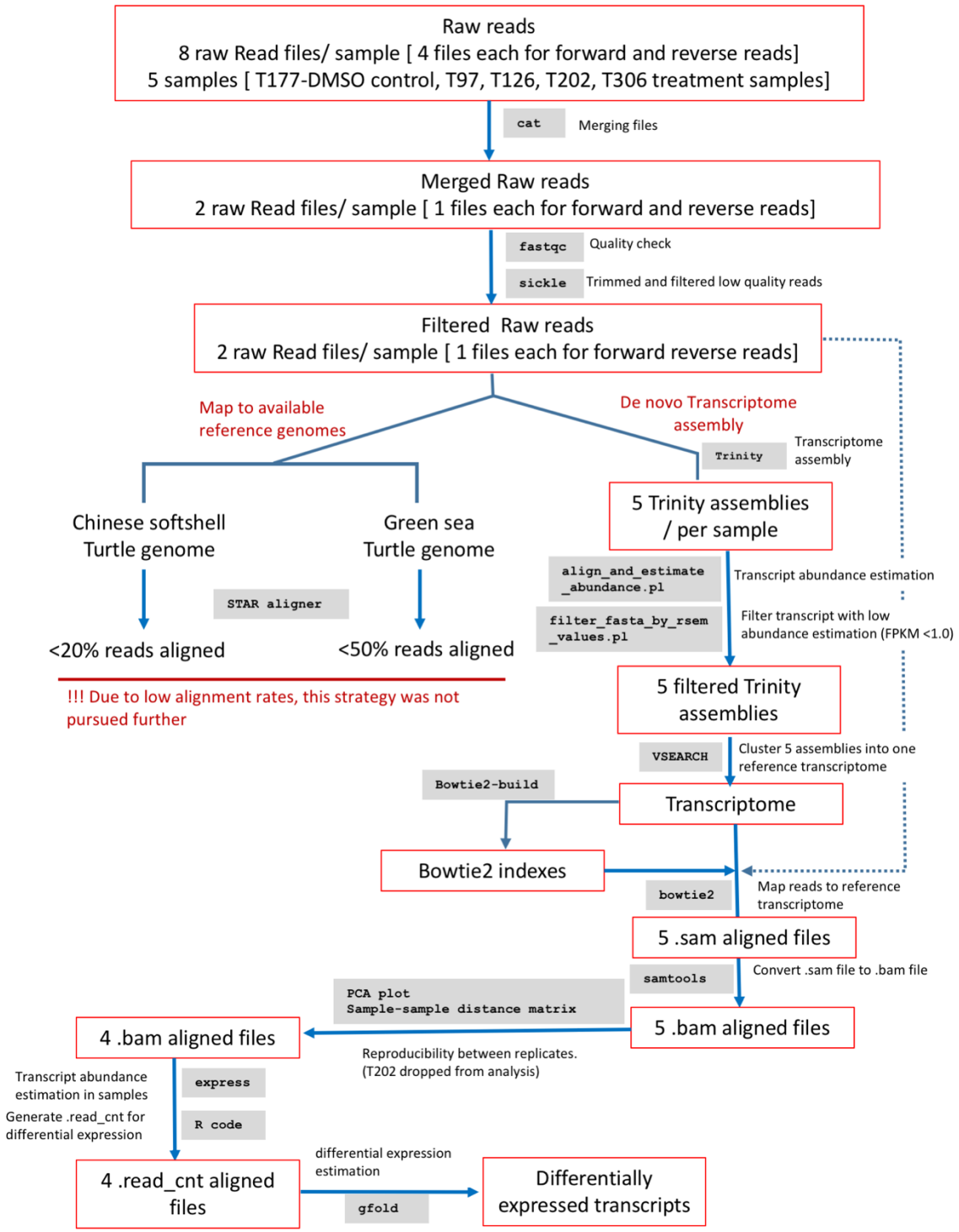
Analysis Outline:



RNA seq Analysis: Tutle yolk tissues following pesticide (Dieldrin) treatment

Neranjan Perera, Kaitlin Aldridge, Vijender Singh, Jill Wegryzn

April 25, 2017

Goal

The aim is to identify transcripts that are differentially expressed in tutle yolk tissues following pesticide (Dieldrin) treatment. In total 5 samples were collected, the control sample was treated with DMSO(T177) and the four treatment samples were treated with pesticide dieldrin (T97, T126, 202 and T306). There is variation in the pesticide crossing (or not crossing) the eggshell and getting to the yolk (so there is some unreliability in this regard). We will be interested in the largest fold-changes at target genes that are being differentially regulated in response

RAW READ Processing

Input Data

Raw Data : Format of file `.fastq`
Loc: `$WORK_DIR/original_fastq`

For each sample the raw sequence reads were split into 8 total files of which 4 corresponds to forward reads(R1) and other 4 for reverse reads(R2).

```
# T126 Sample
# T126_S2_L001_R1_001.fastq  T126_S2_L001_R2_001.fastq  T126_S2_L002_R1_001.f
astq  T126_S2_L002_R2_001.fastq
# T126_S2_L003_R1_001.fastq  T126_S2_L003_R2_001.fastq  T126_S2_L004_R1_001.f
astq  T126_S2_L004_R2_001.fastq
# T177 Sample
# T177_S1_L001_R1_001.fastq  T177_S1_L001_R2_001.fastq  T177_S1_L002_R1_001.f
astq  T177_S1_L002_R2_001.fastq
# T177_S1_L003_R1_001.fastq  T177_S1_L003_R2_001.fastq  T177_S1_L004_R1_001.f
astq  T177_S1_L004_R2_001.fastq
# T202 Sample
# T202_S5_L001_R1_001.fastq  T202_S5_L001_R2_001.fastq  T202_S5_L002_R1_001.f
astq  T202_S5_L002_R2_001.fastq
# T202_S5_L003_R1_001.fastq  T202_S5_L003_R2_001.fastq  T202_S5_L004_R1_001.f
astq  T202_S5_L004_R2_001.fastq
# T306 Sample
# T306_S3_L001_R1_001.fastq  T306_S3_L001_R2_001.fastq  T306_S3_L002_R1_001.f
astq  T306_S3_L002_R2_001.fastq
# T306_S3_L003_R1_001.fastq  T306_S3_L003_R2_001.fastq  T306_S3_L004_R1_001.f
astq  T306_S3_L004_R2_001.fastq
# T97 Sample
# T97_S4_L001_R1_001.fastq  T97_S4_L001_R2_001.fastq  T97_S4_L002_R1_001.fast
```

```
q  T97_S4_L002_R2_001.fastq
# T97_S4_L003_R1_001.fastq  T97_S4_L003_R2_001.fastq  T97_S4_L004_R1_001.fast
q  T97_S4_L004_R2_001.fastq
```

cat commands were used to put the reads from four files into a single file.

```
# cat TXXX_S5_L001_R1_001.fastq TXXX_S5_L002_R1_001.fastq TXXX_S5_L003_R1_001
.fastq TXXX_S5_L004_R1_001.fastq > $WORK_DIR/qc/TXXX_SX.R1.fastq
# cat TXXX_S5_L001_R2_001.fastq TXXX_S5_L002_R2_001.fastq TXXX_S5_L003_R2_001
.fastq TXXX_S5_L004_R2_001.fastq > $WORK_DIR/qc/TXXX_SX.R2.fastq
#
# XXX = 97, 126, 177, 202, 306
```

This leads to two raw reads files per sample, one for the forward reads and other for reverse reads. The new fastq files are located at **$WORK_DIR/qc**. The list of new files is

```
# T126_S2.R1.fastq  T126_S2.R2.fastq
# T177_S1.R1.fastq  T177_S1.R2.fastq
# T202_S5.R1.fastq  T202_S5.R2.fastq
# T306_S3.R1.fastq  T306_S3.R2.fastq
# T97_S4.R1.fastq  T97_S4.R2.fastq
```

The quality of the raw reads is assessed with fastqc tool. The results from fastqc are located at **$WORK_DIR/qc/fastqc_raw**.

```
# fastqc outputs
# multiqc_data
# multiqc_report.html
```

Firefox is ideal to visualise the **multiqc_report.html** result file. If you have X11 (refer : http://bioinformatics.uconn.edu/understanding-the-bbc-bioinformatics-facility-cluster-and-sge/)forwarding enabled then command **$ firefox multiqc_report.html** in your cluster terminal will open the file for inspection.The avearge quality scores at each base over the reads has a PHRED score >=30. However there will be reads in which the quality of bases is below acceptable level, to filter-out such reads **Sickle** was used. The general command that was used to filter the reads was

```
sickle pe -f TXXX_SX.R1.fastq -r TXXX_SX.R2.fastq -t sanger -o TXXX_SX.R1.tri
mmed.fastq -p TXXX_SX.R2.trimmed.fastq -s TXXX_SX.s.fastq -q 30 -l 40 -n
```

code explanation

command | *explanation*
sickle | *execute sickle command*
pe | *paired-end sequence trimming*
-f TXXX_SX.R1.fastq | *paired-end forward fastq file*
-r TXXX_SX.R2.fastq | *paired-end reverse fastq file*
-t sanger | *quality type*
-o TXXX_SX.R1.trimmed.fastq | *trimmed output forward file name*

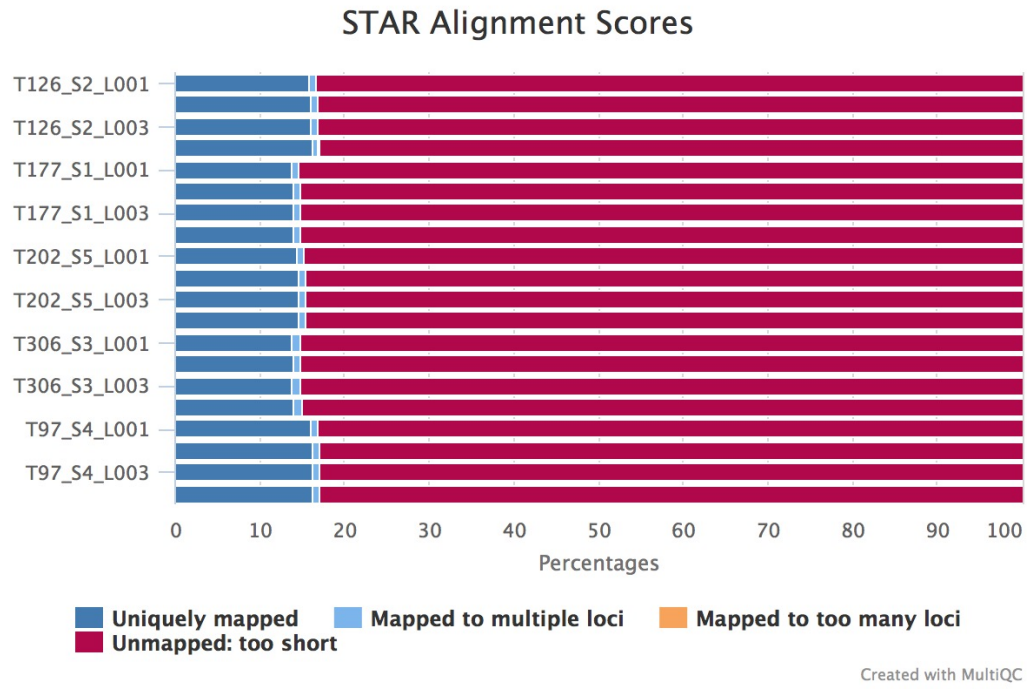
`-p TXXX_SX.R2.trimmed.fastq` | *trimmed output reverse file name*
`-s TXXX_SX.s.fastq` | *trimmed singles file.*
`-q 25` | *minimum quality score cutoff here it is 25.*
`-l 40` | *minimum read length to keep the read (here 40bps)*
`-n` | *Truncate sequences at position of first N*

The trimmed reads from the sickle are stored at **`$WORK_DIR/qc`** and the new files have suffix **`trimmed.fastq`**.

Aligning to existing turtle Genomes.

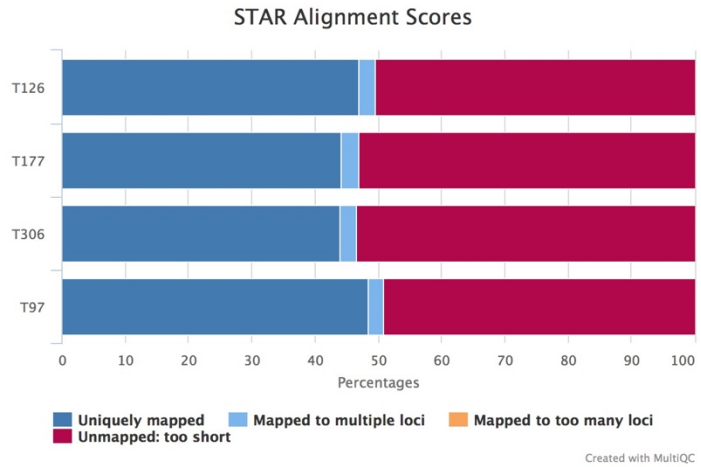
Two turtle genomes, Chinese softshell turtle and green shell turtle, were available which could have potentially served as reference genome provided there is a good alignmnt rate.

The reads were aligned to chinese softshell turtle (*Pelodiscus sinensis*) genome using STAR aligner. The alignment results are projected in the graph below::



As can be seen from the graph that less that 20% of reads alligned to the chinese softshell turtle. Beacuse of this poor allignment rate we did not use chinese softshell turtle genome as reference.

The other turtle genome that was available was green sea turtle (*Chelonia mydas*). A similar mapping approach with STAR aligner gave a better alignment around 50%, but even this is not good enough to serve as reference. The results of alignment from green sea turtle are shown in the graph below:



As the alignment rates of reads with available turtle genommes were below satisfactory levels we decided to do denovo transcriptome assembly using **Trinity**.

Trinity: denovo Transcriptome assembly

The trimmed and quality checked reads were used in Trinity. The Script is located at **$WORK_DIR/Trinity_result/Trinity/scripts/trinity_122116.sh**

```
#!/bin/bash
#$ -N tinity_turtle
#$ -M vijender.singh@uconn.edu
#$ -q highmem.q
#$ -m bea
#$ -S /bin/bash
#$ -cwd
#$ -pe smp 16
#$ -o /home/vsingh/jobsub_out_err/turtle_trinity_$JOB_ID.out
#$ -e /home/vsingh/jobsub_out_err/turtle_trinity_$JOB_ID.err


module load trinity/2.2.0

cp $WORK_DIR/qc/$1_*R1.trimmed.fastq /tempdata3/vjsingh/turtle/
cp $WORK_DIR/qc/$1_*R2.trimmed.fastq /tempdata3/vjsingh/turtle/
#/tempdata3/vjsingh/turtle
mkdir -p /tempdata3/vjsingh/turtle/Trinity/$1_trinity/Trinity

Trinity --seqType fq --max_memory 256G --left /tempdata3/vjsingh/turtle/$1_*.
R1.trimmed.fastq --right /tempdata3/vjsingh/turtle/$1_*.R2.trimmed.fast --CPU
16 --min_contig_length 300 --output /tempdata3/vjsingh/turtle/Trinity/$1_trin
ity --normalize_reads
```

Code explanation

In above script **$1** is a varibale that is sequentialy replaced by T97, T126, T177, T202 and T136. The first 10 lines beginning with **#** refers to resources and environment avriables that are requested at the cluster for Trinity job.

module load trinity/2.2.0 | *Loads Trinity*
cp $WORK_DIR/qc/$1_*R1.trimmed.fastq /tempdata3/vjsingh/turtle/ |
cp $WORK_DIR/qc/$1_*R2.trimmed.fastq /tempdata3/vjsingh/turtle/ | *Copies fastq file form /archive to /tempdata3 for effecient execution of trinity.*
mkdir -p /tempdata3/vjsingh/turtle/Trinity/$1_trinity/Trinity | *Makes directory for each sample to hold output.*
Trinity | *Runs Trinity*
--seqType fq | *input sequence type, fq=fastq*
--max_memory 256G | *Specifies the RAM for running the job*
--left /tempdata3/vjsingh/turtle/$1_*.R1.trimmed.fastq | *Forward read file*
--right /tempdata3/vjsingh/turtle/$1_*.R2.trimmed.fast | *Reverse read file*
--CPU 16 | *no of CPUs requested*
--min_contig_length 300 | *minimum length of contigous sequence* **--output /tempdata3/vjsingh/turtle/Trinity/$1_trinity** | *Output directory*
--normalize_reads | *run in silico normalization of reads.*

The results **Trinity.fasta** from the Trinity runs were saved at **$WORK_DIR/Trinity_result/Trinity/TXXX_trinity/Trinity_assembly**. **XXX= 97, 126, 177, 206 and 306** .
Trinity.fasta contains the Trinity assembled contig in .fasta format. Each of the assemblies were evaluated with help fo perlscript **TrinityStats.pl**. The script is part of Trinity pacakge.

```
perl /opt/bioinformatics/trinity2/util/TrinityStats.pl $WORK_DIR/Trinity_resu
lt/Trinity/TXXX_trinity/Trinity_assembly/Trinity.fasta
```

The results are summarised in the google_document in the sheet Trinity de novo assembly.

	Trinity assembly results				
Trinity	**T126**	**T177**	**T202**	**T306**	**T97**
Total transcripts	459525	455342	451602	366388	398661
Total genes	391824	381534	381737	314217	342182
Average length	908	873	904	845	846
N50	1278	1190	1272	1135	1136
GC%	42.38	42.87	42.27	42.57	42.64

Since the number of contigs were very high (>365000) so the contigs were filtered based on their expression levels. RSEM was used to estimate the abundance of the transcripts.

```
/opt/bioinformatics/trinity2/util/align_and_estimate_abundance.pl \
        --transcripts /tempdata3/vjsingh/turtle/Trinity/$1_trinity/Trinity.fa
sta \
        --seqType fq \
        --left $WORK_DIR/qc/$1_*R1.trimmed.fastq \
        --right $WORK_DIR/qc/$1_*R2.trimmed.fastq \
        --est_method RSEM \
        --aln_method bowtie \
        --thread_count 4 \
        --trinity_mode \
        --output_dir $WORK_DIR/Trinity_result/Trinity/$1_trinity/Trinity/ \
        --prep_reference \
        --output_prefix $1
```

Transcripts whose abundance was filter with threshold value of FPKM value as 1. All Transcripts with FPKM abundance values less than 1 are filtered out.

```
/opt/bioinformatics/trinity2/util/filter_fasta_by_rsem_values.pl \
        --rsem_output=$WORK_DIR/Trinity_result/Trinity/TXXX_trinity/RSEM_abun
dance_results/TXXX.isoforms.results \
        --fasta=$WORK_DIR/Trinity_result/Trinity/TXXX_trinity/Trinity_assembl
y/Trinity.fasta \
        --output=$WORK_DIR/Trinity_result/Trinity/TXXX_trinity/Trinity_filter
edcd /TXXX_Trinity_fpkm_filtered.fasta  \
        --fpkm_cutoff=1
        --isopct_cutoff=1
```

The stats of the filtered Trinity.fasta files is summarised in the google_documnet in the sheet Trinity de novo assembly with coloumns grouped under heading "After Filtering fpkm_cutoff=1,isopct_cutoff=1".

	After Filtering fpkm_cutoff=1,isopct_cutoff=1				
Trinity	**T126**	**T177**	**T202**	**T306**	**T97**
Total transcripts	101141	132948	105427	121913	144012
Total genes	79178	109390	84034	102961	119888
Average length	1333	1115	1268	1066	1076
N50	2066	1671	1948	1598	1648
GC%	42.79	43.26	42.67	42.9	42.92

The number of contigs/transcripts in each sample ranges from 101141 to 144012. In order to get a representative transcriptome for differential expression analysis the filtered fasta files are clustered using vsearch. This was achieved in 2 steps, step first all filtered fasta files are merged into single fasta file and step two. Vsearch programm was run with merged fasta file.

STEP1

```
cat T306_Trinity_fpkm_filtered.fasta T202_Trinity_fpkm_filtered.fasta T177_Tr
inity_fpkm_filtered.fasta T126_Trinity_fpkm_filtered.fasta T97_Trinity_fpkm_f
iltered.fasta >> Trinity_merged.fasta
```

STEP2

```
vsearch --cluster_fast Trinity_merged.fasta \
    --threads 16 \
    --id 0.9 \
    --centroids vsearch_mergered_cluster.fasta \
    --uc clusters.uc
```

The output file is located at **/archive/OBrien_Turtle_RNASeq/Trinity_result/Trinity/cluster_results/vsearch_mergered_cluster.fasta**.

In order to calculate differential expression of transcripts from denovo assembled reference transcriptome generated from last step following steps were followed,

1. **STEP1** : Generate bowtie2 index of transcriptome.
2. **STEP2** : Align reads to reference genome in each sample with **bowtie2**.
3. **STEP3** : Evaluate reproduciblity/variation between samples using PCA and distance-matrix.
4. **STEP4** : Measure expression level of each transcript with express.
5. **STEP5** : Format express output for use in gfold for fold change estimation. R code**.**
6. **STEP6** : Find differentially expressed transcripts with gfold.
7. **STEP7** : Blast and identify the differentially expressed transcripts.

Codes used in each step are given below.

STEP1

```
module load bowtie2
bowtie2-build -f /archive/OBrien_Turtle_RNASeq/Trinity_result/gfold/Trinity_r
esult/Trinity/cluster_results/vsearch_mergered_cluster.fasta turtle_bowtie2_i
ndex
```

STEP2

```
module load bowtie2
bowtie2 -x turtle_bowtie2_index  \
  -1 $WORK_DIR/qc/$1_*R1.trimmed.fastq/TXXX_S2.R1.trimmed.fastq \
  -2 $WORK_DIR/qc/$1_*R1.trimmed.fastq/TXXX_S2.R2.trimmed.fastq  \
```

```
  -S T1XXX_bowtie2.sam

module load samtools/1.3.1
samtools view -bhS TXXX_bowtie2.sam | samtools sort - TXXX_sorted
```

STEP3

```
directory <- "/Users/vijendersingh/Documents/Data_analysis_projects/turtle"
outputPrefix <- "Turtle"
setwd(directory)

express_file_1 <- data.frame(read.table("T97_results.xprs",sep="\t", header =
TRUE))
express_file_2 <- data.frame(read.table("T126_results.xprs",sep="\t", header
= TRUE))
express_file_3 <- data.frame(read.table("T177_results.xprs",sep="\t", header
= TRUE))
express_file_4 <- data.frame(read.table("T202_results.xprs",sep="\t", header
= TRUE))
express_file_5 <- data.frame(read.table("T306_results.xprs",sep="\t", header
= TRUE))

subset1 <- data.frame(express_file_1$target_id, express_file_1$tot_counts)
subset2 <- data.frame(express_file_2$target_id, express_file_2$tot_counts)
subset3 <- data.frame(express_file_3$target_id, express_file_3$tot_counts)
subset4 <- data.frame(express_file_4$target_id, express_file_4$tot_counts)
subset5 <- data.frame(express_file_5$target_id, express_file_5$tot_counts)

sorted1 <- subset1[order(express_file_1$target_id),]
sorted2 <- subset2[order(express_file_2$target_id),]
sorted3 <- subset3[order(express_file_3$target_id),]
sorted4 <- subset4[order(express_file_4$target_id),]
sorted5 <- subset5[order(express_file_5$target_id),]

write.table(sorted1, "T97.txt", sep = "\t", row.names = FALSE, col.names = FA
LSE, quote = FALSE)
write.table(sorted2, "T126.txt", sep = "\t", row.names = FALSE, col.names = F
ALSE, quote = FALSE)
write.table(sorted3, "T177.txt", sep = "\t", row.names = FALSE, col.names = F
ALSE, quote = FALSE)
write.table(sorted4, "T202.txt", sep = "\t", row.names = FALSE, col.names = F
ALSE, quote = FALSE)
write.table(sorted5, "T306.txt", sep = "\t", row.names = FALSE, col.names = F
ALSE, quote = FALSE)

sampleFiles<- c("T97.txt","T126.txt","T177.txt","T202.txt","T306.txt" )
sampleNames <- c("T97_RNAseq","T126_RNAseq","T177_RNAseq","T202_RNAseq","T306
_RNAseq")
sampleCondition <- c("T97","T126","T177","T202","T306")
sampleTable <- data.frame(sampleName = sampleNames, fileName = sampleFiles, c
```

```
ondition = sampleCondition)

treatments= c("T97","T126","T177","T202","T306")

suppressMessages(library("DESeq2"))
ddsHTSeq <- DESeqDataSetFromHTSeqCount(sampleTable = sampleTable,
                                       directory = directory,
                                       design = ~ condition)
colData(ddsHTSeq)$condition <- factor(colData(ddsHTSeq)$condition,
                                      levels = treatments)
dds <- DESeq(ddsHTSeq)


vsd <- varianceStabilizingTransformation(dds, blind=T)
rld <- rlog(dds, blind=T)

#PCA Plot
suppressMessages(library("genefilter"))
suppressMessages(library("ggplot2"))
suppressMessages(library("grDevices"))

rv <- rowVars(assay(rld))
select <- order(rv, decreasing=T)[seq_len(min(500,length(rv)))]
pc <- prcomp(t(assay(vsd)[select,]))

condition <- treatments
scores <- data.frame(pc$x, condition)
(pcaplot <- ggplot(scores, aes(x = PC1, y = PC2, col = (factor(condition))))
  + geom_point(size = 5)
  + ggtitle("Principal Components")
  + scale_colour_brewer(name = " ", palette = "Set1")
  + theme(
    plot.title = element_text(face = 'bold'),
    legend.position = c(.9,.2),
    legend.key = element_rect(fill = 'NA'),
    legend.text = element_text(size = 10, face = "bold"),
    axis.text.y = element_text(colour = "Black"),
    axis.text.x = element_text(colour = "Black"),
    axis.title.x = element_text(face = "bold"),
    axis.title.y = element_text(face = 'bold'),
    panel.grid.major.x = element_blank(),
    panel.grid.major.y = element_blank(),
    panel.grid.minor.x = element_blank(),
    panel.grid.minor.y = element_blank(),
    panel.background = element_rect(color = 'black',fill = NA)
  ))


# Heat map of sample-sample distance
```

```
sampleDists <- dist(t(assay(rld)))
suppressMessages(library("RColorBrewer"))
sampleDistMatrix <- as.matrix(sampleDists)
rownames(sampleDistMatrix) <- paste(rld$condition, rld$type, sep="-")
colnames(sampleDistMatrix) <- NULL
colors <- colorRampPalette( rev(brewer.pal(8, "Blues")) )(255)
heatmap(sampleDistMatrix,
         clustering_distance_rows=sampleDists,
         clustering_distance_cols=sampleDists,
         col=colors)
```

Based on Sample-sample distance matrix and PCA plot it appears that the treatment in sample T202 is not very effective. Sample-sample distance matrix T202 failed to group with other treatment samples and in PCA plot it is closer to T177 in PC1 of variation. PCA plot also shows that the reproducibility between samples was low and may be acounted to uneven treatmenty. Based on which T202 was not considered further in differential expression ananlysis.

FIGURE1

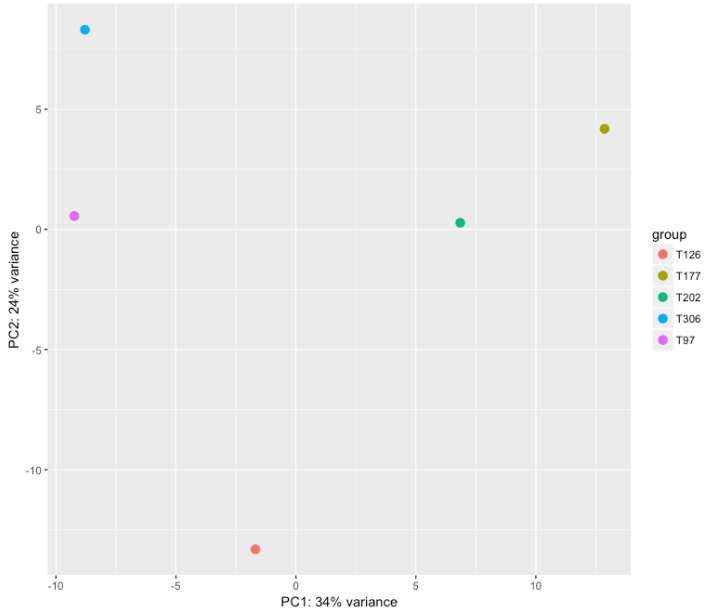
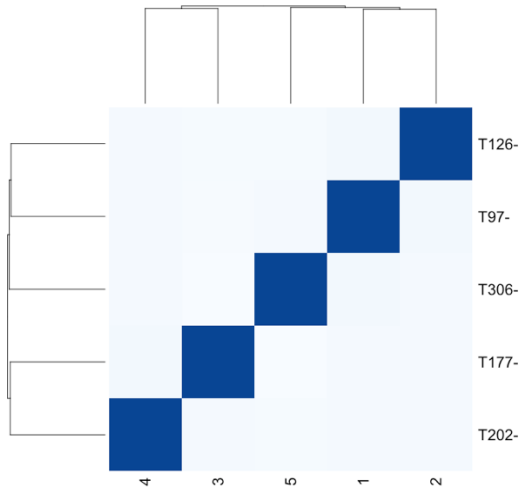


FIGURE2



STEP4

```
express $WORK_DIR/Trinity_result/Trinity/cluster_results/vsearch_mergered_clu
ster.fasta TXXX_sorted.bam -o TXXX_express
```

STEP5

```
express_file_1 <- data.frame(read.table("T177_results.xprs",sep="\t", header
= TRUE))
express_file_2 <- data.frame(read.table("T97_results.xprs",sep="\t", header =
TRUE))
express_file_3 <- data.frame(read.table("T126_results.xprs",sep="\t", header
= TRUE))
express_file_4 <- data.frame(read.table("T306_results.xprs",sep="\t", header
= TRUE))

subset1 <- data.frame(express_file_1$target_id, express_file_1$target_id, exp
ress_file_1$tot_counts, express_file_1$length,express_file_1$fpkm)
subset1 <- data.frame(express_file_1$target_id, express_file_1$target_id, exp
ress_file_1$tot_counts, express_file_1$length,express_file_1$fpkm)
subset1 <- data.frame(express_file_1$target_id, express_file_1$target_id, exp
ress_file_1$tot_counts, express_file_1$length,express_file_1$fpkm)
subset1 <- data.frame(express_file_1$target_id, express_file_1$target_id, exp
ress_file_1$tot_counts, express_file_1$length,express_file_1$fpkm)



subset1 <- rename(subset1,c("express_file_1.target_id"="GeneSymbol","express_
file_1.target_id.1"="GeneName","express_file_1.tot_counts"="Read Count","expr
ess_file_1.length" ="Gene exonlength","express_file_1.fpkm"="RPKM"))

sorted1 <- subset1[order(express_file_1$target_id),]
sorted2 <- subset2[order(express_file_2$target_id),]
```

```
sorted3 <- subset3[order(express_file_3$target_id),]
sorted4 <- subset4[order(express_file_4$target_id),]

write.table(sorted1, "T177.cnt", sep = "\t", row.names = FALSE, col.names = F
ALSE, quote = FALSE)
write.table(sorted2, "T97.cnt", sep = "\t", row.names = FALSE, col.names = FA
LSE, quote = FALSE)
write.table(sorted3, "T126.cnt", sep = "\t", row.names = FALSE, col.names = F
ALSE, quote = FALSE)
write.table(sorted4, "T306.cnt", sep = "\t", row.names = FALSE, col.names = F
ALSE, quote = FALSE)
```

STEP6

```
gfold diff -s1 T97,T126,T306 -s2 T177 -suf .read_cnt -o $WORK_DIR/Trinity_res
ult/gfold/DEseqNormalisation/T97_126_306_VS_T177_DEseqNorm.diff

gfold diff -s1 T97,T126,T306 -s2 T177 -suf .read_cnt -norm count -o $WORK_DIR
/Trinity_result/gfold/countNormalisation/T97_126_306_VS_T177_countNorm.diff
```

Sample output

```
# This file is generated by gfold V1.1.4 on Thu Mar 16 18:10:02 2017
# Normalization constants :
#    T97    484518  1.22474
#    T126   403897  1
#    T306   438018  1.10668
#    T177   412156  1.02988
# The GFOLD value could be considered as a reliable log2 fold change.
# It is positive/negative if the gene is up/down regulated.
# A gene with zero GFOLD value should never be considered as
# differentially expressed. For a comprehensive description of
# GFOLD, please refer to the manual.
#GeneSymbol                      GeneName                GFOLD(0.01)   E-FDR
log2fdc 1stRPKM 2ndRPKM
T126_TRINITY_DN100084_c0_g1_i1  T126_TRINITY_DN100084_c0_g1_i1  0   1   -0.44
1601   7.53901 0
T126_TRINITY_DN100158_c0_g1_i1  T126_TRINITY_DN100158_c0_g1_i1  0   1   0.344
063    4.0825  4.37953
T126_TRINITY_DN100177_c0_g1_i1  T126_TRINITY_DN100177_c0_g1_i1  0   1   -0.03
91536  4.63939 0
T126_TRINITY_DN100177_c1_g1_i1  T126_TRINITY_DN100177_c1_g1_i1  0   1   0.004
15352  3.49838 0
T126_TRINITY_DN100232_c0_g1_i1  T126_TRINITY_DN100232_c0_g1_i1  0   1   0.165
359    6.25644 5.03373
T126_TRINITY_DN100238_c0_g1_i1  T126_TRINITY_DN100238_c0_g1_i1  0   1   -0.44
4182   7.31943 0
T126_TRINITY_DN100263_c0_g1_i1  T126_TRINITY_DN100263_c0_g1_i1  0   1   0.312
583    6.21347 6.66555
```

```
T126_TRINITY_DN100331_c0_g1_i1  T126_TRINITY_DN100331_c0_g1_i1  0   1   0.519
429    2.3197  7.46542
```

The first output file contains 7 columns:

GeneSymbol:
Gene symbols. The order of gene symbol is the same as what appearing in the read count file.

GeneName:
Gene name. The order of gene name is the same as what appearing in the read count file.

GFOLD(value):
GFOLD value for every gene. The GFOLD value could be considered as a reliable log2 fold change. It is positive/negative if the gene is up/down regulated. The main usefulness of GFOLD is to provide a biological meanlingful ranking of the genes. The GFOLD value is zero if the gene doesn't show differential expression. If the log2 fold change is treated as a random variable, a positive GFOLD value x means that the probability of the log2 fold change (2nd/1st) being larger than x is (1 - the parameter specified by -sc); A negative GFOLD value x means that the probability of the log2 fold change (2st/1nd) being smaller than x is (1 - the parameter specified by -sc). If this file is sorted by this column in descending order then genes ranked at the top are differentially up-regulated and genes ranked at the bottom are differentially down-regulated. Note that a gene with GFOLD value 0 should never be considered differentially expressed. However, it doesn't mean that all genes with non-negative GFOLD value are differentially expressed. For taking top differentially expressed genes, the user is responsible for selecting the cutoff. Value in the bracket indicates the significant cutoff for fold change. Default 0.01.

E-FDR:
Empirical FDR based on replicates. It is always 1 when no replicates are available.

log2fdc:
log2 fold change. If no replicate is available, and -acc is T, log2 fold change is based on read counts and normalization constants. Otherwise, log2 fold change is based on the sampled expression level from the posterior distribution.

1stRPKM:
The RPKM for the first condition. It is available only if gene length is available. If multiple replicates are available, the RPKM is calculated simply by summing over replicates. Because RPKM is acturally using sequencing depth as the normalization constant, log2 fold change based on RPKM could be different from the log2fdc field.

2ndRPKM:
The RPKM for the second condition. It is available only if gene length is available. Please refer to 1stRPKM for more information.

STEP7 The Transcripts which had GFOLD value of >1 or <-1 (2 fold change in expression) were identified. The list of such transcripts are beow. The transcripts were blasted to RefSeq data base and link to the top hit is provided in the table below.

```
#GeneSymbol GeneName    GFOLD(0.01) E-FDR   log2fdc 1stRPKM 2ndRPKM BLAST mat
ch
T306_TRINITY_DN182495_c5_g2_i1  T306_TRINITY_DN182495_c5_g2_i1  1.164   0   2
.59571 0.941592    5.30302 https://www.ncbi.nlm.nih.gov/protein/gi%7C11026164
71
T202_TRINITY_DN201895_c12_g2_i6 T202_TRINITY_DN201895_c12_g2_i6 1.11337 0   2
.69724 0.840202    6.18057 https://www.ncbi.nlm.nih.gov/protein/gi%7C10693677
18
T177_TRINITY_DN214142_c4_g1_i3  T177_TRINITY_DN214142_c4_g1_i3  1.03659 0   2
.71875 3.85956 49.6845
T306_TRINITY_DN185419_c0_g4_i2  T306_TRINITY_DN185419_c0_g4_i2  1.02746 0   2
.58503 1.18818 8.03017 https://www.ncbi.nlm.nih.gov/protein/gi%7C1072289864
T306_TRINITY_DN181469_c0_g1_i1  T306_TRINITY_DN181469_c0_g1_i1  1.02281 0   2
.57231 1.4096  11.3412 https://www.ncbi.nlm.nih.gov/protein/gi%7C946605930
```

Transcript sequemces:

```
>T306_TRINITY_DN182495_c5_g2_i1
CCCACTAGAAACGATCTTCTAGAATGGGGAACCGCGACAAGTGAATCACGATGGACTAAG
CGACCAAGAGTTTAAGTGACAGGGTTGTTCGCGTAAATCCCGGCTTAAATTCAAAGTTGA
ATATGTCCGTACGGAGCACTCTAGCACCGGACGAAGGATCGTGGACGGTGGGCGTAGGGT
ACACAGGGTCAGCTGTGTATAAACTCAGTACTGAAGCAGGATTTGGGAGCATTGCAGTCA
CACTCAAGTGGAGTACTCCGCCTTTAATACCCGTATGGGGCAATACTCCGAGGAGGGCGG
GTTTTCAAACGAAACTTGAGTAATACTTTATATGGTCTCAATATACACTTGTGGATGTAG
GGAAACCAGCCATGGGCGGACCCAAGGCGGCGCACCTGACACATGCGGCAAGAGAAAAAA
TCCTTCGGGTGATCTCTGGGTCCATTCCAAAACTATGAATTGCGATTTGTGCTTACCTCG
TGTTGGTTGGCTTTGAACACTAAATACAAACACTACAATGTCTGTTCCAAAATCTCGCGA
ACCTGGCAGCGGAGGAATCCGGCATGTCAATGTTAGGGGAAATGAAACCATCATAATACA
TAGTAGGGGAAGAGGCGGAGGACCTAGACCCGTCTTCCGCGGAAGGGGAAGAGGAAGAGG
TCGAGGCCATCAGGGACCGAAAGGTGAATCCGGGACTTTATCTTCAAACACTTCATTGTT
TGGAGAGGCTGAGACTTTCGCTGGAAAAGTCTCACAATTGGGCACTGAACCCATAAGAGC
TATAGGGCATTCAGTGGGGGGAGTTATAGCTCTCCCGGCAGCTGCTGCTGCATCTGCAAC
ATCTCTAGTACCAGGAATGGAGCCAGTAACTAAAGTGTTGAAAGGATTGGAAATAGGAGC
CAAAGTTGTAGGTGGCGGGGTAATTACCGCACCTGCTGCTGCTACACAAGCCGGCGTGGA
AGGCTTGAATAGAGCGATTGGATCGGCTGTTTCGGCAACAAAAAACTCATTATCAATTGT
AATCTGCTTGATGGCTCTTTTAGTAATGTTGCCTGGAGGAATGACAGAAAAAATTTATGC
CAGTCATGTCTATACTAATGATGAAGTTTGCAGAATAACTAATTGCTGTAATGGAGATGA
GATCTATGGATGCATAACTGGAAGATGTCTTACAGCAATCGGGTGTGTTATCTGTGAGCA
AGCTGACAATCAAGGAATCTGTTGGGAAATGGAGGCATTAGGTCTTTCAACCAGACCTGG
GACCTTGGCAATAACTCACACTGTCCAGCAATGGGCGGGAGTTATTGTTTGGGTTGGAAT
GCTTTGCGACACCTTGGGTTGGGGTGAAGCGTGTTCAGCTGCCTTTGTTTTAGGCGACAT
AGCATCTTTACACTTTTTAACACATCCAAATATAAATATGACATTGGACTGTGAGTTGTA
TGTGTCTGATGGCGCTTCAACTTTTTCTGTATTTTGGAATTTGGCAGGGGCTTCCAAACC
ATTTGAATTATTAACCTATATAGCAAAAATACCATCCGCTATAATATCATTTCTTCTTAC
GGGTGATATGCTTGTATTGTTCTTCATCATTTACAATCTATTTGTTGGGAGAGTGATGAA
GGCTCTTTTGTTAATATTGTTAATTGGAGCCGTTTATGGGAAGCATTGTCCTTTGAATTT
GCCTTATGGACTTAATTGTTCACAATGTGAAATAAAGAATCCAGGACGTTATTCCCATAT
ATACAATTTCTTTGACACTTGCATTTTTGAGAGTACGGCCGGCGCTTGTAACATCACGAG
AGATAAAAAATGTGCAACCGAAGTTGGTACATGGTATAACTACACCGGAACCATTTACAC
AGGAAGCACATTATTTTTCGACGATAAACCCTTTGTGAAAGCTAGGGGGAAATGCCGAAA
```

ATCAATTAAGTCTACAAAGCAACTGAACATCTATTGCGTGGGAAAAGATGGGAAATGGTT
TTACCATTACGGAAATTTCACAGGATTGGAGAACACAAATTTATTATGTACAATCAATTC
ATCATCTGCTCCAGTGGAATCAGTGACGTGCATTTCTGGCGTACCAAGATGTAAGGATGA
GAAAATATATCATGCCAAGTTGTGTAAAGAAAAGAGGAGGTATTACACTTATGAGCGTTG
TGGAGTCACAAATGTGTTATCAGATTTTGCTTATGCTGTTAATTACAAATCTTGGGTCGA
TTGTTATAATAATATGTCACGTGCTGGAGTAATACGTGCAGGATACTATAGTGTTCCATA
TGACACAGGGGTGAAGGCTTCTTGCGATCAAAAGATAGTGTCAGTGAAATGTACAACCAT
TCCAAATGATACACCGCTCATCACTTTACCGATGCCATGTTTCAGTGTGCCATTCACAAC
TAGTAAAACCAGAACGTGTAAAGGATATGCATGGCCTCAGGGTGATATCTTAAGAGGCCT
AGTTTTAGTTGATAAAGTCACCAGAGAATATCAGTTTTTTGCTTCAACACAATCAACTCA
GCAGTACCATCTACTTGAACACACTATATTAGTGTGTATTGTTTTAATGTTAATGGGGAG
TAAATATGTGGTGATTCTTTATTTTTTGTATGCCATGCAGATTTCCCAAATTGGTGCAGC
TTTGCCTAAAACTCTGACTTCAGCCATGGTTGGAGTTTATGCAAGTAATGTAGAATTAAT
GGATTTGTTTTATATTTTAATAATGTTTATTGTAATCATTTATGAAGAAAGAttaatttt
ttttgtattatttaatatcaaaatgtattatcatttattcttttattgtgttattGGGAT
TTATTATCAATTTTGTCCATGGCCTTCCGTTTTAGGAAGCGAAATTTGTGTTTATGTTCC
TGATTTCTGGGGTTCCTATAGAACAGATGTGGTTTTTGGATTATGTATATTTATGATGTG
TTTTTTAACATCATTTATATCTGCGAGGGGAAAAATGATTAGATTAATAATTTTTGCTAG
GGTTCATTATCAAATTCAAAGATTTTCAATTTCTTGTCAAGAAAATTGGATTAGTAATTT
AGAAGGGAAAAAGTATAAGCATGTAATTGCCTTTCTTTTGGGATATTTTTTTCCAAGAGA
AGTGATGCAATGCTTGatttttgtgtttttctttttatttccttgattttGTAATATA
TACTGTTTATTCTTGGTGTTTAATCAAACCAAATATTGTTTTAGCTGTAGAGTTATTAGA
TGTAATAAGTAAAATTAATGATAAAGCAATTTGGTTAATGAAGTGGCAGAAAGTAACAAG
AGAAGCGAAGGTCTCCGTCTTCGATCACTTAGGACATTTATTGGTCGAAACAAAATCAAA
GCTTCAATCAGTTGTCATCTCATTGGACCCAGTTACCATTAAAGAGGAAGAATTACAAAT
CATAATAGACGATACTTATCGTCTTGCATGTGGTGATTTTGTTGATGGAAAGCCAGTTGT
TGCTAGGAAAGGGAAATTAATTCAATTAGGTTGTAATGGAACATTAGCGCAAGGGTTTTT
ACCAACGGCTCCCATAAGGCTATTGGCTTCAAAGAGAGAATCATTATTTTCTGTTATAGG
TTCATCAGCAAGTGGAGTGAGAAAGAAACCTATGACTGGACACATTGCAAATTTGAGAAC
ACCACTAGGTTCCTGTATGGGGTATGGGATAGATGGAGCAATGATGACATGTGTGCATGG
GACCAAGGGCAGGACTGTTGCAACTCCAAAAGGGGCTACATCTGCGTTAGTAGTAGATGC
TCCATCAGACATAGCAACCTATCCGGCGCCATCAGGTTTAGGGATGCTTACCTCTTGTAC
TTCAGGGTGTAGATCTGCATATTATCCAACTAGGGATGGGAAAATATATGAGGTTGTCAA
AACAATAAAACCAGGAGTCTGGAATTTCGTGAGTGCCATACCCTTGAGAGAAGCGAAGGG
AAGTTCTGGAGGCCCACTACTATGTGGAGAAGGAAAGGTTTTAGGGACAATCGTGAGAGT
TGTGGCCACTAGAACCACAGCAGTGGCGGCAATCGTGGTACCTCTTGAGAAGGTTTCTGA
CATGATACCAAACATACAACCAAGCATCATAACCATACCTAAAGTTACTGAAAAATACAC
TGTCGCTTCAATTGTGGCTCCAACTGGAAGCGGAAAATCTACCAGATTGCCTTACAGCTA
TGTGGAAGAAGGGAAAAAAGTCTTGGTTCTAAACCCAAGTGTGGCAACCATACACTCTAT
GGGTCCTTACATGACTCAAAACTATAACATACCCGTAAGTGTCATGCATGGAGGATACTC
AAATGTGATTGGTGAGGCTAAATTAACATATTGCACTTATGGAAAATTTGCTGTGGACCC
AATGATTCTGAAGGGATTTGATGTCGTCATTTGTGACGAATGCCATTCCACGGATTCGAC
ATCAATTCTGGGCATGGGATTGGTGATTTCAAATGCAGAAATGTGTGGAGTTAAATTAGT
ATTATTGGCAACAGCAACACCTCCAGGGACAACAGTATCTCCACACCCCGACATTGAGGA
AGAAGAATTGGGGGAAGAAGGTGATATAGACTTCTATAACAAGAAGTTGTCATCAGTGAA
TTACAAAAAAGGGAGACATCTTATTTTTTGCCACTCAAAAACAGTCTGTGAAGAGCTTGC
TGAAGCCTTTAGATCTGCAGGATGTAACGCCATATTCTATTACAGAGGCATTCCAATATC
TGCAATTCCACAAACAGGAGATGTAGTCATTTGTGCAACTGATGCCTTGATGACAGGGTA
CACGGGAGACTTTGATACAGTAACAGACTGCAACATGAGTACGGCGGAATGCATGACAGT
TGACCTCCAACCCACTTTCACAATAGGAGTGAGAAGCACTGTACACGACACTGTAAACAG

ACTACAAAGGAGAGGGAGGTGTGGTAGAGGAAGGAAAGGAAAATACTTCTACTGCATAAA
AAGAGCAGTACCAAGAGGAAATGTACCTGACTGTAGCGTTCTAGGTGCATATGAGGATGG
AATTCTATGGTACAACATCACTCCAGAGAAAATGACCACACTTTTACAAAACTACTCAAA
TACAGTGGGACTCCCAACAGTAACCTGTGATTTAGAGTTCATGGAAAGGTTAATAGATTT
TTACAAAAAGTTCAGGAAAGCTCCAGAAACTTTAAGAGCCCAACAAATGGGATTCTCACA
CCCATTACTGACCGGAGCACAAGCAGTTGTGTCAGTCAGAGAAAATGCCAAACTACCAGG
AACTACAGGAAGATGGGAAGGTTTCAAGCAAGGCCCCAACAAGTGTCCATTACTTTGCAC
TTTAGACGTCGAAACAATTGACATCCACTACAACAATTTGGAGACCAAAGAATTAGAGAA
AAGACTAGGATATGACAGAAGCGAGGAGGTTTCGTGGACCCTAGCTGCGGCTGCAGGGGT
AACAGTAGGCGTTCTAATATTAAACATCACTGGCAGTCTAGTACCTACAGATGGTCTAAC
AGTTCTTTGTGGGGCTCAAATGGCAGAAAAAGCTTCTACAATTTTCCTCAATGAGGATTT
CCAATTTGAAGAGGAAGTAAGCTGGGACTTGAACGTAGATTTGGGTGCAATAGTACACAA
GGCAGGTACCGCTTATGAAGGCGCTAGGAAGCAGGCTTTAAGAGCAGCAAGATATGTAGG
AGAATTGGGAGTGGAACAAAGAACATCTAGTTTTTTGAGCTCAGCAATACTGGCCCCTCT
TGGGCTAACAGCATGCACATTCTTGTCAGCGACCAACAATCCAGCCTTATCAACAATGAC
AGCCGCTGCAGCAGCTTTAATATCACCACTTGGATTGGGGCAAACCGCAGTTGTAGCTTT
GTTTGGGAGCACCATAATTAGCTGCTTTACATCTTCAAAAACTGGCACCGGCATCGCTAT
CGGAGCTTTGGCTGGTGCAGCAGCATCAGAATTGGGTGTTGCTGAGATGCTCCTTTCAAT
AATTACTGGCTATGGTGCAGCATCAACTTGTGCACATGTTACCTTTGATCTCATGATCGG
AAAGGCTCCATCAGGCCAAGATCTGGCTCAGTGTTTATTGGGACTATCAAATTTAGGAGC
AGCCATAGCAGGTATCACTTGTGGTGTCTTGCTTGCAAAAATGACTAACCCTAGTTCGAG
CAATTGGACTAACAGATTATTAGCAATGATGGTAAAAGGCAACGTGTTGCCAGACTCTTA
TTTTTTGGATGCAAAAAAAGCAGAGGAGCAAATTTCAGCTCTGCTTAACGCCATGACACC
AGTATCACTCTTCAATAGATTAGGGACATGGATGAGGCAAGCTGAAGAAATAGATGTAAG
CCTAGGCATGGAACTACTCCCGGAAGTTTTTAAGGGAGCGTGGGAGTACGGCTCAGAATT
CGTAAGAACAAGACCCTGGGAAAAAATTTGGAGATTCATCTTAATTGTTTGGGATTTTAT
AAAAAATTTAGCAGCAAAATATCTTCCAAGTTTTTCTTTAAAATTACCACTCCTCAATTG
TACAACGAGGGTTGTATCAAAATGGGTAGGGACCGGAAGCGCTGAGTGGACTTGCAATTG
TGGAGCACACATACAAGTAACAATAGCAGATGGGAAGGAACAAACTAGGTCAAGTAGTAA
AAGGTTATGCTTTAGTGCGGCAGGTGGCACTGTATTCAATTTGACTACCAAGTTTTTAGG
GGTGGACCATGTTCACAGAATCAATAATGGGGACATCTACACGTATAGGTATGCCACTGC
TGAAGATTCCTACATTGATCTAATGTTAACTGAATCAGGTGTAGTGATAGTGGGGTCAGA
CATCTTGTACCTAACAGTAGTTGGTGTACTTGAAGCGGTTTCAAGACAACCGCTTCTTGA
AAATGGAGTTCCAATAGCAAAGTGGAGGAAAATACCTGGTAGGAGATACAATACTGGCGC
TATTTTGGACTACAGTGGGACAGCAATCACCCTACCTCATAAAATCAAAGCTGTAACAAC
TAAAACTGACCCAAACGCAGATCCAGGGTCAGATTGGTTTAAAGTGCCCTGGAAAGAAAT
AGGGCCTGTGAAAGTAAATACCCCACAGAAACTTGAGGAAGTAAAATCGACCATTGCGGA
AGAAATGGTCCCCCAAAAGGAAGGAGATGGAAAAAAATTCAACGTAGAGGACTTACCCAC
TCAAGGAACGGTGGAGAAATCAGAGGAAACAACTGAAATTAGTTATGATGAGGCCTTTTA
TGTTGTTGGAAATGAGTCAGAATTGACAGAGTTTGCAAAGTGGATGAATGTGCCTGATTG
GAGATGGGAAGCAAACGTACCATATTATCAATTTGAGCTAATTAGGAATTCAAAATATTA
TGTGCTCACAACTCCAGAAACCTTTGAATTTGAGTTAATTAAAGGGAAAGCTACAGGTGA
AGCGAAGCTTTATGACGCAGGTTTCTCTACAGTGTACAGAATGATGAAGCCAAAAGGACA
GCTTCTGGCAGCAGAAGAAGCTGATGAACTTAAGAAAGAGCAAAATTTGATCACCATATT
TGTCACGAAAGACCCAATAGGAGGAGCAGTCCTGAAGGCGAAGGACCACCTTGTCATAAG
ACCTGATGAGTTTGCAGGACTGTGTGATGGCCTAAAACTACCAAGAACAACCGGAAGAAA
GATGTCTGTAGACTGGTCATATACATGGGACATCATGAATTTTGCCATTACTAGCACCAG
AAATCCAATCATCCCTGTAAGAATGAAGACCGCGTGGCTCAGAGGAGATAGAGCGAATGT
CGCCTTCACGAAGGAGTCGGACTTGGGAGAAAGGTTGAAGAAGGTTACCATCCAGAGACC
TGAGAAATTGCCAGAATTTGTAGTTAGGGCTTATGAAAAAGCAGCAACTCAAGTATCAAA

GATGAATTTACCAGAGTTAACTTATGAAGAAGCAGTGACATTGCTACACAATAAAACAGC
TAAAAGTAGGAAGGGTTATGGAGCAAAAGAAGCAAAAGAGGGGAAAGCTAAAGAGCATGT
GATGGAATTAGAAGCCAAATTATACTCTCCAGAACCATACAGGGAAGATTTCTCCTGTAC
CATAATGCCCAAGCATGAGCTTCAGTTAAATACTGGAAAACCACCAAGAGCAATATGCTA
CCCTGATTTAGGAGTTAGAATACTTGAGAAAATGATTTTGGGAAATGCAGTTAAACAAGT
ACCAGAGGCAATACTTGGAGAATCCTACGGGTTCGTGCCACCAAGAGTTAAGCTCAAAAG
GTTGATTAAATGGTACCAACAAAAAAGAAAACCTTTTGTATTGTCTTTAGATGTATCAAA
AATGGACTCCAATGTAACACCAACTGATATGCAGTTTGAAGCAGAAATCTTCAAGGCCGC
AAGCAGTCCTGGCAACAGAACTGCTATACAACTACTGCATGACAAGGTTTACAAAAGTGG
AGACTTGTATTACAGAGGAAGAAAAGTAGGAACGAGGAATTGTAGAGCCTCAGGAGTACT
TACAACATCAGCTGGAAATACAATATCCCATTACATCAAATTGGAAGCATGTTTAAAAGA
ATTAGGAGTAAAAGCAGATAAAATAATTCAGGGTGATGATGTAACAATGGTTATGGAGAA
GCCTGATAATCCAACAGAATTTGTGGATAAACTAATAGGTTTGTTAAATGCTACAGGGTT
AACCCAACAAGTACCAATGATAGGAAGTAGTTTAGAAGATGTAGGGACCTCAATGATAGT
CAAAAGTGTTCTGGTAGATGATAAGCCTAAATACTACATAGGGAGAGAGCCCTATAAGCC
TTTAGCCACGAGCATGGCTGAAACGACTGAGGGTAATCCGGAATGTAGATGGCTTGGATT
GCTTATAGAGCATTATCCAGAACCTTGGGCTTCAAGAGTTCTAGCTATATCATTACTTAA
TTCATTGTTAGAGTTTAATAAGGAAACCATATCCTTTAAGTATTTAGGAAATATTTATGA
AATTAGTTTAGAAAATTTGCCTTTATTGTTAACTGCTTTGCATGGTGACTGGATTTGGAA
GTTTGAGAACCATCCCAGACAGGTGAACATCACAGCCCAAGCTCTGAAAGATCTAGGATT
GCCACCTTTAAAGGGGTGGTATAGAAGAGCTTTAAGGTTGAGAGTTGCCTGTTTGAGGAA
AGGGGGGATATACAGGTACTTAGCTAAGTACTTGTTGCATTTCACTACTAAAAATGTGGA
ATATCCAAGTCCGTTCCCGTTCAAAGATAAGCAAAATGCCTTAGAATTGTTGTATAAATC
TGTGCCATATTCAAATGTTGTTACATATTTAGAAATAGTTAAGTCTAAGAATTTTGTTTC
TGAAATGTTATTGGGTGCCATTGTTTTTTTCCTTTGTGCTTACATTTTTAAAACTGTGTA
AAGAAAATAAATATCAATAATGTTTTTGGAAGGTGCCGACCAACGGTTGAGTGTTGCGAC
CACTCGTAAAAAAGAGCTTTATTCTCTAAGTTAGAGACTTCAGCGGTGGCAACGCTGTCG
ATTTAAAGCCCTACGTGTGTACACGTAGTAAAAAATATAAAATTTCGTGGTACACCACGG
TAAGGGAC
>T306_TRINITY_DN185419_c0_g4_i2
TGACTACTATAGTGACAGAAACATCAAGGAAAAAGCCCATGCAGTTTCTGTTAAATTGGC
TAGTAAAAAGGTTGGTTCCAAATTAGCTGTAACCTCTAAGGAAAAATTCCTGGCAATGAC
CATGGACAGCTCAATAAAACATCTGCTCATTGGGCAACCACAGTCCAAAAGCTAACAAAA
TATCAAGAAGTATAAAAAGTAGGATAGAAATACTACAAAGAATGTTTTAATATCTTTATG
TAAACTGATGACATGCCCTGATCTTAAATACTGTTCACTTCTGATCAAGCTATCTCATAA
AAGATATAGCAGAAAATGGCTGATTTCAGATATAGATGACAAAAAATGGTTGGGGCAGGG
AAAGACTTCCATATGAATAGAGATTGAAAGGACTGGGAAAGATAACTTTAGAGAAGAGAT
GAATAAGAGAGGACAGTTTAGAAGTGTGCAAAATAATGAATGGTATAGGGAAGGTAGATC
ACACTCTTCAGTTCACCCTTTCTCCTTCTGCAAAAACAAGGGGACAATCAATTAAAACAA
CAGGCAATACATTTAAAACTGATAAGCCAATCCTTAACTGCTTGAGTTTAGGAAGAAAAT
TGCTCTAATGGCAGGTTATTCCATAATTGTCTATTACTGGGATTATTGCATCTTTCTTTG
AAGCATGTGCTACTGGCCATTGTCAGACACTAGACTAGACTGGTTTGATAAGGGATAACA
ATTCCTATTTTCTCATGTGTGTAAAGTAAGTTTACTTTTTAAGGAAAAATAATGAGAGAT
ACTGTGTATGGGTTTTGTGCCTAAAACACTTAAATTATTGAACGGGACACAAAATACTCT
AATCTGGGCCTGTTTAAAACATTGGGCAAGATTCTCTGTCATCAAAGTAAGTGGCACTAT
GACAGTTTACTTCAGCAGAGGATCTGGTCCACTGCACAAACTCTCGTAATGTATGTGTCA
TATACCTTGTCCTGCATCAGTAATAGCTCCCCACCTCCCCCATCCTTATAAACACATTCT
GAAACAATGATCCAAATAGTAGAGTAGCACTATTGTAAGATGTAACATTTTGTACCATGA
CCCTAGAATGAGGTAGCACATTTAATAATAGGGAATCAAAGCTCTCCCCACTCCTCTTCC
TCATTCATTGCAAGTAACATTGACACACCACCGAGTCTATCCCTTATTTTAATAAATGAT

TACAGAATTGTACAGTAATAGTCCTAAGCGCTGCAGTATCAGTAGGCTGAAATATTTTAT
ATCATTTCTGTAGTATTTaaaaaaaaaGGATTACATCAAAACTGTTTCCAAGGCAATGTA
CTATGATGTCACAGAGGATTTATCCTGCTTCCCTTTCCAATTATGATCGAAGAAGAAAGT
TTAGTTTATCACCATGTTAAGGGCCAATAACAAAGAGCCAGAGTCACACTCTCCCAAAGT
TTGATGGTATTTGTCTGGAGTCTTTGTTTGCACCCATCTCTAGTTAAAAGGGAGGAAAAC
TGATCTAGCCCATACAAGCCCAGAGCAATGCTCAGTCTGAAATAGCGGGGGTGGGGAAAA
TGTAGAAAGATAATTATaaaaaaaaCCCAACTTTGGGCTCGTTGGAGCACTGACCGCCTG
TCATATATAATGATTTGCTAGAAACTTGGCAATAACATTCTCCCTTGCGCACCAGCACAG
ATGATGGCAACACAGCTCTGAGGATAGCATTTGTTAAGTGACTGTTATTGCAGAGCCAGT
AAGGGGAGCACATTTGCATGAAAGATTTGTTAAAGTGCTCACAAAATACATAGGCTGTTG
TGTGTGTATTTACAGTTTCTTAGTTTGGAGTAATGACAATAAGGAAAGGTCAGATTCTGG
CTGGCAGTGCACATAAATGCCCATCAATAAATCCATGAGCACAGCATTTAGATGGTAATG
ACATGTAGCTGCTAATGACCTAGGCTGGCTTTGAATTGGTGAGATGGTGATGGAAGTGTT
CATTATCAATCCCCTGAGACACCATGCCCCAAGAACTATGACTGTCACGTCAGACATAAT
GACAGCGTCCCAGGAGAGCCCAAACAGGCACTTTACGTGTTGACATGACTGCAAATGGCC
CAAAGCTTCCAGCGCCATCTTAGTTTTCAATGGTTGCTTAAATCTCACAGGACAAAATCT
GTCTGTCTTACAGATAAATAGAAAGGGGGGAGGAAAACATTTTTCTTAGAGGGCACTTGC
CTACCTAATGAGGTCCAAAATATGACCTACTTTAGAATTGATCATATCTTGTAAAGAATG
TCCTCAGAATCCTAAATAAATATTTTTTAAAGAAGAAAACCTCCTACTGTCCTCCGATAA
AAAAATCACAAGAATGAGCCTGCATTTTGATGCAACAAACTAAGCAATGTTGAGGCTCAT
GTTGACTTCCGCTGCAGTAGAAAGACAATAAAATGCAATGCCTGCTGGAGGACAAGGTTG
AGATGGAATGGAACATTTTCCCCCTATGTTTTTAAACAGCATTTTACATATAGCTCAGAG
AGCAGATAAATCATTATGAGGCTATTAATTCTGACTTGGAATGAGTTCAATATTTGAGTT
TCAGCAACCTCAAAATTCCAGGCTCTCCAGATCTCTAGGTTTGGTTCAGCCCTTTATGGA
AGATGGGGCCATTGTGAAGGTTAGCTATTAAACACCATTAATTTTGGAGGTGCTAAGATT
TAGATTGGGGCCTATTTTGACTACAGAGGTGAACATGCCCCTTTAAACTAGTTTGTGGAA
TACAGTTTTTTAAAAAATACAGTAAGGTTTTATTTTAAGGAAAGAAAAAGAGATGAGACG
AGATAGACTATGGAAAGATGCTTCTCAAAAAAGCTTTAATGAAACTGGCGTCTGGAAATT
CTTCACTCACTGTTGCCTTTCCATAAGAATGTAGGAGCAAGGACTCCTTCTGCTATATTC
CCTGGGGAGCTGTGAGAGGGAGGGTATTTTATGTGTAACATTTTATCTGTTCCCCTGTTT
TGAATAGATTCTCTTTTCATCTAGCTGCAGTGGCAAAGGCTGCATCAGCAGTGGGCAAAG
GGCATCAAAATCCAGTCTTGAACTTACAATACCTGCAACTGTCTGGTGTGTCAACGAATG
TGTTCAAGCACAGTTCACAGGCTGAGGAAAGCGTGTCTTCTGCCTAGGTGACACTCATCC
ACTCTGTATCTTAGACAATTTAAAGGCAACTCCAATCTTAATTTTCATGACCACAGACCA
CACAGCAATTCCCTCAACGACTGCTAACTAGCAGCTTAGAAATATTAAATGGCAGAGGCA
TAACTTGGATCCTTTCCCGATGGACAAAACTCCCCAGTGTGTCACTCCCCAGGATTGTAC
AGGTAAAATGAGGTTAACTAACACTAAGTCTTAGTGCCAGACCAAGAAAGACTCTCTTGG
ACAGTCTGCTCACTCACTCAGGACATGTGCACCGTGATCAGTGGAACTGGAGTCAGAGGG
CCCATGTACCTTGGAGGGACAATCATTGGGGGTCACCAGAAAAACAGGGAATGTGAGGGG
GTTGGGCTGGAAGCAACATAGCCACATAAAGAGAACAGTGCACAGGGTTTAATAAAGATC
CACTTGCTCAACTTAACTTGCATTCAGCTTCACTCTTTGCTATTGAAACAACATGGAAAG
GAGAAGAACCTGAAGTGGATTTTGTCAGTCTTCTCATTCATTAGGAATGTGAAGAAAAAG
CTGAACTAAACCCCTATCCTCATCCTCTGAAAGAGATGGAGTACAGCTCCTGGGTCATTG
GTGAAGTGGCCTGTGTTGTGTTTCAGTGCGGGGACTGGACTGAATGAAGTCTGAAAGGAT
GTCAATAGCTGAGAAAGCTGTAATCAGCAGAGTATGAAATATGGTTTATATTTGAGTTTA
TAAGCCCTGGGTGAGGAAAATTCAGCTGAAACAATGGCTTTCCTGTTCCAAAATAACAAC
AATGGTGGATATTTCAAATGCATAAGCGAGTAATATAACAATACTTTGCACCTCTACCCA
AGGGTCTGAAAGCATATTACATATGTTAATGAGTTATTGGCCTGATTTTCAAAGGTGCTA
AATACTTCCAGCTCCCATCAACTACTTTAAAATCCAAGCATAAACCCTATTAGGCCCCTC
CTGCCATGGTGATAAAGTTAATTATTATCCCCATTTTACAGACAGACAAACTGAGGCATG

ACAAAATTACATGACTTGCCAAAGATCAGACAGCTATTCAGCAGCAGAGTTGGGAACAGA
ACTCAGCTCCCCGGCCACCAAACCCTGTACTTTAACCACCAGTCCATGCAGATTATGCTC
TCATAGATTGCCAGCTCTCCTTACTTAAGTGCGCACACATTCCTCCATGGTCTTAGGGCA
CCAGCTGGCAGCTTCCAATCATTCTTCTTTTAGTTTATCCCTGGGTTAACACTATGGGCG
AGATTCTGCCTCTCTTGCTTATGCTAAGTAGTGTCTTACTACATAATGTACTTCTTGATG
TCAGTCGTGCTACTTGTGTGTGTAACAGGTCATCAGTGAGAGTAAAGGTTTCACATTCTG
GCTCTAATTGAAGGTACCAGTGCTAATAATGATGCTGCTGACTGAGAGTGCGTTGTTATG
GTAATCATGACAGTGAGGTATTTGGGGGCTGATGTTGCTGCTGGGAGAGTTCCCTTGGGA
CCAGAAAGACCTCCTGCTCTGAAGGACATTAGCTCCCTTCCCCTCCAACAAAATTGAGTT
TATATCATATAAACCAGACTGGGGTTGGTGCCTGCCCTGAAAAGCCACTTAACTAGCTGC
CTGGGCTGGAGGTCATTATTAGTAATTATATCTCCACTGAAGACGGATTTGGCTGCCTGC
CAGTCCCAGCTCCTGCCCTTCTCTGACCAGAATGAGGAAACACAGCCAATCAGCAAGAGC
TTTTCATTTCCTTTGCAGGGCCTTGACTGTTCAGCATGAGACCTCTGAAATGTCACCTGA
ATCTCATCTGTGTGTGCAGTATCACTGACCTGAGGCTGCCAGGGTTTTCATGGGGTAGGA
GGGAACAGTGTAGTGGATAACTAGAAGAGAATTCTTCTCCATCTTCAGCTGTGACCTGGT
TTCACAATAGTAAATAATAGTACTTAGTACTTATCCAGTGATTTTCATCCACAGATCTCA
AAGCTCTGTACAAAGATGGGTTAGAATTACTATCCCCACTTTATATATGGGGAAACTGAG
ACAGAGAGATTGGGAACTGGGAATAGCAATGAGATTTCCTGATTCCAATGGACCACAACA
AAAATTGTATCCCTTCTGGCTATGAAGAAAACCTCCCAAGTAACTTAAAGCAGAGTTTAC
TTCCTAAGCCTGCACACAGGCAGACAGAAACAATAGTTCTGAGAGGCAGAACTTACCCCA
TTTAGTTCTATTGCTTGTTAATGCTGAAGTCCAAAGATGACAATTAAAATTTACATATTG
TTAATTATTTCACTGCTGGTCATTTTAGTAGAGACACTGTGGTTTTGAACTGGGGGCAGA
GCTTGTATAGAGTACATTACATTTCCTGCAATTATTGTGGACCCTGGCAGTAGCAGGGGA
GTTAAGAACCATATCTAGCATAGGTTGCTGGCGCATAGGGTTCATATCAGGAAAACACAA
AGGACATGGTGGGCATGAGGAGGTGCAAGTGTGTGAGAGAGAGAATGCCACAGATTCAGT
CCTGGGATTGGGTGAAATAAAATAAGCCCAAACTCAGCTTTCCCTGTCCTAAAACATTAA
GACATTAAACTATGAACTGGTCTGATGTTAGCAATGACATCTTTGGAGAGGTTTATCATG
AACTAGGTTAGTGCTGCCCCTTCTCCATTGCTTCAGCTCTTCCCCTCCATTCAGATCCCA
AATTTGGCTCCTGCCGGATTAGAGGTTTTGTGCAACAGTCTCAACAACAGTTTGAACAGA
GTCTCATGGAATTGGAGATGGGAAAGACACGTTATTTATTTGTAATGCGGTAGCAACTTA
GAGGCCCTGACTCACTGTGCTAAGCATTGTACAAATTTATAACAAAAGGAGGTCCCTGCA
CCGCAGGGCTTACATTCCGAGTCTACAATGAGAGACAACAGGTGCACACAACAGACGGAC
AGGGGGATCACAAGCTAACAGTGCAAGCCTCTCCAATCTGGCTTCCACTTTCTACACTCC
ACTGAAACTGTTTTCAGCGGAGTCTCTGACAACCTCTTCCTTGCTAAATTTCAGGGTGAA
TATTTCATCCTCCTCAGCCTATTGGCTGCCTTAGATATAATTGTCCATCCTCTGCATCTT
GAAATCTTGTCCTCCTTTGGCTTTACACAGTCTATTCCCAGGGCC
>T202_TRINITY_DN201895_c12_g2_i6
TTGTCCAATACTAGGCAGGATTTTCAAAAGTACCTGAGTGATTTAGAAGCATGAGTTCCA
CTGAAAGTTATGTAGATGCTATTGAAAAATCCCACTCAATGTTTTTTGAAAGCttttttt
ttttAAGGTGCTCAGAAACCATGGGGATAGGTTCAGTGCAAGAGCCTAAACAGAGTAGA
ACACGGTGAATATTCAGACTTGACAATTATGGGGGGAAGGGGGAATCTTTGTACATGTAG
ACTAAGTAAATTGAATCTGAAAGAAAAAGATTTAGGCATGGAATAGTACTGTCACTTTTT
TAATACTTTGTTAATTTTCCAGTGAAGTCGAACCACACTTCAAACCAGTAGCTTTACTGT
TCTTCCCCTTGCTACCCCTTTTGTTAAGTGATCTCTGTATTAATTAGTTGATGCTGCCTT
TTTTTCCTGTGAAATTCAAAACAAGTAAACttttttttCCTCCCAGGATTCATTTGTGAT
GTTCTGATGTGCTGTTGAAAGGAAAAGAACAGCGTCTGTTACAGACTGACACTTCAAAAC
GCTCCTGCTAATGTATTTTCAAACTCATTGAATTATCATAGAAATTCATCATGGGAGGAG
CTGTGAGTGCTGGGGAAGACAATGATGATTTAATTGATAACTTGAAAGAAGCTCAGTATA
TTCGCACTGAGAGAGTGGAGCAAGCCTTCAGAGCAATTGACCGTGGTGATTATTATCTAG
AAGGATACAGGGACAATGCTTATAAAGACTTGGCTTGGAAGCATGGAAATATACACTTAT

CAGCACCTTGCATTTATTCTGAAGTCATGGAGGCATTGAAACTTCAACCAGGATTATCTT
TTCTGAATCTGGGTAGCGGAACCGGATATTTGAGTACAATGGTGGGATTAATATTAGGTA
ATTTCTTTTTTATTTTAAACATATATGAAATGTTGGTTACTGAATACAGATTCTTAGATT
AACACGTCAGCATCTTTGCCAGGGTTTTGTGATTTAATGTGTTATAGTTTTTAAGTAATG
TAATGAAAAAAGCATAGAGTTAAGAATTTGGAATGTGCTTTCAGAATATTGTGAAAGAGC
TGCTAGCTAATTCTTAGAAAAACTGTTATTGAAAAGCATTTTTTAAGAATTGCATTTTAA
TGAAAACAATATTTTTGCTGTTTAAGCCTCAGCCACAGATTAACCTTCTGTAGTCACCTG
AAGCTGCAAACCTTTCTGACACCACCGTAATTCCCACACAAGCCAAATGTTGAGTCACAA
AGAATATTATGATTTTCAACAGGGAGTATGGAAACAATTATTTCTAAATTTTTCTGCAGT
TTTTAAGTTCACTAAGGACTCTTAAAGATTGTTGCTCACATAGAAGGCAAGAAATGTATA
TGCTTAGGACACAAAGTAAACTGCTCTTCACAAAATAATCTAGATTAAATAACAGCATTA
GTAAATTTTATGGCAGATATTCACTGCTTAAGACTGGGGTTTTTTTAAGCTAAGTTACCT
GTGTAGTTTACTACTCATCCTCTCAACTGTACTTAAATAAAGCTGTGCTACCAGCTGAAA
AGAAAATGTGGTATGTCGTTGCTAACTGAATTTAATTTTAAAACTATACCTTAGAGAGCA
GAAAATGTACATTAAAATGCATTTGTTTTCAAAATGAAAAGCTAAGCTCTAGTGCTGATT
TTTATAAATATCCTCAGATGTTTTCTACCTATTGTAATAATTTTGCATTACACAACAGAA
TTTCATTTATTTTGGTAGTGGGATGTAAGTGTGACAGCCACATTTTTCTTTCACTAAATA
AATATTTGTATCTCTCAGTATTGTTTTCTAGTTAAGGAAACACTACTTTTACTGCTACAA
CAATCTTACAAAATGTGTACTTTTTCCAGTATATGGAGTTCTTGAAGTCTTGGATTACTC
ATTTTTACAAATAGAGATAATGGAGAGCAGAAAACCATTAAGGGAAGCAGATACTATTCT
GAATGGGAGGAGAATGTGTAACAAATTTAATCTGTCAATTCACTTCCAAGAGGTGTAGGA
TTCCTTCCAATTTAATGTAAACAAAAGTCAAAGAGCTGATTCAGTGCTCCCTCTTGTGGA
AACTGAGATTCGTTTCCTTACCCAATACAATATAGTAAGCAGTAATGCTTCTGGTGAAAT
ATTCATTCATAGGTATAAACTCAAAATTCCAGTACATGCTGTTAAAAGAATTATAGCATT
TCACTTTGTTCATATAATGTACTGTACCTCTTGTGCTAGCACTCTATTTATGAGCTTGGG
GAGTTTTATCTCTATGGGAGATGATAAATATAGGCAATTTTGTATGAGTTCTTCCCCAGT
AGCTGTTAAAGTATAAATATACATATAATATGATTAGAGACTACTTCCTCTTTTGTTTGC
TATTAGCCCTGTGCGCAGGACCATAAATATTTCCTGAGATCATATCAAGAGTGAAGATGA
TTAAATAAAGCCCTTCGTCTTGATATGTATTTGTTTTATTTGGCTAGGGATCCTCTGCAC
CTGCTTCTCTATATGAGGTTGTATTTACCTTATCTAACTAATCTCTGTGGGGTCAGGAGT
GTATGATGTGTGCTGCCTGGCCTGCCGTCTCCTCACCCAGCACAATTCTGTCTCTGAAGA
ATGACCTGTGCTGGTTTCCAGAAAGTGAATGAAGAGAACTAAGTAGTGAAAGTGACTGAT
TTCTCCCCACCTTTTCACATGAGGGGTGAGATCTGCAGCAGAGGGCCTTCTAACTGTGTG
AATCAGCAGACACAACCTGCTCTGCAGTTCTTAACTCAGCATGGTTAGACAGATTTATAG
CAATATTTGAACAAAATATTTTTGAAAAGGAAGGTTTATTTTACTGGTTTTAATAAAGGC
AGCCCCTTTGTGGAAAATTAAACCTTTACACCTTCTCAAAATGGAGCCTATTGTTGCAAT
TTGTACTGTATTGGGTTAGTGTTAGGAAGCCGAAGATTGTTTGTCAATGGTTGGGTATCT
GATGGTTCAGTATAGGGTGGGAGTAGTAGTGTATACCATATTGATGATTTTGTGAATGTT
AATGTGACACTAAATTAATTTGCGGGTGACTTGAAAATGAGGAATTTTTAAAATACAAAG
TGGTCTCTATAACTTGCTACAGTAGTCAGAGATTAAGTTTAATATAAACAGAAGGAAAAT
AATACACACGGAAAGCGTAGTTTACATAATTTTAATGATCAGTACAGACGAATCCTTTGA
TTTCTGTAATTGAGTATTACAAGATATTAGAACTGCAAACACTTTTCCAAGACTTAATAA
TGGTGGTAGATAATTGATAAGGTTAGTATAATCACATTGTGAAGGAAATGAATATTACTC
CAGGAAGTAACACAAAAGTTAGGGTAGTGTCAGTTTCTGTGAATTGGTAAGACTGTTCTG
GAATAGCATATACAGTCATATCCTCAAATTGGAAAAGCTAATGAGAATAGTGTTACCAGC
TCTGTTTTGTTCAATATAACACATAGCAATACTGAAATGTATGTATCATTACACTCCATA
TTTTGGATAGTCATGTTGCTGCAGAAACTAATTGCTGCCTCATAGCACGTTGATCATGAA
AAGTTCTAGTATTACACACACATGTTCCTCATTACCCTGACAGACACAAAATTCAAAAAA
TGTGTTTGATATTCTTCATGATTGTCTTGAAAAGTGGTAAAAATTAATCTGTTTGGAAAT
GTGCATGAATTTTGGAACAGATGGCGGTAACTGTAATTGAAACCAGCTAAGAGTGCAGCA

AAAATACTAGACACATGAAAATAATTTAGGAGGTTAGGTTAGGGGAATGTGAATTCATTT
ATAGTAAATTTATATTTGAAAACCTGCTGTGATTCCCAGGAAAGATTTTTCTAAAATATT
CTGTTTTGAAAACACTGACACCTCTTATAAAAAATGAACAATGCTTACTAGCTACTTCTA
ATTACTAGTCAGAAACTTTTTTGTTAGAATGATATTCTGATGAAGAATACCCTCTTTCTA
TGGGAAAAATTCAGTTTTGCAGAAGAATTTTCATTTTTCCATTGGCAGCTGACATTCTGA
TTTTCCCAATCTGGAAGGTGTTTGTCAATTTCACTTTCTGCTGGCCCAAAGCGCTCTGCT
AGAAGGGCTGAGCAGGAACCAGGGTTGGGAAGCGAGGGCTGGTGGAGTGGCTGGGCTCTC
TGCCTCCCAGGCAATGGGTCTGAAAGGCAGTTTTCTTACAAGAAGATGAAAAACAAtttt
ttgtttttgttttCCTAAAATGGAATATTTTTTGAAATGCCTTCACACACACAATTTCCT
TTTTGCATTTTTGTCCTGGTTCAAGTGTAAATTTGTTTTTAAAAATATTGAAAATTTTTA
CAGAAAGGGATTTCACATTTTACAGCCAGCTTTACTAGTAATGCAAGAGATAGAAAAGGG
AAGCAATGAGCTAAGTTAATAAAGGATGTATGCGTACCAACAGGACACCTTTCAATGCTG
TGAAAGCAGGTATCTCCCTTGTGGAGCACTCTGCCTCCTGTAAGTCTGCCAGCAGGAAAC
CTTTATGAGCAATGCATGTAACGACCTCTTTCGGATCCACTTTGTTCATGGTGTACAAGA
CGTTCTAATTTCATGCTAAGAATTACAGATGTTTAGCATATTGTGGTGATGTTGCACCAT
TAGGTCACGCTTAGTAAATGGAATGTATTAGGTTTTTTTCTCCATAAAAGACAGGCTTTG
CTGCGAATCTGCTATTTAATTTTATCCAGGGTACTTAAACATCAGAAATACTTCTATATT
TTTTTCTTTAAATGCTCACAATACCTCTGTTCTGCAGAAAAAATCCTTACGAATTTCCCT
GGATGAATAATTTTTGTTACTGAAATAATAATTTGAGGATGCTTCTTGTATTAATTAAGT
GTTGGGCTAAATAATTCTAAAAGTAAAGGCTTGGGGCATAGTGATTACTGAATTGATAGA
ATGCAACCTCTACCTGCTGTTTCGTTGAAATGTGCTTTTTTCCTCTTCTCTGCATGGATA
CTAAAATCTCTGTTCTGTGGGCTGCCTATTACATATTGAATTCCTATCTCTCTATGCCCA
CAACAAACATGGCCACCTCTTTACCCTGGCAACCCCCCTAACTCTAGTCTTAATATCTCT
CCTTGCACCTCCCTACTTTTAGAAGAGTGGACTTAGCCCTTTTACCACAAACTGAAATTC
ATTTTCAAGCTGTTTGTTAAAAACATATGACTACTAAATGAGTTGAAAGGATAAGGAGTT
GAGTTAGAAGTGAATGTGGTGGACATGTCAGTGGGATTAATCTGAATAGAGACCTGCACA
AACGCTGTACTGTTGTAGATCAATACTTGTATAGCACTTTAGTAGATTTCTGTGGTCCAG
TTGGCTGAAGGCTAATCCCTCATACGGATTCAGGAAACTGCATAATTGAGAAATTTGTAT
AATTATGGGTATAGATATGCGCATACTATGCTGAATCTCTATGCCATCATCAGTTGAGAT
GATTCCACTTCAATATAATATGCTTTAAATTTACCAACTTTTAATCCTTCTACAAACAAA
TACATTATAACCATCAACACCTGCTGTCTTTAAAAATTTGTGCCACTCTAAAATCTAATT
AAAAGACTAAGAAAAGATCAGACTTACTGCCTTATAAACATTCTCTTATGAAGGCTAATA
AGAAAAATGTATTTGTGCTTTCTTTGCGGTAGTGTGGGCATGTTAGAACAGGAAATAGGG
AGCCAGAACTCCTGGGTTGTACTCCTGGCTTTTTCACTGGTTCACTAAATAATCATGTGC
CTTAGTTTACCCTTCTGTAAAAAGTATAAAAATATCCCTTTTTGTAGCCTTGATGGAGGG
CTTAATCAAATGTTAATAAAAGCACTGTGATACCATTAGATGAAAGGTGTTATACCAGTG
CATAAAACAGTACTTTACTTATAAAAAAACTTTTTTAATTACAGGCCCTTTTGGAATAAA
TCATGGAATAGAGCTTCATTCAGATGTGGTAGAATATGCCAAGGAAAAAATGGAGAGCTT
CATAAAAAATAGTGACAGCTTTGATAAGTAAGTATCTCACTTATCCAGTTCATTTTGCAA
CCAGAAATGTTTTGGATTTTTTTACTGAGGTACAATATTAAGGAATAAATGACATTTTGC
AAATTTTAACATCTTGTTAATGTTCATTATTGGAAATCATATTTGAAGTTTGTTGGACAT
ATTTAAAATGAGGTAAGTGGAAGTTAAAAAATATAGTTTTT
>T306_TRINITY_DN181469_c0_g1_i1
AGAGAACTAAGTTTTCAGTGCATGGTCTGTGCTGTATGGTAGTTATGGGAACAGTGTGTG
GTTTTGGGTAGAAGAGTGGAGGGCATGTATCGTTGGGTGGCTTAACATTTCATTTCCCCG
TTTGGTGGGGTATGCTTCACAACAttttttttCTGTATTAATTATATGGTTACCTGACTC
TTGAACTCAATAGGAGTTTGGTCTGCATAGGAATTGCAAGATCAAGCCCTATTTTGCTAC
ACGTTTGCATAAAGTATGCATCATGTTGAGACTTATTTTAAATTGGTGTACTGTTTGTGA
TAACGGCTTGAAAAAAATGTATAGAATGTTTCTCATATGGTTTTGTAATTGTAAATTCTA
TTCATCTTCTCTAAAGGTCCTTTTAAAGAATCTCTTTAGTGTGTGTACGTACATACGTAC

ACACACATGACCTGATAATTTCTTGGGATATTCTTTGTCTTCATTTATTTATAAGGTAAT
ATGTTATCTGTAGTGATATAGGCTGCCTTCAGTAGGGTTGTGGTGGTAATTTTTTTATGT
TATGTGATCTAAAATAGAAGAGCTCTAGTTGCTCCAAGATATGTTCCAATGGTATAAAAA
TGATATACATTACAGGAGTTTGTCGTTTTTGTCACATTAGATTAGACAGAACCTCAGATG
TGTCAAATGCAGTAAATGAGATGAGAGGCTTCACTTATTCCTAGATTTTGCTAAGAGAAG
GAACACTTTCCTGGTGGGTTTTTTGTTGGTACAAGAATCTCTGATTAGCCTGCCTGGATT
TGGACCTTCAACAATGAAGGAGAAACACCTTGAAAAATGAGTCACTAAAATCTTAATTCA
CTGAAGATTTTCTGACCACTTACTTGAAAAACAGAGAATTGGAGATGCTGCAAATTTGAT
CttttttttGCTCTCCACTTGAATTTTGGGTGTGGTGTTTCCTCCATGAGCTTCAGAGAA
GATTGAATAGTCTCATGTAATTAGATGCAGATTTCCATTATTGGGCATTCTGAAATCCAA
AGATATTTTGTCATCGTGACTGGTCAGAAAAGAACAATTAAAGTTGATAAAACAAAAACC
AGCTAACCAGCCAAGTAGGTAAGGAATATATTTTTGATGCTTTGCTCAATAACAATCTCT
CTATAAAAAATCATGCCTACCTTTTTTATGGCATAGAGGCAAAAGCACAAAACCAAAACT
AGTCAGCCACGAAACATTCCAAGACAGATACCCCAAAGCACTTCAGTAACTTACTTCCTA
ATACCAAGTTTTGAGGTCCCTTCAAGTTTTGGACACAGCAGCACATAAAGGATTTTTAAT
ACAATACATTCAAATATTCAGTTGATAATGGCAAATTACTTTGGAGATCATTGCAGCTAA
TCTATTTGTGGATAGAAAATTGGCAACCTTGCTGAGTGACACATTGTAAACAATTTTATT
CCAACTTaaaaaaaaaaaaaTCCTGTGTCAAGTGTTCCTGAAATCACCCTTGCTTATCTG
TCATAGCTAGCTAAATAATGGTTTACACTTAAACAGTGCATTGTATTTTCTAAGGTTTGT
ACATATATAGTAACATTTACTAATCCTGGTAAGGTCTAAGATATTCAGGGCATTTTAATA
ATCCTAAAGAGGATGAAATGACAGGTGATTTACAGAGTCTCTGAAATTAGTATAATATGG
TTCACTCATTTGACCAGCTGCATTTTAAAAGCATGTGGTTATTTTTAAAGTTTAAGATAA
CTACACAAACATATTTTTACATACATTGATTTCAATTTAAAATTTTCTTTGGATATCAGC
AATACCCTCTCACCCCTCATTTTATTTCTTTATTGTTTaaaaaaaaaaTTGCTAAAGATT
TGTTAATTATCCTTGGAGTAGTATAGTTGTTTCCCCCCATACTCCTTCCCATGAGGGAAA
AACAGCTAGAAAAGAGGGGCATTCCAGGCTCCCCCAGGTTCAAACAGGGTTCCCCAGGAG
GCTATTCCTGTCAGCGATGGACACTCATGGTGCATCCACTGCCTAGGAGAGGCACACATG
CTTCAGAAGTGTATGCATTGCCTCAAGCTCAAGGCATTCACCAAAAAAGGCAGGGAGCTC
AAGTTAAAATTTATTATGATGGAGTTCACTTCGGCCAGCTTCCAACCCTGACTTGGAGAC
CCAGTCTGGGCAGTGATCCCCCTCGCCCCAGATGACAGCTTCAAAGAAACATTCCCTGCT
AACTGTGCAGAGAAAGTTGCTGCATTCCTTGCAAAATGGCTGCCACGAAGCAGGCACCGA
GTTTCCCTACCAGGAAACCAGTTAAAAAGAAAAGGTCCCCAGCATGCTCAATGACATTGG
TACCGACCGGAGCACAGTTGAGTACCTTTGAGTCAGTGGGATCCTCCGGTACCGCTGATT
TTGGTAAAGCCAAGGAGGGCAGGCTTGGAAAGGAGTGGTAGAAATAGGCCTAGCTTCTCC
ACAATGGTACCAACAGAGTTGCTCAGACATGTGGCACTGATTAGATCAACGCTGCAGCAG
ATTATGGCACCACCTACCAATCACTCAACGGAACGCAGAGCCCCCTCGGCACTGTCAACC
GCTTCAACTATAGCGGTACTGCCCCCAATATCTTCATTGTTACCAATGCCCATCCTGATA
CTGCAAGAATTTCATTTCCAGGACTTGCTCGTATCCAAGACACCGGAATCTCAGCTTCTC
AGTACTGATATTGCCCTGGTATGCATCACACCCATAGTGCACTCGGTACCAACTGAATGC
AACCAAACAGGTTAGTGTCTCCCCTTCCTCCCCCCAGTGAGGAGGACAATGAGGGAGGAG
GAGGAGGTTTACTTCACATTAACATACCACAACCCACCTCAACAGGCAGTTACCAGCTCA
GATCCATTTGTGCAGGTGAGATACCAGCCTAGAGGTGATGCACATGGGTGATATGGACAT
TTCCACCACACTGGCCTTACTGGGCTGCATACAGACAGCAATTTTCCAGGCCTCCTAGCA
CCCACAGAGAGAGAATGAGATGCTCCCCATCAGCCTGGGTGGCTAGACCCCCCGAGTCTC
AGGACAAGACATTTGAGGAGCAGGATGAAAGAACTGATAAGGAAGCCACCTGAGCAGCCA
TTTATTTCCTAATCCTCTCTGGACGAGGCTGTCATGTCTCCTCATAAGAACGGCCGTACT
GGGTCAGACCAAAGGTCCATCTAGCCCAG
>T177_TRINITY_DN214142_c4_g1_i3
AGACTCAGCTCCTGACTCACAGCCCTTTTATAGGGCCAGCTGTGGCCTAATTGGGTGTGG
CCCCAGCTGTGGCTGCCTTCCCCATCAGCCTATTCTTATTGGCTCTCAGCCCCAGCCCTC

TCCAAGGGTTGGCTTTTAACTCTTTCAGGGCCAGAGCAGGGTGACTGCCCCACTACACAG
CCAACAGTATTTGGTCCTGCCATGAGGGACCTGGATTAGCAGGGACACCGTGCATCTGGG
TTTTCCCAGACATGGCCTCTGCTTTGGTCCTCTGATCTCATCCAAGTAGATTTCTGAAAT
GTGAAGAAATGTCCAGGAGTTTTCCTTGCTTGTTGCAGATAAGAAAAAAAGTTGTCCGCC
CATCGGATCATTATCCGACAGCGCCTTGTCTCGCCGTGACCTGTATTCCCGGATCCAGGA
GCTAGCCAAACGGGCGTCATCCAAAGAGGACTCGGACAAAGATAAGAAAAAAAGATGTCT
GCCCATCGATTCATTATCCGACAGCGCCTTGTCTCGCCGTGACCTGTGTTCCCAGTTCCA
GGAGCAAGACACACTGTCGTCATCCAAAGAGGACTCGGACAAAGGT