BIN 508: Next Generation Sequence Analysis and Informatics

Assignment 02

By: Taha Ahmad

Student Id: 2546126

Instructor: Dr Yesim Aydin Son

Question 1 A

Untrimmed

General Statistics

Copy table Configure columns	Scatter plot ■ Violin plot	Export as CSV	Showing $^2/_2$ rows and $^6/_6$ columns.	↔ Summarize table		
Sample Name	Dups GC		Avg len	Median len	Failed	Seqs
sample1_1	53.5 %	44.0 %	233 bp	251 bp	18 %	0.0 M
sample1_2	55.4 %	44.0 %	233 bp	251 bp	27 %	0.0 M

Trimmed General Statistics



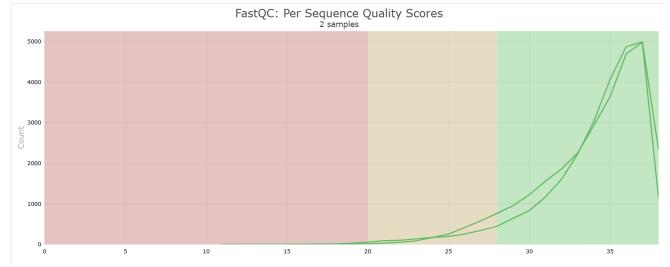
The MultiQC results show clear improvements after trimming—both reads became slightly shorter due to adapter and low-quality base removal, while the reverse read failure rate dropped by 9%. This confirms trimming successfully cleaned the data, boosting reliability for downstream analysis.



Untrimmed forward/reverse reads show good initial quality (Q>30) until ~200bp, with reverse degrading below Q25 at ends—typical for NGS. Post-trimming, reverse's severe drop resolves, though minor decline persists (Q26 at 230-240bp). Forward maintains high quality, confirming trimming effectively addressed end artifacts while preserving data utility.

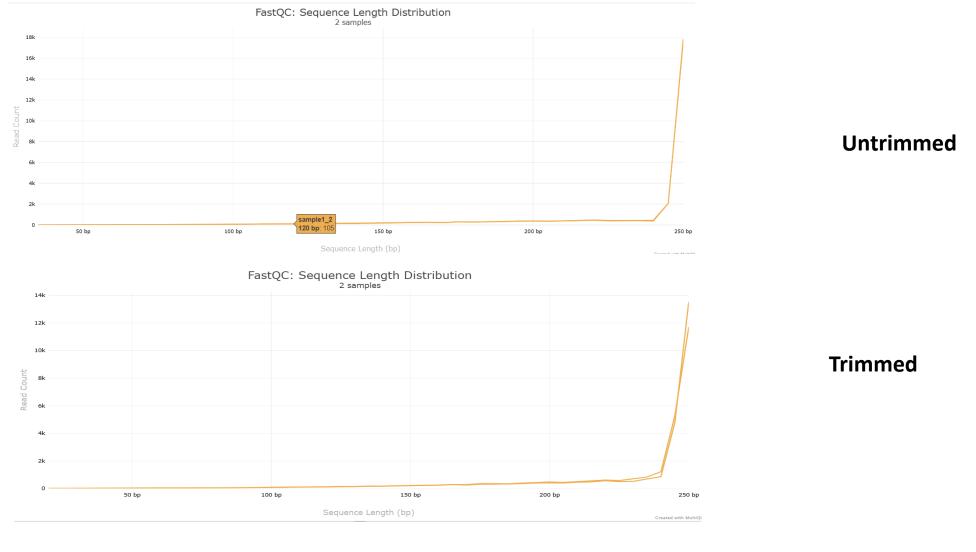


Untrimmed



Trimmed

The per-sequence quality scores show minimal overall changes post-trimming. However, reverse reads (R2) exhibit a slight increase in high-quality sequences: for example, reads with a PHRED score of 37 rose from 4,779 to 4,978 after trimming. This indicates targeted improvement in R2 reliability without major shifts in overall quality profiles.



While overall quality metrics show minimal changes, trimming significantly impacted read lengths. Both R1 and R2 exhibited reduced 250bp reads post-processing (from ~17.5k to 13.5k for R1 and 11.7k for R2). Trimmed data also displays greater variability in sequence lengths compared to raw reads, suggesting selective removal of low-quality/adapter-containing regions. These shifts align with expected trimming behavior, where shorter or fragmented reads are filtered to improve data reliability.

Cutadapt Report

```
(bioinfo) megatron@Taha:/mnt/d/HumanGenome$ cutadapt -q 20 -m 20 -o trimmed_1.fq.gz -p trimmed_2.fq.gz sample1_1.fq.gz sample1_2.fq.gz
This is cutadapt 4.0 with Python 3.9.21
Command line parameters: -q 20 -m 20 -o trimmed_1.fq.qz -p trimmed_2.fq.qz sample1_1.fq.qz sample1_2.fq.qz
Processing paired-end reads on 1 core ...
                               27,606 reads @ 28.2 μs/read; 2.13 M reads/minute
               00:00:00
Finished in 0.78 s (28 \mus/read; 2.11 M reads/minute).
=== Summary ===
Total read pairs processed:
                                        27,606
== Read fate breakdown ==
Pairs that were too short:
                                            14 (0.1%)
Pairs written (passing filters):
                                        27,592 (99.9%)
Total basepairs processed:
                              12,850,084 bp
  Read 1:
              6,424,378 bp
  Read 2:
              6,425,706 bp
Quality-trimmed:
                                 200,520 bp (1.6%)
  Read 1:
                 70,358 bp
                130,162 bp
  Read 2:
Total written (filtered):
                              12,646,002 bp (98.4%)
  Read 1:
              6,350,553 bp
  Read 2:
              6,295,449 bp
```

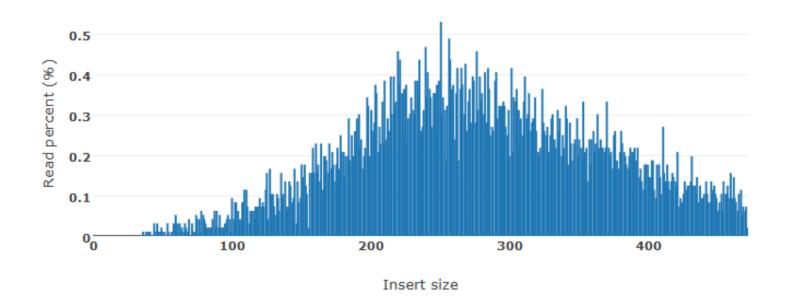
The Cutadapt output reveals R2 required significantly more trimming (130,162bp) compared to R1 (70,358bp), totaling 200,520bp removed due to low quality. An additional 14 bases were discarded for being too short. This aligns perfectly with the MultiQC data showing R2's consistently lower quality, explaining why more aggressive trimming was needed for the reverse reads to achieve quality standards.

fastp report

Summary	
General	
fastp version:	0.23.2 (https://github.com/OpenGene/fastp)
sequencing:	paired end (251 cycles + 251 cycles)
mean length before filtering:	232bp, 232bp
mean length after filtering:	231bp, 231bp
duplication rate:	0.550605%
Insert size peak:	250
Before filtering	
total reads:	55.212000 K
total bases:	12.850084 M
Q20 bases:	11.465327 M (89.223751%)
Q30 bases:	10.838418 M (84.345114%)
GC content:	44.618681%
After filtering	
total reads:	51.718000 K
total bases:	11.990111 M
Q20 bases:	10.909680 M (90.988983%)
Q30 bases:	10.382259 M (86.590183%)
GC content:	44.522524%
Filtering result	
reads passed filters:	51.718000 K (93.671666%)
reads with low quality:	3.376000 K (6.114613%)
reads with too many N:	118 (0.213722%)
reads too short:	0 (0.00000%)

The fastp analysis yielded several key findings. While the total read count decreased by approximately 4,000 after processing, this reduction corresponded with measurable quality improvements - the percentage of bases with Q20+ scores increased from 89% to 90.9%, with similar gains observed for Q30 bases. Notably, read lengths remained largely unchanged and duplication rates stayed exceptionally low at just 0.55%. The filtering removed 6% of reads for failing quality thresholds, along with 0.21% containing excessive ambiguous bases (Ns), though no reads were excluded for being too short. Importantly, the GC content distribution remained stable throughout processing.

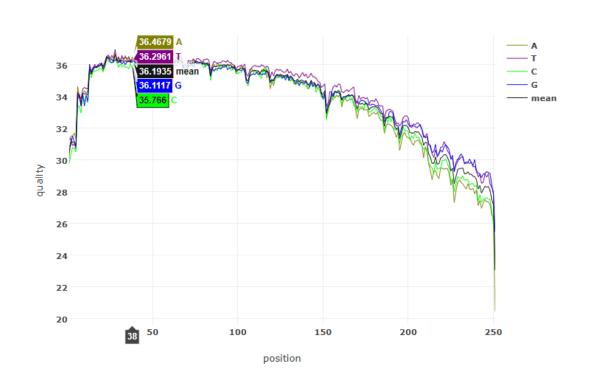
Insert size distribution (18.363523% reads are with unknown length)

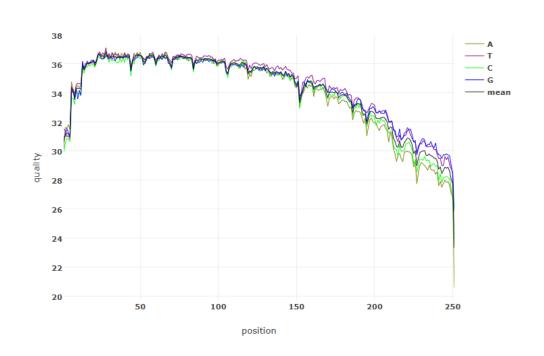


The insert size distribution peaks at 250bp (>0.5% of reads), while 18.36% couldn't be sized - likely fragments <30bp or >470bp based on the detectable range. This indicates some library preparation artifacts exist alongside the dominant fragment population.

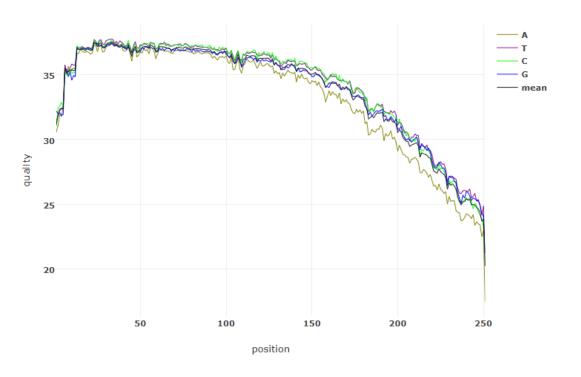


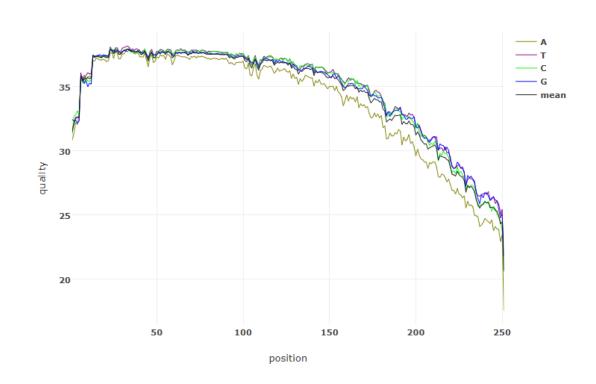






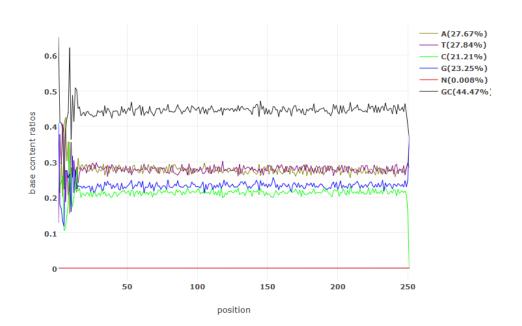
The quality scores across 250bp show nearly identical trends before and after filtering, with both datasets averaging Q27 at the end position. The consistent Y-axis range (20-30) confirms all recorded bases maintain Q20+ scores throughout. This stability indicates the filtering preserved quality while removing problematic reads without altering the remaining data's profile.

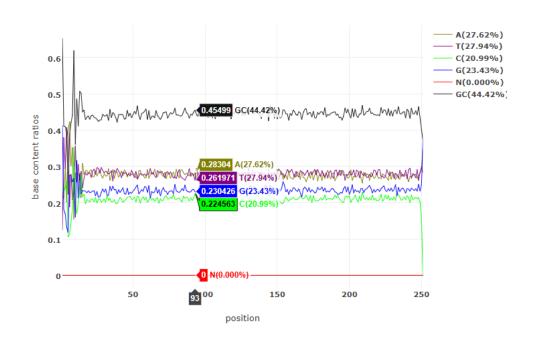




The analysis reveals minimal differences overall, though Read 2 shows a notable quality decline at the 250bp position. Specifically, the "A" nucleotide scores drop below Q20 at this endpoint, pulling down the average. This pattern differs from Read 1's more stable quality profile, highlighting Read 2's greater susceptibility to end-of-read degradation - a common phenomenon in paired-end sequencing that trimming couldn't fully resolve.







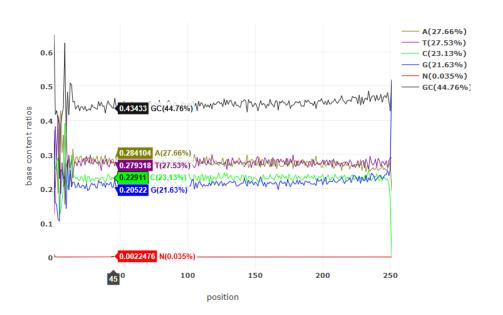
No significant differences in base contents visible in the data before and after filtering

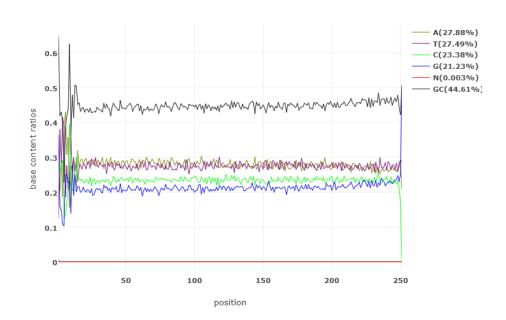
Value of each position will be shown on mouse over.

Value of each position will be shown on mouse over.









No significant differences in base contents visible in the data before and after filtering

Before filtering: read1: KMER counting

After	filtering:	read1:	KMER	counting
HT CET	TITICELING.	reaur.	1/1-III-I/	Countring

AA I	AA	AT	AC	AG	TA	TT	TC	TG	CA	CT	CC	CG	GA	GT	GC	G
														AAAGT		
T				AATAG										AATGT		
														AACGT		
ا ء														AACCT		
λĮ														ATAGT		
I I	ATTAA	ATTAT	ATTAC				ATTTC					ATTOG				
ا ٥	ATCAA								ATCCA						ATCCC	
۱ ء				ATGAG										ATGGT		
λļ														ACAGT		
Ξ.	ACTAA	ACTAT		ACTAG					ACCCA.					ACTGT		
ا ۽			ACCAC	ACCAG							ACGCC			ACCCT		
2														AGAGT		
	RIGHAR	AGAIAL	AGARIC NORTH	AGRAGE 2 CT2 C	ALARTA	ALARIT NOTES	AGRIC	ALERTIC	AGACA	ALAK-I NOROR	NORGO	AGACG	Alakiak 2.0002	AGTGT	REALEC	340
Z	BOCSS	AGCAT	AGCAC		ALCITA											
	ALC:CAR BOCKE						ACCTC					AGCCG			AGCGC	
G A	Aldida			TARAG										AGGGT		
2	TARAR															
-	TATAR	TATAT		TATAG		TACTT	TATTC	TACTO			TATCC			TATCT		TA
<u>.</u>	TACAR	TAGAT	TAGAC		TAGTA	TACIT							TAGGA			TE
à l	TAGAA			TTAAG					TTACA			TTACC				
î	TETAL			TTTAG										TITGT		
÷	TTCAL			TTCAG						TTCCT				TTCGT		
	TTGAA						TTCTC									77
9		TTGAT	TTCAC				TTCTC					TTGCG				
λ I	TCTAA			TCAAG			TCTTC		TCACA					TCTGT	TCAGC	
ċ	TOCAA			TOTAL												100
-	TOGAR	TOGAT					TOCTO				TOGOC	TCCCG				70
à	TCARA	TGAAT	TGAAC		TCATA							203.00	TCCCA			_
÷	TURRE			TCTAG			TGTTC					TGRCG				
ċ	TOTAL			TGCAG										TGCGT		
6				TOGAG								TOGGE				
ă	Cabab	CAAAT											CAAGA			CR
÷	Caraa						CATTC			CATCT						
ė	Caraa			CACAG			CACTO			CACCE						
G				CAGAG			CAGTO							CAGGT		
λ	CTARA	CTARE	CTAAC			CTATT	CTATC		CTACA		CTACC		CTAGA			
î	CTTAA	CTTAT					CTTTC					CTTCG				
ċ				CTCAG			CTCTC					CTCCC				
Ğ														CTCCT		
ž														CCAGT		
î	CCTAA	CCTAT	CCTAC						CCTCA			CCTCG				
ē		CCCAT							CCCCA							
6				COGAG			COCTO			CCGCT				CCCGT		
ă				CGAAG				CCATC			CGACC			CGAGT		
7	CCTAA		CCTAC				CCTTC				CCTCC			CCTCT		
ē	COCAA	CCCAT					COCTC						00000	COCCT	00000	07
6				CCCAC			COGTC						CCCCA	COCCT	COCCC	CV
ă i														CAACT		
-				CATAG					GATCA					CATCT		
2	GACAA	CACAT	GACAC				CACTO			CACCE	CACCC	CACCC	CACCA			CI
G	CACAA	CACAT		CACAC			CACTO					GACCC			GREGGE	CA
λ	GTAAA	GTAAT		GUARG			CTATC				GTACC		GTAGA			
	GTTAA		CTTAC				CITTIC		GTTCA		CTTCC	GTTCG	GITCA	CITCI		GI
		CTCAT	CTCAC	GTCAG		CICIT		CTCTC			CTCCC	GTCCC	GTOGA			GI
T	GTCAA				CTCTA								CTCCA	CTCCT		
T C		GTGAT	CTCAC					GCATC								
T C	GTCAA GTGAA	GTGAT	GTGAC	GCAAG	GCATA	200000										
T C G	GTCAA GTGAA GCAAA	CCAAT	CCAAC	GCAAG					GCTCA	GCTCT	GCTCC					
T C	GTCAA GTGAA GCAAA GCTAA	CCAAT	GCAAC		GCTTA	CCTTT	GCTTC	CCTTC				GCTCG	CCTGA	GCTGT	GCTGC	CC
T C S A	GTCAA GTGAA GCAAA GCTAA	GCRAT	GCARC GCTRC GCCRC	GCAAG GCTAG GCCAG	GCTTA	CCCTT	GCTTC	GCTTG	GCCCA	GCCCT	GCCCC	GCTCG	GCCGA	GCTGT	GCTGC	GC
T G A T C	GTCAA GTGAA GCAAA GCTAA GCCAA	GCRAT GCTAT GCCAT GCGAT	GCAAC GCTAC GCCAC GCGAC	GCAAG GCTAG GCCAG GCGAG	GCTTA GCCTA GCGTA	GCTTT GCCTT GCGTT	GCTTC GCCTC GCGTC	GCTTG GCCTG GCCTG	GCCCA	GCCCT	GCCCC	GCCCG	CCTGA CCCGA CCCGA	GCTGT GCCGT GCGGT	GCTGC GCCGC GCGGC	CC
C G A T C G	GTCAA GTGAA GCAAA GCTAA GCCAA GCGAA GCGAA	GCAAT GCTAT GCCAT GCGAT GCAAT	GCAAC GCTAC GCCAC GCGAC GCGAC	GCAAG GCTAG GCCAG GCGAG GCGAG	GCTTA GCCTA GCGTA GCGTA	GCTTT GCCTT GCGTT GGATT	GCTTC GCCTC GCGTC GGATC	GCTTG GCCTG GCGTG GCATG	GCCCA GCGCA GCACA	GCCCT GCGCT GCACT	GCCCC GCGCC GGACC	GCTCG GCCCG GCACG	GCTGA GCCGA GCGGA GCAGA	GCTGT GCCGT GCGGT GGAGT	GCTGC GCCGC GCGGC GGAGC	GC GC
T C G A	GTCAR GTGAR GCARA GCTAR GCCAR GCGAR GCGAR GCGAR	GCRAT GCTAT GCCAT GCGAT GCRAT GCRAT	GCAAC GCCAC GCCAC GCGAC GCAAC GCCAC	GCAAG GCTAG GCCAG GCGAG GGAAG GGTAG	GCTTA GCCTA GCGTA GGATA GGTTA	GCTTT GCCTT GCGTT GGATT GGTTT	GCTTC GCCTC GCGTC GGATC GGTTC	GCTTG GCCTG GCGTG GGATG GGTTG	GCCCA GCGCA GCACA GCTCA	GCCCT GCGCT GGACT GGTCT	GCCCC GCGCC GGACC GGTCC	GCTCG GCCCG GCACG GCTCG	GCTGA GCCGA GCGGA GGAGA GGTGA	GCTGT GCCGT GCGGT GGAGT	GCTGC GCCGC GCGGC GGAGC GGTGC	00

The k-mer analysis reveals consistent patterns between raw and filtered data, with no significant differences in sequence composition. The heatmap's similar shading intensity across both datasets indicates that filtering preserved the original k-mer distribution. This suggests the quality control process maintained the fundamental sequence characteristics while selectively removing problematic reads. The darkest cells correspond to the most frequent length combinations, which remain stable post-filtering.

After filtering: read2: KMER counting

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

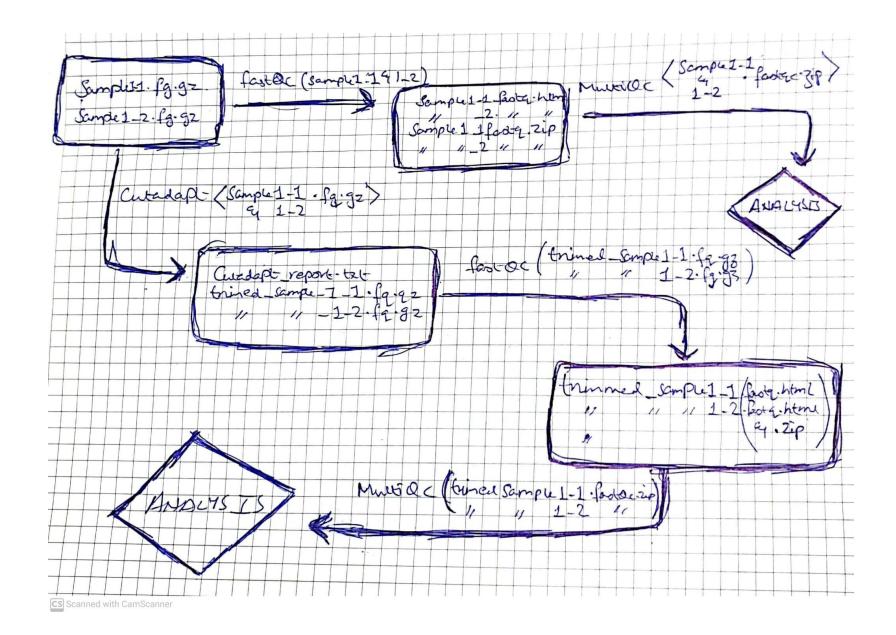
Dati		kground		larger	counts.		ount wi		hown or		over.					
	AA	AT	AC	AG	TA	TT	TC	TG	CA	CT	CC	CG	GA	GT	GC	GG
AAA	AAAAA	AAAAT	ARARC		AAATA	ARATT	AAATC	AAATG		AAACT			AAAGA	AAAGT	AAAGC	AAAGG
AAT	AATAA	AATAT	AATAC	AATAG	AATTA	AATIT	AATTC	AATTG	AATCA	AATCT	AATCC	AATCG	AATCA	AATGT	AATGC	AATGG
AAC	AACAA	AACAT	AACAC	AACAG	AACTA	AACTT			AACCA	AACCT	AACCC	AACCG	AACGA	AACGT	AACGC	ARCCC
AAG	AAGAA	AACAT	AAGAC	AAGAG	AAGTA	AAGTT	AAGTC		AAGCA	AAGCT	AAGCC	AAGCG	AACCA	AAGGT	AAGGC	AAGGG
ATA	ATAAA	ATAAT	ATAAC	ATAAC	ATATA	ATATT	ATATC	ATATC	ATACA	ATACT	ATACC	ATACC	ATAGA	ATACT	ATAGC	ATAGG
ATT	ATTAA	ATTAT	ATTAC	ATTAC	ATTTA	ATTTT	ATTTC	ATTTC	ATTCA	ATTCT	ATTCC	ATTCC	ATTCA	ATTGT	ATTGC	ATTCC
ATC	ATCAA	ATCAT	ATCAC	ATCAC	ATCTA	ATCTT	ATCTC	ATCTC	ATCCA	ATCCT	ATCCC	ATCCC	ATCGA	ATCGT	ATCGC	ATOGG
ATG	ATCAA	ATCAT	ATGAC	ATCAC	ATCTA	ATCIT	ATCTC	ATCTC	ATGCA	ATCCT	ATGCC	ATGCG	ATCCA	ATCCT	ATGGC	ATCCC
ACA	ACAAA	ACAAT	ACAAC	ACAAG	ACATA	ACATT	ACATC	ACATG	ACACA	ACACT	ACACC	ACACG	ACAGA	ACAGT	ACAGC	ACAGG
ACT	ACTAA	ACTAT	ACTAC	ACTAG	ACTTA	ACTIT	ACTTC	ACTTG	ACTCA	ACTCT	ACTCC	ACTCG	ACTGA	ACTGT	ACTGC	ACTGG
ACC	ACCAA	ACCAT	ACCAC	ACCAG	ACCTA	ACCTT	ACCTC	ACCTG	ACCCA	ACCCT	ACCCC	ACCCC	ACCGA	ACCCT	ACCGC	ACCCC
ACG	ACCAA	ACCAT	ACGAC	ACCAC	ACCTA	ACCIT	ACCTC	ACCTC	ACCCA	ACCCT	ACGCC		ACCCA	ACCCT	ACGGC	ACCCC
AGA	AGAAA	ACAAT	ACAAC	AGAAG	ACATA	ACATT	AGATO	ACATC	AGACA	AGACT	AGACC	ACACC	AGAGA	AGAGT	AGAGC	AGAGG
AGT	ACTAA	ACTAT	ACTAC	ACTAC	ACTTA	ACTIT	ACTTC	ACTTC	ACTCA	ACTCT	ACTOC	AGTCG	ACTGA	ACTGT	ACTCC	ACTCC
AGC	AGCAA	AGCAT	AGCAC	ACCAC	ACCTA	AGCTT	ACCTC	ACCTC	AGCCA	AGCCT	AGCCC	AGCCG	AGCGA		AGCGC	AGCGG
AGG	AGGAA	ACCAT	AGGAC	ACCAC	ACCTA	ACCIT	AGGTC	ACCTC	AGGCA	AGGCT	AGGCC	AGGCG	AGGGA	ACCCT	AGGGC	AGGGG
TAA	TAAAA		TAAAC	TAAAC	TAATA	TAATT	TAATC	TAATC	TAACA	TAACT	TARCC	TAACG	TAAGA	TAACT	TAAGC	TAAGG
TAT	TATAA	TATAT	TATAC	TATAC	TATTA	TATTT	TATTC	TATEC	TATCA	TATCT	TATCC	TATES	TATCA	TATCT	TATCC	TATCC
TAC	TACAA	TACAT	TACAC	TACAG	TACTA	TACTT	TACTO	TACTG	TACCA	TACCT	TACCC	TACCG	TACGA	TACGT	TACCC	TACGG
TAG	TAGAA	TAGAT	TAGAC	TAGAG	TACTA	TACTT	TAGTC	TACTC	TAGCA	TAGCT	TAGCC	TACCC	TACCA	TAGGT	TAGGC	TACCC
TTA	TTAAA	TTAAT	TTAAC	TTAAC	TTATA	TTATT	TTATC	TTATC	TTACA	TTACT	TTACC	TTACC	TTAGA	TTACT	TTAGE	TTAGG
TIT	TTTAA	TTTAT	TITAC	TTTAG	TTTTA	TITIT	TETTC	TTTTC	TTTCA	TITCE	TTTCC	TITTES	TITCA	TTTCT	TTTCC	TITCC
TIC	TTCAA	TTCAT	TTCAC	TTCAG	TTCTA	TTCTT	TTCTC	TTCTC	TTCCA	TTCCT	TTCCC	TTOOG	TTCCA	TTCGT	TTCCC	TTCCC
TIG	TTGAA	TTGAT	TTCAC	TTCAC	TTCTA	TIGIT	TREEC	TECEC	TTOCA	TTCCT	TEGEC	TTOCC	TTCCA	TTGGT	TTGGC	TTCCC
TCA	TCAAA	TCAAT	TCAAC	TCAAG	TCATA	TCATT	TCATC	TCATC	TCACA	TCACT	TCACC	TCACG	TCAGA	TCACT	TCAGC	TCAGG
TCT	TCTAA	TCTAT	TCTAC	TCTAG	TCTTA	TCTTT	TCTTC	TCTTC	TCTCA	TCTCT	TCTCC	TCTCG	TCTGA	TCTGT	TCTCC	TCTCC
TCC	TOCAA	TCCAT	TCCAC	TOTAL	TOOTA	TCCTT	TOCTO	TOOTS	TCCCA	TCCCT	TOCOC	TODOG	TCCCA	TOOGT	TOOCC	TCCCC
TCG	TOGAA	TOGAT	TCGAC	TOTAL	TOSTA	TEGIT	TOSTC	TOSTS	TOGGA	TCCCT	TOGOC		TCCCA	TOSST	TOGGC	TCCCC
TGA	TCAAA	TCRAT	TCAAC	TGAAG	TCATA	TCATT	TEATC	TCATC	TCACA	TCACT	TCACC	TCACC	TCACA	TGAGT	TCACC	TCACC
TGT	TCTAA	TCTAT	TGTAC	TGTAG	TCTTA	TCTTT	TETTE	TCTTC	TCTCA	TOTOT	TCTCC	Teres	TCTCA	TCTCT	TCTCC	TCTCC
TGC	TOCAA	TECAT	TGCAC	TOTAL	TOCTA	TGCTT	TECTE	TECTE	TGCCA	TGCCT	TOCCC	TGCCG	TGCGA	TOCCT	TOCOC	TGCGG
TGG	TOGAA	TECAT	TOGAC	TOTAL	TOCTA	TOSTT	TOCTO	TOCTO	TOOCA	TOCCT	TOGEC	TODES	TGGGA	TOSST	TOGGC	TGGGG
CAA	CAAAA	CAAAT	CAAAC	CARAG	CAATA	CAATT	CAATC	CAATC	CAACA	CAACT	CARCC	CAACG	CAAGA	CAACT	CAAGC	CAAGG
CAT	CATAA	CATAT	CATAC	CATAG	CATTA	CATTT	CATTC	CATTG	CATCA	CATCT	CATCC	CATCG	CATGA	CATGT	CATCC	CATGG
CAC	CACAA	CACAT	CACAC	CACAG	CACTA	CACTT	CACTC	CACTG	CACCA	CACCE	CACCC	CACCG	CACGA	CACGT	CACCC	CACGG
CAG	CAGAA	CACAT	CAGAC	CAGAG	CACTA	CACTT	CAGTC	CACTG	CACCA	CACCE	CAGCC	CAGCG	CACCA	CACCT	CAGGC	CAGGG
CTA	CTAAA	CTAAT	CTAAC	CTAAC	CTATA	CTATE	CTATC	CTATC	CTACA	CTACT	CTACC	CTACG	CTAGA	CTACT	CTACC	CTAGG
CIT	CTTAA	CTTAT	CTTAC	CTTAG	CTITA	CITIT	CETTC	CTITC	CTTCA	CITCE	CTTCC	CTTCG	CTTGA	CTTCT	CTTCC	CTTCC
CIC	CTCAA	CTCAT	CTCAC	CTCAG	CTCTA	CTCTT	CTCTC	CTCTC	CTCCA	CTCCT	CTCCC	CTOOL	CTCGA	CTCGT	CTCCC	CTCCC
CIG	CTGAA	CTGAT	CTGAC	CTCAC	CTCTA	CTCTT	CTCTC	CTCTC	CTGCA	CTCCT	CTGCC	CTGCG	CTCCA	CTCCT	CTGGC	CTCCC
CCA	CCAAA	CCAAT	CCAAC	CCAAG	CCATA	CCATT	CCATC	CCATC	CCACA	CCACT	CCACC	CCACG	CCAGA	CCAGT	CCAGC	CCAGG
CCT	CCTAA	CCTAT	CCTAC	CCTAG	CCTTA	CCTTT	CCTTC	CCTTC	CCTCA	CCTCT	CCTCC	CCTCC	CCTGA	CCTGT	CCTGC	CCTCC
CCC	CCCAA	CCCAT	CCCAC	CCCAG	CCCTA	CCCTT	COCTC	CCCTC	CCCCA	CCCCT	ccccc	CCCCG	CCCCA	CCCCT	ccccc	CCCGG
CCG	CCGAA	CCCAT	CCGAC	CCCAG	CCCTA	CCCTT	COCTC	CCCEC	CCCCA	CCCCT	CCCCC		CCCCA	CCCCT	CCGGC	CCCCC
CGA	CGAAA	CCRAT	CCAAC	CCAAC	CCATA	CCATT		CCARC	CCACA	CCACT	CCRCC		CCACA	CCACT	CGAGC	CCACC
CGT	CCTAA	CCTAT	CCTAC	CCTAG	CCTTA	CCTTT	CCTTC	CCITC	CCTCA	CCTCT	CCTCC		CCTCA	CCTCT	CCTCC	CCTCC
CGC	CCCAA	CCCAT	CCCAC	CCCAG	CCCTA	CCCTT	COCTC	CCCTC	CCCCA	CCCCT	CCCCC	000000				CGCGG
CGG	CCCAA	CCCAT	CCCAC	COCAG	COCTA	CECTT	COCTC	COCTO	CCCCA	CCCCT	CCCCC	00000				CGGGG
GAA	CAAAA	CAAAT	GAAAC	GAAAG	CAATA	CAATT	GRATC	GAATG	GAACA	CAACT	CARCC	GAACG			CAAGC	
GAT	CATAA	CATAT	GATAC	GATAG	CATTA	CATTT	CATTC	CATTC	GATCA	CATCE	CATCC		CATCA	CATCT	CATCC	CATCC
GAC	CACAA	CACAT	GACAC	CACAG	CACTA	CACTT	CACTC	CACTG	GACCA	CACCE	CACCC	GROOG	CACCA	CACCT	CACCC	CACCC
GAG	CACAA	CACAT	GAGAC	CACAC	CACTA	CACTT	GRETC	CACTO	CACCA	CACCE	CAGCC	CACCC	CACCA	CACCT	CACCC	CACCC
GTA	CTAAA	GTAAT	CTARC	GTAAG	CTATA	CTATT	CTATC	CTATC	GTACA	CTACT	CTACC	GTACG	CTACA	CTACT	CTACC	GTAGG
CIT	CTTAA	CTTAT	CITAC	CITAG	CTTTA	CITIT	CHTTC	CTITC	GTTCA	CITCE	CTTCC	GTTCG	CITCA	CITCI	CTTCC	CITCC
GTC	CTCAA	CTCAT	CTCAC	GTCAG	CTCTA	CICIT	GTCTC	CTCTC	GTOCA	CTCCT	CTCCC	GTOOG	CICCA	CTCCT	CTCCC	GTCGG
GTG	CTCAA	CTCAT	GTGAC	GTGAG	GTGTA	CICIT	GTGTC	GTGTG	GTGCA	CTCCT	CECCC	CTCCC	CTCCA	CTCCT	CECCC	CTCCC
GCA	GCAAA	CCAAT	GCAAC	GCAAG	GCATA	CCATT	GCATC	GCATG	GCACA	CCACT	GCACC	GCACG	CCACA	CCACT	GCAGC	CCACC
GCT	GCTAA	CCTAT	GCTAC	GCTAG	GCTTA	CCTTT	GCTTC	GCTTG	GCTCA	CCTCT	CCTCC	GCTCG	CCTCA	CCTCT	GCTGC	CCTCC
GCC	GCCAA	GCCAT	GCCAC	GCCAG	CCCTA	CCCTT	GCCTC	GCCTG	GCCCA	CCCCT	CCCCC	G000G	CCCCA	CCCCT	GCCGC	CCCCC
GCG	GOGAA	CCCAT	CCCAC	CCCAG	CCCTA	CCCTT	GOCTO		GEGEA	CCCCT	GCGCC		CCCCA	CCCCT	GCGGC	CCCCC
GGA	GCAAA	CCRAT	CCAAC	GCAAG	CCATA	CCATT	GCATC	CCATC	CCACA	CCACT	CCACC	CCACC	CCACA	CCACT	GCAGC	CCACC
CCT	GCTAA	CCTAT	GGTAC	GGTAG	GGITA	CCTTT	GGTTC	GGITG	GCTCA	CCTCT	CCTCC	CCTCC	CCTCA	CCTCT	CCTCC	CCTCC
GGC	GCCAA	CCCAT	GGCAC	GGCAG	GCCTA	CCCTT	GGCTC	GCCTG	GGCCA	CCCCT	CCCCC	90000	CCCCA	CCCCT	GCCCC	GGCGG
GGG	CCCAA	CCCAT	CCCAC	GCCAG	CCCTA	CCCIT	COCRC	CCCTC	CCCCA	CCCCT	GDGCC	22222	CCCCA	CCCCT	CCCCC	CCCCC

Same as previous analysis

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	GI	GC	GG
AAA	AAAAA	ARART	AAAAC	AAAAG	AAATA	ARATT	AAATC	AAATG	AAACA	ARACT	AAACC	AAACG	ARAGA	AAAGT	AAAGC	AAAGG
AAT	AATAA	AATAT	AATAC	AATAG	AATTA	AATTT	AATTC	AATTG	AATCA	AATCT	AATCC	AATCC	AATGA	AATGT	AATCC	AATCC
AAC	AACAA	AACAT	AACAC	AACAG	AACTA	AACTT	AACTC	AACTG	AACCA	AACCT	AACCC	AACCC	AACGA	AACGT	AACGC	AACCC
AAG	22/222	SECET	aacac	bacac	33/273	AAGTT	AACTYC.	33/07/0	330003	22000	22000	33000	33003	33007	23/000	22000
ATA	ATABA	ATABT	ATABC	DEATA	ATATA	ATATT	ATATC	ATATC	araca	ATACT	ATACC	ATACC	272/22	ATACT	ATACC	ATAGG
				ATTAC	AIRIA	MINII		ALALG	ALALA	MINLI			ALMIAN	HIRIT		ALMala
ATT	ATTAA	ATTAT	ATTAC	ALL LAUS	ATTTA	ATTTT	ATTTC	ATTTG	ATTUA	AFFCT	ATTCC	ATTCC	ATTUA	ATTET	ATTGC	ATTGG
ATC	ATCAA	ATCAT	ATCAC	ATCAG	ATCTA	ATCTT	ATCTC	ATCTG	ATOCA	ATCCT	ATCCC	ATCCC	ATOGA	ATCGT	ATCCC	ATCCC
ATG	ATCAA	ATCAT	ATGAC	ATCAC	ATCTA	ATCTT	ATCTC	ATCTC	ATGCA	ATCCT	ATGCC	ATGCG	ATGGA	ATGGT	ATGGC	ATCCC
ACA.	ACAAA	ACART	acaac	ACAAG	ACATA	ACATT	ACATC	ACATG	acaca	ACACT	ACACC	acacc	acaca	ACACT	ACAGE	ACAGG
ACT	ACTAA	ACTAT	acrac	actac	actea	3CTTT	ACTTC	ACTUC	ACTCA	3/CT/CT	ACTOC	actics	3/07/03	actor	ACTOC	acrec
ACC	20022	ACCAT	30020	20026	30003	3/0075	ACCTC	ACCTIC	3/0003	accer.	200006	3,0000	3,0003	30000	30000	3,0000
	BOCAR	ACCAT	ACCAC.	accas accas	ALCTA	ALLII			ACCCCA.	ALLEL	ACCCC.	ALLLE	3,0003	ALLUGI	NOCCC.	ALLIGIA
ACG	ACCORDING N		ACGAC	200000000000000000000000000000000000000	ACCTA	ACCIT	ACCTC	ACCTC	ACCCA	ACCET	*******		ACCCA	ACCCT	***************************************	ACCOR
AGA	AGAAA	ACAAT	AGAAC	AGAAG	ACATA	ACATT		AGATG	ACACA	ACACT	AGACC	AGACG	ACACA	ACACT	AGAGE	ACACC
AGT	ACTAA	ACTAT	ACTAC	ACTAG	ACTTA	ACTIT	ACTTC	ACTTC	ACTCA	ACTCT	ACTCC	ACTCC	ACTCA	ACTOT	ACTCC	ACTCC
AGC	ACCAA	AGCAT	AGCAC	AGCAG	ACCTA	ACCTT	ACCTC	ACCTC	AGCCA	AGCCT	AGCCC	AGCCC	AGOGA	ACCCT	ACCCC	AGCGG
AGG	accaa	a/ccar	accar	accac	accera	MOCTE	ACCTC	accerc	acces	accor.	30000	acces	3/0003	accer.	accer	20000
TAA	Tabab	TABLE	TRABE	Tabac	TARTA	TANTE	TARTC	TARTC	Thaca	TABCT	TARCC	TARCC	F33.03	TARCT	Tabor	F3.3.00
			TATAC		IARIA	IAAII		IMMIG	IMMLA	IAALI		TATICC	I HARLAN.	IMMI	TATION	I JUNEAU
TAT	TATAA	TATAT		TATAG	TATTA	TATTY	TATTC	TATES	TATUA	TATCT	TATCC	20000	TATUA	TATET	TATISC	TATGE
TAC	TACAA	TACAT	TACAC	TACAG	TACTA	TACTT	TACTC	TACTC	TACCA	TACCT	TACCC	TACCC	TACCA	TACCT	TACCC	TACCC
TAG	TAGAA	TACAT	TAGAC	TAGAG	TAGTA	TACTT	TACTO	TACTC	TAGCA	TACCT	TAGCC	TACCC	TAGGA	TAGGT	TAGGE	TACCC
TTA	TTAAA	TTAAT	TTAAC	TTAAC	TTATA	TTATE	TTATC	TTATC	TTACA	TTACT	TTACC	TTACC	TTAGA	TTAGT	TTACC	TTAGG
TIT	TTTAN	TETTAT	TITAC	TTTAC	TTTTA	TITLE	TTTTTC	TTTTC	TTTTCA	TTTCT	TETOC	TTTTCC	TTTTCA	TTTCT	TTTTCC	TTTTCC
TTC	TTCAA	TICAL	mmana	mmon.o	mmoms.		mnono	meana	mmoon	BECOME	menon	777000	mmoon.	magan	TECOTO .	
			TILAL	TICAL	TICIA	111011	THERE	11010	TILLE	111001	TRUCK	11006	TILLER	THUGH	TICUL	TILLE
TIG	TTCAR	TTCAT	TTCAC	TTCAC	TTGTA	TIGHT	TREEC	TECHC	TIBLE	TICCI	TEGEC	TITECC	TICCA	TIGGT	TTGGC	TTOCC
TCA	TCAAA	TCAAT	TCAAC	TCAAG	TCATA	TCATT	TCATC	TCATG	TCACA	TCACT	TCACC	TCACC	TCACA	TCACT	TCACC	TCAGG
TCT	TCTAA	TCTAT	TCTAC	TCTAG	TCTTA	TCTTT	TCTTC	TCTTC	TCTCA	TCTCT	TCTCC	TCTCC	TCTCA	TCTCT	TCTCC	TCTCC
TCC	TOCAA	TOCAT	TCCAC	TOCAG	TCCTA	TCCTT	TOCTO						23	TOOGT	TOOGC	TCCCC
TCG	TOGAA	TOCAT	TOTAL	TOGAG	TOTA	DOCTO.	TOGEC	Ħ TC∧	TC: 42	204			22	TOCOT	TOOR	TODGC
TGA	TCAAA	TCAAT	TCAAC	TCAAC	TONES.	DOLDER.	TCATC	H ICA	NTC: 12	.39 I				TOLOT	TCACC	TCAGG
	2.000.000	ILEMALL	IGARAL	Itanina	ILANIA	IGNII		Н					-	I latter I	Itaklat	Lande
TGT	TCTAR	TOTAL	TCTAC	TGTAG	TETTA	TOTT	Terre	ш aq∠	1926 t	imes a	as mea	an valı	10 🏝	TOTOT	TETEC	TGTGG
TGC	TOCAR	TGCAT	TGCAC	TOCAG	TOCTA	TGCTT	TOCTO	J.J-	1220 1	IIIIC3 C	13 IIIC	in van	uc 🚉	TOCCT	TOCOC	TCCCC
TGG	TOCAR	TECAT	TGGAC	TOCAC	TOSTA	TGGTT	TOCTO							TOCCT	TOCCC	TCCCC
CAA	CAAAA	CARAT	CAAAC	CARAG	CAATA	CAATT	CAATC	CAATG	CAACA	CAACT	CARCC	CAACG	CAACA	CAAGT	CAAGC	CAAGG
CAA			CAAAC	CARAG	CARTA	CANTT	CARTC	CARTG	CARCA	CAACT			CARGA	CAAGT	CARGO	CARGG
CAT	CATAA	CATAT	CATAC	CATAG	CARTA	CATTT	CATTC	CARTG	CATCA	CATCT	CATCC	CATCG	CATCA		CARGO	CAAGG
CAT	CATAA	CATAT	CATAC		CAATA CATTA CACTA	CATTT	CATTC	CAATG CATTG CACTG	CATCA CACCA	CATCT	CACCC	CATCG			CATGC	CATGG
CAC	CATAA CACAA CACAA	CATAT CACAT CAGAT	CATAC CACAC CAGAC	CATAG	CAATA CAITA CACTA CAGTA	CATTT CACTT CACTT	CATTC	CAATG CATTG CACTG CAGTG	CATCA	CATCT	CATCC CACCC CAGCC	CATCG	CATCA		CAAGC CATGC CACGC CAGGC	CATGG CACGG CAGGG
CAT CAC CAG CTA	CATAR CACAR CAGAR CTAAR	CATAT	CATAC	CATAG	CAATA CATTA CACTA CAGTA CTATA	CATTT	CATTC	CAATG CATTG CACTG CAGTG CTATG	CATCA CACCA	CATCT	CACCC	CATCG	CATCA		CATGC	CATGG
CAT CAC CAG CTA CTT	CATAA CACAA CACAA	CATAT CACAT CAGAT	CATAC CACAC CAGAC	CATAG	CAATA CACTA CACTA CAGTA CTATA CTITA	CATTT CACTT CACTT	CATTC	CAATG CATTG CACTG CAGTG CTATG CTITG	CATCA CACCA	CATCT	CATCC CACCC CAGCC	CATCG	CATCA		CATGC	CATGG CACGG CAGGG
CAT CAC CAG CTA	CATAR CACAR CAGAR CTAAR	CATAT CACAT CACAT CAGAT CTAAT	CATAC CACAC CAGAC CTAAC	CATAG	CAATA CATTA CACTA CAGTA CTATA CTATA CTCTA	CATTT CACTT CAGTT CTATT	CATTC	CAATG CATTG CACTG CAGTG CTATG CTTTG CTCTG	CATCA CACCA	CATCT	CATCC CACCC CAGCC	CATCG	CATCA		CATGC	CATGG CACGG CAGGG CTAGG
CAT CAC CAG CTA CTT CTC	CATAR CACAR CAGAR CTAAR CTTAA	CATAT CACAT CAGAT CTAAT CTTAT	CATAC CACAC CAGAC CTAAC CTAAC	CATAG	CAATA CATTA CACTA CAGTA CTATA CTATA CTCTA CTCTA	CATTT CACTT CAGTT CTATT	CATTC	CAATG CATTG CACTG CAGTG CTATG CTTTG CTCTG	CATCA CACCA	CATCT	CATCC CACCC CAGCC	CATCG	CATCA		CATGC	CATGG CACGG CAGGG CTAGG
CAT CAC CAG CTA CTT CTC CTG	CATAA CACAA CAGAA CTAAA CTTAA CTCAA CTCAA	CATAT CACAT CAGAT CTAAT CTTAT CTCAT CTGAT	CATAC CACAC CAGAC CTAAC CTTAC CTCAC CTGAC	CATAG	CAATA CATTA CACTA CACTA CAGTA CTATA CTATA CTGTA CTGTA CTGTA	CATTT CACTT CAGTT CAGTT CTATT CTATT CTTTT	CATTC CACTC CAGTC CTATC CTTTC CTCTC	CAATG CATTG CACTG CACTG CAGTG CTATG CTATG CTCTG CTCTG	CATCA CACCA	CATCT	CATCC CACCC CAGCC CTACC CTTCC CTCCC CTCCC	CATCG CACCG CAGCG CTACG CTTCG CTCCG	CATCA		CATGC	CATGG CACGG CAGGG CTAGG
CAT CAC CAG CTA CTT CTC CTG CTG	CATAA CACAA CAGAA CTAAA CTTAA CTCAA CTGAA CTGAA	CATAT CACAT CAGAT CTAAT CTTAT CTCAT CTGAT CTGAT	CATAC CACAC CAGAC CTAAC CTTAC CTCAC CTGAC CTGAC	CATAG CACAG CAGAG CTAAG CTTAG CTCAG CTGAG CCAAG	CAATA CACTA CACTA CACTA CAGTA CTATA CTATA CTCTA CTCTA CTCTA CCATA	CATTT CACTT CAGTT CAGTT CTATT CTATT CTTTT	CATTC CACTC CAGTC CTATC CTTTC CTCTC CTCTC CTGTC	CAATG CACTG CACTG CACTG CACTG CTATG CTATG CTATG CTCTG CTCTG CCATG	CATCA CACCA	CATCT	CATOC CACOC CAGOC CTACC CTTOC CTCCC CTCCC CTGCC	CATCG CACCG CACCG CTACG CTTCG CTTCG CTCCG CTCCG CTCCG	CATCA		CATGC	CATGG CACGG CAGGG CTAGG
CAT CAG CTA CTT CTC CTG CCA CCT	CATAA CACAA CACAA CAGAA CTTAA CTTAA CTGAA CTGAA CCTAA	CATAT CACAT	CATAC CACAC CAGAC CTAAC CTTAC CTCAC CTGAC CCGAC CCAAC	CATAG CACAG CAGAG CTAAG CTTAG CTCAG CTGAG	CAATA CACTA CACTA CACTA CACTA CTATA CTATA CTCTA CTCTA CTCTA CCTTA CCTTA CCTTA	CATTT CACTT CAGTT CAGTT CTATT CTTTT CTCTT CTGTT CCATT CCATT	CATTC CACTC CACTC CAGTC CTATC CTTTC CTCTC CTCTC CCATC CCATC	CAATG CATTG CACTG CACTG CACTG CTATG CTATG CTCTG CTCTG CTCTG CCATG CCTTG	CATCA CACCA	CATCT CACCT CACCT CACCT CTACT CTTCT CTCCT CTCCT CTCCT CCTCCT CCTCCT	CATOC CACOC CAGOC CTACC CTTOC CTCCC CTCCC CTGCC CCACC CCACC	CATCG CACCG CACCG CTACG CTTCG CTCCG CTCCG CTCCG CTCCG CTCCG	CATCA		CATGC	CATGG CACGG CAGGG CTAGG
CAT CAC CAG CTA CTT CTC CTG CCA CCT CCC	CATAR CACAR CACAR CACAR CTAR CTTAR CTCAR CTCAR CTCAR CCTAR CCTAR CCTAR CCTAR CCTAR	CATAT CACAT CACAT CACAT CACAT CTAAT CTTAT CTCAT CTCAT CCTAT CCTAT CCTAT CCCAT	CATAC CACAC CACAC CTAAC CTAAC CTCAC CTCAC CCCAC CCCAC	CATAG CACAG CAGAG CTAAG CTTAG CTCAG CTGAG CCAAG	CAATA CACTA	CATTT CACTT CAGTT CAGTT CTATT CTATT CTTTT	CATTC CACTC CACTC CTATC CTATC CTTC CTCTC CTCTC CCATC CCATC CCTTC	CAATG CACTG CACTG CACTG CACTG CACTG CTATG CTCTG CTCTG CTCTG CCATG CCATG CCCTG	CATCA CACCA	CATCT	CATCC CACCC CACCC CTACC CTCCC CTCCC CTCCC CTCCC CCCCC CCCCC CCCCC CCCCC	CATCG CACCG CACCG CTACG CTTCG CTCCG CTCCG CTCCG CTGCG	CATCA		CATGC	CARGG CARGG CARGG CARGG CTAGG CTTAGG
CAT CAC CAG CTA CTT CTC CTG CCA CCT CCC CCC	CATAR CACAR CAGAR CTRAR CTTAR CTTAR CTCAR CTCAR CTCAR CCTAR CCTAR CCTAR CCTAR CCTAR CCTAR CCTAR CCTAR CCCAR CCCAR	CATAT CACAT	CATAC CACAC CACAC CTARC CTTAC CTCAC CTCAC CCARC CCARC CCCAC CCCAC	CATAG CACAG CAGAG CTAAG CTTAG CTCAG CTGAG CCAAG	CAATA CACTA CACTA CACTA CACTA CTATA CTATA CTCTA CTCTA CCTTA CCTTA CCTTA CCTTA CCTTA CCCTA CCCTA	CATTT CACTT CAGTT CIATT CITTT CICTT CICTT COATT COATT COCTT COCTT	CATTC CACTC CACTC CAGTC CTATC CTTTC CTCTC CTCTC CCATC CCATC	CATTG CACTG CAGTG CTATG CTATG CTCTG CTCTG CTCTG CCATG CCATG CCTTG CCCTG CCCTG	CATCA CACCA	CATCT CACCT CACCT CACCT CTACT CTTCT CTCCT CTCCT CTCCT CCTCCT CCTCCT	CATCC CACCC CACCC CTACC CTTCC CTCCC CTCCC CTCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC	CATCG CACCG CACCG CTACG CTTCG CTCCG CTCCG CTCCG CTCCG CTCCG	CATCA		CATGC	CATCG CATCG CATCG CATCGC CTTGGC CTTGGG
CAT CAC CAG CTA CTT CTC CTG CCA CCT CCA CCT CCC CCG CCG CCG	CATAR CACAR CACAR CACAR CTRAR CTTAR CTCAR CTCAR CCCAR	CATAT CACAT CACAT CACAT CTAAT CTCAT CTCAT CTCAT CTCAT CCCAT CCCAT CCCAT CCCAT CCCAT	CATAC CACAC CACAC CTAAC CTAAC CTCAC CTCAC CCCAC CCCAC	CATAG CACAG CAGAG CTAAG CTTAG CTCAG CTGAG CCAAG	CAATA CACTA CACTA CACTA CACTA CTCTA CTCTA CTCTA CCATA	CATTT CACTT CAGTT CAGTT CTATT CTTTT CTCTT CTGTT CCATT CCATT	CATTC CACTC CACTC CTATC CTATC CTTC CTCTC CTCTC CCATC CCATC CCTTC	CAATG CACTG CACTG CACTG CACTG CTCTG CTCTG CTCTG CCATG CCATG CCATG CCATG CCATG CCATG CCATG CCATG CCATG	CATCA CACCA	CATCT CACCT CACCT CACCT CTACT CTTCT CTCCT CTCCT CTCCT CCTCCT CCTCCT	CATCC CACCC CACCC CTACC CTCCC CTCCC CTCCC CTCCC CCCCC CCCCC CCCCC CCCCC	CATCG CACCG CACCG CTACG CTTCG CTCCG CTCCG CTCCG CTCCG CTCCG	CATCA		CATGC	CARGG CARGG CARGG CARGG CTAGG CTTAGG
CAT CAC CAG CTA CTT CTC CTG CCA CCT CCC CCC	CATAR CACAR CAGAR CTRAR CTTAR CTTAR CTCAR CTCAR CTCAR CCTAR CCTAR CCTAR CCTAR CCTAR CCTAR CCTAR CCTAR CCCAR CCCAR	CATAT CACAT CACAT CACAT CACAT CTAAT CTTAT CTCAT CTCAT CCTAT CCTAT CCTAT CCCAT	CATAC CACAC CACAC CTARC CTTAC CTCAC CTCAC CCARC CCARC CCCAC CCCAC	CATAG CACAG CAGAG CTAAG CTTAG CTCAG CTGAG CCAAG	CAATA CACTA CACTA CACTA CACTA CTATA CTOTA CTOTA COCTA CCATA CCOTA CCOTA CCOTA CCGTA CCGTA CGGTA CGGTA	CATTT CACTT CAGTT CIATT CITTT CICTT CICTT COATT COATT COCTT COCTT	CATTC CACTC CACTC CTATC CTATC CTTC CTCTC CTCTC CCATC CCATC CCTTC	CATTG CACTG CAGTG CTATG CTATG CTCTG CTCTG CTCTG CCATG CCATG CCTTG CCCTG CCCTG	CATCA CACCA	CATCT CACCT CACCT CACCT CTACT CTTCT CTCCT CTCCT CTCCT CCTCCT CCTCCT	CATCC CACCC CACCC CTACC CTTCC CTCCC CTCCC CTCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC	CATCG CACCG CACCG CTACG CTTCG CTCCG CTCCG CTCCG CTCCG CTCCG	CATCA		CATGC	CATCG CATCG CATCG CATCGC CTTGGC CTTGGG
CAT CAC CAG CTA CTT CTC CTG CCA CCT CCA CCT CCC CCG CCG CCG	CATAR CACAR CACAR CACAR CTRAR CTTAR CTCAR CTCAR CCCAR	CATAT CACAT CACAT CACAT CTAAT CTCAT CTCAT CTCAT CTCAT CCCAT CCCAT CCCAT CCCAT CCCAT	CATAC CACAC CTARC CTARC CTTAC CTCAC CTCAC CCARC CCARC CCCAC CCCAC CCCAC	CATAG CACAG CAGAG CTAAG CTTAG CTCAG CTGAG CCAAG	CAATA CACTA CACTA CACTA CACTA CTOTA CTOTA CTOTA CCOTA	CATTT CACTT CACTT CACTT CIATT CITTT CICTT CICTT COCTT CCCTT CCCTT CCCTT CCCTT CCCTT CCCTT CCCTT	CATTC CACTC CACTC CTATC CTATC CTCTC CTCTC CCCTC CCCTC CCCTC CCCTC CCCTC	CATTG CACTG CAGTG CTATG CTTTG CTCTG CTCTG CCTGTG CCTTG CCTTG CCTTG CCTTG CCTTG CCTTG CCTTG	CATCA CACCA	CATCT CACCT CACCT CACCT CTACT CTTCT CTCCT CTCCT CTCCT CCTCCT CCTCCT	CATCC CACCC CACCC CTACC CTTCC CTCCC CTCCC CTCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCC	CATCG CACCG CACCG CTACG CTTCG CTCCG CTCCG CTCCG CTCCG CTCCG	CATCA		CATGC	CATCG CATCG CATCG CATCGC CTTGGC CTTGGG
CAT CAC CAG CTA CTT CTC CTG CCA CCT CCC CCG CCG CCG CCG CCG CCG CCG CCG	CATAR CACAR CACAR CACAR CATAR	CATAT CACAT CAGAT CAGAT CTAAT CTCAT CTCAT CTCAT CCTAT CCTAT CCTAT CCCAT CCCAT CCCAT CCCAT CCCAT CCCAT CCCAT CCCAT	CATAC CACAC CAGAC CTAAC CTTAC CTCAC CTCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC	CATAG CACAG CAGAG CTAAG CTTAG CTCAG CTGAG CCAAG	CARTA CACTTA CACTTA CACTTA CTGTA CTGTA CTGTA CCGTA CCGTA CCGTA CCGTA CCGTA CCGTA CCGTA CCGTA CCGTA CGGTA	CATTT CACTT CAGTT CTATT CTATT CTGTT CTGTT CTGTT CCATT CCATT CCATT CCGTT CCGTT CCGTT CCGTT CCGTT CCGTT	CATTC CACTC CACTC CACTC CTATC CTATC CTCTC CTCTC CCATC CCATC CCTTC CCCTC CCCTC CCCTC CCATC CCCTC CCCTC CCCTC CCCTC	CATTS CACTG CAGTG CTATG CTATG CTCTG CTCTG CTCTG CCTTG	CATCA CACCA	CATCT CACCT CACCT CACCT CTACT CTTCT CTCCT CTCCT CTCCT CCTCCT CCTCCT	CATCC CACCC CACCC CACCC CTACC CTTCC CTCCC CTCCC CCTCC CCTCC CCCC CCCCC CCCCC CCCCC CCACC CCACC	CATCG CACCG CACCG CACCG CACCG CTACG CTCCG CTCCG CTCCG CCCCG CCCCCG CCCCCG CCCCCG CCCCCC	CATCA		CATGC	CATCG CATCG CATCG CATCGC CTTGGC CTTGGG
CAT CAC CAG CTA CTT CTC CTG CCA CCT CCC CCG CCG CCG CCG CCG CCG CCG CCG	CATAR CACAR CACAR CACAR CACAR CTARA CTCAR CTCAR CTCAR CCCAR	CATAT CACAT CAGAT CAGAT CATAT	CATAC CACAC CACAC CACAC CATAC	CATAG CACAG CAGAG CAGAG CAGAG CATAG	CARTA CACTA CACTA CACTA CACTA CTATA CTOTA CTOTA CCATA CCATA CCATA CCCTA CCATA CCCTA	CATTT CACTT CAGTT CTATT CTATT CTCTT CTCTT CTCTT CCATT CCATT CCATT CCCTT CCGTT CCGTT CCGTT CCGTT CGGTT CGGTT CGGTT CGGTT	CATTC CACTC CAGTC CTATC CTATC CTATC CTCTC CTCTC CCATC CCATC CCATC CCCTC CCGTC CCGTC CCGTC CCGTC CCGTC CCGTC CCGTC CCGTC	CATTG CACTG CACTG CACTG CATTG CTATG CTTTG CTCTG CCATG	CATCA CACCA	CATCT CACCT CACCT CACCT CTACT CTTCT CTCCT CTCCT CTCCT CCTCCT CCTCCT	CATCC CACCC CACCC CACCC CTTCC CTCCC CTCCC CTCCC CCACC	CATCG CACCG CACCG CACCG CTACG CTACG CTACG CTACG CTACG CTACG CTACG CACCG	CATCA CROCA CROCA CROCA CTOCA CTOCA CTOCA CTOCA CCOCA CCOCA CCOCA CCOCA CCOCA CCOCA CCOCA CCOCA CCOCCA	CAUGI CAUGI CAUGI CAUGI CAUGI CAUGI CTUGI CTUGI CTUGI COUGI	CATIC CATIC CACIC CAGGC CTAGC CTTGC CTGGC CTGGC CTGGC CTGGC CTGGC CCTGGC	CARGO
CAT CAC CAG CTT CTC CTG CCA CCT CCC CCG CCC CCG CCG CCGA CGT CGG CGA CGT CGG CGG CGA	CATAR CACAR CACAR CACAR CATAR CTTAR CTTAR CTCAR CTCAR CCCAR CCAR CCCAR CCAR CCCAR CCAR	CATAT CACAT CACAT CACAT CATAT CTTAT CTCAT CTCAT CCCAT	CATAC CACAC CACAC CATAC	CATAG CACAG CACAG CACAG CATAG	CATA CATA CATA CACTA CACTA CACTA CACTA CTATA CTATA CTATA CTATA CTATA CCATA	CATTT CACTT CACTT CACTT CTATT CTTT CTCTT CTCTT CCCTT C	CATTC CACTC CACTC CACTC CTATC CTTTC CTCTC CCATC CAATC CAATC CAATC CAATC	CATTG CACTG CACTG CACTG CATTG CTATG CTTTG CTCTG CCATG	CATCA CACCA	CATCT CACCT CACCT CACCT CTACT CTTCT CTTCT CTCCT CCCCT CCCT CCCCT CCCT CCCCT CCCT CCCCT CCCT	CATCC CACCC CACCC CACCC CTTCC CTTCC CTCCC CTCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCC	CATCS CACCS CACCS CACCS CTACS CACCS	CATCA CACCA	CALGIT CALGIT CALGIT CALGIT CTLGIT CTLGIT CTLGIT CTLGIT CCLGIT CCLGIT CCLGIT CCLGIT CCLGIT CCLGIT CCLGIT CCLGIT CCLGIT CGLGIT CG	CATACO CATACO CAGACO CATACO CA	CARGO
CAT CAC CAC CAC CTA CTT CTC CTC CCA CCA	CATAR CACRA CACRA CACRA CACRA CACRA CATRA	CATAT CACAT CACAT CACAT CACAT CATAT	CATAC CACAC	CATAG CACAG CAGAG CAGAG CAGAG CAGAG CATAG CATAG CATAG CATAG CATAG CAGAG	CATTA CATTA CACTTA CACTTA CACTTA CACTTA CTATTA CTOTTA CTOTTA CCOTTA COTTA COTTA COTTA	CATTI CACTI CACTI CACTI CACTI CACTI CIATI CIATI CIATI CCATI CATI	CATTC CACTC CACTC CACTC CACTC CTATC CTATC CTATC CTATC CCATC CCATC CCATC CCATC CCATC CCATC CCATC CCATC CATTC	CATTG CACTG CACTG CACTG CATTG CTATG CTTTG CTCTG CCATG	CATCA CACCA CACCA CACCA CTACA CTACA CTACA CTACA CTACA CCACA CACCA CCACA CACCA	CATCT CACCT CACCT CACCT CTACT CTACT CTCCT CTCCT CCCCT	CATOC CACOC CACOC CTACC CTACC CTACC CTACC CTACC CTACC CTACC CTACC CCACC CACC CCACC C	CATCS CACCS CACCS CACCS CTACS CTACS CTCCS CTCCS CTCCS CCCCS CCCCCS CCCCCS CCCCS CCCCS CCCCCS CCCCCS CCCCCS CCCCCS CCCCCS CCCCCS CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCC	CATCA CACCA CACCA CACCA CACCA CATCA	CATGT CATGT CACGT CAGGT CTAGT CTGGT CTGGT CCTGGT CCTGGT CCGGT CCGGT CCGGT CCGGT CCGGT CCGGT CGGGT CGGG	CATGG CAGGG CAGGG CAGGG CTAGG CTTGG CTGGG CTGGG CTGGG CCGGGG CCGGG CCGGG CCGGG CGGGG	CARGO
CAT CAC CAC CAC CAT CTT CTC CTG CCC CCG CCG CCG CCG CCG CGG CG	CATAR CACRA CACRA CACRA CACRA CACRA CATRA	CATH CACH CACH CACH CTAN CTIAN CTIAN CTIAN CCAN CCAN CCAN CCAN CCAN CCAN CCAN C	CATAC CACAC CACAC CTAAC CTTAC CTGAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCC CCCC CCCAC CCCC CCC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC C	CATAG CACAG CACAG CACAG CATAG	CATTA CATTA CACTTA CACTTA CACTTA CTATTA CTOTTA CCTATA CACTTA CACTTA	CATTI CACTI CACTI CACTI CACTI CIGIT CIGIT CIGIT COATI COTTI COTTI COTTI COTTI COGTI	CATTC CACTC CACTC CACTC CACTC CTATC CTOTC CTOTC CCATC CATC CATC CATC CACC CATC	CATTG CACTG CACTG CACTG CATTG CTATG CTTTG CTCTG CCATG	CATCA CACCA	CATCT CACCT	CATCC CACCC CACCC CACCC CTACC CTOCC CTOCC CTOCC CCACC CACCC CAACC CAACC CAACC CAACC	CATCS CACCS	CATCA CACCA	CARTET CACET CACET CACET CACET CTACT CTIGT CTOST CCOST CSACT CSACT CSACT CSACT CACCT	CATTOC CACTOC CACTOC CTATAC CTTOC CTCOC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCAC CCACC CCACC CCAC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCAC	CARGG
CAT CAC CAC CAC CTA CTT CTC CTC CCA CCA	CATAR CACRA CACRA CACRA CACRA CACRA CATRA	CATAT CACAT CACAT CACAT CACAT CATAT	CATAC CACAC	CATAG CACAG CAGAG CAGAG CAGAG CAGAG CATAG CATAG CATAG CATAG CATAG CAGAG	CARTA CAITA CACTA CACTA CACTA CACTA CTOTTA CTOTTA CCATA CACTA CACTA CACTA CACTA CACTA	CATTI CACTI CACTI CACTI CACTI CACTI CIATI CIATI CIATI CCATI CATI	CATTC CACTC CACTC CACTC CACTC CTATC CTATC CTATC CTATC CCATC CCATC CCATC CCATC CCATC CCATC CCATC CCATC CATTC	CATTG CACTG CACTG CACTG CATTG CTATG CTTTG CTCTG CCATG	CATCA CACCA CACCA CACCA CTACA CTACA CTACA CTACA CTACA CCACA CACCA CCACA CACCA	CATCT CACCT CACCT CACCT CTACT CTACT CTCCT CTCCT CCCCT	CATOC CACOC CACOC CTACC CTACC CTACC CTACC CTACC CTACC CTACC CTACC CCACC CACC CCACC C	CATCS CACCS CACCS CACCS CTACS CTACS CTCCS CTCCS CTCCS CCCCS CCCCCS CCCCCS CCCCS CCCCS CCCCCS CCCCCS CCCCCS CCCCCS CCCCCS CCCCCS CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCC	CATCA CACCA CACCA CACCA CACCA CATCA	CATGT CATGT CACGT CAGGT CTAGT CTGGT CTGGT CCTGGT CCTGGT CCGGT CCGGT CCGGT CCGGT CCGGT CCGGT CGGGT CGGG	CATGG CAGGG CAGGG CAGGG CTAGG CTTGG CTGGG CTGGG CTGGG CCGGGG CCGGG CCGGG CCGGG CGGGG	CARGO
CAT CAC CAC CAC CAT CTT CTC CTG CCC CCG CCG CCG CCG CCG CGG CG	CATAR CACRA CACRA CACRA CACRA CACRA CATRA	CATH CACH CACH CACH CTAN CTIAN CTIAN CTIAN CCAN CCAN CCAN CCAN CCAN CCAN CCAN C	CATAC CACAC CACAC CTAAC CTTAC CTGAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCC CCCC CCCAC CCCC CCC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC C	CATAG CACAG CACAG CACAG CTTAG CTCAG CTCAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG	CARTA CARTA CACTA CACTA CACTA CTOTTA CTOTTA CTOTTA CCOTTA	CATTI CACTI CACTI CACTI CACTI CIGIT CIGIT CIGIT COATI COTTI COTTI COTTI COTTI COGTI	CATTC CACTC CACTC CACTC CACTC CTATC CTOTC CTOTC CCATC CATC CATC CATC CACC CATC	CATTG CACTG CACTG CACTG CATTG CTATG CTTTG CTCTG CCATG	CATCA CACCA CACCA CTACA CTOCA CTOCA CTOCA COCCA CCACCA CCACCA CCACCA CCACCA CCACCA CCACCA	CATCT CACCT	CATCC CACCC CACCC CACCC CTACC CTOCC CTOCC CTOCC CCACC CACCC CAACC CAACC CAACC CAACC	CATCS CACCS	CATCA CACCA	CARTET CACET CACET CACET CACET CTACT CTIGT CTOST CCOST CSACT CSACT CSACT CSACT CACCT	CATTOC CACTOC CACTOC CTATAC CTTOC CTCOC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCAC CCACC CCACC CCAC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCAC	CARGG
CAT CAC CAC CAC CTA CTT CTC CTC CCC CCC	CATRA CACRA CACRA CACRA CACRA CACRA CACRA CATRA CACRA CATRA CACRA CATRA CACRA CATRA CACRA CACRA CATRA CACRA CATRA CACRA CATRA	CATH CACHT CACHT CACHT CTANT CTOM CTOM CCAMT CCAMT CCAMT CCAMT CCAMT CCAMT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CACHT CAC	CATAC CACAC CACAC CTAAC CTTAC CTGAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCC CCCC CCCAC CCCC CCC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC CCCAC C	CATAG CACAG CACAG CACAG CTTAG CTCAG CTCAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG	CATTA CATTA CACTTA CACTTA CACTTA CTOTTA CTOTTA CCATTA CACTTA CACT	CATTI CACTI	CATTC CACTC CACTC CACTC CACTC CTATC CTOTC CTOTC CCATC CATC CATC CATC CACC CATC	CATTG CACTG CACTG CACTG CATTG CTATG CTTTG CTCTG CCATG	CATCA CACCA CACCA CTACA CTOCA CTOCA CTOCA COCCA CCACCA CCACCA CCACCA CCACCA CCACCA CCACCA	CATCT CACCT	CATCC CACCC CACCC CACCC CTACC CTOCC CTOCC CTOCC CCACC CACCC CAACC CAACC CAACC CAACC	CATCS CACCS	CATCA CACCA	CARTET CACET CACET CACET CACET CTACT CTIGT CTOST CCOST CSACT CSACT CSACT CSACT CACCT	CATTOC CACTOC CACTOC CTATAC CTTOC CTCOC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCAC CCACC CCACC CCAC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCAC	CARCS
CAT CAC CAG CTA CTT CTC CTG CCA CCT CCC CCG CGA CCG CGA CGG GGA CGG GGA GGT GGC GGG GGG GGA GGT GGC GGG GGG GGA GGT GGC GGG GGG GGG GGG GGG GGG GGG GGG	CATAR CACER	CARM CAGM CHAR CHAR CTIME CTOM CTOM CTOM COM COM COM COM COM COM COM COM COM C	CATAC CACAC CAGAC CACAC	CATAG CACAG CACAG CACAG CTTAG CTCAG CTCAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG	CARTA CACTA CACTA CACTA CACTA CACTA CTOTA CTOTA CTOTA CCOTA COTTA COTTA COTTA CACTA	CATTI CACTI CACTI CACTI CAGTI CAGTI CTATI CTOTI CCOTI CCATI CCATI CCATI CCATI CCATI CCATI CCATI CCATI CCATI CGATI CGATI CGATI CGATI CGATI CGATI CGATI CATI CATI CATI CATI CATI CATI CATI C	CATTC CACTC CACTC CACTC CACTC CTATC CTOTC CTOTC CCATC CATC CATC CATC CACC CATC	CATTG CACTG CACTG CACTG CATTG CTATG CTTTG CTCTG CCATG	CATCA CACCA CACCA CTACA CTOCA CTOCA CTOCA COCCA CCACCA CCACCA CCACCA CCACCA CCACCA CCACCA	CATCT CACCT	CATOC CACCC CACCC CACCC CTACC CTOCC CTOCC CTOCC CCACCC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CACCC	CATCS CACCS CACCS CACCS CTACS CTACS CTOCS CTOCS CTOCS CCACS CACCS	CATCA CACCA	CARTET CACET CACET CACET CACET CTACT CTIGT CTOST CCOST CSACT CSACT CSACT CSACT CACCT	CATTOC CACTOC CACTOC CTATAC CTTOC CTCOC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCAC CCACC CCACC CCAC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCAC	CARCE
CAT CAC CAG CATA CTT CTC CTG CCA CCAT CCC CCG CCGA CGAT CGC CGG CGAA CGAT CGC CGG CGC CGG CGG CGC CGG CGG CGC CGG CGG CGC CGG CGG CGG CGC CGG	CATRA CACRA CACRA CACRA CACRA CACRA CACRA CATRA	CARM CACM CACM CACM CACM CACM CACM CACM	CATAC CACAC	CATAG CACAG CACAG CACAG CTTAG CTCAG CTCAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG	CATTA CATTA CACTTA CACTTA CACTTA CTOTTA CTOTTA CCATTA CACTTA CTCTA CTCTA	CATTI CACTI CACT CACT	CATTC CACTC CACTC CACTC CACTC CTATC CTOTC CTOTC CCATC CATC CATC CATC CACC CATC	CATTG CACTG CACTG CACTG CATTG CTATG CTTTG CTCTG CCATG	CATCA CACCA CACCA CTACA CTOCA CTOCA CTOCA COCCA CCACCA CCACCA CCACCA CCACCA CCACCA CCACCA	CATCT CACCT	CATCC CACCC CACCC CACCC CTACC CTOCC CTOCC CTOCC CCACC CACCC CAACC CAACC CAACC CAACC	CATCS CACCS	CATCA CACCA	CARTET CACET CACET CACET CACET CTACT CTIGT CTOST CCOST CSACT CSACT CSACT CSACT CACCT	CATTOC CACTOC CACTOC CACTOC CTACC CTTOC CTCOC CCTOC CC	CARGG
CAT CAC CAG CTA CTT CTC CTG CCA CCT CCG CCG CCG CGG CGA GAT GAC GAT GAC GAT GAC GAT GAC GAC GAT GAC	CATAA CACAA	CATH CACHT CACHT CTANT CTOM CTOM CTOM COM COM COM COM COM COM COM COM COM C	CATAC CACAC CAGAC CACAC CAGAC CACAC	CATAG CACAG CACAG CACAG CTTAG CTCAG CTCAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG	CARTA CACTA CACTA CACTA CACTA CTOTA CTOTA CTOTA CCOTA COTTA COTTA COTTA COTTA COTTA COTTA CTOTA CTOTA CTOTA	CATTI CACTI CACTI CACTI CAGTI CAGTI CTATI CTOTI CCOTI CCATI CCATI CCATI CCATI CCATI CCATI CCATI CCATI CCATI CGATI CGATI CGATI CGATI CGATI CGATI CGATI CATI CATI CATI CATI CATI CATI CATI C	CATTC CACTC CACTC CAGTC	CATTG CACTG CACTG CACTG CATTG CTATG CTTTG CTCTG CCATG	CATCA CACCA CACCA CTACA CTOCA CTOCA CTOCA COCCA CCACCA CCACCA CCACCA CCACCA CCACCA CCACCA	CATCT CACCT	CATCC CACCC CACC CACCC C	CATCS CACCS CACCS CACCS CTACS CTACS CTOCS CTOCS CTOCS CCACS CACCS	CATCA CACCA	CARTET CACET CACET CACET CACET CTACT CTIGT CTOST CCOST CSACT CSACT CSACT CSACT CACCT	CATTOC CACTOC CACTOC CACTOC CTACC CTTOC CTCOC CCTOC CC	CATCC CATCC CACCC CACC CACCC C
CAT CAC CAG CTA CTT CTC CTG CCC CCG CCC CCG CGA CGT CGG GAA GAT GAC GAG GAT GAC GAG GAT GAC GAG GAC GAG GAT GAC GAG GAC GAC	CATRA CACRA	CATAT CACAT	CATAC CACAC	CATAG CACAG CACAG CACAG CTTAG CTCAG CTCAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG	CANTA CACTA	CATTI CACTI CACTI CACTI CAGTI CAGTI CTATI CTOTI CCOTI CCATI CCATI CCATI CCATI CCATI CCATI CCATI CCATI CCATI CGATI CGATI CGATI CGATI CGATI CGATI CGATI CATI CATI CATI CATI CATI CATI CATI C	CATTC CACTC CACTC CACTC CACTC CTATC CTOTC CTOTC CCATC CATC CATC CATC CACC CATC	CATTG CACTG CACTG CACTG CATTG CTATG CTTTG CTCTG CCATG	CATCA CACCA CACCA CTACA CTOCA CTOCA CTOCA COCCA CCACCA CCACCA CCACCA CCACCA CCACCA CCACCA	CATCT CACCT	CATOC CACCC CACCC CACCC CTACC CTOCC CTOCC CTOCC CCACCC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CCACC CACCC	CATCS CACCS CACCS CACCS CTACS CTACS CTOCS CTOCS CTOCS CCACS CACCS	CATCA CACCA	CARTET CACET CACET CACET CACET CTACT CTIGT CTOST CCOST CSACT CSACT CSACT CSACT CACCT	CATTOC CACTOC CACTOC CACTOC CTACC CTTOC CTCOC CCTOC CC	CARGG
CAT CAC CAG CTA CTT CTC CTG CCA CCT CCG CCG CCG CGG CGA GAT GAC GAT GAC GAT GAC GAT GAC GAC GAT GAC	CATAA CACAA	CATH CACHT CACHT CTANT CTOM CTOM CTOM COM COM COM COM COM COM COM COM COM C	CATAC CACAC CAGAC CACAC CAGAC CACAC	CATAG CACAG CACAG CACAG CTTAG CTCAG CTCAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG	CANTA CACTA	CATTI CACTI CACTI CACTI CAGTI CAGTI CTATI CTOTI CCOTI CCATI CCATI CCATI CCATI CCATI CCATI CCATI CCATI CCATI CGATI CGATI CGATI CGATI CGATI CGATI CGATI CATI CATI CATI CATI CATI CATI CATI C	CATTC CACTC CACTC CAGTC	CATTG CACTG CACTG CACTG CATTG CTATG CTTTG CTCTG CCATG	CATCA CACCA CACCA CTACA CTOCA CTOCA CTOCA COCCA CCACCA CCACCA CCACCA CCACCA CCACCA CCACCA	CATCT CACCT	CATCC CACCC CACC CACCC C	CATCS CACCS CACCS CACCS CTACS CTACS CTOCS CTOCS CTOCS CCACS CACCS	CATCA CACCA	CARTET CACET CACET CACET CACET CTACT CTIGT CTOST CCOST CSACT CSACT CSACT CSACT CACCT	CATTOC CACTOC CACTOC CACTOC CTACC CTTOC CTCOC CCTOC CC	CATCC CATCC CACCC CACC CACCC C
CAT CAC CAG CTA CTT CTC CTG CCC CCG CCC CCG CGA CGT CGG GAA GAT GAC GAG GAT GAC GAG GAC GAC	CATRA CACRA	CATAT CACAT	CATAC CACAC	CATAG CACAG CACAG CACAG CTTAG CTCAG CTCAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG CCAAG	CANTA CACTA	CATTI CACTI CACTI CACTI CAGTI CAGTI CTATI CTOTI CCOTI CCATI CCATI CCATI CCATI CCATI CCATI CCATI CCATI CCATI CGATI CGATI CGATI CGATI CGATI CGATI CGATI CATI CATI CATI CATI CATI CATI CATI C	CATTC CACTC	CATTG CACTG CACTG CACTG CATTG CTATG CTTTG CTCTG CCATG	CATCA CACCA CACCA CTACA CTOCA CTOCA CTOCA COCCA CCACCA CCACCA CCACCA CCACCA CCACCA CCACCA	CATCT CACCT	CATCC CACCC CACC CACCC C	CATCS CACCS CACCS CACCS CACCS CTACS CTOCS CTOCS CTOCS CCCS CCCS CCCS CCCS	CATCA CACCA	CARTET CACET CACET CACET CACET CTACT CTIGT CTOST CCOST CSACT CSACT CSACT CSACT CACCT	CATTOC CACTOC CACTOC CACTOC CTACC CTTOC CTCOC CCTOC CC	CATCC CATCC CACCC CACC CACCC C
CAT CAC CAG CTA CTC CTC CTC CCC CCC CCC CCC CCC CCC	CATAA CACAA	CATAT CACAT	CATAC CROSS	CATAG CACAG	CANTA CACTA	CATTI CACTI CACTI CACTI CAGTI CAGTI CTATI CTOTI CCOTI CCATI CCATI CCATI CCATI CCATI CCATI CCATI CCATI CCATI CGATI CGATI CGATI CGATI CGATI CGATI CGATI CATI CATI CATI CATI CATI CATI CATI C	CATTC CACTC	CATTG CACTG CACTG CACTG CATTG CTATG CTTTG CTCTG CCATG	CATCA CACCA	CATCT CACCT	CATOC CACCO	CATCS CACCS CACCS CACCS CACCS CTACS CTOCS CTOCS CTOCS CCCS CCCS CCCS CCCS	CATCA CACCA	CARTET CACET CACET CACET CACET CTACT CTIGT CTOST CCOST CSACT CSACT CSACT CSACT CACCT	CATTOC CACTOC CACTOC CACTOC CTACC CTTOC CTCOC CCTOC CC	CATCC CATCC CACCC CACC CACCC C
CAT CAC CAG CTTA CTTC CTCG CCCT CCCC CCCA CCCT CCCC CCCA CCCA	CATAA CACAA	CATAT CACAT	CATAC CACAC	CATAG CACAG	CARTA CACTA	CATTI CACTI CACTI CACTI CAGTI CAGTI CTATI CTOTI CCOTI CCATI CCATI CCATI CCATI CCATI CCATI CCATI CCATI CCATI CGATI CGATI CGATI CGATI CGATI CGATI CGATI CATI CATI CATI CATI CATI CATI CATI C	CATTC CACTC	CATTG CACTG CACTG CACTG CATTG CTATG CTTTG CTCTG CCATG	CATCA CACCA	CATCT CACCT	CATOC CACCO	CATCS CACCS CACCS CACCS CACCS CTACS CTOCS CTOCS CTOCS CCCS CCCS CCCS CCCS	CATCA CACCA	CARTET CACET CACET CACET CACET CTACT CTIGT CTOST CCOST CSACT CSACT CSACT CSACT CACCT	CATTOC CACTOC CACTOC CACTOC CTACC CTTOC CTCOC CCTOC CC	CATOS CHASS CHASS CTASS CTOSS CTOSS CTOSS CTOSS CCOSS
CAT CAC CAG CTA CTC CTC CTC CCC CCC CCC CCC CCC CCC	CATAR CACAR	CATAT CACAT	CATAC CACAC	CATAG CACAG	CANTA CACTA CACTA CACTA CACTA CACTA CACTA CATTA CACTA	CATT CACTT C	CATTC CACTC	CATTG CACTG CACTG CACTG CATTG CTATG CTTTG CTCTG CCATG	CATCA CACCA	CATCT CACCT	CATOC CACCO	CATCS CACCS	CATCA CACCA	CARTET CACET CACET CACET CACET CTACT CTIGT CTOST CCOST CSACT CSACT CSACT CSACT CACCT	CATTOC CACTOC CACTOC CACTOC CTACC CTTOC CTCOC CCTOC CC	CARGO
CAT CAC CAG CTA CTC CTC CTC CTC CCC CCC CCC CCC CCC	CATAR CATAR CACAR CACAR CACAR CATAR CATAR CATAR CACAR	CALM CALM CALM CALM CALM CTAM CTAM CTAM COM COM COM COM COM COM COM COM COM CO	CATAC CATAC CATAC CATAC CTAC CTAC CTAC	CATAG CACAG	CARTA CACTA	CATTI CASTI CASTI CASTI CASTI CASTI CICII CICII CICII COSTI	CATTC CACTC	CATTG CACTG CACTG CACTG CATTG CTATG CTTTG CTCTG CCATG	CATCA CACCA	CATCT CACCT	CATOC CAROC	CATCS CACCS	CATCA CACCA	CARTET CACET CACET CACET CACET CTACT CTIGT CTOST CCOST CSACT CSACT CSACT CSACT CACCT	CATTOC CACTOC CACTOC CACTOC CTACC CTTOC CTCOC CCTOC CC	CATAGO CALOGO CA
CAT CAC CAG CTA CTC CTC CTC CTC CCC CCG CCG CGG CGA CGT CGC CGC CGC CGC CGC CGC CGC CGC CGC	CATAR CACAR CACAR CTARA CTARA CTARA CTORA CTORA CTORA COCRA	CATAT CACAT	CATAC CARAC CARAC CTARC CTARC CTARC CTARC CTARC CCARC	CATAG CACAG	CARTA CACTA CACTA CACTA CACTA CACTA CACTA CATTA CACTA	CASTI CACTI	CATTIC CACTIC CA	CATTO CACTO	CATCA CACCA	CATCT CACCT	CATOC CHARGE CTANCE CTANCE CTOCC CTOCC CTANCE CTOCC CCACC CC	CATCS CACCS	CATCA CACCA	CARTET CACET CACET CACET CACET CTACET CTICST CTOCST CCOCET	CATTOC CACTOC CACTOC CACTOC CTACC CTTOC CTCOC CCTOC CC	CALCAGE CALCAG
CAT CAC CAG CTA CTC CTC CTC CTC CCC CCC CCC CCC CCC	CATAR CATAR CACAR CACAR CACAR CATAR CATAR CATAR CACAR	CALM CALM CALM CALM CALM CTAM CTAM CTAM COM COM COM COM COM COM COM COM COM CO	CATAC CATAC CATAC CATAC CTAC CTAC CTAC	CATAG CACAG	CARTA CACTA	CATTI CASTI CASTI CASTI CASTI CASTI CICII CICII CICII COSTI	CATTC CACTC	CATTG CACTG CACTG CACTG CATTG CTATG CTTTG CTCTG CCATG	CATCA CACCA	CATCT CACCT	CATOC CAROC	CATCS CACCS	CATCA CACCA	CARTET CACET CACET CACET CACET CTACET CTICST CTOCST CCOCET	CATTOC CACTOC CACTOC CACTOC CTACC CTTOC CTCOC CCTOC CC	CHROS



WORKFLOW 1A



When working with clean data, fastp is convenient because it handles trimming and basic quality checks all at once, which saves time and keeps things tidy. But if I'm aiming for a more in-depth look at quality metrics or need to compare multiple runs in detail, I'd lean toward using FastQC followed by Cutadapt and summarized with MultiQC. This combination gives better insights into subtle differences, especially when data quality isn't ideal. In our case, since duplication levels were already low, fastp didn't highlight much variation, but with noisier data, those differences would likely stand out more according to my understanding.

Cutadapt Report

```
(bioinfo) megatron@Taha:/mnt/d/HumanGenome$ cutadapt -q 20 -m 20 -o sample2_trimmed_1.fq.gz -p sample2_trimmed_2.fq.gz sample2_1.fastqsanger sample2_2.fastqsanger
This is cutadapt 4.0 with Python 3.9.21
Command line parameters: -q 20 -m 20 -o sample2_trimmed_1.fq.qz -p sample2_trimmed_2.fq.qz sample2_1.fastqsanqer sample2_2.fastqsanqer
Processing paired-end reads on 1 core ...
                             339,276 reads @ 21.5 μs/read; 2.80 M reads/minutee
              00:00:07
Finished in 7.29 s (21 \mus/read; 2.79 M reads/minute).
=== Summary ===
Total read pairs processed:
                                      339,276
== Read fate breakdown ==
                                       10,175 (3.0%)
Pairs that were too short:
Pairs written (passing filters):
                                      329,101 (97.0%)
Total basepairs processed:
                             33,927,600 bp
 Read 1:
            16,963,800 bp
            16,963,800 bp
 Read 2:
Quality-trimmed:
                                807,738 bp (2.4%)
 Read 1:
               234,433 bp
               573,305 bp
  Read 2:
                             32,600,871 bp (96.1%)
Total written (filtered):
            16,310,200 bp
  Read 1:
 Read 2: 16.290.671 bp
```

The Cutadapt trimming removed 3% of reads for being too short and trimmed 2.4% due to low quality. These modest percentages indicate the original data was relatively clean, with the majority of reads meeting length and quality thresholds.

Untrimmed

♣ Copy table	₩ Configure columns	₌i∎ Scatter plot	≧ Violin plot	Export as CSV	Showing 2/2	Showing 2l_2 rows and 6l_6 columns.						
Sample Nam	e		Dups		GC	Avg len	Median len	Failed	Seqs			
sample2_1			9.396		50.0%	50 bp	50 bp	0 %	0.3M			
sample2_2			8.5%		50.0%	50 bp	50 bp	0 %	0.3 M			

Trimmed

General Statistics

♣ Copy table	Scatter plot	Export as CSV	Showing ² / ₂ rows and ⁶ / ₆ colur	↔ Summarize table			
Sample Name	Dups	GC	Avg len	Median len	Failed	Seqs	
sample2_trimmed_1	8.8%	50.0%	50 bp	50 bp	0%	0.3 M	
sample2_trimmed_2	8.1%	50.0%	50 bp	50 bp	0%	0.3 M	

The general statistics comparison shows that other than duplication and 3.9% trimmed bases, the GC % and length has not changed



The mean quality scores confirm high-quality reads both before and after trimming, with noticeable improvement post-processing. While both datasets maintained strong baseline quality, the trimming enhanced scores further without altering the fundamental profile

```
(bioinfo) megatron@Taha:/mnt/d/HumanGenome$ head sample2_aligned.sam
@SQ
        SN:NC_000001.11 LN:248956422
@SQ
        SN:NT_187361.1 LN:175055
@SQ
        SN:NT_187362.1
                        LN:32032
@SQ
        SN:NT_187363.1 LN:127682
@SQ
        SN:NT_187364.1
                       LN:66860
@SQ
        SN:NT_187365.1
                        LN:40176
@SQ
        SN:NT_187366.1 LN:42210
@SQ
        SN:NT_187367.1
                       LN:176043
        SN:NT_187368.1
@SQ
                       LN:40745
@SQ
        SN:NT_187369.1
                       LN:41717
```

After QC, the alignment was performed.

Above is a screenshot of the files generated after indexing and aligning, along with a snippet of the .sam file

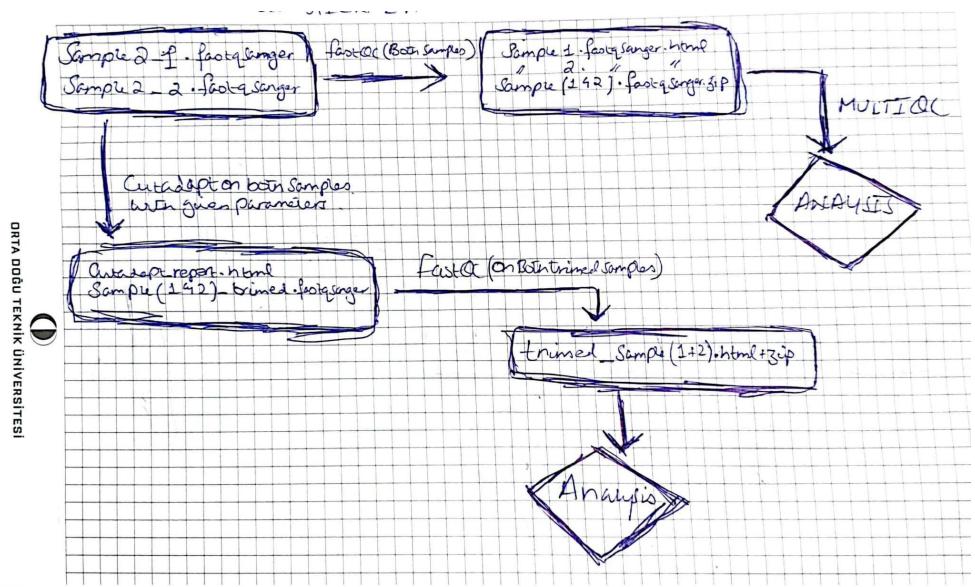
```
SN:NC_012920.1 LN:16569
       ID:bwa PN:bwa VN:0.7.17-r1188 CL:bwa mem -t 8_GCF_000001405.40_GRCh38.p14_genomic.fna sample2_trimmed_1.fq.gz sample2_trimmed_2.fq.gz
      AS1794_0001_FC61K0J:8:49:15990:3926#0
                                             83
                                                                                   60
                                                                                           50M
                                                                                                         120169669
                                                                                                                         -231
                                                                                                                                TCCTTTTGAATTTTGTACTGTATGAATGTATAAC
 AACTGAAACAAACAG
                                                                                  MD:Z:50 MC:Z:50M
                                                                                                         AS:i:50 XS:i:20
                      i:0
HWUSI-EAS1794_0001_FC61KOJ:8:49:15990:3926#0
                                             163
                                                    NC_000012.12
                                                                   120169669
                                                                                   40
                                                                                           50M
                                                                                                         120169850
                                                                                                                         231
                                                                                                                                AATTTTTGTACTTTTAGTAAAGACGGGGTTTCAC
                                                                           NM:i:1 MD:Z:38T11
                                                                                                                 AS:i:45 XS:i:45
CATGCTGGCCAGAATG
                      EHIHIIIIHIIIFIIHIHIHGIIHDIHHIIIIIGIIIIIIGIIIIEII
                                                                                                  MC:Z:50M
                                                                                   21
                                                                                                         175179057
                                                                                                                         -196
                                                                                                                                AGACACCATTCTGGGCCCCATTAAATACCGTA
HWUSI-EAS1794_0001_FC61KOJ:8:106:17677:4998#0
                                                    NC_000002.12
                                                                    175179212
                                                                                          41M9S
                      FGGGFGEGGGGGGEHEBHGGGGGGGGGFGHHHHGHGEHHHDHHHHHHHH
                                                                                                         AS:i:41 XS:i:37 XA:Z:NC_000006.12,-7331554,47M3S,2;
GAGCCCTCTCCAGTCC
                                                                           NM:i:0 MD:Z:41 MC:Z:16S34M
                                                                   175179057
                                                                                          16S34M =
                                                                                                         175179212
                                                                                                                                CTTCAGCGAAGGGTTTCTGGCATAACCAATGATA
HWUSI-EAS1794_0001_FC61KOJ:8:106:17677:4998#0
                                            163
                                                    NC_000002.12
                                                                                                                         196
                                                                                                         AS:i:34 XS:i:36 XA:Z:NC_000006.12,+7331394,9S41M,1;
AGGCTGCCAAAGACTG
                      IIIFFIGIIIGGIIIIDIIFIIIIIIGHIFIGIDEIHGIGG<FEFG@<DG
                                                                           NM:i:0 MD:Z:34 MC:Z:41M9S
HWUSI-EAS1794_0001_FC61KOJ:8:12:4514:3966#0
                                                                    92940599
                                                                                           50M
                                                                                                          92940749
                                                                                                                         200
                                                                                                                                 GAGGCTCCCGCGCCTATCCTCACCTAGGTGACTT
                                             99
                                                    NC_000014.9
                                                                                                         AS:i:50 XS:i:50 XA:Z:NT_187601.1,+55161,50M,0;
TATGGAAAGGCAGGCT
                      HDFEHHHHHGGHHHHHHHHHHHHHHHHH
                                                                           NM:i:0 MD:Z:50 MC:Z:50M
HWUSI-EAS1794_0001_FC61K0J:8:12:4514:3966#0
                                                                    92940749
                                                                                          50M
                                                                                                                         -200
                                             147
                                                    NC_000014.9
                                                                                   0
                                                                                                         92940599
                                                                                                                                TTGCCCAAATCACAGCACAAATCCTCAGCACAGG
CGCCATCGGCTCGGGC
                      HGHHFFCHEEGGHHIHEBIIHFHGEIIGIIIIII@IGHEDIDIIIIII
                                                                           NM:i:0 MD:Z:50 MC:Z:50M
                                                                                                          AS:i:50 XS:i:50 XA:Z:NT_187601.1,-55311,50M,0;
HWUSI-EAS1794_0001_FC61KOJ:8:38:8158:19548#0
                                             65
                                                    NC_000003.12
                                                                   101794419
                                                                                   60
                                                                                          50M
                                                                                                  NC_000017.11
                                                                                                                 64502186
                                                                                                                                        GGGTAGGACATTCAGGGAGTTTCTGG
CTGTGAAGAAGATGGAAGGTAGGA
                              GGGBGIIHIIIIIIIIIHIFIIIIIIIIIIIIIIIIIHIGEEGGBGGGE
                                                                                   NM:i:0
                                                                                          MD:Z:50 MC:Z:40M
                                                                                                                 AS:i:50 XS:i:0
HWUSI-EAS1794_0001_FC61KOJ:8:38:8158:19548#0
                                             129
                                                                    64502186
                                                                                          40M
                                                                                                  NC_000003.12
                                                                                                                 101794419
                                                                                                                                        GTTGACTCTTGTCACCATGGATACCC
                                                    NC_000017.11
                                                                                   60
ATGGCAGGCCACCC IIIIIIIIIIIIHGIIEGIGBGIIIGIBIIHHIIBD>DDD
                                                                                   MC:Z:50M
                                                                                                  AS:i:39 XS:i:0
                                                            NM:i:1
                                                                   MD:Z:39T0
HWUSI-EAS1794 0001 FC61KOJ:8:106:18976:17584#0 81
                                                                   23310392
                                                                                                                         -15325 ACAGACAGGGGTCTTCGGAGTCAAGATTGCTGTG
                                                    NC_000022.11
                                                                                   0
                                                                                           43M7S
                                                                                                 =
                                                                                                         23295110
GTCACCAAGAGAGAGA
                      NM:i:0 MD:Z:43 MC:Z:46M4S
                                                                                                         AS:i:43 XS:i:43 XA:Z:NC_000022.11,-24645704,43M7S,0;NC_000
022.11,-22631029,43M7S,0;
                                                                                   60
HWUSI-EAS1794_0001_FC61KOJ:8:106:18976:17584#0 161
                                                    NC_000022.11
                                                                   23295110
                                                                                          46M4S
                                                                                                         23310392
                                                                                                                         15325
                                                                                                                                GGACGGCGAGAGCACGGACAGACTCATGGGGAAG
                      IIIIIIHHIIHIHIIGIIIGEBHIIFIIIIIIIIGIIIIHHBIHHIHEII
                                                                                                         AS:i:46 XS:i:0
GGCCAGGTCCAGCTGG
                                                                           NM:i:0 MD:Z:46 MC:Z:43M7S
HWUSI-EAS1794_0001_FC61K0J:8:12:5094:3287#0
                                                                   128184805
                                                                                          50M
                                                                                                         128186598
                                                                                                                         1843
                                                                                                                                GGTTAAATGAGTTATTAAAGGAACATTCAAGCAC
                                             99
                                                    NC_000005.10
                                                                                   60
AGCTAATATTATTGTC
                      NM:i:0 MD:Z:50 MC:Z:50M
                                                                                                         AS:i:50 XS:i:19
                                                                                                         128184805
HWUSI-EAS1794_0001_FC61KOJ:8:12:5094:3287#0
                                             147
                                                    NC_000005.10
                                                                   128186598
                                                                                   60
                                                                                          50M
                                                                                                                         -1843
                                                                                                                                GAATCATCAGAGTGTCCTTACCTTCTATTCATAA
ATGTTCTATACAGTGG
                      NM:i:0 MD:Z:50 MC:Z:50M
                                                                                                         AS:i:50 XS:i:19
```

2.b) position 120,169,850 on chromosome 12. (highlighted in yellow)

```
(bioinfo) megatron@Taha:/mnt/d/HumanGenome$ head -n 5 sample2_sorted_depth.txt
NC_012920.1
                8648
                        1028
NC_012920.1
                8647
                        1012
NC_012920.1
                8650
                        992
NC_012920.1
                8649
                        989
NC_012920.1
                8643
                        975
```

2.c) NC_012920.1 which corresponds to chrM, reads aligned the most

Q2) QC WorkFlow



TEKNIK ÜNIVERSITESI

Linux History

```
megatron@Taha:~$ sudo apt update
megatron@Taha:~$ source ~/miniconda3/bin/activate
(base) megatron@Taha:~$ conda activate bioinfo
(bioinfo) megatron@Taha:~$ conda env list
(bioinfo) megatron@Taha:/mnt/d/HumanGenome/bin508-03$ mkdir raw fastqc
(bioinfo) megatron@Taha:/mnt/d/HumanGenome$ fastqc sample1 1.fq.gz sample1 2.fq.gz -o raw fastqc/
(bioinfo) megatron@Taha:/mnt/d/HumanGenome$ multiqc raw fastqc/ -o raw multiqc/
(bioinfo) megatron@Taha:/mnt/d/HumanGenome$ cutadapt -q 20 -m 20 -o trimmed 1.fq.gz -p trimmed 2.fq.gz sample1 1.fq.gz sample1 2.fq.gz
(bioinfo) megatron@Taha:/mnt/d/HumanGenome$ mkdir trimmed fastqc
(bioinfo) megatron@Taha:/mnt/d/HumanGenome$ fastqc trimmed 1.fq.gz trimmed 2.fq.gz -o trimmed fastqc/
(bioinfo) megatron@Taha:/mnt/d/HumanGenome$ multiqc trimmed fastqc/ -o trimmed multiqc/
(bioinfo) megatron@Taha:/mnt/d/HumanGenome$ fastp -i sample1 1.fq.gz -I sample1 2.fq.gz -o fastp 1.fq.gz -O fastp 2.fq.gz --qualified quality phred 20
 -length_required 20 --html fastp report.html
(bioinfo) megatron@Taha:/mnt/d/HumanGenome$ mkdir sample2 raw fastqc
(bioinfo) megatron@Taha:/mnt/d/HumanGenome$ fastqc sample2 1.fastqsanger sample2 2.fastqsanger -o sample2 raw fastqc/
```

```
(bioinfo) megatron@Taha:/mnt/d/HumanGenome$ cutadapt -q 20 -m 20 -o sample2 trimmed 1.fq.gz -p sample2 trimmed 2.fq.gz sample2 1.fastqsanger
    sample2 2.fastqsanger
     (bioinfo) megatron@Taha:/mnt/d/HumanGenome$ mkdir sample2 trimmed fastqc
    (bioinfo) megatron@Taha:/mnt/d/HumanGenome$ fastqc sample2 trimmed 1.fq.gz sample2 trimmed 2.fq.gz -o sample2 trimmed fastqc/
     (bioinfo) megatron@Taha:/mnt/d/HumanGenome$ multiqc sample2 trimmed fastqc/ -o sample2 trimmed multiqc/
     (bioinfo) megatron@Taha:/mnt/d/HumanGenome$ bwa mem -t 8 -R "@RG\tID:sample2\tSM:sample2\tPL:ILLUMINA" /mnt/d/HumanGenome/
    GCF 000001405.40 GRCh38.p14 genomic.fna sample2 trimmed 1.fq.gz sample2 trimmed 2.fq.gz > sample2 aligned.sam
     (bioinfo) megatron@Taha:/mnt/d/HumanGenome$ head sample2 aligned.sam
     (bioinfo) megatron@Taha:/mnt/d/HumanGenome$ mkdir -p sample2 q2c results/alignment sample2 q2c results/depth
     (bioinfo) megatron@Taha:/mnt/d/HumanGenome$ samtools sort sample2 aligned.sam -o sample2 q2c results/alignment/sample2 q2c sorted.bam
     (bioinfo) megatron@Taha:/mnt/d/HumanGenome$ samtools depth sample2 q2c results/alignment/sample2 q2c sorted.bam > sample2 q2c results/depth/
    sample2 q2c depth.txt
     (bioinfo) megatron@Taha:/mnt/d/HumanGenome$ sort -k3,3nr sample2 q2c results/depth/sample2 q2c depth.txt > sample2 q2c results/depth/
    sample2_q2c_sorted_depth.txt
     (bioinfo) megatron@Taha:/mnt/d/HumanGenome$ head -n 5 sample2_q2c_results/depth/sample2_q2c_sorted_depth.txt
81
```