snpTree - a web-server to identify and construct SNP trees from whole genome sequence data. - DTU Orbit (30/08/2017)

snpTree - a web-server to identify and construct SNP trees from whole genome sequence data.

Background
The advances and decreasing economical cost of whole genome sequencing (WGS), will soon make this technology available for routine infectious disease epidemiology. In epidemiological studies, outbreak isolates have very little diversity and require extensive genomic analysis to differentiate and classify isolates. One of the successfully and broadly used methods is analysis of single nucleotide polymorphisms (SNPs). Currently, there are different tools and methods to identify SNPs including various options and cut-off values. Furthermore, all current methods require bioinformatic skills. Thus, we lack a standard and simple automatic tool to determine SNPs and construct phylogenetic tree from WGS data.

Results
Here we introduce snpTree, a server for online-automatic SNPs analysis. This tool is composed of different SNPs analysis suites, perl and python scripts. snpTree can identify SNPs and construct phylogenetic trees from WGS as well as from assembled genomes or contigs. WGS data in fastq format are aligned to reference genomes by BWA while contigs in fasta format are processed by Nucmer. SNPs are concatenated based on position on reference genome and a tree is constructed from concatenated SNPs using FastTree and a perl script. The online server was implemented by HTML, Java and python script.

The server was evaluated using four published bacterial WGS data sets (V. cholerae, S. aureus CC398, S. Typhimurium and M. tuberculosis). The evaluation results for the first three cases was consistent and concordant for both raw reads and assembled genomes. In the latter case the original publication involved extensive filtering of SNPs, which could not be repeated using snpTree.

Conclusions
The snpTree server is an easy to use option for rapid standardised and automatic SNP analysis in epidemiological studies also for users with limited bioinformatic experience. The web server is freely accessible at http://www.cbs.dtu.dk/services/snpTree-1.0/.

General information
State: Published
Organisations: National Food Institute, Division of Epidemiology and Microbial Genomics, Center for Systems Microbiology, Department of Systems Biology, Center for Biological Sequence Analysis, Division of Microbiology and Risk Assessment
Authors: Leekitcharoenphon, P. (Intern), Kaas, R. S. (Intern), Thomsen, M. C. F. (Intern), Rundsten, C. F. (Intern), Rasmussen, S. (Intern), Aarestrup, F. M. (Intern)
Number of pages: 8
Pages: S6
Publication date: 2012
Main Research Area: Technical/natural sciences

Publication information
Journal: B M C Genomics
Volume: 13
Issue number: Suppl 7
ISSN (Print): 1471-2164
Ratings:
BFI (2017): BFI-level 1
Web of Science (2017): Indexed Yes
BFI (2016): BFI-level 1
Scopus rating (2016): CiteScore 4.05 SJR 2.065 SNIP 1.122
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 1
Scopus rating (2015): SJR 2.287 SNIP 1.172 CiteScore 4.3
Web of Science (2015): Indexed yes
BFI (2014): BFI-level 1
Scopus rating (2014): SJR 2.297 SNIP 1.205 CiteScore 4.18
Web of Science (2014): Indexed yes
BFI (2013): BFI-level 1
Scopus rating (2013): SJR 2.141 SNIP 1.174 CiteScore 4.39
ISI indexed (2013): ISI indexed yes
Web of Science (2013): Indexed yes
BFI (2012): BFI-level 1
Scopus rating (2012): SJR 2.181 SNIP 1.225 CiteScore 4.61