Eric Almeida

Analysis of *Staphylococcus aureus* Gene Expression with Co-colonization with *Corynebacterium striatum*

Staphylococcus aureus is a fast emerging pathogen that is becoming severe in its methicillin-resistant strains known as methicillin-resistant *S. aureus* (MRSA) (CDC, 2005-2013). *S. aureus* is known to be a part of the normal human microbiome and is found in the nasal cavity in more than a quarter of the U.S. population (Gorwitz el al., 2008). Since *S. aureus* is a common member in the human microbiome, there needs to be more research to understand how this commensal organism becomes pathogenic and how to prevent the transition. This research would be very prominent in the control of pathogenicity of MRSA and other *S. aureus* stains.

S. aureus is found to conjugate with other organisms in the microbiome such as *Corynebacterium spp.* (Uehara el al., 2000). Interactions with other commensal organisms may facilitate and/or regulate the gene expression or growth of *S. aureus* with *Corynebacterium spp.* (Ramsey et al., 2016). Specifically, *Corynebacterium striatum* was used in this research to see its effects on *S. aureus'* gene expression. It has been previously shown that *Corynebacterium striatum* can down-regulate virulence such as toxins and protease production and up-regulate epithelial cell-adhesion activity (Ramsey et al., 2016). In this project, I wanted to analyze the RNA-seq data obtained in Matthew Ramsey's research on *S. aureus'* shift toward commensalism.

Methods

The RNA-seq data contained a mono- and co-culture of *S. aureus* and *C. striatum*. There were biological replicates of these conditions. I was given the raw reads from sequencing using single-end 50bp reads on an Illumina HiSeq 2500 system. I processed these reads in Trimmomatic where I selected for a minimum base pair length of 36 and increasing the quality score by trimming out adapters and low quality reads.

```
java –jar trimmomatic-0.36.jar SE –phred33 input.fq ./trimmed/output.fq
ILLUMINACLIP:TruSeq3-SE:2:30:10 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15
MINLEN:36
```

I used Bowtie2 to align the trimmed reads to a reference genome of *S. aureus*. I was only interested in the analysis of *S. aureus* gene expression and did not analyze *C. striatum*. I generated SAM files from Bowtie2 which were later sorted by SAMtools for later analysis. These SAM files were used in a python program called htseq-count. This program uses SAM files to align to a reference annotated genome. Each read that aligns to a gene on the reference annotated genome is counted. Htseq-count names reads that map over two genes as ambiguous and they are not counted. The output of this program is a list of genes and the amount of reads that aligned to each gene.

```
bowtie2 –x genome.fna –q –k 1 –U trimmed/output.fq –S output.sam


samtools sort –O sam output.sam –o output.sorted.sam


htseq –m intersection-nonempty –t CDS –i ID output.sorted.sam genome.gff >
output_HTSEQcount.txt
```

To analyze the data from htseq-count, a package within R-studio called DESeq2 was used to visualize the data. I created two heat maps, an MA plot, and a PCA plot.

```
source('http://bioconductor.org/biocLite.R')
```

```
biocLite('DESeq2')
library('DESeq2')
library('RColorBrewer')
library('gplots')
directory<-'~/Desktop/BIO594_FinalProject/RNAFinal'
HTS_Files<-grep('treated',list.files(directory),value=TRUE)
SACondition<-c('treated','treated','untreated','untreated')
SATable<-data.frame(sampleName=HTS_Files, fileName=HTS_Files,
condition=SACondition)
ddsHTSeq<-DESeqDataSetFromHTSeqCount(sampleTable=SATable,
directory=directory,    design=~condition)
colData(ddsHTSeq)$condition<-factor(colData(ddsHTSeq)$condition,
      levels=c('untreated','treated'))
dds<-DESeq(ddsHTSeq)
plotMA(dds,ylim=c(-4,4),main='Number of S. aureus Gene Counts')
      dev.copy(png,'deseq2_MAplot.png')
      dev.off()
vsd<-varianceStabilizingTransformation(dds, blind=TRUE)
select <- order(rowMeans(counts(dds,normalized=TRUE)),decreasing=TRUE)[1:30]
hmcol<- colorRampPalette(brewer.pal(9, 'GnBu'))(100)
heatmap.2(assay(vsd)[select,], col = hmcol,
          Rowv = FALSE, Colv = FALSE, scale='none',
          dendrogram='none', trace='none', margin=c(10, 6))
dev.copy(pdf,'DESeq2_heatmap3')
dev.off()
distsRL <- dist(t(assay(rld)))
mat<- as.matrix(distsRL)
rownames(mat) <- colnames(mat) <- with(colData(dds),
                                       paste(condition,HTS_Files, sep=' : '))
```

```
hc <- hclust(distsRL)
heatmap.2(mat, Rowv=as.dendrogram(hc),
          symm=TRUE, trace='none',
          col = rev(hmcol), margin=c(13, 13))
dev.copy(pdf,'deseq2_heatmaps_samplebysample.pdf')
dev.off()
rld<- rlogTransformation(dds, blind=TRUE)
print(plotPCA(rld, intgroup=c('condition')))
dev.copy(pdf,'deseq2_pca.pdf')
dev.off()
```

Results

First, the question of whether the co-culturing of *S. aureus* and *C. striatum* produces a change in gene expression must be answered and whether these changes are significant. I first wanted to observe the fold changes of all the genes and see which genes are differentially expressed. In Figure 1, the red dots depict the genes that are differentially expressed when *S. aureus* is co-cultured with *C. striatum*. As you can see, there are multiple genes that are significantly up- or down-regulated. *C. striatum* causes a change in gene expression, but these gene changes might not necessarily have anything to do with *S. aureus'* virulence and commensalism.

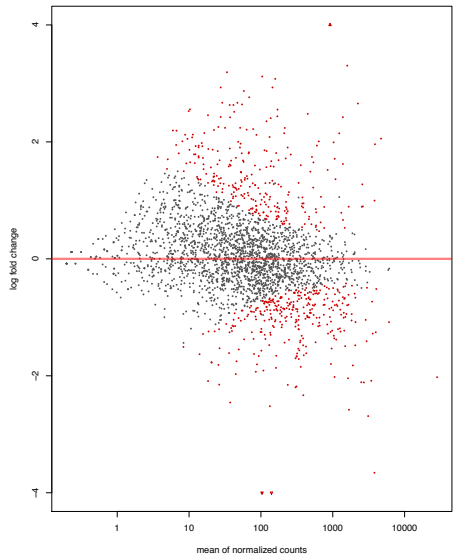

Figure 1.

Figure 2 is a heat map that I made to see how similar the mono-culture control is to the co-culture treatment. Figure 2 shows that the mono-culture controls are similar to each other. The co-cultures are similar to each other, as well. There seems to be no similarities between the mono- and co-culture. Figure 3 is showing the clustering of the 4 different samples (2 monoculture, 2 co-culture). Again, figure 3 recapitulates the idea that the untreated (control) cluster together because they are similar while the treated (co-culture) cluster together.

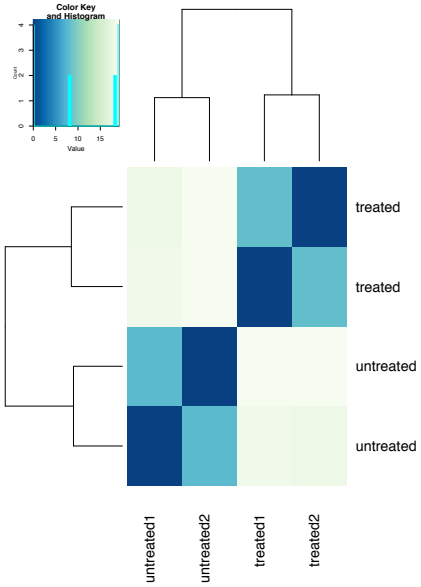

Figure 2.

Figure 4 is another heat map that focuses on some of the genes that are differentially expressed. It compares the mono-culture to the co-culture. Most of the genes in figure 4 show down-regulation when in co-culture. The mono-culture of *S. aureus* shows that the majority of the genes depicted in the heat map are more expressed when they are alone and less expressed when *C. striatum* is around. The next step for me is identifying all the genes that are up- or down-regulated and see if they have a role in virulence or commensalism.

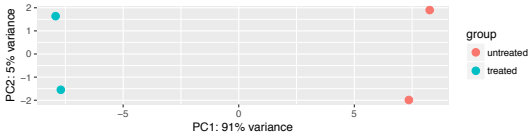

Figure 3.

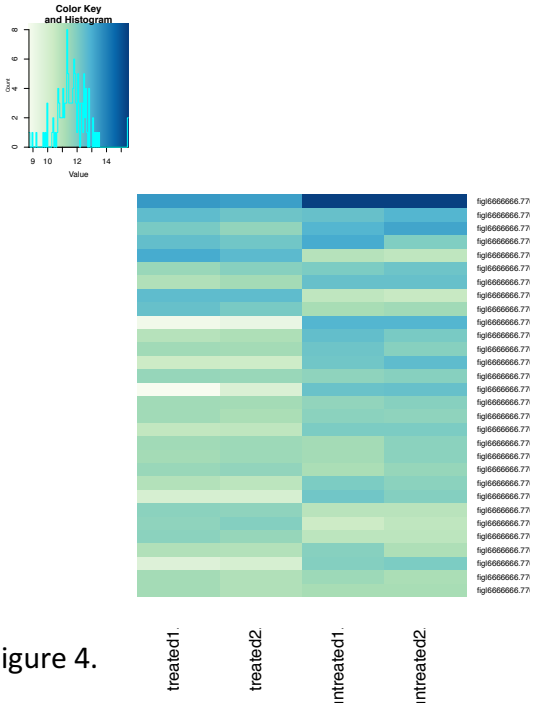

Figure 4.

Bibliography

CDC (2005–2013). Active Bacterial Core Surveillance Report, Emerging Infectious Program Network, Methicillin-Resistant Staphylococcus aureus. Atlanta, GA: Centers for Disease Control and Prevention.

Gorwitz, R. J., Kruszon-Moran, D., Mcallister, S. K., Mcquillan, G., Mcdougal, L. K., Fosheim, G. E., et al. (2008). Changes in the prevalence of nasal colonization with Staphylococcus aureus in the United States, 2001-2004. J. Infect. Dis. 197, 1226–1234. doi: 10.1086/533494

Ramsey, Matthew M., Marcelo O. Freire, Rebecca A. Gabrilska, Kendra P. Rumbaugh, and Katherine P. Lemon. "Staphylococcus Aureus Shifts toward Commensalism in Response to Corynebacterium Species." Frontiers in Microbiology 7 (2016): n. pag. Web.

Uehara, Y., Nakama, H., Agematsu, K., Uchida, M., Kawakami, Y., Abdul Fattah, A. S., et al. (2000). Bacterial interference among nasal inhabitants: eradication of Staphylococcus aureus from nasal cavities by artificial implantation of Corynebacterium sp. J. Hosp. Infect. 44, 127–133. doi: 10.1053/jhin.1999.0680