

## 1 Answers

### 1.1 1. The Pan Genome

Q1: b

Q2: a

### 1.2 2. Preparing the data

Q3: b

Q4: c

### 1.3 3. Performing QC on your data

Q5: If the data is bad going in, the results will be bad coming out.

Q6: Most commonly there is contamination, or the genomes are too fragmented.

### 1.4 4. Running Roary

Q7: c

Q8: b

### 1.5 5. Exploring the results

Q9: b

Q10: c

Q11: c

### 1.6 6. Visualising the results with phandango

Q12: lytN

Q13: No (it is only present in sample 1 and 2)