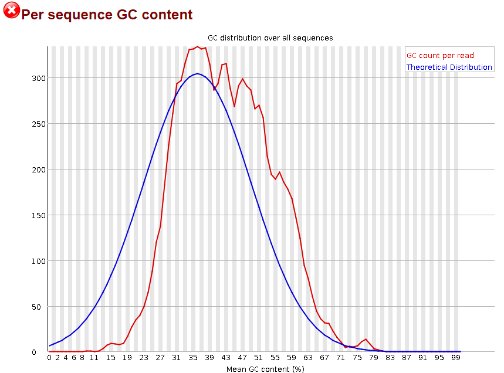
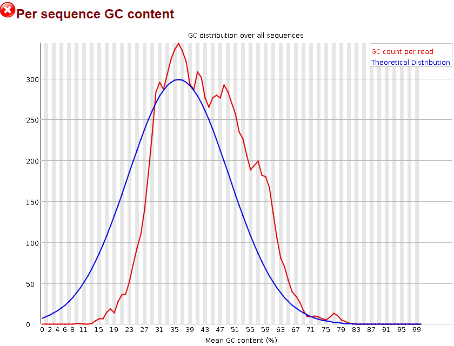
chr2\_CM000492.1\_1000000-1100000

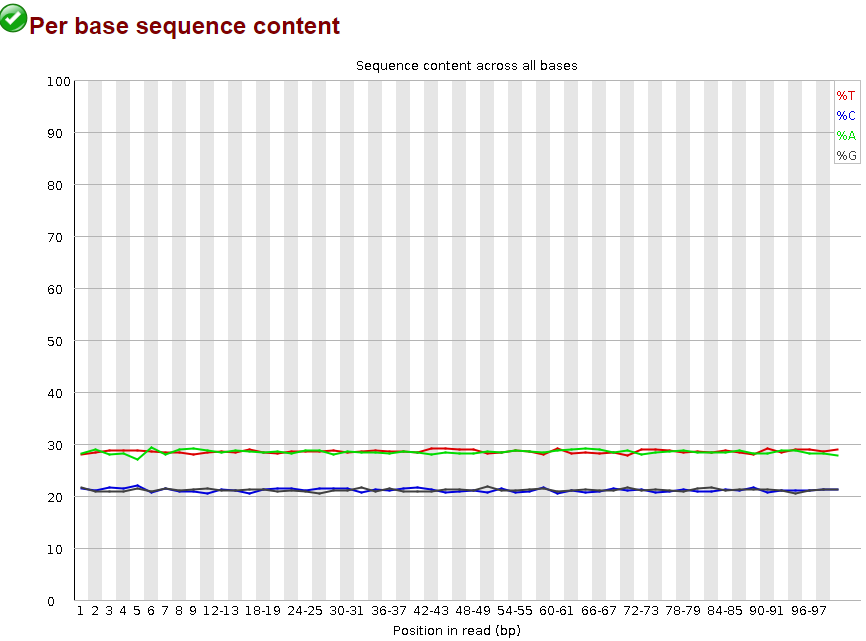
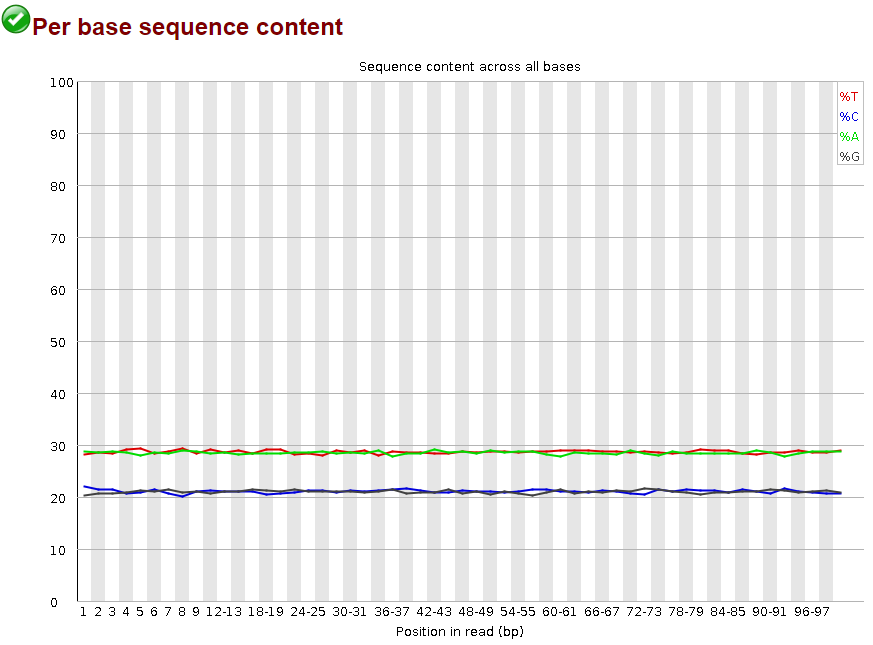
Python script

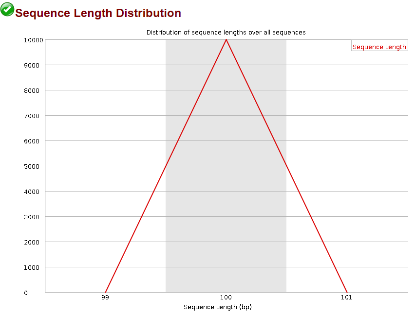
chr2\_CM000492.1\_1000000-1100000

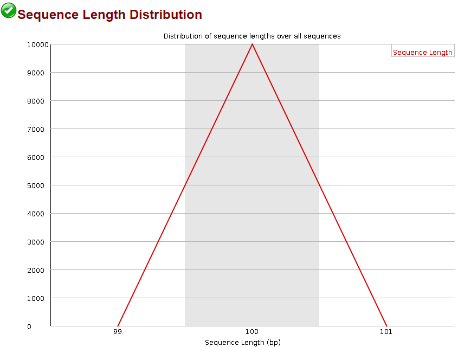
ART

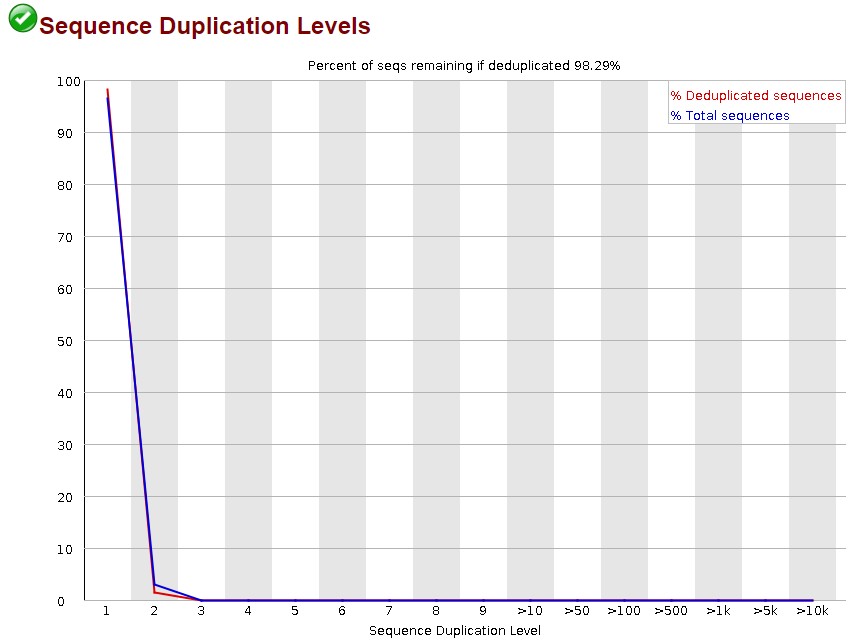


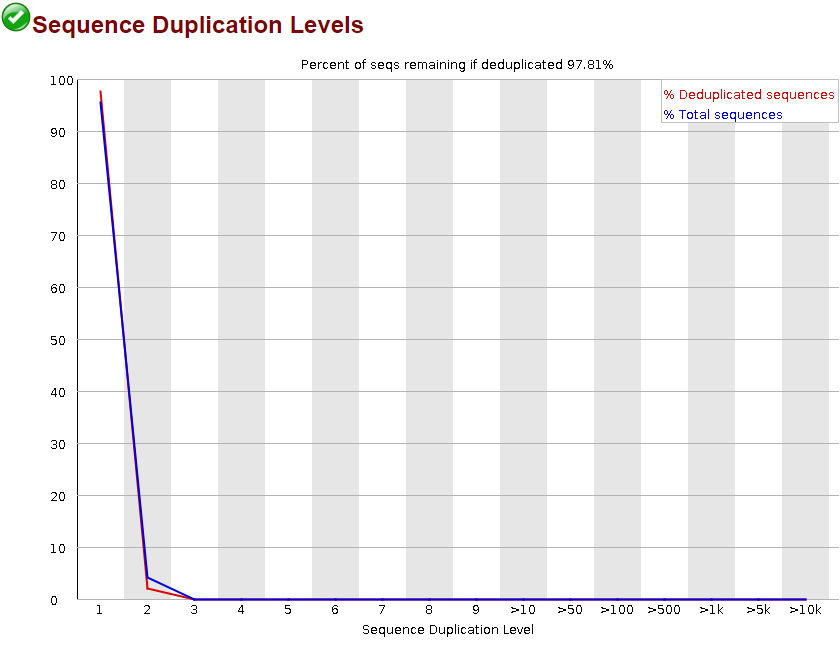




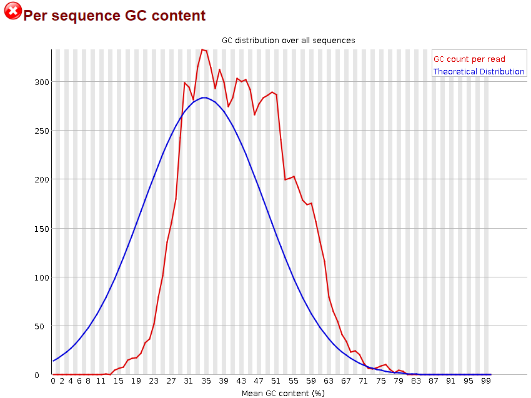
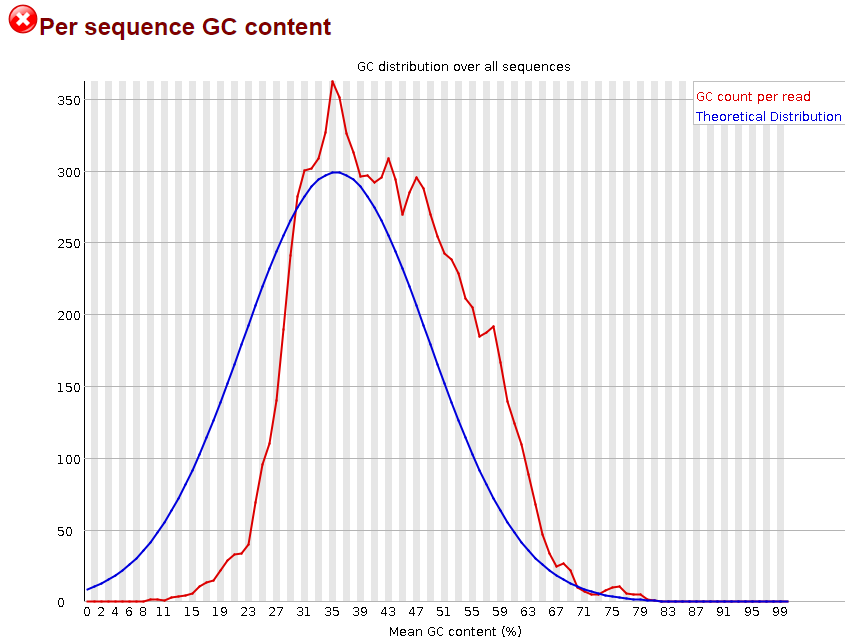
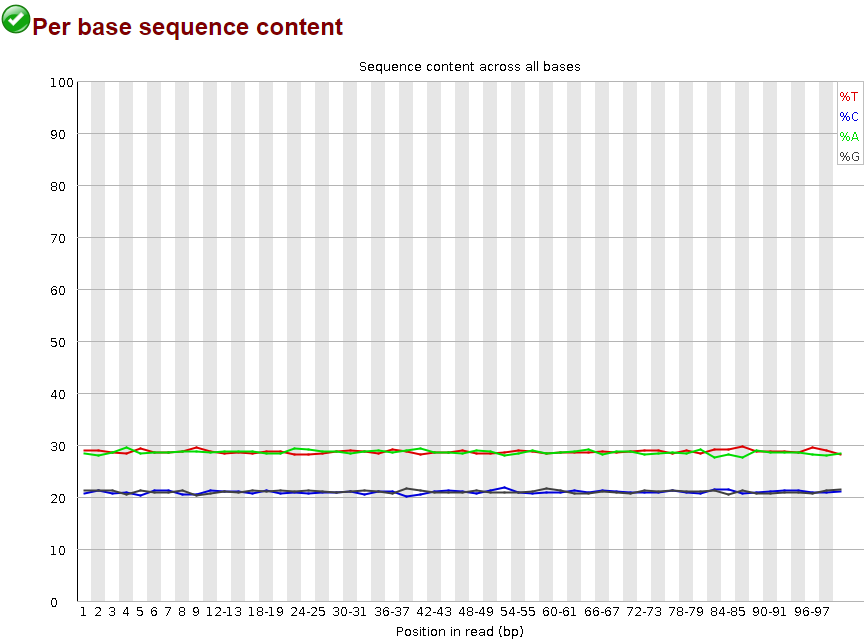
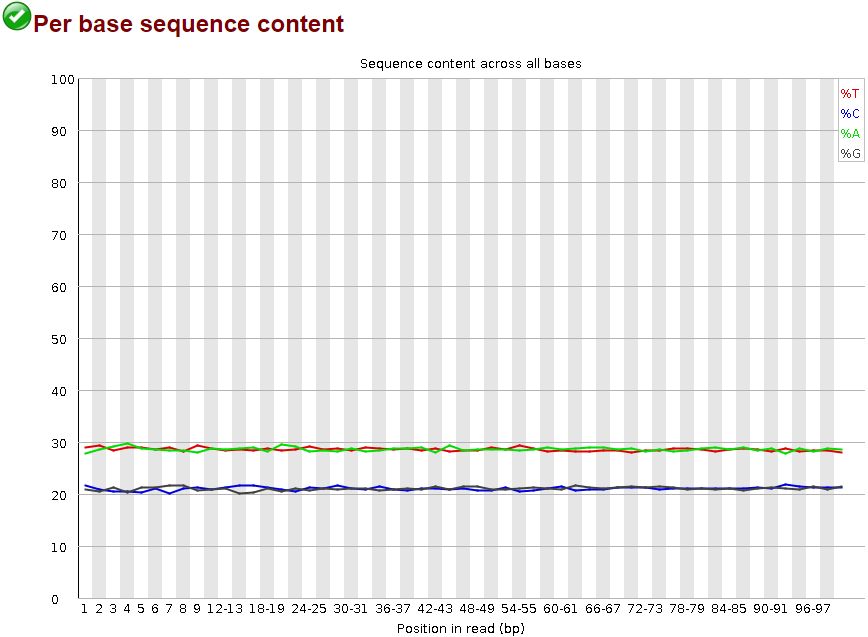
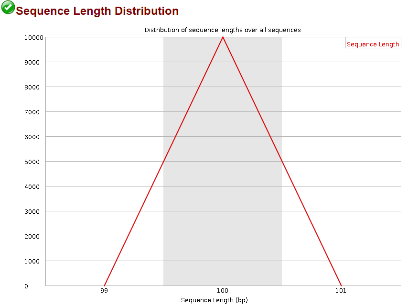
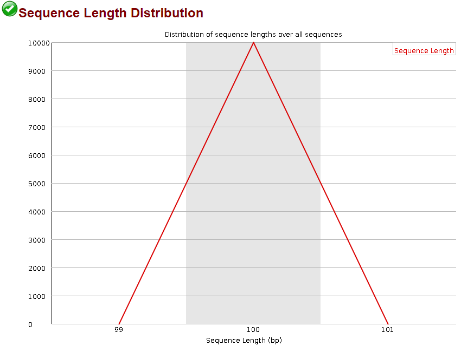
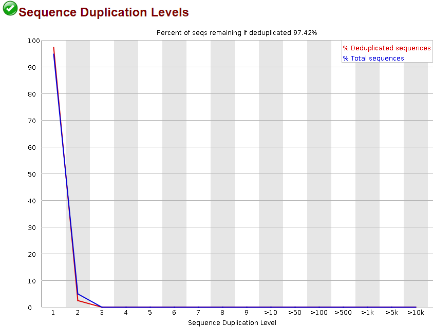
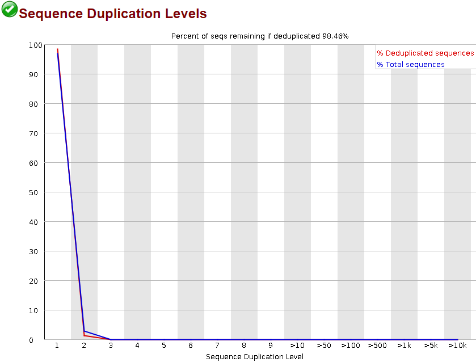








Per sequence GC content is about the same in the ART and my simulation results with the same maximum as the theoretical distribution, but both results are shifted to the right a bit.The per base sequence content is the same in both results. The sequence length distribution is the same for both at every sequence at 100 bp. The sequence duplication is the same for both results. There was no N content for either results as well as no overrepresented sequences or Kmer content.



Per sequence GC content is about the same in the ART and my simulation results, with my simulation having the same maximum as the theoretical distribution and the ART results, but with more distribution at a higher mean GC content %. The per base sequence content is the same in both results. The sequence length distribution is the same for both at every sequence at 100 bp. The sequence duplication is the same for both results. There was no N content for either results as well as no overrepresented sequences or Kmer content.

chr2\_CM000463.1\_1000000-1100000

ART

chr2\_CM000463.1\_1000000-1100000

Python script