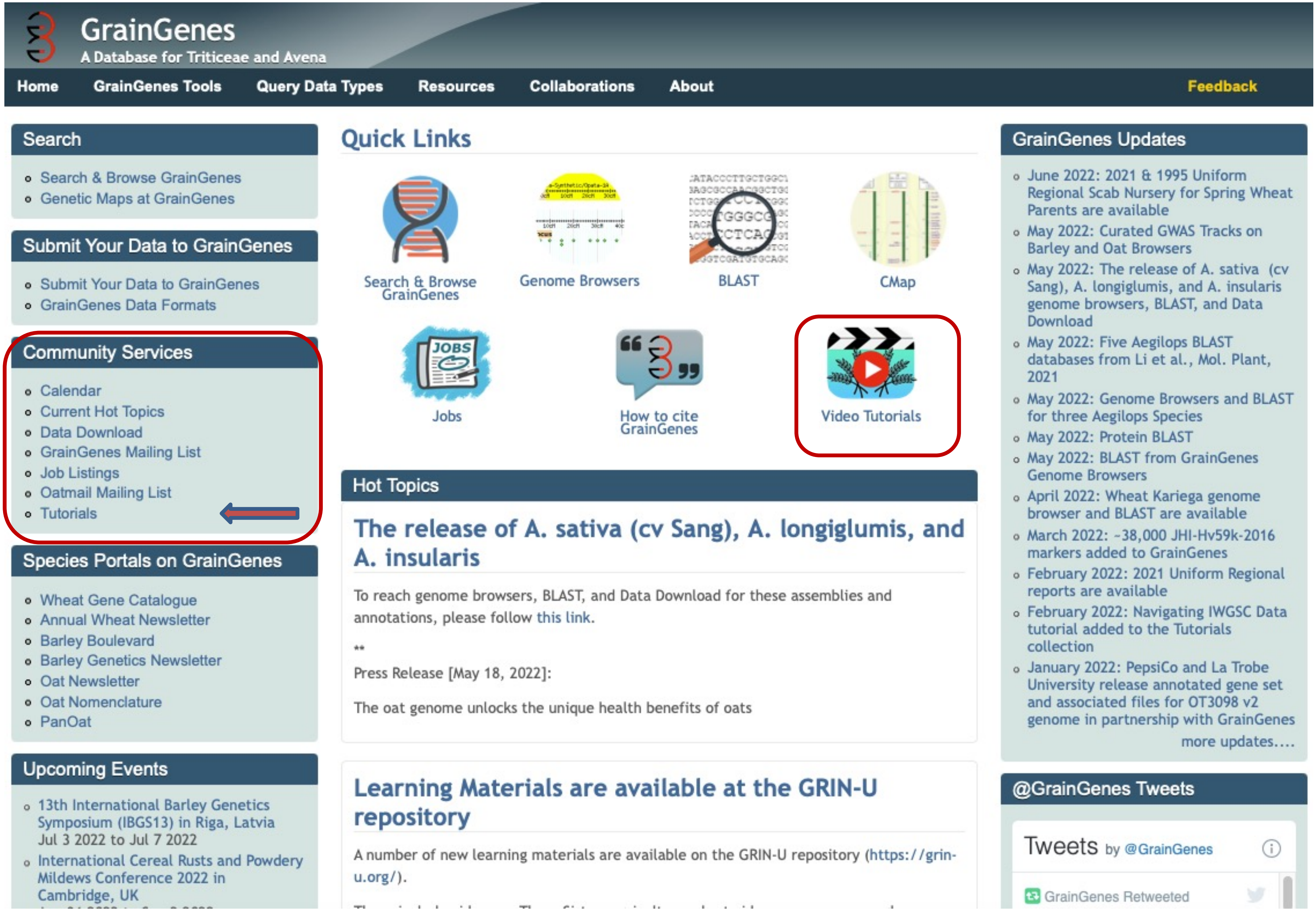
- Please keep to the topic of oats.
- Anything deemed to be spam will not be posted.
- Announcements concerning innovative new products may be allowed, but overt sales pitches for commercial products will not be posted.
- Healthy discussion is encouraged, but messages containing inflammatory statements, derogatory remarks, foul language, etc., will not be posted.

E-mail Addresses: This list is available on request from the List Administrator.

Contact Details

* First Name:
* Last Name:
Title/Position:
Organization:
Web Page URL:
* Email:

GrainGenes - Tutorials



Search

- Search & Browse GrainGenes
- Genetic Maps at GrainGenes

Submit Your Data to GrainGenes

- Submit Your Data to GrainGenes
- GrainGenes Data Formats

Community Services

- Calendar
- Current Hot Topics
- Data Download
- GrainGenes Mailing List
- Job Listings
- Oatmail Mailing List
- Tutorials

Species Portals on GrainGenes

- Wheat Gene Catalogue
- Annual Wheat Newsletter
- Barley Boulevard
- Barley Genetics Newsletter
- Oat Newsletter
- Oat Nomenclature
- PanOat

Upcoming Events

- 13th International Barley Genetics Symposium (IBGS13) in Riga, Latvia
  Jul 3 2022 to Jul 7 2022
- International Cereal Rusts and Powdery Mildews Conference 2022 in Cambridge, UK

Quick Links

Hot Topics

The release of A. sativa (cv Sang), A. longiglumis, and A. insularis

To reach genome browsers, BLAST, and Data Download for these assemblies and annotations, please follow this link.

**

Press Release [May 18, 2022]:

The oat genome unlocks the unique health benefits of oats

Learning Materials are available at the GRIN-U repository

A number of new learning materials are available on the GRIN-U repository (https://grin-u.org/).

GrainGenes Updates

- June 2022: 2021 & 1995 Uniform Regional Scab Nursery for Spring Wheat Parents are available
- May 2022: Curated GWAS Tracks on Barley and Oat Browsers
- May 2022: The release of A. sativa (cv Sang), A. longiglumis, and A. insularis genome browsers, BLAST, and Data Download
- May 2022: Five Aegilops BLAST databases from Li et al., Mol. Plant, 2021
- May 2022: Genome Browsers and BLAST for three Aegilops Species
- May 2022: Protein BLAST
- May 2022: BLAST from GrainGenes Genome Browsers
- April 2022: Wheat Kariega genome browser and BLAST are available
- March 2022: ~38,000 JHI-Hv59k-2016 markers added to GrainGenes
- February 2022: 2021 Uniform Regional reports are available
- February 2022: Navigating IWGSC Data tutorial added to the Tutorials collection
- January 2022: PepsiCo and La Trobe University release annotated gene set and associated files for OT3098 v2 genome in partnership with GrainGenes

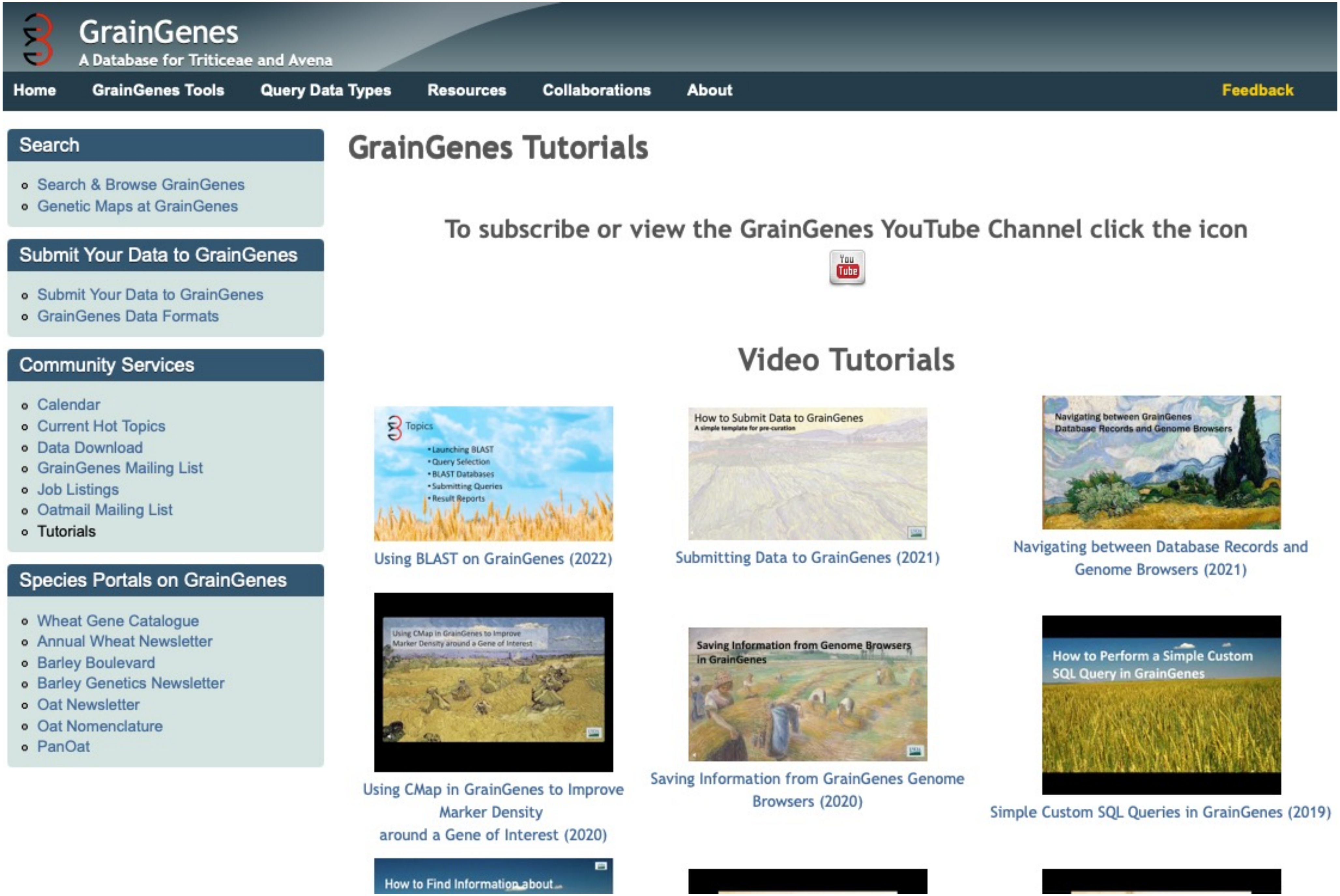
more updates....

@GrainGenes Tweets

Tweets by @GrainGenes

GrainGenes Retweeted

GrainGenes – Tutorials



Search

- Search & Browse GrainGenes
- Genetic Maps at GrainGenes

Submit Your Data to GrainGenes

- Submit Your Data to GrainGenes
- GrainGenes Data Formats

Community Services

- Calendar
- Current Hot Topics
- Data Download
- GrainGenes Mailing List
- Job Listings
- Oatmail Mailing List
- Tutorials

Species Portals on GrainGenes

- Wheat Gene Catalogue
- Annual Wheat Newsletter
- Barley Boulevard
- Barley Genetics Newsletter
- Oat Newsletter
- Oat Nomenclature
- PanOat

GrainGenes Tutorials

To subscribe or view the GrainGenes YouTube Channel click the icon

Video Tutorials

Using BLAST on GrainGenes (2022)

Submitting Data to GrainGenes (2021)

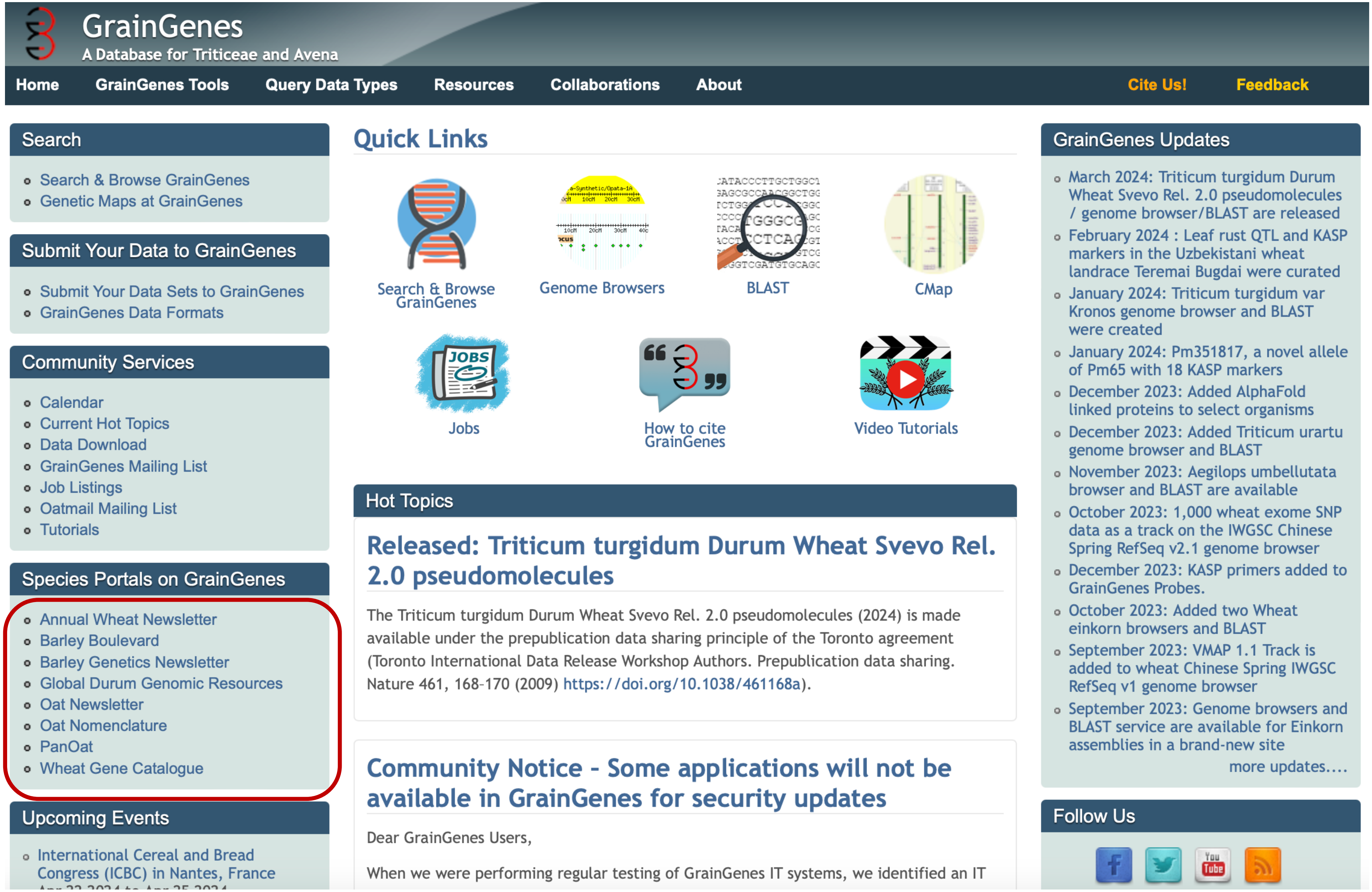
Navigating between Database Records and Genome Browsers (2021)

Using CMap in GrainGenes to Improve Marker Density
around a Gene of Interest (2020)

Saving Information from GrainGenes Genome Browsers (2020)

Simple Custom SQL Queries in GrainGenes (2019)

GrainGenes – Species Portal



Search

- Search & Browse GrainGenes
- Genetic Maps at GrainGenes

Submit Your Data to GrainGenes

- Submit Your Data Sets to GrainGenes
- GrainGenes Data Formats

Community Services

- Calendar
- Current Hot Topics
- Data Download
- GrainGenes Mailing List
- Job Listings
- Oatmail Mailing List
- Tutorials

Species Portals on GrainGenes

- Annual Wheat Newsletter
- Barley Boulevard
- Barley Genetics Newsletter
- Global Durum Genomic Resources
- Oat Newsletter
- Oat Nomenclature
- PanOat
- Wheat Gene Catalogue

Upcoming Events

- International Cereal and Bread Congress (ICBC) in Nantes, France

Quick Links

Hot Topics

Released: Triticum turgidum Durum Wheat Svevo Rel. 2.0 pseudomolecules

The Triticum turgidum Durum Wheat Svevo Rel. 2.0 pseudomolecules (2024) is made available under the prepublication data sharing principle of the Toronto agreement (Toronto International Data Release Workshop Authors. Prepublication data sharing. Nature 461, 168-170 (2009) https://doi.org/10.1038/461168a).

Community Notice - Some applications will not be available in GrainGenes for security updates

Dear GrainGenes Users,

When we were performing regular testing of GrainGenes IT systems, we identified an IT

GrainGenes Updates

- March 2024: Triticum turgidum Durum Wheat Svevo Rel. 2.0 pseudomolecules / genome browser/BLAST are released
- February 2024 : Leaf rust QTL and KASP markers in the Uzbekistani wheat landrace Teremai Bugdai were curated
- January 2024: Triticum turgidum var Kronos genome browser and BLAST were created
- January 2024: Pm351817, a novel allele of Pm65 with 18 KASP markers
- December 2023: Added AlphaFold linked proteins to select organisms
- December 2023: Added Triticum urartu genome browser and BLAST
- November 2023: Aegilops umbellutata browser and BLAST are available
- October 2023: 1,000 wheat exome SNP data as a track on the IWGSC Chinese Spring RefSeq v2.1 genome browser
- December 2023: KASP primers added to GrainGenes Probes.
- October 2023: Added two Wheat einkorn browsers and BLAST
- September 2023: VMAP 1.1 Track is added to wheat Chinese Spring IWGSC RefSeq v1 genome browser
- September 2023: Genome browsers and BLAST service are available for Einkorn assemblies in a brand-new site

more updates....

Follow Us

DATA CONTENT AND TOOLS

FAIR Principles

Increasing the utility of data - “Chez Panisse”

Findable

Accessible

Interoperable

Reusable

- ★ *15* Guiding Principles published in 2016
- ★ Guidance for ***scientific data management and stewardship***
- ★ Promote ***maximal use*** of research data

SCIENTIFIC DATA

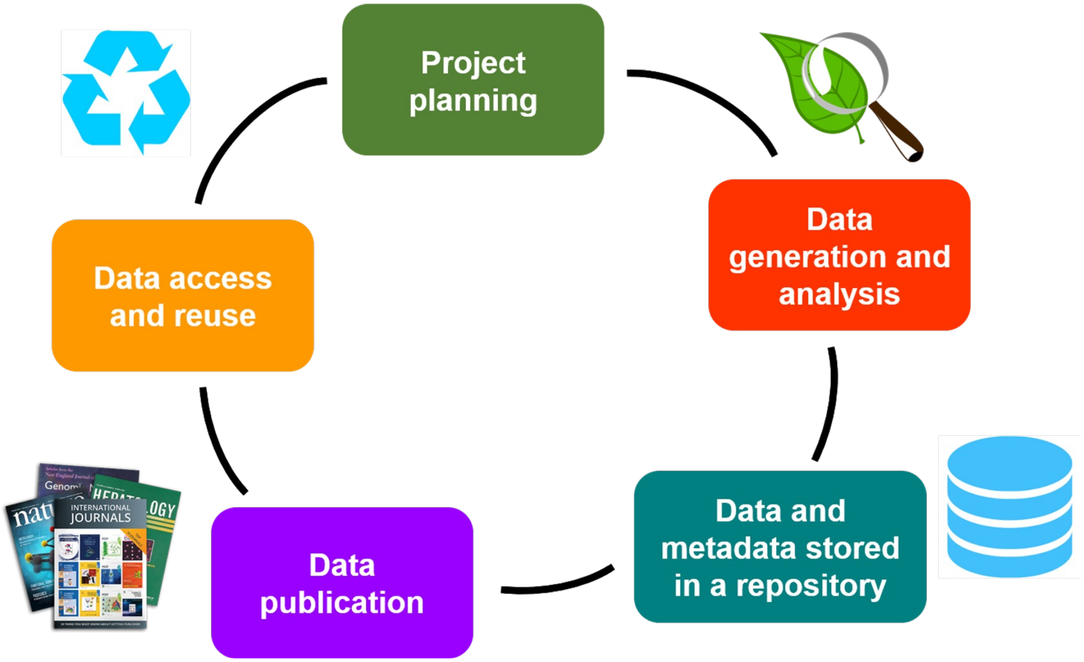
Amended: Addendum

OPEN

SUBJECT CATEGORIES
» Research data
» Publication characteristics

Comment: The FAIR Guiding Principles for scientific data management and stewardship

Mark D. Wilkinson *et al.*#



Project planning

Data generation and analysis

Data and metadata stored in a repository

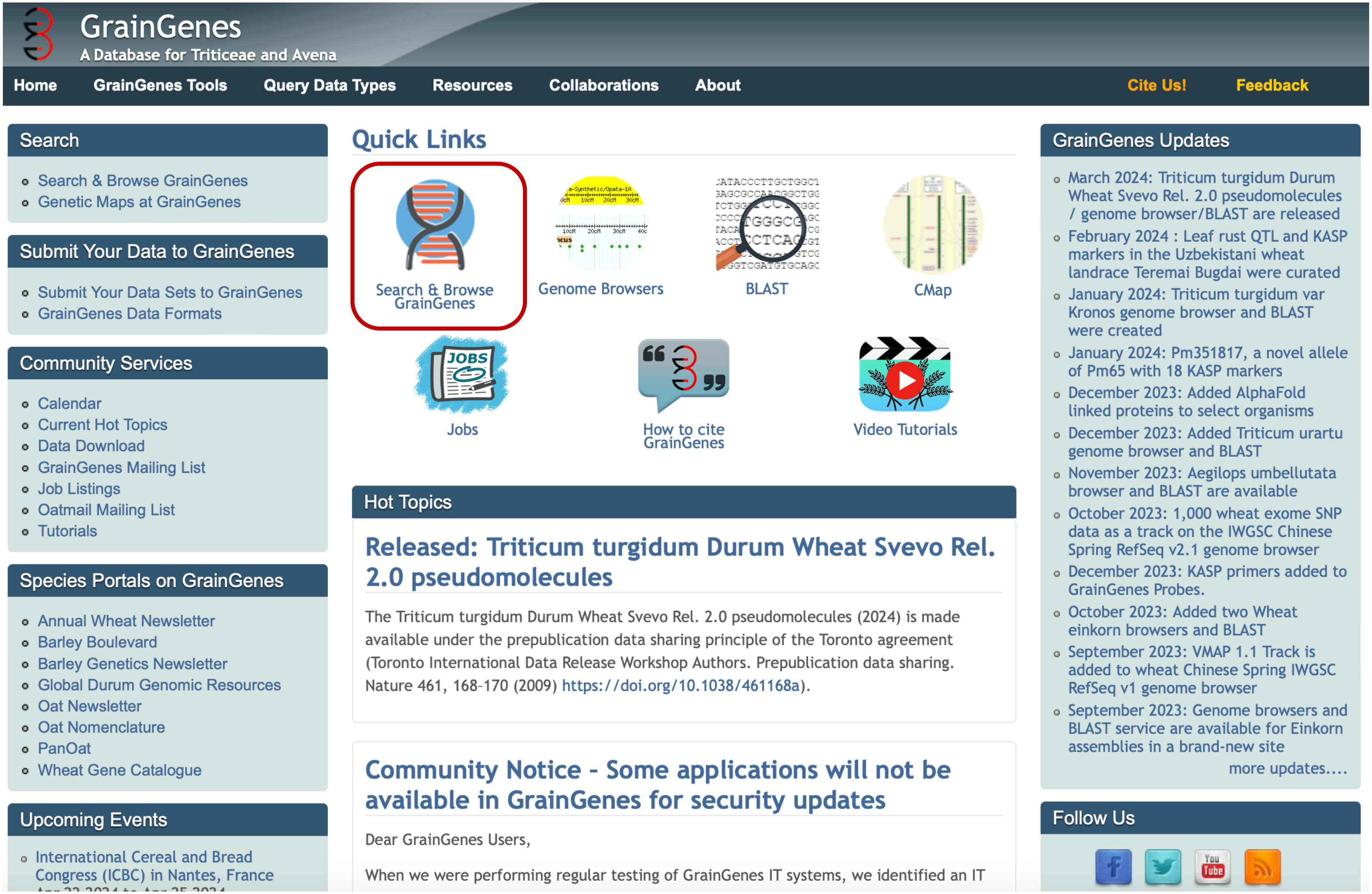
Data publication

Data access and reuse

Equity issues need to be taken into account

Original slides provided by AgBioData

GrainGenes – Search for Data



Search

- Search & Browse GrainGenes
- Genetic Maps at GrainGenes

Submit Your Data to GrainGenes

- Submit Your Data Sets to GrainGenes
- GrainGenes Data Formats

Community Services

- Calendar
- Current Hot Topics
- Data Download
- GrainGenes Mailing List
- Job Listings
- Oatmail Mailing List
- Tutorials

Species Portals on GrainGenes

- Annual Wheat Newsletter
- Barley Boulevard
- Barley Genetics Newsletter
- Global Durum Genomic Resources
- Oat Newsletter
- Oat Nomenclature
- PanOat
- Wheat Gene Catalogue

Upcoming Events

- International Cereal and Bread Congress (ICBC) in Nantes, France

Quick Links

Hot Topics

Released: Triticum turgidum Durum Wheat Svevo Rel. 2.0 pseudomolecules

The Triticum turgidum Durum Wheat Svevo Rel. 2.0 pseudomolecules (2024) is made available under the prepublication data sharing principle of the Toronto agreement (Toronto International Data Release Workshop Authors. Prepublication data sharing. Nature 461, 168-170 (2009) https://doi.org/10.1038/461168a).

Community Notice - Some applications will not be available in GrainGenes for security updates

Dear GrainGenes Users,

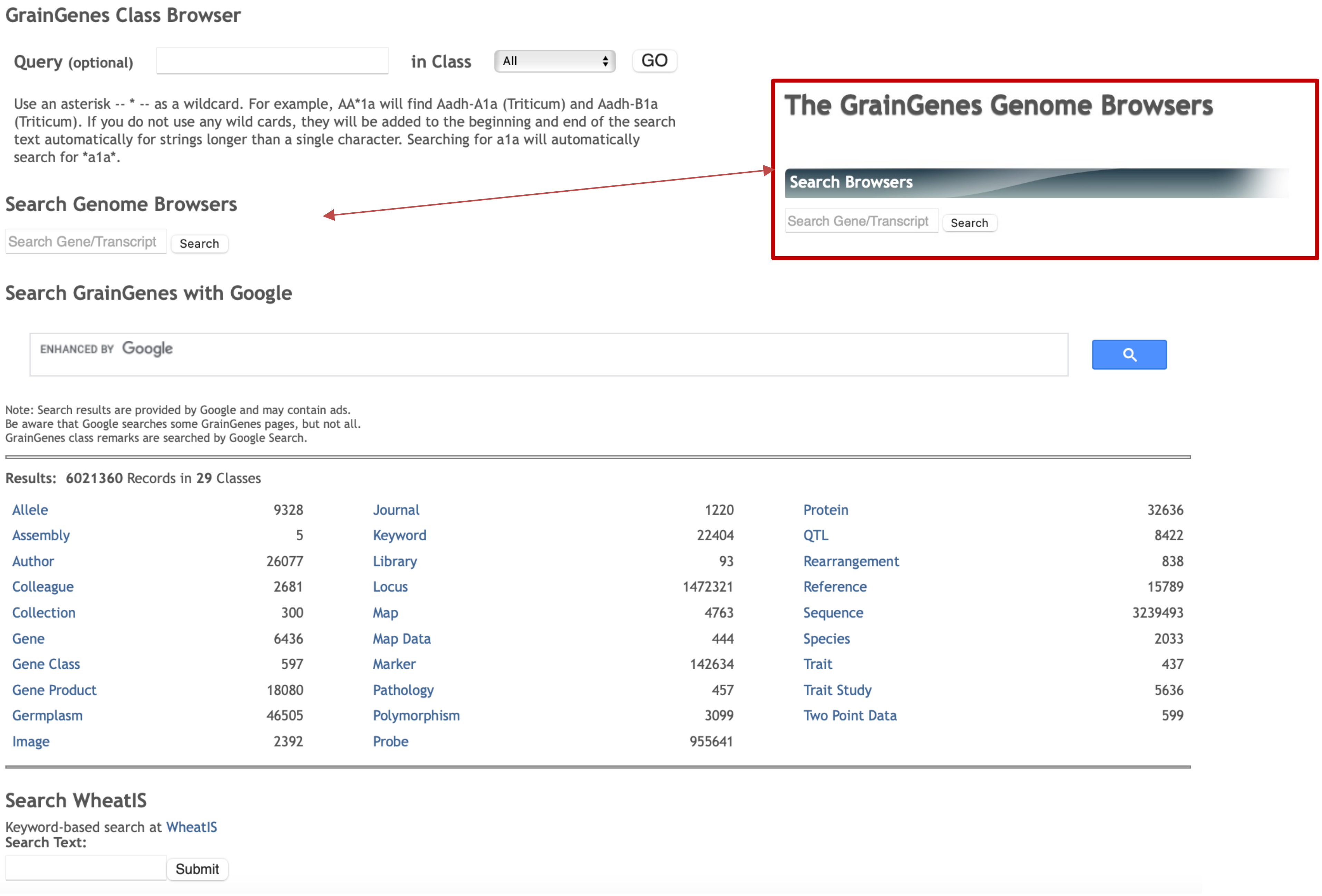
When we were performing regular testing of GrainGenes IT systems, we identified an IT

GrainGenes Updates

- March 2024: Triticum turgidum Durum Wheat Svevo Rel. 2.0 pseudomolecules / genome browser/BLAST are released
- February 2024 : Leaf rust QTL and KASP markers in the Uzbekistani wheat landrace Teremai Bugdai were curated
- January 2024: Triticum turgidum var Kronos genome browser and BLAST were created
- January 2024: Pm351817, a novel allele of Pm65 with 18 KASP markers
- December 2023: Added AlphaFold linked proteins to select organisms
- December 2023: Added Triticum urartu genome browser and BLAST
- November 2023: Aegilops umbellutata browser and BLAST are available
- October 2023: 1,000 wheat exome SNP data as a track on the IWGSC Chinese Spring RefSeq v2.1 genome browser
- December 2023: KASP primers added to GrainGenes Probes.
- October 2023: Added two Wheat einkorn browsers and BLAST
- September 2023: VMAP 1.1 Track is added to wheat Chinese Spring IWGSC RefSeq v1 genome browser
- September 2023: Genome browsers and BLAST service are available for Einkorn assemblies in a brand-new site

more updates....

GrainGenes – Search for Data



GrainGenes Class Browser

Use an asterisk -- * -- as a wildcard. For example, AA*1a will find Aadh-A1a (Triticum) and Aadh-B1a (Triticum). If you do not use any wild cards, they will be added to the beginning and end of the search text automatically for strings longer than a single character. Searching for a1a will automatically search for *a1a*.

Search Genome Browsers

Search GrainGenes with Google

Note: Search results are provided by Google and may contain ads.
Be aware that Google searches some GrainGenes pages, but not all.
GrainGenes class remarks are searched by Google Search.

Results: 6021360 Records in **29** Classes

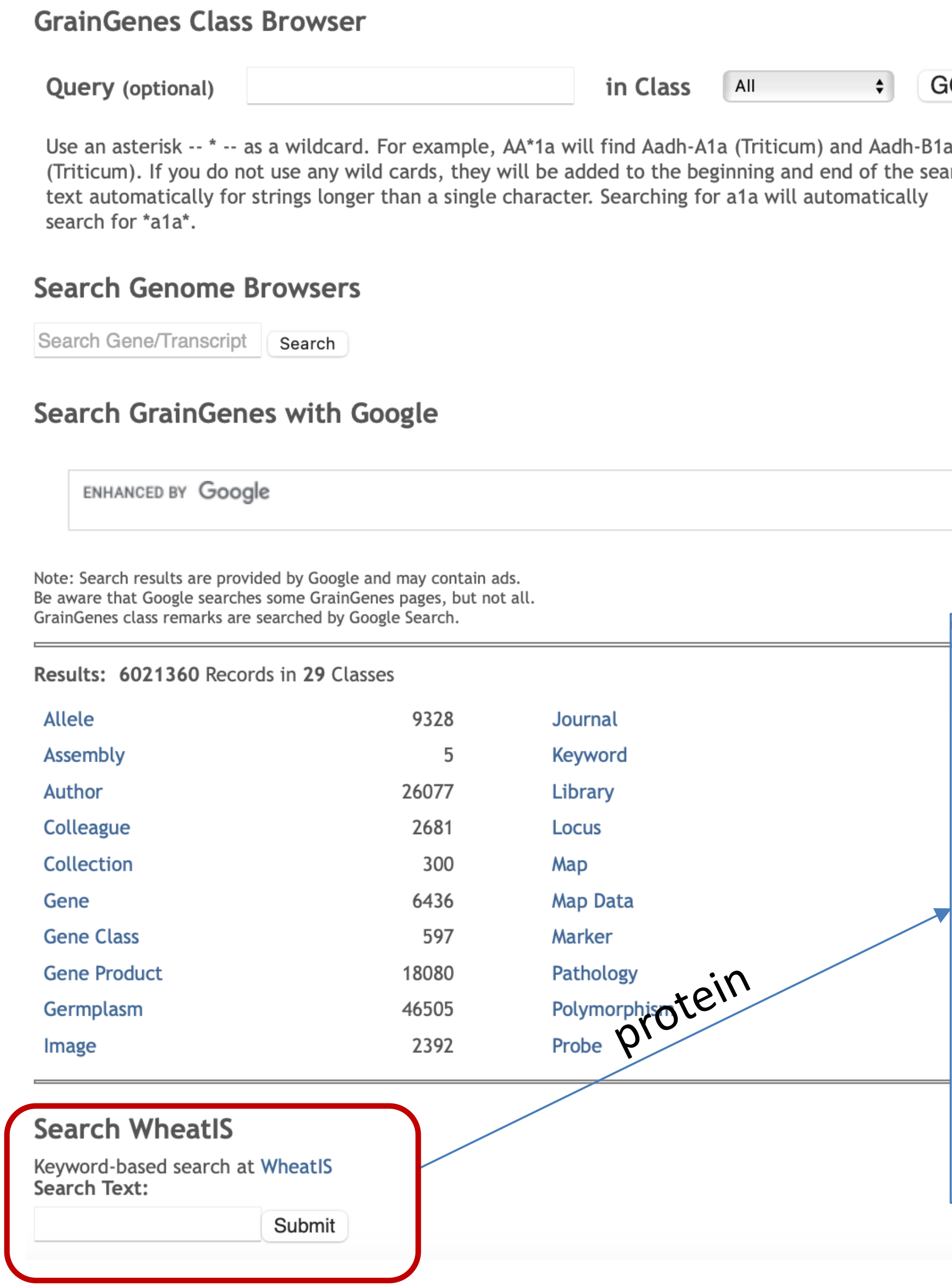
Allele	9328	Journal	1220	Protein	32636
Assembly	5	Keyword	22404	QTL	8422
Author	26077	Library	93	Rearrangement	838
Colleague	2681	Locus	1472321	Reference	15789
Collection	300	Map	4763	Sequence	3239493
Gene	6436	Map Data	444	Species	2033
Gene Class	597	Marker	142634	Trait	437
Gene Product	18080	Pathology	457	Trait Study	5636
Germplasm	46505	Polymorphism	3099	Two Point Data	599
Image	2392	Probe	955641		

Search WheatIS

Keyword-based search at WheatIS
Search Text:

The GrainGenes Genome Browsers

GrainGenes – Linked Search



GrainGenes Class Browser

Query (optional) in Class All GO

Use an asterisk -- * -- as a wildcard. For example, AA*1a will find Aadh-A1a (Triticum) and Aadh-B1a (Triticum). If you do not use any wild cards, they will be added to the beginning and end of the search text automatically for strings longer than a single character. Searching for a1a will automatically search for *a1a*.

Search Genome Browsers

Search Gene/Transcript Search

Search GrainGenes with Google

ENHANCED BY Google

Note: Search results are provided by Google and may contain ads.
Be aware that Google searches some GrainGenes pages, but not all.
GrainGenes class remarks are searched by Google Search.

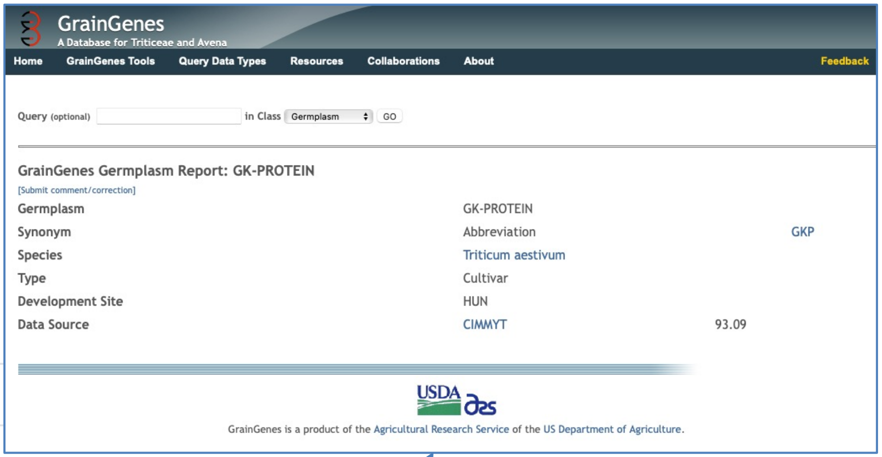
Results: 6021360 Records in **29** Classes

Allele	9328	Journal
Assembly	5	Keyword
Author	26077	Library
Colleague	2681	Locus
Collection	300	Map
Gene	6436	Map Data
Gene Class	597	Marker
Gene Product	18080	Pathology
Germplasm	46505	Polymorphis
Image	2392	Probe

Search WheatIS

Keyword-based search at WheatIS
Search Text:

Submit



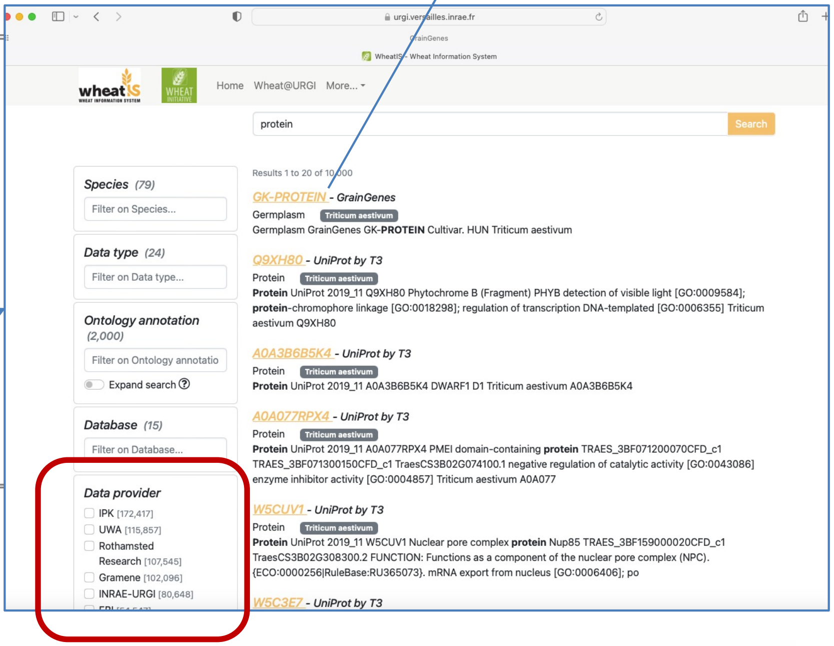
Query (optional) in Class Germplasm GO

GrainGenes Germplasm Report: GK-PROTEIN

[Submit comment/correction]

Germplasm	GK-PROTEIN	
Synonym	Abbreviation	GKP
Species	Triticum aestivum	
Type	Cultivar	
Development Site	HUN	
Data Source	CIMMYT	93.09

GrainGenes is a product of the Agricultural Research Service of the US Department of Agriculture.



protein

Species (79)
Filter on Species...

Data type (24)
Filter on Data type...

Ontology annotation (2,000)
Filter on Ontology annotatio
Expand search

Database (15)
Filter on Database...

Data provider
- ☐ IPK [172,417]
- ☐ UWA [115,857]
- ☐ Rothamsted Research [107,545]
- ☐ Gramene [102,096]
- ☐ INRAE-URGI [80,648]

Results 1 to 20 of 10,000

GK-PROTEIN - *GrainGenes*
Germplasm **Triticum aestivum**
Germplasm GrainGenes GK-**PROTEIN** Cultivar. HUN Triticum aestivum

Q9XH80 - *UniProt by T3*
Protein **Triticum aestivum**
Protein UniProt 2019_11 Q9XH80 Phytochrome B (Fragment) PHYB detection of visible light [GO:0009584]; **protein**-chromophore linkage [GO:0018298]; regulation of transcription DNA-templated [GO:0006355] Triticum aestivum Q9XH80

A0A3B6B5K4 - *UniProt by T3*
Protein **Triticum aestivum**
Protein UniProt 2019_11 A0A3B6B5K4 DWARF1 D1 Triticum aestivum A0A3B6B5K4

A0A077RPX4 - *UniProt by T3*
Protein **Triticum aestivum**
Protein UniProt 2019_11 A0A077RPX4 PMEI domain-containing **protein** TRAES_3BF071200070CFD_c1 TRAES_3BF071300150CFD_c1 TraesCS3B02G074100.1 negative regulation of catalytic activity [GO:0043086] enzyme inhibitor activity [GO:0004857] Triticum aestivum A0A077

W5CUV1 - *UniProt by T3*
Protein **Triticum aestivum**
Protein UniProt 2019_11 W5CUV1 Nuclear pore complex **protein** Nup85 TRAES_3BF159000020CFD_c1 TraesCS3B02G308300.2 FUNCTION: Functions as a component of the nuclear pore complex (NPC). {ECO:0000256|RuleBase:RU365073}. mRNA export from nucleus [GO:0006406]; po

W5C3E7 - *UniProt by T3*

Data Veracity <-> Curation – Genetic examples

One of the strengths of GrainGenes

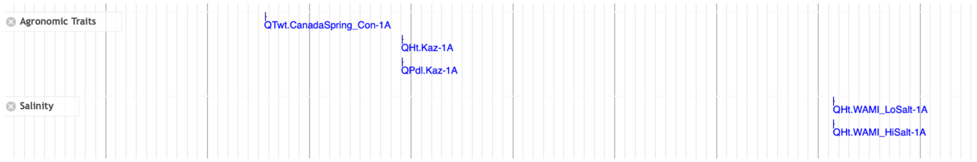
Stripe Rust QTL curated from the Vavilov wheat diversity panel

- Forty-seven Stripe rust QTL from Jambuthenne et al. (2022) were mapped on the IWGSC v1 assembly with DArT markers, 14 of which overlapped with previously described genes.

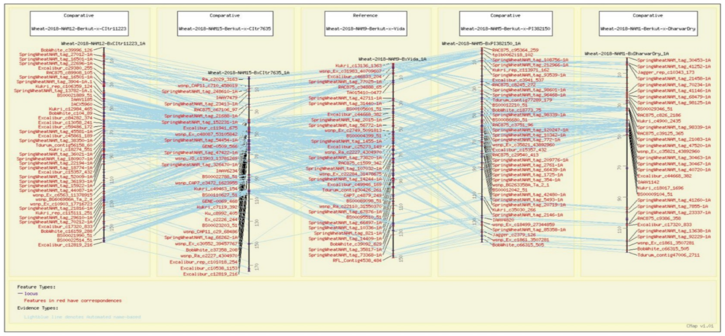
New diagnostic PCR markers for Yr27, Yr28, Sr46, and Rg-B1 in wheat. Diagnostic marker data were contributed to GrainGenes by Naveenkumar Athiyannan via the GrainGenes Data Template forms. The papers below were curated for new diagnostic PCR primers for Yr27, Yr28, Sr46, and Rg-B1.

- Athiyannan N et al. (2022) Haplotype variants of the stripe rust resistance gene Yr28 in Aegilops tauschii. Theoretical and Applied Genetics online.
- Athiyannan N et al. (2022) Haplotype variants of Sr46 in Aegilops tauschii, the diploid D genome progenitor of wheat. Theoretical and Applied Genetics 135:2627-2639.
- Athiyannan N et al. (2022) Long-read genome sequencing of bread wheat facilitates disease resistance gene cloning. Nature Genetics 54:227-231.
- Abrouk M et al. (2021) Population genomics and haplotype analysis in spelt and bread wheat identifies a gene regulating glume color. Communications Biology 4:375.

GWAS results for pathology, salinity, and agronomic traits added to the IWGSC Chinese Spring v1 browser (Genetic-Genomic contexts)



Genome Browser



CMap

Genome Browsers

Wheat

- Triticum turgidum var Kronos, Seong et al. (2023)
- Aegilops umbellulata genome assembly; Abrouk et al. (2023)
- Wheat einkorn 2 assemblies; Ahmed et al. (2023)
- Wheat cultivar Sonmez genome assembly; Akpinar et al. (2022)
- Wheat cultivar Renan genome assembly; Aury et al. (2022)
- Wheat cultivar Attraktion genome assembly; Kale et al. (2022)
- Aegilops longisimma genome assembly; Avni et al. (2022)
- Aegilops speltoides genome assembly; Avni et al. (2022)
- Aegilops sharonensis genome assembly; Avni et al. (2022)
- Wheat Kariega genome assembly; Athiyannan et al. (2022)
- Wheat Fielder genome assembly; Sato et al. (2021)
- Aegilops tauschii Aet v5.0 genome assembly; Wang et al. (2021)
- Aegilops tauschii (wheat) 4 assemblies; Zhou et al. (2021)
- Wheat Chinese Spring IWGSC RefSeq v2.1 genome assembly (2021)
- 10+ Wheat Genome Project (2020)
- Triticum turgidum Durum Wheat Svevo Rel. 1.0 (2019)
- Wild Emmer Wheat Zavitan WEWSeq v2.0 genome assembly (2019)
- Wheat Chinese Spring IWGSC RefSeq v1.0 genome assembly (2018)
- Triticum urartu genome assembly (2018)
- Aegilops tauschii Aet v4.0 genome assembly (2017)
- Wild Emmer Wheat Zavitan WEWSeq v1.0 genome assembly (2017)
- Hexaploid Wheat Pangenome, Montenegro et al. (2017)

Barley

- Barley Morex V3 (2021)
- Barley Pangenome Project - 19 assemblies (2020)
- Barley MorexV2 assembled by TRITEX (2019)
- Barley MorexV1, IBSCv2 genome assembly (2017)

Oat

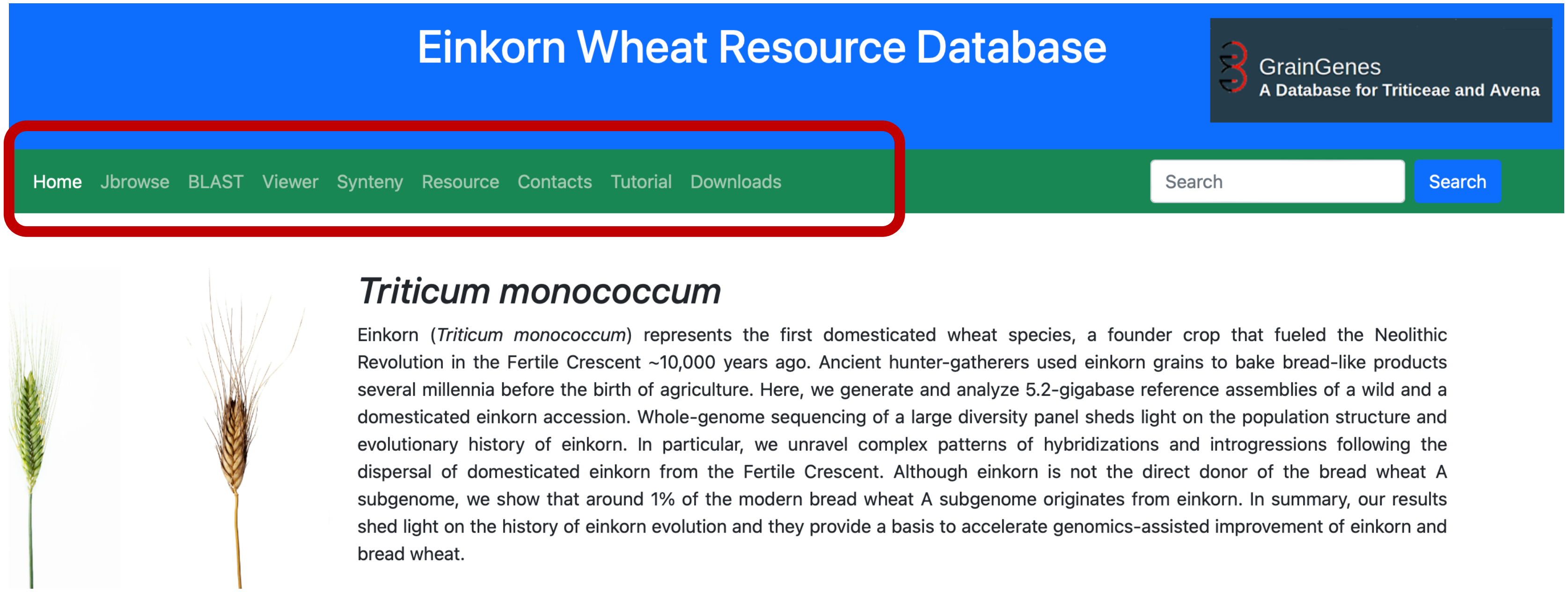
- A. sativa ssp. nuda cv. Sanfensan Genome, Peng et al. (2022)
- Avena insularis (CN 108634), Peng et al. (2022)
- Avena longiglumis (CN 58139) , Peng et al. (2022)

- **IWGSC v1, v2**
- **Wheat 10+**

<u>**Upcoming**</u>

- **Svevo2**: available under the Toronto Agreement
- **PanOat**: under embargo
- **Barley Pangenome**: Currently building

Einkorn Pangenome
(https://wheat.pw.usda.gov/GG3/pangenome)



Triticum monococcum

Einkorn (*Triticum monococcum*) represents the first domesticated wheat species, a founder crop that fueled the Neolithic Revolution in the Fertile Crescent ~10,000 years ago. Ancient hunter-gatherers used einkorn grains to bake bread-like products several millennia before the birth of agriculture. Here, we generate and analyze 5.2-gigabase reference assemblies of a wild and a domesticated einkorn accession. Whole-genome sequencing of a large diversity panel sheds light on the population structure and evolutionary history of einkorn. In particular, we unravel complex patterns of hybridizations and introgressions following the dispersal of domesticated einkorn from the Fertile Crescent. Although einkorn is not the direct donor of the bread wheat A subgenome, we show that around 1% of the modern bread wheat A subgenome originates from einkorn. In summary, our results shed light on the history of einkorn evolution and they provide a basis to accelerate genomics-assisted improvement of einkorn and bread wheat.

Article

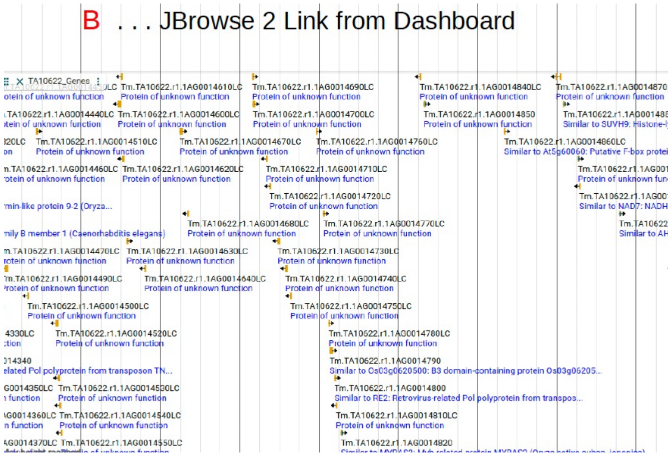
Einkorn genomics sheds light on history of the oldest domesticated wheat

https://doi.org/10.1038/s41586-023-06389-7
Received: 16 October 2022
Accepted: 29 June 2023

Hanin Ibrahim Ahmed[1,2,11], Matthias Heuberger[3,11], Adam Schoen[4,11], Dal-Hoe Koo[5], Jesus Quiroz-Chavez[6], Laxman Adhikari[1,2], John Raupp[5], Stéphane Cauet[7], Nathalie Rodde[7], Charlotte Cravero[7], Caroline Callot[7], Gerard R. Lazo[8], Nagarajan Kathiresan[9], Parva K. Sharma[4], Ian Moot[4], Inderjit Singh Yadav[4], Lovepreet Singh[4], Gautam Saripalli[4], Nidhi Rawat[4], Raju Datla[10], Naveenkumar Athiyannan[1,2], Ricardo H. Ramirez-Gonzalez[6], Cristobal Uauy[6], Thomas Wicker[3], Vijay K. Tiwari[4]✉, Michael Abrouk[1,2]✉, Jesse Poland[1,2]✉ & Simon G. Krattinger[1,2]✉

Einkorn (*Triticum monococcum*) was the first domesticated wheat species, and was central to the birth of agriculture and the Neolithic Revolution in the Fertile Crescent around 10,000 years ago[1,2]. Here we generate and analyse 5.2-Gb genome assemblies for wild and domesticated einkorn, including completely assembled centromeres. Einkorn centromeres are highly dynamic, showing evidence of ancient and recent centromere shifts caused by structural rearrangements. Whole-genome sequencing analysis of a diversity panel uncovered the population structure and evolutionary history of einkorn, revealing complex patterns of hybridizations and introgressions after the dispersal of domesticated einkorn from the Fertile Crescent. We also show that around 1% of the modern bread wheat (*Triticum aestivum*) A subgenome originates from einkorn. These resources and findings highlight the history of einkorn evolution and provide a basis to accelerate the genomics-assisted improvement of einkorn and bread wheat.

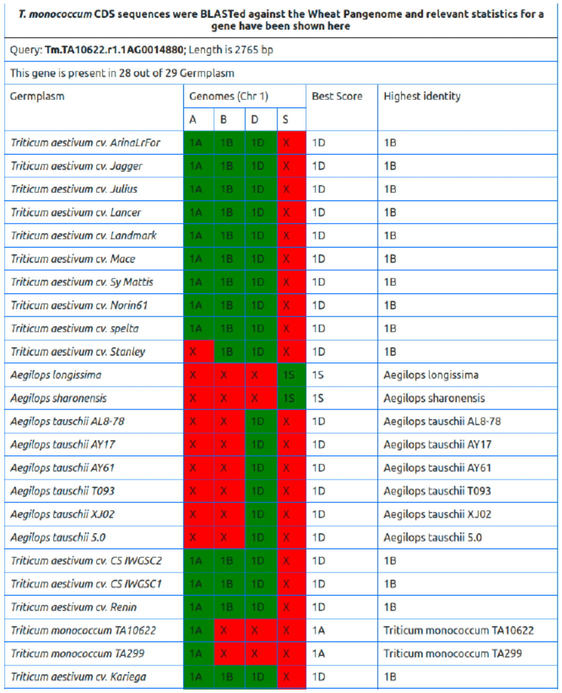
Einkorn Pangenome
(https://wheat.pw.usda.gov/GG3/pangenome)



B . . . JBrowse 2 Link from Dashboard

C

. . . BLAST results Link from dashboard

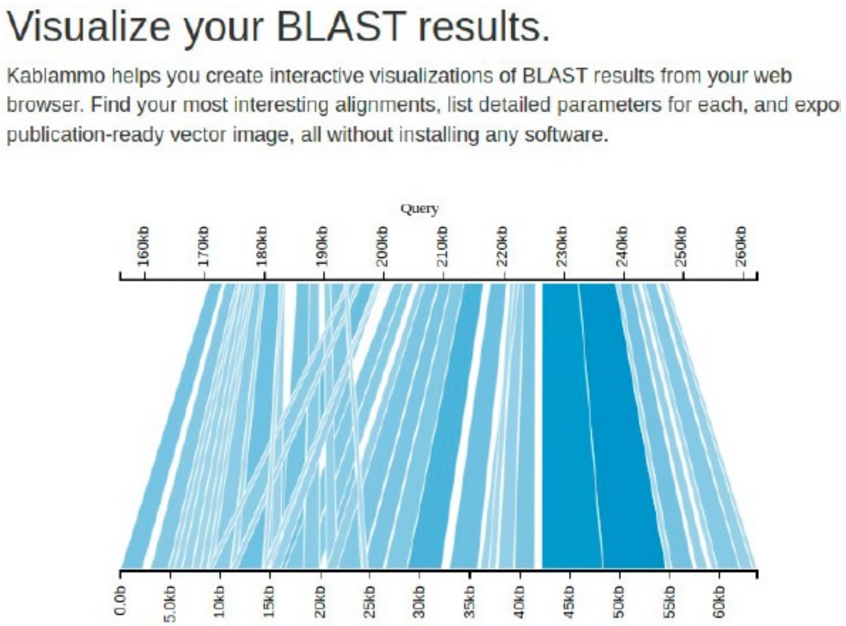


T. monococcum **CDS sequences were BLASTed against the Wheat Pangenome and relevant statistics for a gene have been shown here**

Query: **Tm.TA10622.r1.1AG0014880;** Length is 2765 bp

This gene is present in 28 out of 29 Germplasm

Germplasm	Genomes (Chr 1)				Best Score	Highest identity
	A	B	D	S		
Triticum aestivum cv. ArinaLrFor	1A	1B	1D	X	1D	1B
Triticum aestivum cv. Jagger	1A	1B	1D	X	1D	1B
Triticum aestivum cv. Julius	1A	1B	1D	X	1D	1B
Triticum aestivum cv. Lancer	1A	1B	1D	X	1D	1B
Triticum aestivum cv. Landmark	1A	1B	1D	X	1D	1B
Triticum aestivum cv. Mace	1A	1B	1D	X	1D	1B
Triticum aestivum cv. Sy Mattis	1A	1B	1D	X	1D	1B
Triticum aestivum cv. Norin61	1A	1B	1D	X	1D	1B
Triticum aestivum cv. spelta	1A	1B	1D	X	1D	1B
Triticum aestivum cv. Stanley	X	1B	1D	X	1D	1B
Aegilops longissima	X	X	X	1S	1S	Aegilops longissima
Aegilops sharonensis	X	X	X	1S	1S	Aegilops sharonensis
Aegilops tauschii AL8-78	X	X	1D	X	1D	Aegilops tauschii AL8-78
Aegilops tauschii AY17	X	X	1D	X	1D	Aegilops tauschii AY17
Aegilops tauschii AY61	X	X	1D	X	1D	Aegilops tauschii AY61
Aegilops tauschii T093	X	X	1D	X	1D	Aegilops tauschii T093
Aegilops tauschii XJ02	X	X	1D	X	1D	Aegilops tauschii XJ02
Aegilops tauschii 5.0	X	X	1D	X	1D	Aegilops tauschii 5.0
Triticum aestivum cv. CS IWGSC2	1A	1B	1D	X	1D	1B
Triticum aestivum cv. CS IWGSC1	1A	1B	1D	X	1D	1B
Triticum aestivum cv. Renin	1A	1B	1D	X	1D	1B
Triticum monococcum TA10622	1A	X	X	X	1A	Triticum monococcum TA10622
Triticum monococcum TA299	1A	X	X	X	1A	Triticum monococcum TA299
Triticum aestivum cv. Kariega	1A	1B	1D	X	1D	1B

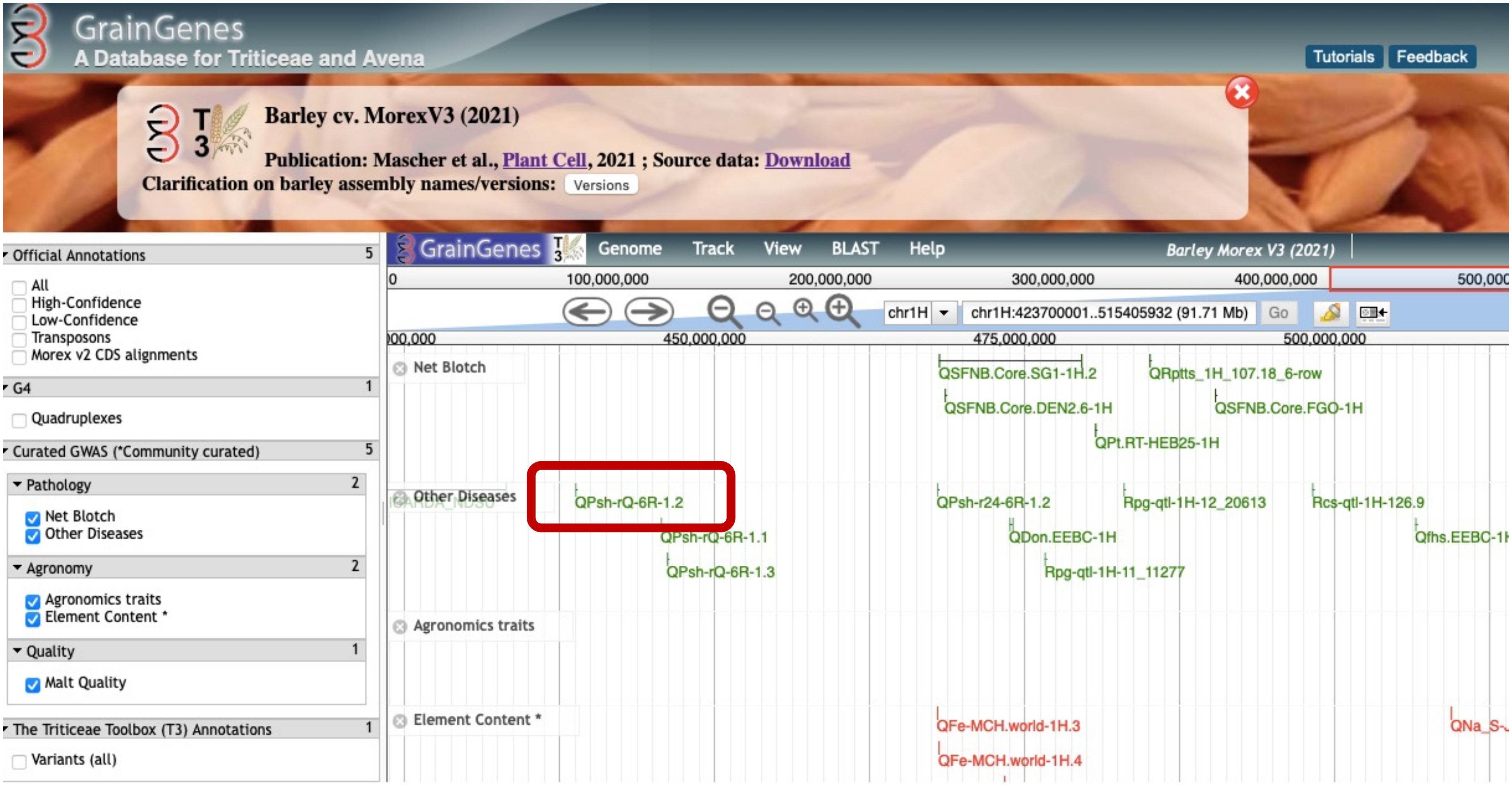


Visualize your BLAST results.

Kablammo helps you create interactive visualizations of BLAST results from your web browser. Find your most interesting alignments, list detailed parameters for each, and expo publication-ready vector image, all without installing any software.

Accusyn Visualizations

GrainGenes – GWAS QTL tracks



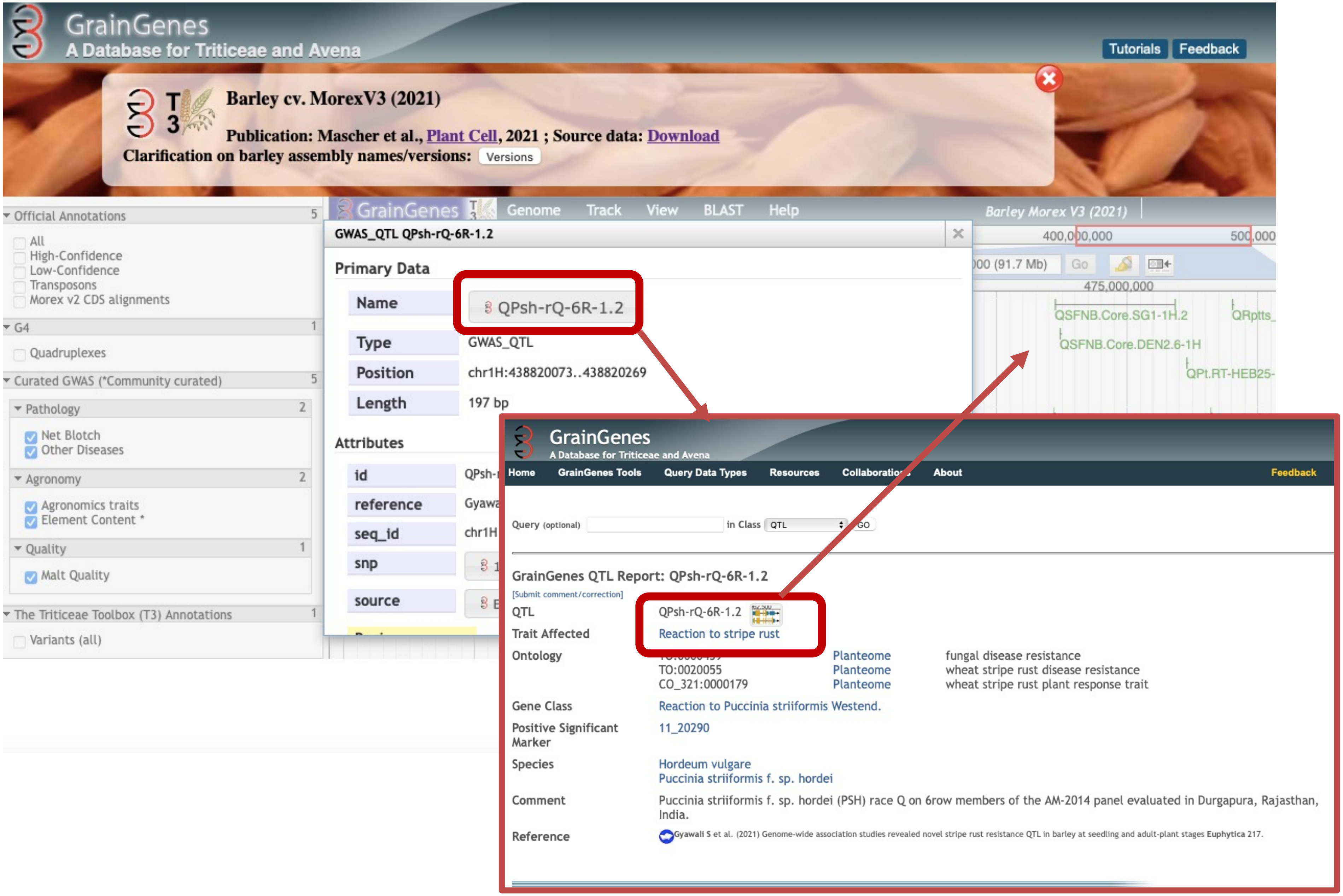
Barley cv. MorexV3 (2021)

Publication: Mascher et al., Plant Cell, 2021 ; Source data: Download

Clarification on barley assembly names/versions: Versions

chr1H:423700001..515405932 (91.71 Mb)

GrainGenes – GWAS QTL tracks



Barley cv. MorexV3 (2021)

Publication: Mascher et al., Plant Cell, 2021 ; Source data: Download

Clarification on barley assembly names/versions:

GWAS_QTL QPsh-rQ-6R-1.2

Primary Data

Name	QPsh-rQ-6R-1.2
Type	GWAS_QTL
Position	chr1H:438820073..438820269
Length	197 bp

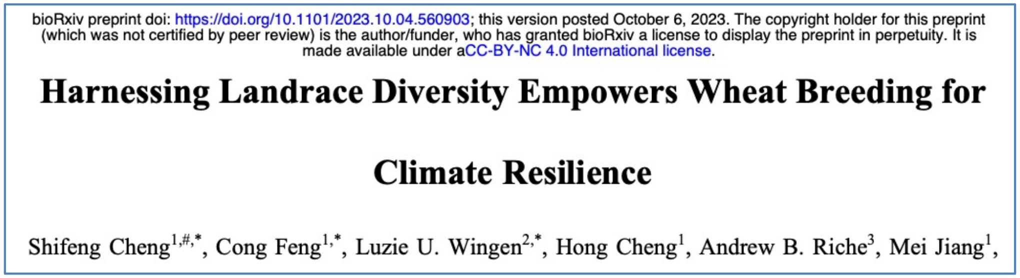
Attributes

id	QPsh-
reference	Gyawa
seq_id	chr1H
snp	1
source	E

GrainGenes QTL Report: QPsh-rQ-6R-1.2

QTL	QPsh-rQ-6R-1.2		
Trait Affected	Reaction to stripe rust		
Ontology		Planteome	fungal disease resistance
	TO:0020055	Planteome	wheat stripe rust disease resistance
	CO_321:0000179	Planteome	wheat stripe rust plant response trait
Gene Class	Reaction to Puccinia striiformis Westend.		
Positive Significant Marker	11_20290		
Species	Hordeum vulgare Puccinia striiformis f. sp. hordei		
Comment	Puccinia striiformis f. sp. hordei (PSH) race Q on 6row members of the AM-2014 panel evaluated in Durgapura, Rajasthan, India.		
Reference	Gyawali S et al. (2021) Genome-wide association studies revealed novel stripe rust resistance QTL in barley at seedling and adult-plant stages **Euphytica** 217.		

Watkins Collection +



Harnessing Landrace Diversity Empowers Wheat Breeding for Climate Resilience

Shifeng Cheng[1,#,*], Cong Feng[1,*], Luzie U. Wingen[2,*], Hong Cheng[1], Andrew B. Riche[3], Mei Jiang[1],

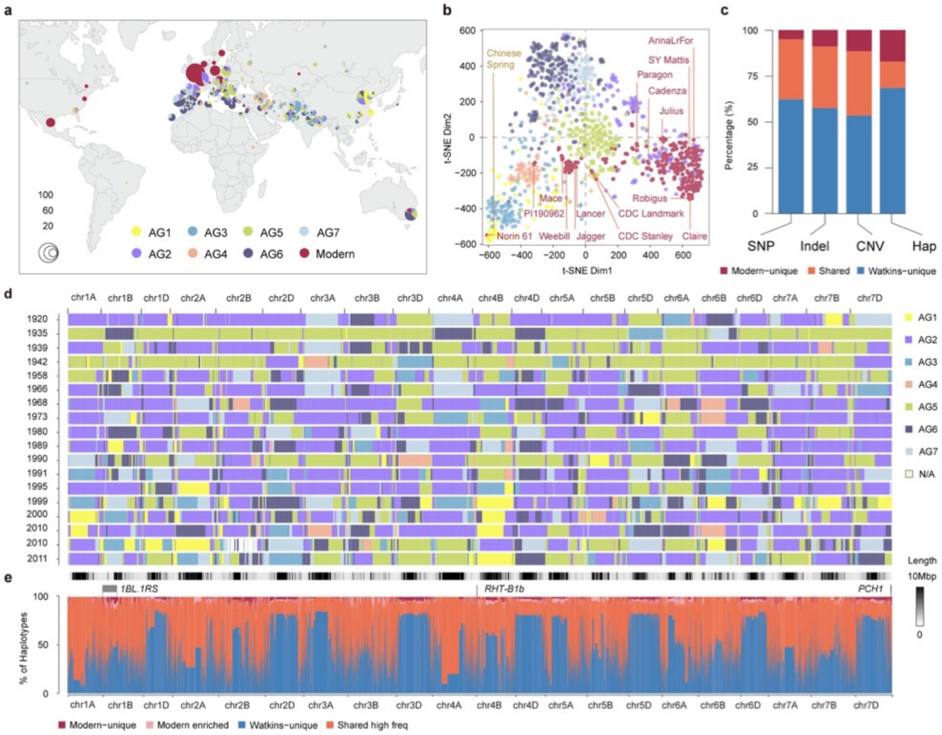
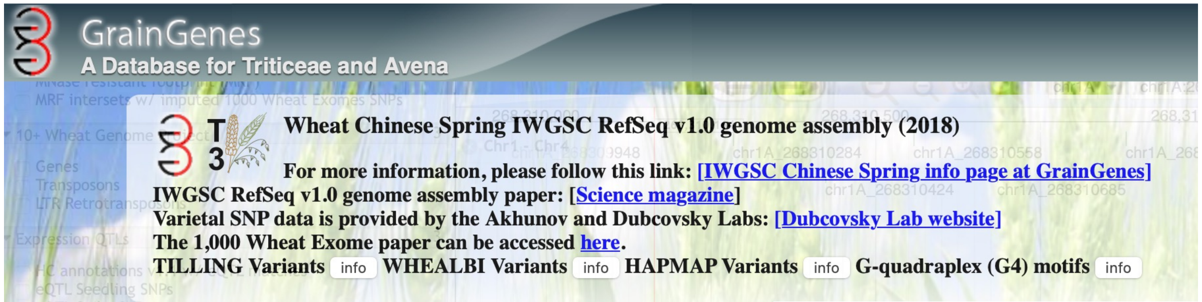


Fig. 1 | Genomic variants in Watkins landraces compared to modern wheat. a, Geographical distribution of all accessions, including the entire Watkins collection (*n* = 827) and modern wheat cultivars (*n* = 224; comprising 208 cultivars sequenced in this study and 16 previously described modern wheat cultivars including Chinese Spring). The seven ancestral groups (AG1–7) derived from Watkins and modern wheat are colour-coded. **b,** t-SNE plot based on the 10M SNPs shared by different ancestral

Watkins Collection Tracks



GrainGenes
A Database for Triticeae and Avena

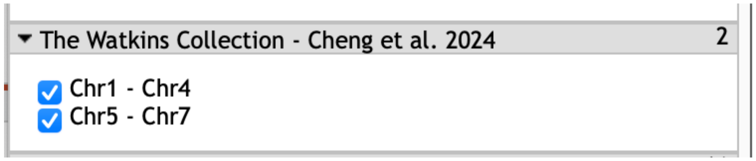
Wheat Chinese Spring IWGSC RefSeq v1.0 genome assembly (2018)

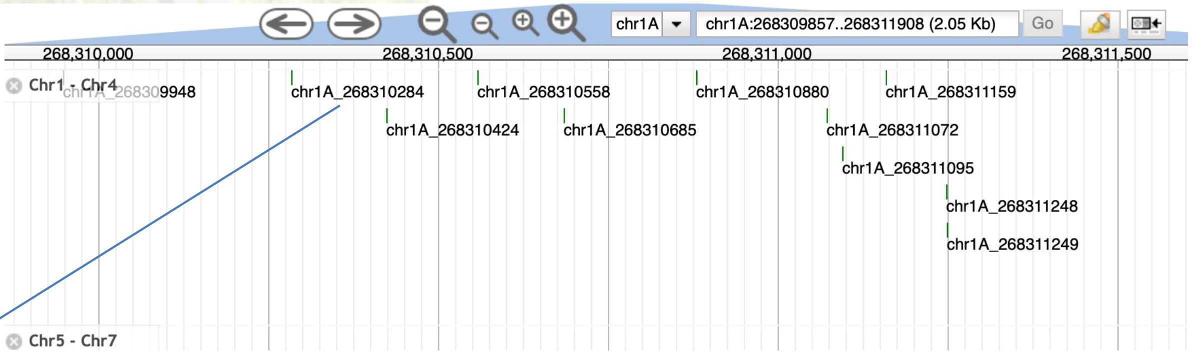
For more information, please follow this link: [IWGSC Chinese Spring info page at GrainGenes]

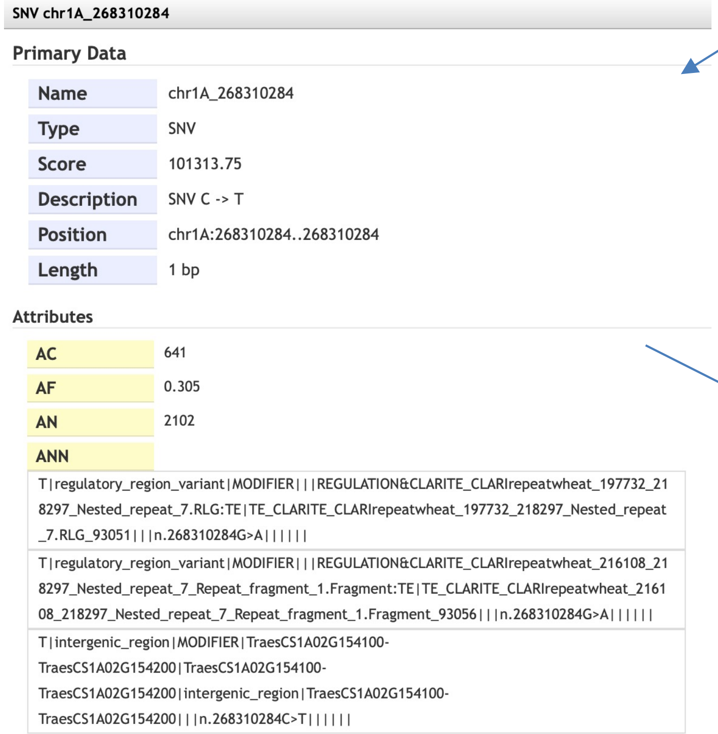
IWGSC RefSeq v1.0 genome assembly paper: [Science magazine]

Varietal SNP data is provided by the Akhunov and Dubcovsky Labs: [Dubcovsky Lab website]

The 1,000 Wheat Exome paper can be accessed here.







SNV chr1A_268310284

Primary Data

Name	chr1A_268310284
Type	SNV
Score	101313.75
Description	SNV C -> T
Position	chr1A:268310284..268310284
Length	1 bp

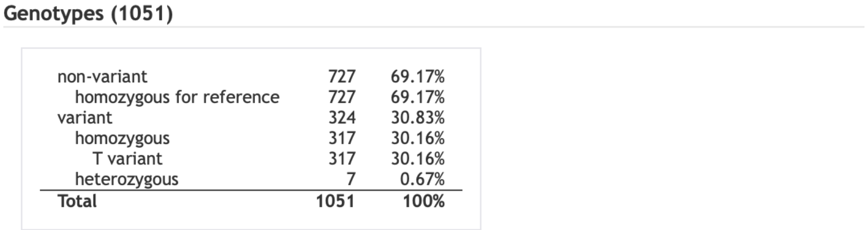
Attributes

AC	641
AF	0.305
AN	2102
ANN	

T|regulatory_region_variant|MODIFIER|||REGULATION&CLARITE_CLARIrepeatwheat_197732_218297_Nested_repeat_7.RLG:TE|TE_CLARITE_CLARIrepeatwheat_197732_218297_Nested_repeat_7.RLG_93051|||n.268310284G>A||||||

T|regulatory_region_variant|MODIFIER|||REGULATION&CLARITE_CLARIrepeatwheat_216108_218297_Nested_repeat_7_Repeat_fragment_1.Fragment:TE|TE_CLARITE_CLARIrepeatwheat_216108_218297_Nested_repeat_7_Repeat_fragment_1.Fragment_93056|||n.268310284G>A||||||

T|intergenic_region|MODIFIER|TraesCS1A02G154100-TraesCS1A02G154200|TraesCS1A02G154100-TraesCS1A02G154200|intergenic_region|TraesCS1A02G154100-TraesCS1A02G154200|||n.268310284C>T||||||

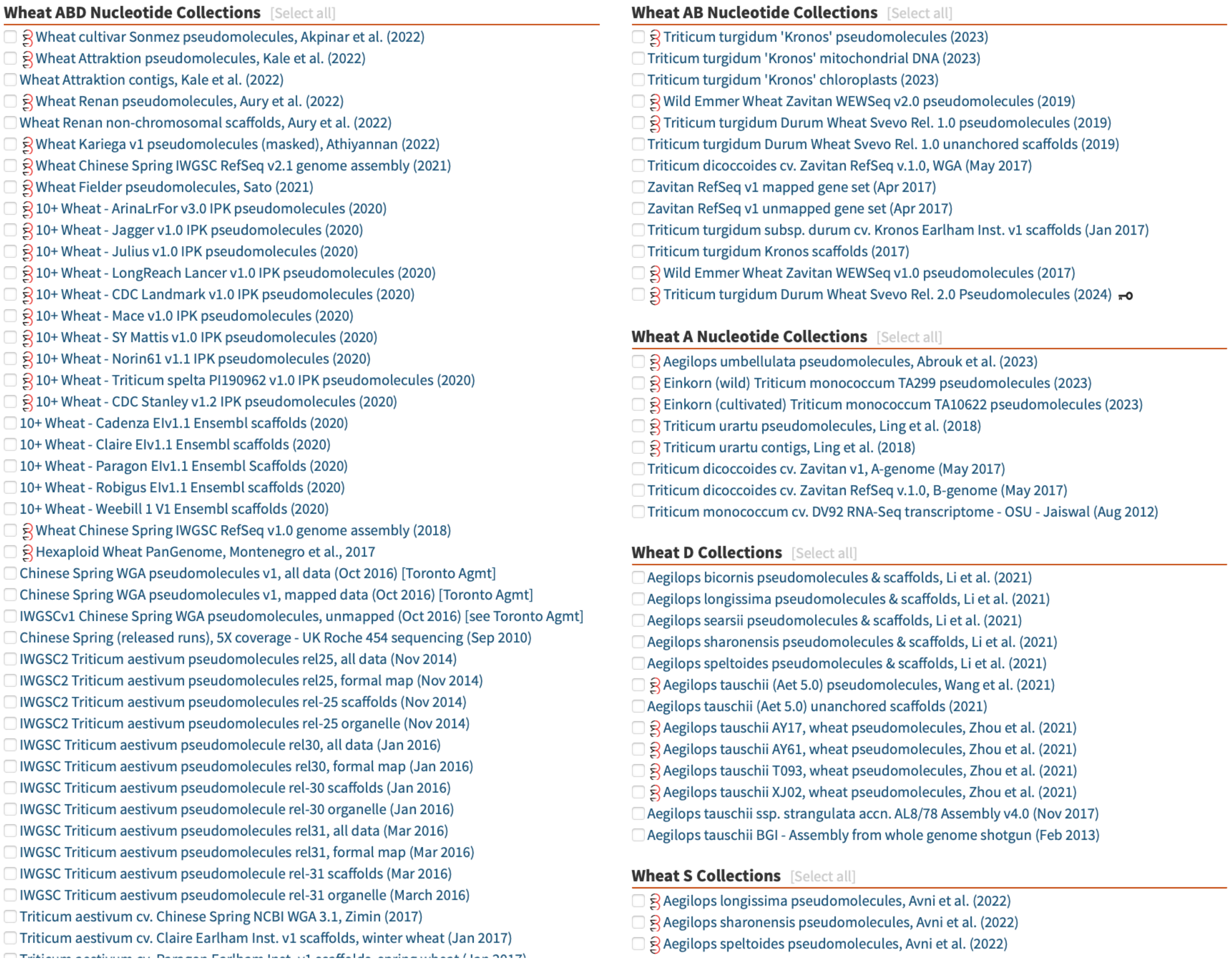


Genotypes (1051)

non-variant	727	69.17%
homozygous for reference	727	69.17%
variant	324	30.83%
homozygous	317	30.16%
T variant	317	30.16%
heterozygous	7	0.67%
Total	**1051**	**100%**

Name	GT	AD	DP	GQ	PL
Agron	ref (C) / ref (C)	13 0	13	36	0 36 540
Alba	ref (C) / ref (C)	10 0	10	30	0 30 394
Albatross	ref (C) / ref (C)	11 0	11	33	0 33 421
Alcedo	ref (C) / ref (C)	5 0	5	15	0 15 156
Alert	ref (C) /	11 0	11	33	0 33

BLAST – Nucleotide - wheat



Wheat ABD Nucleotide Collections [Select all]

- ☐ Wheat cultivar Sonmez pseudomolecules, Akpinar et al. (2022)
- ☐ Wheat Attraktion pseudomolecules, Kale et al. (2022)
- ☐ Wheat Attraktion contigs, Kale et al. (2022)
- ☐ Wheat Renan pseudomolecules, Aury et al. (2022)
- ☐ Wheat Renan non-chromosomal scaffolds, Aury et al. (2022)
- ☐ Wheat Kariega v1 pseudomolecules (masked), Athiyannan (2022)
- ☐ Wheat Chinese Spring IWGSC RefSeq v2.1 genome assembly (2021)
- ☐ Wheat Fielder pseudomolecules, Sato (2021)
- ☐ 10+ Wheat - ArinaLrFor v3.0 IPK pseudomolecules (2020)
- ☐ 10+ Wheat - Jagger v1.0 IPK pseudomolecules (2020)
- ☐ 10+ Wheat - Julius v1.0 IPK pseudomolecules (2020)
- ☐ 10+ Wheat - LongReach Lancer v1.0 IPK pseudomolecules (2020)
- ☐ 10+ Wheat - CDC Landmark v1.0 IPK pseudomolecules (2020)
- ☐ 10+ Wheat - Mace v1.0 IPK pseudomolecules (2020)
- ☐ 10+ Wheat - SY Mattis v1.0 IPK pseudomolecules (2020)
- ☐ 10+ Wheat - Norin61 v1.1 IPK pseudomolecules (2020)
- ☐ 10+ Wheat - Triticum spelta PI190962 v1.0 IPK pseudomolecules (2020)
- ☐ 10+ Wheat - CDC Stanley v1.2 IPK pseudomolecules (2020)
- ☐ 10+ Wheat - Cadenza EIv1.1 Ensembl scaffolds (2020)
- ☐ 10+ Wheat - Claire EIv1.1 Ensembl scaffolds (2020)
- ☐ 10+ Wheat - Paragon EIv1.1 Ensembl Scaffolds (2020)
- ☐ 10+ Wheat - Robigus EIv1.1 Ensembl scaffolds (2020)
- ☐ 10+ Wheat - Weebill 1 V1 Ensembl scaffolds (2020)
- ☐ Wheat Chinese Spring IWGSC RefSeq v1.0 genome assembly (2018)
- ☐ Hexaploid Wheat PanGenome, Montenegro et al., 2017
- ☐ Chinese Spring WGA pseudomolecules v1, all data (Oct 2016) [Toronto Agmt]
- ☐ Chinese Spring WGA pseudomolecules v1, mapped data (Oct 2016) [Toronto Agmt]
- ☐ IWGSCv1 Chinese Spring WGA pseudomolecules, unmapped (Oct 2016) [see Toronto Agmt]
- ☐ Chinese Spring (released runs), 5X coverage - UK Roche 454 sequencing (Sep 2010)
- ☐ IWGSC2 Triticum aestivum pseudomolecules rel25, all data (Nov 2014)
- ☐ IWGSC2 Triticum aestivum pseudomolecules rel25, formal map (Nov 2014)
- ☐ IWGSC2 Triticum aestivum pseudomolecules rel-25 scaffolds (Nov 2014)
- ☐ IWGSC2 Triticum aestivum pseudomolecules rel-25 organelle (Nov 2014)
- ☐ IWGSC Triticum aestivum pseudomolecule rel30, all data (Jan 2016)
- ☐ IWGSC Triticum aestivum pseudomolecules rel30, formal map (Jan 2016)
- ☐ IWGSC Triticum aestivum pseudomolecule rel-30 scaffolds (Jan 2016)
- ☐ IWGSC Triticum aestivum pseudomolecule rel-30 organelle (Jan 2016)
- ☐ IWGSC Triticum aestivum pseudomolecules rel31, all data (Mar 2016)
- ☐ IWGSC Triticum aestivum pseudomolecules rel31, formal map (Mar 2016)
- ☐ IWGSC Triticum aestivum pseudomolecule rel-31 scaffolds (Mar 2016)
- ☐ IWGSC Triticum aestivum pseudomolecule rel-31 organelle (March 2016)
- ☐ Triticum aestivum cv. Chinese Spring NCBI WGA 3.1, Zimin (2017)
- ☐ Triticum aestivum cv. Claire Earlham Inst. v1 scaffolds, winter wheat (Jan 2017)
- ☐ Triticum aestivum cv. Paragon Earlham Inst. v1 scaffolds, spring wheat (Jan 2017)

Wheat AB Nucleotide Collections [Select all]

- ☐ Triticum turgidum 'Kronos' pseudomolecules (2023)
- ☐ Triticum turgidum 'Kronos' mitochondrial DNA (2023)
- ☐ Triticum turgidum 'Kronos' chloroplasts (2023)
- ☐ Wild Emmer Wheat Zavitan WEWSeq v2.0 pseudomolecules (2019)
- ☐ Triticum turgidum Durum Wheat Svevo Rel. 1.0 pseudomolecules (2019)
- ☐ Triticum turgidum Durum Wheat Svevo Rel. 1.0 unanchored scaffolds (2019)
- ☐ Triticum dicoccoides cv. Zavitan RefSeq v.1.0, WGA (May 2017)
- ☐ Zavitan RefSeq v1 mapped gene set (Apr 2017)
- ☐ Zavitan RefSeq v1 unmapped gene set (Apr 2017)
- ☐ Triticum turgidum subsp. durum cv. Kronos Earlham Inst. v1 scaffolds (Jan 2017)
- ☐ Triticum turgidum Kronos scaffolds (2017)
- ☐ Wild Emmer Wheat Zavitan WEWSeq v1.0 pseudomolecules (2017)
- ☐ Triticum turgidum Durum Wheat Svevo Rel. 2.0 Pseudomolecules (2024)

Wheat A Nucleotide Collections [Select all]

- ☐ Aegilops umbellulata pseudomolecules, Abrouk et al. (2023)
- ☐ Einkorn (wild) Triticum monococcum TA299 pseudomolecules (2023)
- ☐ Einkorn (cultivated) Triticum monococcum TA10622 pseudomolecules (2023)
- ☐ Triticum urartu pseudomolecules, Ling et al. (2018)
- ☐ Triticum urartu contigs, Ling et al. (2018)
- ☐ Triticum dicoccoides cv. Zavitan v1, A-genome (May 2017)
- ☐ Triticum dicoccoides cv. Zavitan RefSeq v.1.0, B-genome (May 2017)
- ☐ Triticum monococcum cv. DV92 RNA-Seq transcriptome - OSU - Jaiswal (Aug 2012)

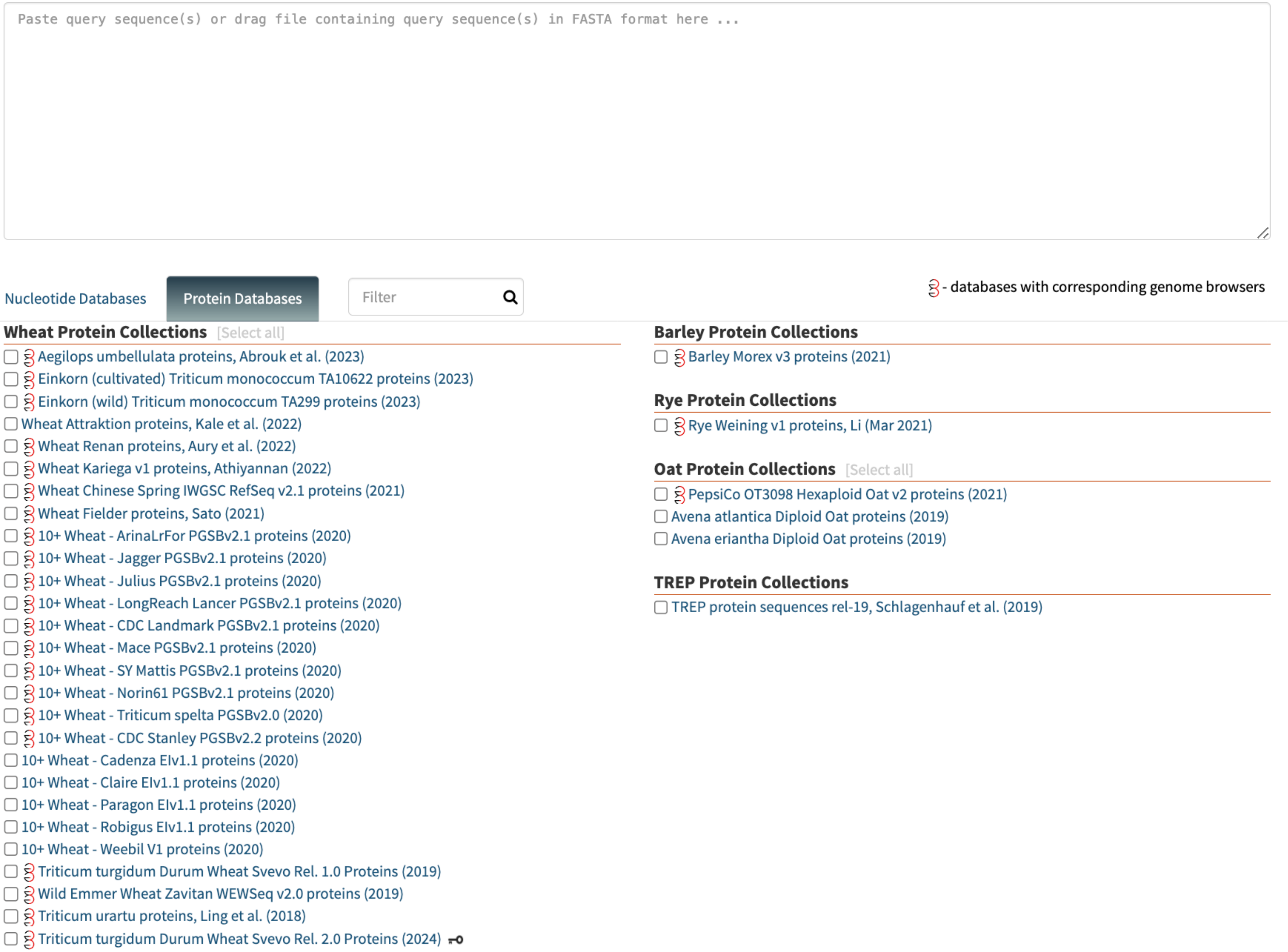
Wheat D Collections [Select all]

- ☐ Aegilops bicornis pseudomolecules & scaffolds, Li et al. (2021)
- ☐ Aegilops longissima pseudomolecules & scaffolds, Li et al. (2021)
- ☐ Aegilops searsii pseudomolecules & scaffolds, Li et al. (2021)
- ☐ Aegilops sharonensis pseudomolecules & scaffolds, Li et al. (2021)
- ☐ Aegilops speltoides pseudomolecules & scaffolds, Li et al. (2021)
- ☐ Aegilops tauschii (Aet 5.0) pseudomolecules, Wang et al. (2021)
- ☐ Aegilops tauschii (Aet 5.0) unanchored scaffolds (2021)
- ☐ Aegilops tauschii AY17, wheat pseudomolecules, Zhou et al. (2021)
- ☐ Aegilops tauschii AY61, wheat pseudomolecules, Zhou et al. (2021)
- ☐ Aegilops tauschii T093, wheat pseudomolecules, Zhou et al. (2021)
- ☐ Aegilops tauschii XJ02, wheat pseudomolecules, Zhou et al. (2021)
- ☐ Aegilops tauschii ssp. strangulata accn. AL8/78 Assembly v4.0 (Nov 2017)
- ☐ Aegilops tauschii BGI - Assembly from whole genome shotgun (Feb 2013)

Wheat S Collections [Select all]

- ☐ Aegilops longissima pseudomolecules, Avni et al. (2022)
- ☐ Aegilops sharonensis pseudomolecules, Avni et al. (2022)
- ☐ Aegilops speltoides pseudomolecules, Avni et al. (2022)

BLAST - Protein



🧬 - databases with corresponding genome browsers

Wheat Protein Collections [Select all]

- ☐ 🧬 Aegilops umbellulata proteins, Abrouk et al. (2023)
- ☐ 🧬 Einkorn (cultivated) Triticum monococcum TA10622 proteins (2023)
- ☐ 🧬 Einkorn (wild) Triticum monococcum TA299 proteins (2023)
- ☐ Wheat Attraktion proteins, Kale et al. (2022)
- ☐ 🧬 Wheat Renan proteins, Aury et al. (2022)
- ☐ 🧬 Wheat Kariega v1 proteins, Athiyannan (2022)
- ☐ 🧬 Wheat Chinese Spring IWGSC RefSeq v2.1 proteins (2021)
- ☐ 🧬 Wheat Fielder proteins, Sato (2021)
- ☐ 🧬 10+ Wheat - ArinaLrFor PGSBv2.1 proteins (2020)
- ☐ 🧬 10+ Wheat - Jagger PGSBv2.1 proteins (2020)
- ☐ 🧬 10+ Wheat - Julius PGSBv2.1 proteins (2020)
- ☐ 🧬 10+ Wheat - LongReach Lancer PGSBv2.1 proteins (2020)
- ☐ 🧬 10+ Wheat - CDC Landmark PGSBv2.1 proteins (2020)
- ☐ 🧬 10+ Wheat - Mace PGSBv2.1 proteins (2020)
- ☐ 🧬 10+ Wheat - SY Mattis PGSBv2.1 proteins (2020)
- ☐ 🧬 10+ Wheat - Norin61 PGSBv2.1 proteins (2020)
- ☐ 🧬 10+ Wheat - Triticum spelta PGSBv2.0 (2020)
- ☐ 🧬 10+ Wheat - CDC Stanley PGSBv2.2 proteins (2020)
- ☐ 10+ Wheat - Cadenza EIv1.1 proteins (2020)
- ☐ 10+ Wheat - Claire EIv1.1 proteins (2020)
- ☐ 10+ Wheat - Paragon EIv1.1 proteins (2020)
- ☐ 10+ Wheat - Robigus EIv1.1 proteins (2020)
- ☐ 10+ Wheat - Weebil V1 proteins (2020)
- ☐ 🧬 Triticum turgidum Durum Wheat Svevo Rel. 1.0 Proteins (2019)
- ☐ 🧬 Wild Emmer Wheat Zavitan WEWSeq v2.0 proteins (2019)
- ☐ 🧬 Triticum urartu proteins, Ling et al. (2018)
- ☐ 🧬 Triticum turgidum Durum Wheat Svevo Rel. 2.0 Proteins (2024) 🔑

Barley Protein Collections

- ☐ 🧬 Barley Morex v3 proteins (2021)

Rye Protein Collections

- ☐ 🧬 Rye Weining v1 proteins, Li (Mar 2021)

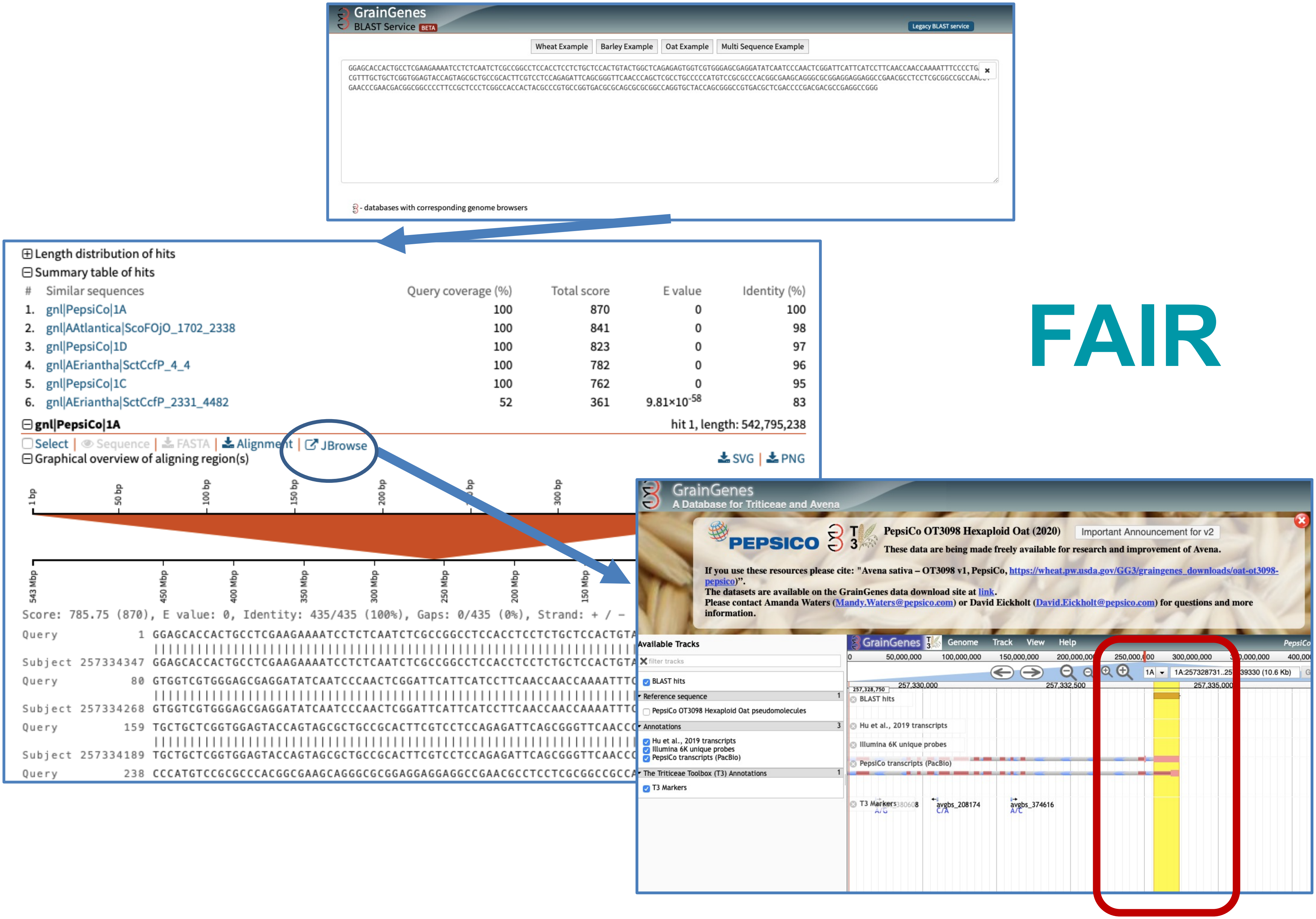
Oat Protein Collections [Select all]

- ☐ 🧬 PepsiCo OT3098 Hexaploid Oat v2 proteins (2021)
- ☐ Avena atlantica Diploid Oat proteins (2019)
- ☐ Avena eriantha Diploid Oat proteins (2019)

TREP Protein Collections

- ☐ TREP protein sequences rel-19, Schlagenhauf et al. (2019)

BLAST – Linkages to Browser



GrainGenes
BLAST Service BETA
Legacy BLAST service

Wheat Example | Barley Example | Oat Example | Multi Sequence Example

```
GGAGCACCACTGCCTCGAAGAAAATCCTCTCAATCTCGCCGGCCTCCACCTCCTCTGCTCCACTGTACTGGCTCAGAGAGTGGTCGTGGGAGCGAGGATATCAATCCCAACTCGGATTCATTCATCCTTCAACCAACCAAAATTTCCCCTG
CGTTTGCTGCTCGGTGGAGTACCAGTAGCGCTGCCGCACTTCGTCCTCCAGAGATTCAGCGGGTTCAACCCAGCTCGCCTGCCCCCATGTCCGCGCCCACGGCGAAGCAGGGCGCGGAGGAGGAGGCCGAACGCCTCCTCGCGGCCGCCAA
GAACCCGAACGACGGCGGCCCCTTCCGCTCCCTCGGCCACCACTACGCCCGTGCCGGTGACGCGCAGCGCGCGGCCAGGTGCTACCAGCGGGCCGTGACGCTCGACCCCGACGACGCCGAGGCCGGG
```

- databases with corresponding genome browsers

⊞ Length distribution of hits

⊟ Summary table of hits

#	Similar sequences	Query coverage (%)	Total score	E value	Identity (%)
1.	gnl|PepsiCo|1A	100	870	0	100
2.	gnl|AAtlantica|ScoFOjO_1702_2338	100	841	0	98
3.	gnl|PepsiCo|1D	100	823	0	97
4.	gnl|AEriantha|SctCcfP_4_4	100	782	0	96
5.	gnl|PepsiCo|1C	100	762	0	95
6.	gnl|AEriantha|SctCcfP_2331_4482	52	361	9.81×10^{-58}	83

⊟ **gnl|PepsiCo|1A** hit 1, length: 542,795,238

☐ Select | Sequence | FASTA | Alignment | JBrowse

⊟ Graphical overview of aligning region(s) SVG | PNG

```
Score: 785.75 (870), E value: 0, Identity: 435/435 (100%), Gaps: 0/435 (0%), Strand: + / -
Query          1 GGAGCACCACTGCCTCGAAGAAAATCCTCTCAATCTCGCCGGCCTCCACCTCCTCTGCTCCACTGTA
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Subject 257334347 GGAGCACCACTGCCTCGAAGAAAATCCTCTCAATCTCGCCGGCCTCCACCTCCTCTGCTCCACTGTA
Query         80 GTGGTCGTGGGAGCGAGGATATCAATCCCAACTCGGATTCATTCATCCTTCAACCAACCAAAATTTC
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Subject 257334268 GTGGTCGTGGGAGCGAGGATATCAATCCCAACTCGGATTCATTCATCCTTCAACCAACCAAAATTTC
Query        159 TGCTGCTCGGTGGAGTACCAGTAGCGCTGCCGCACTTCGTCCTCCAGAGATTCAGCGGGTTCAACCC
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Subject 257334189 TGCTGCTCGGTGGAGTACCAGTAGCGCTGCCGCACTTCGTCCTCCAGAGATTCAGCGGGTTCAACCC
Query        238 CCCATGTCCGCGCCCACGGCGAAGCAGGGCGCGGAGGAGGAGGCCGAACGCCTCCTCGCGGCCGCCA
```

GrainGenes
A Database for Triticeae and Avena

PepsiCo OT3098 Hexaploid Oat (2020) Important Announcement for v2

These data are being made freely available for research and improvement of Avena.

If you use these resources please cite: "Avena sativa – OT3098 v1, PepsiCo, https://wheat.pw.usda.gov/GG3/graingenes_downloads/oat-ot3098-pepsico)".
The datasets are available on the GrainGenes data download site at link.
Please contact Amanda Waters (Mandy.Waters@pepsico.com) or David Eickholt (David.Eickholt@pepsico.com) for questions and more information.

Available Tracks
filter tracks

☑ BLAST hits

Reference sequence 1

☐ PepsiCo OT3098 Hexaploid Oat pseudomolecules

Annotations 3

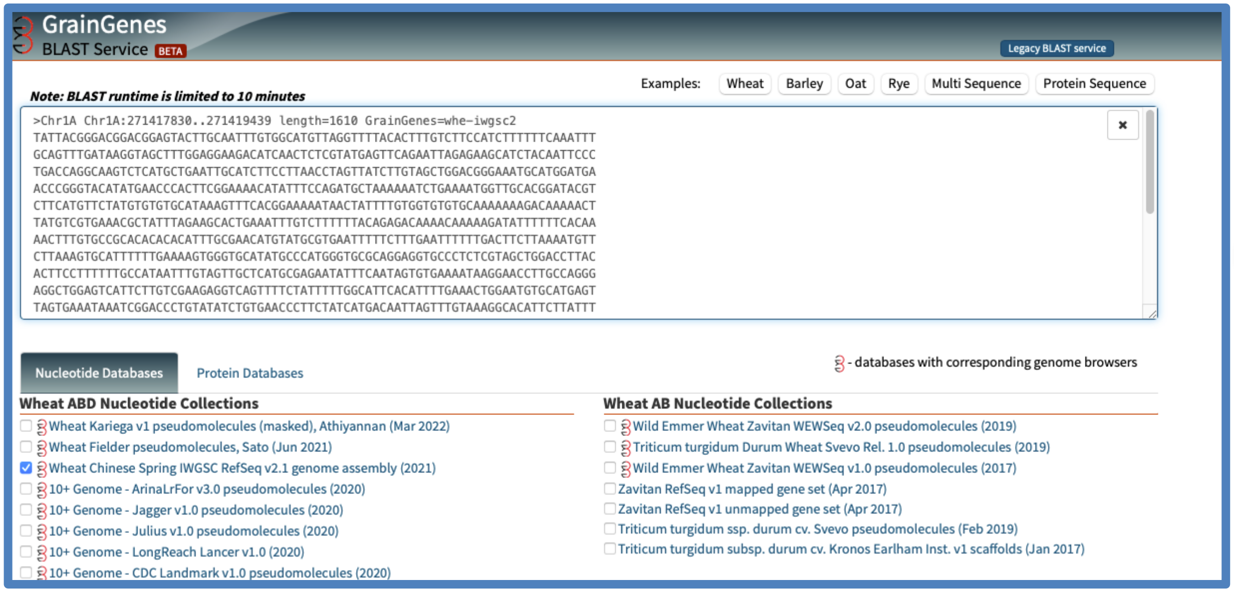
☑ Hu et al., 2019 transcripts
☑ Illumina 6K unique probes
☑ PepsiCo transcripts (PacBio)

The Triticeae Toolbox (T3) Annotations 1

☑ T3 Markers

FAIR

Genome Browser to BLAST



Note: BLAST runtime is limited to 10 minutes

```
>Chr1A Chr1A:271417830..271419439 length=1610 GrainGenes=whe-iwgsc2
TATTACGGGACGGACGGAGTACTTGCAATTTGTGGCATGTTAGGTTTTACACTTTGTCTTCCATCTTTTTTCAAATTT
GCAGTTTGATAAGGTAGCTTTGGAGGAAGACATCAACTCTCGTATGAGTTCAGAATTAGAGAAGCATCTACAATTCCC
TGACCAGGCAAGTCTCATGCTGAATTGCATCTTCCTTAACCTAGTTATCTTGTAGCTGGACGGGAAATGCATGGATGA
ACCCGGGTACATATGAACCCACTTCGGAAAACATATTTCCAGATGCTAAAAAATCTGAAAATGGTTGCACGGATACGT
CTTCATGTTCTATGTGTGTGCATAAAGTTTCACGGAAAAATAACTATTTTGTGGTGTGTGCAAAAAAAGACAAAAACT
TATGTCGTGAAACGCTATTTAGAAGCACTGAAATTTGTCTTTTTTACAGAGACAAAACAAAAAGATATTTTTTCACAA
AACTTTGTGCCGCACACACACATTTGCGAACATGTATGCGTGAATTTTTCTTTGAATTTTTTGACTTCTTAAAATGTT
CTTAAAGTGCATTTTTTTGAAAAGTGGGTGCATATGCCCATGGGTGCGCAGGAGGTGCCCTCTCGTAGCTGGACCTTAC
ACTTCCTTTTTTGCCATAATTTGTAGTTGCTCATGCGAGAATATTTCAATAGTGTGAAAATAAGGAACCTTGCCAGGG
AGGCTGGAGTCATTCTTGTCGAAGAGGTCAGTTTTCTATTTTTGGCATTCACATTTTGAACTGGAATGTGCATGAGT
TAGTGAAATAAATCGGACCCTGTATATCTGTGAACCCTTCTATCATGACAATTAGTTTGTAAAGGCACATTCTTATTT
```

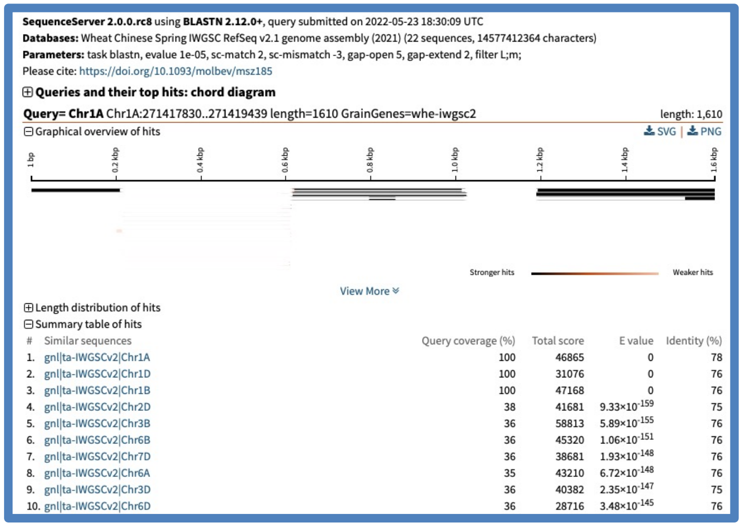
- databases with corresponding genome browsers

Wheat ABD Nucleotide Collections

☐ Wheat Kariega v1 pseudomolecules (masked), Athiyannan (Mar 2022)
☐ Wheat Fielder pseudomolecules, Sato (Jun 2021)
☑ Wheat Chinese Spring IWGSC RefSeq v2.1 genome assembly (2021)
☐ 10+ Genome - ArinaLrFor v3.0 pseudomolecules (2020)
☐ 10+ Genome - Jagger v1.0 pseudomolecules (2020)
☐ 10+ Genome - Julius v1.0 pseudomolecules (2020)
☐ 10+ Genome - LongReach Lancer v1.0 (2020)
☐ 10+ Genome - CDC Landmark v1.0 pseudomolecules (2020)

Wheat AB Nucleotide Collections

☐ Wild Emmer Wheat Zavitan WEWSeq v2.0 pseudomolecules (2019)
☐ Triticum turgidum Durum Wheat Svevo Rel. 1.0 pseudomolecules (2019)
☐ Wild Emmer Wheat Zavitan WEWSeq v1.0 pseudomolecules (2017)
☐ Zavitan RefSeq v1 mapped gene set (Apr 2017)
☐ Zavitan RefSeq v1 unmapped gene set (Apr 2017)
☐ Triticum turgidum ssp. durum cv. Svevo pseudomolecules (Feb 2019)
☐ Triticum turgidum subsp. durum cv. Kronos Earlham Inst. v1 scaffolds (Jan 2017)



SequenceServer 2.0.0.rc8 using **BLASTN 2.12.0+**, query submitted on 2022-05-23 18:30:09 UTC

Databases: Wheat Chinese Spring IWGSC RefSeq v2.1 genome assembly (2021) (22 sequences, 14577412364 characters)

Parameters: task blastn, evalue 1e-05, sc-match 2, sc-mismatch -3, gap-open 5, gap-extend 2, filter L;m;

Please cite: https://doi.org/10.1093/molbev/msz185

⊞ **Queries and their top hits: chord diagram**

Query= Chr1A Chr1A:271417830..271419439 length=1610 GrainGenes=whe-iwgsc2 length: 1,610

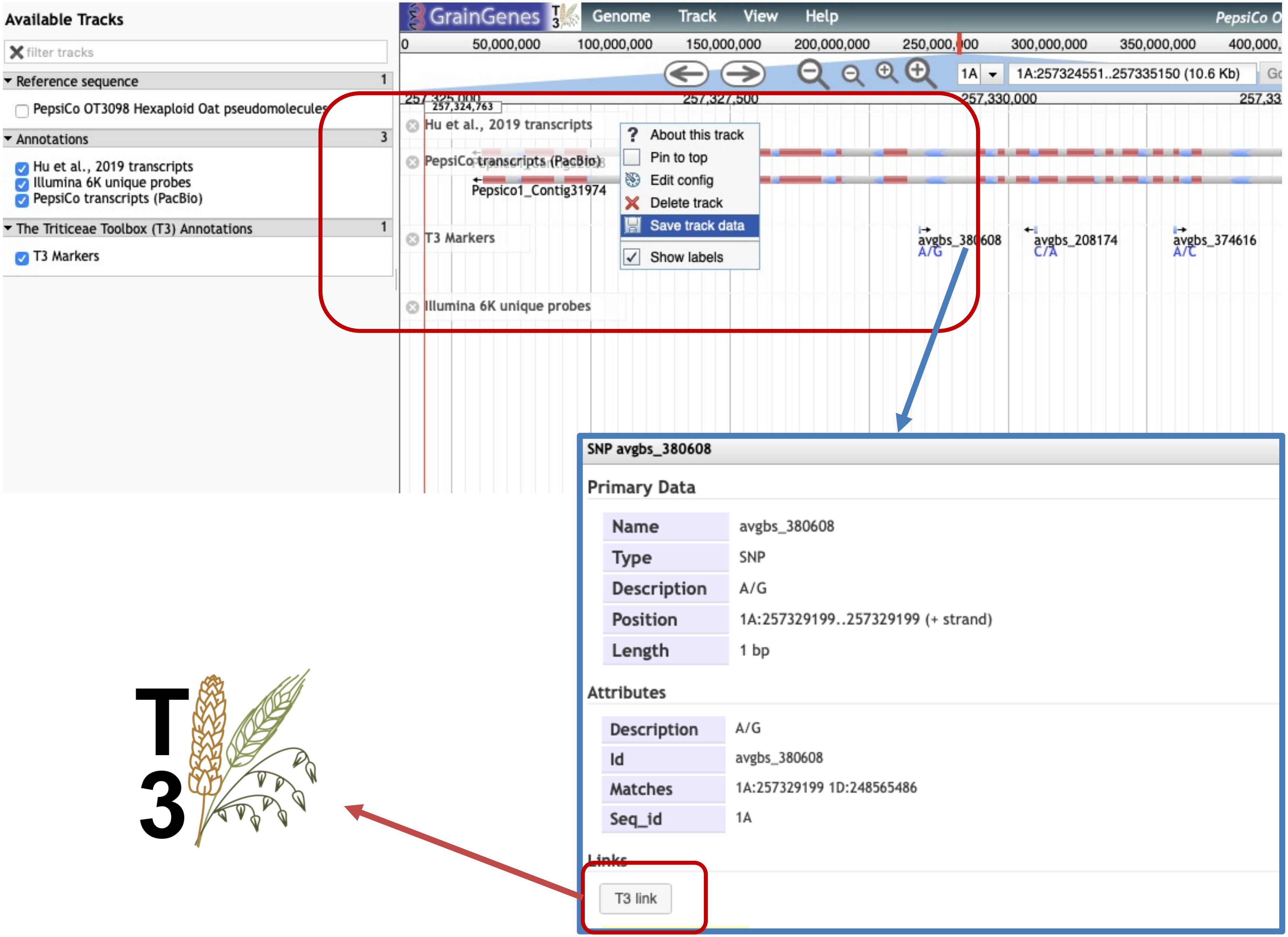
⊟ Graphical overview of hits

⊞ Length distribution of hits

⊟ Summary table of hits

#	Similar sequences	Query coverage (%)	Total score	E value	Identity (%)
1.	gnl|ta-IWGSCv2|Chr1A	100	46865	0	78
2.	gnl|ta-IWGSCv2|Chr1D	100	31076	0	76
3.	gnl|ta-IWGSCv2|Chr1B	100	47168	0	76
4.	gnl|ta-IWGSCv2|Chr2D	38	41681	9.33×10^{-159}	75
5.	gnl|ta-IWGSCv2|Chr3B	36	58813	5.89×10^{-155}	76
6.	gnl|ta-IWGSCv2|Chr6B	36	45320	1.06×10^{-151}	76
7.	gnl|ta-IWGSCv2|Chr7D	36	38681	1.93×10^{-148}	76
8.	gnl|ta-IWGSCv2|Chr6A	35	43210	6.72×10^{-148}	76
9.	gnl|ta-IWGSCv2|Chr3D	36	40382	2.35×10^{-147}	75
10.	gnl|ta-IWGSCv2|Chr6D	36	28716	3.48×10^{-145}	76

GrainGenes – Links to other Resources



SNP avgbs_380608

Primary Data

Name	avgbs_380608
Type	SNP
Description	A/G
Position	1A:257329199..257329199 (+ strand)
Length	1 bp

Attributes

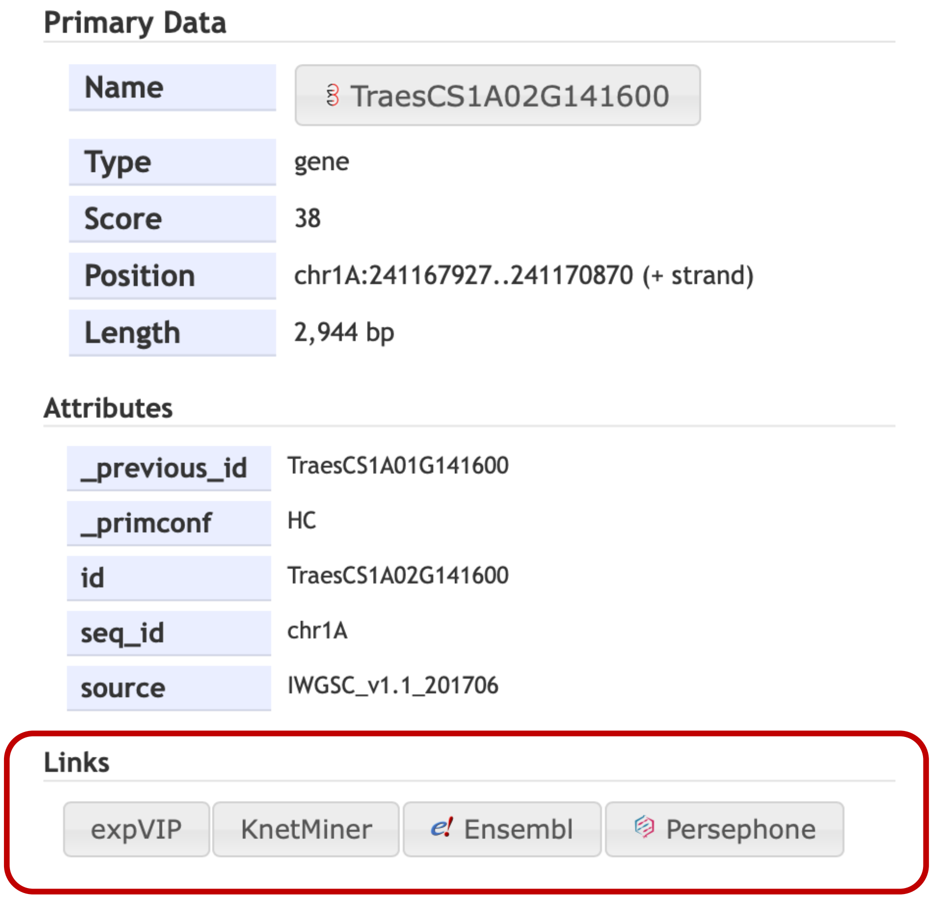
Description	A/G
Id	avgbs_380608
Matches	1A:257329199 1D:248565486
Seq_id	1A

Links

T3 link

GrainGenes – Links to other Resources

Wheat Chinese Spring IWGSC RefSeq v1.0 genome assembly (2018)



Primary Data

Name	TraesCS1A02G141600
Type	gene
Score	38
Position	chr1A:241167927..241170870 (+ strand)
Length	2,944 bp

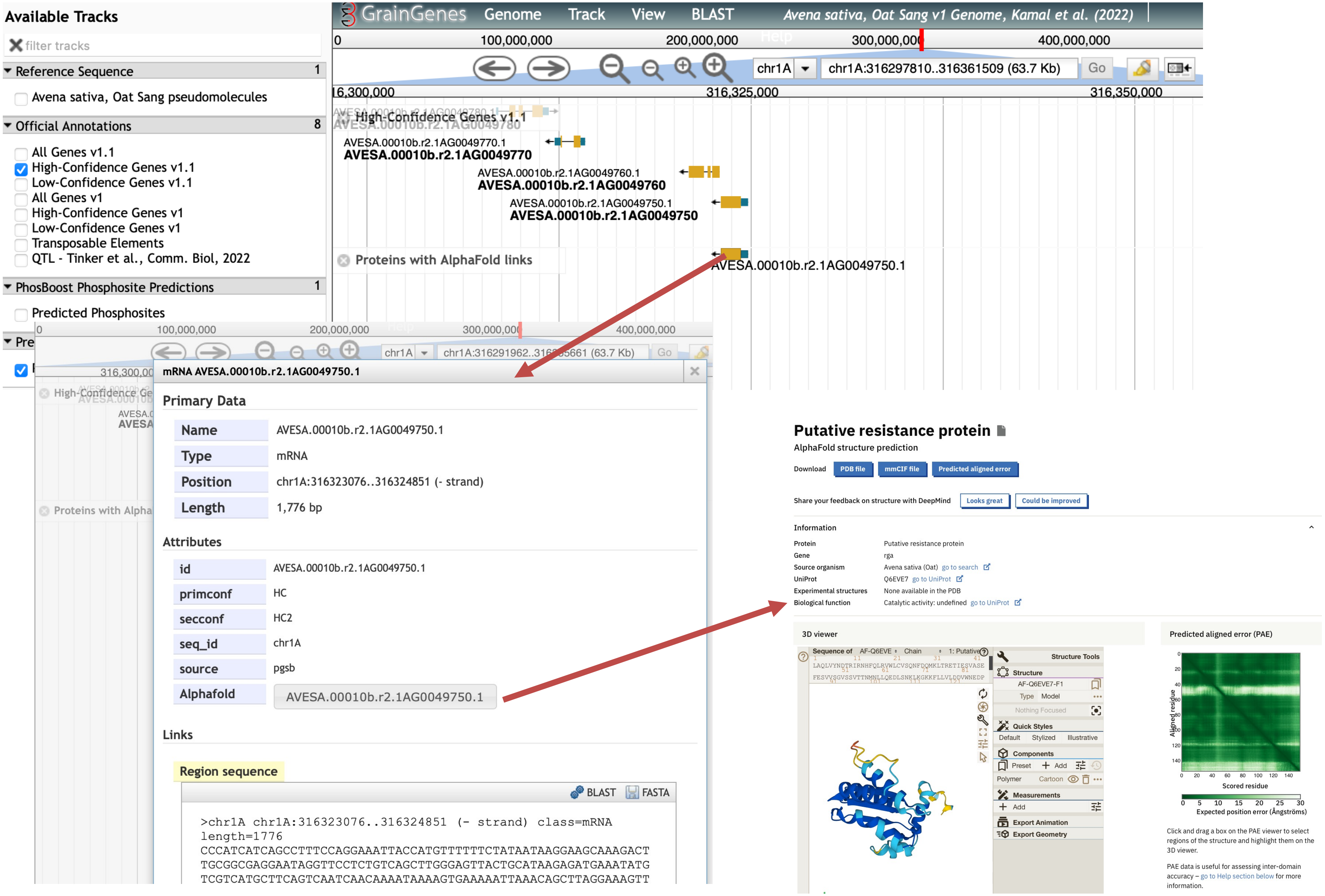
Attributes

_previous_id	TraesCS1A01G141600
_primconf	HC
id	TraesCS1A02G141600
seq_id	chr1A
source	IWGSC_v1.1_201706

Links

Genome Browser – AlphaFold/Protein 3D structure links

(Chinese Spring IWGSC v1 | Durum Wheat | Aegilops v4 | Triticum urartu | Barley Morex v3 | Oat Sang)



mRNA AVESA.00010b.r2.1AG0049750.1

Primary Data

Name	AVESA.00010b.r2.1AG0049750.1
Type	mRNA
Position	chr1A:316323076..316324851 (- strand)
Length	1,776 bp

Attributes

id	AVESA.00010b.r2.1AG0049750.1
primconf	HC
secconf	HC2
seq_id	chr1A
source	pgsb
Alphafold	AVESA.00010b.r2.1AG0049750.1

Links

Region sequence

BLAST FASTA

```
>chr1A chr1A:316323076..316324851 (- strand) class=mRNA
length=1776
CCCATCATCAGCCTTTCCAGGAAATTACCATGTTTTTTCTATAATAAGGAAGCAAAGACT
TGCGGCGAGGAATAGGTTCCTCTGTCAGCTTGGGAGTTACTGCATAAGAGATGAAATATG
TCGTCATGCTTCAGTCAATCAACAAAATAAAAGTGAAAAATTAAACAGCTTAGGAAAGTT
```

Putative resistance protein

AlphaFold structure prediction

Download PDB file mmCIF file Predicted aligned error

Share your feedback on structure with DeepMind Looks great Could be improved

Information

Protein	Putative resistance protein
Gene	rga
Source organism	Avena sativa (Oat) go to search
UniProt	Q6EVE7 go to UniProt
Experimental structures	None available in the PDB
Biological function	Catalytic activity: undefined go to UniProt

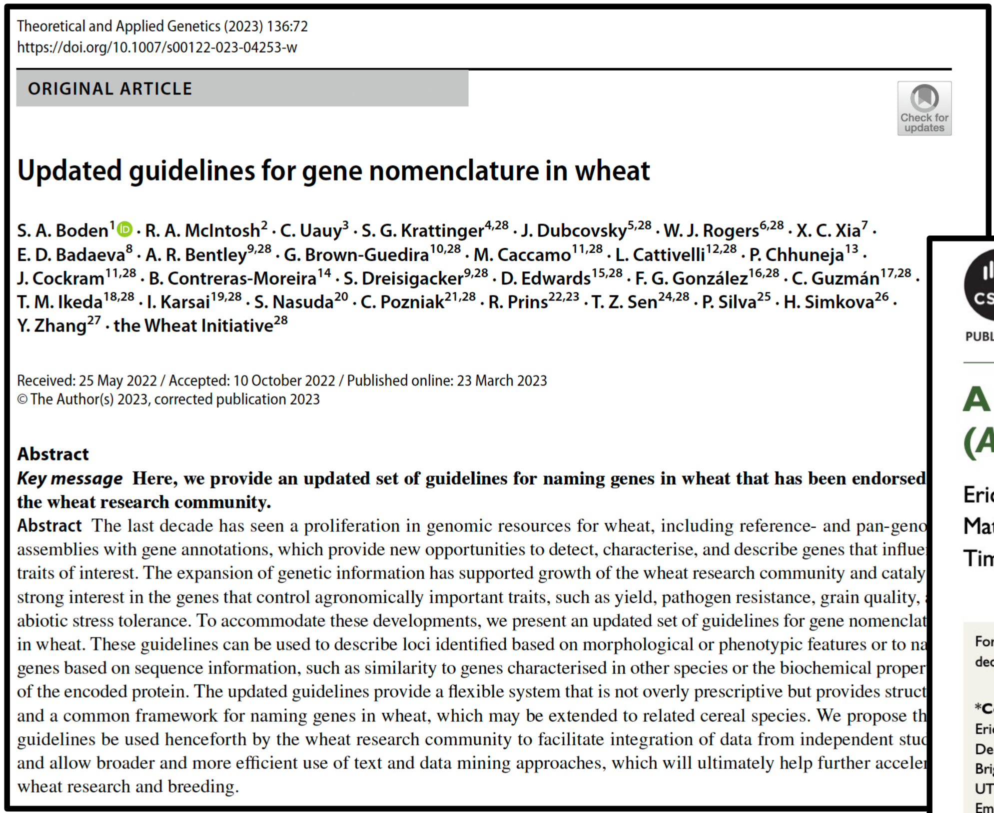
3D viewer

```
Sequence of  AF-Q6EVE  Chain  1: Putative
1         11        21        31        41
LAQLVYNDTRIRNHFQLRVWLCVSQNFDQMKLTRETIESVASE
51        61        71        81
FESVVSGVSSVTTNMNLLQEDLSNKLKGKKFLLVLDDVWNEDP
91        101       111       121
```

Click and drag a box on the PAE viewer to select regions of the structure and highlight them on the 3D viewer.

PAE data is useful for assessing inter-domain accuracy – go to Help section below for more information.

Nomenclature & Data Standards



Theoretical and Applied Genetics (2023) 136:72
https://doi.org/10.1007/s00122-023-04253-w

ORIGINAL ARTICLE

Check for updates

Updated guidelines for gene nomenclature in wheat

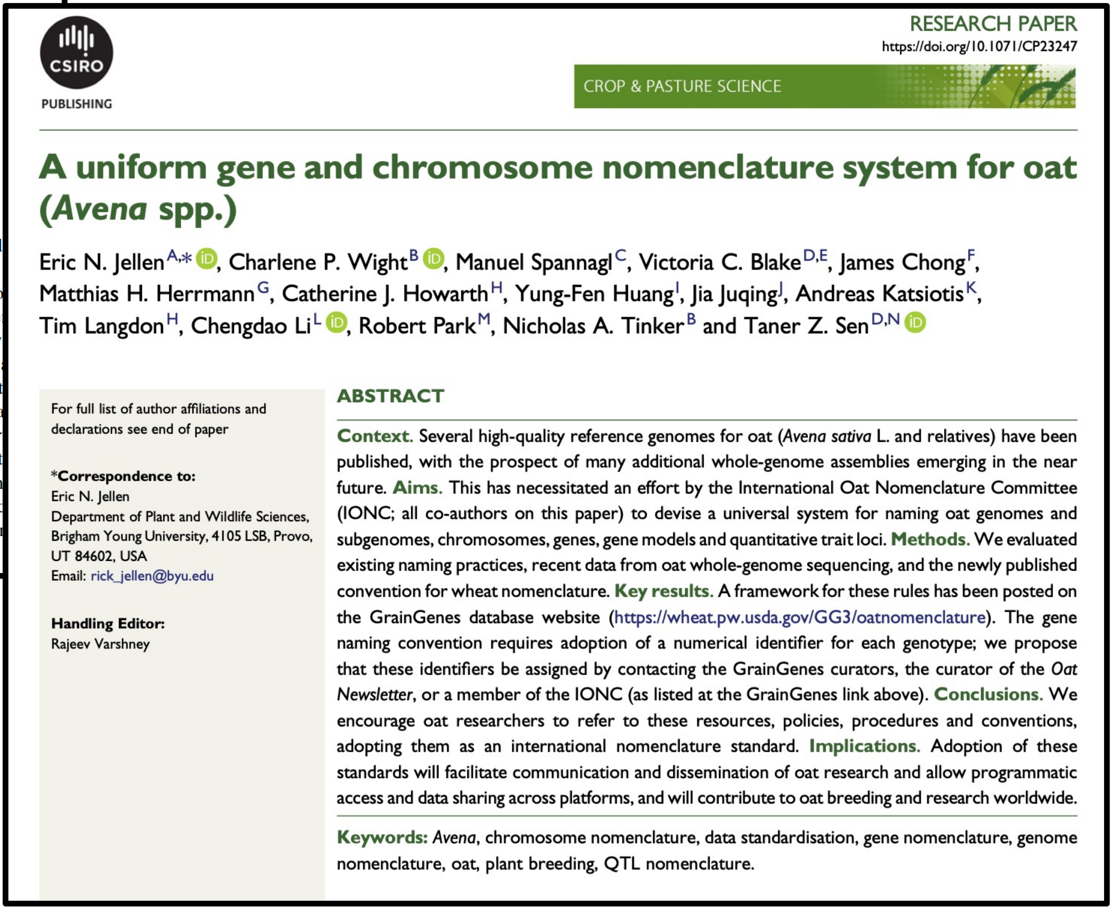
S. A. Boden[1] · R. A. McIntosh[2] · C. Uauy[3] · S. G. Krattinger[4,28] · J. Dubcovsky[5,28] · W. J. Rogers[6,28] · X. C. Xia[7] · E. D. Badaeva[8] · A. R. Bentley[9,28] · G. Brown-Guedira[10,28] · M. Caccamo[11,28] · L. Cattivelli[12,28] · P. Chhuneja[13] · J. Cockram[11,28] · B. Contreras-Moreira[14] · S. Dreisigacker[9,28] · D. Edwards[15,28] · F. G. González[16,28] · C. Guzmán[17,28] · T. M. Ikeda[18,28] · I. Karsai[19,28] · S. Nasuda[20] · C. Pozniak[21,28] · R. Prins[22,23] · T. Z. Sen[24,28] · P. Silva[25] · H. Simkova[26] · Y. Zhang[27] · the Wheat Initiative[28]

Received: 25 May 2022 / Accepted: 10 October 2022 / Published online: 23 March 2023
© The Author(s) 2023, corrected publication 2023

Abstract

Key message **Here, we provide an updated set of guidelines for naming genes in wheat that has been endorsed the wheat research community.**

Abstract The last decade has seen a proliferation in genomic resources for wheat, including reference- and pan-geno assemblies with gene annotations, which provide new opportunities to detect, characterise, and describe genes that influe traits of interest. The expansion of genetic information has supported growth of the wheat research community and cataly strong interest in the genes that control agronomically important traits, such as yield, pathogen resistance, grain quality, abiotic stress tolerance. To accommodate these developments, we present an updated set of guidelines for gene nomenclat in wheat. These guidelines can be used to describe loci identified based on morphological or phenotypic features or to na genes based on sequence information, such as similarity to genes characterised in other species or the biochemical proper of the encoded protein. The updated guidelines provide a flexible system that is not overly prescriptive but provides struct and a common framework for naming genes in wheat, which may be extended to related cereal species. We propose th guidelines be used henceforth by the wheat research community to facilitate integration of data from independent stuc and allow broader and more efficient use of text and data mining approaches, which will ultimately help further acceler wheat research and breeding.



CSIRO PUBLISHING

RESEARCH PAPER
https://doi.org/10.1071/CP23247

CROP & PASTURE SCIENCE

A uniform gene and chromosome nomenclature system for oat (*Avena* spp.)

Eric N. Jellen[A,*], Charlene P. Wight[B], Manuel Spannagl[C], Victoria C. Blake[D,E], James Chong[F], Matthias H. Herrmann[G], Catherine J. Howarth[H], Yung-Fen Huang[I], Jia Juqing[J], Andreas Katsiotis[K], Tim Langdon[H], Chengdao Li[L], Robert Park[M], Nicholas A. Tinker[B] and Taner Z. Sen[D,N]

For full list of author affiliations and declarations see end of paper

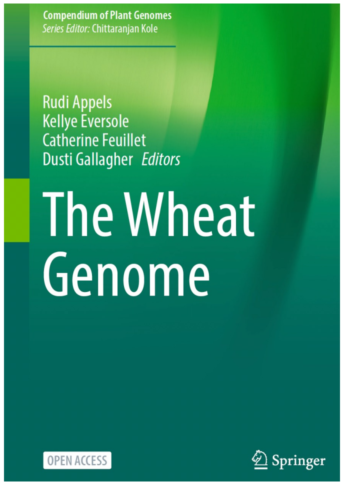
***Correspondence to:**
Eric N. Jellen
Department of Plant and Wildlife Sciences, Brigham Young University, 4105 LSB, Provo, UT 84602, USA
Email: rick_jellen@byu.edu

Handling Editor:
Rajeev Varshney

ABSTRACT

Context. Several high-quality reference genomes for oat (*Avena sativa* L. and relatives) have been published, with the prospect of many additional whole-genome assemblies emerging in the near future. **Aims.** This has necessitated an effort by the International Oat Nomenclature Committee (IONC; all co-authors on this paper) to devise a universal system for naming oat genomes and subgenomes, chromosomes, genes, gene models and quantitative trait loci. **Methods.** We evaluated existing naming practices, recent data from oat whole-genome sequencing, and the newly published convention for wheat nomenclature. **Key results.** A framework for these rules has been posted on the GrainGenes database website (https://wheat.pw.usda.gov/GG3/oatnomenclature). The gene naming convention requires adoption of a numerical identifier for each genotype; we propose that these identifiers be assigned by contacting the GrainGenes curators, the curator of the *Oat Newsletter*, or a member of the IONC (as listed at the GrainGenes link above). **Conclusions.** We encourage oat researchers to refer to these resources, policies, procedures and conventions, adopting them as an international nomenclature standard. **Implications.** Adoption of these standards will facilitate communication and dissemination of oat research and allow programmatic access and data sharing across platforms, and will contribute to oat breeding and research worldwide.

Keywords: *Avena*, chromosome nomenclature, data standardisation, gene nomenclature, genome nomenclature, oat, plant breeding, QTL nomenclature.



Compendium of Plant Genomes
Series Editor: Chittaranjan Kole

Rudi Appels
Kellye Eversole
Catherine Feuillet
Dusti Gallagher *Editors*

The Wheat Genome

OPEN ACCESS

Springer

FAIR

Wheat Data Integration and FAIRification: IWGSC, GrainGenes, Ensembl and Other Data Repositories

2

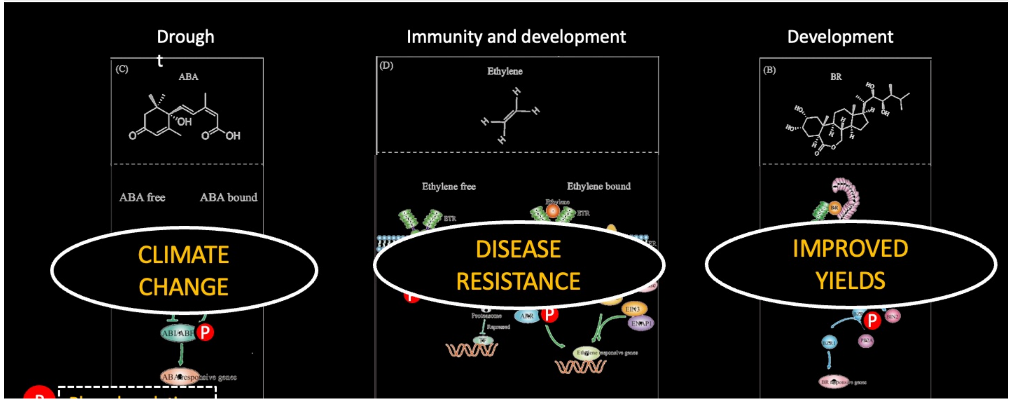
Michael Alaux, Sarah Dyer and Taner Z. Sen

- Please use the guidelines ("Chez Panisse")
- More work on protein naming (Cristobal Uauy taking the lead)
- Genotype abbreviations will be/are in GrainGenes

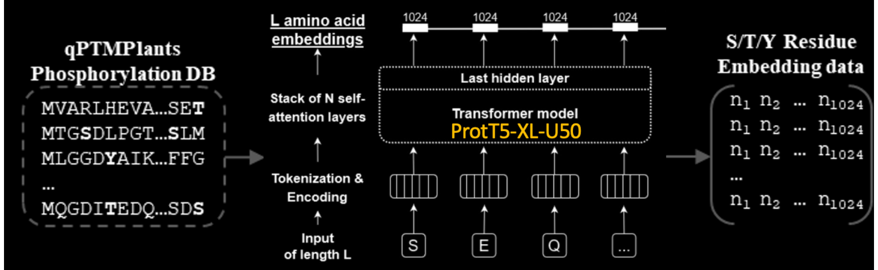
RESEARCH

Research: Phosphorylation Prediction

Protein phosphorylation plays an important role in all aspects of plant development, immunity and stress response



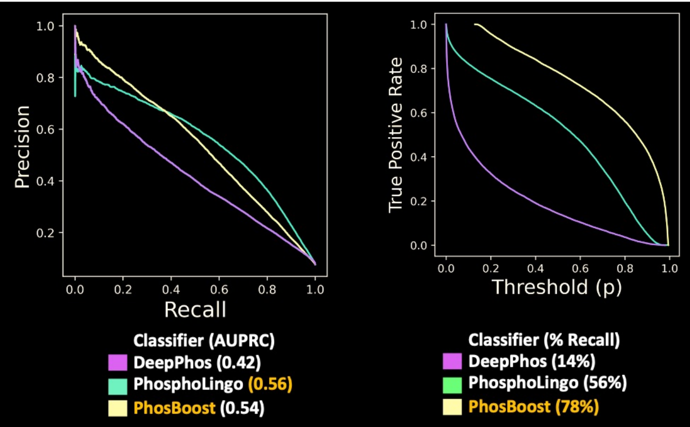
Using Protein Language Models (PLMs) to generate input data for a protein phosphorylation machine learning classifier



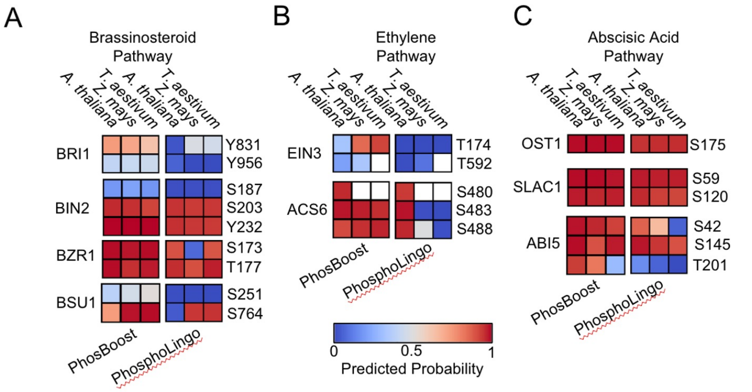
Protein Language Models (PLMs) capture functional and structural information for each amino-acid

qPTMplants: Xue (2022) Nucl. Acids Res.
ProtT5: Elnaggar *et al.* (2021) IEEE TPAMI

PhosBoost performs better than DeepPhos and is comparable to PhosphoLingo while providing higher recall



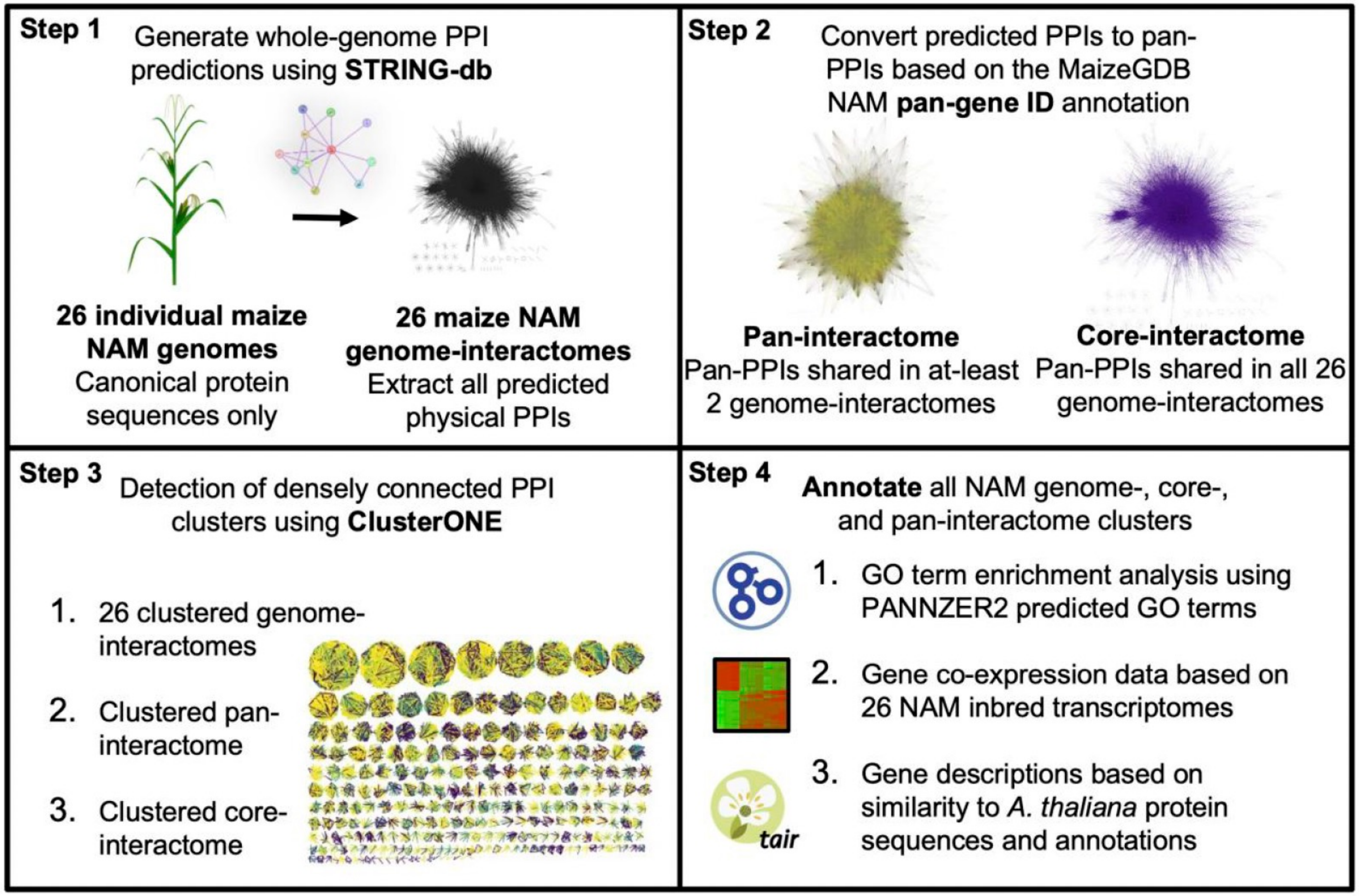
Higher recall comes at the tradeoff of reduced precision (more False Positives)



PhosBoost predicts more function phosphosites and their conserved putative functional phosphosites in distantly related species.

Poretsky, Andorf, Sen, Plant Direct, 2023

Research: Pan-Interactome Networks in Maize



Step 1 Generate whole-genome PPI predictions using **STRING-db**

26 individual maize NAM genomes
Canonical protein sequences only

26 maize NAM genome-interactomes
Extract all predicted physical PPIs

Step 2 Convert predicted PPIs to pan-PPIs based on the MaizeGDB NAM **pan-gene ID** annotation

Pan-interactome
Pan-PPIs shared in at-least 2 genome-interactomes

Core-interactome
Pan-PPIs shared in all 26 genome-interactomes

Step 3 Detection of densely connected PPI clusters using **ClusterONE**

1. 26 clustered genome-interactomes
2. Clustered pan-interactome
3. Clustered core-interactome

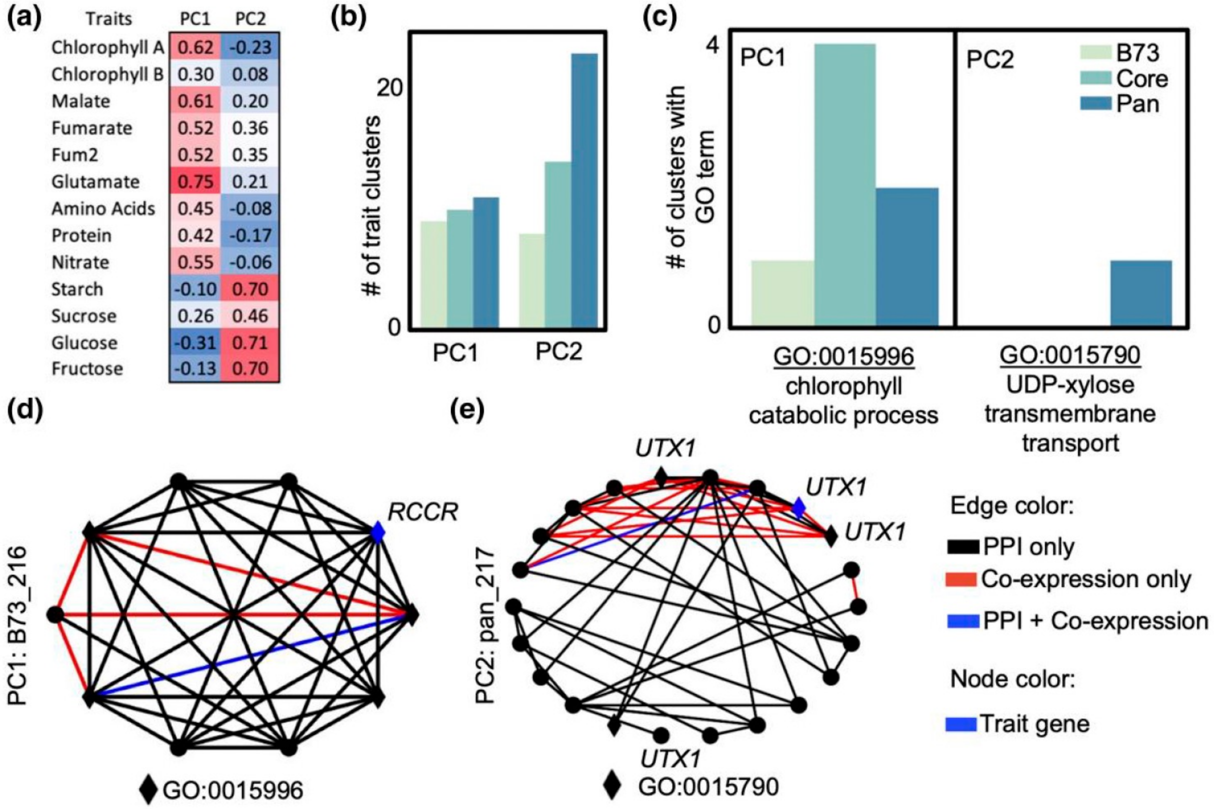
Step 4 **Annotate** all NAM genome-, core-, and pan-interactome clusters

1. GO term enrichment analysis using PANNZER2 predicted GO terms
2. Gene co-expression data based on 26 NAM inbred transcriptomes
3. Gene descriptions based on similarity to *A. thaliana* protein sequences and annotations

- Our method captures substantially more functionally enriched clusters with unique GO term annotations.

Poretsky, Cagirici, Andorf, Sen, G3, 2024

Research: Pan-Interactome Networks in Maize



Traits	PC1	PC2
Chlorophyll A	0.62	-0.23
Chlorophyll B	0.30	0.08
Malate	0.61	0.20
Fumarate	0.52	0.36
Fum2	0.52	0.35
Glutamate	0.75	0.21
Amino Acids	0.45	-0.08
Protein	0.42	-0.17
Nitrate	0.55	-0.06
Starch	-0.10	0.70
Sucrose	0.26	0.46
Glucose	-0.31	0.71
Fructose	-0.13	0.70

- Simple integration of supporting information such as gene coexpression and gene description annotations with the predicted interactomes, increases the breadth of genomic annotations that can be included in the pan-interactome analysis.

Using functional enrichment to annotate PPI clusters can be used for putative protein function prediction and prioritization of trait-specific candidate gene sets.

Poretsky, Cagirici, Andorf, Sen, G3, 2024

The GrainGenes Team

- **Taner Sen** – Lead Scientist
- **Gerard Lazo** – Scientist
- **Victoria Blake** – Curator
- **Steve Michel** – System and Interface Admin.
- **Eric Yao** – Browser & BLAST
- **Elly Poretsky** – Postdoc
- **Parva Sharma** - Postdoc

PARTNERS

(Some) Global Partnerships

- **Agriculture and Agri-Food Canada** (Nick Tinker and Charlene Wight) for oat data curation
- **PanOat Consortium**

- **Tetraploid Wheat Consortium**
- **Pepsico and Corteva**
- **ARS Databases**

The Triticeae Toolbox

MaizeGDB
Maize Genetics and Genomics Database

- **AgBioData**
- **BrAPI**

AgBioData
Toward enhanced genomics, genetics and breeding research outcomes through standardization of practices and protocols across agricultural databases

- **International collaborations** with Italy (durum wheat), UK, Germany, Turkey, Israel (wild emmer), among others; **Wheat Gene Catalogue** (Bob McIntosh – Australia; Japan, USA), **IWGSC, Ensembl, IOC for IWC**
- **International Wheat Information System** (global) – **URGI** GrainGenes QTL; germplasm; genetic map data sets are discoverable and searchable at **WheatIS**. **FAIDARE**

Take home message - 1

- We have extraordinary global challenges when it comes to agriculture, shaped by climate change

- We need to put more efforts into research to ensure a sustainable and nutritious agriculture that provides foods to billions

- We need to make sure data is FAIR and linked

- GrainGenes strives to serve the small grains research communities through curated and linked data, tools, community building efforts, and research

Take home message - 2

GrainGenes: why do we do what we do?

Wheat Barley Rye Oat

food – fodder – beverage - > Calories and Nutrition

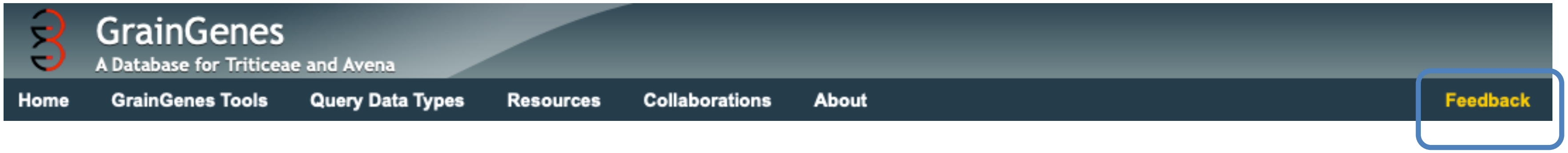
We also help bring fun!

GrainGenes
A Database for Triticeae and Avena

Please cite us: Yao et al., Database 2022

- We appreciate
  - Data generators
  - GrainGenes Liaison Committee

contact us for your needs



USDA
American Malting Barley Association
Berkeley University of California
University of Maryland
PEPSICO
Montana State University
CORTEVA agriscience
Agriculture and Agri-Food Canada
IPK Leibniz Institute
IWGS International Wheat Genome Sequencing Consortium
AgBioData
Toward enhanced genomics, genetics, and breeding research outcomes through standardization of practices and protocols across agricultural databases
wheatIS Wheat Information System