



Database update

GrainGenes: centralized small grain resources and digital platform for geneticists and breeders

Victoria C. Blake^{1,†}, Margaret R. Woodhouse^{1,†}, Gerard R. Lazo¹, Sarah G. Odell^{1,2}, Charlene P. Wight³, Nicholas A. Tinker³, Yi Wang¹, Yong Q. Gu¹, Clay L. Birkett⁴, Jean-Luc Jannink^{4,5}, Dave E. Matthews⁴, David L. Hane¹, Steve L. Michel¹, Eric Yao^{1,6} and Taner Z. Sen^{ID 1,7,*}

¹Western Regional Research Center, Crop Improvement and Genetics Research Unit, United States Department of Agriculture—Agricultural Research Service, 800 Buchanan St., Albany, CA 94710, USA, ²Department of Plant Sciences, University of California, Davis, CA 95616, USA; ³Ottawa Research and Development Centre, Agriculture and Agri-Food Canada, 960 Carling Ave., Ottawa, ON K1A 0C6, Canada, ⁴Robert Holley Center, United States Department of Agriculture—Agricultural Research Service, 538 Tower Rd., Ithaca, NY 14853, USA, ⁵Section of Plant Breeding and Genetics, Cornell University, Ithaca, NY 14853, USA, ⁶Department of Bioengineering, University of California, Berkeley, Berkeley, CA 94720, USA and ⁷Department of Genetics, Development, and Cell Biology, Iowa State University, Ames, IA 50011, USA

*Corresponding author: Tel.: 510-559-5982; Fax: 510-559-5818; Email: taner.sen@ars.usda.gov

†These authors contributed equally to this work.

Citation details: Blake, V.C., Woodhouse, M.R., Lazo, G.R. *et al.* GrainGenes: centralized small grain resources and digital platform for geneticists and breeders. *Database* (2019) Vol. 2019: article ID baz065; doi:10.1093/database/baz065

Received 7 November 2018; Revised 18 April 2019; Accepted 22 April 2019

Abstract

GrainGenes (<https://wheat.pw.usda.gov> or <https://graingenes.org>) is an international centralized repository for curated, peer-reviewed datasets useful to researchers working on wheat, barley, rye and oat. GrainGenes manages genomic, genetic, germplasm and phenotypic datasets through a dynamically generated web interface for facilitated data discovery. Since 1992, GrainGenes has served geneticists and breeders in both the public and private sectors on six continents. Recently, several new datasets were curated into the database along with new tools for analysis. The GrainGenes homepage was enhanced by making it more visually intuitive and by adding links to commonly used pages. Several genome assemblies and genomic tracks are displayed through the genome browsers at GrainGenes, including the *Triticum aestivum* (bread wheat) cv. ‘Chinese Spring’ IWGSC RefSeq v1.0 genome assembly, the *Aegilops tauschii* (D genome progenitor) Aet v4.0 genome assembly, the *Triticum turgidum ssp. dicoccoides* (wild emmer wheat) cv. ‘Zavitan’ WEWSeq v.1.0 genome assembly, a *T. aestivum* (bread wheat) pangenome, the *Hordeum vulgare* (barley) cv. ‘Morex’ IBSC genome assembly, the *Secale cereale* (rye) select ‘Lo7’ assembly, a partial hexaploid *Avena sativa* (oat) assembly and the *Triticum durum* cv. ‘Svevo’ (durum wheat) RefSeq Release 1.0 assembly. New

genetic maps and markers were added and can be displayed through CMAP. Quantitative trait loci, genetic maps and genes from the Wheat Gene Catalogue are indexed and linked through the Wheat Information System (WheatIS) portal. Training videos were created to help users query and reach the data they need. GSP (Genome Specific Primers) and PIECE2 (Plant Intron Exon Comparison and Evolution) tools were implemented and are available to use. As more small grains reference sequences become available, GrainGenes will play an increasingly vital role in helping researchers improve crops.

Database URL: <https://wheat.pw.usda.gov> or <https://graingenes.org>

Introduction

Since 1992, GrainGenes has served as an international centralized hub for peer-reviewed biological data for small grains researchers working on wheat, barley, rye and oat. Its content includes genetic maps, markers, traits and phenotypes (1, 2). With a range of query and visualization tools, including a Structured Query Language (SQL) interface, GrainGenes serves information to enable breeders and geneticists to improve small grain varieties and to better understand the biology of important crops at a molecular level. Since wheat alone constitutes 20% of the calories and 20% of proteins consumed by humans (3), access to the wealth of data in crop databases like GrainGenes contributes to worldwide food security.

Reflecting the global importance of small grains and crop diversity, the userbase of GrainGenes is distributed across six continents. More than half of the GrainGenes users are located in three countries: the USA, China and India. These users are harnessing the resources at GrainGenes to understand the genetic and genomic bases of traits and to develop plants that can combat the continuing challenges from negative biotic and abiotic conditions, such as drought, flood, diseases and pests.

To serve its users, GrainGenes prioritizes and curates peer-reviewed small grains datasets and displays them through dynamically generated web pages. These pages link to tools at GrainGenes and to other biological databases where users can learn more about their biological entities of interest (e.g. genes, markers, genomic regions). Here we describe the most recent additions to the curated data content and tools at GrainGenes that are available to our users. New datasets include results from genomic experiments, characterizations of genes conferring disease resistance and other plant traits and genome sequences from hexaploid ‘Chinese Spring’ (bread wheat), tetraploid ‘Zavitan’ (wild emmer wheat) and ‘Morex’ (barley). New tools include genome browsers for assembled genomes and an expanded collection of BLAST databases. We also improved our front page to make it more visual and intuitive for better user experience.

The GrainGenes portal

The GrainGenes website serves as an information hub for the small grains community and includes linked data sets, job listings and upcoming events. The web interface of GrainGenes is coded in HTML, PHP, Perl and JavaScript, is supported by a MySQL backend database and is wrapped by an instance of the Drupal (drupal.org) content management system. According to Google Analytics, in the 2018 calendar year, GrainGenes was used by 26 433 users (based on unique IPs) in 69 858 sessions and provided 460 291 page views. The main data types at GrainGenes are DNA sequences, quantitative trait loci (QTLs), genetic maps, germplasm, genes, genetic markers (loci, probes, alleles), trait studies, proteins, images and colleague records, among others (Supplementary Figure 1, as of January 2019). GrainGenes has curated information for many genera, including *Triticum* (354), *Aegilops* (50), *Secale* (40) and *Avena* (147). The numbers in parentheses include various species, historical names and aneuploids.

Enhanced homepage

The GrainGenes team has recently made extensive enhancements to the GrainGenes Homepage to improve content, ensure consistent visual communication and provide means for user feedback and interaction (Figure 1). Recent improvements to this interface include the addition of links for data download, access points for the GrainGenes Grains and OatMail mailing lists and tutorials. The visibility of ‘Species Portals’ on the front page has also been improved. These include the Annual Wheat Newsletter, the Barley Genetics Newsletter and the Oat Newsletter. The ‘GrainGenes Updates’ have been itemized and dated, and new ‘Quick Links’ have been added for the database browser, the GrainGenes genome browsers and the revamped CMap page (4,5; <https://wheat.pw.usda.gov/GG3/CMAP/>). To improve communication with our users, a ‘Feedback’ button has now been added to the header of each page, promoting a 24-hour response time to assure users that their feedback is valued. The GrainGenes team

The screenshot shows the GrainGenes homepage with the following sections:

- Header:** GrainGenes logo, "A Database for Triticeae and Avena", and navigation links: Home, GrainGenes Tools, Query Data Types, Resources, Collaborations, About, and Feedback.
- Search:** Links to "Browse GrainGenes" and "Search Website".
- Community Services:** Links to "Calendar", "Current Hot Topics", "Data Download", "Database Information", "GrainGenes Mailing List", "Job Listings", "Oatmail Mailing List", "Submit Data", and "Tutorials".
- Species Portals on GrainGenes:** Links to "Annual Wheat Newsletter", "Barley Boulevard", "Barley Genetics Newsletter", and "Oat Newsletter".
- Upcoming Events:** Lists events such as the "15th Gatersleben Research Conference, Applied Bioinformatics in Crops, Gatersleben, Germany" (Mar 18-20, 2019) and the "1st International Wheat Congress, Saskatoon, Saskatchewan, Canada" (Jul 21-26, 2019).
- About GrainGenes:** Includes the USDA logo.
- Quick Links:** "Browse GrainGenes", "Genome Browsers", and "CMap".
- Hot Topics:** "Durum Wheat (cv. Svevo) RefSeq Release 1.0 at GrainGenes". Text: "In collaboration with the International Durum Wheat Genome Sequencing Consortium, GrainGenes is releasing the Durum Wheat (cv. Svevo) RefSeq Rel. 1.0 and its annotations. The manuscript for the Svevo genome reference is accepted for publication in Nature Genetics. Please cite the manuscript for your publications."
- All IWGSC reference sequence resources now publicly available:** Text: "The International Wheat Genome Sequencing Consortium is pleased to announce that all data related to the reference sequence of bread wheat, IWGSC RefSeq v1.0, are available without restriction. You can reach the Science article for the assembly here. The following resources are now publicly available, without registration, at the IWGSC Data Repository hosted by URGI-INRA:"
- A new mobile app for plant breeding and genetics : Verify:** Text: "Researchers at Kansas State University and Cornell University have developed and released an Android application called 'Verify' that will help plant breeders by decreasing the amount of time and energy it takes to manage samples and seed stocks. Verify makes integrating and utilizing barcodes a simple process by combining an"
- GrainGenes Updates:** A list of recent updates including "March 2019: Durum QTL added to GrainGenes", "February 2019: Legacy oat maps online at GrainGenes", "February 2019: Durum Wheat (cv. Svevo) RefSeq Release 1.0 is available at GrainGenes", "December 2018: Barley Genetics Newsletter v47 (2017) is available at GrainGenes", "November 2018: GrainGenes has an active GitHub account to share scripts", "October 2018: Spring Wheat Nested Association (NAM) Maps are in GrainGenes", "September 2018: Germplasm from the Global Tetraploid Wheat Collection added to GrainGenes and WheatIS", "August 2018: 2017 Uniform Regional Scab Nursery for Spring Wheat Parents Nursery report is available", "July 2018: Wheat Gene Catalogue with links to Komugi is added to WheatIS", "July 2018: 2017 Uniform Regional Hard Red Spring Wheat Nursery data is available", "July 2018: Genome browser at GrainGenes for the IWGSC Chinese Spring Browser is now publicly available", and "June 2018: Oatmail Mailing List is created".
- @GrainGenes Tweets:** A tweet from @GrainGenes: "Durum Wheat (cv. Svevo) RefSeq Release 1.0 is available at GrainGenes: wheat.pw.usda.gov/GG3/node/759".

Figure 1. The enhanced GrainGenes homepage.

has a social media presence on Twitter and Facebook to share current information and open direct communication channels with users of these platforms.

Genomic data visualization

Driven by lower sequencing and computational costs (6), more genomes are being sequenced and assembled. Stephens *et al.* (7) predict that one zetta-base (i.e. 10^{21}) of sequence information will be generated by 2025. Consequently, GrainGenes is transitioning into a genome-centric database, focusing on genomic data integration, visualization and management. In the past few years, GrainGenes has imported multiple small grains genome assemblies that are served using JBrowse-based genome browsers. These genome assemblies include the *Triticum aestivum* (bread wheat) cv. Chinese Spring IWGSC RefSeq v1.0 genome assembly (8), the *Aegilops tauschii* (D genome progenitor) Aet v4.0 genome assembly (9), the *Triticum turgidum ssp. dicoccoides* (wild emmer wheat) cv. Zavitan WEWSeq v1.0 genome assembly (10), a *T. aestivum* (bread wheat) pangenome (11), the *Hordeum vulgare* (barley) cv. Morex IBSC genome assembly (12), the *Secale cereale* (rye)

select 'Lo7' assembly (13), the *Triticum durum* cv. 'Svevo' (durum wheat) RefSeq Release 1.0 assembly (unpublished) and a hexaploid *Avena sativa* (oat) assembly (14; note that this is not a whole genome assembly). Tracks showing single nucleotide polymorphism variation on the GrainGenes *H. vulgare* (barley) cv. Morex IBSC Genome Browser were created by the Triticeae Toolbox (T3) team (15; Figure 2).

New curated genetic content

Since 2016, several sets of high-impact genetic markers and maps have been added to the GrainGenes database (16–27; Table 1). These data are displayed through a CMap (5) instance, thus adding value to the data content at GrainGenes for genetic maps, QTLs, loci, genes, DNA sequences and germplasm. Table 1 is a brief summary of newly curated genetic data sets (28).

Collecting this information has involved developing and maintaining close collaborations with many groups, for example MASWheat (<http://maswheat.ucdavis.edu>), an extensively curated site maintained at UC Davis by Jorge Dubcovsky and Marcelo Soria that assists small grains researchers in the use of marker-assisted selection for a

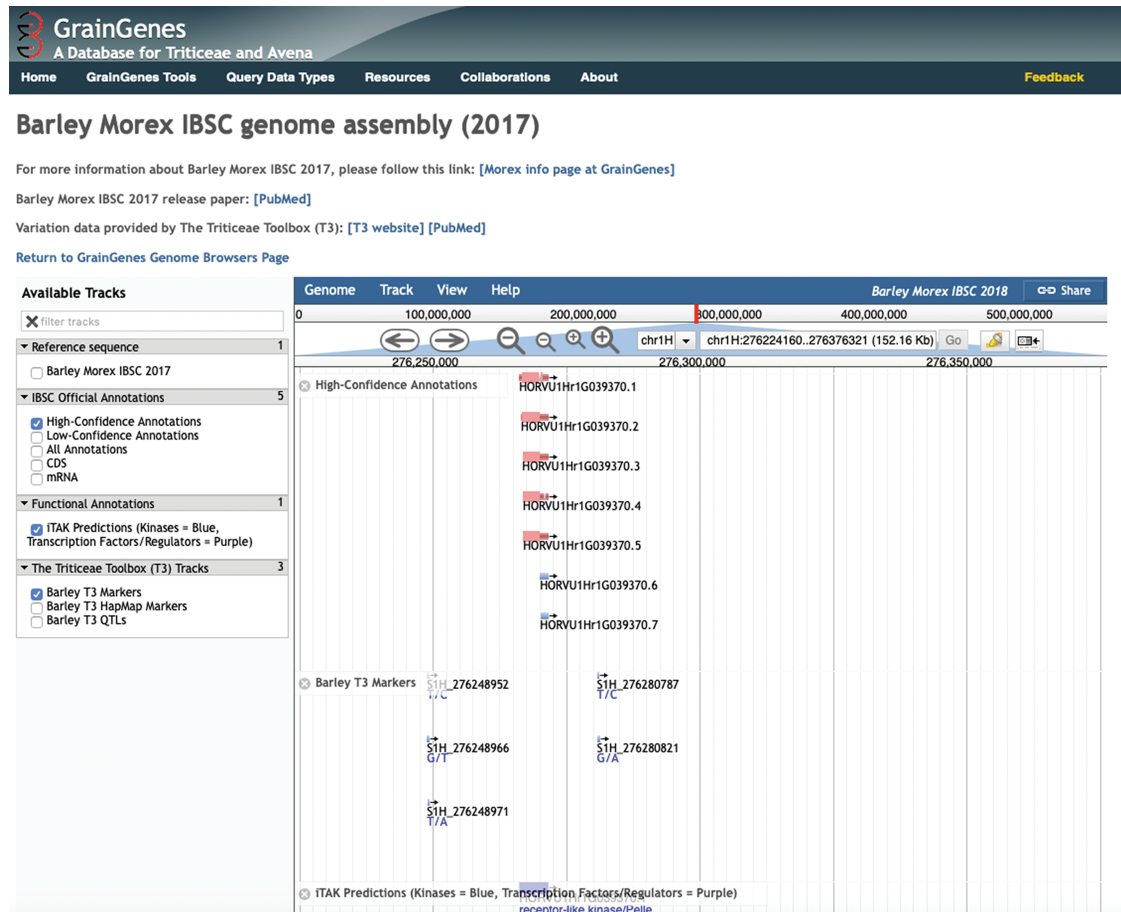


Figure 2. A view from the Barley Morex genome browser.

Table 1. A brief summary of newly curated genetic data sets in GrainGenes

Species	Project	reference
Wheat	MASWheat project: new additions include genes implemented in leaf rust (Lr19), stem rust (Sr25) and quality traits (semolina color gene Y)	http://maswheat.ucdavis.edu
Wheat	Three Pistil (Pis1) QTL Map: Pis1 directs the plant to produce three seeds where normal wheat produces one seed.	(27)
Wheat	Tetraploid wheat consensus map	(25)
Wheat	Spring Wheat Nested Association Maps	(24)
Wheat (durum)	Mapping of carotenoid genes and QTL	(20)
Wheat (durum)	National Small Grains Collection: durum wheat stem rust data	(19)
Wheat (durum)	Durum wheat pre-harvest sprouting data	(18)
<i>Aegilops umbellulata</i>	Consensus maps from two F2 <i>A. umbellulata</i> populations from Manisa, Usak, Denizli and Diyarbakir in Turkey	(22)
Barley	Barley 9K iSelect SNP data	(21)
Rye	Rye consensus map	(23)
Oat	2016 Oat consensus map	(17)
Oat	2018 Oat consensus map	(16)
Oat	Partial maps of Pendek39/Pendek48 and Pendek48/Pendek38 NIL populations, segregating for crown rust resistance	(26)

wide range of traits. We curated data sets for the leaf rust resistance LR genes (e.g. *Lr19*), the stem rust resistance SR genes (e.g. *Sr25*) and the stripe rust Yr resistance genes (e.g. *Yr5*). A complete list can be found here: <https://wheat.pw.usda.gov/GG3/node/657>.

A more recent collaboration has been established with Agriculture and Agri-Food Canada and the oat community. We added two new oat (*A. sativa*) map data sets to GrainGenes. These are from the Pendek 39×Pendek 48 Oat-2004-P39×P48 and Pendek 48×Pendek 38 Oat-2004-P48×P38 populations (29). We have begun to curate and include additional ‘fragmentary’ maps from the historical literature and more recent studies as well. New consensus maps for oat have also been added (30–34).

In addition, to reconcile marker names, GrainGenes has embarked on a major effort to add probes, marker types and locus orthology groups to all existing historical markers in GrainGenes, thus enabling CMap to create more correspondences among existing maps in GrainGenes through common markers. Also, GrainGenes is working with the AgBioData Consortium to develop common data standards (35). More information about recent updates to the GrainGenes database can be found here: <https://wheat.pw.usda.gov/GG3/GGupdates>.

New tools and resources at GrainGenes: GSP and PIECE2

GrainGenes has added the web-based tool GSP (Genome Specific Primers), designed to identify genome-specific targets in polyploid species (<https://probes.pw.usda.gov/GSP/>; 36), to its existing tool set (Supplementary Figure 2). When users enter sequence(s), GSP calculates and predicts primer sets that will amplify polymerase chain reaction product(s) from a specific subgenome of a polyploid species, such as *T. aestivum*, *Panicum virgatum* or *Gossypium hirsutum*. PIECE version 2 (Plant Intron Exon Comparison and Evolution; 37), a database for studying plant gene structure and evolution, is now a part of GrainGenes (<http://probes.pw.usda.gov/piece>) as well. We encourage users who would like to learn more about GSP and PIECE2 to read the following publications, which describe the tools in detail and provide use cases that demonstrate how these tools can be useful for research (36, 37).

Data discovery through the Wheat Information System

Wheat Information System (WheatIS) is an international collaborative website led by researchers at Unité de Recherche Génomique Info (URGI) under Institut National de la Recherche Agronomique in Versailles, France.

GrainGenes is part of the WheatIS Expert Working Group that consists of representatives from universities, governments and industries with the goal of sharing information and best practices for storing, analyzing and displaying wheat genomic sequence data (38, 39). URGI is the largest contributor to WheatIS, followed by Ensembl Plants (40). The full list of collaborators and contributors can be found at <http://wheatis.org/Collaborators.php> and in Alaux *et al.* (39). WheatIS indexes wheat data sets from globally distributed databases to make them searchable at wheatis.org and provides links back to databases. As a result, instead of visiting multiple databases separately, users can use the WheatIS website to query and find information distributed across different databases in one single location. The data providers have several options for their data to be indexed at WheatIS, including (i) sending a comma-separated value file to the developers through the WheatIS website and (ii) installing and configuring an Apache Solr server (<http://lucene.apache.org/solr/>) and creating custom-formatted data sets to make them searchable at WheatIS. The data formats can be easily obtained by contacting WheatIS personnel.

As a service to the global small grains community, GrainGenes indexed 548 QTLs, 91 genetic maps, 10 physical maps and 14411 germplasm records at WheatIS with links back to appropriate GrainGenes pages. In addition, GrainGenes collaborated with the Wheat Gene Catalogue team to index 3119 genes from the Catalogue at WheatIS with links to the appropriate pages at the Komugi database, located in Japan (<https://shigen.nig.ac.jp/wheat/komugi/genes/symbolClassList.jsp>).

Training and public outreach

While GrainGenes attempts to provide intuitive interfaces, gaining the knowledge required for the most effective use of the GrainGenes database and its related tools is critical to our users. To this end, we created the GrainGenes YouTube Channel (<https://wheat.pw.usda.gov/GG3/tutorials>). To date, three training videos have been posted on YouTube for various tools and features offered at GrainGenes: (i) ‘Navigating the GrainGenes homepage’, (ii) a database browser demo and (iii) a demo on ‘Obtaining FASTA sequences from Markers’. In addition, PDF slides converted from the video tutorials are also available, as well as slides on how to use the genome browsers, CMap and the Quick Queries tool. In accordance with the GrainGenes project mandate, the GrainGenes team will continue to create training videos for the users.

Important scientific articles and breakthroughs in the area of small grains research are shared on the GrainGenes homepage. To improve public outreach, social media

accounts have been created for GrainGenes on Facebook and Twitter to (i) increase the reach of GrainGenes updates, (ii) attract new users, (iii) inform users about new tools and (iv) broadcast news of new data at GrainGenes. Important scientific articles and breakthroughs in the area of small grains research are shared on the GrainGenes homepage. GrainGenes also provides a free service for the announcement of relevant job openings and meetings. Such public outreach efforts help GrainGenes to continue serving as a hub for small grains communication.

Conclusion and future steps

With the decreasing cost of genome sequencing technologies and the improvement of assembly algorithms, we expect more small grains genomes to be sequenced and assembled in the near future. In addition, new methods such as genotyping-by-sequencing (41, 42) are creating large amounts of data, reflecting the vast diversity and complexity of small grains genomes. More pangenomes that integrate genome assemblies with diversity datasets are also expected to become available. The integration of large amounts of data will be increasingly challenging, but will also be increasingly valuable. Because biological repositories are tasked with managing, curating, storing, querying and visualizing an ever-increasing volume and diversity of data, more resources will be required to store, analyze and link these datasets, and more efficient web-based tools must be implemented. GrainGenes recognizes these challenges. Our future objectives include developing semi-automated curation workflows, implementing improved web-based visualization and query tools and developing computational pipelines to create integrated views. GrainGenes will continuously seek feedback from its users and the GrainGenes Liaison Committee to enhance GrainGenes' value for the broader small grains community.

Supplementary data

Supplementary data are available at *Database* Online.

Acknowledgements

GrainGenes is supported by the United States Department of Agriculture—Agricultural Research Service (USDA-ARS) under the CRIS project 2030-21000-024-00D. Mention of trade names or commercial products is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the USDA. USDA is an equal opportunity provider and employer. The authors would like to thank Ann Blechl for critically reading the manuscript and providing useful suggestions. GrainGenes benefits significantly from the oversight provided by the members of the GrainGenes Liaison Committee: Mark Sorrells (Chair), Jorge Dubcovsky, Catherine Howarth, Kevin Smith and Roger Wise.

Conflict of interest. None declared.

References

1. Matthews,D.E., Carollo,V.L., Lazo,G.R. *et al.* (2003) GrainGenes, the genome database for small-grain crops. *Nucleic Acids Res.*, **31**, 183–186.
2. Carollo,V., Matthews,D.E., Lazo,G.R. *et al.* (2005) GrainGenes 2.0. An improved resource for the small-grains community. *Plant Physiol.*, **139**, 643–651.
3. Shiferaw,B., Smale,M., Braun,H.-J. *et al.* (2013) Crops that feed the world 10. Past successes and future challenges to the role played by wheat in global food security. *Food Secur.*, **5**, 291–317.
4. Faga,B. (2007) Installing and configuring CMap. *Curr. Protoc. Bioinformatics*, Chapter 9, Unit 9.8.1–9.8.30.
5. Youens-Clark,K., Faga,B., Yap,I.V. *et al.* (2009) CMap 1.01: a comparative mapping application for the internet. *Bioinformatics*, **25**, 3040–3042.
6. Muir,P., Li,S., Lou,S. *et al.* (2016) The real cost of sequencing: scaling computation to keep pace with data generation. *Genome Biol.*, **17**, 53.
7. Stephens,Z.D., Lee,S.Y., Faghri,F. *et al.* (2015) Big data: astronomical or genomics? *PLoS Biol.*, **13**, e1002195.
8. International Wheat Genome Sequencing Consortium, IWGSC RefSeq principal investigators, Appels,R. *et al.* (2018) Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science*, **361**, 661.
9. Luo,M.C., Gu,Y.Q., Puiu,D. *et al.* (2017) Genome sequence of the progenitor of the wheat D genome *Aegilops tauschii*. *Nature*, **551**, 498–502.
10. Avni,R., Nave,M., Barad,O. *et al.* (2017) Wild emmer genome architecture and diversity elucidate wheat evolution and domestication. *Science*, **357**, 93–97.
11. Montenegro,J.D., Golicz,A.A., Bayer,P.E. *et al.* (2017) The pangenome of hexaploid bread wheat. *Plant J.*, **90**, 1007–1013.
12. Beier,S., Himmelbach,A., Colmsee,C. *et al.* (2017) Construction of a map-based reference genome sequence for barley, *Hordeum vulgare* L. *Sci. Data*, **4**, 170044.
13. Bauer,E., Schmutzer,T., Barilar,I. *et al.* (2017) Towards a whole-genome sequence for rye (*Secale cereale* L.). *Plant J.*, **89**, 853–869.
14. Oliver,R.E., Tinker,N.A., Lazo,G.R. *et al.* (2013) SNP discovery and chromosome anchoring provide the first physically-anchored hexaploid oat map and reveal synteny with model species. *PLoS One*, **8**, e58068.
15. Blake,V.C., Birkett,C., Matthews,D.E. *et al.* (2016) The Triticeae toolbox: combining phenotype and genotype data to advance small-grains breeding. *Plant Genome*, **9**. [10.3835/plantgenome2014.12.0099](https://doi.org/10.3835/plantgenome2014.12.0099).
16. Bekele,W.A., Wight,C.P., Chao,S. *et al.* (2018) Haplotype-based genotyping-by-sequencing in oat genome research. *Plant Biotechnol. J.*, **16**, 1452–1463.
17. Chaffin,A.S., Huang,Y.F., Smith,S. *et al.* (2016) A consensus map in cultivated Hexaploid oat reveals conserved grass synteny with substantial subgenome rearrangement. *Plant Genome*, **9**. [10.3835/plantgenome2015.10.0102](https://doi.org/10.3835/plantgenome2015.10.0102).
18. Chao,S., Elias,E., Benschler,D. *et al.* (2016) Genetic mapping of major-effect seed dormancy quantitative trait loci on chromosome 2B using recombinant substitution lines in tetraploid wheat. *Crop. Sci.*, **56**, 59–72.

19. Chao,S., Rouse,M.N., Acevedo,M. *et al.* (2017) Evaluation of genetic diversity and host resistance to stem rust in USDA NSGC durum wheat accessions. *Plant Genome*, **10**. 10.3835/plantgenome2016.07.0071.
20. Colasuonno,P., Lozito,M.L., Marcotuli,I. *et al.* (2017) The carotenoid biosynthetic and catabolic genes in wheat and their association with yellow pigments. *BMC Genomics*, **18**, 122.
21. Comadran,J., Kilian,B., Russell,J. *et al.* (2012) Natural variation in a homolog of *Antirrhinum* CENTRORADIALIS contributed to spring growth habit and environmental adaptation in cultivated barley. *Nat. Genet.*, **44**, 1388–1392.
22. Edae,E.A., Olivera,P.D., Jin,Y. *et al.* (2017) Genotyping-by-sequencing facilitates a high-density consensus linkage map for *Aegilops umbellulata*, a wild relative of cultivated wheat. *G3 (Bethesda)*, **7**, 1551–1561.
23. Gustafson,J.P., Ma,X.F., Korzun,V. *et al.* (2009) A consensus map of rye integrating mapping data from five mapping populations. *Theor. Appl. Genet.*, **118**, 793–800.
24. Jordan,K.W., Wang,S., He,F. *et al.* (2018) The genetic architecture of genome-wide recombination rate variation in allopolyploid wheat revealed by nested association mapping. *Plant J.*, **95**, 1039–1054.
25. Maccaferri,M., Ricci,A., Salvi,S. *et al.* (2015) A high-density, SNP-based consensus map of tetraploid wheat as a bridge to integrate durum and bread wheat genomics and breeding. *Plant Biotechnol. J.*, **13**, 648–663.
26. Wight,C.P., O'Donoghue,L.S., Chong,J. *et al.* (2005) Discovery, localization, and sequence characterization of molecular markers for the crown rust resistance genes *Pc38*, *Pc39*, and *Pc48* in cultivated oat (*Avena sativa* L.). *Mol. Breed.*, **14**, 349–361.
27. Yang,Z., Chen,Z., Peng,Z. *et al.* (2017) Development of a high-density linkage map and mapping of the three-pistil gene (*Pis1*) in wheat using GBS markers. *BMC Genomics*, **18**, 567.
28. Odell,S.G., Lazo,G.R., Woodhouse,M.R. *et al.* (2017) The art of curation at a biological database: principles and application. *Curr. Plant. Biol.*, **11–12**, 2–11.
29. Wight,C.P., O'Donoghue,L.S., Chong,J. *et al.* (2004) Discovery, localization, and sequence characterization of molecular markers for the crown rust resistance genes *Pc38*, *Pc39*, and *Pc48* in cultivated oat (*Avena sativa* L.). *Mol. Breed.*, **14**, 349–361.
30. Wight,C.P., Kibite,S., Tinker,N.A. *et al.* (2006) Identification of molecular markers for aluminium tolerance in diploid oat through comparative mapping and QTL analysis. *Theor. Appl. Genet.*, **112**, 222–231.
31. Rooney,W.L. (1994) Identification of RFLP markers linked to crown rust resistance genes *pc 91* and *pc* in oat. *Crop. Sci.*, **34**, 940–944.
32. Groh,S., Zacharias,A., Kianian,S.F. *et al.* (2001) Comparative AFLP mapping in two hexaploid oat populations. *Theor. Appl. Genet.*, **102**, 876–884.
33. Kianian,S.F., Phillips,R.L., Rines,H.W. *et al.* (2000) Quantitative trait loci influencing β -glucan content in oat (*Avena sativa*, $2n = 6x = 42$). *Theor. Appl. Genet.*, **101**, 1039–1048.
34. Kianian,S.F., Egli,M.A., Phillips,R.L. *et al.* (1999) Association of a major groat oil content QTL and an acetyl-CoA carboxylase gene in oat. *Theor. Appl. Genet.*, **98**, 884–994.
35. Harper,L., Campbell,J., Cannon,E.K.S. *et al.* (2018) AgBioData consortium recommendations for sustainable genomics and genetics databases for agriculture. *Database (Oxford)*, **2018**.
36. Wang,Y., Tiwari,V.K., Rawat,N. *et al.* (2016) GSP: a web-based platform for designing genome-specific primers in polyploids. *Bioinformatics*, **32**, 2382–2383.
37. Wang,Y., Xu,L., Thilmony,R. *et al.* (2017) PIECE 2.0: an update for the plant gene structure comparison and evolution database. *Nucleic Acids Res.*, **45**, 1015–1020.
38. Dzale Yeumo,E., Alaux,M., Arnaud,E. *et al.* (2017) Developing data interoperability using standards: a wheat community use case. *F1000Res*, **6**, 1843.
39. Alaux,M., Rogers,J., Letellier,T. *et al.* (2018) Linking the international wheat genome sequencing consortium bread wheat reference genome sequence to wheat genetic and phenomic data. *Genome Biol.*, **19**, 111.
40. Bolser,D.M., Staines,D.M., Perry,E. *et al.* (2017) Ensembl plants: integrating tools for visualizing, mining, and analyzing plant genomic data. *Methods Mol. Biol.*, **1533**, 1–31.
41. Deschamps,S., Llaca,V. and May,G.D. (2012) Genotyping-by-sequencing in plants. *Biology (Basel)*, **1**, 460–483.
42. Elshire,R.J., Glaubitz,J.C., Sun,Q. *et al.* (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One*, **6**, e19379.