Background: Whole-genome sequencing (WGS) projects provide short read nucleotide sequences from nuclear and possibly organelle DNA depending on the source of origin. Mitochondrial DNA is present in animals and fungi, while plants contain DNA from both mitochondria and chloroplasts. Current techniques for separating organelle reads from nuclear reads in WGS data require full reference or partial seed sequences for assembling. Results: Norgal (de Novo ORGAneLle extractor) avoids this requirement by identifying a high frequency subset of k-mers that are predominantly of mitochondrial origin and performing a de novo assembly on a subset of reads that contains these k-mers. The method was applied to WGS data from a panda, brown algae seaweed, butterfly and filamentous fungus. We were able to extract full circular mitochondrial genomes and obtained sequence identities to the reference sequences in the range from 98.5 to 99.5%. We also assembled the chloroplasts of grape vines and cucumbers using Norgal together with seed-based de novo assemblers. Conclusion: Norgal is a pipeline that can extract and assemble full or partial mitochondrial and chloroplast genomes from WGS short reads without prior knowledge. The program is available at: https://bitbucket.org/kosaidtu/norgal.

General information
State: Published
Organisations: Department of Bio and Health Informatics, Metagenomics
Contributors: Al-Nakeeb, K. A. A., Petersen, T. N., Sicheritz-Pontén, T.
Number of pages: 7
Publication date: 2017
Peer-reviewed: Yes

Publication information
Journal: B M C Bioinformatics
Volume: 18
Issue number: 1
Article number: 510
ISSN (Print): 1471-2105
Ratings:
BFI (2018): BFI-level 1
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 1
Scopus rating (2017): CiteScore 2.49 SJR 1.479 SNIP 0.878
Web of Science (2017): Impact factor 2.213
Web of Science (2017): Indexed yes
BFI (2016): BFI-level 1
Scopus rating (2016): CiteScore 2.54 SJR 1.581 SNIP 0.974
Web of Science (2016): Impact factor 2.448
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 1
Scopus rating (2015): CiteScore 2.77 SJR 1.737 SNIP 1.079
Web of Science (2015): Impact factor 2.435
Web of Science (2015): Indexed yes
BFI (2014): BFI-level 1
Scopus rating (2014): CiteScore 2.91 SJR 1.916 SNIP 1.185
Web of Science (2014): Impact factor 2.576
Web of Science (2014): Indexed yes
BFI (2013): BFI-level 1
Scopus rating (2013): CiteScore 3.38 SJR 1.999 SNIP 1.323
Web of Science (2013): Impact factor 2.672
ISI indexed (2013): ISI indexed yes
Web of Science (2013): Indexed yes
BFI (2012): BFI-level 1
Scopus rating (2012): CiteScore 3.24 SJR 1.9 SNIP 1.145
Web of Science (2012): Impact factor 3.024
ISI indexed (2012): ISI indexed yes
Web of Science (2012): Indexed yes
BFI (2011): BFI-level 1
Scopus rating (2011): CiteScore 3.34 SJR 1.662 SNIP 1.19
Web of Science (2011): Impact factor 2.751
ISI indexed (2011): ISI indexed yes
Web of Science (2011): Indexed yes
BFI (2010): BFI-level 1
Scopus rating (2010): SJR 1.775 SNIP 1.13
Web of Science (2010): Impact factor 3.029
Web of Science (2010): Indexed yes
BFI (2009): BFI-level 1
Scopus rating (2009): SJR 1.893 SNIP 1.295
Web of Science (2009): Indexed yes
BFI (2008): BFI-level 1
Scopus rating (2008): SJR 1.951 SNIP 1.13
Web of Science (2008): Indexed yes
Scopus rating (2007): SJR 1.973 SNIP 1.12
Web of Science (2007): Indexed yes
Scopus rating (2006): SJR 1.913 SNIP 1.21
Scopus rating (2005): SJR 2.635 SNIP 1.61
Web of Science (2005): Indexed yes
Scopus rating (2004): SJR 3.304 SNIP 1.723
Web of Science (2004): Indexed yes
Scopus rating (2003): SJR 3.063 SNIP 1.229
Web of Science (2003): Indexed yes
Scopus rating (2002): SJR 24.693 SNIP 1.02
Scopus rating (2001): SJR 0.527 SNIP 0.457
Original language: English
Keywords: De novo assembly, K-mer, Mitochondrial dna, Next-generation sequencing
Electronic versions:
10.1186_2Fs12859_017_1927_y.pdf
DOIs:
10.1186/s12859-017-1927-y

Bibliographical note
Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated.
Source: FindIt
Source-ID: 2393584814
Research output: Research - peer-review › Journal article – Annual report year: 2017