Cluster Analysis Keck Local

Dr. Ramsingh, Anthony R. Colombo

April 7, 2015

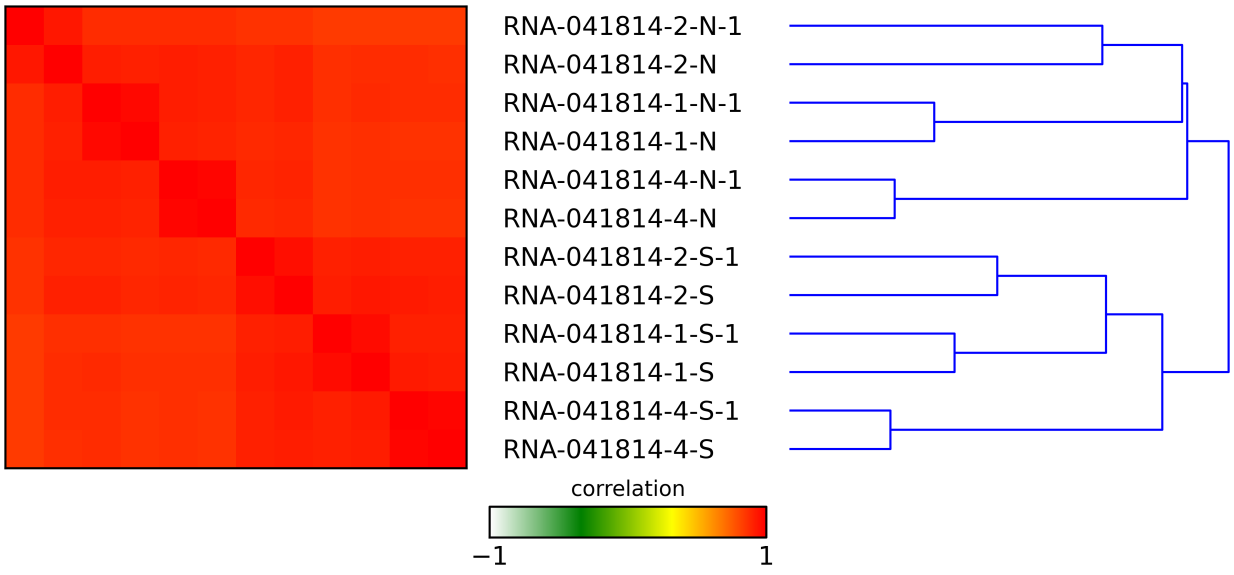
Contents

1 Introduction

This report was done using Partek for preliminary analysis; the clear separation between the PCA plot and Normal vs. Senescent intimates that the samples are not convoluted. The report does not indicate that the sample numbers suffice for good biological proof.

2 RNA Express Basespace

From running basespace native app, RNA Express shows high correlation between UCLA samples



3 Basespace TopHat2 Native App, filtered Fold change plot

Using a native application, I filtered out insignificant genes. There is a gene list in a CSV file format included with this report

4 PCA Plots

The PCA plots assign individual transcripts of gene based on the Expectation/Maximization (E/M) (Xing et.al) algorithm that is present in the reads. PGS presented gene-level data RPKM normalized counts mapped to genes in columns.

The transcript-level data shows a normalized count of sequenced reads mapped to each transcript with read counts per transcript as columns, and samples as rows.

The PCA plot of the Exons are the mapped reads over exotic regions.

5 ANOVA significant Gene List

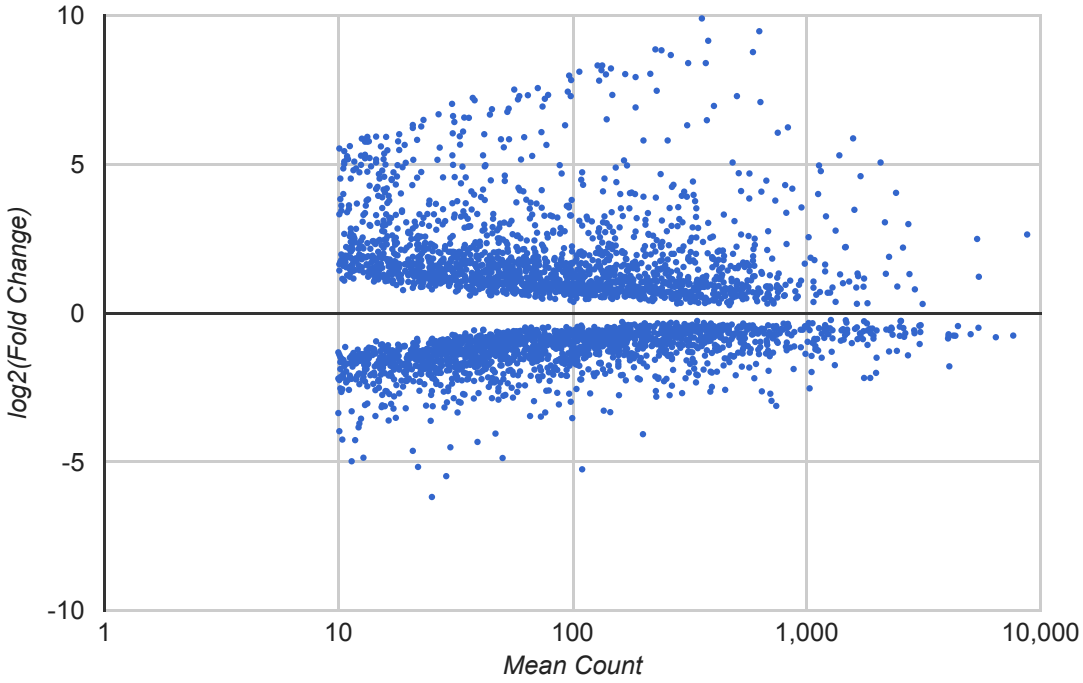
An ANOVA spreadsheet was created and available. Note - the filtered Fold Change plot from basespace is not the same gene list from Partek; completely different algorithm used. From Norris software ANNOVA, the highest ranked gene is HERPUD1 which is associated with UBQLN1, and UBQLN2 that associated with encoding proteins that increase the level amyloid-beta proteins following over expression. From the ANOVA list I did not find UBQLN1 or UBQLN2 included, and it was filtered out. However using Basespaces Log2 volcano plot I find HERPUD1, ranked much lower, and UBQLN1 , ranked nearly at the bottom.

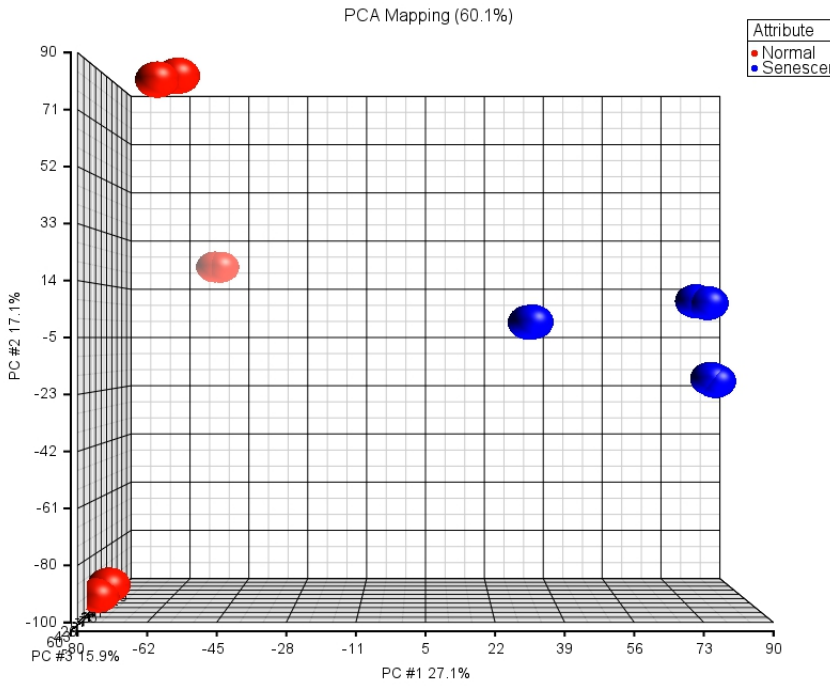
From the log2 fold change plot I have the following data

Name	Mean Count	Fold Change Value	Std. Err	Q val
HERPUD1	1700	4.59	.251	0.00
UBQLN1	374	-.437	1.81	.0438

6 Cluster Diagram

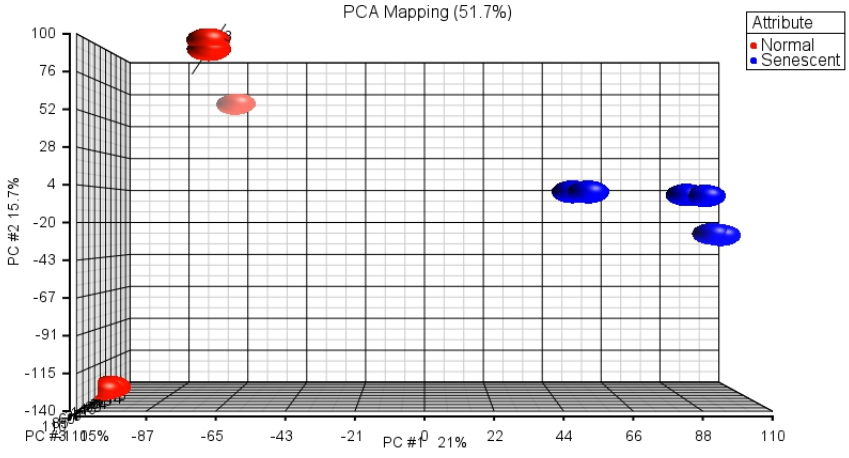
From the ANOVA gene list , we can create a cluster diagram. The dendrograms show clear separation clusters of Normal and Senescent. The images also show clear blocks of Normal down regulated versus Senescent up regulated. This is positive indication


10
5
0
-5
-10
log2(Fold Change)
1
10
100
1,000
10,000
Mean Count



Plot of Normalized Reads Over Genes.jpeg

Figure 2: Mapped Genes PCA



Normalized PCA Plot.jpeg

Figure 3: Mapped Transcripts PCA

Exons PCA.jpeg

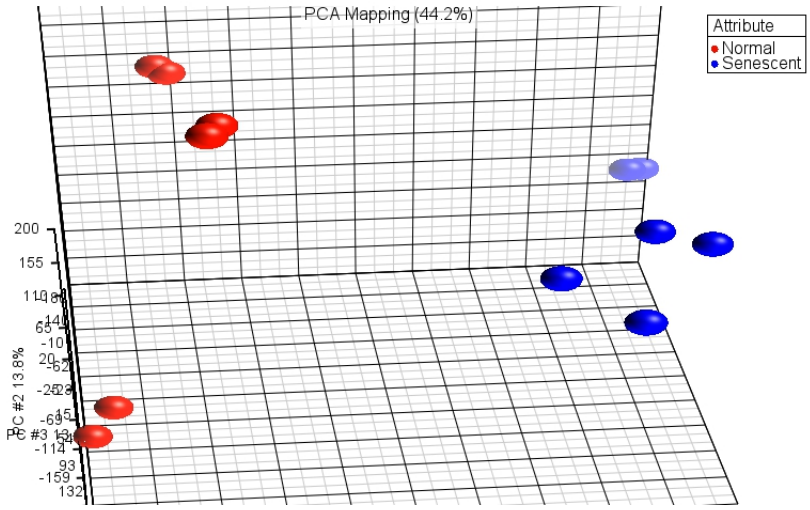
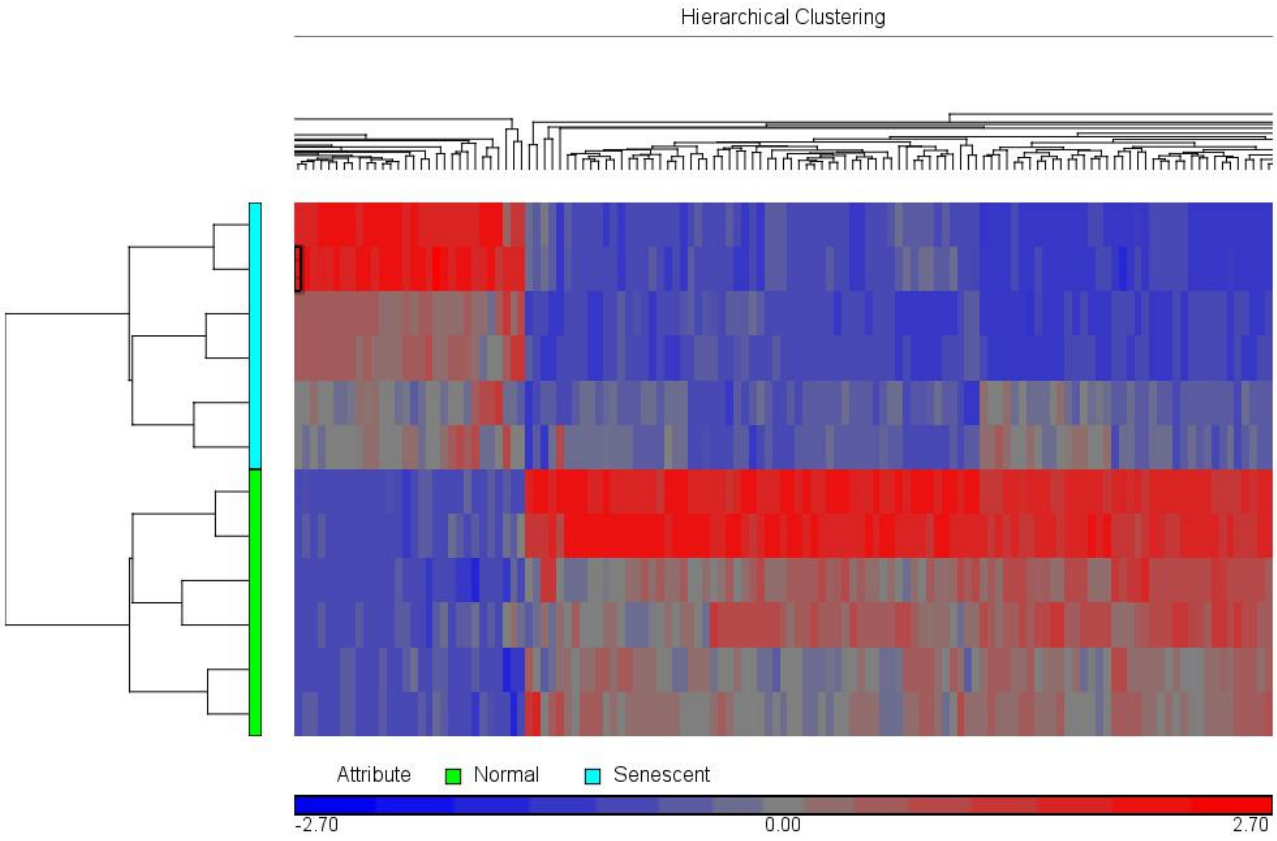
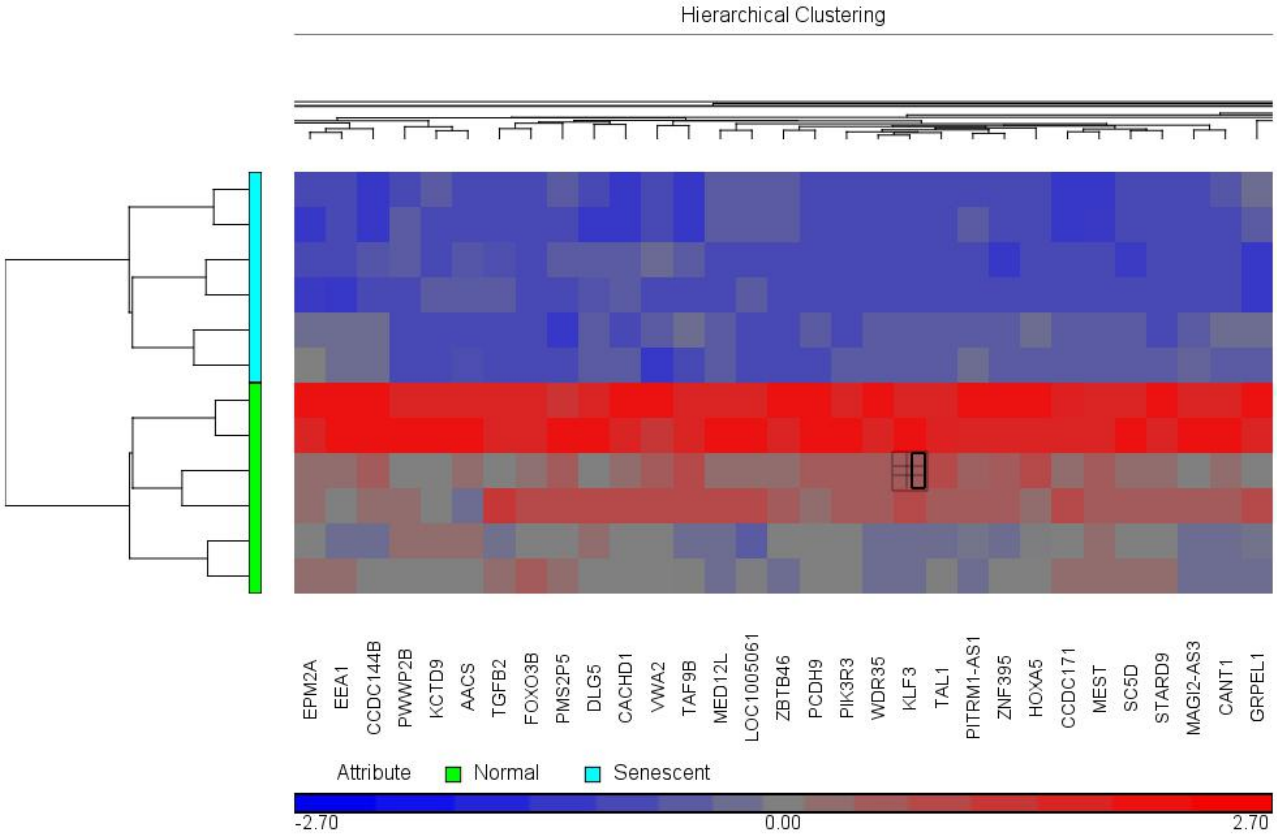
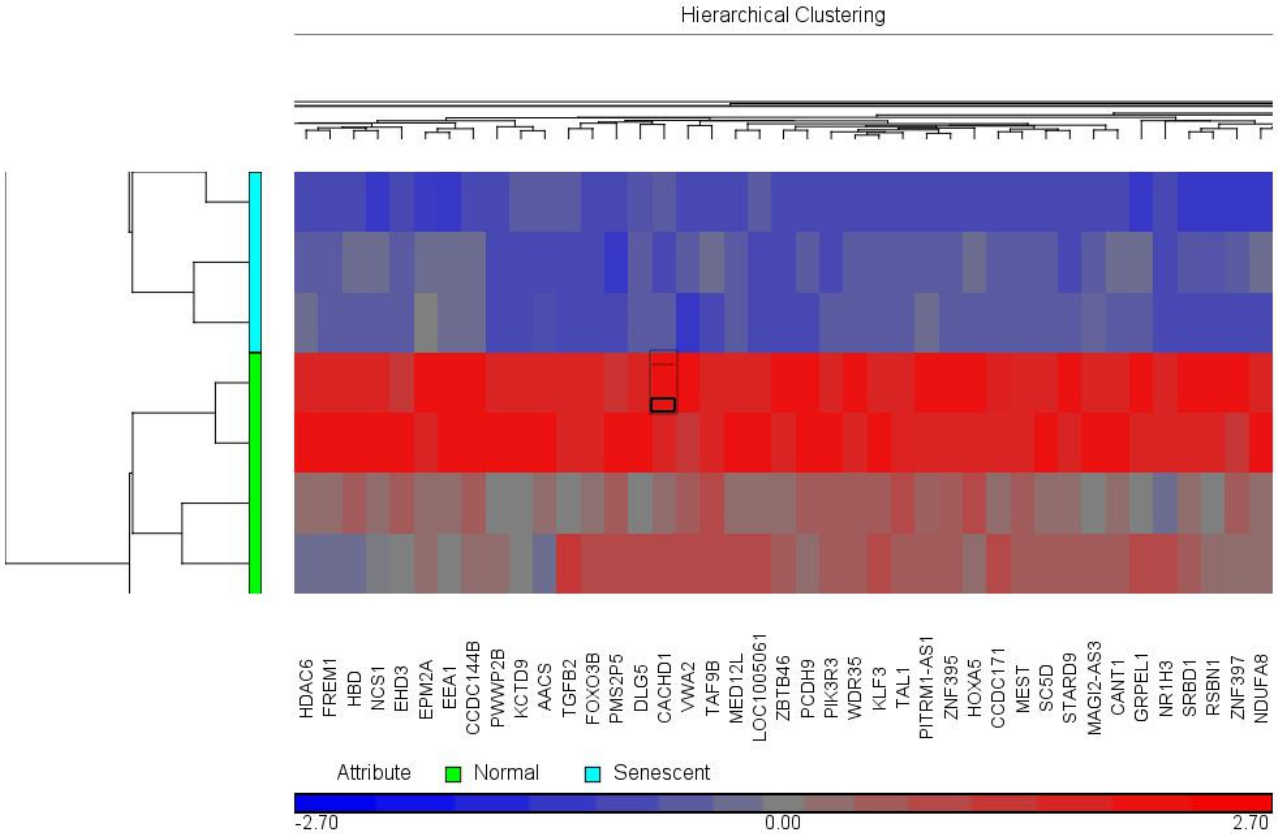


Figure 4: Mapped Exonic Regions PCA


Hierarchical Clustering
Attribute
Normal
Senescent
-2.70
0.00
2.70


Hierarchical Clustering
EPM2A
EEA1
CCDC144B
PWWP2B
KCTD9
AACS
TGFB2
FOXO3B
PMS2P5
DLG5
CACHD1
VWA2
TAF9B
MED12L
LOC1005061
ZBTB46
PCDH9
PIK3R3
WDR35
KLF3
TAL1
PITRM1-AS1
ZNF395
HOXA5
CCDC171
MEST
SC5D
STARD9
MAGI2-AS3
CANT1
GRPEL1
Attribute
Normal
Senescent
-2.70
0.00
2.70


Hierarchical Clustering
HDAC6
FREM1
HBD
NCS1
EHD3
EPM2A
EEA1
CCDC144B
PWWP2B
KCTD9
AACS
TGFB2
FOXO3B
PMS2P5
DLG5
CACHD1
VWA2
TAF9B
MED12L
LOC100506l
ZBTB46
PCDH9
PIK3R3
WDR35
KLF3
TAL1
PITRM1-AS1
ZNF395
HOXA5
CCDC171
MEST
SC5D
STARD9
MAGI2-AS3
CANT1
GRPEL1
NR1H3
SRBD1
RSBN1
ZNF397
NDUFA8
Attribute
Normal
Senescent
-2.70
0.00
2.70