labibi Documentation

Release 0.9.0

C. Titus Brown

October 02, 2015

Contents

1 Helpful instructions: **3**

2 Protocols: **5**

2.1 mRNAseq assembly: the Eel Pond Protocol 5

2.2 Metagenome assembly: the Kalamazoo Protocol 5

3 Additional information **7**

4 Funding **9**

5 TODO: **11**

version 0.8.5 (unreleased)

This is a set of protocols for doing genomic data analysis – specifically, de novo mRNAseq assembly and de novo metagenome assembly – in the cloud.

The latest released version of these protocols can always be found at:

http://khmer-protocols.readthedocs.org/

If you need to reference these protocols, please cite:

```
Brown, C. Titus; Sheneman, Leigh; Scott, Camille; Crusoe, Michael;
Rosenthal, Josh; Howe, Adina Chuang (2013): khmer-protocols
documentation.
figshare. http://dx.doi.org/10.6084/m9.figshare.878460
```

CHAPTER 1

Helpful instructions:

- `amazon/index` – information on Amazon instances
- `../docs/command-line` – command-line conventions for the khmer protocols

CHAPTER 2

Protocols:

2.1 mRNAseq assembly: the Eel Pond Protocol

`mrnaseq/index`

This is a lightweight protocol for assembling up to a few hundred million mRNAseq reads, annotating the resulting assembly, and doing differential expression with RSEM.

2.2 Metagenome assembly: the Kalamazoo Protocol

`metagenomics/index`

This is a protocol for assembling low- and medium-diversity metagenomes. Marine sediment and soil data sets may not be assemblable in the cloud just yet.

CHAPTER 3

Additional information

Need help? Either post comments on the bottom of each page, OR sign up for the mailing list.

Have you used these protocols in a scientific publication? We'll have citation instructions up soon.

CHAPTER 4

Funding

khmer-protocols development has largely been supported by AFRI Competitive Grant no. 2010-65205-20361 from the USDA NIFA, and Award Number R25HG006243 from the National Institutes of Health, both to C. Titus Brown. We now have continuing support from the National Human Genome Research Institute of the National Institutes of Health under Award Number R01HG007513, also to C. Titus Brown.

CTB's work on the Eel Pond mRNAseq tutorial was enabled by his 2013 summer research work at the Marine Biological Laboratory, funded by the Burr and Susie Steinbach Award and the Laura and Arthur Colwin Endowed Summer Research Fellowship Fund

CHAPTER 5

TODO:

- remove/transition stuff from the angus site.
- add sfg/stanford: http://sfg.stanford.edu/
- send to biostar-ninjas