

MP5k

Summary

General

fastp version:	0.19.6 (https://github.com/OpenGene/fastp)
sequencing:	paired end (151 cycles + 151 cycles)
mean length before filtering:	118bp, 119bp
mean length after filtering:	118bp, 119bp
duplication rate:	79.973156%
Insert size peak:	0

Before filtering

total reads:	144.453478 M
total bases:	17.207941 G
Q20 bases:	16.327942 G (94.886089%)
Q30 bases:	15.139712 G (87.980961%)
GC content:	42.858371%

After filtering

total reads:	138.605294 M
total bases:	16.481603 G
Q20 bases:	15.816498 G (95.964562%)
Q30 bases:	14.733968 G (89.396452%)
GC content:	42.823141%

Filtering result

reads passed filters:	138.605294 M (95.951510%)
reads with low quality:	5.687612 M (3.937331%)
reads with too many N:	1.332000 K (0.000922%)
reads too short:	134.194000 K (0.092898%)
reads with low complexity:	25.046000 K (0.017338%)

Adapters

Adapter or bad ligation of read1

The input has little adapter percentage (~0.016648%), probably it's trimmed before.

Sequence	Occurrences
all adapter sequences	29939

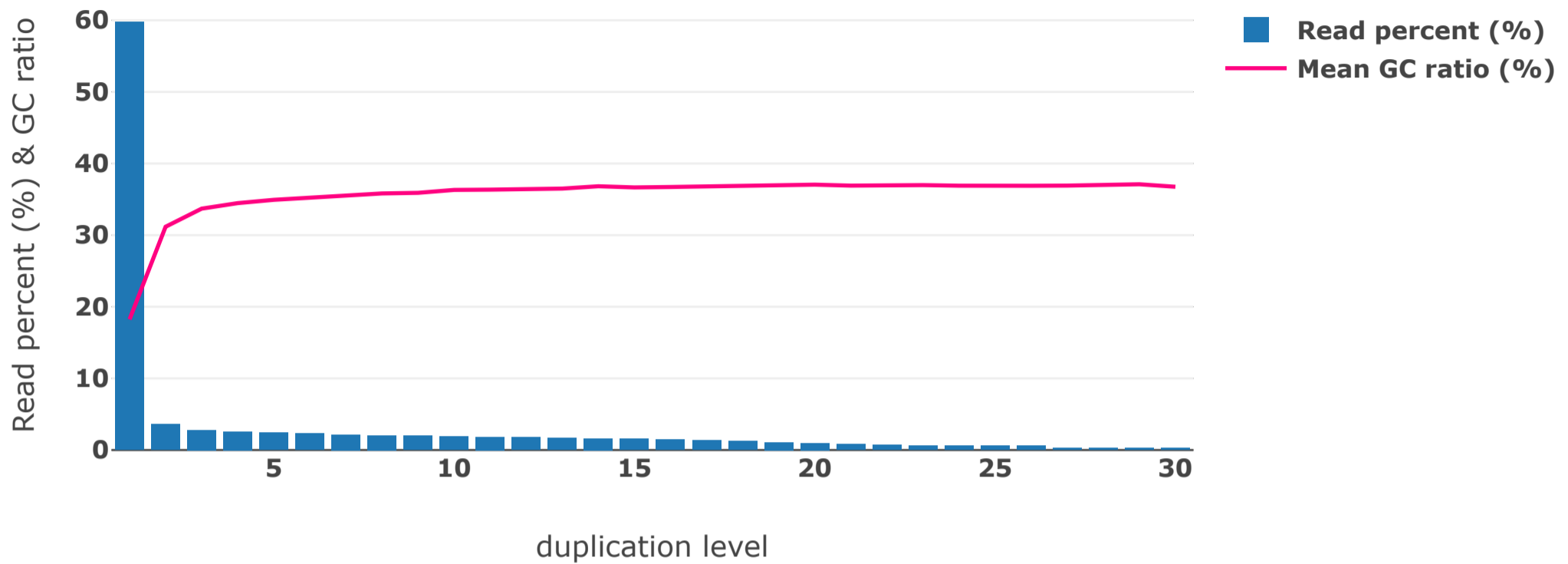
Adapter or bad ligation of read2

The input has little adapter percentage (~0.035893%), probably it's trimmed before.

Sequence	Occurrences
----------	-------------

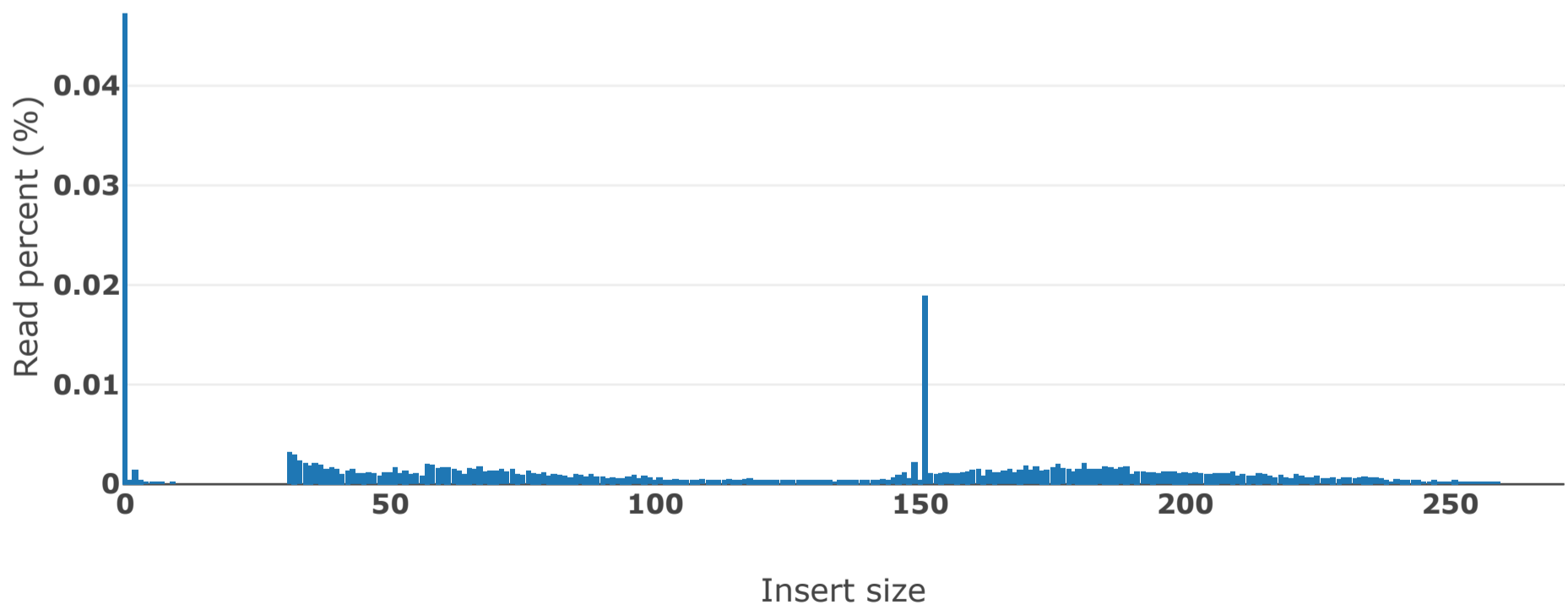
Duplication

duplication rate (79.973156%)



Insert size estimation

Insert size distribution (99.714437% reads are with unknown length)

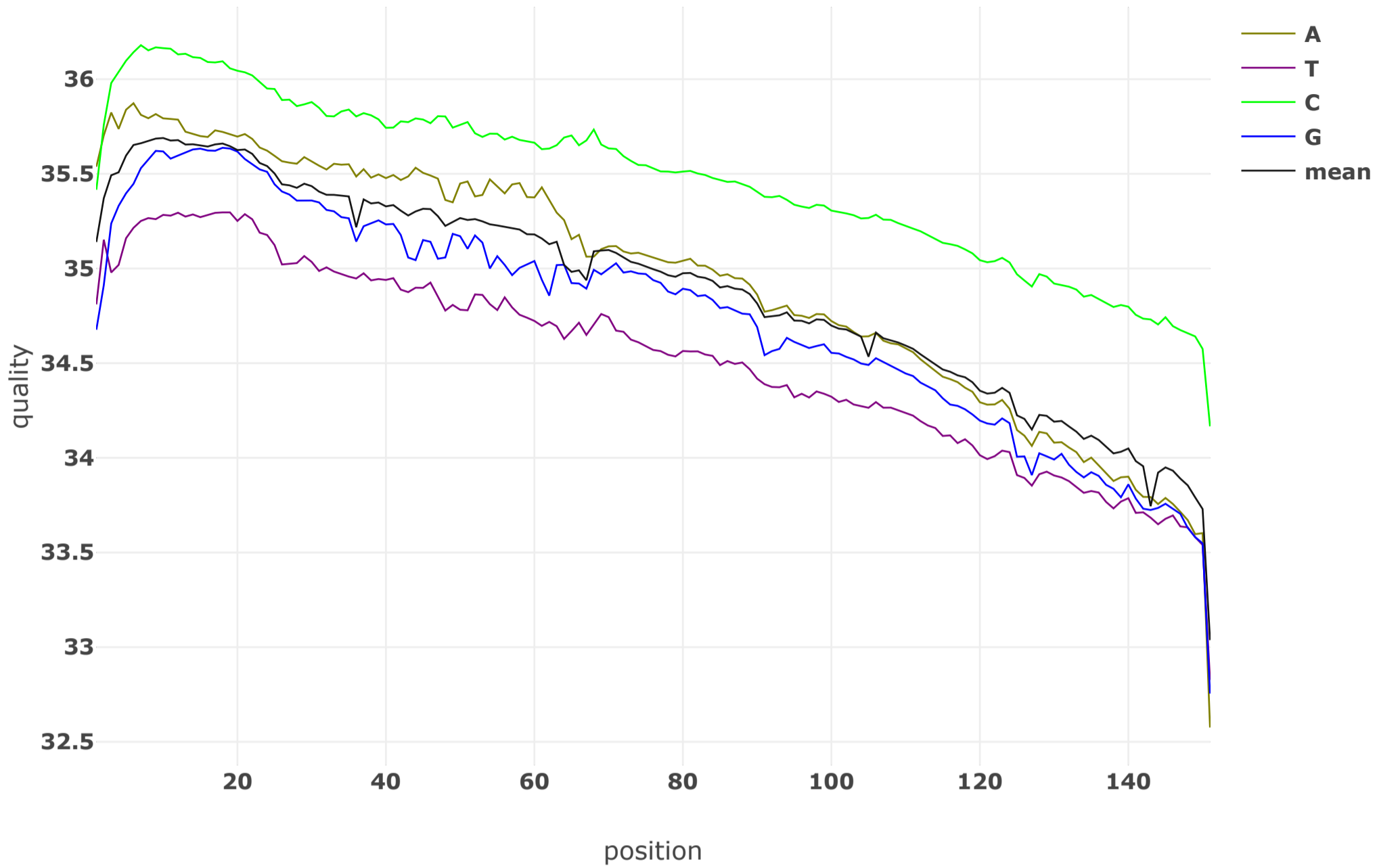


This estimation is based on paired-end overlap analysis, and there are 99.714437% reads found not overlapped. The nonoverlapped read pairs may have insert size <30 or >272, or contain too much sequencing errors to be detected as overlapped.

Before filtering

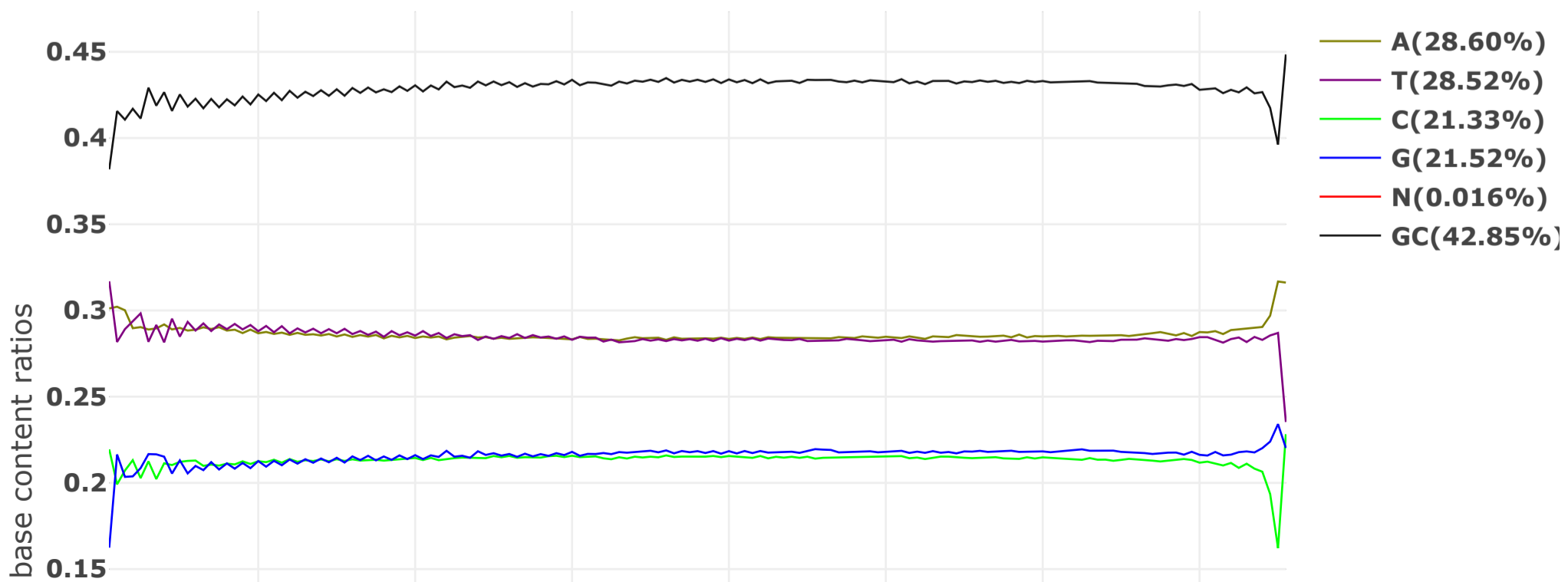
Before filtering: read1: quality

Value of each position will be shown on mouse over.



Before filtering: read1: base contents

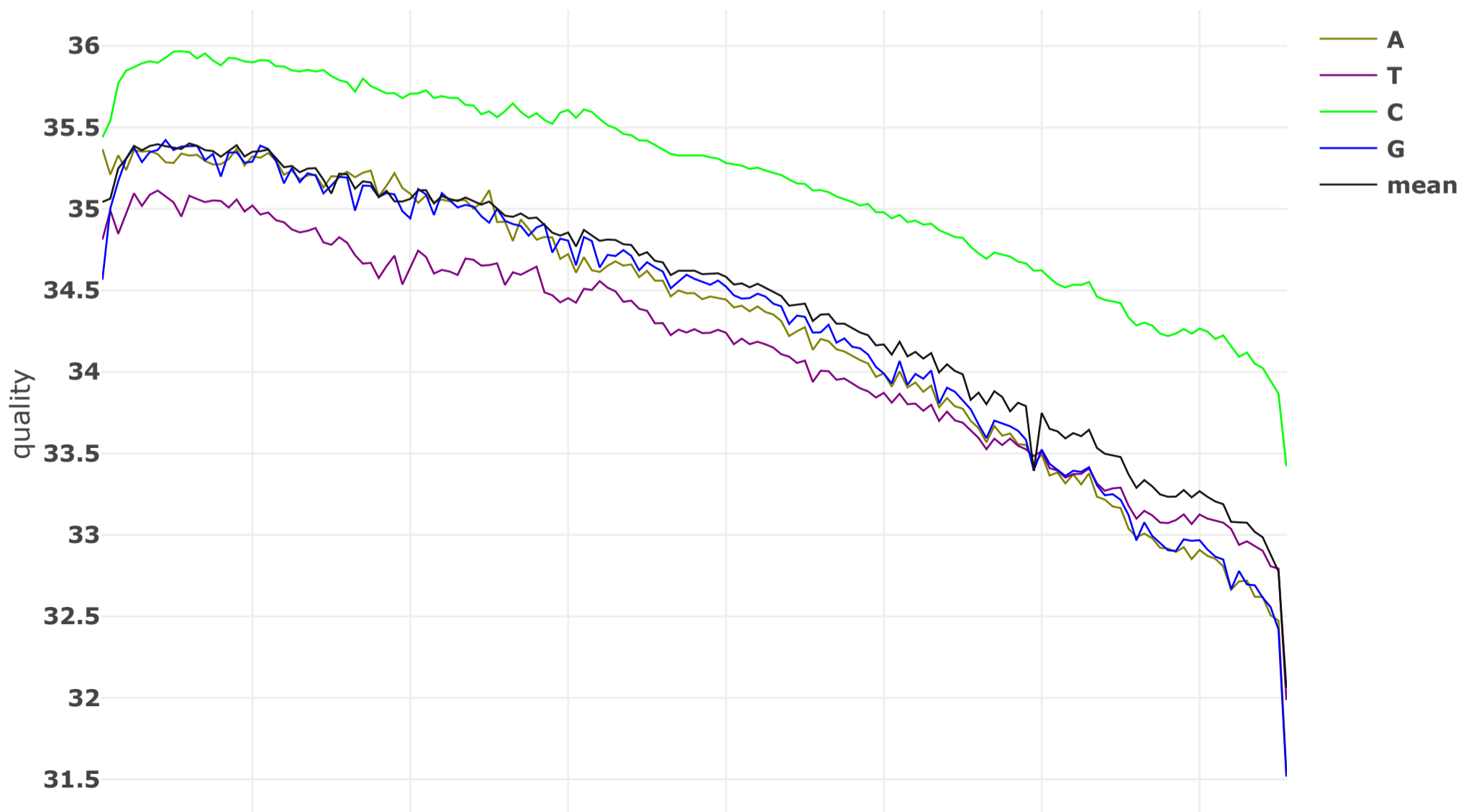
Value of each position will be shown on mouse over.



ACACACACACACACACACACACACACACACACACACAC	10310 (0.096409%)	
ACACACACACACACACACACACACACACACACACACAG	253 (0.002366%)	
AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	3715 (0.034739%)	
ATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	2110 (0.019731%)	
CACACACACACACACACACACACACACACACACACAAA	319 (0.002983%)	
CACACACACACACACACACACACACACACACACACACA	9176 (0.085805%)	
CTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1708 (0.015972%)	
CTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	5166 (0.048307%)	
GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGA	3035 (0.028380%)	
GCACACACACACACACACACACACACACACACACACAC	5913 (0.055293%)	
GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	9588 (0.089658%)	
TACACACACACACACACACACACACACACACACACACA	5429 (0.050767%)	
TATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG	3451 (0.032270%)	
TCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTC	2430 (0.022723%)	
TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG	6144 (0.057453%)	

Before filtering: read2: quality

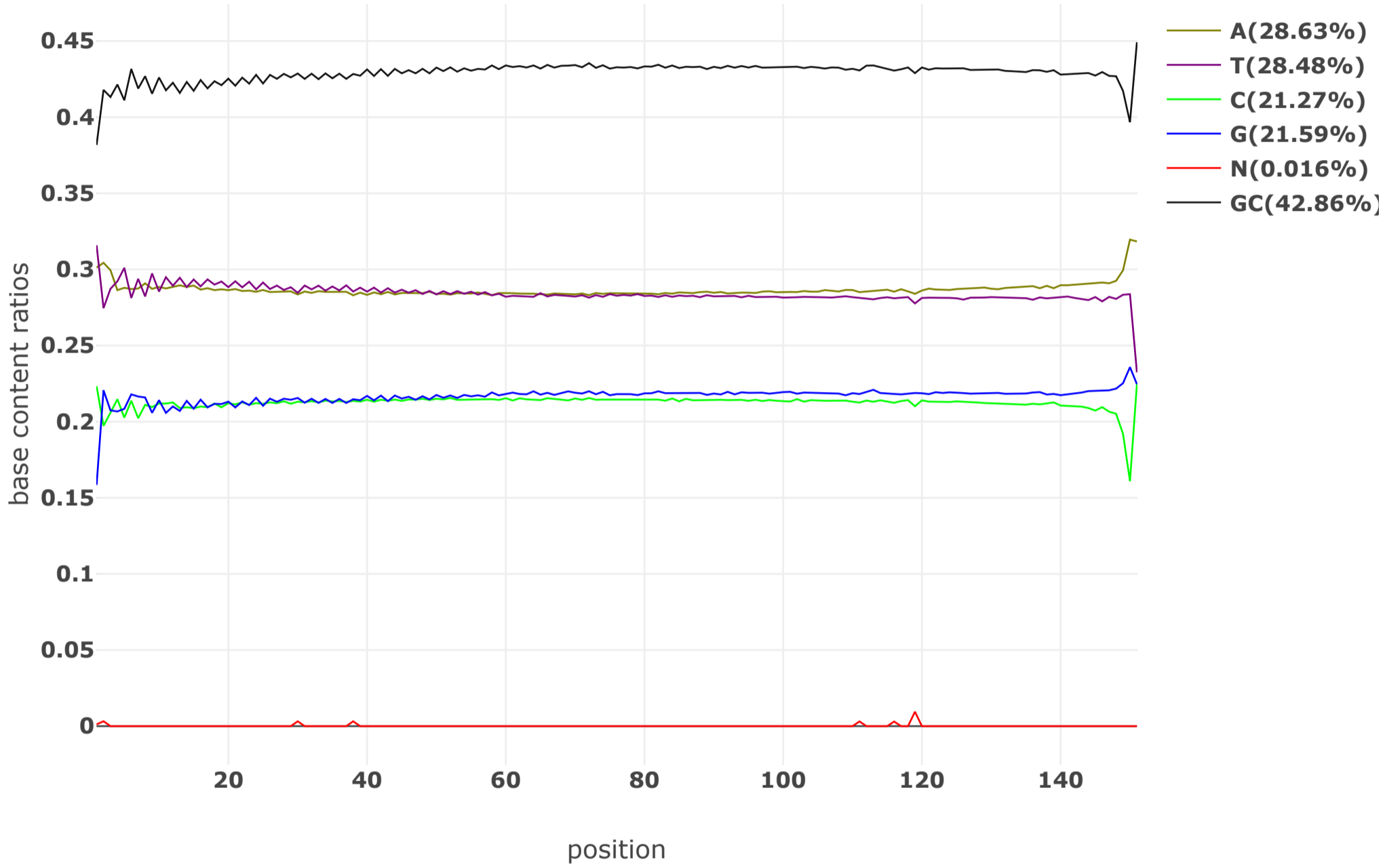
Value of each position will be shown on mouse over.



position

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.

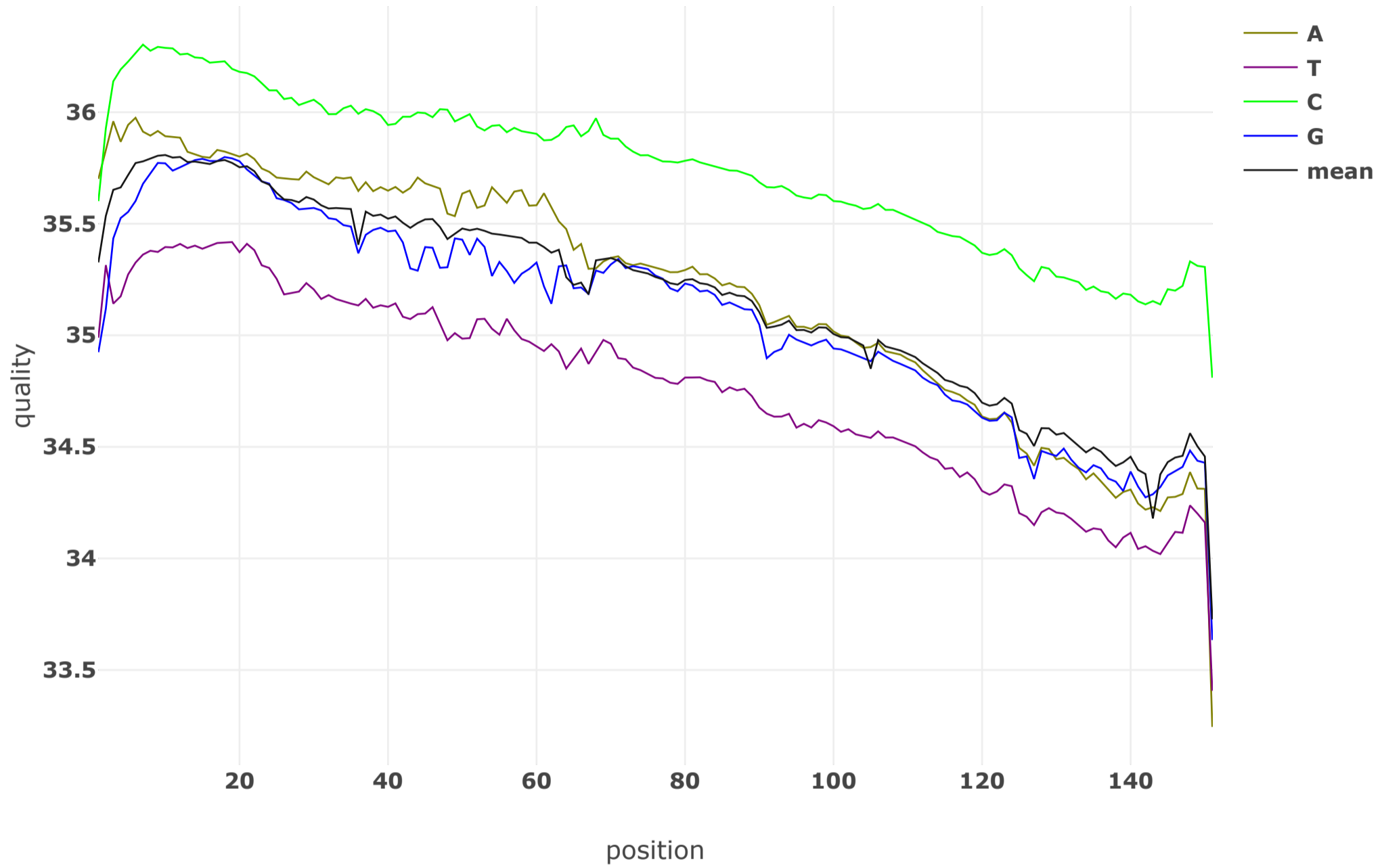
	AA	AT	AC	AG	TA	TT	TC	TG	CA	CT	CC	CG	GA	GT	GC	GG
AAA	AAAAA	AAAAT	AAAAC	AAAAG	AAATA	AAATT	AAATC	AAATG	AAACA	AAACT	AAACC	AAACG	AAAGA	AAAGT	AAAGC	AAAGG
AAT	AATAA	AATAT	AATAC	AATAG	AATTA	AATTT	AATTC	AATTG	AATCA	AATCT	AATCC	AATCG	AATGA	AATGT	AATGC	AATGG
AAC	AACAA	AACAT	AACAC	AACAG	AACTA	AACTT	AACTC	AACTG	AACCA	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	ATTTT	ATTTG	ATTCA	ATTCT	ATTCC	ATTTC	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
AGC	AGCAA	AGCAT	AGCAC	AGCAG	AGCTA	AGCTT	AGCTC	AGCTG	AGCCA	AGCCT	AGCCC	AGCCG	AGCGA	AGCGT	AGCGC	AGCGG
AGG	AGGAA	AGGAT	AGGAC	AGGAG	AGGTA	AGGTT	AGGTC	AGGTG	AGGCA	AGGCT	AGGCC	AGGCG	AGGGA	AGGGT	AGGGC	AGGGG
TAA	TAAAA	TAAAT	TAAAC	TAAAG	TAATA	TAATT	TAATC	TAATG	TAACA	TAACT	TAACC	TAACG	TAAGA	TAAGT	TAAGC	TAAGG
TAT	TATAA	TATAT	TATAC	TATAG	TATTA	TATTT	TATTC	TATTG	TATCA	TATCT	TATCC	TATCG	TATGA	TATGT	TATGC	TATGG
TAC	TACAA	TACAT	TACAC	TACAG	TACTA	TACTT	TACTC	TACTG	TACCA	TACCT	TACCC	TACCG	TACGA	TACGT	TACGC	TACGG
TAG	TAGAA	TAGAT	TAGAC	TAGAG	TAGTA	TAGTT	TAGTC	TAGTG	TAGCA	TAGCT	TAGCC	TAGCG	TAGGA	TAGGT	TAGGC	TAGGG
TTA	TTAAA	TTAAT	TTAAC	TTAAG	TTATA	TTATT	TTATC	TTATG	TTACA	TTACT	TTACC	TTACG	TTAGA	TTAGT	TTAGC	TTAGG
TTT	TTTAA	TTTAT	TTTAC	TTTAG	TTTTA	TTTTT	TTTTT	TTTTG	TTTCA	TTTCT	TTTCC	TTTCG	TTTGA	TTTGT	TTTGC	TTTGG
TTC	TTCAA	TTCAT	TTCAC	TTCAG	TTCTA	TTCTT	TTCTC	TTCTG	TTCCA	TTCCT	TTCCC	TTCCG	TTCGA	TTCGT	TTCGC	TTCGG
TTG	TTGAA	TTGAT	TTGAC	TTGAG	TTGTA	TTGTT	TTGTC	TTGTG	TTGCA	TTGCT	TTGCC	TTGCG	TTGGA	TTGGT	TTGGC	TTGGG



After filtering

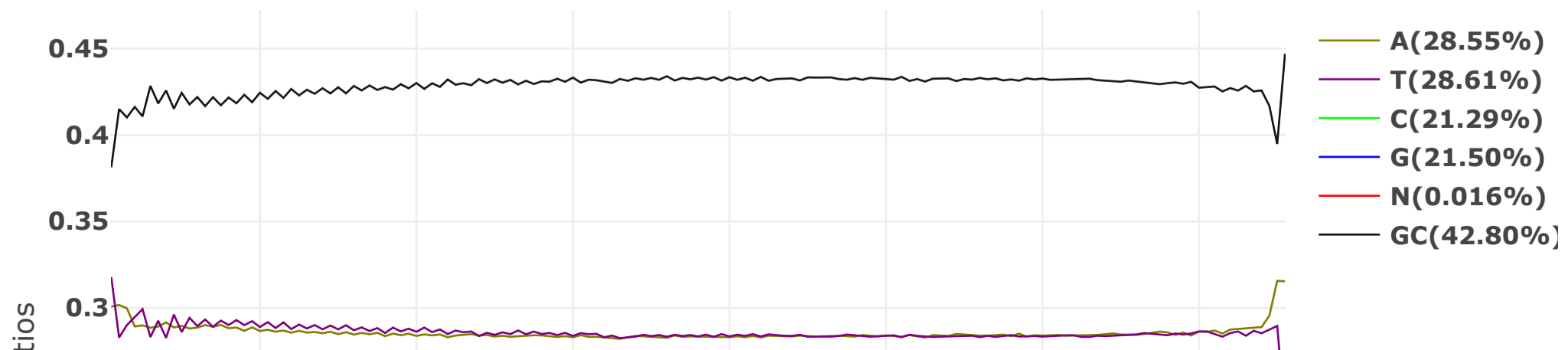
After filtering: read1: quality

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.



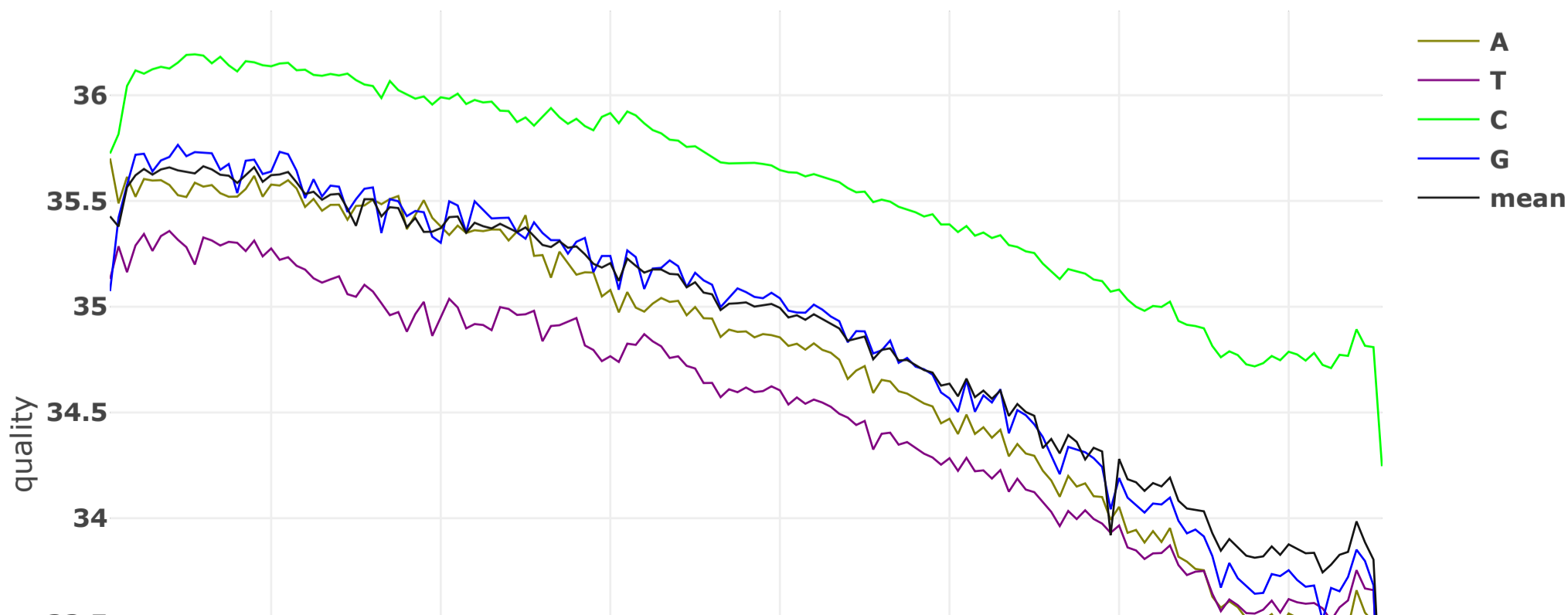
After filtering: read1: overrepresented sequences

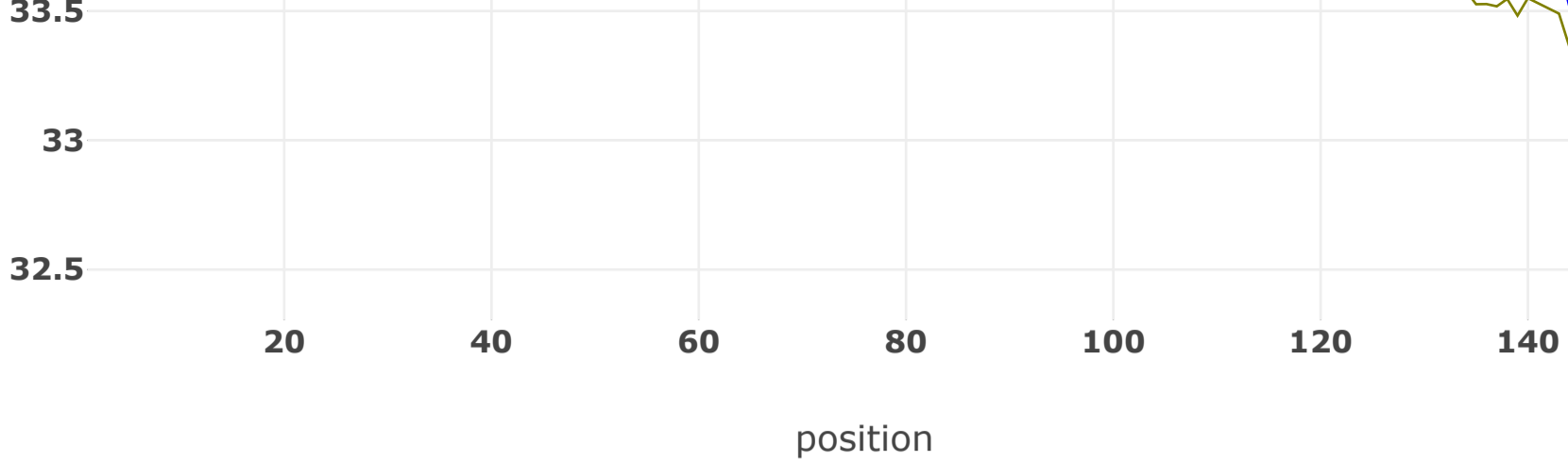
Sampling rate: 1 / 20

overrepresented sequence	count (% of bases)	distribution: cycle 1 ~ cycle 151
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	14 (0.000136%)	
ACACACACACACACACACACACACACACACAA	320 (0.003117%)	
ACACACACACACACACACACACACACACACACAC	9437 (0.091935%)	
ACACACACACACACACACACACACACACACACAG	241 (0.002348%)	
AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	3560 (0.034681%)	
ATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1986 (0.019348%)	
CACACACACACACACACACACACACACACACACAAA	281 (0.002737%)	
CACACACACACACACACACACACACACACACACACA	8398 (0.081813%)	
CTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1466 (0.014282%)	
CTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	4803 (0.046791%)	
GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGA	2908 (0.028330%)	
GCACACACACACACACACACACACACACACACACAC	5416 (0.052763%)	
GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	8922 (0.086918%)	
TACACACACACACACACACACACACACACACACACA	5105 (0.049733%)	
TATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG	3140 (0.030590%)	
TCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTC	2177 (0.021208%)	
TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG	5444 (0.053035%)	

After filtering: read2: quality

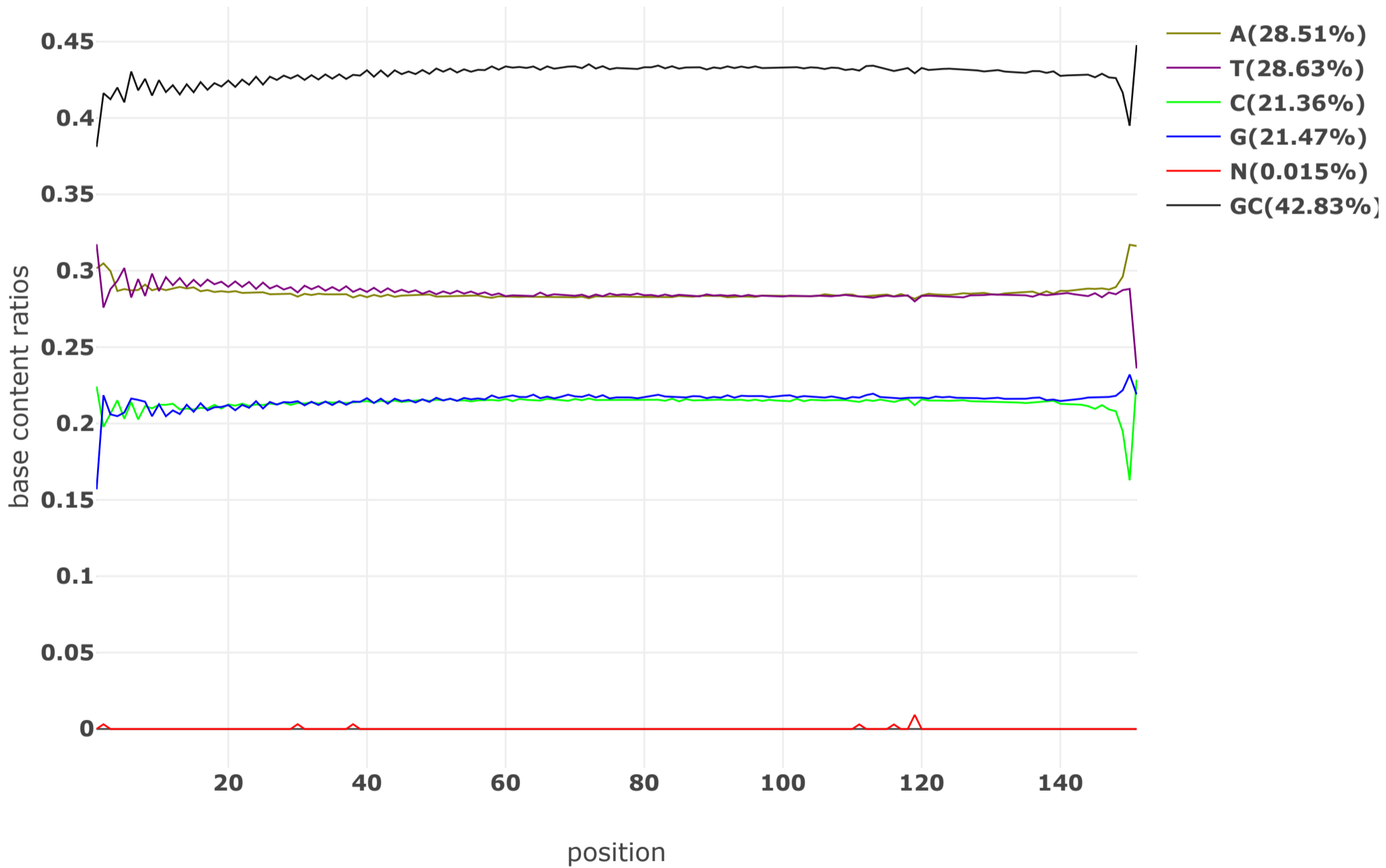
Value of each position will be shown on mouse over.





After filtering: read2: base contents





Value of each position will be shown on mouse over.



After 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	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	ATTTT	ATTTG	ATTTA	ATTTT	ATTTT	ATTTG	ATTTA	ATTTT	ATTTT	ATTTG
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	AC CCT	ACCCC	ACCCG	ACCGA	ACCGT	ACCGC	ACCGG

TCACACACACACACACACACACACACACACACACACAC	4217 (0.040795%)	
TCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTC	2978 (0.028809%)	
TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGAG	272 (0.002631%)	
TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG	11986 (0.115951%)	

```
fastp -i /work/frr6/SHAD/NXTRIM/MP5k_R1.mp.fastq.gz -I /work/frr6/SHAD/NXTRIM/MP5k_R2.mp.fastq.gz -o MP5k_F.trimmed.fq.gz -O
MP5k_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=MP5k.json --html=MP5k.html --report_title=MP5k --thread=8
```

fastp 0.19.6, at 2019-01-18 11:12:36