Jabba: Hybrid Error Correction of Long Sequencing Reads

Giles Miclotte, Mahdi Heydari, Piet Demeester,
Pieter Audenaert, and Jan Fostier

Ghent University, Department of Information Technology - iMinds,
Internet Based Communication Networks and Services (IBCN)
Gaston Crommenlaan 8 (Bus 201), B-9050 Gent, Belgium
`www.ibcn.intec.ugent.be`

Background. Third generation sequencing techniques produce longer reads with higher error rates than second generation methods. While the improved read lengths can provide useful information for downstream analysis, the higher error rates can complicate the required mapping or alignment. Hybrid strategies have been proposed to correct the long reads using accurate short reads. Mapping short reads on long reads may eliminate up to 99% of all errors in bacterial datasets, however this requires significant amounts of computing resources. Mapping the long reads on a k-mer frequencies based de Bruijn graph is significantly more efficient, but loses some accuracy on larger genomes.

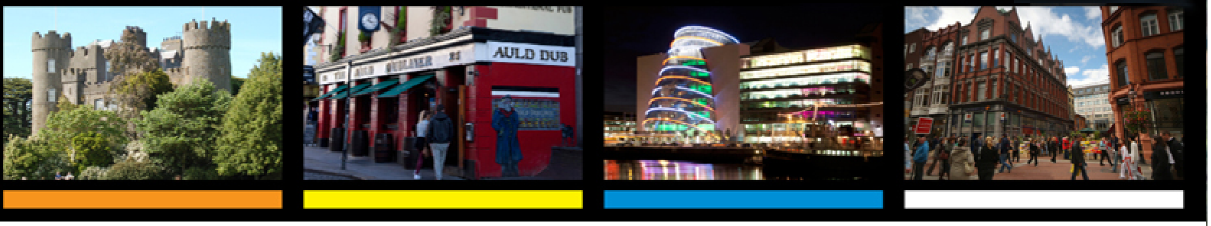
Description. We present Jabba, a hybrid method to correct long reads by mapping them on a corrected de Bruijn graph. First, accurate second generation reads are used to build a de Bruijn graph, which is then corrected based on standard topological graph correction methods. Finally a path in the graph is then found by using maximal exact matches between a long erroneous read and the nodes of the de Bruijn graph. This path then dictates the corrected sequence.

Conclusions. Jabba achieves comparable gain to other available tools for bacteria and other small genomes. For larger genomes Jabba keeps performing well, while others either can not practically handle these at all, or only at significantly reduced gain.

DUBLIN ISMB ECCB 2015 JULY 10-14

23rd Annual International Conference on Intelligent Systems for Molecular Biology
14th European Conference on Computational Biology

An Official Conference of iSCB International Society for Computational Biology



Google™ Custom Search

REGISTRATION JOIN ISCB KEY DATES NEWS

- HOME
- CAREER FAIR
- CONFERENCE ACCOMMODATION
- GENERAL INFO
- PROGRAM
- EXHIBITORS & SPONSORS
- SUBMIT
- COMMITTEES
- CONTACT
- PROMOTION
- PRESS PASS

Talks Available

ISMB/ECCB talks available* at: http://www.iscb.org/ismb-mm/media-ismbeccb2015
(*not all presenters provide permission to make talks available)

Award Winners

Congratulations to the ISMB/ECCB 2015 Award winners

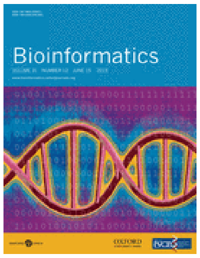
Delegate Survey

Let us know what you think, complete the delegate survey here

PDF Conference Program Also Available

Vote for you favorite Art & Science Entry

Special Issue of Bioinformatics Featuring ISMB/ECCB 2015



Exhibitors

PLOS · QIAGEN Bioinformatics · Oxford University Press · sbv IMPROVER Systems Biology Verification
BioMedBridges · The Royal Society Publishing · Cambridge University Press · elixir
EMBL-EBI · iSCB International Society for Computational Biology · F1000Research Open for Science · iSCB Student Council Symposium 2015

ISMB/ECCB 2015 appreciates the generous support of the following sponsors:

Bronze Sponsors:

QIAGEN Bioinformatics

F1000Research Open for Science

GOBLET
Global Organisation for Bioinformatics
Learning, Education & Training

ECCB 2016
The Hague
September 3-7

CRC Press
Taylor & Francis Group

BioMed Central
The Open Access Publisher

Jalview

TimeLogic®

AL
Active Learning
Publishers

Springer

amazon
web services

DNA genotek