

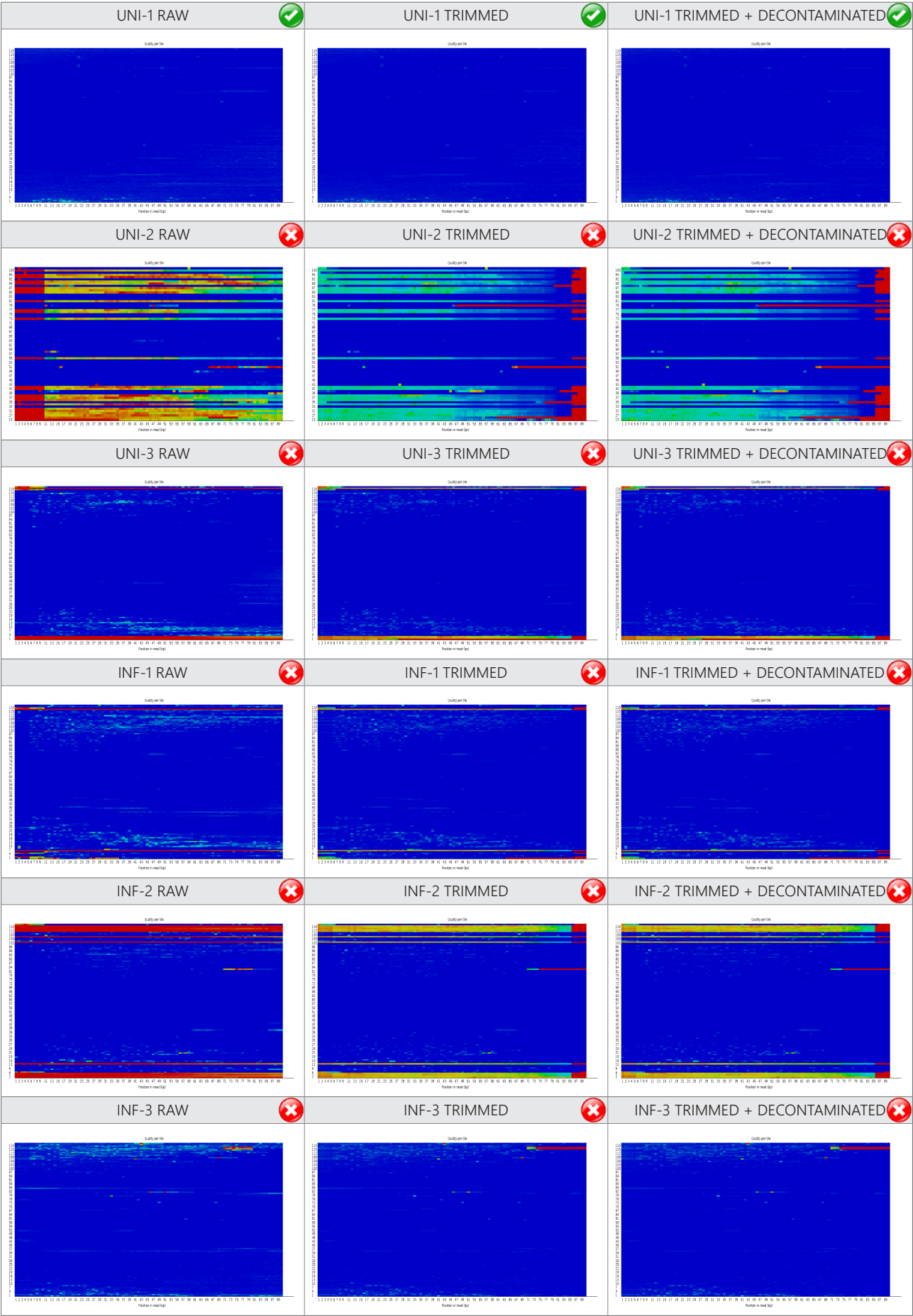
Basic statistics

UNI-1 RAW 		UNI-1 TRIMMED 		UNI-1 TRIMMED + DECONTAMINATED 	
Filename	uni1.fastq	Filename	uni1-t.fastq	Filename	uni1-td.fastq
File type	Conventional base calls	File type	Conventional base calls	File type	Conventional base calls
Encoding	Sanger / Illumina 1.9	Encoding	Sanger / Illumina 1.9	Encoding	Sanger / Illumina 1.9
Total Sequences	49786224	Total Sequences	49591551	Total Sequences	45923260
Flagg. as poor quality	0	Flagg. as poor quality	0	Flagg. as poor quality	0
Sequence length	90	Sequence length	10-90	Sequence length	13-90
%GC	47	%GC	46	%GC	46
UNI-2 RAW 		UNI-2 TRIMMED 		UNI-2 TRIMMED + DECONTAMINATED 	
Filename	uni2.fastq	Filename	uni2-t.fastq	Filename	uni2-td.fastq
File type	Conventional base calls	File type	Conventional base calls	File type	Conventional base calls
Encoding	Sanger / Illumina 1.9	Encoding	Sanger / Illumina 1.9	Encoding	Sanger / Illumina 1.9
Total Sequences	14187100	Total Sequences	13227308	Total Sequences	11792587
Flagg. as poor quality	0	Flagg. as poor quality	0	Flagg. as poor quality	0
Sequence length	90	Sequence length	10-90	Sequence length	13-90
%GC	44	%GC	43	%GC	44
UNI-3 RAW 		UNI-3 TRIMMED 		UNI-3 TRIMMED + DECONTAMINATED 	
Filename	uni3.fastq	Filename	uni3-t.fastq	Filename	uni3-td.fastq
File type	Conventional base calls	File type	Conventional base calls	File type	Conventional base calls
Encoding	Sanger / Illumina 1.9	Encoding	Sanger / Illumina 1.9	Encoding	Sanger / Illumina 1.9
Total Sequences	41276761	Total Sequences	40899111	Total Sequences	38404042
Flagg. as poor quality	0	Flagg. as poor quality	0	Flagg. as poor quality	0
Sequence length	90	Sequence length	10-90	Sequence length	13-90
%GC	46	%GC	45	%GC	46
INF-1 RAW 		INF-1 TRIMMED 		INF-1 TRIMMED + DECONTAMINATED 	
Filename	inf1.fastq	Filename	inf1-t.fastq	Filename	inf1-td.fastq
File type	Conventional base calls	File type	Conventional base calls	File type	Conventional base calls
Encoding	Sanger / Illumina 1.9	Encoding	Sanger / Illumina 1.9	Encoding	Sanger / Illumina 1.9
Total Sequences	38463425	Total Sequences	38161005	Total Sequences	35838467
Flagg. as poor quality	0	Flagg. as poor quality	0	Flagg. as poor quality	0
Sequence length	90	Sequence length	10-90	Sequence length	13-90
%GC	46	%GC	45	%GC	46
INF-2 RAW 		INF-2 TRIMMED 		INF-2 TRIMMED + DECONTAMINATED 	
Filename	inf2.fastq	Filename	inf2-t.fastq	Filename	inf2-td.fastq
File type	Conventional base calls	File type	Conventional base calls	File type	Conventional base calls
Encoding	Sanger / Illumina 1.9	Encoding	Sanger / Illumina 1.9	Encoding	Sanger / Illumina 1.9
Total Sequences	33518174	Total Sequences	32846310	Total Sequences	29857785
Flagg. as poor quality	0	Flagg. as poor quality	0	Flagg. as poor quality	0
Sequence length	90	Sequence length	10-90	Sequence length	13-90
%GC	44	%GC	43	%GC	44
INF-3 RAW 		INF-3 TRIMMED 		INF-3 TRIMMED + DECONTAMINATED 	
Filename	inf3.fastq	Filename	inf3-t.fastq	Filename	inf3-td.fastq
File type	Conventional base calls	File type	Conventional base calls	File type	Conventional base calls
Encoding	Sanger / Illumina 1.9	Encoding	Sanger / Illumina 1.9	Encoding	Sanger / Illumina 1.9
Total Sequences	47206685	Total Sequences	46978115	Total Sequences	43190157
Flagg. as poor quality	0	Flagg. as poor quality	0	Flagg. as poor quality	0
Sequence length	90	Sequence length	10-90	Sequence length	13-90
%GC	44	%GC	43	%GC	44

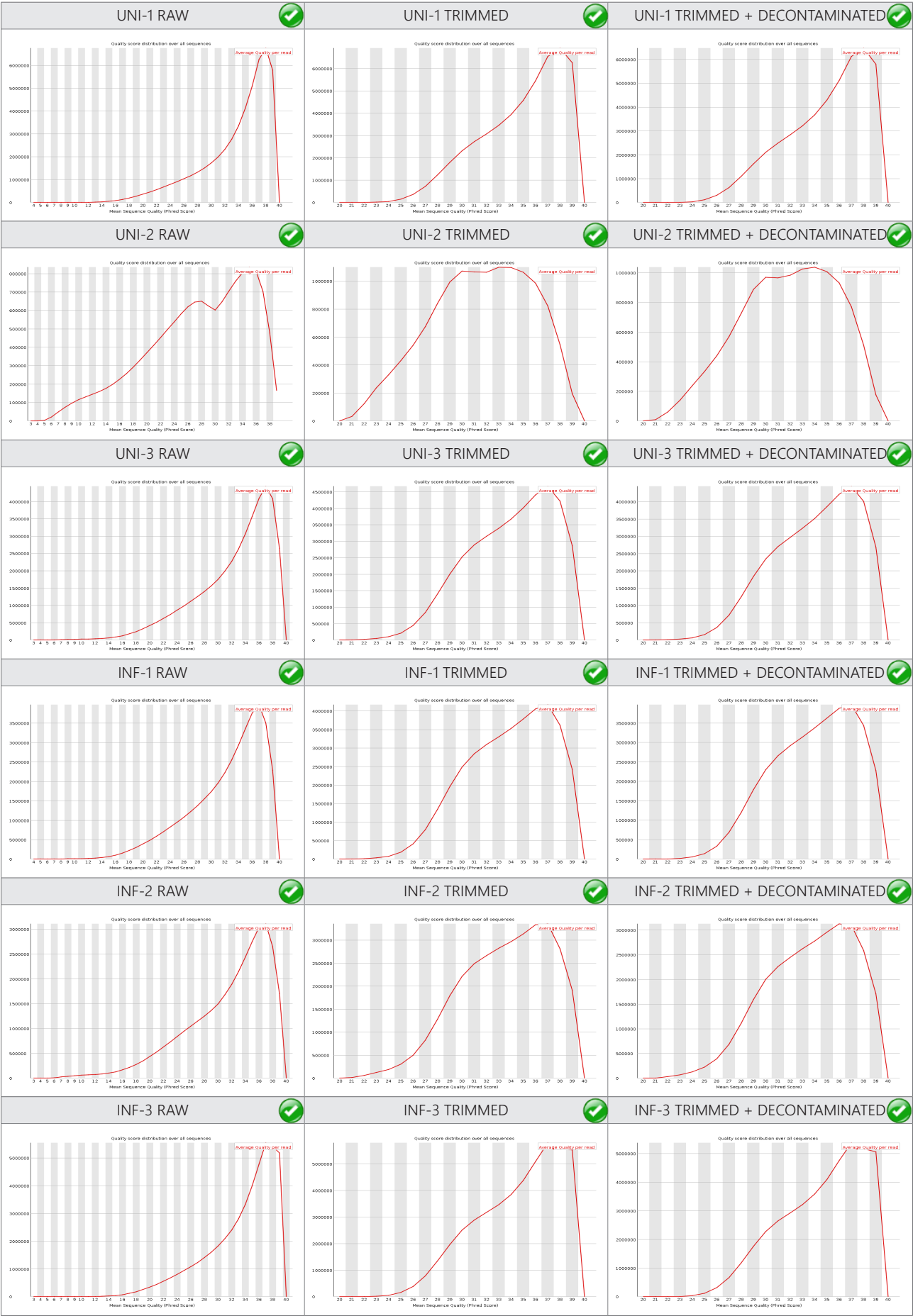
Per base sequence quality



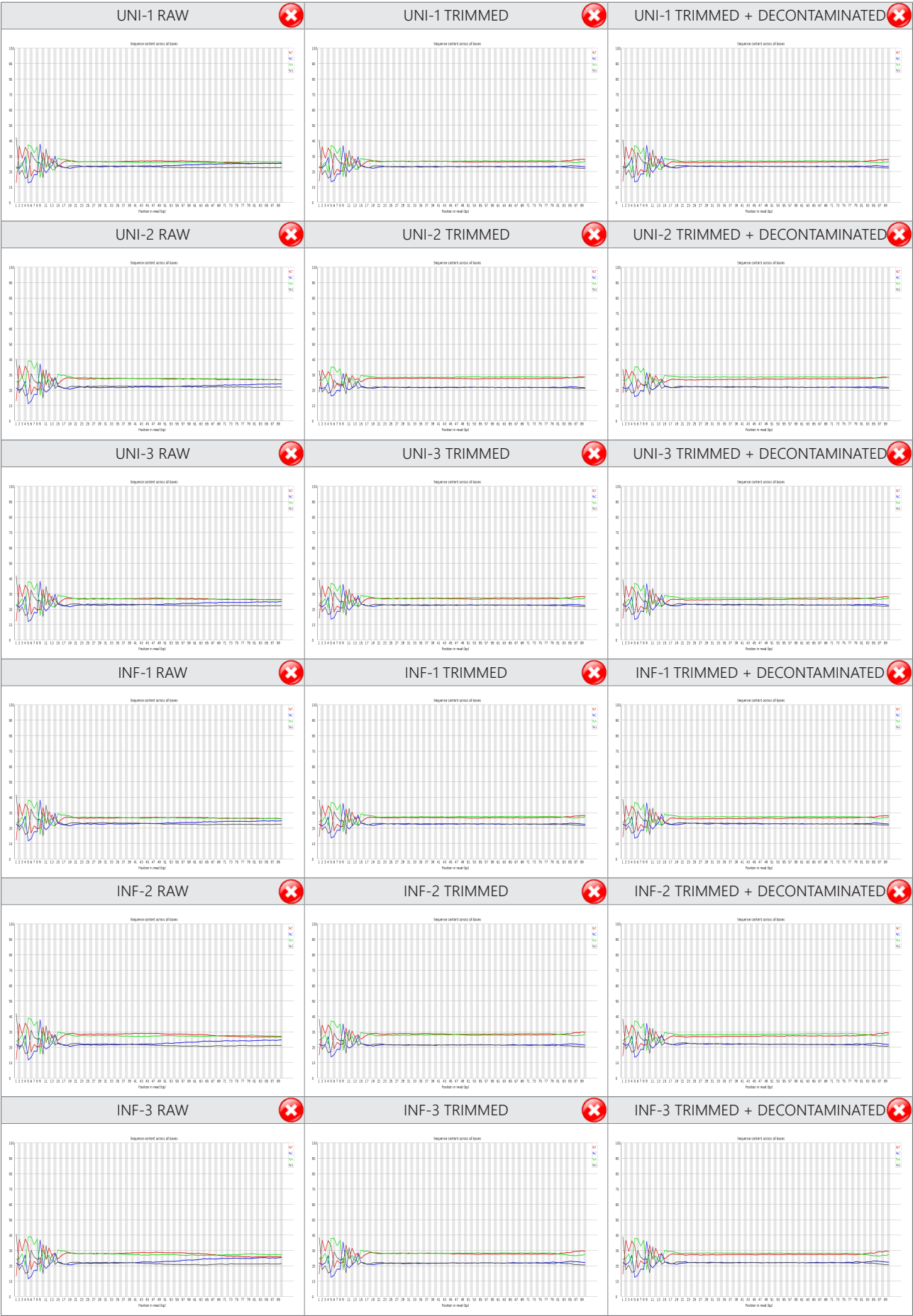
Per tile sequence quality



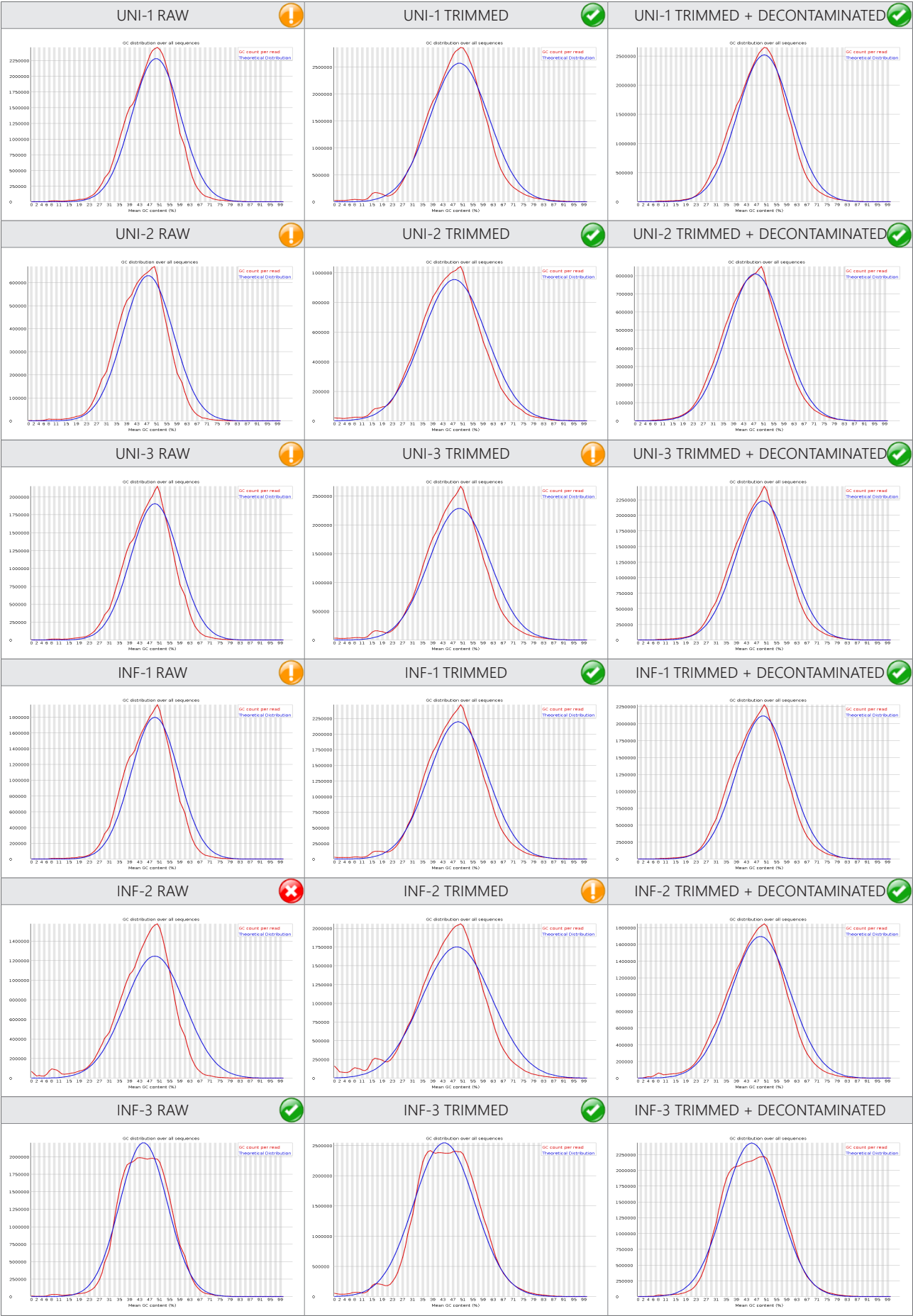
Per sequence quality scores



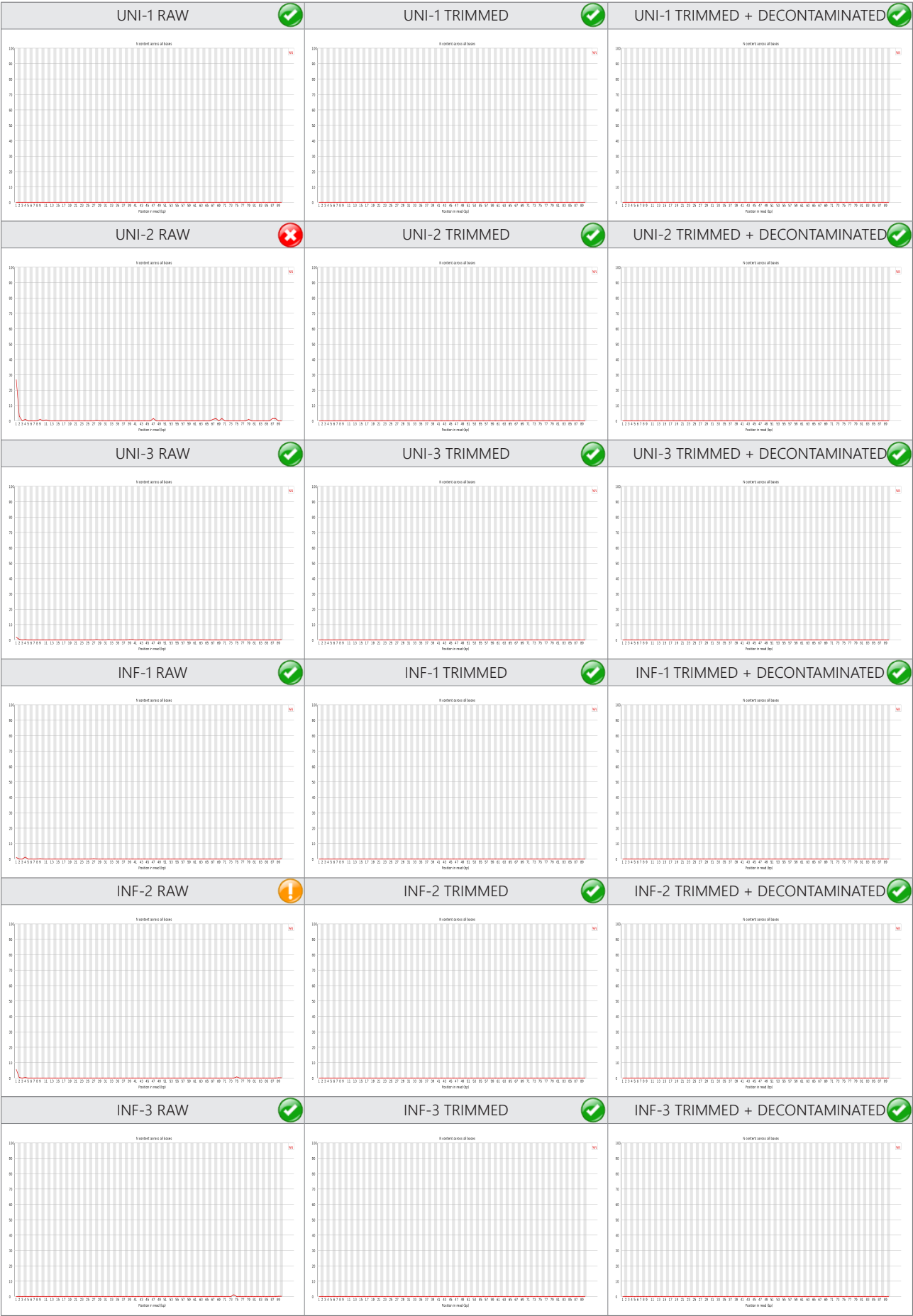
Per base sequence content



Per sequence GC content



Per base N content



Sequence length distribution



Sequence duplication levels



Overrepresented sequences

UNI-1 RAW 	UNI-1 TRIMMED 	UNI-1 TRIMMED + DECONTAMINATED 
No overrepresented sequences	No overrepresented sequences	No overrepresented sequences
UNI-2 RAW 	UNI-2 TRIMMED 	UNI-2 TRIMMED + DECONTAMINATED 
No overrepresented sequences	No overrepresented sequences	No overrepresented sequences
UNI-3 RAW 	UNI-3 TRIMMED 	UNI-3 TRIMMED + DECONTAMINATED 
No overrepresented sequences	No overrepresented sequences	No overrepresented sequences
INF-1 RAW 	INF-1 TRIMMED 	INF-1 TRIMMED + DECONTAMINATED 
No overrepresented sequences	No overrepresented sequences	No overrepresented sequences
INF-2 RAW 	INF-2 TRIMMED 	INF-2 TRIMMED + DECONTAMINATED 
No overrepresented sequences	No overrepresented sequences	No overrepresented sequences
INF-3 RAW	INF-3 TRIMMED 	INF-3 TRIMMED + DECONTAMINATED 
GGTATCAACGCAGAGTAC(T) ₃₁ 59845 0.12677229930464298 No Hit GTATCAACGCAGAGTAC(T) ₃₃ 58522 0.1239697301346197 No Hit	No overrepresented sequences	No overrepresented sequences

Adapter content

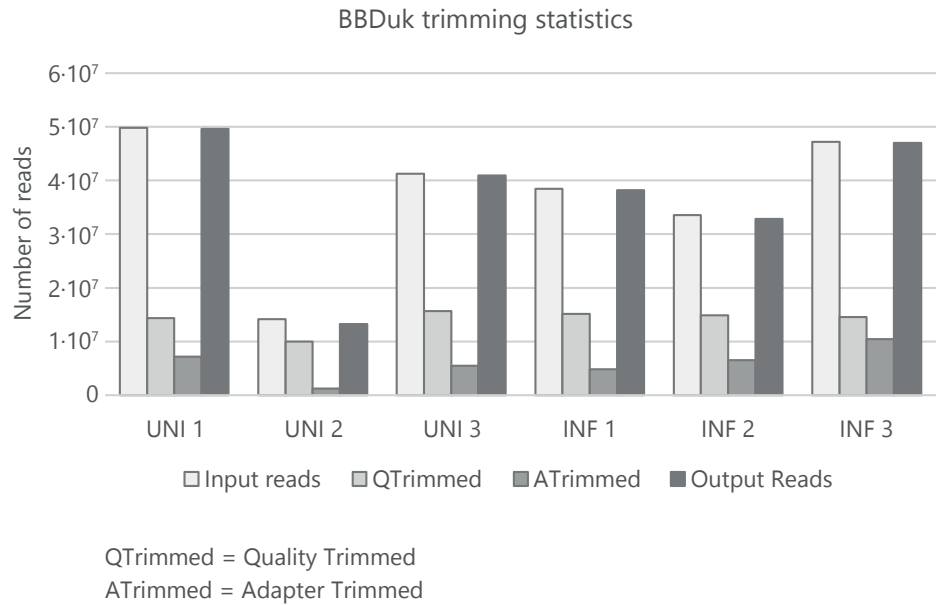


Kmer content



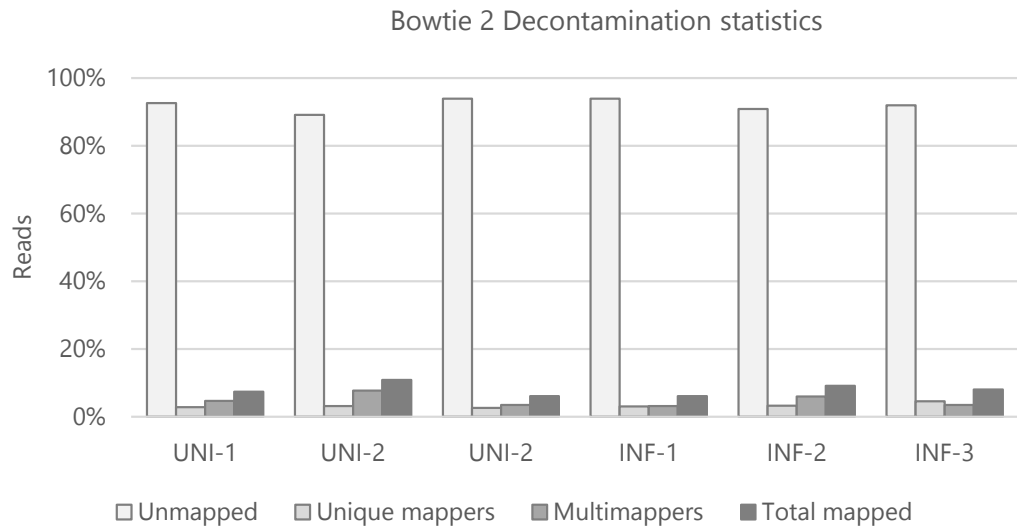
Trimming statistics:

```
BBduk run parameters:  
k=28 hdist=1 mink=13 tbo ktrim=r qtrim=rl trimq=20 ref=./resources/nextera.fa
```



Decontamination statistics:

```
reads mapped with bowtie to metagenome 'contaminants'
```



Mapping statistics:
STAR - multimappers discarded

