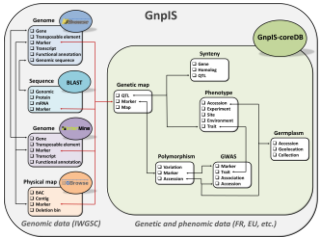
JBrowse & Apollo for manual annotation of wheat

Anne-Françoise Adam-Blondon,
Michael Alaux

IWGSC manual and functional annotation workshop



IWGSC Data Repository

Seq Repository

You are here : Home / Home Wheat / Seq Repository

Seq Repository

Click on a chromosome to download, BLAST or display the sequences.

News, Access status, etc. are detailed in the left menu.

Chr1A

Chromosome 1A:
- IWGSC BLAST (**IWGSC RefSeq**, Survey sequence, TGACv1, 3B reference, etc.).
- Download assemblies .
- Download annotations .
- Display IWGSC RefSeq v1.0 browser .
- Display physical maps browser (click on the chromosome arm).
- Order BAC clones: TaaCsp1AShA , TaaCsp1ALhA , TaaCsp1ALhB .

Sequence Repository page .
IWGSC page .

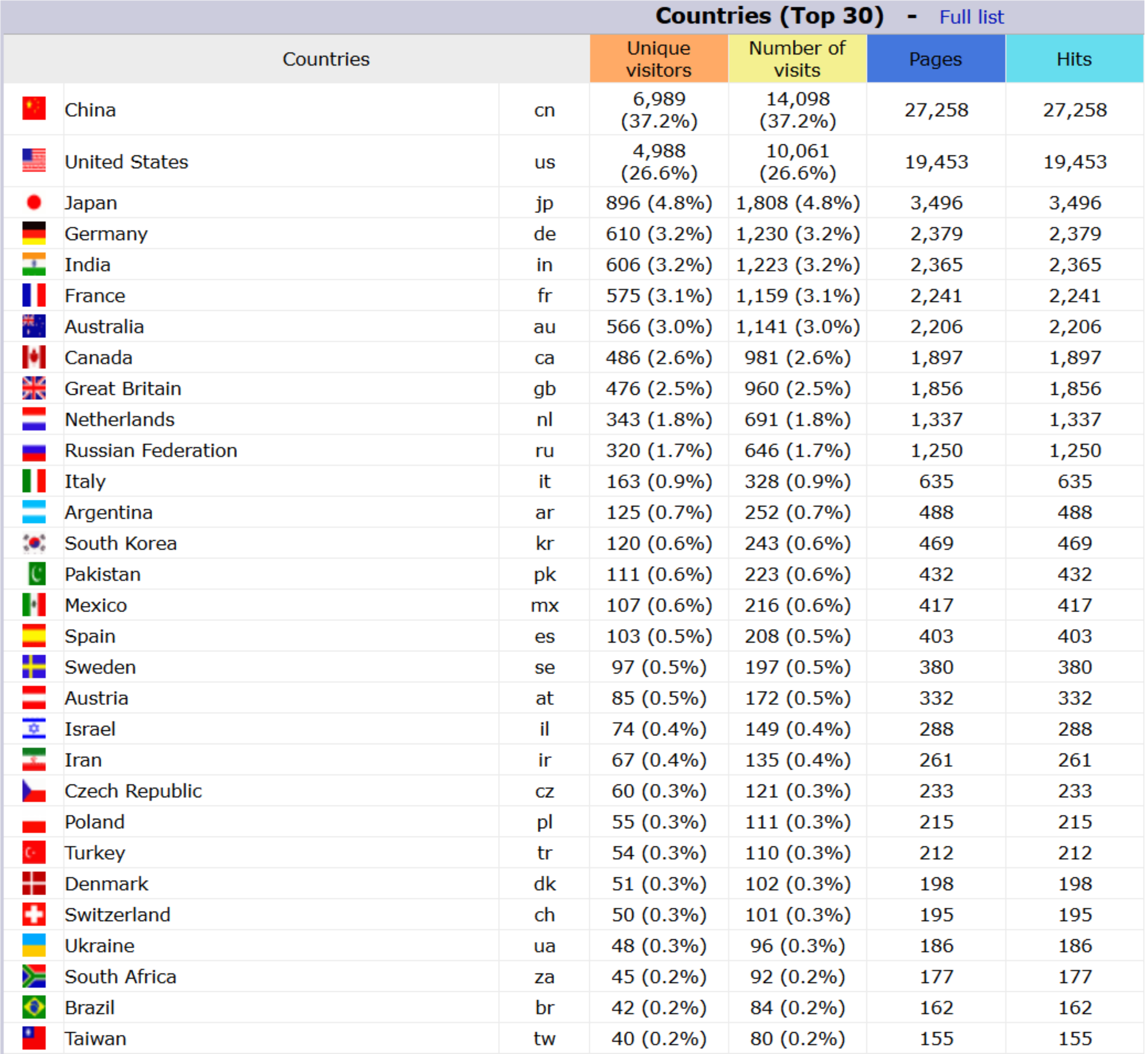
http://wheat-urgi.versailles.inra.fr/Seq-Repository

Usage metrics

IWGSC Data Repository resources	2019
Data Repository website visits	37 866
Downloads	43 530
Genome browsers visits	100 182
BLAST searches	145 638

- All statistics show increases except the Blasts (slightly lower than in 2018)
- The downloads concern essentially the V2 assembly

IWGSC Data Repository website visits by country: an international resource



Countries (Top 30) - Full list

Countries		Unique visitors	Number of visits	Pages	Hits
China	cn	6,989 (37.2%)	14,098 (37.2%)	27,258	27,258
United States	us	4,988 (26.6%)	10,061 (26.6%)	19,453	19,453
Japan	jp	896 (4.8%)	1,808 (4.8%)	3,496	3,496
Germany	de	610 (3.2%)	1,230 (3.2%)	2,379	2,379
India	in	606 (3.2%)	1,223 (3.2%)	2,365	2,365
France	fr	575 (3.1%)	1,159 (3.1%)	2,241	2,241
Australia	au	566 (3.0%)	1,141 (3.0%)	2,206	2,206
Canada	ca	486 (2.6%)	981 (2.6%)	1,897	1,897
Great Britain	gb	476 (2.5%)	960 (2.5%)	1,856	1,856
Netherlands	nl	343 (1.8%)	691 (1.8%)	1,337	1,337
Russian Federation	ru	320 (1.7%)	646 (1.7%)	1,250	1,250
Italy	it	163 (0.9%)	328 (0.9%)	635	635
Argentina	ar	125 (0.7%)	252 (0.7%)	488	488
South Korea	kr	120 (0.6%)	243 (0.6%)	469	469
Pakistan	pk	111 (0.6%)	223 (0.6%)	432	432
Mexico	mx	107 (0.6%)	216 (0.6%)	417	417
Spain	es	103 (0.5%)	208 (0.5%)	403	403
Sweden	se	97 (0.5%)	197 (0.5%)	380	380
Austria	at	85 (0.5%)	172 (0.5%)	332	332
Israel	il	74 (0.4%)	149 (0.4%)	288	288
Iran	ir	67 (0.4%)	135 (0.4%)	261	261
Czech Republic	cz	60 (0.3%)	121 (0.3%)	233	233
Poland	pl	55 (0.3%)	111 (0.3%)	215	215
Turkey	tr	54 (0.3%)	110 (0.3%)	212	212
Denmark	dk	51 (0.3%)	102 (0.3%)	198	198
Switzerland	ch	50 (0.3%)	101 (0.3%)	195	195
Ukraine	ua	48 (0.3%)	96 (0.3%)	186	186
South Africa	za	45 (0.2%)	92 (0.2%)	177	177
Brazil	br	42 (0.2%)	84 (0.2%)	162	162
Taiwan	tw	40 (0.2%)	80 (0.2%)	155	155

IWGSC RefSeq v1.0

- All data are in open access
- New annotation data in 2019:
  - GBS SNP calls (Poland and Muehlbauer labs)
  - 280k markers from the BreedWheat project (E. Paux)
  - tRNA (H. Budak)
  - 1000 wheat exomes (E. Akhunov)

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

IWGSC RefSeq v1.0

- Download assembly and annotations (v1.0 and v1.1)

```
[DIR] Parent Directory                                          -
[ ]   iwgsc_refseqv1.0_all_chromosomes.zip          04-May-2017 12:31  8.0G
[TXT] iwgsc_refseqv1.0_all_chromosomes.zip.md5.txt   05-Dec-2017 15:00   71
[ ]   iwgsc_refseqv1.0_chr1A.fsa.zip                04-May-2017 12:14 169M
[TXT] iwgsc_refseqv1.0_chr1A.fsa.zip.md5.txt         05-Dec-2017 15:00   65
[ ]   iwgsc_refseqv1.0_chr1B.fsa.zip                04-May-2017 12:15 197M
[TXT] iwgsc_refseqv1.0_chr1B.fsa.zip.
[ ]   iwgsc_refseqv1.0_chr1D.fsa.zip
[TXT] iwgsc_refseqv1.0_chr1D.fsa.zip.
```

```
[ ]   iwgsc_refseqv1.0_1000_wheat_exomes.zip                            09-Jul-2019 11:40  2.1G
[TXT] iwgsc_refseqv1.0_1000_wheat_exomes.zip.md5.txt                    09-Jul-2019 11:49    72
[ ]   iwgsc_refseqv1.0_FunctionalAnnotation_v1.zip                      06-Dec-2017 14:49   19M
[TXT] iwgsc_refseqv1.0_FunctionalAnnotation_v1.zip.md5.txt              25-Jan-2018 16:36    81
[ ]   iwgsc_refseqv1.0_HighConf_2017Mar13.gff3.zip                      24-Mar-2017 17:26   11M
[TXT] iwgsc_refseqv1.0_HighConf_2017Mar13.gff3.zip.md5.txt              05-Dec-2017 15:01    79
[ ]   iwgsc_refseqv1.0_HighConf_CDS_2017Mar13.fa.zip                    24-Mar-2017 17:27   43M
[TXT] iwgsc_refseqv1.0_HighConf_CDS_2017Mar13.fa.zip.md5.txt            05-Dec-2017 15:01    81
[ ]   iwgsc_refseqv1.0_HighConf_PROTEIN_2017Mar13.fa.zip                24-Mar-2017 17:27   26M
[TXT] iwgsc_refseqv1.0_HighConf_PROTEIN_2017Mar13.fa.zip.md5.txt        05-Dec-2017 15:01    85
[ ]   iwgsc_refseqv1.0_HighConf_REPR_CDS_2017Apr03.fa.zip               11-Apr-2017 15:45   40M
[TXT] iwgsc_refseqv1.0_HighConf_REPR_CDS_2017Apr03.fa.zip.md5.txt       05-Dec-2017 15:01    86
[ ]   iwgsc_refseqv1.0_HighConf_REPR_PROTEIN_2017Apr03.fa.zip           11-Apr-2017 15:45   25M
[TXT] iwgsc_refseqv1.0_HighConf_REPR_PROTEIN_2017Apr03.fa.zip.md5.txt   05-Dec-2017 15:01    90
[ ]   iwgsc_refseqv1.0_HighConf_UTR_2017May05.gff3.zip                  09-May-2017 09:53   18M
[TXT] iwgsc_refseqv1.0_HighConf_UTR_2017May05.gff3.zip.md5.txt          05-Dec-2017 15:01    83
[ ]   iwgsc_refseqv1.0_ISBP.fa.zip                                      03-Jan-2018 18:06  209M
[TXT] iwgsc_refseqv1.0_ISBP.fa.zip.md5.txt                              05-Dec-2017 15:01    64
```

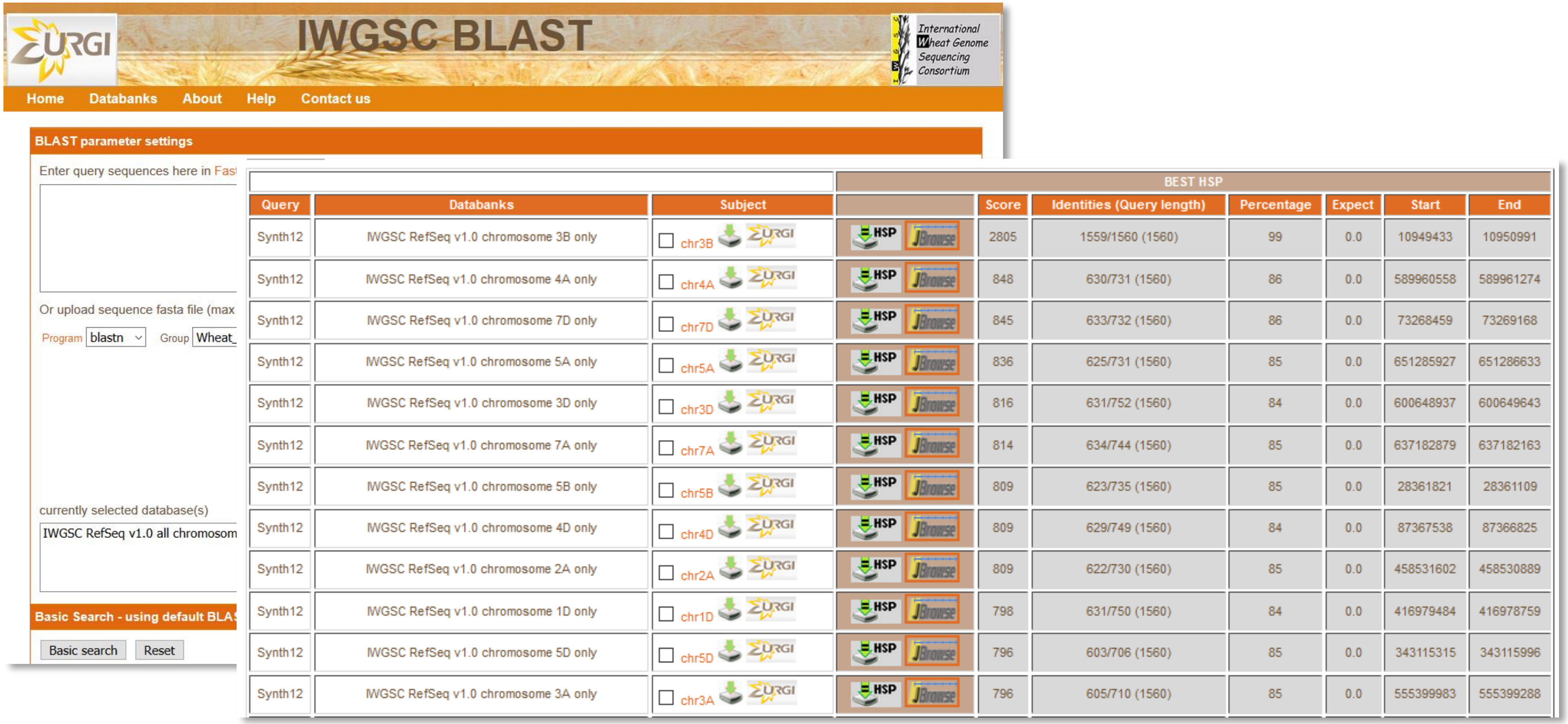
http://wheat-urgi.versailles.inra.fr/Seq-Repository

IWGSC RefSeq v1.0

OPEN ACCESS

- BLAST
  - New bank available: durum wheat Svevo
  - Addition of 10 genomes as soon as open

10+ GENOME PROJECT



BLAST parameter settings

Enter query sequences here in Fast

Or upload sequence fasta file (max

Program blastn Group Wheat_

currently selected database(s)

IWGSC RefSeq v1.0 all chromosom

Basic Search - using default BLAS

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

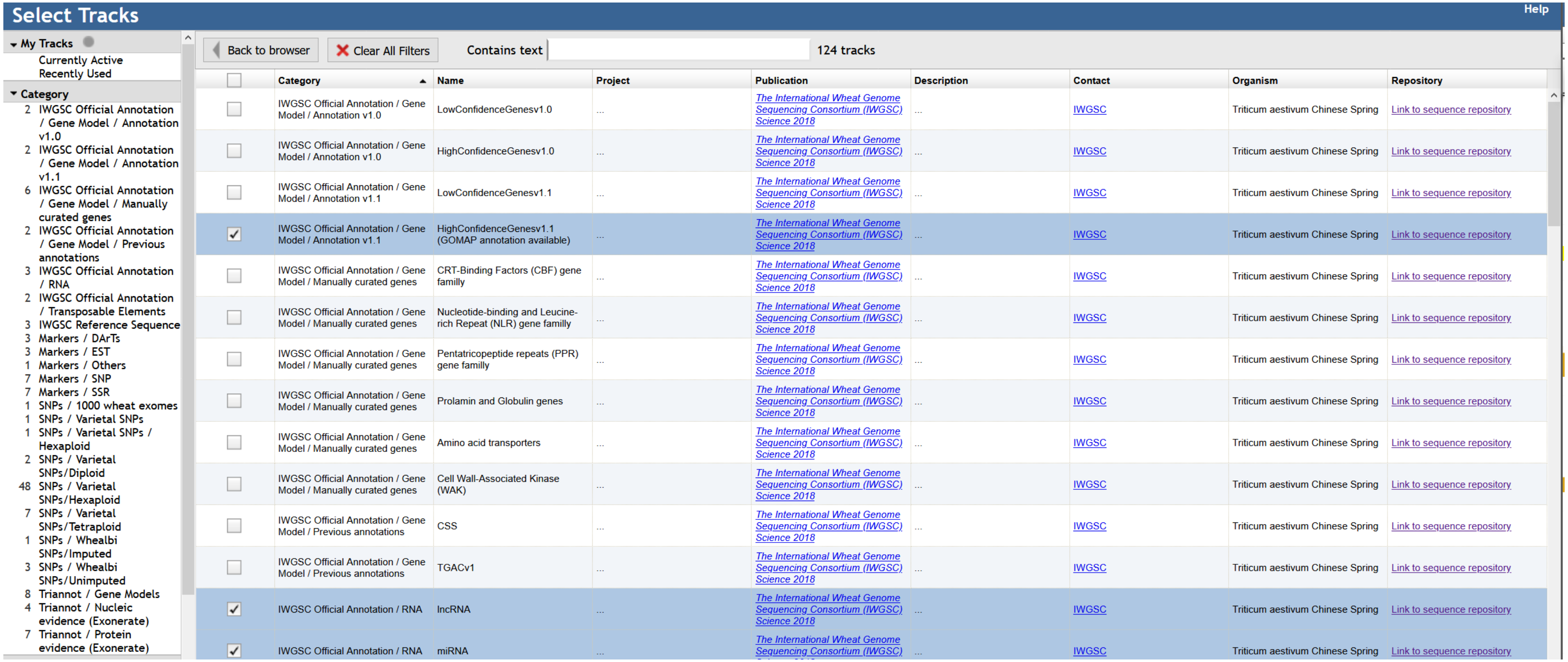
URGI

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

IWGSC RefSeq v1.0

OPEN ACCESS

- JBrowse
  - **Tracks selector tool:** one button => 124 tracks with more metadata. Facets and possibility to do text search to ease the selection of tracks



2 IWGSC Official Annotation / Gene Model / Annotation v1.0
2 IWGSC Official Annotation / Gene Model / Annotation v1.1
6 IWGSC Official Annotation / Gene Model / Manually curated genes
2 IWGSC Official Annotation / Gene Model / Previous annotations
3 IWGSC Official Annotation / RNA
2 IWGSC Official Annotation / Transposable Elements
3 IWGSC Reference Sequence
3 Markers / DArTs
3 Markers / EST
1 Markers / Others
7 Markers / SNP
7 Markers / SSR
1 SNPs / 1000 wheat exomes
1 SNPs / Varietal SNPs
1 SNPs / Varietal SNPs / Hexaploid
2 SNPs / Varietal SNPs/Diploid
48 SNPs / Varietal SNPs/Hexaploid
7 SNPs / Varietal SNPs/Tetraploid
1 SNPs / Whealbi SNPs/Imputed
3 SNPs / Whealbi SNPs/Unimputed
8 Triannot / Gene Models
4 Triannot / Nucleic evidence (Exonerate)
7 Triannot / Protein evidence (Exonerate)

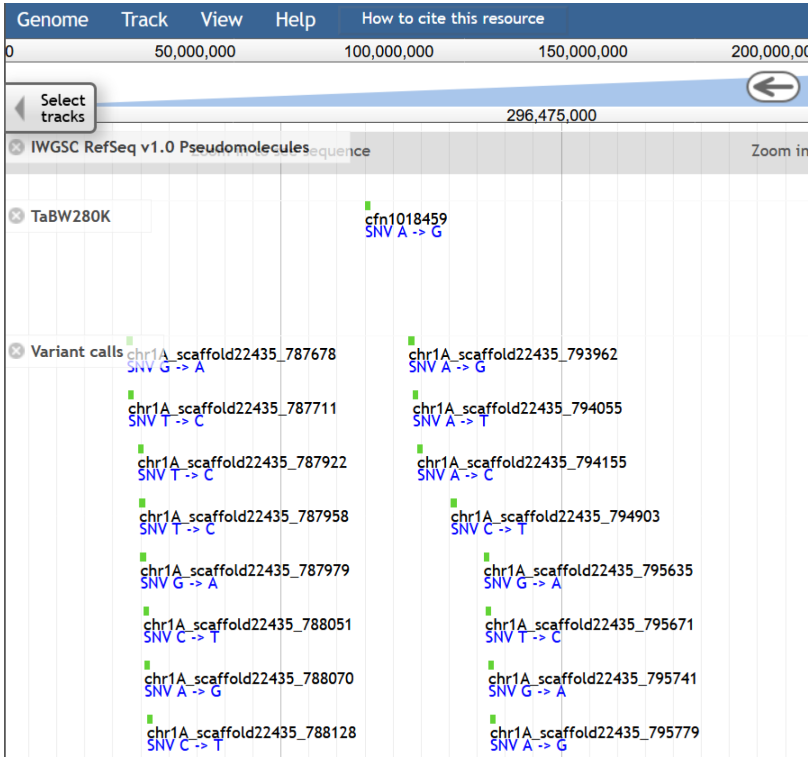
☐	Category	Name	Project	Publication	Description	Contact	Organism	Repository
☐	IWGSC Official Annotation / Gene Model / Annotation v1.0	LowConfidenceGenesv1.0	...	The International Wheat Genome Sequencing Consortium (IWGSC) Science 2018	...	IWGSC	Triticum aestivum Chinese Spring	Link to sequence repository
☐	IWGSC Official Annotation / Gene Model / Annotation v1.0	HighConfidenceGenesv1.0	...	The International Wheat Genome Sequencing Consortium (IWGSC) Science 2018	...	IWGSC	Triticum aestivum Chinese Spring	Link to sequence repository
☐	IWGSC Official Annotation / Gene Model / Annotation v1.1	LowConfidenceGenesv1.1	...	The International Wheat Genome Sequencing Consortium (IWGSC) Science 2018	...	IWGSC	Triticum aestivum Chinese Spring	Link to sequence repository
☑	IWGSC Official Annotation / Gene Model / Annotation v1.1	HighConfidenceGenesv1.1 (GOMAP annotation available)	...	The International Wheat Genome Sequencing Consortium (IWGSC) Science 2018	...	IWGSC	Triticum aestivum Chinese Spring	Link to sequence repository
☐	IWGSC Official Annotation / Gene Model / Manually curated genes	CRT-Binding Factors (CBF) gene familly	...	The International Wheat Genome Sequencing Consortium (IWGSC) Science 2018	...	IWGSC	Triticum aestivum Chinese Spring	Link to sequence repository
☐	IWGSC Official Annotation / Gene Model / Manually curated genes	Nucleotide-binding and Leucine-rich Repeat (NLR) gene familly	...	The International Wheat Genome Sequencing Consortium (IWGSC) Science 2018	...	IWGSC	Triticum aestivum Chinese Spring	Link to sequence repository
☐	IWGSC Official Annotation / Gene Model / Manually curated genes	Pentatricopeptide repeats (PPR) gene familly	...	The International Wheat Genome Sequencing Consortium (IWGSC) Science 2018	...	IWGSC	Triticum aestivum Chinese Spring	Link to sequence repository
☐	IWGSC Official Annotation / Gene Model / Manually curated genes	Prolamin and Globulin genes	...	The International Wheat Genome Sequencing Consortium (IWGSC) Science 2018	...	IWGSC	Triticum aestivum Chinese Spring	Link to sequence repository
☐	IWGSC Official Annotation / Gene Model / Manually curated genes	Amino acid transporters	...	The International Wheat Genome Sequencing Consortium (IWGSC) Science 2018	...	IWGSC	Triticum aestivum Chinese Spring	Link to sequence repository
☐	IWGSC Official Annotation / Gene Model / Manually curated genes	Cell Wall-Associated Kinase (WAK)	...	The International Wheat Genome Sequencing Consortium (IWGSC) Science 2018	...	IWGSC	Triticum aestivum Chinese Spring	Link to sequence repository
☐	IWGSC Official Annotation / Gene Model / Previous annotations	CSS	...	The International Wheat Genome Sequencing Consortium (IWGSC) Science 2018	...	IWGSC	Triticum aestivum Chinese Spring	Link to sequence repository
☐	IWGSC Official Annotation / Gene Model / Previous annotations	TGACv1	...	The International Wheat Genome Sequencing Consortium (IWGSC) Science 2018	...	IWGSC	Triticum aestivum Chinese Spring	Link to sequence repository
☑	IWGSC Official Annotation / RNA	lncRNA	...	The International Wheat Genome Sequencing Consortium (IWGSC) Science 2018	...	IWGSC	Triticum aestivum Chinese Spring	Link to sequence repository
☑	IWGSC Official Annotation / RNA	miRNA	...	The International Wheat Genome Sequencing Consortium (IWGSC)	...	IWGSC	Triticum aestivum Chinese Spring	Link to sequence repository

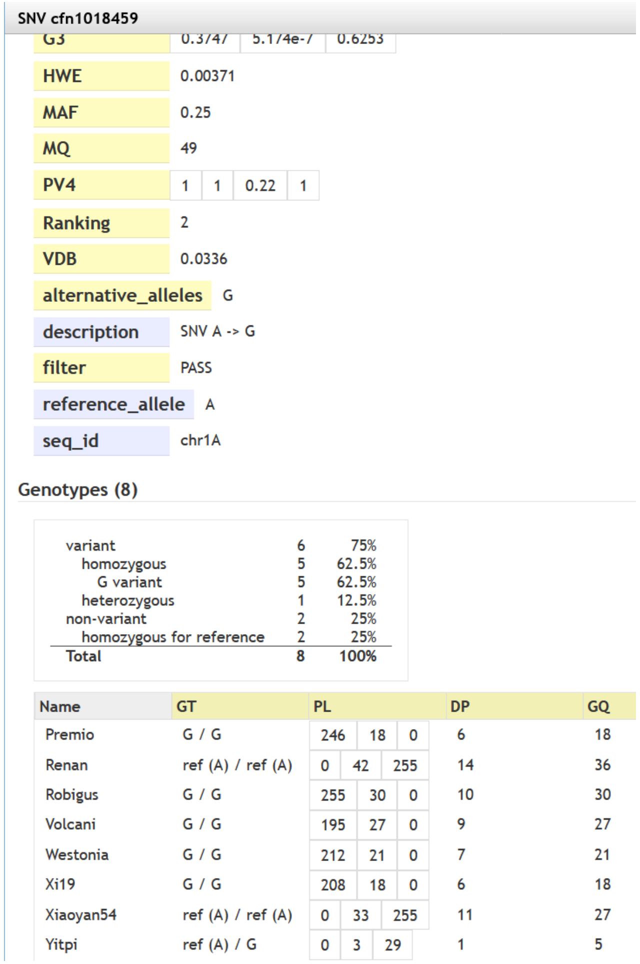
URGI

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

IWGSC RefSeq v1.0

- JBrowse
  - New tracks of variants, with MAF, genotypes, …





SNV cfn1018459

G3	0.3747	5.174e-7	0.6253
HWE	0.00371		
MAF	0.25		
MQ	49		
PV4	1	1	0.22	1
Ranking	2		
VDB	0.0336		
alternative_alleles	G		
description	SNV A -> G		
filter	PASS		
reference_allele	A		
seq_id	chr1A		

Genotypes (8)

variant	6	75%
homozygous	5	62.5%
G variant	5	62.5%
heterozygous	1	12.5%
non-variant	2	25%
homozygous for reference	2	25%
Total	**8**	**100%**

Name	GT	PL			DP	GQ
Premio	G / G	246	18	0	6	18
Renan	ref (A) / ref (A)	0	42	255	14	36
Robigus	G / G	255	30	0	10	30
Volcani	G / G	195	27	0	9	27
Westonia	G / G	212	21	0	7	21
Xi19	G / G	208	18	0	6	18
Xiaoyan54	ref (A) / ref (A)	0	33	255	11	27
Yitpi	ref (A) / G	0	3	29	1	5

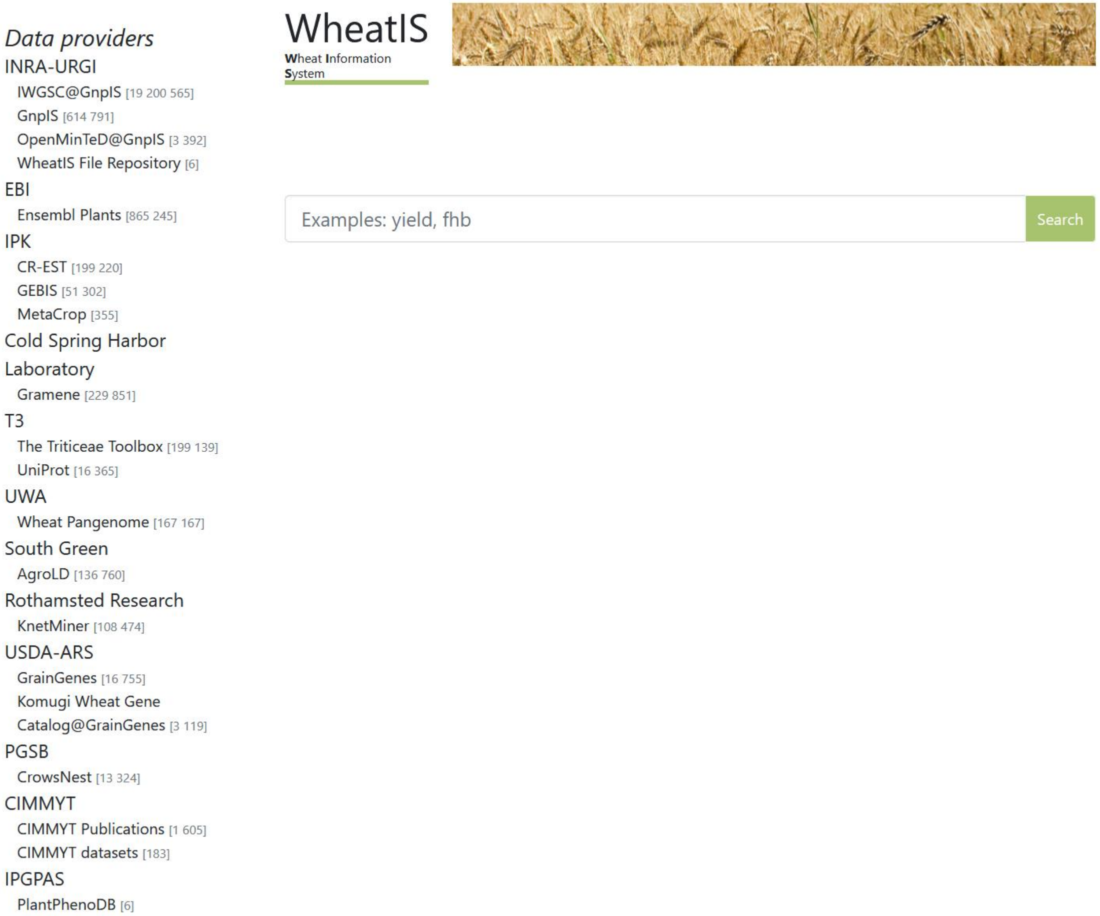
http://wheat-urgi.versailles.inra.fr/Seq-Repository

IWGSC RefSeq v1.0

OPEN ACCESS

- WheatIS data discovery
  - All the IWGSC open access data are searchable

IWGSC@GnpIS [19 200 565]



WheatIS
Wheat Information System

Data providers

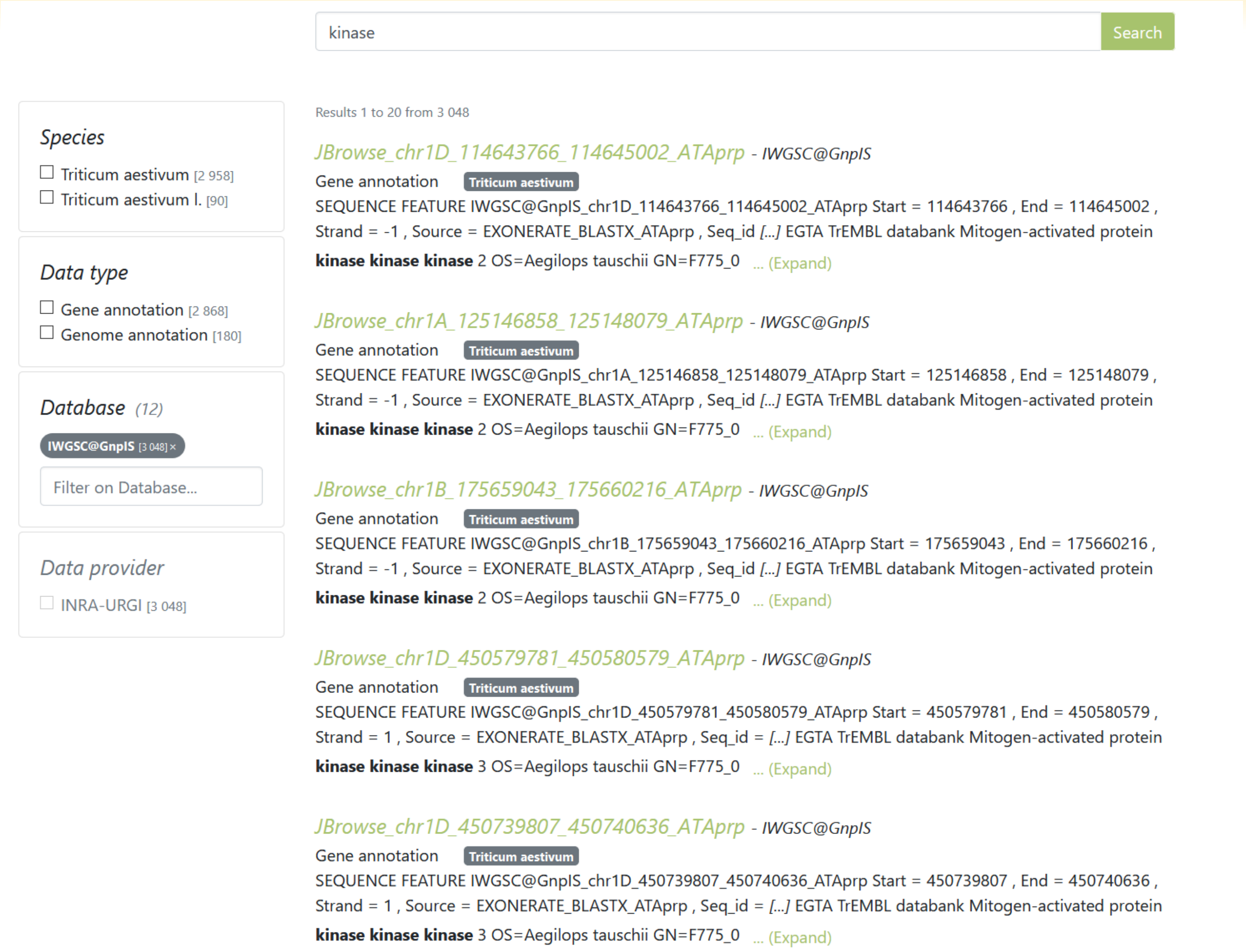
- INRA-URGI
  - IWGSC@GnpIS [19 200 565]
  - GnpIS [614 791]
  - OpenMinTeD@GnpIS [3 392]
  - WheatIS File Repository [6]
- EBI
  - Ensembl Plants [865 245]
- IPK
  - CR-EST [199 220]
  - GEBIS [51 302]
  - MetaCrop [355]
- Cold Spring Harbor Laboratory
  - Gramene [229 851]
- T3
  - The Triticeae Toolbox [199 139]
  - UniProt [16 365]
- UWA
  - Wheat Pangenome [167 167]
- South Green
  - AgroLD [136 760]
- Rothamsted Research
  - KnetMiner [108 474]
- USDA-ARS
  - GrainGenes [16 755]
  - Komugi Wheat Gene Catalog@GrainGenes [3 119]
- PGSB
  - CrowsNest [13 324]
- CIMMYT
  - CIMMYT Publications [1 605]
  - CIMMYT datasets [183]
- IPGPAS
  - PlantPhenoDB [6]

URGI

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

IWGSC RefSeq v1.0

OPEN ACCESS



kinase

Search

Species

☐ Triticum aestivum [2 958]
☐ Triticum aestivum l. [90]

Data type

☐ Gene annotation [2 868]
☐ Genome annotation [180]

Database (12)

IWGSC@GnpIS [3 048] ×

Filter on Database...

Data provider

☐ INRA-URGI [3 048]

Results 1 to 20 from 3 048

JBrowse_chr1D_114643766_114645002_ATAprp - *IWGSC@GnpIS*

Gene annotation **Triticum aestivum**

SEQUENCE FEATURE IWGSC@GnpIS_chr1D_114643766_114645002_ATAprp Start = 114643766 , End = 114645002 , Strand = -1 , Source = EXONERATE_BLASTX_ATAprp , Seq_id *[...]* EGTA TrEMBL databank Mitogen-activated protein **kinase kinase kinase** 2 OS=Aegilops tauschii GN=F775_0 ... (Expand)

JBrowse_chr1A_125146858_125148079_ATAprp - *IWGSC@GnpIS*

Gene annotation **Triticum aestivum**

SEQUENCE FEATURE IWGSC@GnpIS_chr1A_125146858_125148079_ATAprp Start = 125146858 , End = 125148079 , Strand = -1 , Source = EXONERATE_BLASTX_ATAprp , Seq_id *[...]* EGTA TrEMBL databank Mitogen-activated protein **kinase kinase kinase** 2 OS=Aegilops tauschii GN=F775_0 ... (Expand)

JBrowse_chr1B_175659043_175660216_ATAprp - *IWGSC@GnpIS*

Gene annotation **Triticum aestivum**

SEQUENCE FEATURE IWGSC@GnpIS_chr1B_175659043_175660216_ATAprp Start = 175659043 , End = 175660216 , Strand = -1 , Source = EXONERATE_BLASTX_ATAprp , Seq_id *[...]* EGTA TrEMBL databank Mitogen-activated protein **kinase kinase kinase** 2 OS=Aegilops tauschii GN=F775_0 ... (Expand)

JBrowse_chr1D_450579781_450580579_ATAprp - *IWGSC@GnpIS*

Gene annotation **Triticum aestivum**

SEQUENCE FEATURE IWGSC@GnpIS_chr1D_450579781_450580579_ATAprp Start = 450579781 , End = 450580579 , Strand = 1 , Source = EXONERATE_BLASTX_ATAprp , Seq_id = *[...]* EGTA TrEMBL databank Mitogen-activated protein **kinase kinase kinase** 3 OS=Aegilops tauschii GN=F775_0 ... (Expand)

JBrowse_chr1D_450739807_450740636_ATAprp - *IWGSC@GnpIS*

Gene annotation **Triticum aestivum**

SEQUENCE FEATURE IWGSC@GnpIS_chr1D_450739807_450740636_ATAprp Start = 450739807 , End = 450740636 , Strand = 1 , Source = EXONERATE_BLASTX_ATAprp , Seq_id = *[...]* EGTA TrEMBL databank Mitogen-activated protein **kinase kinase kinase** 3 OS=Aegilops tauschii GN=F775_0 ... (Expand)

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

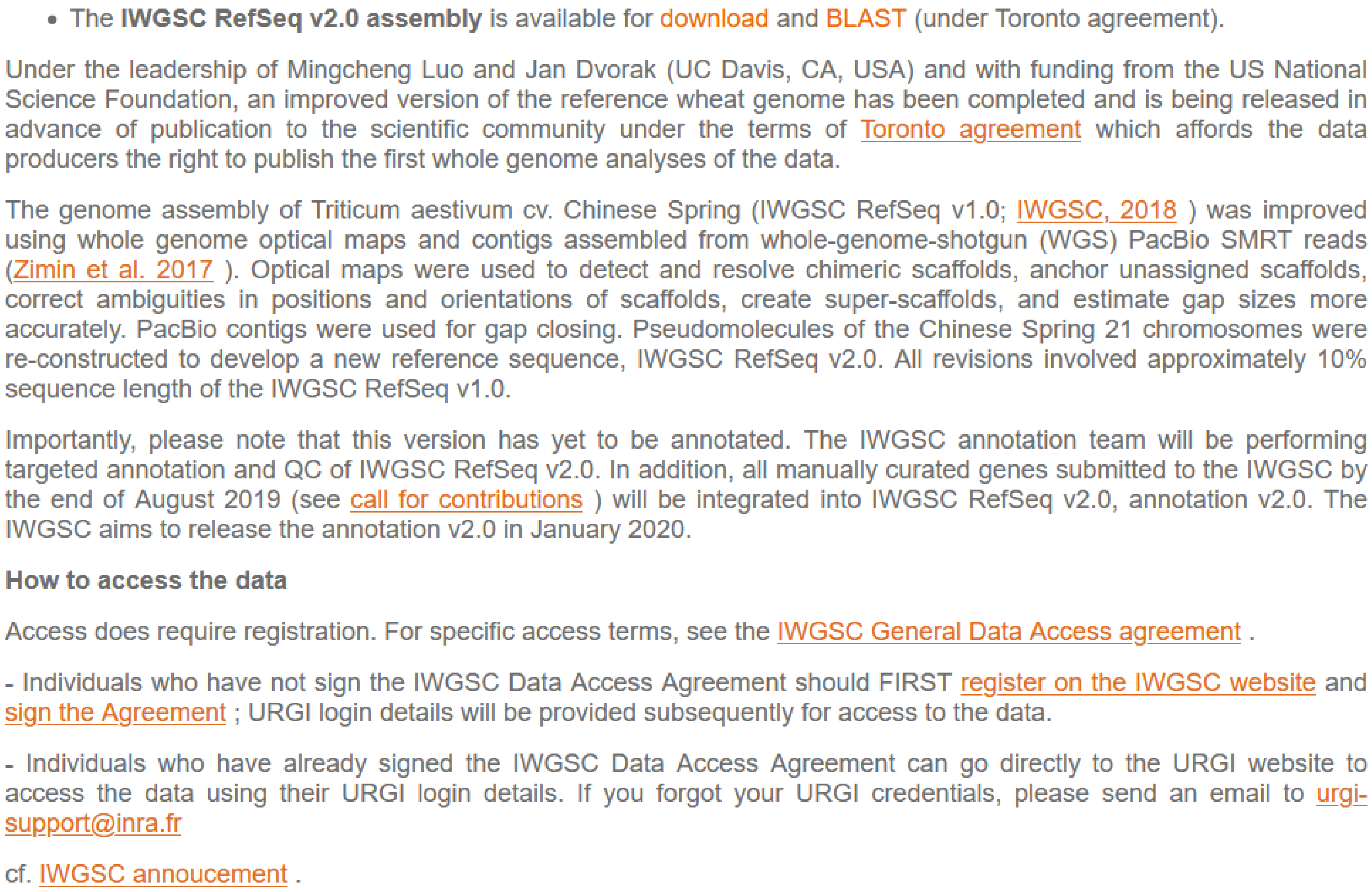
New registered access data

- Sonmez genome: elite Turkish cultivar (H. Budak)
- Exome data from 47 lines (H. Budak)

Early access for IWGSC Sponsors during 6 months → available in open access in June 2020.

IWGSC RefSeq v2.0

- Data under the terms of the Toronto agreement



- The **IWGSC RefSeq v2.0 assembly** is available for download and BLAST (under Toronto agreement).

Under the leadership of Mingcheng Luo and Jan Dvorak (UC Davis, CA, USA) and with funding from the US National Science Foundation, an improved version of the reference wheat genome has been completed and is being released in advance of publication to the scientific community under the terms of Toronto agreement which affords the data producers the right to publish the first whole genome analyses of the data.

The genome assembly of Triticum aestivum cv. Chinese Spring (IWGSC RefSeq v1.0; IWGSC, 2018) was improved using whole genome optical maps and contigs assembled from whole-genome-shotgun (WGS) PacBio SMRT reads (Zimin et al. 2017). Optical maps were used to detect and resolve chimeric scaffolds, anchor unassigned scaffolds, correct ambiguities in positions and orientations of scaffolds, create super-scaffolds, and estimate gap sizes more accurately. PacBio contigs were used for gap closing. Pseudomolecules of the Chinese Spring 21 chromosomes were re-constructed to develop a new reference sequence, IWGSC RefSeq v2.0. All revisions involved approximately 10% sequence length of the IWGSC RefSeq v1.0.

Importantly, please note that this version has yet to be annotated. The IWGSC annotation team will be performing targeted annotation and QC of IWGSC RefSeq v2.0. In addition, all manually curated genes submitted to the IWGSC by the end of August 2019 (see call for contributions) will be integrated into IWGSC RefSeq v2.0, annotation v2.0. The IWGSC aims to release the annotation v2.0 in January 2020.

How to access the data

Access does require registration. For specific access terms, see the IWGSC General Data Access agreement .

- Individuals who have not sign the IWGSC Data Access Agreement should FIRST register on the IWGSC website and sign the Agreement ; URGI login details will be provided subsequently for access to the data.

- Individuals who have already signed the IWGSC Data Access Agreement can go directly to the URGI website to access the data using their URGI login details. If you forgot your URGI credentials, please send an email to urgi-support@inra.fr

cf. IWGSC annoucement .

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

IWGSC RefSeq v2.0

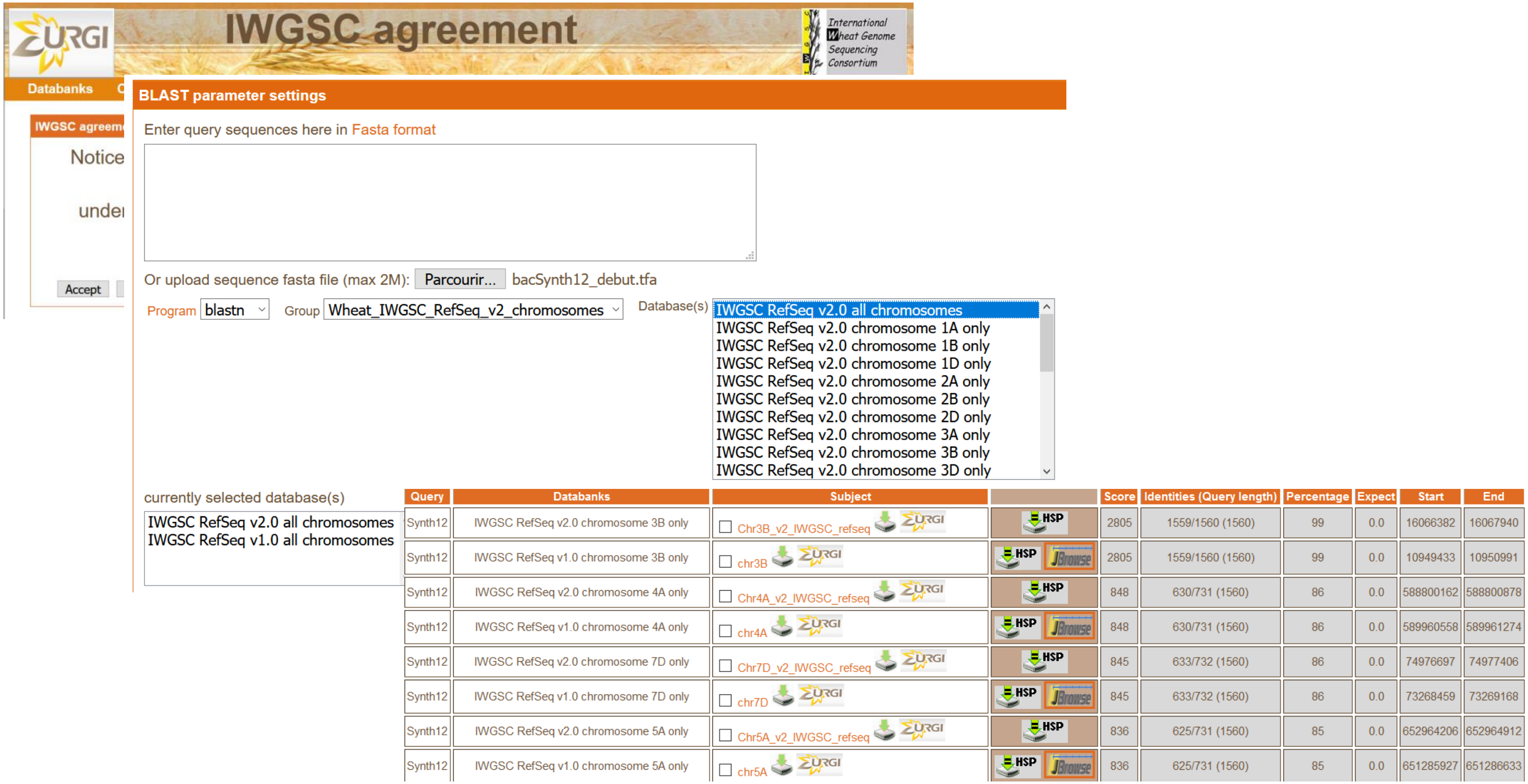
- Download assembly

[DIR]	Parent Directory		-
[ ]	iwgsc_refseqv2.0_README.pdf	18-Jul-2019 16:56	119K
[ ]	iwgsc_refseqv2.0_all_chromosomes.fa.zip	18-Jul-2019 16:23	4.1G
[TXT]	iwgsc_refseqv2.0_all_chromosomes.fa.zip.md5.txt	18-Jul-2019 16:23	73
[ ]	iwgsc_refseqv2.0_chr1A.fa.zip	18-Jul-2019 15:13	171M
[TXT]	iwgsc_refseqv2.0_chr1A.fa.zip.md5.txt	18-Jul-2019 15:13	63
[ ]	iwgsc_refseqv2.0_chr1B.fa.zip	18-Jul-2019 15:14	200M
[TXT]	iwgsc_refseqv2.0_chr1B.fa.zip.md5.txt	18-Jul-2019 15:14	63
[ ]	iwgsc_refseqv2.0_chr1D.fa.zip	18-Jul-2019 15:14	142M
[TXT]	iwgsc_refseqv2.0_chr1D.fa.zip.md5.txt	18-Jul-2019 15:14	63
[ ]	iwgsc_refseqv2.0_chr2A.fa.zip	18-Jul-2019 15:15	226M
[TXT]	iwgsc_refseqv2.0_chr2A.fa.zip.md5.txt	18-Jul-2019 15:15	63
[ ]	iwgsc_refseqv2.0_chr2B.fa.zip	18-Jul-2019 15:16	233M
[TXT]	iwgsc_refseqv2.0_chr2B.fa.zip.md5.txt	18-Jul-2019 15:16	63
[ ]	iwgsc_refseqv2.0_chr2D.fa.zip	18-Jul-2019 15:16	188M
[TXT]	iwgsc_refseqv2.0_chr2D.fa.zip.md5.txt	18-Jul-2019 15:16	63

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

IWGSC RefSeq v2.0

- BLAST: currently link to Jbrowse only for V1



BLAST parameter settings

Enter query sequences here in Fasta format

Or upload sequence fasta file (max 2M): Parcourir... bacSynth12_debut.tfa

Program blastn Group Wheat_IWGSC_RefSeq_v2_chromosomes Database(s)

IWGSC RefSeq v2.0 all chromosomes
IWGSC RefSeq v2.0 chromosome 1A only
IWGSC RefSeq v2.0 chromosome 1B only
IWGSC RefSeq v2.0 chromosome 1D only
IWGSC RefSeq v2.0 chromosome 2A only
IWGSC RefSeq v2.0 chromosome 2B only
IWGSC RefSeq v2.0 chromosome 2D only
IWGSC RefSeq v2.0 chromosome 3A only
IWGSC RefSeq v2.0 chromosome 3B only
IWGSC RefSeq v2.0 chromosome 3D only

currently selected database(s)

IWGSC RefSeq v2.0 all chromosomes
IWGSC RefSeq v1.0 all chromosomes

Query	Databanks	Subject		Score	Identities (Query length)	Percentage	Expect	Start	End
Synth12	IWGSC RefSeq v2.0 chromosome 3B only	☐ Chr3B_v2_IWGSC_refseq	HSP	2805	1559/1560 (1560)	99	0.0	16066382	16067940
Synth12	IWGSC RefSeq v1.0 chromosome 3B only	☐ chr3B	HSP JBrowse	2805	1559/1560 (1560)	99	0.0	10949433	10950991
Synth12	IWGSC RefSeq v2.0 chromosome 4A only	☐ Chr4A_v2_IWGSC_refseq	HSP	848	630/731 (1560)	86	0.0	588800162	588800878
Synth12	IWGSC RefSeq v1.0 chromosome 4A only	☐ chr4A	HSP JBrowse	848	630/731 (1560)	86	0.0	589960558	589961274
Synth12	IWGSC RefSeq v2.0 chromosome 7D only	☐ Chr7D_v2_IWGSC_refseq	HSP	845	633/732 (1560)	86	0.0	74976697	74977406
Synth12	IWGSC RefSeq v1.0 chromosome 7D only	☐ chr7D	HSP JBrowse	845	633/732 (1560)	86	0.0	73268459	73269168
Synth12	IWGSC RefSeq v2.0 chromosome 5A only	☐ Chr5A_v2_IWGSC_refseq	HSP	836	625/731 (1560)	85	0.0	652964206	652964912
Synth12	IWGSC RefSeq v1.0 chromosome 5A only	☐ chr5A	HSP JBrowse	836	625/731 (1560)	85	0.0	651285927	651286633

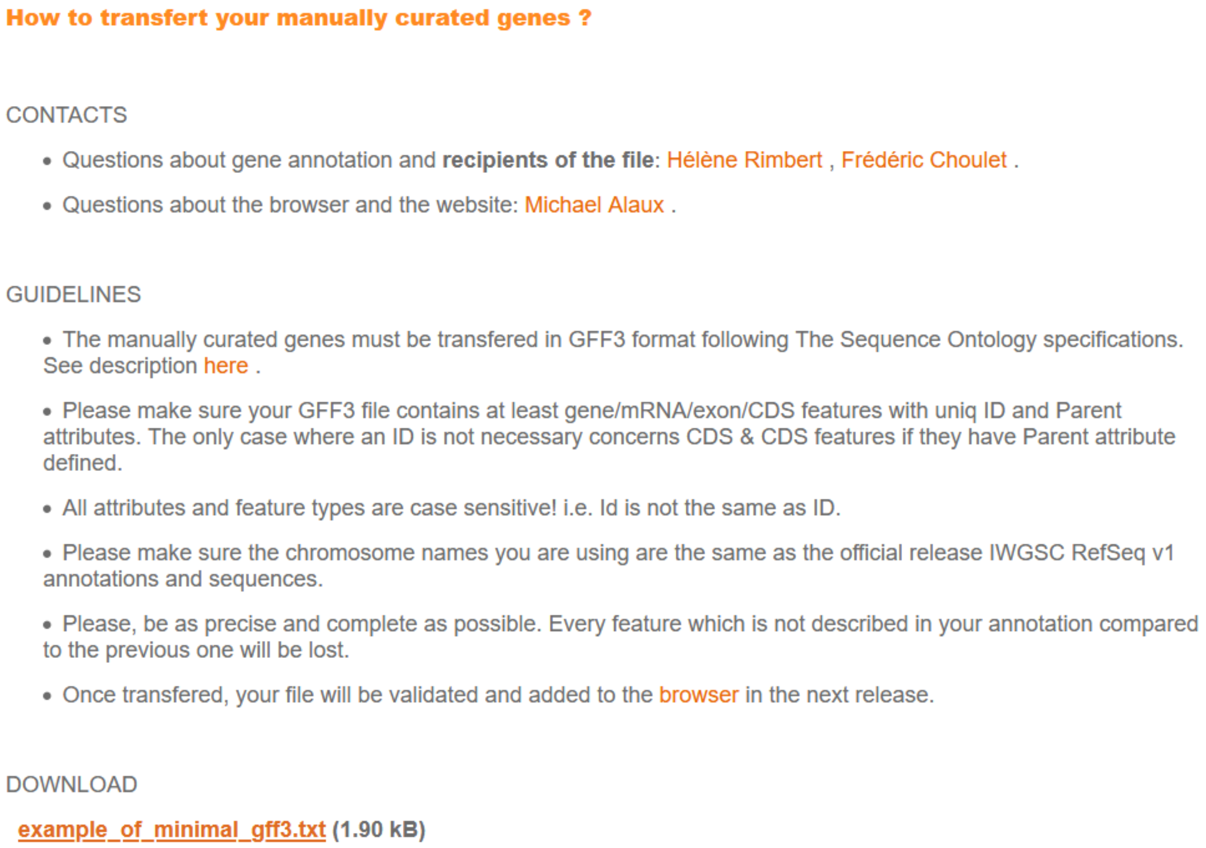
http://wheat-urgi.versailles.inra.fr/Seq-Repository

IWGSC RefSeq v2.0

- Annotation v2.0 to come (F. Choulet)
- A JBrowse will be set-up
    - Linked with the BLAST
    - Linked with the IWGSC RefSeq v1.0

IWGSC RefSeq v2.0

- Manual curation
    - Some teams already sent their manual curations of gene families to GDEC following the process decribed below:



How to transfert your manually curated genes ?

CONTACTS

- Questions about gene annotation and **recipients of the file**: Hélène Rimbert , Frédéric Choulet .
- Questions about the browser and the website: Michael Alaux .

GUIDELINES

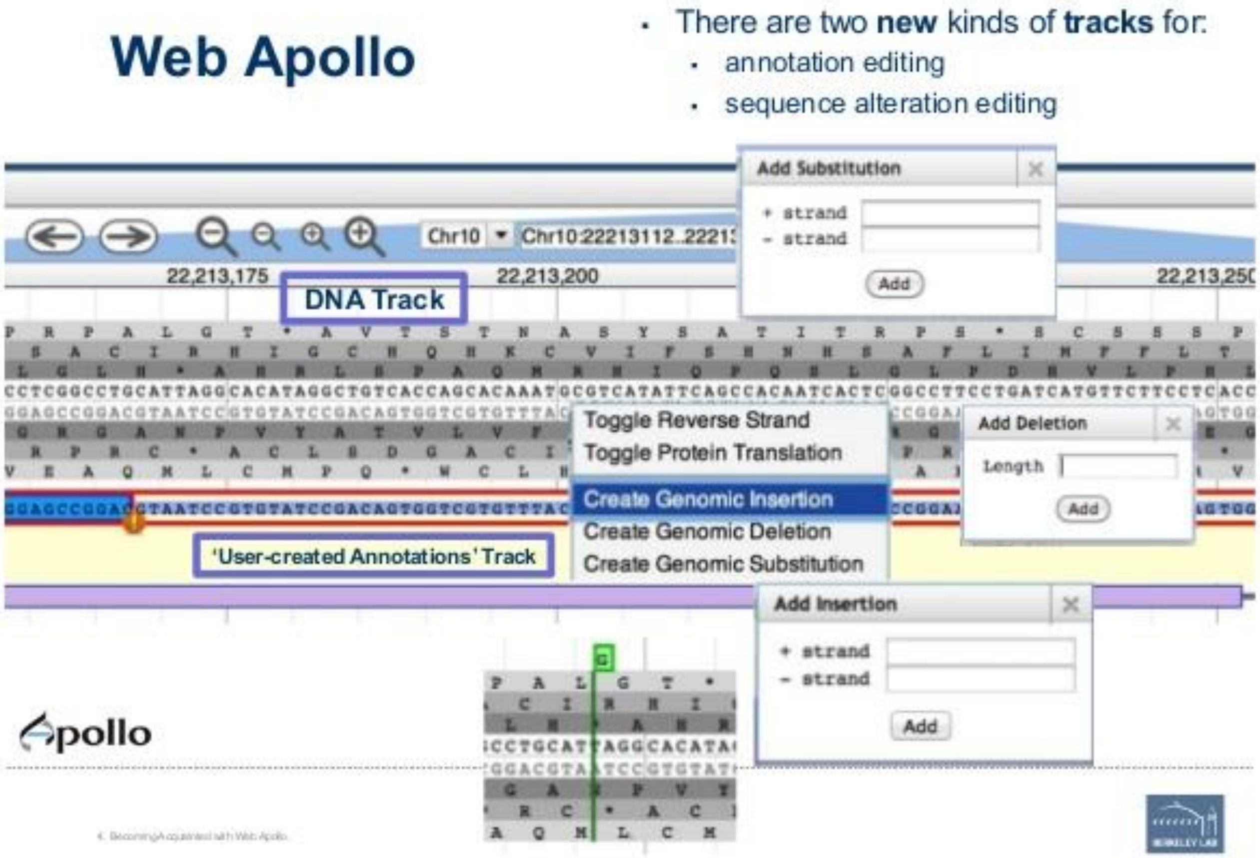
- The manually curated genes must be transfered in GFF3 format following The Sequence Ontology specifications. See description here .
- Please make sure your GFF3 file contains at least gene/mRNA/exon/CDS features with uniq ID and Parent attributes. The only case where an ID is not necessary concerns CDS & CDS features if they have Parent attribute defined.
- All attributes and feature types are case sensitive! i.e. Id is not the same as ID.
- Please make sure the chromosome names you are using are the same as the official release IWGSC RefSeq v1 annotations and sequences.
- Please, be as precise and complete as possible. Every feature which is not described in your annotation compared to the previous one will be lost.
- Once transfered, your file will be validated and added to the browser in the next release.

DOWNLOAD

example_of_minimal_gff3.txt (1.90 kB)

IWGSC RefSeq v2.0

- If other people are still interested to do manual curation, we can facilitate this process by settiing-up an Apollo tool



Web Apollo

- There are two **new** kinds of **tracks** for:
  - annotation editing
  - sequence alteration editing

Acknowledgements

URGI

Alaux M.
Flores R.
Letellier T.
Alfama F.
Pommier C.
Mohellibi N.
Kimmel E.
Michotey C.
Francillonne N.
Quesneville H.
Adam-Blondon A.F.

International Wheat Genome Sequencing Consortium

Eversole K.
Caugant I.

IWGSC Coordinating Committee

IWGSC Sequencing and Analysis team

All data providers

GDEC
Génétique Diversité Ecophysiologie des Céréales

Choulet F.
Rimbert H.
Leroy P.
Salse J.
Paux E.

INRAE Breed wheat Whealbi WHEAT INITIATIVE RDA RESEARCH DATA ALLIANCE openMINTED Open Mining Infrastructure for Text & Data elixir

ANR INVESTISSEMENTS D'AVENIR

URGI

Questions

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

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

Linking the International Wheat Genome Sequencin…

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

genomebiology.biomedcentral.com