

MP10k_pe

Summary

General

fastp version:	0.19.6 (https://github.com/OpenGene/fastp)
sequencing:	paired end (151 cycles + 151 cycles)
mean length before filtering:	111bp, 112bp
mean length after filtering:	117bp, 115bp
duplication rate:	45.214608%
Insert size peak:	151

Before filtering

total reads:	185.190824 M
total bases:	20.795627 G
Q20 bases:	19.431985 G (93.442648%)
Q30 bases:	17.831500 G (85.746393%)
GC content:	43.291273%

After filtering

total reads:	153.613798 M
total bases:	17.835487 G
Q20 bases:	16.975903 G (95.180481%)
Q30 bases:	15.675142 G (87.887375%)
GC content:	43.037997%

Filtering result

reads passed filters:	153.613798 M (82.948925%)
reads with low quality:	12.159456 M (6.565906%)
reads with too many N:	1.128000 K (0.000609%)
reads too short:	19.079170 M (10.302438%)
reads with low complexity:	337.272000 K (0.182121%)

Adapters

Adapter or bad ligation of read1

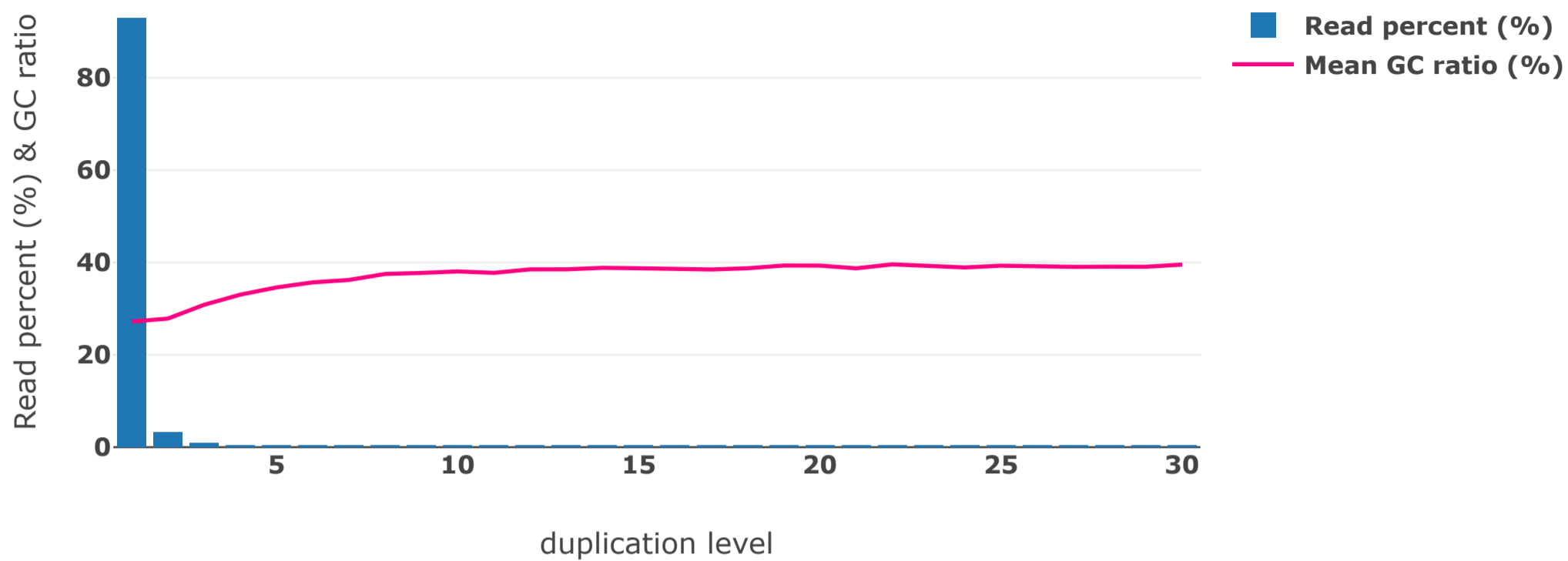
Sequence	Occurrences
all adapter sequences	9733269

Adapter or bad ligation of read2

Sequence	Occurrences
all adapter sequences	11069653

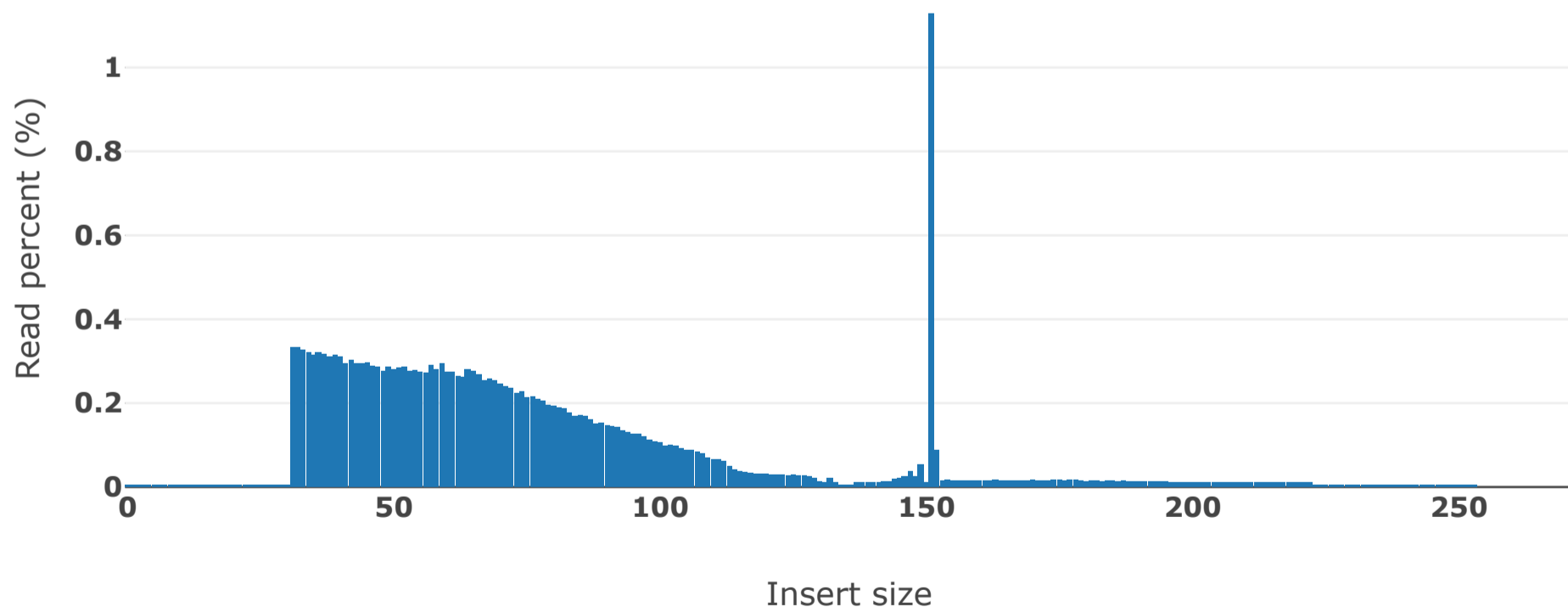
Duplication

duplication rate (45.214608%)



Insert size estimation

Insert size distribution (79.330470% reads are with unknown length)

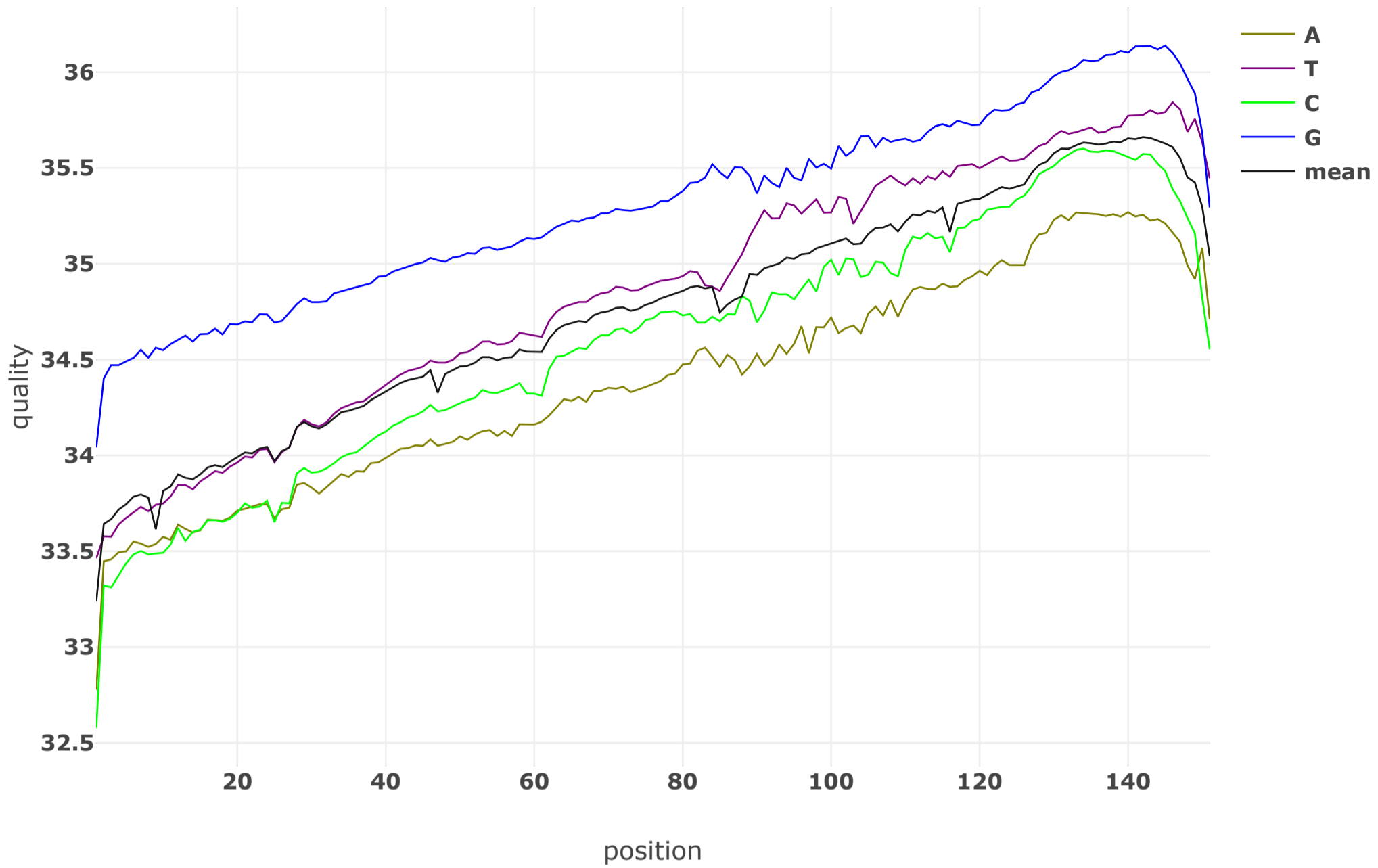


This estimation is based on paired-end overlap analysis, and there are 79.330470% reads found not overlapped. The nonoverlapped read pairs may have insert size <30 or >272, or contain too much sequencing errors to be detected as overlapped.

Before filtering

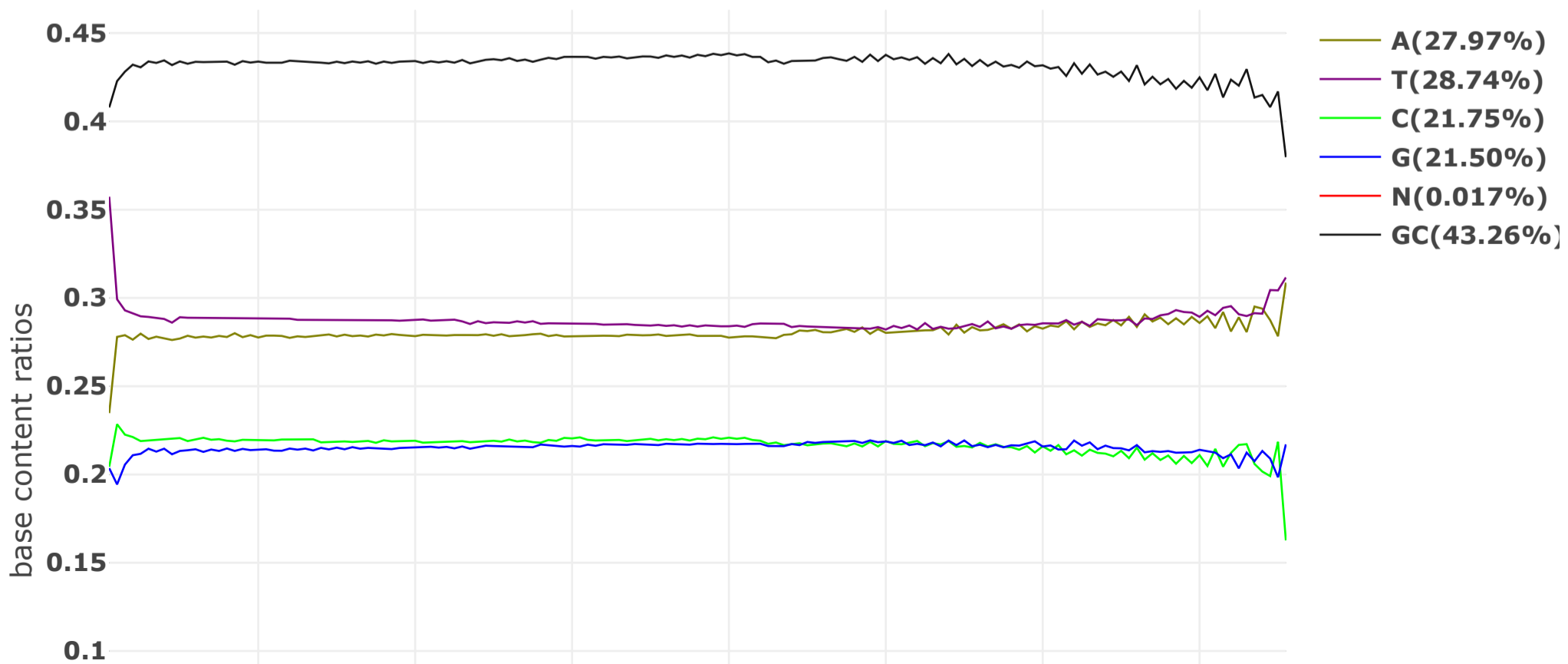
Before filtering: read1: quality

Value of each position will be shown on mouse over.



Before filtering: read1: base contents

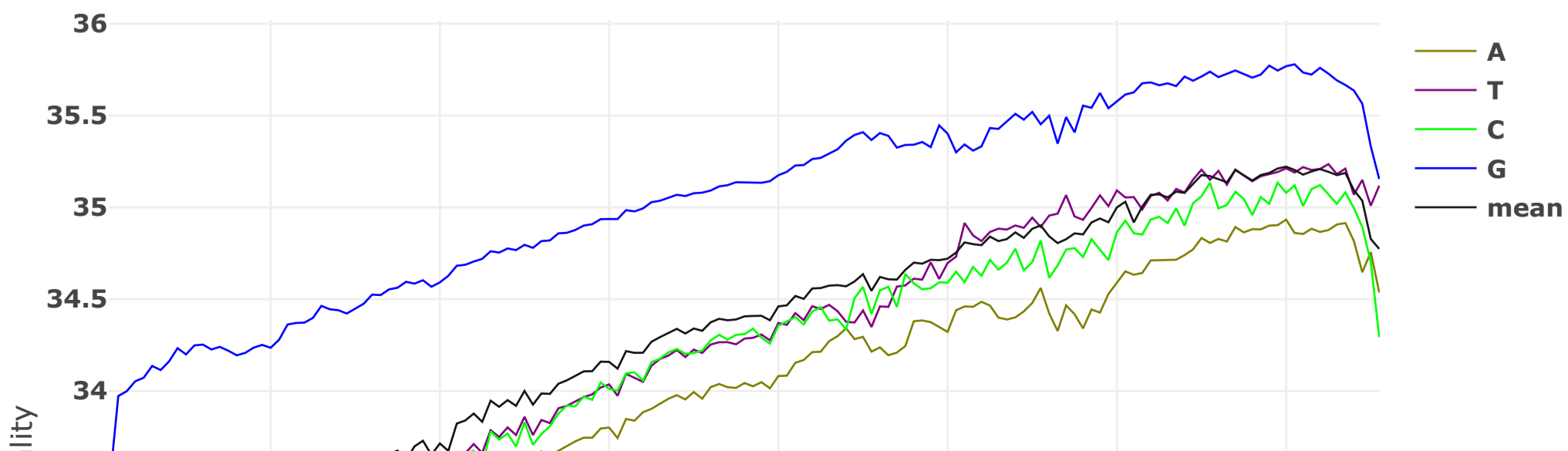
Value of each position will be shown on mouse over.



ACACACACACACACACACACACACACACACACACAT	349 (0.002700%)	
AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	2198 (0.017007%)	
CACACACACACACACACACACACACACACACACACA	10100 (0.078147%)	
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	19432 (0.150353%)	
CC	771 (0.0014914%)	
CC	1384 (0.026771%)	
CC	19 (0.000368%)	
CC	1245 (0.024083%)	
CC	14 (0.000271%)	
CC	6 (0.000116%)	
CTCACACACACACACACACACACACACACACACACA	6999 (0.054154%)	
CTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	3774 (0.029201%)	
CTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	8326 (0.064421%)	
GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGA	2615 (0.020233%)	
GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGC	459 (0.003551%)	
GT	17253 (0.133493%)	
TCACACACACACACACACACACACACACACACACAC	2679 (0.020728%)	
TCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTC	4898 (0.037898%)	
TGA	254 (0.001965%)	
TG	7888 (0.061032%)	
TTGT	3641 (0.028172%)	
TTG	10276 (0.079509%)	

Before filtering: read2: quality

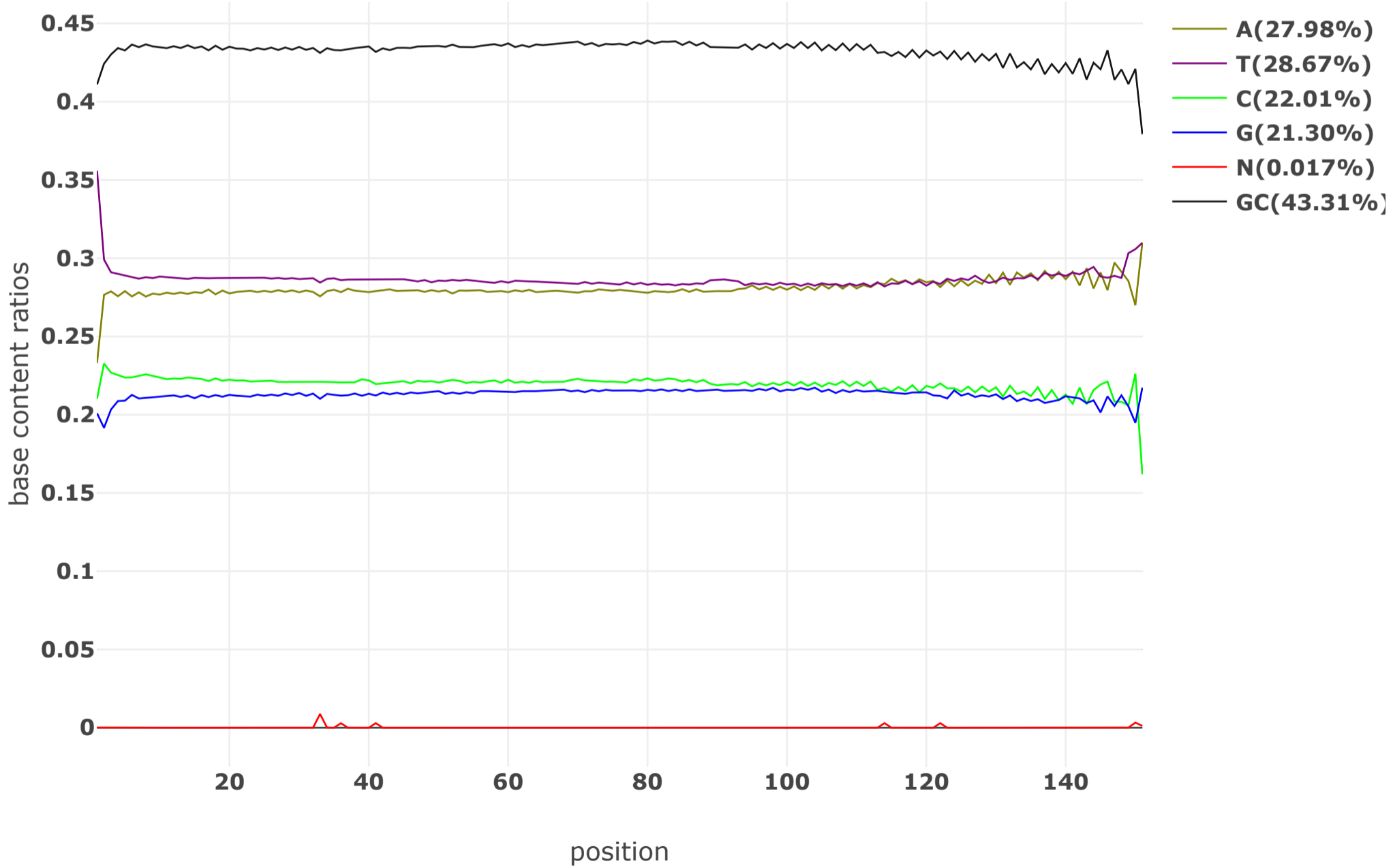
Value of each position will be shown on mouse over.





Before filtering: read2: base contents

Value of each position will be shown on mouse over.



Before filtering: read2: KMER counting

Darker background means larger counts. The count will be shown on mouse over.

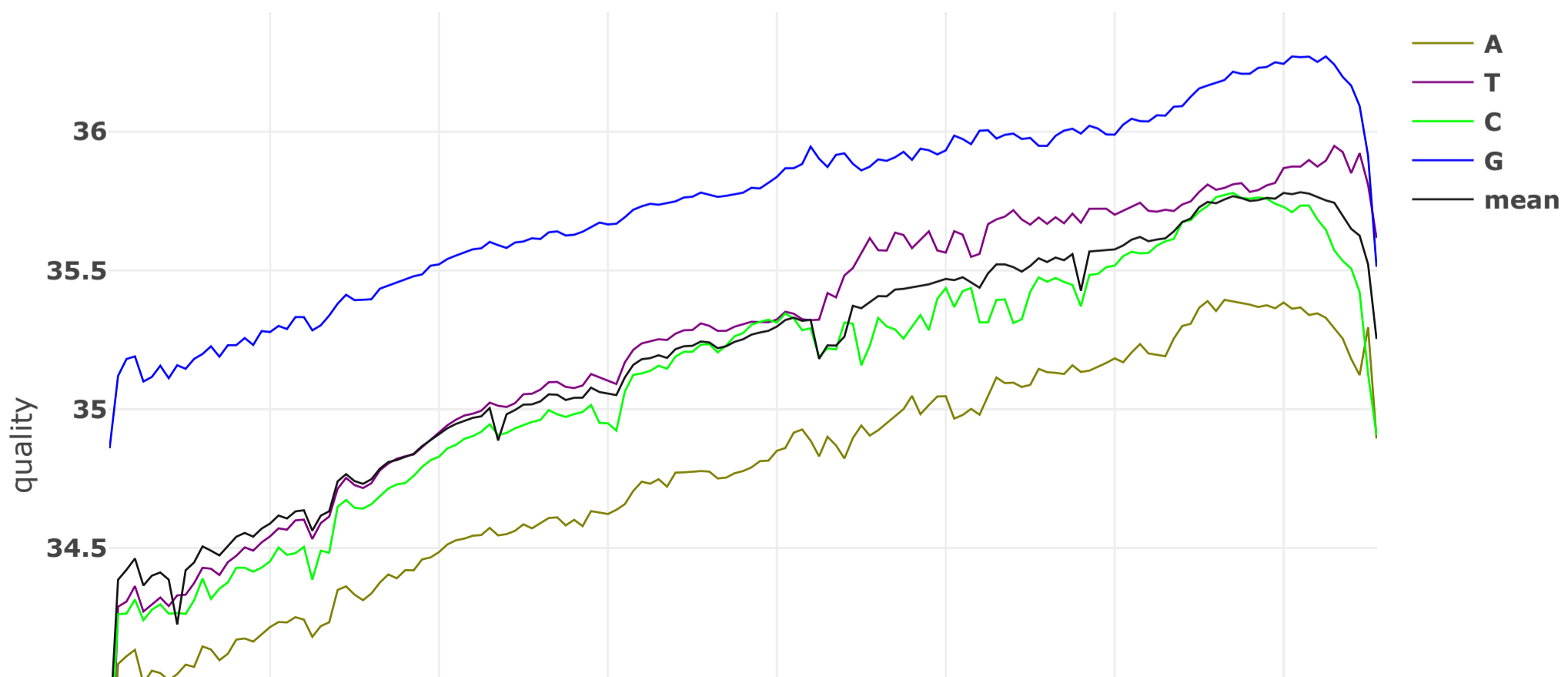
	AA	AT	AC	AG	TA	TT	TC	TG	CA	CT	CC	CG	GA	GT	GC	GG
AAA	AAAAA	AAAAT	AAAAC	AAAAG	AAATA	AAATT	AAATC	AAATG	AAACA	AAACT	AAACC	AAACG	AAAGA	AAAGT	AAAGC	AAAGG
AAT	AATAA	AATAT	AATAC	AATAG	AATTA	AATTT	AATTC	AATTG	AATCA	AATCT	AATCC	AATCG	AATGA	AATGT	AATGC	AATGG
AAC	AACAA	AACAT	AACAC	AACAG	AAC TA	AAC TT	AAC TC	AAC TG	AAC CA	AAC CT	AAC CC	AAC CG	AAC GA	AAC GT	AAC GC	AAC GG

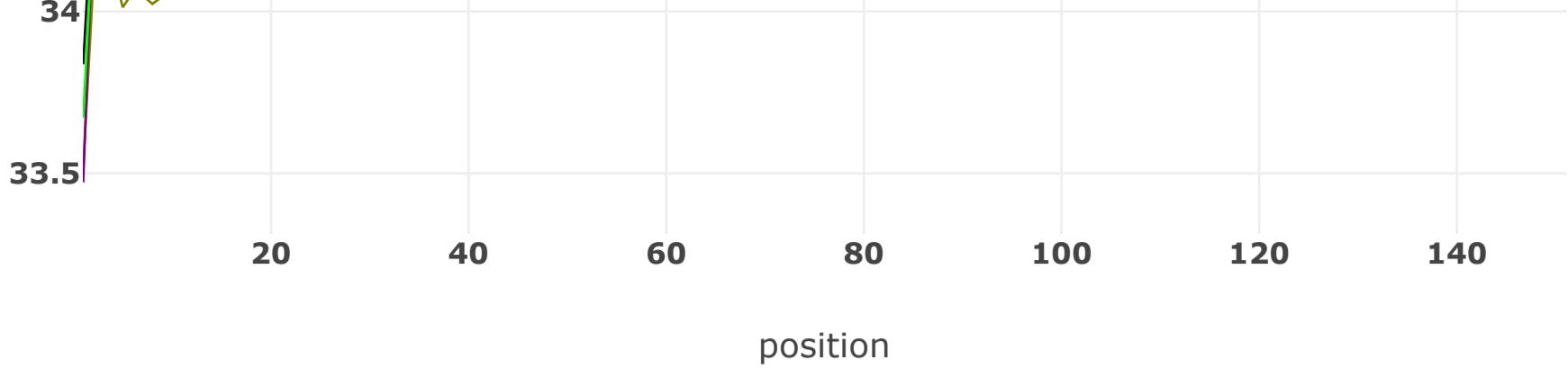
CTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	2823 (0.021599%)	
GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGA	3201 (0.024491%)	
GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGC	472 (0.003611%)	
GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	15692 (0.120059%)	
GTGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	4614 (0.035302%)	
GTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	822 (0.006289%)	
TATCTATCTATCTATCTATCTATCTATCTATCTATCTATC	530 (0.004055%)	
TCACACACACACACACACACACACACACACACACACAC	2459 (0.018814%)	
TCTATCTATCTATCTATCTATCTATCTATCTATCTATCTA	161 (0.001232%)	
TCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTC	4497 (0.034406%)	
TCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG	5530 (0.042310%)	
TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGA	197 (0.001507%)	
TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG	6641 (0.050810%)	
TGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG	1447 (0.011071%)	
TTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	3241 (0.024797%)	
TTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG	4072 (0.031155%)	
TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	973 (0.007444%)	

After filtering

After filtering: read1: quality

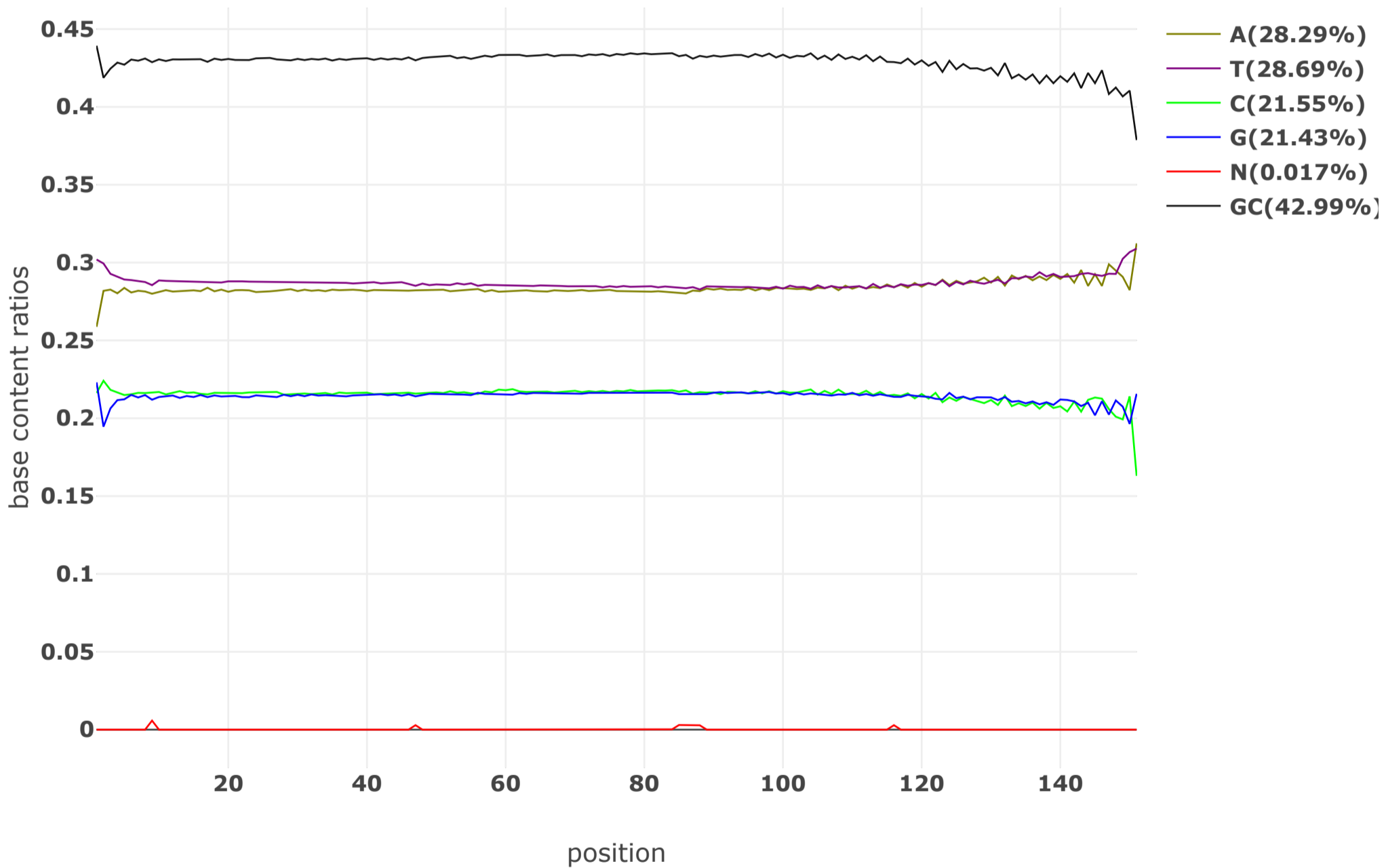
Value of each position will be shown on mouse over.





After filtering: read1: base contents

Value of each position will be shown on mouse over.



After filtering: read1: KMER counting

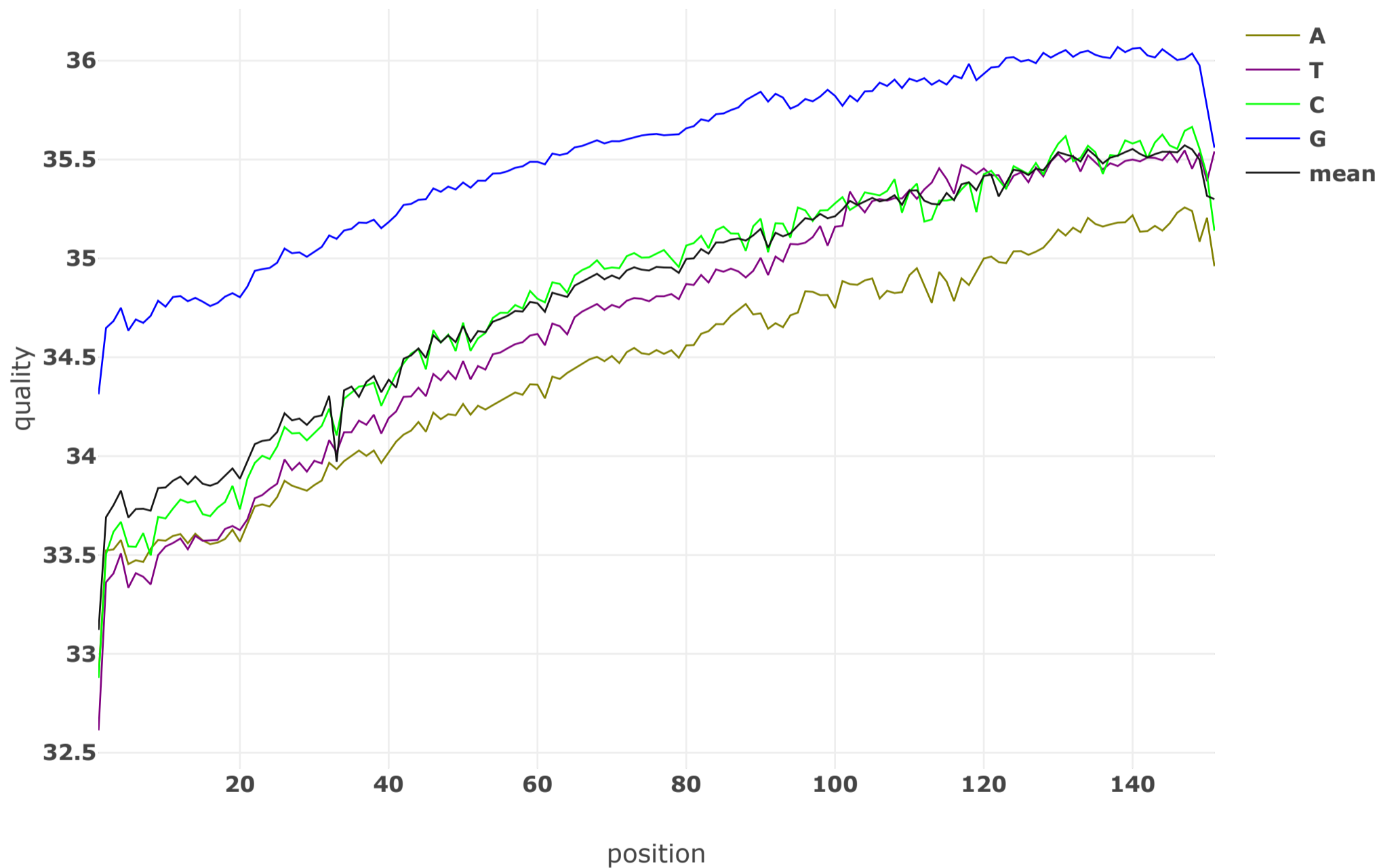
Darker background means larger counts. The count will be shown on mouse over.

	AA	AT	AC	AG	TA	TT	TC	TG	CA	CT	CC	CG	GA	GT	GC	GG
AAA	AAAAA	AAAAT	AAAAC	AAAAG	AAATA	AAATT	AAATC	AAATG	AAACA	AAACT	AAACC	AAACG	AAAGA	AAAGT	AAAGC	AAAGG
AAT	AATAA	AATAT	AATAC	AATAG	AATTA	AATTT	AATTC	AATTG	AATCA	AATCT	AATCC	AATCG	AATGA	AATGT	AATGC	AATGG
AAC	AACAA	AACAT	AACAC	AACAG	AACTA	AACTT	AACTC	AACTG	AACCA	AACCT	AACCC	AACCG	AACGA	AACGT	AACGC	AACGG
AAG	AAGAA	AAGAT	AAGAC	AAGAG	AAGTA	AAGTT	AAGTC	AAGTG	AAGCA	AAGCT	AAGCC	AAGCG	AAGGA	AAGGT	AAGGC	AAGGG
ATA	ATAAA	ATAAT	ATAAC	ATAAG	ATATA	ATATT	ATATC	ATATG	ATACA	ATACT	ATACC	ATACG	ATAGA	ATAGT	ATAGC	ATAGG
ATT	ATTAA	ATTAT	ATTAC	ATTAG	ATTTA	ATTTT	ATTTC	ATTTG	ATTCA	ATTCT	ATTCC	ATTCCG	ATTGA	ATTGT	ATTGC	ATTGG
ATC	ATCAA	ATCAT	ATCAC	ATCAG	ATCTA	ATCTT	ATCTC	ATCTG	ATCCA	ATCCT	ATCCC	ATCCG	ATCGA	ATCGT	ATCGC	ATCGG
ATG	ATGAA	ATGAT	ATGAC	ATGAG	ATGTA	ATGTT	ATGTC	ATGTG	ATGCA	ATGCT	ATGCC	ATGCG	ATGGA	ATGGT	ATGGC	ATGGG
ACA	ACAAA	ACAAT	ACAAC	ACAAG	ACATA	ACATT	ACATC	ACATG	ACACA	ACACT	ACACC	ACACG	ACAGA	ACAGT	ACAGC	ACAGG
ACT	ACTAA	ACTAT	ACTAC	ACTAG	ACTTA	ACTTT	ACTTC	ACTTG	ACTCA	ACTCT	ACTCC	ACTCG	ACTGA	ACTGT	ACTGC	ACTGG
ACC	ACCAA	ACCAT	ACCAC	ACCAG	ACCTA	ACCTT	ACCTC	ACCTG	ACCCA	ACCTT	ACCCC	ACCCG	ACCGA	ACCGT	ACCGC	ACCGG
ACG	ACGAA	ACGAT	ACGAC	ACGAG	ACGTA	ACGTT	ACGTC	ACGTG	ACGCA	ACGCT	ACGCC	ACGCG	ACGGA	ACGGT	ACGGC	ACGGG
AGA	AGAAA	AGAAT	AGAAC	AGAAG	AGATA	AGATT	AGATC	AGATG	AGACA	AGACT	AGACC	AGACG	AGAGA	AGAGT	AGAGC	AGAGG
AGT	AGTAA	AGTAT	AGTAC	AGTAG	AGTTA	AGTTT	AGTTC	AGTTG	AGTCA	AGTCT	AGTCC	AGTCG	AGTGA	AGTGT	AGTGC	AGTGG

TGTA	199 (0.001771%)	
TG	6609 (0.058816%)	
TTGT	2589 (0.023040%)	
TTTG	7157 (0.063693%)	

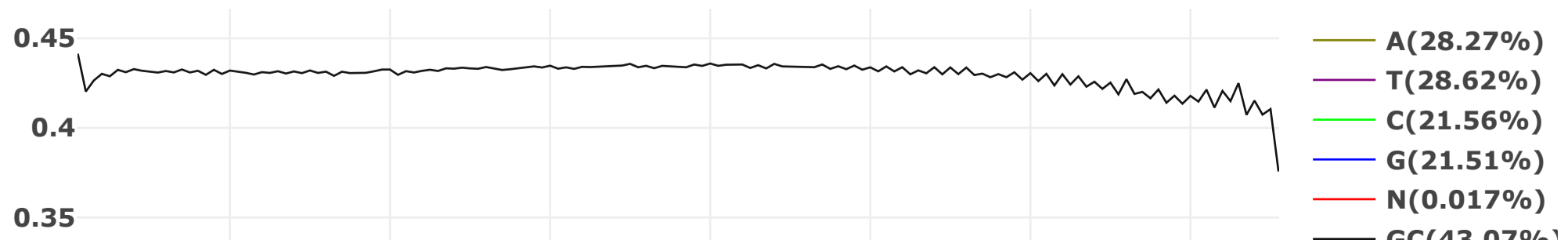
After filtering: read2: quality

Value of each position will be shown on mouse over.



After filtering: read2: base contents

Value of each position will be shown on mouse over.



GCC	GCCAA	GCCAT	GCCAC	GCCAG	GCCTA	GCCTT	GCCTC	GCCTG	GCCCA	GCCCT	GCCCC	GCCCG	GCCGA	GCCGT	GCCGC	GCCGG
GCG	GCGAA	GCGAT	GCGAC	GCGAG	GCGTA	GCGTT	GCGTC	GCGTG	GCGCA	GCGCT	GCGCC	GCGCG	GCGGA	GCGGT	GCGGC	GCGGG
GGA	GGAAA	GGAAT	GGAAC	GGAAG	GGATA	GGATT	GGATC	GGATG	GGACA	GGACT	GGACC	GGACG	GGAGA	GGAGT	GGAGC	GGAGG
GGT	GGTAA	GGTAT	GGTAC	GGTAG	GGTTA	GGTTT	GGTTC	GGTTG	GGTCA	GGTCT	GGTCC	GGTCG	GGTGA	GGTGT	GGTGC	GGTGG
GGC	GGCAA	GGCAT	GGCAC	GGCAG	GGCTA	GGCTT	GGCTC	GGCTG	GGCCA	GGCCT	GGCCC	GGCCG	GGCGA	GGCGT	GGCGC	GGCGG
GGG	GGGAA	GGGAT	GGGAC	GGGAG	GGGTA	GGGTT	GGGTC	GGGTG	GGGCA	GGGCT	GGGCC	GGGCG	GGGGA	GGGGT	GGGGC	GGGGG

After filtering: read2: overrepresented sequences

Sampling rate: 1 / 20

overrepresented sequence	count (% of bases)	distribution: cycle 1 ~ cycle 151
AC	12683 (0.114699%)	
AG	2534 (0.022916%)	
ATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCT	107 (0.000968%)	
CA	7352 (0.066488%)	
CC	429 (0.003880%)	
CTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTAT	66 (0.000597%)	
CTCA	5430 (0.049107%)	
CT	2864 (0.025901%)	
CTGT	2195 (0.019851%)	
GA	2539 (0.022962%)	
GTGC	333 (0.003012%)	
GT	12889 (0.116562%)	
GTGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	3609 (0.032638%)	
GTTTGT	598 (0.005408%)	
TATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATC	443 (0.004006%)	
TCAC	1786 (0.016152%)	
TCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTA	123 (0.001112%)	
TC	3466 (0.031345%)	
TCTG	4319 (0.039059%)	
TGTA	164 (0.001483%)	
TG	6067 (0.054867%)	
TGTTTG	1068 (0.009659%)	
TTGT	2443 (0.022093%)	
TTTG	2962 (0.026787%)	
TT	31 (0.000280%)	

```
fastp -i /work/frr6/SHAD/NXTRIM/MP10k_R1.pe.fastq.gz -I /work/frr6/SHAD/NXTRIM/MP10k_R2.pe.fastq.gz -o MP10k_pe_F.trimmed.fq.gz
-o MP10k_pe_R.trimmed.fq.gz --detect_adapter_for_pe --cut_front --cut_tail --cut_window_size=4 --cut_mean_quality=20 --
qualified_quality_phred=20 --unqualified_percent_limit=30 --n_base_limit=5 --length_required=50 --low_complexity_filter --
complexity_threshold=30 --overrepresentation_analysis --json=MP10k_pe.json --html=MP10k_pe.html --report_title=MP10k_pe --
thread=8
```

fastp 0.19.6, at 2019-01-18 11:26:58

