Supplementary Data

Synteny Portal: a web-based application portal for synteny block analysis

Jongin Lee, Woon-young Hong, Minah Cho, Mikang Sim, Daehwan Lee, Younhee Ko, and Jaebum Kim

Server specification and sustainability of our web server

The Synteny Portal web server is installed and running in the DELL server with 16 cores of the Intel (R) Xeon (R) 2.40 GHz CPU, 96 GB memory, and 40 TB disk storage. We are using a job scheduler and time-consuming tasks requested by users are handled by the job scheduler. Therefore, our web server can continue to serve its functions to users even with a moderate number of simultaneous user connections.

Synteny block construction

To construct synteny blocks, we used a program in the inferCars software (http://www.bx.psu.edu/miller_lab/car/). The program takes prebuilt chain/net files as input. The chain/net files (https://genome.ucsc.edu/FAQ/FAQformat.html) are ones of widely used alignment formats. In the UCSC genome browser, the chain/net files of many species are already available, and we downloaded and used them to run the synteny block construction program. The program can be easily run by specifying the list of species and a resolution (the minimum synteny block size).

The motivation of the development of our web server was the reuse of the whole-genome alignments from the UCSC genome browser database in order to reduce user-side burden for the construction of whole-genome sequence alignments. The reasons we chose inferCars as a tool for the construction of synteny blocks are followings: 1) chain/net files are widely used in the UCSC genome browser database and thus, they are one of easily accessible formats. inferCars is designed to support these formats. 2) inferCars is one of the widely used programs in many studies since its publication in 2006 (Genome Research). Note that this manuscript has been cited more than 200 times.

BLAST search

To search for reference genome sequences by using a given DNA or protein sequence, we first created a BLAST database for each reference species from its genome sequence. Given a user-provided sequence, we run BLASTN (for a DNA sequence) or TBLASTN (for a protein sequence) against the prebuilt BLAST database with default parameters, and the top-matched result is retrieved. With the coordinates in the BLAST result, overlapped syntey blocks with our results are identified and shown in a result webpage.

Table S1. Tool comparison results

Software Name	User Interface	Visualization of results	Multiple-species comparison	Fully functional without user-provided data	Linear plot support	Circular plot support	Raw synteny file saving/downloading	High-resolution image saving/downloading	URL
C-Sibelia	GUI (web server)	Yes	No	No	No	Yes	Yes	Yes	http://etool.me/software/csibelia
Cinteny	GUI (web server)	Yes	No	Yes	Yes	No	Yes	No	http://cinteny.cchmc.org
CoGe	GUI (web server)	Yes	No	Yes	Yes	No	Yes	Yes	https://genomevolution.org/coge/
CYNTENATOR	CLI	No	Yes	No	No	No	Yes	No	http://www.bioinformatics.org/cyntenator/wiki/
DRIMM-synteny	CLI	No	Yes	No	No	No	Yes	No	http://bix.ucsd.edu/projects/drimm/
Ensembl Synteny View	GUI (web server)	Yes	No	Yes	Yes	No	Yes	Yes	http://www.ensembl.org/index.html
Gbrowser Syn	GUI (web install)	Yes	Yes	No	Yes	No	Yes	No	http://gmod.org/wiki/GBrowse_syn
GenomeMatcher	GUI (stand-alone)	Yes	No	No	Yes	No	Yes	Yes	http://www.ige.tohoku.ac.jp/joho/gmProject/gmhome.html
Genomicus	GUI (web server)	Yes	Yes	Yes	Yes	No	Yes	Yes	http://www.genomicus.biologie.ens.fr/genomicus-83.01/cgi-bin/search.pl
GIMM-synteny	GUI (web server)	Yes	Yes	No	No	No	No	No	http://grimm.ucsd.edu/cgi-bin/grimm.cgi
GSV	GUI (web server)	Yes	No	No	Yes	No	Yes	No	http://cas-bioinfo.cas.unt.edu/gsv/homepage.php
i-ADHoRe3.0	CLI	Yes	Yes	No	Yes	No	Yes	Yes	http://bioinformatics.psb.ugent.be/software/details/i--ADHoRe3.0
inferCars	CLI	No	Yes	No	No	No	Yes	No	http://www.bx.psu.edu/miller_lab/car/
Mauve	GUI (stand-alone)	Yes	Yes	No	Yes	No	Yes	Yes	http://darlinglab.org/mauve/mauve.html
mGSV	GUI (web server)	Yes	Yes	No	Yes	Yes	Yes	No	http://cas-bioinfo.cas.unt.edu/mgsv/
MizBee	GUI (stand-alone)	Yes	No	No	Yes	Yes	Yes	No	http://www.cs.utah.edu/~miriah/mizbee/Overview.html
Osfinder	CLI	Yes	Yes	No	Yes	No	Yes	Yes	http://osfinder.dna.bio.keio.ac.jp/index.html
Sibelia	CLI	Yes	Yes	No	No	Yes	Yes	Yes	http://bioinf.spbau.ru/sibelia
Sybil	GUI (web install)	Yes	Yes	No	Yes	No	Yes	Yes	http://sybil.sourceforge.net
SyMAP	GUI (stand-alone)	Yes	Yes	No	Yes	Yes	Yes	Yes	http://www.symapdb.org
Synteny Portal	GUI (web server)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	http://bioinfo.konkuk.ac.kr/synteny_portal/
VISTA	GUI (web server)	Yes	No	Yes	Yes	No	No	No	http://pipeline.lbl.gov/cgi-bin/gateway2

Description of each column

User Interface

GUI (stand-alone): graphical user interface, stand-alone program (users need to download and install it)
GUI (web server): graphical user interface, directly usable web server
GUI (web install): graphical user interface, web server installation package (users need to download and install it)
CLI: command-line interface

Visualization of results

Some CLI-based programs also support stand-alone viewer programs or files that can be easily visualized by web browsers or the Circos program

Multiple-species comparison

Whether users can compare multiple species or not. "No" indicats the program only supports pairwise comparison.

Fully functional without user-provided data

If yes, tools can work without user-provided data (sequences or alignments).

Raw synteny file saving/downloading

Whether users can save/download text files that contain synteny information or not.

High-resolution image saving/downloading

We assumed that if CLI-based programs support the visualization of results, then they also support the high-resolution image saving/downloading.