

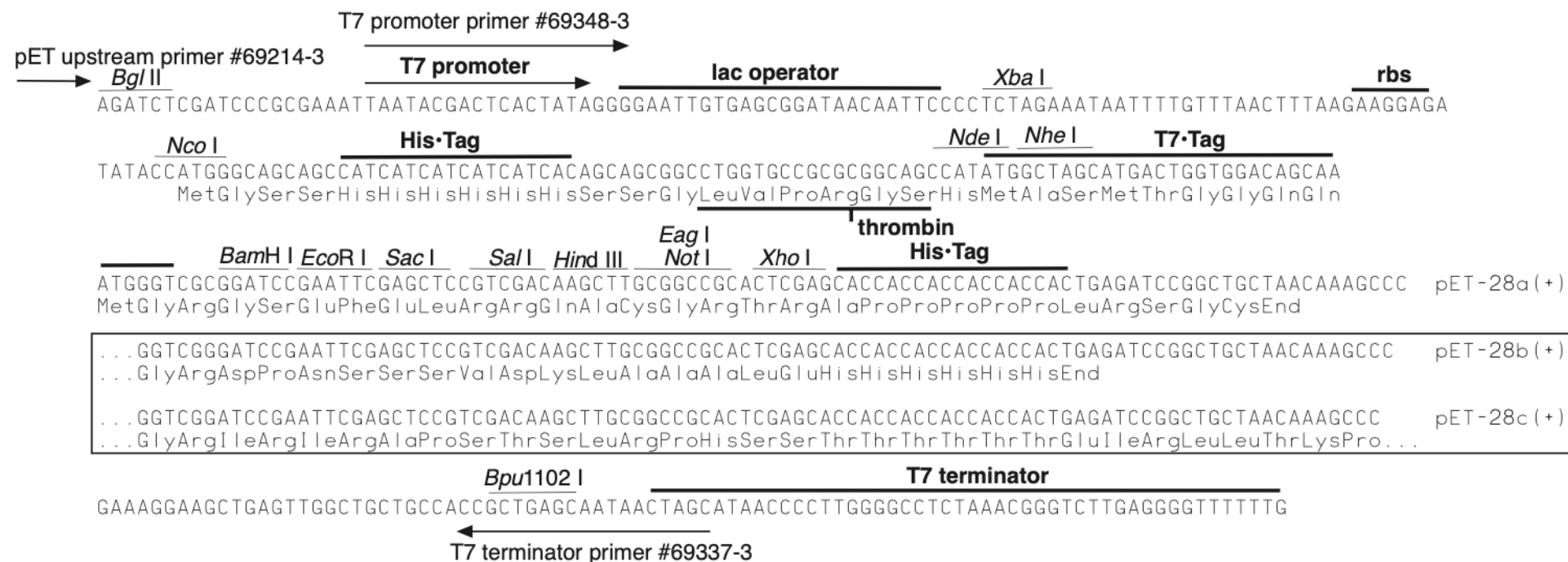
The background of the slide is a dark blue/black field filled with numerous thin, vertical lines of varying heights and colors. The colors include bright green, cyan, magenta, and light blue, creating a dynamic, data-driven visual effect reminiscent of a DNA microarray or sequencing gel.

# **ANALYZING DNA SEQUENCING DATA**

---

JUNE 14, 2023

# FIRST STEP – KNOW YOUR PLASMID!



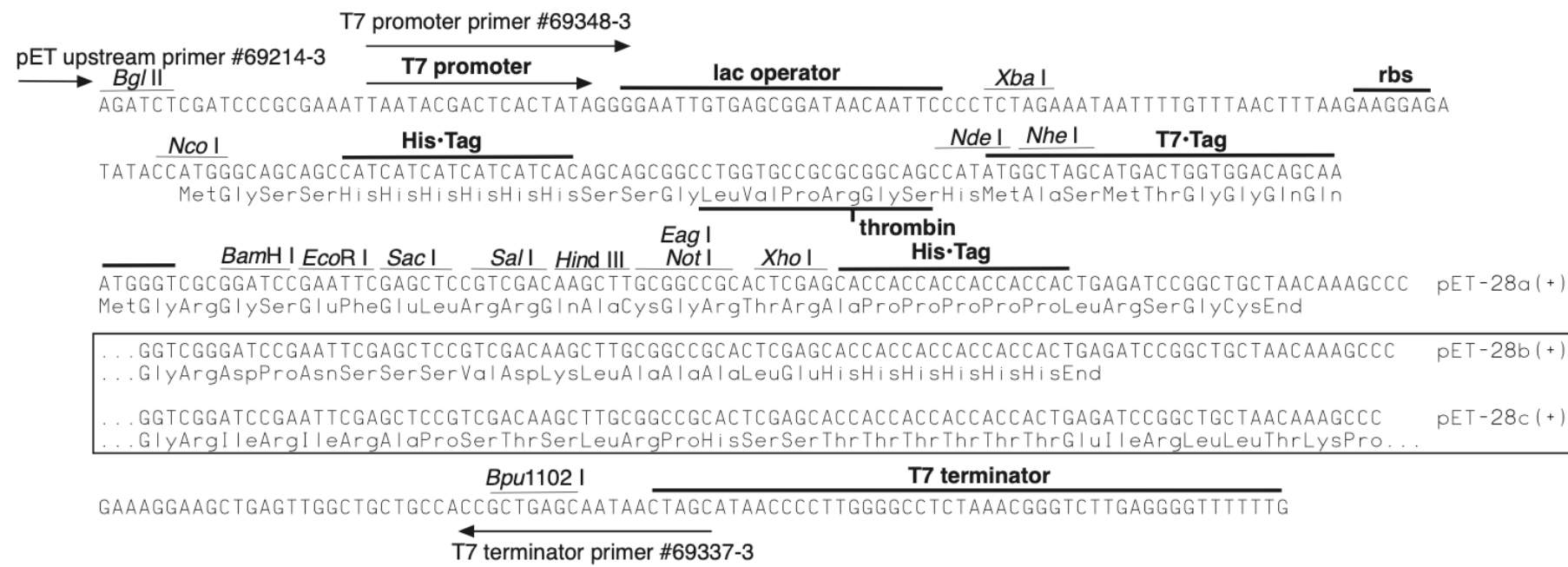
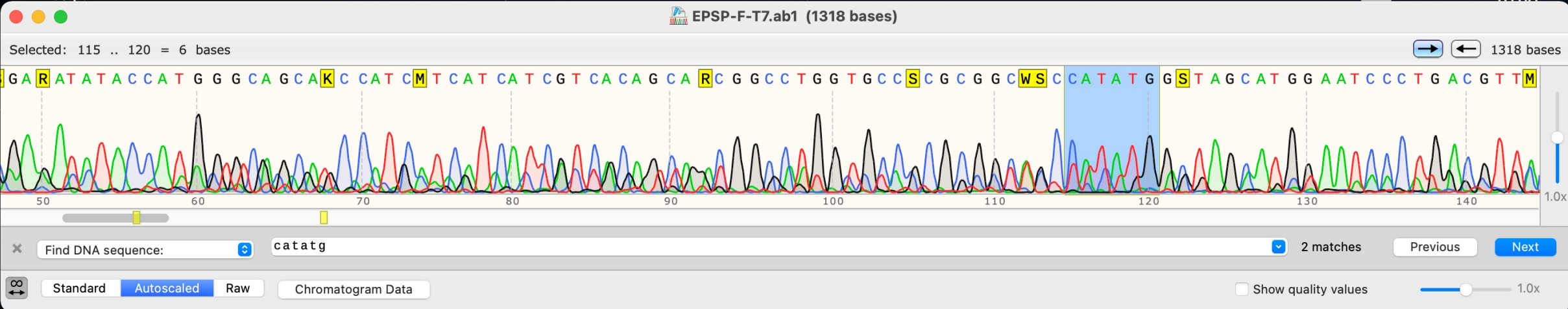
**pET-28a-c(+)** cloning/expression region

**Novagen • ORDERING 800-526-7319 • TECHNICAL SUPPORT 800-207-0144**

For an insert cloned as NdeI/XhoI fragment:

- Look for the ...GGCAGCCATATG from the T7 forward sequencing reaction
- Look for the ...GTGGTGCTCGAG from the T7 reverse sequencing reaction

Look for the ...GGCAGCCATATG from the T7 forward sequencing reaction



pET-28a-c(+) cloning/expression region

Can manually call some ambiguous reads



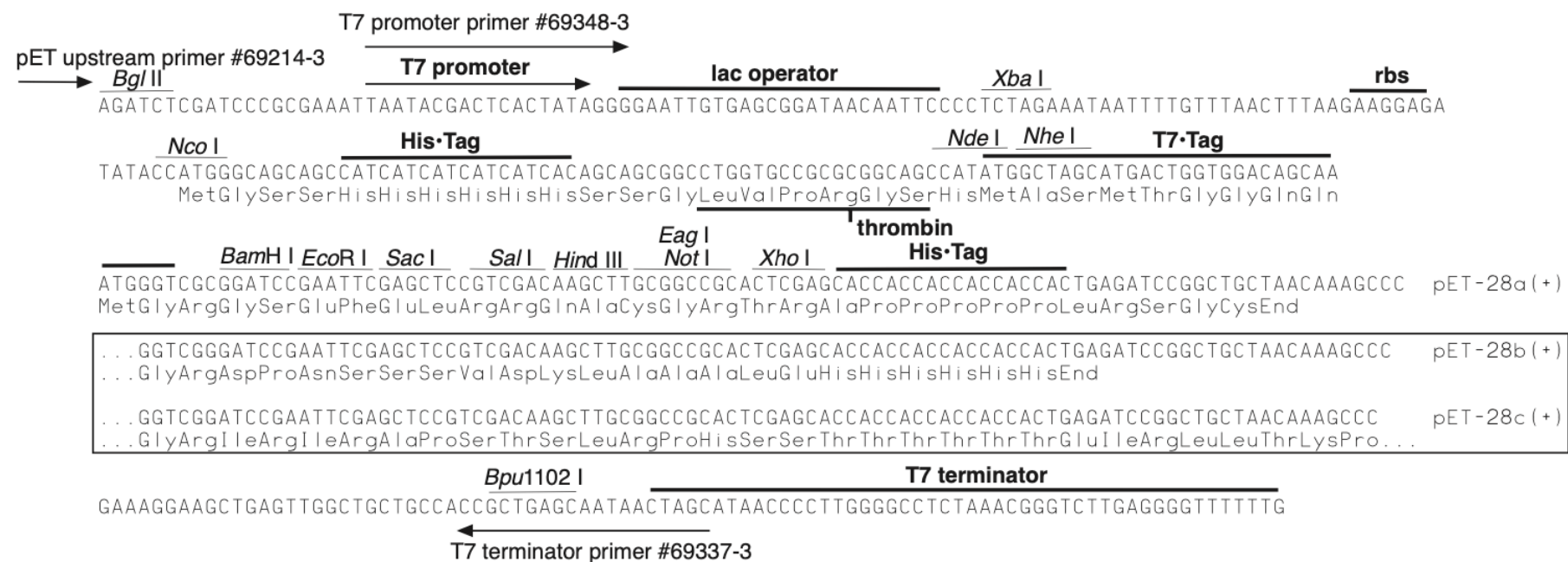
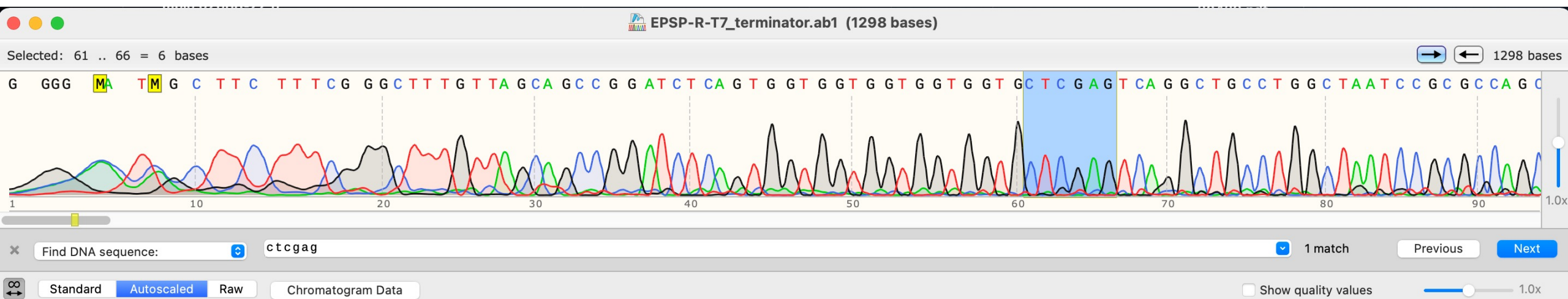
**CALL AS MANY AMBIGUOUS BASES AS YOU CAN...SAVE A "GENE-FWD" FILE**



The image shows a terminal window with a dark background and a title bar that reads "EPSP-Fwd". The terminal displays a long DNA sequence in all caps, consisting of 20 lines of text. The sequence is a single forward strand, as indicated by the title. The sequence starts with "GKKGAMSGTAATTYTCGYKCTAGAATCATTTTGT" and ends with "CAGAGTCCGATCAMTAGATAMCGCRWGCAAKCGTGTACTRCGTCGKGGCGYGGCATGCGAWMTSAGCTASGTACGCATCCTYGAGYMTGCGA".

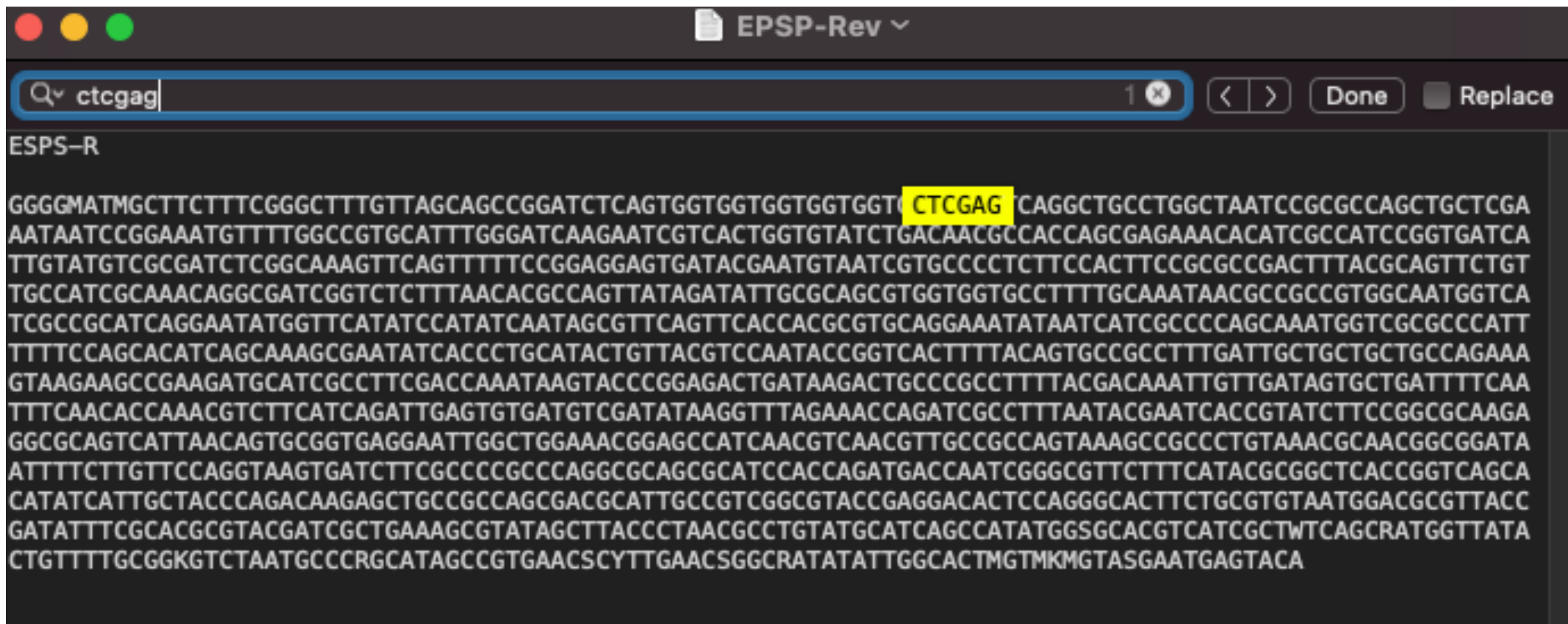
```
EPSP-Fwd
EPSP-F
GKKGAMSGTAATTYTCGYKCTAGAATCATTTTGT
TTACTTTAASAASGARATATACCATGGGCAGCAKCCATCMT
CATCATCGTCACAGCARCGGCCTGGT
GCCGCGCGGCAGC
CATATGGGTAGCATGGAATCCCTGACGTTACATCCCATCGCTCGTGT
CGATGGCACTATTAATCTGCCCGGTTCCAAGAGCGTTTCTAACC
GCGCTTTATT
GCTGGCGGCATTAGCACACGGCAAAACAGTATTAACCAATCTGCTGGATAGCGATGACGTGCGCCATATGCTGAATGCATTAACAGCGTTAGGGGTAAGCT
ATACGCTTTCAGCCGATCGTACGCGTTGCGAAATTATCGGTAACGGCGGTCCATTACACGCAGAAGGTGCCCTGGAGTTGTTCTCGGTAACGCCGGAACG
GCAATGCGTCCGCTGGCGGCAGCTCTTTGTCTGGGTAGCAATGATATTGTGCTGACCGGTGAGCCGCGTATGAAAGAACGCCCGATTGGTCATCTGGTGGA
TGCGCTGCGCCTGGGCGGGGCGAAGATCACTTACCTGGAACAAGAAAATTATCCGCCGTTGCGTTTACAGGGCGGCTTTACTGGCGGCAACGTTGACGTTG
ATGGCTCCGTTTCCAGCCAATTCCTCACCGCACTGTTAATGACTGCGCCTCTTGCGCCGGAAGATACGGTGATTTCGTATTAAAGGCGATCTGGTTTCTAAA
CCTTATATCGACATCACACTCAATCTGATGAAGACGTTTGGTGTTGAAATTGAAAATCAGCACTATCAACAATTTGTCGTAAAAGGCGGGCAGTCTTATCA
GTCTCCGGGTACTTATTTGGTCGAAAGCGATGCATCTTCGGCTTCTTACTTTCTGGCAGCAGCAGCAATCAAAGGCGGCACTGTAAAGTGACCGGTATTGG
ACGTAACAGTATGCAGGGTGATATTCGCTTTGCTGATGTGCTGGAAAAATGGCGCGACCATTTGCTGGGGCGATGATTATATTTCTGACGCGTGTGAAC
TGACGCTATTGATATGGATATGAACCATATCCTGATGCGGCGATGACATTGGCYMCGGCGGCGTWTGCAAGGAYCAMCMGCTGCCAWATCTATACTGG
CGKGTTAARAGAACGAWYSCYGYTKSCATGCAMCAAAYKMSTAAGTCGCSCGAGTGAAAAGGRGWCAATAACTTCATWCCTCCCTCGGARAACKAACTTTG
CAGAGTCCGATCAMTAGATAMCGCRWGCAAKCGTGTACTRCGTCGKGGCGYGGCATGCGAWMTSAGCTASGTACGCATCCTYGAGYMTGCGA
```

Look for the ...GTGGTGCTCGAG from the T7 reverse sequencing reaction



pET-28a-c(+) cloning/expression region

CALL AS MANY AMBIGUOUS BASES AS YOU CAN...SAVE A "GENE-REV" FILE



The screenshot shows a text editor window titled "EPSP-Rev". At the top, there are three colored window control buttons (red, yellow, green). Below the title bar is a search bar containing the text "ctcgag". To the right of the search bar are navigation buttons: a left arrow, a right arrow, a "Done" button, and a "Replace" button with a checkbox. The main text area displays a DNA sequence labeled "ESPS-R". The sequence is a single line of text with many bases. One instance of the sequence "CTCGAG" is highlighted in yellow. The sequence starts with "GGGGMATMGCTTCTTTCGGGCTTTGTTAGCAGCCGGATCTCAGTGGTGGTGGTGGT" and ends with "CTGTTTTGCGGKGTCTAATGCCRCGATAGCCGTGAACSCYTTGAACSGGCRATATATTGGCACTMGTMKMGTA SGAATGAGTACA".

```
GGGGMATMGCTTCTTTCGGGCTTTGTTAGCAGCCGGATCTCAGTGGTGGTGGTGGTCTCGAG CAGGCTGCCTGGCTAATCCGCGCCAGCTGCTCGA
AATAATCCGGAAATGTTTTGGCCGTGCATTTGGGATCAAGAATCGTCACTGGTGTATCTGACAACGCCACCAGCGAGAAACACATCGCCATCCGGTGATCA
TTGTATGTCGCGATCTCGGCAAAGTTCAGTTTTTCCGGAGGAGTGATACGAATGTAATCGTGCCCTCTTCCACTTCCGCGCCGACTTTACGCAGTTCTGT
TGCCATCGCAAACAGGCGATCGGTCTCTTTAACACGCCAGTTATAGATATTGCGCAGCGTGGTGGTGCCTTTTGCAAATAACGCCGCCGTGGCAATGGTCA
TCGCCGCATCAGGAATATGGTTCATATCCATATCAATAGCGTTCAGTTCACCACGCGTGCAGGAAATATAATCATCGCCCCAGCAAATGGTCGCGCCCATT
TTTTCCAGCACATCAGCAAAGCGAATATCACCTGCTACTGTTACGTCCAATACCGGTCACCTTTACAGTGCCGCCTTTGATTGCTGCTGCTGCCAGAAA
GTAAGAAGCCGAAGATGCATCGCCTTCGACCAAATAAGTACCCGGAGACTGATAAGACTGCCCGCCTTTTACGACAAATTGTTGATAGTGCTGATTTTCAA
TTTCAACACCAAACGTCTTCATCAGATTGAGTGTGATGTCGATATAAGGTTTAGAAACCAGATCGCCTTTAATACGAATCACCGTATCTTCCGGCGCAAGA
GGCGCAGTCATTAACAGTGCGGTGAGGAATTGGCTGGAAACGGAGCCATCAACGTCAACGTTGCCGCCAGTAAAGCCGCCCTGTAAACGCAACGGCGGATA
ATTTTCTTGTTCCAGGTAAGTGATCTTCGCCCCGCCAGGCGCAGCGCATCCACCAGATGACCAATCGGGCGTTCTTTCATACGCGGCTCACCGGTCAGCA
CATATCATTGCTACCCAGACAAGAGCTGCCGCCAGCGACGCATTGCCGTCGGCGTACCGAGGACACTCCAGGGCACTTCTGCGTGTAATGGACGCGTTACC
GATATTTGCGACGCGTACGATCGCTGAAAGCGTATAGCTTACCCTAACGCCTGTATGCATCAGCCATATGSGGCACGTCATCGCTWTCAGCRATGGTTATA
CTGTTTTGCGGKGTCTAATGCCRCGATAGCCGTGAACSCYTTGAACSGGCRATATATTGGCACTMGTMKMGTA SGAATGAGTACA
```



MAKE  
REVERSE  
COMPLEMENT

Reverse Complement

genscript.com/sms2/rev\_comp.html

GenScript  
The Biology CRO

Format Conversion

-Combine FASTA

-EMBL to FASTA

-EMBL Feature Extractor

-EMBL Trans Extractor

-Filter DNA

-Filter Protein

-GenBank to FASTA

-GenBank Feature Extractor

-GenBank Trans Extractor

-One to Three

-Range Extractor DNA

-Range Extractor Protein

-Reverse Complement

-Three to One

Sequence Analysis

-Codon Plot

-Codon Usage

-CpG Islands

-DNA Pattern Find

-DNA Stats

-Fuzzy Search DNA

-Fuzzy Search Protein

-Ident and Sim

-Multi Rev Trans

-Mutate for Digest

-ORF Finder

-Pairwise Align DNA

-Pairwise Align Protein

-PCR Primer Stats

-PCR Products

-Protein Isoelectric Point

-Protein Molecular Weight

-Protein Pattern Find

-Protein Stats

-Restriction Digest

-Restriction Summary

-Reverse Translate

-Translate

Sequence Figures

-Color Align Conservation

-Color Align Properties

-Group DNA

-Group Protein

-Primer Map

-Restriction Map

-Translation Map

Random Sequences

-Mutate DNA

-Mutate Protein

-Random Coding DNA

-Random DNA Sequence

-Random Protein Sequence

-Sample DNA

Sequence Manipulation Suite:

Reverse Complement

Reverse Complement converts a DNA sequence into its reverse, complement, or reverse-complement counterpart. The entire IUPAC DNA alphabet is supported, and the case of each input sequence character is maintained. You may want to work with the reverse-complement of a sequence if it contains an ORF on the reverse strand.

Paste the raw sequence or one or more FASTA sequences into the text area below. Input limit is 100000 characters.

>Sample sequence 1

GGGGMATMGCTTCTTTTCGGGCTTTGTTAGCAGCCGGATCTCAGTGGTGGTGGTGGTGGTGC

TTCGAGTCAGGCTGCCTGGCTAATCCGCGCCAGCTGCTCGA

CCGTGCATTGGGATCAAGAATCGTCACTGGTGTATCTGA

CATCGCCATCCGGTGATCATTGTATGTCGCGATCTCGGC

GTGATACGAATGTAATCGTGCCCTCTTCCACTTCCGCG

Please check the browser compatibility page before

Submit

Clear

Reset

• reverse-complement

\*This page requires JavaScript. See browser comp

\*You can mirror this page or use it off-line.

GenScript Corporation

• Gene Synthesis for only \$0.28/bp. More ...

• Monoclonal antibody: starting from \$3,800. More ...

• PolyExpress™ antibody: starting from \$499. More ...

SMS - Output

Reverse Complement results

>Sample sequence 1 reverse complement

TGTACTCATTCTSTACKMKACKAGTGCCAATATATYGCCSGTTCAARGSGTTCACGGCTAT

GCYGGGCATTAGACMCCGCAAAACAGTATAACCATYGCTGAWAGCGATGACGTGCSCCAT

ATGGCTGATGCATACAGGCGTTAGGGTAAGCTATACGCTTTCAGCGATCGTACGCGTGCG

AAATATCGGTAACGCGTCCATTACACGCAGAAGTGCCCTGGAGTGCTCTCGGTACGCCGA

CGGCAATGCGTCGCTGGCGGCAGCTCTTGCTCTGGGTAGCAATGATATGTGCTGACCGGTG

AGCCGCGTATGAAAGAACGCCCATTGGTCATCTGGTGGATGCGCTGCGCCTGGGCGGGG

CGAAGATCACTTACCTGGAACAAGAAAATATCCGCCGTTGCGTTTACAGGGCGGCTTAA

CTGGCGGCAACGTTGACGTTGATGGCTCCGTTTCCAGCCAATTCCTCACCACGACTGTAA

TGACTGCGCCTCTTGCGCCGGAAGATACGGTGATTCGTATTAAAGGCGATCTGGTTTCTA

AACCTTATATCGACATCACACTCAATCTGATGAAGACGTTTGGTGTTGAAATTGAAAATC

AGCACTATCAACAATTTGTCGTAAGGCGGGCAGTCTTATCAGTCTCCGGGTACTTATT

TGGTCTGAAGGCGATGCATCTTCGGCTTCTTACTTCTGGCAGCAGCAGCAATCAAAGCGG

GCACTGTAAAAGTGACCGGTATTGGACGTAACAGTATGCAGGGTGATATTTCGCTTTGCTG

ATGTGCTGGAATAAATGGGCGCAGCATTGCTGGGGCGATGATTATATTTCTGCACGC

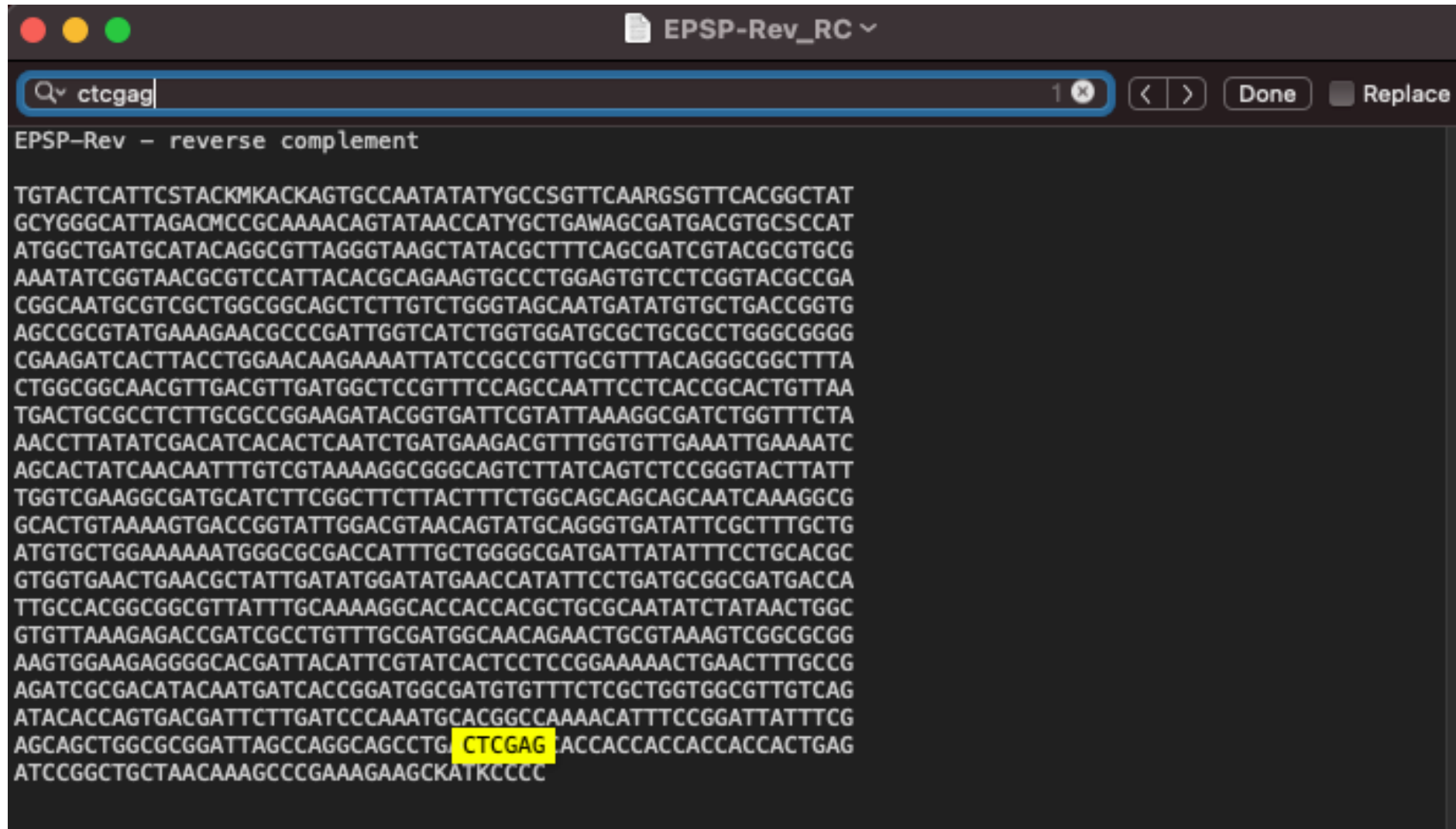
GTGGTGAAGTGAACGCTATTGATATGGATATGAACCATATTCCTGATGCGGCGATGACCA

TTGCCACGGCGCGTTATTTGCAAAAGGCACCACCGCTGCGCAATATCTATAACTGGC

GTGTTAAAGAGACCGATCGCCTGTTTGGCGATGGCAACAGAACTGCGTAAAGTCGGCGCGG

AAGTGGAAGAGGGGACGATTACATTCGTATCACTCCTCCGGAATACTGAACCTTGCCG

# MAKE REVERSE COMPLEMENT AND SAVE A “GENE-REV\_RC” FILE



```
EPSP-Rev - reverse complement










TGTA CTCA TTCSTA CKMKACKAGTGCCAATATATYGCCSGTTCAARGSGTTCACGGCTAT
GCYGGGCATTAGACMCCGCAAACAGTATAACCATYGCTGAWAGCGATGACGTGCSCCAT
ATGGCTGATGCATACAGGCGTTAGGGTAAGCTATACGCTTTCAGCGATCGTACGCGTGCG
AAATATCGGTAACGCGTCCATTACACGCAGAAGTGCCCTGGAGTGTCTCGGTACGCCGA
CGGCAATGCGTCGCTGGCGGCAGCTCTTGTCTGGGTAGCAATGATATGTGCTGACCGGTG
AGCCGCGTATGAAAGAACGCCCGATTGGTCATCTGGTGGATGCGCTGCGCCTGGGCGGGG
CGAAGATCACTTACCTGGAACAAGAAAATTATCCGCCGTTGCGTTTACAGGGCGGCTTTA
CTGGCGGCAACGTTGACGTTGATGGCTCCGTTTCCAGCCAATTCCTCACC GCACTGTAA
TGACTGCGCCTCTTGCGCCGGAAGATACGGTGATTCTGATTAAGGCGATCTGGTTTCTA
AACCTTATATCGACATCACACTCAATCTGATGAAGACGTTTGGTGTGAAATTGAAATC
AGCACTATCAACAATTTGTCGTAAAAGGCGGGCAGTCTTATCAGTCTCCGGGTACTTATT
TGGTCGAAGGCGATGCATCTTCGGCTTCTTACTTTCTGGCAGCAGCAGCAATCAAAGGCG
GCACTGTAAAAGTGACCGGTATTGGACGTAACAGTATGCAGGGTGATATTCGCTTTGCTG
ATGTGCTGGA AAAAATGGGCGCGACCATTTGCTGGGGCGATGATTATATTTCTGCACGC
GTGGTGA ACTGAACGCTATTGATATGGATATGAACCATATTCCTGATGCGGCGATGACCA
TTGCCACGGCGGCGTTATTTGCAAAGGCACCACCACGCTGCGCAATATCTATAACTGGC
GTGTTAAAGAGACCGATCGCCTGTTTGCGATGGCAACAGAACTGCGTAAAGTCGGCGCGG
AAGTGGAAGAGGGGACGATTACATTCGTATCACTCCTCCGGAAAACTGAACTTTGCCG
AGATCGCGACATACAATGATCACCGGATGGCGATGTGTTTCTCGCTGGTGGCGTTGTCAG
ATACACCA GTGACGATTCTTGATCCCAAATGCACGGCCAAAACATTTCCGGATTATTTG
AGCAGCTGGCGCGGATTAGCCAGGCAGCCTG CTCGAG ACCACCACCACCACCTGAG
ATCCGGCTGCTAACAAAGCCCGAAAGAAGCKATKCCC
```



# ALIGN THE “GENE-FWD” AND “GENE-REV\_RC” SEQUENCES

← → ↺

ebi.ac.uk/Tools/msa/clustalo/



Input form

Web services

Help & Documentation

Bioinformatics Tools FAQ

Feedback

STEP 1 - Enter your input sequences

Enter or paste a set of

DNA

sequences in any supported format:

TTGATATGGATATGAAAGCATATGCTGATGCGGGGATGAGATGAGTGGGTTGGGGGGGGGTTWTTTGGAGAGATGAMOMMMGGTGGCAWATGTATAGTGGGAGTTAA  
RAGAACGAWYSCYGYTKSCATGCAMCAAAYKMSTAAGTCGCSCGAGTGAAAAGGRGWCAATAACTTCATWCCTCCCTCGGARAACKAACTTTGCAGAGTC  
CGATCAMTAGATAMCGCRWGGCAAKCGTGTACTRCGTGCGKGGCGYGGCATGCGAWMTSAGCTASGTACGCATCCTYGAGYMTGCGA

>Rev  
TGTAATCATTCTACKMKACKAGTGCCAATATATYGCCSGTTCAARGSGTTCACGGCTAT  
GCYGGGCATTAGACMCCGCAAAACAGTATAACCATYGCTGAWAGCGATGACGTGCSCCAT  
ATGGCTGATGCATACAGGCGTTAGGGTAAGCTATACGCTTTCAGCGATCGTACGCGTGCG  
AAATATCCCTAAGCCCTGCATTACAGCCAGAACTCCGCTCCACTCTGCTCCCTAGCCCA

Or, upload a file: 

Choose File

 No file chosen

Use a example sequence

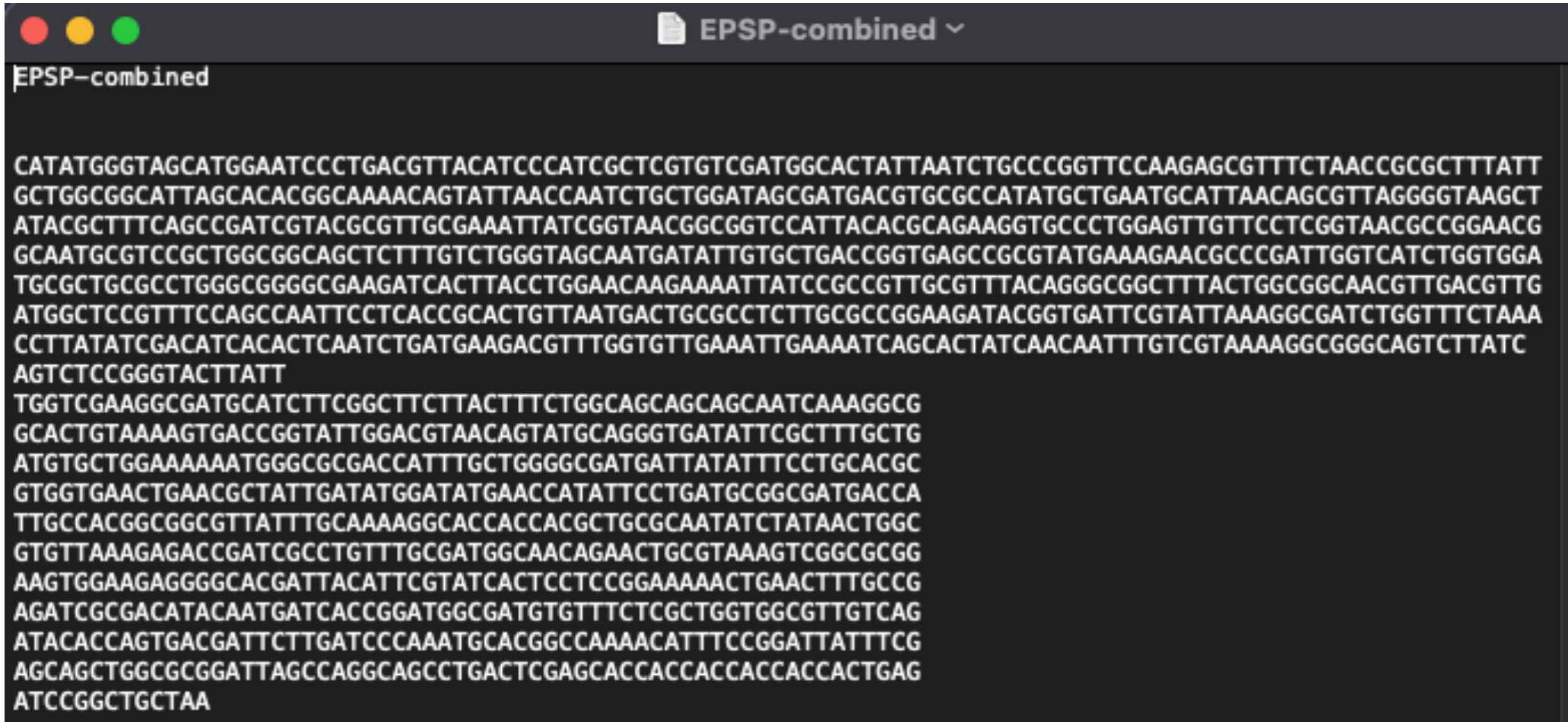
Clear sequence

See more example inputs

# FIND OVERLAPPING REGION TO JOIN FWD AND REV SEQUENCES

align_EPSP.clustal_num			CLUSTAL 0(1.2.4) multiple sequence alignment	
Fwd	CATATGGGTAGCATGGAATCCCTGACGTTACATCCCATCGTCGTGTCGATGGCACTATT	60		
Rev	-----TGTACTATTCTACKMKA	19		
			* ** * **	
Fwd	AATCTGCCCGTTCCAAGAGCGTTTCTAACCGCGTTTATTGCTGCCGGGATTAGCACAC	120		
Rev	CKAGTGCC---AATATATYGCCSGTTCAARGSGTTACGCG---CTATGCGGGCATTAGACMC	76		
	*** * * * * * * * * * * * * * * *		** ** * ** *	
Fwd	GGCAAAACAGTATTAAACCAATCTGCTGGATAGCGATGACGTGCGCCATATGCTGAATGCA	180		
Rev	CGCAAAACAGTATAAC-----CATYCTGAWAGCGATGACGTGCGCCATATGGCTGATGCA	132		
	***** * * * * * * * * * * * * * * *			
Fwd	TTAACAGCGTTAGGGTAAGCTATACGCTTTTCAGCCGATCGTACGCGTTGCGAAATTATC	240		
Rev	TACA---GGCGTTAGGGTAAGCTATACGCTTTTCAGCGAT---CGTACGCGTGCAGAA---TATC	187		
	* * * * * * * * * * * * * * * * * * *		** **	
Fwd	GGTAACGGCGGTCCATTACACGCAGAAGTGCCCTGGAGTTGTTCTCGGTAAACGCCGA	300		
Rev	GGT---AACGCGTCCATTACACGCAGAAGTGCCCT---GGAGTGTCT---CTCGG---TACGCCG	239		
	*** * * * * * * * * * * * * * * *			
Fwd	ACGGCAATGCGTCCGCTGGCGGCAGCTCTTTGCTGGGTAGCAATGATATTGTGTGACC	360		
Rev	ACGGCAATGCG---TCGCTGGCGGCAGC---TCTTGTCTGGGTG---GCAATGATATGTGTGACC	296		
	***** * * * * * * * * * * * * * * *			
Fwd	GGTGAGCCGCGTATGAAAGAACGCCCGATTGGTCATCTGCTGGATGCGCTGCGCTGGGC	420		
Rev	GGTGAGCCGCGTATGAAAGAACGCCCGATTGGTCATCTGCTGGATGCGCTGCGCTGGGC	356		
	*****			
Fwd	GGGGCGAAGATCACTTACCTGGAACAAGAAAATTATCCGCCGTTGCGTTTACAGGGCGGC	480		
Rev	GGGGCGAAGATCACTTACCTGGAACAAGAAAATTATCCGCCGTTGCGTTTACAGGGCGGC	416		
	*****			
Fwd	TTTACTGGCGCAACGTTGACGTTGATGGCTCCGTTTCCAGCCAATTCTCACCGCACTG	540		
Rev	TTTACTGGCGCAACGTTGACGTTGATGGCTCCGTTTCCAGCCAATTCTCACCGCACTG	476		
	*****			
Fwd	TTAATGACTGCGCTCTTGCGCCGGAAGATACGGTGATTGCTGATTAAGGCGATCTGCTT	600		
Rev	TTAATGACTGCGCTCTTGCGCCGGAAGATACGGTGATTGCTGATTAAGGCGATCTGCTT	536		
	*****			
Fwd	TCTAAACCTTATATCGACATCACACTCAATCTGATGAAGACGTTTGGTGTGAAATTGAA	660		
Rev	TCTAAACCTTATATCGACATCACACTCAATCTGATGAAGACGTTTGGTGTGAAATTGAA	596		
	*****			
Fwd	AATCAGCACTATCAACAATTTGTGCTAAAAGGCGGGCAGTCTTATCAGTCTCCGGGTACT	720		
Rev	AATCAGCACTATCAACAATTTGTGCTAAAAGGCGGGCAGTCTTATCAGTCTCCGGGTACT	656		
	*****			
Fwd	TATTTGGTGAAGGCGATGCATCTTCGGCTTCTTACTTTCTGGCAGCAGCAGCAATCAAA	780		
Rev	TATTTGGTGAAGGCGATGCATCTTCGGCTTCTTACTTTCTGGCAGCAGCAGCAATCAAA	716		
	*****			
Fwd	GGCGGCACTGTAA---AGTGACCGGTATTGGACGTAACAGTATGCAGGGTGATATTGCTTT	839		
Rev	GGCGGCACTGTAAAGTGACCGGTATTGGACGTAACAGTATGCAGGGTGATATTGCTTT	776		
	*****			
Fwd	GCTGATGTCTGAAAAAT---GGCGCGACCATTTGCTGGGCGGATGATTATATTTCTGCG	897		
Rev	GCTGATGTCTGAAAAATGGCGCGACCATTTGCTGGGCGGATGATTATATTTCTGCG	836		
	*****			
Fwd	ACGCGTGTGAACG---ACGCTATTGATATGGATATGAACCATAT---CCTGATGCGGCGATG	954		
Rev	ACGCGTGTGAACGTAACGCTATTGATATGGATATGAACCATATCCTGATGCGGCGATG	896		
	***** * *****			
Fwd	ACATTGGCYMCGCGCGGTWTTTGCAAGGAYCAMMMGC-----TGCCAWATCTAT---AC-	1007		
Rev	ACCATTGCCACGCGCGGTATTTGCAAAAGGCAACACGCTGCGCAATATCTATAAC	956		
	** * * * * * * * * * * * * * * *		*** ** * *	
Fwd	TGGCGGTAAARAG---AACGAWSCYGYTKSCATGCAMC---AAAYKMSTA-----A---GTC	1057		
Rev	TGGCGGTAAAGAGACCGATCGCCTGTTGCGATGGCAACAGAACTGCGTAAAGTCGGC	1016		
	***** * * * * * * * * * * * * * * *			
Fwd	GCSCGAGTGAAAAGGRGW---CAATAACTTCATWC---CTCCC---TCGGARAACKAACTTT	1110		
Rev	GCGGAAGTGGAAGAGGGGACGATTACATTGATCACTCTCCGAAAAAAGTGAACCTTT	1076		
	** * * * * * * * * * * * * * * *			

# MAKE “GENE-COMBINED” FILE (START TO STOP CODON IS FINE)



```
EPSP-combined

CATATGGGTAGCATGGAATCCCTGACGTTACATCCCATCGCTCGTGTGCGATGGCACTATTAATCTGCCCGGTTCCAAGAGCGTTTCTAACCGCGCTTTATT
GCTGGCGGCATTAGCACACGGCAAACAGTATTAACCAATCTGCTGGATAGCGATGACGTGCGCCATATGCTGAATGCATTAACAGCGTTAGGGGTAAGCT
ATACGCTTTTCAGCCGATCGTACGCGTTGCGAAATTATCGGTAACGGCGGTCCATTACACGCAGAAGGTGCCCTGGAGTTGTTCTCGGTAACGCCGGAACG
GCAATGCGTCCGCTGGCGGCAGCTCTTTGTCTGGGTAGCAATGATATTGTGCTGACCGGTGAGCCGCGTATGAAAGAACGCCCGATTGGTCATCTGGTGGA
TGCGCTGCGCCTGGGCGGGGCGAAGATCACTTACCTGGAACAAGAAAATTATCCGCCGTTGCGTTTACAGGGCGGCTTTACTGGCGGCAACGTTGACGTTG
ATGGCTCCGTTTCCAGCCAATTCCTCACCGCACTGTTAATGACTGCGCCTCTTGCGCCGGAAGATACGGTGATTGATTAAAGGCGATCTGGTTTCTAAA
CCTTATATCGACATCACACTCAATCTGATGAAGACGTTTGGTGTGAAATTGAAAATCAGCACTATCAACAATTTGTCGTAAAAGGCGGGCAGTCTTATC
AGTCTCCGGGTACTTATT
TGGTCGAAGGCGATGCATCTTCGGCTTCTTACTTTCTGGCAGCAGCAGCAATCAAAGGCG
GCACTGTAAAAGTGACCGGTATTGGACGTAAACAGTATGCAGGGTGATATTCGCTTTGCTG
ATGTGCTGGAAAAAATGGGCGCGACCATTTGCTGGGGCGATGATTATATTTCTGCACGC
GTGGTGAAGTGAACGCTATTGATATGGATATGAACCATATTCCTGATGCGGCGATGACCA
TTGCCACGGCGGCGTTATTTGCAAAGGCACCACCACGCTGCGCAATATCTATAACTGGC
GTGTTAAAGAGACCGATCGCCTGTTTGCGATGGCAACAGAACTGCGTAAAGTCGGCGCGG
AAGTGGAAGAGGGGCACGATTACATTCGTATCACTCCTCCGGAAAACTGAACTTTGCCG
AGATCGCGACATACAATGATCACCGGATGGCGATGTGTTTCTCGCTGGTGGCGTTGTCAG
ATACACCAGTGACGATTCTTGATCCCAAATGCACGGCCAAACATTTCCGGATTATTTG
AGCAGCTGGCGCGGATTAGCCAGGCAGCCTGACTCGAGCACCACCACCACCACCTGAG
ATCCGGCTGCTAA
```



# TRANSLATE (EXPASY TRANSLATE), AND SAVE TRANSLATED PROTEIN FILE

● ● ●

EPSP\_seq\_translated

MGSMESLTLHPIARVDGTINLPGSKSVSNRALLAALAHGKTVLTNLLDSDDVRHMLNALTALGVSYTLSADRTRCEIIGNGGPLHAEGALELFLGNAGTAMRPLAAALCLGSNDIVLTGEPKMKERPIGHLVDALRLGGAKITYLEQENYPPLRLQGGFTGGNVDVDGVSQFLTALLMTAPLAPEDTVIRIKGDLVSKPYIDITLNLTKTFGVEIENQHYQQFVVKGGQSYQSPGTYLVEGDASSASYFLAAAAIKGGTVKVTGIGRNSMQGDIRFADVLEKMGATICWGDDYISCTRGE LNAIDMDMNHIPDAAMTIATAALFAKGTTLRNINWVRVKETDRLFAMATELRKVGAEVEEGHDYIRITPPEKLNFAEIATYNDHRMAMCFSLVALSDTPVTILDPKCTAKTFPDYFEQLARISQAA

Download

GenPept

Graphics

3-phosphoshikimate 1-carboxyvinyltransferase [Escherichia coli]

Sequence ID: [EEZ8503789.1](#) Length: 427 Number of Matches: 1

Range 1: 1 to 427 [GenPept](#) [Graphics](#)

Next Match Previous Match

Score	Expect	Method	Identities	Positives	Gaps
881 bits(2277)	0.0	Compositional matrix adjust.	427/427(100%)	427/427(100%)	0/427(0%)
Query 4	MESLTLHPIARVDGTINLPGSKSVSNRALLAALAHGKTVLTNLLDSDDVRHMLNALTAL				63
Sbjct 1	MESLTLHPIARVDGTINLPGSKSVSNRALLAALAHGKTVLTNLLDSDDVRHMLNALTAL				60
Query 64	GVSYTLSADRTRCEIIGNGGPLHAEGALELFLGNAGTAMRPLAAALCLGSNDIVLTGEPK				123
Sbjct 61	GVSYTLSADRTRCEIIGNGGPLHAEGALELFLGNAGTAMRPLAAALCLGSNDIVLTGEPK				120
Query 124	MKERPIGHLVDALRLGGAKITYLEQENYPPLRLQGGFTGGNVDVDGVSQFLTALLMTA				183
Sbjct 121	MKERPIGHLVDALRLGGAKITYLEQENYPPLRLQGGFTGGNVDVDGVSQFLTALLMTA				180
Query 184	PLAPEDTVIRIKGDLVSKPYIDITLNLTKTFGVEIENQHYQQFVVKGGQSYQSPGTYLVE				243
Sbjct 181	PLAPEDTVIRIKGDLVSKPYIDITLNLTKTFGVEIENQHYQQFVVKGGQSYQSPGTYLVE				240
Query 244	GDASSASYFLAAAAIKGGTVKVTGIGRNSMQGDIRFADVLEKMGATICWGDDYISCTRGE				303
Sbjct 241	GDASSASYFLAAAAIKGGTVKVTGIGRNSMQGDIRFADVLEKMGATICWGDDYISCTRGE				300
Query 304	LNAIDMDMNHIPDAAMTIATAALFAKGTTLRNINWVRVKETDRLFAMATELRKVGAEVE				363
Sbjct 301	LNAIDMDMNHIPDAAMTIATAALFAKGTTLRNINWVRVKETDRLFAMATELRKVGAEVE				360
Query 364	EGHDYIRITPPEKLNFAEIATYNDHRMAMCFSLVALSDTPVTILDPKCTAKTFPDYFEQL				423
Sbjct 361	EGHDYIRITPPEKLNFAEIATYNDHRMAMCFSLVALSDTPVTILDPKCTAKTFPDYFEQL				420
Query 424	ARISQAA	430			
Sbjct 421	ARISQAA	427			

ALIGN WITH EXPECTED (E.G. CLUSTAL O), OR  
SIMPLY BLAST