



Project No. **283496**

transPLANT

Trans-national Infrastructure for Plant Genomic Science

Instrument: **Combination of Collaborative Project and Coordination and Support Action**

Thematic Priority: FP7-INFRASTRUCTURES-2011-2

D8.1

Datasets with associations available and integrated into visualisation interfaces

Due date of deliverable: August 31, 2012
Actual submission date: September 7, 2012

Start date of project: 1.9.2011

Duration: 48 months

Organisation name of lead contractor for this deliverable: KeyGene

Project co-funded by the European Commission within the Seventh Framework Programme (2011-2014)		
Dissemination Level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Contributor

KeyGene

Introduction

Deliverable reference number: D8.1

The goal of this task is to link genome sequences with all types of sequence associated data such as genetic markers and other sequence based data. Four species were identified to serve as role models to implement this integration: Rice, Maize, Soybean and *Brassica rapa*. In a next step, an inventory of public resources was made and data was collected for these four species. These include existing genome browsers for the four species, at multiple locations, the sequence databases (EMBL, NCBI) and other types of public data, e.g. for genetic markers and maps. These data was combined and filtered and built into a GBrowse visualization tool.

Methods

Data for each of the four species was collected from different sources, combined and checked on nomenclature. If possible, corresponding entries were identified and discrepancies were eliminated. Genetic maps were gathered and integrated using CMAP. GBrowse is used for visualization. Perl scripts were used for integration, filtering and reformatting.

Results (if applicable, interactions with other workpackages)

Data are visualized in a GBrowse and consist of genomic sequence data, mostly on a pseudo-chromosome level, gene information (orientation, splicing, function), annotated information such as transposons and restriction sites, clone information (BACs) and genetic markers and maps. The GBrowises are hosted on a KeyGene server and a link to this server is provided on the transPLANT website. The GBrowises are available through the transplant website <http://transplantdb.eu/resources> and directly from the KeyGene url:

<http://transplant.keygene.com/>

Screenshots of the GBrowse are included: Fig 1-3.

Publications

n.a.

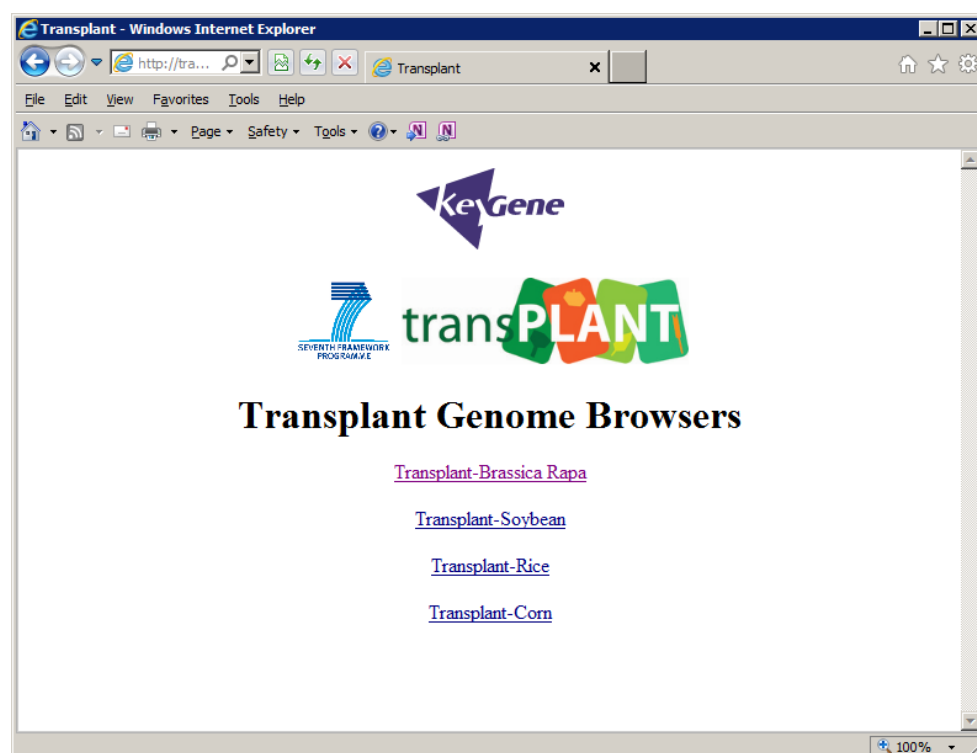


Fig1. Main page for the GBrowises

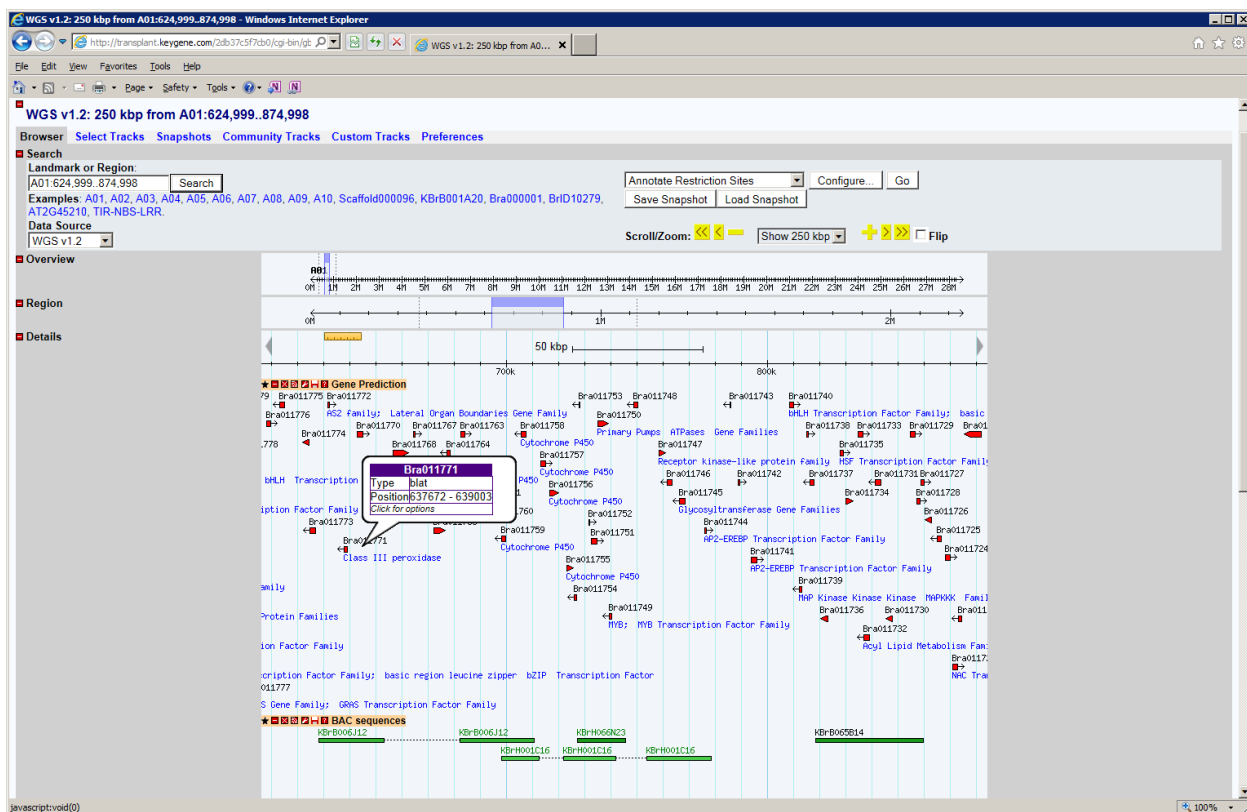


Fig 2. Example of the *Brassica rapa* GBrowse showing gene annotation and BAC clones

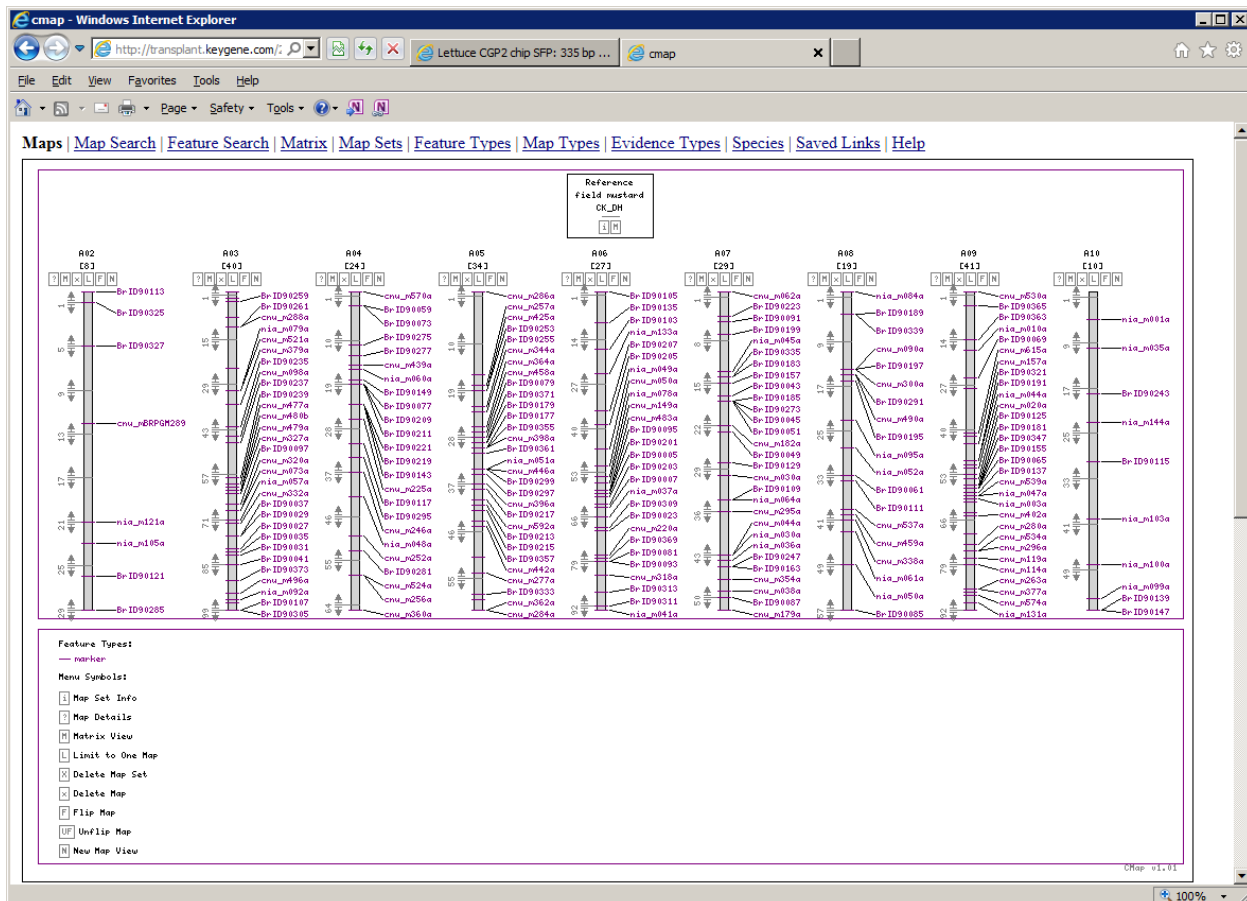


Fig 3. Example of the comparative genetic map viewer for *Brassica rapa*