A number of comparative studies of genomic data have shown that genomics tools, including MultiPipMaker (Schwartz et al., 2003), and the UCSC genome browser (Karolchik et al., 2008), are well studied with comparative evolutionary information in genomic data. Conservation patterns along genome sequences are recognized as those defined at NCBI, the system automatically fetches the annotation data from NCBI through the TogoWS system (http://www.ncbi.nlm.nih.gov/Sitemap/samplerecord.html) or the gff data format (http://www.sequenceontology.org/gff3.shtml).

In the case where the sequence names in the uploaded alignment data are recognized as those defined at NCBI, the system automatically fetches the annotation data from NCBI through the TogoWS system (http://togows.dbcls.jp/).

Users can move about the entire dot plot by dragging the comparison pane and zoom it in/out with the mouse wheel in the same manner as the Google Maps. The navigation controls, placed...
Fig. 1. The main page of CGAS, consisting with comparison and annotation panes. This example shows a comparison of the regions including the opsin loci in the X chromosomes of human and chimpanzee (A: single mode) and macaque (B: dual mode). Red and blue dots/lines in the comparison pane represent the BLAST hits in the same/different directions between the query and the subject.

at the upper-left corner of the comparison pane, may be used to navigate. Diagonal movements are possible using the navigation controls. Thanks to the Google Maps API, the annotation panes are updated synchronously with the comparison pane.

The names of the annotated objects are shown in a small box, which can be expanded by moving the mouse cursor on it and shows annotation details along with the link to the original data at NCBI. If the mouse cursor is within the comparison pane, the position of the cursor in each of the query and the subject(s) is indicated by a red thin line in the annotation pane with the position information in base pairs. The current version of CGAS is capable of displaying two subjects side by side (Fig. 1B). Two comparison panes are synchronously aligned to the query. This functionality allows the users to compare the differences in collinearity between the subjects.

CGAS provides a unique function to align the query and the subject annotation panes based on the putative orthology relationships between the sequences compared. If the user double-clicks in the subject or query panes, the most similar region of the clicked region in the query or subject sequence is centered in the pane, and the comparison pane is updated accordingly. This is useful for tracking the significant hits.

The uploaded data are indexed with user’s login ID and retained in the server for a week after the last access or until the user deletes them.

3 IMPLEMENTATION

The CGAS is a web server with the Ruby on Rails framework with JavaScript. The interface within the dot plot viewer and the annotation viewer was implemented using Google Maps API version 2. Interpretation of biological data and interactive entry fetching from public database were achieved by the BioRuby library (http://bioruby.org/) and the TogoWS system. The documentation of CGAS is provided at http://cgas.ist.hokudai.ac.jp/help/.

ACKNOWLEDGEMENTS

The authors wish to thank the developers of BioRuby, Ruby on Rails, and Google Maps API.

Funding: Grant-in-Aid for Global COE program ‘Center for Next-Generation Information Technology based on Knowledge Discovery and Knowledge Federation’ from the MEXT, Japan.

Conflict of Interest: none declared.

REFERENCES

Abril, J. et al. (2003) gff2aplot: Plotting sequence comparisons. Bioinformatics, 19, 2477–2479.
Altschul, S. et al. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res., 25, 3389–3402.
Engels, R. et al. (2006) Combo: a whole genome comparative browser. Bioinformatics, 22, 1782–1783.
Jareborg, N. and Durbin, R. (2000) Alfresco—a workbench for comparative genomic sequence analysis. Genome Res., 10, 1148–1157.
Karolchik, D. et al. (2008) The UCSC Genome Browser Database: 2008 update. Nucleic Acids Res., 36, D773–D779.
Schwartz, S. et al. (2000) PipMaker—a web server for aligning two genomic DNA sequences. Genome Res., 10, 577–586.
Schwartz, S. et al. (2003) MultiPipMaker and supporting tools: alignments and analysis of multiple genomic DNA sequences. Nucleic Acids Res., 31, 3518–3524.
Sommer, A. and Durbin, R. (1995) A dot-matrix program with dynamic threshold control suited for genomic DNA and protein sequence analysis. Gene, 167, GC1–GC10.
von Mering, C. et al. (2003) STRING: a database of predicted functional associations between proteins. Nucleic Acids Res., 31, 258–261.
Wilson, M. et al. (2001) Comparative analysis of the gene-dense ACHE/TFR2 region on human chromosome 7q22 with the orthologous region on mouse chromosome 5. Nucleic Acids Res., 29, 1352–1365.