

QC

Answers to questions

- 1. Which positions in the reads have a median phred-score above 28 (very good quality calls) in each sample?**

Sample	Read 1	Read 2
HG00097	2-101	2-101
HG00100	1-107	1-107
HG00101	1-101	1-101

- 2. Red x-marks indicate failed QC tests. Have any of the samples failed the Per Base Sequence Quality test?**

HG00097 Read 1 and Read 2

- 3. Why?**

This module will raise a failure if the lower quartile for any base is less than 5 or if the median for any base is less than 20.

In sample HG00097 the first base have mead quality of 15, which is less than the threshold for raising a failure. An option could be to trim off the first base, but since the quality values are taken into account in variant calling it is OK to keep this one base with lower quality in the reads.

- 4. How long are the reads?**

Sample	Read 1	Read 2
HG00097	101	101
HG00100	100-108	100-108
HG00101	101	101

- 5. Orange exclamation marks indicate QC warnings. Have any of the samples raised warnings in the Sequence Length Distribution test?**

Yes, HG100100

- 6. Why?**

This module will raise a warning if all sequences are not the same length.
For HG100100 most reads are 100 bp, but a subset of the reads are 108 bp.