Comparative Genomes Viewer

ID - 204

Sales Gabriele[1], Molineris Ivan[1], Caselle Michele[1]

[1]Department of Theoretical Physics, Università degli Studi, Torino

Motivation

The amount of information about genomes, both complete sequences and annotations, has been exponentially increasing in the last few years. As a result there is the need for tools providing a graphical representation of such informations being both comprehensive and intuitive.

Visual representation is especially important in the comparative genomics field since it should provide a combined view of informations belonging to different genomes.

Most of the existing tools are limited in this respect as they concentrate the attention on a single genome at a time (conservation histograms) or compress alignment representations in a single dimension. Therefore we took example from dot-plots (actually using the full 2D space) and improved them both in term of rendering speed and information content.

Methods

We start by performing whole genome high sensitivity alignments using the BLAST software (Washington University implementation).

We treat each alignment as a segment in a bidimensional plane whose sides correspond to the two chromosomes being aligned (as in a dot-plot). We cluster bidimensionally these segments and use the resulting hierarchical structure to store them efficiently in a MySQL database; this form has the added benefit of making queries extremely fast (up to 40 times with respect to the naive format).

Given a pair of regions in chromosomal coordinates, our system recovers all the overlapping alignments (segments) and presents them graphically; the background of the image is coloured (using horizontal and vertical stripes) depending on the functional annotations.

Moreover we recover gene annotations and other tracks from the UCSC genome browser and we show them at both sides of our plot.

Results

We developed a web based tool called Comparative Geomes Viewer (CGV): it integrates the expressive power of dot-plots (allowing an intuitive representation of alignments between two regions, both at small and big scales) with the richness of informations typical of existing genome browsers.

We give access to our system through a web based interface allowing the user to update the representation interactively and in real time, for example by easy to use (click and drag) shifting and zooming features.

Our tool is also embeddable in external web pages by means of a straightforward AJAX API.

Availability: http://to444xl.to.infn.it/cgv

Image: http://to444xl.to.infn.it/cgv/bits2007_abstract.png

Email: sales@to.infn.it