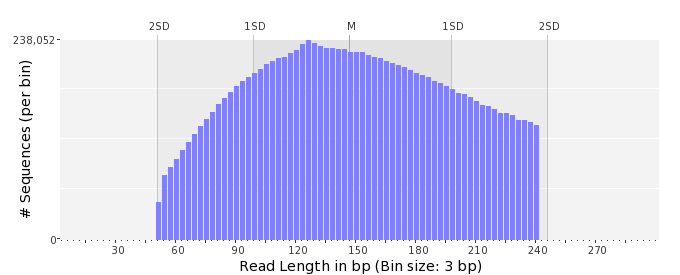
**F3**

**Input Information**

|  |  |
| --- | --- |
| Input file(s): | **F\_3.assembled.fastq** |
| Input format(s): | **FASTQ** |
| # Sequences: | **11,589,014** |
| Total bases: | **1,701,765,604** |

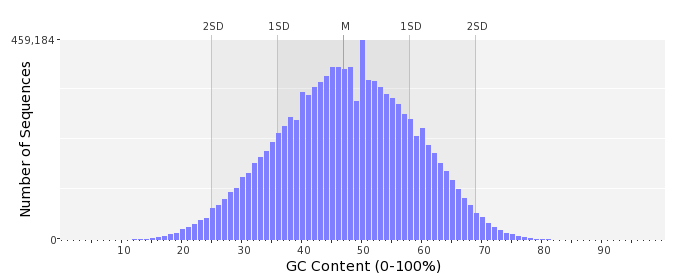
**Length Distribution**

|  |  |
| --- | --- |
| Mean sequence length: | **146.84 ± 49.53 bp** |
| Minimum length: | **50 bp** |
| Maximum length: | **240 bp** |
| Length range: | **191 bp** |
| Mode length: | **124 bp with 79,600 sequences** |

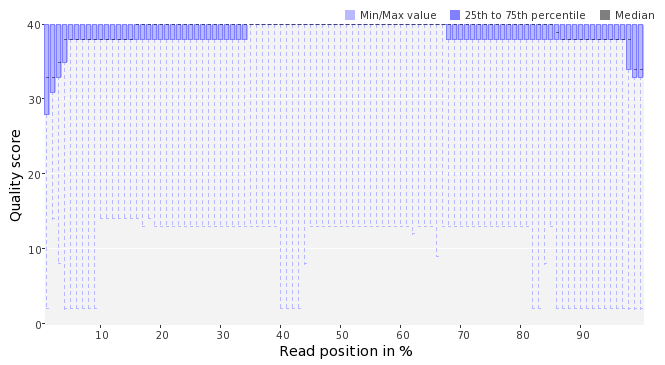
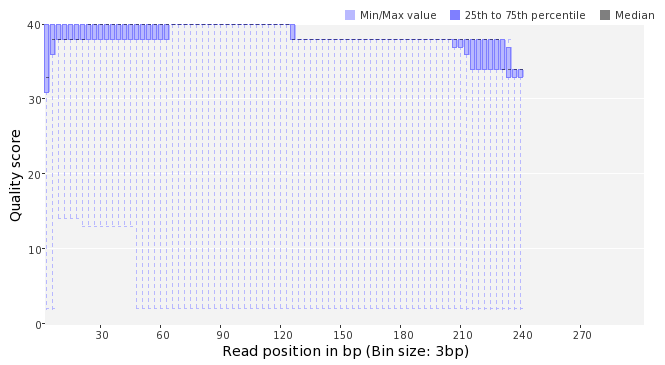
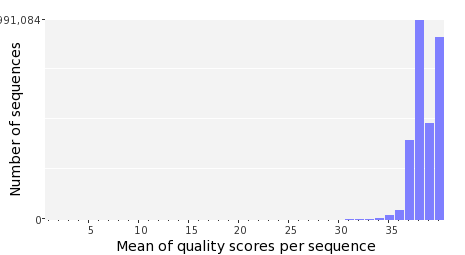


**GC Content Distribution**

|  |  |
| --- | --- |
| Mean GC content: | **47.23 ± 11.25 %** |
| Minimum GC content: | **0 %** |
| Maximum GC content: | **100 %** |
| GC content range: | **101 %** |
| Mode GC content: | **50 % with 459,181 sequences** |

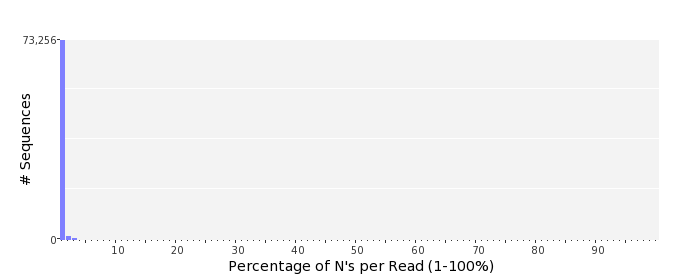


**Base Quality Distribution**

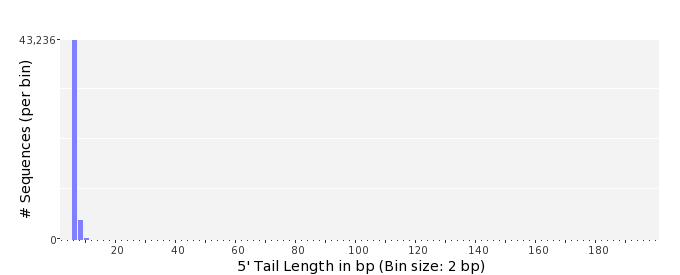
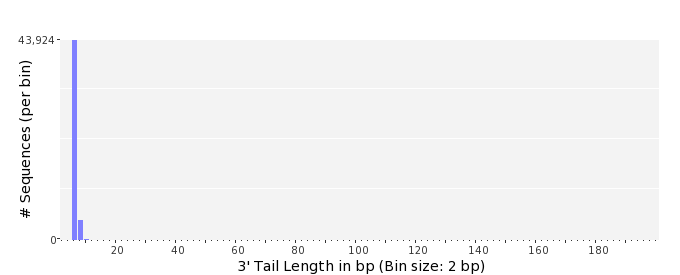
**Occurence of N**

|  |  |
| --- | --- |
| Sequences with N: | **75,507  (0.65 %)** |
| Max percentage of Ns per sequence: | **5 %** |



**Poly-A/T Tails**

|  |  |  |
| --- | --- | --- |
|  | **5'-end** | **3'-end** |
| Sequences with tail: | **48,017  (0.41 %)** | **48,700  (0.42 %)** |
| Maximum tail length: | **125** | **125** |

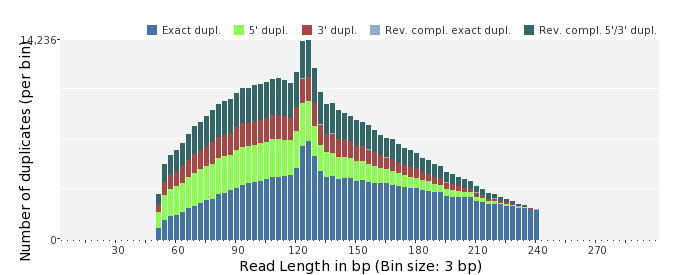
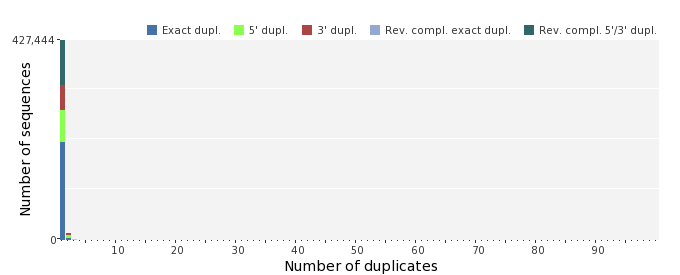
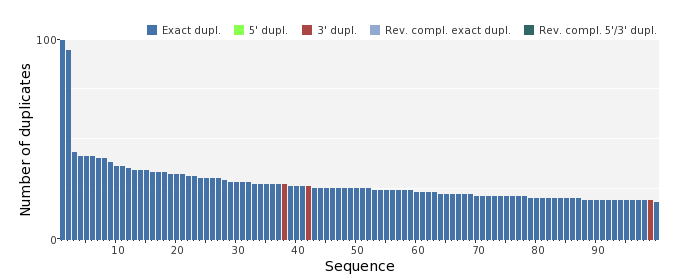
**Tag Sequence Check**

|  |  |  |
| --- | --- | --- |
|  | **5'-end** | **3'-end** |
| Probability of tag sequence: | **0 %** | **0 %** |
| GSMIDs or RLMIDs: | **none** |  |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | ... |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | 1 |  |  |  | 5 |  |  |  |  | 10 |  |  |  |  | 15 |  |  |  |  | 20 |  | 20 |  |  |  |  | 15 |  |  |  |  | 10 |  |  |  |  | 5 |  |  |  | 1 |
|  | **Position from Sequence Ends** | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

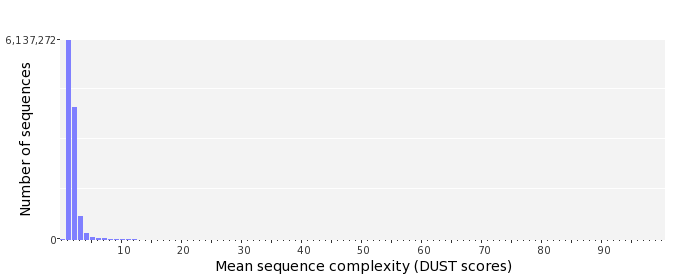
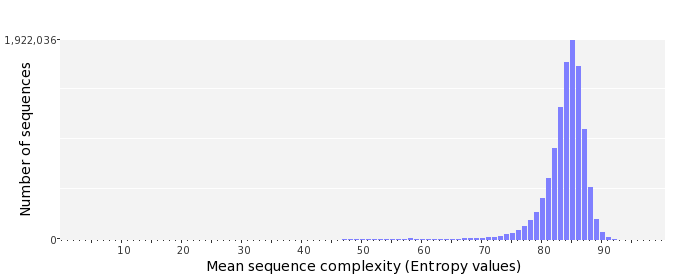
**Sequence Duplication**

|  |  |  |
| --- | --- | --- |
|  | **# Sequences** | **Max duplicates** |
| Exact duplicates: | **225,363  (1.94 %)** | **100** |
| Exact duplicates with reverse complements: | **562  (0.00 %)** | **1** |
| 5' duplicates | **93,169  (0.80 %)** | **13** |
| 3' duplicates | **62,107  (0.54 %)** | **28** |
| 5'/3' duplicates with reverse complements | **97,593  (0.84 %)** | **3** |
| Total: | **478,794  (4.13 %)** | **-** |

**Sequence Complexity**

|  |  |  |
| --- | --- | --- |
|  | **Value** | **Sequence** |
| Minimum DUST score: | **0** | CACTTGCTTCAAGTGATTCAGCAACAGCCTTCATAGATGTGCCAGCGGGTTGAATACCCT GCGCTCTGGCATCACGGTAAGCCGATAGTTCCGACTCCCATTT |
| Maximum DUST score: | **99** | TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT TTTTTTTTTTTTTTTTTTTTTTTT |
| Minimum Entropy value: | **0** | TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT TTTTT |
| Maximum Entropy value: | **97** | GGCATCAAGGTAAACGCTAGAATGCAGTCCTCTGACCCGGATGTTTACGC |

**Dinucleotide Odds Ratios**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **AA/TT** | **AC/GT** | **AG/CT** | **AT** | **CA/TG** | **CC/GG** | **CG** | **GA/TC** | **GC** | **TA** |
| Odds ratio | **1.1443** | **0.8496** | **0.9637** | **1.0410** | **1.0849** | **1.0073** | **0.9375** | **1.0390** | **1.1391** | **0.7070** |

