

Metagenomic Binning

Saeed Keshani

ECE480

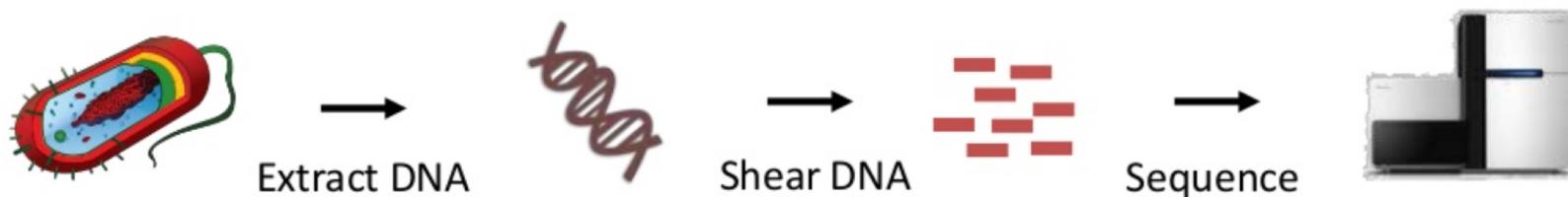
Winter 2017

Overview

- Introduction to Sequencing and Assembly
- Binning and Assembly
- Binning Methods
- Binning Steps
- MetaBAT:
 - Options
- MaxBin
 - Options
- Comparing the Results MaxBin vs MetaBAT and other Methods

Introduction

Wet lab work

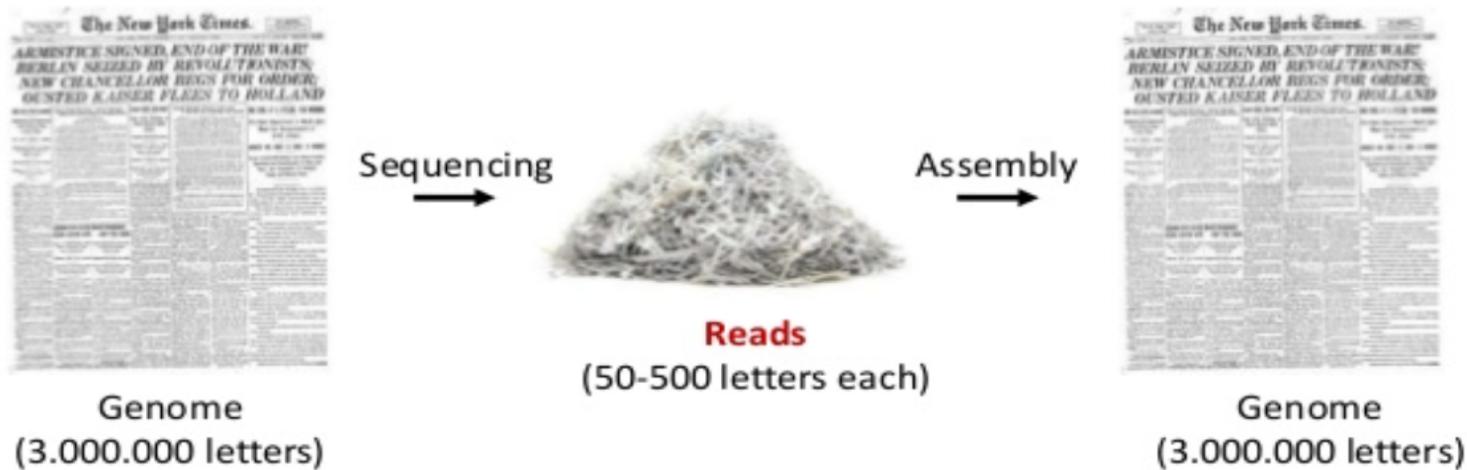


Bioinformatics



Binning and Assembly

- Metagenomic Binnig



Binning and Assembly

- Binning: Metagenomics binning is a classification of the sequence data according to the contributing species. Algorithms:
 - Similarity-based binning.
 - Compositional binning algorithms.

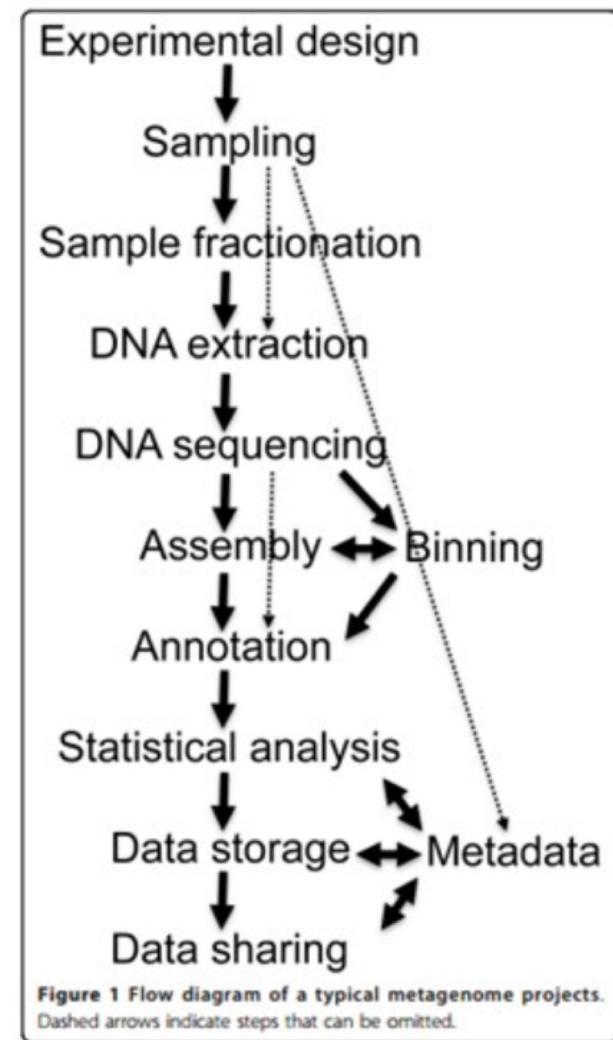
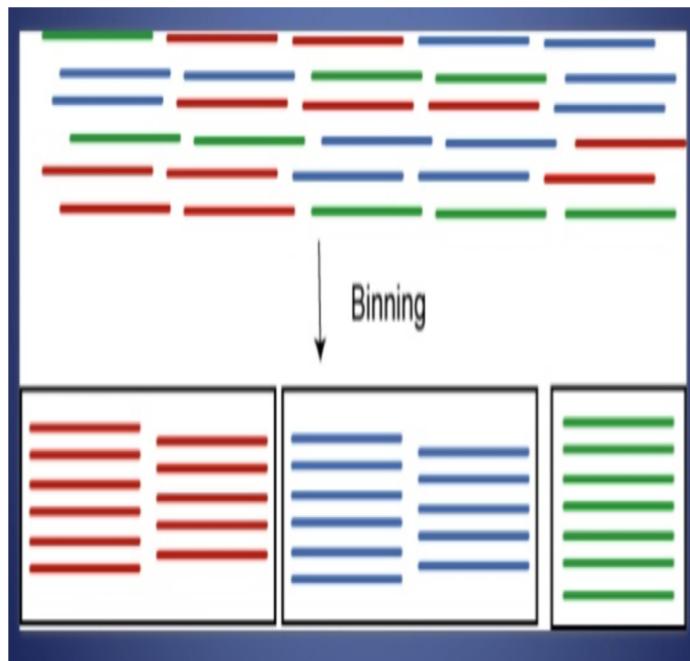


Fig source: Thomas et al. Microbial Informatics and Experimentation 2012, 2:3

Binning Methods

- Supervised binning:
 - Used known genome rely on sequence homology or sequence composition similarity
- Unsupervised binning:
 - Discriminative sequence composition or species co-abundance

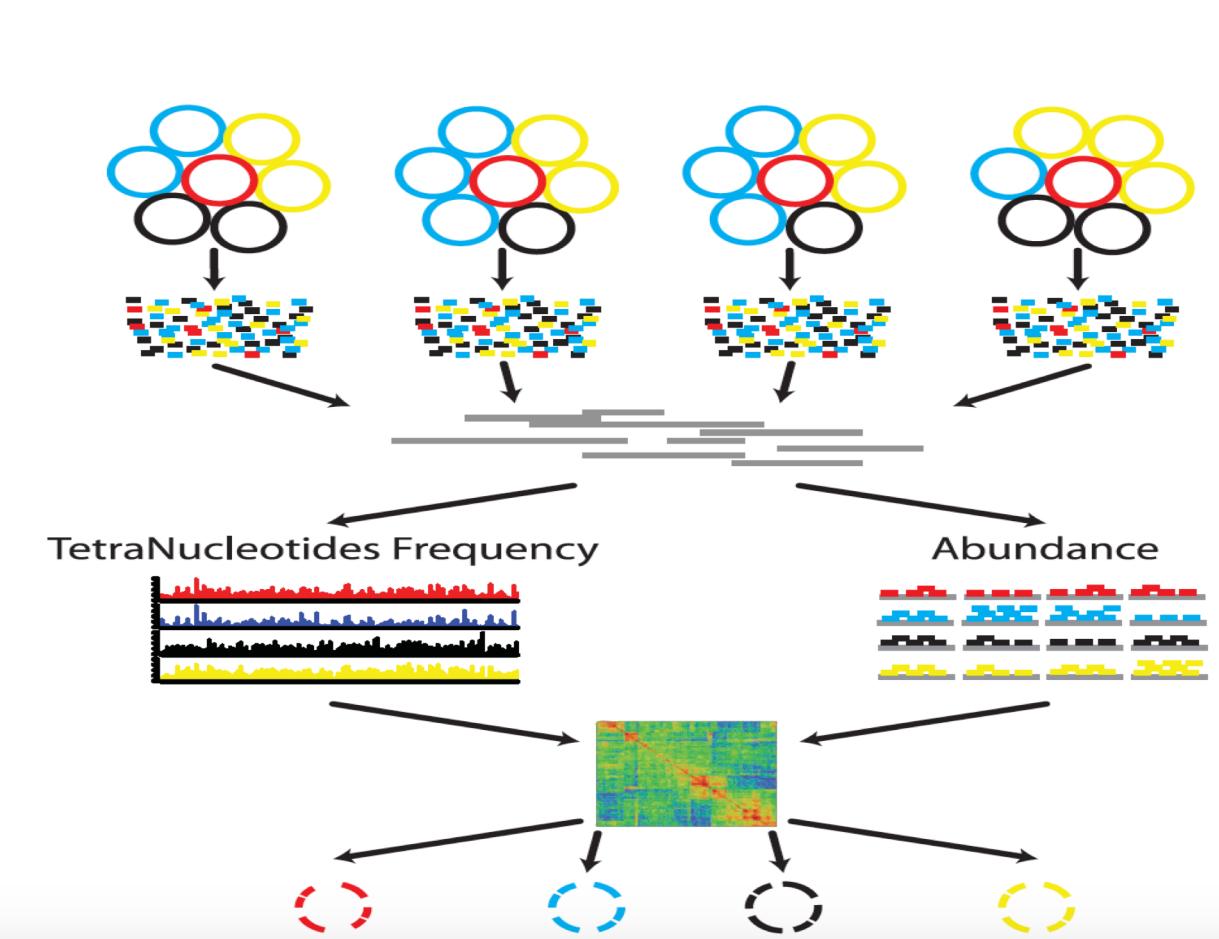
Binning Methods

- GroopM
- MaxBin
- CONCOCT
- MetaBAT
- Canopy

Prepare metadata

- Preassembly:
 - Using tools such as velvet or IDBA
- Mapping Metadata:
 - BWA (<http://bio-bwa.sourceforge.net/bwa.shtml>)
 - Samtool (<http://www.htslib.org/doc/samtools.html>)
- Choosing tools for binning.

MetaBat



<https://peerj.com/articles/1165/>

MetaBat

- <https://bitbucket.org/berkeleylab/metabat/overview>

MaxBin

- http://downloads.jbei.org/data/microbial_communities/MaxBin/MaxBin.html

Results Comparison

	MetaBAT	Canopy	CONCOCT	MaxBin	GroopM ^b
Number of bins identified (>200 kb)	234	223	260	168	335
Number of genomes detected (Precision > 9 & Recall > .3)	130	96	64	39	28
Wall time (16 cores; 32 hyper-threads)	00:03:36	00:02:31 ^a	82:19:53	06:49:39	12:19:12
Peak memory usage (for binning step)	3.0G	1.6G ^a	7G	5.8G	6.3G

References

- <https://peerj.com/articles/1165/>
- http://downloads.jbei.org/data/microbial_communities/MaxBin/MaxBin.html
- <https://bitbucket.org/berkeleylab/metabat/overview>
- <http://bio-bwa.sourceforge.net/bwa.shtml>
- <http://www.htslib.org/doc/samtools.html>
- <http://www.slideshare.net/MadsAlbertsen/2013-1029-albertsen-genomics-metagenomics>