A high density GBS map of bread wheat and its application for genetic improvement of the crop

Sukhwinder-Singh
Wheat-Lead Seeds of Discovery (suk.singh@cgiar.org)

International Maize and Wheat Improvement Center (CIMMYT),
Apdo. Postal 6-641, 06600, Mexico DF, Mexico

PAG: Jan. 10, 2015 Posters # 0594 &P0596

SeeD Wheat - Team Players

Sukhwinder-Singh
Parshant Vikram
Huihui Li
Carolina Saint-Pierre
Tom Payne
Juan Burgueño
Carolina Sansaloni
César Petroli
Gilberto Salinas
Peter Wenzl
Karina , Laura, Rosaura, Jose
Luis, Jessica, Cynthia, Penny

Matthew Reynolds: *heat/drought*
Pawan Singh: *Diseases*
Carlos/Velu: *Quality*
Ravi Prakash Singh: *Elite lines and populations*
Deepmala: *Allele mining*
Ivan Ortíz-Monasterio: *P efficiency*
Marta Lopes: *heat/drought in WW wheat*

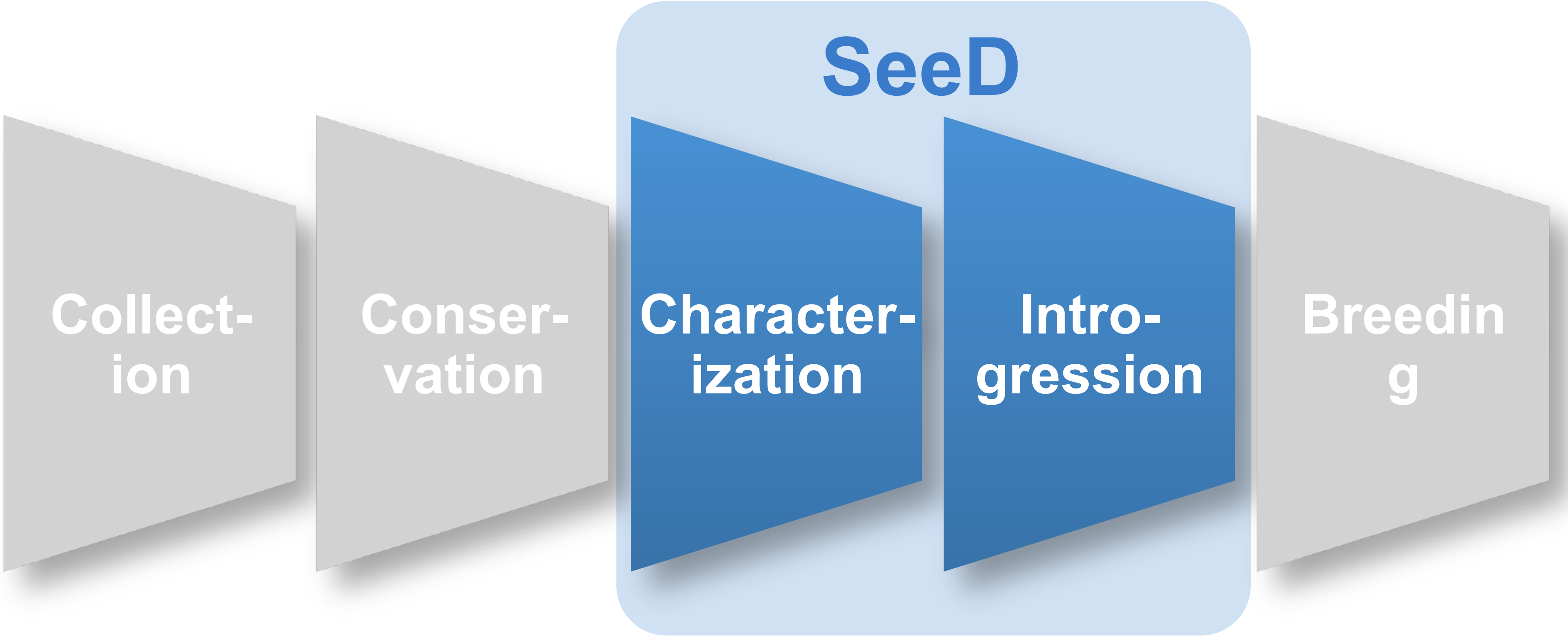
Pedro Figueroa (INIFAP): *heat/drought*
Guillermo Fuentes (INIFAP)*: karnal bunt*
Andrzej Kilian (DArT): *genotyping, software*
David Marshall (JHI): *software tools*

Kevin Pixley: *Director GRP,* **Hans Braun:** *Director GWP*

OUTLINE

- Over-view: SeeD-Wheat
- High density GBS map
- GWAS: examples
- Conclusion

The Scope of SeeD



SeeD

Collect-
ion

Conser-
vation

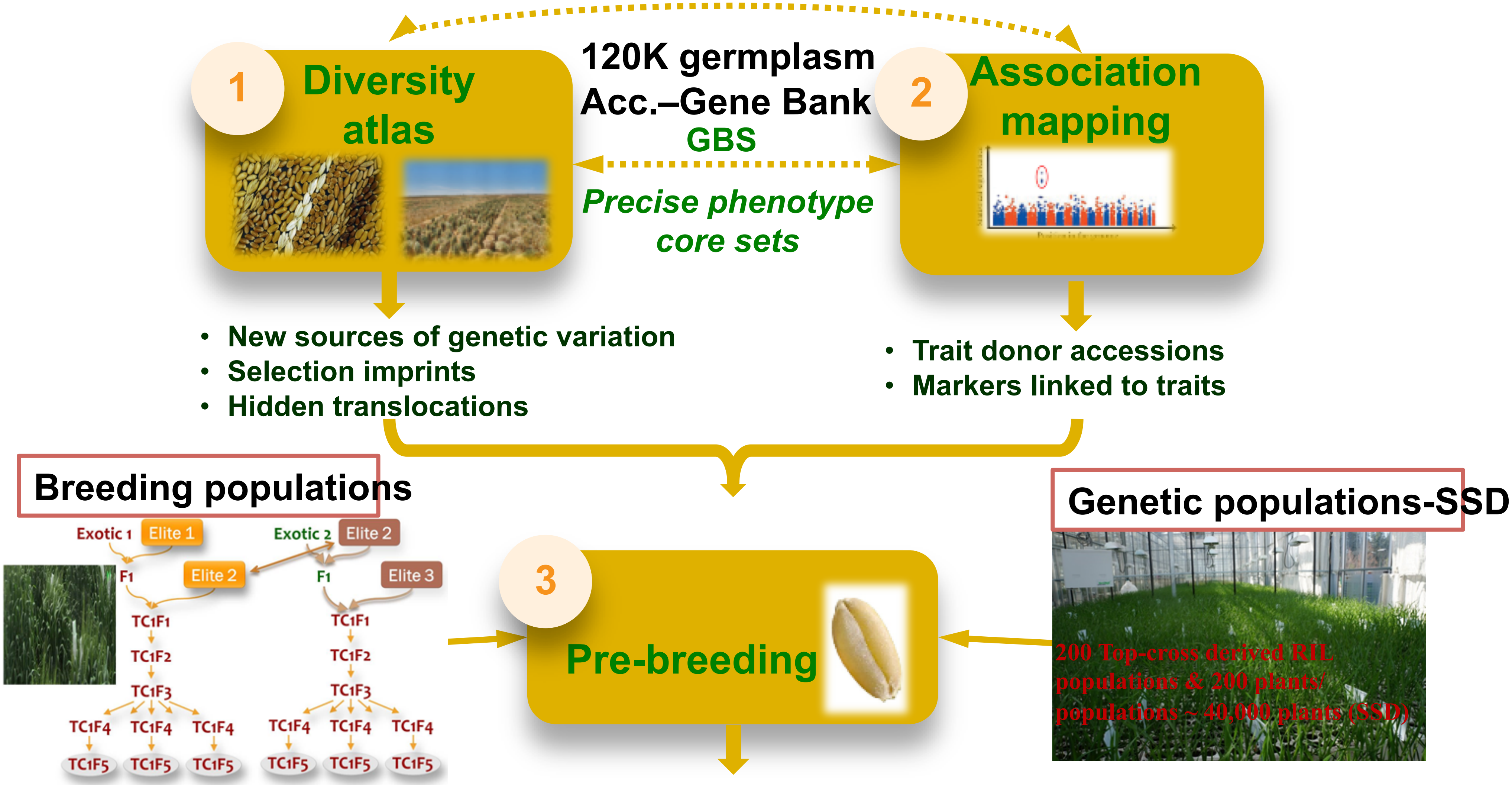
Character-
ization

Intro-
gression

Breedin
g

Strategic focus on "characterization <u>for use in breeding</u>" and pre-breeding (priority traits = breeding targets)

SeeD-Wheat: Overview



1 **Diversity atlas**

120K germplasm
Acc.–Gene Bank
GBS

Precise phenotype core sets

2 **Association mapping**

- New sources of genetic variation
- Selection imprints
- Hidden translocations

- Trait donor accessions
- Markers linked to traits

Breeding populations

Genetic populations-SSD

200 Top-cross derived RIL populations & 200 plants/ populations ~ 40,000 plants (SSD)

3 **Pre-breeding**

- New beneficial alleles for priority traits in elite genetic backgrounds
- Molecular markers linked to beneficial alleles and statistical models for estimating breeding values to accelerate genetic progress in breeding programs

Resource for genomic selection & nested mapping

Anchoring GBS tags on chromosomes

Population	Number of markers	Number of individuals
PBW343 x Kingbird	13123	198
PBW343 x MUU	6936	148
PBW343 x K.Swara	18612	109

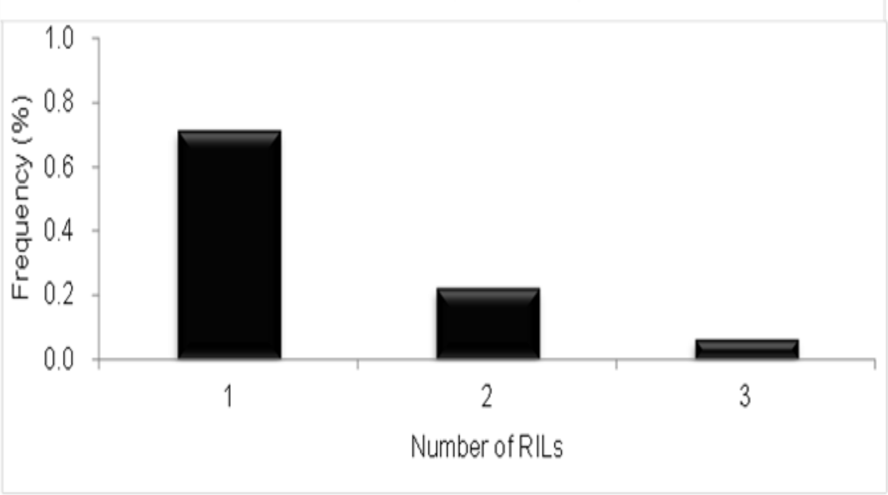
Genetic linkage mapping to construct a consensus map containing 28644 GBS markers.

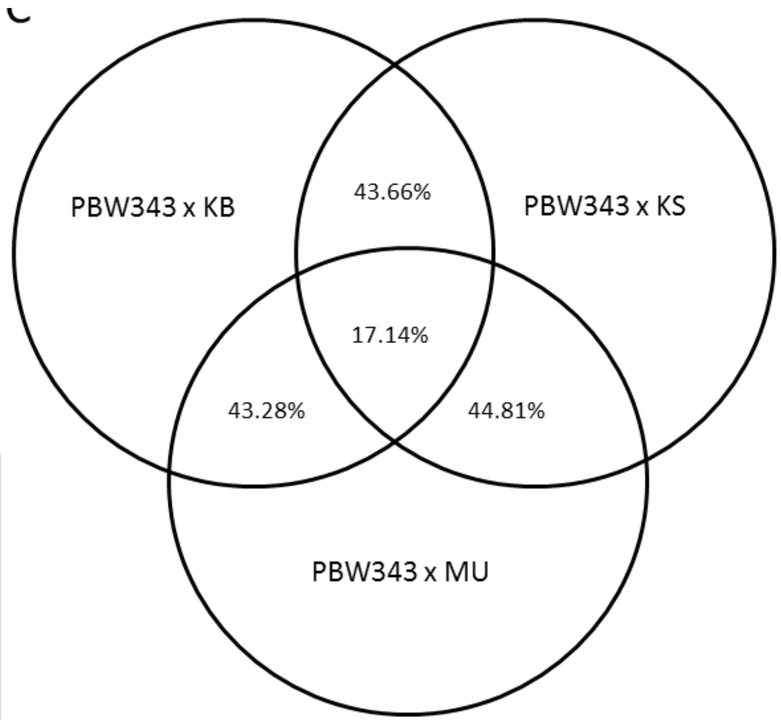
Consensus map – 3757 unique positions

Total length of the map – 3302.5 cM

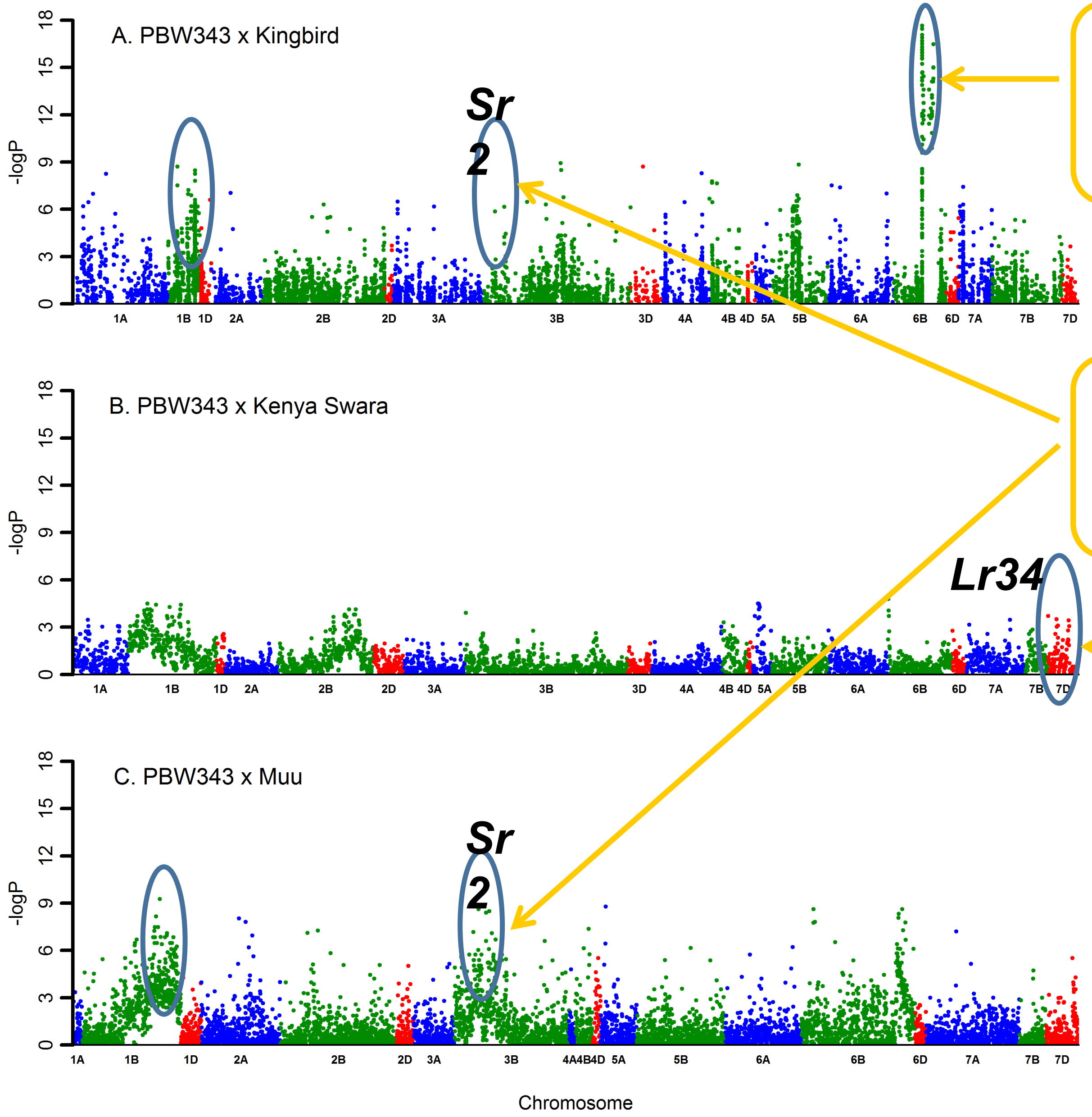
CIMMYT MR

Proportion of common GBS markers across three populations





Distribution of segregation distortion loci across three RIL mapping populations



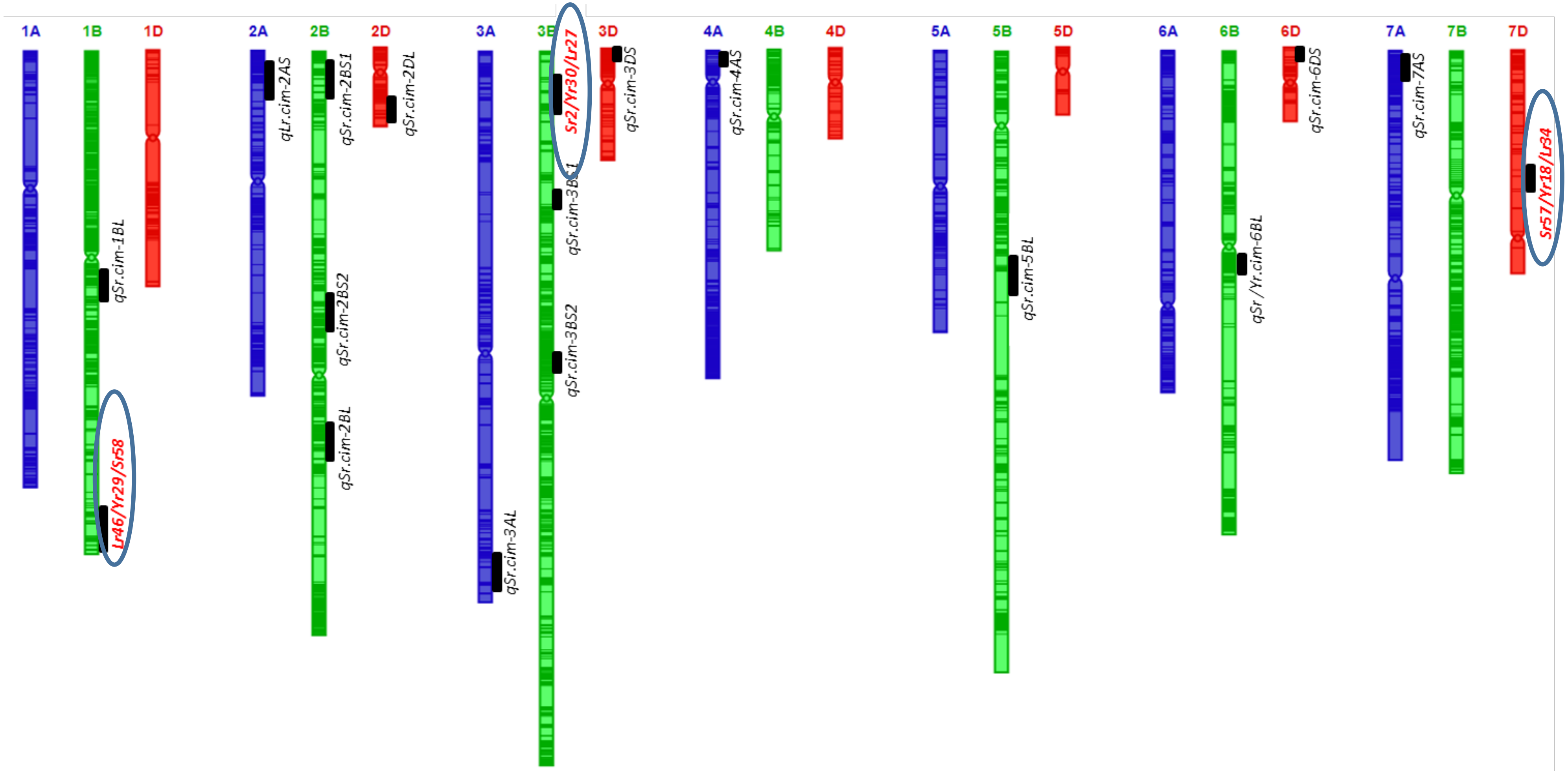
selection favored alleles from PBW343

selection favored alleles from PBW343

selection favored alleles from Kenya Swara

Validation of GBS map & Genetic dissection of complex traits in wheat

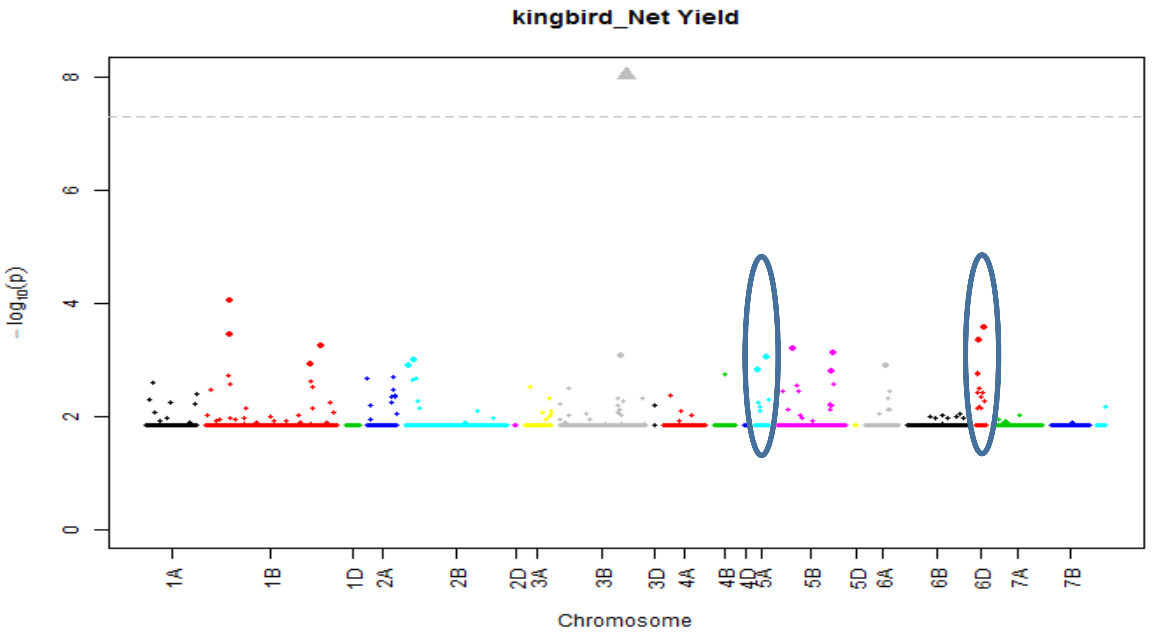
Identified chromosomal regions (QTLs) harboring APR to wheat rusts

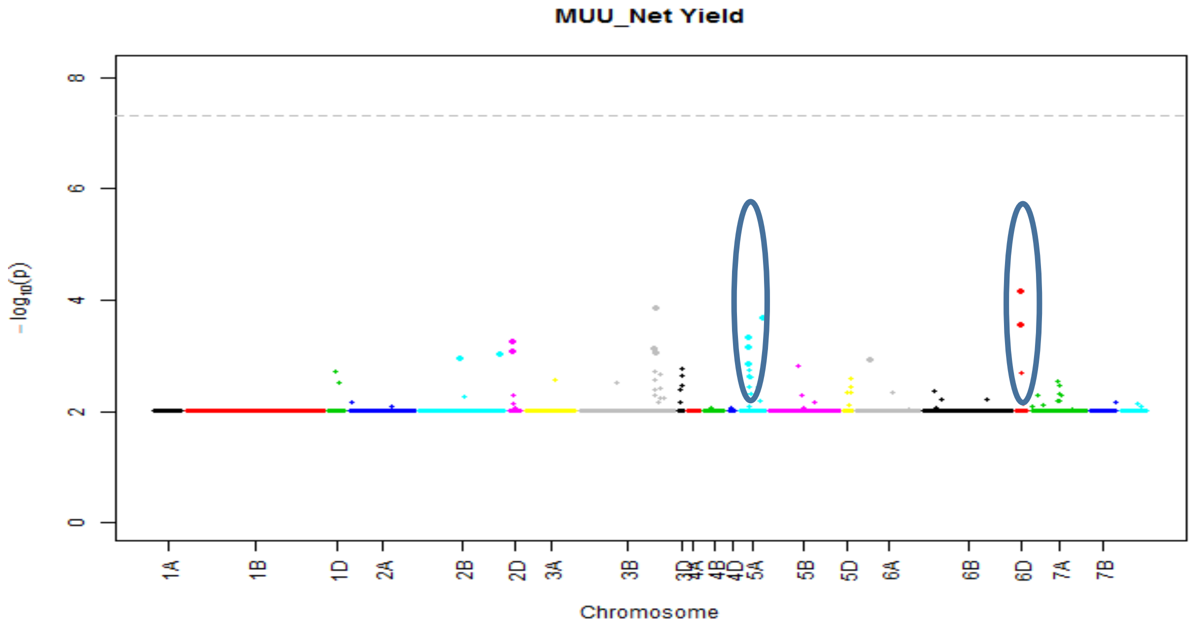


Identified chromosomal regions (QTLs) harboring TS resistance-PBW343/Kingbird population

Trait	Chro		Position	Left-Marker	Right-Marker	LOD	PVE(%)	Add
TS	5B		263	1100643|F|0	1069940|F|0	21.6214	50.6367	-0.3248
TS	5A		124	2326786|F|0	1864591|F|0	4.4252	7.537	0.1238
TS	1B		93	3025657|F|0	1130776|F|0	4.1395	7.0445	-0.1201
TS	3A		56	1101113|F|0	3029754|F|0	4.6102	7.9568	-0.1275
TS	7A		165	1226998|F|0	1128339|F|0	13.0078	25.6901	0.2291

Genomic associations- Yield in PBW343 derived populations





Blast the GBS markers against the genome sequence of Chinese Spring

- BLASTN 2.2.26
- E-value < 1e-4

SPECIAL SECTION *SLICING THE WHEAT GENOME*

A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome

The International Wheat Genome Sequencing Consortium (IWGSC)

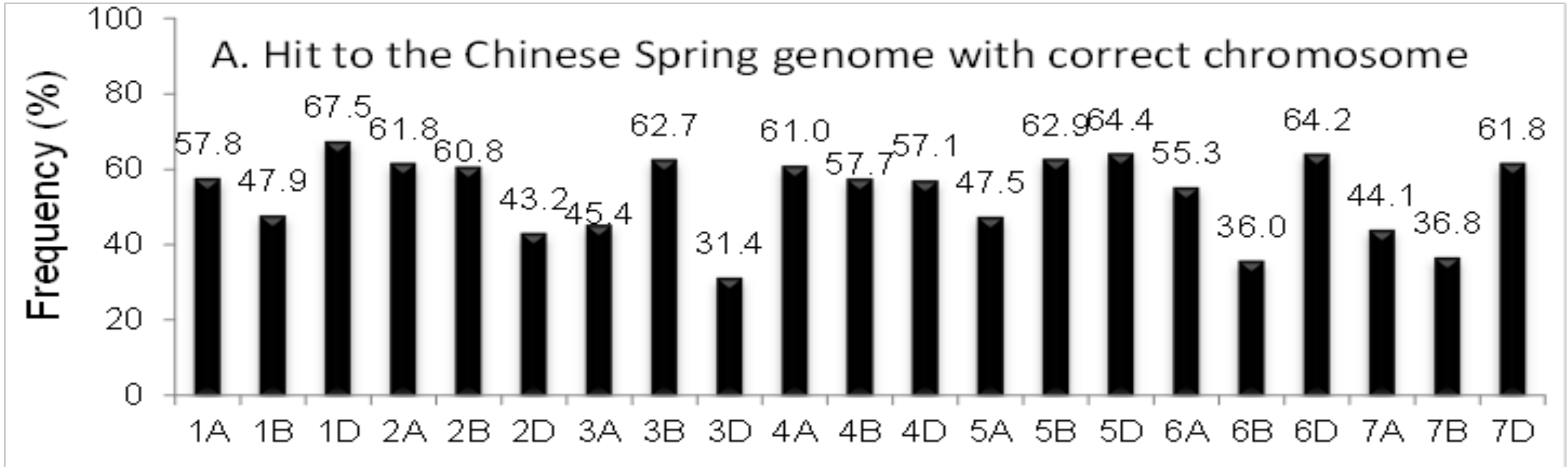
An ordered draft sequence of the 17-gigabase hexaploid bread wheat (*Triticum aestivum*) genome has been produced by sequencing isolated chromosome arms. We have annotated 124,201 gene loci distributed nearly evenly across the homeologous chromosomes and subgenomes. Comparative gene analysis of wheat subgenomes and extant diploid and tetraploid wheat relatives showed that high sequence similarity and structural conservation are retained, with limited gene loss, after polyploidization. However, across the genomes there was evidence of dynamic gene gain, loss, and duplication since the divergence of the wheat lineages. A high degree of transcriptional autonomy and no global dominance was found for the subgenomes. These insights into the genome biology of a polyploid crop provide a springboard for faster gene isolation, rapid genetic marker development, and precise breeding to meet the needs of increasing food demand worldwide.

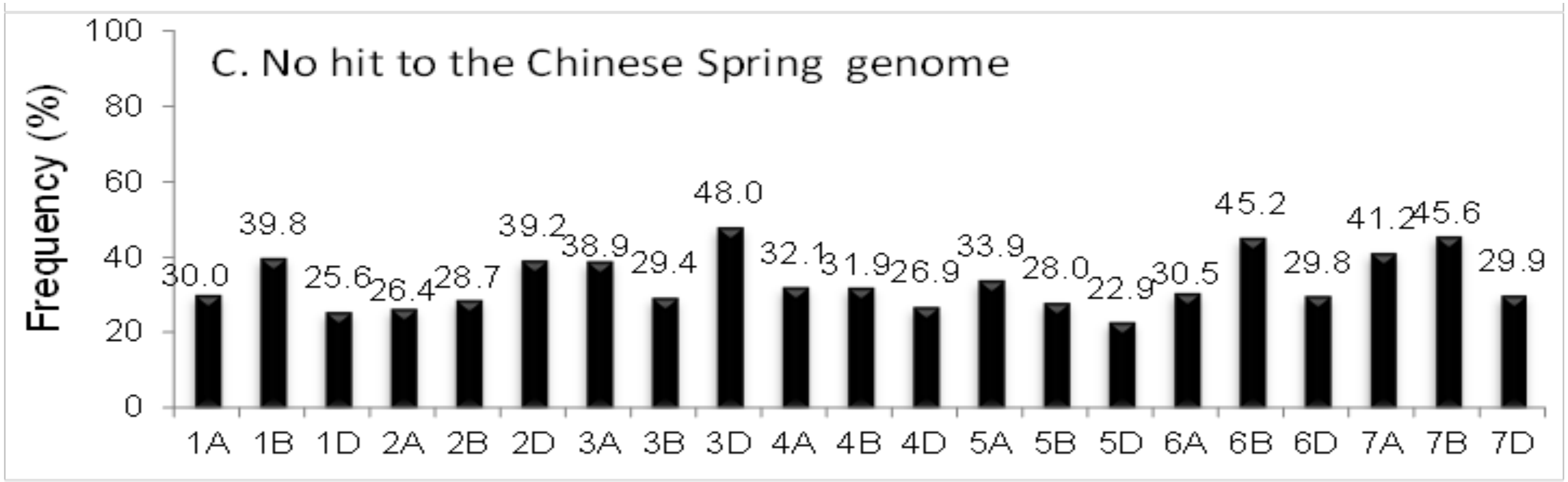
Lists of authors and affiliations are available in the full article online.

Corresponding author: K. X. Mayer, e-mail: k.mayer@helmholtz-muenchen.de

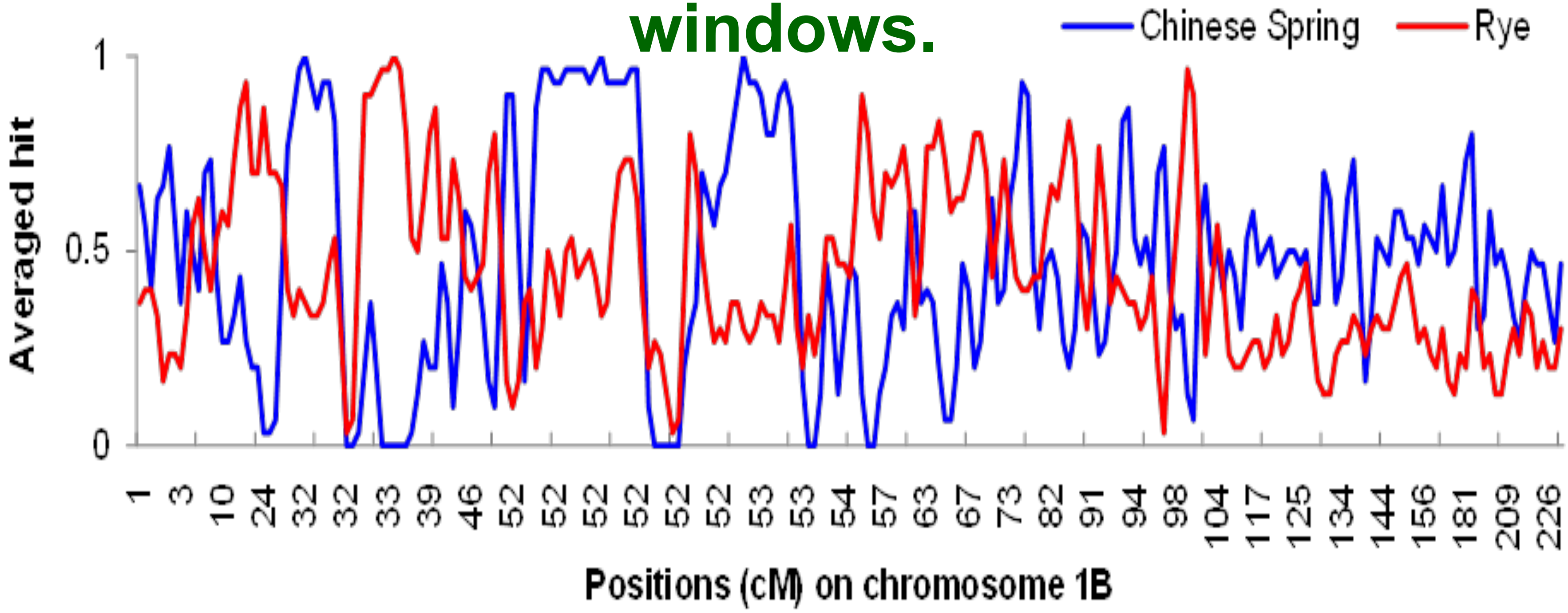
Read the full article at http://dx.doi.org/10.1126/science.1251788

Frequency of the sequences of GBS markers -in the Chinese Spring genome.



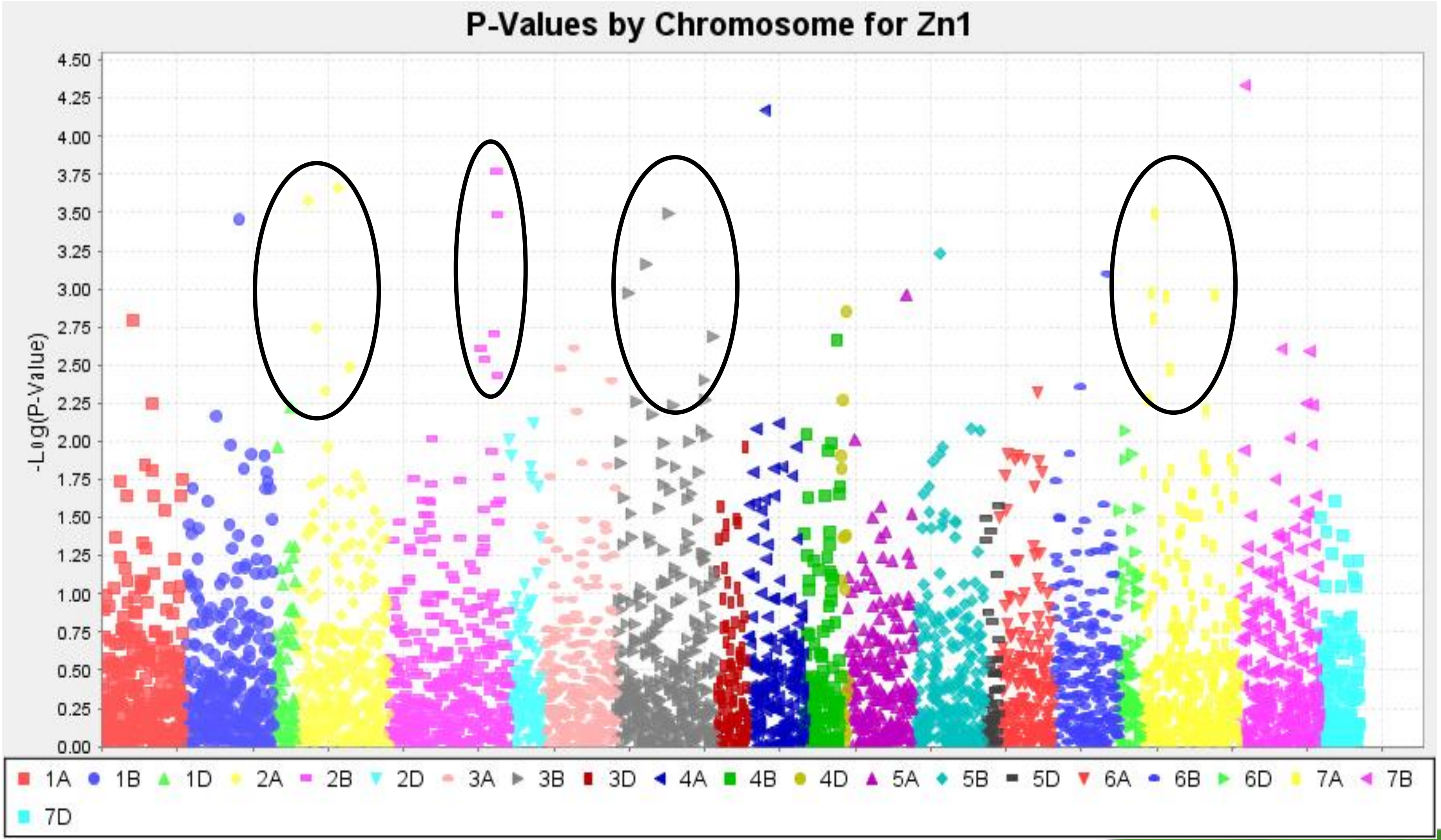


Averaged hits to the Chinese Spring genome, and rye genome on across chromosome 1B within sliding window with 30 markers in length and 15 markers overlapped between neighboring windows.



Clear enrichment for rye hits indicating known IB/IR translocation

Iranian landrace: GWAS for Zn



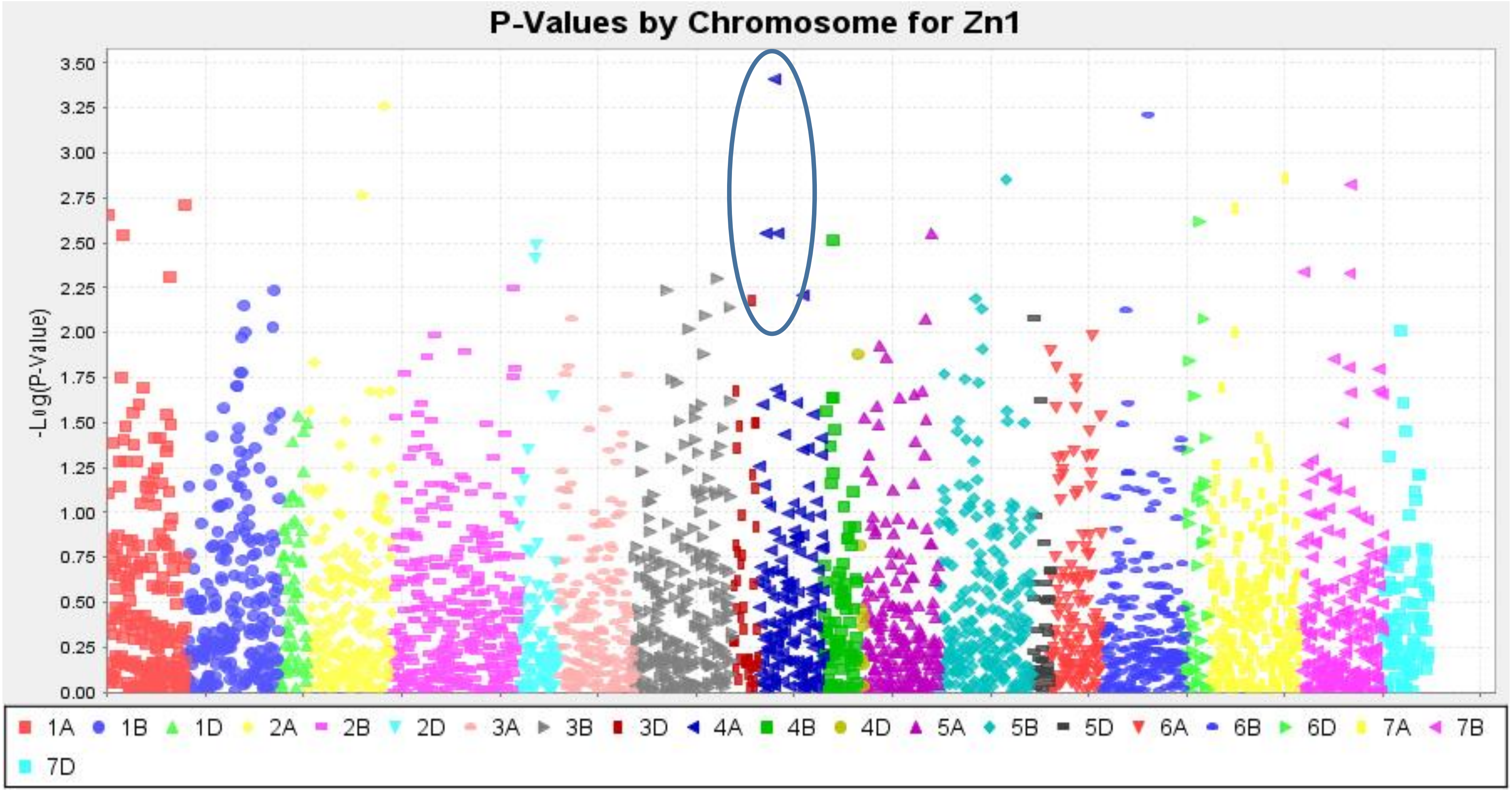
3A

1125892|F|0--31:A>G
1092240|F|0--18:C>T
999525|F|0--36:C>T

4A

1115172|F|0--19:G>T
1228831|F|0--47:C>T
1076151|F|0--41:C>G

Mexican landrace: GWAS for Zn



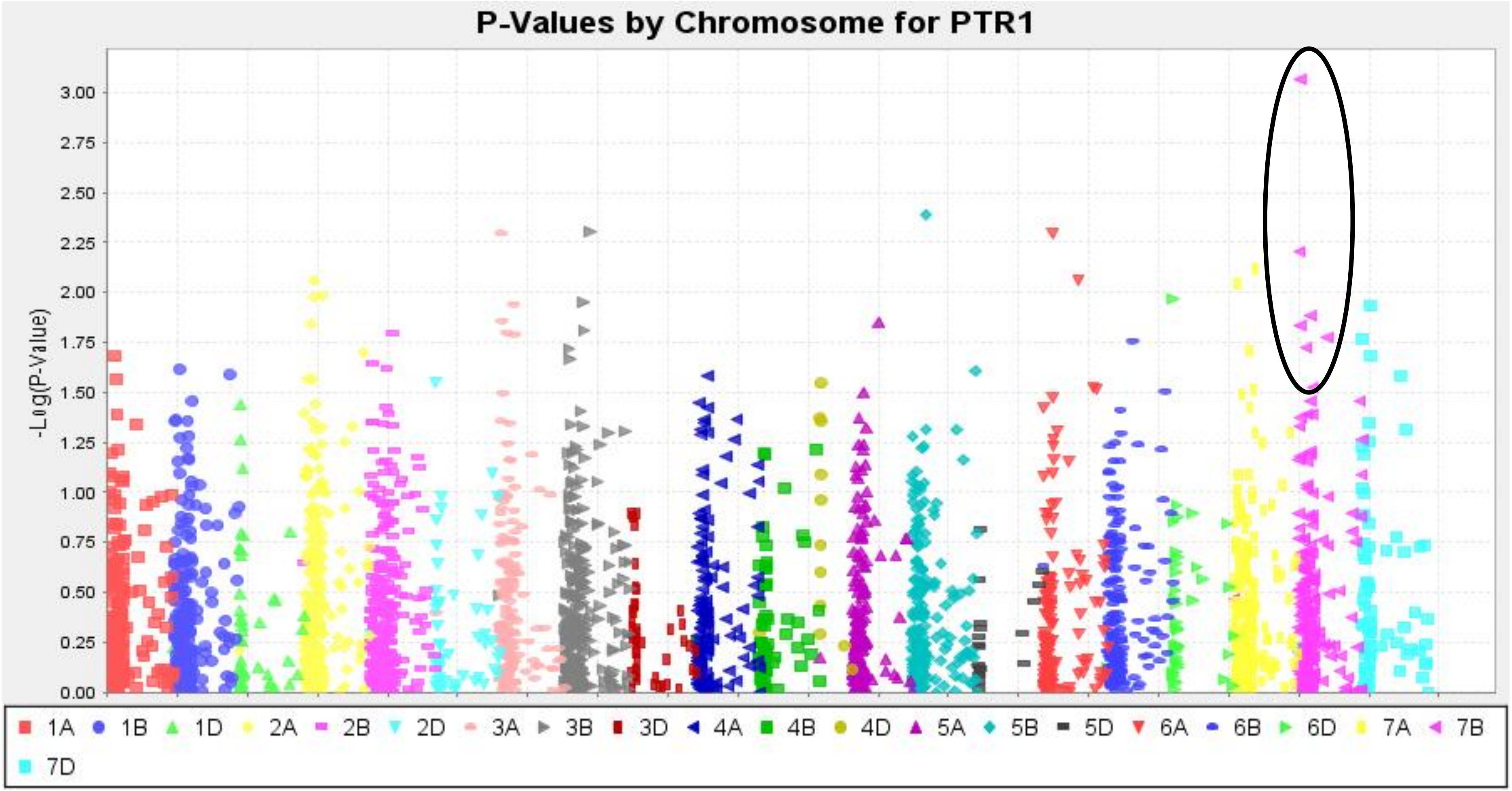
4A

1267597|F|0
1270934|F|0
1271661|F|0

5B

1107049|F|0--21:C>G
982062|F|0--55:G>A
2267710|F|0--66:G>C

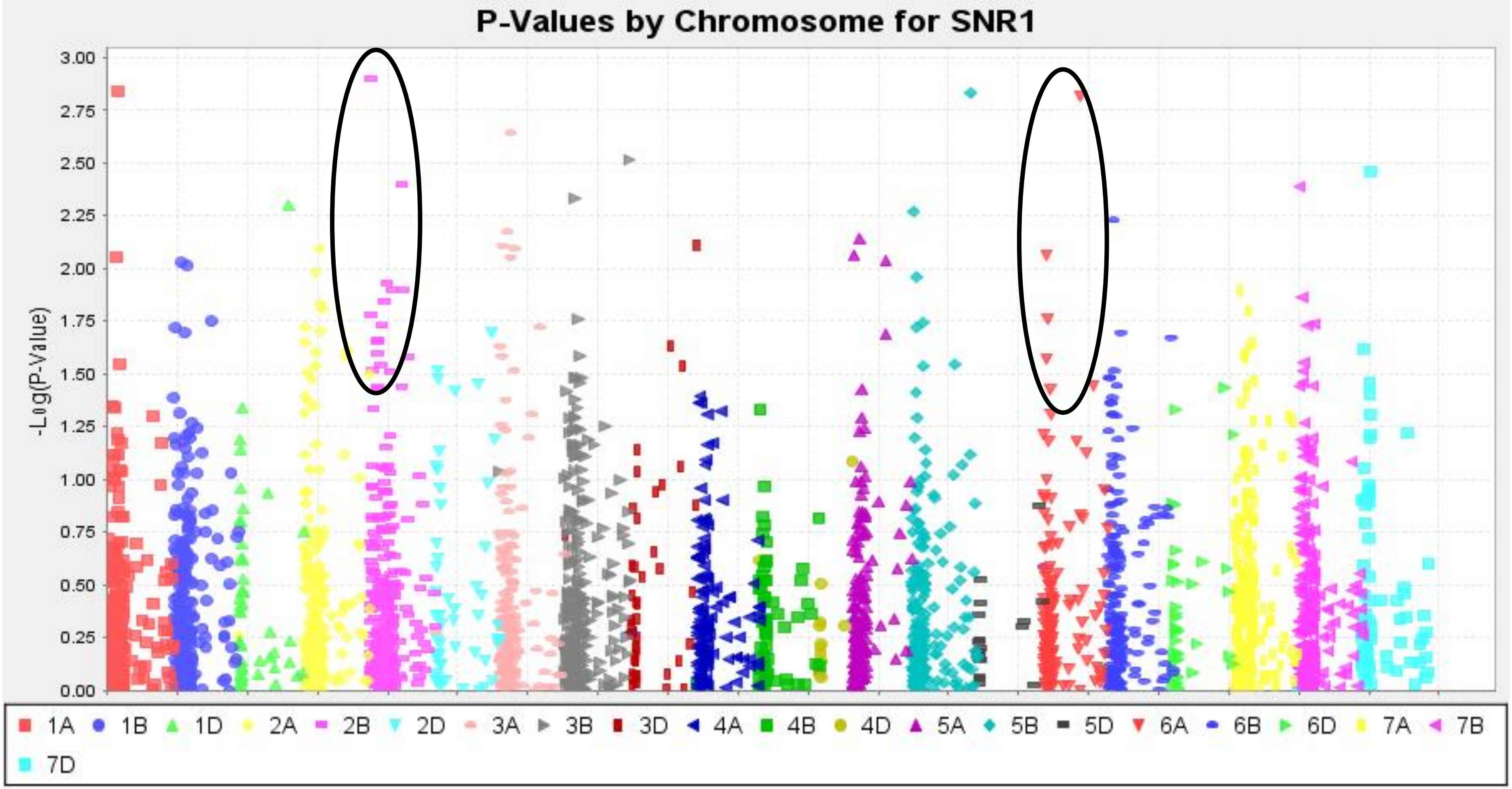
Iranian landrace: GWAS for Tan spot



7B

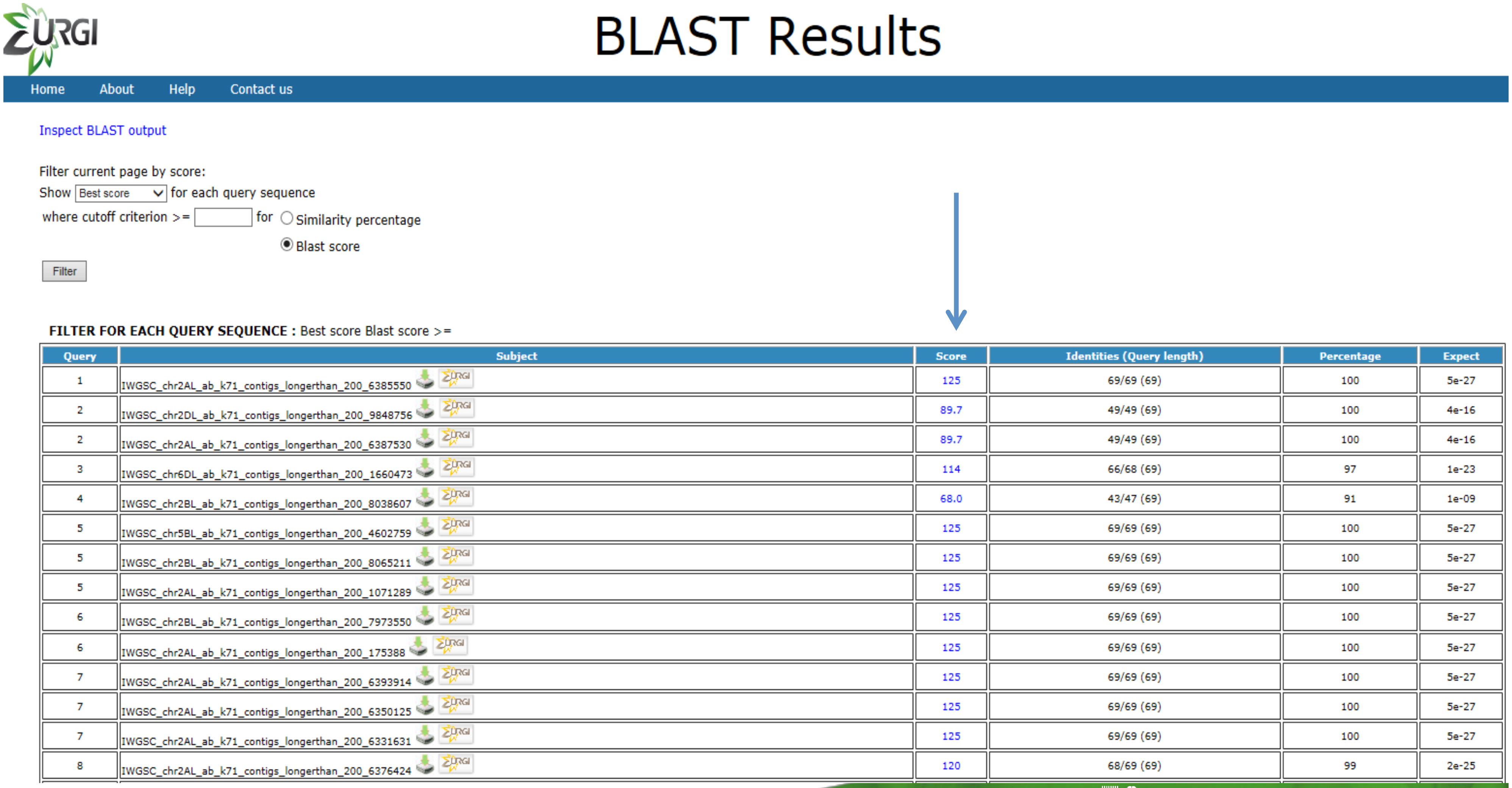
987702|F|0--31:G>A
2279071|F|0--23:C>T

Iranian landrace: GWAS for SBN



2B	1107425|F|0--6:G>C 1119788|F|0
6A	1108888|F|0 1102743|F|0--51:A>T

GBS query sequences blasted in IWGSC and filtered using best score criterion for each sequence



BLAST Results

Inspect BLAST output

Filter current page by score:

Show Best score for each query sequence

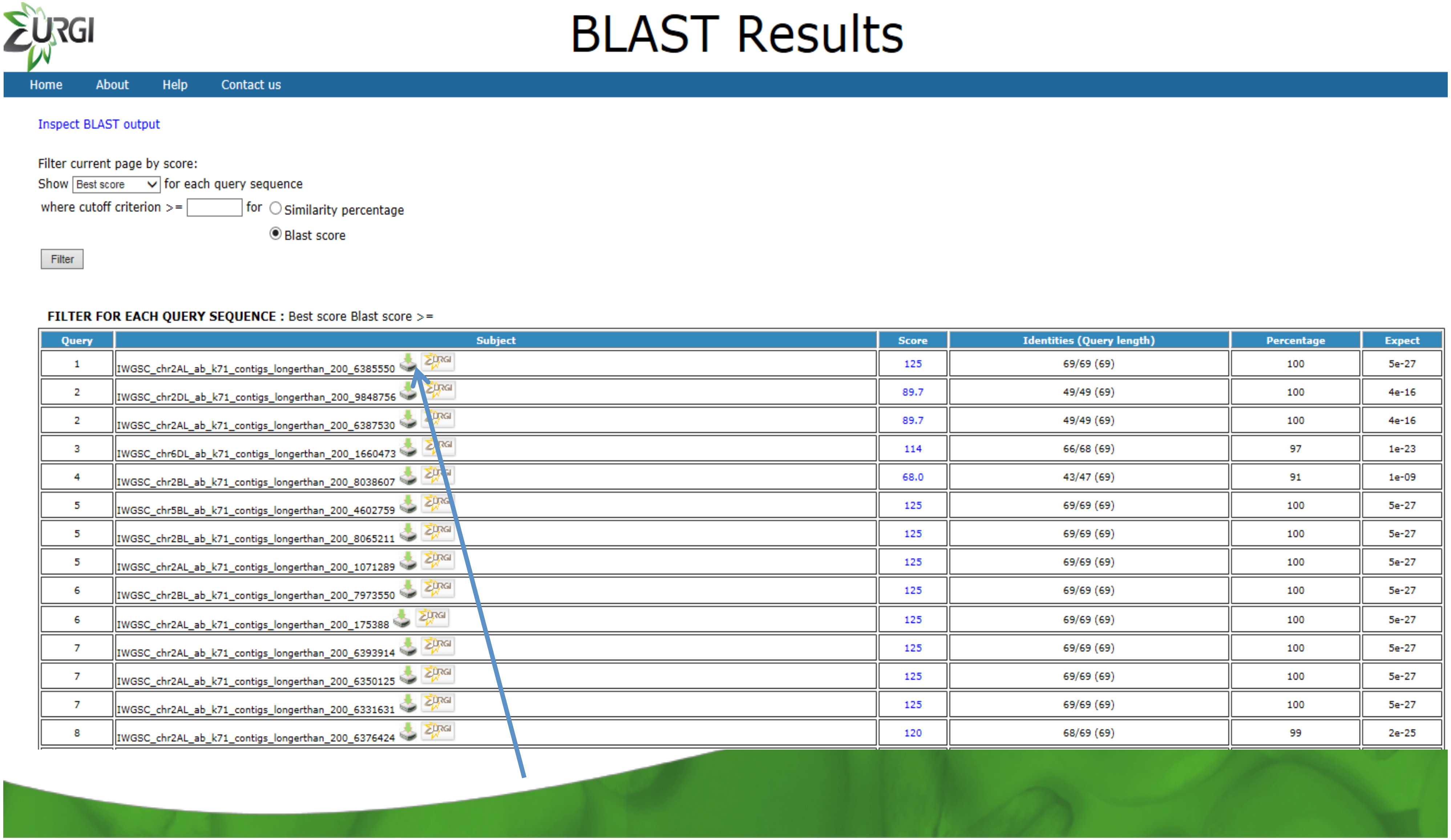
where cutoff criterion >= for ○ Similarity percentage

● Blast score

FILTER FOR EACH QUERY SEQUENCE : Best score Blast score >=

Query	Subject	Score	Identities (Query length)	Percentage	Expect
1	IWGSC_chr2AL_ab_k71_contigs_longerthan_200_6385550	125	69/69 (69)	100	5e-27
2	IWGSC_chr2DL_ab_k71_contigs_longerthan_200_9848756	89.7	49/49 (69)	100	4e-16
2	IWGSC_chr2AL_ab_k71_contigs_longerthan_200_6387530	89.7	49/49 (69)	100	4e-16
3	IWGSC_chr6DL_ab_k71_contigs_longerthan_200_1660473	114	66/68 (69)	97	1e-23
4	IWGSC_chr2BL_ab_k71_contigs_longerthan_200_8038607	68.0	43/47 (69)	91	1e-09
5	IWGSC_chr5BL_ab_k71_contigs_longerthan_200_4602759	125	69/69 (69)	100	5e-27
5	IWGSC_chr2BL_ab_k71_contigs_longerthan_200_8065211	125	69/69 (69)	100	5e-27
5	IWGSC_chr2AL_ab_k71_contigs_longerthan_200_1071289	125	69/69 (69)	100	5e-27
6	IWGSC_chr2BL_ab_k71_contigs_longerthan_200_7973550	125	69/69 (69)	100	5e-27
6	IWGSC_chr2AL_ab_k71_contigs_longerthan_200_175388	125	69/69 (69)	100	5e-27
7	IWGSC_chr2AL_ab_k71_contigs_longerthan_200_6393914	125	69/69 (69)	100	5e-27
7	IWGSC_chr2AL_ab_k71_contigs_longerthan_200_6350125	125	69/69 (69)	100	5e-27
7	IWGSC_chr2AL_ab_k71_contigs_longerthan_200_6331631	125	69/69 (69)	100	5e-27
8	IWGSC_chr2AL_ab_k71_contigs_longerthan_200_6376424	120	68/69 (69)	99	2e-25

Sequences with highest Blast score and with 100% identity were downloaded for each sequence and blasted in NCBI



BLAST Results

Inspect BLAST output

Filter current page by score:

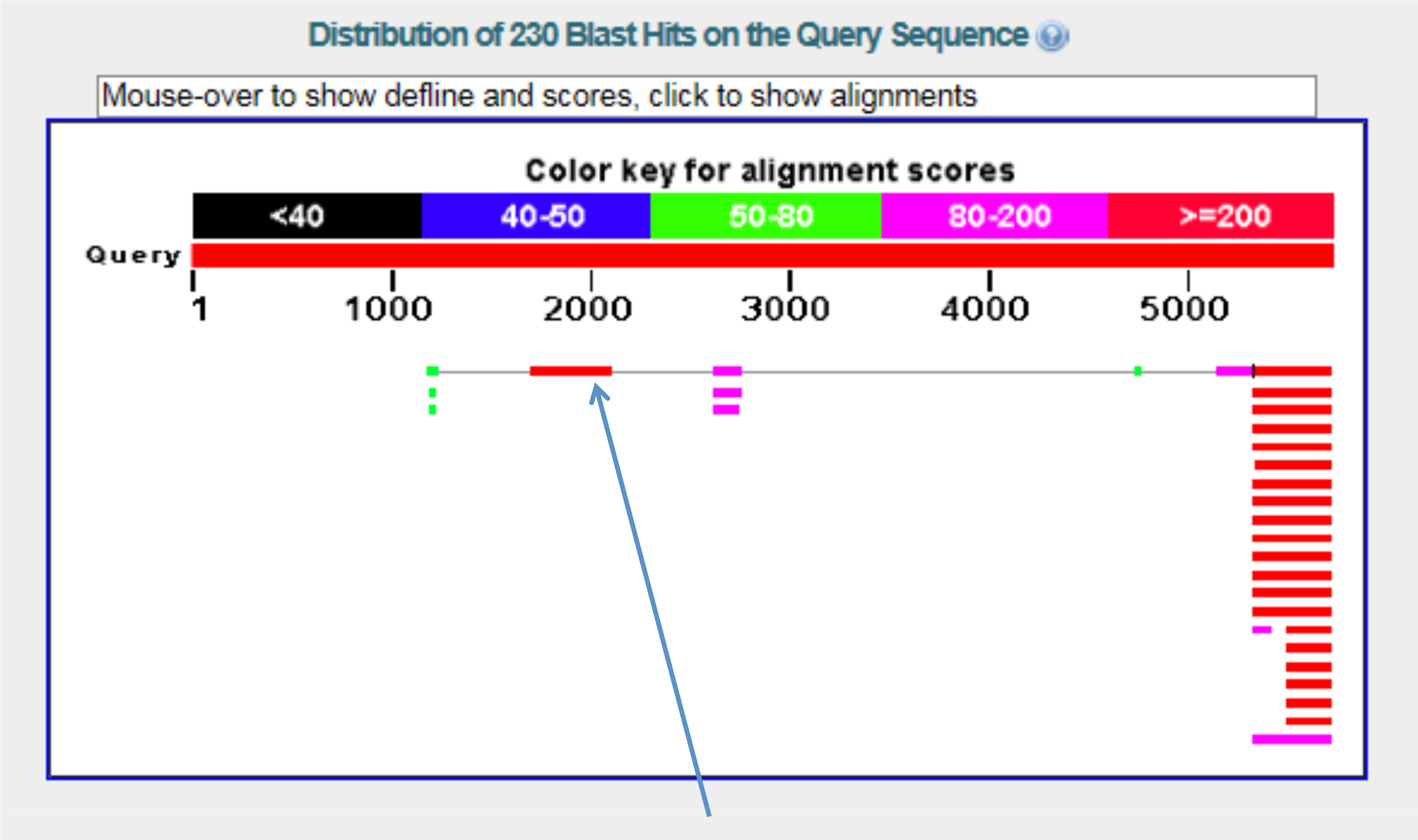
Show Best score for each query sequence

where cutoff criterion >= for ○ Similarity percentage

● Blast score

FILTER FOR EACH QUERY SEQUENCE : Best score Blast score >=

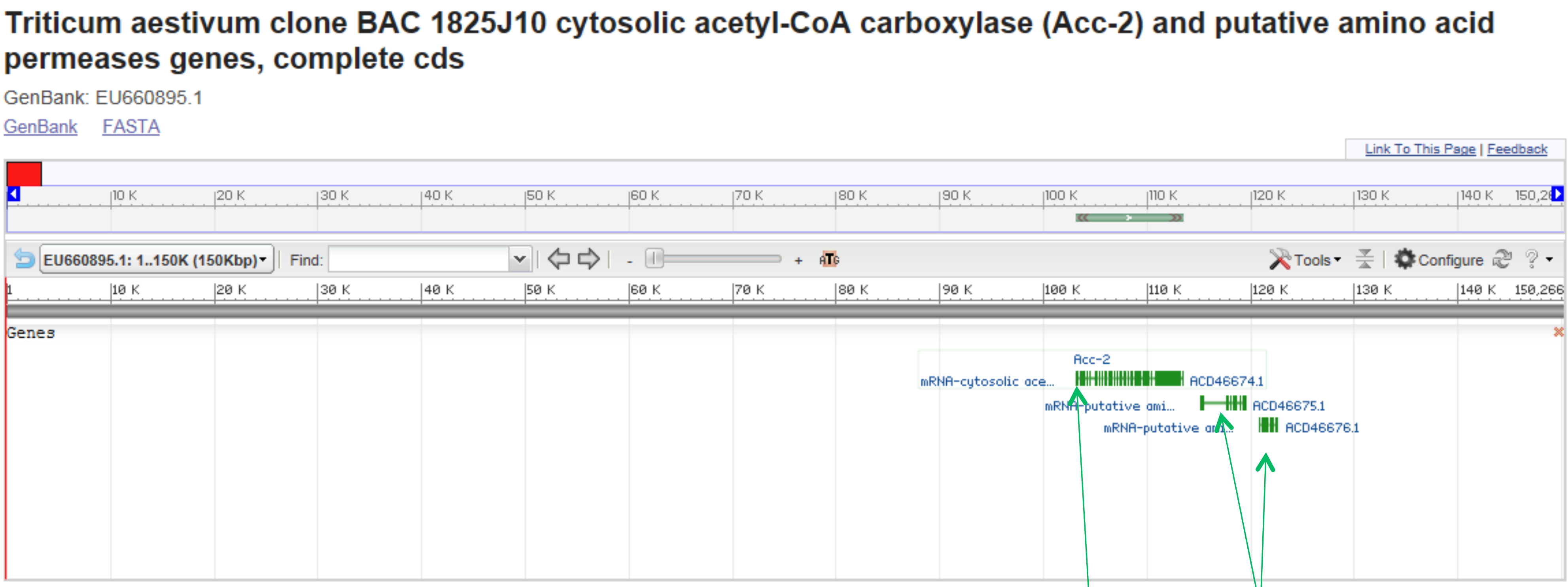
Query	Subject	Score	Identities (Query length)	Percentage	Expect
1	IWGSC_chr2AL_ab_k71_contigs_longerthan_200_6385550	125	69/69 (69)	100	5e-27
2	IWGSC_chr2DL_ab_k71_contigs_longerthan_200_9848756	89.7	49/49 (69)	100	4e-16
2	IWGSC_chr2AL_ab_k71_contigs_longerthan_200_6387530	89.7	49/49 (69)	100	4e-16
3	IWGSC_chr6DL_ab_k71_contigs_longerthan_200_1660473	114	66/68 (69)	97	1e-23
4	IWGSC_chr2BL_ab_k71_contigs_longerthan_200_8038607	68.0	43/47 (69)	91	1e-09
5	IWGSC_chr5BL_ab_k71_contigs_longerthan_200_4602759	125	69/69 (69)	100	5e-27
5	IWGSC_chr2BL_ab_k71_contigs_longerthan_200_8065211	125	69/69 (69)	100	5e-27
5	IWGSC_chr2AL_ab_k71_contigs_longerthan_200_1071289	125	69/69 (69)	100	5e-27
6	IWGSC_chr2BL_ab_k71_contigs_longerthan_200_7973550	125	69/69 (69)	100	5e-27
6	IWGSC_chr2AL_ab_k71_contigs_longerthan_200_175388	125	69/69 (69)	100	5e-27
7	IWGSC_chr2AL_ab_k71_contigs_longerthan_200_6393914	125	69/69 (69)	100	5e-27
7	IWGSC_chr2AL_ab_k71_contigs_longerthan_200_6350125	125	69/69 (69)	100	5e-27
7	IWGSC_chr2AL_ab_k71_contigs_longerthan_200_6331631	125	69/69 (69)	100	5e-27
8	IWGSC_chr2AL_ab_k71_contigs_longerthan_200_6376424	120	68/69 (69)	99	2e-25



Distribution of 230 Blast Hits on the Query Sequence

Mouse-over to show defline and scores, click to show alignments

Region containing GBS query sequence



Triticum aestivum clone BAC 1825J10 cytosolic acetyl-CoA carboxylase (Acc-2) and putative amino acid permeases genes, complete cds

GenBank: EU660895.1

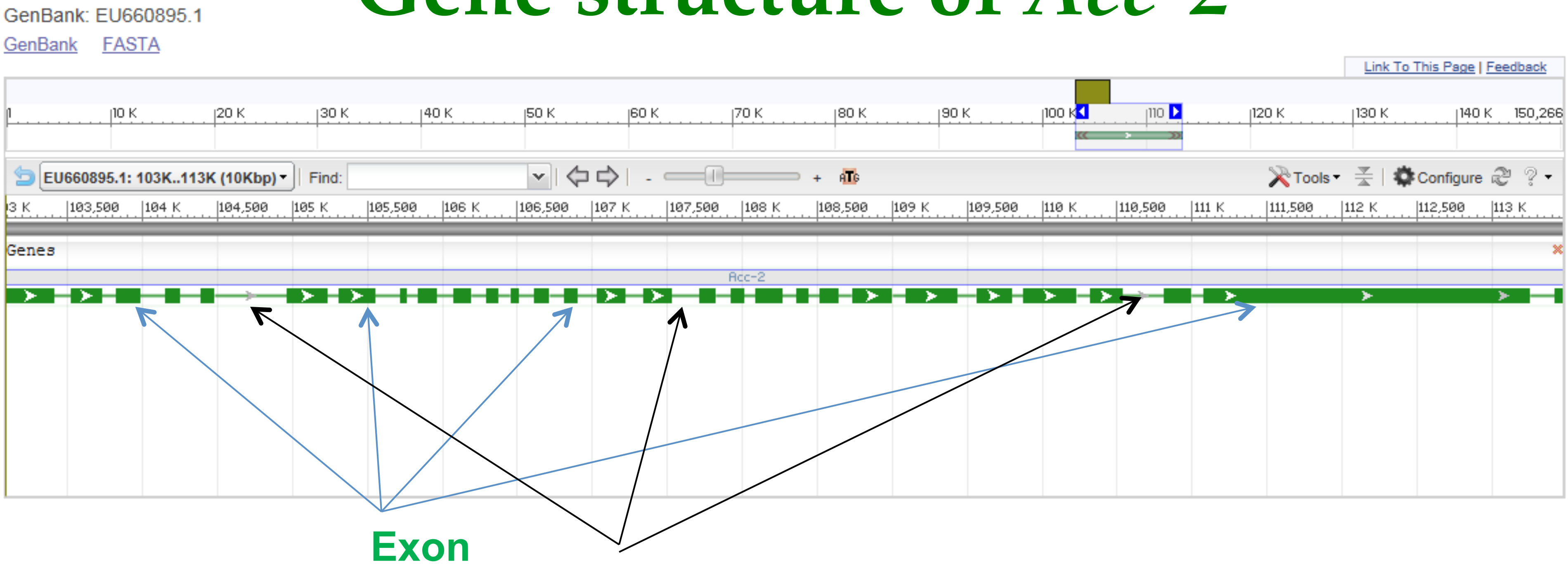
GenBank FASTA

Acetyl-CoA carboxylase (*Acc-2*)

Amino acid permease (*AAP*)

BAC clone containing putative gene(s) identified; putative acetyl-CoA carboxylase (*Acc-2*) and amino acid permeases (*AAP*)

Gene structure of *Acc-2*

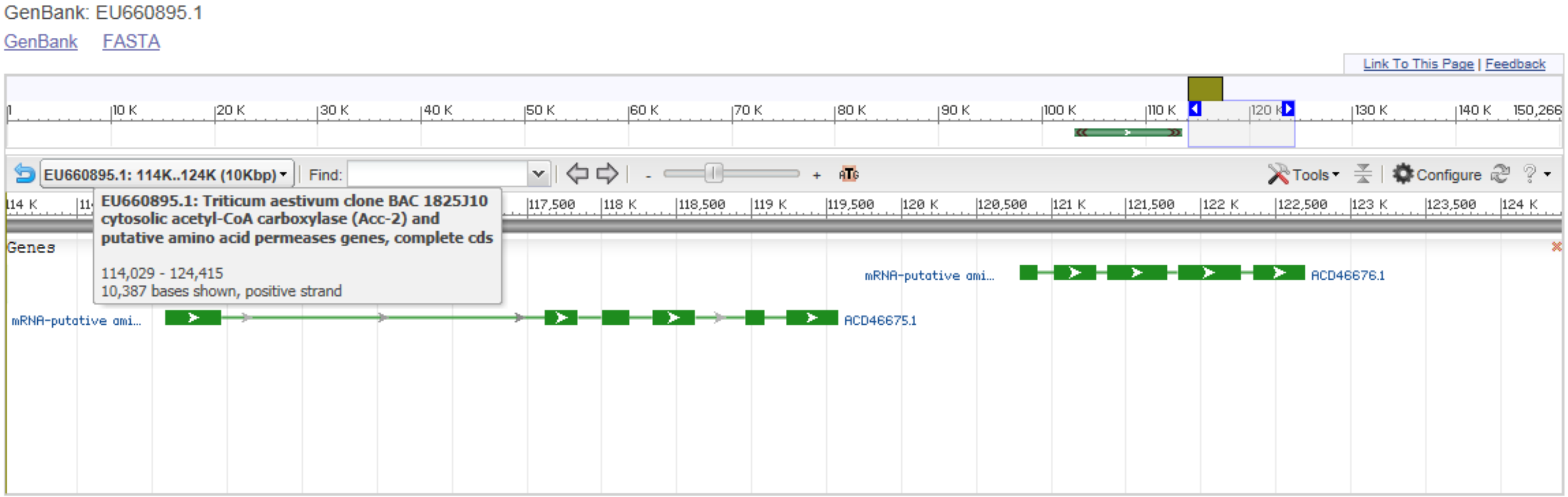


Exons

Introns

The gene Acc-2 has 29 exons and 28 introns. *Acc-2* is associated with grain yield, grain harvest index and panicle yield under irrigated and drought stress conditions in pearl millet (Sehgal et al. 2015 - PLOS One)

Gene structure of AAPs



GenBank: EU660895.1

GenBank FASTA

EU660895.1: 114K..124K (10Kbp)

Find:

EU660895.1: Triticum aestivum clone BAC 1825J10 cytosolic acetyl-CoA carboxylase (Acc-2) and putative amino acid permeases genes, complete cds

114,029 - 124,415

10,387 bases shown, positive strand

The AAP genes have 5 or 6 exons and 4 or 5 introns. *AAPs* are the integral membrane proteins which mediate the transport of amino acids across cellular membranes in higher plants, and play an indispensable role in various processes of plant growth and development, including long distance amino acid transport, response to pathogen and abiotic stresses (Zhao et al. 2012 PLOS One).

Conclusion

- High-density genome profiles, generated using the GBS platform, have the potential to stimulate genetic improvement of wheat

- Validated known genes controlling disease resistance; and detected new genetic regions for resistance.

- Facilitate genetic dissection of complex trait in wheat and genomics-assisted breeding.