UFABC
Bacharelado em Biotecnologia

Construção do Vetor Recombinante

Proteínas recombinantes
Sergio Daishi Sasaki
(NHZ6006-18)

Universidade Federal do ABC

Etapas construção primers

- 1 – Verificar a sequência de DNA do gene de interesse se não há sítios de restrição para as enzimas que serão utilizadas. Por análise de restrição.
- 2 – Construção do primer sense – FW. Cópia da sequência 5´ da fita codificadora acrescida do sítio de restrição e outros nucleotídeos complementares, se necessário.
- 3 – Construção do primer reverso. Cópia de sequência 5´da fica reversa complementar da fica codificadora, acrescida do sítio de restrição e outros nucleotídeos complementares, se necessário.
- 4 – Verificar se o fragmento a ser inserido no vetor estará "in frame" com o sistema de expressão para produzir a proteína corretamente.
- 5 – Checar os primers – Tm, formação de hairpin, complementaridade entre primers.

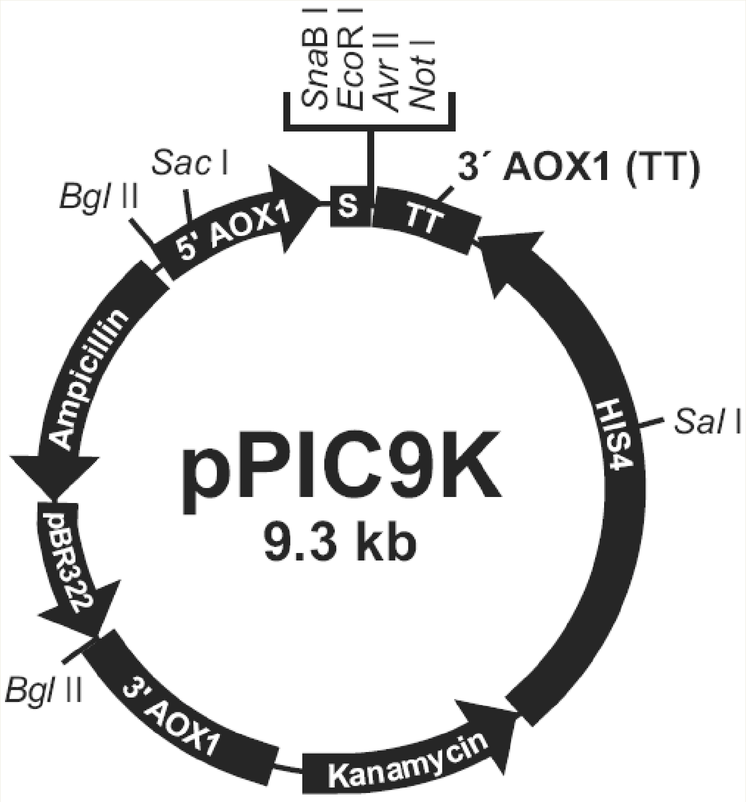
Exemplo: Clonar a proteína Boophilina no vetor pPIC9K

Boophilus microplus mRNA for boophilin, isoform H2
GenBank: AJ304447.1

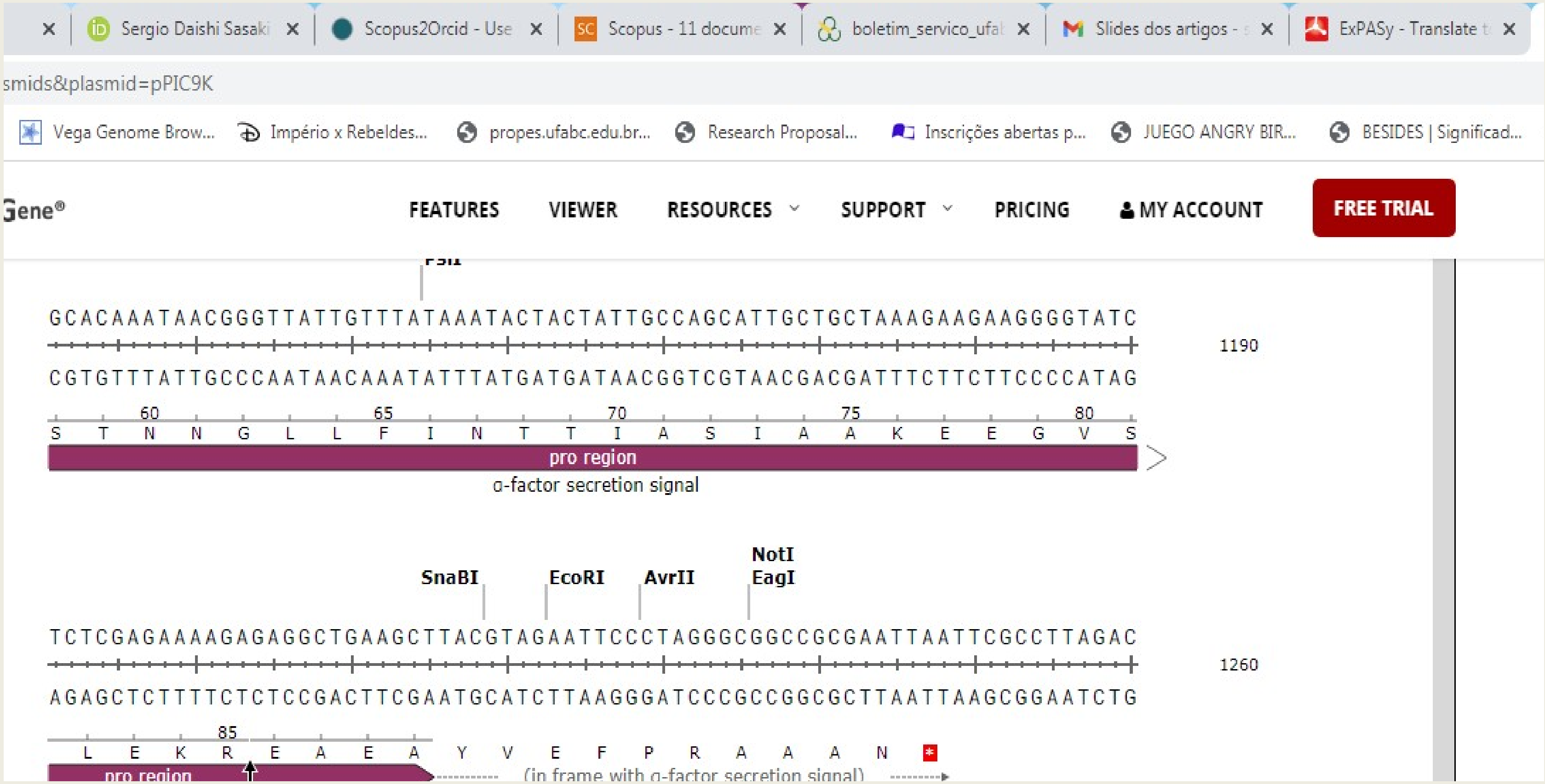
GenBank Graphics

```
>AJ304447.1 Boophilus microplus mRNA for boophilin, isoform H2
ATTGGCCAGGCACTCAAGATGAAGTGCATCATTCTTCTAGCCGTTCTAGGAACAGCGTTTGCGCAGAGAA
ATGGATTCTGCCGACTGCCGGCAGATGAAGGCATCTGCAAGGCCCTCATACCTCGCTTCTACTTCAACAC
TGAAACCGGAAAATGTACCATGTTTTCCTATGGAGGCTGTGGAGGCAACGAGAACAACTTCGAGACCATA
GAGGAGTGTCAAAAGGCGTGCGGAGCTCCGGAGCGAGTGAACGACTTCGAAAGCGCCGATTTCAAGACTG
GCTGTGAGCCGGCGGCCGACAGCGGCTCTTGCGCTGGTCAACTGGAGCGCTGGTTTTACAATGTTCAATC
GGGAGAGTGCGAGACATTCGTCTACGGGGGCTGCGGGGGCAACGACAACAATTACGAAAGCGAGGAGGAG
TGTGAACTCGTCTGCAAGAACATGTAATTATCCATACCTAGCGGACTTCAAAGTGCCTTCCCTCATGGAT
CGCATCATGCGTCATAAAGATACCGAAGTGTACCTTGTGAATAAAATGAA
```

Vetor pPIC9K



Vetor pPIC9K

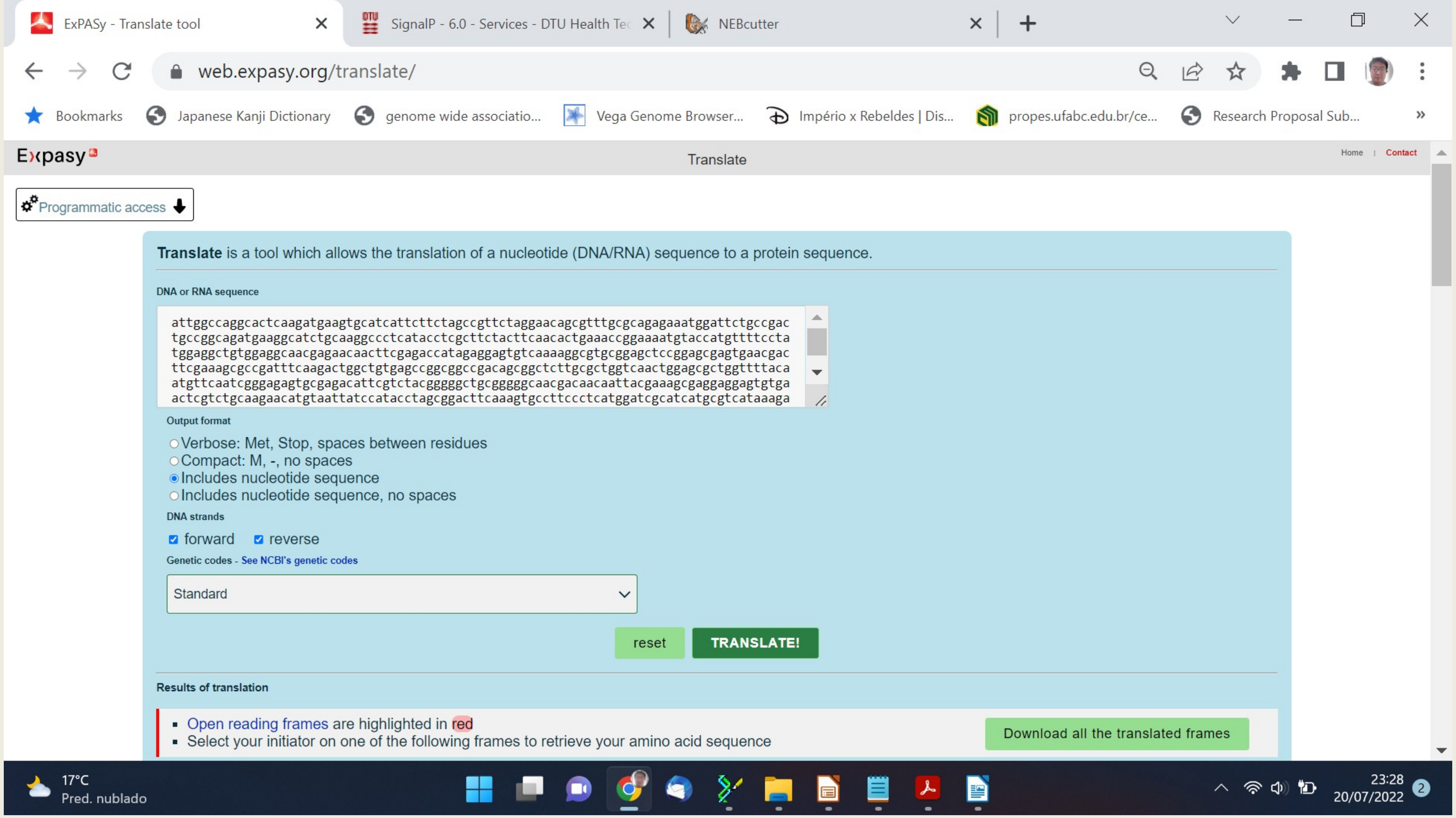


```
GCACAAATAACGGGTTATTGTTTATAAATACTACTATTGCCAGCATTGCTGCTAAAGAAGAAGGGGTATC
CGTGTTTATTGCCCAATAACAAATATTTATGATGATAACGGTCGTAACGACGATTTCTTCTTCCCCATAG
S  T  N  N  G  L  L  F  I  N  T  T  I  A  S  I  A  A  K  E  E  G  V  S

TCTCGAGAAAAGAGAGGCTGAAGCTTACGTAGAATTCCCTAGGGCGGCCGCGAATTAATTCGCCTTAGAC
AGAGCTCTTTTCTCTCCGACTTCGAATGCATCTTAAGGGATCCCGCCGGCGCTTAATTAAGCGGAATCTG
L  E  K  R  E  A  E  A  Y  V  E  F  P  R  A  A  A  N  *
```

Tradução da sequência com ferramenta EXPASY translate:

https://web.expasy.org/translate/



Translate is a tool which allows the translation of a nucleotide (DNA/RNA) sequence to a protein sequence.

DNA or RNA sequence

```
attggccaggcactcaagatgaagtgcatcattcttctagccgttctaggaacagcgtttgcgcagagaaatggattctgccgac
tgccggcagatgaaggcatctgcaaggccctcatacctcgcttctacttcaacactgaaaccggaaaatgtaccatgttttccta
tggaggctgtggaggcaacgagaacaacttcgagaccatagaggagtgtcaaaaggcgtgcggagctccggagcgagtgaacgac
ttcgaaagcgccgatttcaagactggctgtgagccggcggccgacagcggctcttgcgctggtcaactggagcgctggttttaca
atgttcaatcgggagagtgcgagacattcgtctacgggggctgcgggggcaacgacaacaattacgaaagcgaggaggagtgtga
actcgtctgcaagaacatgtaattatccatacctagcggacttcaaagtgccttccctcatggatcgcatcatgcgtcataaaga
```

Output format

○ Verbose: Met, Stop, spaces between residues
○ Compact: M, -, no spaces
● Includes nucleotide sequence
○ Includes nucleotide sequence, no spaces

DNA strands

☑ forward ☑ reverse

Genetic codes - See NCBI's genetic codes

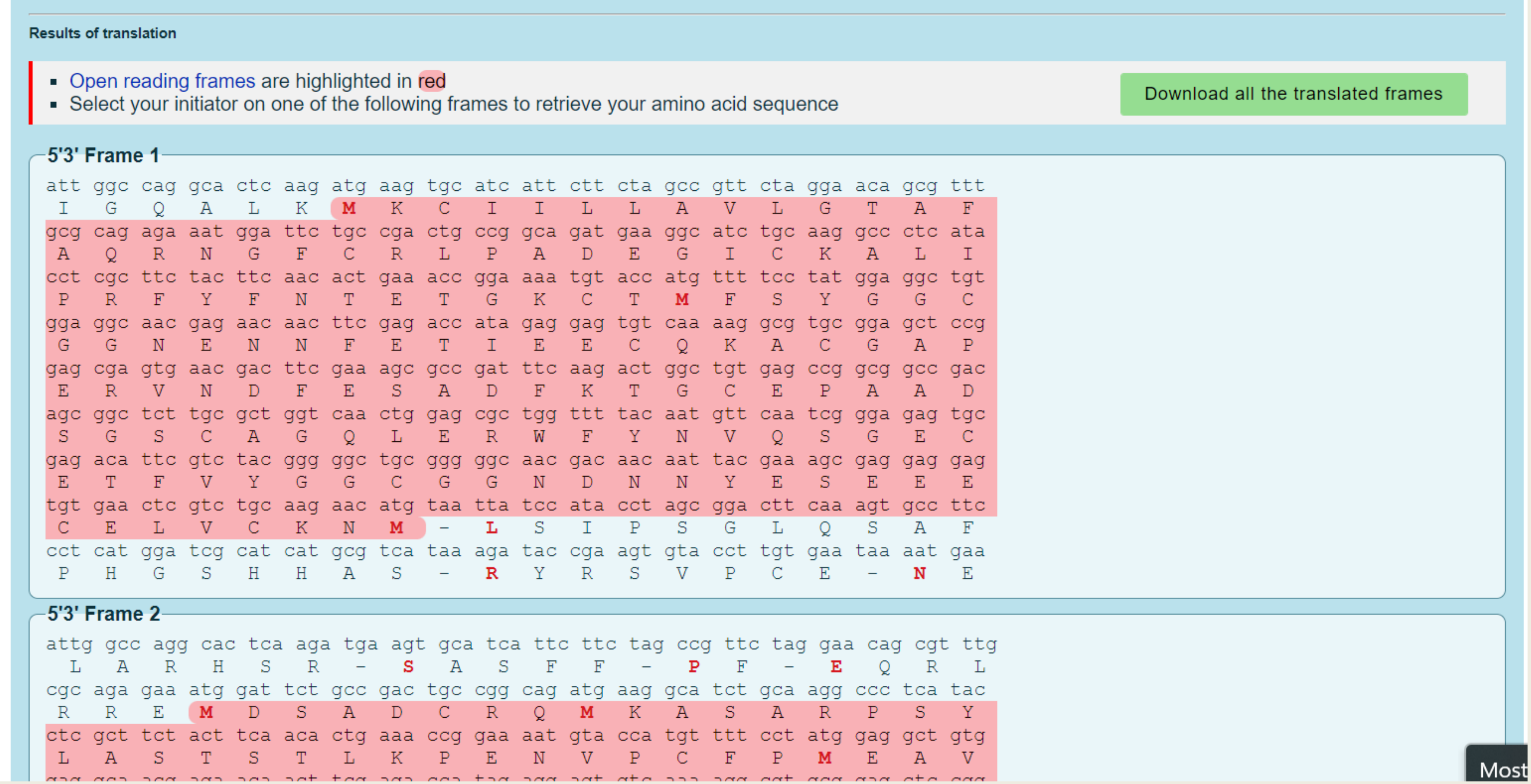
Standard

Results of translation

- Open reading frames are highlighted in red
- Select your initiator on one of the following frames to retrieve your amino acid sequence

Tradução da sequência com ferramenta EXPASY translate:

https://web.expasy.org/translate/



Results of translation

- Open reading frames are highlighted in red
- Select your initiator on one of the following frames to retrieve your amino acid sequence

5'3' Frame 1

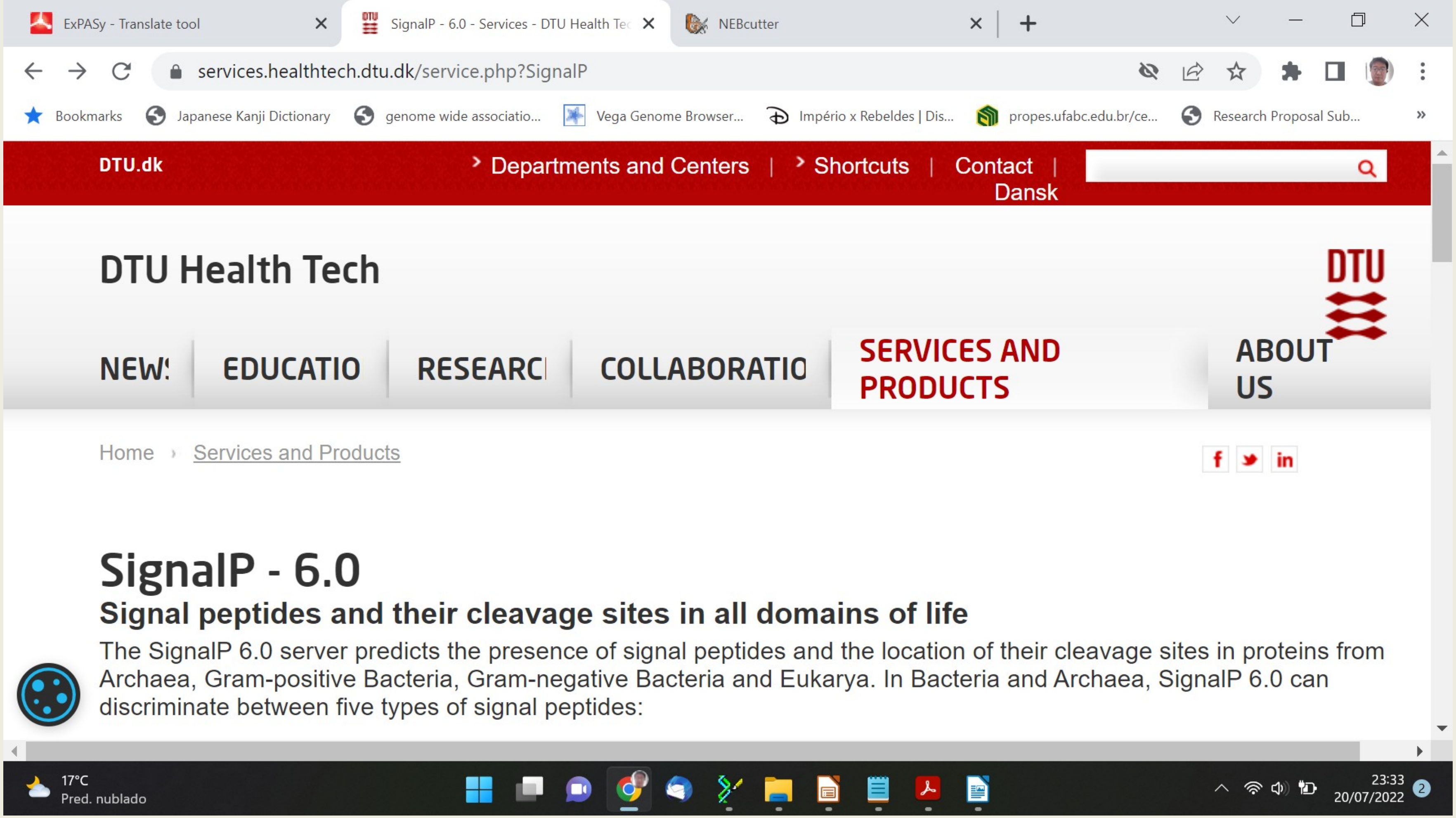
```
att ggc cag gca ctc aag atg aag tgc atc att ctt cta gcc gtt cta gga aca gcg ttt
 I   G   Q   A   L   K   M   K   C   I   I   L   L   A   V   L   G   T   A   F
gcg cag aga aat gga ttc tgc cga ctg ccg gca gat gaa ggc atc tgc aag gcc ctc ata
 A   Q   R   N   G   F   C   R   L   P   A   D   E   G   I   C   K   A   L   I
cct cgc ttc tac ttc aac act gaa acc gga aaa tgt acc atg ttt tcc tat gga ggc tgt
 P   R   F   Y   F   N   T   E   T   G   K   C   T   M   F   S   Y   G   G   C
gga ggc aac gag aac aac ttc gag acc ata gag gag tgt caa aag gcg tgc gga gct ccg
 G   G   N   E   N   N   F   E   T   I   E   E   C   Q   K   A   C   G   A   P
gag cga gtg aac gac ttc gaa agc gcc gat ttc aag act ggc tgt gag ccg gcg gcc gac
 E   R   V   N   D   F   E   S   A   D   F   K   T   G   C   E   P   A   A   D
agc ggc tct tgc gct ggt caa ctg gag cgc tgg ttt tac aat gtt caa tcg gga gag tgc
 S   G   S   C   A   G   Q   L   E   R   W   F   Y   N   V   Q   S   G   E   C
gag aca ttc gtc tac ggg ggc tgc ggg ggc aac gac aac aat tac gaa agc gag gag gag
 E   T   F   V   Y   G   G   C   G   G   N   D   N   N   Y   E   S   E   E   E
tgt gaa ctc gtc tgc aag aac atg taa tta tcc ata cct agc gga ctt caa agt gcc ttc
 C   E   L   V   C   K   N   M   -   L   S   I   P   S   G   L   Q   S   A   F
cct cat gga tcg cat cat gcg tca taa aga tac cga agt gta cct tgt gaa taa aat gaa
 P   H   G   S   H   H   A   S   -   R   Y   R   S   V   P   C   E   -   N   E
```

5'3' Frame 2

```
attg gcc agg cac tca aga tga agt gca tca ttc ttc tag ccg ttc tag gaa cag cgt ttg
  L   A   R   H   S   R   -   S   A   S   F   F   -   P   F   -   E   Q   R   L
cgc aga gaa atg gat tct gcc gac tgc cgg cag atg aag gca tct gca agg ccc tca tac
 R   R   E   M   D   S   A   D   C   R   Q   M   K   A   S   A   R   P   S   Y
ctc gct tct act tca aca ctg aaa ccg gaa aat gta cca tgt ttt cct atg gag gct gtg
 L   A   S   T   S   T   L   K   P   E   N   V   P   C   F   P   M   E   A   V
```

Análise se a proteína tem peptídeo sinal por meio do uso do Signal P.

https://services.healthtech.dtu.dk/service.php?SignalP



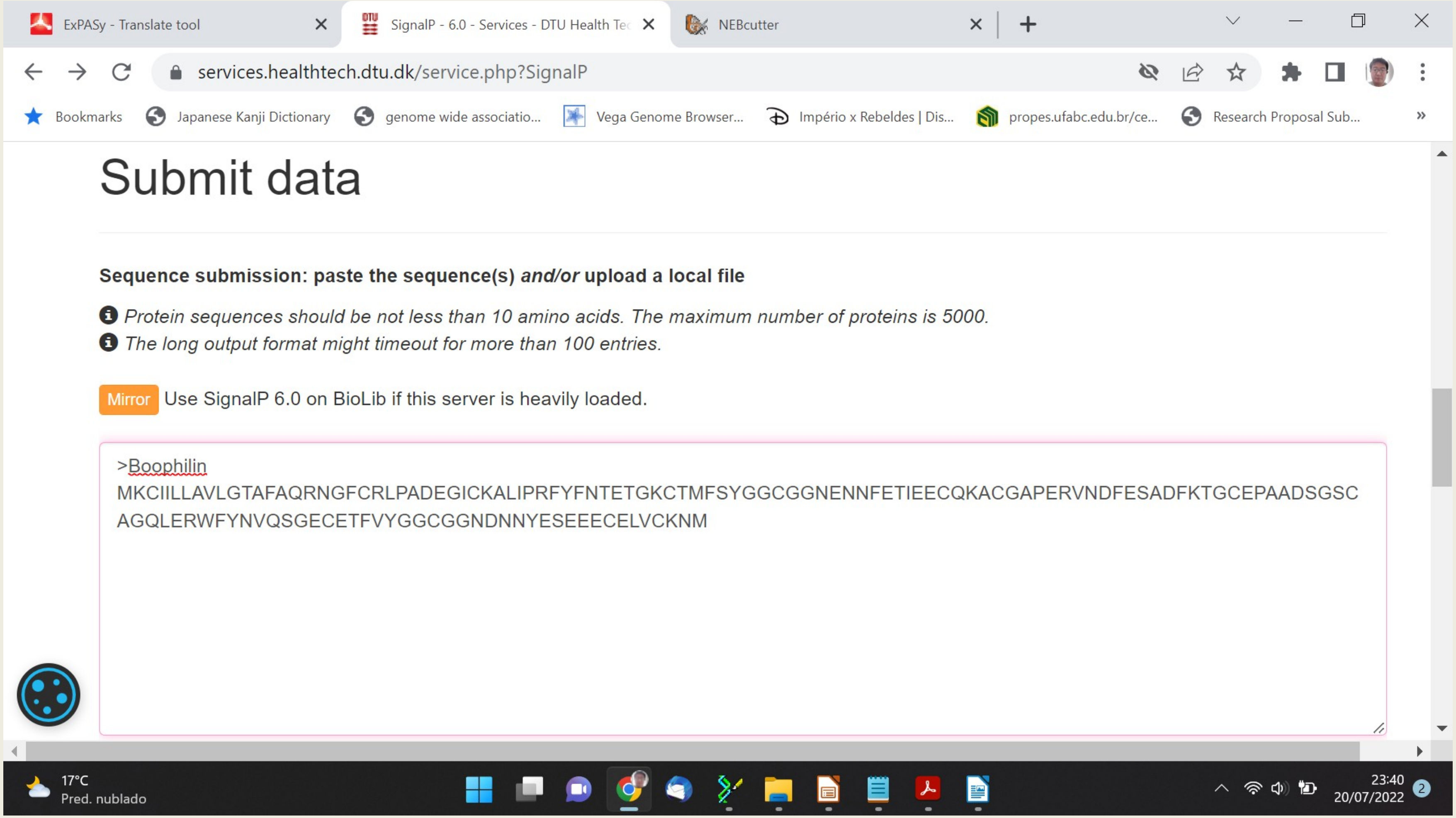
SignalP - 6.0

Signal peptides and their cleavage sites in all domains of life

The SignalP 6.0 server predicts the presence of signal peptides and the location of their cleavage sites in proteins from Archaea, Gram-positive Bacteria, Gram-negative Bacteria and Eukarya. In Bacteria and Archaea, SignalP 6.0 can discriminate between five types of signal peptides:

Análise se a proteína tem peptídeo sinal por meio do uso do Signal P.

https://services.healthtech.dtu.dk/service.php?SignalP



Submit data

Sequence submission: paste the sequence(s) *and/or* upload a local file

Protein sequences should be not less than 10 amino acids. The maximum number of proteins is 5000.

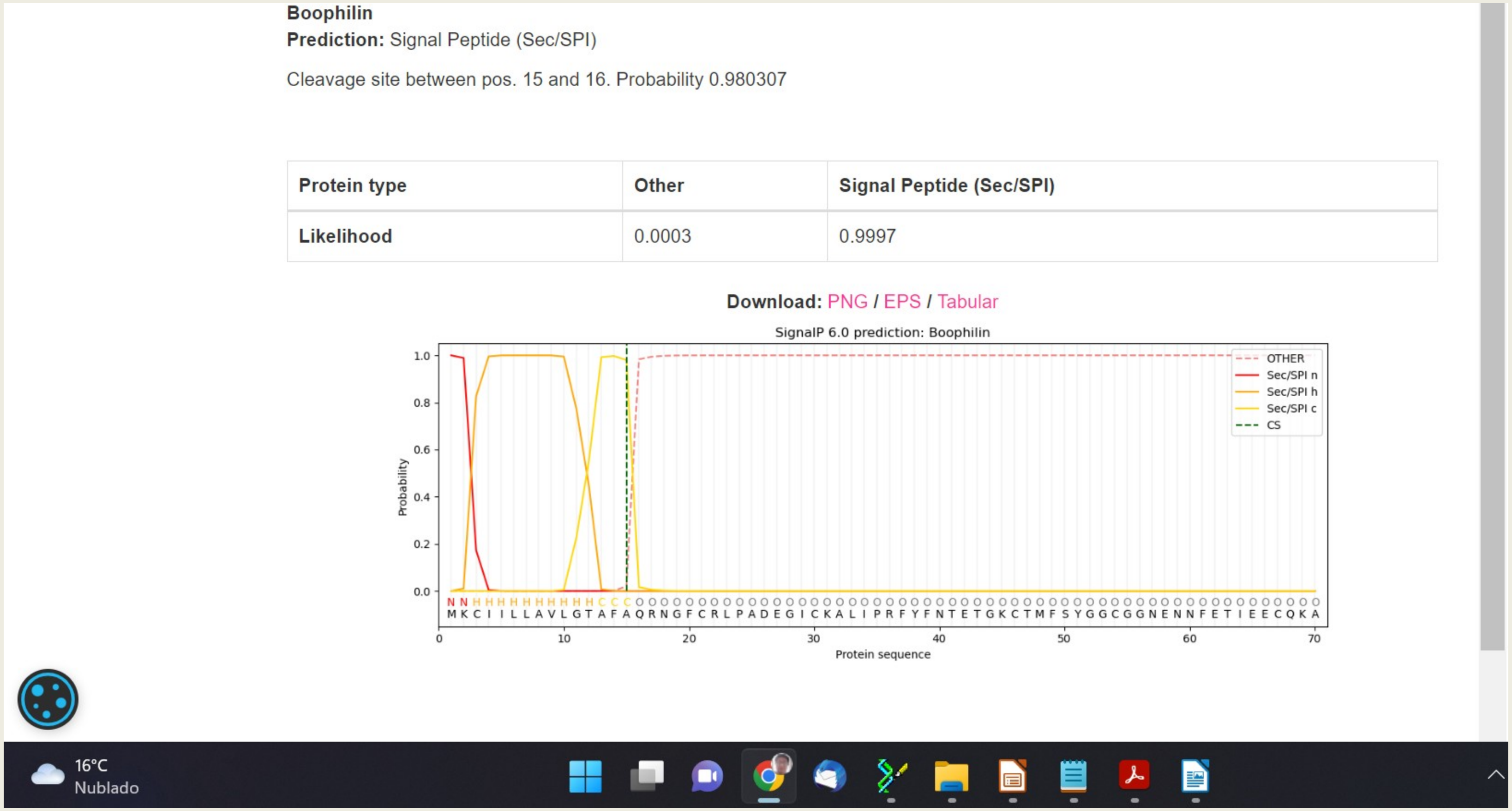
The long output format might timeout for more than 100 entries.

Mirror Use SignalP 6.0 on BioLib if this server is heavily loaded.

```
>Boophilin
MKCIILLAVLGTAFAQRNGFCRLPADEGICKALIPRFYFNTETGKCTMFSYGGCGGNENNFETIEECQKACGAPERVNDFESADFKTGCEPAADSGSC
AGQLERWFYNVQSGECETFVYGGCGGNDNNYESEEECELVCKNM
```

Análise se a proteína tem peptídeo sinal por meio do uso do Signal P.

https://services.healthtech.dtu.dk/service.php?SignalP



Boophilin
Prediction: Signal Peptide (Sec/SPI)

Cleavage site between pos. 15 and 16. Probability 0.980307

Protein type	Other	Signal Peptide (Sec/SPI)
Likelihood	0.0003	0.9997

Download: PNG / EPS / Tabular

Seleciona-se a região de interesse a ser clonada:

```
cag aga aat gga ttc tgc cga ctg ccg gca gat gaa ggc atc tgc aag gcc ctc ata
 Q   R   N   G   F   C   R   L   P   A   D   E   G   I   C   K   A   L   I
cct cgc ttc tac ttc aac act gaa acc gga aaa tgt acc atg ttt tcc tat gga ggc tgt
 P   R   F   Y   F   N   T   E   T   G   K   C   T   M   F   S   Y   G   G   C
gga ggc aac gag aac aac ttc gag acc ata gag gag tgt caa aag gcg tgc gga gct ccg
 G   G   N   E   N   N   F   E   T   I   E   E   C   Q   K   A   C   G   A   P
gag cga gtg aac gac ttc gaa agc gcc gat ttc aag act ggc tgt gag ccg gcg gcc gac
 E   R   V   N   D   F   E   S   A   D   F   K   T   G   C   E   P   A   A   D
agc ggc tct tgc gct ggt caa ctg gag cgc tgg ttt tac aat gtt caa tcg gga gag tgc
 S   G   S   C   A   G   Q   L   E   R   W   F   Y   N   V   Q   S   G   E   C
gag aca ttc gtc tac ggg ggc tgc ggg ggc aac gac aac aat tac gaa agc gag gag gag
 E   T   F   V   Y   G   G   C   G   G   N   D   N   N   Y   E   S   E   E   E
tgt gaa ctc gtc tgc aag aac atg taa
 C   E   L   V   C   K   N   M
```

Apenas os nucleotídeos

cag aga aat gga ttc tgc cga ctg ccg gca gat gaa ggc atc tgc aag gcc ctc ata

cct cgc ttc tac ttc aac act gaa acc gga aaa tgt acc atg ttt tcc tat gga ggc tgt

gga ggc aac gag aac aac ttc gag acc ata gag gag tgt caa aag gcg tgc gga gct ccg

gag cga gtg aac gac ttc gaa agc gcc gat ttc aag act ggc tgt gag ccg gcg gcc gac

agc ggc tct tgc gct ggt caa ctg gag cgc tgg ttt tac aat gtt caa tcg gga gag tgc

gag aca ttc gtc tac ggg ggc tgc ggg ggc aac gac aac aat tac gaa agc gag gag gag

tgt gaa ctc gtc tgc aag aac atg taa

Estabelecer a região dos primers:

cag aga aat gga ttc tgc cga ctg ccg gca gat gaa ggc atc tgc aag gcc ctc ata

cct cgc ttc tac ttc aac act gaa acc gga aaa tgt acc atg ttt tcc tat gga ggc tgt

gga ggc aac gag aac aac ttc gag acc ata gag gag tgt caa aag gcg tgc gga gct ccg

gag cga gtg aac gac ttc gaa agc gcc gat ttc aag act ggc tgt gag ccg gcg gcc gac

agc ggc tct tgc gct ggt caa ctg gag cgc tgg ttt tac aat gtt caa tcg gga gag tgc

gag aca ttc gtc tac ggg ggc tgc ggg ggc aac gac aac aat tac gaa agc gag gag gag

tgt gaa ctc gtc tgc aag aac atg taa

Analisar as regiões selecionadas

https://www.genscript.com/sms2/rev_comp.html (para obter sequência reversa complementar)

https://www.thermofisher.com/br/en/home/brands/thermo-scientific/molecular-biology/molecular-biology-learning-center/molecular-biology-resource-library/thermo-scientific-web-tools/tm-calculator.html

https://m4.igenetech.com/hairpin/a261de23-1897-47e0-b83f-e38c2e562c6d

Forward:
5´cag aga aat gga ttc tgc c 3´

Tm: 53oC

Pré-reverso – fase 1:
5´c gtc tgc aag aac atg taa3´

>reverse primer reverse complement
5´ttacatgttcttgcagacg 3´

Tm: 52,2oC

Sítios de restrição escolhidos

Enzyme	Sequence
SnaBI	TAC↑↓GTA

Enzyme	Sequence	Overhang	Properties
NotI	GC↓GGCC↑GC	5′ GGCC	

Adicionar os sítios de restrição

Forward:

5´agcc TACGTA cag aga aat gga ttc tgc c 3´

Tm:65,4oC

Reverso:

5´ ttatt CGCGGCCGC tta cat gtt ctt gca gac g 3´

Tm:69,8oC

Analisar se formam hairpin ou tem autoclomplementaridade

https://www.thermofisher.com/br/en/home/brands/thermo-scientific/molecular-biology/molecular-biology-learning-center/molecular-biology-resource-library/thermo-scientific-web-tools/multiple-primer-analyzer.html

Forward:
5´agcc TACGTA cag aga aat gga ttc tgc c 3´

Tm:65,4oC

Reverso:

5´ ttatt CGCGGCCGC tta cat gtt ctt gca gac g 3´

Tm:69,8oC

Outros programas
DNAMAN
BIOEDIT

https://eurofinsgenomics.eu/en/ecom/tools/oligo-analysis/

https://www.sigmaaldrich.com/BR/pt/technical-documents/technical-article/genomics/pcr/oligo-evaluator-for-tm-calculation-primer-analysis?srsltid=AfmBOopVrLXMbjU_vN5RGOeaViN4XrXm_7IQnaIJWG5Kkv7BgHs4qzNd