PathogenFinder - Distinguishing Friend from Foe Using Bacterial Whole Genome Sequence Data.

Although the majority of bacteria are harmless or even beneficial to their host, others are highly virulent and can cause serious diseases, and even death. Due to the constantly decreasing cost of high-throughput sequencing there are now many completely sequenced genomes available from both human pathogenic and innocuous strains. The data can be used to identify gene families that correlate with pathogenicity and to develop tools to predict the pathogenicity of newly sequenced strains, investigations that previously were mainly done by means of more expensive and time consuming experimental approaches. We describe PathogenFinder (http://cge.cbs.dtu.dk/services/PathogenFinder/), a web-server for the prediction of bacterial pathogenicity by analysing the input proteome, genome, or raw reads provided by the user. The method relies on groups of proteins, created without regard to their annotated function or known involvement in pathogenicity. The method has been built to work with all taxonomic groups of bacteria and using the entire training-set, achieved an accuracy of 88.6% on an independent test-set, by correctly classifying 398 out of 449 completely sequenced bacteria. The approach here proposed is not biased on sets of genes known to be associated with pathogenicity, thus the approach could aid the discovery of novel pathogenicity factors. Furthermore the pathogenicity prediction web-server could be used to isolate the potential pathogenic features of both known and unknown strains.

General information

State: Published
Organisations: Department of Systems Biology, Center for Biological Sequence Analysis, Immunological Bioinformatics, Integrative Systems Biology, National Food Institute, Division of Epidemiology and Microbial Genomics
Contributors: Cosentino, S., Larsen, M. V., Aarestrup, F. M., Lund, O.
Pages: e77302
Publication date: 2013
Peer-reviewed: Yes

Publication information

Journal: PloS one.
Volume: 8
Issue number: 10
ISSN (Print): 1932-6203
Ratings:
Web of Science (2019): Indexed yes
BFI (2018): BFI-level 1
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 1
Scopus rating (2017): CiteScore 3.01 SJR 1.164 SNIP 1.111
Web of Science (2017): Indexed yes
BFI (2016): BFI-level 1
Scopus rating (2016): CiteScore 3.11 SJR 1.236 SNIP 1.101
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 1
Scopus rating (2015): CiteScore 3.32 SJR 1.427 SNIP 1.136
Web of Science (2015): Indexed yes
BFI (2014): BFI-level 1
Scopus rating (2014): CiteScore 3.54 SJR 1.559 SNIP 1.148
Web of Science (2014): Indexed yes
BFI (2013): BFI-level 1
Scopus rating (2013): CiteScore 3.94 SJR 1.772 SNIP 1.153
ISI indexed (2013): ISI indexed yes
Web of Science (2013): Indexed yes
BFI (2012): BFI-level 1
Scopus rating (2012): CiteScore 4.15 SJR 1.982 SNIP 1.156
Web of Science (2012): Impact factor 3.73
ISI indexed (2012): ISI indexed yes
Web of Science (2012): Indexed yes
BFI (2011): BFI-level 1
Scopus rating (2011): CiteScore 4.58 SJR 2.425 SNIP 1.233
Web of Science (2011): Impact factor 4.092