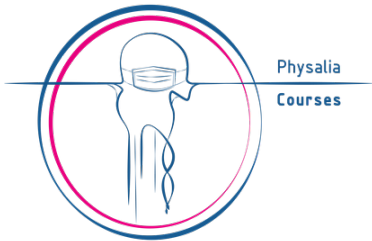
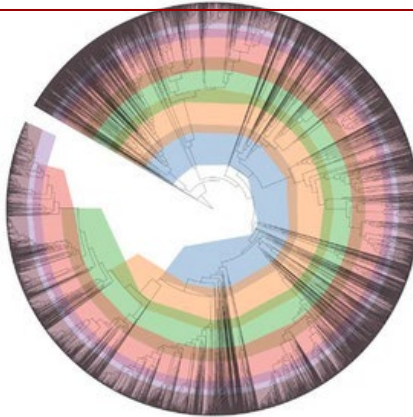


ENVIRONMENTAL METAGENOMICS

Physalia course, online, 11-15 November 2024

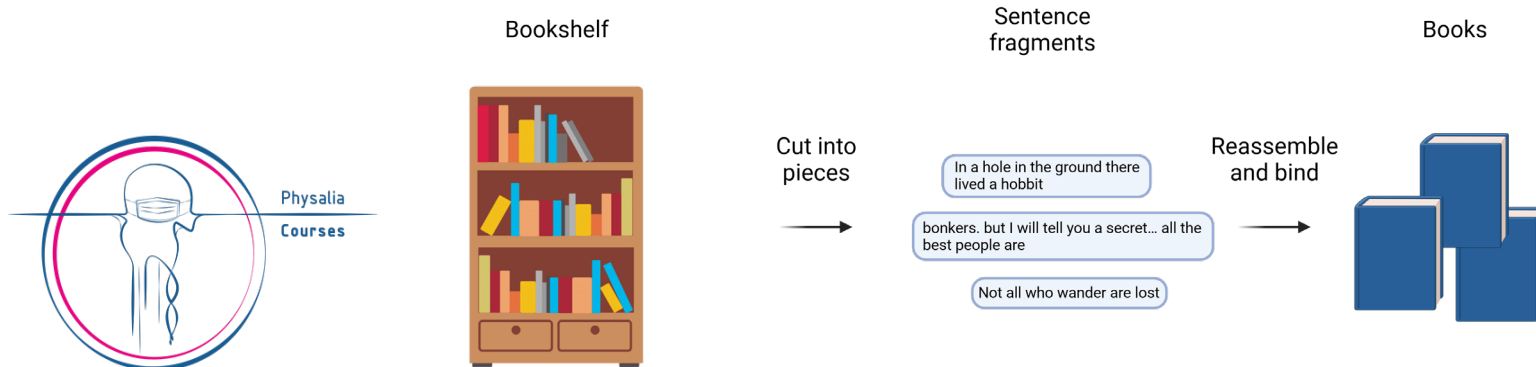
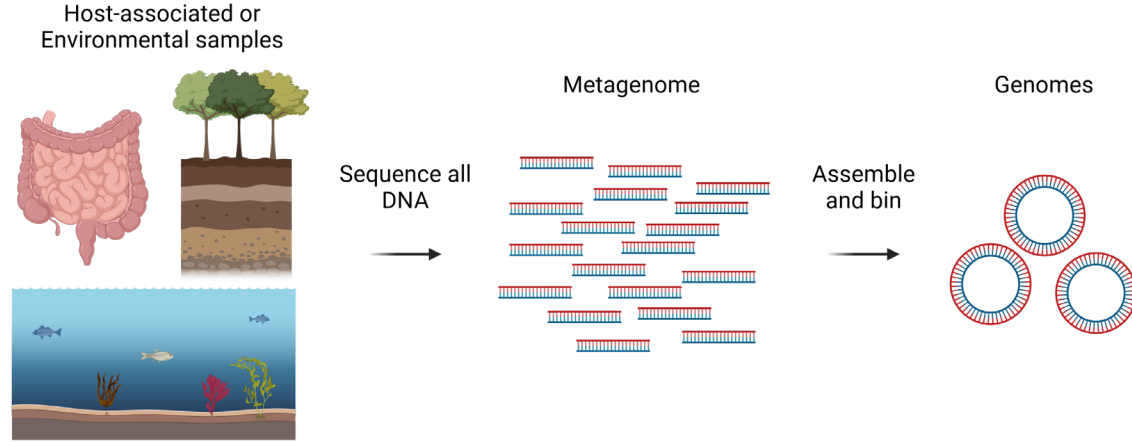
MAG binning

Nikolay Oskolkov, Lund University, NBIS SciLifeLab
Samuel Aroney, Queensland University of Technology



NB: original course material courtesy:
Dr. Antti Karkman, University of Helsinki
Dr. Igor Pessi, Finnish Environment Institute (SYKE)
As. Prof. Luis Pedro Coelho

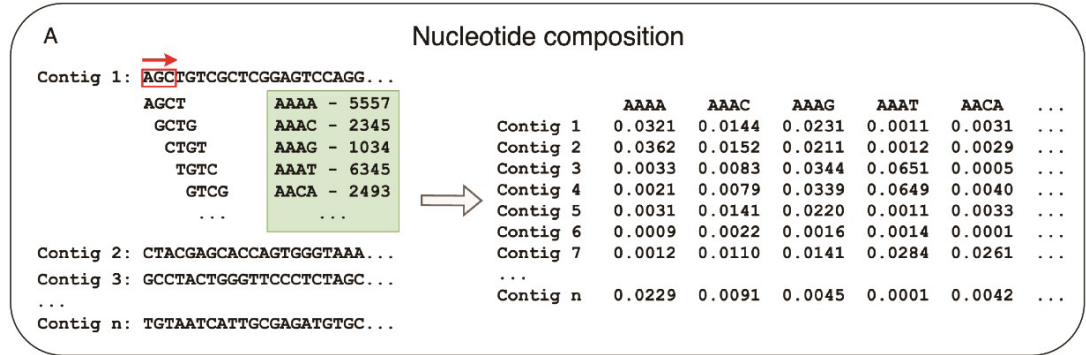
Binning: collating contigs into collections



Standard binning strategies

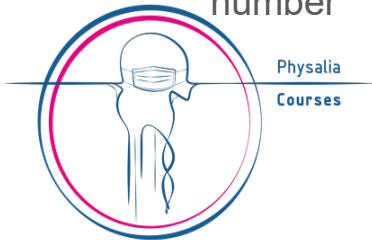
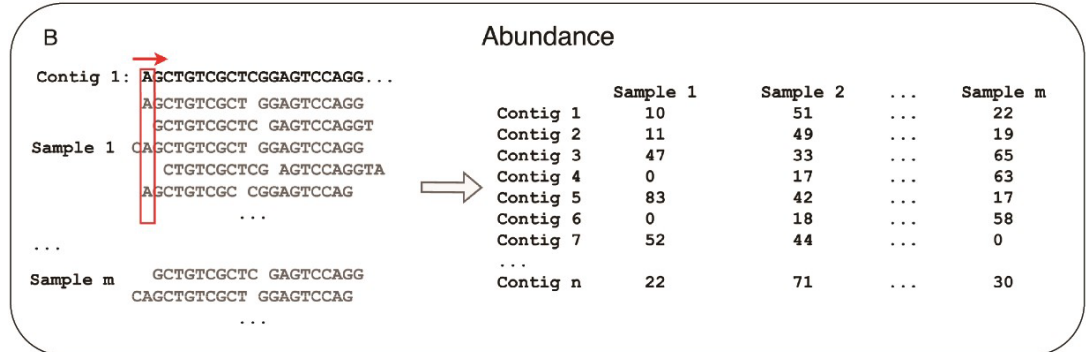
- Nucleotide composition

- Microbes tend to have similar %GC and codon choice across genome



- Abundance

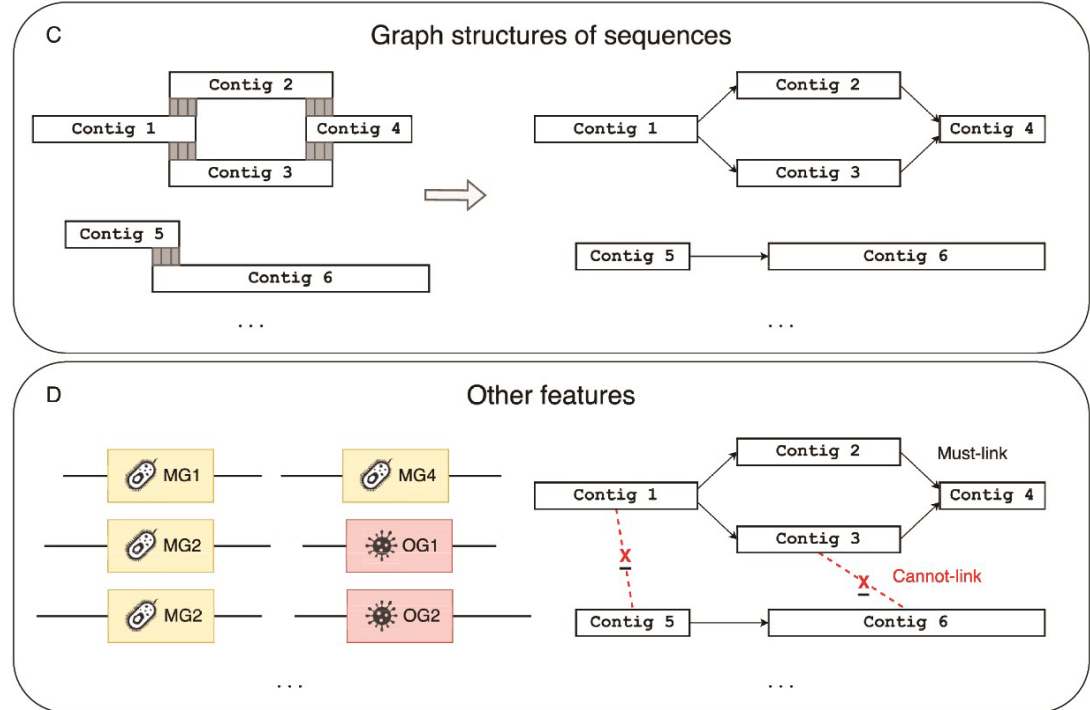
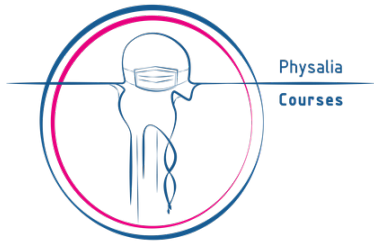
- The whole genome should be present at similar copy number



Physalia
Courses

More binning strategies

- Graph structures
 - Assembly graph structures can be reused
- Other features
 - Methylation patterns
 - Single-copy marker genes
 - Machine learning inputs



All the binners

Binning metagenomic contigs by coverage and composition

Johannes Alneberg^{1,8}, Brynjar Smári Bjarnason^{1,8}, Ino de Bruijn^{1,2}, Melanie Schirmer³, Joshua Quick^{4,5}, Umer Z Ijaz³, Leo Lahti^{6,7}, Nicholas J Loman⁴, Anders F Andersson^{1,9} & Christopher Quince^{3,9}

MetaBAT, an efficient tool for accurately reconstructing single genomes from

MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies

Dongwan D. Kang¹, Feng Li², Edward Kirton¹, Ashleigh Thomas¹, Rob Egan¹, Hong An² and Zhong Wang^{1,3,4}

Effective binning of metagenomic contigs using contrastive multi-view representation learning

Received: 28 June 2023

Ziye Wang¹, Ronghui You¹, Haitao Han¹, Wei Liu¹, Fengzhu Sun² &

Shanfeng Zhu^{1,3,4,5,6}

Accepted: 7 December 2023

A de novo metagenomic binning method for large datasets across different environments

Shaojun Pan^{1,2}, Chengkai Zhu^{1,2,3}, Xing-Ming Zhao^{1,2,4,5} & Luis Pedro Coelho^{1,2}

MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets

Yu-Wei Wu^{1,2,*}, Blake A

Improved metagenome binning and assembly using deep variational autoencoders

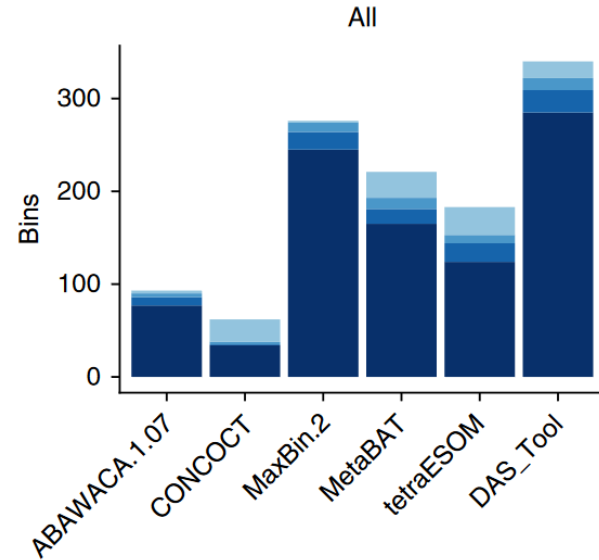
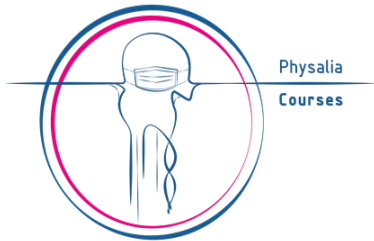
Jakob Nybo Nissen^{1,2}, Joachim Johansen², Rosa Lundbye Allesøe², Casper Kaae Sønderby³, Jose Juan Almagro Armenteros¹, Christopher Heje Grønbech^{3,4}, Lars Juhl Jensen², Henrik Bjørn Nielsen⁵, Thomas Nordahl Petersen⁵, Ole Winther^{3,4,7} and Simon Rasmussen²

SemiBin2: self-supervised contrastive learning leads to better MAGs for short- and long-read sequencing

Shaojun Pan^{1,2}, Xing-Ming Zhao^{1,2,3,4,*}, Luis Pedro Coelho^{1,2,*}

Ensemble binning

- Combining bins from multiple binner
- DAS Tool
- MetaWRAP
- MetaBinner
- BASALT



Aviary: Our Assembly + Binning pipeline

- Snakemake pipeline
- Assembly: metaSPAdes or MegaHit
 - Particularly good at hybrid short+long read
- Binning
 - MetaBAT1/2, VAMB, SemiBin1/2, CONCOCT, MaxBin2, Rosella
 - DAS Tool

