

MCF10A RNA-Seq Analysis

Sebastian Kurscheid

26 August 2016

Contents

1	Introduction	1
1.1	R Markdown	1
1.2	Including Plots	2

1 Introduction

Purpose of the analysis:

- Determine differential expression in human breast epithelial cell lines MCF10A/MCF10Aca1a
- Following conditions were used
 - MCF10A WT vs MCF10A TGFb-treated (induction of EMT)
 - MCF10A WT vs MCF10A H2A.Z knock-down (induction of EMT/effects of siRNAi mediated gene silencing)
 - MCF10A WT vs MCF10ca1a WT (normal breast epithelial cells vs RAS-transformde, cancerous cells)
 - MFC10Aca1a WT vs MCF10Aca1a shZ (effects of siRNAi mediated gene silencing)

Two biological replicates were used for each condition and libraries were sequenced on an Illumina Next-Seq 500 sequencer using Illumina TruSeq protocol, 76bp PE. The sequencing reads were pre-processed and aligned to the human reference genome hg38-based Ensembl Version 84 transcriptome annotation using STAR. Actual transcript quantification was performed using kallisto against an index built from Ensembl 84 (all cDNAs + ncRNAs). The sequencing data processing pipeline is implemented using snakemake <https://bitbucket.org/snakemake/> and can be found at <https://github.com/JCSMR-Tremethick-Lab/Breast>.

1.1 R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

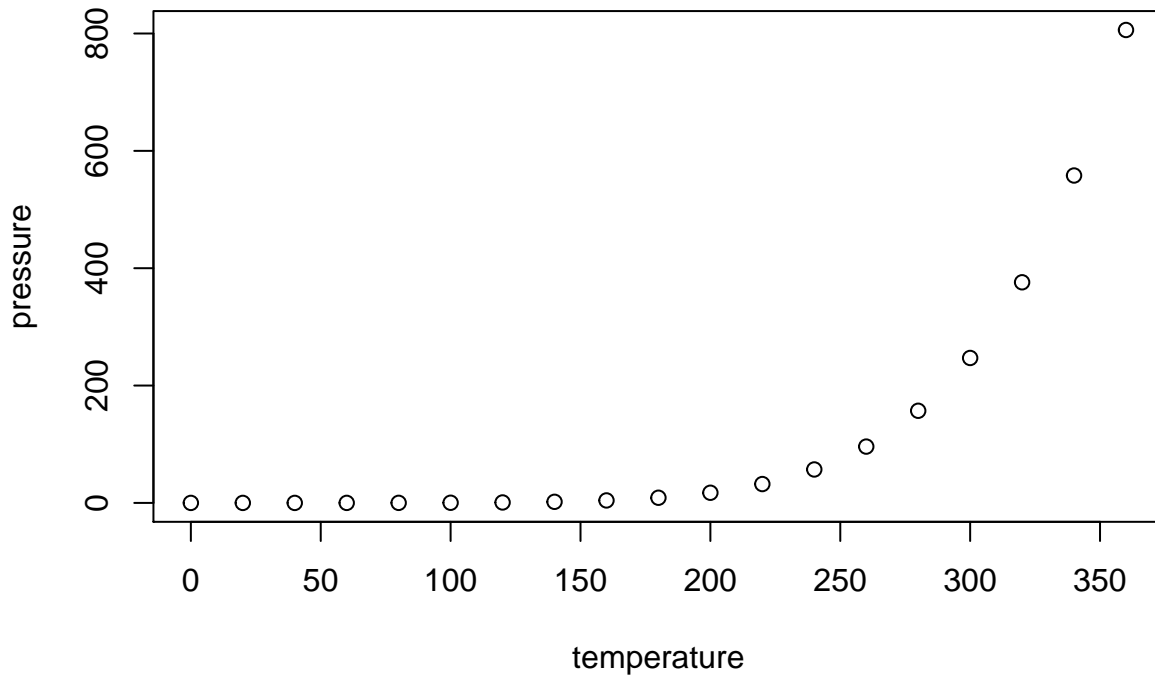
When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```
summary(cars)
```

```
##           speed           dist
##  Min.   : 4.0    Min.   : 2.00
## 1st Qu.:12.0    1st Qu.:26.00
## Median :15.0    Median :36.00
## Mean   :15.4    Mean   :42.98
## 3rd Qu.:19.0    3rd Qu.:56.00
## Max.   :25.0    Max.   :120.00
```

1.2 Including Plots

You can also embed plots, for example:



Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.