

# RNASeq Module

Trushar Shah

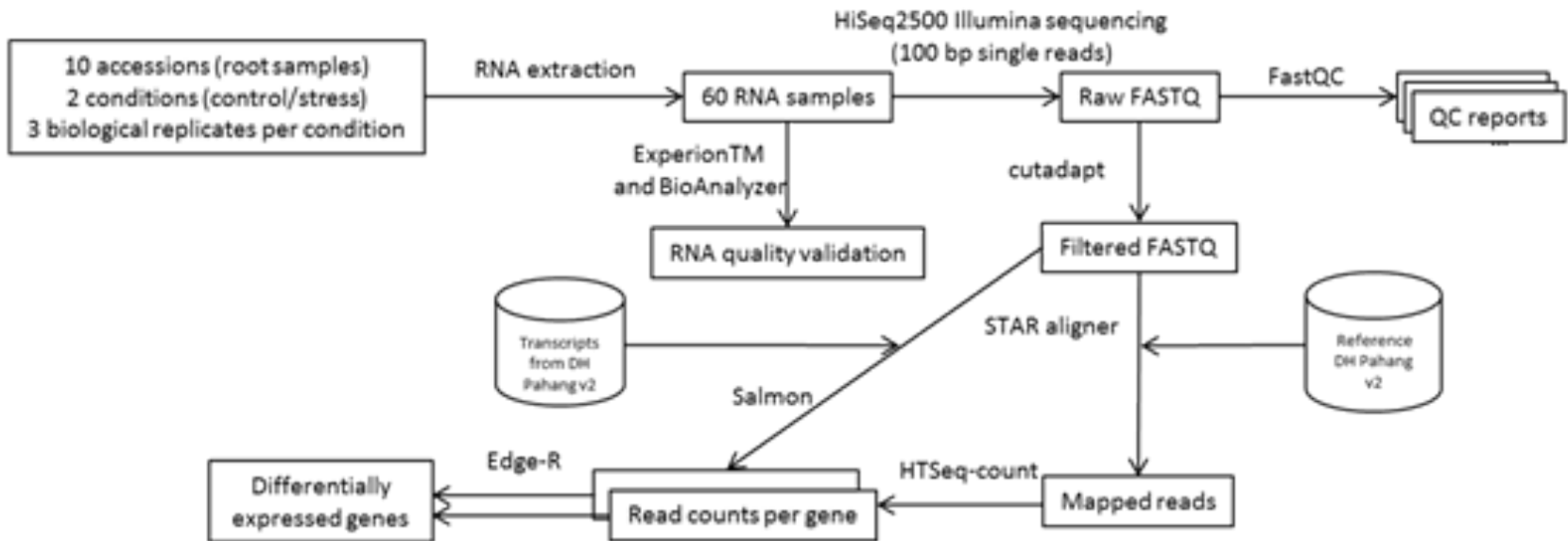
IITA

20/06/18

# Plan for today

- Go through an RNASeq experiment using R and DESeq2 (part of the Bioconductor suite of packages)
- Look at a visualisation tool Mapman for looking at pathways based on mapping to known organisms or GO

# Overview of an RNASeq Expt



Courtesy M. Rouard (Bioversity International)

# Description of Experiment

- The goal of this project was to identify genes that are differentially expressed in the two different banana varieties ( *k and m* ) in their response to bacterial infection
-

# Analysis – QC and mapping

- The sequences were trimmed for quality using Trimmomatic and aligned against the reference genome using Bowtie2/ Tophat. The reference genome and annotations were downloaded from Phytozome v10, *Musa acuminata* annotation version 304.
- Tophat now superseded by HISAT2

# Analysis – DESeq (Theory)

- Library normalisation – calculates a scaling factor to account for difference in library sizes or differences in library composition
- Dispersion – When calculating differential expression across groups we also need to know the within group variability. Needs a large no. of replicates or borrows information from similarly expressed genes.
- Statistical test – Generalised Linear Model, negative binomial distribution used for count data
- Multiple testing problem – can get good values by chance as we are doing many many test