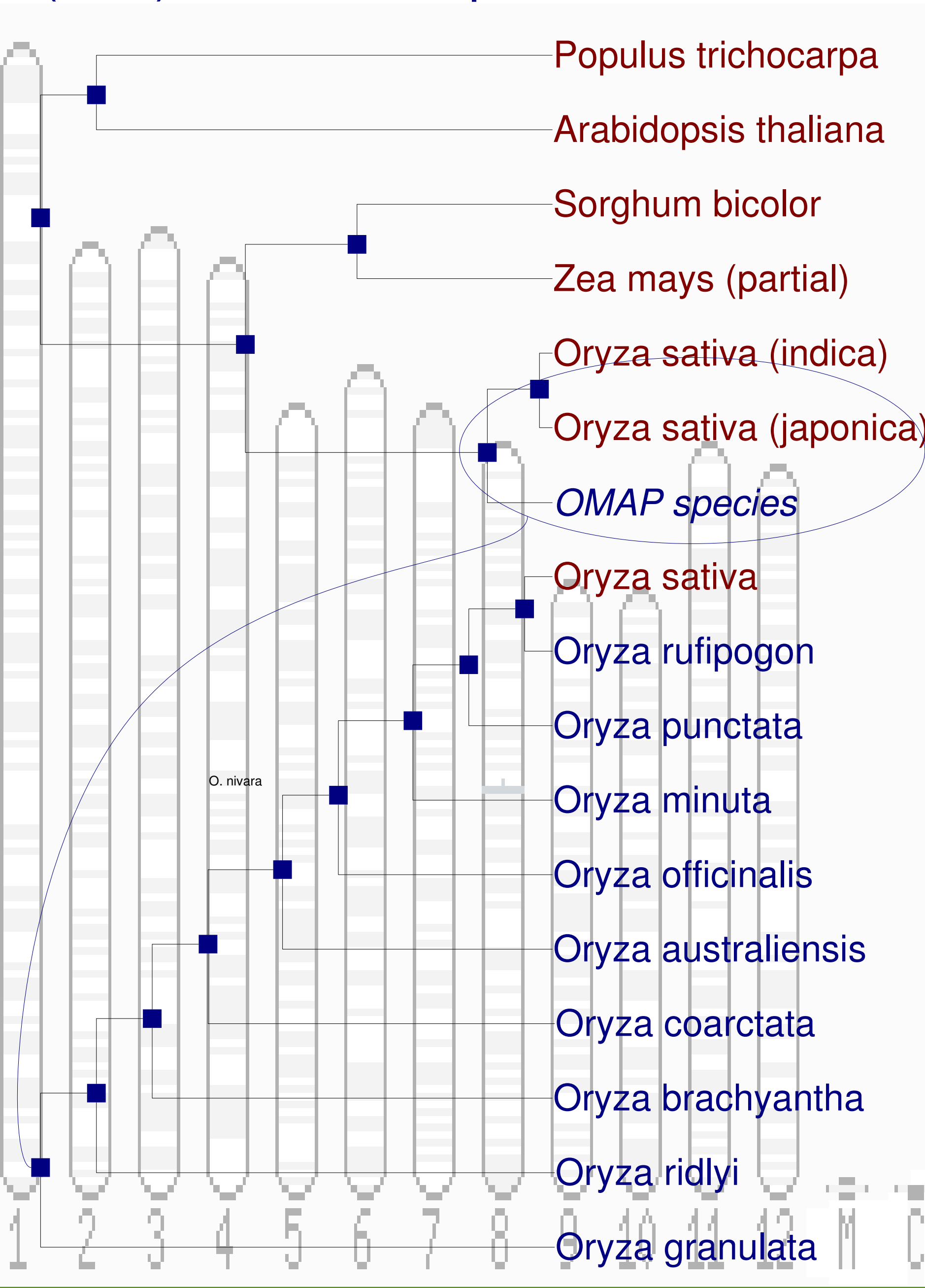# GRAMENE

Comparative Genome Analyses in Gramene

Spooner, William (whs@ebi.ac.uk); Wei, Sharon; Zhao, Wei; Hurwitz, Bonnie; Liang, Chengzhi; Stein, Lincoln; Ware, Doreen

With the sequenced genomes of rice, *Arabidopsis*, poplar and *Sorghum* in the public domain, and the sequencing of maize well under way, the multi-species comparative analysis of plant genome sequences has recently become feasible. The number of plant species covered by whole-genome FPC maps is also increasing; through mapping of marker and clone-end sequences to the sequenced genomes these maps form an additional resource for analysis of genomic synteny.

FingerPrint Contig (FPC) Projection

An atomated method has been developed to project the tiling paths from FPC maps onto sequenced genomes

i. Align clone end sequences from the FPC map clone library to the sequenced genome,

ii. Clones are projected where alignments of both clone ends are consistent in orientation and distance,

iii. FPContigs are projected to cover regions where at least 3 overlapping clones have been projected.

Gramene uses 'Compara' (developed by the Ensembl project, http://www.ensembl.org) as its framework for comparative genome analysis. This consists of database schemas, data analysis pipelines, and visualisation tools. 'Compara' products include; gene trees (with associated orthologs/paralogs), whole-genome DNA alignmnents, and syntenic blocks. Data are currently available from the Gramene Genome Browser and BioMart query interfaces

Plants with Sequence (red) and Physical (blue) Genome Maps (distances not to scale)

Whole-Genome Sequence Alignments

Alignments between two sequenced genomes are performed using either translated BLAT or BLASTZ/Chain/Net depending on phylogenetic distance between the species
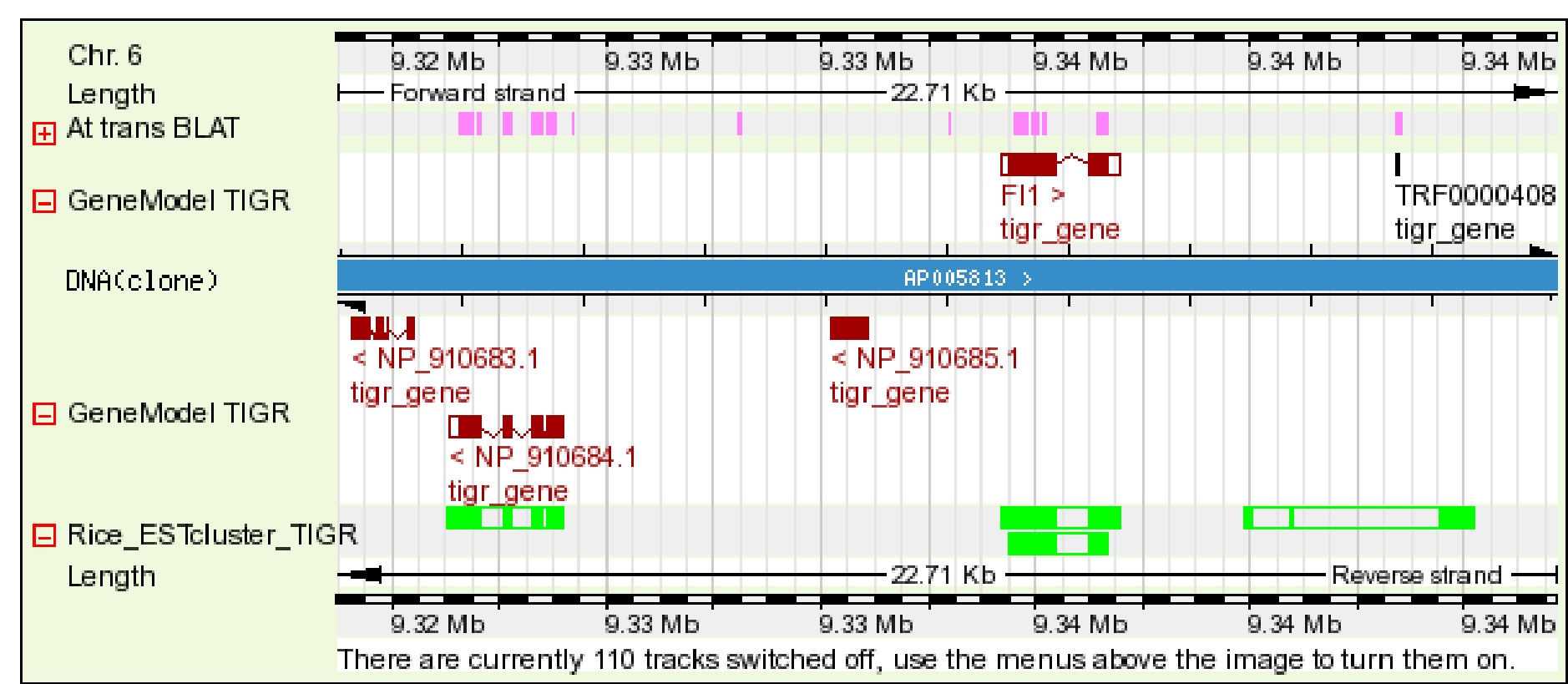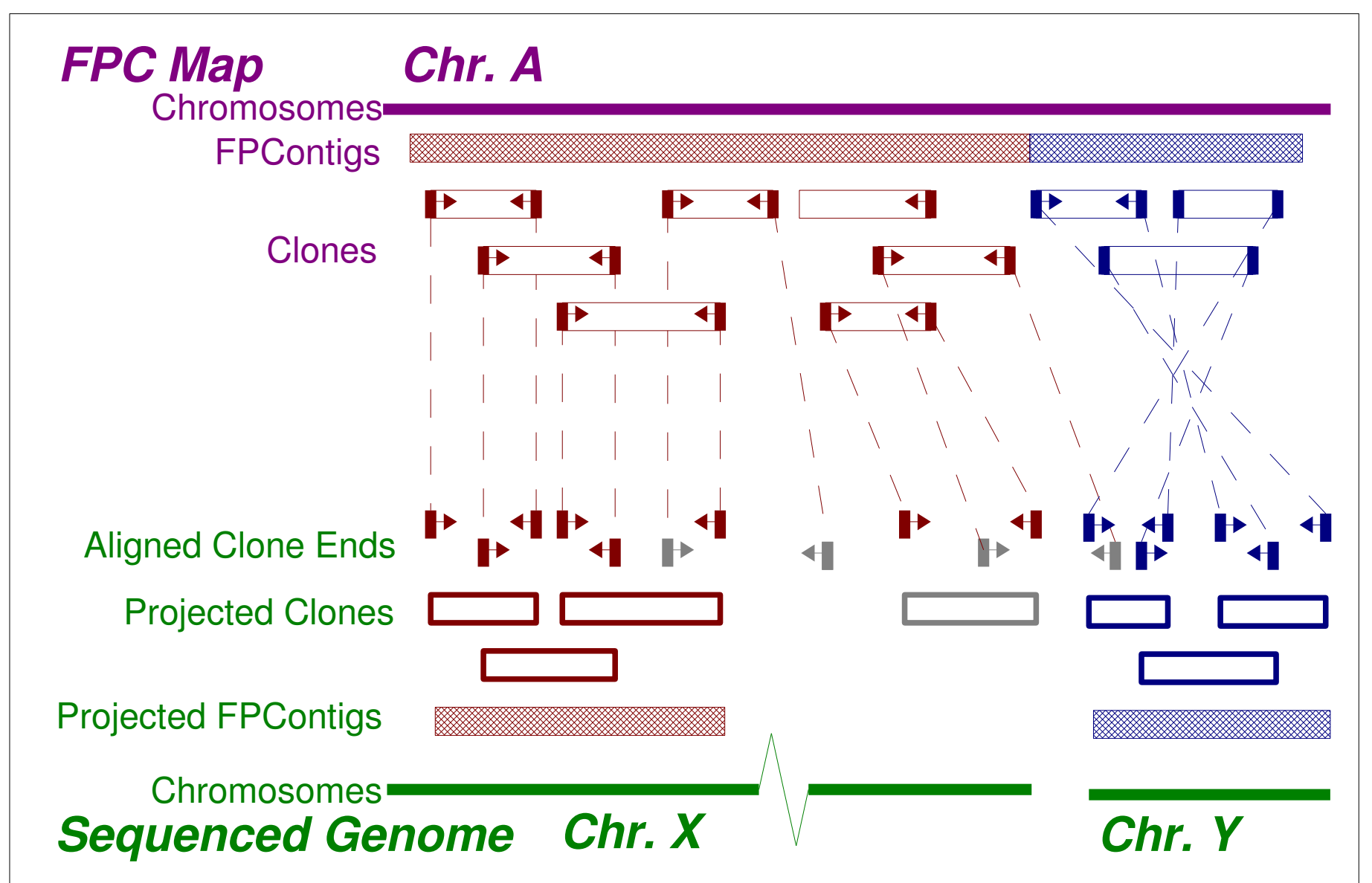

A 22kbase region of Chr 6 the O. sativa genome. The top (pink) track indicates regions that align to the A. thaliana genome. Other tracks show predicted genes, and alignments of rice ESTs to the genome.
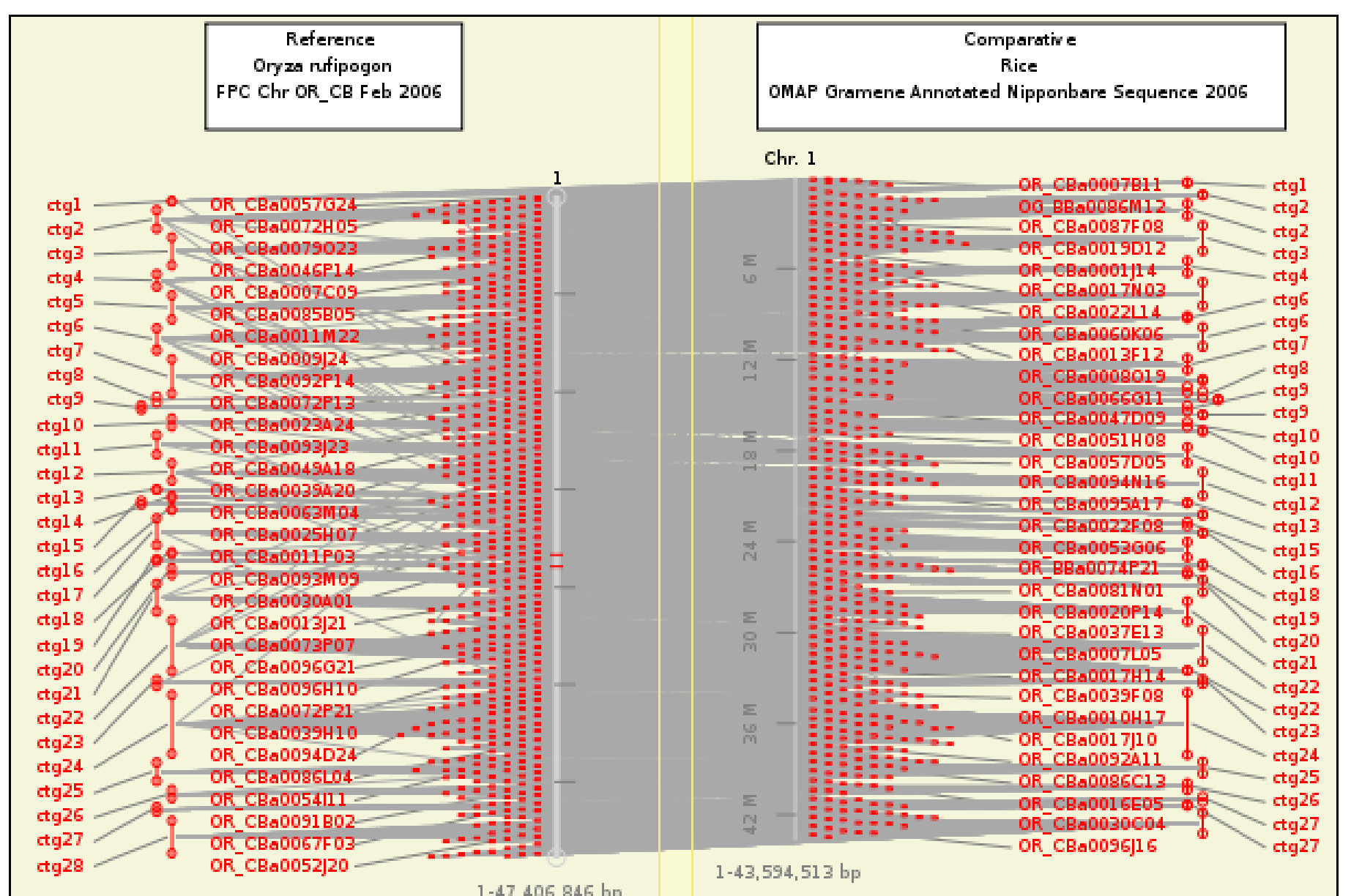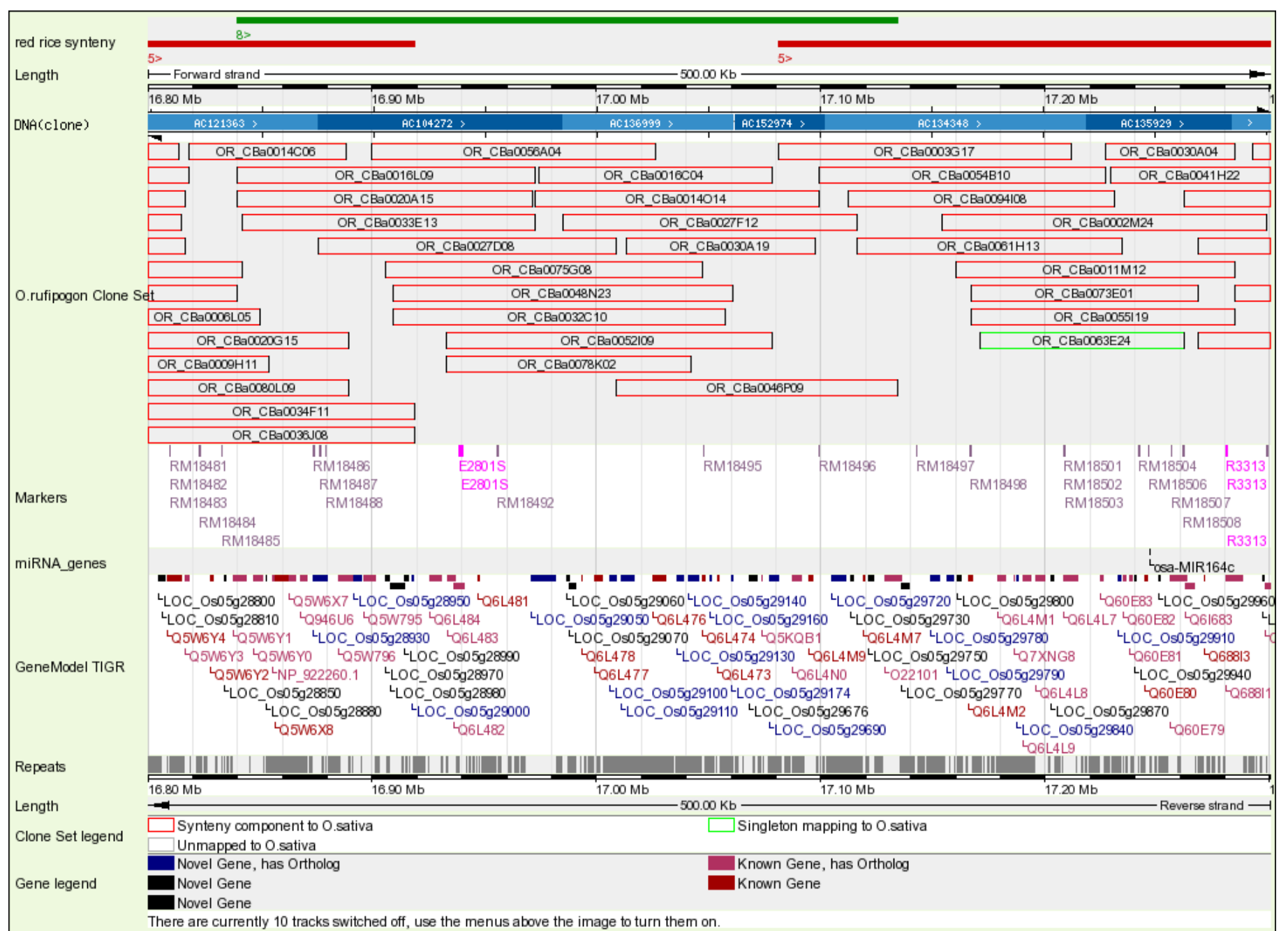

Schematic representation of the FPC Projection Method

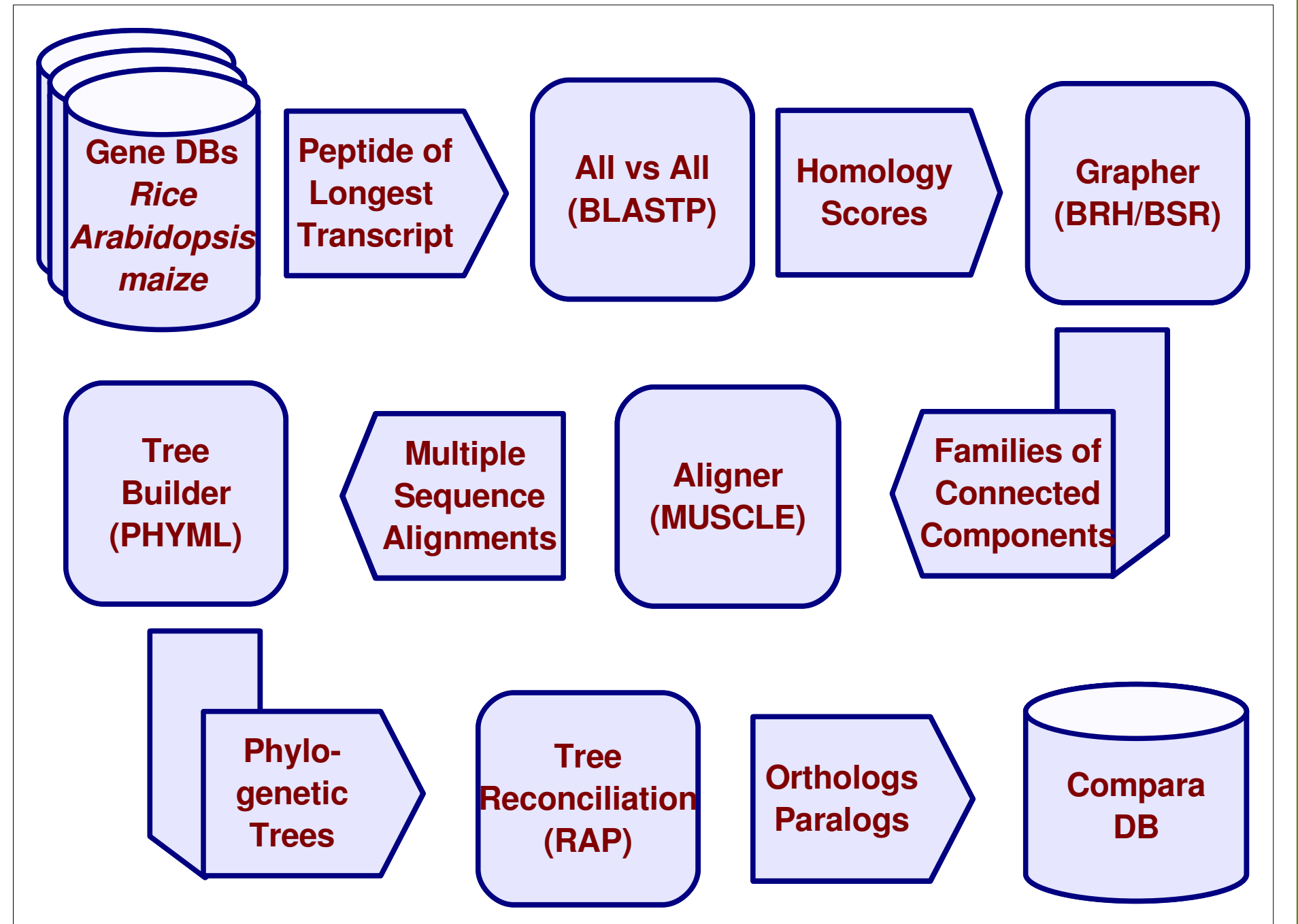*FPC Projection between O. nivara chr1 and O. sativa chr1*

A 5Mbase region of the genome assembly for O. sativa Chr 5 showing a projection of the the BAC clone tiling paths from O. rufipogon.
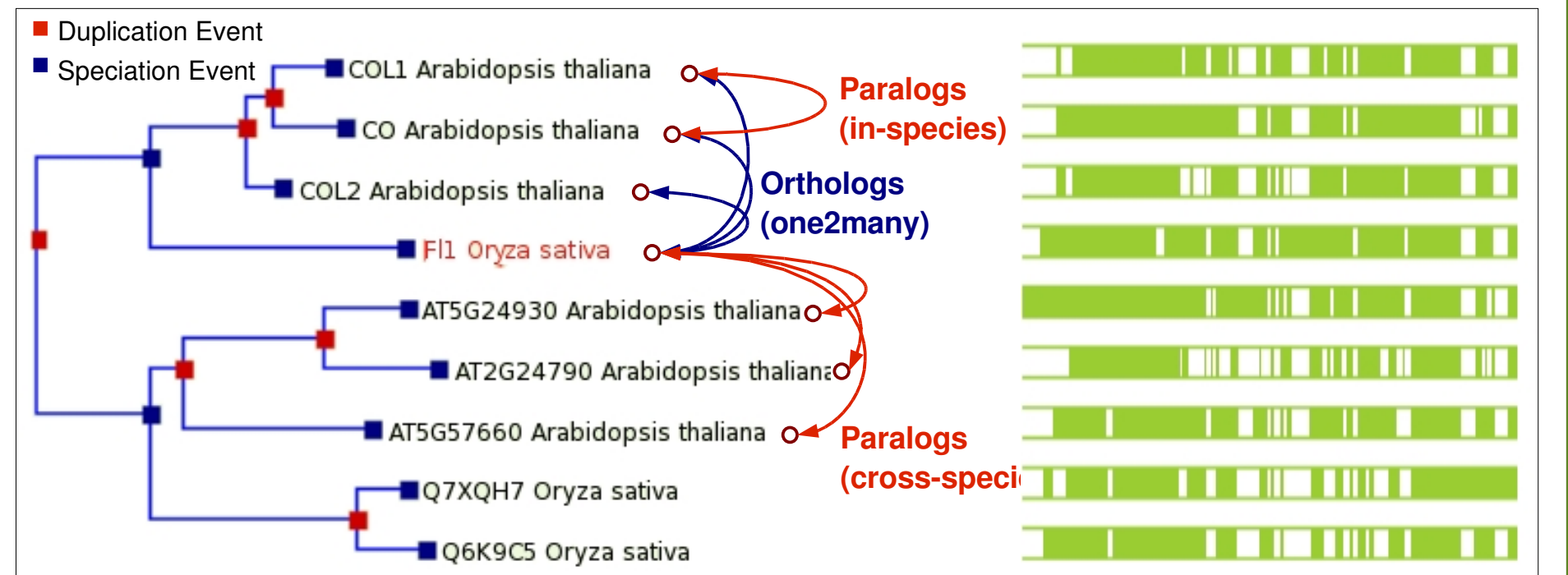
Gene Trees; Pipeline

Gene orthologs and paralogs are determined from maximum likelihood phylogenetic trees based on multiple alignments of homologous gene clusters.
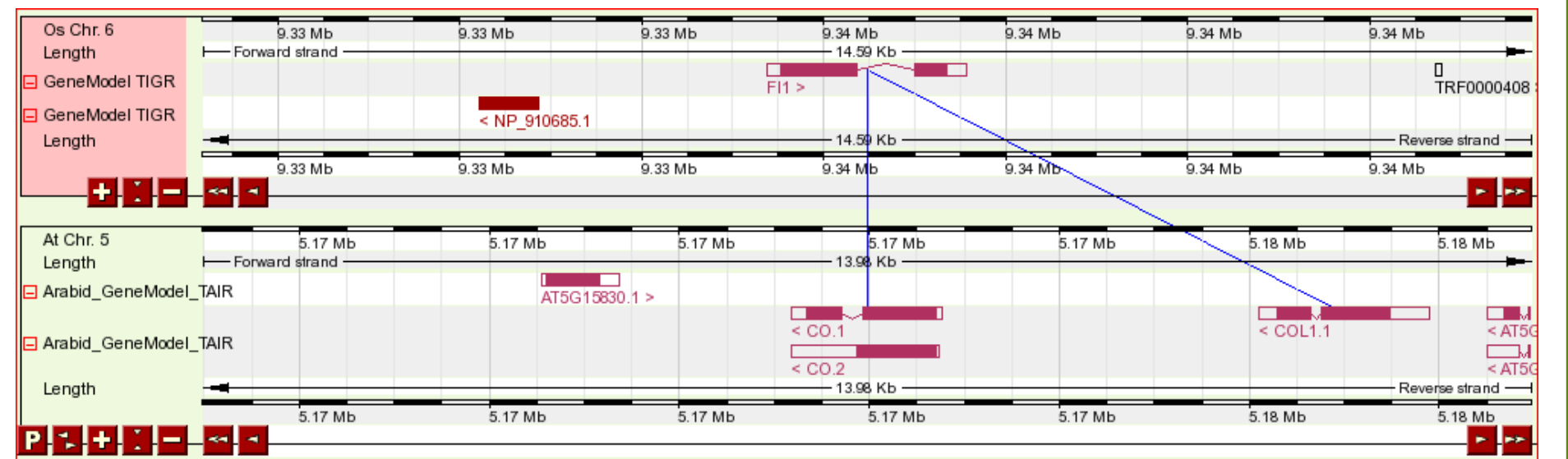

Schematic representation of the "Ensembl gene tree" pipeline

Orthologs can be visualised in 'tree' or "genonic" context;

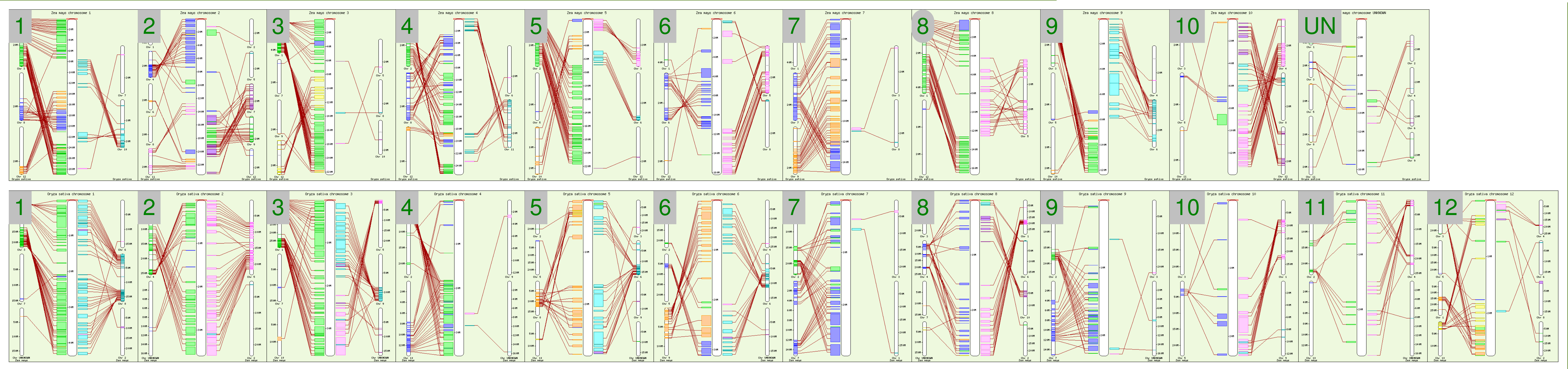*Annotated gene tree centred on rice gene Fl1 (LOC_Os06g16370).*

Genomic context of rice gene Fl1, and two Arabidopsis orthologs, CO and COL1.

Cross-genome Synteny

Syntenic blocks between sequenced genomes and FPContig maps are determined from conserved order/orientation of aligned sequences.

The following example shows maze vs. rice synteny (top) and rice vs. maize synteny (bottom) for each chromosome in the respective genomes