Fingerprint-based Minimal Tiling Path (FMTP)

Summary

FMTP is a software tool to compute the MTP of a physical map based purely on restriction fingerprint data (and the contigs). FMTP completely ignores the ordering of clones obtained by the physical map algorithm.

Features

- Does not require sequence or ordering of clones
- Generates MTPs with higher coverage than MTPs by conventional tools
- Does not require any commercial tools/libraries
- Compatible on Linux and Mac OS X.

Download

To download latest version of FMTP and sample data files click FMTP_v1.zip.

How to install FMTP

Software requirements

- Linux/Unix or Mac OS X operating systems
- C++ compiler
- Boost C++ Library
- GLPK (GNU Linear Programming Kit)
- Perl

To install FMTP

Open a terminal window and type

unzip FMTP_v1.zip

Type

cd /absolute/path/FMTP_v1/source/

Modify the the path of the Boost C++ library in **Makefile-MODIFY_BEFORE_USING**.

Change the filename of **Makefile-MODIFY_BEFORE_USING** to **Makefile**

Type

make

to compile FMTP.

To run FMTP, enter **run_dir** directory and type

./run_script_FMTP

to run FMTP on the sample data set. You could change the parameter in the script files to run it on another data set.

FMTP Parameters

FMTP consists of two modules, namely MST and ILP. Most of the parameters are common but there are some parameters unique to each module. Below is the list of parameters and their description. Please note that the scripts in the run_dir directory already use the default parameters to run FMTP. You just need to modify these scripts if you want to change any parameters.

ILP_Module Parameters

-c ctg_BAC_clone file
-f Fingerprinting Method.. 1: HICF, 2: Agarose
-s size file
-b bury threshold (default 80)
-B cooperative bury threshold (default 80)

-t sulston_score_threshold.. Fragments of clones will be matched if their sulston score is lower than this threshold..(Default values agarose:1e-10, HICF:1e-25)
-T Tolerance (rice:7, barley:3, cowpea:5)..
-G Gellength (rice: 3300, barley:18000, cowpea: 36,000)..
-o output_file base file (without extension).. Linear Programming Model file (for GLPK, in MathProg language)
-O partial ctg_2_mtp_file
-v verbose level min 0 max 4

MST_Module Parameters

-c ctg_BAC_clone file
-f Fingerprinting Method.. 1: HICF, 2: Agarose
-s size file
-b bury threshold (default 80)
-B cooperative bury threshold (default 80)
-t sulston_score_threshold.. Fragments of clones will be matched if their sulston score is lower than this threshold..(Default values agarose:1e-2, HICF:1e-5)
-T Tolerance (rice:7, barley:3, cowpea:5)..
-G Gellength (rice: 3300, barley:18000, cowpea: 36,000)..
-o output_file base file (without extension).. Linear Programming Model file (for GLPK, in MathProg language)
-O output ctg 2 ordered_clones (OPTIONAL)
-e putative ovErlapping pairs File.. when MST is built.. it's check which clones are overlapping and which are not.. putative overlapping pairs are stored here..(OPTIONAL)
-n putative non-overlapping pairs File.. when MST is built, it's check which clones are overlappig and which are not.. putative non-overlapping pairs are stored here... (OPTIONAL)
-v verbose level min 0 max 4

Copyright

For any questions please contact Serdar Bozdag via bozdags@mail.nih.gov

How to cite

If you use our tool in your project, please cite the following publication.

S. Bozdag, T. J. Close, S. Lonardi, "Computing the minimal tiling path from a physical map by integer linear programming", Proceedings of Workshop on Algorithms in Bioinformatics (WABI'08), pp 148-161