Insights from 20 years of bacterial genome sequencing

Since the first two complete bacterial genome sequences were published in 1995, the science of bacteria has dramatically changed. Using third-generation DNA sequencing, it is possible to completely sequence a bacterial genome in a few hours and identify some types of methylation sites along the genome as well. Sequencing of bacterial genome sequences is now a standard procedure, and the information from tens of thousands of bacterial genomes has had a major impact on our views of the bacterial world. In this review, we explore a series of questions to highlight some insights that comparative genomics has produced. To date, there are genome sequences available from 50 different bacterial phyla and 11 different archaeal phyla. However, the distribution is quite skewed towards a few phyla that contain model organisms. But the breadth is continuing to improve, with projects dedicated to filling in less characterized taxonomic groups. The clustered regularly interspaced short palindromic repeats (CRISPR)-Cas system provides bacteria with immunity against viruses, which outnumber bacteria by tenfold. How fast can we go? Second-generation sequencing has produced a large number of draft genomes (close to 90 % of bacterial genomes in GenBank are currently not complete); third-generation sequencing can potentially produce a finished genome in a few hours, and at the same time provide methylation sites along the entire chromosome. The diversity of bacterial communities is extensive as is evident from the genome sequences available from 50 different bacterial phyla and 11 different archaeal phyla. Genome sequencing can help in classifying an organism, and in the case where multiple genomes of the same species are available, it is possible to calculate the pan- and core genomes; comparison of more than 2000 Escherichia coli genomes finds an E. coli core genome of about 3100 gene families and a total of about 89,000 different gene families. Why do we care about bacterial genome sequencing? There are many practical applications, such as genome-scale metabolic modeling, biosurveillance, bioforensics, and infectious disease epidemiology. In the near future, high-throughput sequencing of patient metagenomic samples could revolutionize medicine in terms of speed and accuracy of finding pathogens and knowing how to treat them.

General information
State: Published
Organisations: Department of Systems Biology, Agricultural and Environmental Proteomics, Center for Biological Sequence Analysis, Oak Ridge National Laboratory, Molecular Microbiology and Genomics Consultants
Number of pages: 21
Pages: 141-161
Publication date: 2015
Peer-reviewed: Yes

Publication information
Journal: Functional & Integrative Genomics
Volume: 15
Issue number: 2
ISSN (Print): 1438-793X
Ratings:
BFI (2018): BFI-level 1
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 1
Scopus rating (2017): CiteScore 3.52 SJR 1.41 SNIP 0.986
Web of Science (2017): Impact factor 3.889
Web of Science (2017): Indexed yes
BFI (2016): BFI-level 1
Scopus rating (2016): CiteScore 3.37 SJR 1.288 SNIP 0.907
Web of Science (2016): Impact factor 3.496
BFI (2015): BFI-level 1
Scopus rating (2015): CiteScore 2.56 SJR 1.045 SNIP 0.725
Web of Science (2015): Impact factor 2.265
Web of Science (2015): Indexed yes
BFI (2014): BFI-level 1
Scopus rating (2014): CiteScore 2.85 SJR 1.195 SNIP 0.914
Web of Science (2014): Impact factor 2.479
BFI (2013): BFI-level 1
Scopus rating (2013): CiteScore 3.08 SJR 1.187 SNIP 0.922
Web of Science (2013): Impact factor 2.691
ISI indexed (2013): ISI indexed yes
BFI (2012): BFI-level 1  
Scopus rating (2012): CiteScore 3.68 SJR 1.517 SNIP 1.139  
Web of Science (2012): Impact factor 3.292  
ISI indexed (2012): ISI indexed yes  
BFI (2011): BFI-level 1  
Scopus rating (2011): CiteScore 3.18 SJR 1.387 SNIP 0.903  
Web of Science (2011): Impact factor 2.842  
ISI indexed (2011): ISI indexed yes  
BFI (2010): BFI-level 1  
Scopus rating (2010): SJR 1.552 SNIP 0.995  
Web of Science (2010): Impact factor 3.397  
BFI (2009): BFI-level 1  
Scopus rating (2009): SJR 2.103 SNIP 1.302  
Web of Science (2009): Indexed yes  
BFI (2008): BFI-level 1  
Scopus rating (2008): SJR 1.748 SNIP 0.955  
Web of Science (2008): Indexed yes  
Scopus rating (2007): SJR 2.398 SNIP 1.409  
Scopus rating (2006): SJR 2.041 SNIP 1.199  
Scopus rating (2005): SJR 2.273 SNIP 1.096  
Scopus rating (2004): SJR 2.17 SNIP 0.798  
Scopus rating (2003): SJR 1.798 SNIP 0.598  
Scopus rating (2002): SJR 1.413 SNIP 0.624  
Scopus rating (2001): SJR 0.523 SNIP 0.389  

Original language: English  
Keywords: Bacteria, Comparative genomics, Bacterial genomes, Metagenomics, Core-genome, Pan-genome, Next-generation sequencing  
Electronic versions:  
2015_433.pdf  
DOIs:  
10.1007/s10142-015-0433-4  

**Bibliographical note**  
This article is published with open access at Springerlink.com  
Source: FindIt  
Source-ID: 274336943  
Research output: Research - peer-review > Journal article – Annual report year: 2015