

## Legend

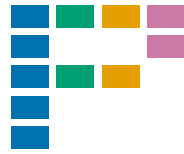
- Yes
- Partially
- No

SEISMIC-RNA  
DanceMapper  
DRACO  
DREEM

## Explanation for "Partially"

### Platform and Installation

- Runs on Linux
- Runs on macOS (both Intel and ARM)
- Install with Git or from GitHub
- Install with Conda or Mamba
- Install dependencies automatically



### Inputs

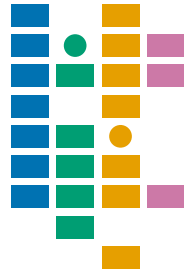
- Accepts raw reads in FASTQ format
- Accepts raw reads in gzipped FASTQ format
- Accepts aligned reads in BAM format
- Handles paired-end reads
- Handles spliced reads (i.e. introns removed)



Only if mate pairs are merged first  
Must input reads in BAM format

### Data Selection and Filtering

- Separate plus- and minus-strands
- Trim adapters and low-quality base calls
- Exclude low-quality base calls from counts
- Exclude specific types of mutations
- Exclude positions highly mutated in untreated
- Exclude arbitrary positions
- Filter out low-quality positions and reads
- Filter out reads that align too far from primers
- Downsample reads to even out coverage



Trims low quality, but not adapters

Must create a mask file manually

### Data Merging

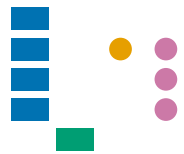
- Pool replicates *without* copying raw data
- Join regions *without* counting overlaps twice
- Join regions *after* clustering them individually



Pools counts, but not reads

### Correction of Artifacts and Biases

- Correct bias from local alignment
- Handle ambiguous insertions/deletions
- Correct bias from mutations too close
- Calculate read bias (jackpotting)
- Correct errors in reference sequences



Misses complex ambiguities  
Handles only PCR amplicons  
Graphs, but does not quantify

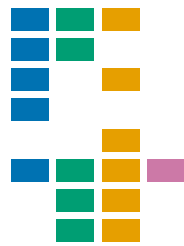
### Clustering

- Cluster short RNAs or short regions of RNAs
- Cluster long RNAs in their entirety



### Data Analysis and Normalization

- Calculate correlations between samples
- Calculate correlations between positions
- Calculate ROC curves and AUC-ROC
- Calculate Gini index
- Calculate Shannon entropy
- Normalize reactivities using a percentile
- Normalize reactivities using untreated sample
- Normalize reactivities per type of nucleotide



### RNA Structure Modeling

- Predict RNA structures using reactivities
- Predict pseudoknots
- Optimize SHAPE slope/intercept parameters



### Outputs

- Output results as tabular data (e.g. CSV files)
- Output results as graphs
- Draw RNA structures
- Write reports of parameters and results
- Write log files



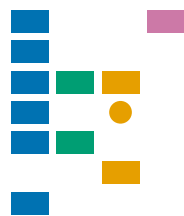
### Simulation

- Simulate raw FASTQ files
- Simulate preprocessed data



### Performance and User Friendliness

- Run entire workflow with one command
- Process entire directories with one command
- Preprocess samples/references in parallel
- Cluster samples/references in parallel
- Build index for each reference automatically
- Download prebuilt index for each reference
- Includes Python interface (importable module)



Parallelizes only references