EDGE Version: 1.0

Project Start: 2014 Oct 22 17:56:49

Output Dir: /opt/apps/edge/edge_ui/EDGE_output/Ecoli_10x

Reference: reference.fasta

Features:

QC	Host Removal	Assembly	Annotation	Taxonomy Classification	Primer Design
V	V	V	V	V	V

Inputs:

Inputs	Reads	Bases	Avg_Len
Ecoli_10x.1.fastq	165563	25000013	151
Ecoli_10x.2.fastq	165563	25000013	151

QC stats

Before Trimming
Reads #: 331126
Total bases: 50000026
Reads Length: 151.00

After Trimming
Reads #: 331026 (99.97 %)
Total bases: 49984926 (99.97 %)
Mean Reads Length: 151.00
  Paired Reads #: 330930 (99.97 %)
  Paired total bases: 49970430 (99.97 %)
  Unpaired Reads #: 96 (0.03 %)
  Unpaired total bases: 14496 (0.03 %)

Discarded reads #: 100 (0.03 %)
Trimmed bases: 15100 (0.03 %)
  Reads Filtered by length cutoff (50 bp): 0 (0.00 %)
  Bases Filtered by length cutoff: 0 (0.00 %)
  Reads Filtered by continuous base "N" (2): 0 (0.00 %)
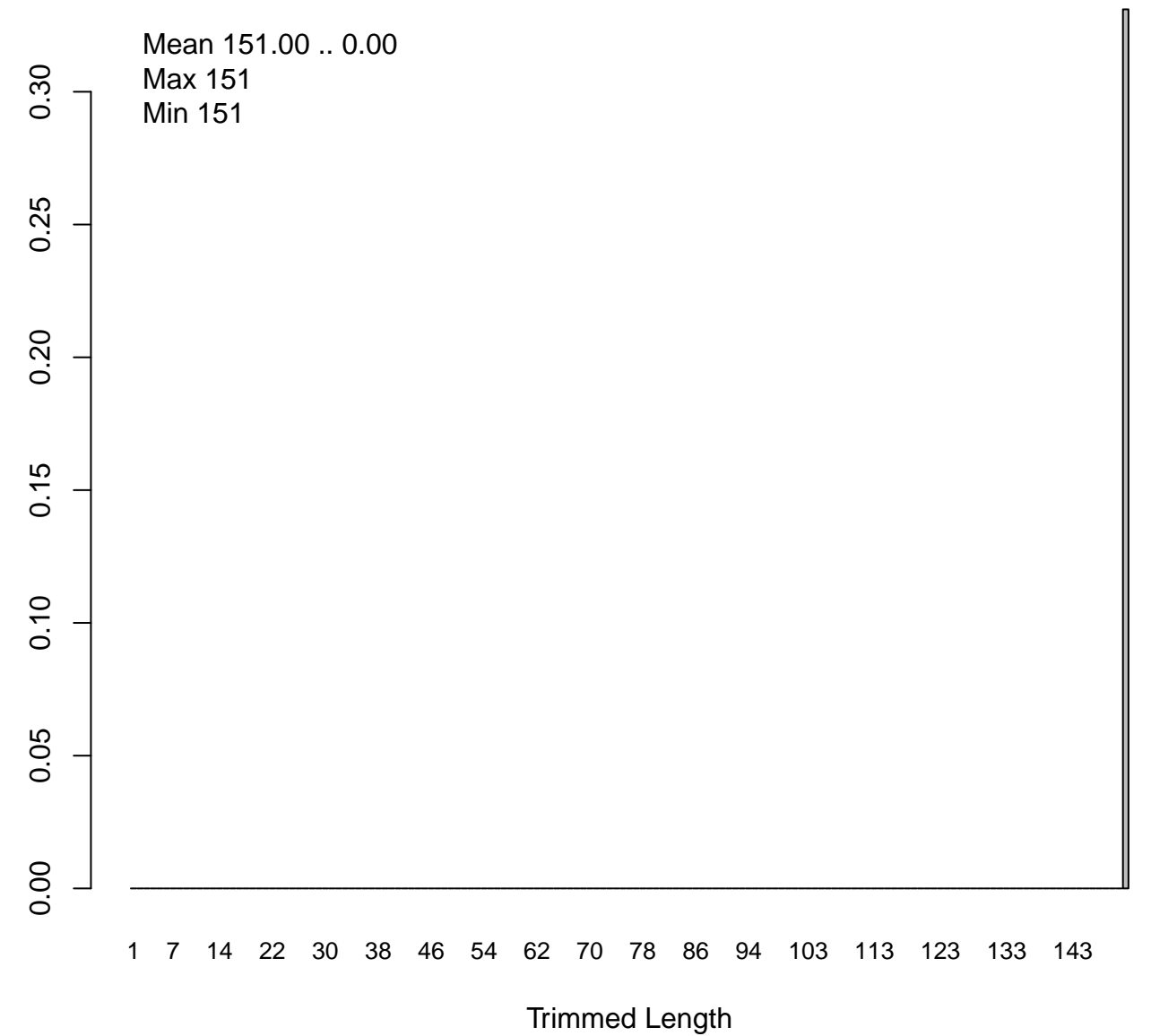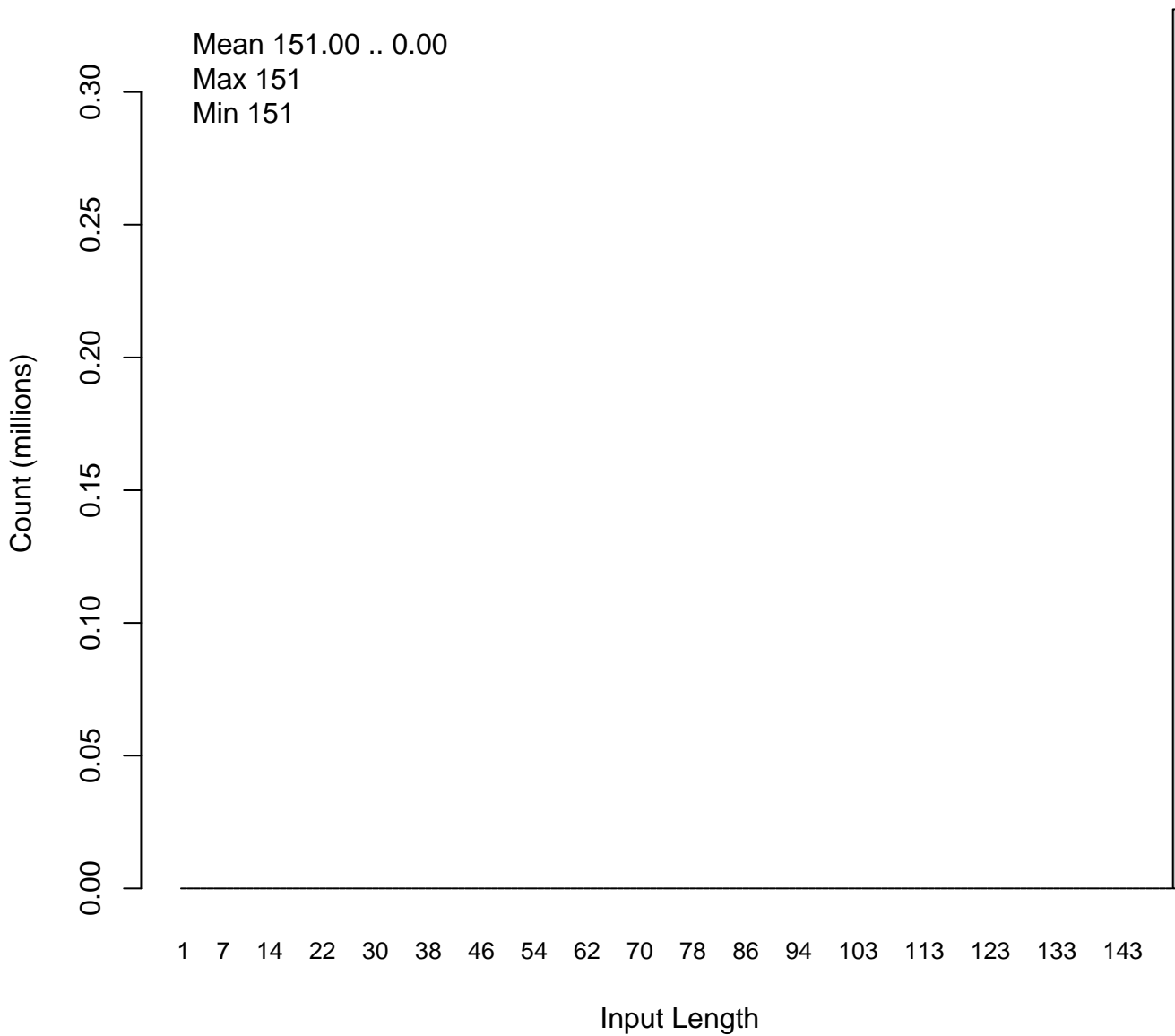 Bases Filtered by continuous base "N": 0 (0.00 %)
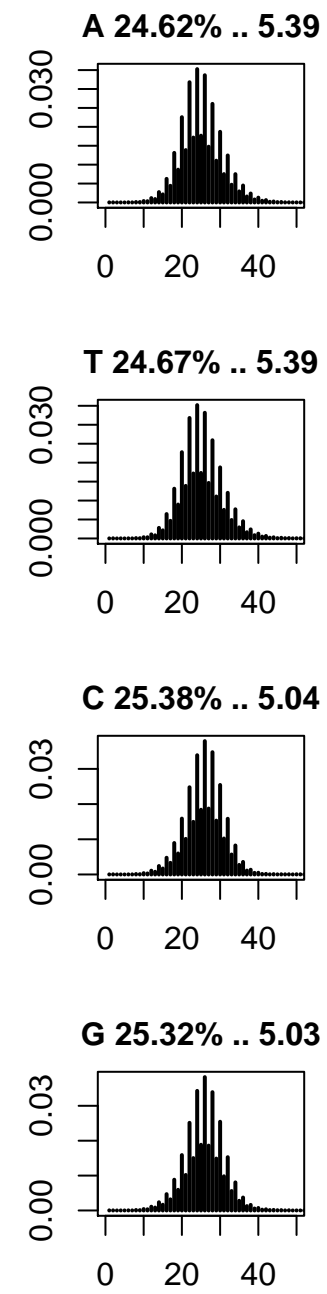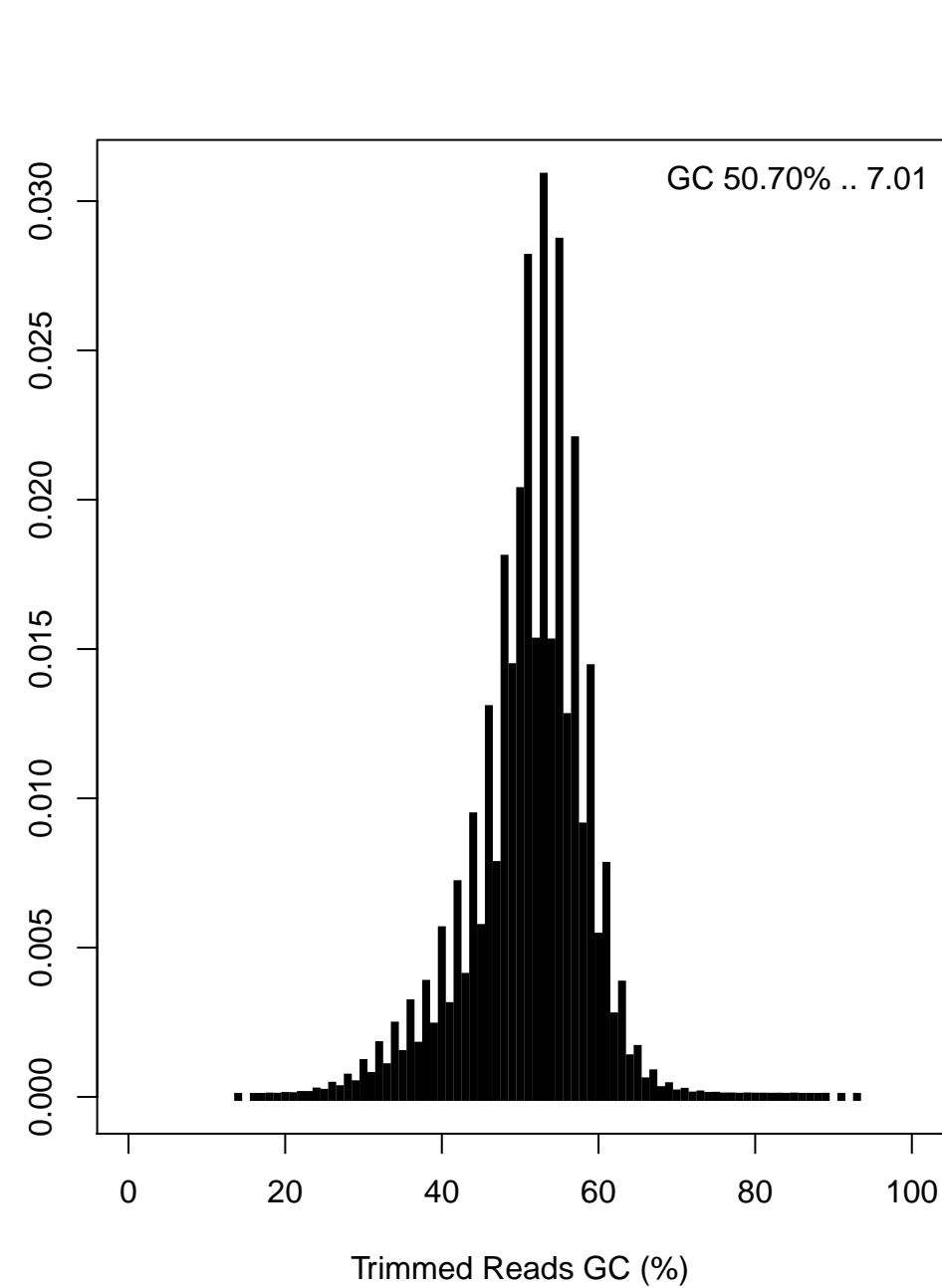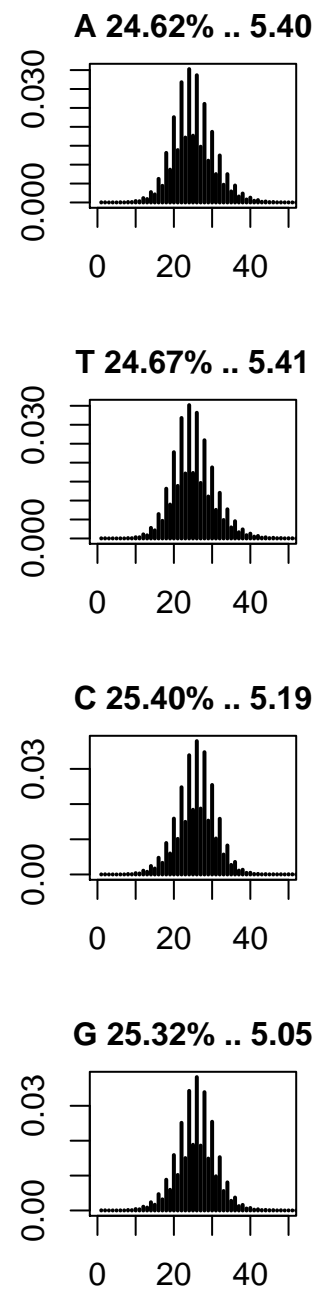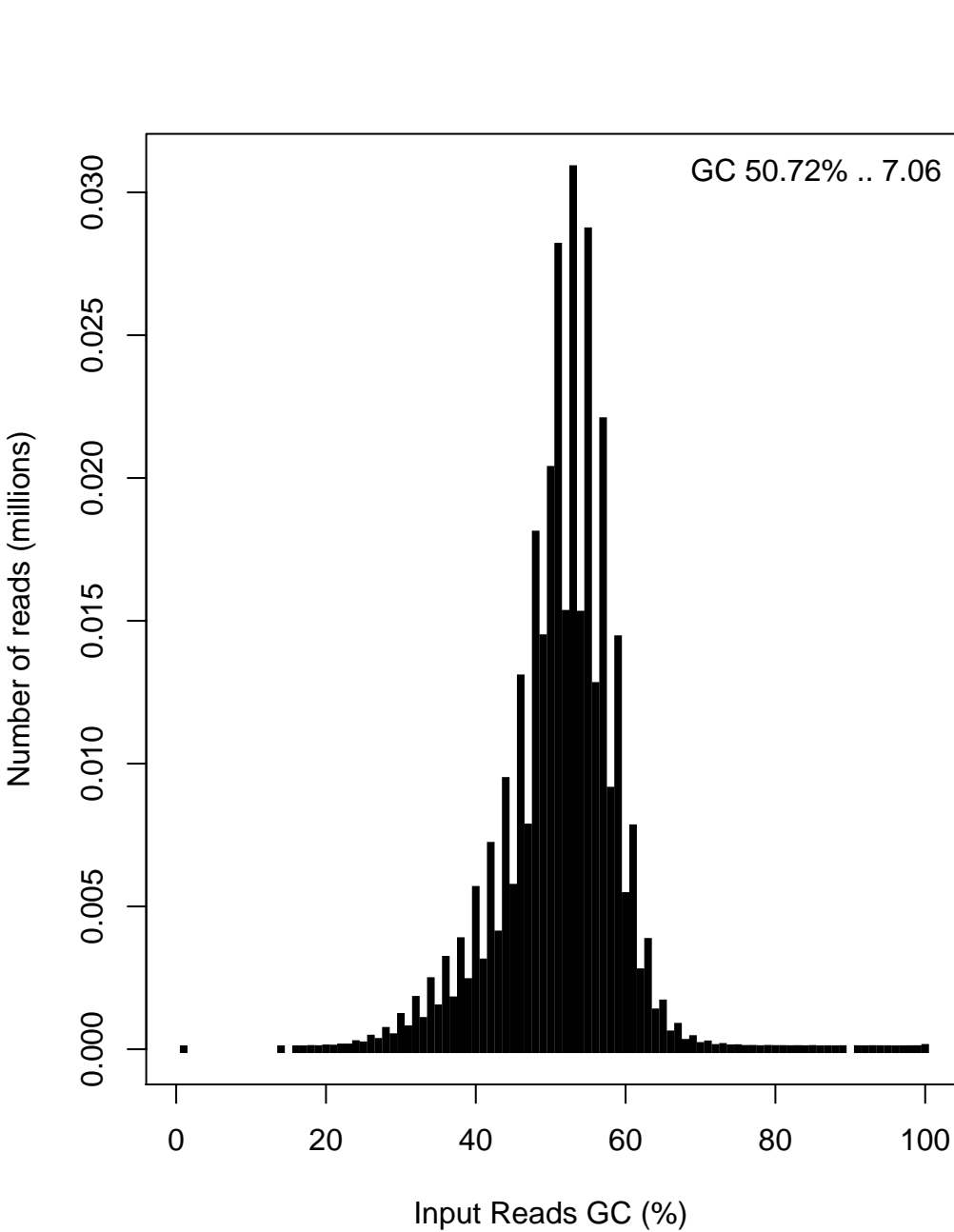 Reads Filtered by low complexity ratio (0.8): 100 (0.03 %)
  Bases Filtered by low complexity ratio: 15100 (0.03 %)
  Reads Trimmed by quality (5.0): 0 (0.00 %)
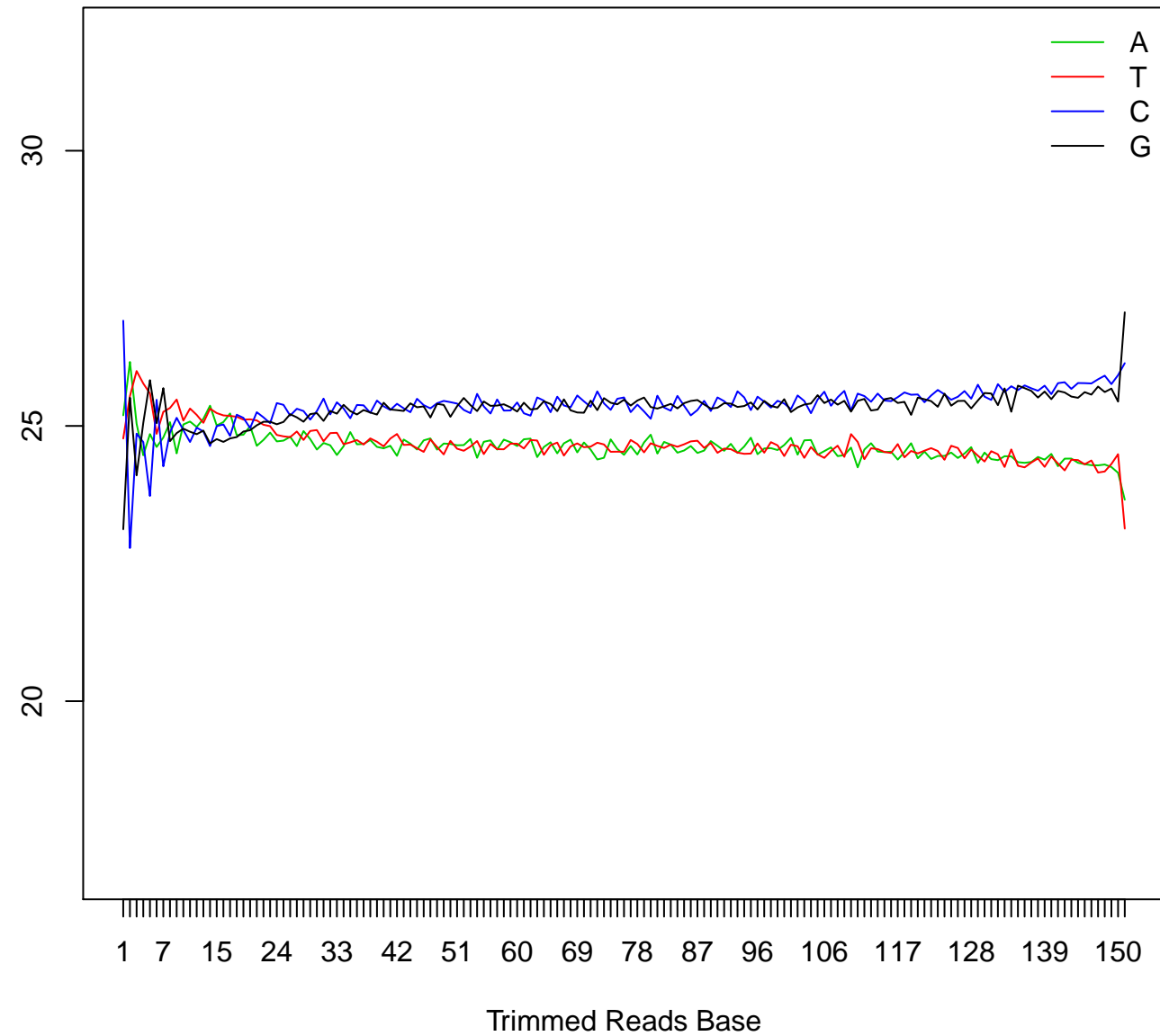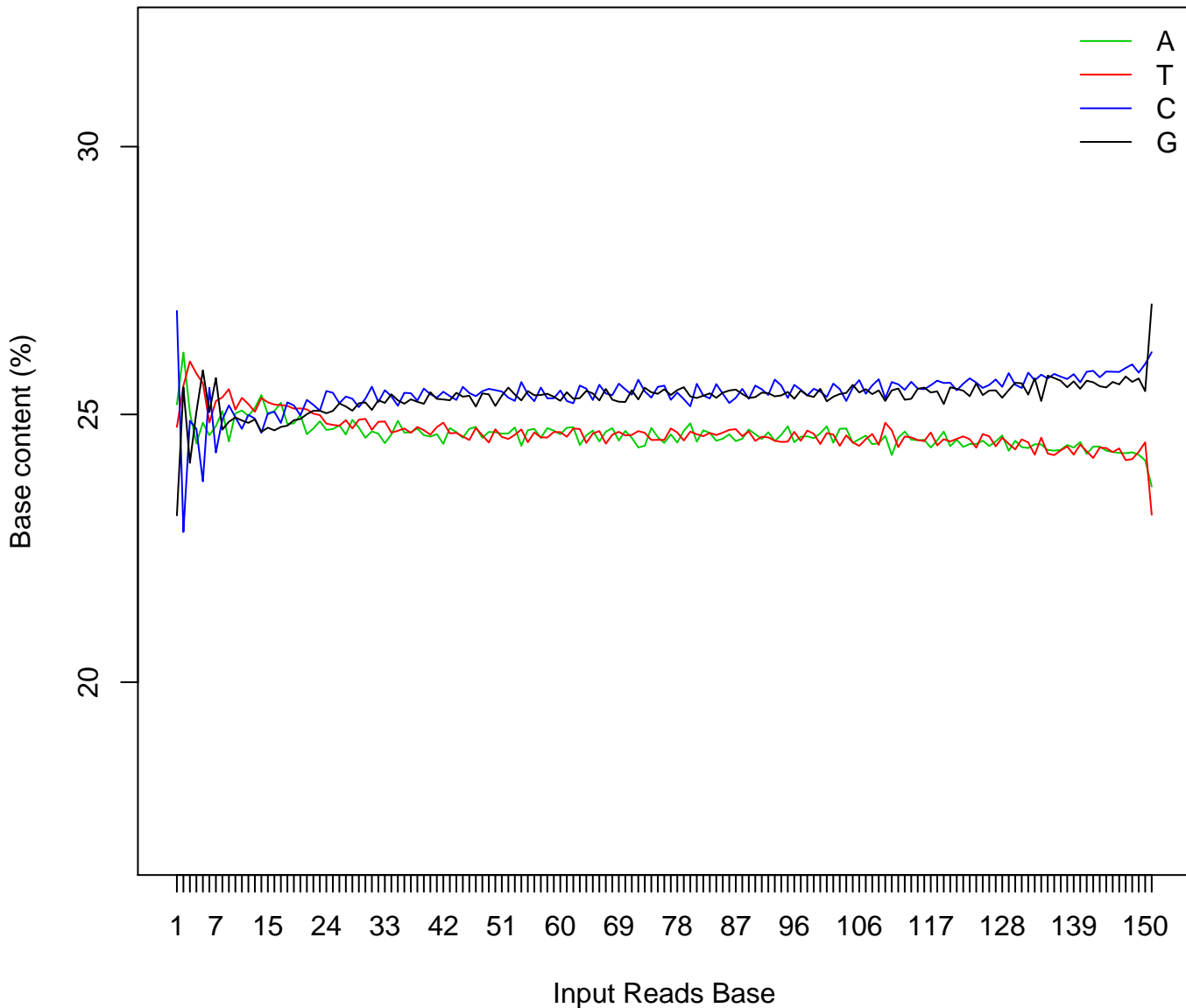 Bases Trimmed by quality: 0 (0.00 %)

Reads Length Histogram

Mean 151.00 .. 0.00
Max 151
Min 151

Mean 151.00 .. 0.00
Max 151
Min 151

Count (millions)

0.30
0.25
0.20
0.15
0.10
0.05
0.00

1   7   14   22   30   38   46   54   62   70   78   86   94   103   113   123   133   143

Input Length

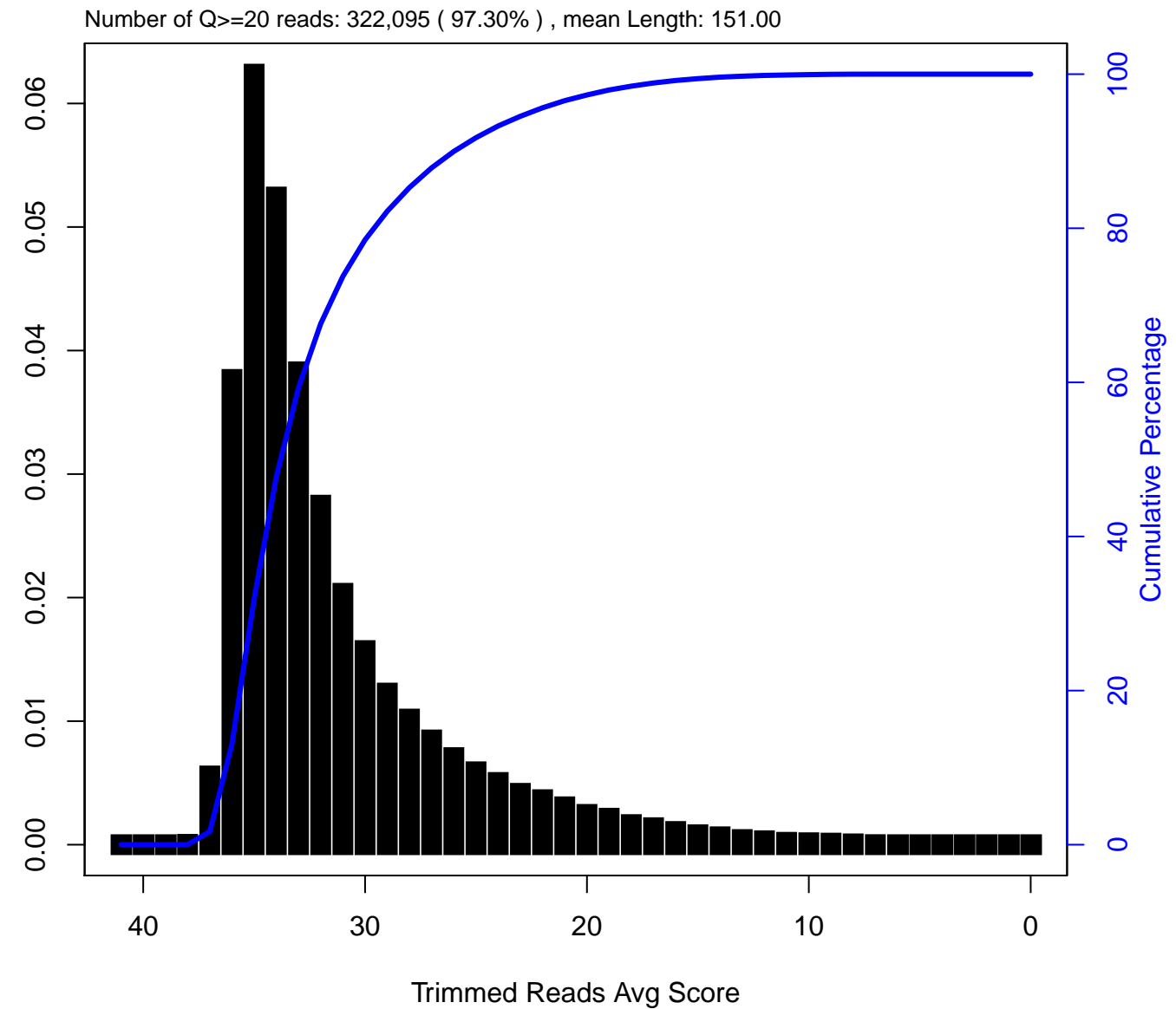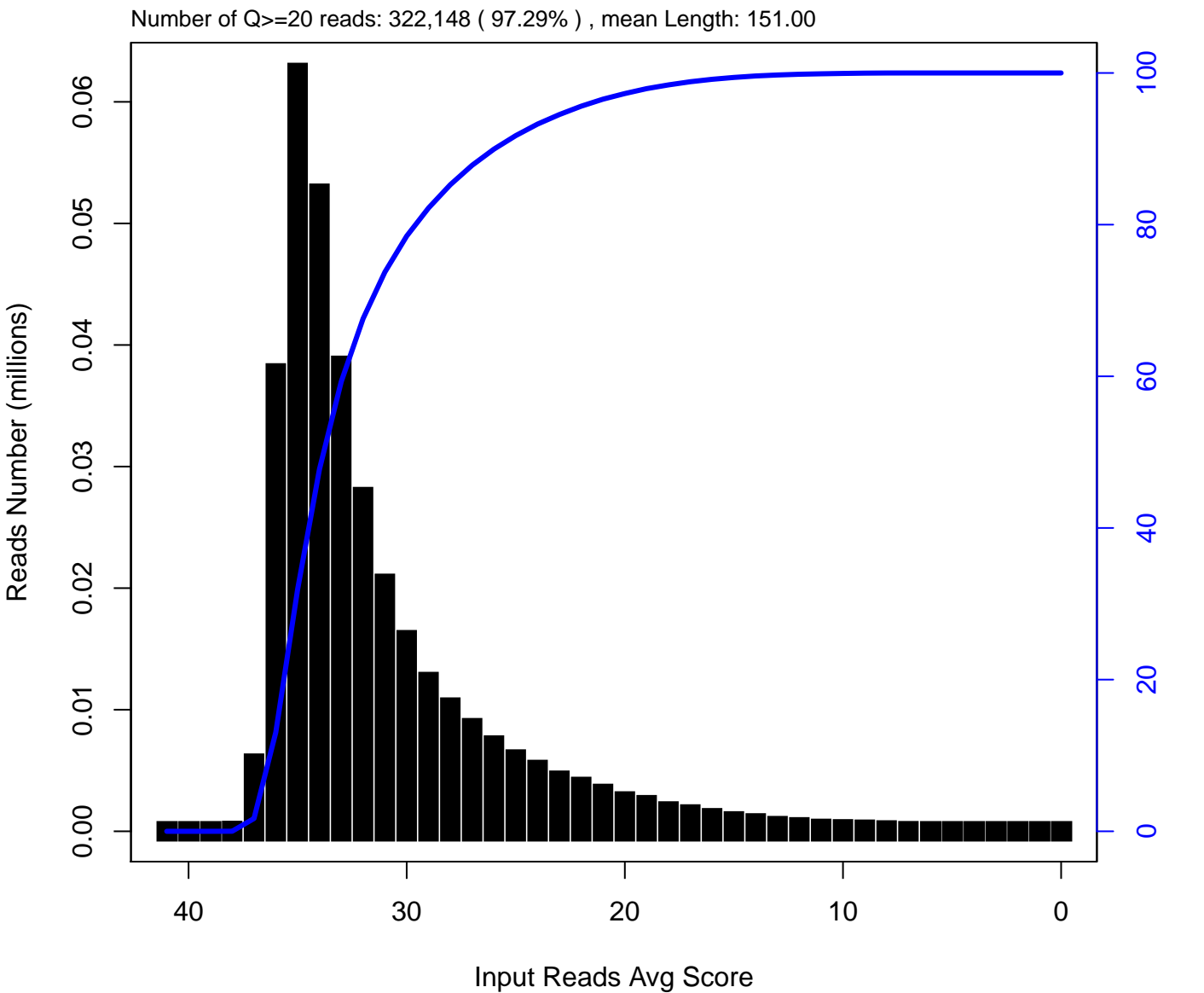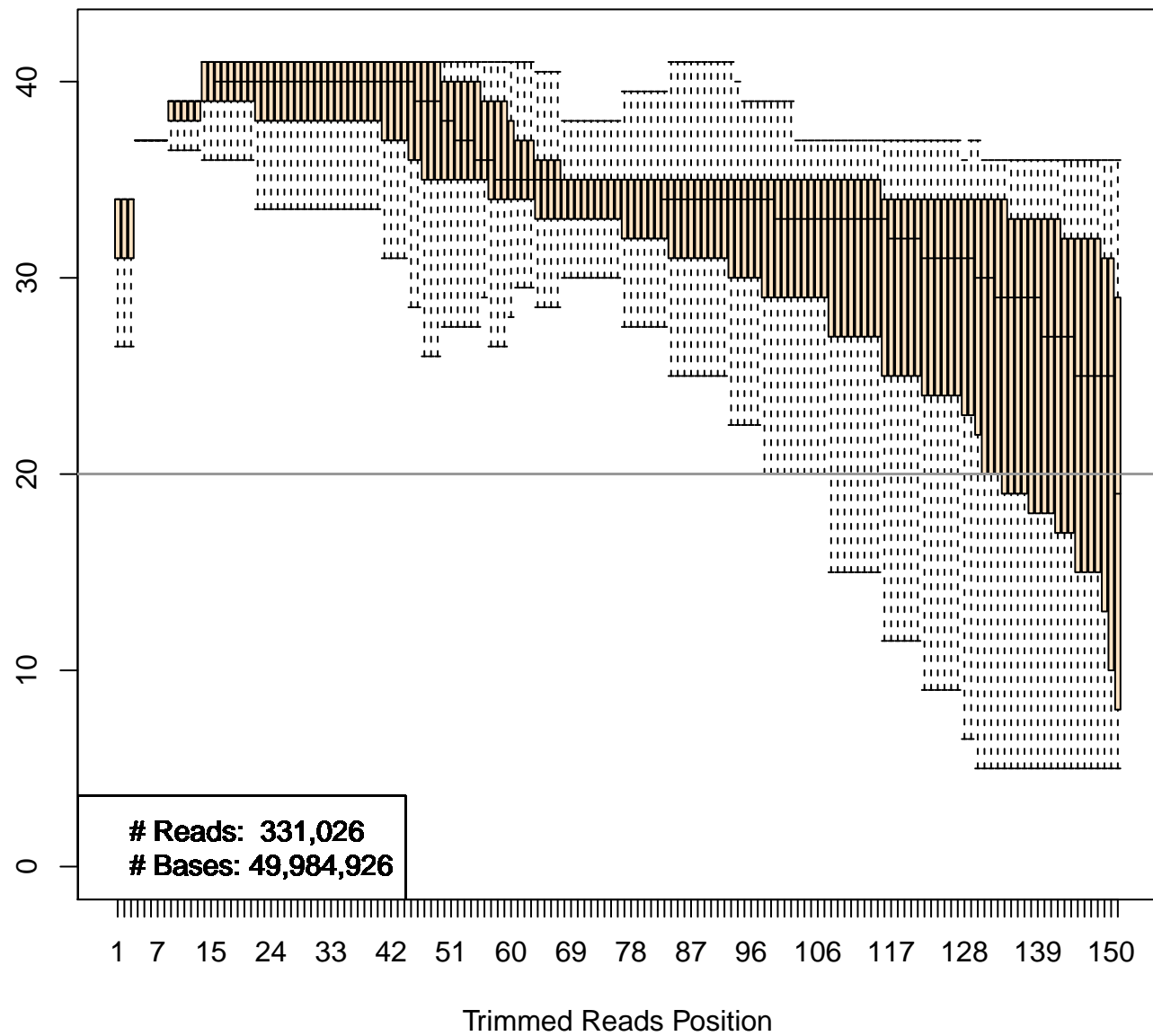1   7   14   22   30   38   46   54   62   70   78   86   94   103   113   123   133   143

Trimmed Length

Nucleotide Content Per Cycle

Reads Average Quality Histogram

Number of Q>=20 reads: 322,148 (97.29%) , mean Length: 151.00

Number of Q>=20 reads: 322,095 (97.30%) , mean Length: 151.00

Quality Boxplot Per Cycle
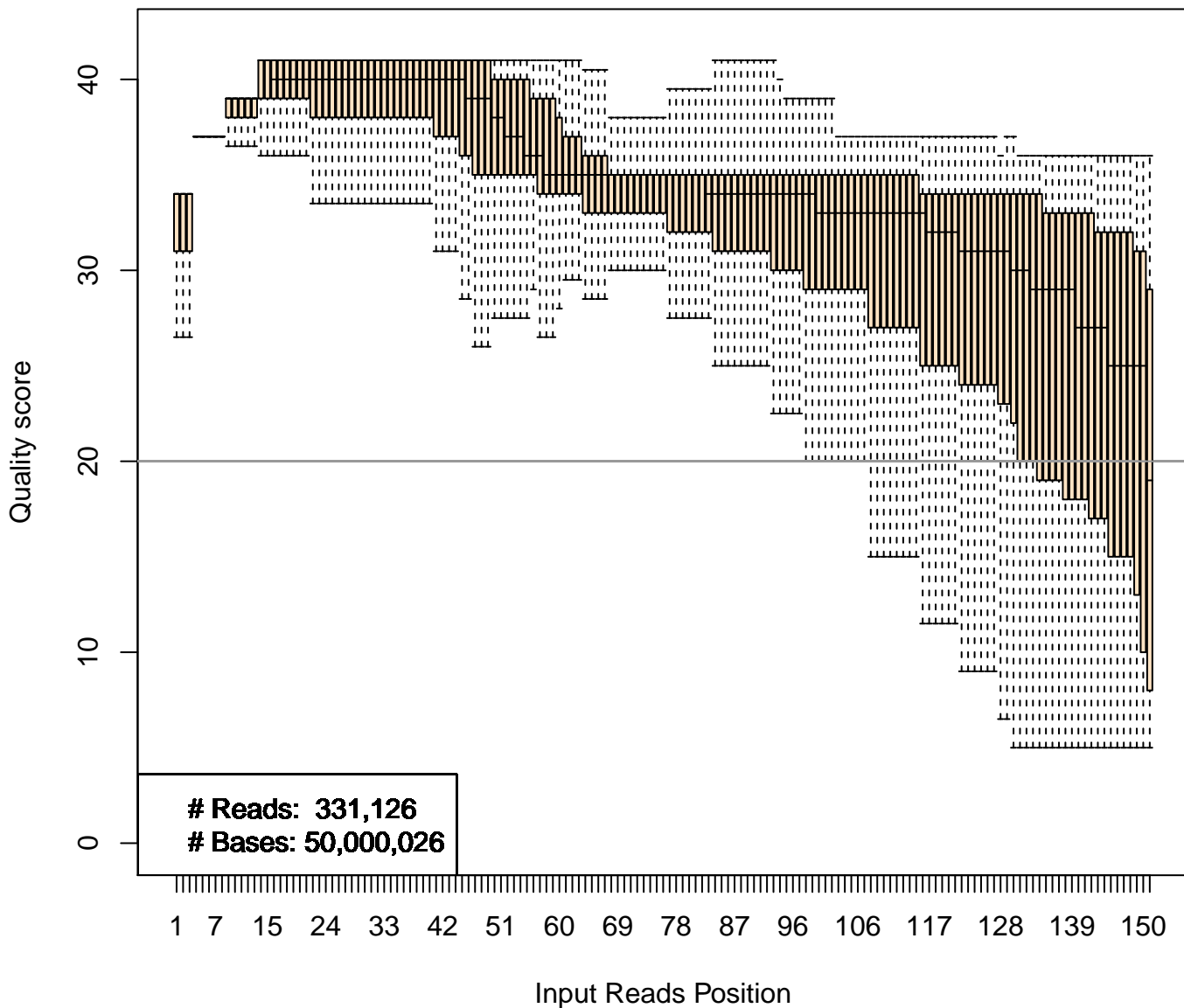
Quality 3D plot. (Position vs. Score vs. Frequency)
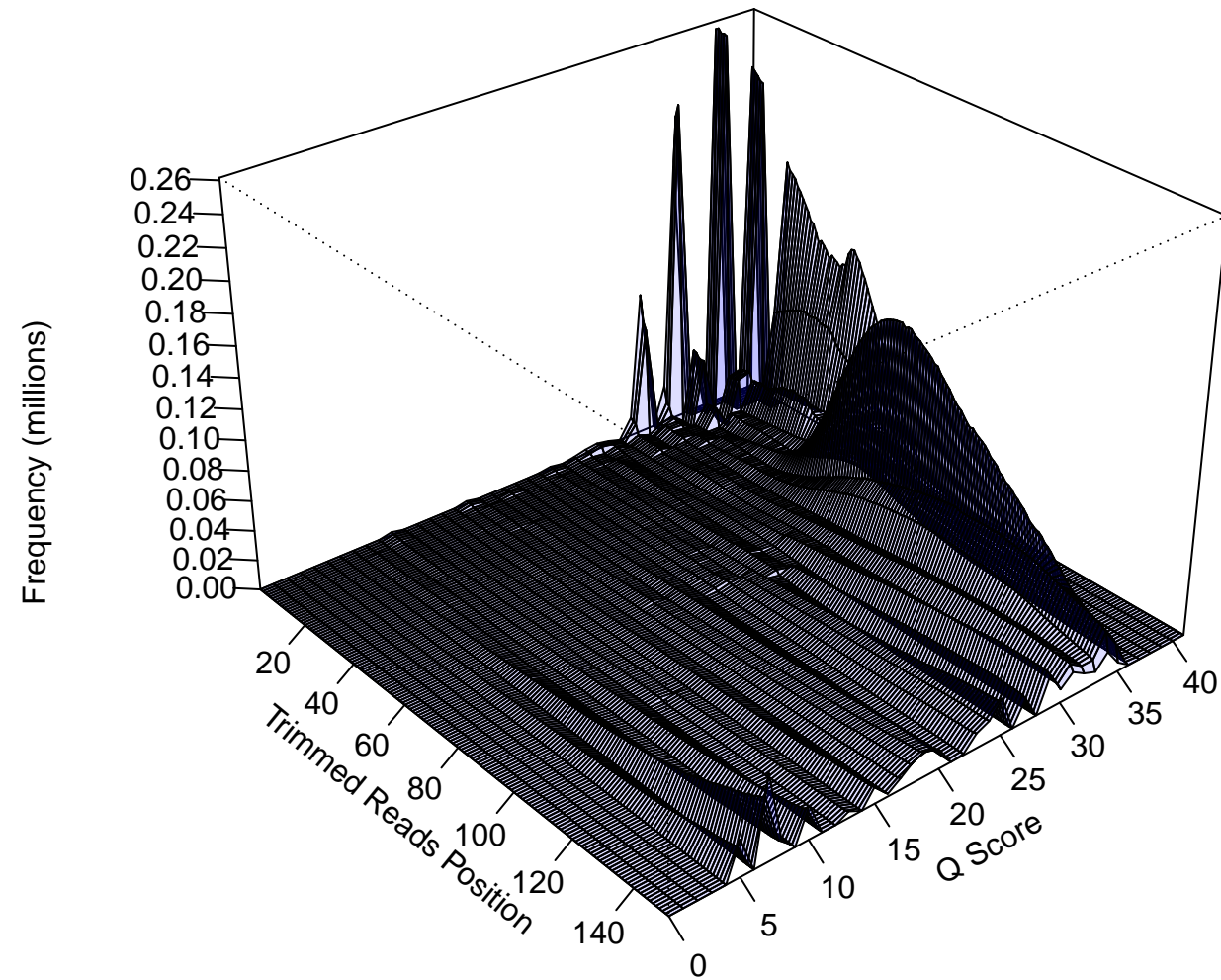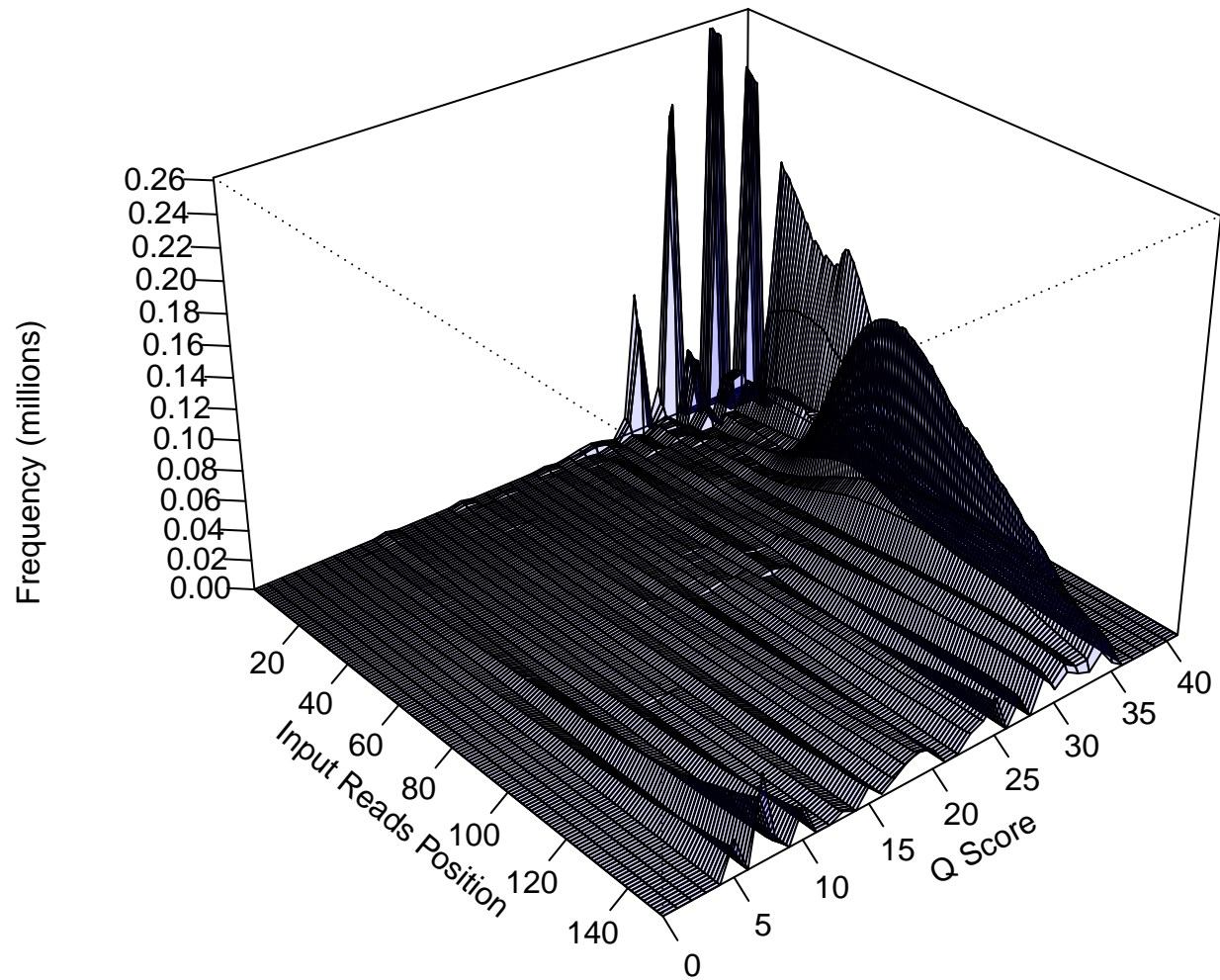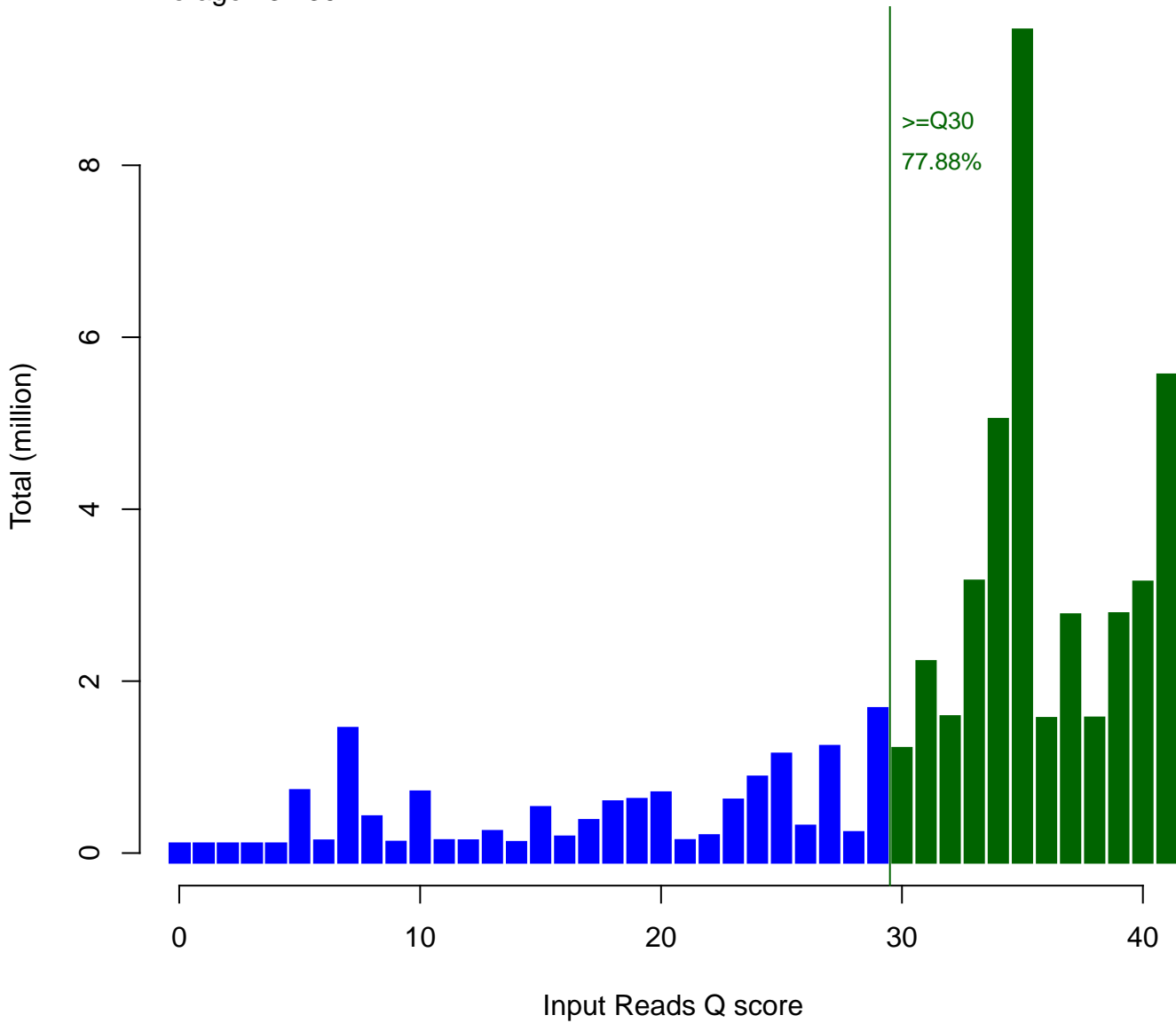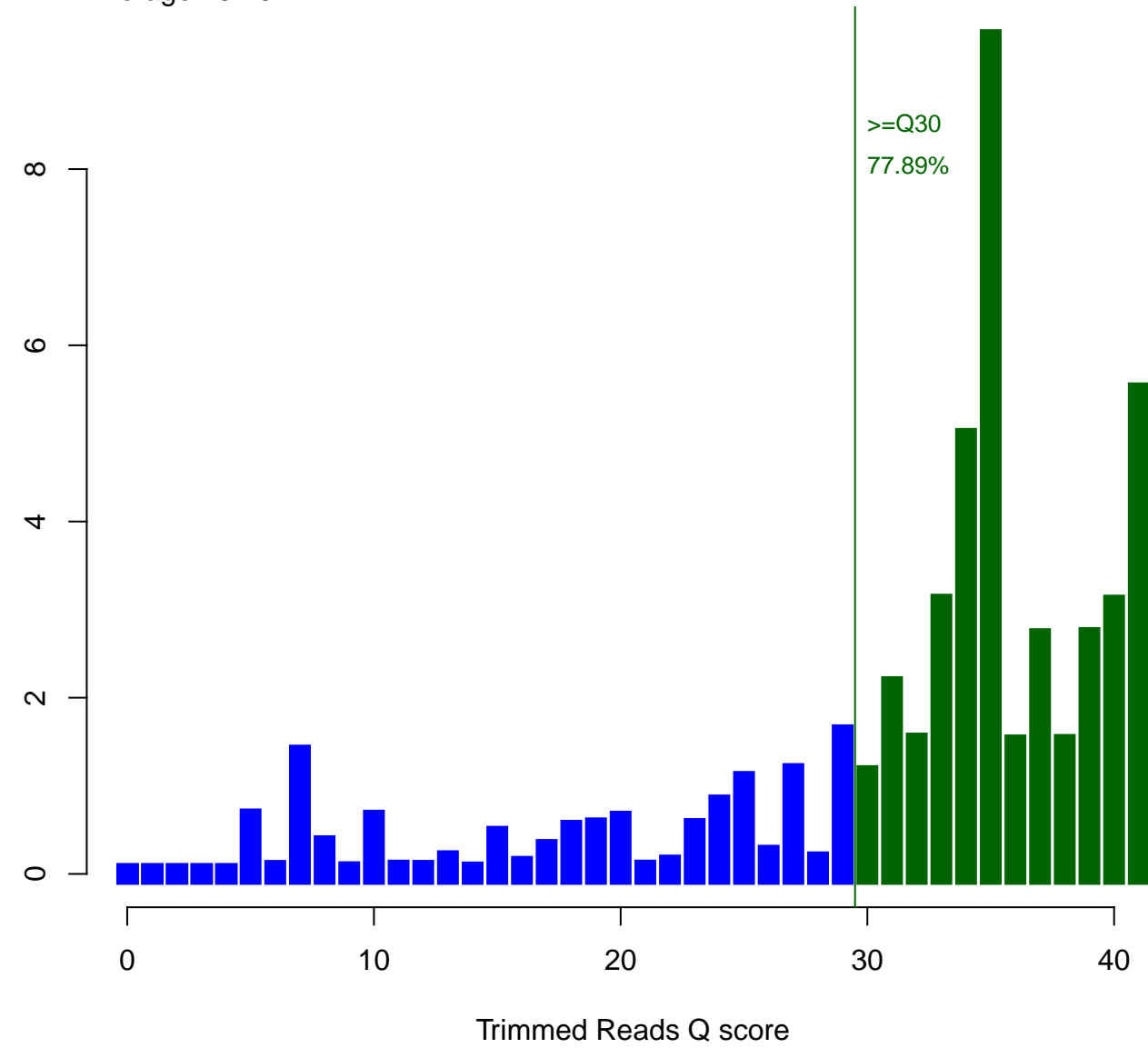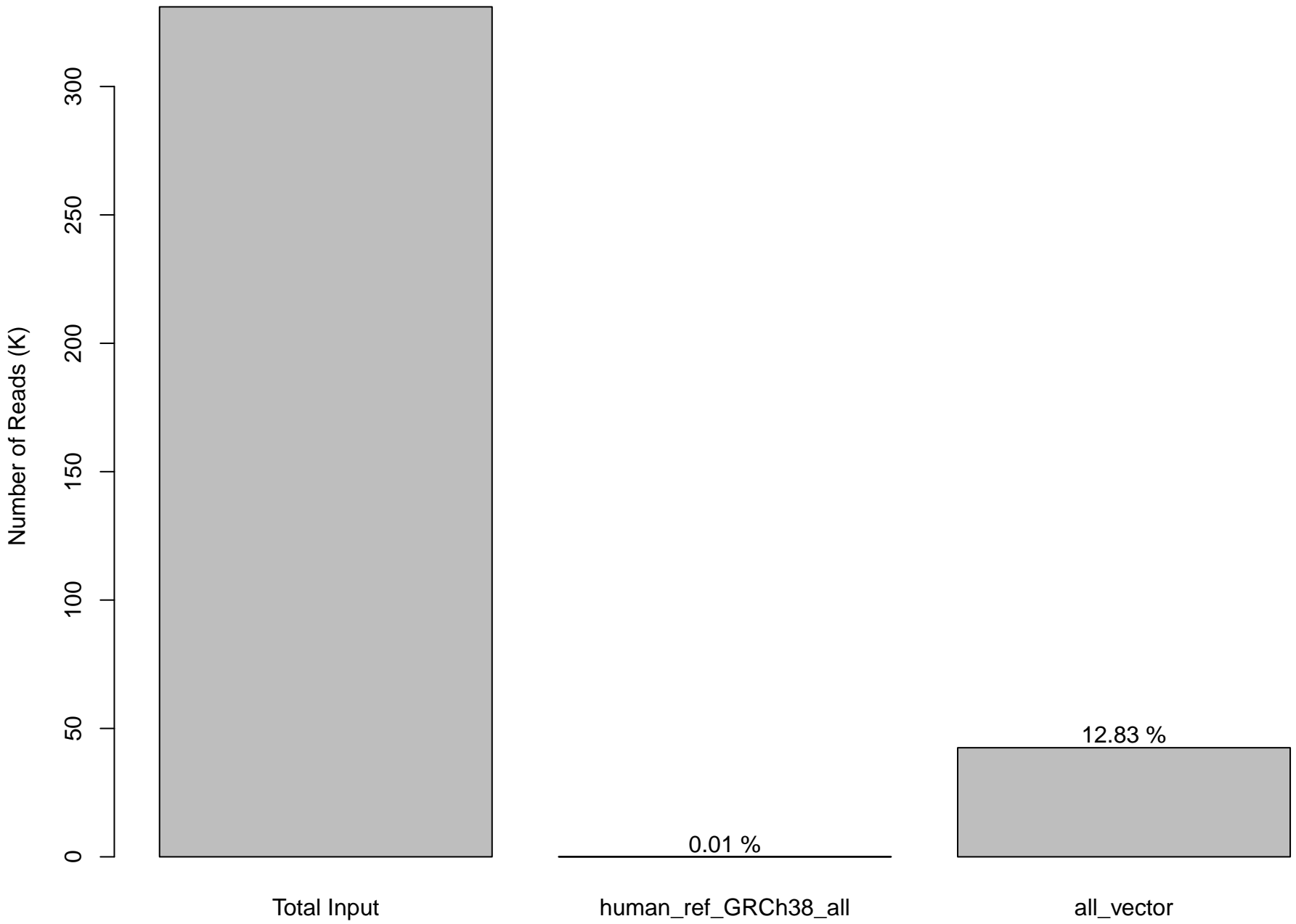
Host Removal

Length Distribution

1,367 Contigs
N50 6128
Max 36,181
Min 267
Median 1592
Total Base 4,228,447

Frequency
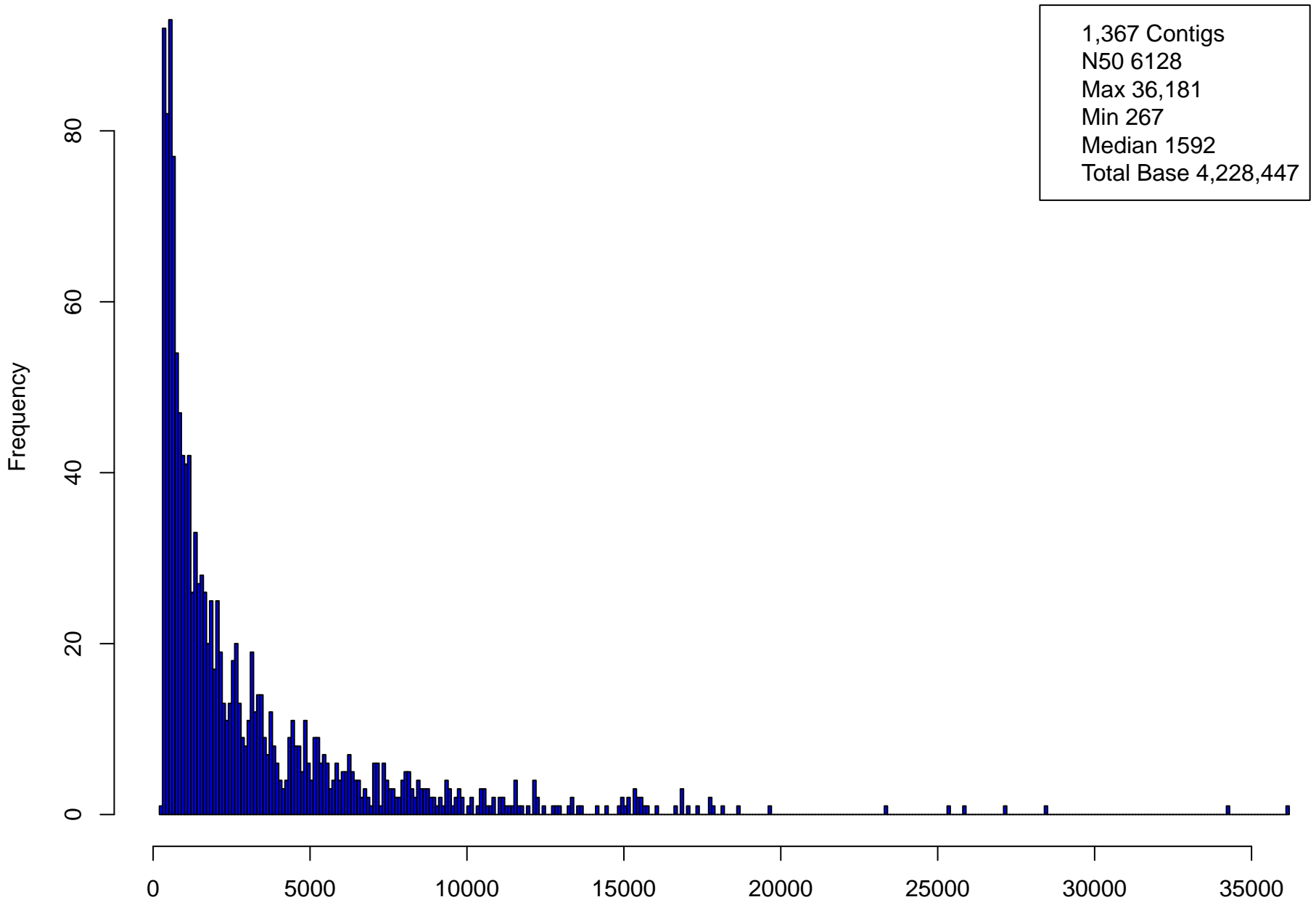
Length (bp)
/opt/apps/edge/edge_ui/EDGE_output/Ecoli_10x/AssemblyBasedAnalysis/contigs.fa

GC Histogram

Mean 51.73 %
Std 3.76

# of contigs

GC (%)
/opt/apps/edge/edge_ui/EDGE_output/Ecoli_10x/AssemblyBasedAnalysis/contigs.fa
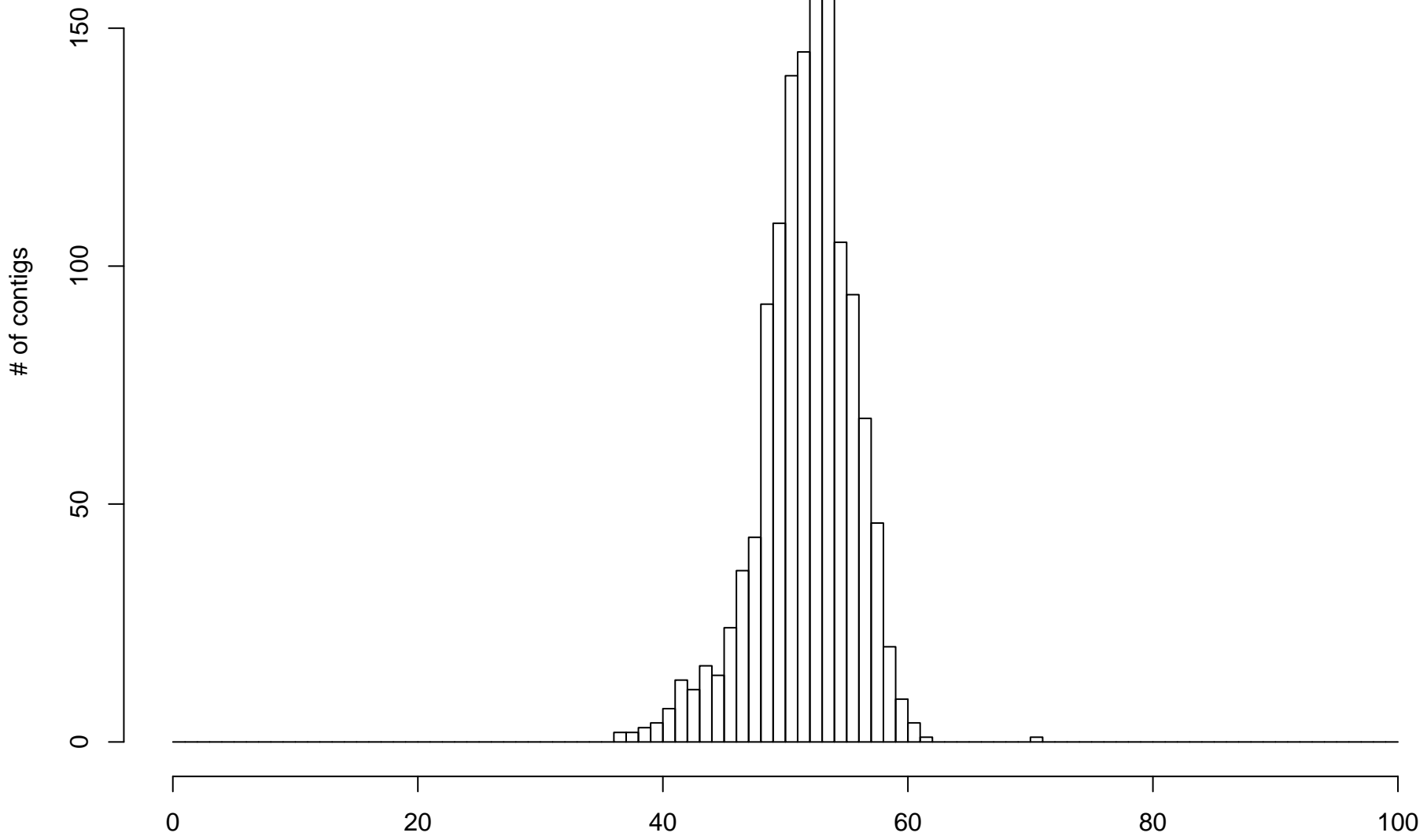
Cumulative Length

Contigs are ordered from largest to smallest.

Size (Mbp)

# of Contig

Nx Length

N50: 6.128

Nx

Size (Kbp)
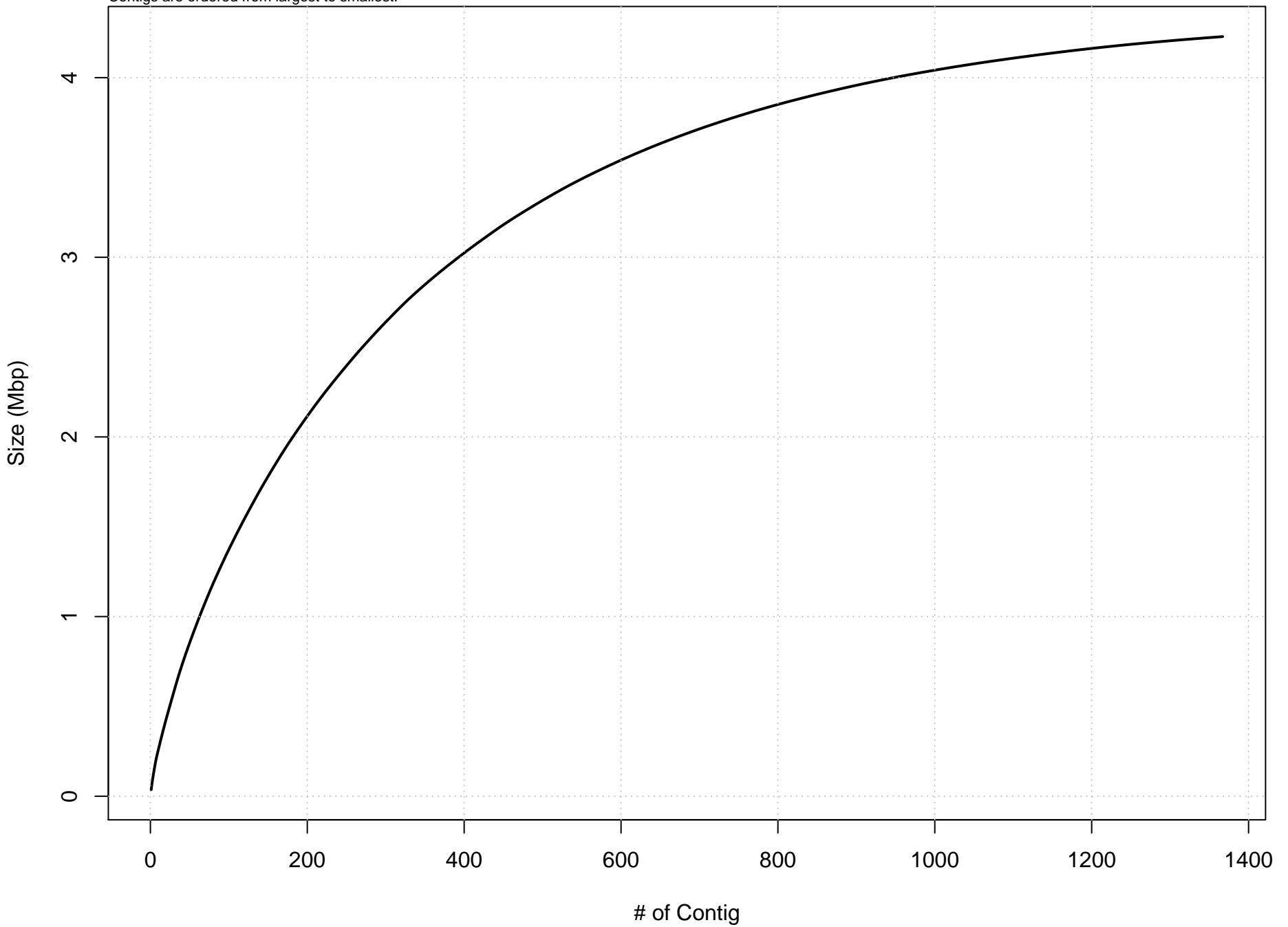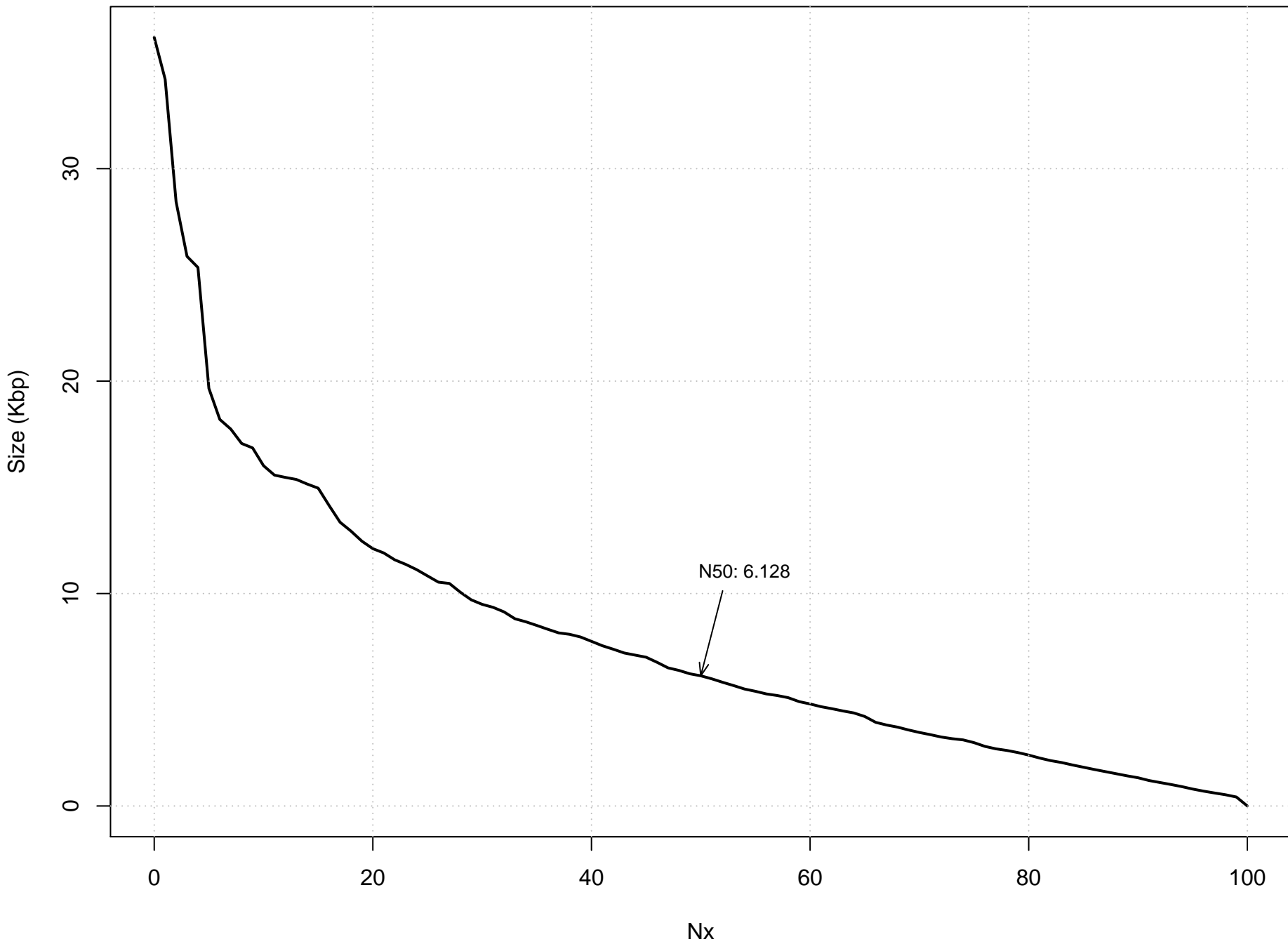
Mapping Reads to Contigs

288505 + 0 in total (QC–passed reads + QC–failed reads)

0 + 0 duplicates

279127 + 0 mapped (96.75%:–nan%)

267988 + 0 paired in sequencing

133994 + 0 read1

133994 + 0 read2

200434 + 0 properly paired (74.79%:–nan%)

258146 + 0 with itself and mate mapped

3524 + 0 singletons (1.31%:–nan%)

2746 + 0 with mate mapped to a different chr

2410 + 0 with mate mapped to a different chr (mapQ>=5)

Avg_coverage_fold: 7.7774

Coverage: 97.1961%

Contigs Average Fold Coverage vs. Contigs Length
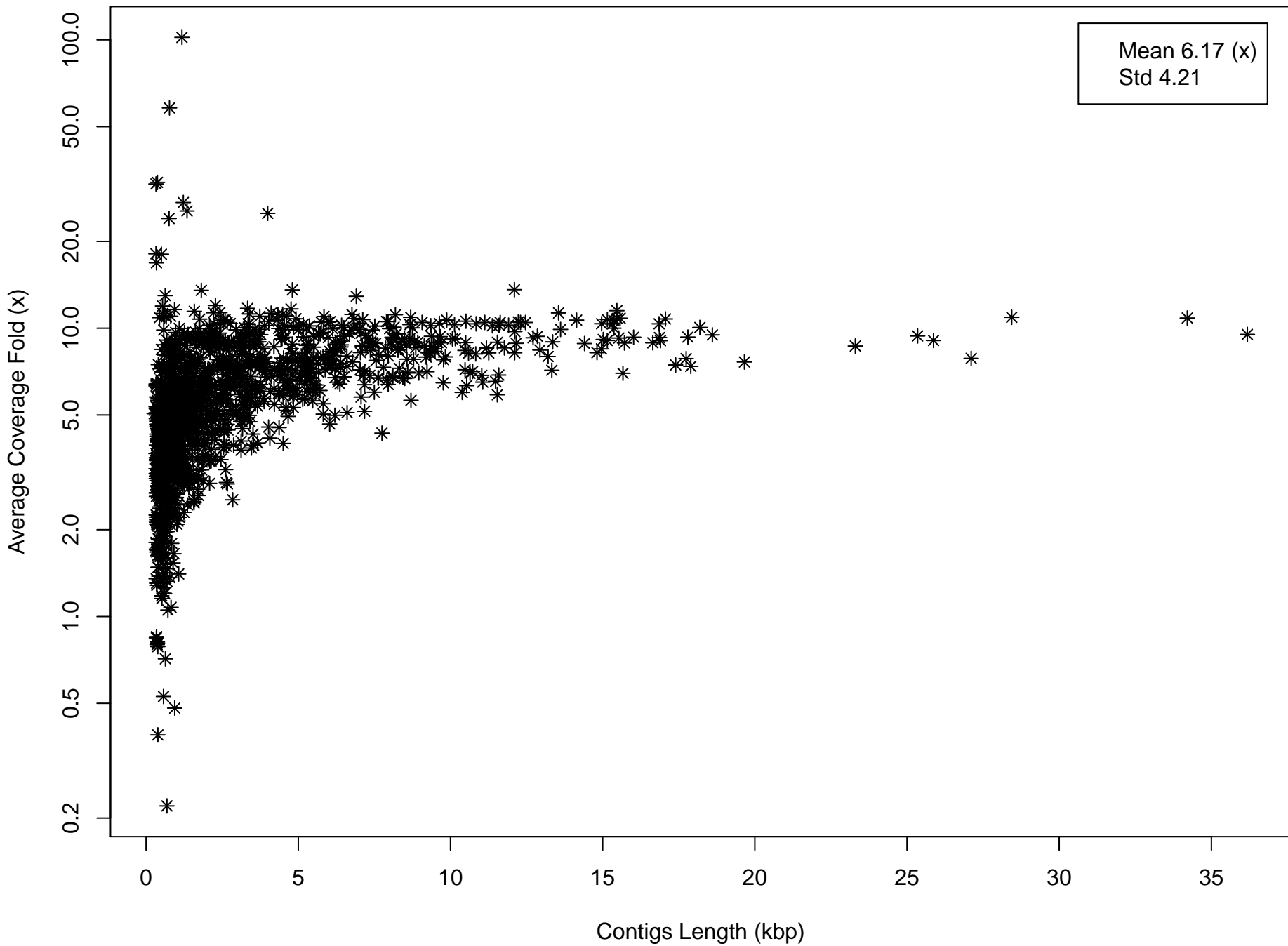
Contigs Coverage vs. Contigs Length

Overall Cov 97.20 %

Contigs Coverage (%)

Contigs Length (bp)
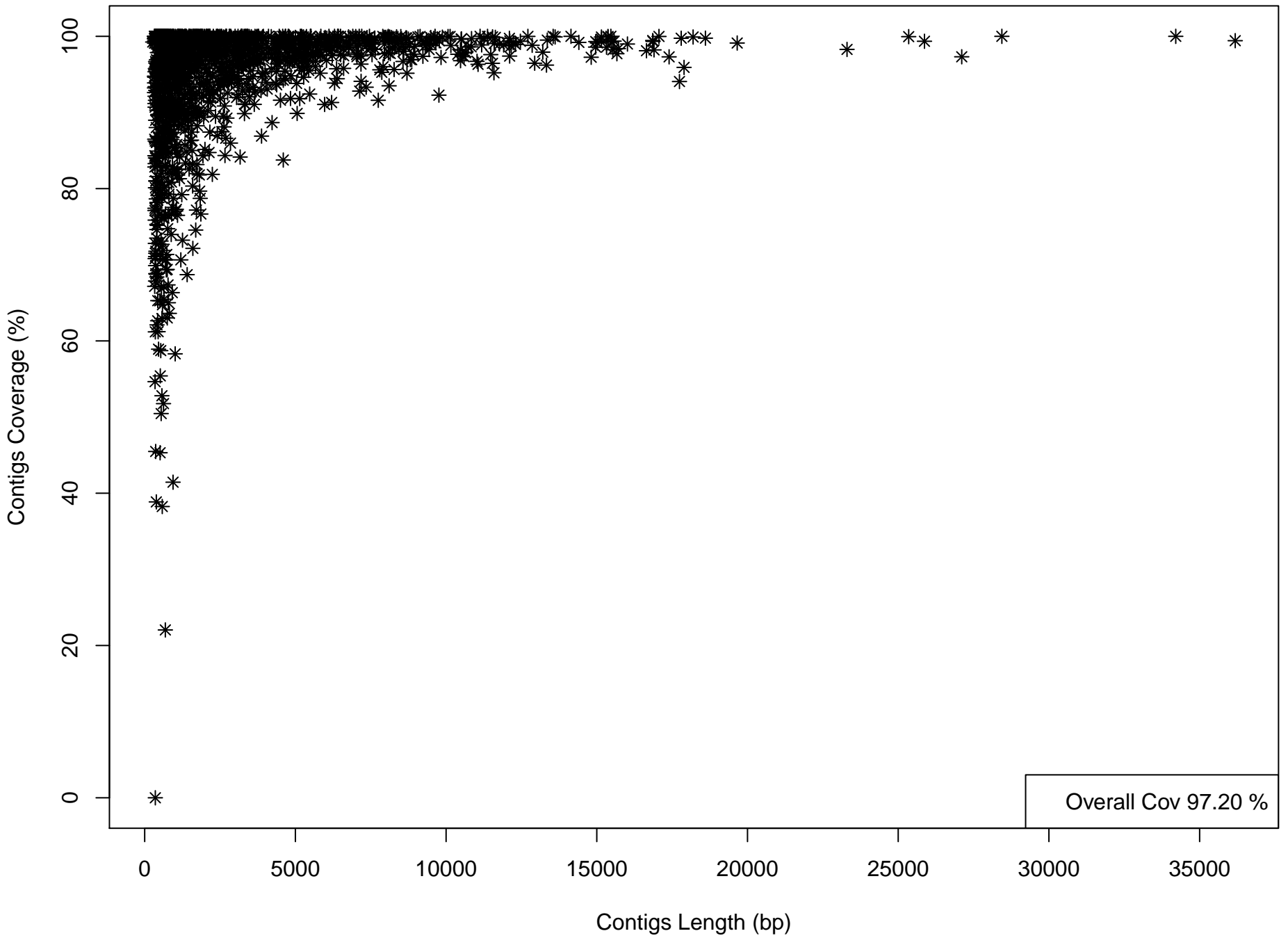
Contigs Average Fold Coverage vs. GC
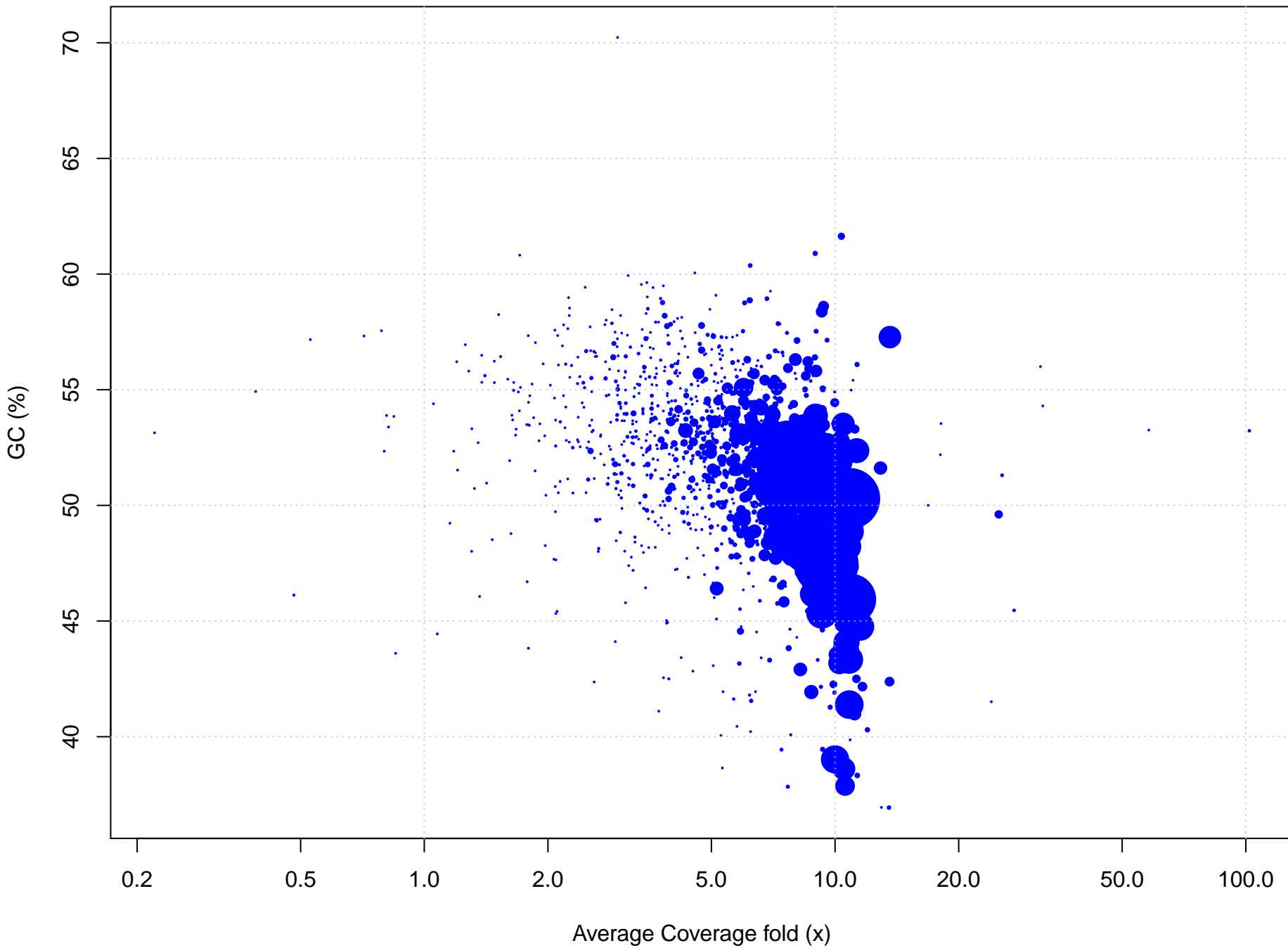
ANNOTATION STATS – Ecoli_10x (1367 contigs, 4228447bp)

Feature count
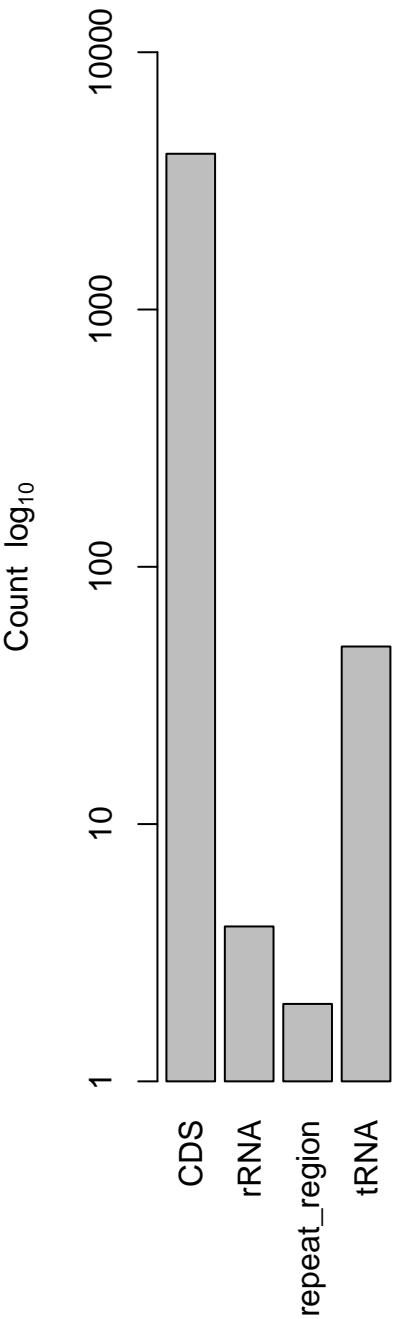
Distribution of protein size

CDS: 4031
– CDS (Function assigned): 2822
– CDS (hypothetical/putative): 1209
 * hypothetical: 989
 * putative: 204
 * putative lipoprotein: 11
 * putative membrane: 4
– EC_NUMBER: 1422
 * EC1: Dehydrogenase, oxidase: 294
 * EC2: Transaminase, kinase: 440
 * EC3: Lipase, amylase, peptidase: 400
 * EC4: Decarboxylase: 134
 * EC5: Isomerase, mutase: 87
 * EC6: Synthetase: 67
– GENE: 2832
– INFERENCE: 4031
 * CLUSTERS: 298
 * HAMAP: 62
 * Pfam: 159
 * Prodigal: 4031
 * UniProtKB: 2811

Mapping Contigs to Reference

Mapping criteria

Aligned portion should be >= 85.00 % identity

Total_reads: 1367

Unused Contigs#: 0 (0.00 %)

Avg_coverage_fold: 0.9621

Reference_Coverage: 92.2247%

Number of SNPs: 251

Number of INDELs: 17
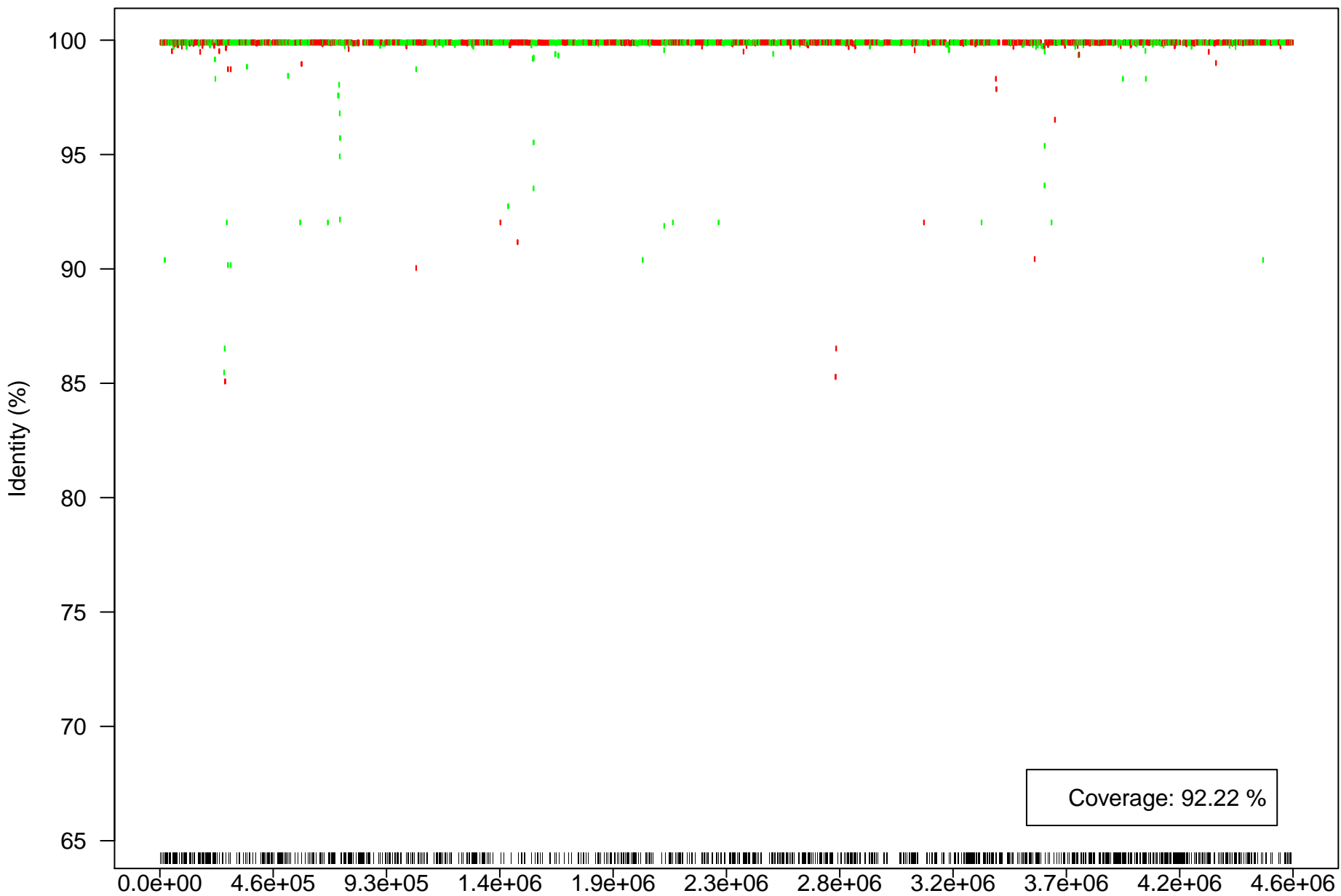
Mapping Reads to Reference

288505 + 0 in total (QC−passed reads + QC−failed reads)

0 + 0 duplicates

285649 + 0 mapped (99.01%:−nan%)

267988 + 0 paired in sequencing

133994 + 0 read1

133994 + 0 read2

204742 + 0 properly paired (76.40%:−nan%)

263854 + 0 with itself and mate mapped

1474 + 0 singletons (0.55%:−nan%)

0 + 0 with mate mapped to a different chr

0 + 0 with mate mapped to a different chr (mapQ>=5)

readsToRef_NC_000913

Average fold: 7.324 sd 4.6

Frequency

Coverage(fold)

Coverage: 94.46 % ; Average fold: 7.32x +/− 4.60 ; Gaps: 2494 (256943 bp) ;

Color scale legend: 10^{-1}, 10^{0}, 10^{1}, 10^{2}
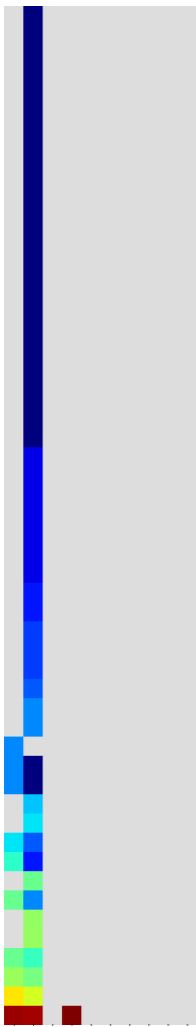
Row labels (top to bottom):
- Sphaerochaeta coccoides
- Caviid herpesvirus 2
- Bordetella bronchiseptica
- Modestobacter marinus
- Mycobacterium ulcerans
- Citrobacter rodentium
- Anabaena cylindrica
- Cercopithecine herpesvirus 2
- Shigella phage SfIV
- Rhizobium sp. IRBG74
- Shewanella sp. ANA-3
- Sphingomonas sp. MM-1
- Dill cryptic virus 1
- Crinalium epipsammum
- Rhodothermus marinus
- Dulcamara mottle virus
- Enterobacteria phage lambda
- Paenibacillus mucilaginosus
- Streptomyces albus
- Primula malacoides virus 1
- Enterobacteriaceae bacterium strain FGI 57
- Aeromonas hydrophila
- Cyprinid herpesvirus 1
- Hepatitis C virus
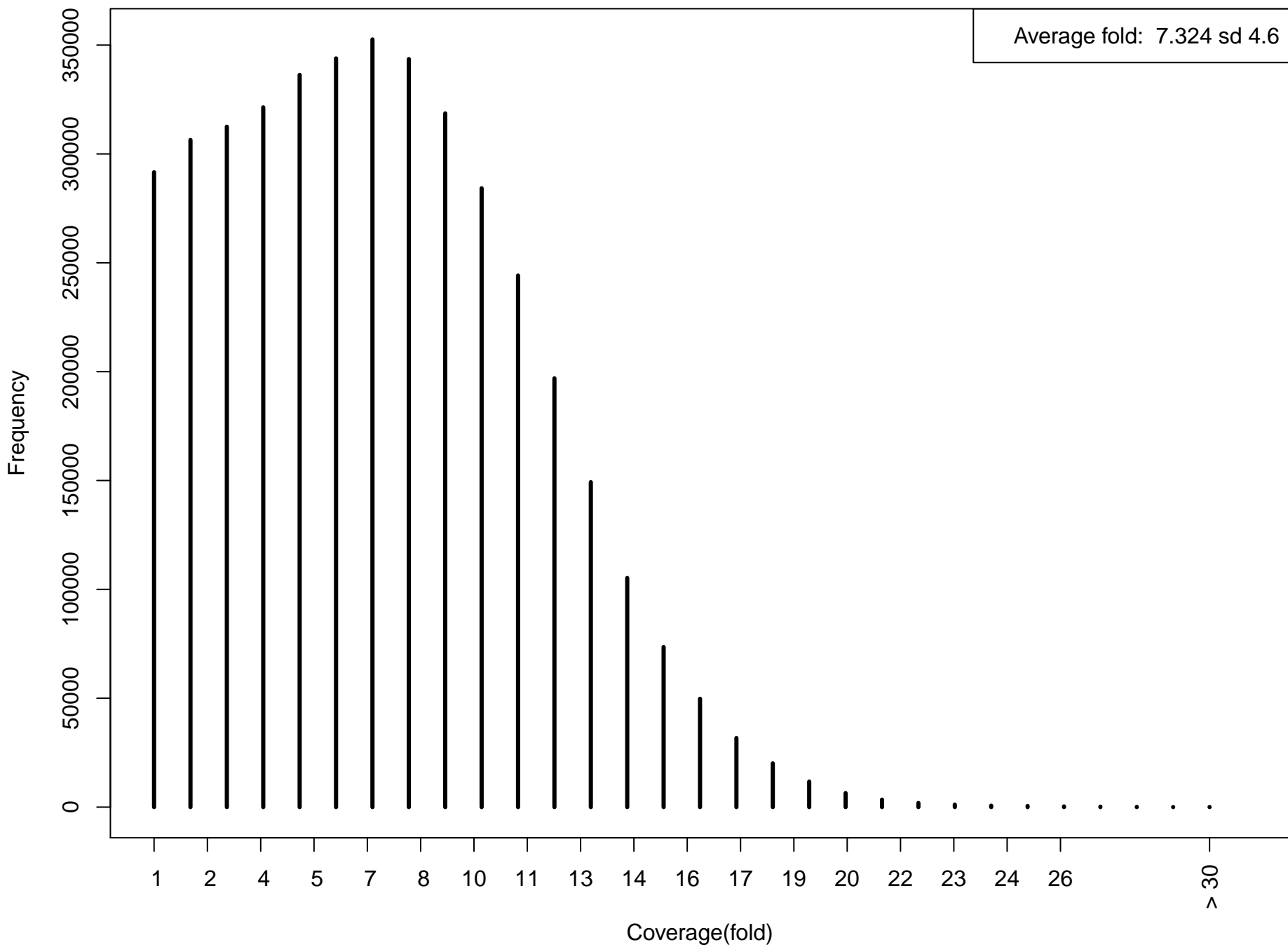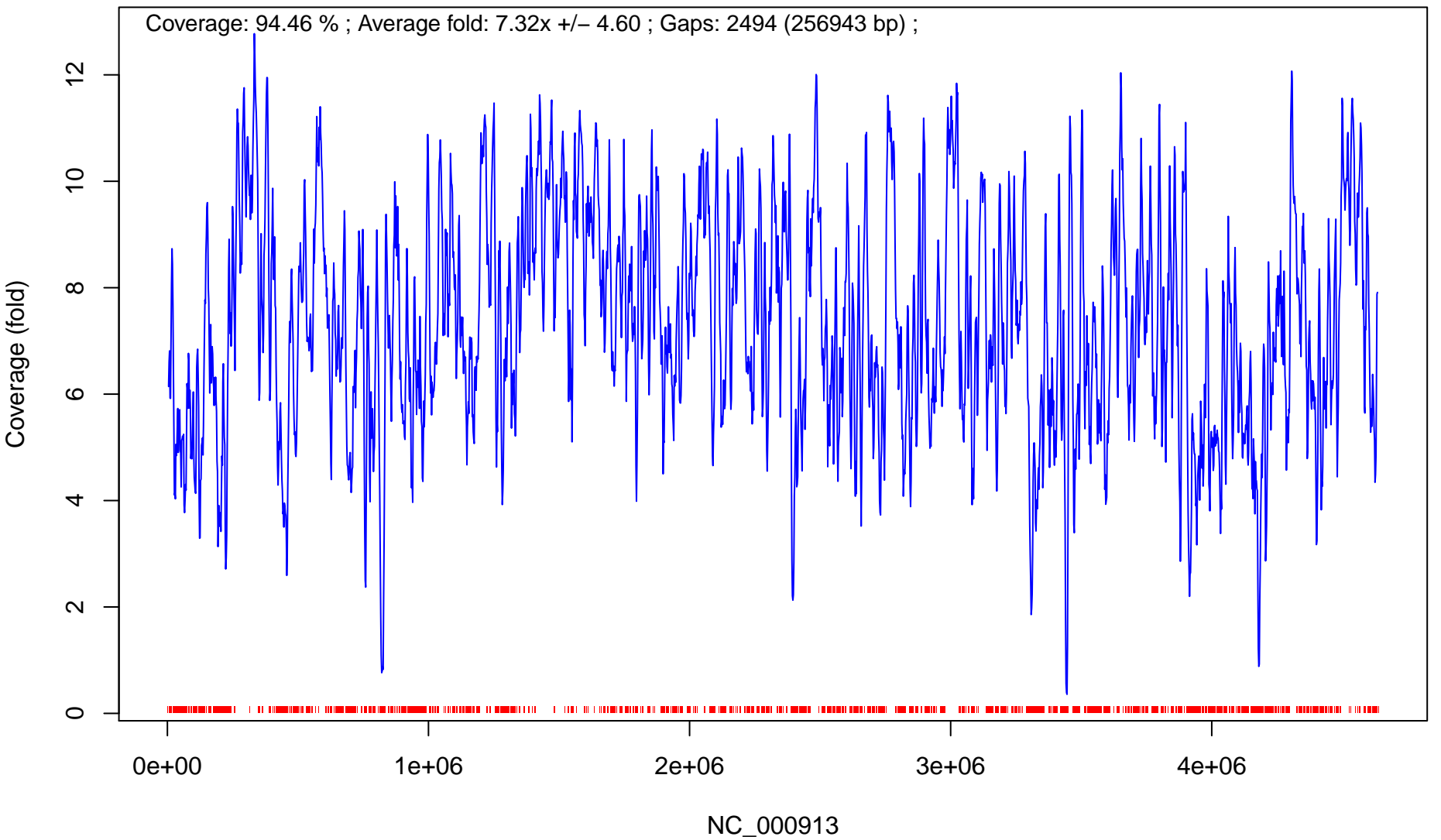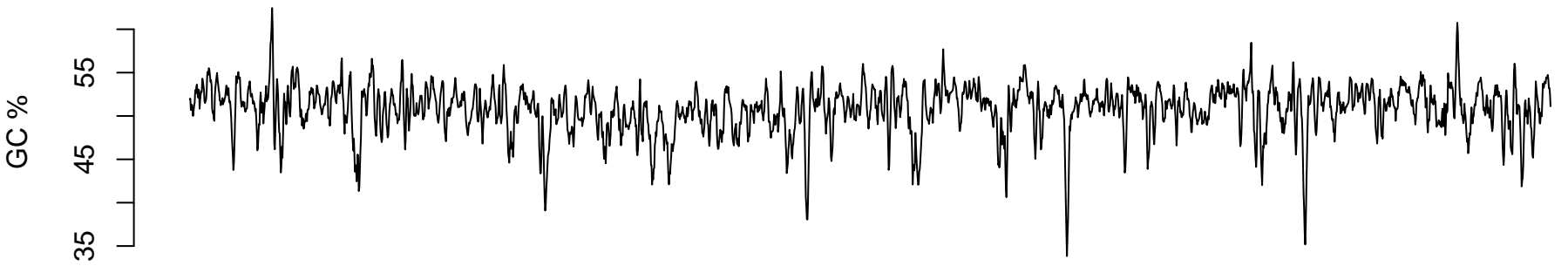- Klebsiella pneumoniae
- Pseudomonas sp. TKP
- Phytophthora infestans RNA virus 1
- Mycoplasma hyopneumoniae
- Streptomyces cattleya
- Stackebrandtia nassauensis
- Methylobacterium extorquens
- Pleurotus ostreatus virus 1
- Sorangium cellulosum
- Treponema primitia
- Streptomyces bingchenggensis
- Mycobacterium tuberculosis
- Escherichia fergusonii
- Salmonella enterica
- Citrobacter koseri
- Enterobacteria phage phiX174 sensu lato
- Enterobacter aerogenes
- Encephalomyocarditis virus
- Shigella dysenteriae
- Burkholderia mallei
- Delftia sp. Cs1-4
- Shigella boydii
- Delftia acidovorans
- Shigella flexneri
- Alteromonas macleodii
- Streptococcus suis
- Shigella sonnei
- Achromobacter xylosoxidans
- Escherichia coli

Column labels (left to right):
- kraken_mini
- bwa
- gottcha-genDB-b
- gottcha-speDB-b
- gottcha-strDB-b
- gottcha-genDB-v
- gottcha-speDB-v
- gottcha-strDB-v
- metaphlan
- metaphyler-srv
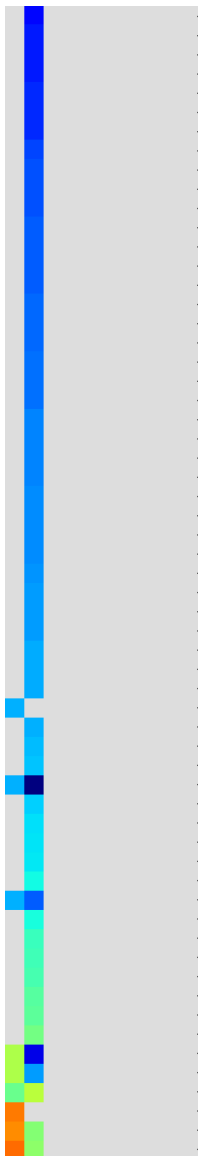
Escherichia coli IAI39
Shigella flexneri 5 str. 8401
Shigella flexneri 2a str. 301
Escherichia coli APEC O1
Escherichia coli IHE3034
Escherichia coli O127:H6 str. E2348/69
Escherichia coli JJ1886
Escherichia coli IAI1
Shigella flexneri 2002017
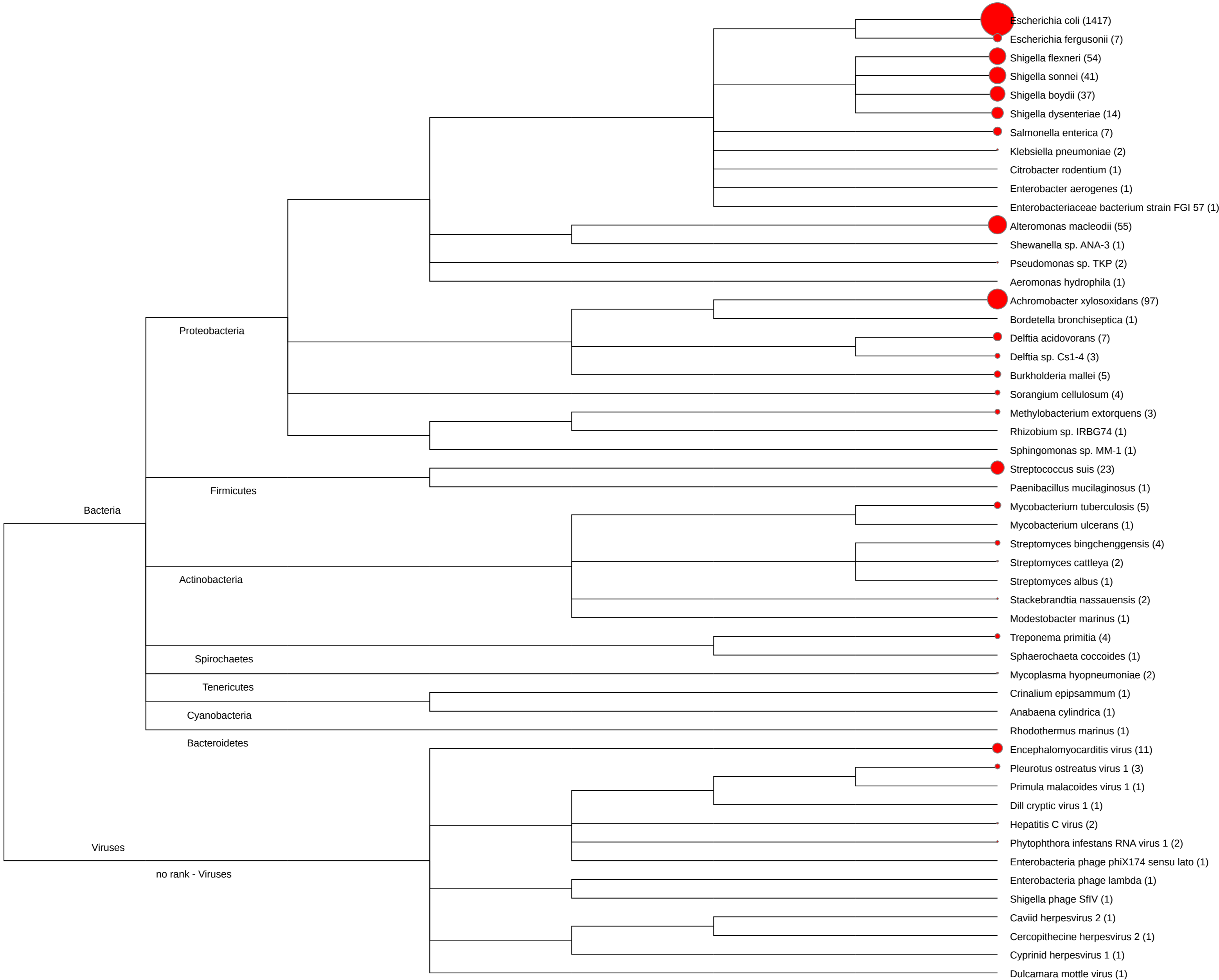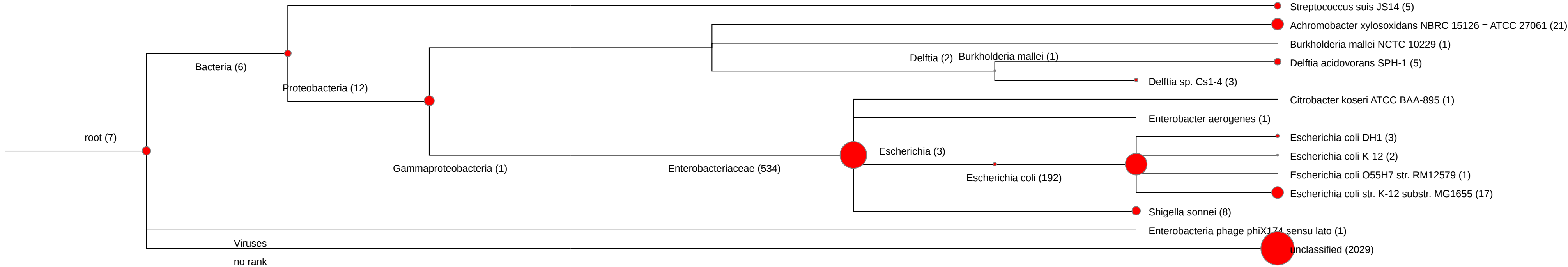Escherichia coli str. 'clone D i14'
Shigella dysenteriae Sd197
Escherichia coli 55989
Escherichia coli E24377A
Shigella boydii CDC 3083-94
Escherichia coli O157:H7 str. EDL933
Escherichia coli O104:H4 str. 2011C-3493
Escherichia coli O55:H7 str. CB9615
Escherichia coli NA114
Escherichia coli LY180
Escherichia coli O104:H4 str. 2009EL-2050
Escherichia coli O157:H7 str. TW14359
Escherichia coli APEC O78
Escherichia coli S88
Escherichia coli O157:H7 str. Sakai
Escherichia coli UMN026
Escherichia coli O111:H- str. 11128
Escherichia coli Xuzhou21
Shigella sonnei 53G
Shigella flexneri 2a str. 2457T
Shigella sonnei Ss046
Escherichia coli O157:H7 str. EC4115
Shigella boydii Sb227
Escherichia coli O26:H11 str. 11368
Escherichia coli O103:H2 str. 12009
Escherichia coli HS
Escherichia coli 'BL21-Gold(DE3)pLysS AG'
Citrobacter koseri ATCC BAA-895
Escherichia coli ATCC 8739
Escherichia coli O104:H4 str. 2009EL-2071
Escherichia coli B str. REL606
Burkholderia mallei NCTC 10229
Escherichia coli SE11
Escherichia coli UMNK88
Escherichia coli ETEC H10407
Escherichia coli P12b
Escherichia coli W
Escherichia coli O55:H7 str. RM12579
Escherichia coli KO11FL
Escherichia coli str. K-12 substr. MDS42
Alteromonas macleodii str. 'Ionian Sea U8'
Escherichia coli BL21(DE3)
Escherichia coli BW2952
Escherichia coli str. K-12 substr. W3110
Escherichia coli str. K-12 substr. DH10B
Delftia acidovorans SPH-1
Streptococcus suis JS14
Escherichia coli DH1
Escherichia coli K-12
Escherichia coli str. K-12 substr. MG1655
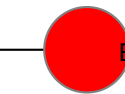Achromobacter xylosoxidans NBRC 15126 = ATCC 27061

Escherichia coli (1417)
Escherichia fergusonii (7)
Shigella flexneri (54)
Shigella sonnei (41)
Shigella boydii (37)
Shigella dysenteriae (14)
Salmonella enterica (7)
Klebsiella pneumoniae (2)
Citrobacter rodentium (1)
Enterobacter aerogenes (1)
Enterobacteriaceae bacterium strain FGI 57 (1)
Alteromonas macleodii (55)
Shewanella sp. ANA-3 (1)
Pseudomonas sp. TKP (2)
Aeromonas hydrophila (1)
Achromobacter xylosoxidans (97)
Bordetella bronchiseptica (1)
Delftia acidovorans (7)
Delftia sp. Cs1-4 (3)
Burkholderia mallei (5)
Sorangium cellulosum (4)
Methylobacterium extorquens (3)
Rhizobium sp. IRBG74 (1)
Sphingomonas sp. MM-1 (1)
Streptococcus suis (23)
Paenibacillus mucilaginosus (1)
Mycobacterium tuberculosis (5)
Mycobacterium ulcerans (1)
Streptomyces bingchenggensis (4)
Streptomyces cattleya (2)
Streptomyces albus (1)
Stackebrandtia nassauensis (2)
Modestobacter marinus (1)
Treponema primitia (4)
Sphaerochaeta coccoides (1)
Mycoplasma hyopneumoniae (2)
Crinalium epipsammum (1)
Anabaena cylindrica (1)
Rhodothermus marinus (1)
Encephalomyocarditis virus (11)
Pleurotus ostreatus virus 1 (3)
Primula malacoides virus 1 (1)
Dill cryptic virus 1 (1)
Hepatitis C virus (2)
Phytophthora infestans RNA virus 1 (2)
Enterobacteria phage phiX174 sensu lato (1)
Enterobacteria phage lambda (1)
Shigella phage SfIV (1)
Caviid herpesvirus 2 (1)
Cercopithecine herpesvirus 2 (1)
Cyprinid herpesvirus 1 (1)
Dulcamara mottle virus (1)

Proteobacteria

Firmicutes

Actinobacteria

Spirochaetes

Tenericutes

Cyanobacteria

Bacteroidetes

Bacteria

Viruses
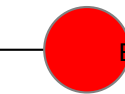
no rank - Viruses

Proteobacteria

Gammaproteobacteria

Escherichia coli BW2952 (1.29)

Proteobacteria

Gammaproteobacteria

Escherichia coli BW2952 (1.29)

PCR Assay Validation

```
Validate primers binding to Contigs
Primer Pair Forward and Reverse
PCR success!
Primer Forward Alignment to (Ecoli_10x_221)

 r     TGAGGGACCTGCGAAGCCCGGAT
       ||||||||||||| |||||||||
 p     TGAGGGACCTGCGAGGCCCGGAT

Primer Reverse Alignment to (Ecoli_10x_221)

 r     TTCATCGTGAATCCATTAGACTT
       |||||||||||||||||||||||
 p     TTCATCGTGAATCCATTAGACTT

The primers amplify Ecoli_10x_221 from 4336 to 4958, with size 623
----------
Validate primers binding to Reference
Primer Pair Forward and Reverse
PCR success!
Primer Forward Alignment to (NC_000913)

 r     ATCCGGGCCTCGCAGGTCCCTCA
       |||||||||||||||||||||||
 p     ATCCGGGCCTCGCAGGTCCCTCA

Primer Reverse Alignment to (NC_000913)

 r     AAGTCTAATGGATTCACGATGAA
       |||||||||||||||||||||||
 p     AAGTCTAATGGATTCACGATGAA

The primers amplify NC_000913 from 3957949 to 3958571, with size 623
----------
```