A

Data

MicrobeCensus

Shotmap

AGS estimates

Per-gene-family counts

Average family lengths

B

Per-gene-family counts

Negative binomial fit

NB parameters

Over-dispersion (mode)

Mean counts

Simulated counts

Resampling

Optional

Normalization for AGS and AFL

Fit linear model

Residual log RPKG

Resampled residual log RPKG

Resampled residual variances

Simulated residual variances

Center and scale resampled residual variances using negative-binomial simulation

Obtain per-gene-family null distribution

p-values

Receive per-gene-family null distribution

Optional

Obtain per-gene-family null distribution