



**HAL**  
open science

## The IWGSC Data repository and wheat data resources hosted at URGI: Overview and perspectives

Michael Alaux, Jane Rogers, Thomas Letellier, Raphaël Flores, Cyril Pommier, Nacer Mohellibi, Sophie Durand, Erik Kimmel, Célia Michotey, Mikaël Loaec, et al.

### ► To cite this version:

Michael Alaux, Jane Rogers, Thomas Letellier, Raphaël Flores, Cyril Pommier, et al.. The IWGSC Data repository and wheat data resources hosted at URGI: Overview and perspectives. PAG XXVI - Plant and Animal Genome Conference, Jan 2018, San Diego, United States. pp.32 slides. hal-02788401

**HAL Id: hal-02788401**

**<https://hal.inrae.fr/hal-02788401v1>**

Submitted on 5 Jun 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



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

Michael Alaux



## IWGSC Data Repository

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



Michael Alaux

## 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 (4A, 5B);
- MTP BAC WGP™ 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 CaxRn 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**.

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



Michael Alaux

## Assemblies

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

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



Michael Alaux

## 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 (E1), 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 PGSC 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 will take no more than 2 business days for your URGI account to be established but occasionally it may 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.

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



Michael Alaux

## 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



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

Michael Alaux

## Annotations

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



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

Michael Alaux

## Access to the IWGC RefSeq v1.0 data

- Under Toronto agreement (IWGC general access agreement)

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

→ will be in open access once published

### How to access IWGC RefSeq v1.0 data?

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

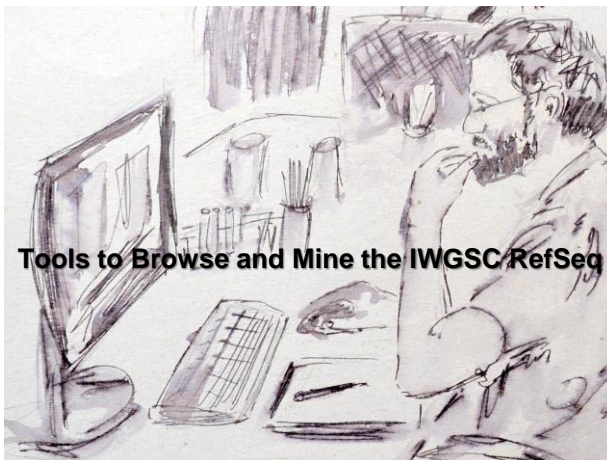
- Individuals who have not signed the IWGC Data Access Agreement should **FIRST register on the IWGC website** and sign the Agreement. URGI login details will be provided subsequently by email for access to the data. Typically, this it will take no more than 2 business days for your URGI account to be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

- Individuals who have already signed the IWGC Data Access Agreement can go directly to the URGI website to access the data using their URGI login details.



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

Michael Alaux



## Download

### IWGC RefSeq 1.0 assembly

[ ] <a href="#">iwgc_refseq1_0_all_chromosomes.zip</a>	04-May-2017 12:31 8.00	[ ] <a href="#">iwgc_refseq1_0_chr1A.fasta.zip</a>	04-May-2017 12:23 202M
[TXT] <a href="#">iwgc_refseq1_0_all_chromosomes.zip.md5.txt</a>	05-Dec-2017 15:00 71	[TXT] <a href="#">iwgc_refseq1_0_chr1A.fasta.zip.md5.txt</a>	05-Dec-2017 15:00 65
[ ] <a href="#">iwgc_refseq1_0_chr1A.fasta.zip</a>	04-May-2017 12:14 169M	[ ] <a href="#">iwgc_refseq1_0_chr1B.fasta.zip</a>	04-May-2017 12:24 203M
[TXT] <a href="#">iwgc_refseq1_0_chr1A.fasta.zip.md5.txt</a>	05-Dec-2017 15:00 65	[TXT] <a href="#">iwgc_refseq1_0_chr1B.fasta.zip.md5.txt</a>	05-Dec-2017 15:00 65
[ ] <a href="#">iwgc_refseq1_0_chr1B.fasta.zip</a>	04-May-2017 12:15 197M	[ ] <a href="#">iwgc_refseq1_0_chr1C.fasta.zip</a>	04-May-2017 12:25 140M
[TXT] <a href="#">iwgc_refseq1_0_chr1B.fasta.zip.md5.txt</a>	05-Dec-2017 15:00 65	[TXT] <a href="#">iwgc_refseq1_0_chr1C.fasta.zip.md5.txt</a>	05-Dec-2017 15:00 65
[ ] <a href="#">iwgc_refseq1_0_chr1D.fasta.zip</a>	04-May-2017 12:15 140M	[ ] <a href="#">iwgc_refseq1_0_chr1E.fasta.zip</a>	04-May-2017 12:23 176M
[TXT] <a href="#">iwgc_refseq1_0_chr1D.fasta.zip.md5.txt</a>	05-Dec-2017 15:00 65	[TXT] <a href="#">iwgc_refseq1_0_chr1E.fasta.zip.md5.txt</a>	05-Dec-2017 15:00 65
[ ] <a href="#">iwgc_refseq1_0_chr1E.fasta.zip</a>	04-May-2017 12:16 222M	[ ] <a href="#">iwgc_refseq1_0_chr1F.fasta.zip</a>	04-May-2017 12:26 205M
[TXT] <a href="#">iwgc_refseq1_0_chr1E.fasta.zip.md5.txt</a>	05-Dec-2017 15:00 65	[TXT] <a href="#">iwgc_refseq1_0_chr1F.fasta.zip.md5.txt</a>	05-Dec-2017 15:00 65
[ ] <a href="#">iwgc_refseq1_0_chr1F.fasta.zip</a>	04-May-2017 12:17 229M	[ ] <a href="#">iwgc_refseq1_0_chr1G.fasta.zip</a>	04-May-2017 12:27 134M
[TXT] <a href="#">iwgc_refseq1_0_chr1F.fasta.zip.md5.txt</a>	05-Dec-2017 15:00 65	[TXT] <a href="#">iwgc_refseq1_0_chr1G.fasta.zip.md5.txt</a>	05-Dec-2017 15:00 65
[ ] <a href="#">iwgc_refseq1_0_chr1G.fasta.zip</a>	04-May-2017 12:18 183M	[ ] <a href="#">iwgc_refseq1_0_chr1H.fasta.zip</a>	04-May-2017 12:27 209M
[TXT] <a href="#">iwgc_refseq1_0_chr1G.fasta.zip.md5.txt</a>	05-Dec-2017 15:00 65	[TXT] <a href="#">iwgc_refseq1_0_chr1H.fasta.zip.md5.txt</a>	05-Dec-2017 15:00 65
[ ] <a href="#">iwgc_refseq1_0_chr1H.fasta.zip</a>	04-May-2017 12:18 214M	[ ] <a href="#">iwgc_refseq1_0_chr1I.fasta.zip</a>	04-May-2017 12:29 214M
[TXT] <a href="#">iwgc_refseq1_0_chr1H.fasta.zip.md5.txt</a>	05-Dec-2017 15:00 65	[TXT] <a href="#">iwgc_refseq1_0_chr1I.fasta.zip.md5.txt</a>	05-Dec-2017 15:00 65
[ ] <a href="#">iwgc_refseq1_0_chr1I.fasta.zip</a>	04-May-2017 12:19 237M	[ ] <a href="#">iwgc_refseq1_0_chr1J.fasta.zip</a>	04-May-2017 12:29 181M
[TXT] <a href="#">iwgc_refseq1_0_chr1I.fasta.zip.md5.txt</a>	05-Dec-2017 15:00 65	[TXT] <a href="#">iwgc_refseq1_0_chr1J.fasta.zip.md5.txt</a>	05-Dec-2017 15:00 65
[ ] <a href="#">iwgc_refseq1_0_chr1J.fasta.zip</a>	04-May-2017 12:20 175M	[ ] <a href="#">iwgc_refseq1_0_chr1K.fasta.zip</a>	04-May-2017 12:29 118M
[TXT] <a href="#">iwgc_refseq1_0_chr1J.fasta.zip.md5.txt</a>	05-Dec-2017 15:00 65	[TXT] <a href="#">iwgc_refseq1_0_chr1K.fasta.zip.md5.txt</a>	05-Dec-2017 15:00 65
[ ] <a href="#">iwgc_refseq1_0_chr1K.fasta.zip</a>	04-May-2017 12:21 212M	[ ] <a href="#">iwgc_refseq1_0_chr1L.fasta.zip</a>	04-May-2017 12:29 118M
[TXT] <a href="#">iwgc_refseq1_0_chr1K.fasta.zip.md5.txt</a>	05-Dec-2017 15:00 65	[TXT] <a href="#">iwgc_refseq1_0_chr1L.fasta.zip.md5.txt</a>	05-Dec-2017 15:00 65
[ ] <a href="#">iwgc_refseq1_0_chr1L.fasta.zip</a>	04-May-2017 12:22 162M	[ ] <a href="#">iwgc_refseq1_0_chr1M.fasta.zip</a>	04-May-2017 12:29 118M
[TXT] <a href="#">iwgc_refseq1_0_chr1L.fasta.zip.md5.txt</a>	05-Dec-2017 15:00 65	[TXT] <a href="#">iwgc_refseq1_0_chr1M.fasta.zip.md5.txt</a>	05-Dec-2017 15:00 65
[ ] <a href="#">iwgc_refseq1_0_chr1M.fasta.zip</a>	04-May-2017 12:22 144M	[ ] <a href="#">iwgc_refseq1_0_chr1N.fasta.zip</a>	04-May-2017 12:29 118M
[TXT] <a href="#">iwgc_refseq1_0_chr1M.fasta.zip.md5.txt</a>	05-Dec-2017 15:00 65	[TXT] <a href="#">iwgc_refseq1_0_chr1N.fasta.zip.md5.txt</a>	05-Dec-2017 15:00 65



[https://urgi.versailles.inra.fr/download/iwgc/IWGC\\_RefSeq\\_Assemblies/v1.0/](https://urgi.versailles.inra.fr/download/iwgc/IWGC_RefSeq_Assemblies/v1.0/)

Michael Alaux

## Download

### IWGC RefSeq v1.0 annotation

[ ] <a href="#">iwgc_refseq1_0_annotation_genes_highconf_20170401.zip</a>	04-Dec-2017 14:49 108M	[ ] <a href="#">iwgc_refseq1_0_annotation_genes_lowconf_20170401.zip</a>	24-Mar-2017 17:27 108M
[TXT] <a href="#">iwgc_refseq1_0_annotation_genes_highconf_20170401.zip.md5.txt</a>	04-Dec-2017 14:49 82	[TXT] <a href="#">iwgc_refseq1_0_annotation_genes_lowconf_20170401.zip.md5.txt</a>	04-Dec-2017 15:00 80
[ ] <a href="#">iwgc_refseq1_0_annotation_genes_lowconf_20170401.zip</a>	24-Mar-2017 17:26 108M	[ ] <a href="#">iwgc_refseq1_0_annotation_genes_highconf_20170401.zip</a>	04-Dec-2017 14:49 108M
[TXT] <a href="#">iwgc_refseq1_0_annotation_genes_lowconf_20170401.zip.md5.txt</a>	05-Dec-2017 15:00 79	[TXT] <a href="#">iwgc_refseq1_0_annotation_genes_highconf_20170401.zip.md5.txt</a>	04-Dec-2017 15:00 79
[ ] <a href="#">iwgc_refseq1_0_annotation_genes_highconf_20170401.zip</a>	24-Mar-2017 17:27 108M	[ ] <a href="#">iwgc_refseq1_0_annotation_genes_lowconf_20170401.zip</a>	04-Dec-2017 14:49 108M
[TXT] <a href="#">iwgc_refseq1_0_annotation_genes_highconf_20170401.zip.md5.txt</a>	05-Dec-2017 15:00 81	[TXT] <a href="#">iwgc_refseq1_0_annotation_genes_lowconf_20170401.zip.md5.txt</a>	04-Dec-2017 15:00 80
[ ] <a href="#">iwgc_refseq1_0_annotation_genes_lowconf_20170401.zip</a>	24-Mar-2017 17:27 108M	[ ] <a href="#">iwgc_refseq1_0_annotation_genes_highconf_20170401.zip</a>	04-Dec-2017 14:49 108M
[TXT] <a href="#">iwgc_refseq1_0_annotation_genes_lowconf_20170401.zip.md5.txt</a>	05-Dec-2017 15:00 80	[TXT] <a href="#">iwgc_refseq1_0_annotation_genes_highconf_20170401.zip.md5.txt</a>	04-Dec-2017 15:00 80
[ ] <a href="#">iwgc_refseq1_0_annotation_genes_highconf_20170401.zip</a>	24-Mar-2017 17:27 108M	[ ] <a href="#">iwgc_refseq1_0_annotation_genes_lowconf_20170401.zip</a>	04-Dec-2017 14:49 108M
[TXT] <a href="#">iwgc_refseq1_0_annotation_genes_highconf_20170401.zip.md5.txt</a>	05-Dec-2017 15:00 80	[TXT] <a href="#">iwgc_refseq1_0_annotation_genes_lowconf_20170401.zip.md5.txt</a>	04-Dec-2017 15:00 80
[ ] <a href="#">iwgc_refseq1_0_annotation_genes_lowconf_20170401.zip</a>	24-Mar-2017 17:27 108M	[ ] <a href="#">iwgc_refseq1_0_annotation_genes_highconf_20170401.zip</a>	04-Dec-2017 14:49 108M
[TXT] <a href="#">iwgc_refseq1_0_annotation_genes_lowconf_20170401.zip.md5.txt</a>	05-Dec-2017 15:00 80	[TXT] <a href="#">iwgc_refseq1_0_annotation_genes_highconf_20170401.zip.md5.txt</a>	04-Dec-2017 15:00 80
[ ] <a href="#">iwgc_refseq1_0_annotation_genes_highconf_20170401.zip</a>	24-Mar-2017 17:27 108M	[ ] <a href="#">iwgc_refseq1_0_annotation_genes_lowconf_20170401.zip</a>	04-Dec-2017 14:49 108M
[TXT] <a href="#">iwgc_refseq1_0_annotation_genes_highconf_20170401.zip.md5.txt</a>	05-Dec-2017 15:00 80	[TXT] <a href="#">iwgc_refseq1_0_annotation_genes_lowconf_20170401.zip.md5.txt</a>	04-Dec-2017 15:00 80
[ ] <a href="#">iwgc_refseq1_0_annotation_genes_lowconf_20170401.zip</a>	24-Mar-2017 17:27 108M	[ ] <a href="#">iwgc_refseq1_0_annotation_genes_highconf_20170401.zip</a>	04-Dec-2017 14:49 108M
[TXT] <a href="#">iwgc_refseq1_0_annotation_genes_lowconf_20170401.zip.md5.txt</a>	05-Dec-2017 15:00 80	[TXT] <a href="#">iwgc_refseq1_0_annotation_genes_highconf_20170401.zip.md5.txt</a>	04-Dec-2017 15:00 80
[ ] <a href="#">iwgc_refseq1_0_annotation_genes_highconf_20170401.zip</a>	24-Mar-2017 17:27 108M	[ ] <a href="#">iwgc_refseq1_0_annotation_genes_lowconf_20170401.zip</a>	04-Dec-2017 14:49 108M
[TXT] <a href="#">iwgc_refseq1_0_annotation_genes_highconf_20170401.zip.md5.txt</a>	05-Dec-2017 15:00 80	[TXT] <a href="#">iwgc_refseq1_0_annotation_genes_lowconf_20170401.zip.md5.txt</a>	04-Dec-2017 15:00 80



[https://urgi.versailles.inra.fr/download/iwgc/IWGC\\_RefSeq\\_Annotations/v1.0/](https://urgi.versailles.inra.fr/download/iwgc/IWGC_RefSeq_Annotations/v1.0/)

Michael Alaux

## BLAST

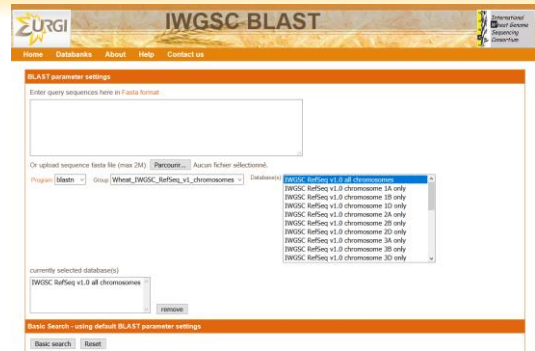
- BLAST dedicated to IWGSC general access agreement:  
[https://urgi.versailles.inra.fr/blast\\_iwgsc/?dbgroup=wheat\\_iwgsc\\_refseq\\_v1\\_chromosomes&program=blastn](https://urgi.versailles.inra.fr/blast_iwgsc/?dbgroup=wheat_iwgsc_refseq_v1_chromosomes&program=blastn)
- Allow to BLAST all the available assemblies in one time including the **IWGSC RefSeq v1.0**
- 476k BLAST searches performed in 2017

👍👎🔥 476k



Michael Alaux

## BLAST



[https://urgi.versailles.inra.fr/blast\\_iwgsc/?dbgroup=wheat\\_iwgsc\\_refseq\\_v1\\_chromosomes&program=blastn](https://urgi.versailles.inra.fr/blast_iwgsc/?dbgroup=wheat_iwgsc_refseq_v1_chromosomes&program=blastn)

ael Alaux

## BLAST

Query	Database	Subject	Score	MaxIdent (Query length)	Percentage	Expect	Start	End
SpwH12	IWGSC RefSeq v1.0 chromosome 1B only	Wheat_1B_112500000	955	9589168 (1580)	99	0.0	7094433	10992891
SpwH12	IWGSC RefSeq v1.0 chromosome 1A only	Wheat_1A_112500000	940	8307731 (1580)	99	0.0	73284459	73288124
SpwH12	IWGSC RefSeq v1.0 chromosome 1D only	Wheat_1D_112500000	898	8337732 (1580)	98	0.0	73284459	73288124
SpwH12	IWGSC RefSeq v1.0 chromosome 5A only	Wheat_5A_112500000	838	6201731 (1580)	95	0.0	60128827	60128827
SpwH12	IWGSC RefSeq v1.0 chromosome 5D only	Wheat_5D_112500000	818	6201732 (1580)	94	0.0	60064897	60064897
SpwH12	IWGSC RefSeq v1.0 chromosome 1D only	Wheat_1D_112500000	814	8247734 (1580)	95	0.0	63714278	63714278
SpwH12	IWGSC RefSeq v1.0 chromosome 1B only	Wheat_1B_112500000	809	6201730 (1580)	95	0.0	30581401	30581401
SpwH12	IWGSC RefSeq v1.0 chromosome 4D only	Wheat_4D_112500000	809	6201749 (1580)	94	0.0	87387930	87388229
SpwH12	IWGSC RefSeq v1.0 chromosome 1D only	Wheat_1D_112500000	799	6201730 (1580)	93	0.0	49383192	49383289
SpwH12	IWGSC RefSeq v1.0 chromosome 1D only	Wheat_1D_112500000	798	6201730 (1580)	94	0.0	41887484	41887578
SpwH12	IWGSC RefSeq v1.0 chromosome 1D only	Wheat_1D_112500000	796	6201730 (1580)	95	0.0	34015215	34015286
SpwH12	IWGSC RefSeq v1.0 chromosome 5A only	Wheat_5A_112500000	796	6001730 (1580)	95	0.0	58150983	58150983

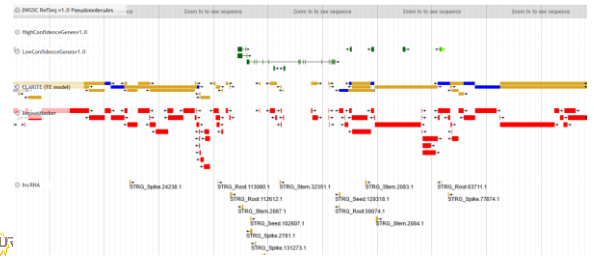
- Link to download the matching sequence
- Link to display the JBrowse zoomed in the matching region



Michael Alaux

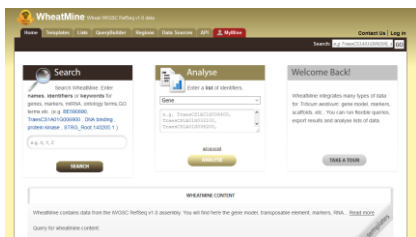
## JBrowse

- IWGSC RefSeq v1.0 JBrowse available under general access agreement:  
[https://urgi.versailles.inra.fr/browseiwgsc/gmod\\_jbrowse/?data=myData%2FIWGSC\\_RefSeq\\_v1.0](https://urgi.versailles.inra.fr/browseiwgsc/gmod_jbrowse/?data=myData%2FIWGSC_RefSeq_v1.0)



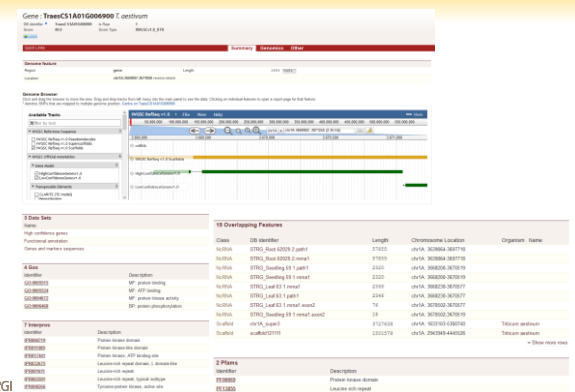
## WheatMine

- IWGSC RefSeq v1.0 InterMine available under general access agreement:  
<https://urgi.versailles.inra.fr/WheatMine>



Michael Alaux

## WheatMine



## WheatIS search

- IWGSC data can be discovered using the WheatIS search:

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



Michael Alaux

## WheatIS search

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

ID	Source	Type	Taxon	Description
WheatIS.chr1A_105233074_105233074_mRNA	IWGSC@GnPS	SEQUENCE FEATURE	Triticum aestivum	Start = 10223074 [ ]
WheatIS.chr1A_10159696_10159696_mRNA	IWGSC@GnPS	SEQUENCE FEATURE	Triticum aestivum	Start = 10159694 [ ]
WheatIS.chr1A_10159696_10159696_mRNA	IWGSC@GnPS	SEQUENCE FEATURE	Triticum aestivum	Start = 10159690 [ ]
WheatIS.chr1A_10231918_10231918_mRNA	IWGSC@GnPS	SEQUENCE FEATURE	Triticum aestivum	Start = 10231918 [ ]
WheatIS.chr1A_10242497_10242497_mRNA	IWGSC@GnPS	SEQUENCE FEATURE	Triticum aestivum	Start = 10242497 [ ]
WheatIS.chr1A_10453921_10453921_mRNA	IWGSC@GnPS	SEQUENCE FEATURE	Triticum aestivum	Start = 10453921 [ ]
WheatIS.chr1A_10453921_10453921_mRNA	IWGSC@GnPS	SEQUENCE FEATURE	Triticum aestivum	Start = 10453921 [ ]
WheatIS.chr1A_10561944_10561944_mRNA	IWGSC@GnPS	SEQUENCE FEATURE	Triticum aestivum	Start = 10561944 [ ]
WheatIS.chr1A_10747409_10747409_mRNA	IWGSC@GnPS	SEQUENCE FEATURE	Triticum aestivum	Start = 10747409 [ ]
WheatIS.chr1A_10807135_10807135_mRNA	IWGSC@GnPS	SEQUENCE FEATURE	Triticum aestivum	Start = 10807135 [ ]

URGI (UP):

- IWGSC@GnPS: 16385772

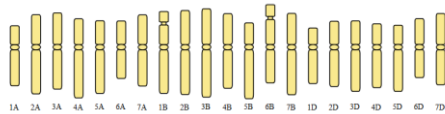


<http://www.wheatis.org/Search.php>

Michael Alaux

## Physical maps

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



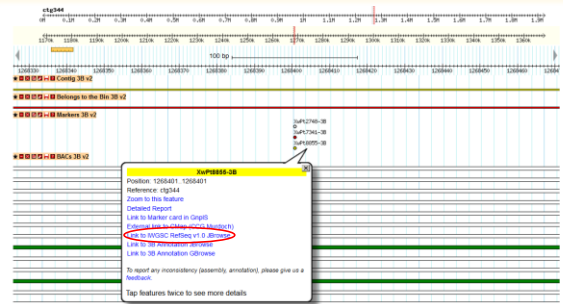
- Link to the IWGSC RefSeq v1.0



[https://urgi.versailles.inra.fr/gb2/gbrowse/wheat\\_phys\\_pub/](https://urgi.versailles.inra.fr/gb2/gbrowse/wheat_phys_pub/)

Michael Alaux

## Physical maps



[https://urgi.versailles.inra.fr/gb2/gbrowse/wheat\\_phys\\_pub/](https://urgi.versailles.inra.fr/gb2/gbrowse/wheat_phys_pub/)

Michael Alaux



## GnPS-coreDB

- Big data generated from french and european projects



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



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

Michael Alaux

## GnplS-coreDB

- Wheat data overview:

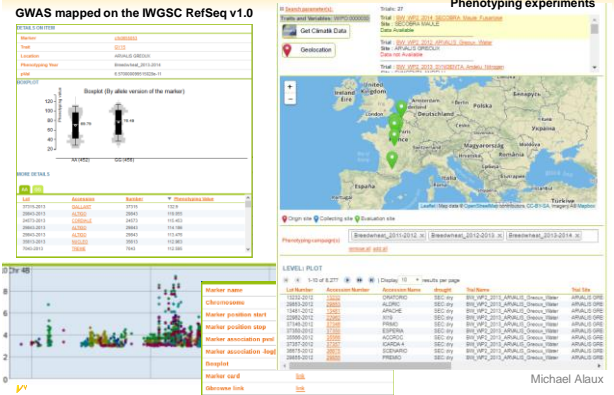
Thematic	Object	#Total	#Open access	#Restricted access to projects
Genetic Resources	Taxon	56	56	0
	Accession	12839	10016	2823
Genetic Maps	Map	30	29	1
	Marker	704822	34164	670658
SNP discovery	OTL	749	465	284
	In Silico Analysis	11	9	2
Genotyping (high throughput)	Sequence Variation	134904	55362	79542
	SNP, indel	724132	95	724037
Phenotyping	Experiment	22	1	21
	Sample	8229	42	8216
	Marker	668543	0	668543
GWAS	Trial	853	821	32
	Sample	3660	2985	901
Phenotyping	Variable	291	91	200
	Observation	8	527981	633057
GWAS	Analysis	1555	43	1512
	Sample	2365	1839	526
Phenotyping	Variable	359	37	322
	Marker	123866	4109	119757
GWAS	Association	824217	48596	775621



<http://wheat-urqi.versailles.inra.fr/>

Michael Alaux

## Genetic and phenomic data



Michael Alaux

## 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)



Michael Alaux

## Integration of the sequence to genetic and phenomic data

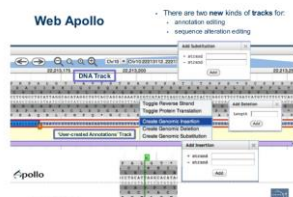
- WheatIS search
  - Gene (WheatMine)
  - Marker (GnpIS-coreDB)
  - GWAS (GnpIS-coreDB)
  - Phenotyping experiments (GnpIS-coreDB)



Michael Alaux

## Perspectives

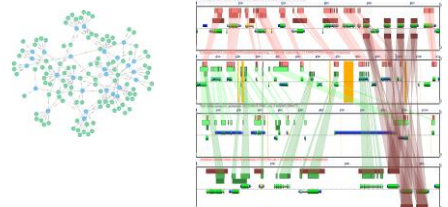
- Management of new/curated versions of the IWGSC RefSeq annotation



Michael Alaux

## Perspectives

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



Michael Alaux

## Acknowledgements



Alaux M.  
Letellier T.  
Flores R.  
Alfama F.  
Pommier C.  
Mohellibi N.  
Durand S.  
Kimmel E.  
Michotey C.  
Guerche C.  
Loaec M.  
Jamiloux 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

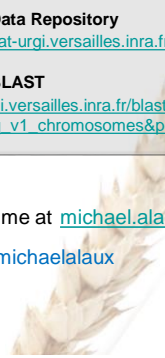
<http://wheat-urgi.versailles.inra.fr/Seq-Repository/Assemblies>

### IWGSC BLAST

[https://urgi.versailles.inra.fr/blast\\_iwgsc/?dbgroup=wheat\\_iwgsc\\_refseq\\_v1\\_chromosomes&program=blastn](https://urgi.versailles.inra.fr/blast_iwgsc/?dbgroup=wheat_iwgsc_refseq_v1_chromosomes&program=blastn)

Contact me at [michael.alaux@inra.fr](mailto:michael.alaux@inra.fr)

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



Michael Alaux



Michael Alaux