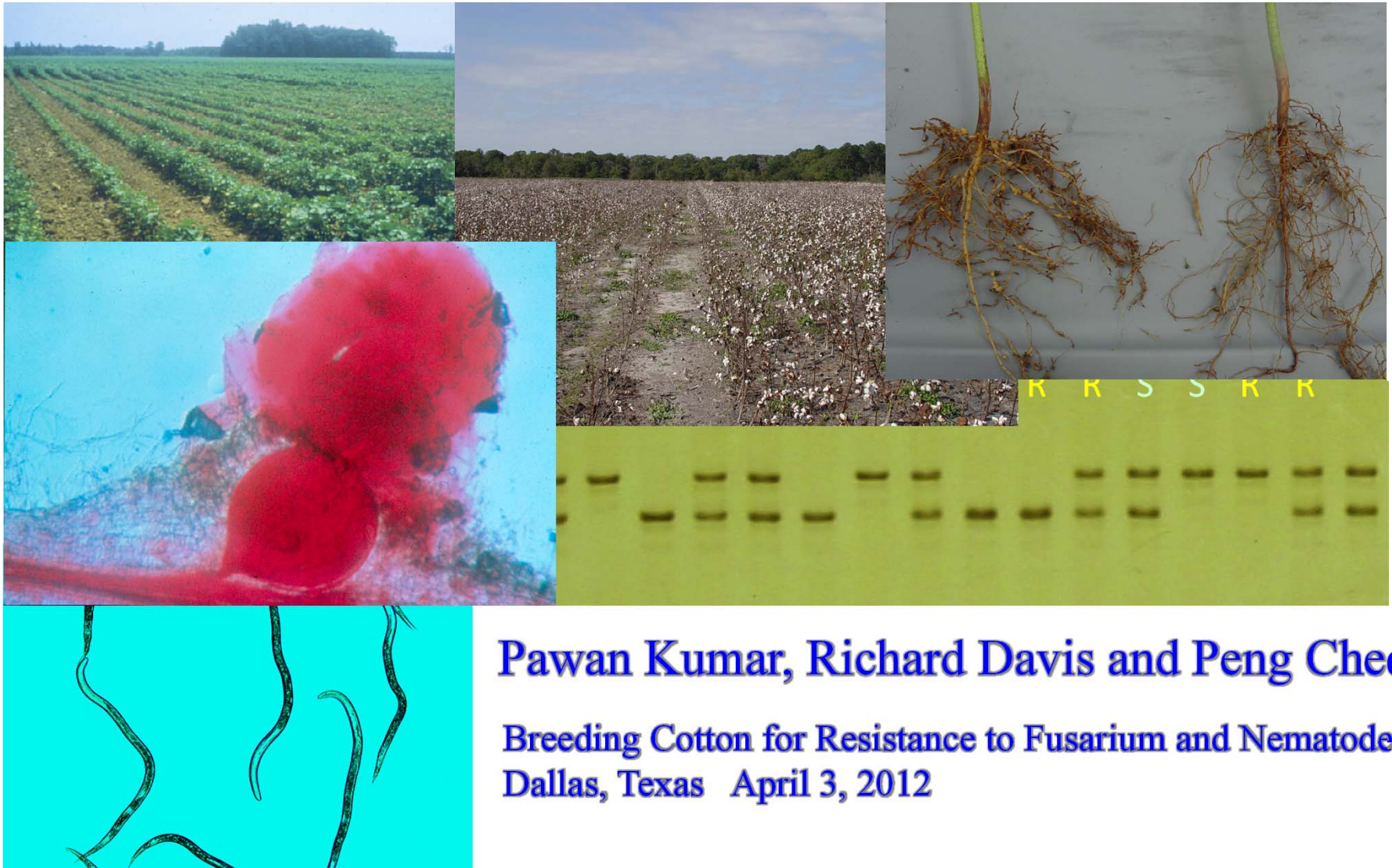
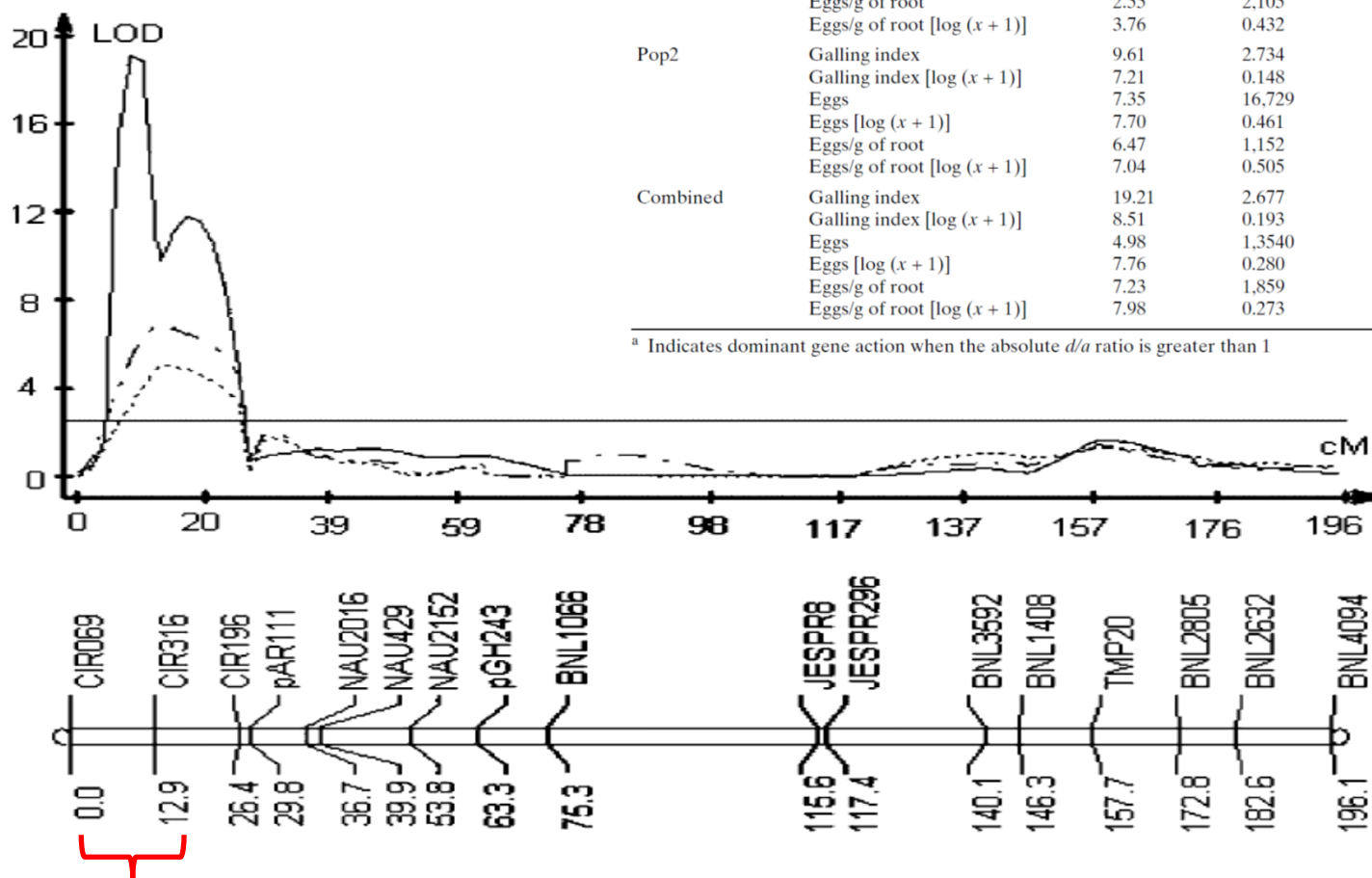


Utilizing the *G. raimondii* Genome Sequence for Mapping RKN Resistance



Pawan Kumar, Richard Davis and Peng Chee

Breeding Cotton for Resistance to Fusarium and Nematodes
Dallas, Texas April 3, 2012



RKN resistance QTL

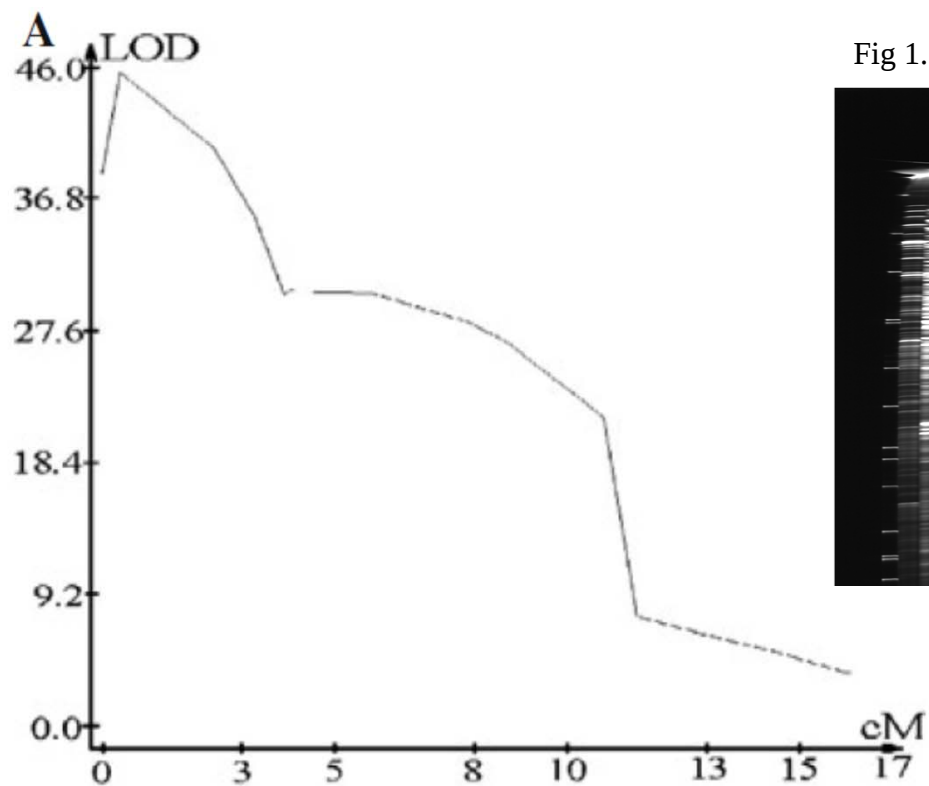
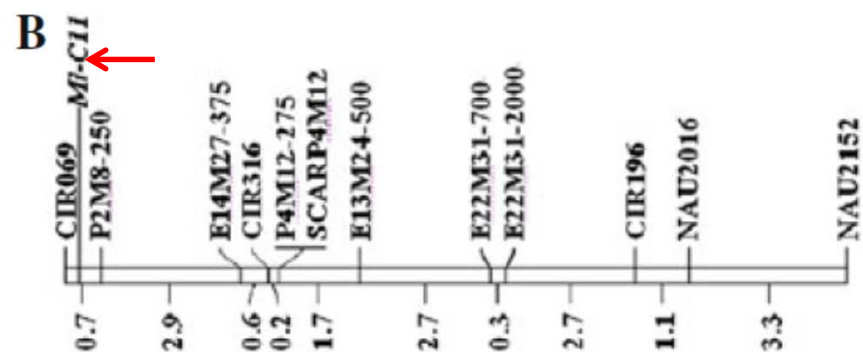
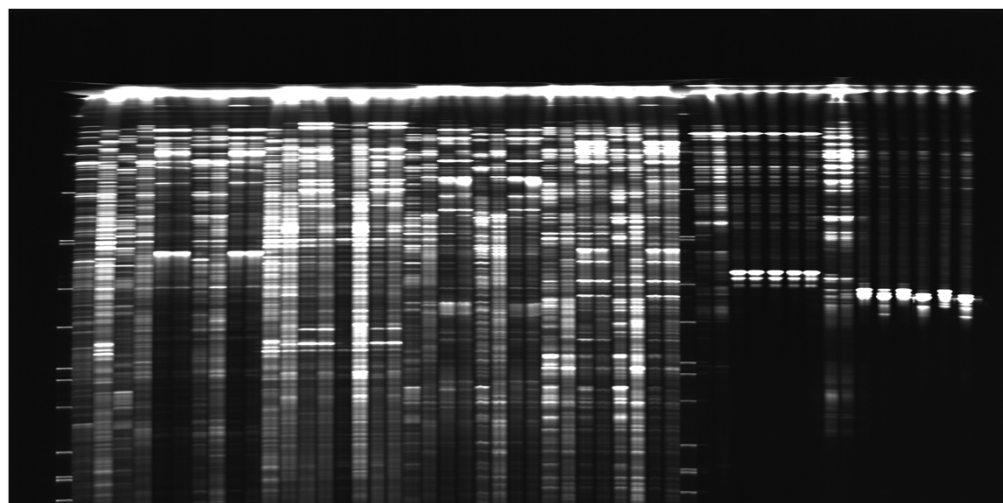
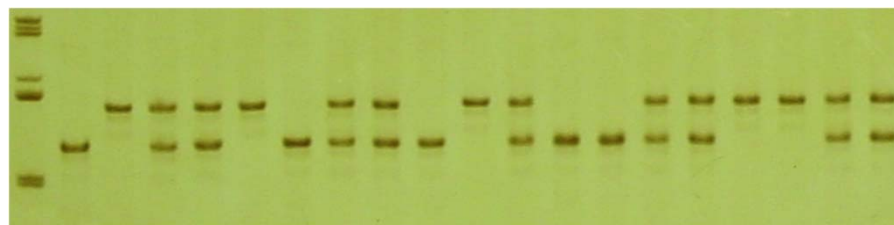


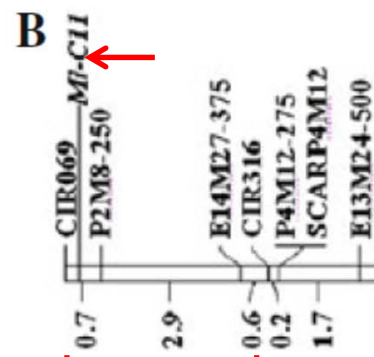
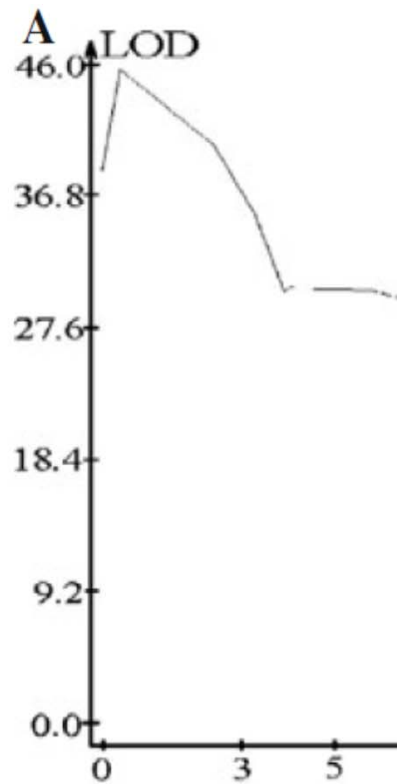
Fig 1. Polymorphisms analysis using AFLP markers



RKN resistance QTL

Fig 3. Polymorphisms analysis of SCAR marker P4M12 in mapping parents and individuals.

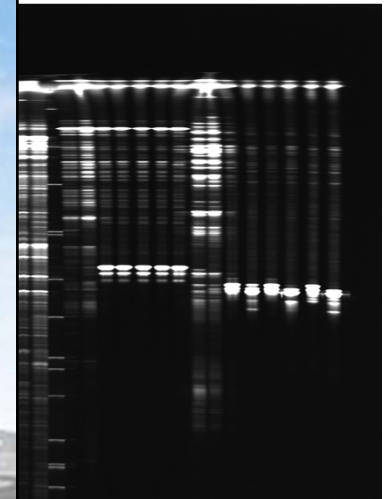




RKN resistance QTL



kers



marker P4M12 in mapping



Shen et al 2010 TAG



PRESSROOM

► Press Releases

/ 2012

2011

2010

2009

2008

► Video Archives

► Did You Know?

► Media Contacts

► CottonLifestyleMonitor.com

'Gold-Standard' Cotton Genome Sequence Will Advance Fiber, Fuel, and Food Applications

Thursday January 05, 2012

New York, NY

An international consortium, led by Professor Andrew Paterson of the University of Georgia, has made publicly available the first 'gold-standard' genome sequence for cotton. Cotton was among the first plants studied at the molecular level, and the sequence obtained by Paterson and his team is the culmination of a 20-plus year effort in the analysis of cotton genes and genomic DNA. This critical sequence will be invaluable to better understanding and optimizing the production and sustainability of the cotton plant.

The research effort of Paterson and others gained momentum in 2007 when a proposal from 22 leading cotton scientists representing the world's seven largest cotton-producing nations was approved by the United States Department of Energy (DOE) Joint Genome Institute (JGI) Community Sequencing Program. The study established the strategy that was used for 'gold-standard' sequencing of the New World cotton progenitor, *Gossypium raimondii*, which was chosen by the worldwide cotton community to be the first of 50 cotton species to be sequenced. "This achievement, and the ongoing research community annotations of our cotton genome, will speed continued improvement of cotton production and help sustain one of the world's largest industries," said Professor Paterson.

The cotton sequence is among the highest-quality flowering plant sequences yet produced. A novel strategy integrating "next-generation" and conventional sequencing methods was used. Critical to the effort was information about the cotton hereditary blueprint, which had been accumulated over more than 20 years of research funded by the U.S. National Science Foundation, the U.S. Department of Agriculture, Cotton Incorporated, and other public and private agencies.

Plant Genomics Program

Capturing Light to Fuel Our Future



Announcements

- Jan 16, 2012

Phytozome v8.0 has been released!

- Oct 07, 2011

Release of *Thellungiella halophila* genome and Biomart sequence issue fixed.

Releases

- December 22, 2011

Asterochloris sp. Cqr/DA1pho v1.0

- August 22, 2011

Micromonas pusilla CCMP1545 v3.0

- January 8, 2010

Cucumis sativus v1.0

- January 8, 2010

Mimulus guttatus v1.0

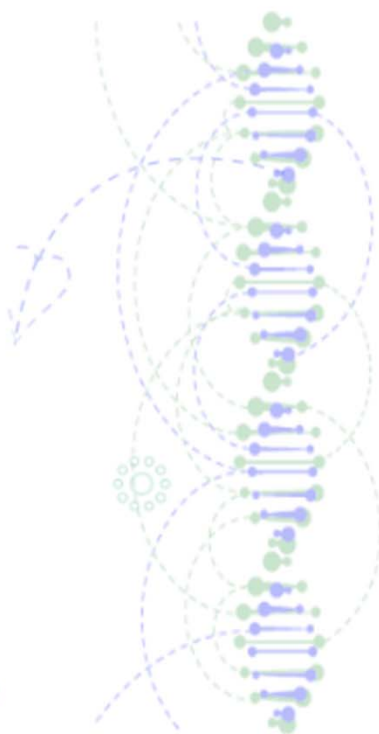
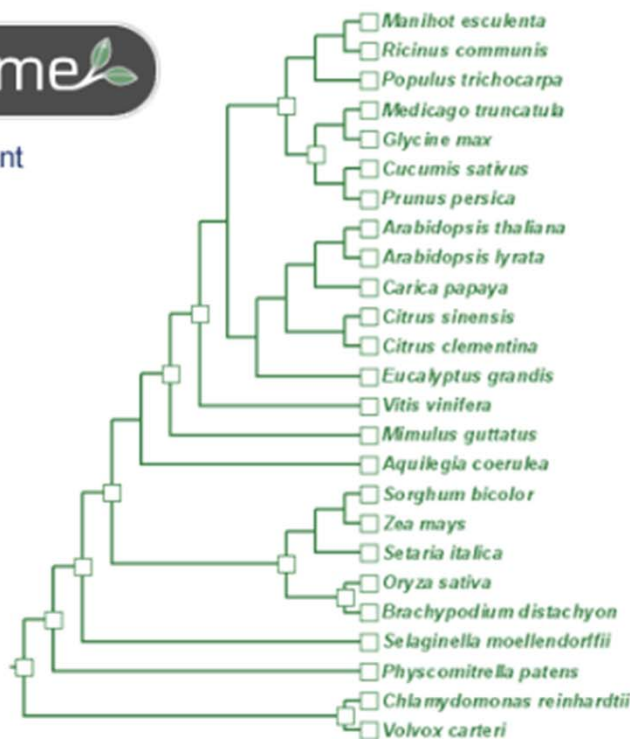
- December 18, 2009

Coccomyxa sp. C-169 v2.0

[more >>](#)

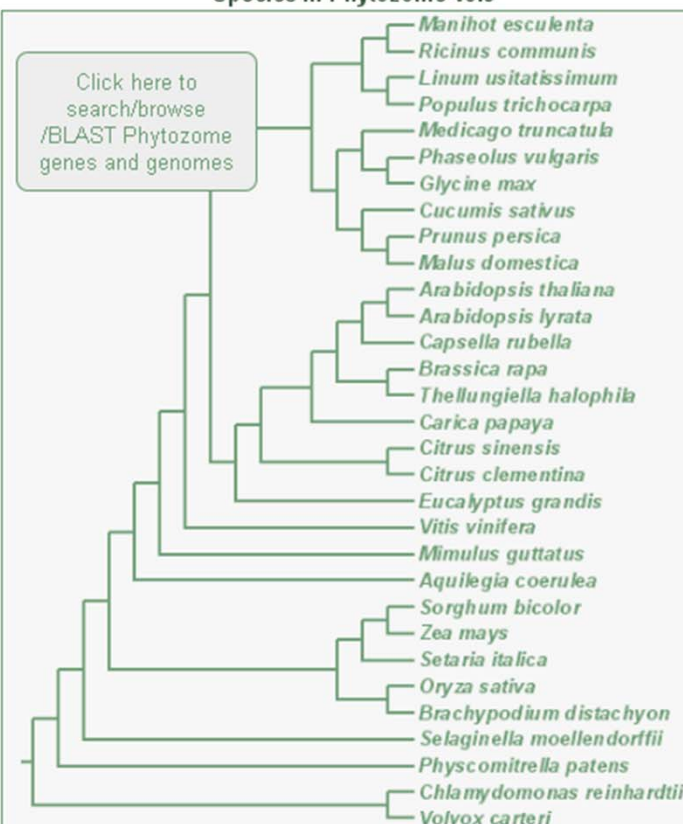
phytozome

a tool for green plant comparative genomics



Species in Phytozome v8.0

Click here to
search/browse
/BLAST Phytozome
genes and genomes



Welcome to Phytozome!

Phytozome is a joint project of the Department of Energy's Joint Genome Institute and the Center for Integrative Genomics to facilitate comparative genomic studies amongst green plants. Families of orthologous and paralogous genes that represent the modern descendants of ancestral gene sets are constructed at key phylogenetic nodes. These families allow easy access to clade specific orthology/paralogy relationships as well as clade specific genes and gene expansions. As of release v8.0, Phytozome provides access to thirty-one sequenced and annotated green plant genomes which have been clustered into gene families at eleven evolutionarily significant nodes. Where possible, each gene has been annotated with PFAM, KOG, KEGG, and PANTHER assignments, and publicly available annotations from RefSeq, UniProt, TAIR, JGI are hyper-linked and searchable.

Using Phytozome: video tutorials

Announcements

[14 Feb 2012] **A service outage for filesystem maintenance is currently scheduled for Wednesday, February 15th, 2012 from approximately 7 AM to 2 PM Pacific Standard Time.**

[16 Jan 2012] **Phytozome v8.0 has been released!**

[12 Dec 2011] **A service outage for filesystem maintenance is currently scheduled for Tuesday, December 13th, 2011 from approximately 12 AM to 3 PM Pacific Standard Time.**

News

[16 Jan 2012] **Phytozome v8.0 has been released!**

The latest version of Phytozome has been released. Version 8 includes 21 genomes assembled and/or annotated at the JGI, and 9

[12 Dec 2011] **Service Outage Tuesday, December 13th, 2011**

A service outage for filesystem maintenance is currently scheduled for Tuesday, December 13th, 2011 from approximately 12 AM to 3 PM

[07 Oct 2011] **Release of Thellungiella halophila genome and Biomart sequence issue fixed.**

The version 1.0 genome annotation and assembly of Thellungiella

[05 Aug 2011] **Service Outage Tuesday, August 9th, 2011**

A service outage for filesystem maintenance is currently scheduled for Tuesday, August 9th, 2011 from approximately 7 AM to 5 PM Pacific

:: Go to the news page to view all the news ::

Early release genomes

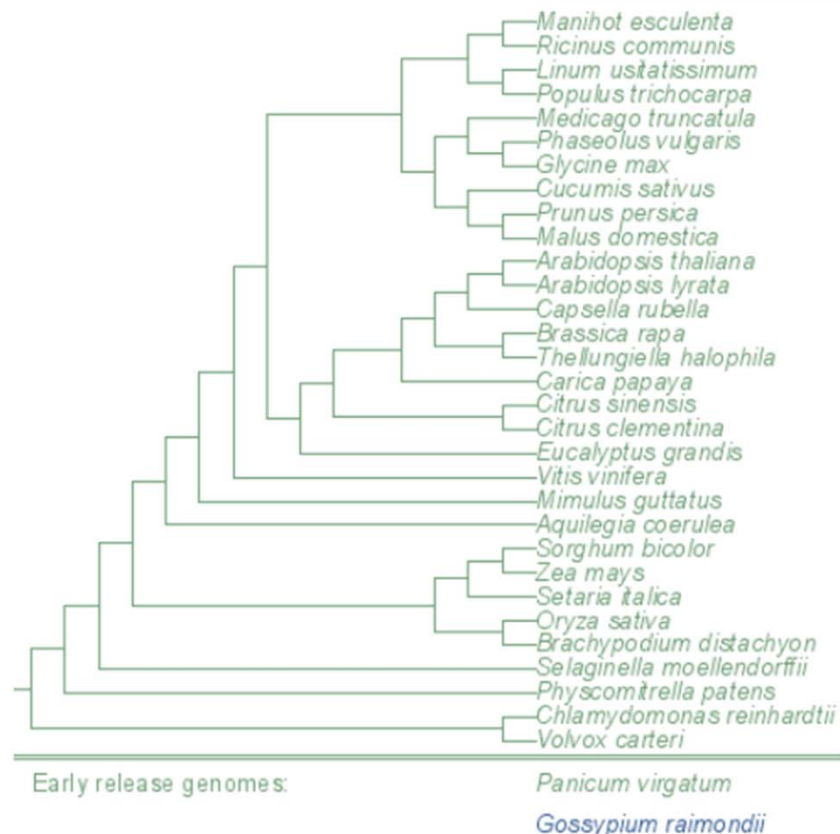
- o Panicum virgatum
- o Gossypium raimondii

Info

- o Release notes
- o Plant Genomics resources

***Gossypium raimondii* (cotton)**Browse
GenomeBLAST
GenomeDownload
data**About the genome:**

- **Statistics**
- **FAQ**
 - What can I do with the *Gossypium raimondii* dataset?

**Statistics**

This is release of assembly version v1.0 of the Cotton D genome.

Genome

Scaffold total: 1,448

Contig total: 28,582

Scaffold sequence total: 763.8 Mb

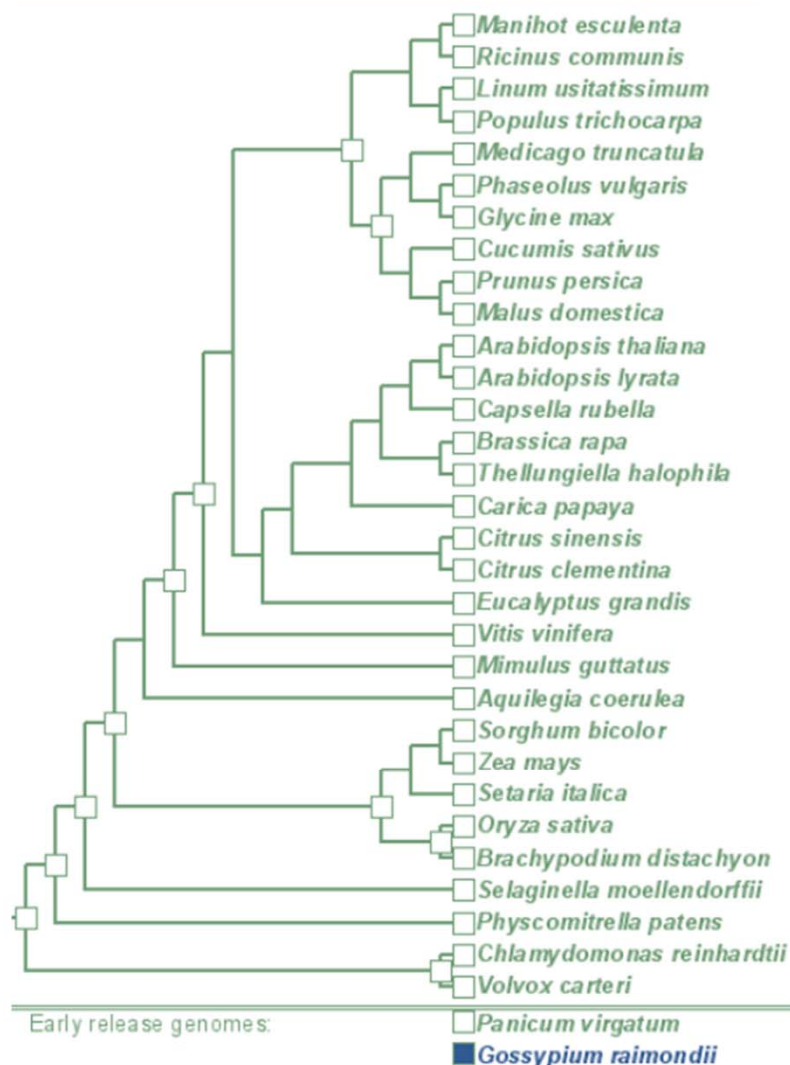
Contig sequence total: 749.4 Mb (-> 1.9% gap)

Scaffold N50 (L50) = 6 (62.2 Mb)

Contig N50 (L50) = 1,596 (135.1kb)

35 scaffolds are > 50kb in size, representing approximately 98.3% of the genome

Gossypium raimondii selected



Select all species

2. Choose a tool:

Keyword search

expand

BLAST search

submit

Target *Gossypium raimondii*

BLAST program BLASTN - nucleotide query to nucleotide db

Query name (optional)

Query sequence ☒ enter manually ☐ upload file

Enter your Query Sequence

- ☒ Show results in browser (default)
☐ Notify by email when job completes (long jobs)

► Algorithm parameters

BLAT search

expand

Genome browser

Launch

Info page

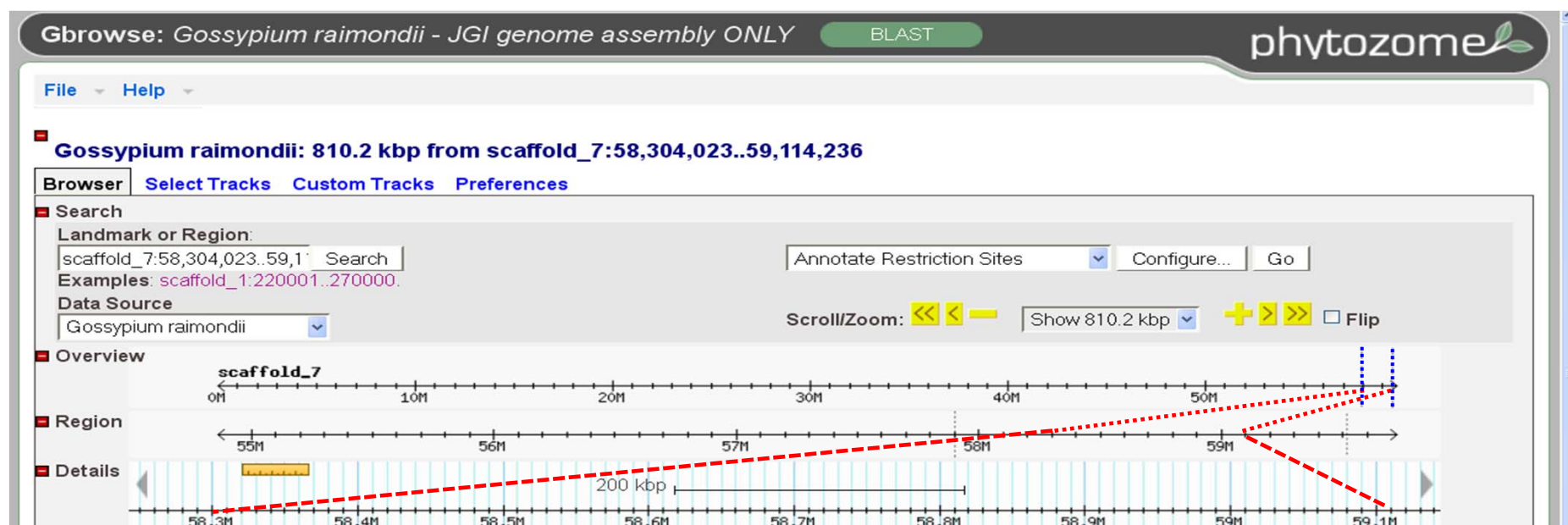
Get info

Bulk data

Get data

BLAST results- sequences of CIR316 and CIR069

Chromosome 11 = Scaffold-7 of *Gossypium raimondii* (59.7Mb)



CIR316

CIR069

810 Kbp



File Help

Gossypium raimondii: 810.2 kbp from scaffold_7:58,304,023..59,114,236 Target Region only

Browser Select Tracks Custom Tracks Preferences

Search

Landmark or Region:

scaffold_7:58,304,023..59,114,236 Search

Examples: scaffold_1:220001..270000

Data Source

Gossypium raimondii

Annotate Restriction Sites

Configure...

Go

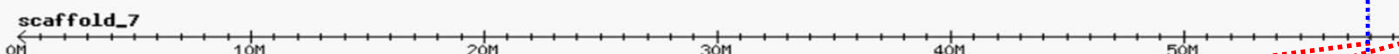
Scroll/Zoom:

Show 810.2 kbp

+ > >>

Flip

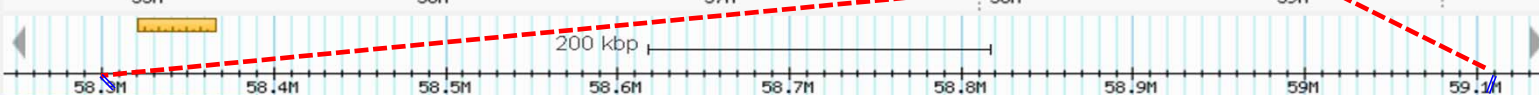
Overview



Region



Details



CIR316

CIR069

Gossypium raimondii: 2.724 Mbp from scaffold_7:57,000,000..59,724,023 Whole Telomeric Region (2.7Mb)

Browser Select Tracks Custom Tracks Preferences

Search

Landmark or Region:

scaffold_7:57,000,000..59,724,023 Search

Examples: scaffold_1:220001..270000

Data Source

Gossypium raimondii

Annotate Restriction Sites

Configure...

Go

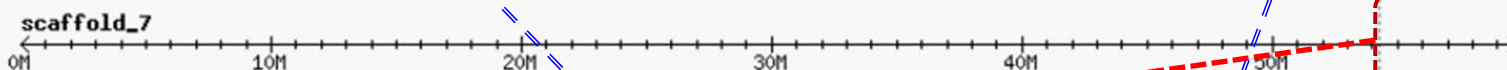
Scroll/Zoom:

Show 2.724 Mbp

+ > >>

Flip

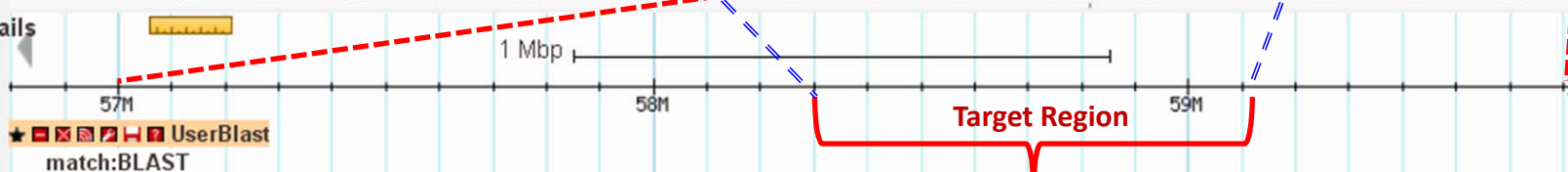
Overview



Region

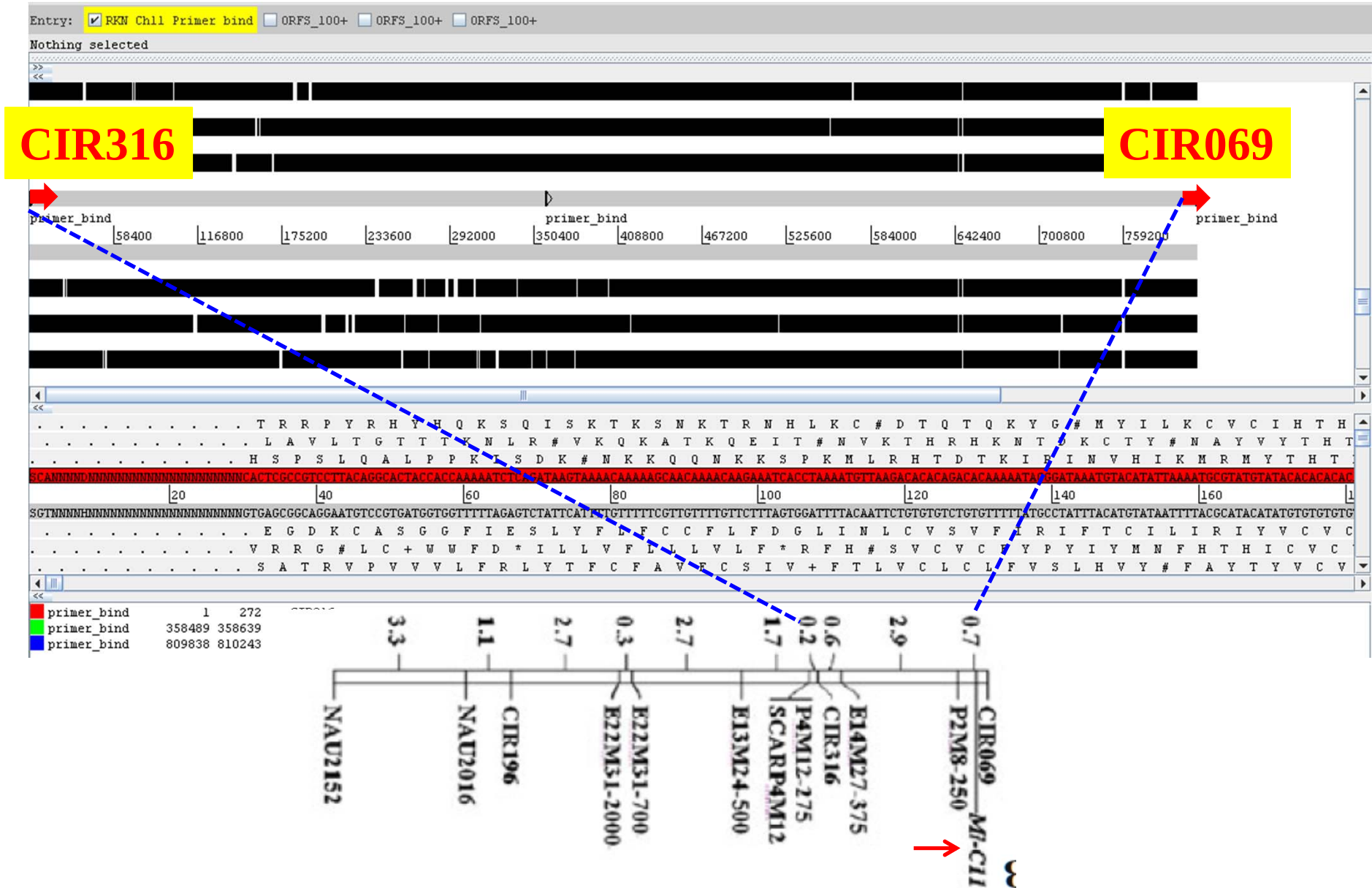


Details

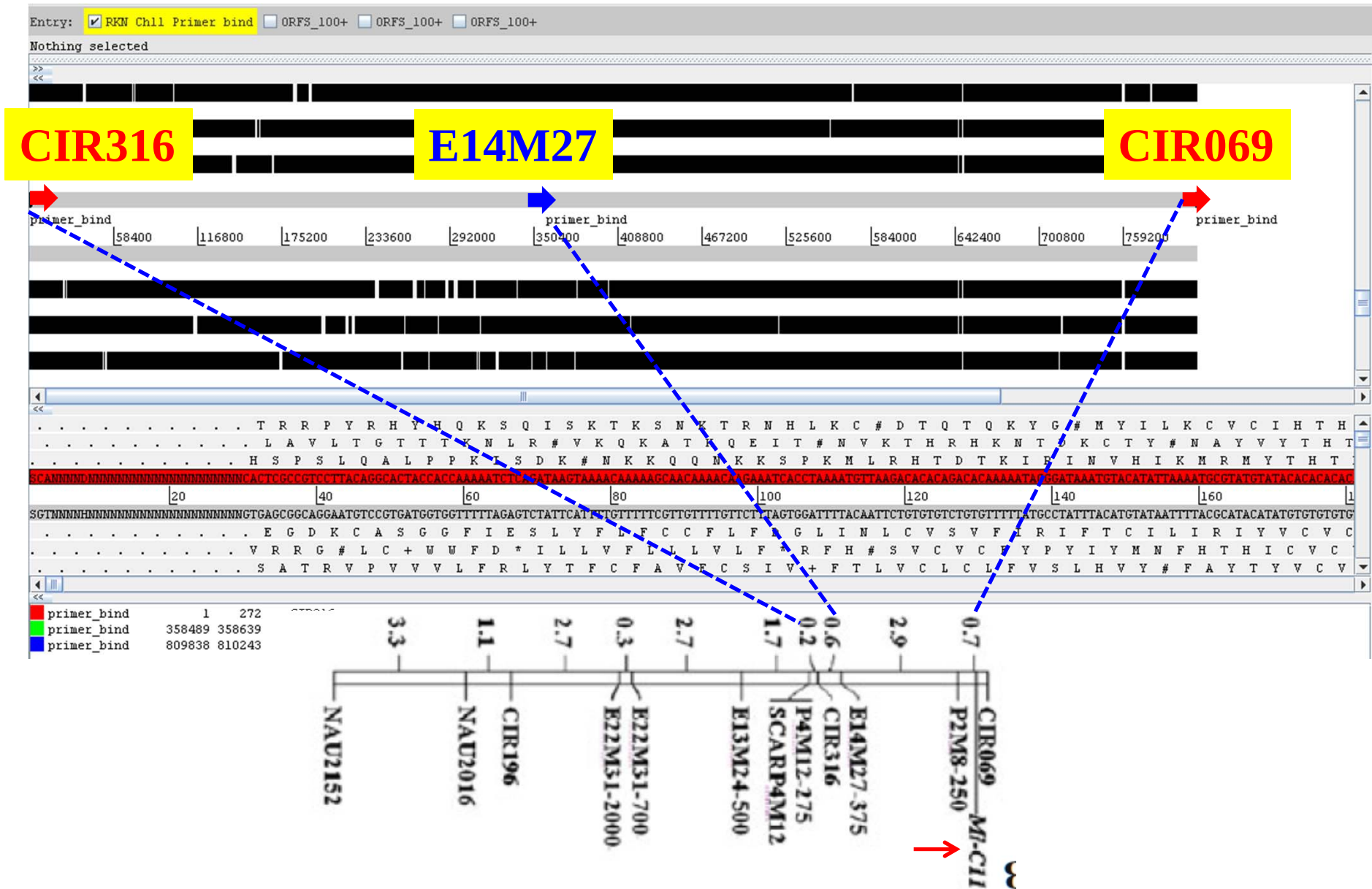


Target Region

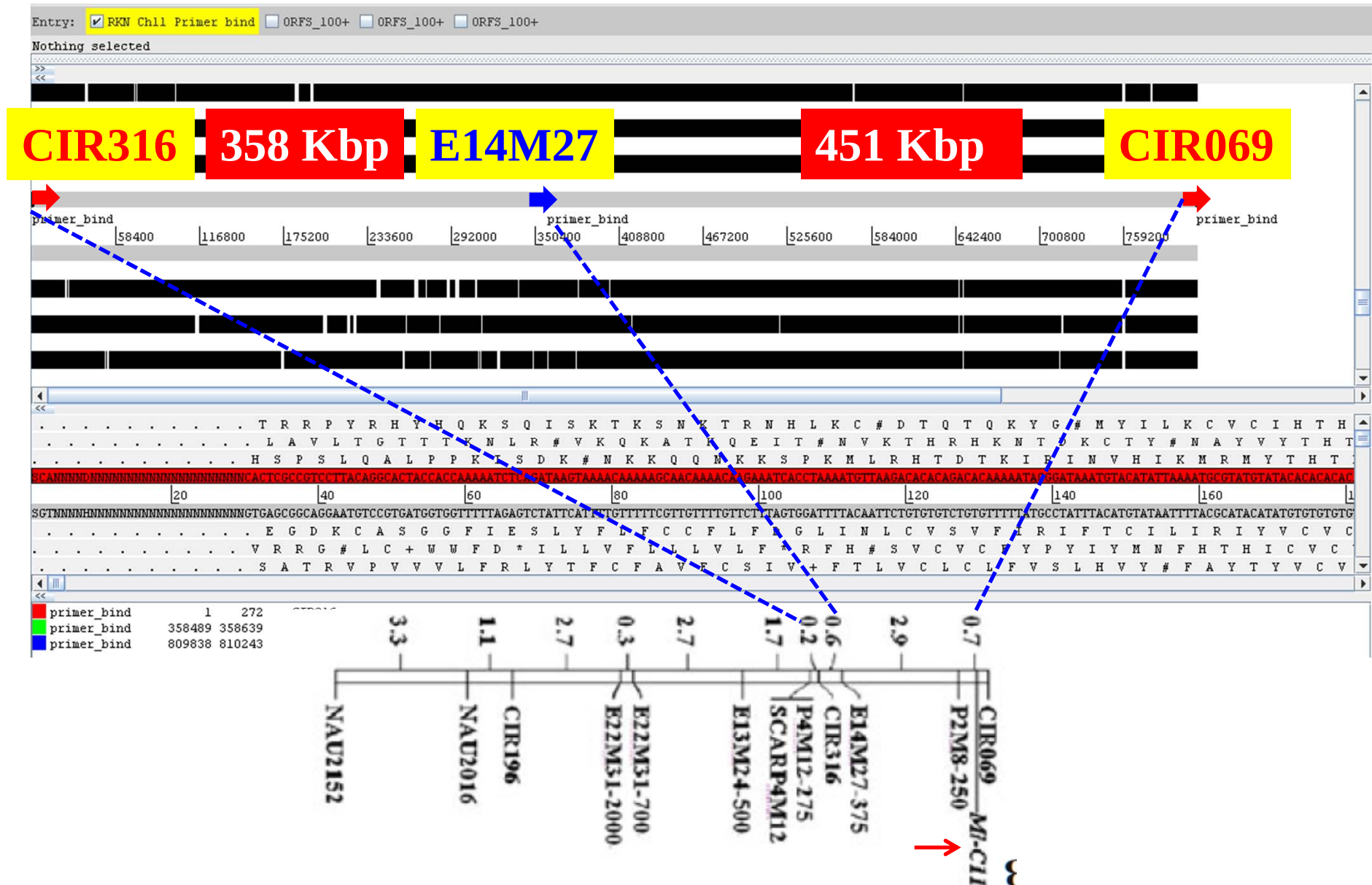
Sequence location of the AFLP marker (E14M27)



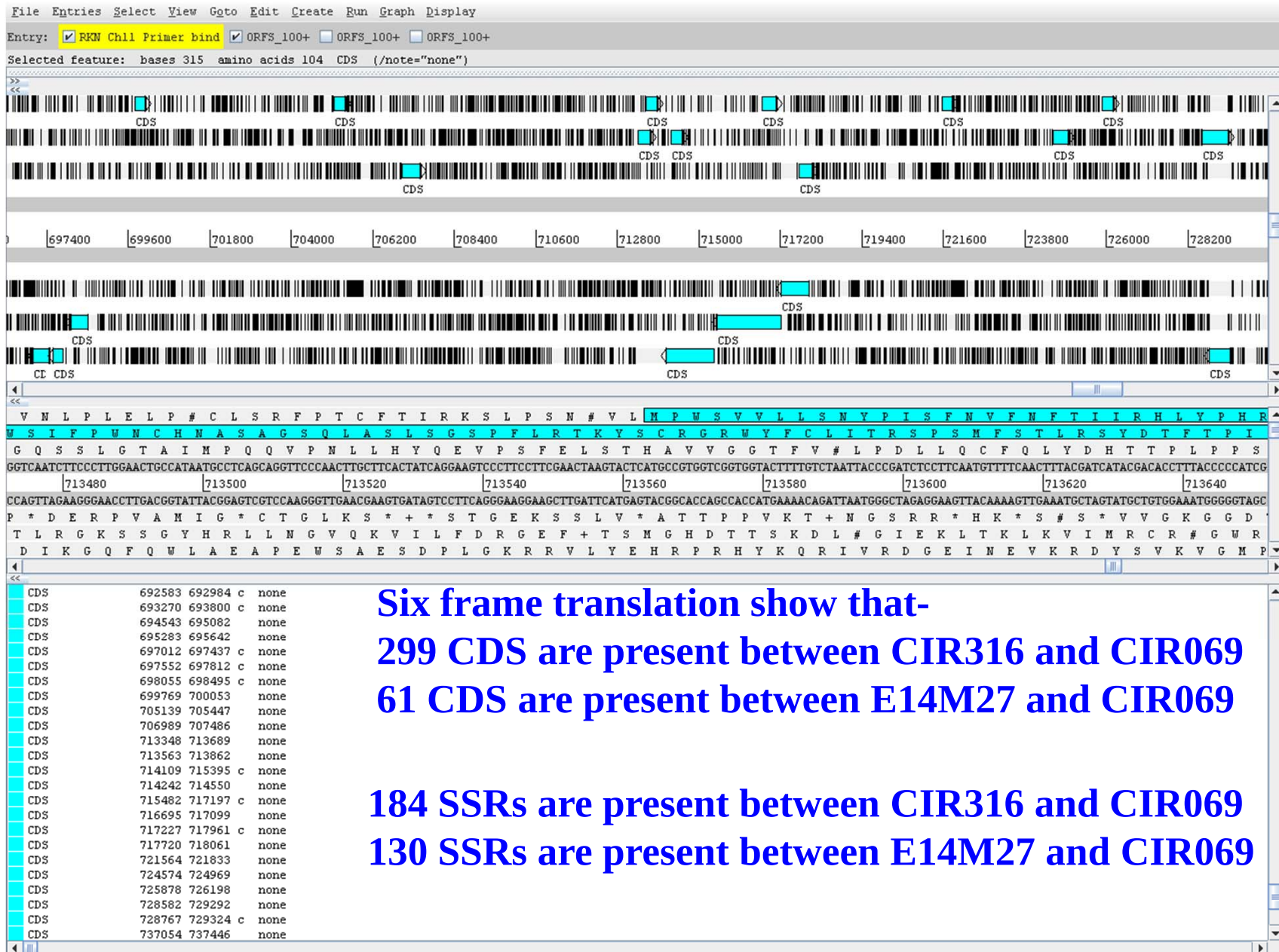
Sequence location of the new AFLP marker (E14M27)



Sequence location of the new AFLP marker (E14M27)



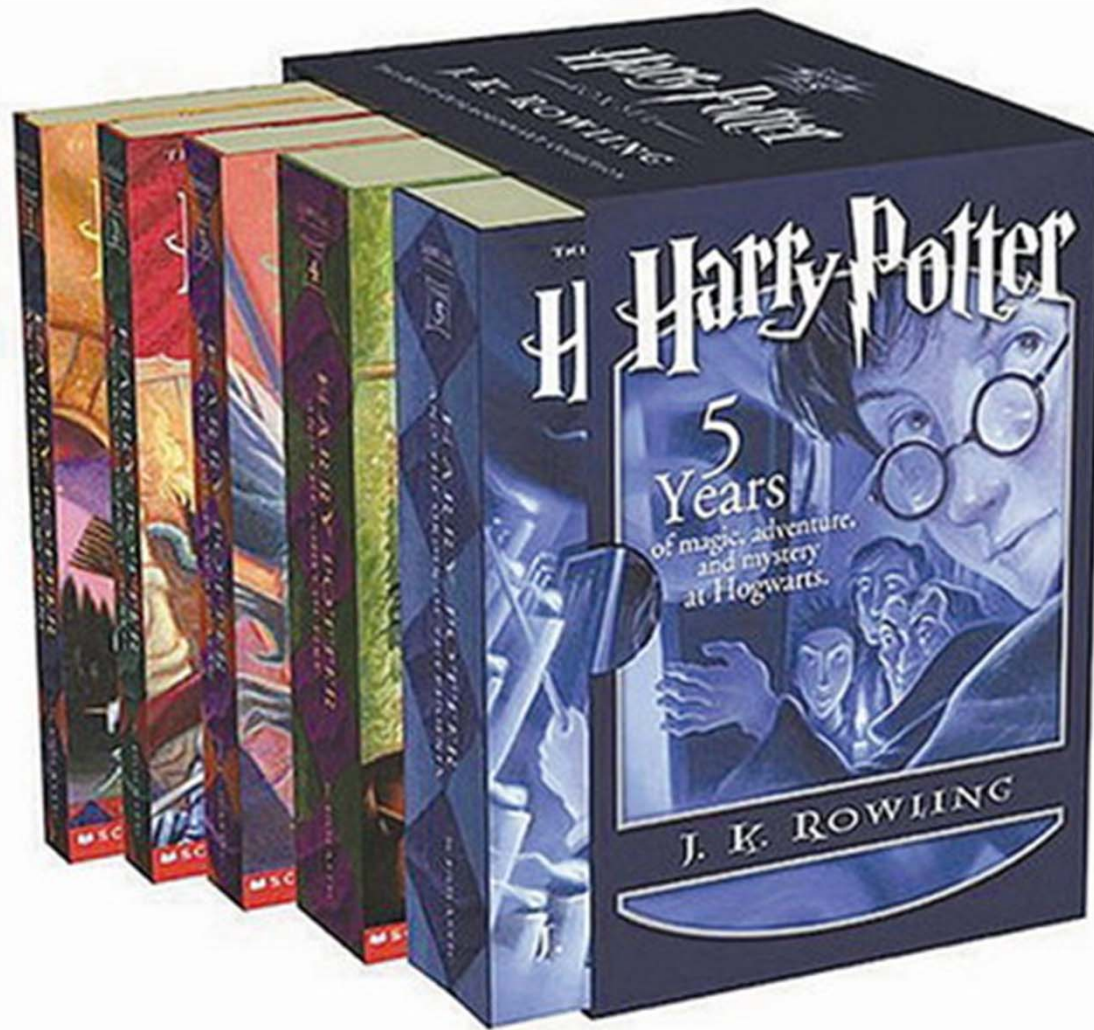
CDS (Coding DNA Sequences) or the ORF (Open Reading Frames) in this region



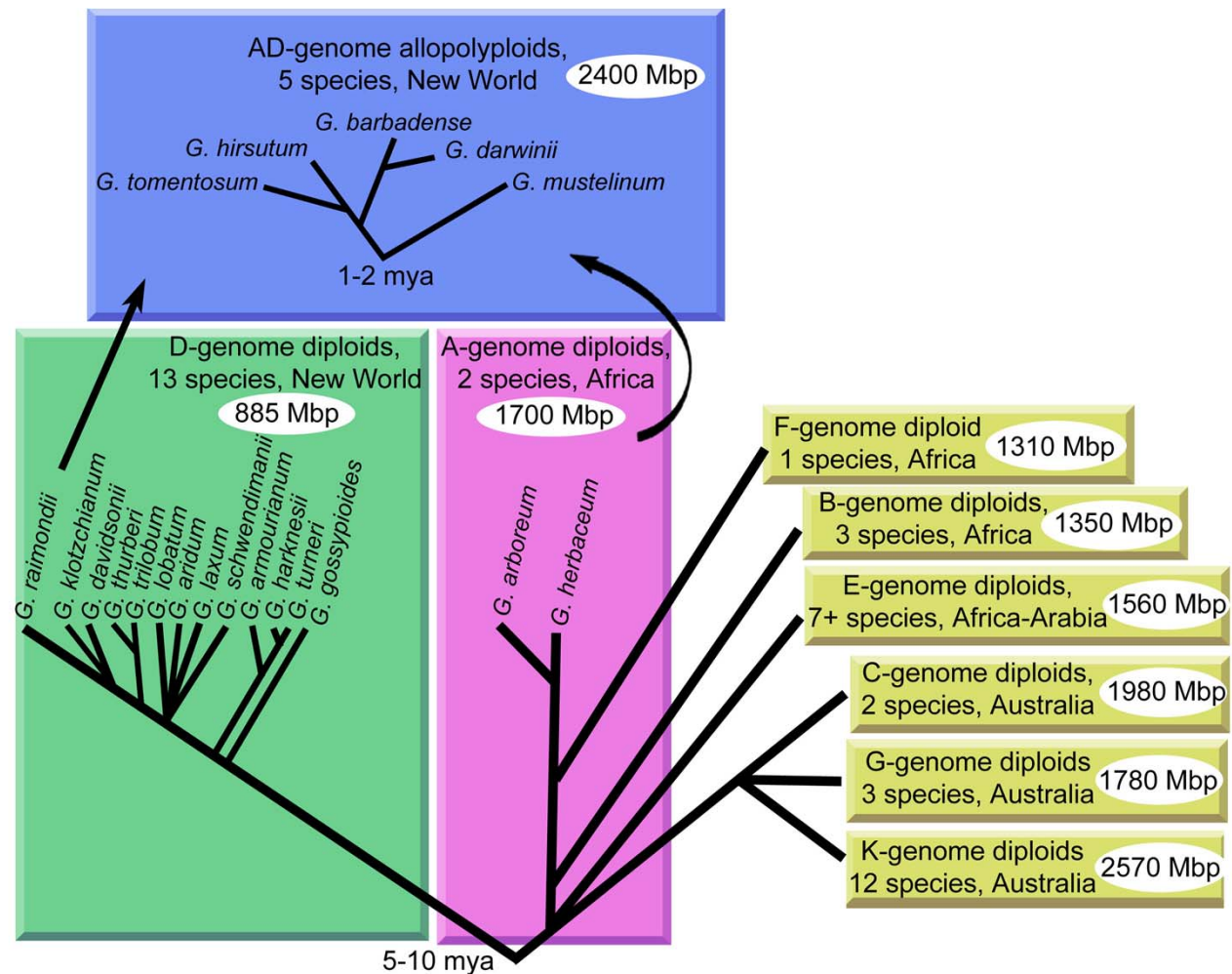
Six frame translation show that-
299 CDS are present between CIR316 and CIR069
61 CDS are present between E14M27 and CIR069

184 SSRs are present between CIR316 and CIR069
130 SSRs are present between E14M27 and CIR069

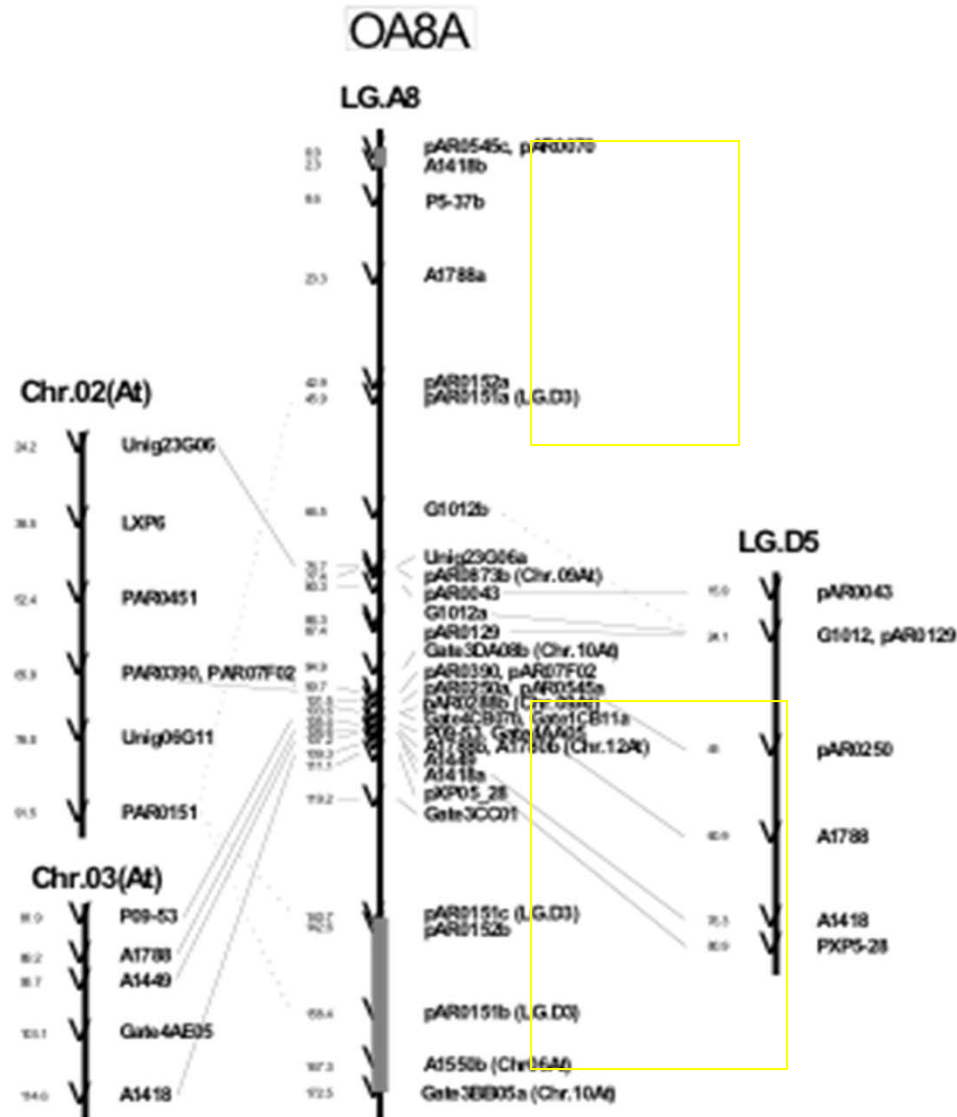
Challenges: 810 Kbp between flanking markers



Challenges: A genome (1.75Gb) is about twice the size of D genome (880Mb)

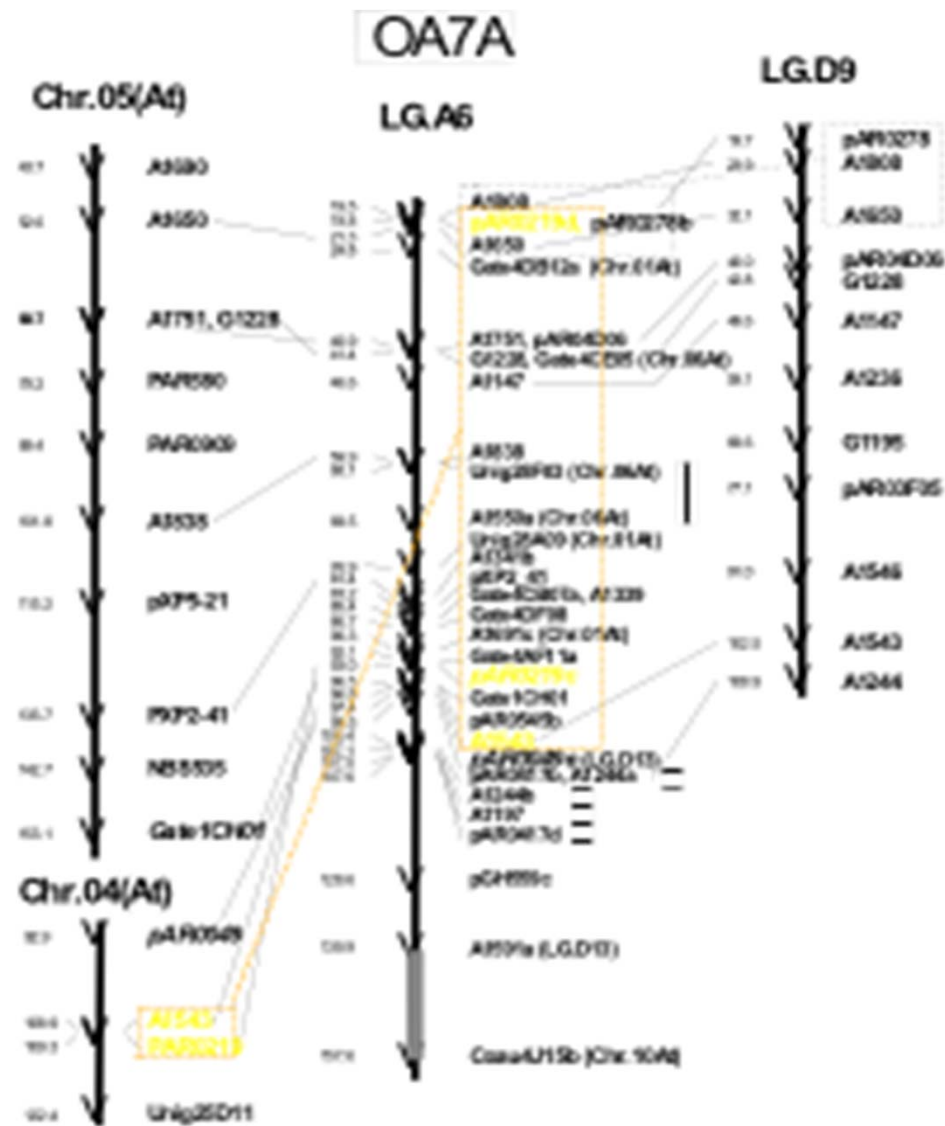


Challenges: chromosomal structural rearrangements between D vs At genome



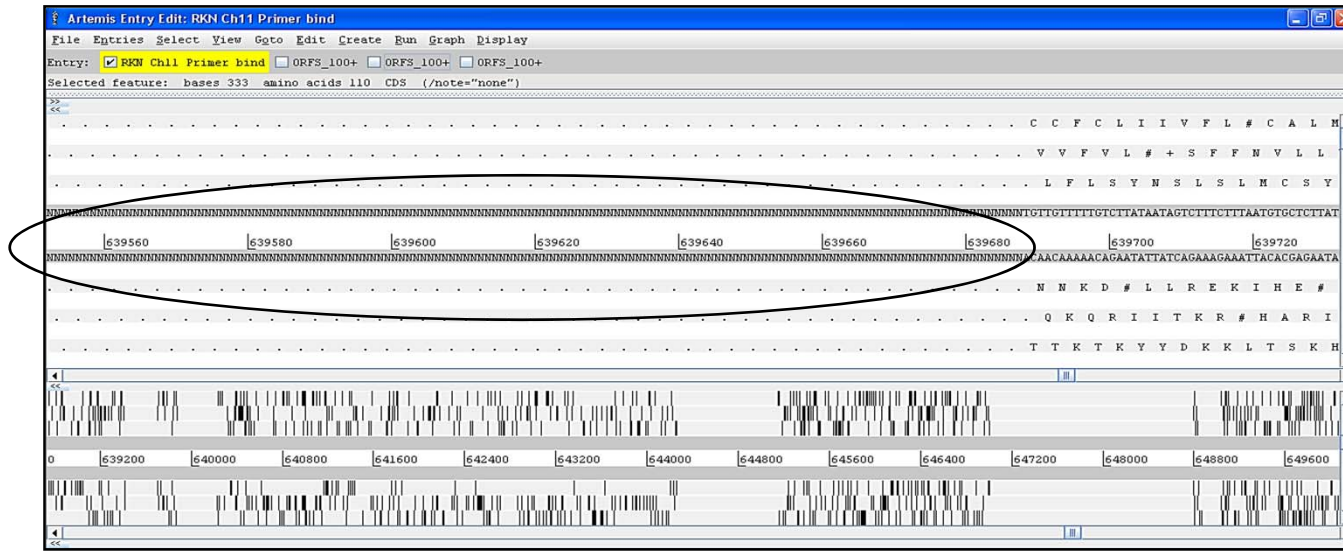
Desai et al. 2006

Challenges: chromosomal structural rearrangements
between D vs At genome

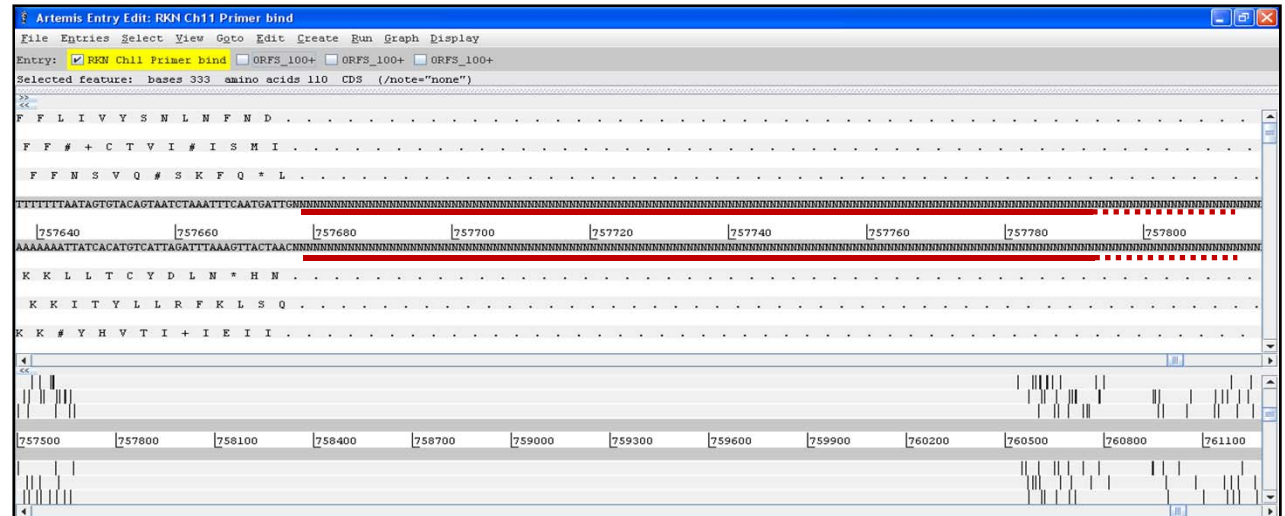


Desai et al. 2006

Challenges: no annotation and gaps in the sequence

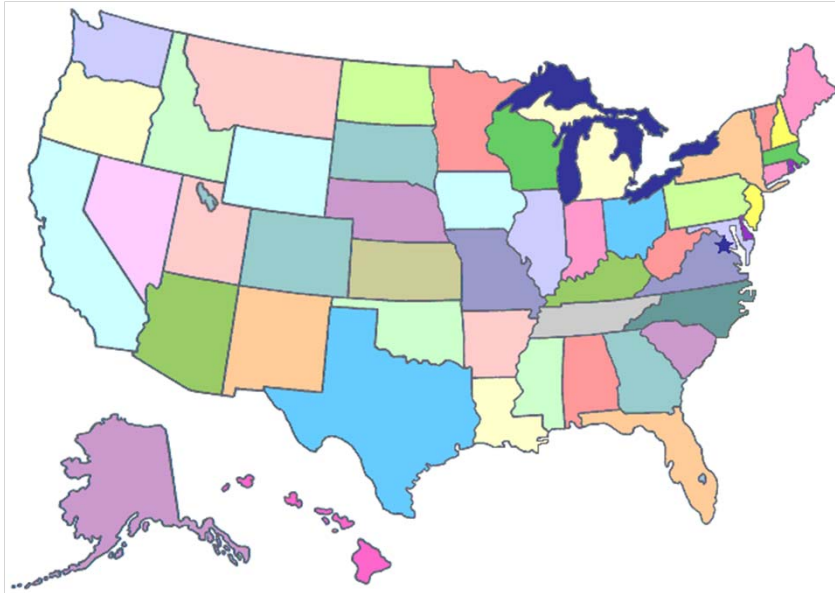


>1 Kb Gap



~ 3 Kb Gap

SUMMARY



START



2010



NOW