HelmholtzZentrum münchen
German Research Center for Environmental Health

Triticeae genome resources @ MIPS PlantsDB

Manuel Spannagl

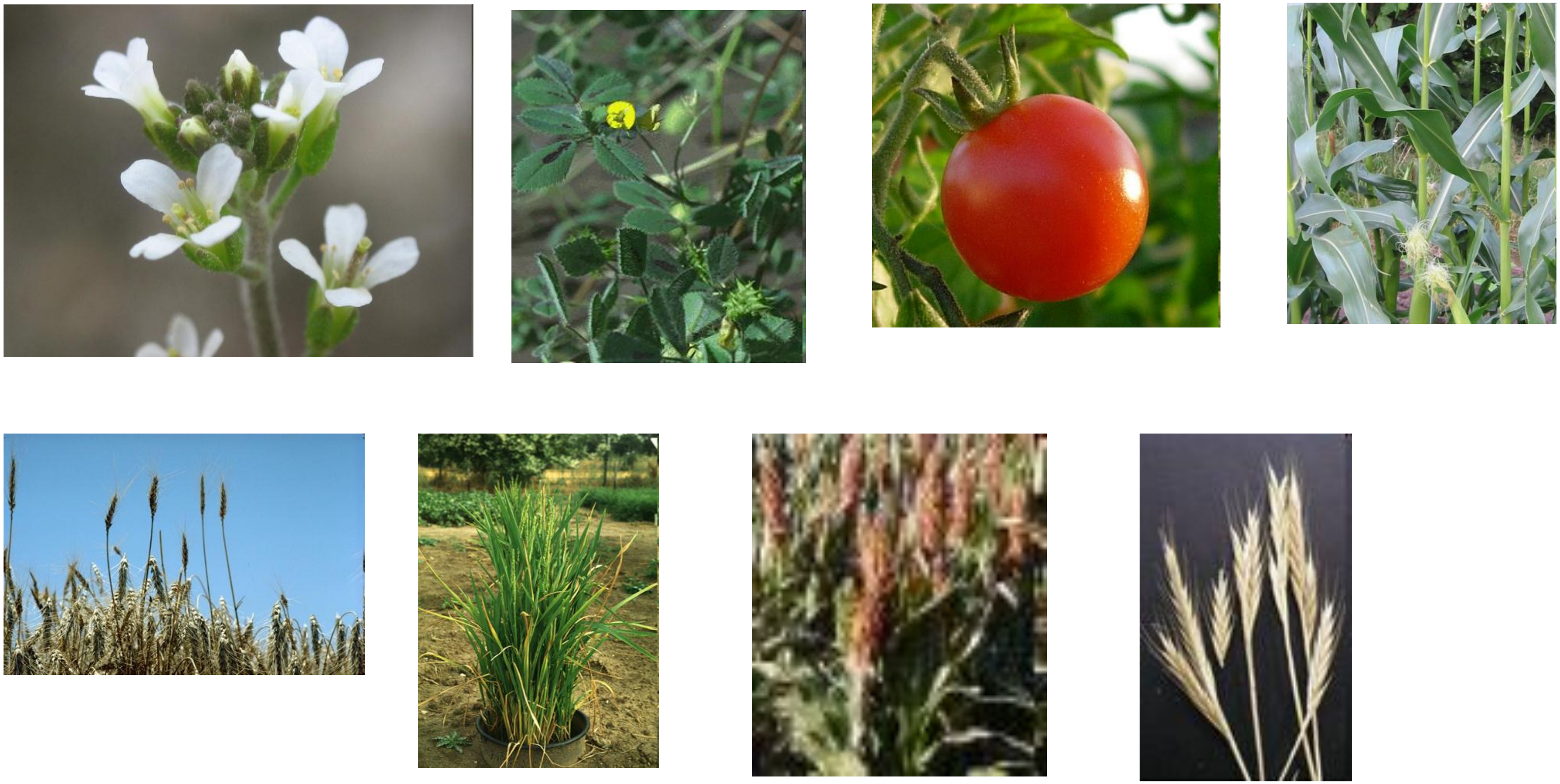
transPLANT workshop – Versailles Nov 2012

What is MIPS PlantsDB?

Generic database schema+system for the integration, management and (comparative) analysis of plant genome data

- modular design, extendable
- re-usable, flexible schema
- strong inter-connection with external databases
- standardized data retrieval
- tools available

MIPS PlantsDB – data content (Nov 2012)



…and many more internal instances.

What is stored in MIPS PlantsDB?

- whole genome sequences: pseudo-chromosome sequences if available, otherwise sequence contigs (also raw or assembled reads)
- gene and transcript annotation on these sequences
- non-coding transcripts/RNA, regulatory elements
- markers and genetic and physical maps
- repetitive elements and its annotation
- gene supporting evidences: fl-cDNAs, ESTs etc.
- transcriptomics data: RNA-seq
- comparative genome analysis results

What is NOT (yet) stored in MIPS PlantsDB?

- primary sequence data
- re-sequencing data and its results (under construction)
- phenotypic information
- metabolic/pathway information
- no community annotation interface

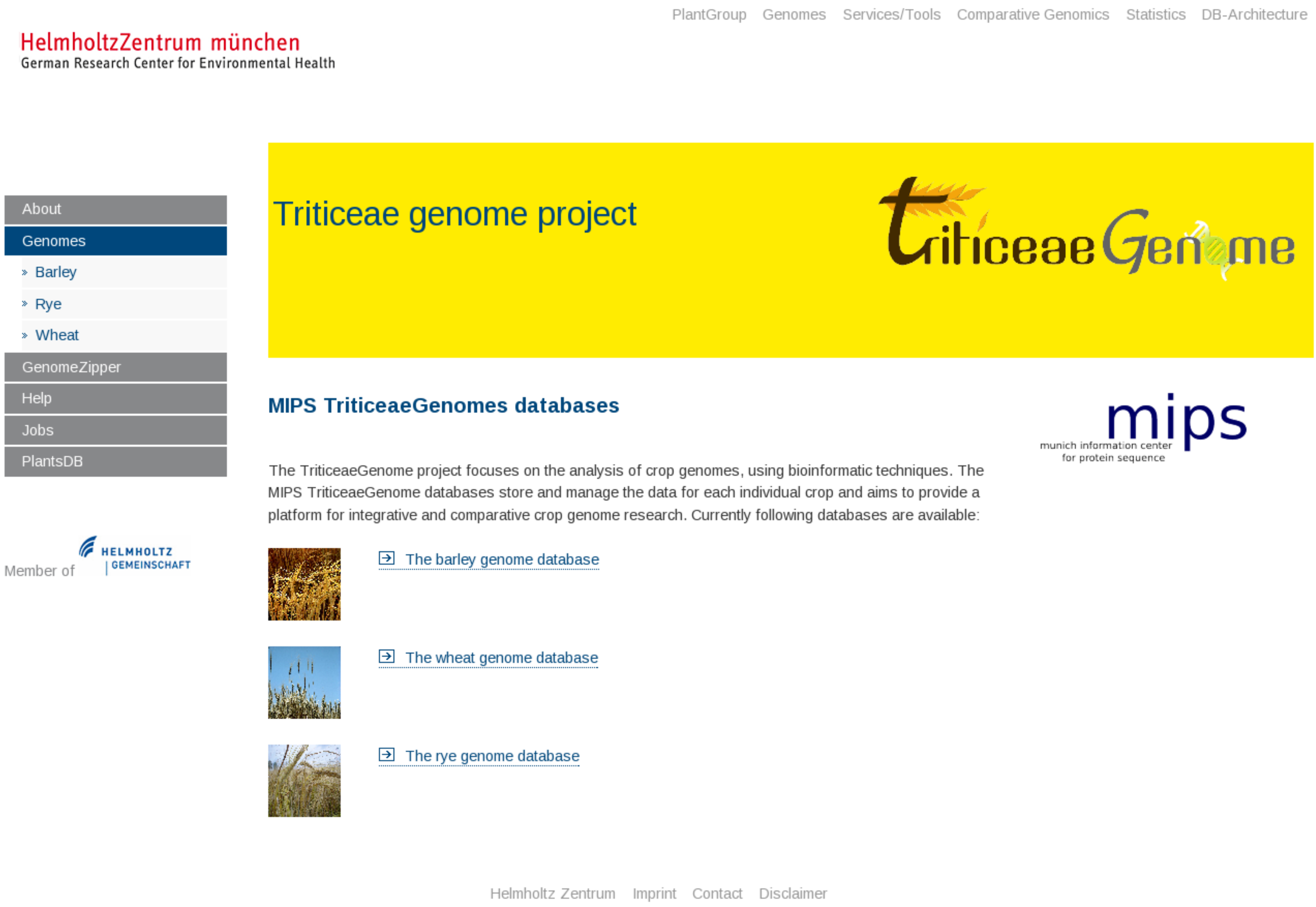
Data content in MIPS PlantsDB triticeae instances - barley

- Morex WGS sequence assembly (+ add. varieties)
- Gene annotation
- Genome zipper results
- Marker maps, genetic and physical maps, gene anchoring
- Synteny to Brachy, Sorghum and Rice – CrowsNest
- Orthologous gene families
- BLAST server @ IPK
- Batch download via FTP
- …

Data content in MIPS PlantsDB triticeae instances - wheat

- UK 5x 454 Triticum aestivum WGS sequence assembly
- wheat gene sub-assemblies for ~ 20,000 grass orthologous gene representatives
- Sub-genome classification (A,B,D) for sub-assemblies
- BLAST server to search OGRs
- Genome zipper results for flow-sorted IWGSC wheat sequences (coming soon)
- Batch download via FTP
- …

http://mips.helmholtz-muenchen.de/plant/triticeae/



MIPS TriticeaeGenomes databases

The TriticeaeGenome project focuses on the analysis of crop genomes, using bioinformatic techniques. The MIPS TriticeaeGenome databases store and manage the data for each individual crop and aims to provide a platform for integrative and comparative crop genome research. Currently following databases are available:

- The barley genome database
- The wheat genome database
- The rye genome database

Use case 1

I have a gene model in Brachypodium (**Bradi2g39900.1**), what is the corresponding region, elements and sequence in barley?

-> GenomeZipper

Use case 2

Where can I search the barley genes for all „**Alpha-glucosidase**“ proteins?

-> barley annotation page

Building a comprehensive barley gene set

RNA-seq gene loci

85,549 gene loci
142,763 transcripts

Hv fl-cDNAs

28,592 FL-cDNAs

CD-Hit Clustering (98% id)

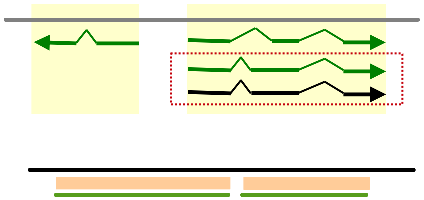
Non. Red. Hv fl-cDNAs

23,349

ORF Predictor

CDS + Protein sequences

CDS + Protein sequences



1) Genomic (structure based) clustering
clustering of mapped FL-cDNAs with predicted gene loci

2) Sequence homology based clustering
sequence search using BLAST comparisons

Barley working gene set

79,379 gene clusters

Gene annnotation
(human readable descriptions)

Expression analysis
(significantly differentially expressed genes)

Alternative splicing
(alternative splicing between tissues)

Gene family analysis
(Expanded / contracted gene families)

Use case 3

I have a rice gene (**Os04g0206700**) and wondering if/what barley gene/region is syntenic to?

->MIPS CrowsNest

MIPS CrowsNest: Features

- Display genetic and pysical maps

- Display pre-calculated syntenic regions

- Browse from macro- to micro-synteny levels

- Integrate add. information such as gene distribution for selected gene families, Ka/Ks distributions etc.

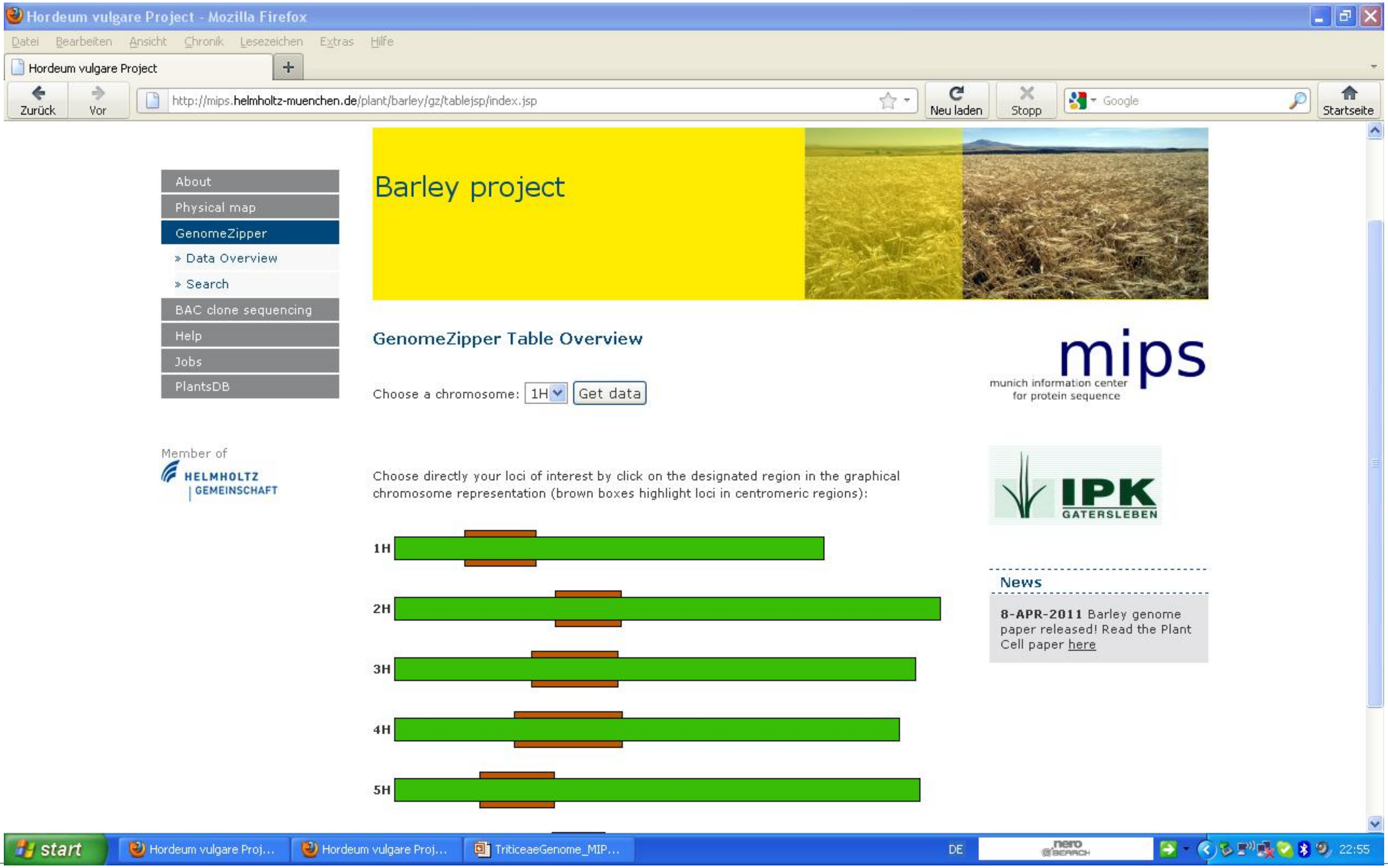
- Display density plots and heatmaps in association with other features

Use case 4

I have a brachypodium gene (**Bradi1g04150.1**) and searching for corresponding wheat genic sub-assembly sequences

-> UK wheat 454 survey section

GenomeZipper online (1)



GenomeZipper Table Overview

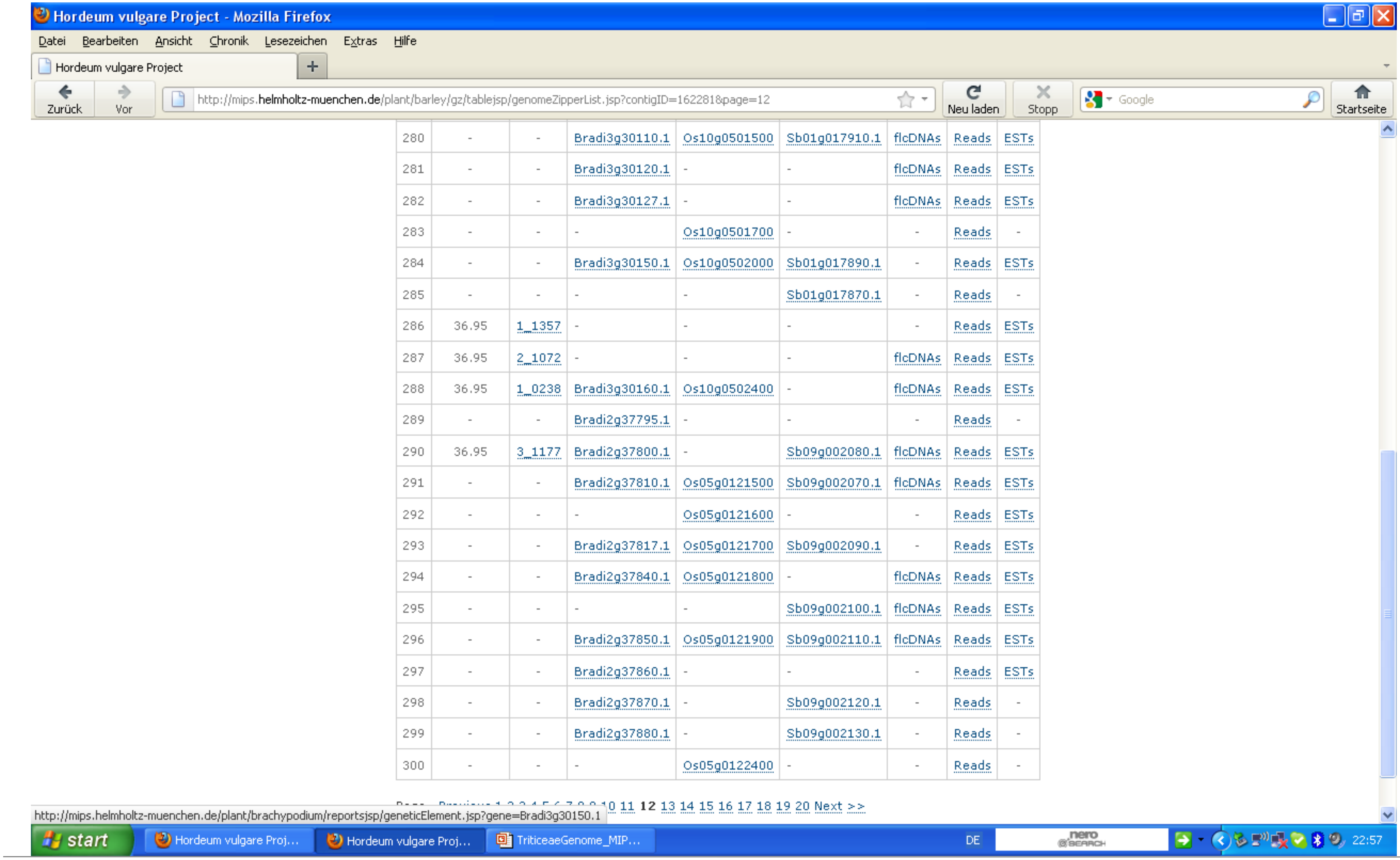
Choose a chromosome: 1H

Choose directly your loci of interest by click on the designated region in the graphical chromosome representation (brown boxes highlight loci in centromeric regions):

News

8-APR-2011 Barley genome paper released! Read the Plant Cell paper here

GenomeZipper online (2)



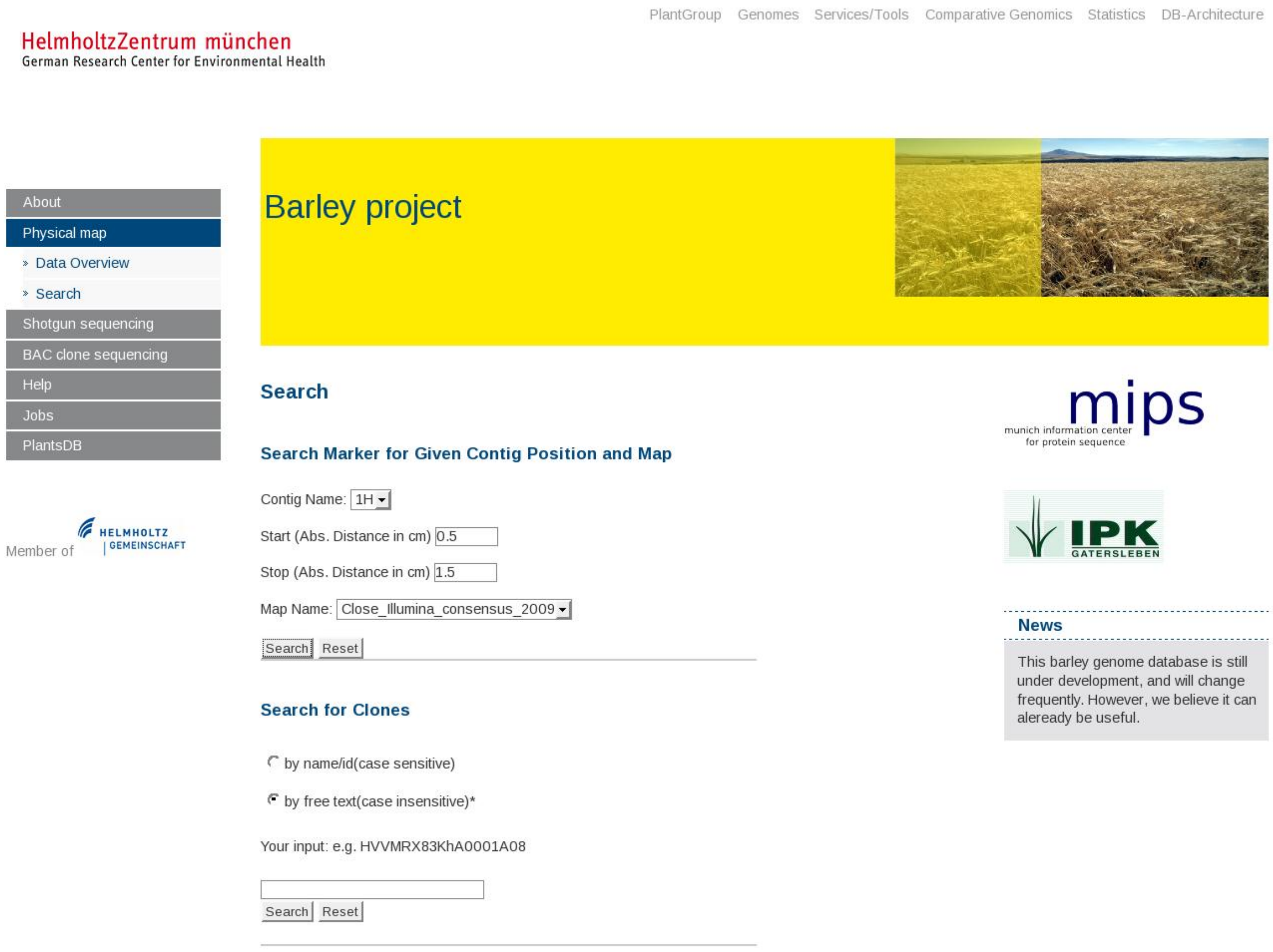
http://mips.helmholtz-muenchen.de/plant/barley/gz/tablejsp/genomeZipperList.jsp?contigID=162281&page=12

280	-	-	Bradi3g30110.1	Os10g0501500	Sb01g017910.1	flcDNAs	Reads	ESTs
281	-	-	Bradi3g30120.1	-	-	flcDNAs	Reads	ESTs
282	-	-	Bradi3g30127.1	-	-	flcDNAs	Reads	ESTs
283	-	-	-	Os10g0501700	-	-	Reads	-
284	-	-	Bradi3g30150.1	Os10g0502000	Sb01g017890.1	-	Reads	ESTs
285	-	-	-	-	Sb01g017870.1	-	Reads	-
286	36.95	1_1357	-	-	-	-	Reads	ESTs
287	36.95	2_1072	-	-	-	flcDNAs	Reads	ESTs
288	36.95	1_0238	Bradi3g30160.1	Os10g0502400	-	flcDNAs	Reads	ESTs
289	-	-	Bradi2g37795.1	-	-	-	Reads	-
290	36.95	3_1177	Bradi2g37800.1	-	Sb09g002080.1	flcDNAs	Reads	ESTs
291	-	-	Bradi2g37810.1	Os05g0121500	Sb09g002070.1	flcDNAs	Reads	ESTs
292	-	-	-	Os05g0121600	-	-	Reads	ESTs
293	-	-	Bradi2g37817.1	Os05g0121700	Sb09g002090.1	-	Reads	ESTs
294	-	-	Bradi2g37840.1	Os05g0121800	-	flcDNAs	Reads	ESTs
295	-	-	-	-	Sb09g002100.1	flcDNAs	Reads	ESTs
296	-	-	Bradi2g37850.1	Os05g0121900	Sb09g002110.1	flcDNAs	Reads	ESTs
297	-	-	Bradi2g37860.1	-	-	-	Reads	ESTs
298	-	-	Bradi2g37870.1	-	Sb09g002120.1	-	Reads	-
299	-	-	Bradi2g37880.1	-	Sb09g002130.1	-	Reads	-
300	-	-	-	Os05g0122400	-	-	Reads	-

Page Previous 1 2 3 4 5 6 7 8 9 10 11 **12** 13 14 15 16 17 18 19 20 Next >>

http://mips.helmholtz-muenchen.de/plant/brachypodium/reportsjsp/geneticElement.jsp?gene=Bradi3g30150.1

Marker search and data availability



Barley project

Search

Search Marker for Given Contig Position and Map

Contig Name: 1H

Start (Abs. Distance in cm) 0.5

Stop (Abs. Distance in cm) 1.5

Map Name: Close_Illumina_consensus_2009

Search for Clones

○ by name/id(case sensitive)

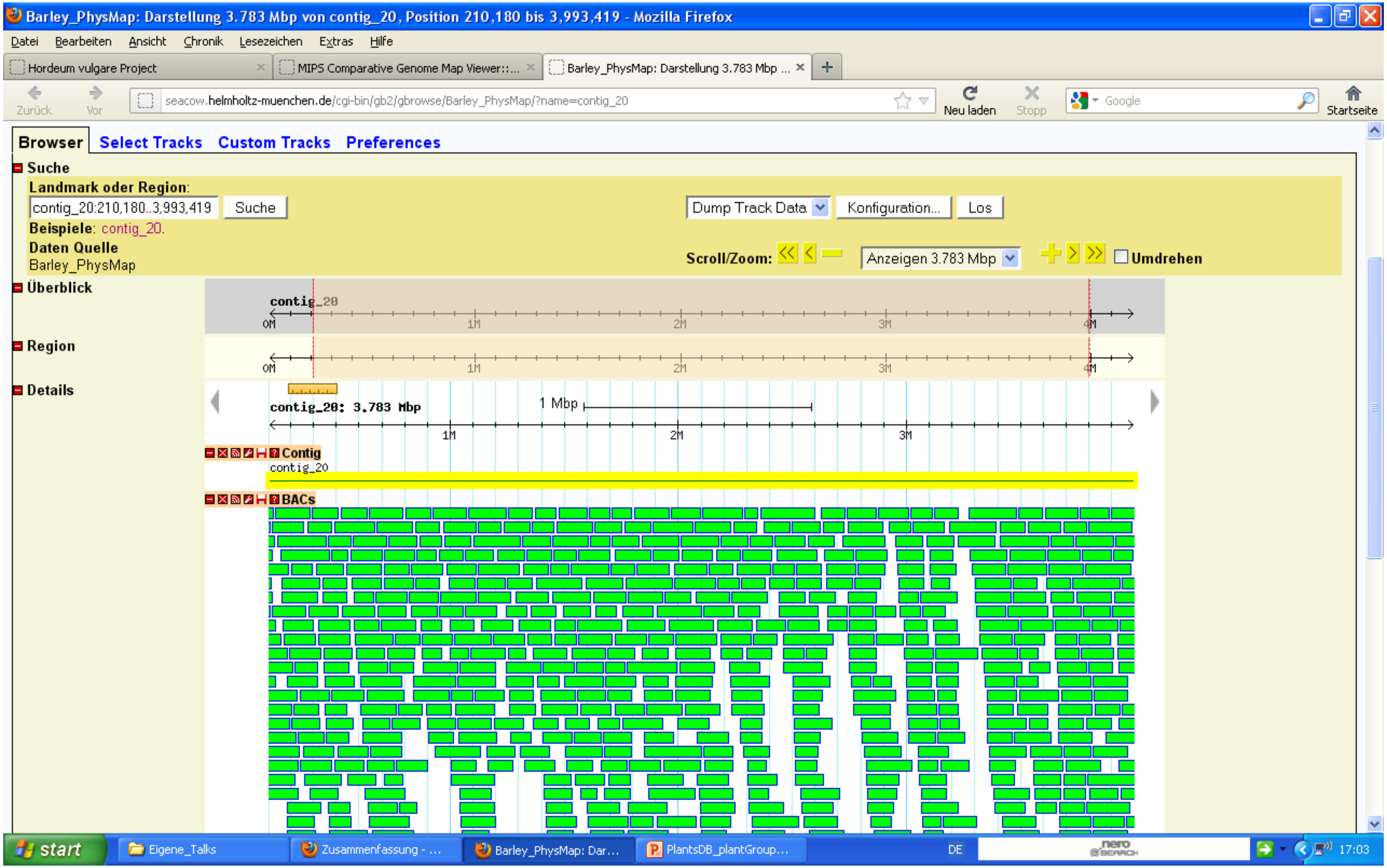
● by free text(case insensitive)*

Your input: e.g. HVVMRX83KhA0001A08

News

This barley genome database is still under development, and will change frequently. However, we believe it can aleready be useful.

Barley physical map visualization



Barley_PhysMap: Darstellung 3.783 Mbp von contig_20, Position 210,180 bis 3,993,419 - Mozilla Firefox

seacow.helmholtz-muenchen.de/cgi-bin/gb2/gbrowse/Barley_PhysMap/?name=contig_20

Suche

Landmark oder Region: contig_20:210,180..3,993,419

Beispiele: contig_20.

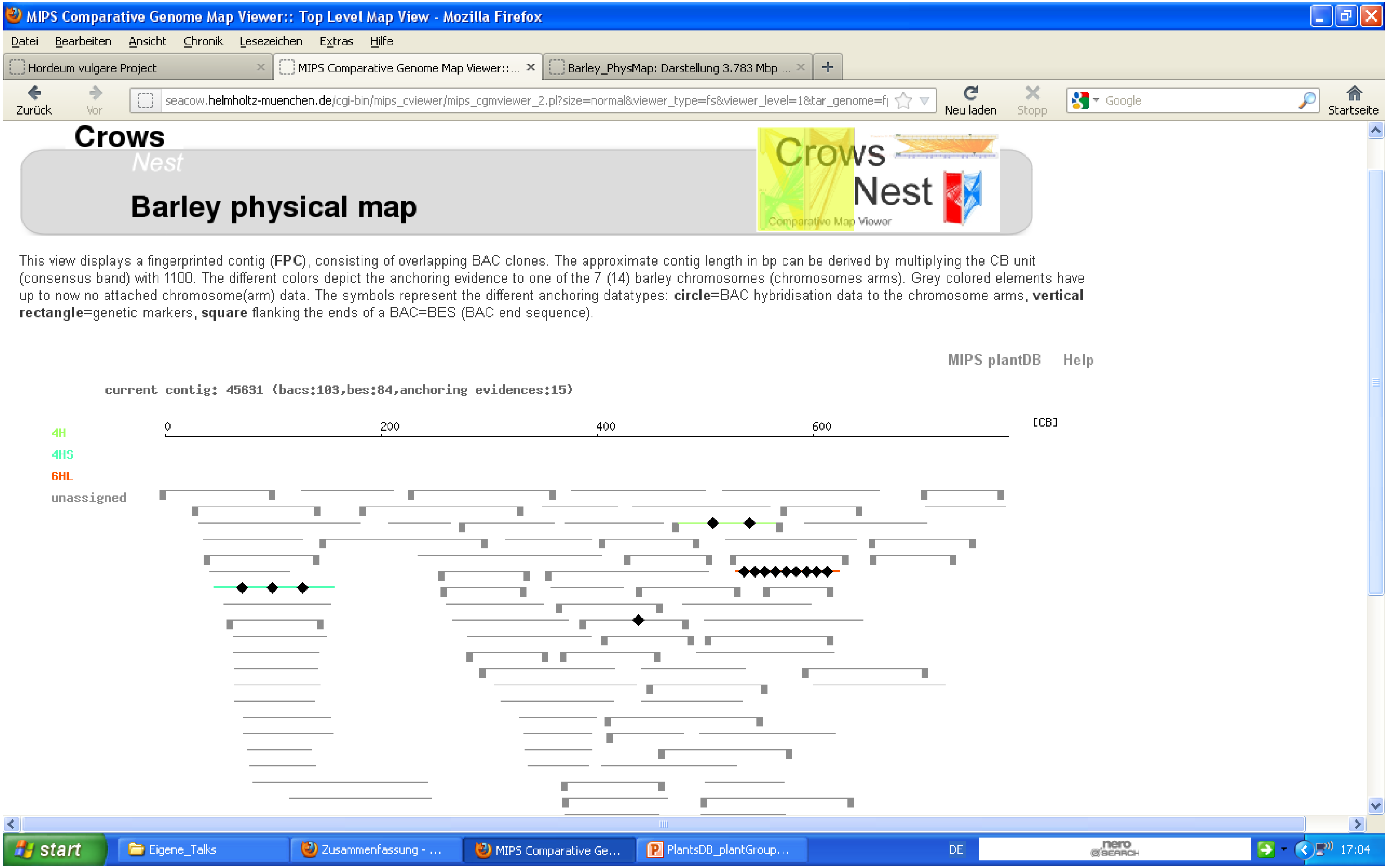
Daten Quelle

Barley_PhysMap

Scroll/Zoom: Anzeigen 3.783 Mbp ☐ **Umdrehen**

contig_20: 3.783 Mbp

Barley physical map visualization

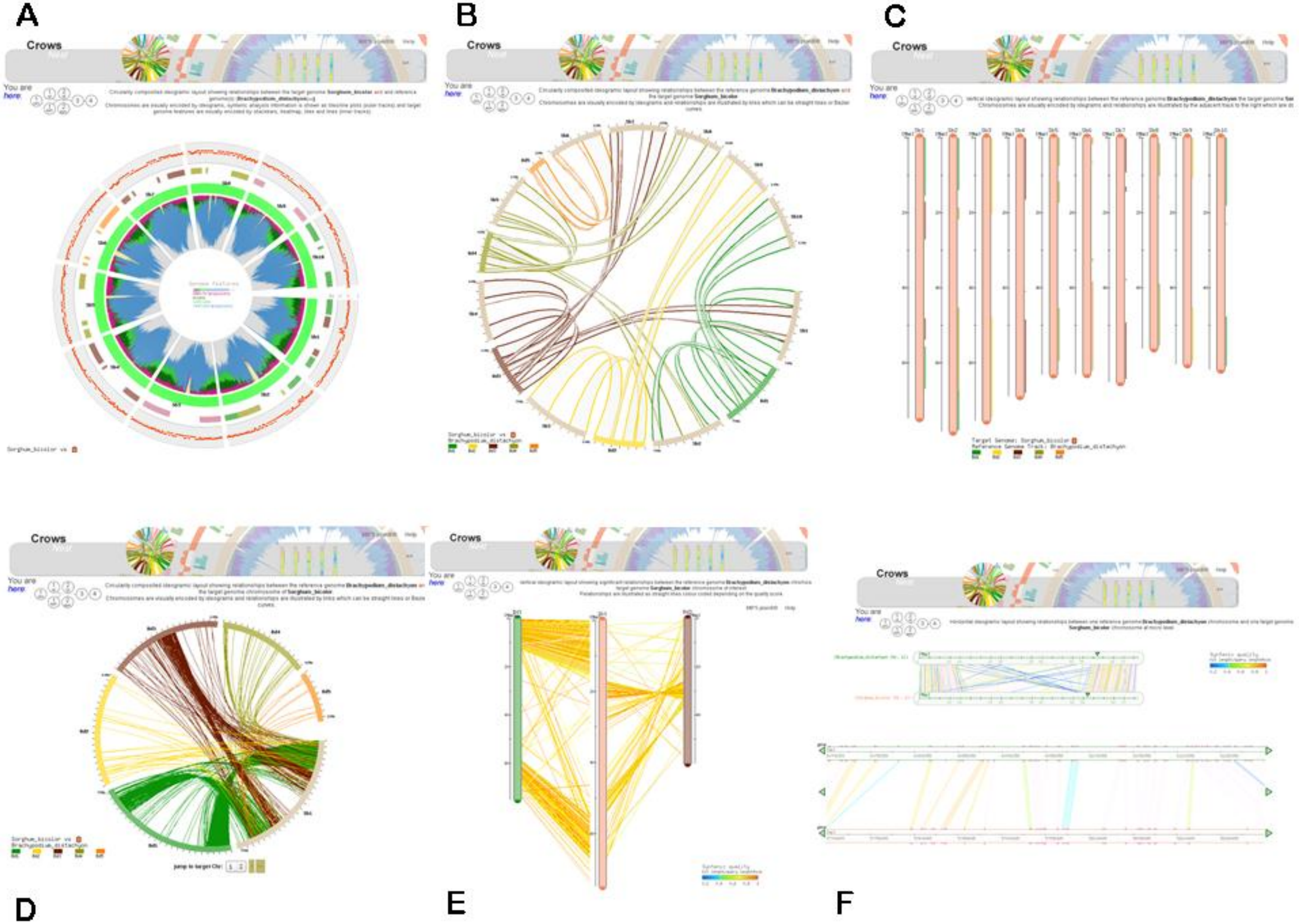


Barley physical map

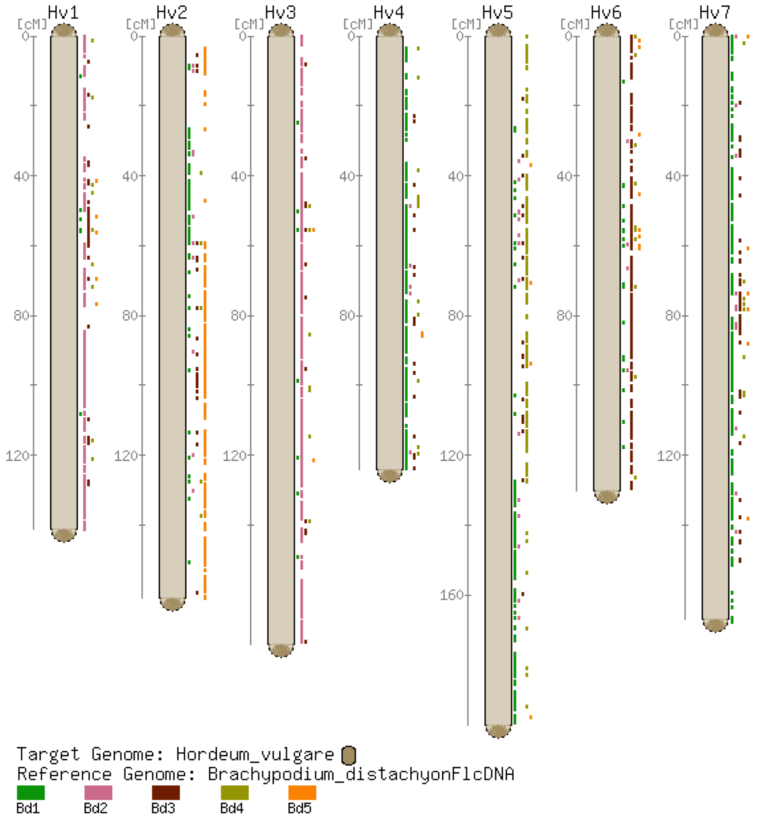
This view displays a fingerprinted contig (**FPC**), consisting of overlapping BAC clones. The approximate contig length in bp can be derived by multiplying the CB unit (consensus band) with 1100. The different colors depict the anchoring evidence to one of the 7 (14) barley chromosomes (chromosomes arms). Grey colored elements have up to now no attached chromosome(arm) data. The symbols represent the different anchoring datatypes: **circle**=BAC hybridisation data to the chromosome arms, **vertical rectangle**=genetic markers, **square** flanking the ends of a BAC=BES (BAC end sequence).

current contig: 45631 (bacs:103,bes:84,anchoring evidences:15)

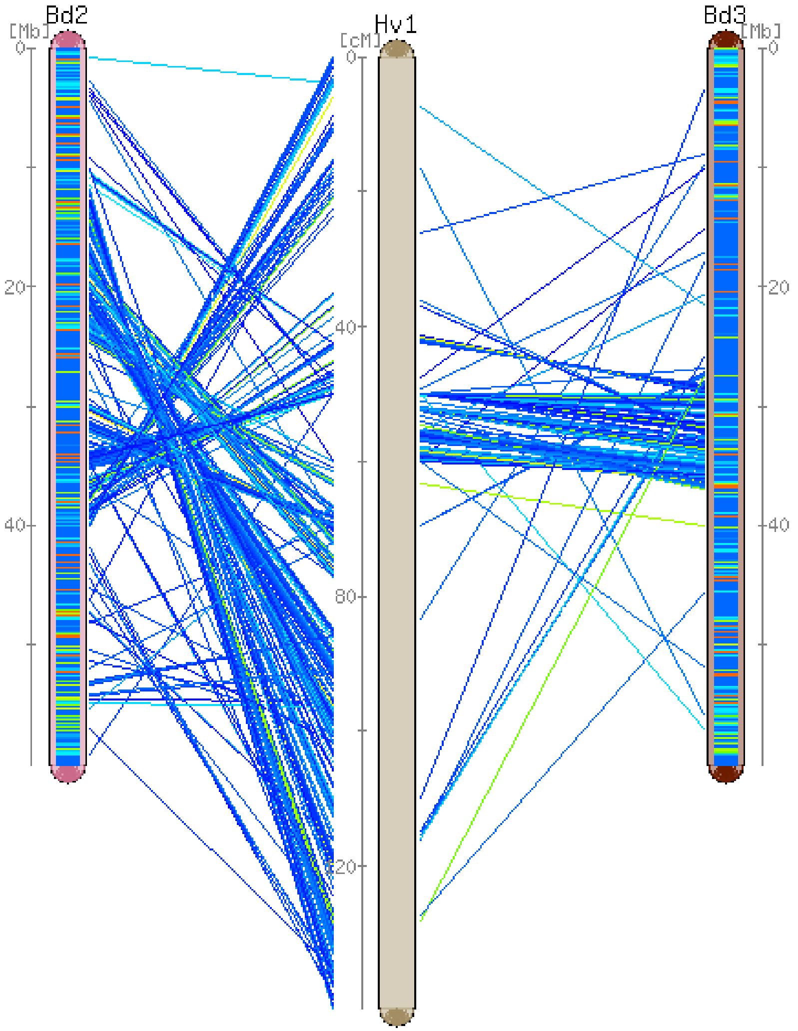
CrowsNest – a tool for comparative genomics



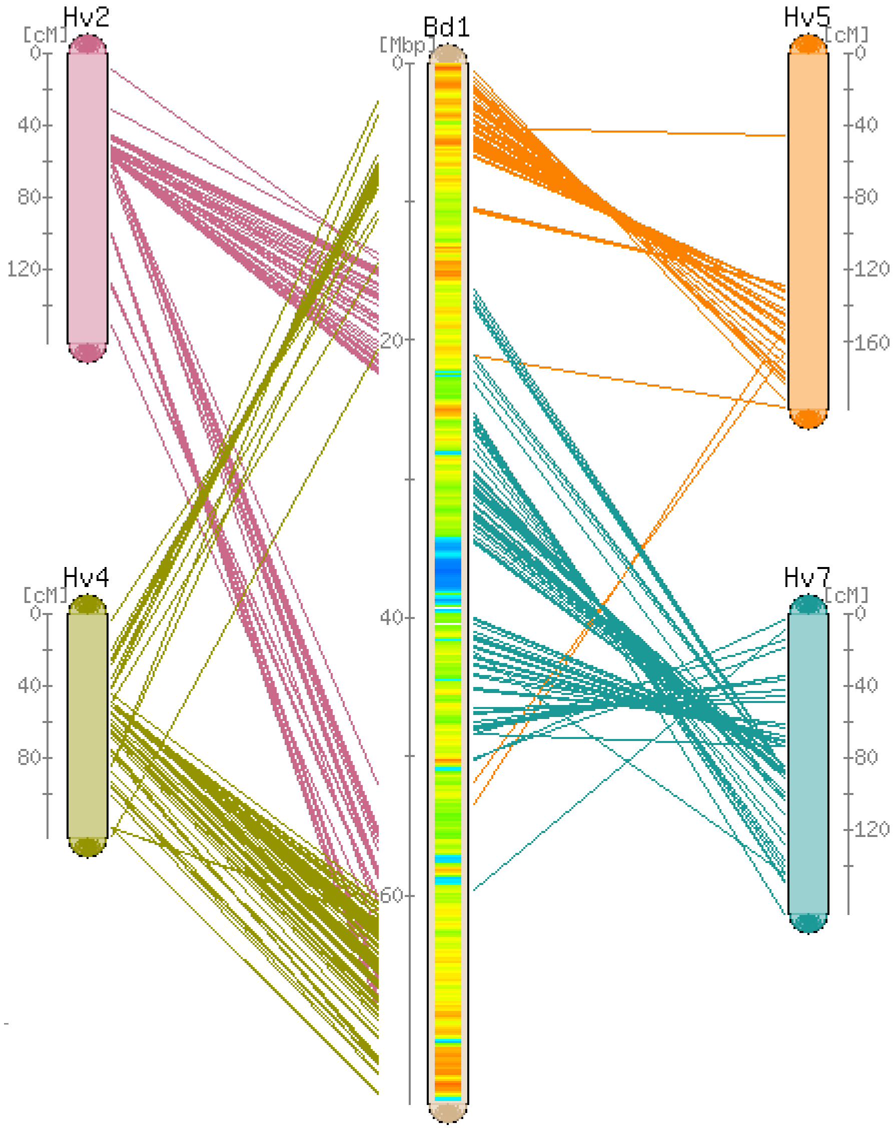
CrowsNest – a visualization tool for synteny

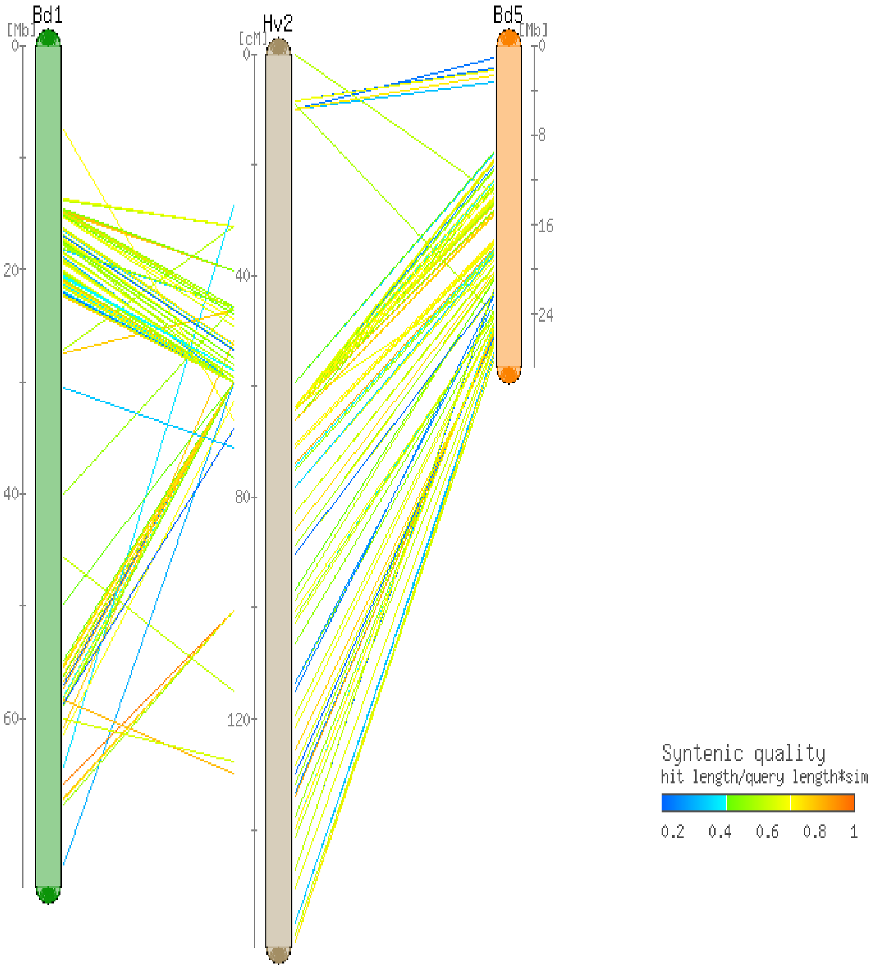


CrowsNest – a visualization tool for synteny

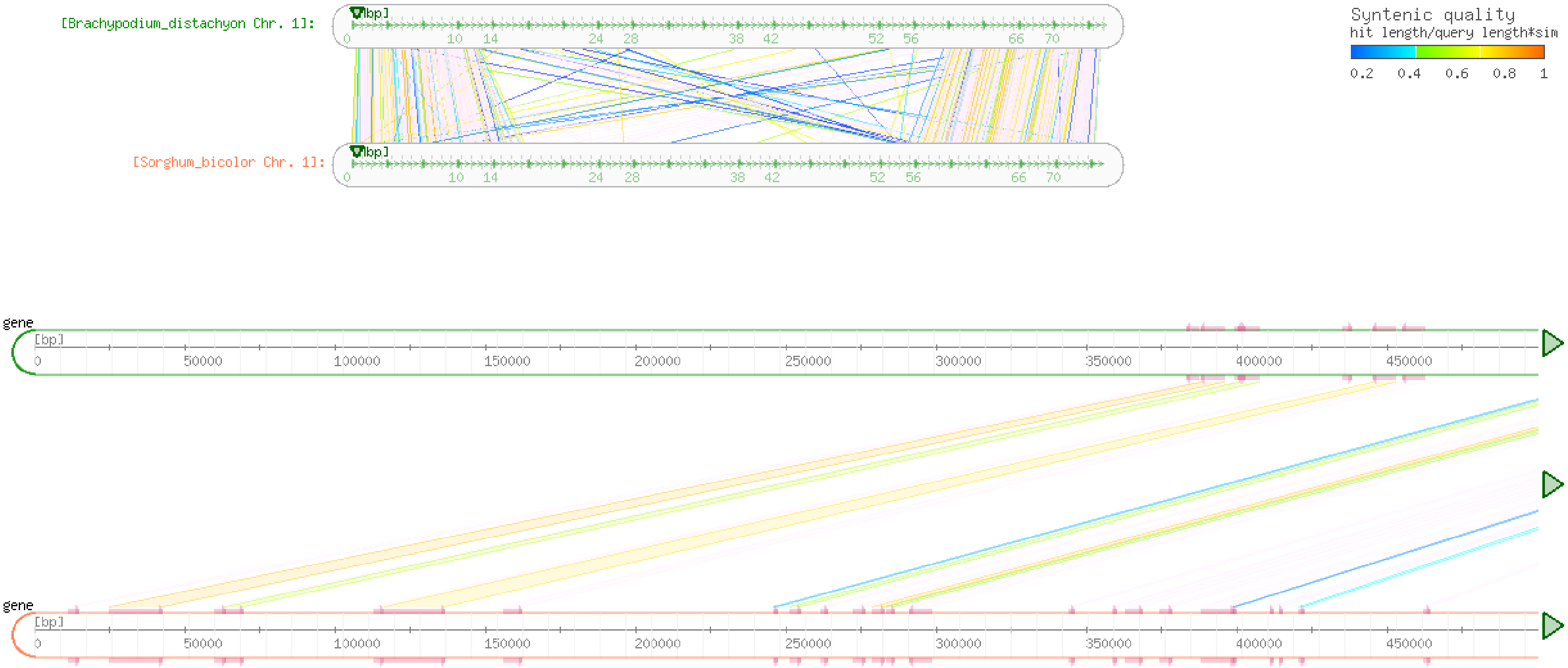


Homolgy between Hordeum to Brachypodium, Sorghum and Oryza





CrowsNest – microsynteny view



The Cooks
- this menu couldn‘t be served without:

IPK GATERSLEBEN

Nils Stein
Andreas Graner
B Steuernagel
Uwe Scholz
Ruvini A.

IEB

Hana Simkova
Pavla Suchankova
Jaroslav Dolezel
Miroslav Valerik

The James Hutton Institute

Robbie Waugh
Pete Hedley
Hui Liu
David Marshall

NIAS/Okayama
Takashi Matsumoto
Tsuyoshi Tanaka
Takeshi Itoh

Kaz Sato et al.

IWGSC

FLI Jena
Stefan Taudien
Marius Felder
Andreas Petzold
Matthias Platzer

Univ Cordoba
Pilar Hernandez

INRA

Cathrine Feulliet
URGI Versailles

TGAC
The Genome Analysis Centre™
BBSRC EEDA Greater Norwich Development Partnership

gabi

Bundesministerium
für Bildung
und Forschung

transPLANT

SEVENTH FRAMEWORK PROGRAMME

HelmholtzZentrum münchen
German Research Center for Environmental Health

The Cooks
- this menu couldn‘t be served without:

John Innes Center
Mike Bevan

University Liverpool
Neil Hall
Rachel Brenchley
Anthony Hall

University Bristol
Keith Edwards
Gary Barker

EBI
Paul Kersey
Dan Bolser

The MIPS team…