1. Bioinformatics Analysis

1.1 Data Quality Control

|  |
| --- |
| **Data Quality Control** |
| Distribution of Sequencing Error Rate |
| Distribution of A/T/G/C Base Content |
| Data Filtering: Removal of reads with adaptor contamination, uncertain nucleotides(N)with > 10% content and nucleotides of low quality |
| Statistic Summary of Data Quality: Amount of data output, Error rate, Q20%, Q30% and GC content, etc |

1.2 Mapping Only Analysis

|  |
| --- |
| **Only Mapping Analysis** |
| Distribution of Sequencing Error Rate |
| Distribution of A/T/G/C Base Content |
| Data Filtering: Removal of reads with adaptor contamination, uncertain nucleotides(N)with > 10% content and nucleotides of low quality |
| Statistic Summary of Data Quality: Amount of data output, Error rate, Q20%, Q30% and GC content, etc |
| Mapping Clean Reads to Reference Genome |

Note: there might be multiple Data Quality Control presented in the tables which does not mean repetition of quality control; regularly data quality control is considered to be one-time required.