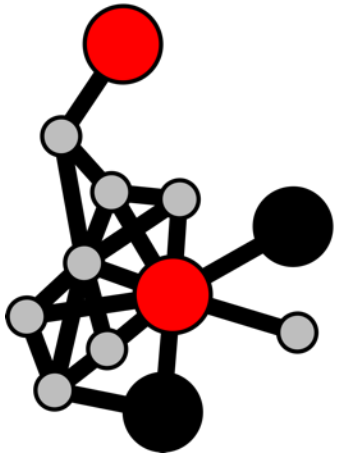
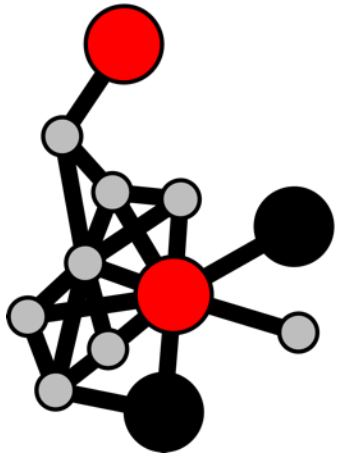
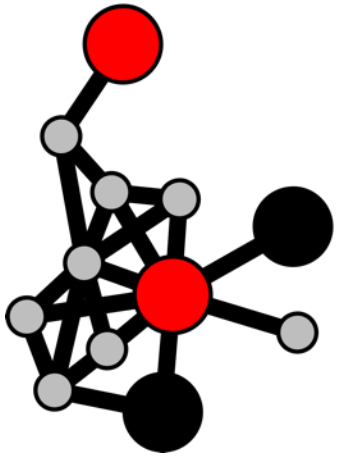
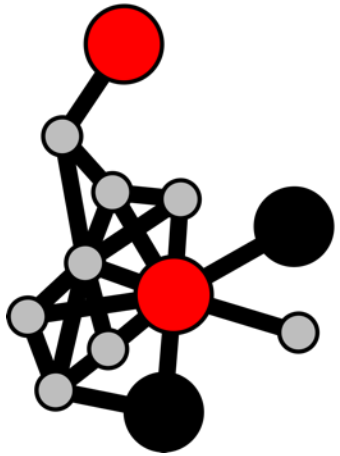




Transcriptomics lesson

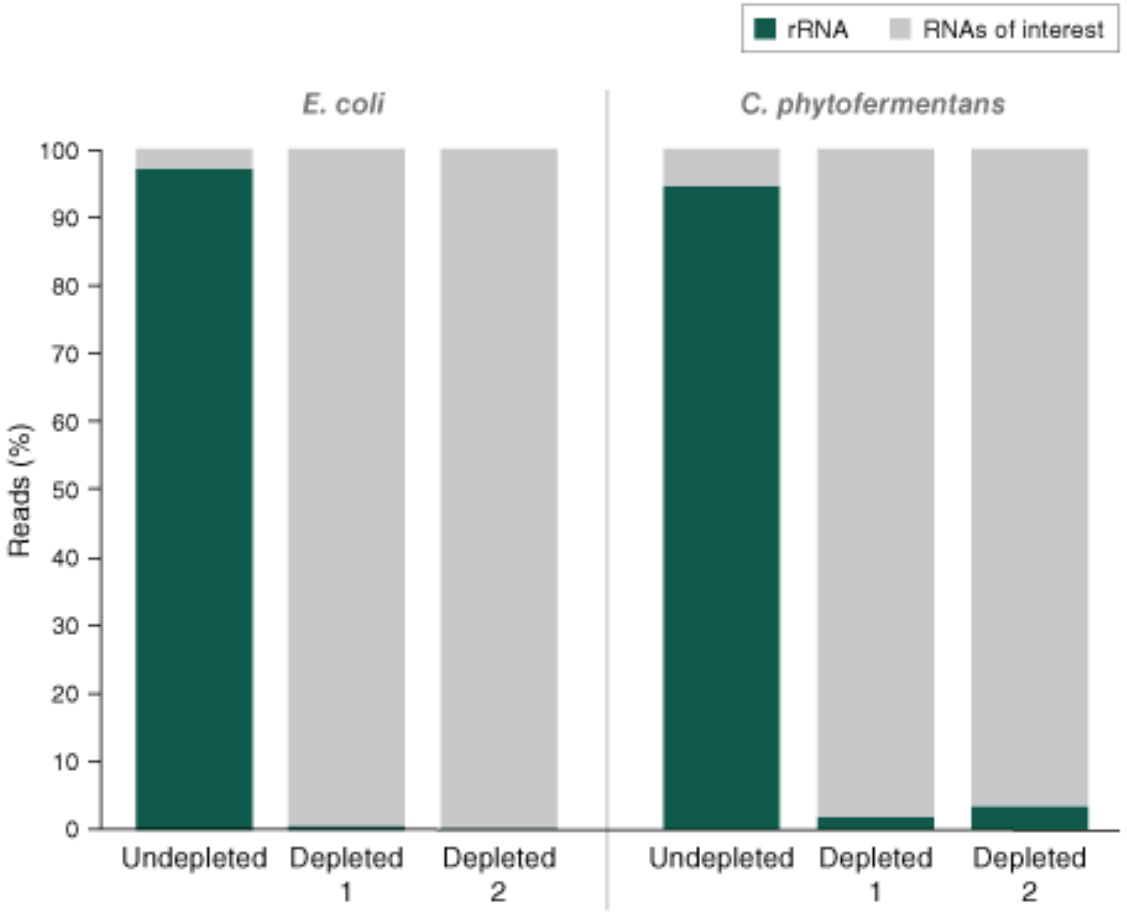




rRNA depletion:

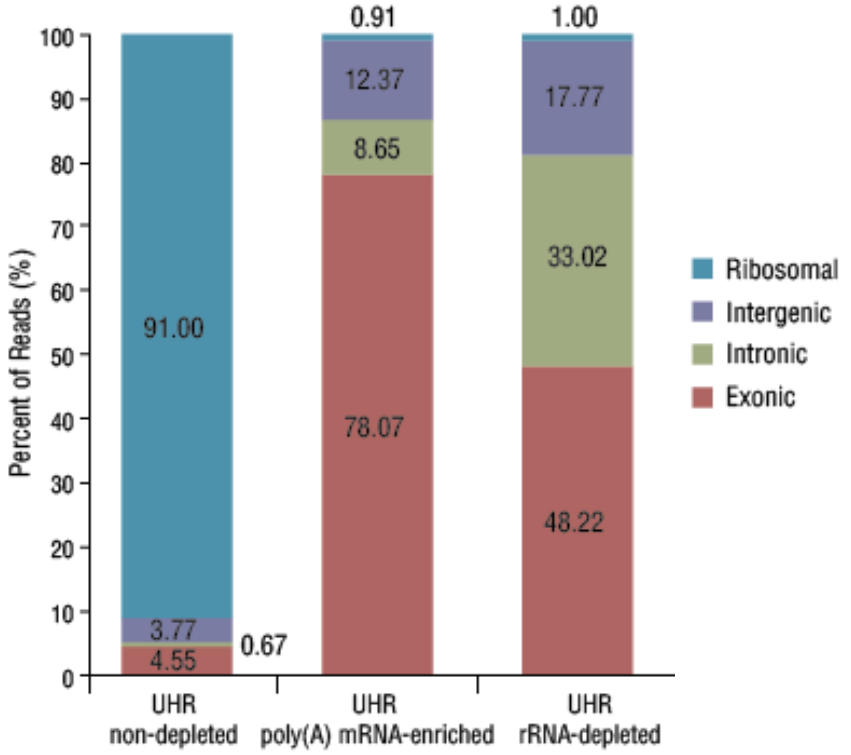
Wet-lab and in silico approaches

Why remove rRNA from you RNA samples?



New England Biolabs

Human RNA

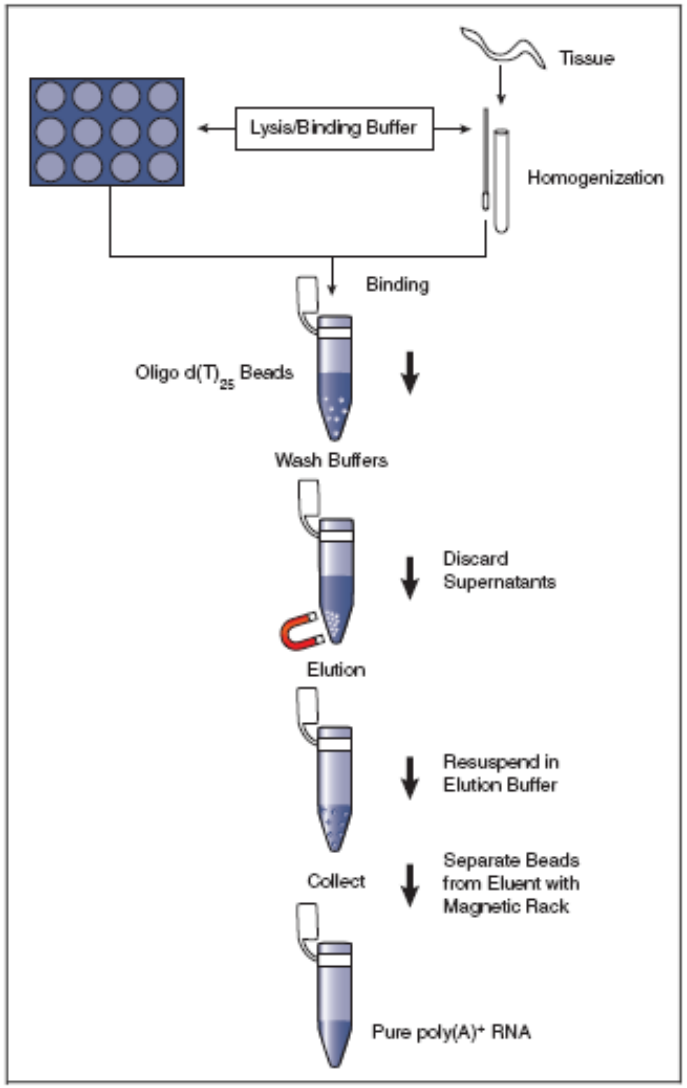


New England Biolabs

Why waste money sequencing rRNA, which makes up >90% of all RNA

NEBNext rRNA depletion using RNase H

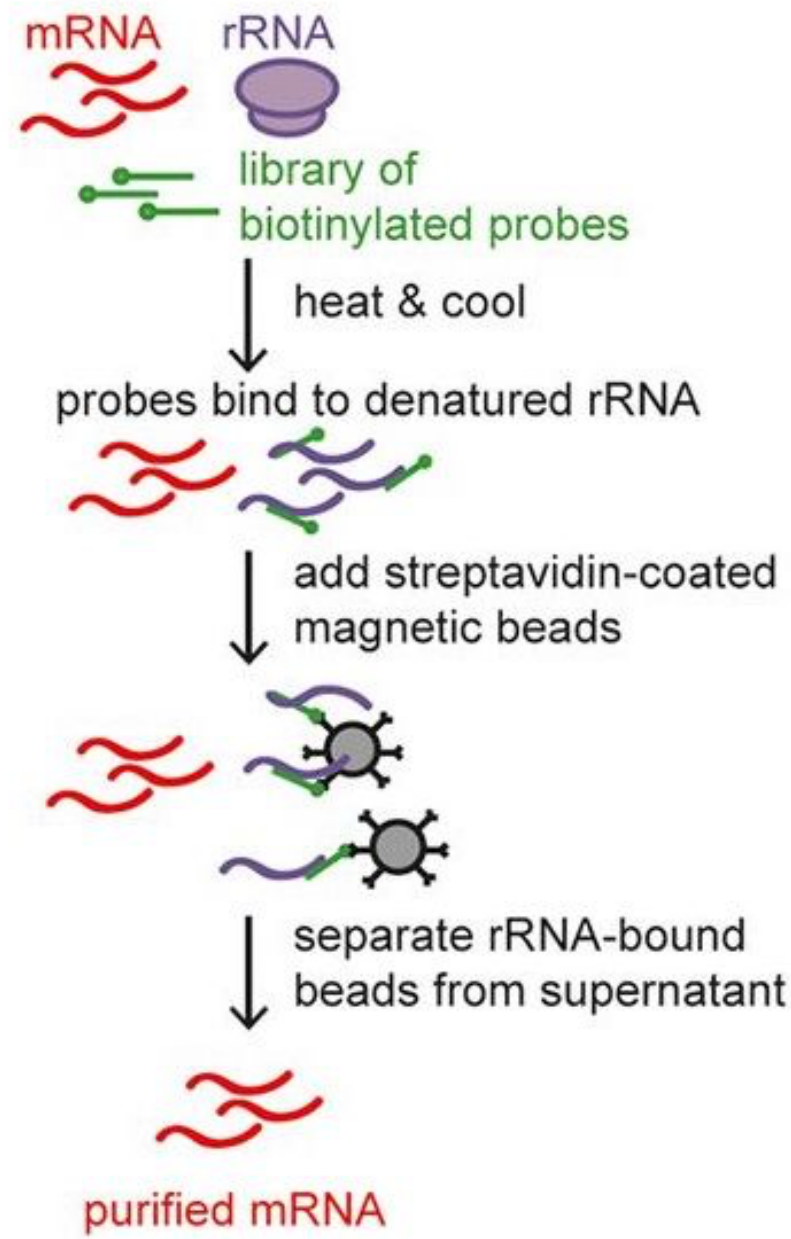
NEBNext mRNA enrichment using magnetic beads (for eukaryotic RNA only)



AMERICAN SOCIETY FOR MICROBIOLOGY | mBio®

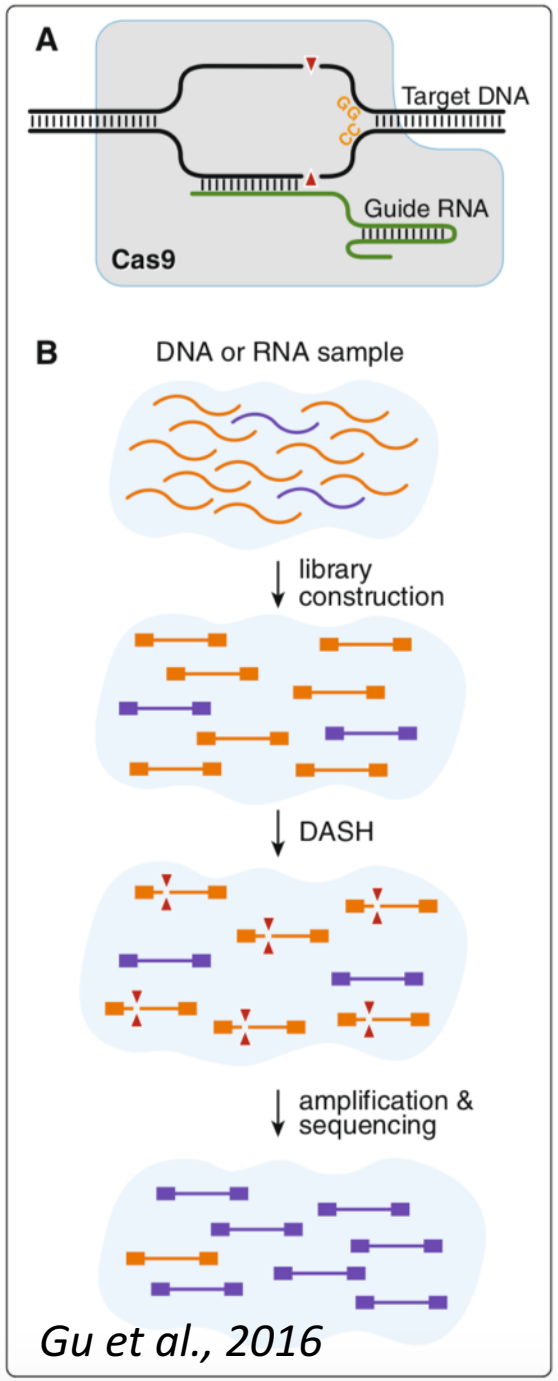
A Simple, Cost-Effective, and Robust Method for rRNA Depletion in RNA-Sequencing Studies

Peter H. Culviner,[a] Chantal K. Guegler,[a] Michael T. Laub[a,b]



Depletion of Abundant Sequences by Hybridization (DASH): using Cas9 to remove unwanted high-abundance species in sequencing libraries and molecular counting applications

W. Gu[1†], E. D. Crawford[2,3†], B. D. O'Donovan[4], M. R. Wilson[5], E. D. Chow[6], H. Retallack[7] and J. L. DeRisi[2,3*]



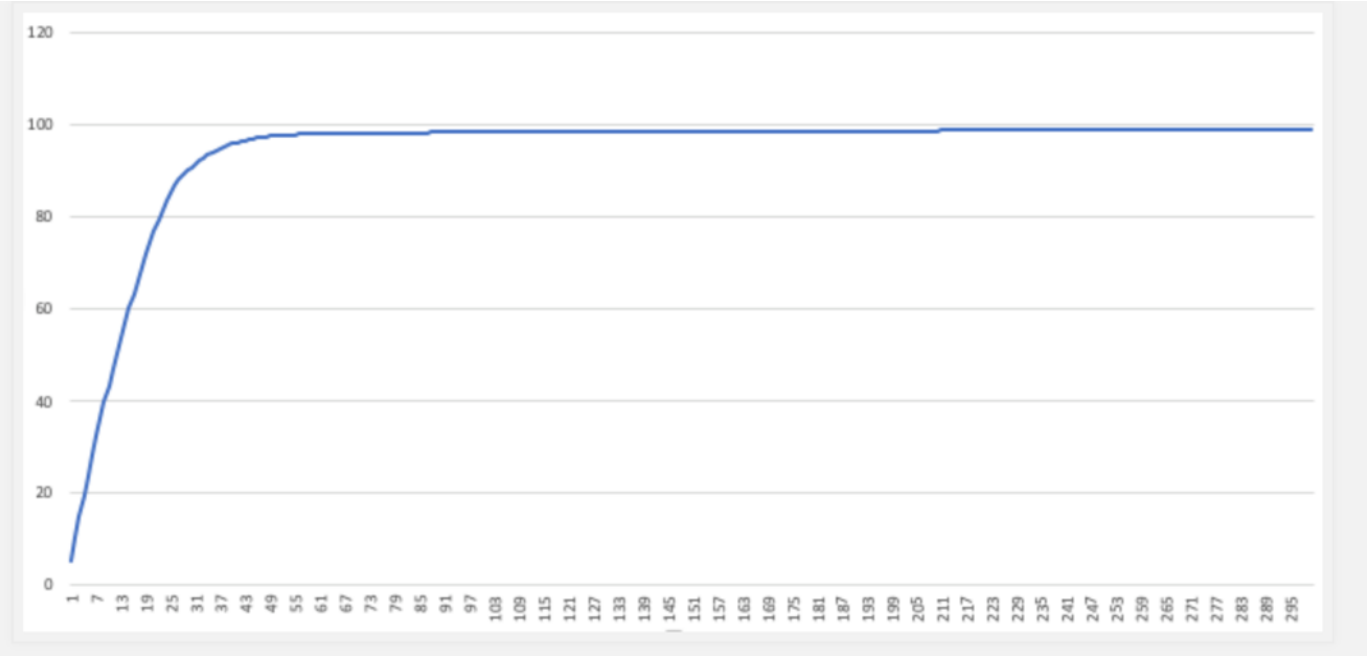
Gu et al., 2016

How does DASH work?

Shallow sequencing to identify abundant rRNA reads, and mRNA reads to avoid

"DASHit" software to identify "off-target" mRNA and "on-target" rRNA sites

Synthesize "on-target" guide RNAs to remove rRNA (or other highly abundant seqs) → **Works on cDNA!!!**



# of on-target sites

So you've completed RNA-seq, now what?

1. Your rRNA-depletion worked, but you still likely have some degree of rRNA in your reads. To avoid potential biases in mapping estimates, you should still remove the remaining rRNA reads.
2. Your rRNA-depletion did not work well for some reason, so you would like to try to salvage your experiment by removing all the rRNA reads from the data.

BIOINFORMATICS ORIGINAL PAPER

Vol. 28 no. 24 2012, pages 3211–3217
doi:10.1093/bioinformatics/bts611

Sequence analysis

Advance Access publication October 15, 2012

SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data

Evguenia Kopylova[1,2,*], Laurent Noé[1,2] and Hélène Touzet[1,2]

[1]LIFL (UMR CNRS 8022 Université Lille 1) and [2]Inria Lille Nord-Europe, 59655 Villeneuve d'Ascq, France

Associate Editor: Ivo Hofacker

Software	Illumina					Roche 454				
	rRNA	Run time	Latency	Memory (%)	Sensitivity (%)	rRNA	Run time	Latency	Memory (%)	Sensitivity (%)
SortMeRNA	998615	1 min 39 s	1×	8.5	99.861	299979	1 min 43 s	1×	6.3	99.993
riboPicker	558607	18 min 45 s	11×	6.8	55.860	123024	18 min 36 s	11×	5.6	41.008
riboPicker*	999941	6 h 33 min	238×	35.3	99.994	299999	9 h	314×	34	99.999
BLASTN	995322	23 h 52 min	868×	3.0	99.532	299978	18 h 35 min	649×	1.4	99.992
Meta-RNA	983332	2 h	72×	33.3	98.333	299980	1 h 57 min	68×	12.9	99.993
rRNASelector	974118	1 h 47 min	64×	17.4	97.411	299976	2 h	70×	7	99.992
SSU-ALIGN	971221	6 h 49 min	248×	0.1	97.122	299902	5 h 50 min	204×	0.1	99.967

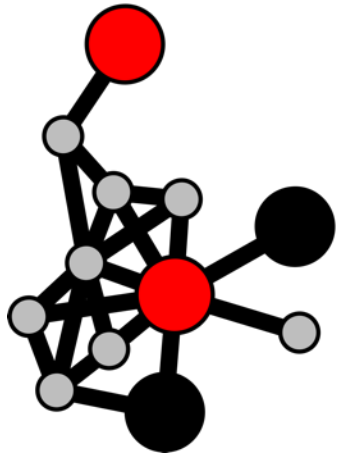
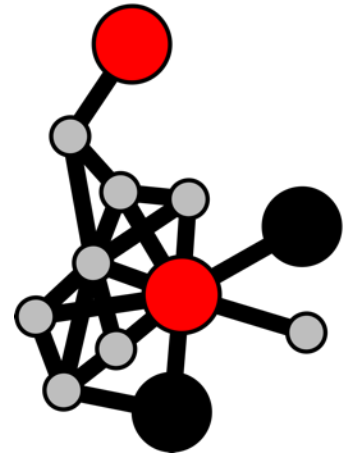
One million of MetaSim-simulated Illumina (100 *nt*) and 300 000 Roche 454 (≥200 *nt*) rRNA reads against a representative 16 S rRNA database of 7659 sequences.

- Provide single-end or paired reads
- Recommended pipeline:

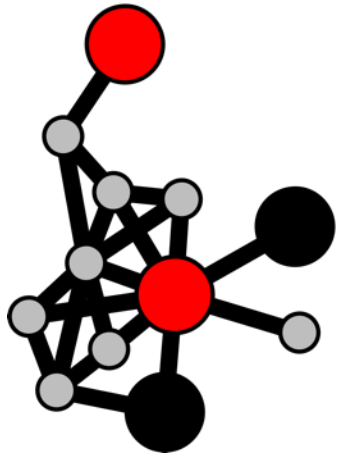
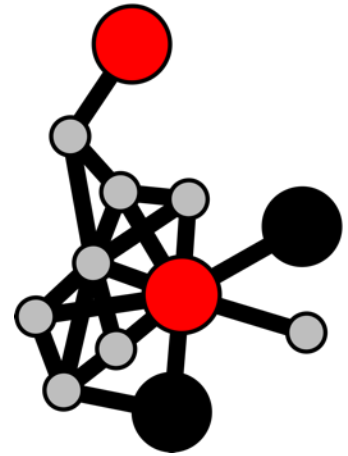
1. Quality-trimming (e.g. Trimmomatic)
2. Combine forward and reverse (e.g. Flash)
3. **rRNA-depletion SortMeRNA**

- SortMeRNA uses a set of rRNA reference files

1. Archaeal 16S
2. Archaea 23S
3. Bacterial 16S
4. Bacterial 23S
5. Eukaryotic 18S
6. Eukaryotic 28S
7. 5S

Onto the Jupyter Binder tutorial!

Arkadiy-Garber / **bcvn-binder-sortmerna**

Unwatch 1 | Star 0 | Fork 0

Code | Issues 0 | Pull requests 0 | Actions | Projects 0 | Wiki | Security 0 | Insights | Settings

Lecture and tutorial on wet-lab and in silico approaches to rRNA depletion for RNA seq

Edit

Manage topics

20 commits | 1 branch | 0 packages | 0 releases | 1 contributor

Branch: master | New pull request | Create new file | Upload files | Find file | Clone or download

Arkadiy-Garber Update environment.yml — Latest commit a9e7680 7 days ago

rRNA_databases/sample	sample reads with index	9 days ago
reads	paired reads samples	7 days ago
.bashrc	Update .bashrc	9 days ago
README.md	Update README.md	7 days ago
environment.yml	Update environment.yml	7 days ago
postBuild	Create postBuild	9 days ago
rRNA depletion.pdf	Add files via upload	7 days ago
runtime.txt	Create runtime.txt	9 days ago

README.md

Binder for BVCN sortmerna lesson

Initially forked from here. Thank you to the awesome binder team!

Part of the Bioinformatics Virtual Coordination Network :)

Walkthrough