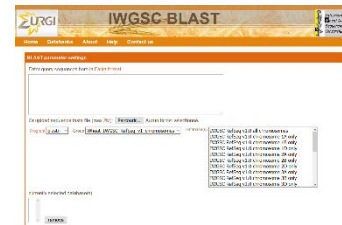


The IWGSC data repository and wheat data resources hosted at URGI: overview and perspectives

Michael Alaux





IWGSC Data Repository Overview

IWGSC Data Repository

The screenshot displays the IWGSC Data Repository website. On the left is a vertical navigation menu with orange buttons for: Create an account, News, Access Status, Assemblies, Annotations, BLAST, BAC Libraries, Physical maps, Genetic maps, 3B reference sequence, Expression, Variations, Publication (IWGSC), and FAQ and support. The main content area is titled 'Seq Repository' and features the IWGSC logo and a text prompt: 'Click on a chromosome to download, BLAST or display the sequences. News, Access status, etc. are detailed in the left menu.' Below this are chromosome icons labeled 1A, 2A, 3A, 4A, 5A, 6A, 7A in the top row and 1B, 2B, 3B in the bottom row. A browser window is overlaid on the 3A chromosome, showing the URL 'https://wheat-urgi.versailles.inra.fr/Chr3B' and the 'WHEAT URGI' logo. The browser content displays 'Chr3B' and a list of links for 'Chromosome 3B': Public BLAST, IWGSC BLAST, Download assemblies, Download annotations, Display IWGSC RefSeq v1.0 browser, Display physical maps browser, and Order BAC clones (TaaCsp3BFhA, TaaCsp3BFhB).

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

Assemblies

- **IWGSC RefSeq v1.0 Assembly**

- The [IWGSC RefSeq v1.0](#) (the first version of the reference sequence of the bread wheat variety *Chinese Spring*) assembly is now available for [download](#) , [BLAST](#) and display in a [browser](#) .

The pre-publication data are being made available under the IWGSC [General Data Access Agreement](#) which is consistent with the [Toronto Agreement](#) and that grants the IWGSC the right to publish the first global analyses of the data. This includes descriptions of whole chromosome or genome-level analyses of genes, gene families, repetitive elements, and comparisons with other organisms.

The **IWGSC RefSeq v1.0 assembly** is an integration of the IWGSC WGA v0.4 – made available in June 2016 – with IWGSC chromosome-based and other resources, including but not limited to:

- Physical maps for all chromosomes;
- Sequenced BACs for 8 chromosomes (1A, 1B, 3B, 3D, 6B, 7A, 7B, 7D) and partial MTP BAC sequences for 2 chromosome arms (4AL, 5BS);
- MTP BAC WGPTM sequence tags for all chromosomes, except 3B;
- BioNano optical maps (7A, 7B, 7DS);
- Alignment to RH maps (D chromosomes); and
- GBS map of the SynOp RIL population CsXRn genetic map (INRA).

With the addition of the resources that have been developed by IWGSC members over the past few years, the quality of the assembly increased substantially. When compared with IWGSC WGA v0.4, the chromosomal scaffold/ superscaffold N50 increased from 7.0 Mb to 22.8 Mb.

The data are available for [BLAST searches](#) and can be [downloaded](#) .

Assemblies

- Other assemblies available:
 - IWGSC WGA v0.4
 - IWGSC survey sequence (all versions)
 - TGAC v1
 - Other wheat species

Annotations

- **IWGSC RefSeq v1.0 Annotation**

- **IWGSC RefSeq v1.0 annotation** is available for [download](#) and display in a [browser](#) .

The pre-publication data are being made available under the IWGSC [General Data Access Agreement](#) which is consistent with the [Toronto Agreement](#) and that grants the IWGSC the right to publish the first global analyses of the data. This includes descriptions of whole chromosome or genome-level analyses of genes, gene families, repetitive elements, and comparisons with other organisms.

The IWGSC RefSeq v1.0 annotation includes gene models generated by integrating predictions made by INRA-GDEC using Triannot and PGSB using their customised pipeline (previously MIPS pipeline). The integration was undertaken by the Earlham institute (EI), who have also added UTRs to the gene models where supporting data are available. Gene models have been assigned to high confidence (HC) or low confidence (LC) classes based on completeness, similarity to genes represented in protein and DNA databases and repeat content. The automated assignment of functional annotation to genes has been generated by PGSB based on AHRD parameters. In addition, annotated transposable elements (TEs) and non-coding RNAs are available. More information about the annotation data is provided in the [README file](#) .

How to access the data?

Access does require registration and agreeing to respect the right of the IWGSC to publish first. For specific access terms, see the [IWGSC General Data Access agreement](#) .

- Individuals who have not signed the IWGSC Data Access Agreement should **FIRST** [register on the IWGSC website](#) and sign the Agreement; URGI login details will be provided subsequently by email for access to the data. Typically, this it will take no more than 2 business days for your URGI account to be established but occasionally it make take up to a week.
- 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.

Annotations

- Data available:
 - Genes: HighConf and LowConf
 - Functional annotation
 - Transposable elements
 - Markers: ISBP, SNP, DArT, SSR, EST, etc.
 - ncRNAs: miRNA, lncRNA
 - RH maps
 - GBS maps
 - Optical maps

Annotations

- IWGSC RefSeq v1.1 Annotation
 - Refers to the same assembly: the IWGSC RefSeq v1.0 Assembly
 - Data will be available upon publication:
 - Genes: HighConf and LowConf
 - RNA-seq mapping

Access to the IWGSC RefSeq v1.0 data

- Under Toronto agreement (IWGSC general access agreement)

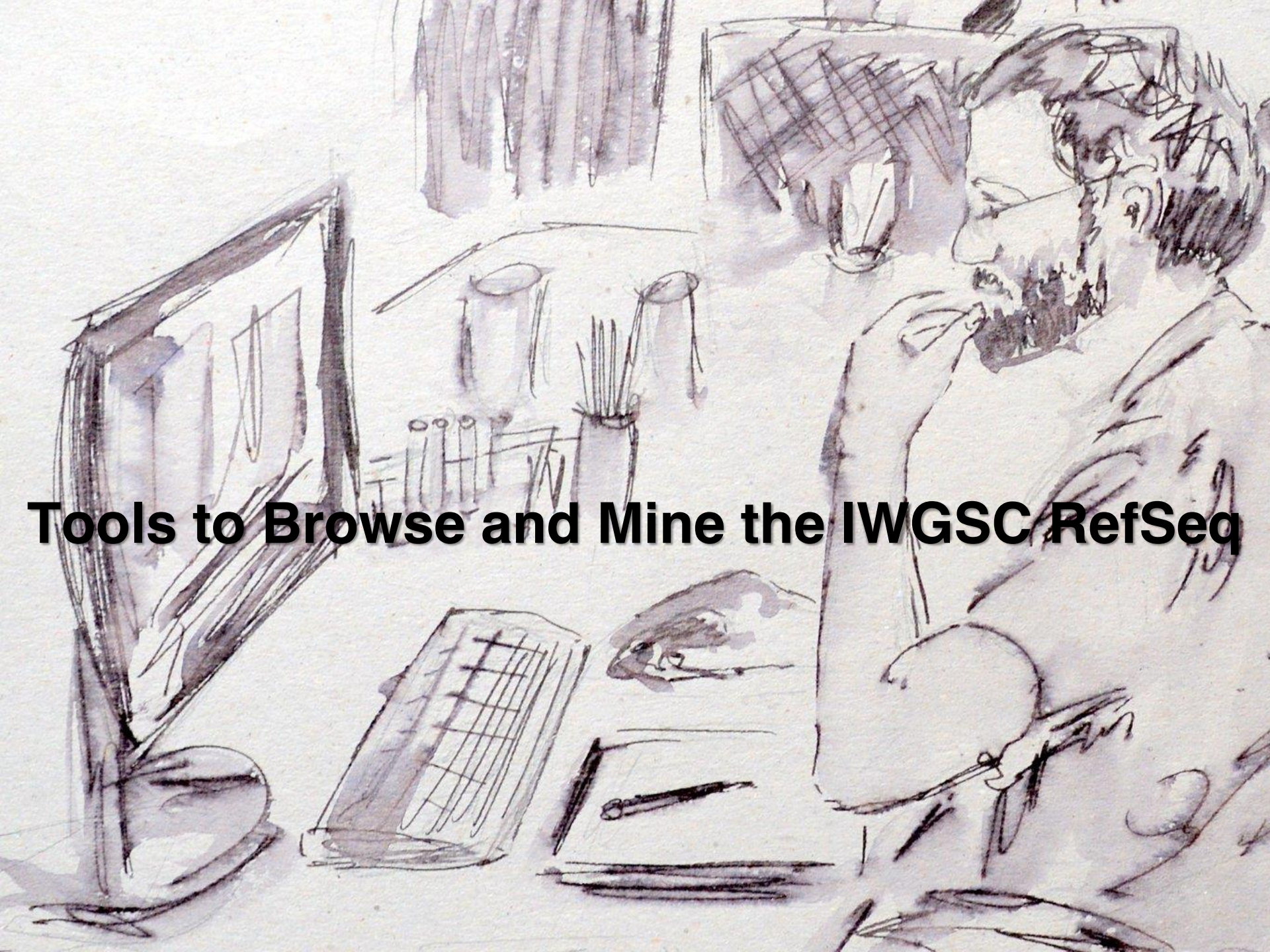
<http://www.wheatgenome.org/Tools-and-Resources>

→ will be in open access once published

How to access IWGSC RefSeq v1.0 data?

Access does require registration and agreeing to respect the right of the IWGSC to publish first. For specific access terms, see the [IWGSC General Data Access agreement](#) .

- Individuals who have not signed the IWGSC Data Access Agreement should **FIRST** [register on the IWGSC website](#) and sign the Agreement; URGI login details will be provided subsequently by email for access to the data. Typically, this it will take no more than 2 business days for your URGI account to be established but occasionally it make take up to a week.
- 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.



Tools to Browse and Mine the IWGSC RefSeq

Download

IWGSC RefSeq 1.0 assembly

[]	iwgsc_refseqv1.0_all_chromosomes.zip	04-May-2017 12:31	8.0G	[]	iwgsc_refseqv1.0_chr5A.fsa.zip	04-May-2017 12:23	202M
[TXT]	iwgsc_refseqv1.0_all_chromosomes.zip.md5.txt	05-Dec-2017 15:00	71	[TXT]	iwgsc_refseqv1.0_chr5A.fsa.zip.md5.txt	05-Dec-2017 15:00	65
[]	iwgsc_refseqv1.0_chr1A.fsa.zip	04-May-2017 12:14	169M	[]	iwgsc_refseqv1.0_chr5B.fsa.zip	04-May-2017 12:24	203M
[TXT]	iwgsc_refseqv1.0_chr1A.fsa.zip.md5.txt	05-Dec-2017 15:00	65	[TXT]	iwgsc_refseqv1.0_chr5B.fsa.zip.md5.txt	05-Dec-2017 15:00	65
[]	iwgsc_refseqv1.0_chr1B.fsa.zip	04-May-2017 12:15	197M	[]	iwgsc_refseqv1.0_chr5D.fsa.zip	04-May-2017 12:25	160M
[TXT]	iwgsc_refseqv1.0_chr1B.fsa.zip.md5.txt	05-Dec-2017 15:00	65	[TXT]	iwgsc_refseqv1.0_chr5D.fsa.zip.md5.txt	05-Dec-2017 15:00	65
[]	iwgsc_refseqv1.0_chr1D.fsa.zip	04-May-2017 12:15	140M	[]	iwgsc_refseqv1.0_chr6A.fsa.zip	04-May-2017 12:25	176M
[TXT]	iwgsc_refseqv1.0_chr1D.fsa.zip.md5.txt	05-Dec-2017 15:00	65	[TXT]	iwgsc_refseqv1.0_chr6A.fsa.zip.md5.txt	05-Dec-2017 15:00	65
[]	iwgsc_refseqv1.0_chr2A.fsa.zip	04-May-2017 12:16	222M	[]	iwgsc_refseqv1.0_chr6B.fsa.zip	04-May-2017 12:26	205M
[TXT]	iwgsc_refseqv1.0_chr2A.fsa.zip.md5.txt	05-Dec-2017 15:00	65	[TXT]	iwgsc_refseqv1.0_chr6B.fsa.zip.md5.txt	05-Dec-2017 15:00	65
[]	iwgsc_refseqv1.0_chr2B.fsa.zip	04-May-2017 12:17	229M	[]	iwgsc_refseqv1.0_chr6D.fsa.zip	04-May-2017 12:27	134M
[TXT]	iwgsc_refseqv1.0_chr2B.fsa.zip.md5.txt	05-Dec-2017 15:00	65	[TXT]	iwgsc_refseqv1.0_chr6D.fsa.zip.md5.txt	05-Dec-2017 15:00	65
[]	iwgsc_refseqv1.0_chr2D.fsa.zip	04-May-2017 12:18	185M	[]	iwgsc_refseqv1.0_chr7A.fsa.zip	04-May-2017 12:27	209M
[TXT]	iwgsc_refseqv1.0_chr2D.fsa.zip.md5.txt	05-Dec-2017 15:00	65	[TXT]	iwgsc_refseqv1.0_chr7A.fsa.zip.md5.txt	05-Dec-2017 15:00	65
[]	iwgsc_refseqv1.0_chr3A.fsa.zip	04-May-2017 12:18	214M	[]	iwgsc_refseqv1.0_chr7B.fsa.zip	04-May-2017 12:28	214M
[TXT]	iwgsc_refseqv1.0_chr3A.fsa.zip.md5.txt	05-Dec-2017 15:00	65	[TXT]	iwgsc_refseqv1.0_chr7B.fsa.zip.md5.txt	05-Dec-2017 15:00	65
[]	iwgsc_refseqv1.0_chr3B.fsa.zip	04-May-2017 12:19	237M	[]	iwgsc_refseqv1.0_chr7D.fsa.zip	04-May-2017 12:29	181M
[TXT]	iwgsc_refseqv1.0_chr3B.fsa.zip.md5.txt	05-Dec-2017 15:00	65	[TXT]	iwgsc_refseqv1.0_chr7D.fsa.zip.md5.txt	05-Dec-2017 15:00	65
[]	iwgsc_refseqv1.0_chr3D.fsa.zip	04-May-2017 12:20	175M	[]	iwgsc_refseqv1.0_chrUn.fsa.zip	04-May-2017 12:29	118M
[TXT]	iwgsc_refseqv1.0_chr3D.fsa.zip.md5.txt	05-Dec-2017 15:00	65	[TXT]	iwgsc_refseqv1.0_chrUn.fsa.zip.md5.txt	05-Dec-2017 15:00	65
[]	iwgsc_refseqv1.0_chr4A.fsa.zip	04-May-2017 12:21	212M				
[TXT]	iwgsc_refseqv1.0_chr4A.fsa.zip.md5.txt	05-Dec-2017 15:00	65				
[]	iwgsc_refseqv1.0_chr4B.fsa.zip	04-May-2017 12:22	192M				
[TXT]	iwgsc_refseqv1.0_chr4B.fsa.zip.md5.txt	05-Dec-2017 15:00	65				
[]	iwgsc_refseqv1.0_chr4D.fsa.zip	04-May-2017 12:22	144M				
[TXT]	iwgsc_refseqv1.0_chr4D.fsa.zip.md5.txt	05-Dec-2017 15:00	65				

Download

IWGSC RefSeq v1.0 annotation

[]	iwgsc_refseqv1.0_FunctionalAnnotation_v1.zip	06-Dec-2017 14:49	19M	[]	iwgsc_refseqv1.0_Marker_mapping_summary_2017Mar13.zip	24-Mar-2017 17:27	39M
[TXT]	iwgsc_refseqv1.0_FunctionalAnnotation_v1.zip.md5.txt	06-Dec-2017 14:49	82	[TXT]	iwgsc_refseqv1.0_Marker_mapping_summary_2017Mar13.zip.md5.txt	05-Dec-2017 15:01	88
[]	iwgsc_refseqv1.0_HighConf_2017Mar13.gff3.zip	24-Mar-2017 17:26	11M	[]	iwgsc_refseqv1.0_PGSB_annotation_files.zip	04-Dec-2017 18:04	252M
[TXT]	iwgsc_refseqv1.0_HighConf_2017Mar13.gff3.zip.md5.txt	05-Dec-2017 15:01	79	[TXT]	iwgsc_refseqv1.0_PGSB_annotation_files.zip.md5.txt	05-Dec-2017 15:01	75
[]	iwgsc_refseqv1.0_HighConf_CDS_2017Mar13.fa.zip	24-Mar-2017 17:27	43M	[]	iwgsc_refseqv1.0_README.pdf	05-Jan-2018 11:58	311K
[TXT]	iwgsc_refseqv1.0_HighConf_CDS_2017Mar13.fa.zip.md5.txt	05-Dec-2017 15:01	81	[]	iwgsc_refseqv1.0_RH_map.zip	27-Nov-2017 17:38	2.1M
[]	iwgsc_refseqv1.0_HighConf_PROTEIN_2017Mar13.fa.zip	24-Mar-2017 17:27	26M	[TXT]	iwgsc_refseqv1.0_RH_map.zip.md5.txt	05-Dec-2017 15:01	63
[TXT]	iwgsc_refseqv1.0_HighConf_PROTEIN_2017Mar13.fa.zip.md5.txt	05-Dec-2017 15:01	85	[]	iwgsc_refseqv1.0_SynOpRIL993_GBS.zip	09-Oct-2017 11:19	485K
[]	iwgsc_refseqv1.0_HighConf_REPR_CDS_2017Apr03.fa.zip	11-Apr-2017 15:45	40M	[TXT]	iwgsc_refseqv1.0_SynOpRIL993_GBS.zip.md5.txt	05-Dec-2017 15:01	69
[TXT]	iwgsc_refseqv1.0_HighConf_REPR_CDS_2017Apr03.fa.zip.md5.txt	05-Dec-2017 15:01	86	[]	iwgsc_refseqv1.0_TransposableElements_2017Mar13.gff3.zip	24-Mar-2017 17:27	139M
[]	iwgsc_refseqv1.0_HighConf_REPR_PROTEIN_2017Apr03.fa.zip	11-Apr-2017 15:45	25M	[TXT]	iwgsc_refseqv1.0_TransposableElements_2017Mar13.gff3.zip.md5.txt	05-Dec-2017 15:01	91
[TXT]	iwgsc_refseqv1.0_HighConf_REPR_PROTEIN_2017Apr03.fa.zip.md5.txt	05-Dec-2017 15:01	90	[]	iwgsc_refseqv1.0_WheatIbI_GWAS.zip	06-Dec-2017 11:10	226M
[]	iwgsc_refseqv1.0_HighConf_UTR_2017May05.gff3.zip	09-May-2017 09:53	18M	[TXT]	iwgsc_refseqv1.0_WheatIbI_GWAS.zip.md5.txt	06-Dec-2017 11:10	66
[TXT]	iwgsc_refseqv1.0_HighConf_UTR_2017May05.gff3.zip.md5.txt	05-Dec-2017 15:01	83	[]	iwgsc_refseqv1.0_lncRNA_2017Dec13.gff3.zip	19-Dec-2017 16:53	112M
[]	iwgsc_refseqv1.0_ISBP.fa.zip	03-Jan-2018 18:06	209M	[TXT]	iwgsc_refseqv1.0_lncRNA_2017Dec13.gff3.zip.md5.txt	19-Dec-2017 16:53	78
[TXT]	iwgsc_refseqv1.0_ISBP.fa.zip.md5.txt	05-Dec-2017 15:01	64	[]	iwgsc_refseqv1.0_miRNA_2017May26.gff3.zip	29-May-2017 11:17	347K
[]	iwgsc_refseqv1.0_LowConf_2017Mar13.gff3.zip	24-Mar-2017 17:27	7.3M	[TXT]	iwgsc_refseqv1.0_miRNA_2017May26.gff3.zip.md5.txt	05-Dec-2017 15:01	76
[TXT]	iwgsc_refseqv1.0_LowConf_2017Mar13.gff3.zip.md5.txt	05-Dec-2017 15:01	78	[]	iwgsc_refseqv1.0_miRNA_mature-miRNA_sequences_2017May26.fsa.zip	29-May-2017 11:17	57K
[]	iwgsc_refseqv1.0_LowConf_CDS_2017Mar13.fa.zip	24-Mar-2017 17:27	30M	[TXT]	iwgsc_refseqv1.0_miRNA_mature-miRNA_sequences_2017May26.fsa.zip.md5.txt	05-Dec-2017 15:01	98
[TXT]	iwgsc_refseqv1.0_LowConf_CDS_2017Mar13.fa.zip.md5.txt	05-Dec-2017 15:01	80	[]	iwgsc_refseqv1.0_miRNA_pre-miRNA_sequences_2017May26.fsa.zip	29-May-2017 11:17	287K
[]	iwgsc_refseqv1.0_LowConf_PROTEIN_2017Mar13.fa.zip	24-Mar-2017 17:27	18M	[TXT]	iwgsc_refseqv1.0_miRNA_pre-miRNA_sequences_2017May26.fsa.zip.md5.txt	05-Dec-2017 15:01	95
[TXT]	iwgsc_refseqv1.0_LowConf_PROTEIN_2017Mar13.fa.zip.md5.txt	05-Dec-2017 15:01	84	[]	iwgsc_refseqv1.0_optical_maps_group7.zip	27-Nov-2017 17:38	4.6M
[]	iwgsc_refseqv1.0_LowConf_REPR_CDS_2017Apr03.fa.zip	11-Apr-2017 15:45	30M	[TXT]	iwgsc_refseqv1.0_optical_maps_group7.zip.md5.txt	05-Dec-2017 15:01	76
[TXT]	iwgsc_refseqv1.0_LowConf_REPR_CDS_2017Apr03.fa.zip.md5.txt	05-Dec-2017 15:01	85	[]	iwgsc_refseqv1.0_vs_TGACv1.refmap.zip	26-Jun-2017 13:39	8.9M
[]	iwgsc_refseqv1.0_LowConf_REPR_PROTEIN_2017Apr03.fa.zip	11-Apr-2017 15:45	18M	[TXT]	iwgsc_refseqv1.0_vs_TGACv1.refmap.zip.md5.txt	05-Dec-2017 15:01	74
[TXT]	iwgsc_refseqv1.0_LowConf_REPR_PROTEIN_2017Apr03.fa.zip.md5.txt	05-Dec-2017 15:01	89	[]	iwgsc_refseqv1.0_vs_TGACv1.stats.zip	26-Jun-2017 13:39	1.0K
[]	iwgsc_refseqv1.0_LowConf_UTR_2017May05.gff3.zip	09-May-2017 09:53	10M	[TXT]	iwgsc_refseqv1.0_vs_TGACv1.stats.zip.md5.txt	05-Dec-2017 15:01	71
[TXT]	iwgsc_refseqv1.0_LowConf_UTR_2017May05.gff3.zip.md5.txt	05-Dec-2017 15:01	82	[]	iwgsc_refseqv1.0_vs_TGACv1.tmap.zip	26-Jun-2017 13:39	8.9M
				[TXT]	iwgsc_refseqv1.0_vs_TGACv1.tmap.zip.md5.txt	05-Dec-2017 15:01	70

BLAST

- BLAST dedicated to IWGSC general access agreement:

https://urgi.versailles.inra.fr/blast_iwgsc/?dbgroup=wheat_iwgsc_refseq_v1_chromosomes&program=blastn

- Allow to BLAST all the available assemblies in one time including the **IWGSC RefSeq v1.0**
- 476k BLAST searches performed in 2017

 476k

BLAST



IWGSC BLAST



[Home](#) [Databanks](#) [About](#) [Help](#) [Contact us](#)

BLAST parameter settings

Enter query sequences here in [Fasta format](#)

Or upload sequence fasta file (max 2M): [Parcourir...](#) Aucun fichier sélectionné.

Program

Group

Database(s)

- IWGSC RefSeq v1.0 all chromosomes
- IWGSC RefSeq v1.0 chromosome 1A only
- IWGSC RefSeq v1.0 chromosome 1B only
- IWGSC RefSeq v1.0 chromosome 1D only
- IWGSC RefSeq v1.0 chromosome 2A only
- IWGSC RefSeq v1.0 chromosome 2B only
- IWGSC RefSeq v1.0 chromosome 2D only
- IWGSC RefSeq v1.0 chromosome 3A only
- IWGSC RefSeq v1.0 chromosome 3B only
- IWGSC RefSeq v1.0 chromosome 3D only

currently selected database(s)

[remove](#)





































Basic Search - using default BLAST parameter settings

[Basic search](#)

[Reset](#)



BLAST

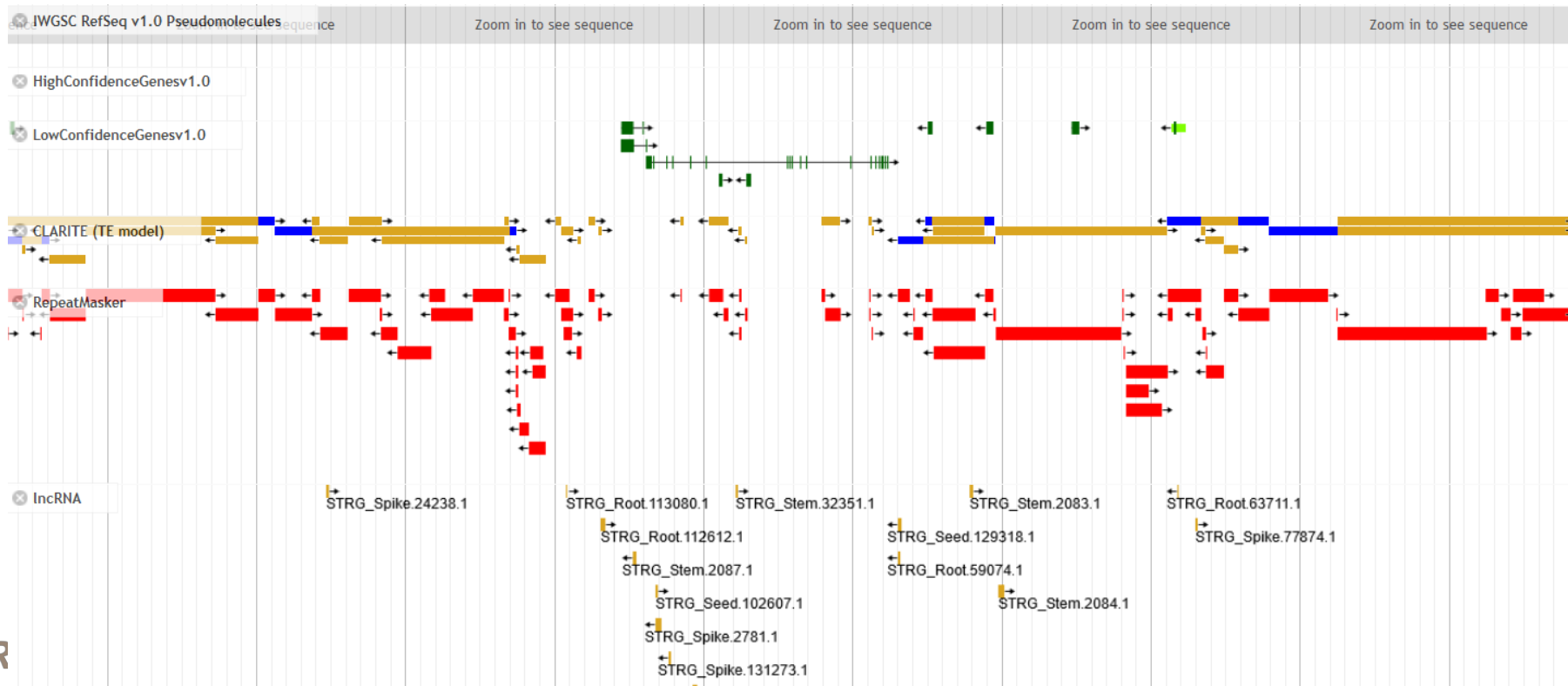
Query	Databanks	Subject	BEST HSP						
			Score	Identities (Query length)	Percentage	Expect	Start	End	
Synth12	IWGSC RefSeq v1.0 chromosome 3B only	<input type="checkbox"/> chr3B 	 	2805	1559/1560 (1560)	99	0.0	10949433	10950991
Synth12	IWGSC RefSeq v1.0 chromosome 4A only	<input type="checkbox"/> chr4A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

- Link to download the matching sequence
- Link to display the JBrowse zoomed in the matching region

JBrowse

- IWGSC RefSeq v1.0 JBrowse available under general access agreement:

https://urgi.versailles.inra.fr/jbrowseiwgsc/gmod_jbrowse/?data=myData%2FIWGSC_RefSeq_v1.0



WheatMine

- IWGSC RefSeq v1.0 InterMine available under general access agreement:

<https://urgi.versailles.inra.fr/WheatMine>

The screenshot shows the WheatMine web application interface. At the top, there is a navigation bar with links for Home, Templates, Lists, QueryBuilder, Regions, Data Sources, API, and MyMine. A search bar on the right contains the text "e.g. TraesCS1A01G000200, s" and a "GO" button. Below the navigation bar, there are three main sections: "Search", "Analyse", and "Welcome Back!". The "Search" section has a magnifying glass icon and a text input field with the placeholder "e.g. X, Y, Z" and a "SEARCH" button. The "Analyse" section has a document icon and a text input field with the placeholder "e.g. TraesCS1A01G006400, TraesCS1A01G033100, TraesCS1A01G034200," and an "ANALYSE" button. The "Welcome Back!" section has a "TAKE A TOUR" button. At the bottom, there is a "WHEATMINE CONTENT" section with a "Read more" link and a "Query for wheatmine content:" field. A "templates" label is visible in the bottom right corner of the content area.

WheatMine

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

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

[ENRIS](#)

Quick Links: [Summary](#) [Genomics](#) [Other](#)

Genome feature

Region: **gene** Length: 1954 [FASTA...](#)
 Location: **chr1A:3669097-3671050** reverse strand

Genome Browser

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

3 Data Sets

Name
High confidence genes
Functional annotation
Genes and markers sequences

4 Gos

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

7 Interpro

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

15 Overlapping Features

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

[Show more rows](#)

2 Pfams

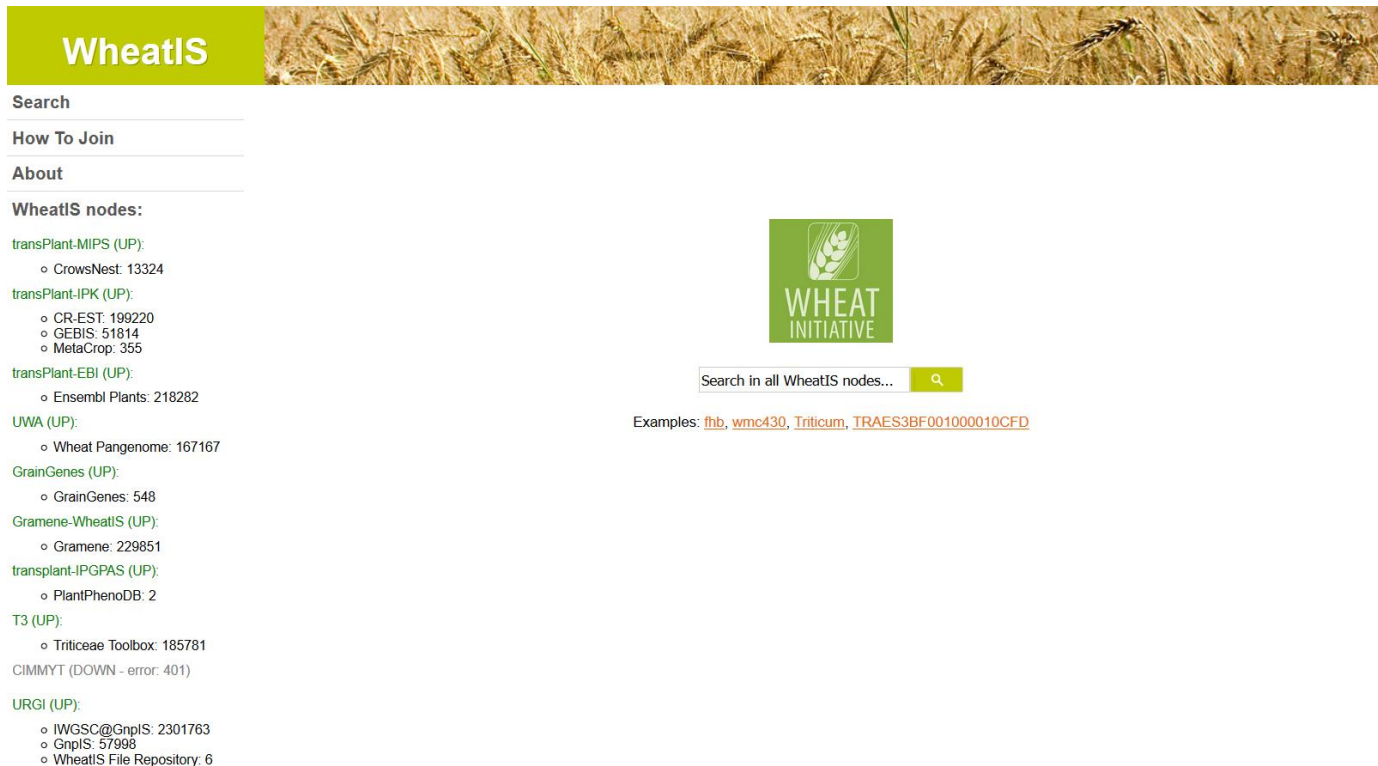
Identifier	Description
PF00069	Protein kinase domain
PF13855	Leucine rich repeat



WheatIS search

- IWGSC data can be discovered using the WheatIS search:

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



The screenshot shows the WheatIS search interface. On the left, there is a navigation menu with links for Search, How To Join, and About. Below this is a list of WheatIS nodes, each with a sub-link and a count. On the right, there is a search bar with the text "Search in all WheatIS nodes..." and a magnifying glass icon. Below the search bar, there are examples of search terms: fhb, wmc430, Triticum, and TRAES3BF001000010CFD. The background of the interface features a green header with the text "WheatIS" and a photograph of a wheat field.

WheatIS

Search

How To Join

About

WheatIS nodes:

- transPlant-MIPS (UP):
 - CrowsNest: 13324
- transPlant-IPK (UP):
 - CR-EST: 199220
 - GEBIS: 51814
 - MetaCrop: 355
- transPlant-EBI (UP):
 - Ensembl Plants: 218282
- UWA (UP):
 - Wheat Pangenome: 167167
- GrainGenes (UP):
 - GrainGenes: 548
- Gramene-WheatIS (UP):
 - Gramene: 229851
- transplant-IPGPAS (UP):
 - PlantPhenoDB: 2
- T3 (UP):
 - Triticeae Toolbox: 185781
- CIMMYT (DOWN - error: 401)
- URGI (UP):
 - IWGSC@GnplS: 2301763
 - GnplS: 57998
 - WheatIS File Repository: 6

WHEAT INITIATIVE

Search in all WheatIS nodes...

Examples: [fbh](#), [wmc430](#), [Triticum](#), [TRAES3BF001000010CFD](#)

WheatIS search

- IWGSC RefSeq v1.0 data will be added as soon as they are in open access

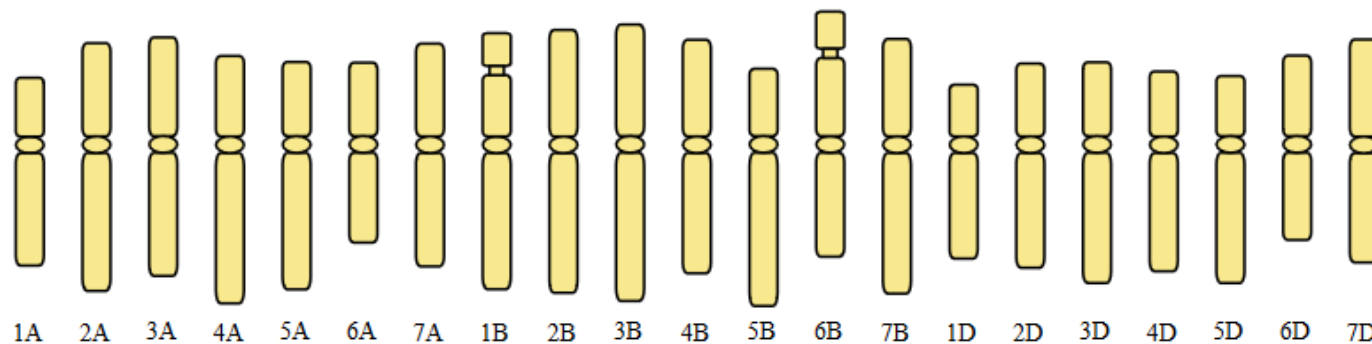
ID	Source	Type	Taxon	Description
JBrowse chr1A 100233074 100233167 miRNA	IWGSC@GnpIS	SEQUENCE FEATURE	Triticum aestivum	Start = 100233074 [...]
JBrowse chr1A 101106604 101106809 miRNA	IWGSC@GnpIS	SEQUENCE FEATURE	Triticum aestivum	Start = 101106604 [...]
JBrowse chr1A 10158696 10158807 miRNA	IWGSC@GnpIS	SEQUENCE FEATURE	Triticum aestivum	Start = 10158696 [...]
JBrowse chr1A 10231918 10232011 miRNA	IWGSC@GnpIS	SEQUENCE FEATURE	Triticum aestivum	Start = 10231918 [...]
JBrowse chr1A 102424297 102424428 miRNA	IWGSC@GnpIS	SEQUENCE FEATURE	Triticum aestivum	Start = 102424297 [...]
JBrowse chr1A 104839251 104839344 miRNA	IWGSC@GnpIS	SEQUENCE FEATURE	Triticum aestivum	Start = 104839251 [...]
JBrowse chr1A 104839691 104839784 miRNA	IWGSC@GnpIS	SEQUENCE FEATURE	Triticum aestivum	Start = 104839691 [...]
JBrowse chr1A 105619444 105619537 miRNA	IWGSC@GnpIS	SEQUENCE FEATURE	Triticum aestivum	Start = 105619444 [...]
JBrowse chr1A 107474939 107475091 miRNA	IWGSC@GnpIS	SEQUENCE FEATURE	Triticum aestivum	Start = 107474939 [...]
JBrowse chr1A 108071355 108071486 miRNA	IWGSC@GnpIS	SEQUENCE FEATURE	Triticum aestivum	Start = 108071355 [...]

URGI (UP):

- IWGSC@GnpIS: 16385772

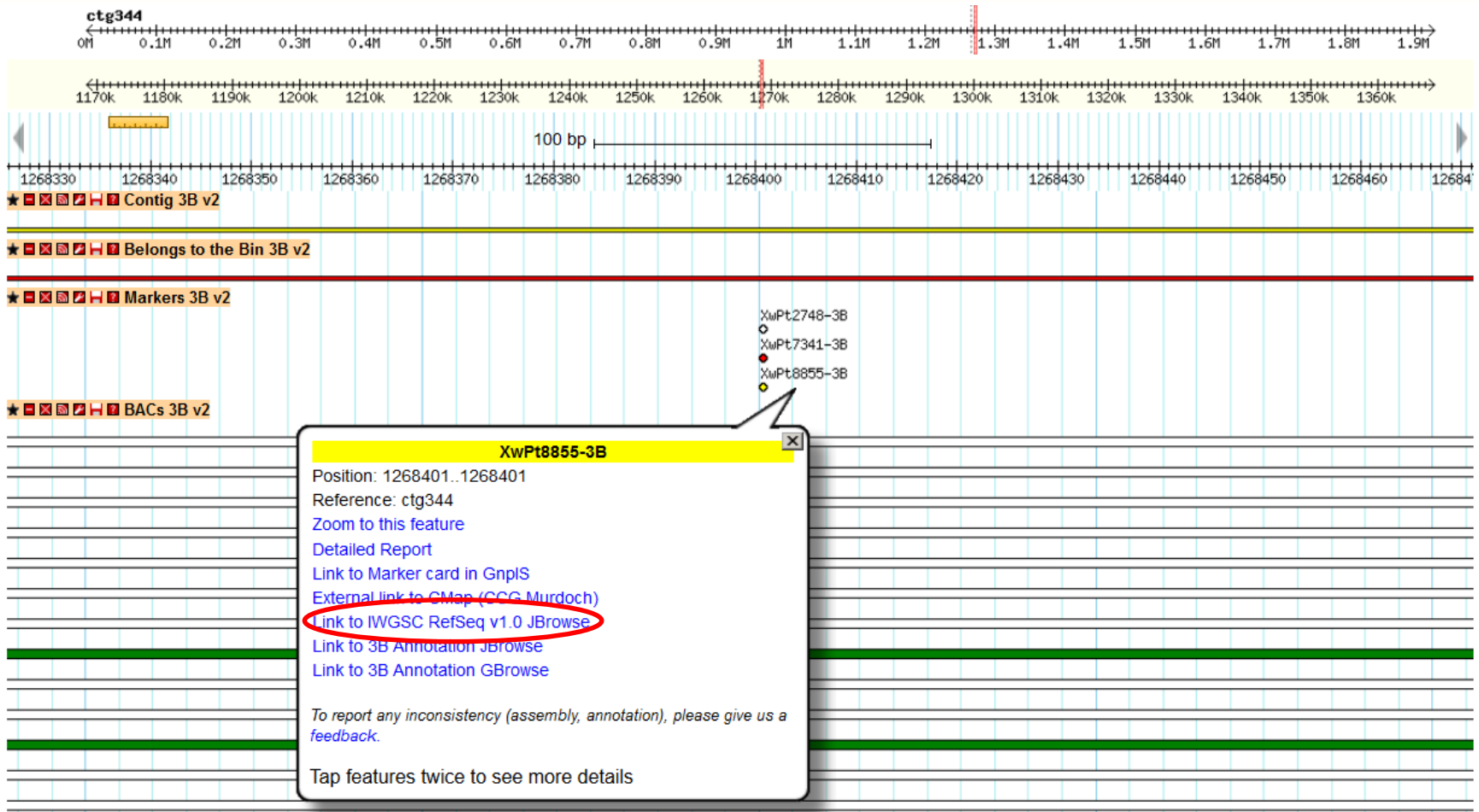
Physical maps

- New versions of 7DS, 1BS added
- All the physical maps are available to display



- Link to the IWGSC RefSeq v1.0

Physical maps





Other Wheat data resources hosted at URGI

GnplS-coreDB

- Big data generated from french and european projects

Triticaceae Genome

Wheatbi

Breed wheat

- These data are integrated in the GnplS-coreDB developed at URGI (part of the GnplS Information System).

GnplS-coreDB

- Wheat data overview:

Thematic	Object	#Total	#Open access	#Restricted access to projects
Genetic Resources	Taxon	56	56	0
	Accession	12839	10016	2823
Genetic Maps	Map	30	29	1
	Marker	704822	34164	670658
	QTL	749	465	284
SNP discovery	In Silico Analysis	11	9	2
	Sequence Variation	134904	55362	79542
	SNP, indel	724132	95	724037
Genotyping (high throughput)	Experiment	22	1	21
	Sample	8229	42	8216
	Marker	668543	0	668543
Phenotyping	Trial	853	821	32
	Sample	3660	2985	901
	Variable	291	91	200
	Observation	116103	8	527981
GWAS	Analysis	1555	43	1512
	Sample	2365	1839	526
	Variable	359	37	322
	Marker	123866	4109	119757
	Association	824217	48596	775621

Genetic and phenomic data

GWAS mapped on the IWGSC RefSeq v1.0

DETAILS ON ITEM

Marker	cfn0655053
Trait	QY15
Location	ARVALIS GREOUX
Phenotyping Year	Breadwheat_2013-2014
pVal	6.570000099515028e-11

BOXPLOT

MORE DETAILS

AA **GG**

Lot	Accession	Number	Phenotyping Value
37315-2013	GALLANT	37315	132.9
29843-2013	ALTIGO	29843	119.055
24573-2013	CORDIALE	24573	115.453
29843-2013	ALTIGO	29843	114.186
29843-2013	ALTIGO	29843	113.476
35613-2013	NUCLEO	35613	112.963
7043-2013	TREME	7043	112.595

Phenotyping experiments

Search parameter(s):

Traits and Variables: [WIPO:0000030:](#)

[Get Climat Data](#)

[Geolocation](#)

Trials: 27

Trial : [BW_WP2_2014_SECOBRA Maule Fusariose](#)
Site : SECOBRA MAULE
Data Available

Trial : [BW_WP2_2012_ARVALIS Greoux Water](#)
Site : ARVALIS GREOUX
Data not Available

Trial : [BW_WP2_2013_SYNGENTA Andelu Nitrogen](#)
Site : SYNGENTA ANDELU

Origin site Collecting site Evaluation site

Phenotyping campaign(s)

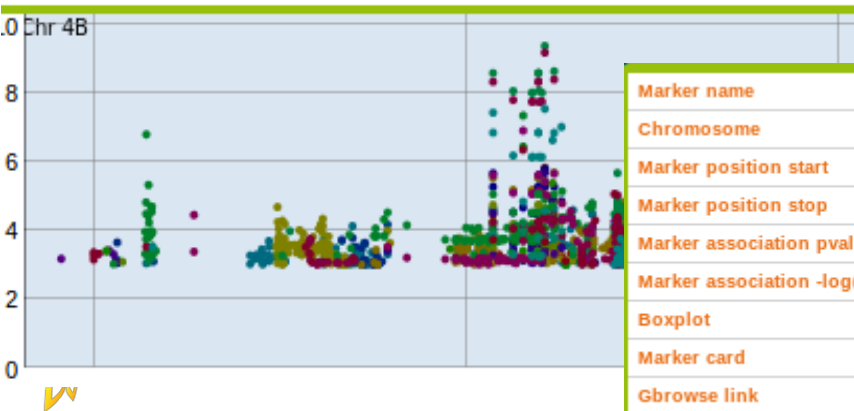
[Breadwheat_2011-2012](#) [Breadwheat_2012-2013](#) [Breadwheat_2013-2014](#)

[remove all](#) [add all](#)

LEVEL: PLOT

1-10 of 8,277 results per page

Lot Number	Accession Number	Accession Name	drought	Trial Name	Trial Site
13232-2012	13232	ORATORIO	SEC: dry	BW_WP2_2013_ARVALIS_Greoux_Water	ARVALIS GRE
29853-2012	29853	ALDRIC	SEC: dry	BW_WP2_2013_ARVALIS_Greoux_Water	ARVALIS GRE
13481-2012	13481	APACHE	SEC: dry	BW_WP2_2013_ARVALIS_Greoux_Water	ARVALIS GRE
22982-2012	22982	XI19	SEC: dry	BW_WP2_2013_ARVALIS_Greoux_Water	ARVALIS GRE
37348-2012	37348	PRIMO	SEC: dry	BW_WP2_2013_ARVALIS_Greoux_Water	ARVALIS GRE
37350-2012	37350	ESPERIA	SEC: dry	BW_WP2_2013_ARVALIS_Greoux_Water	ARVALIS GRE
35586-2012	35586	ACCROC	SEC: dry	BW_WP2_2013_ARVALIS_Greoux_Water	ARVALIS GRE
37357-2012	37357	ICARDA 4	SEC: dry	BW_WP2_2013_ARVALIS_Greoux_Water	ARVALIS GRE
36675-2012	36675	SCENARIO	SEC: dry	BW_WP2_2013_ARVALIS_Greoux_Water	ARVALIS GRE
29855-2012	29855	PREMIO	SEC: dry	BW_WP2_2013_ARVALIS_Greoux_Water	ARVALIS GRE



Integration of the sequence to genetic and phenomic data

- cf. **Alaux *et al.*** companion paper that will be submitted soon.
- Use cases:
 - BLAST
 - Gene (JBrowse)
 - Marker (GnpIS-coreDB)
 - QTLs (GnpIS-coreDB)
 - Phenotyping experiments (GnpIS-coreDB)

Integration of the sequence to genetic and phenomic data

- WheatIS search
 - Gene (WheatMine)
 - Marker (GnpIS-coreDB)
 - GWAS (GnpIS-coreDB)
 - Phenotyping experiments (GnpIS-coreDB)

Perspectives

- Management of new/curated versions of the IWGSC RefSeq annotation

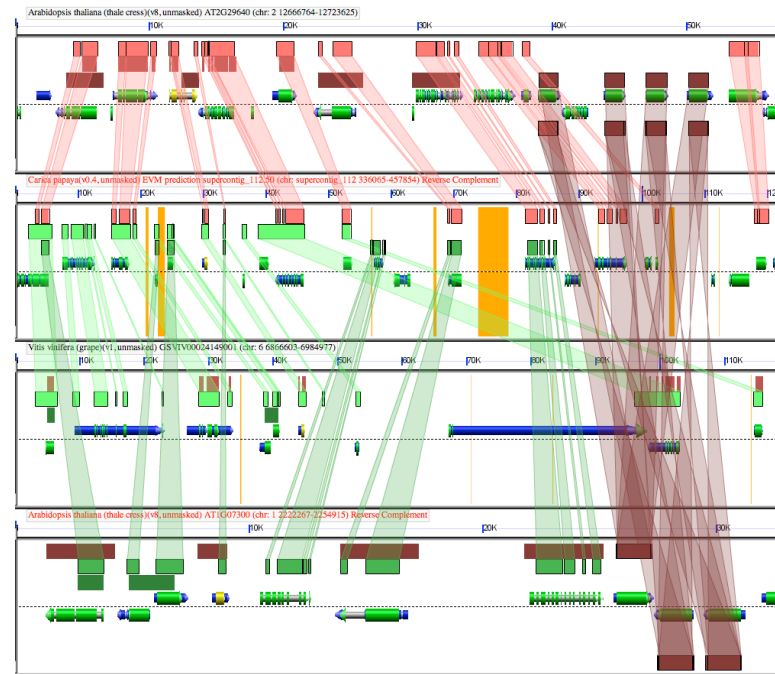
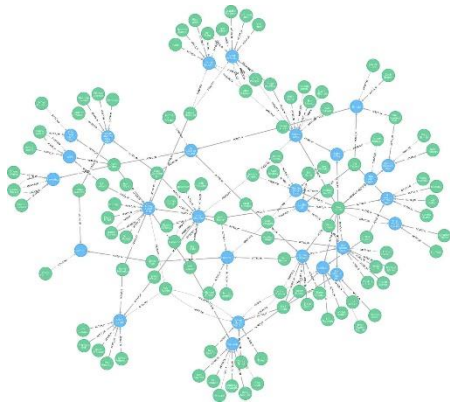
Web Apollo

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

The screenshot displays the Web Apollo genome browser interface. At the top, navigation controls and a search bar show 'Chr10' and coordinates '22,213,175' to '22,213,200'. Below this are several tracks: a 'DNA Track' showing raw sequence, a 'Toggle Reverse Strand' and 'Toggle Protein Translation' track, and a 'User-created Annotations' Track. Three pop-up windows are visible: 'Add Substitution', 'Add Deletion', and 'Add Insertion', each with fields for '+ strand' and '- strand' and an 'Add' button. The 'Apollo' logo is in the bottom left, and a 'WELLSY' logo is in the bottom right.

Perspectives

- Bioinformatics challenges of a Pan-genome
 - Use new technologies to handle and display large amount of linked data



Acknowledgements



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

Choulet F.
Rimbert H.
Leroy P.
Guilhot N.
Paux E.

Rogers J.
Caugant I.
Eversole K.

IWGSC Coordinating
Committee

IWGSC Sequencing and
Analysis team

All data providers



Questions

IWGSC Data Repository

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

IWGSC BLAST

https://urgi.versailles.inra.fr/blast_iwgsc/?dbgroup=wheat_iwgsc_refseq_v1_chromosomes&program=blastn

Contact me at michael.alaux@inra.fr

 [@michaelalaux](https://twitter.com/michaelalaux)