A New BLAST-based Gbrowse Plugin

Augusto Vellozo, Francisco Lobo, Michel Yamagishi

Embrapa, Embrapa, Embrapa

Nowadays, there are available many freely genome browsers and in order to choose one to be adopted in a project, many criteria [1,2] must be taken into account. GBrowse is one of the most popular genome browsers. It has some interesting qualities, such as an active development community and the possibility to attach customizable modules (plugins) to add new functionalities. In this work, we present a new plugin to search sequences against the genome using the BLAST tool. Furthermore, we have fixed some bugs at GBrowse 2.48 BLAT plugin. These two fully functional plugins enable to query biological sequences (DNA or Amino Acids) against the whole genome. The hit returned by these plugins is displayed as a feature inside a new track. In the case that more than one hit is returned, the plugins show a list of the hits where the user can click on a hit to view it as a feature inside a new track on the related genome region. All steps are done within the GBrowse and the user doesn't need to change the window or tab in the web browser. Despite the active community, the original BLAT plugin provided into the GBrowse 2.48 does not work properly. In fact, all plugins of the type "finder" does not work. We fixed that bugs to enable the use of the BLAT plugin and improved the plugin results filter. In the original Blat plugin, the hits are sorted by the natural order of Blat output (strand and sequence target name) and the user give a maximum number of hits to be shown, but if the query returns more hits than selected by the user then the best hit may not be shown. Using the program pslReps (found into the Blat package) we show, according Blat score, just the more significant results. We developed also a new plugin to search sequences in the genome using the Blast program. The user give the sequences to query, the maximum e-value and the maximum number of sequences to return. The default Blast program output is already sorted by the e-value of the hits. As all steps (searching and viewing) are done into the same application, searching for similar regions of a given sequence became easier and user friendly with these plugins. References: [1] Lacroix, Thomas et al. 2011. CompaGB: An open framework for genome browsers comparison. [2] Taner Z. Sen et al. 2010. Choosing a genome browser for a Model Organism Database: surveying the Maize community.

Keywords: GBrowse BLAST BLAT

Concentration area: Integration