

small RNA analysis - Nodulation dataset

Vikas Gupta, Dennis Berg Holt, Niraj Shah, Stig U. Andersen and Katharina Markmann

November 14, 2013

Contents

1. Introduction	2
2. Aim	2
3. Dataset	3
3.1 Genotypes	3
3.2 FastQC	4
3.3 Read Counts	5

1. Introduction

Lotus japonicus is considered as one of the model plant to study symbiotic interaction with rhizobium bacteria. In this study, we aim to reveal the molecular mechanism governing the successful symbiotic relation. Our main focus is understand the role of nodulation genes such as **Snf1**, **Sst1**, **Sen1** in the nodule and root tissue type. We have analysed the data using an open source tool **ShortRan**.

2. Aim

Dennis has created a set of hypothesis that would be potentially interesting to confirm in lights of publication strategy (Table 2).

Suggestion for comparisons	Question to answer
Gifu wt roots vs Gifu wt nods	Organogenesis/Infection
Gifu wt roots vs snf1 uninfected nodules	Organogenesis only
snf1 uninf. nods (spont.) vs snf1 inf. nods	Infection only
sst1 roots vs sst1 nods	Fix- infection/Organogenesis
Gifu wt nods vs sst1 nods	Fix-/early senescence
sen1 roots vs sen1 nods	Fix- infection/Organogenesis
Gifu wt nods vs sen1 nods	Fix-/early senescence
sst1 nods vs sen1 nods	Overlaps and differences in these mut nods with very similar phenotype

Table 1: Hypothesis to be tested

3. Dataset

3.1 Genotypes

A total of 8 genotypes were sequenced, each with two additional replicate (Table 1). Genxpro technology was used to do the sequencing of small RNA samples.

Library	Genotype	Tissue
KM45	Gifu wt	root
KM46	Gifu wt	root
KM47	Gifu wt	root
KM48	Gifu wt	nod
KM49	Gifu wt	nod
KM50	Gifu wt	nod
KM51	snf1	nod (spont.)
KM52	snf1	nod (spont.)
KM53	snf1	nod (spont.)
KM54	snf1	nod
KM55	snf1	nod
KM56	snf1	nod
KM57	sst1	root
KM58	sst1	root
KM59	sst1	root
KM60	sst1	nod
KM61	sst1	nod
KM62	sst1	nod
KM63	sen1	root
KM64	sen1	root
KM65	sen1	root
KM66	sen1	nod
KM67	sen1	nod
KM68	sen1	nod

Table 2: Genotype sample sequenced in the study

3..2 FastQC

We ran FastQC on the raw fastq files to check for adaptor and there were no signatures were found. We also looked at the sequencing quality on 3' end of the read and quality seems to be quite high hence reads were trimmed. A summary from the library KM-63 is shown below.



Basic Statistics

Measure	Value
Filename	KM63-miRvana.miRNA.fastq.noDuplicates
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	13049137
Filtered Sequences	0
Sequence length	16-38
%GC	41

Figure 1: KM-63 Raw read abundance

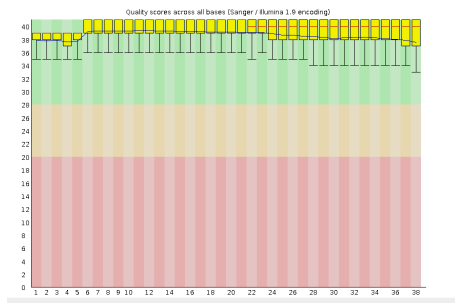


Figure 2: KM 63 Sequencing Quality

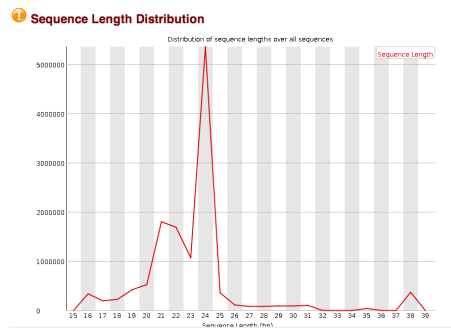


Figure 3: KM 63 Size distribution

3.3 Read Counts

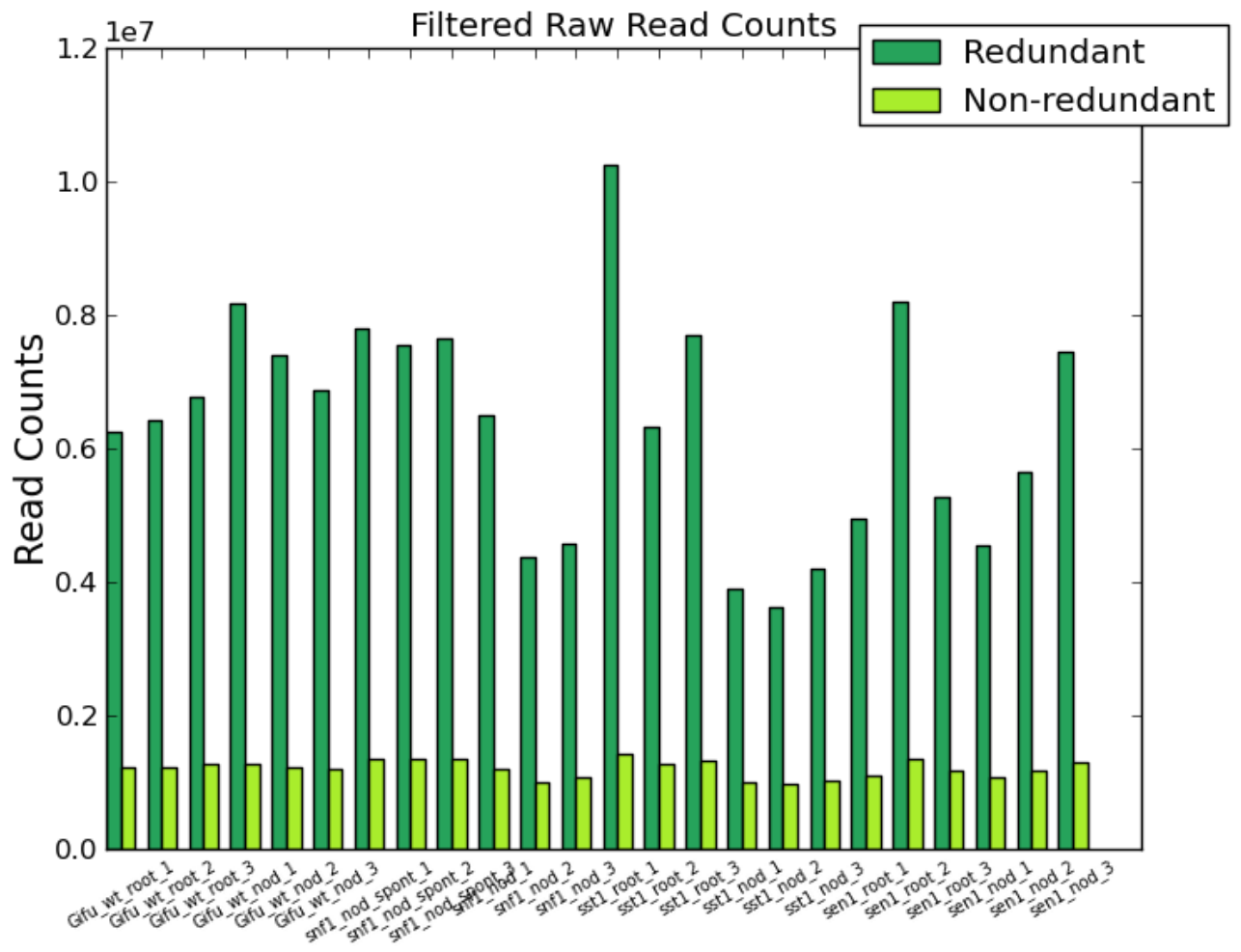


Figure 4: Raw read abundance