FOCUS:
HOW DOES GENE EXPRESSION CHANGE AT DIFFERENT RpoS LEVELS?

Chris Bickel, Science (2016)

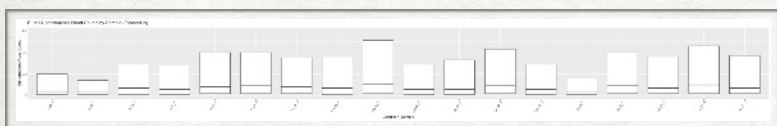The protein RpoS regulates the general stress response of the bacteria.

GOAL: FIND PATTERNS

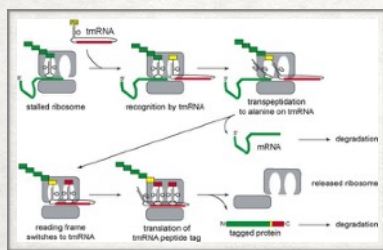What are the patterns of gene expression across all RpoS-regulated genes as RpoS level changes?

GOAL: COMPARE TO OLD DATA

Dr. Stoebel, Dr Hardin and their lab conducted a similar investigation with three levels of RpoS. Now, with six RpoS levels, we wish to compare.

PROCESS

1) Perform quality control.
2) Normalize all counts.
3) Find the regulon using differential expression analysis.
4) Design profiles and sort genes using Pearson correlation.
5) K-means clustering; Spearman correlation as dissimilarity metric.
6) K-medoids (PAM) clustering; Spearman correlation as dissimilarity metric.
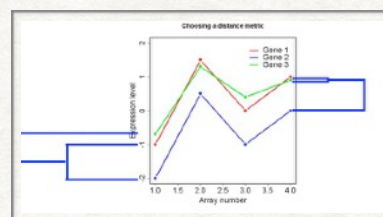
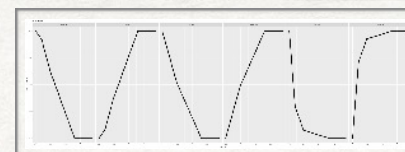Dr. Elena Marchiori, Institute for Computing and Information Science, Radboud University
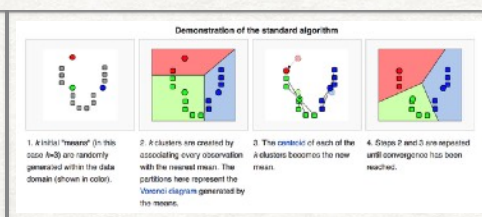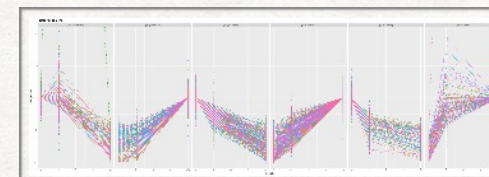

K-means Algorithm, Wikipedia

RpoS-Regulated Gene Expression Patterns in *E. coli*
DR. DAN STOEBEL, HARVEY MUDD COLLEGE
DR. JO HARDIN, POMONA COLLEGE
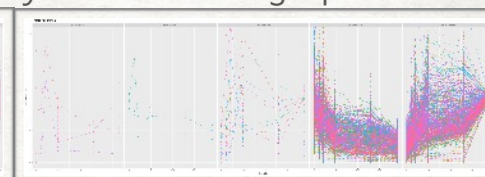### MADISON HOBBS, SCRIPPS COLLEGE

METHOD 1: DESIGN PROFILES

Design 6 profiles, modeled after the profiles presented by Dr. Hardin and Dr. Stoebel's lab, to group genes by sensitivity to increasing RpoS.
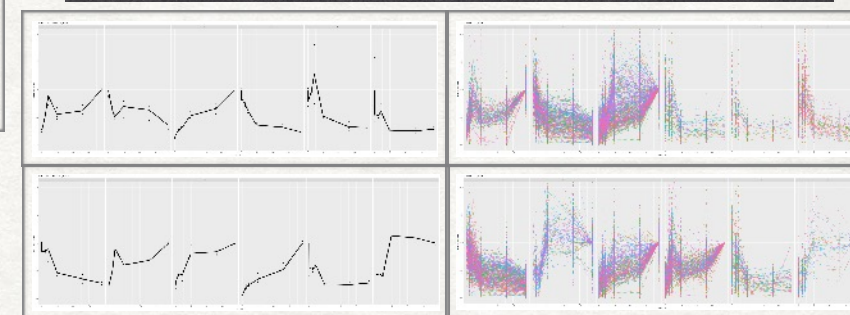

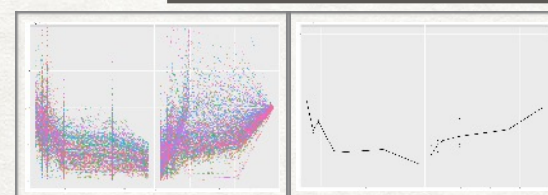old data set: profile assignment with 3 RpoS levels   new data set: profile assignment with 6 RpoS levels
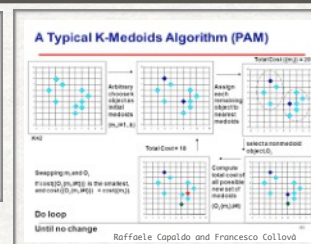
METHOD 2: K-MEANS CLUSTERING

Adjusted Rand Index for K-means #1 vs. K-means #2: 0.515
The cluster centers change each time through the algorithm, even with the same value for k.

CAVEATS AND OPEN QUESTIONS

The vast majority of genes have very low counts with a few genes in each sample dominating.

Keiler KC (2007) Curr Opin Microbiol 10:169-75

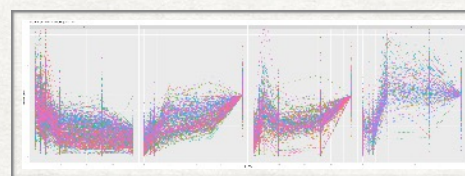ssrA (tmRNA) has very high counts, and further investigation would be interesting.

METHOD 3: PARTITIONING AROUND MEDOIDS CLUSTERING (K-MEDOIDS)

average silhouette width: 0.834

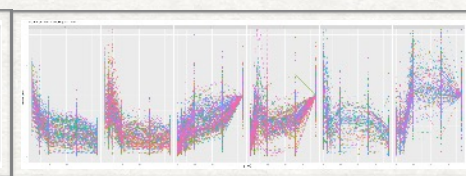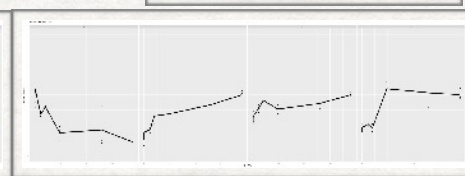A Typical K-Medoids Algorithm (PAM)
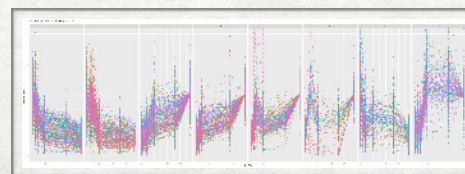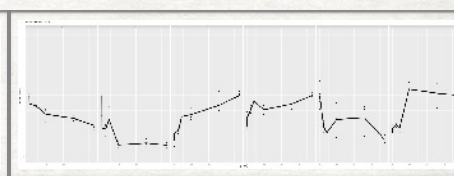Raffaele Capoldo and Francesco Collovà

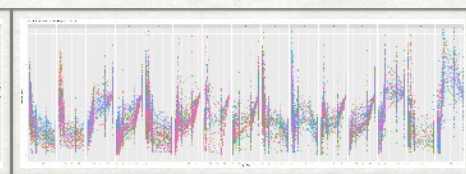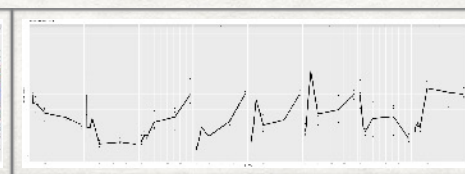average silhouette width: 0.568

average silhouette width: 0.561

average silhouette width: 0.322, ARI with profiles: 0.369

average silhouette width: 0.256

average silhouette width: 0.203

Howard Hughes Medical Institute

SCRIPPS THE WOMEN'S COLLEGE · CLAREMONT

Pomona College

HARVEY MUDD COLLEGE