Освоение компьютерных программ анализа нуклеотидных и аминокислотных последовательностей: SnapGene (GeneRunner; Vector NTI 11; CLC Sequence Viewer 8)

Преподаватель: старший преподаватель
кафедры молекулярной биологии и генетики,
PhD, Смекенов И.Т.
Дисциплина: Рекомбинация ДНК

(*Лекция 9*)

Цель

Изучить компьютерные программы для анализа нуклеотидных и аминокислотных последовательностей с целью повышения эффективности молекулярно-биологических исследований.

Задачи

1.Ознакомиться с функциональными возможностями программы SnapGene для визуализации и аннотирования ДНК-клонирования.

2.Рассмотреть программу GeneRunner и её применение для анализа нуклеотидных последовательностей.

3.Изучить возможности Vector NTI 11 для работы с последовательностями и проектирования векторов.

4.Оценить функционал CLC Sequence Viewer 8 для многократного выравнивания последовательностей и анализа данных.

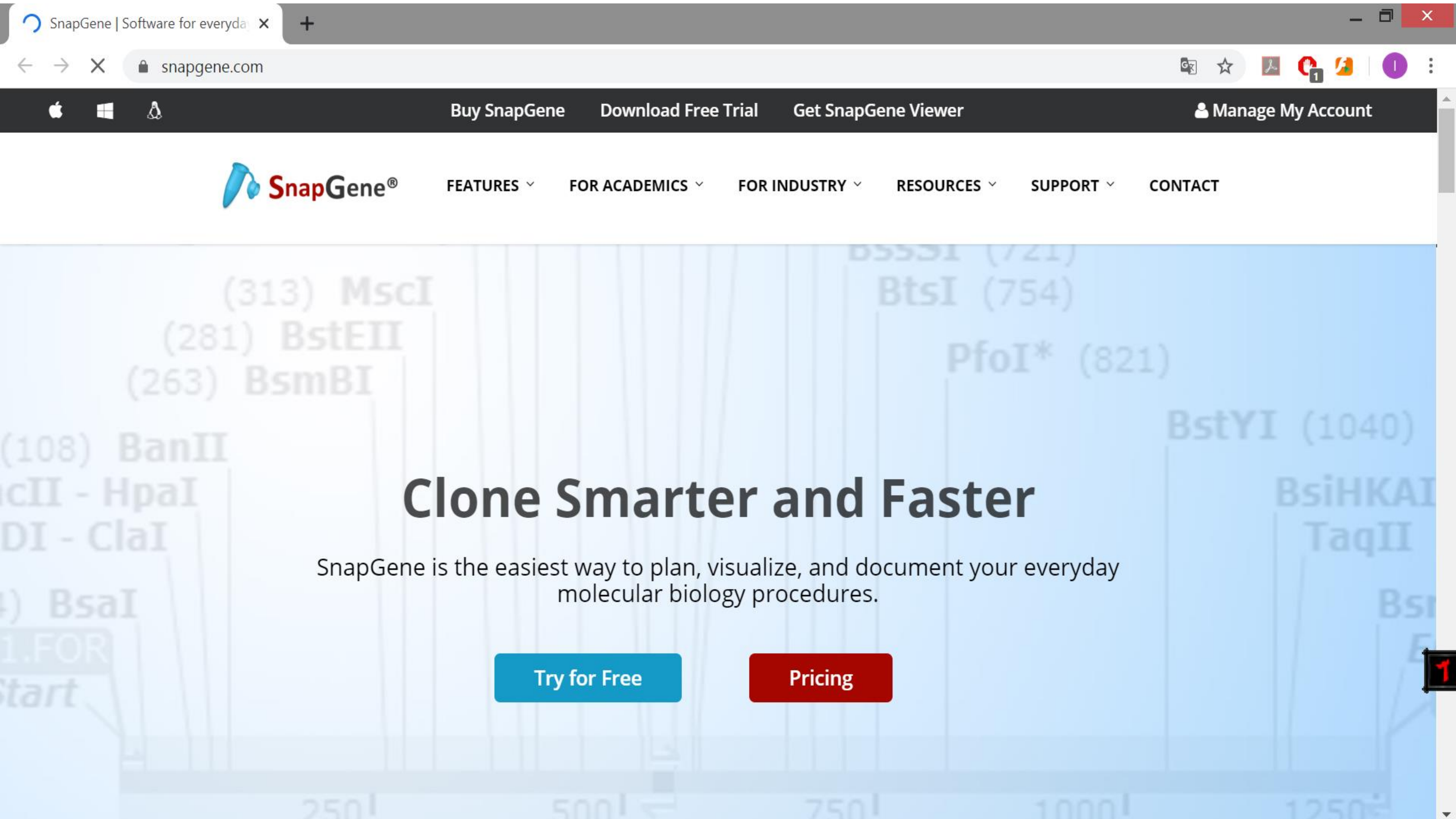
Ключевые слова: *анализ нуклеотидных последовательностей, SnapGene, GeneRunner, Vector NTI 11, CLC Sequence Viewer 8, аминокислотные последовательности, визуализация последовательностей*

SnapGENE – клонируй умнее и быстрее

- это самый простой способ планирования, визуализирования и документирования ваших повседневных процедур молекулярной биологии.

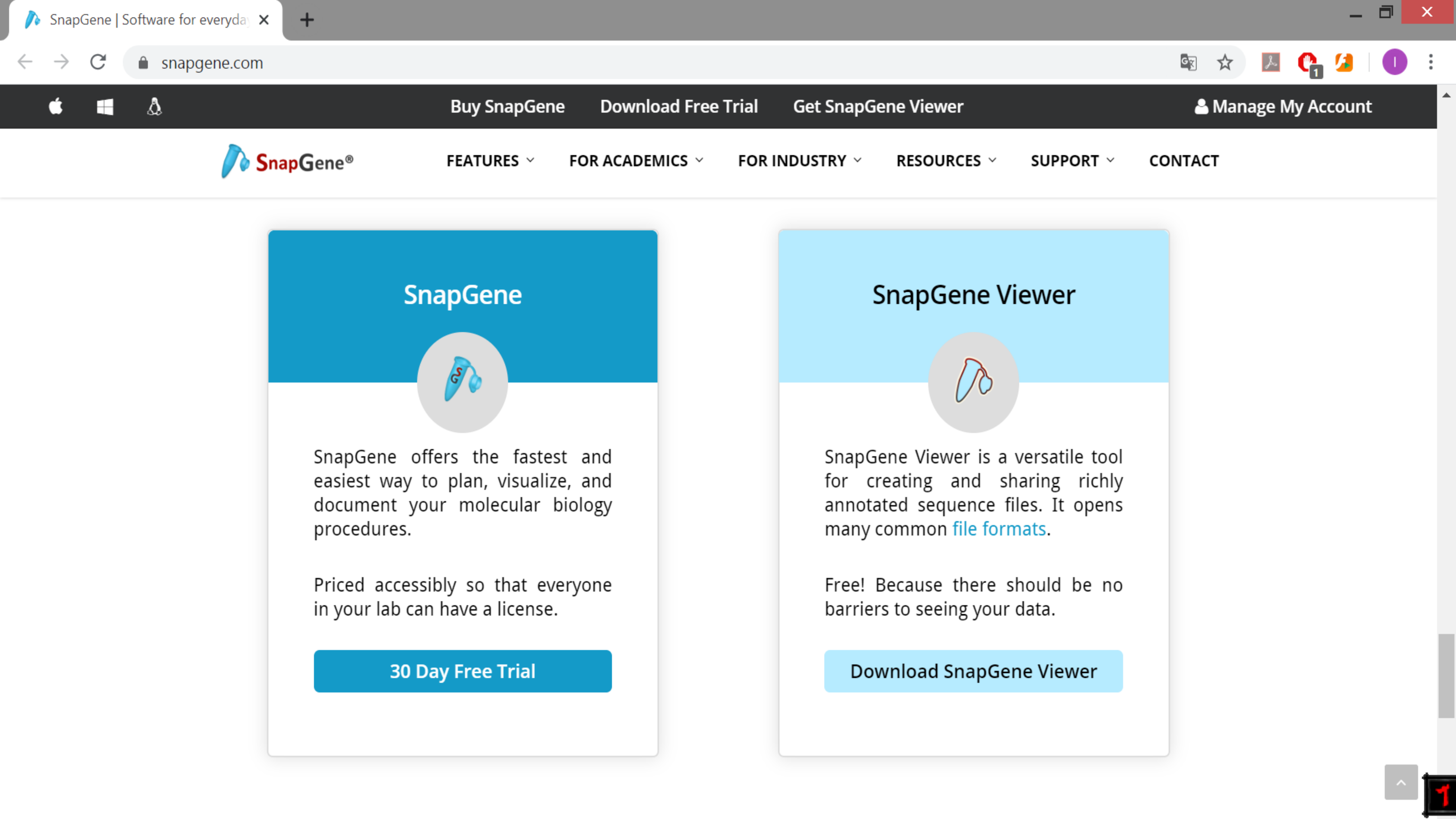
SnapGene®

Software for molecular biology



Clone Smarter and Faster

SnapGene is the easiest way to plan, visualize, and document your everyday molecular biology procedures.



SnapGene

SnapGene offers the fastest and easiest way to plan, visualize, and document your molecular biology procedures.

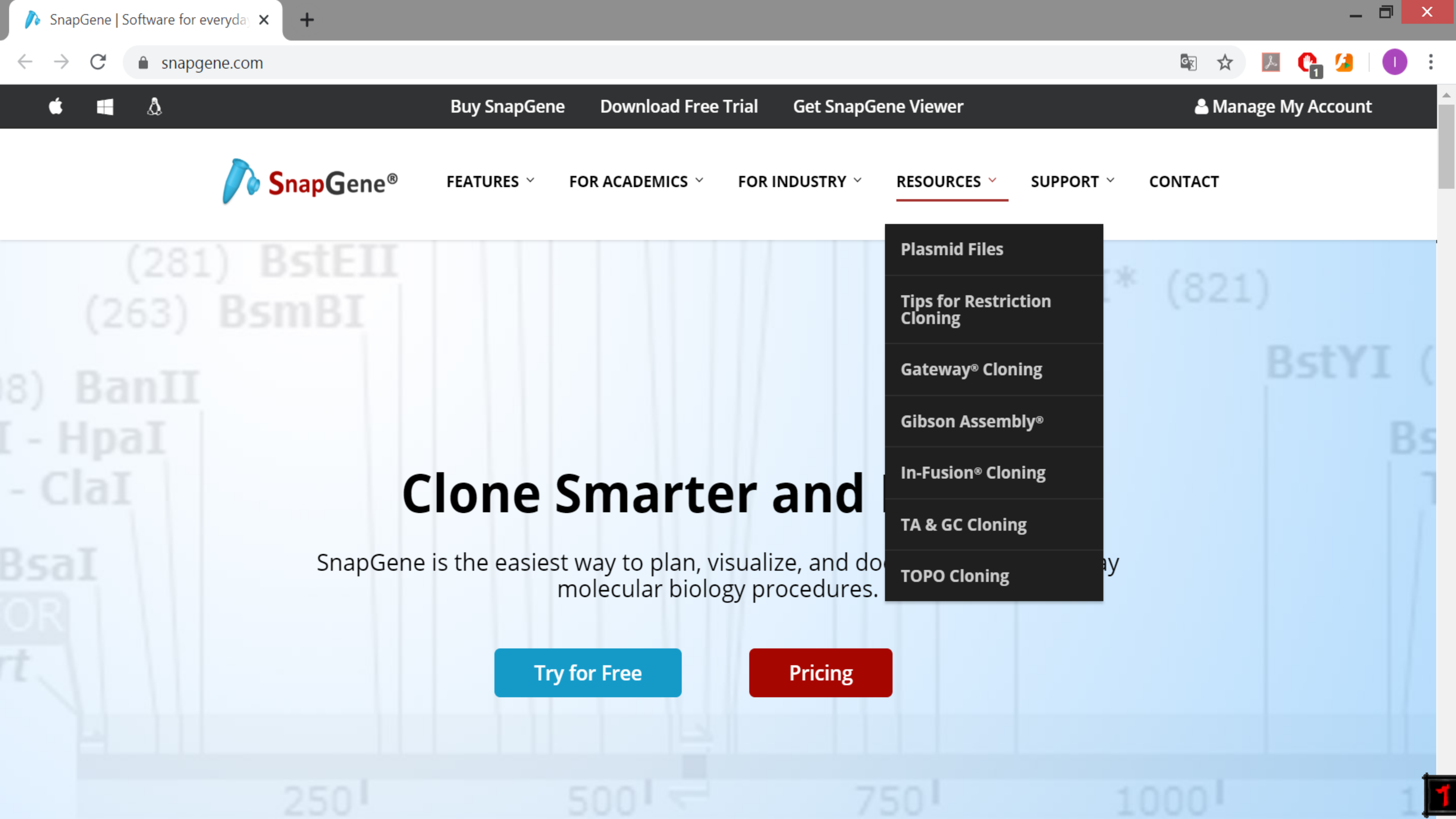
Priced accessibly so that everyone in your lab can have a license.

SnapGene Viewer

SnapGene Viewer is a versatile tool for creating and sharing richly annotated sequence files. It opens many common file formats.

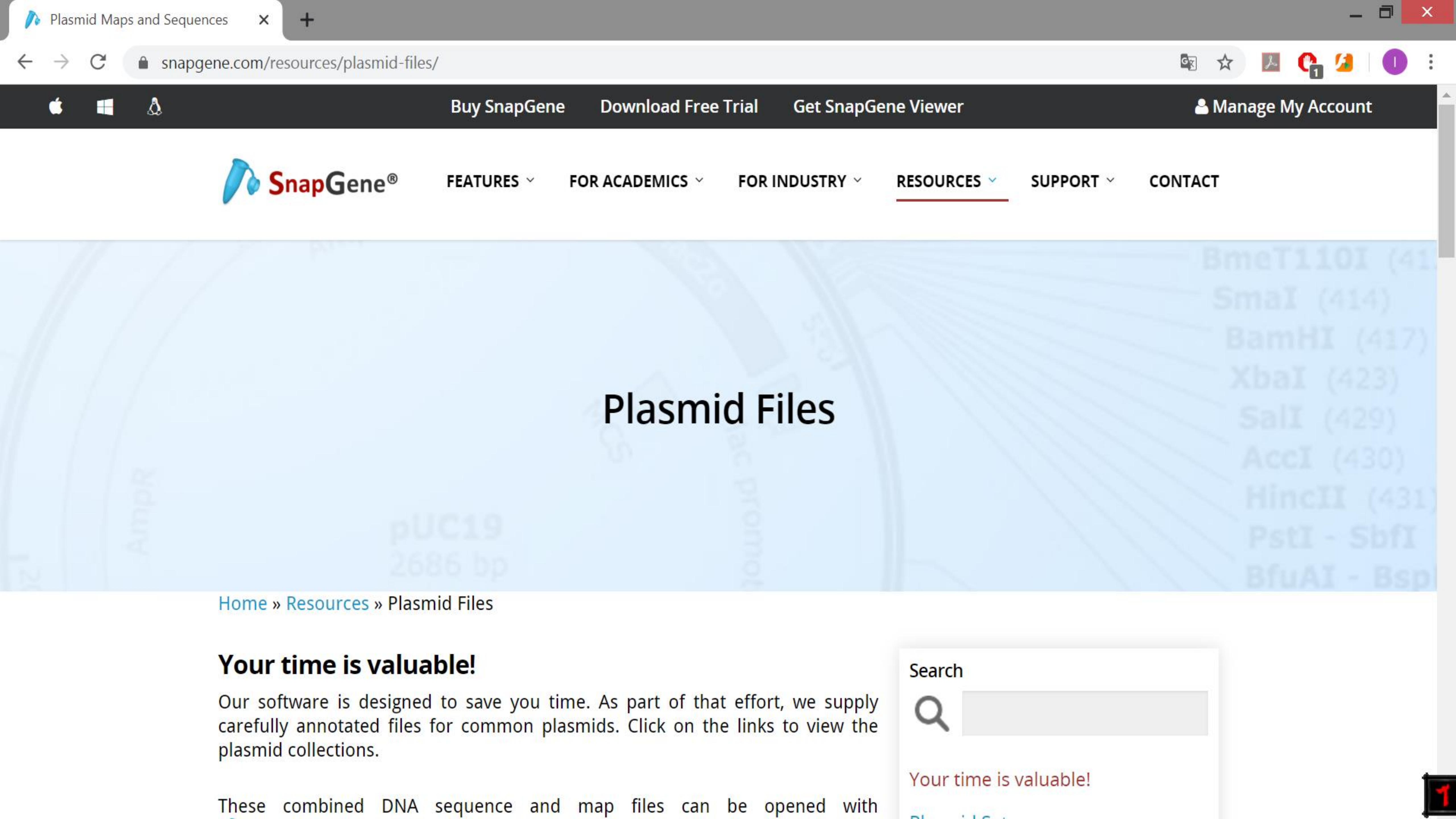
Free! Because there should be no barriers to seeing your data.


Этот компьютер
Новая папка
Корзина
График оплаты 20...
ЭЦП Жанар пароль Аа...
Ахметова Жанар Нұр...
Темірғали Иманали И...
жанар перевод
4x_1172_pl...
SnapGene
SnapGene™ 1.1.3
New DNA File...
Open...
Open Recent
Import from GenBank...
Reference
Help
Exit



Clone Smarter and

SnapGene is the easiest way to plan, visualize, and do … ay
molecular biology procedures.



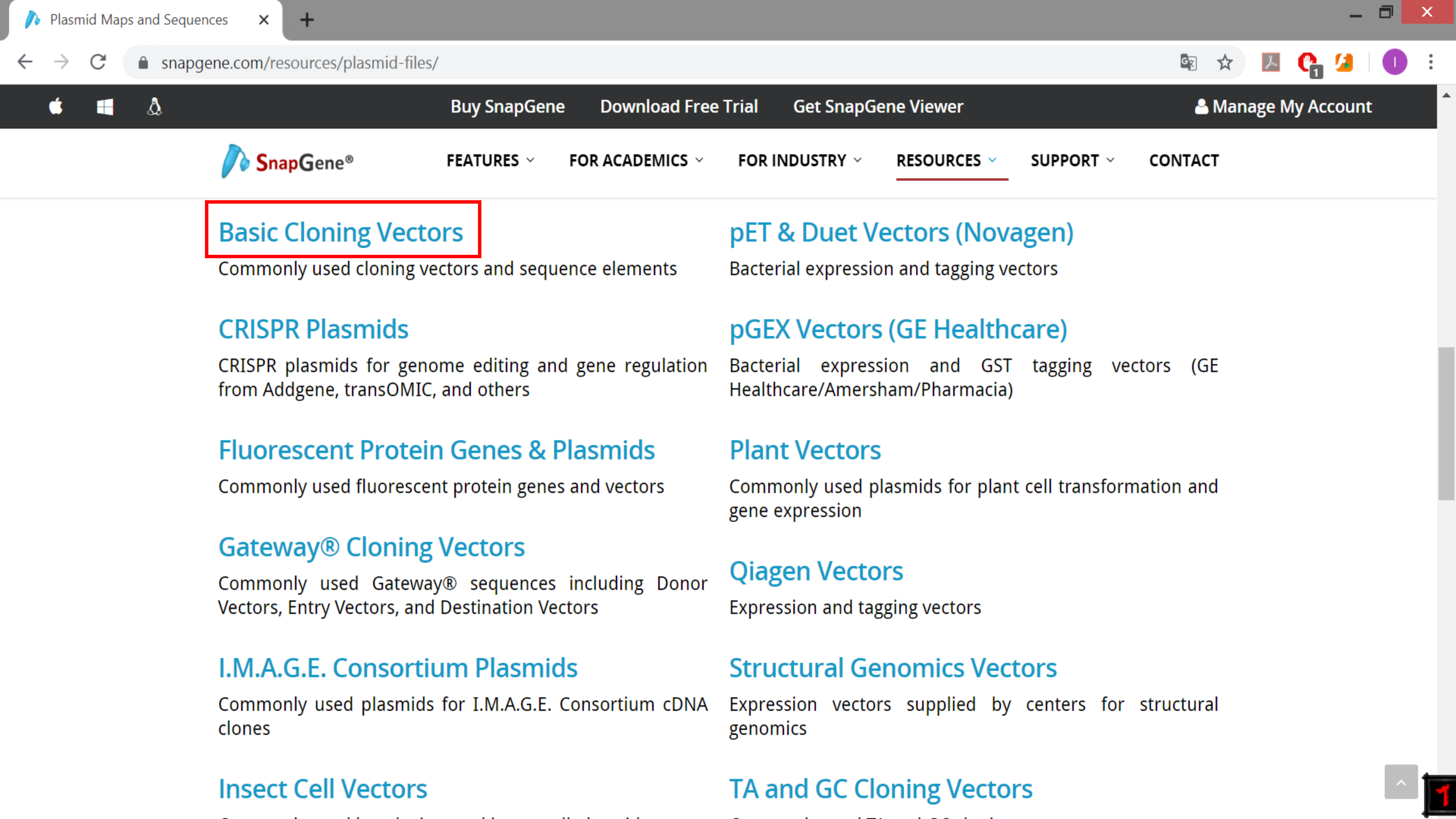
Plasmid Files

Home » Resources » Plasmid Files

Your time is valuable!

Our software is designed to save you time. As part of that effort, we supply carefully annotated files for common plasmids. Click on the links to view the plasmid collections.

These combined DNA sequence and map files can be opened with



Basic Cloning Vectors

Commonly used cloning vectors and sequence elements

CRISPR Plasmids

CRISPR plasmids for genome editing and gene regulation from Addgene, transOMIC, and others

Fluorescent Protein Genes & Plasmids

Commonly used fluorescent protein genes and vectors

Gateway® Cloning Vectors

Commonly used Gateway® sequences including Donor Vectors, Entry Vectors, and Destination Vectors

I.M.A.G.E. Consortium Plasmids

Commonly used plasmids for I.M.A.G.E. Consortium cDNA clones

Insect Cell Vectors

pET & Duet Vectors (Novagen)

Bacterial expression and tagging vectors

pGEX Vectors (GE Healthcare)

Bacterial expression and GST tagging vectors (GE Healthcare/Amersham/Pharmacia)

Plant Vectors

Commonly used plasmids for plant cell transformation and gene expression

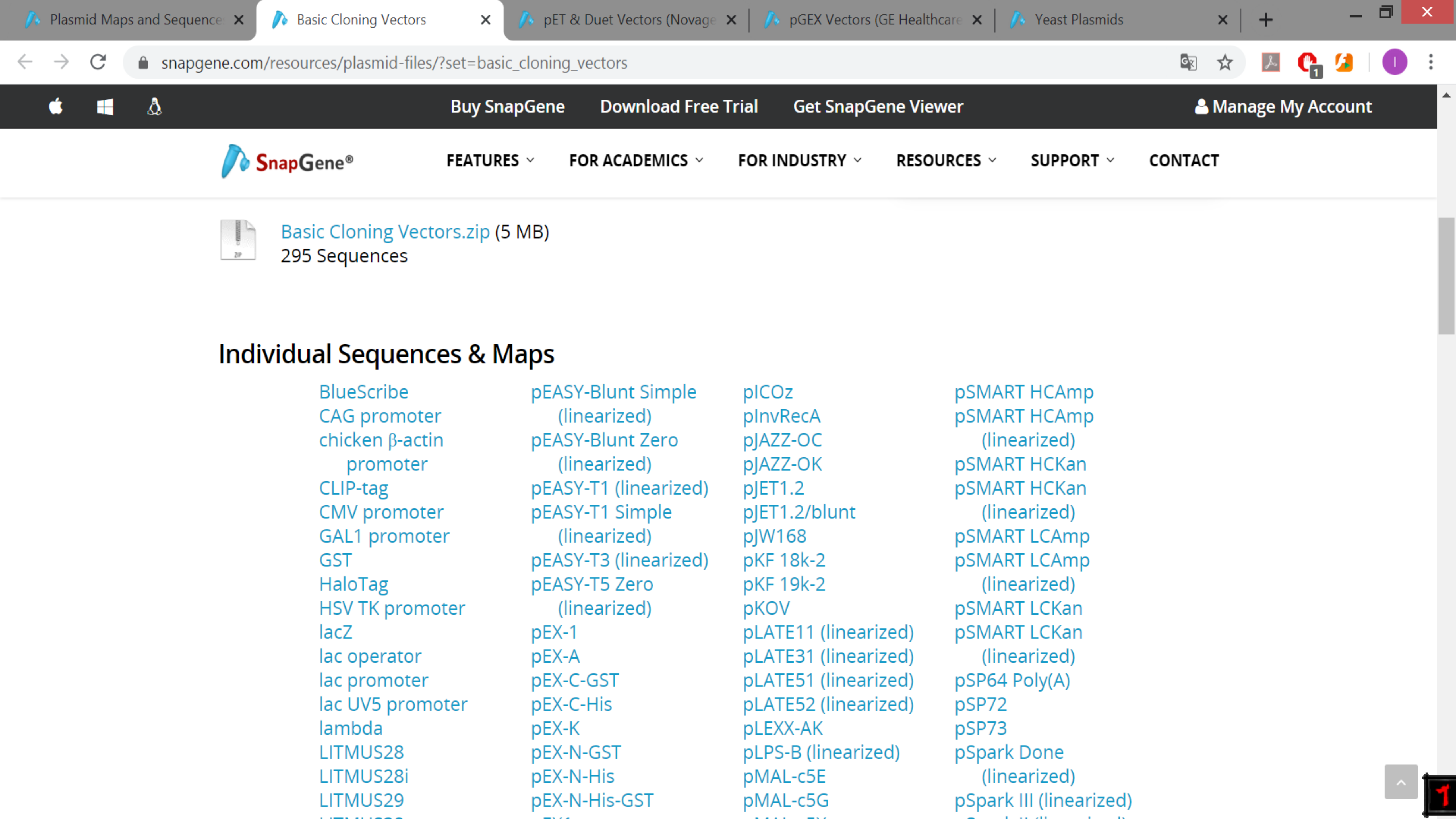
Qiagen Vectors

Expression and tagging vectors

Structural Genomics Vectors

Expression vectors supplied by centers for structural genomics

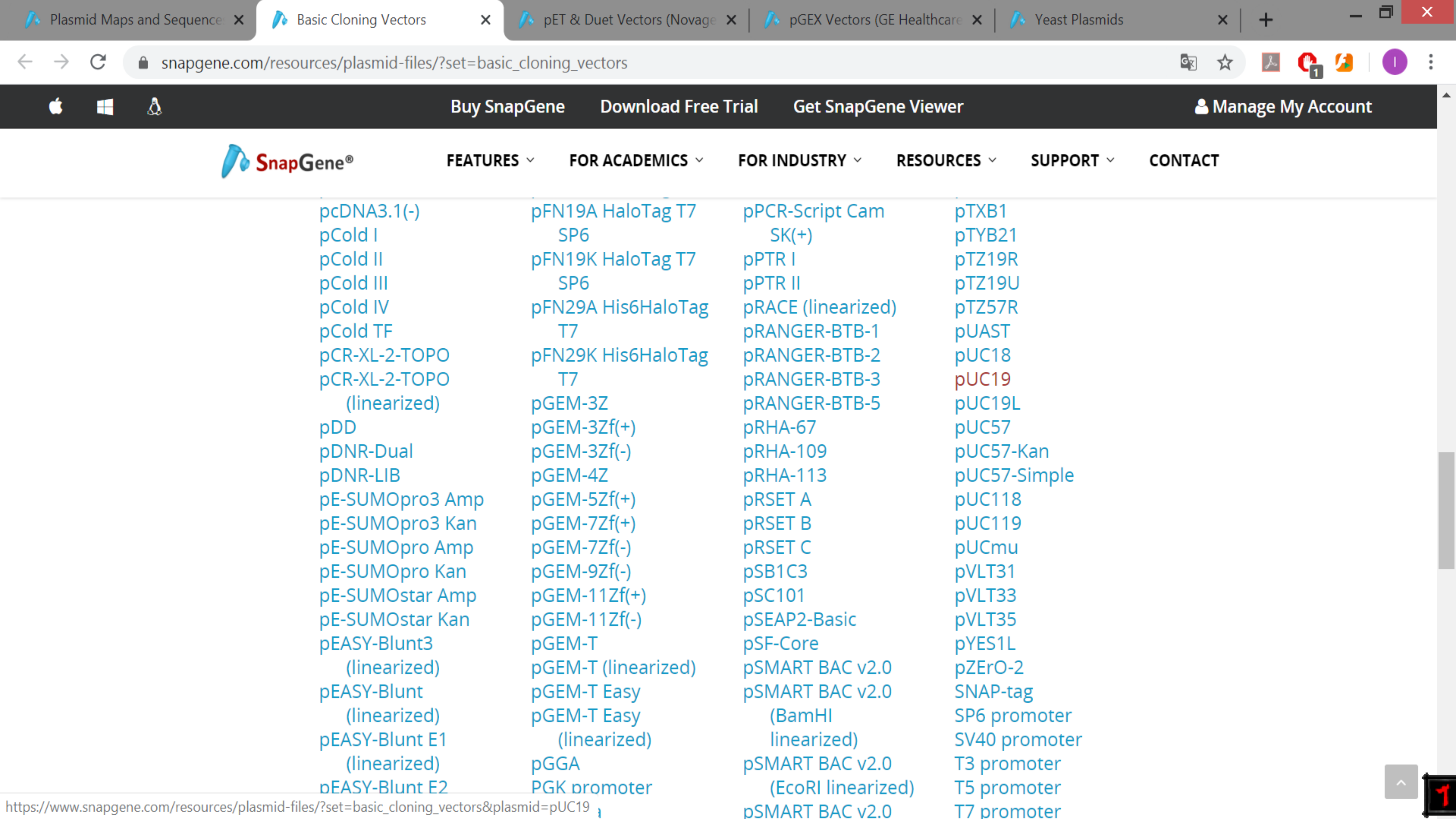
TA and GC Cloning Vectors



Basic Cloning Vectors.zip (5 MB)
295 Sequences

Individual Sequences & Maps

- BlueScribe
- CAG promoter
- chicken β-actin promoter
- CLIP-tag
- CMV promoter
- GAL1 promoter
- GST
- HaloTag
- HSV TK promoter
- lacZ
- lac operator
- lac promoter
- lac UV5 promoter
- lambda
- LITMUS28
- LITMUS28i
- LITMUS29
- pEASY-Blunt Simple (linearized)
- pEASY-Blunt Zero (linearized)
- pEASY-T1 (linearized)
- pEASY-T1 Simple (linearized)
- pEASY-T3 (linearized)
- pEASY-T5 Zero (linearized)
- pEX-1
- pEX-A
- pEX-C-GST
- pEX-C-His
- pEX-K
- pEX-N-GST
- pEX-N-His
- pEX-N-His-GST
- pICOz
- pInvRecA
- pJAZZ-OC
- pJAZZ-OK
- pJET1.2
- pJET1.2/blunt
- pJW168
- pKF 18k-2
- pKF 19k-2
- pKOV
- pLATE11 (linearized)
- pLATE31 (linearized)
- pLATE51 (linearized)
- pLATE52 (linearized)
- pLEXX-AK
- pLPS-B (linearized)
- pMAL-c5E
- pMAL-c5G
- pSMART HCAmp
- pSMART HCAmp (linearized)
- pSMART HCKan
- pSMART HCKan (linearized)
- pSMART LCAmp
- pSMART LCAmp (linearized)
- pSMART LCKan
- pSMART LCKan (linearized)
- pSP64 Poly(A)
- pSP72
- pSP73
- pSpark Done (linearized)
- pSpark III (linearized)



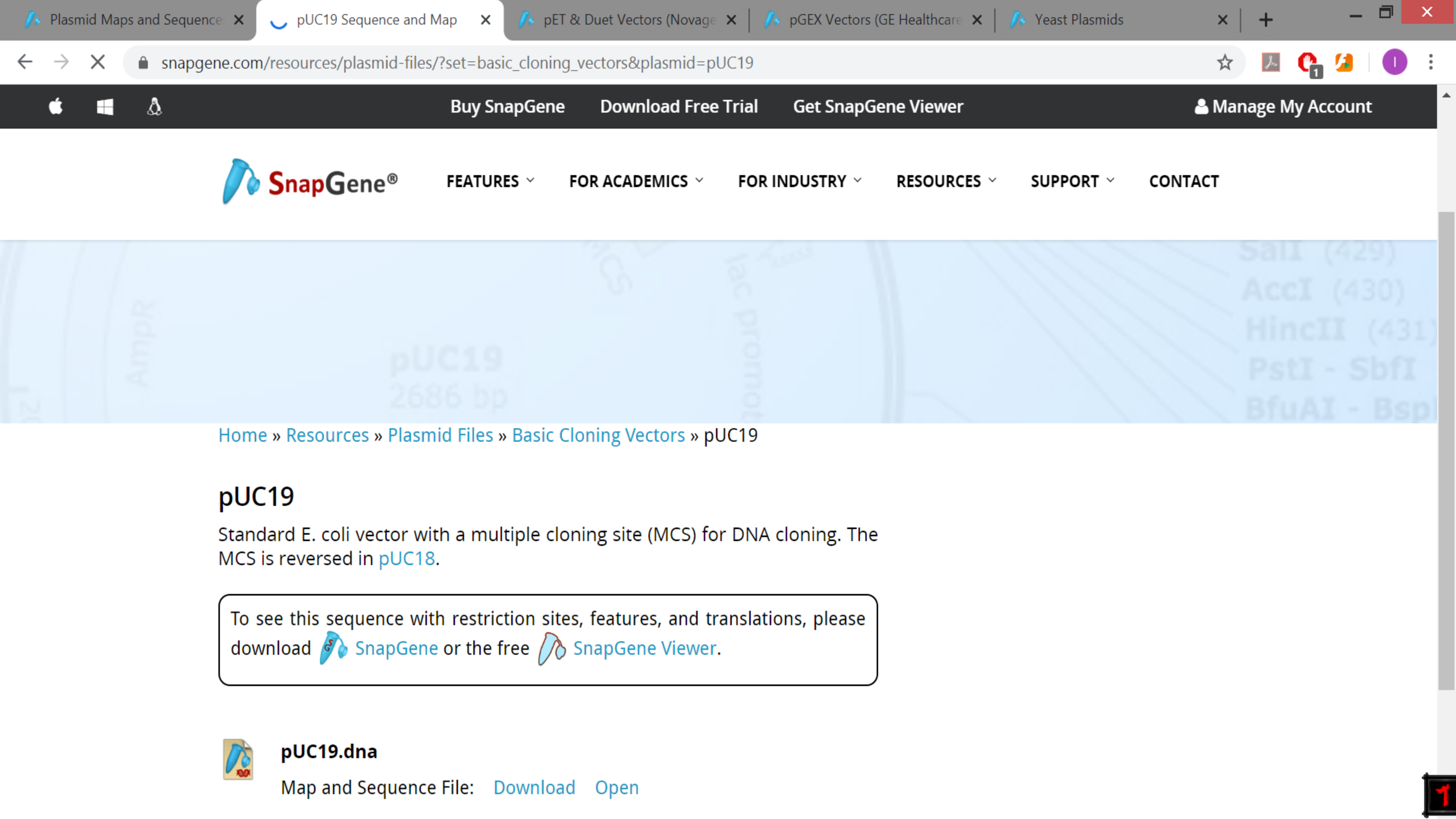
pcDNA3.1(-)
pCold I
pCold II
pCold III
pCold IV
pCold TF
pCR-XL-2-TOPO
pCR-XL-2-TOPO (linearized)
pDD
pDNR-Dual
pDNR-LIB
pE-SUMOpro3 Amp
pE-SUMOpro3 Kan
pE-SUMOpro Amp
pE-SUMOpro Kan
pE-SUMOstar Amp
pE-SUMOstar Kan
pEASY-Blunt3 (linearized)
pEASY-Blunt (linearized)
pEASY-Blunt E1 (linearized)
pEASY-Blunt E2

pFN19A HaloTag T7 SP6
pFN19K HaloTag T7 SP6
pFN29A His6HaloTag T7
pFN29K His6HaloTag T7
pGEM-3Z
pGEM-3Zf(+)
pGEM-3Zf(-)
pGEM-4Z
pGEM-5Zf(+)
pGEM-7Zf(+)
pGEM-7Zf(-)
pGEM-9Zf(-)
pGEM-11Zf(+)
pGEM-11Zf(-)
pGEM-T
pGEM-T (linearized)
pGEM-T Easy
pGEM-T Easy (linearized)
pGGA
PGK promoter

pPCR-Script Cam SK(+)
pPTR I
pPTR II
pRACE (linearized)
pRANGER-BTB-1
pRANGER-BTB-2
pRANGER-BTB-3
pRANGER-BTB-5
pRHA-67
pRHA-109
pRHA-113
pRSET A
pRSET B
pRSET C
pSB1C3
pSC101
pSEAP2-Basic
pSF-Core
pSMART BAC v2.0
pSMART BAC v2.0 (BamHI linearized)
pSMART BAC v2.0 (EcoRI linearized)
pSMART BAC v2.0

pTXB1
pTYB21
pTZ19R
pTZ19U
pTZ57R
pUAST
pUC18
pUC19
pUC19L
pUC57
pUC57-Kan
pUC57-Simple
pUC118
pUC119
pUCmu
pVLT31
pVLT33
pVLT35
pYES1L
pZErO-2
SNAP-tag
SP6 promoter
SV40 promoter
T3 promoter
T5 promoter
T7 promoter

https://www.snapgene.com/resources/plasmid-files/?set=basic_cloning_vectors&plasmid=pUC19



Home » Resources » Plasmid Files » Basic Cloning Vectors » pUC19

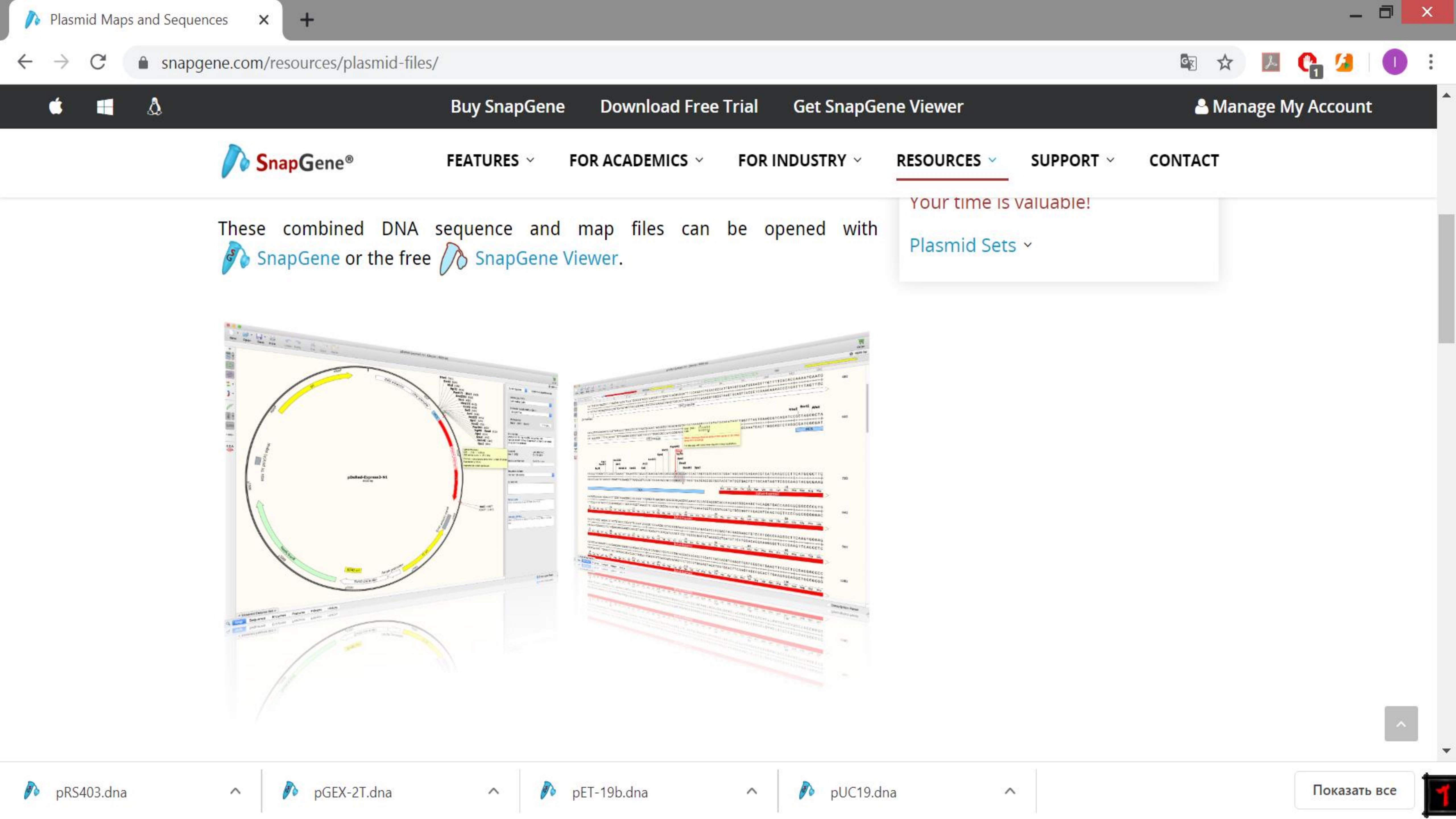
pUC19

Standard E. coli vector with a multiple cloning site (MCS) for DNA cloning. The MCS is reversed in pUC18.

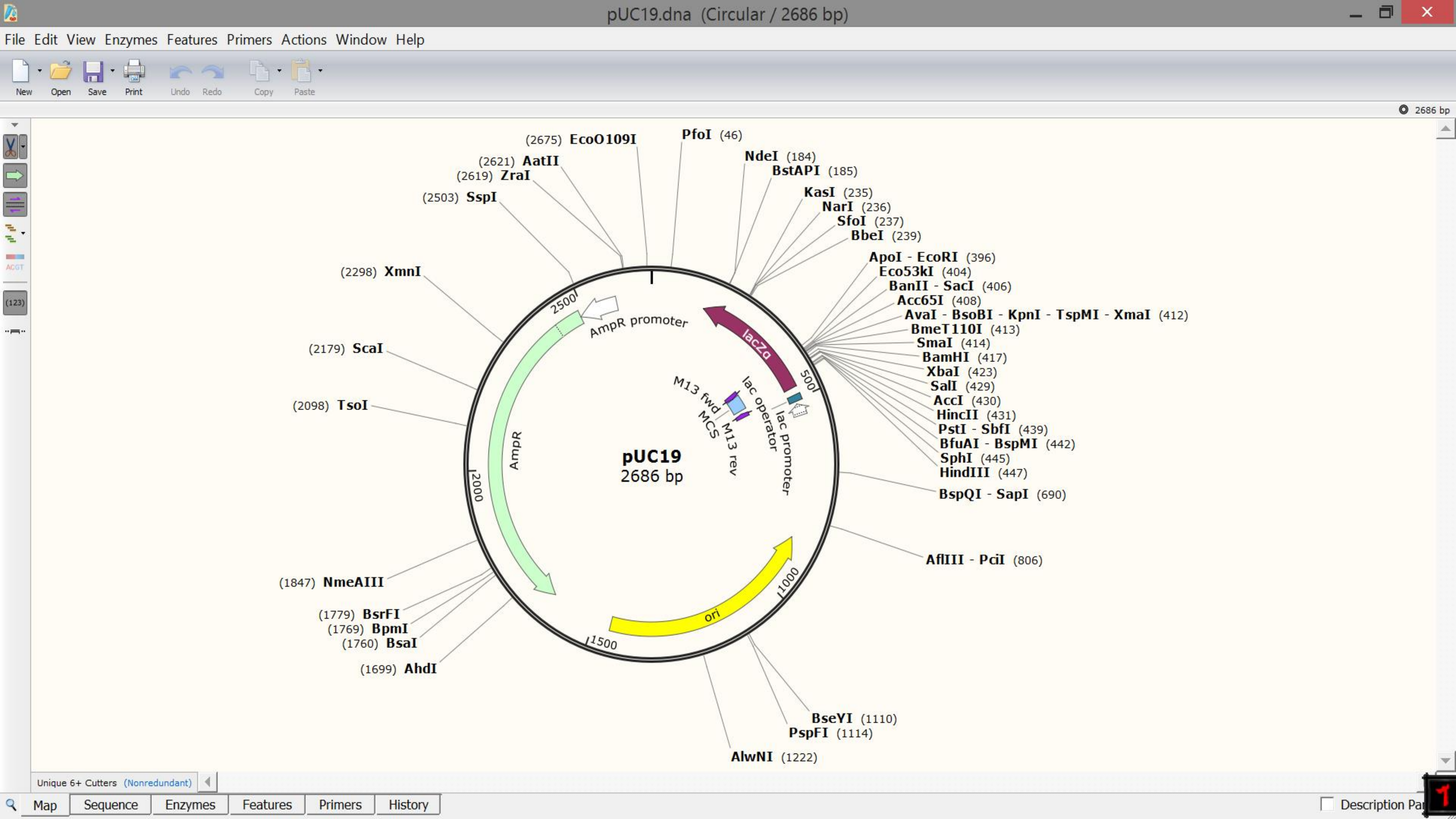
To see this sequence with restriction sites, features, and translations, please download SnapGene or the free SnapGene Viewer.

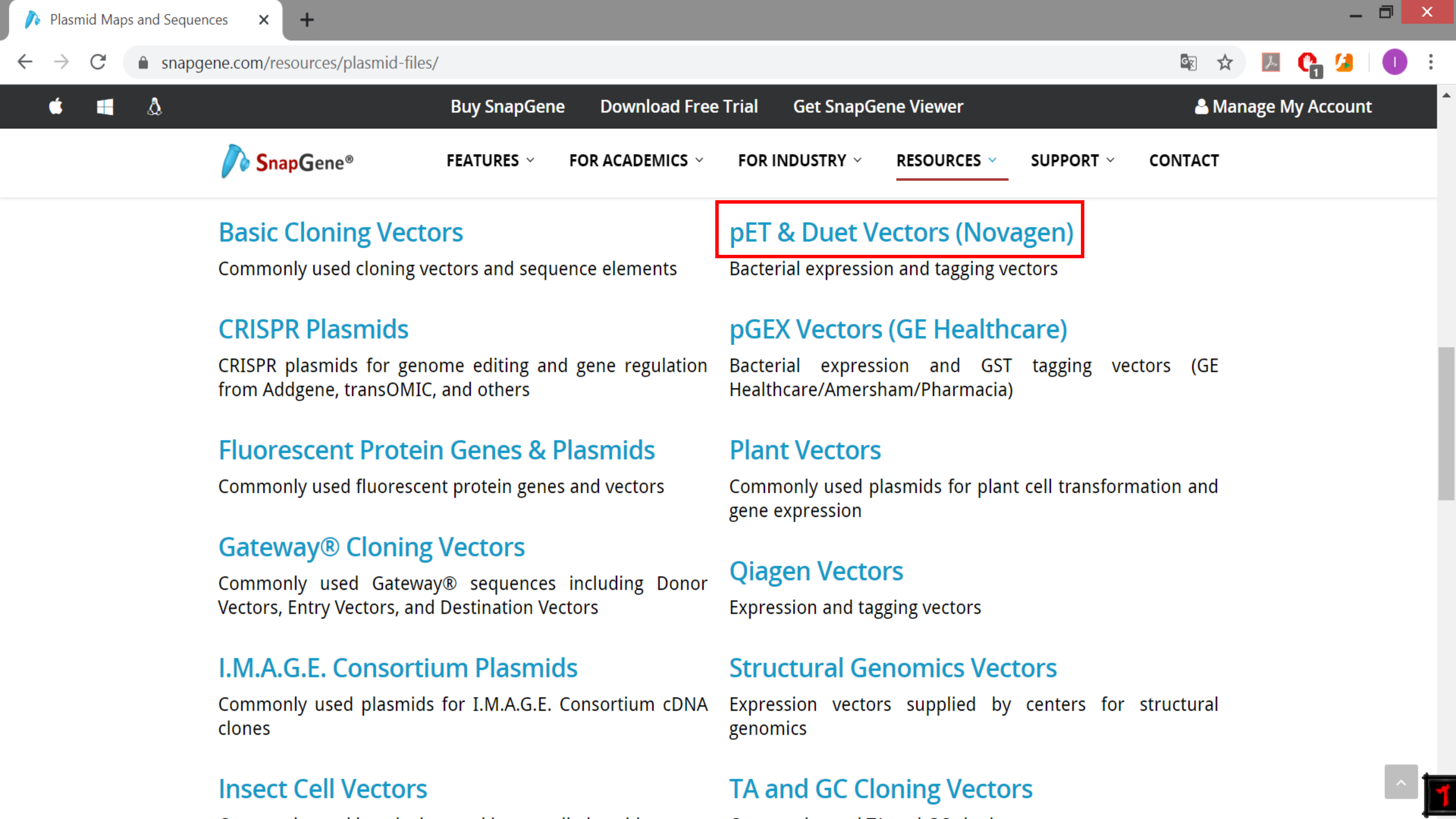
pUC19.dna

Map and Sequence File: Download Open



These combined DNA sequence and map files can be opened with SnapGene or the free SnapGene Viewer.


pUC19.dna (Circular / 2686 bp)
File Edit View Enzymes Features Primers Actions Window Help
New
Open
Save
Print
Undo
Redo
Copy
Paste
2686 bp
(123)
(2675) EcoO109I
PfoI (46)
NdeI (184)
BstAPI (185)
(2621) AatII
(2619) ZraI
(2503) SspI
KasI (235)
NarI (236)
SfoI (237)
BbeI (239)
ApoI - EcoRI (396)
Eco53kI (404)
BanII - SacI (406)
Acc65I (408)
AvaI - BsoBI - KpnI - TspMI - XmaI (412)
BmeT110I (413)
SmaI (414)
BamHI (417)
XbaI (423)
SalI (429)
AccI (430)
HincII (431)
PstI - SbfI (439)
BfuAI - BspMI (442)
SphI (445)
HindIII (447)
BspQI - SapI (690)
AflIII - PciI (806)
(2298) XmnI
(2179) ScaI
(2098) TsoI
(1847) NmeAIII
(1779) BsrFI
(1769) BpmI
(1760) BsaI
(1699) AhdI
BseYI (1110)
PspFI (1114)
AlwNI (1222)
AmpR promoter
lacZα
M13 fwd
MCS
M13 rev
lac operator
lac promoter
AmpR
ori
pUC19
2686 bp
2500
500
2000
1000
1500
Unique 6+ Cutters (Nonredundant)
Map
Sequence
Enzymes
Features
Primers
History
Description Pa



Basic Cloning Vectors

Commonly used cloning vectors and sequence elements

pET & Duet Vectors (Novagen)

Bacterial expression and tagging vectors

CRISPR Plasmids

CRISPR plasmids for genome editing and gene regulation from Addgene, transOMIC, and others

pGEX Vectors (GE Healthcare)

Bacterial expression and GST tagging vectors (GE Healthcare/Amersham/Pharmacia)

Fluorescent Protein Genes & Plasmids

Commonly used fluorescent protein genes and vectors

Plant Vectors

Commonly used plasmids for plant cell transformation and gene expression

Gateway® Cloning Vectors

Commonly used Gateway® sequences including Donor Vectors, Entry Vectors, and Destination Vectors

Qiagen Vectors

Expression and tagging vectors

I.M.A.G.E. Consortium Plasmids

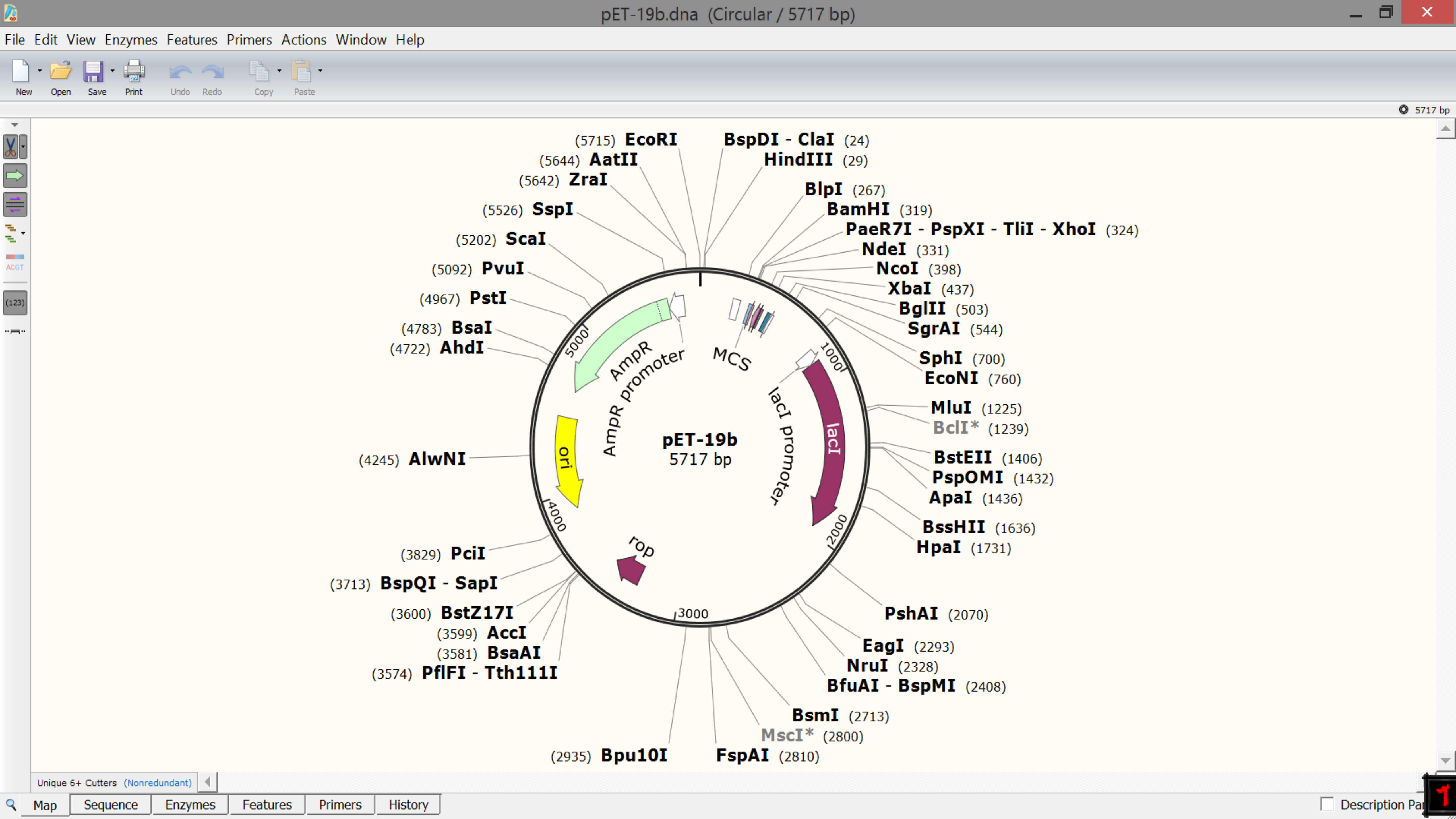
Commonly used plasmids for I.M.A.G.E. Consortium cDNA clones

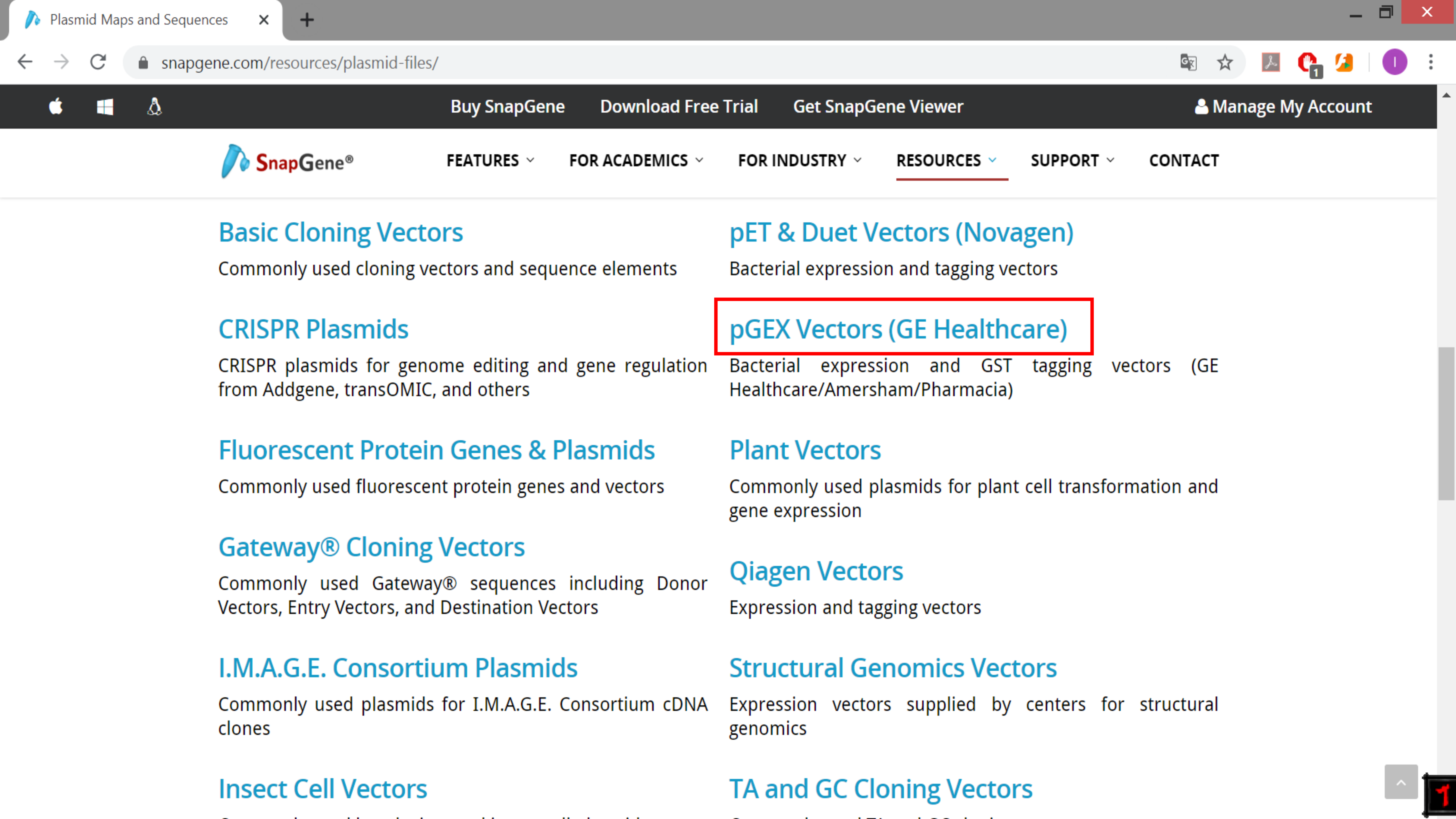
Structural Genomics Vectors

Expression vectors supplied by centers for structural genomics

Insect Cell Vectors

TA and GC Cloning Vectors


pET-19b.dna (Circular / 5717 bp)
File Edit View Enzymes Features Primers Actions Window Help
New Open Save Print Undo Redo Copy Paste
5717 bp
(5715) EcoRI
(5644) AatII
(5642) ZraI
(5526) SspI
(5202) ScaI
(5092) PvuI
(4967) PstI
(4783) BsaI
(4722) AhdI
(4245) AlwNI
(3829) PciI
(3713) BspQI - SapI
(3600) BstZ17I
(3599) AccI
(3581) BsaAI
(3574) PflFI - Tth111I
(2935) Bpu10I
BspDI - ClaI (24)
HindIII (29)
BlpI (267)
BamHI (319)
PaeR7I - PspXI - TliI - XhoI (324)
NdeI (331)
NcoI (398)
XbaI (437)
BglII (503)
SgrAI (544)
SphI (700)
EcoNI (760)
MluI (1225)
BclI* (1239)
BstEII (1406)
PspOMI (1432)
ApaI (1436)
BssHII (1636)
HpaI (1731)
PshAI (2070)
EagI (2293)
NruI (2328)
BfuAI - BspMI (2408)
BsmI (2713)
MscI* (2800)
FspAI (2810)
pET-19b
5717 bp
AmpR
AmpR promoter
MCS
lacI
lacI promoter
ori
rop
1000
2000
3000
4000
5000
Unique 6+ Cutters (Nonredundant)
Map Sequence Enzymes Features Primers History
Description Pa



Basic Cloning Vectors

Commonly used cloning vectors and sequence elements

CRISPR Plasmids

CRISPR plasmids for genome editing and gene regulation from Addgene, transOMIC, and others

Fluorescent Protein Genes & Plasmids

Commonly used fluorescent protein genes and vectors

Gateway® Cloning Vectors

Commonly used Gateway® sequences including Donor Vectors, Entry Vectors, and Destination Vectors

I.M.A.G.E. Consortium Plasmids

Commonly used plasmids for I.M.A.G.E. Consortium cDNA clones

Insect Cell Vectors

pET & Duet Vectors (Novagen)

Bacterial expression and tagging vectors

pGEX Vectors (GE Healthcare)

Bacterial expression and GST tagging vectors (GE Healthcare/Amersham/Pharmacia)

Plant Vectors

Commonly used plasmids for plant cell transformation and gene expression

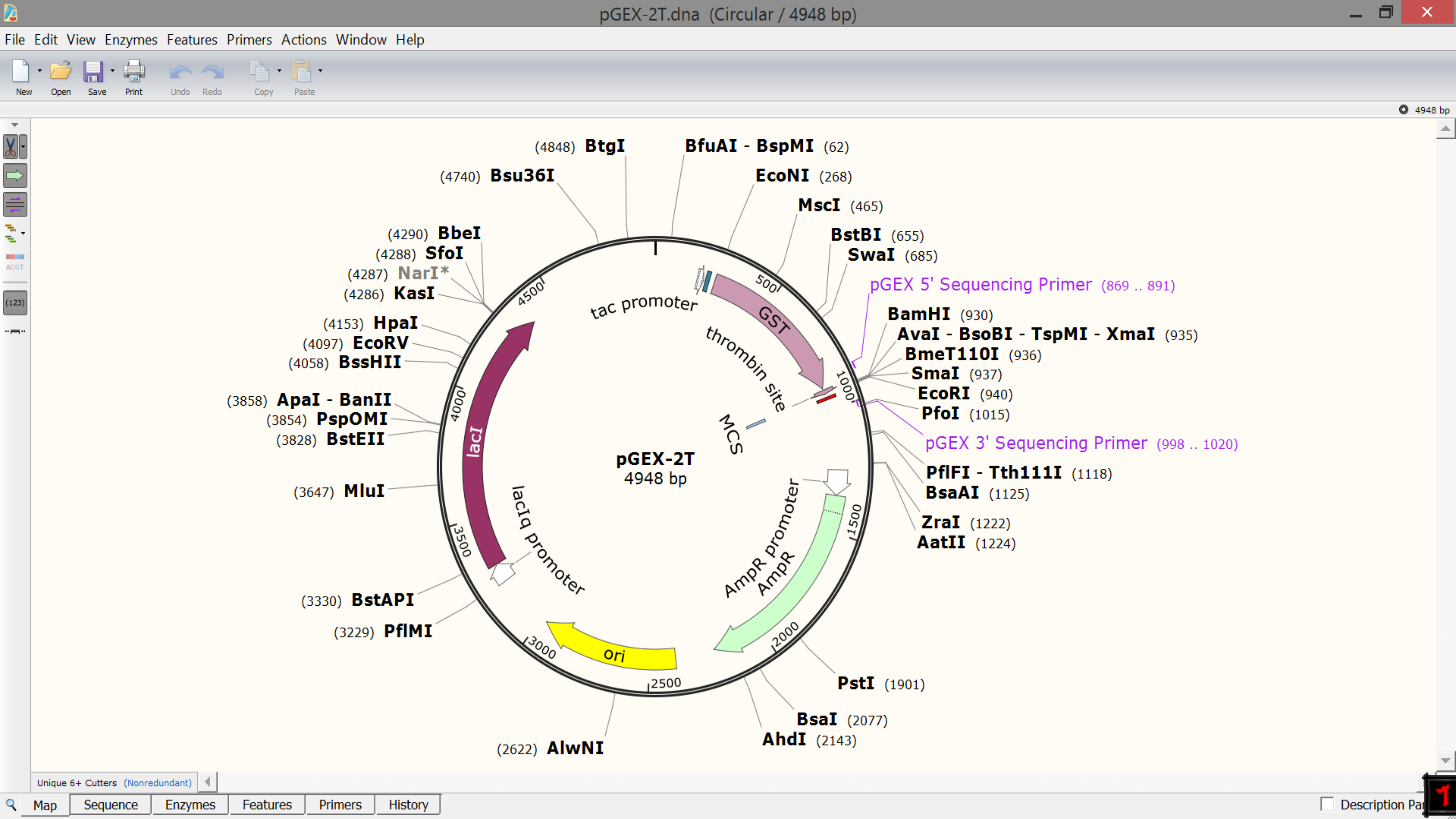
Qiagen Vectors

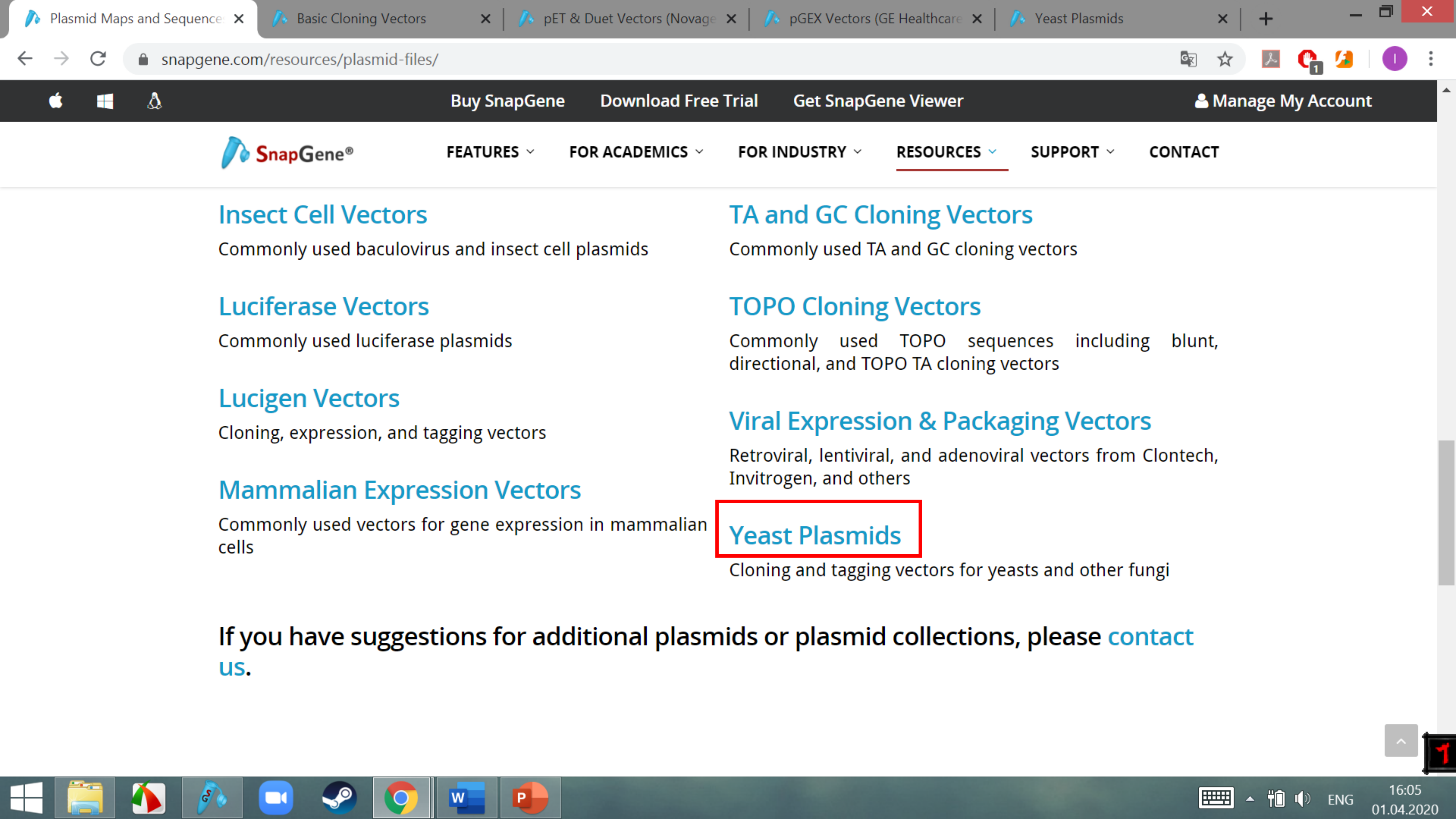
Expression and tagging vectors

Structural Genomics Vectors

Expression vectors supplied by centers for structural genomics

TA and GC Cloning Vectors


pGEX-2T.dna (Circular / 4948 bp)
File Edit View Enzymes Features Primers Actions Window Help
New Open Save Print Undo Redo Copy Paste
4948 bp
(123)
ACGT
(4848) BtgI
BfuAI - BspMI (62)
(4740) Bsu36I
EcoNI (268)
MscI (465)
BstBI (655)
SwaI (685)
(4290) BbeI
(4288) SfoI
(4287) NarI*
(4286) KasI
pGEX 5' Sequencing Primer (869 .. 891)
BamHI (930)
AvaI - BsoBI - TspMI - XmaI (935)
BmeT110I (936)
SmaI (937)
EcoRI (940)
PfoI (1015)
pGEX 3' Sequencing Primer (998 .. 1020)
(4153) HpaI
(4097) EcoRV
(4058) BssHII
(3858) ApaI - BanII
(3854) PspOMI
(3828) BstEII
(3647) MluI
PflFI - Tth111I (1118)
BsaAI (1125)
ZraI (1222)
AatII (1224)
(3330) BstAPI
(3229) PflMI
PstI (1901)
BsaI (2077)
AhdI (2143)
(2622) AlwNI
tac promoter
GST
thrombin site
MCS
lacI
lacIq promoter
AmpR promoter
AmpR
ori
pGEX-2T
4948 bp
500
1000
1500
2000
2500
3000
3500
4000
4500
Unique 6+ Cutters (Nonredundant)
Map Sequence Enzymes Features Primers History
Description Pa



Insect Cell Vectors

Commonly used baculovirus and insect cell plasmids

Luciferase Vectors

Commonly used luciferase plasmids

Lucigen Vectors

Cloning, expression, and tagging vectors

Mammalian Expression Vectors

Commonly used vectors for gene expression in mammalian cells

TA and GC Cloning Vectors

Commonly used TA and GC cloning vectors

TOPO Cloning Vectors

Commonly used TOPO sequences including blunt, directional, and TOPO TA cloning vectors

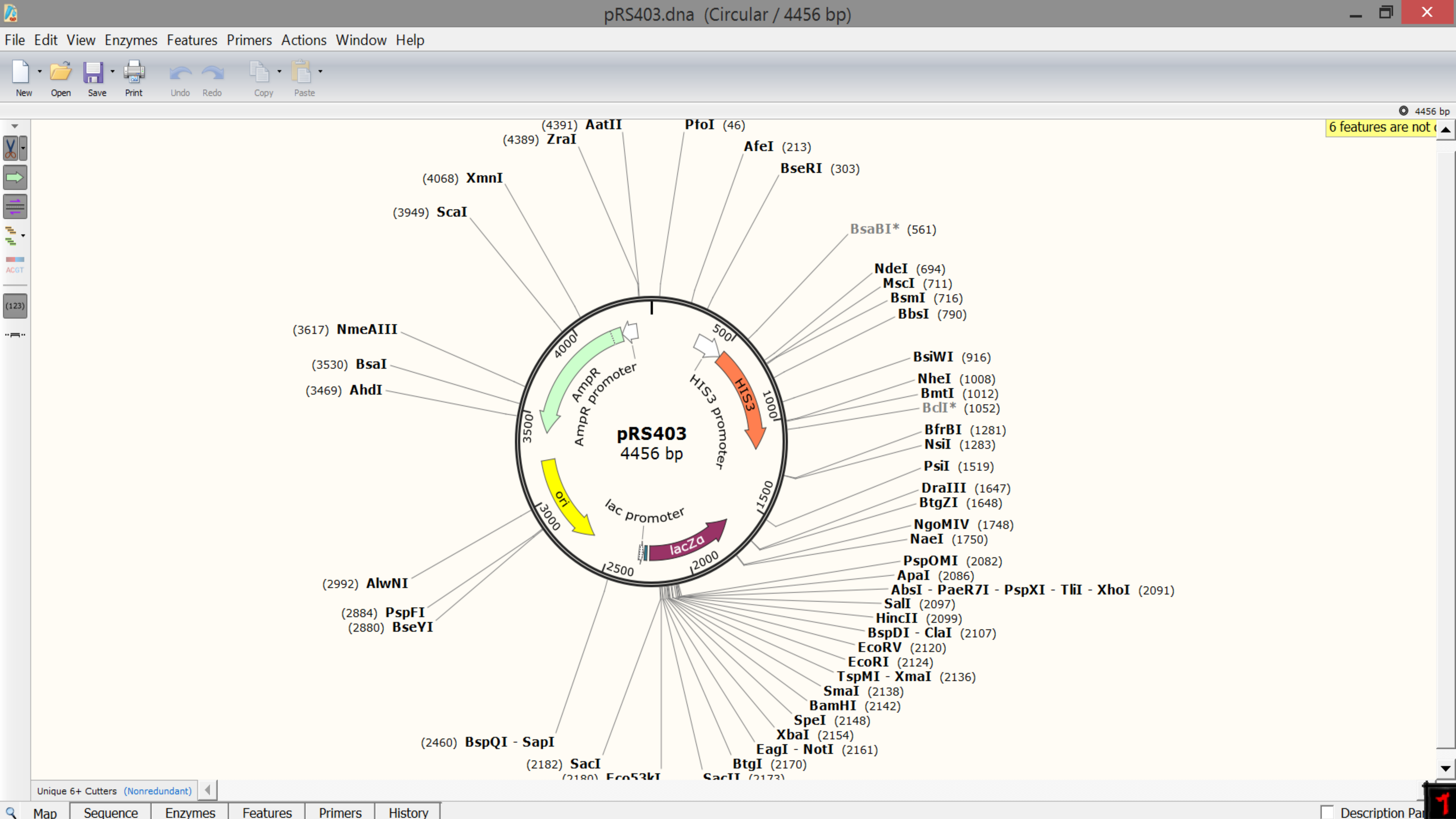
Viral Expression & Packaging Vectors

Retroviral, lentiviral, and adenoviral vectors from Clontech, Invitrogen, and others

Yeast Plasmids

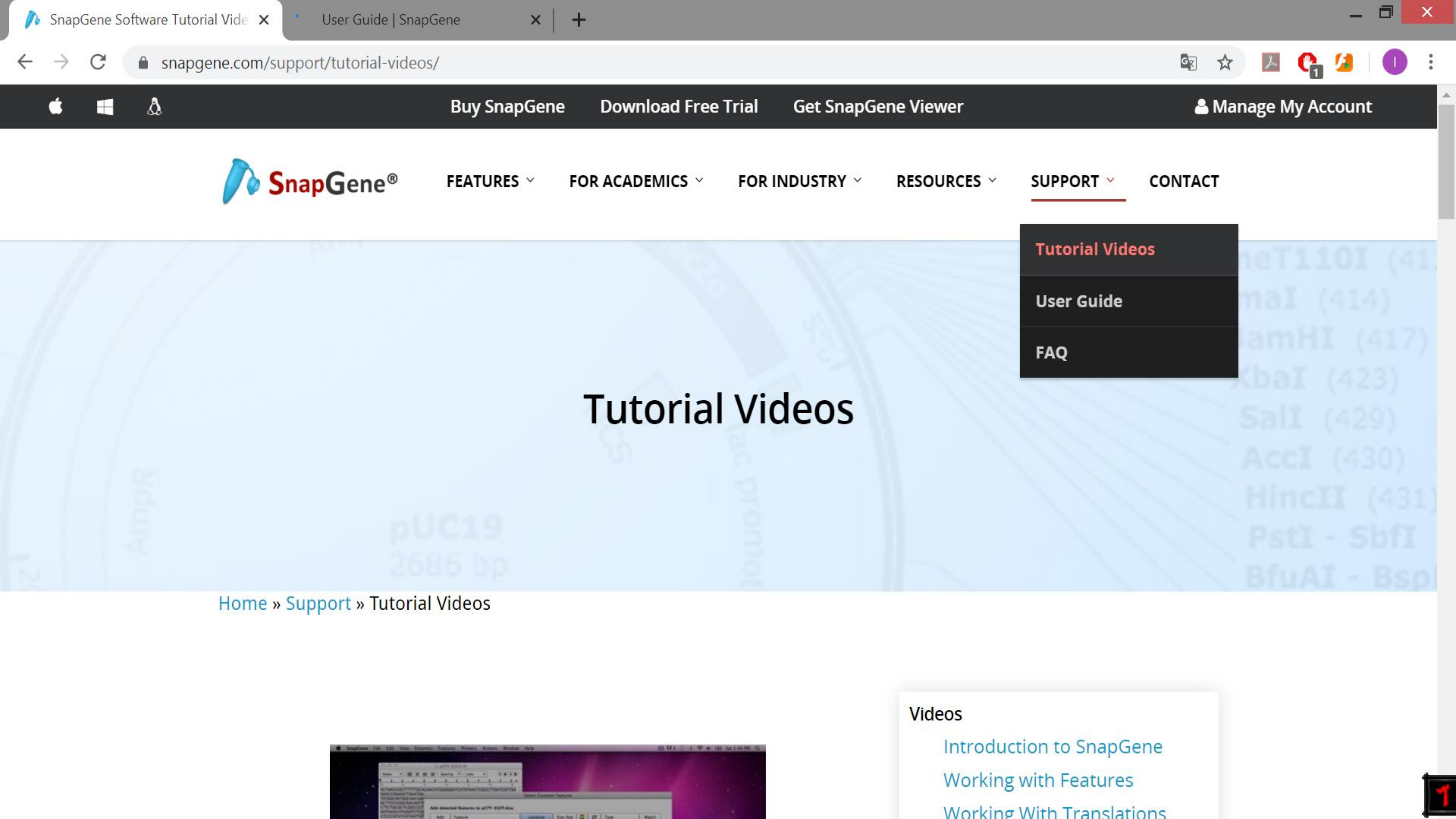
Cloning and tagging vectors for yeasts and other fungi

If you have suggestions for additional plasmids or plasmid collections, please contact us.



pRS403.dna (Circular / 4456 bp)

6 features are not

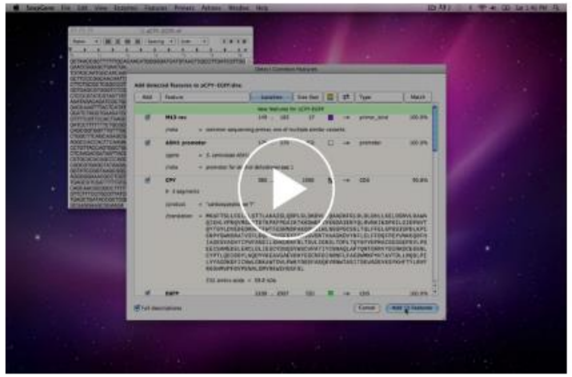


Tutorial Videos

Home » Support » Tutorial Videos

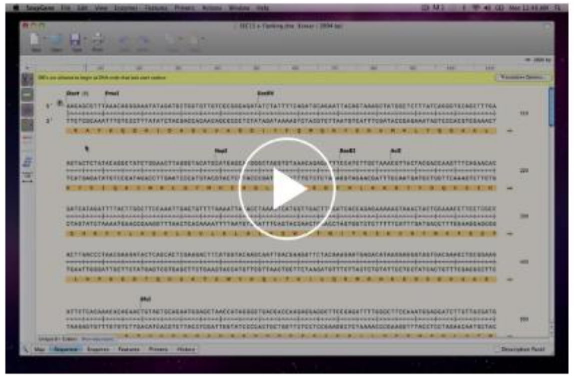
Videos

- Introduction to SnapGene
- Working with Features
- Working With Translations



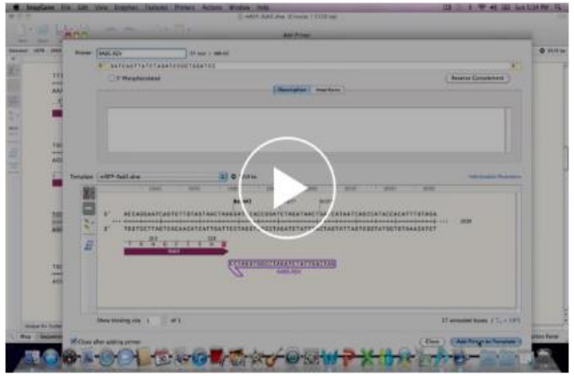
Working with Features

Learn how to create and edit DNA features with SnapGene or SnapGene Viewer, and how to customize the automatic annotation of common features.



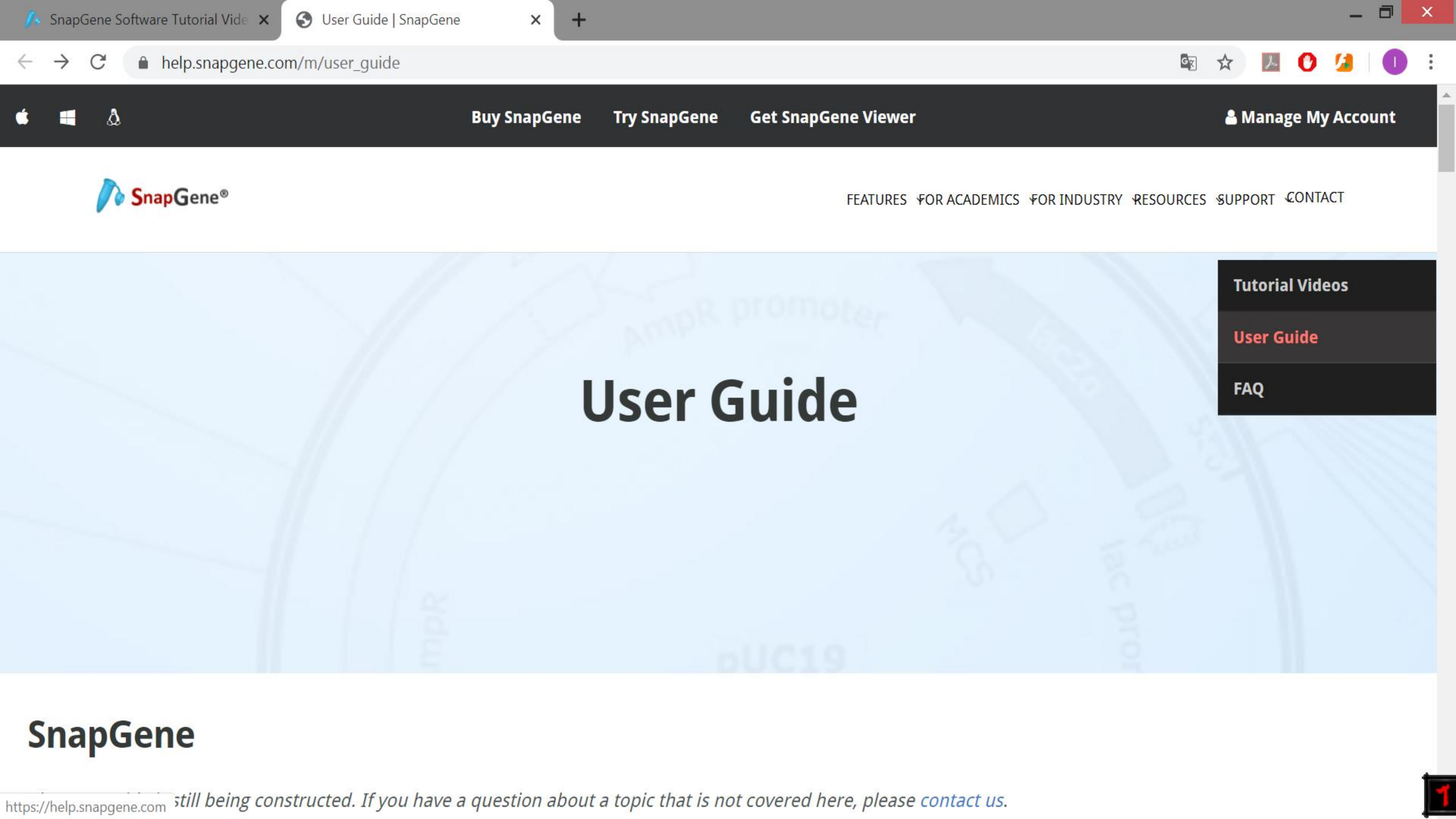
Working with Translations

Learn how SnapGene or SnapGene Viewer can help you to view and edit translated features, open reading frames (ORFs), and whole-sequence translations.



Primers, PCR, and Mutagenesis

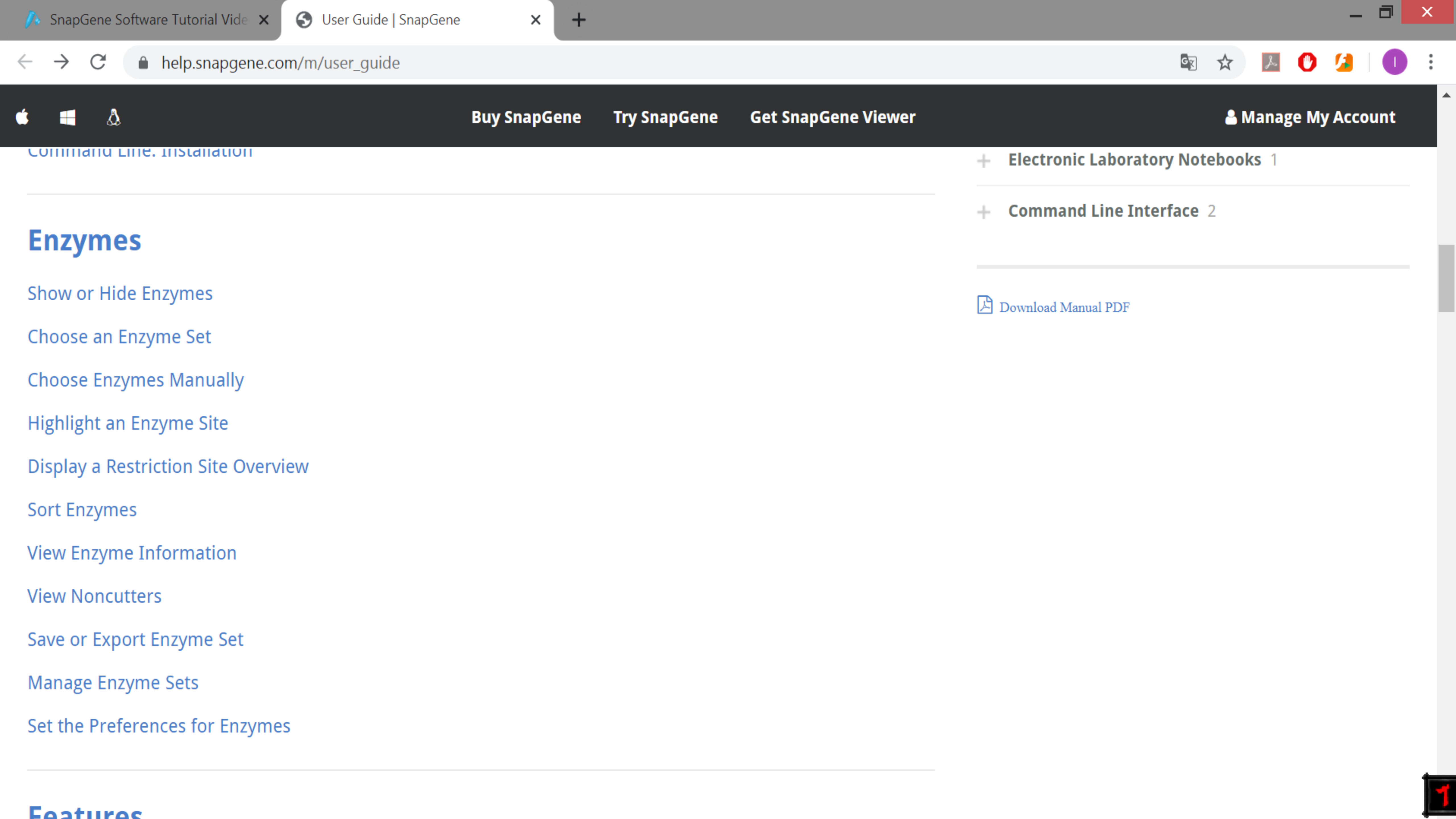
Learn how to make primers, and to simulate PCR and primer-directed mutagenesis.



User Guide

SnapGene

still being constructed. If you have a question about a topic that is not covered here, please contact us.



Command Line: Installation

Enzymes

Show or Hide Enzymes

Choose an Enzyme Set

Choose Enzymes Manually

Highlight an Enzyme Site

Display a Restriction Site Overview

Sort Enzymes

View Enzyme Information

View Noncutters

Save or Export Enzyme Set

Manage Enzyme Sets

Set the Preferences for Enzymes

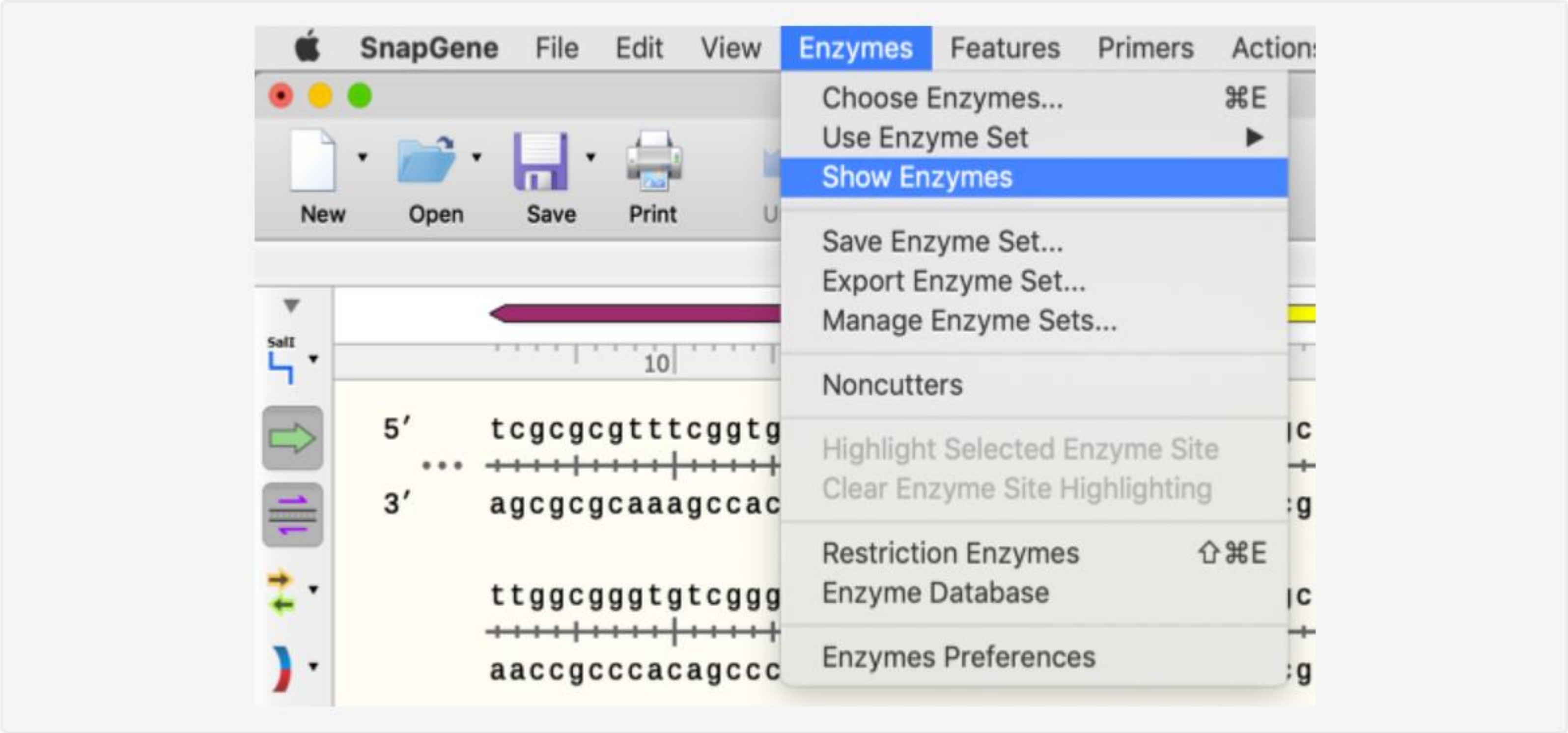
+ **Electronic Laboratory Notebooks** 1

+ **Command Line Interface** 2

Download Manual PDF

Features

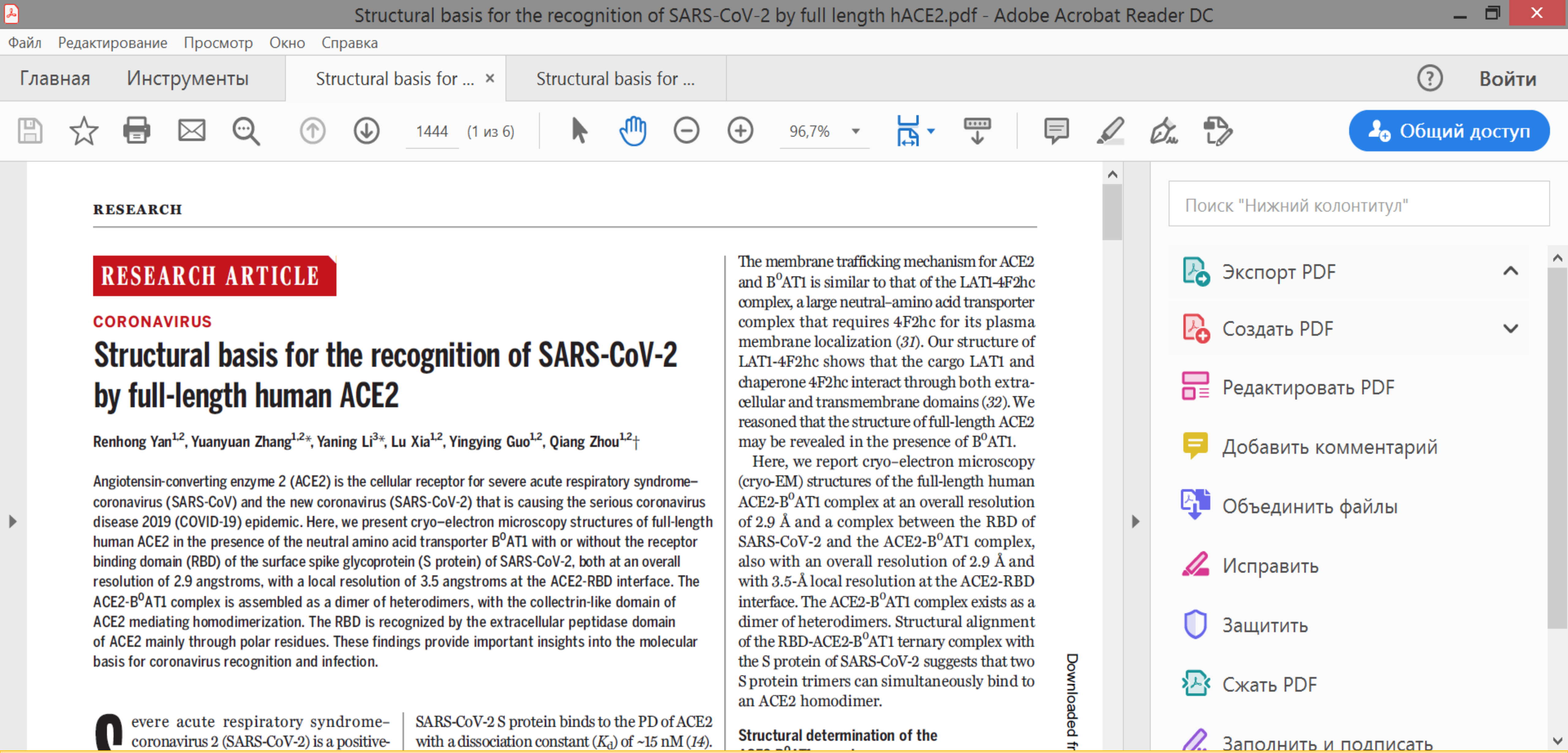
Show the Enzymes



To show the enzymes, click **Enzymes** → **Show Enzymes**.

Hide the Enzymes

Highlight an Enzyme Site
Display a Restriction Site Overview
Sort Enzymes
View Enzyme Information
View Noncutters
Save or Export Enzyme Set
Manage Enzyme Sets
Set the Preferences for Enzymes
+ Features 21
+ Primers 7
+ Translations 6
+ Colors 4
+ Searching 3



RESEARCH

RESEARCH ARTICLE

CORONAVIRUS

Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2

Renhong Yan[1,2], Yuanyuan Zhang[1,2]*, Yaning Li[3]*, Lu Xia[1,2], Yingying Guo[1,2], Qiang Zhou[1,2]†

Angiotensin-converting enzyme 2 (ACE2) is the cellular receptor for severe acute respiratory syndrome–coronavirus (SARS-CoV) and the new coronavirus (SARS-CoV-2) that is causing the serious coronavirus disease 2019 (COVID-19) epidemic. Here, we present cryo–electron microscopy structures of full-length human ACE2 in the presence of the neutral amino acid transporter B^{0}AT1 with or without the receptor binding domain (RBD) of the surface spike glycoprotein (S protein) of SARS-CoV-2, both at an overall resolution of 2.9 angstroms, with a local resolution of 3.5 angstroms at the ACE2-RBD interface. The ACE2-B^{0}AT1 complex is assembled as a dimer of heterodimers, with the collectrin-like domain of ACE2 mediating homodimerization. The RBD is recognized by the extracellular peptidase domain of ACE2 mainly through polar residues. These findings provide important insights into the molecular basis for coronavirus recognition and infection.

Severe acute respiratory syndrome–coronavirus 2 (SARS-CoV-2) is a positive-

SARS-CoV-2 S protein binds to the PD of ACE2 with a dissociation constant (K_d) of ~15 nM (14).

The membrane trafficking mechanism for ACE2 and B^{0}AT1 is similar to that of the LAT1-4F2hc complex, a large neutral–amino acid transporter complex that requires 4F2hc for its plasma membrane localization (31). Our structure of LAT1-4F2hc shows that the cargo LAT1 and chaperone 4F2hc interact through both extracellular and transmembrane domains (32). We reasoned that the structure of full-length ACE2 may be revealed in the presence of B^{0}AT1.

Here, we report cryo–electron microscopy (cryo-EM) structures of the full-length human ACE2-B^{0}AT1 complex at an overall resolution of 2.9 Å and a complex between the RBD of SARS-CoV-2 and the ACE2-B^{0}AT1 complex, also with an overall resolution of 2.9 Å and with 3.5-Å local resolution at the ACE2-RBD interface. The ACE2-B^{0}AT1 complex exists as a dimer of heterodimers. Structural alignment of the RBD-ACE2-B^{0}AT1 ternary complex with the S protein of SARS-CoV-2 suggests that two S protein trimers can simultaneously bind to an ACE2 homodimer.

Structural determination of the

Ангиотензин-превращающий фермент 2 (ACE2) является клеточным рецептором для тяжелого острого респираторного синдрома - коронавируса (SARS-CoV) и нового коронавируса (SARS-CoV-2), который вызывает серьезную эпидемию коронавирусной болезни 2019 (COVID-19).

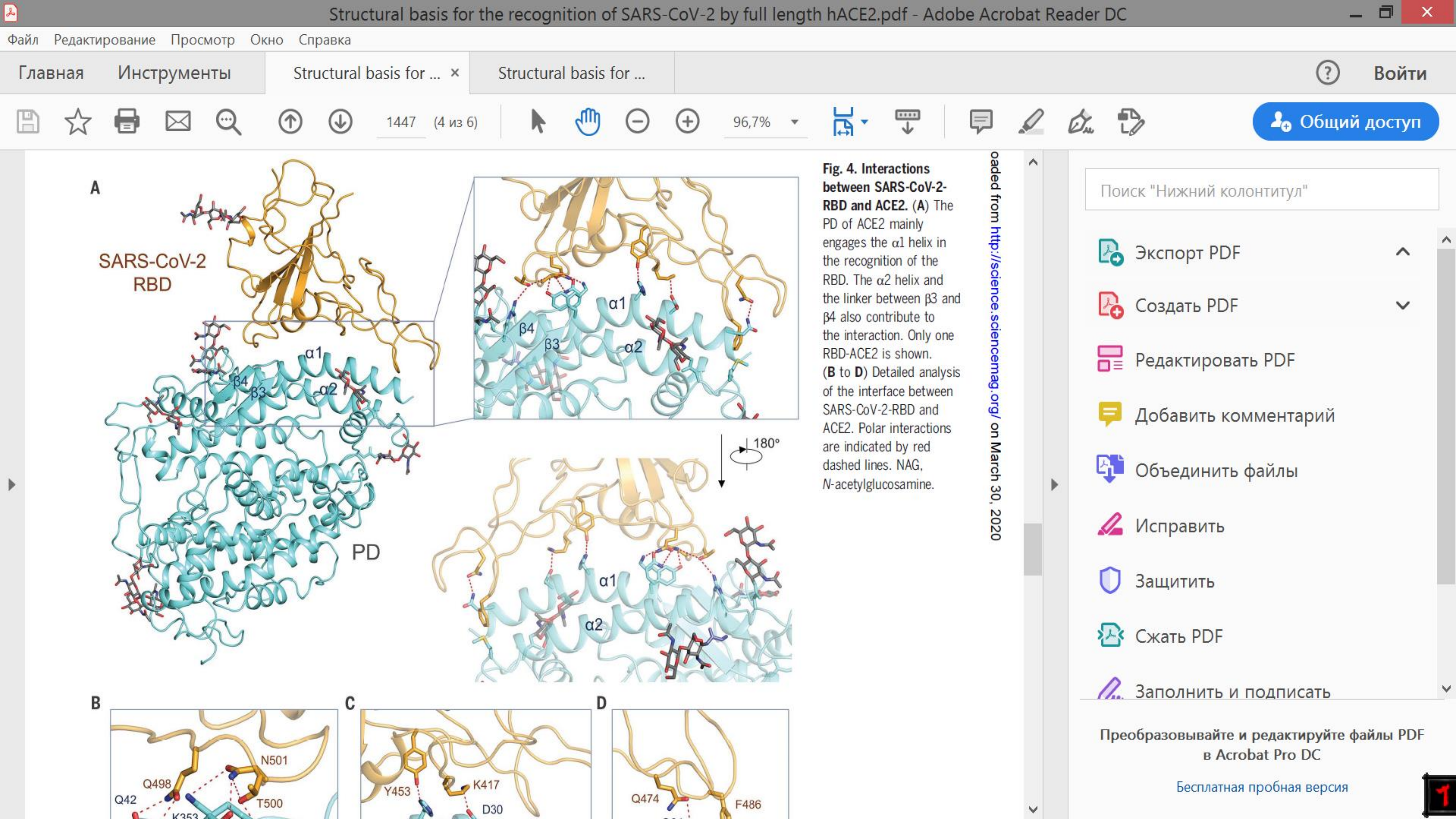
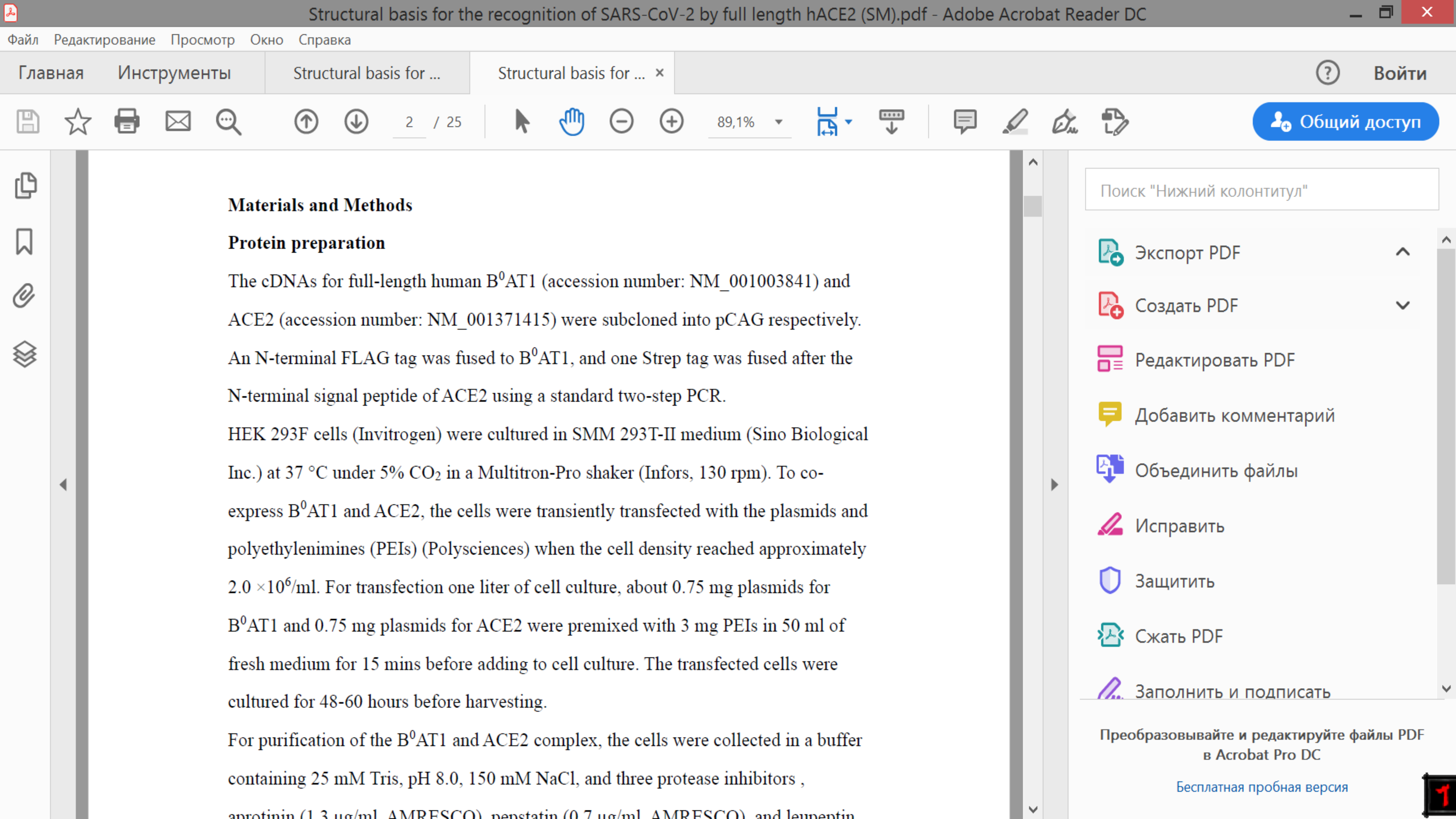


Fig. 4. Interactions between SARS-CoV-2-RBD and ACE2. (**A**) The PD of ACE2 mainly engages the α1 helix in the recognition of the RBD. The α2 helix and the linker between β3 and β4 also contribute to the interaction. Only one RBD-ACE2 is shown. (**B** to **D**) Detailed analysis of the interface between SARS-CoV-2-RBD and ACE2. Polar interactions are indicated by red dashed lines. NAG, *N*-acetylglucosamine.

Преобразовывайте и редактируйте файлы PDF в Acrobat Pro DC

Бесплатная пробная версия



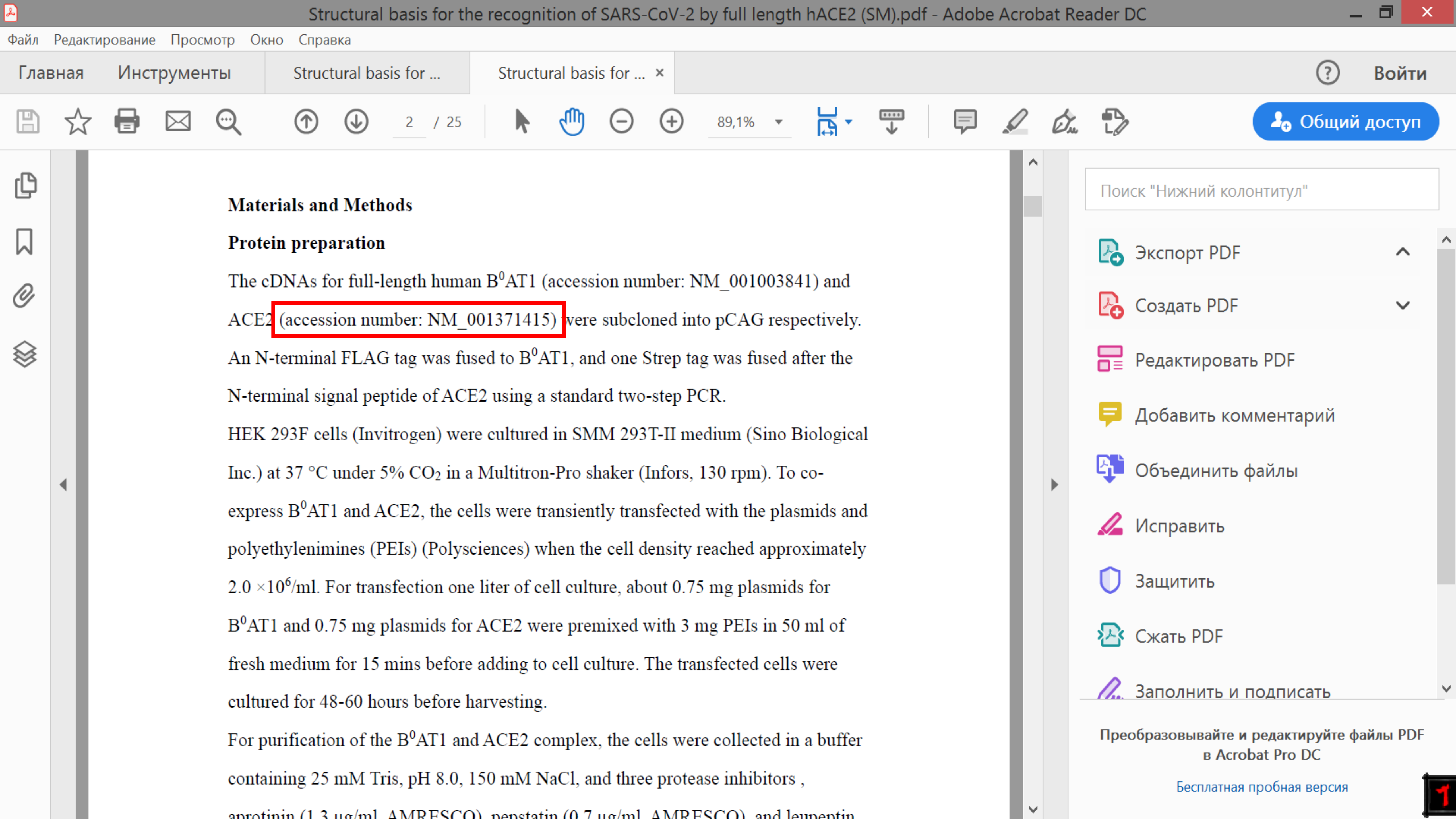
Materials and Methods

Protein preparation

The cDNAs for full-length human B^0AT1 (accession number: NM_001003841) and ACE2 (accession number: NM_001371415) were subcloned into pCAG respectively. An N-terminal FLAG tag was fused to B^0AT1, and one Strep tag was fused after the N-terminal signal peptide of ACE2 using a standard two-step PCR.

HEK 293F cells (Invitrogen) were cultured in SMM 293T-II medium (Sino Biological Inc.) at 37 °C under 5% CO_2 in a Multitron-Pro shaker (Infors, 130 rpm). To co-express B^0AT1 and ACE2, the cells were transiently transfected with the plasmids and polyethylenimines (PEIs) (Polysciences) when the cell density reached approximately 2.0×10^6/ml. For transfection one liter of cell culture, about 0.75 mg plasmids for B^0AT1 and 0.75 mg plasmids for ACE2 were premixed with 3 mg PEIs in 50 ml of fresh medium for 15 mins before adding to cell culture. The transfected cells were cultured for 48-60 hours before harvesting.

For purification of the B^0AT1 and ACE2 complex, the cells were collected in a buffer containing 25 mM Tris, pH 8.0, 150 mM NaCl, and three protease inhibitors , aprotinin (1.3 μg/ml, AMRESCO), pepstatin (0.7 μg/ml, AMRESCO), and leupeptin



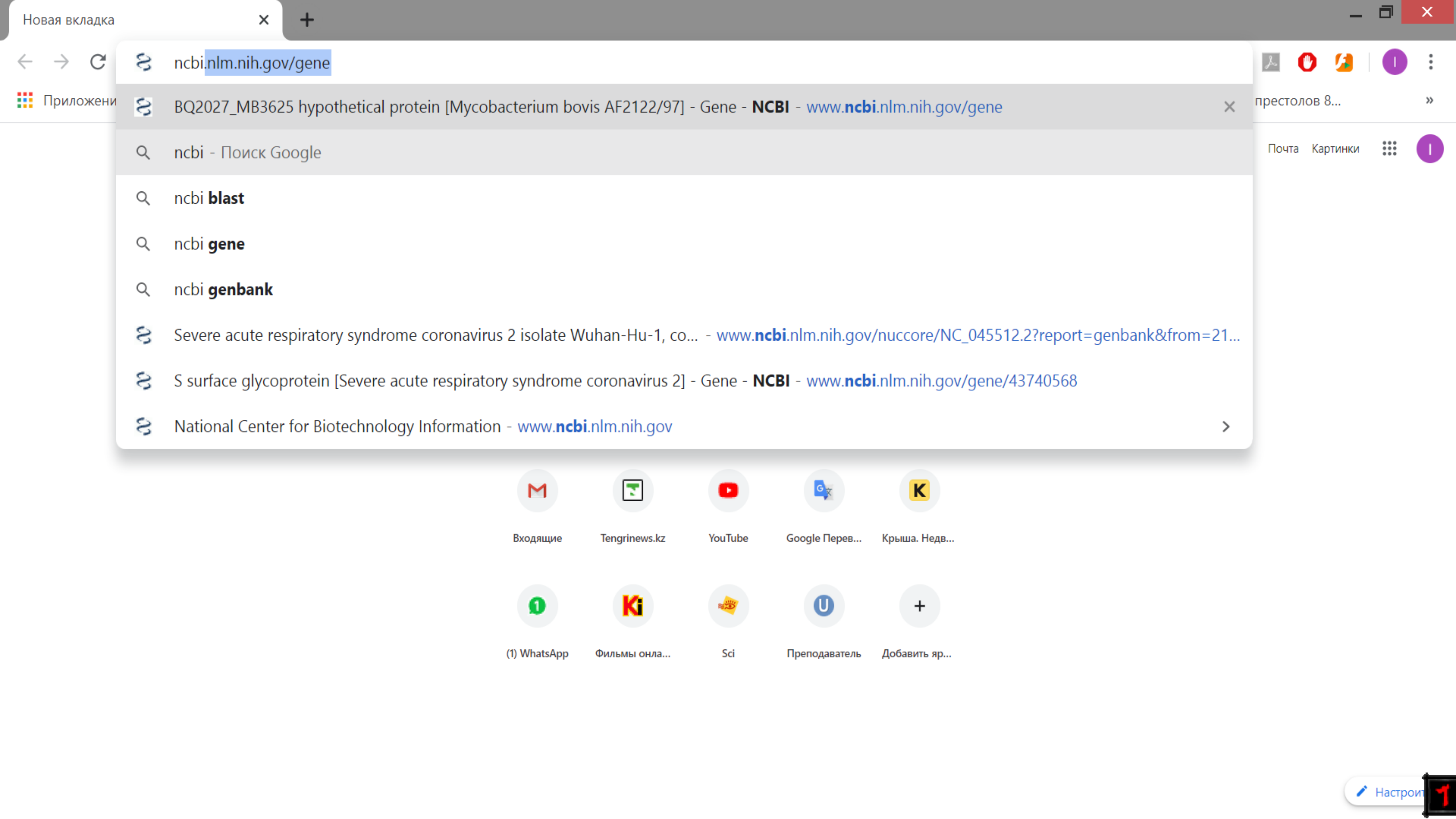
Materials and Methods

Protein preparation

The cDNAs for full-length human B^{0}AT1 (accession number: NM_001003841) and ACE2 (accession number: NM_001371415) were subcloned into pCAG respectively. An N-terminal FLAG tag was fused to B^{0}AT1, and one Strep tag was fused after the N-terminal signal peptide of ACE2 using a standard two-step PCR.

HEK 293F cells (Invitrogen) were cultured in SMM 293T-II medium (Sino Biological Inc.) at 37 °C under 5% CO$_2$ in a Multitron-Pro shaker (Infors, 130 rpm). To co-express B^{0}AT1 and ACE2, the cells were transiently transfected with the plasmids and polyethylenimines (PEIs) (Polysciences) when the cell density reached approximately 2.0×10^6/ml. For transfection one liter of cell culture, about 0.75 mg plasmids for B^{0}AT1 and 0.75 mg plasmids for ACE2 were premixed with 3 mg PEIs in 50 ml of fresh medium for 15 mins before adding to cell culture. The transfected cells were cultured for 48-60 hours before harvesting.

For purification of the B^{0}AT1 and ACE2 complex, the cells were collected in a buffer containing 25 mM Tris, pH 8.0, 150 mM NaCl, and three protease inhibitors , aprotinin (1.3 μg/ml, AMRESCO), pepstatin (0.7 μg/ml, AMRESCO), and leupeptin



ncbi.nlm.nih.gov/gene

BQ2027_MB3625 hypothetical protein [Mycobacterium bovis AF2122/97] - Gene - **NCBI** - www.**ncbi**.nlm.nih.gov/gene

ncbi - Поиск Google

ncbi **blast**

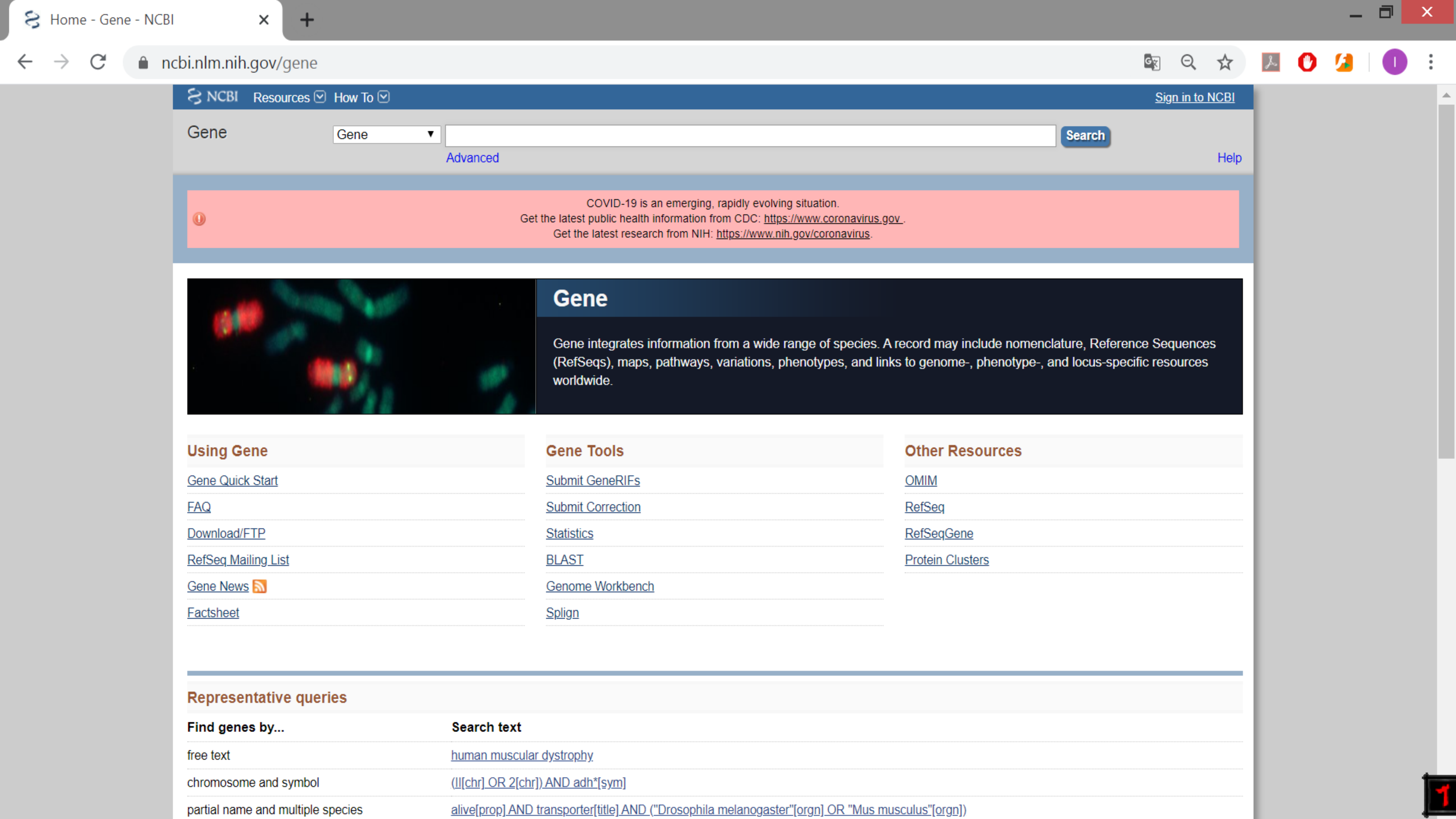
ncbi **gene**

ncbi **genbank**

Severe acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1, co... - www.**ncbi**.nlm.nih.gov/nuccore/NC_045512.2?report=genbank&from=21...

S surface glycoprotein [Severe acute respiratory syndrome coronavirus 2] - Gene - **NCBI** - www.**ncbi**.nlm.nih.gov/gene/43740568

National Center for Biotechnology Information - www.**ncbi**.nlm.nih.gov



COVID-19 is an emerging, rapidly evolving situation.
Get the latest public health information from CDC: https://www.coronavirus.gov .
Get the latest research from NIH: https://www.nih.gov/coronavirus.

Gene

Gene integrates information from a wide range of species. A record may include nomenclature, Reference Sequences (RefSeqs), maps, pathways, variations, phenotypes, and links to genome-, phenotype-, and locus-specific resources worldwide.

Using Gene

- Gene Quick Start
- FAQ
- Download/FTP
- RefSeq Mailing List
- Gene News
- Factsheet

Gene Tools

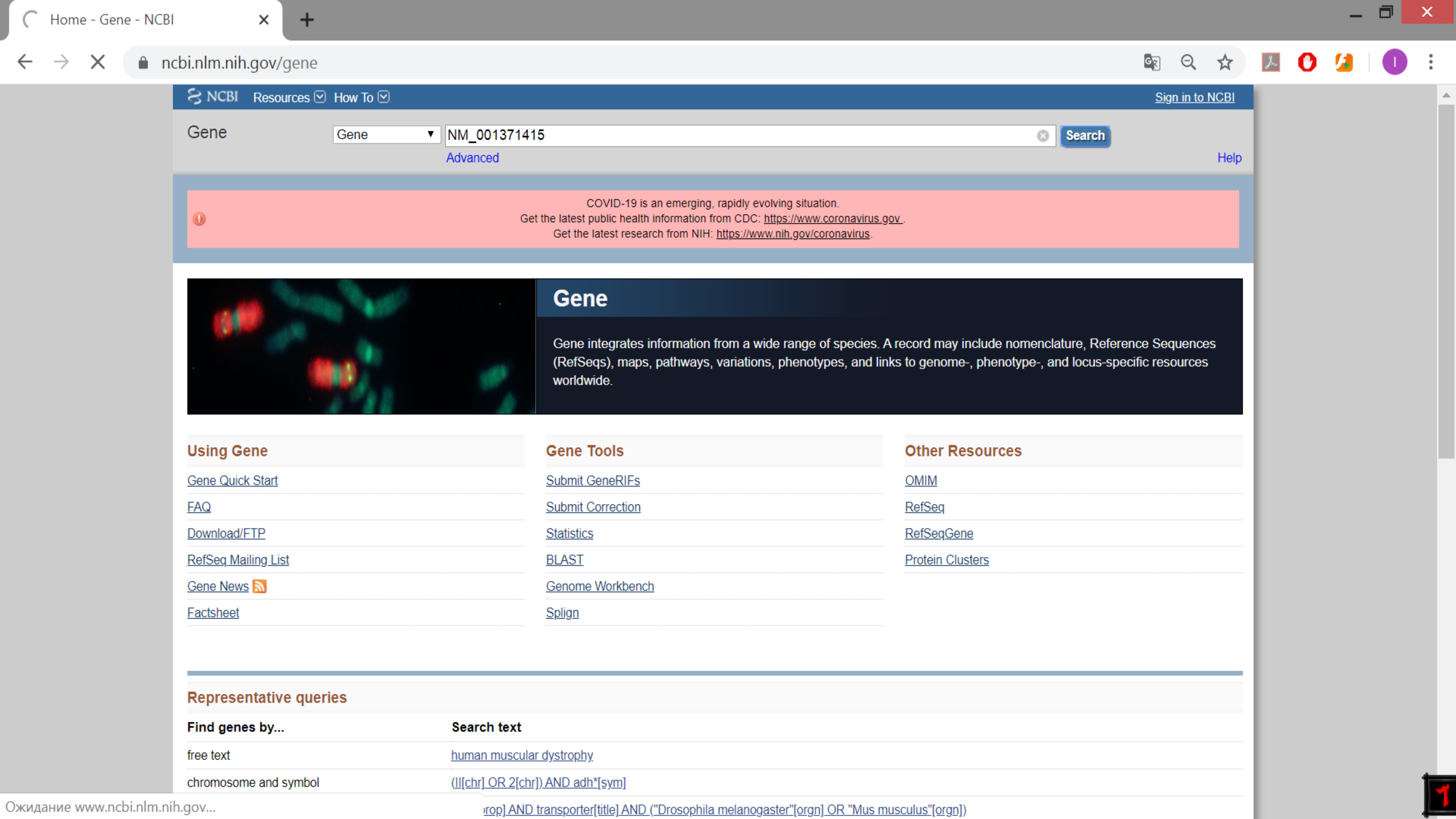
- Submit GeneRIFs
- Submit Correction
- Statistics
- BLAST
- Genome Workbench
- Splign

Other Resources

- OMIM
- RefSeq
- RefSeqGene
- Protein Clusters

Representative queries

Find genes by...	Search text
free text	human muscular dystrophy
chromosome and symbol	(II[chr] OR 2[chr]) AND adh*[sym]
partial name and multiple species	alive[prop] AND transporter[title] AND ("Drosophila melanogaster"[orgn] OR "Mus musculus"[orgn])



COVID-19 is an emerging, rapidly evolving situation.
Get the latest public health information from CDC: https://www.coronavirus.gov .
Get the latest research from NIH: https://www.nih.gov/coronavirus.

Gene

Gene integrates information from a wide range of species. A record may include nomenclature, Reference Sequences (RefSeqs), maps, pathways, variations, phenotypes, and links to genome-, phenotype-, and locus-specific resources worldwide.

Using Gene

- Gene Quick Start
- FAQ
- Download/FTP
- RefSeq Mailing List
- Gene News
- Factsheet

Gene Tools

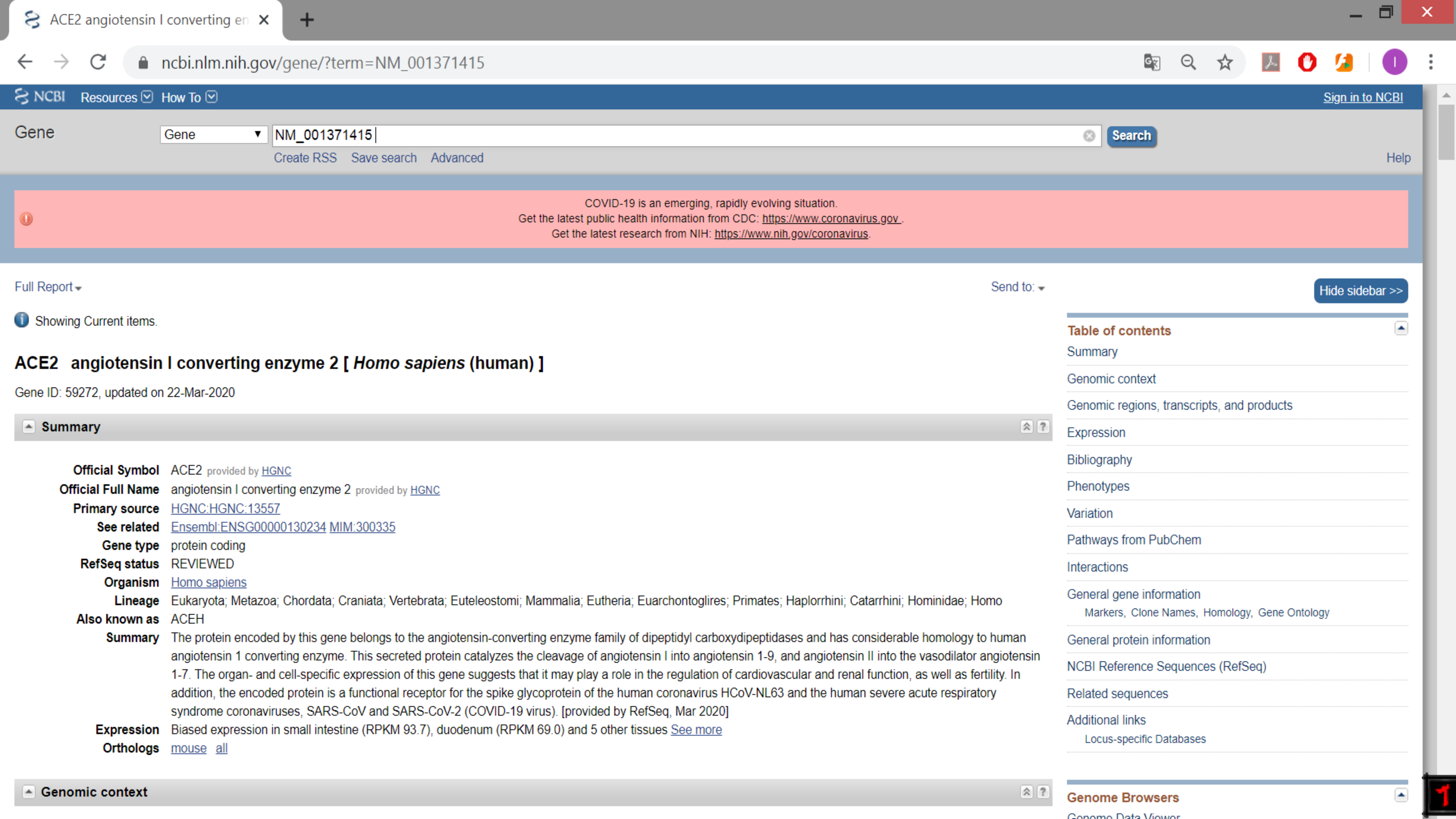
- Submit GeneRIFs
- Submit Correction
- Statistics
- BLAST
- Genome Workbench
- Splign

Other Resources

- OMIM
- RefSeq
- RefSeqGene
- Protein Clusters

Representative queries

Find genes by...	Search text
free text	human muscular dystrophy
chromosome and symbol	(II[chr] OR 2[chr]) AND adh*[sym]
	rop] AND transporter[title] AND ("Drosophila melanogaster"[orgn] OR "Mus musculus"[orgn])



COVID-19 is an emerging, rapidly evolving situation.
Get the latest public health information from CDC: https://www.coronavirus.gov .
Get the latest research from NIH: https://www.nih.gov/coronavirus.

Showing Current items.

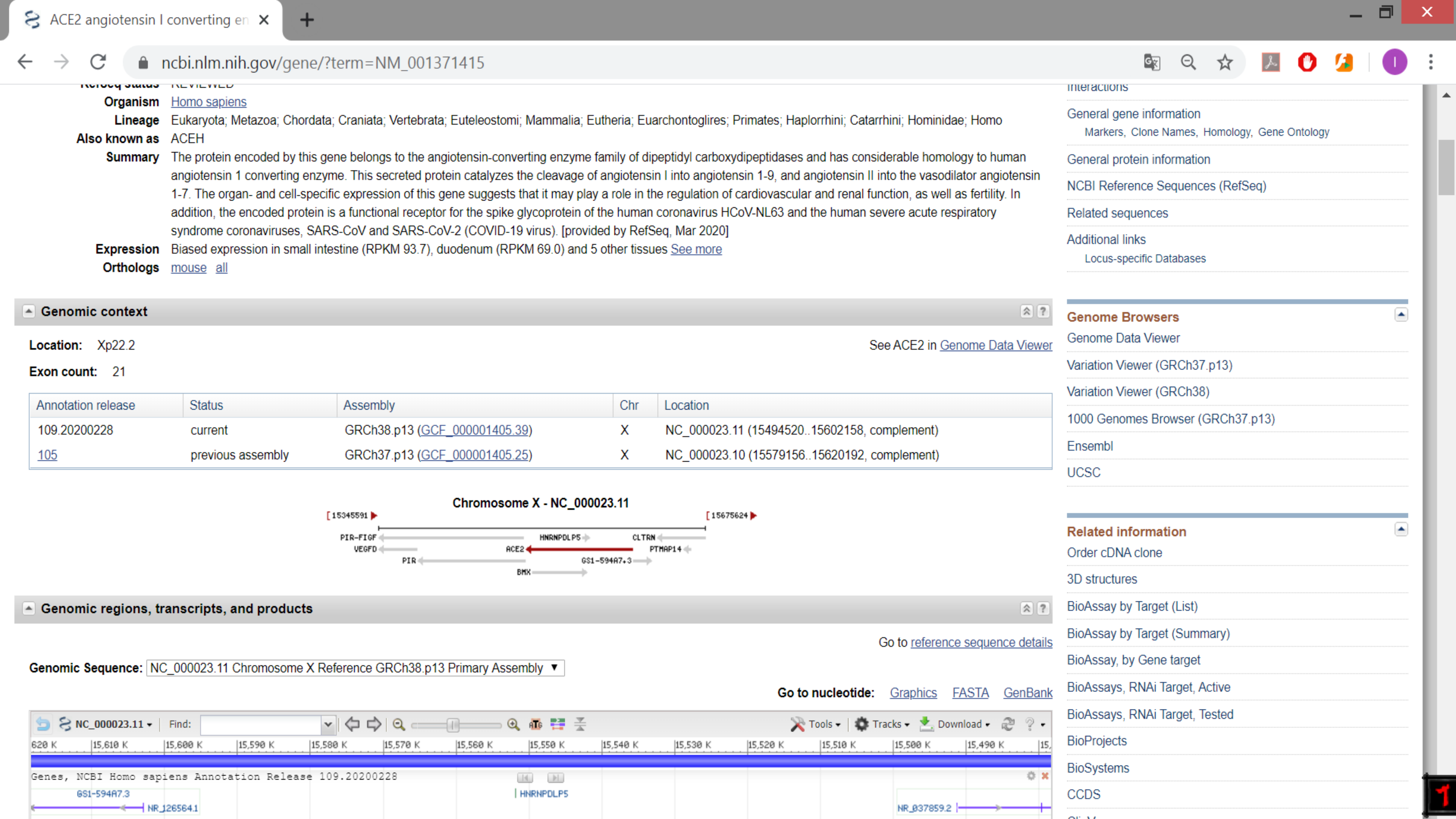
ACE2 angiotensin I converting enzyme 2 [*Homo sapiens* (human)]

Gene ID: 59272, updated on 22-Mar-2020

Summary

Official Symbol	ACE2 provided by HGNC
Official Full Name	angiotensin I converting enzyme 2 provided by HGNC
Primary source	HGNC:HGNC:13557
See related	Ensembl:ENSG00000130234 MIM:300335
Gene type	protein coding
RefSeq status	REVIEWED
Organism	Homo sapiens
Lineage	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo
Also known as	ACEH
Summary	The protein encoded by this gene belongs to the angiotensin-converting enzyme family of dipeptidyl carboxydipeptidases and has considerable homology to human angiotensin 1 converting enzyme. This secreted protein catalyzes the cleavage of angiotensin I into angiotensin 1-9, and angiotensin II into the vasodilator angiotensin 1-7. The organ- and cell-specific expression of this gene suggests that it may play a role in the regulation of cardiovascular and renal function, as well as fertility. In addition, the encoded protein is a functional receptor for the spike glycoprotein of the human coronavirus HCoV-NL63 and the human severe acute respiratory syndrome coronaviruses, SARS-CoV and SARS-CoV-2 (COVID-19 virus). [provided by RefSeq, Mar 2020]
Expression	Biased expression in small intestine (RPKM 93.7), duodenum (RPKM 69.0) and 5 other tissues See more
Orthologs	mouse all

Genomic context



Organism Homo sapiens

Lineage Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo

Also known as ACEH

Summary The protein encoded by this gene belongs to the angiotensin-converting enzyme family of dipeptidyl carboxydipeptidases and has considerable homology to human angiotensin 1 converting enzyme. This secreted protein catalyzes the cleavage of angiotensin I into angiotensin 1-9, and angiotensin II into the vasodilator angiotensin 1-7. The organ- and cell-specific expression of this gene suggests that it may play a role in the regulation of cardiovascular and renal function, as well as fertility. In addition, the encoded protein is a functional receptor for the spike glycoprotein of the human coronavirus HCoV-NL63 and the human severe acute respiratory syndrome coronaviruses, SARS-CoV and SARS-CoV-2 (COVID-19 virus). [provided by RefSeq, Mar 2020]

Expression Biased expression in small intestine (RPKM 93.7), duodenum (RPKM 69.0) and 5 other tissues See more

Orthologs mouse all

Genomic context

Location: Xp22.2

See ACE2 in Genome Data Viewer

Exon count: 21

Annotation release	Status	Assembly	Chr	Location
109.20200228	current	GRCh38.p13 (GCF_000001405.39)	X	NC_000023.11 (15494520..15602158, complement)
105	previous assembly	GRCh37.p13 (GCF_000001405.25)	X	NC_000023.10 (15579156..15620192, complement)

Chromosome X - NC_000023.11

Genomic regions, transcripts, and products

Go to reference sequence details

Genomic Sequence: NC_000023.11 Chromosome X Reference GRCh38.p13 Primary Assembly

Go to nucleotide: Graphics FASTA GenBank

Interactions

General gene information

Markers, Clone Names, Homology, Gene Ontology

General protein information

NCBI Reference Sequences (RefSeq)

Related sequences

Additional links

Locus-specific Databases

Genome Browsers

Genome Data Viewer

Variation Viewer (GRCh37.p13)

Variation Viewer (GRCh38)

1000 Genomes Browser (GRCh37.p13)

Ensembl

UCSC

Related information

Order cDNA clone

3D structures

BioAssay by Target (List)

BioAssay by Target (Summary)

BioAssay, by Gene target

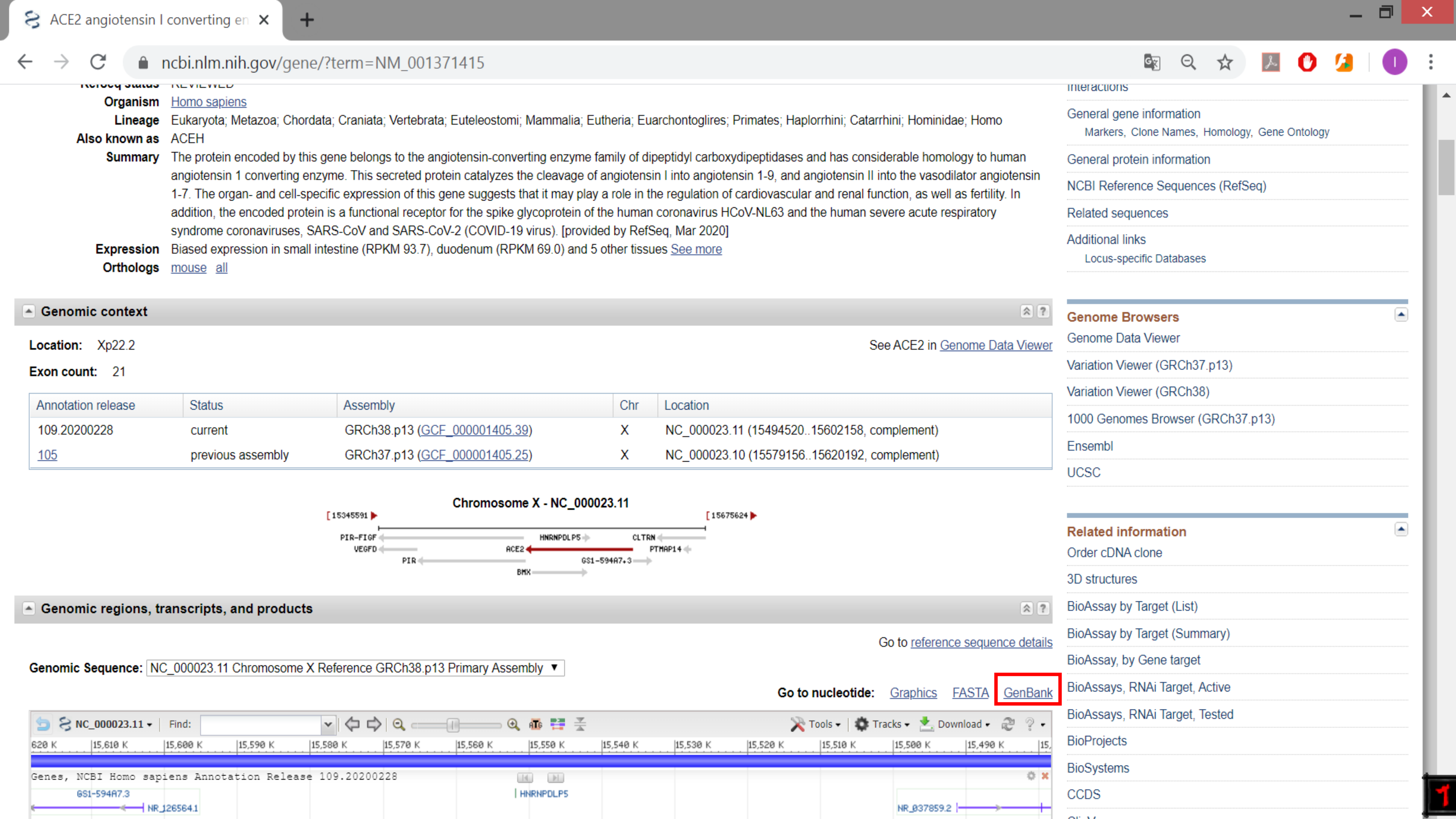
BioAssays, RNAi Target, Active

BioAssays, RNAi Target, Tested

BioProjects

BioSystems

CCDS



Organism Homo sapiens

Lineage Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo

Also known as ACEH

Summary The protein encoded by this gene belongs to the angiotensin-converting enzyme family of dipeptidyl carboxydipeptidases and has considerable homology to human angiotensin 1 converting enzyme. This secreted protein catalyzes the cleavage of angiotensin I into angiotensin 1-9, and angiotensin II into the vasodilator angiotensin 1-7. The organ- and cell-specific expression of this gene suggests that it may play a role in the regulation of cardiovascular and renal function, as well as fertility. In addition, the encoded protein is a functional receptor for the spike glycoprotein of the human coronavirus HCoV-NL63 and the human severe acute respiratory syndrome coronaviruses, SARS-CoV and SARS-CoV-2 (COVID-19 virus). [provided by RefSeq, Mar 2020]

Expression Biased expression in small intestine (RPKM 93.7), duodenum (RPKM 69.0) and 5 other tissues See more

Orthologs mouse all

Genomic context

Location: Xp22.2

See ACE2 in Genome Data Viewer

Exon count: 21

Annotation release	Status	Assembly	Chr	Location
109.20200228	current	GRCh38.p13 (GCF_000001405.39)	X	NC_000023.11 (15494520..15602158, complement)
105	previous assembly	GRCh37.p13 (GCF_000001405.25)	X	NC_000023.10 (15579156..15620192, complement)

Chromosome X - NC_000023.11

Genomic regions, transcripts, and products

Go to reference sequence details

Genomic Sequence: NC_000023.11 Chromosome X Reference GRCh38.p13 Primary Assembly

Go to nucleotide: Graphics FASTA GenBank

Interactions

General gene information

Markers, Clone Names, Homology, Gene Ontology

General protein information

NCBI Reference Sequences (RefSeq)

Related sequences

Additional links

Locus-specific Databases

Genome Browsers

Genome Data Viewer

Variation Viewer (GRCh37.p13)

Variation Viewer (GRCh38)

1000 Genomes Browser (GRCh37.p13)

Ensembl

UCSC

Related information

Order cDNA clone

3D structures

BioAssay by Target (List)

BioAssay by Target (Summary)

BioAssay, by Gene target

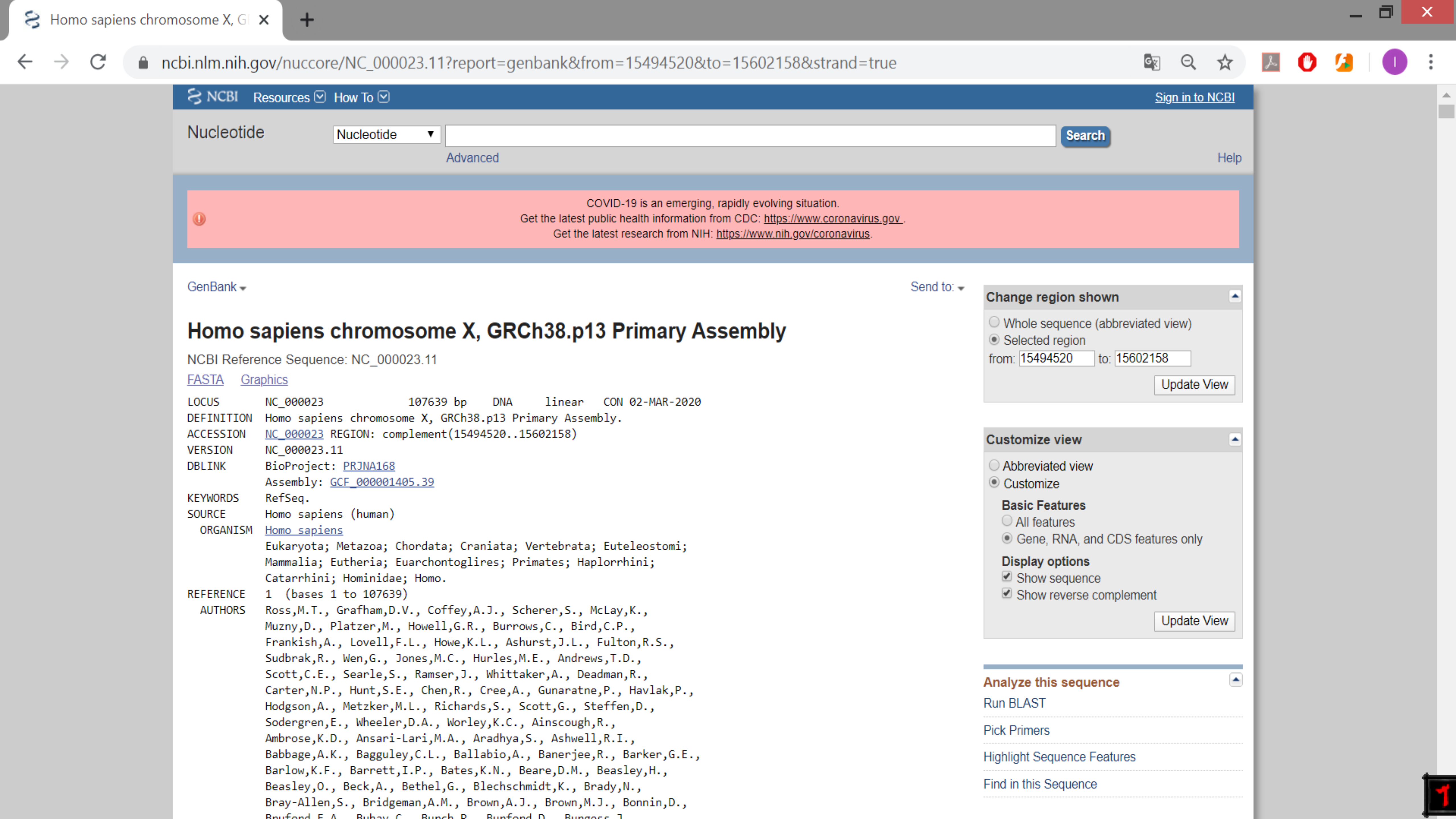
BioAssays, RNAi Target, Active

BioAssays, RNAi Target, Tested

BioProjects

BioSystems

CCDS



COVID-19 is an emerging, rapidly evolving situation.
Get the latest public health information from CDC: https://www.coronavirus.gov .
Get the latest research from NIH: https://www.nih.gov/coronavirus.

Homo sapiens chromosome X, GRCh38.p13 Primary Assembly

NCBI Reference Sequence: NC_000023.11

```
LOCUS       NC_000023          107639 bp    DNA     linear   CON 02-MAR-2020
DEFINITION  Homo sapiens chromosome X, GRCh38.p13 Primary Assembly.
ACCESSION   NC_000023 REGION: complement(15494520..15602158)
VERSION     NC_000023.11
DBLINK      BioProject: PRJNA168
            Assembly: GCF_000001405.39
KEYWORDS    RefSeq.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;
            Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 107639)
  AUTHORS   Ross,M.T., Grafham,D.V., Coffey,A.J., Scherer,S., McLay,K.,
            Muzny,D., Platzer,M., Howell,G.R., Burrows,C., Bird,C.P.,
            Frankish,A., Lovell,F.L., Howe,K.L., Ashurst,J.L., Fulton,R.S.,
            Sudbrak,R., Wen,G., Jones,M.C., Hurles,M.E., Andrews,T.D.,
            Scott,C.E., Searle,S., Ramser,J., Whittaker,A., Deadman,R.,
            Carter,N.P., Hunt,S.E., Chen,R., Cree,A., Gunaratne,P., Havlak,P.,
            Hodgson,A., Metzker,M.L., Richards,S., Scott,G., Steffen,D.,
            Sodergren,E., Wheeler,D.A., Worley,K.C., Ainscough,R.,
            Ambrose,K.D., Ansari-Lari,M.A., Aradhya,S., Ashwell,R.I.,
            Babbage,A.K., Bagguley,C.L., Ballabio,A., Banerjee,R., Barker,G.E.,
            Barlow,K.F., Barrett,I.P., Bates,K.N., Beare,D.M., Beasley,H.,
            Beasley,O., Beck,A., Bethel,G., Blechschmidt,K., Brady,N.,
            Bray-Allen,S., Bridgeman,A.M., Brown,A.J., Brown,M.J., Bonnin,D.,
```

Change region shown

○ Whole sequence (abbreviated view)
● Selected region
from: 15494520 to: 15602158

Customize view

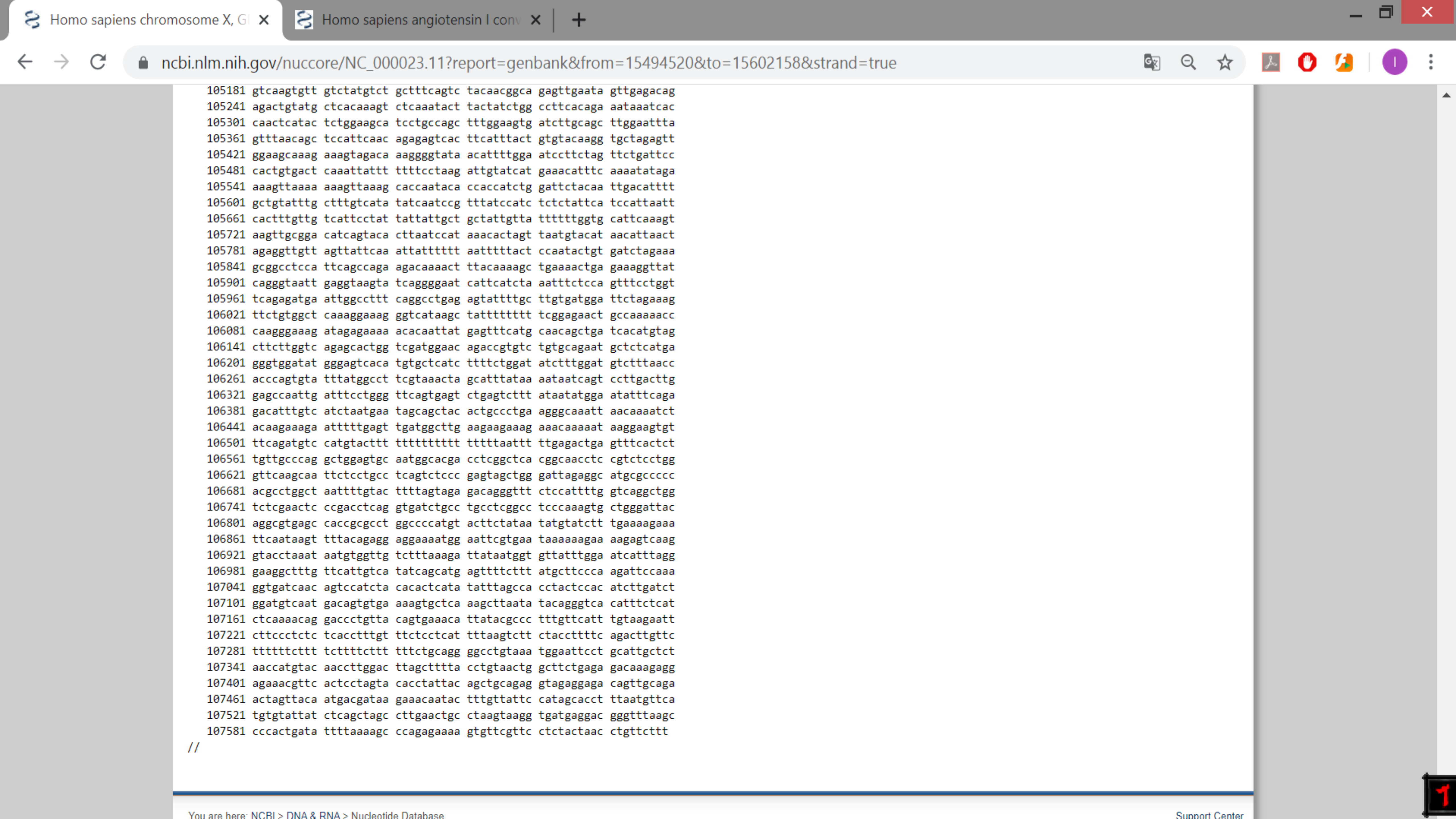
○ Abbreviated view
● Customize

Basic Features

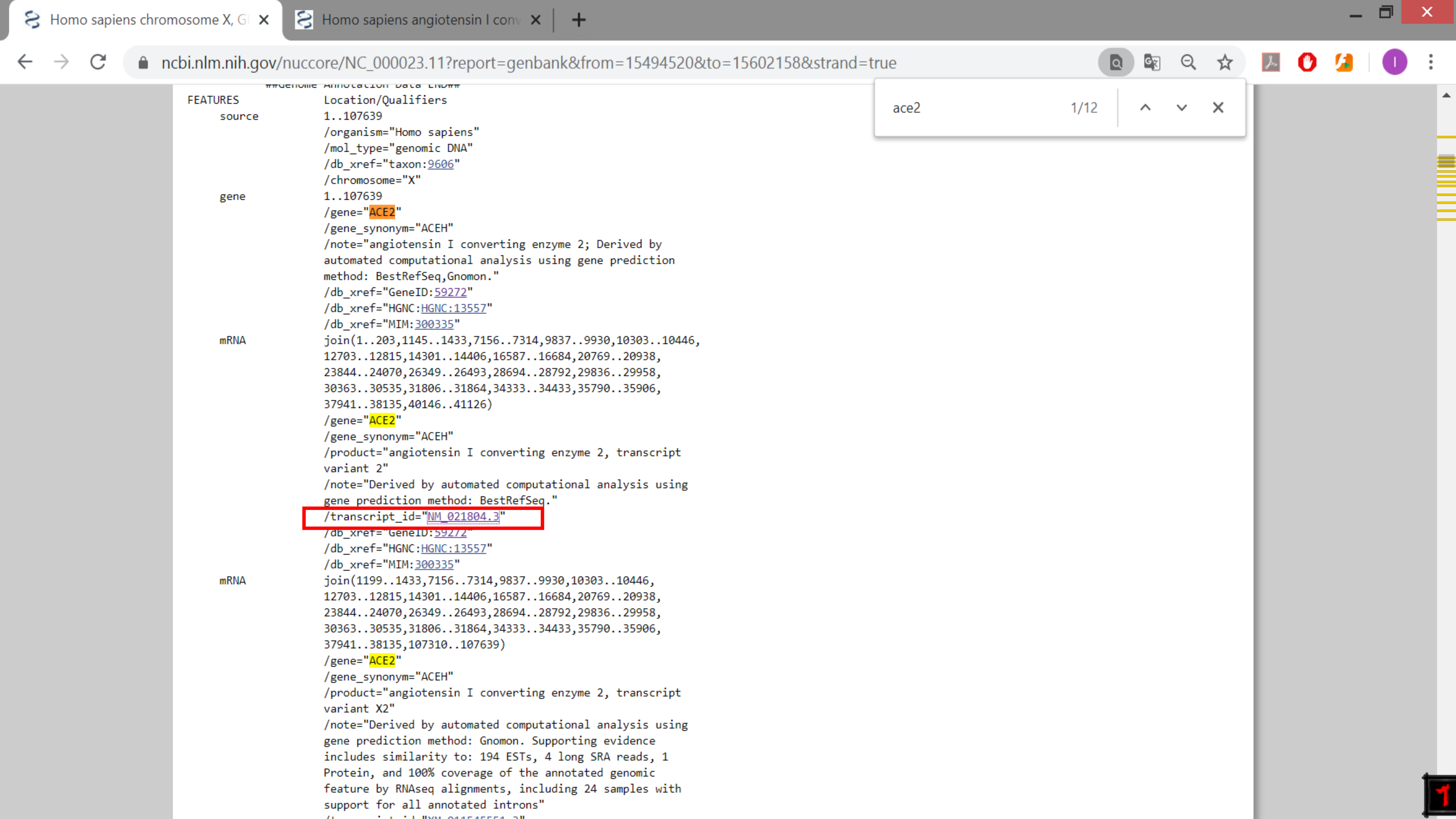
○ All features
● Gene, RNA, and CDS features only

Display options

☑ Show sequence
☑ Show reverse complement

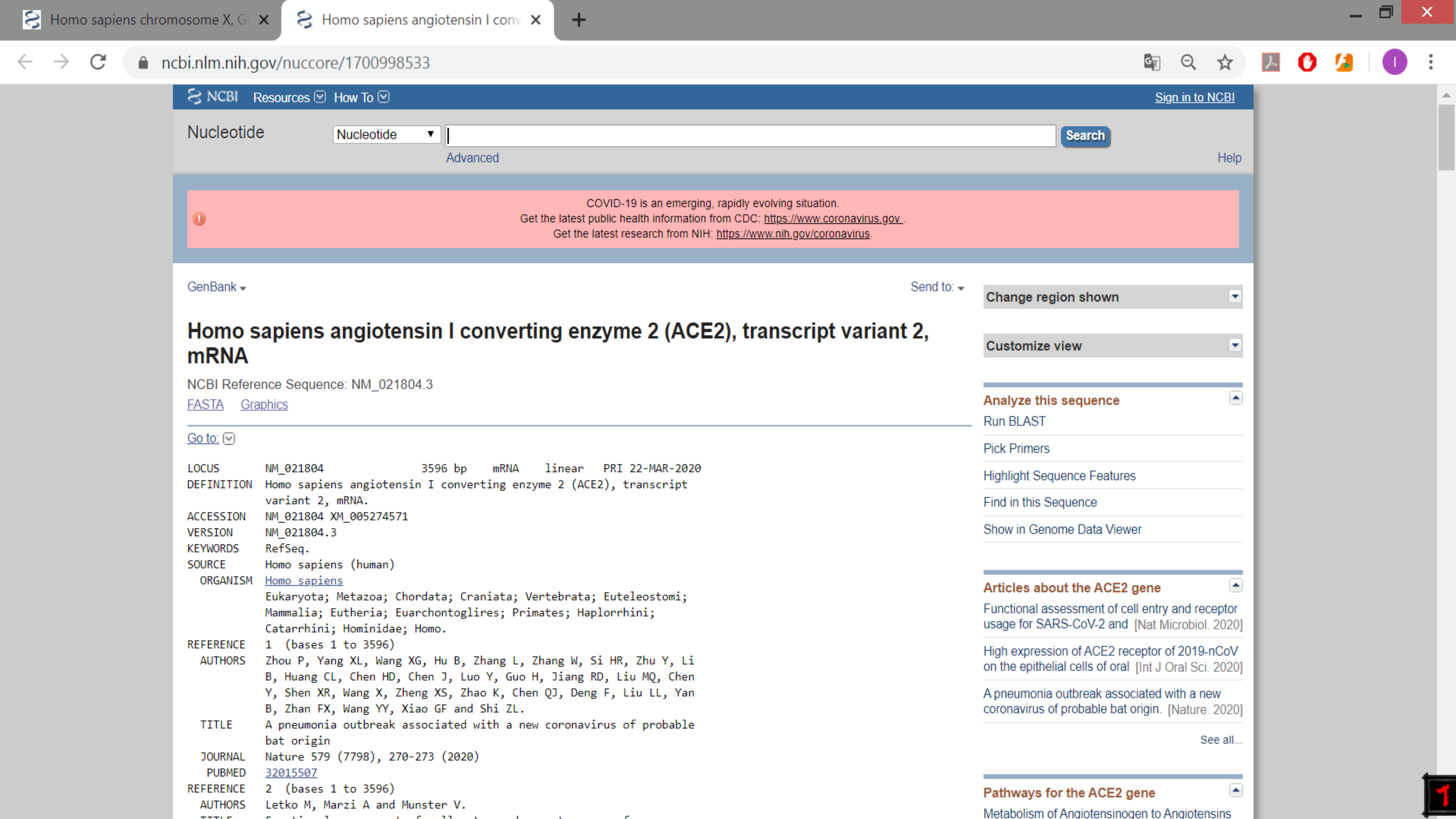


```
   105181 gtcaagtgtt gtctatgtct gctttcagtc tacaacggca gagttgaata gttgagacag
   105241 agactgtatg ctcacaaagt ctcaaatact tactatctgg ccttcacaga aataaatcac
   105301 caactcatac tctggaagca tcctgccagc tttggaagtg atcttgcagc ttggaattta
   105361 gtttaacagc tccattcaac agagagtcac ttcatttact gtgtacaagg tgctagagtt
   105421 ggaagcaaag aaagtagaca aaggggtata acattttgga atccttctag ttctgattcc
   105481 cactgtgact caaattattt ttttcctaag attgtatcat gaaacatttc aaaatataga
   105541 aaagttaaaa aaagttaaag caccaataca ccaccatctg gattctacaa ttgacatttt
   105601 gctgtatttg ctttgtcata tatcaatccg tttatccatc tctctattca tccattaatt
   105661 cactttgttg tcattcctat tattattgct gctattgtta ttttttggtg cattcaaagt
   105721 aagttgcgga catcagtaca cttaatccat aaacactagt taatgtacat aacattaact
   105781 agaggttgtt agttattcaa attatttttt aatttttact ccaatactgt gatctagaaa
   105841 gcggcctcca ttcagccaga agacaaaact ttacaaaagc tgaaaactga gaaaggttat
   105901 cagggtaatt gaggtaagta tcaggggaat cattcatcta aatttctcca gtttcctggt
   105961 tcagagatga attggccttt caggcctgag agtattttgc ttgtgatgga ttctagaaag
   106021 ttctgtggct caaaggaaag ggtcataagc tatttttttt tcggagaact gccaaaaacc
   106081 caagggaaag atagagaaaa acacaattat gagtttcatg caacagctga tcacatgtag
   106141 cttcttggtc agagcactgg tcgatggaac agaccgtgtc tgtgcagaat gctctcatga
   106201 gggtggatat gggagtcaca tgtgctcatc ttttctggat atctttggat gtctttaacc
   106261 acccagtgta tttatggcct tcgtaaacta gcatttataa aataatcagt ccttgacttg
   106321 gagccaattg atttcctggg ttcagtgagt ctgagtcttt ataatatgga atatttcaga
   106381 gacatttgtc atctaatgaa tagcagctac actgccctga agggcaaatt aacaaaatct
   106441 acaagaaaga atttttgagt tgatggcttg aagaagaaag aaacaaaaat aaggaagtgt
   106501 ttcagatgtc catgtacttt tttttttttt tttttaattt ttgagactga gtttcactct
   106561 tgttgcccag gctggagtgc aatggcacga cctcggctca cggcaacctc cgtctcctgg
   106621 gttcaagcaa ttctcctgcc tcagtctccc gagtagctgg gattagaggc atgcgccccc
   106681 acgcctggct aattttgtac ttttagtaga gacagggttt ctccattttg gtcaggctgg
   106741 tctcgaactc ccgacctcag gtgatctgcc tgcctcggcc tcccaaagtg ctgggattac
   106801 aggcgtgagc caccgcgcct ggccccatgt acttctataa tatgtatctt tgaaaagaaa
   106861 ttcaataagt tttacagagg aggaaaatgg aattcgtgaa taaaaaagaa aagagtcaag
   106921 gtacctaaat aatgtggttg tctttaaaga ttataatggt gttatttgga atcatttagg
   106981 gaaggctttg ttcattgtca tatcagcatg agttttcttt atgcttccca agattccaaa
   107041 ggtgatcaac agtccatcta cacactcata tatttagcca cctactccac atcttgatct
   107101 ggatgtcaat gacagtgtga aaagtgctca aagcttaata tacagggtca catttctcat
   107161 ctcaaaacag gaccctgtta cagtgaaaca ttatacgccc tttgttcatt tgtaagaatt
   107221 cttccctctc tcacctttgt ttctcctcat tttaagtctt ctaccttttc agacttgttc
   107281 tttttctttt tcttttcttt tttctgcagg ggcctgtaaa tggaattcct gcattgctct
   107341 aaccatgtac aaccttggac ttagctttta cctgtaactg gcttctgaga gacaaagagg
   107401 agaaacgttc actcctagta cacctattac agctgcagag gtagaggaga cagttgcaga
   107461 actagttaca atgacgataa gaaacaatac tttgttattc catagcacct ttaatgttca
   107521 tgtgtattat ctcagctagc cttgaactgc ctaagtaagg tgatgaggac gggtttaagc
   107581 cccactgata ttttaaaagc ccagagaaaa gtgttcgttc ctctactaac ctgttcttt
//
```



ncbi.nlm.nih.gov/nuccore/NC_000023.11?report=genbank&from=15494520&to=15602158&strand=true

```
FEATURES             Location/Qualifiers
     source          1..107639
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="X"
     gene            1..107639
                     /gene="ACE2"
                     /gene_synonym="ACEH"
                     /note="angiotensin I converting enzyme 2; Derived by
                     automated computational analysis using gene prediction
                     method: BestRefSeq,Gnomon."
                     /db_xref="GeneID:59272"
                     /db_xref="HGNC:HGNC:13557"
                     /db_xref="MIM:300335"
     mRNA            join(1..203,1145..1433,7156..7314,9837..9930,10303..10446,
                     12703..12815,14301..14406,16587..16684,20769..20938,
                     23844..24070,26349..26493,28694..28792,29836..29958,
                     30363..30535,31806..31864,34333..34433,35790..35906,
                     37941..38135,40146..41126)
                     /gene="ACE2"
                     /gene_synonym="ACEH"
                     /product="angiotensin I converting enzyme 2, transcript
                     variant 2"
                     /note="Derived by automated computational analysis using
                     gene prediction method: BestRefSeq."
                     /transcript_id="NM_021804.3"
                     /db_xref="GeneID:59272"
                     /db_xref="HGNC:HGNC:13557"
                     /db_xref="MIM:300335"
     mRNA            join(1199..1433,7156..7314,9837..9930,10303..10446,
                     12703..12815,14301..14406,16587..16684,20769..20938,
                     23844..24070,26349..26493,28694..28792,29836..29958,
                     30363..30535,31806..31864,34333..34433,35790..35906,
                     37941..38135,107310..107639)
                     /gene="ACE2"
                     /gene_synonym="ACEH"
                     /product="angiotensin I converting enzyme 2, transcript
                     variant X2"
                     /note="Derived by automated computational analysis using
                     gene prediction method: Gnomon. Supporting evidence
                     includes similarity to: 194 ESTs, 4 long SRA reads, 1
                     Protein, and 100% coverage of the annotated genomic
                     feature by RNAseq alignments, including 24 samples with
                     support for all annotated introns"
```



COVID-19 is an emerging, rapidly evolving situation.
Get the latest public health information from CDC: https://www.coronavirus.gov .
Get the latest research from NIH: https://www.nih.gov/coronavirus.

Homo sapiens angiotensin I converting enzyme 2 (ACE2), transcript variant 2, mRNA

NCBI Reference Sequence: NM_021804.3

```
LOCUS       NM_021804               3596 bp    mRNA    linear   PRI 22-MAR-2020
DEFINITION  Homo sapiens angiotensin I converting enzyme 2 (ACE2), transcript
            variant 2, mRNA.
ACCESSION   NM_021804 XM_005274571
VERSION     NM_021804.3
KEYWORDS    RefSeq.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;
            Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3596)
  AUTHORS   Zhou P, Yang XL, Wang XG, Hu B, Zhang L, Zhang W, Si HR, Zhu Y, Li
            B, Huang CL, Chen HD, Chen J, Luo Y, Guo H, Jiang RD, Liu MQ, Chen
            Y, Shen XR, Wang X, Zheng XS, Zhao K, Chen QJ, Deng F, Liu LL, Yan
            B, Zhan FX, Wang YY, Xiao GF and Shi ZL.
  TITLE     A pneumonia outbreak associated with a new coronavirus of probable
            bat origin
  JOURNAL   Nature 579 (7798), 270-273 (2020)
   PUBMED   32015507
REFERENCE   2  (bases 1 to 3596)
  AUTHORS   Letko M, Marzi A and Munster V.
```

Analyze this sequence

Articles about the ACE2 gene

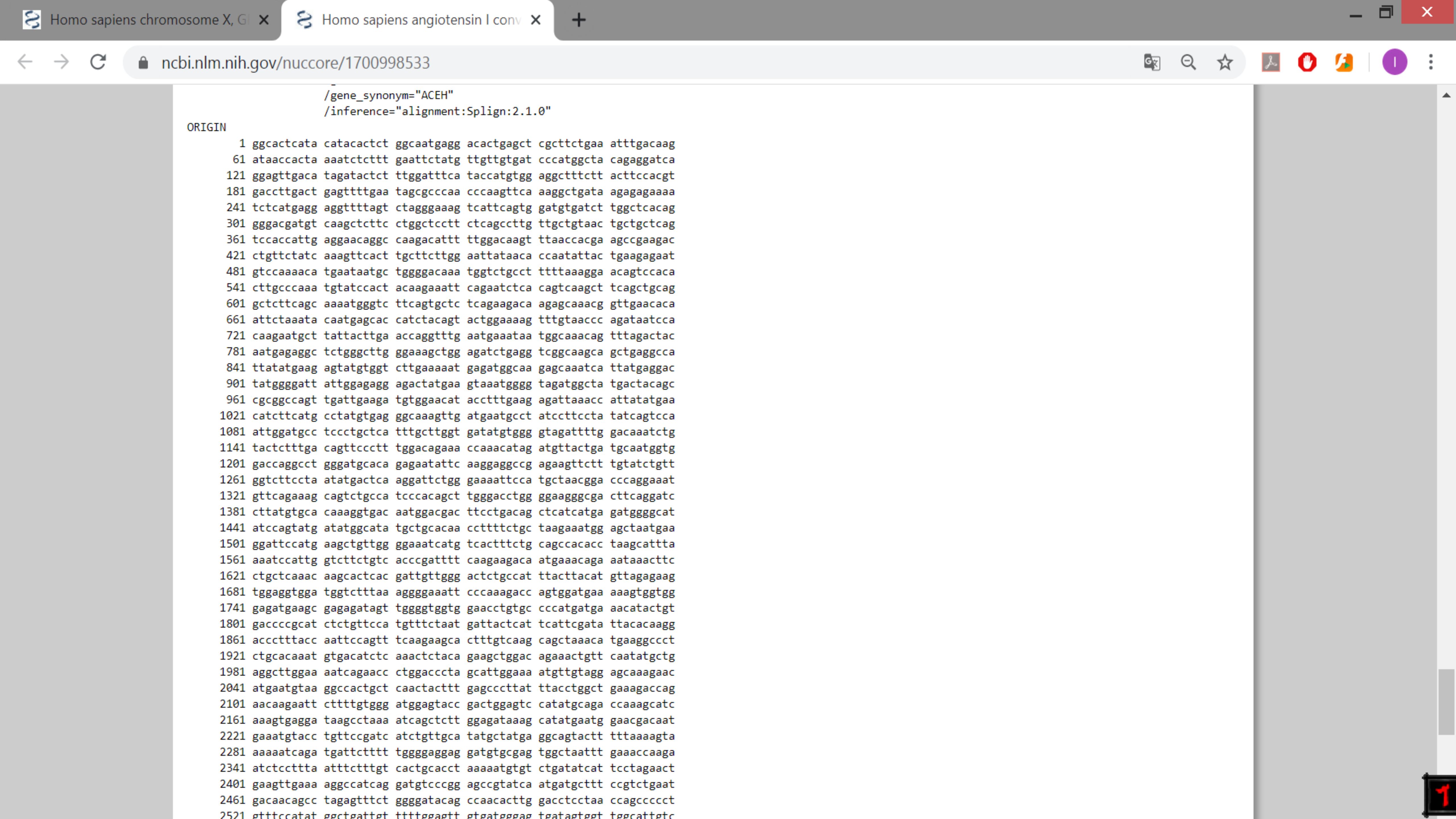
Functional assessment of cell entry and receptor usage for SARS-CoV-2 and [Nat Microbiol. 2020]

High expression of ACE2 receptor of 2019-nCoV on the epithelial cells of oral [Int J Oral Sci. 2020]

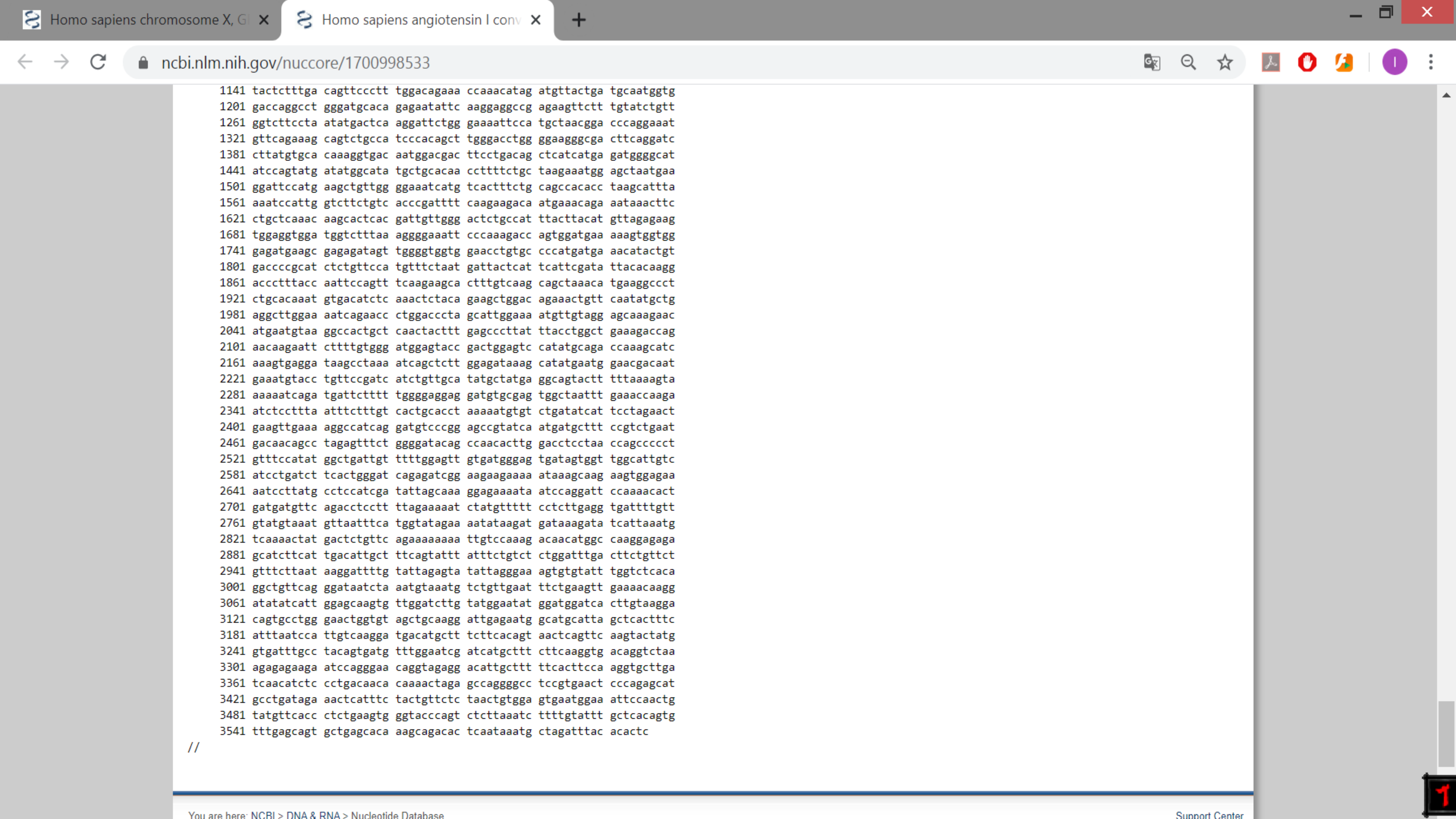
A pneumonia outbreak associated with a new coronavirus of probable bat origin. [Nature. 2020]

Pathways for the ACE2 gene

Metabolism of Angiotensinogen to Angiotensins



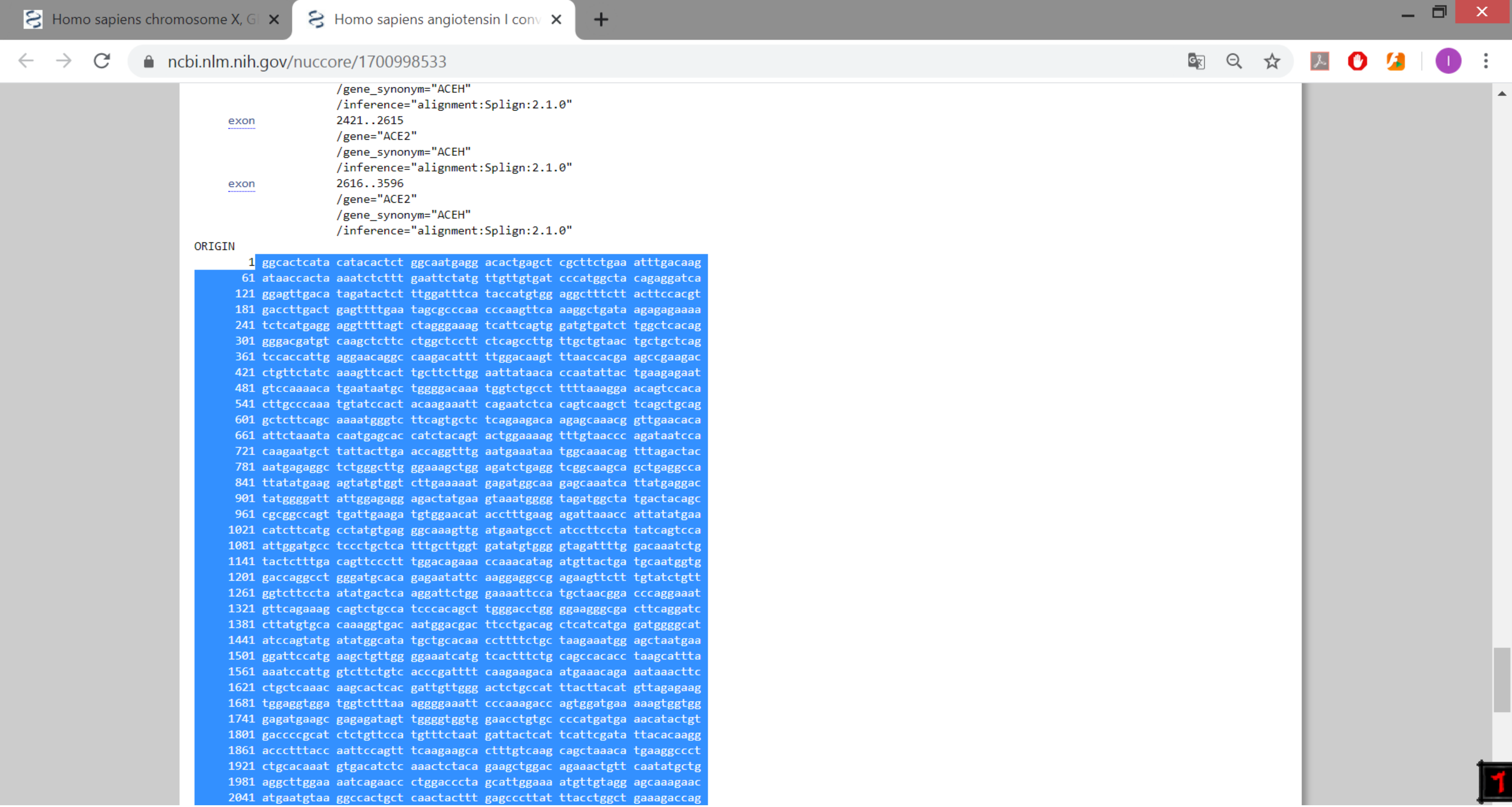
```
                     /gene_synonym="ACEH"
                     /inference="alignment:Splign:2.1.0"
ORIGIN
        1 ggcactcata catacactct ggcaatgagg acactgagct cgcttctgaa atttgacaag
       61 ataaccacta aaatctcttt gaattctatg ttgttgtgat cccatggcta cagaggatca
      121 ggagttgaca tagatactct ttggatttca taccatgtgg aggctttctt acttccacgt
      181 gaccttgact gagttttgaa tagcgcccaa cccaagttca aaggctgata agagagaaaa
      241 tctcatgagg aggttttagt ctagggaaag tcattcagtg gatgtgatct tggctcacag
      301 gggacgatgt caagctcttc ctggctcctt ctcagccttg ttgctgtaac tgctgctcag
      361 tccaccattg aggaacaggc caagacattt ttggacaagt ttaaccacga agccgaagac
      421 ctgttctatc aaagttcact tgcttcttgg aattataaca ccaatattac tgaagagaat
      481 gtccaaaaca tgaataatgc tggggacaaa tggtctgcct ttttaaagga acagtccaca
      541 cttgcccaaa tgtatccact acaagaaatt cagaatctca cagtcaagct tcagctgcag
      601 gctcttcagc aaaatgggtc ttcagtgctc tcagaagaca agagcaaacg gttgaacaca
      661 attctaaata caatgagcac catctacagt actggaaaag tttgtaaccc agataatcca
      721 caagaatgct tattacttga accaggtttg aatgaaataa tggcaaacag tttagactac
      781 aatgagaggc tctgggcttg ggaaagctgg agatctgagg tcggcaagca gctgaggcca
      841 ttatatgaag agtatgtggt cttgaaaaat gagatggcaa gagcaaatca ttatgaggac
      901 tatggggatt attggagagg agactatgaa gtaaatgggg tagatggcta tgactacagc
      961 cgcggccagt tgattgaaga tgtggaacat acctttgaag agattaaacc attatatgaa
     1021 catcttcatg cctatgtgag ggcaaagttg atgaatgcct atccttccta tatcagtcca
     1081 attggatgcc tccctgctca tttgcttggt gatatgtggg gtagattttg gacaaatctg
     1141 tactctttga cagttccctt tggacagaaa ccaaacatag atgttactga tgcaatggtg
     1201 gaccaggcct gggatgcaca gagaatattc aaggaggccg agaagttctt tgtatctgtt
     1261 ggtcttccta atatgactca aggattctgg gaaaattcca tgctaacgga cccaggaaat
     1321 gttcagaaag cagtctgcca tcccacagct tgggacctgg ggaagggcga cttcaggatc
     1381 cttatgtgca caaaggtgac aatggacgac ttcctgacag ctcatcatga gatggggcat
     1441 atccagtatg atatggcata tgctgcacaa ccttttctgc taagaaatgg agctaatgaa
     1501 ggattccatg aagctgttgg ggaaatcatg tcactttctg cagccacacc taagcattta
     1561 aaatccattg gtcttctgtc acccgatttt caagaagaca atgaaacaga aataaacttc
     1621 ctgctcaaac aagcactcac gattgttggg actctgccat ttacttacat gttagagaag
     1681 tggaggtgga tggtctttaa aggggaaatt cccaaagacc agtggatgaa aaagtggtgg
     1741 gagatgaagc gagagatagt tggggtggtg gaacctgtgc cccatgatga aacatactgt
     1801 gaccccgcat ctctgttcca tgtttctaat gattactcat tcattcgata ttacacaagg
     1861 accctttacc aattccagtt tcaagaagca ctttgtcaag cagctaaaca tgaaggccct
     1921 ctgcacaaat gtgacatctc aaactctaca gaagctggac agaaactgtt caatatgctg
     1981 aggcttggaa aatcagaacc ctggacccta gcattggaaa atgttgtagg agcaaagaac
     2041 atgaatgtaa ggccactgct caactacttt gagcccttat ttacctggct gaaagaccag
     2101 aacaagaatt cttttgtggg atggagtacc gactggagtc catatgcaga ccaaagcatc
     2161 aaagtgagga taagcctaaa atcagctctt ggagataaag catatgaatg gaacgacaat
     2221 gaaatgtacc tgttccgatc atctgttgca tatgctatga ggcagtactt tttaaaagta
     2281 aaaaatcaga tgattctttt tggggaggag gatgtgcgag tggctaattt gaaaccaaga
     2341 atctccttta atttctttgt cactgcacct aaaaatgtgt ctgatatcat tcctagaact
     2401 gaagttgaaa aggccatcag gatgtcccgg agccgtatca atgatgcttt ccgtctgaat
     2461 gacaacagcc tagagtttct ggggatacag ccaacacttg gacctcctaa ccagcccct
     2521 gtttccatat ggctgattgt ttttggagtt gtgatgggag tgatagtggt tggcattgtc
```



```
     1141 tactctttga cagttccctt tggacagaaa ccaaacatag atgttactga tgcaatggtg
     1201 gaccaggcct gggatgcaca gagaatattc aaggaggccg agaagttctt tgtatctgtt
     1261 ggtcttccta atatgactca aggattctgg gaaaattcca tgctaacgga cccaggaaat
     1321 gttcagaaag cagtctgcca tcccacagct tgggacctgg ggaagggcga cttcaggatc
     1381 cttatgtgca caaaggtgac aatggacgac ttcctgacag ctcatcatga gatggggcat
     1441 atccagtatg atatggcata tgctgcacaa ccttttctgc taagaaatgg agctaatgaa
     1501 ggattccatg aagctgttgg ggaaatcatg tcactttctg cagccacacc taagcattta
     1561 aaatccattg gtcttctgtc acccgatttt caagaagaca atgaaacaga aataaacttc
     1621 ctgctcaaac aagcactcac gattgttggg actctgccat ttacttacat gttagagaag
     1681 tggaggtgga tggtctttaa aggggaaatt cccaaagacc agtggatgaa aaagtggtgg
     1741 gagatgaagc gagagatagt tggggtggtg gaacctgtgc cccatgatga aacatactgt
     1801 gaccccgcat ctctgttcca tgtttctaat gattactcat tcattcgata ttacacaagg
     1861 accctttacc aattccagtt tcaagaagca ctttgtcaag cagctaaaca tgaaggccct
     1921 ctgcacaaat gtgacatctc aaactctaca gaagctggac agaaactgtt caatatgctg
     1981 aggcttggaa aatcagaacc ctggacccta gcattggaaa atgttgtagg agcaaagaac
     2041 atgaatgtaa ggccactgct caactacttt gagcccttat ttacctggct gaaagaccag
     2101 aacaagaatt cttttgtggg atggagtacc gactggagtc catatgcaga ccaaagcatc
     2161 aaagtgagga taagcctaaa atcagctctt ggagataaag catatgaatg gaacgacaat
     2221 gaaatgtacc tgttccgatc atctgttgca tatgctatga ggcagtactt tttaaaagta
     2281 aaaaatcaga tgattctttt tggggaggag gatgtgcgag tggctaattt gaaaccaaga
     2341 atctccttta atttctttgt cactgcacct aaaaatgtgt ctgatatcat tcctagaact
     2401 gaagttgaaa aggccatcag gatgtcccgg agccgtatca atgatgcttt ccgtctgaat
     2461 gacaacagcc tagagtttct ggggatacag ccaacacttg gacctcctaa ccagcccccct
     2521 gtttccatat ggctgattgt ttttggagtt gtgatgggag tgatagtggt tggcattgtc
     2581 atcctgatct tcactgggat cagagatcgg aagaagaaaa ataaagcaag aagtggagaa
     2641 aatccttatg cctccatcga tattagcaaa ggagaaaata atccaggatt ccaaaacact
     2701 gatgatgttc agacctcctt ttagaaaaat ctatgttttt cctcttgagg tgattttgtt
     2761 gtatgtaaat gttaatttca tggtatagaa aatataagat gataaagata tcattaaatg
     2821 tcaaaactat gactctgttc agaaaaaaaa ttgtccaaag acaacatggc caaggagaga
     2881 gcatcttcat tgacattgct ttcagtattt atttctgtct ctggatttga cttctgttct
     2941 gtttcttaat aaggattttg tattagagta tattagggaa agtgtgtatt tggtctcaca
     3001 ggctgttcag ggataatcta aatgtaaatg tctgttgaat ttctgaagtt gaaaacaagg
     3061 atatatcatt ggagcaagtg ttggatcttg tatggaatat ggatggatca cttgtaagga
     3121 cagtgcctgg gaactggtgt agctgcaagg attgagaatg gcatgcatta gctcactttc
     3181 atttaatcca ttgtcaagga tgacatgctt tcttcacagt aactcagttc aagtactatg
     3241 gtgatttgcc tacagtgatg tttggaatcg atcatgcttt cttcaaggtg acaggtctaa
     3301 agagagaaga atccagggaa caggtagagg acattgcttt ttcacttcca aggtgcttga
     3361 tcaacatctc cctgacaaca caaaactaga gccaggggcc tccgtgaact cccagagcat
     3421 gcctgataga aactcatttc tactgttctc taactgtgga gtgaatggaa attccaactg
     3481 tatgttcacc ctctgaagtg ggtacccagt ctcttaaatc ttttgtattt gctcacagtg
     3541 tttgagcagt gctgagcaca aagcagacac tcaataaatg ctagatttac acactc
//
```

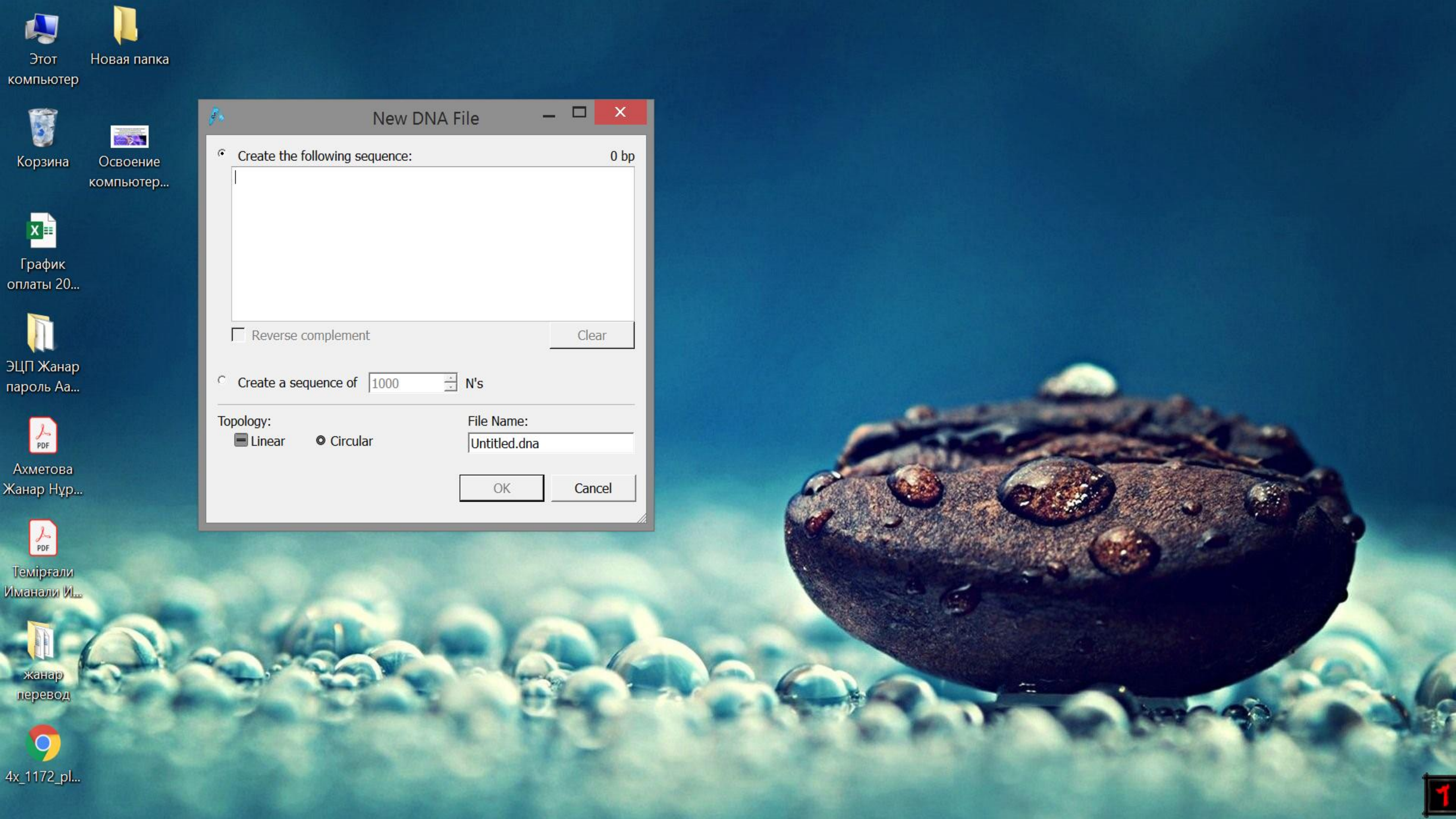
You are here: NCBI > DNA & RNA > Nucleotide Database

Support Center



```
                     /gene_synonym="ACEH"
                     /inference="alignment:Splign:2.1.0"
     exon            2421..2615
                     /gene="ACE2"
                     /gene_synonym="ACEH"
                     /inference="alignment:Splign:2.1.0"
     exon            2616..3596
                     /gene="ACE2"
                     /gene_synonym="ACEH"
                     /inference="alignment:Splign:2.1.0"
ORIGIN
        1 ggcactcata catacactct ggcaatgagg acactgagct cgcttctgaa atttgacaag
       61 ataaccacta aaatctcttt gaattctatg ttgttgtgat cccatggcta cagaggatca
      121 ggagttgaca tagatactct ttggatttca taccatgtgg aggctttctt acttccacgt
      181 gaccttgact gagttttgaa tagcgcccaa cccaagttca aaggctgata agagagaaaa
      241 tctcatgagg aggttttagt ctagggaaag tcattcagtg gatgtgatct tggctcacag
      301 gggacgatgt caagctcttc ctggctcctt ctcagccttg ttgctgtaac tgctgctcag
      361 tccaccattg aggaacaggc caagacattt ttggacaagt ttaaccacga agccgaagac
      421 ctgttctatc aaagttcact tgcttcttgg aattataaca ccaatattac tgaagagaat
      481 gtccaaaaca tgaataatgc tggggacaaa tggtctgcct ttttaaagga acagtccaca
      541 cttgcccaaa tgtatccact acaagaaatt cagaatctca cagtcaagct tcagctgcag
      601 gctcttcagc aaaatgggtc ttcagtgctc tcagaagaca agagcaaacg gttgaacaca
      661 attctaaata caatgagcac catctacagt actggaaaag tttgtaaccc agataatcca
      721 caagaatgct tattacttga accaggtttg aatgaaataa tggcaaacag tttagactac
      781 aatgagaggc tctgggcttg ggaaagctgg agatctgagg tcggcaagca gctgaggcca
      841 ttatatgaag agtatgtggt cttgaaaaat gagatggcaa gagcaaatca ttatgaggac
      901 tatggggatt attggagagg agactatgaa gtaaatgggg tagatggcta tgactacagc
      961 cgcggccagt tgattgaaga tgtggaacat acctttgaag agattaaacc attatatgaa
     1021 catcttcatg cctatgtgag ggcaaagttg atgaatgcct atccttccta tatcagtcca
     1081 attggatgcc tccctgctca tttgcttggt gatatgtggg gtagattttg gacaaatctg
     1141 tactctttga cagttccctt tggacagaaa ccaaacatag atgttactga tgcaatggtg
     1201 gaccaggcct gggatgcaca gagaatattc aaggaggccg agaagttctt tgtatctgtt
     1261 ggtcttccta atatgactca aggattctgg gaaaattcca tgctaacgga cccaggaaat
     1321 gttcagaaag cagtctgcca tcccacagct tgggacctgg ggaagggcga cttcaggatc
     1381 cttatgtgca caaaggtgac aatggacgac ttcctgacag ctcatcatga gatggggcat
     1441 atccagtatg atatggcata tgctgcacaa ccttttctgc taagaaatgg agctaatgaa
     1501 ggattccatg aagctgttgg ggaaatcatg tcactttctg cagccacacc taagcattta
     1561 aaatccattg gtcttctgtc acccgatttt caagaagaca atgaaacaga aataaacttc
     1621 ctgctcaaac aagcactcac gattgttggg actctgccat ttacttacat gttagagaag
     1681 tggaggtgga tggtctttaa aggggaaatt cccaaagacc agtggatgaa aaagtggtgg
     1741 gagatgaagc gagagatagt tggggtggtg gaacctgtgc cccatgatga aacatactgt
     1801 gaccccgcat ctctgttcca tgtttctaat gattactcat tcattcgata ttacacaagg
     1861 accctttacc aattccagtt tcaagaagca ctttgtcaag cagctaaaca tgaaggccct
     1921 ctgcacaaat gtgacatctc aaactctaca gaagctggac agaaactgtt caatatgctg
     1981 aggcttggaa aatcagaacc ctggacccta gcattggaaa atgttgtagg agcaaagaac
     2041 atgaatgtaa ggccactgct caactacttt gagcccttat ttacctggct gaaagaccag
```


Этот компьютер
Новая папка
Корзина
График оплаты 20...
ЭЦП Жанар пароль Аа...
PDF
Ахметова Жанар Нұр...
PDF
Теміргали Иманали И...
жанар перевод
4x_1172_pl...
SnapGene
SnapGene™ 1.1.3
New DNA File...
Open...
Open Recent
Import from GenBank...
Reference
Help
Exit



New DNA File

● Create the following sequence: 0 bp

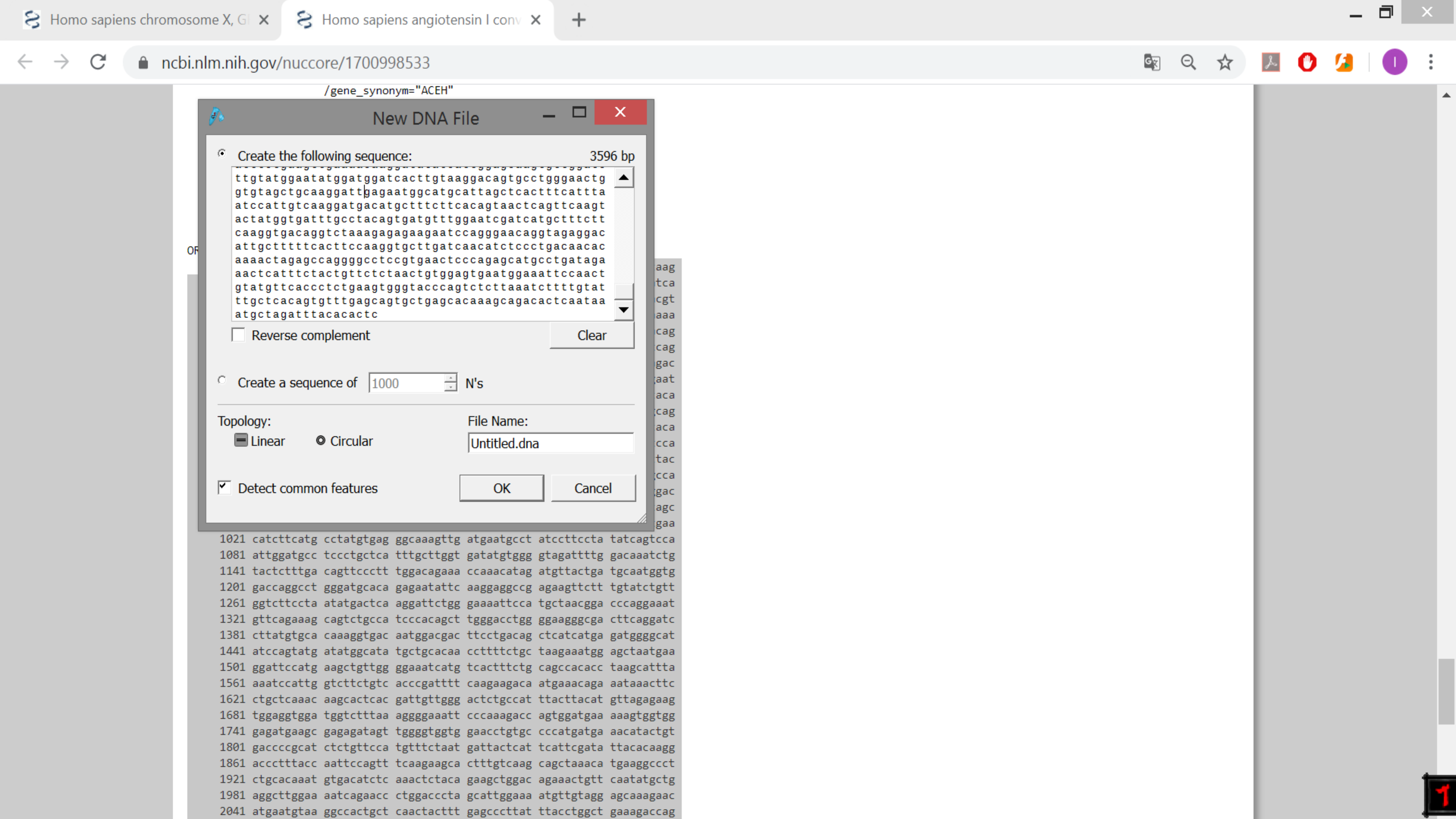
☐ Reverse complement

○ Create a sequence of 1000 N's

Topology:

○ Linear ● Circular

File Name: Untitled.dna



```
/gene_synonym="ACEH"
```

New DNA File

● Create the following sequence: 3596 bp

```
ttgtatggaatatggatggatcacttgtaaggacagtgcctgggaactg
gtgtagctgcaaggattgagaatggcatgcattagctcactttcattta
atccattgtcaaggatgacatgctttcttcacagtaactcagttcaagt
actatggtgatttgcctacagtgatgtttggaatcgatcatgctttctt
caaggtgacaggtctaaagagagaagaatccagggaacaggtagaggac
attgcttttcacttccaaggtgcttgatcaacatctccctgacaacac
aaaactagagccaggggcctccgtgaactcccagagcatgcctgataga
aactcatttctactgttctctaactgtggagtgaatggaaattccaact
gtatgttcaccctctgaagtgggtacccagtctcttaaatcttttgtat
ttgctcacagtgtttgagcagtgctgagcacaaagcagacactcaataa
atgctagatttacacactc
```

☐ Reverse complement

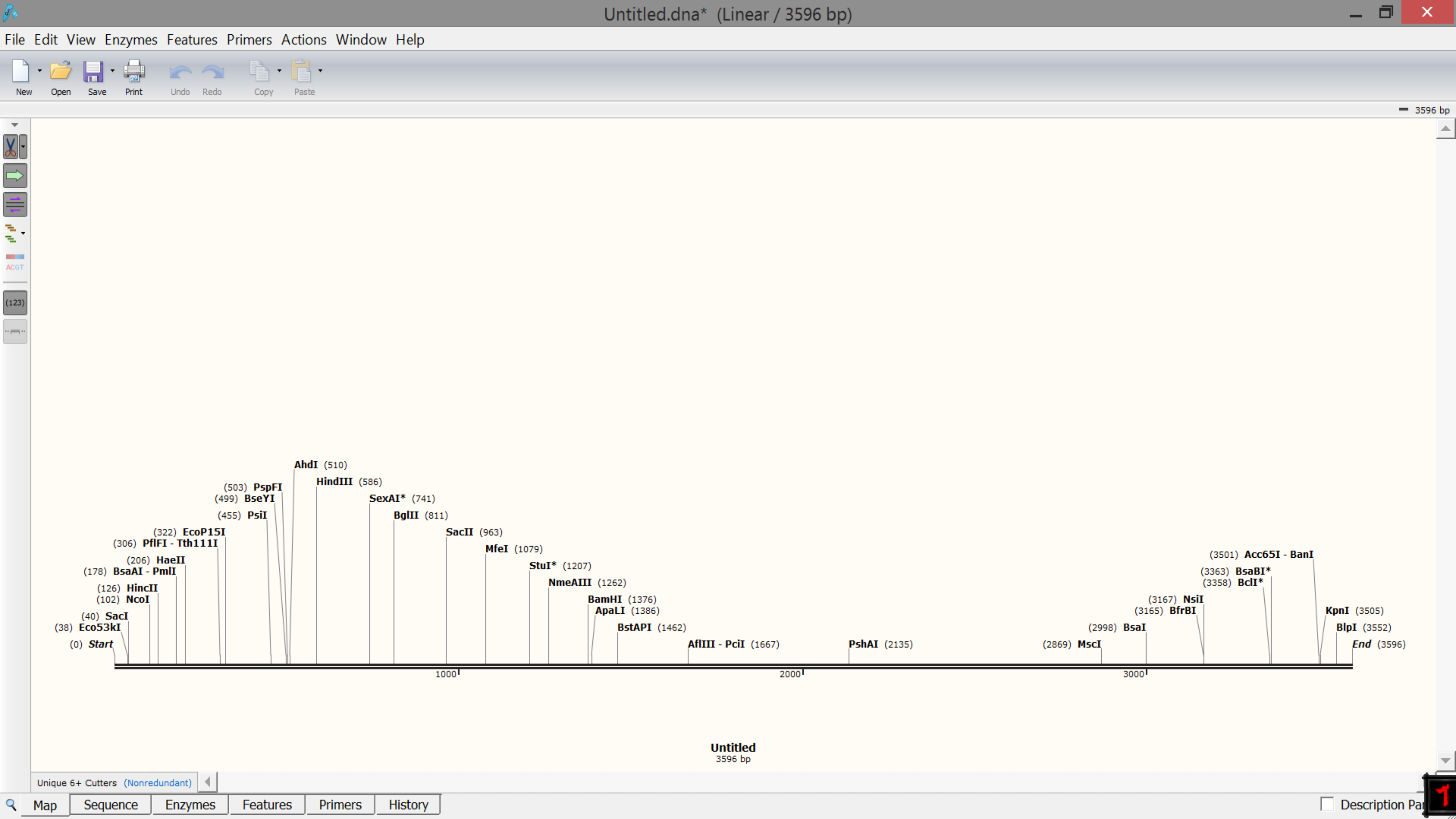
○ Create a sequence of 1000 N's

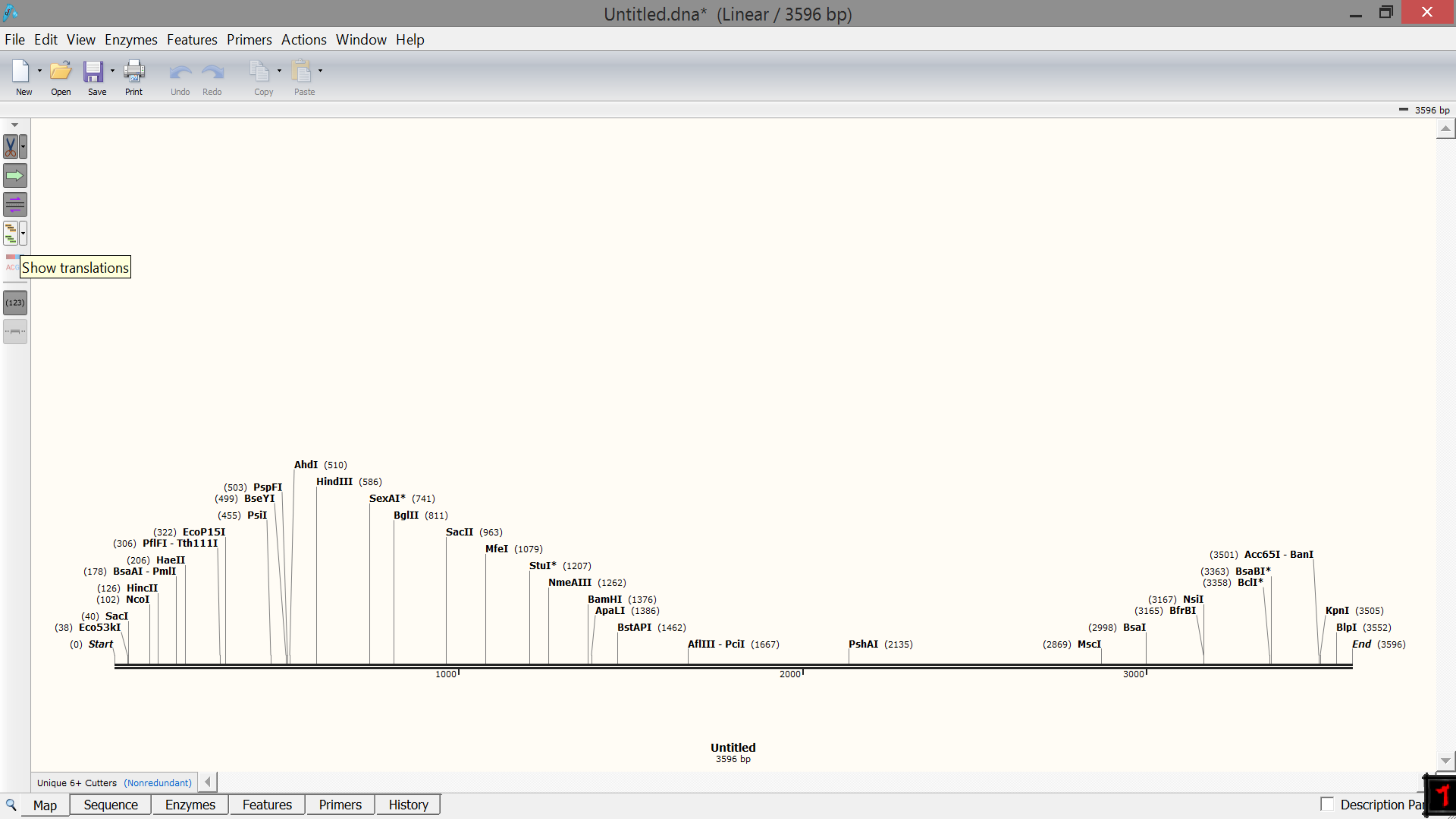
Topology: ☐ Linear ● Circular

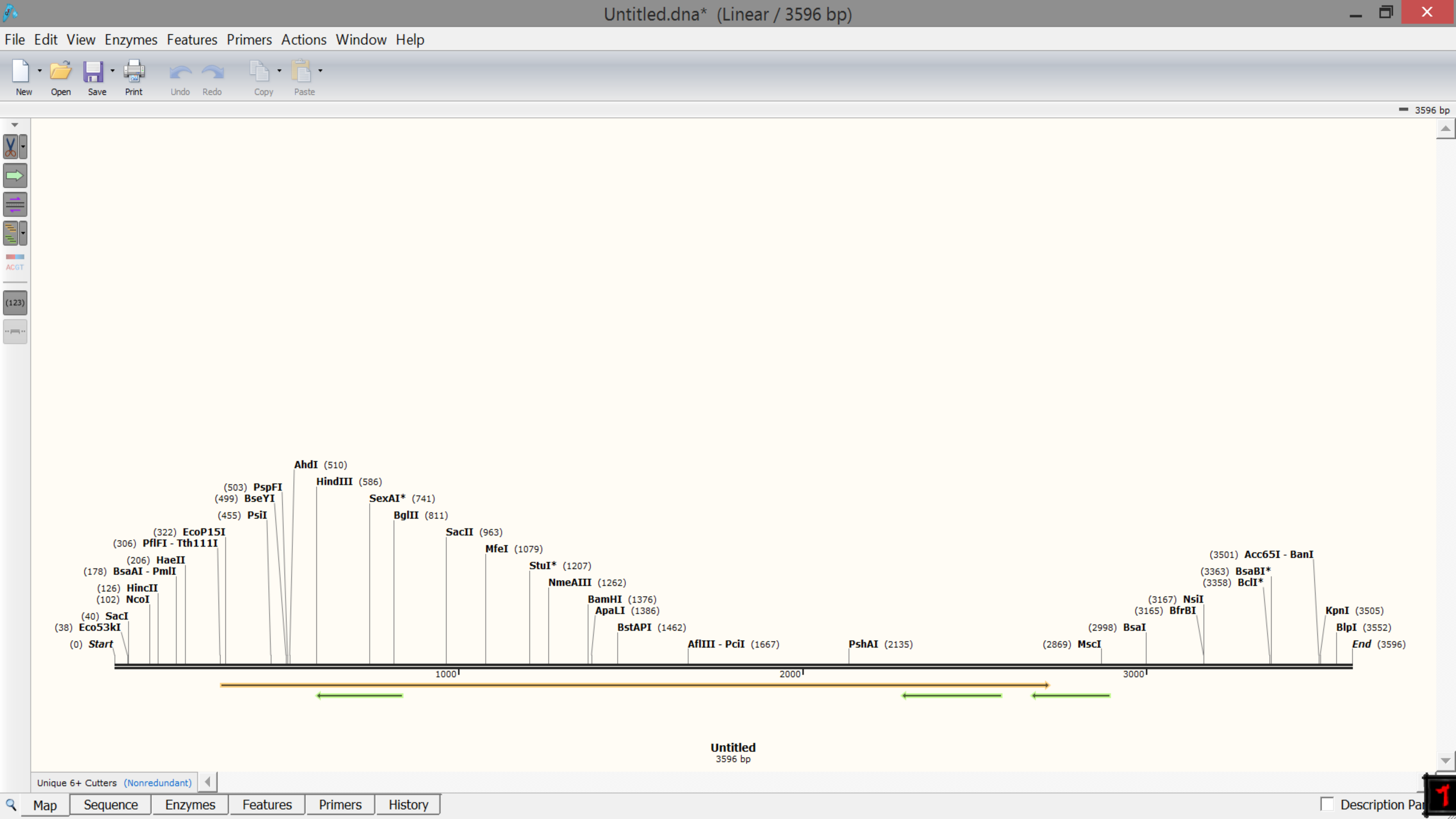
File Name: Untitled.dna

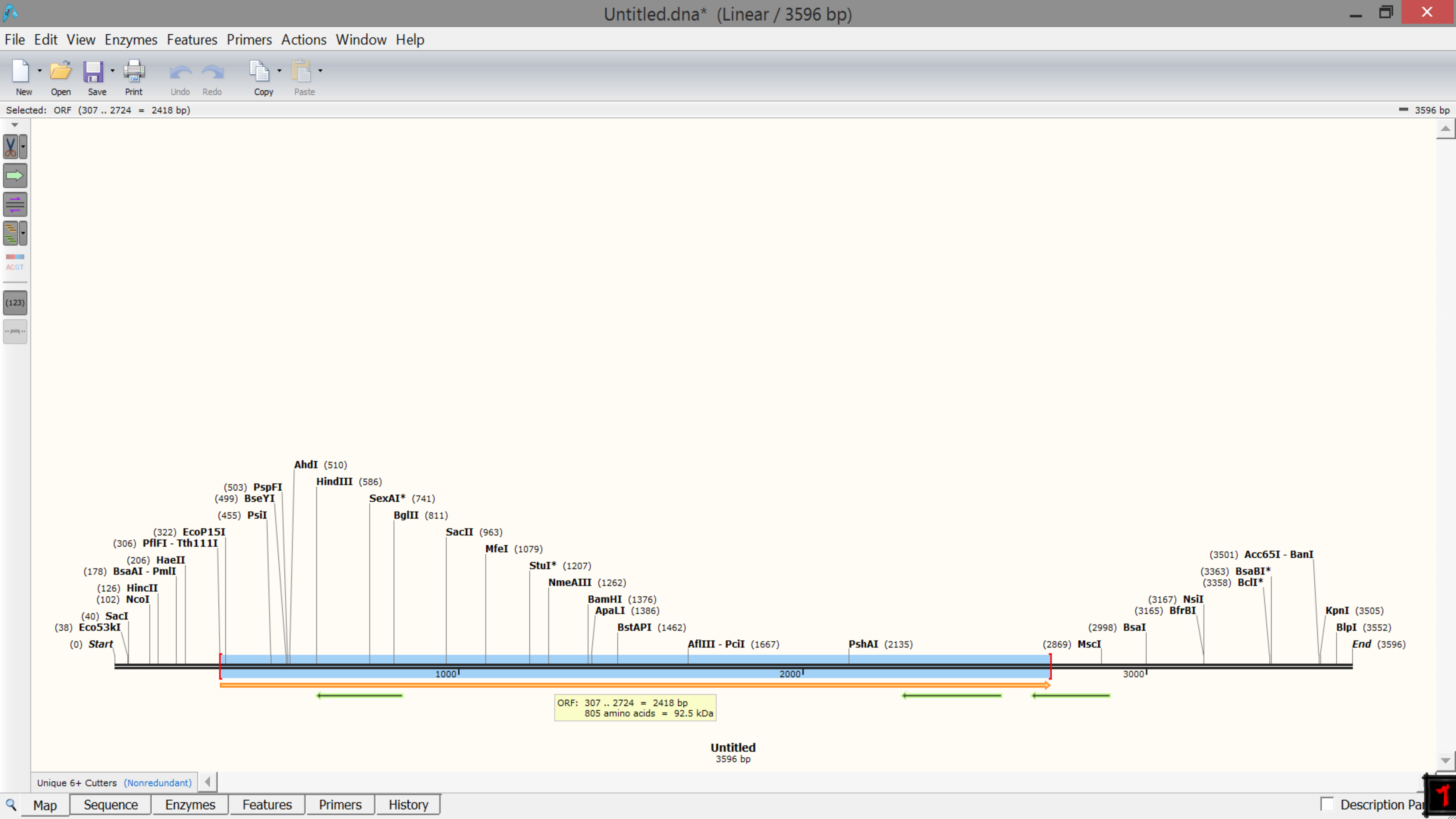
☑ Detect common features

```
1021 catcttcatg cctatgtgag ggcaaagttg atgaatgcct atccttccta tatcagtcca
1081 attggatgcc tccctgctca tttgcttggt gatatgtggg gtagattttg gacaaatctg
1141 tactctttga cagttccctt tggacagaaa ccaaacatag atgttactga tgcaatggtg
1201 gaccaggcct gggatgcaca gagaatattc aaggaggccg agaagttctt tgtatctgtt
1261 ggtcttccta atatgactca aggattctgg gaaaattcca tgctaacgga cccaggaaat
1321 gttcagaaag cagtctgcca tcccacagct tgggacctgg ggaagggcga cttcaggatc
1381 cttatgtgca caaaggtgac aatggacgac ttcctgacag ctcatcatga gatggggcat
1441 atccagtatg atatggcata tgctgcacaa ccttttctgc taagaaatgg agctaatgaa
1501 ggattccatg aagctgttgg ggaaatcatg tcactttctg cagccacacc taagcattta
1561 aaatccattg gtcttctgtc acccgatttt caagaagaca atgaaacaga aataaacttc
1621 ctgctcaaac aagcactcac gattgttggg actctgccat ttacttacat gttagagaag
1681 tggaggtgga tggtctttaa aggggaaatt cccaaagacc agtggatgaa aaagtggtgg
1741 gagatgaagc gagagatagt tggggtggtg gaacctgtgc cccatgatga aacatactgt
1801 gaccccgcat ctctgttcca tgtttctaat gattactcat tcattcgata ttacacaagg
1861 accctttacc aattccagtt tcaagaagca ctttgtcaag cagctaaaca tgaaggccct
1921 ctgcacaaat gtgacatctc aaactctaca gaagctggac agaaactgtt caatatgctg
1981 aggcttggaa aatcagaacc ctggacccta gcattggaaa atgttgtagg agcaaagaac
2041 atgaatgtaa ggccactgct caactacttt gagcccttat ttacctggct gaaagaccag
```


Untitled.dna* (Linear / 3596 bp)
File
Edit
View
Enzymes
Features
Primers
Actions
Window
Help
New
Open
Save
Print
Undo
Redo
Copy
Paste
3596 bp
ACGT
(123)
AhdI (510)
HindIII (586)
(503) PspFI
(499) BseYI
SexAI* (741)
(455) PsiI
BglII (811)
(322) EcoP15I
SacII (963)
(306) PflFI - Tth111I
MfeI (1079)
(206) HaeII
StuI* (1207)
(178) BsaAI - PmlI
NmeAIII (1262)
(126) HincII
(102) NcoI
BamHI (1376)
ApaLI (1386)
(40) SacI
(38) Eco53kI
BstAPI (1462)
(0) Start
AflIII - PciI (1667)
PshAI (2135)
(2869) MscI
(2998) BsaI
(3167) NsiI
(3165) BfrBI
(3363) BsaBI*
(3358) BclI*
(3501) Acc65I - BanI
KpnI (3505)
BlpI (3552)
End (3596)
1000
2000
3000
Untitled
3596 bp
Unique 6+ Cutters (Nonredundant)
Map
Sequence
Enzymes
Features
Primers
History
Description Pa


Untitled.dna* (Linear / 3596 bp)
File
Edit
View
Enzymes
Features
Primers
Actions
Window
Help
New
Open
Save
Print
Undo
Redo
Copy
Paste
3596 bp
Show translations
(123)
AhdI (510)
HindIII (586)
(503) PspFI
(499) BseYI
SexAI* (741)
(455) PsiI
BglII (811)
(322) EcoP15I
SacII (963)
(306) PflFI - Tth111I
MfeI (1079)
(206) HaeII
StuI* (1207)
(178) BsaAI - PmlI
NmeAIII (1262)
(126) HincII
BamHI (1376)
(102) NcoI
ApaLI (1386)
(40) SacI
BstAPI (1462)
(38) Eco53kI
(0) Start
AflIII - PciI (1667)
PshAI (2135)
(2869) MscI
(2998) BsaI
(3167) NsiI
(3165) BfrBI
(3501) Acc65I - BanI
(3363) BsaBI*
(3358) BclI*
KpnI (3505)
BlpI (3552)
End (3596)
1000
2000
3000
Untitled
3596 bp
Unique 6+ Cutters (Nonredundant)
Map
Sequence
Enzymes
Features
Primers
History
Description Pa

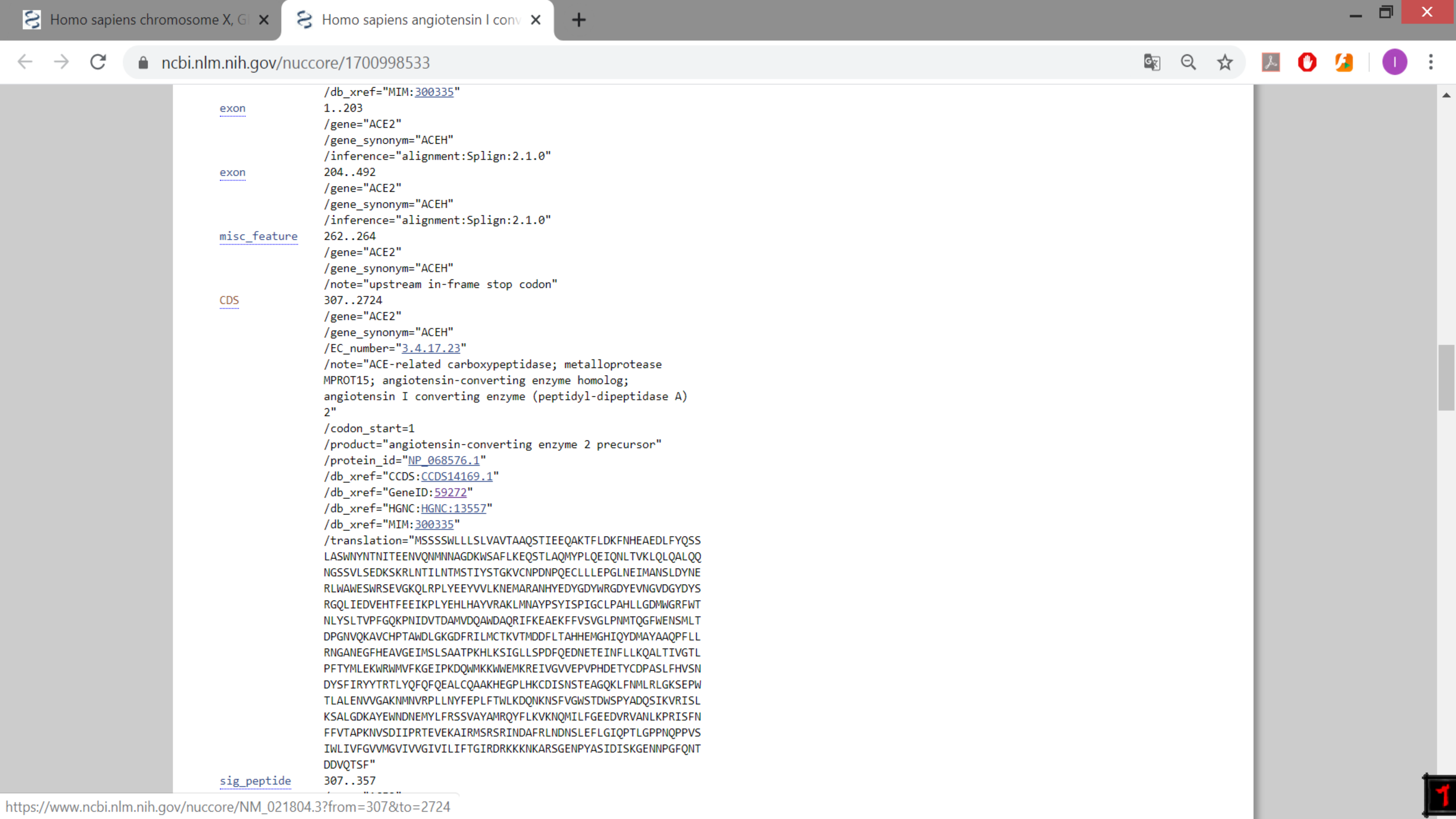

Untitled.dna* (Linear / 3596 bp)
File Edit View Enzymes Features Primers Actions Window Help
New Open Save Print Undo Redo Copy Paste
3596 bp
ACGT
(123)
AhdI (510)
HindIII (586)
(503) PspFI
(499) BseYI
SexAI* (741)
(455) PsiI
BglII (811)
(322) EcoP15I
SacII (963)
(306) PflFI - Tth111I
MfeI (1079)
(206) HaeII
StuI* (1207)
(178) BsaAI - PmlI
NmeAIII (1262)
(126) HincII
BamHI (1376)
(102) NcoI
ApaLI (1386)
(40) SacI
BstAPI (1462)
(38) Eco53kI
(0) Start
AflIII - PciI (1667)
PshAI (2135)
(2869) MscI
(2998) BsaI
(3165) BfrBI
(3167) NsiI
(3358) BclI*
(3363) BsaBI*
(3501) Acc65I - BanI
KpnI (3505)
BlpI (3552)
End (3596)
1000
2000
3000
Untitled
3596 bp
Unique 6+ Cutters (Nonredundant)
Map Sequence Enzymes Features Primers History
Description Pa



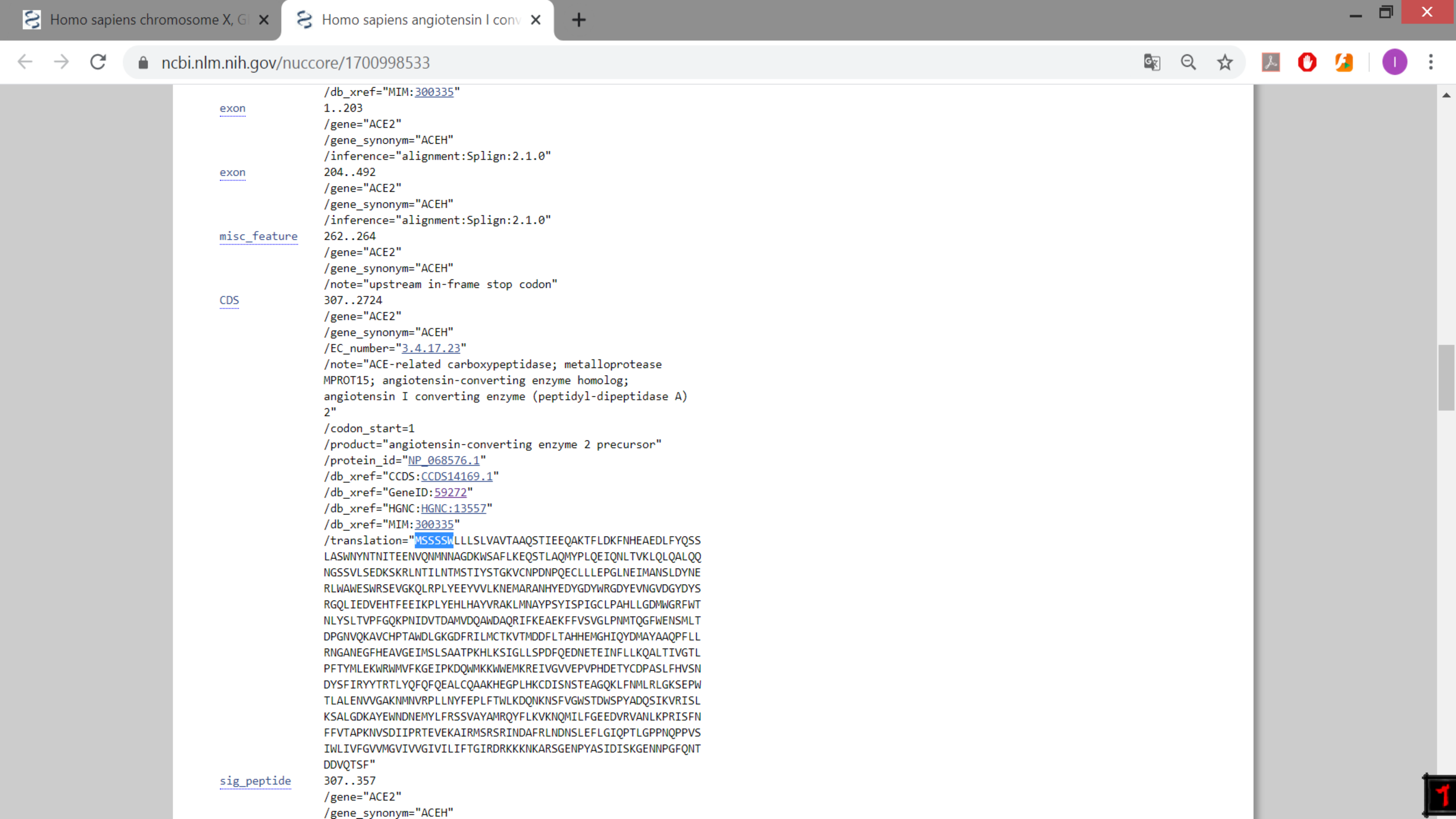
Untitled.dna* (Linear / 3596 bp)

Selected: ORF (307 .. 2724 = 2418 bp)

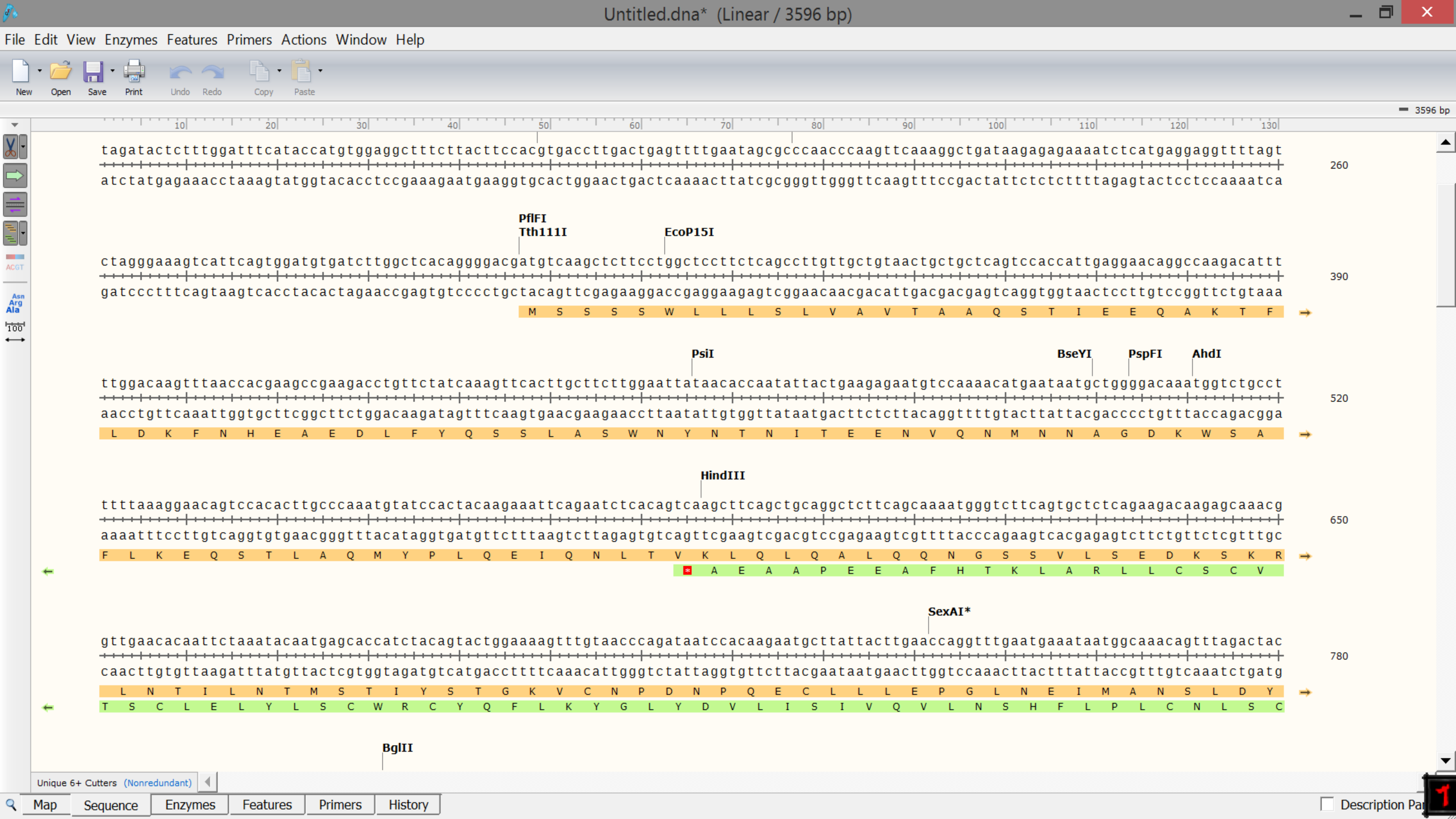
ORF: 307 .. 2724 = 2418 bp
805 amino acids = 92.5 kDa



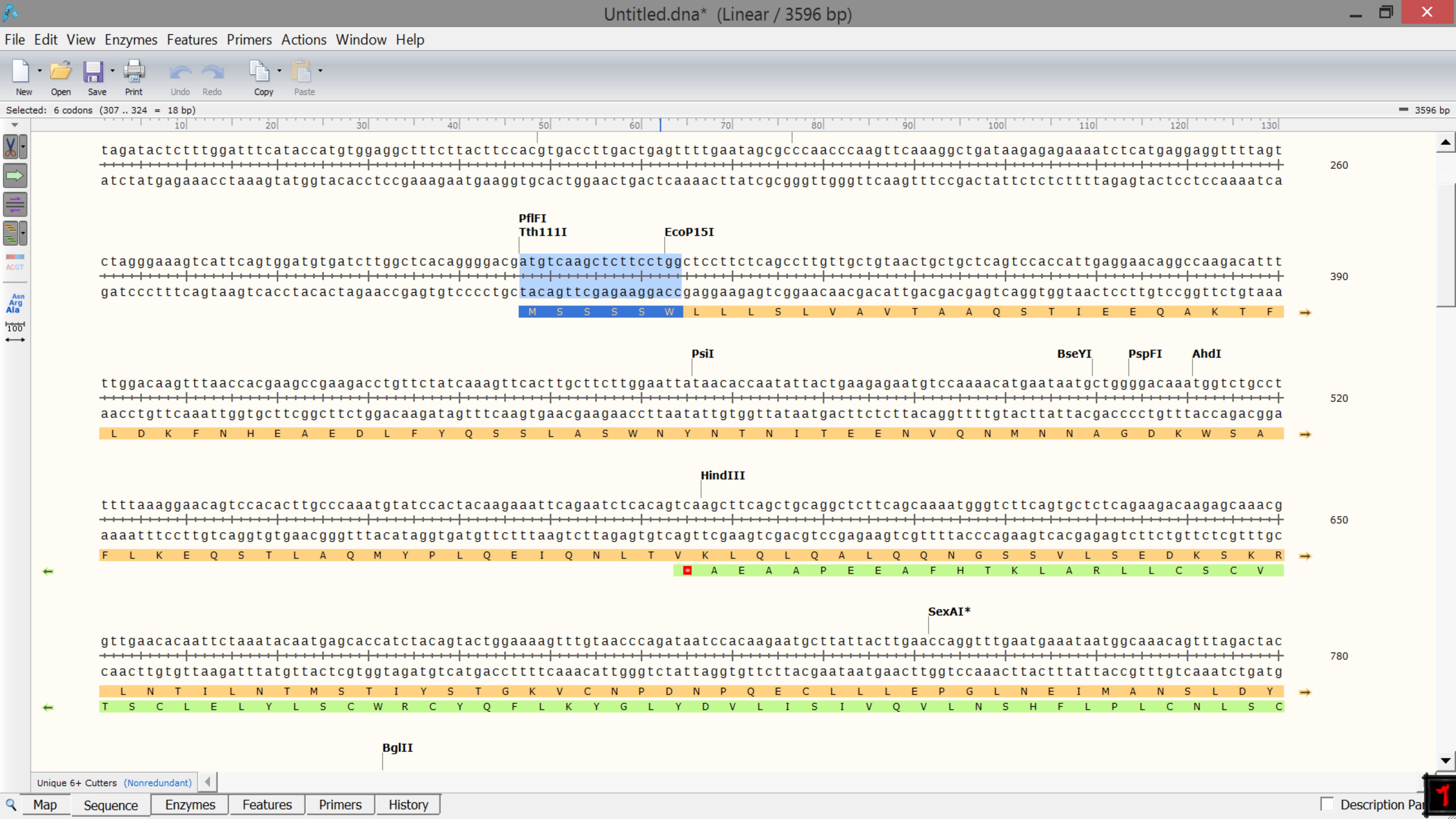
```
                     /db_xref="MIM:300335"
     exon            1..203
                     /gene="ACE2"
                     /gene_synonym="ACEH"
                     /inference="alignment:Splign:2.1.0"
     exon            204..492
                     /gene="ACE2"
                     /gene_synonym="ACEH"
                     /inference="alignment:Splign:2.1.0"
     misc_feature    262..264
                     /gene="ACE2"
                     /gene_synonym="ACEH"
                     /note="upstream in-frame stop codon"
     CDS             307..2724
                     /gene="ACE2"
                     /gene_synonym="ACEH"
                     /EC_number="3.4.17.23"
                     /note="ACE-related carboxypeptidase; metalloprotease
                     MPROT15; angiotensin-converting enzyme homolog;
                     angiotensin I converting enzyme (peptidyl-dipeptidase A)
                     2"
                     /codon_start=1
                     /product="angiotensin-converting enzyme 2 precursor"
                     /protein_id="NP_068576.1"
                     /db_xref="CCDS:CCDS14169.1"
                     /db_xref="GeneID:59272"
                     /db_xref="HGNC:HGNC:13557"
                     /db_xref="MIM:300335"
                     /translation="MSSSSWLLLSLVAVTAAQSTIEEQAKTFLDKFNHEAEDLFYQSS
                     LASWNYNTNITEENVQNMNNAGDKWSAFLKEQSTLAQMYPLQEIQNLTVKLQLQALQQ
                     NGSSVLSEDKSKRLNTILNTMSTIYSTGKVCNPDNPQECLLLEPGLNEIMANSLDYNE
                     RLWAWESWRSEVGKQLRPLYEEYVVLKNEMARANHYEDYGDYWRGDYEVNGVDGYDYS
                     RGQLIEDVEHTFEEIKPLYEHLHAYVRAKLMNAYPSYISPIGCLPAHLLGDMWGRFWT
                     NLYSLTVPFGQKPNIDVTDAMVDQAWDAQRIFKEAEKFFVSVGLPNMTQGFWENSMLT
                     DPGNVQKAVCHPTAWDLGKGDFRILMCTKVTMDDFLTAHHEMGHIQYDMAYAAQPFLL
                     RNGANEGFHEAVGEIMSLSAATPKHLKSIGLLSPDFQEDNETEINFLLKQALTIVGTL
                     PFTYMLEKWRWMVFKGEIPKDQWMKKWWEMKREIVGVVEPVPHDETYCDPASLFHVSN
                     DYSFIRYYTRTLYQFQFQEALCQAAKHEGPLHKCDISNSTEAGQKLFNMLRLGKSEPW
                     TLALENVVGAKNMNVRPLLNYFEPLFTWLKDQNKNSFVGWSTDWSPYADQSIKVRISL
                     KSALGDKAYEWNDNEMYLFRSSVAYAMRQYFLKVKNQMILFGEEDVRVANLKPRISFN
                     FFVTAPKNVSDIIPRTEVEKAIRMSRSRINDAFRLNDNSLEFLGIQPTLGPPNQPPVS
                     IWLIVFGVVMGVIVVGIVILIFTGIRDRKKKNKARSGENPYASIDISKGENNPGFQNT
                     DDVQTSF"
     sig_peptide     307..357
```



```
                     /db_xref="MIM:300335"
     exon            1..203
                     /gene="ACE2"
                     /gene_synonym="ACEH"
                     /inference="alignment:Splign:2.1.0"
     exon            204..492
                     /gene="ACE2"
                     /gene_synonym="ACEH"
                     /inference="alignment:Splign:2.1.0"
     misc_feature    262..264
                     /gene="ACE2"
                     /gene_synonym="ACEH"
                     /note="upstream in-frame stop codon"
     CDS             307..2724
                     /gene="ACE2"
                     /gene_synonym="ACEH"
                     /EC_number="3.4.17.23"
                     /note="ACE-related carboxypeptidase; metalloprotease
                     MPROT15; angiotensin-converting enzyme homolog;
                     angiotensin I converting enzyme (peptidyl-dipeptidase A)
                     2"
                     /codon_start=1
                     /product="angiotensin-converting enzyme 2 precursor"
                     /protein_id="NP_068576.1"
                     /db_xref="CCDS:CCDS14169.1"
                     /db_xref="GeneID:59272"
                     /db_xref="HGNC:HGNC:13557"
                     /db_xref="MIM:300335"
                     /translation="MSSSSWLLLSLVAVTAAQSTIEEQAKTFLDKFNHEAEDLFYQSS
                     LASWNYNTNITEENVQNMNNAGDKWSAFLKEQSTLAQMYPLQEIQNLTVKLQLQALQQ
                     NGSSVLSEDKSKRLNTILNTMSTIYSTGKVCNPDNPQECLLLEPGLNEIMANSLDYNE
                     RLWAWESWRSEVGKQLRPLYEEYVVLKNEMARANHYEDYGDYWRGDYEVNGVDGYDYS
                     RGQLIEDVEHTFEEIKPLYEHLHAYVRAKLMNAYPSYISPIGCLPAHLLGDMWGRFWT
                     NLYSLTVPFGQKPNIDVTDAMVDQAWDAQRIFKEAEKFFVSVGLPNMTQGFWENSMLT
                     DPGNVQKAVCHPTAWDLGKGDFRILMCTKVTMDDFLTAHHEMGHIQYDMAYAAQPFLL
                     RNGANEGFHEAVGEIMSLSAATPKHLKSIGLLSPDFQEDNETEINFLLKQALTIVGTL
                     PFTYMLEKWRWMVFKGEIPKDQWMKKWWEMKREIVGVVEPVPHDETYCDPASLFHVSN
                     DYSFIRYYTRTLYQFQFQEALCQAAKHEGPLHKCDISNSTEAGQKLFNMLRLGKSEPW
                     TLALENVVGAKNMNVRPLLNYFEPLFTWLKDQNKNSFVGWSTDWSPYADQSIKVRISL
                     KSALGDKAYEWNDNEMYLFRSSVAYAMRQYFLKVKNQMILFGEEDVRVANLKPRISFN
                     FFVTAPKNVSDIIPRTEVEKAIRMSRSRINDAFRLNDNSLEFLGIQPTLGPPNQPPVS
                     IWLIVFGVVMGVIVVGIVILIFTGIRDRKKKNKARSGENPYASIDISKGENNPGFQNT
                     DDVQTSF"
     sig_peptide     307..357
                     /gene="ACE2"
                     /gene_synonym="ACEH"
```



```
         10        20        30        40        50        60        70        80        90        100       110       120       130

tagatactctttggatttcataccatgtggaggctttcttacttccacgtgaccttgactgagttttgaatagcgcccaacccaagttcaaaggctgataagagagaaaatctcatgaggaggttttagt
atctatgagaaacctaaagtatggtacacctccgaaagaatgaaggtgcactggaactgactcaaaacttatcgcgggttgggttcaagtttccgactattctctcttttagagtactcctccaaaatca   260

                                              PflFI
                                              Tth111I                     EcoP15I
ctagggaaagtcattcagtggatgtgatcttggctcacaggggacgatgtcaagctcttcctggctccttctcagccttgttgctgtaactgctgctcagtccaccattgaggaacaggccaagacattt
gatccctttcagtaagtcacctacactagaaccgagtgtcccctgctacagttcgagaaggaccgaggaagagtcggaacaacgacattgacgacgagtcaggtggtaactccttgtccggttctgtaaa   390
                                              M  S  S  S  S  W  L  L  L  S  L  V  A  V  T  A  A  Q  S  T  I  E  E  Q  A  K  T  F  →

                                                                            PsiI                                       BseYI       PspFI      AhdI
ttggacaagtttaaccacgaagccgaagacctgttctatcaaagttcacttgcttcttggaattataacaccaatattactgaagagaatgtccaaaacatgaataatgctggggacaaatggtctgcct
aacctgttcaaattggtgcttcggcttctggacaagatagtttcaagtgaacgaagaaccttaatattgtggttataatgacttctcttacaggttttgtacttattacgacccctgtttaccagacgga   520
 L  D  K  F  N  H  E  A  E  D  L  F  Y  Q  S  S  L  A  S  W  N  Y  N  T  N  I  T  E  E  N  V  Q  N  M  N  N  A  G  D  K  W  S  A  →

                                                                             HindIII
ttttaaaggaacagtccacacttgcccaaatgtatccactacaagaaattcagaatctcacagtcaagcttcagctgcaggctcttcagcaaaatgggtcttcagtgctctcagaagacaagagcaaacg
aaaatttccttgtcaggtgtgaacgggtttacataggtgatgttctttaagtcttagagtgtcagttcgaagtcgacgtccgagaagtcgttttacccagaagtcacgagagtcttctgttctcgtttgc   650
F  L  K  E  Q  S  T  L  A  Q  M  Y  P  L  Q  E  I  Q  N  L  T  V  K  L  Q  L  Q  A  L  Q  Q  N  G  S  S  V  L  S  E  D  K  S  K  R  →
←                                                                  *  A  E  A  A  P  E  E  A  F  H  T  K  L  A  R  L  L  C  S  C  V

                                                                                                       SexAI*
gttgaacacaattctaaatacaatgagcaccatctacagtactggaaaagtttgtaacccagataatccacaagaatgcttattacttgaaccaggtttgaatgaaataatggcaaacagtttagactac
caacttgtgttaagatttatgttactcgtggtagatgtcatgaccttttcaaacattgggtctattaggtgttcttacgaataatgaacttggtccaaacttactttattaccgtttgtcaaatctgatg   780
  L  N  T  I  L  N  T  M  S  T  I  Y  S  T  G  K  V  C  N  P  D  N  P  Q  E  C  L  L  L  E  P  G  L  N  E  I  M  A  N  S  L  D  Y  →
← T  S  C  L  E  L  Y  L  S  C  W  R  C  Y  Q  F  L  K  Y  G  L  Y  D  V  L  I  S  I  V  Q  V  L  N  S  H  F  L  P  L  C  N  L  S  C

                              BglII
```

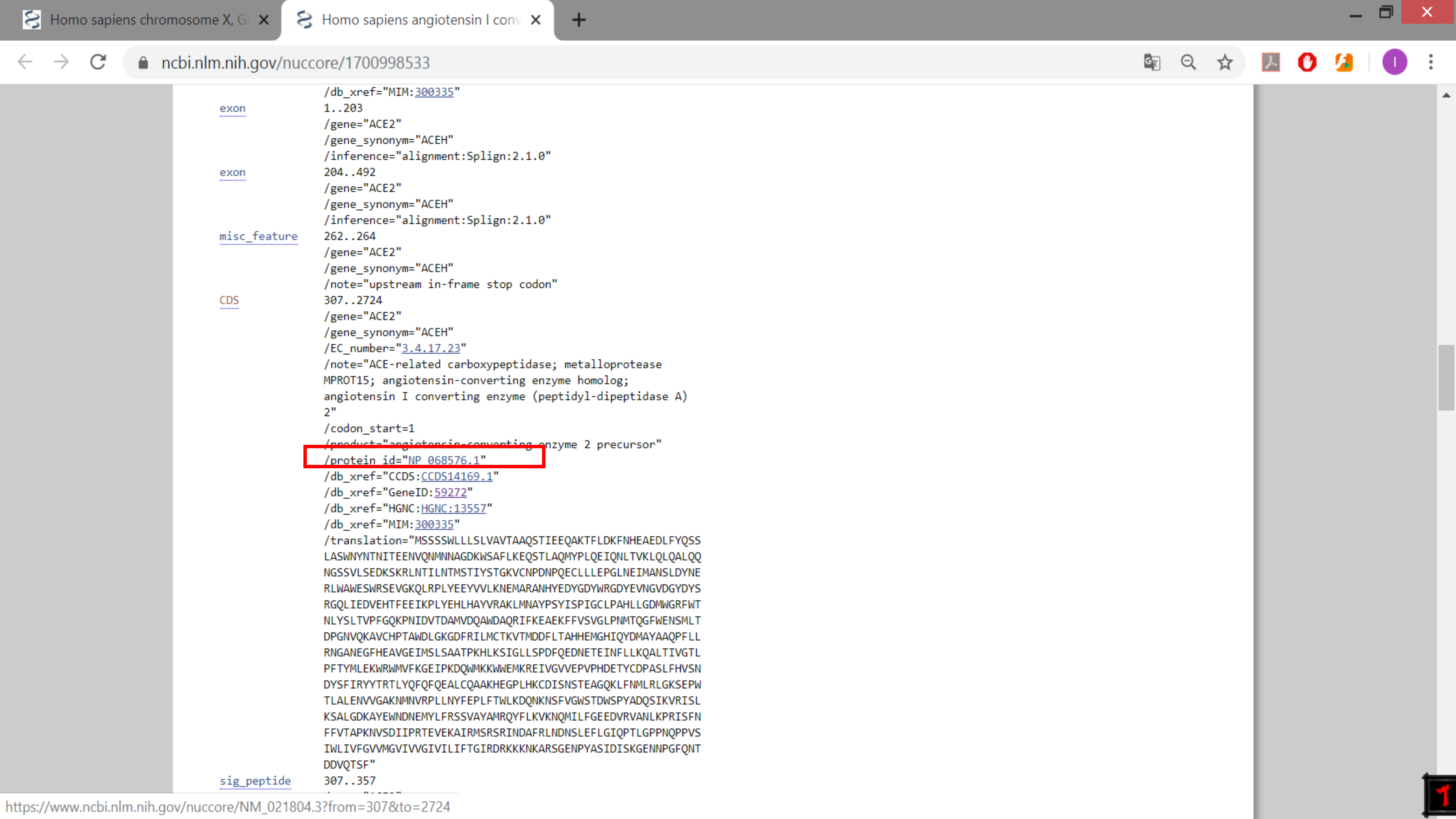


Untitled.dna* (Linear / 3596 bp)

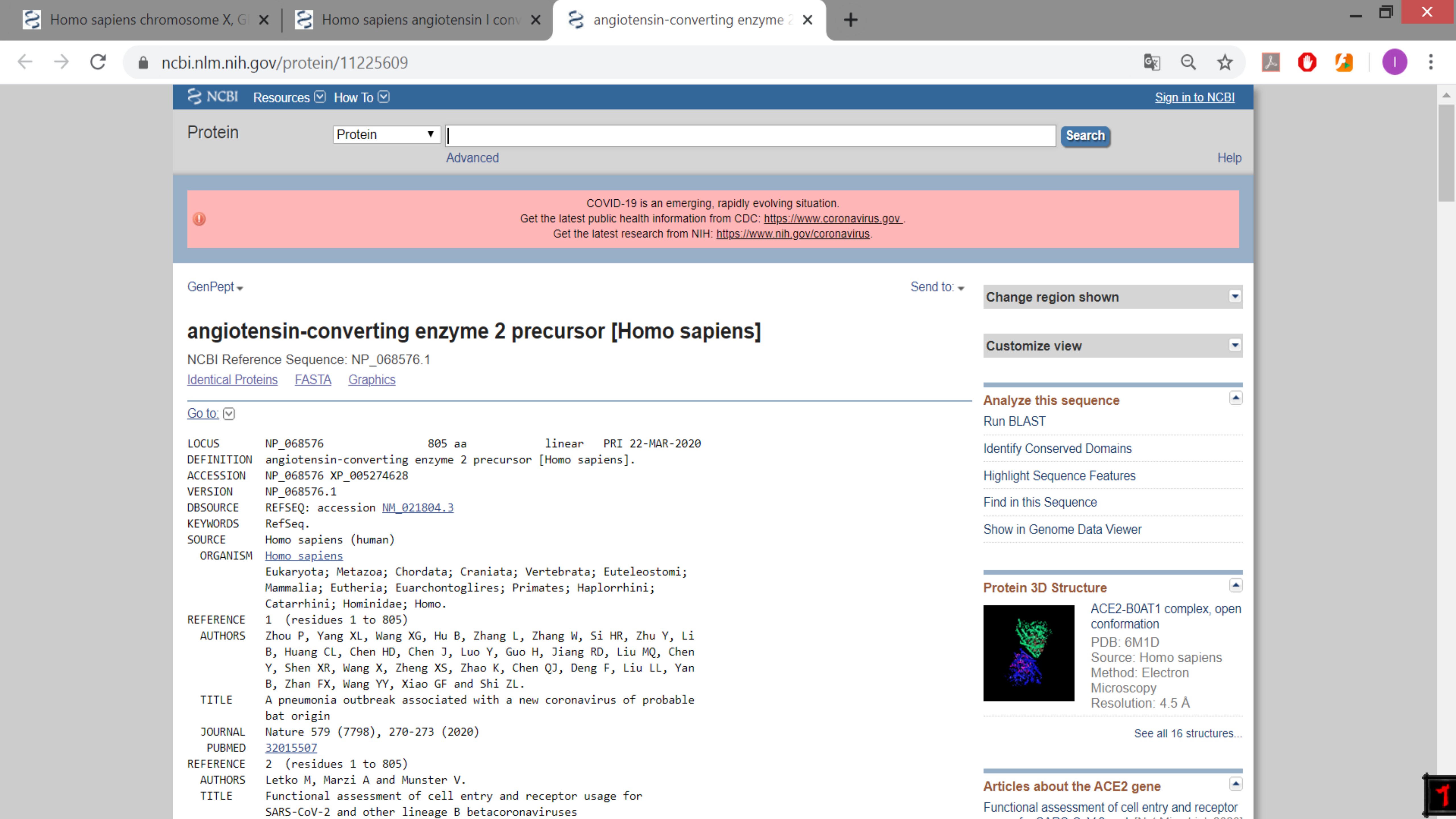
Selected: 6 codons (307 .. 324 = 18 bp)

3596 bp

```
tagatactctttggatttcataccatgtggaggctttcttacttccacgtgaccttgactgagttttgaatagcgcccaacccaagttcaaaggctgataagagagaaaatctcatgaggaggttttagt
atctatgagaaacctaaagtatggtacacctccgaaagaatgaaggtgcactggaactgactcaaaacttatcgcgggttgggttcaagtttccgactattctctcttttagagtactcctccaaaatca     260

PflFI
Tth111I                                                  EcoP15I
ctagggaaagtcattcagtggatgtgatcttggctcacaggggacgatgtcaagctcttcctggctccttctcagccttgttgctgtaactgctgctcagtccaccattgaggaacaggccaagacattt
gatccctttcagtaagtcacctacactagaaccgagtgtcccctgctacagttcgagaaggaccgaggaagagtcggaacaacgacattgacgacgagtcaggtggtaactccttgtccggttctgtaaa     390
M  S  S  S  S  W  L  L  L  S  L  V  A  V  T  A  A  Q  S  T  I  E  E  Q  A  K  T  F

PsiI   BseYI   PspFI   AhdI
ttggacaagtttaaccacgaagccgaagacctgttctatcaaagttcacttgcttcttggaattataacaccaatattactgaagagaatgtccaaaacatgaataatgctggggacaaatggtctgcct
aacctgttcaaattggtgcttcggcttctggacaagatagtttcaagtgaacgaagaaccttaatattgtggttataatgacttctcttacaggttttgtacttattacgacccctgtttaccagacgga     520
L  D  K  F  N  H  E  A  E  D  L  F  Y  Q  S  S  L  A  S  W  N  Y  N  T  N  I  T  E  E  N  V  Q  N  M  N  N  A  G  D  K  W  S  A

HindIII
ttttaaaggaacagtccacacttgcccaaatgtatccactacaagaaattcagaatctcacagtcaagcttcagctgcaggctcttcagcaaaatgggtcttcagtgctctcagaagacaagagcaaacg
aaaatttccttgtcaggtgtgaacgggtttacataggtgatgtttctttaagtcttagagtgtcagttcgaagtcgacgtccgagaagtcgttttacccagaagtcacgagagtcttctgttctcgtttgc     650
F  L  K  E  Q  S  T  L  A  Q  M  Y  P  L  Q  E  I  Q  N  L  T  V  K  L  Q  L  Q  A  L  Q  Q  N  G  S  S  V  L  S  E  D  K  S  K  R
                                                                   *  A  E  A  A  P  E  E  A  F  H  T  K  L  A  R  L  L  C  S  C  V

SexAI*
gttgaacacaattctaaatacaatgagcaccatctacagtactggaaaagtttgtaacccagataatccacaagaatgcttattacttgaaccaggtttgaatgaaataatggcaaacagtttagactac
caacttgtgttaagatttatgttactcgtggtagatgtcatgaccttttcaaacattgggtctattaggtgttcttacgaataatgaacttggtccaaacttactttattaccgtttgtcaaatctgatg     780
L  N  T  I  L  N  T  M  S  T  I  Y  S  T  G  K  V  C  N  P  D  N  P  Q  E  C  L  L  L  E  P  G  L  N  E  I  M  A  N  S  L  D  Y
T  S  C  L  E  L  Y  L  S  C  W  R  C  Y  Q  F  L  K  Y  G  L  Y  D  V  L  I  S  I  V  Q  V  L  N  S  H  F  L  P  L  C  N  L  S  C

BglII
```



```
                     /db_xref="MIM:300335"
     exon            1..203
                     /gene="ACE2"
                     /gene_synonym="ACEH"
                     /inference="alignment:Splign:2.1.0"
     exon            204..492
                     /gene="ACE2"
                     /gene_synonym="ACEH"
                     /inference="alignment:Splign:2.1.0"
     misc_feature    262..264
                     /gene="ACE2"
                     /gene_synonym="ACEH"
                     /note="upstream in-frame stop codon"
     CDS             307..2724
                     /gene="ACE2"
                     /gene_synonym="ACEH"
                     /EC_number="3.4.17.23"
                     /note="ACE-related carboxypeptidase; metalloprotease
                     MPROT15; angiotensin-converting enzyme homolog;
                     angiotensin I converting enzyme (peptidyl-dipeptidase A)
                     2"
                     /codon_start=1
                     /product="angiotensin-converting enzyme 2 precursor"
                     /protein_id="NP_068576.1"
                     /db_xref="CCDS:CCDS14169.1"
                     /db_xref="GeneID:59272"
                     /db_xref="HGNC:HGNC:13557"
                     /db_xref="MIM:300335"
                     /translation="MSSSSWLLLSLVAVTAAQSTIEEQAKTFLDKFNHEAEDLFYQSS
                     LASWNYNTNITEENVQNMNNAGDKWSAFLKEQSTLAQMYPLQEIQNLTVKLQLQALQQ
                     NGSSVLSEDKSKRLNTILNTMSTIYSTGKVCNPDNPQECLLLEPGLNEIMANSLDYNE
                     RLWAWESWRSEVGKQLRPLYEEYVVLKNEMARANHYEDYGDYWRGDYEVNGVDGYDYS
                     RGQLIEDVEHTFEEIKPLYEHLHAYVRAKLMNAYPSYISPIGCLPAHLLGDMWGRFWT
                     NLYSLTVPFGQKPNIDVTDAMVDQAWDAQRIFKEAEKFFVSVGLPNMTQGFWENSMLT
                     DPGNVQKAVCHPTAWDLGKGDFRILMCTKVTMDDFLTAHHEMGHIQYDMAYAAQPFLL
                     RNGANEGFHEAVGEIMSLSAATPKHLKSIGLLSPDFQEDNETEINFLLKQALTIVGTL
                     PFTYMLEKWRWMVFKGEIPKDQWMKKWWEMKREIVGVVEPVPHDETYCDPASLFHVSN
                     DYSFIRYYTRTLYQFQFQEALCQAAKHEGPLHKCDISNSTEAGQKLFNMLRLGKSEPW
                     TLALENVVGAKNMNVRPLLNYFEPLFTWLKDQNKNSFVGWSTDWSPYADQSIKVRISL
                     KSALGDKAYEWNDNEMYLFRSSVAYAMRQYFLKVKNQMILFGEEDVRVANLKPRISFN
                     FFVTAPKNVSDIIPRTEVEKAIRMSRSRINDAFRLNDNSLEFLGIQPTLGPPNQPPVS
                     IWLIVFGVVMGVIVVGIVILIFTGIRDRKKKNKARSGENPYASIDISKGENNPGFQNT
                     DDVQTSF"
     sig_peptide     307..357
```



COVID-19 is an emerging, rapidly evolving situation.
Get the latest public health information from CDC: https://www.coronavirus.gov .
Get the latest research from NIH: https://www.nih.gov/coronavirus.

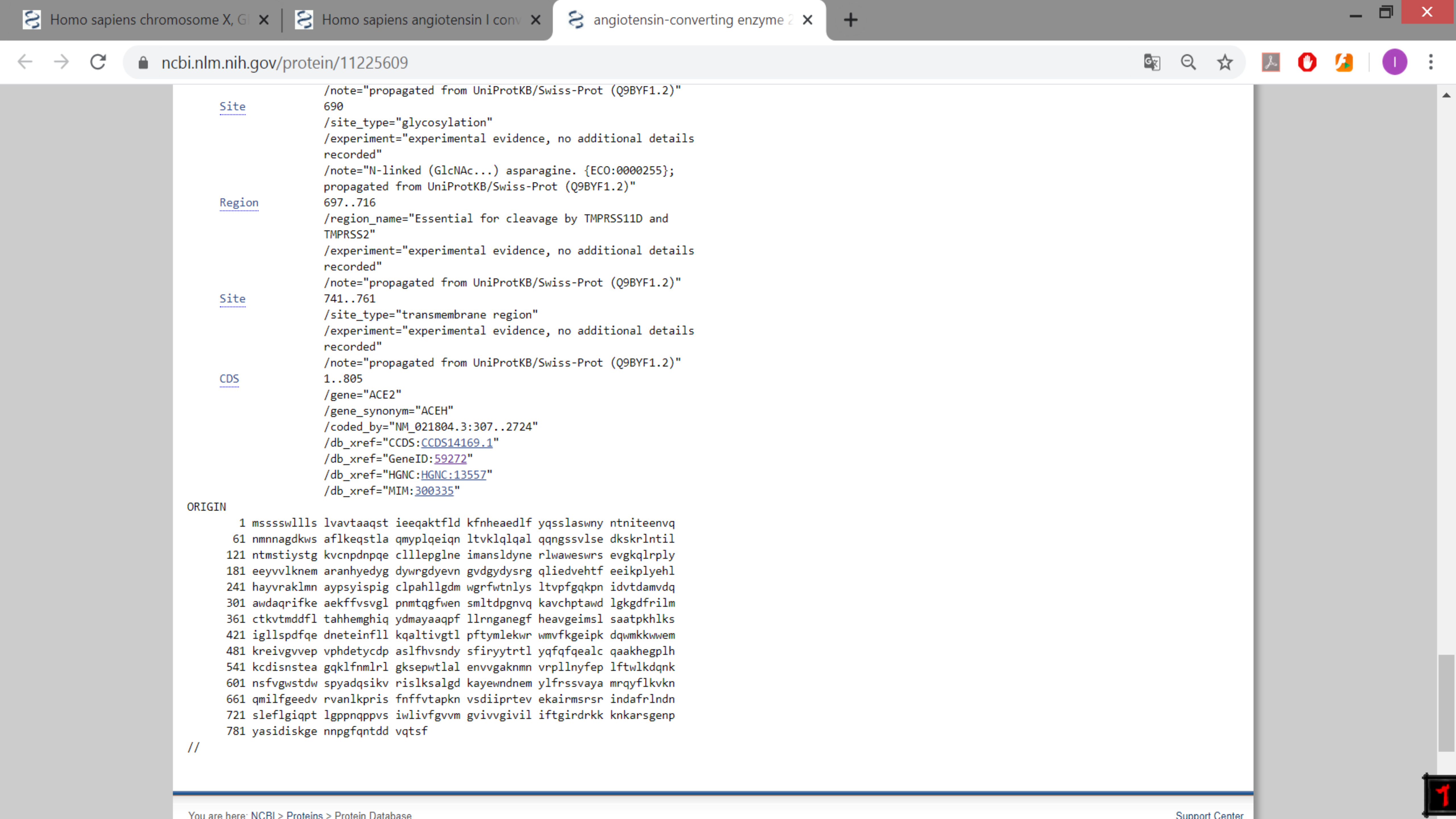
angiotensin-converting enzyme 2 precursor [Homo sapiens]

NCBI Reference Sequence: NP_068576.1

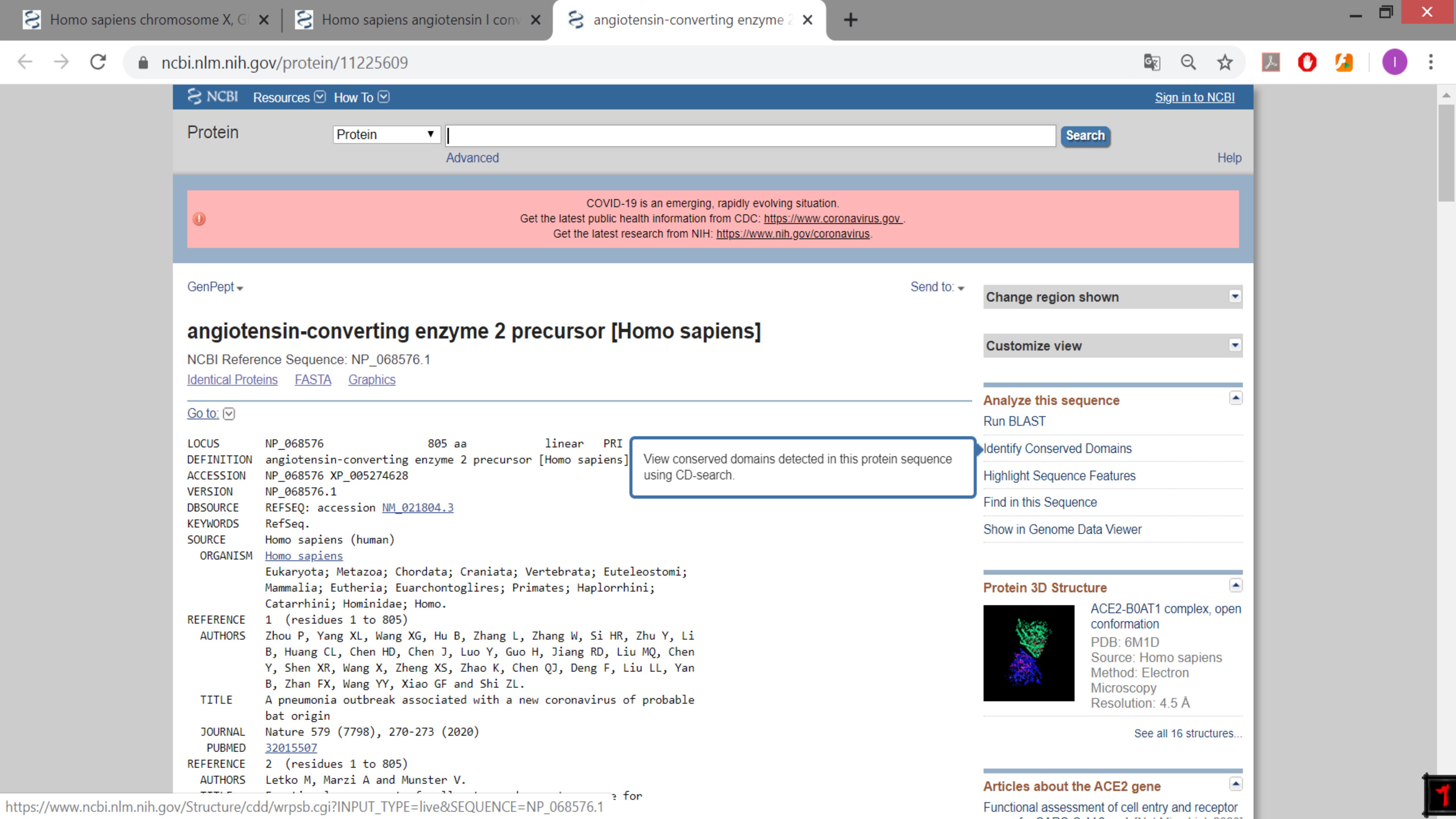
```
LOCUS       NP_068576                805 aa            linear   PRI 22-MAR-2020
DEFINITION  angiotensin-converting enzyme 2 precursor [Homo sapiens].
ACCESSION   NP_068576 XP_005274628
VERSION     NP_068576.1
DBSOURCE    REFSEQ: accession NM_021804.3
KEYWORDS    RefSeq.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;
            Catarrhini; Hominidae; Homo.
REFERENCE   1  (residues 1 to 805)
  AUTHORS   Zhou P, Yang XL, Wang XG, Hu B, Zhang L, Zhang W, Si HR, Zhu Y, Li
            B, Huang CL, Chen HD, Chen J, Luo Y, Guo H, Jiang RD, Liu MQ, Chen
            Y, Shen XR, Wang X, Zheng XS, Zhao K, Chen QJ, Deng F, Liu LL, Yan
            B, Zhan FX, Wang YY, Xiao GF and Shi ZL.
  TITLE     A pneumonia outbreak associated with a new coronavirus of probable
            bat origin
  JOURNAL   Nature 579 (7798), 270-273 (2020)
   PUBMED   32015507
REFERENCE   2  (residues 1 to 805)
  AUTHORS   Letko M, Marzi A and Munster V.
  TITLE     Functional assessment of cell entry and receptor usage for
            SARS-CoV-2 and other lineage B betacoronaviruses
```

ACE2-B0AT1 complex, open conformation
PDB: 6M1D
Source: Homo sapiens
Method: Electron Microscopy
Resolution: 4.5 Å

Functional assessment of cell entry and receptor



```
                     /note="propagated from UniProtKB/Swiss-Prot (Q9BYF1.2)"
     Site            690
                     /site_type="glycosylation"
                     /experiment="experimental evidence, no additional details
                     recorded"
                     /note="N-linked (GlcNAc...) asparagine. {ECO:0000255};
                     propagated from UniProtKB/Swiss-Prot (Q9BYF1.2)"
     Region          697..716
                     /region_name="Essential for cleavage by TMPRSS11D and
                     TMPRSS2"
                     /experiment="experimental evidence, no additional details
                     recorded"
                     /note="propagated from UniProtKB/Swiss-Prot (Q9BYF1.2)"
     Site            741..761
                     /site_type="transmembrane region"
                     /experiment="experimental evidence, no additional details
                     recorded"
                     /note="propagated from UniProtKB/Swiss-Prot (Q9BYF1.2)"
     CDS             1..805
                     /gene="ACE2"
                     /gene_synonym="ACEH"
                     /coded_by="NM_021804.3:307..2724"
                     /db_xref="CCDS:CCDS14169.1"
                     /db_xref="GeneID:59272"
                     /db_xref="HGNC:HGNC:13557"
                     /db_xref="MIM:300335"
ORIGIN
        1 msssswllls lvavtaaqst ieeqaktfld kfnheaedlf yqsslaswny ntniteenvq
       61 nmnnagdkws aflkeqstla qmyplqeiqn ltvklqlqal qqngssvlse dkskrlntil
      121 ntmstiystg kvcnpdnpqe clllepglne imansldyne rlwaweswrs evgkqlrply
      181 eeyvvlknem aranhyedyg dywrgdyevn gvdgydysrg qliedvehtf eeikplyehl
      241 hayvraklmn aypsyispig clpahllgdm wgrfwtnlys ltvpfgqkpn idvtdamvdq
      301 awdaqrifke aekffvsvgl pnmtqgfwen smltdpgnvq kavchptawd lgkgdfrilm
      361 ctkvtmddfl tahhemghiq ydmayaaqpf llrnganegf heavgeimsl saatpkhlks
      421 igllspdfqe dneteinfll kqaltivgtl pftymlekwr wmvfkgeipk dqwmkkwwem
      481 kreivgvvep vphdetycdp aslfhvsndy sfiryytrtl yqfqfqealc qaakhegplh
      541 kcdisnstea gqklfnmlrl gksepwtlal envvgaknmn vrpllnyfep lftwlkdqnk
      601 nsfvgwstdw spyadqsikv rislksalgd kayewndnem ylfrssvaya mrqyflkvkn
      661 qmilfgeedv rvanlkpris fnffvtapkn vsdiiprtev ekairmsrsr indafrlndn
      721 sleflgiqpt lgppnqppvs iwlivfgvvm gvivvgivil iftgirdrkk knkarsgenp
      781 yasidiskge nnpgfqntdd vqtsf
//
```



COVID-19 is an emerging, rapidly evolving situation.
Get the latest public health information from CDC: https://www.coronavirus.gov .
Get the latest research from NIH: https://www.nih.gov/coronavirus.

angiotensin-converting enzyme 2 precursor [Homo sapiens]

NCBI Reference Sequence: NP_068576.1

Identical Proteins FASTA Graphics

```
LOCUS       NP_068576                805 aa            linear   PRI
DEFINITION  angiotensin-converting enzyme 2 precursor [Homo sapiens]
ACCESSION   NP_068576 XP_005274628
VERSION     NP_068576.1
DBSOURCE    REFSEQ: accession NM_021804.3
KEYWORDS    RefSeq.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;
            Catarrhini; Hominidae; Homo.
REFERENCE   1  (residues 1 to 805)
  AUTHORS   Zhou P, Yang XL, Wang XG, Hu B, Zhang L, Zhang W, Si HR, Zhu Y, Li
            B, Huang CL, Chen HD, Chen J, Luo Y, Guo H, Jiang RD, Liu MQ, Chen
            Y, Shen XR, Wang X, Zheng XS, Zhao K, Chen QJ, Deng F, Liu LL, Yan
            B, Zhan FX, Wang YY, Xiao GF and Shi ZL.
  TITLE     A pneumonia outbreak associated with a new coronavirus of probable
            bat origin
  JOURNAL   Nature 579 (7798), 270-273 (2020)
   PUBMED   32015507
REFERENCE   2  (residues 1 to 805)
  AUTHORS   Letko M, Marzi A and Munster V.
                                                          for
```

View conserved domains detected in this protein sequence using CD-search.

Analyze this sequence

Run BLAST
Identify Conserved Domains
Highlight Sequence Features
Find in this Sequence
Show in Genome Data Viewer

Protein 3D Structure

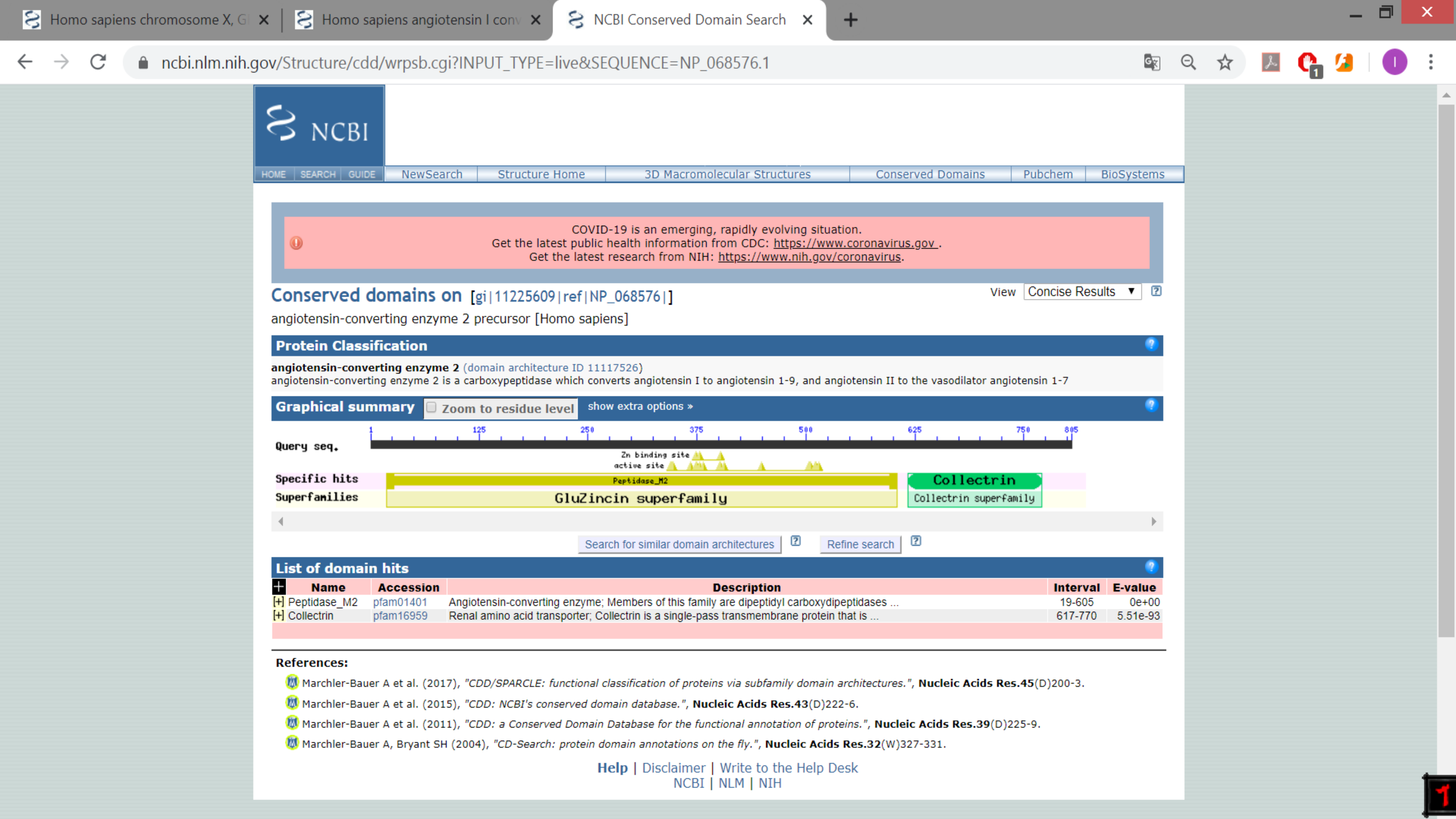
ACE2-B0AT1 complex, open conformation
PDB: 6M1D
Source: Homo sapiens
Method: Electron Microscopy
Resolution: 4.5 Å

See all 16 structures...

Articles about the ACE2 gene

Functional assessment of cell entry and receptor

https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi?INPUT_TYPE=live&SEQUENCE=NP_068576.1



COVID-19 is an emerging, rapidly evolving situation.
Get the latest public health information from CDC: https://www.coronavirus.gov .
Get the latest research from NIH: https://www.nih.gov/coronavirus.

Conserved domains on [gi|11225609|ref|NP_068576|]

View Concise Results ▼

angiotensin-converting enzyme 2 precursor [Homo sapiens]

Protein Classification

angiotensin-converting enzyme 2 (domain architecture ID 11117526)
angiotensin-converting enzyme 2 is a carboxypeptidase which converts angiotensin I to angiotensin 1-9, and angiotensin II to the vasodilator angiotensin 1-7

Graphical summary

☐ Zoom to residue level show extra options »

List of domain hits

Name	Accession	Description	Interval	E-value
[+] Peptidase_M2	pfam01401	Angiotensin-converting enzyme; Members of this family are dipeptidyl carboxydipeptidases ...	19-605	0e+00
[+] Collectrin	pfam16959	Renal amino acid transporter; Collectrin is a single-pass transmembrane protein that is ...	617-770	5.51e-93

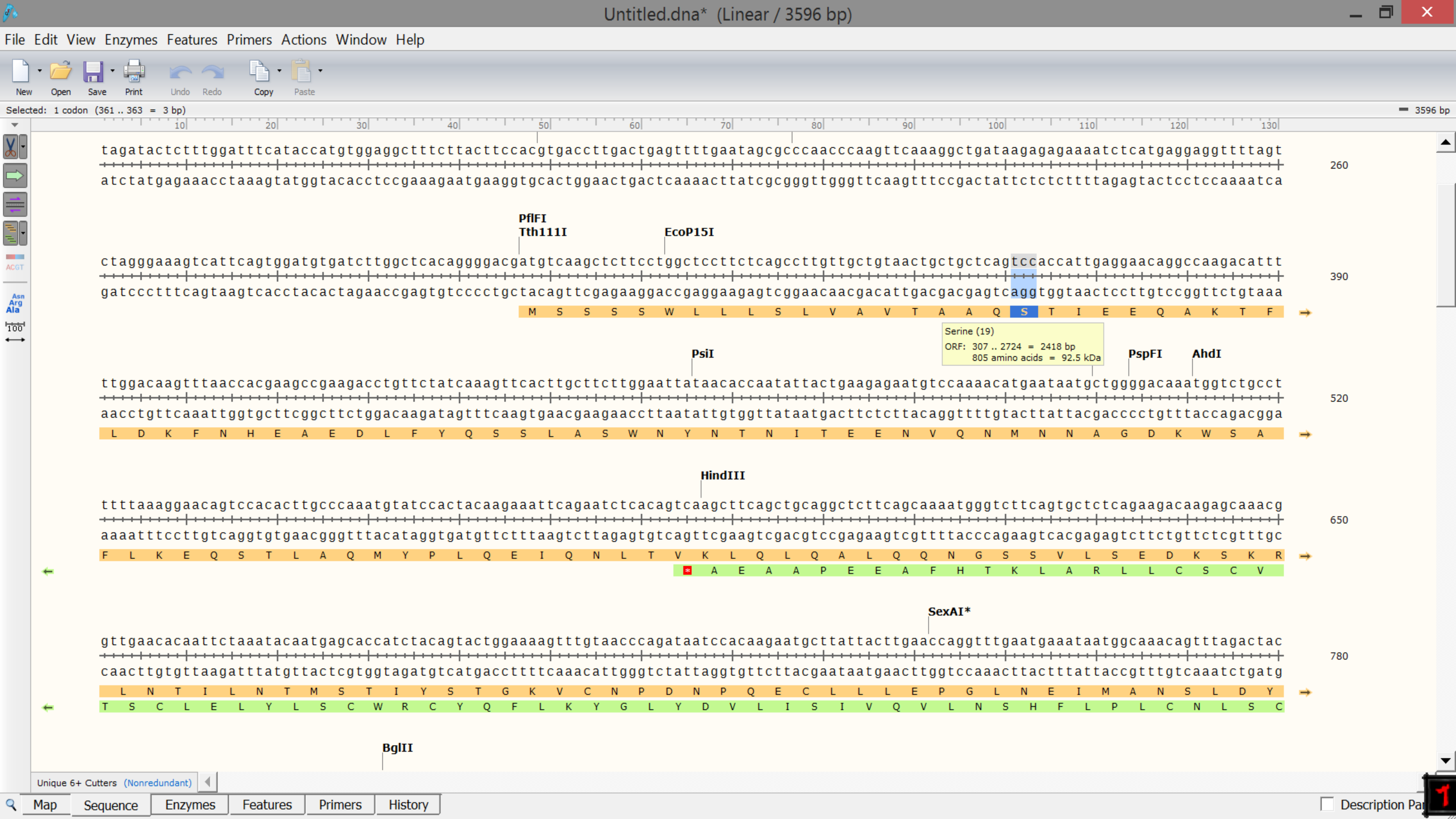
References:

Marchler-Bauer A et al. (2017), *"CDD/SPARCLE: functional classification of proteins via subfamily domain architectures."*, **Nucleic Acids Res.45**(D)200-3.

Marchler-Bauer A et al. (2015), *"CDD: NCBI's conserved domain database."*, **Nucleic Acids Res.43**(D)222-6.

Marchler-Bauer A et al. (2011), *"CDD: a Conserved Domain Database for the functional annotation of proteins."*, **Nucleic Acids Res.39**(D)225-9.

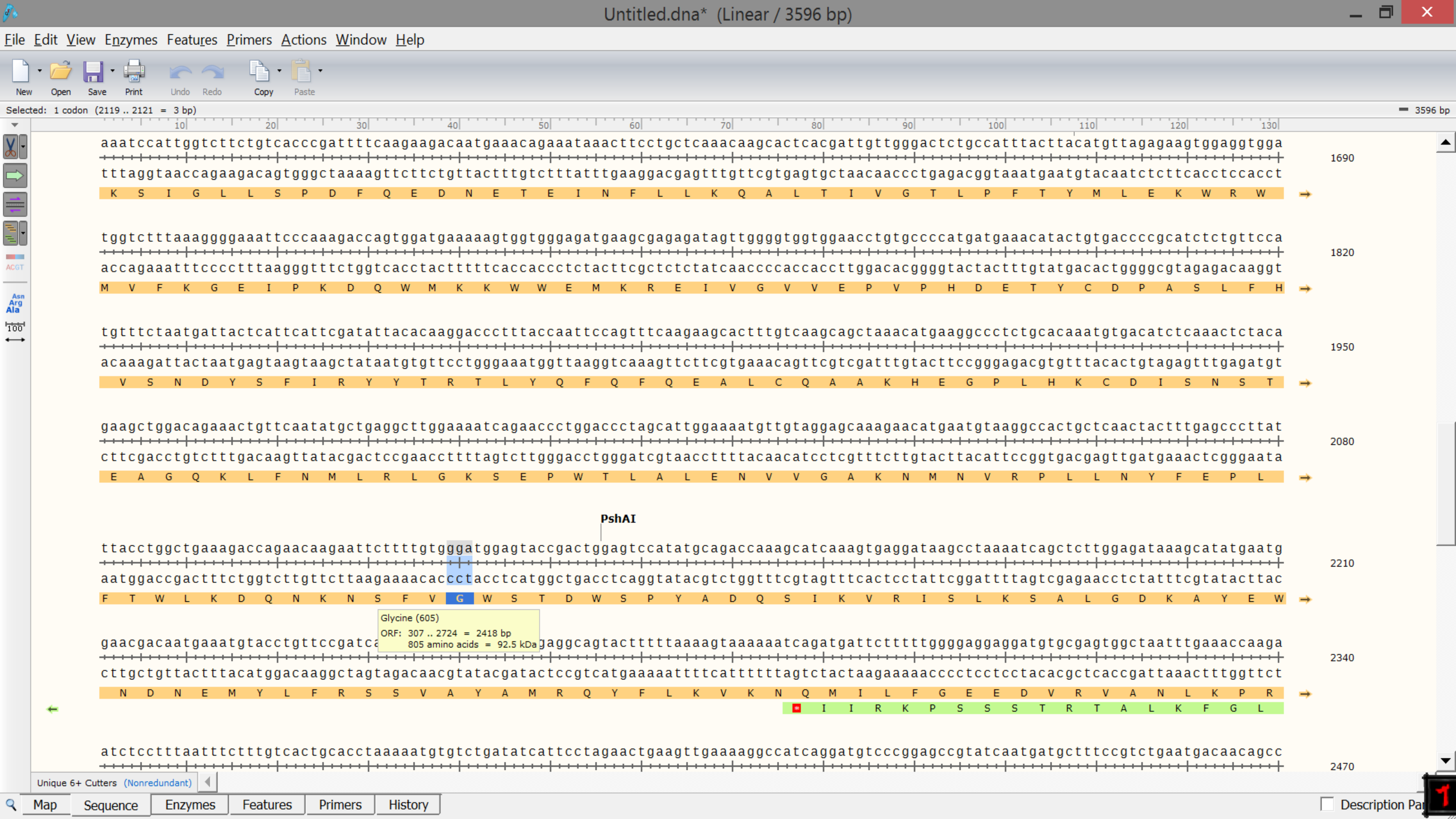
Marchler-Bauer A, Bryant SH (2004), *"CD-Search: protein domain annotations on the fly."*, **Nucleic Acids Res.32**(W)327-331.



Untitled.dna* (Linear / 3596 bp)

Selected: 1 codon (361 .. 363 = 3 bp)   3596 bp

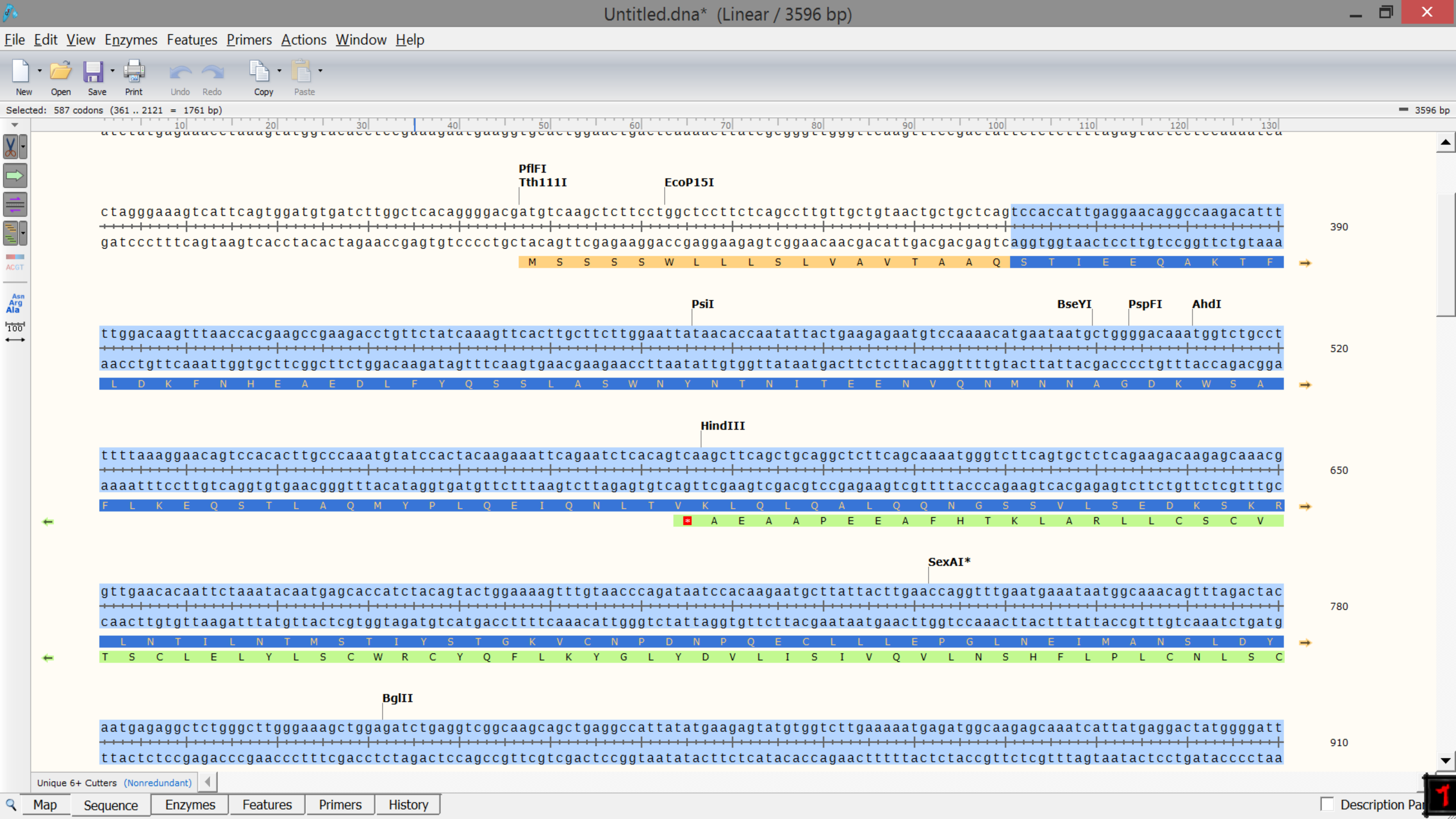
```
tagatactctttggatttcataccatgtggaggctttcttacttccacgtgaccttgactgagttttgaatagcgcccaacccaagttcaaaggctgataagagagaaaatctcatgaggaggttttagt
atctatgagaaacctaaagtatggtacacctccgaaagaatgaaggtgcactggaactgactcaaaacttatcgcgggttgggttcaagtttccgactattctctcttttagagtactcctccaaaatca   260

PflFI
Tth111I   EcoP15I
ctagggaaagtcattcagtggatgtgatcttggctcacaggggacgatgtcaagctcttcctggctccttctcagccttgttgctgtaactgctgctcagtccaccattgaggaacaggccaagacattt
gatccctttcagtaagtcacctacactagaaccgagtgtcccctgctacagttcgagaaggaccgaggaagagtcggaacaacgacattgacgacgagtcaggtggtaactccttgtccggttctgtaaa   390
M S S S S W L L L S L V A V T A A Q S T I E E Q A K T F

Serine (19)
ORF: 307 .. 2724 = 2418 bp
805 amino acids = 92.5 kDa

PsiI   PspFI   AhdI
ttggacaagtttaaccacgaagccgaagacctgttctatcaaagttcacttgcttcttggaattataacaccaatattactgaagagaatgtccaaaacatgaataatgctggggacaaatggtctgcct
aacctgttcaaattggtgcttcggcttctggacaagatagtttcaagtgaacgaagaaccttaatattgtggttataatgacttctcttacaggttttgtacttattacgacccctgtttaccagacgga   520
L D K F N H E A E D L F Y Q S S L A S W N Y N T N I T E E N V Q N M N N A G D K W S A

HindIII
ttttaaaggaacagtccacacttgcccaaatgtatccactacaagaaattcagaatctcacagtcaagcttcagctgcaggctcttcagcaaaatgggtcttcagtgctctcagaagacaagagcaaacg
aaaatttccttgtcaggtgtgaacgggtttacataggtgatgttctttaagtcttagagtgtcagttcgaagtcgacgtccgagaagtcgttttacccagaagtcacgagagtcttctgttctcgtttgc   650
F L K E Q S T L A Q M Y P L Q E I Q N L T V K L Q L Q A L Q Q N G S S V L S E D K S K R
* A E A A P E E A F H T K L A R L L C S C V

SexAI*
gttgaacacaattctaaatacaatgagcaccatctacagtactggaaaagtttgtaacccagataatccacaagaatgcttattacttgaaccaggtttgaatgaaataatggcaaacagtttagactac
caacttgtgttaagatttatgttactcgtggtagatgtcatgaccttttcaaacattgggtctattaggtgttcttacgaataatgaacttggtccaaacttactttattaccgtttgtcaaatctgatg   780
L N T I L N T M S T I Y S T G K V C N P D N P Q E C L L L E P G L N E I M A N S L D Y
T S C L E L Y L S C W R C Y Q F L K Y G L Y D V L I S I V Q V L N S H F L P L C N L S C

BglII
```



Selected: 1 codon (2119 .. 2121 = 3 bp)

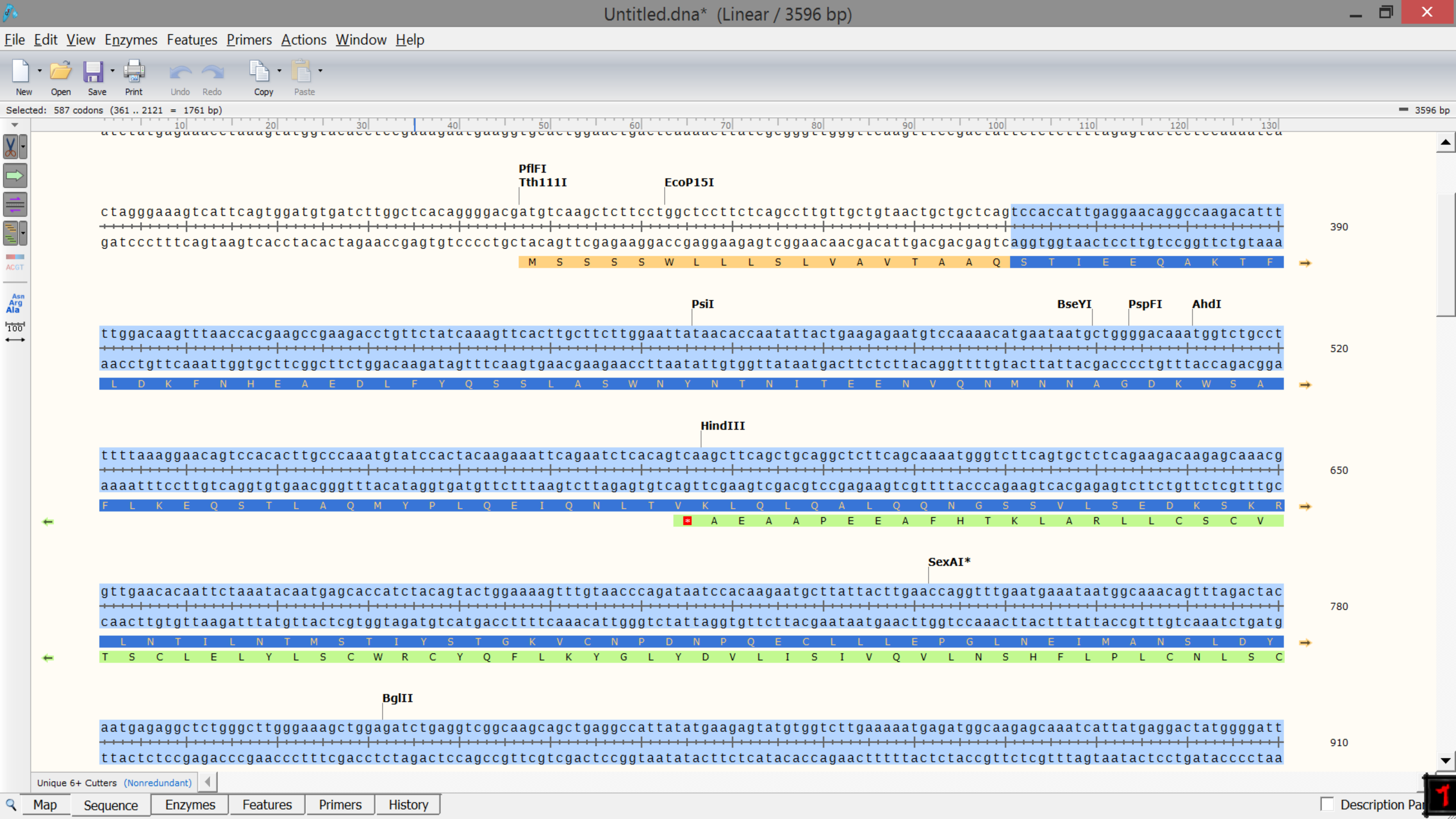
```
aaatccattggtcttctgtcacccgattttcaagaagacaatgaaacagaaataaacttcctgctcaaacaagcactcacgattgttgggactctgccatttacttacatgttagagaagtggaggtgga
tttaggtaaccagaagacagtgggctaaaagttcttctgttactttgtctttatttgaaggacgagtttgttcgtgagtgctaacaaccctgagacggtaaatgaatgtacaatctcttcacctccacct   1690
 K  S  I  G  L  L  S  P  D  F  Q  E  D  N  E  T  E  I  N  F  L  L  K  Q  A  L  T  I  V  G  T  L  P  F  T  Y  M  L  E  K  W  R  W  →

tggtctttaaaggggaaattcccaaagaccagtggatgaaaaagtggtgggagatgaagcgagagatagttggggtggtggaacctgtgccccatgatgaaacatactgtgaccccgcatctctgttcca
accagaaatttcccctttaagggtttctggtcacctactttttcaccaccctctacttcgctctctatcaaccccaccaccttggacacggggtactactttgtatgacactggggcgtagagacaaggt   1820
M  V  F  K  G  E  I  P  K  D  Q  W  M  K  K  W  W  E  M  K  R  E  I  V  G  V  V  E  P  V  P  H  D  E  T  Y  C  D  P  A  S  L  F  H  →

tgtttctaatgattactcattcattcgatattacacaaggaccctttaccaattccagtttcaagaagcactttgtcaagcagctaaacatgaaggccctctgcacaaatgtgacatctcaaactctaca
acaaagattactaatgagtaagtaagctataatgtgttcctgggaaatggttaaggtcaaagttcttcgtgaaacagttcgtcgatttgtacttccgggagacgtgtttacactgtagagtttgagatgt   1950
  V  S  N  D  Y  S  F  I  R  Y  Y  T  R  T  L  Y  Q  F  Q  F  Q  E  A  L  C  Q  A  A  K  H  E  G  P  L  H  K  C  D  I  S  N  S  T  →

gaagctggacagaaactgttcaatatgctgaggcttggaaaatcagaaccctggaccctagcattggaaaatgttgtaggagcaaagaacatgaatgtaaggccactgctcaactactttgagcccttat
cttcgacctgtctttgacaagttatacgactccgaaccttttagtcttgggacctgggatcgtaaccttttacaacatcctcgtttcttgtacttacattccggtgacgagttgatgaaactcgggaata   2080
 E  A  G  Q  K  L  F  N  M  L  R  L  G  K  S  E  P  W  T  L  A  L  E  N  V  V  G  A  K  N  M  N  V  R  P  L  L  N  Y  F  E  P  L  →

                                                                         PshAI
ttacctggctgaaagaccagaacaagaattcttttgtggatggagtaccgactggagtccatatgcagaccaaagcatcaaagtgaggataagcctaaaatcagctcttggagataaagcatatgaatg
aatggaccgactttctggtcttgttcttaagaaaacacctacctcatggctgacctcaggtatacgtctggtttcgtagtttcactcctattcggattttagtcgagaacctctatttcgtatacttac   2210
F  T  W  L  K  D  Q  N  K  N  S  F  V  G  W  S  T  D  W  S  P  Y  A  D  Q  S  I  K  V  R  I  S  L  K  S  A  L  G  D  K  A  Y  E  W  →

gaacgacaatgaaatgtacctgttccgatca                      gaggcagtacttttaaaagtaaaaaatcagatgattcttttggggaggaggatgtgcgagtggctaatttgaaaccaaga
cttgctgttactttacatggacaaggctagtagacaacgtatacgatactccgtcatgaaaaatttcatttttagtctactaagaaaaacccctcctcctacacgctcaccgattaaactttggttct   2340
   N  D  N  E  M  Y  L  F  R  S  S  V  A  Y  A  M  R  Q  Y  F  L  K  V  K  N  Q  M  I  L  F  G  E  D  V  R  V  A  N  L  K  P  R  →
←                                                                                 ■  I  I  R  K  P  S  S  S  T  R  T  A  L  K  F  G  L

atctcctttaatttctttgtcactgcacctaaaaatgtgtctgatatcattcctagaactgaagttgaaaaggccatcaggatgtcccggagccgtatcaatgatgctttccgtctgaatgacaacagcc
                                                                                                                                       2470
```

Glycine (605)
ORF: 307 .. 2724 = 2418 bp
805 amino acids = 92.5 kDa



Selected: 587 codons (361 .. 2121 = 1761 bp)

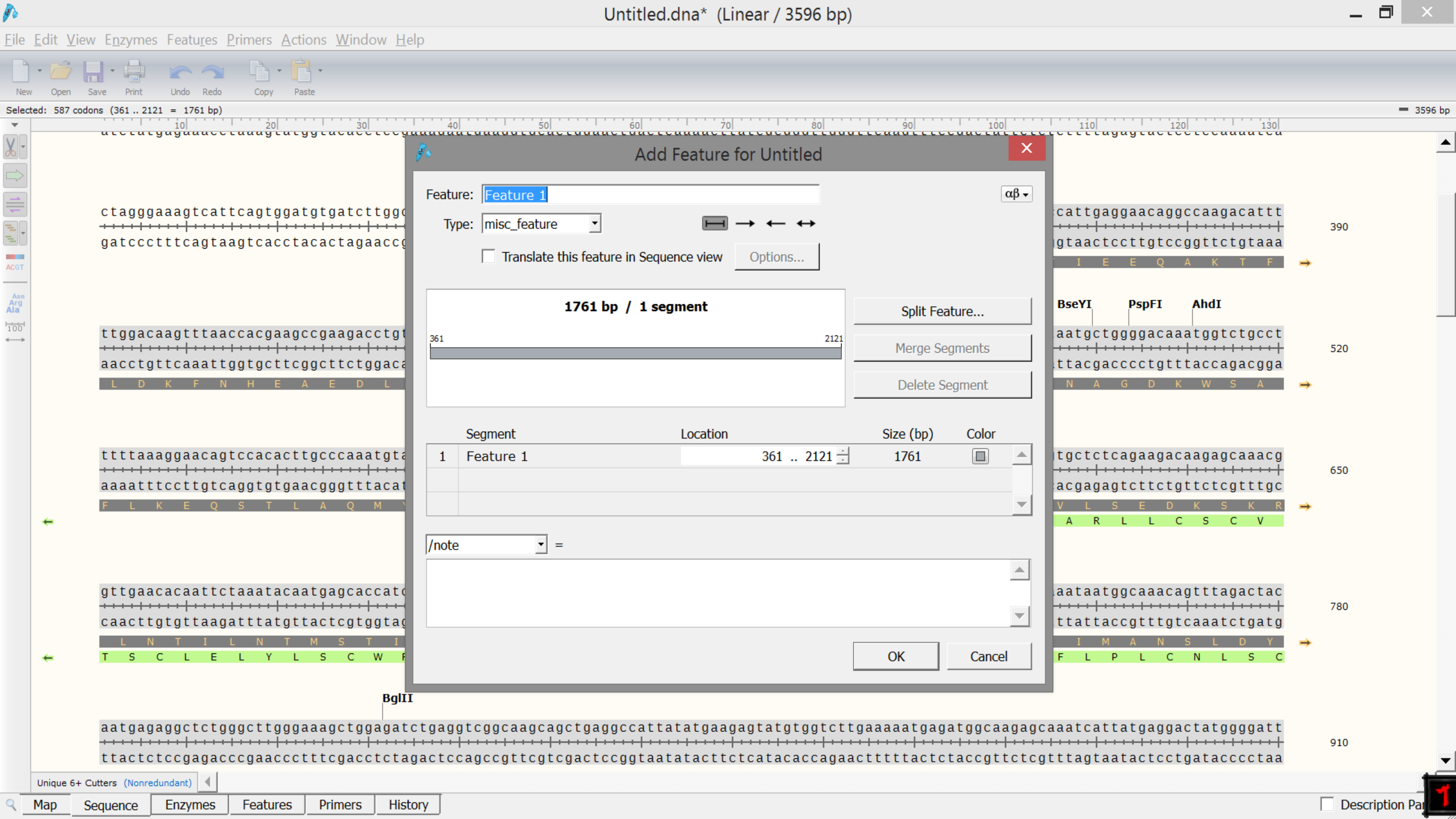
```
                                                   PflFI
                                                   Tth111I         EcoP15I
ctagggaaagtcattcagtggatgtgatcttggctcacaggggacgatgtcaagctcttcctggctccttctcagccttgttgctgtaactgctgctcagtccaccattgaggaacaggccaagacattt
gatccctttcagtaagtcacctacactagaaccgagtgtcccctgctacagttcgagaaggaccgaggaagagtcggaacaacgacattgacgacgagtcaggtggtaactccttgtccggttctgtaaa   390
                                              M  S  S  S  S  W  L  L  L  S  L  V  A  V  T  A  A  Q  S  T  I  E  E  Q  A  K  T  F

                                                                        PsiI                                    BseYI  PspFI  AhdI
ttggacaagtttaaccacgaagccgaagacctgttctatcaaagttcacttgcttcttggaattataacaccaatattactgaagagaatgtccaaaacatgaataatgctggggacaaatggtctgcct
aacctgttcaaattggtgcttcggcttctggacaagatagtttcaagtgaacgaagaaccttaatattgtggttataatgacttctcttacaggttttgtacttattacgacccctgtttaccagacgga   520
 L  D  K  F  N  H  E  A  E  D  L  F  Y  Q  S  S  L  A  S  W  N  Y  N  T  N  I  T  E  E  N  V  Q  N  M  N  N  A  G  D  K  W  S  A

                                                                         HindIII
ttttaaaggaacagtccacacttgcccaaatgtatccactacaagaaattcagaatctcacagtcaagcttcagctgcaggctcttcagcaaaatgggtcttcagtgctctcagaagacaagagcaaacg
aaaatttccttgtcaggtgtgaacgggtttacataggtgatgttctttaagtcttagagtgtcagttcgaagtcgacgtccgagaagtcgttttacccagaagtcacgagagtcttctgttctcgtttgc   650
 F  L  K  E  Q  S  T  L  A  Q  M  Y  P  L  Q  E  I  Q  N  L  T  V  K  L  Q  L  Q  A  L  Q  Q  N  G  S  S  V  L  S  E  D  K  S  K  R
                                                                      *  A  E  A  A  P  E  E  A  F  H  T  K  L  A  R  L  L  C  S  C  V

                                                                                             SexAI*
gttgaacacaattctaaatacaatgagcaccatctacagtactggaaaagtttgtaacccagataatccacaagaatgcttattacttgaaccaggtttgaatgaaataatggcaaacagtttagactac
caacttgtgttaagatttatgttactcgtggtagatgtcatgaccttttcaaacattgggtctattaggtgttcttacgaataatgaacttggtccaaacttactttattaccgtttgtcaaatctgatg   780
  L  N  T  I  L  N  T  M  S  T  I  Y  S  T  G  K  V  C  N  P  D  N  P  Q  E  C  L  L  L  E  P  G  L  N  E  I  M  A  N  S  L  D  Y
T  S  C  L  E  L  Y  L  S  C  W  R  C  Y  Q  F  L  K  Y  G  L  Y  D  V  L  I  S  I  V  Q  V  L  N  S  H  F  L  P  L  C  N  L  S  C

                               BglII
aatgagaggctctgggcttgggaaagctggagatctgaggtcggcaagcagctgaggccattatatgaagagtatgtggtcttgaaaaatgagatggcaagagcaaatcattatgaggactatggggatt
ttactctccgagacccgaaccctttcgacctctagactccagccgttcgtcgactccggtaatatacttctcatacaccagaactttttactctaccgttctcgtttagtaatactcctgataccccctaa   910
```



Selected: 587 codons (361 .. 2121 = 1761 bp)

3596 bp

```
                                             PflFI
                                             Tth111I      EcoP15I
ctagggaaagtcattcagtggatgtgatcttggctcacaggggacgatgtcaagctcttcctggctccttctcagccttgttgctgtaactgctgctcagtccaccattgaggaacaggccaagacattt   390
gatccctttcagtaagtcacctacactagaaccgagtgtcccctgctacagttcgagaaggaccgaggaagagtcggaacaacgacattgacgacgagtcaggtggtaactccttgtccggttctgtaaa
                                             M  S  S  S  S  W  L  L  L  S  L  V  A  V  T  A  A  Q  S  T  I  E  E  Q  A  K  T  F  →

                                                                PsiI                                                 BseYI   PspFI  AhdI
ttggacaagtttaaccacgaagccgaagacctgttctatcaaagttcacttgcttcttggaattataacaccaatattactgaagagaatgtccaaaacatgaataatgctggggacaaatggtctgcct   520
aacctgttcaaattggtgcttcggcttctggacaagatagtttcaagtgaacgaagaaccttaatattgtggttataatgacttctcttacaggttttgtacttattacgacccctgtttaccagacgga
 L  D  K  F  N  H  E  A  E  D  L  F  Y  Q  S  S  L  A  S  W  N  Y  N  T  N  I  T  E  E  N  V  Q  N  M  N  N  A  G  D  K  W  S  A  →

                                                                 HindIII
ttttaaaggaacagtccacacttgcccaaatgtatccactacaagaaattcagaatctcacagtcaagcttcagctgcaggctcttcagcaaaatgggtcttcagtgctctcagaagacaagagcaaacg   650
aaaatttccttgtcaggtgtgaacgggtttacataggtgatgttctttaagtcttagagtgtcagttcgaagtcgacgtccgagaagtcgttttacccagaagtcacgagagtcttctgttctcgtttgc
F  L  K  E  Q  S  T  L  A  Q  M  Y  P  L  Q  E  I  Q  N  L  T  V  K  L  Q  L  Q  A  L  Q  Q  N  G  S  S  V  L  S  E  D  K  S  K  R  →
←                                                              *  A  E  A  A  P  E  E  A  F  H  T  K  L  A  R  L  L  C  S  C  V

                                                                                    SexAI*
gttgaacacaattctaaatacaatgagcaccatctacagtactggaaaagtttgtaacccagataatccacaagaatgcttattacttgaaccaggtttgaatgaaataatggcaaacagtttagactac   780
caacttgtgttaagatttatgttactcgtggtagatgtcatgaccttttcaaacattgggtctattaggtgttcttacgaataatgaacttggtccaaacttactttattaccgtttgtcaaatctgatg
  L  N  T  I  L  N  T  M  S  T  I  Y  S  T  G  K  V  C  N  P  D  N  P  Q  E  C  L  L  L  E  P  G  L  N  E  I  M  A  N  S  L  D  Y  →
← T  S  C  L  E  L  Y  L  S  C  W  R  C  Y  Q  F  L  K  Y  G  L  Y  D  V  L  I  S  I  V  Q  V  L  N  S  H  F  L  P  L  C  N  L  S  C

                               BglII
aatgagaggctctgggcttgggaaagctggagatctgaggtcggcaagcagctgaggccattatatgaagagtatgtggtcttgaaaaatgagatggcaagagcaaatcattatgaggactatggggatt   910
ttactctccgagacccgaaccctttcgacctctagactccagccgttcgtcgactccggtaatatacttctcatacaccagaacttttttactctaccgttctcgtttagtaatactcctgatacccctaa
```



Untitled.dna* (Linear / 3596 bp)

Selected: 587 codons (361 .. 2121 = 1761 bp)

Add Feature for Untitled

Feature: Feature 1

Type: misc_feature

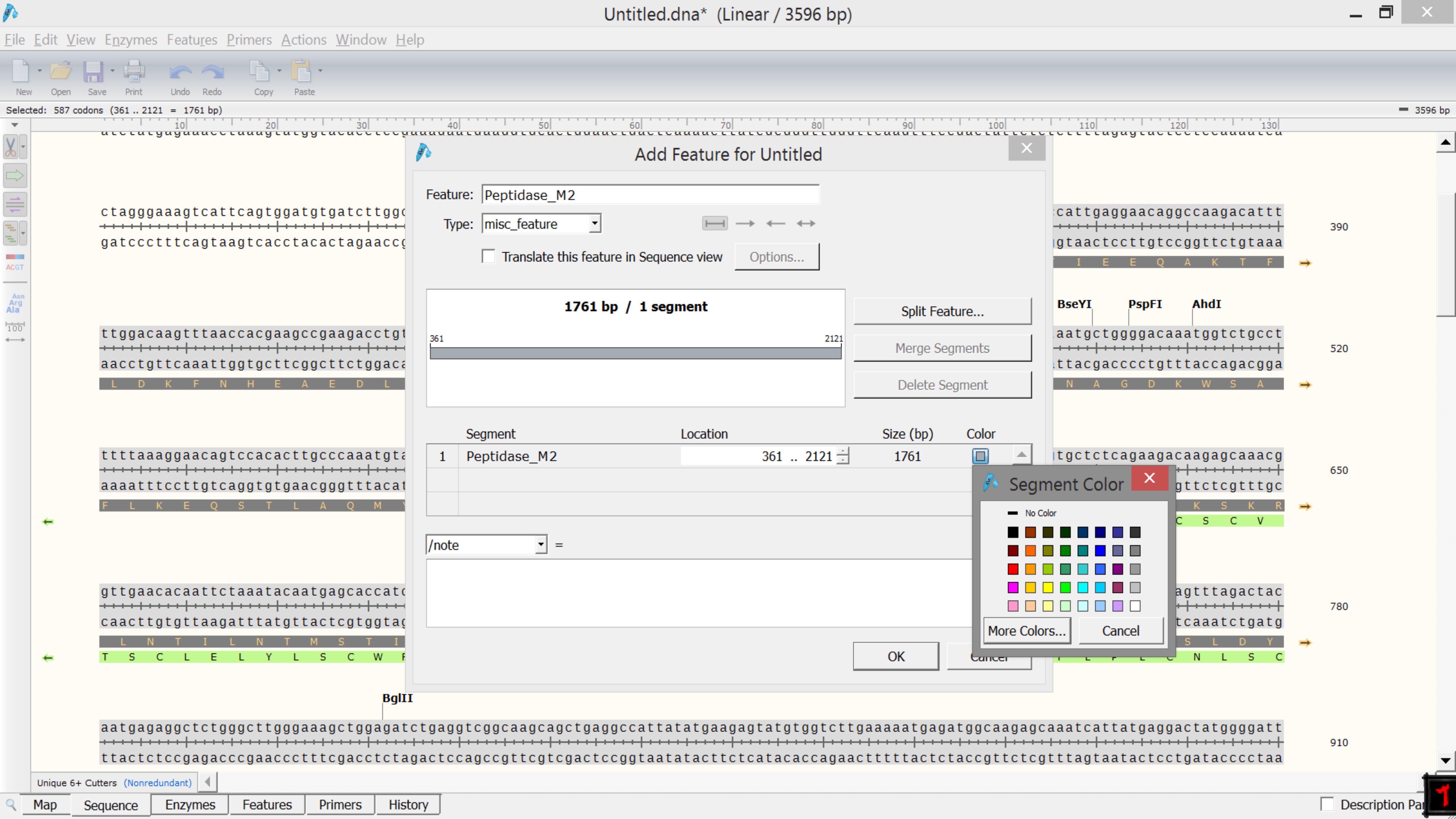
☐ Translate this feature in Sequence view

1761 bp / 1 segment

361 … 2121

	Segment	Location	Size (bp)	Color
1	Feature 1	361 .. 2121	1761	

/note =



Untitled.dna* (Linear / 3596 bp)

Selected: 587 codons (361 .. 2121 = 1761 bp)

Add Feature for Untitled

Feature: Peptidase_M2

Type: misc_feature

☐ Translate this feature in Sequence view

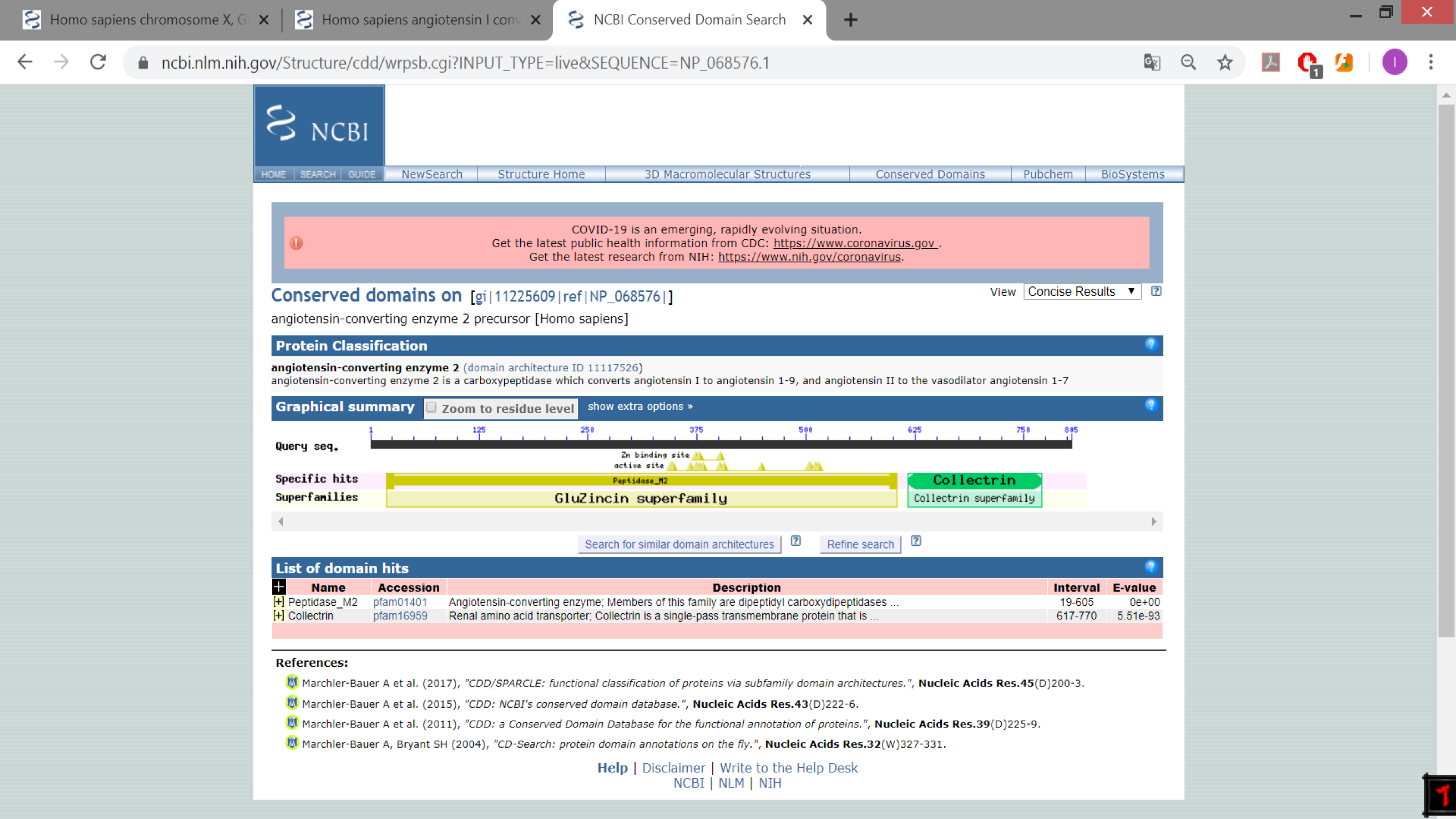
1761 bp / 1 segment

361 2121

	Segment	Location	Size (bp)	Color
1	Peptidase_M2	361 .. 2121	1761	▢

/note =

Segment Color



COVID-19 is an emerging, rapidly evolving situation.
Get the latest public health information from CDC: https://www.coronavirus.gov .
Get the latest research from NIH: https://www.nih.gov/coronavirus.

Conserved domains on [gi|11225609|ref|NP_068576|]

angiotensin-converting enzyme 2 precursor [Homo sapiens]

View: Concise Results ▼

Protein Classification

angiotensin-converting enzyme 2 (domain architecture ID 11117526)
angiotensin-converting enzyme 2 is a carboxypeptidase which converts angiotensin I to angiotensin 1-9, and angiotensin II to the vasodilator angiotensin 1-7

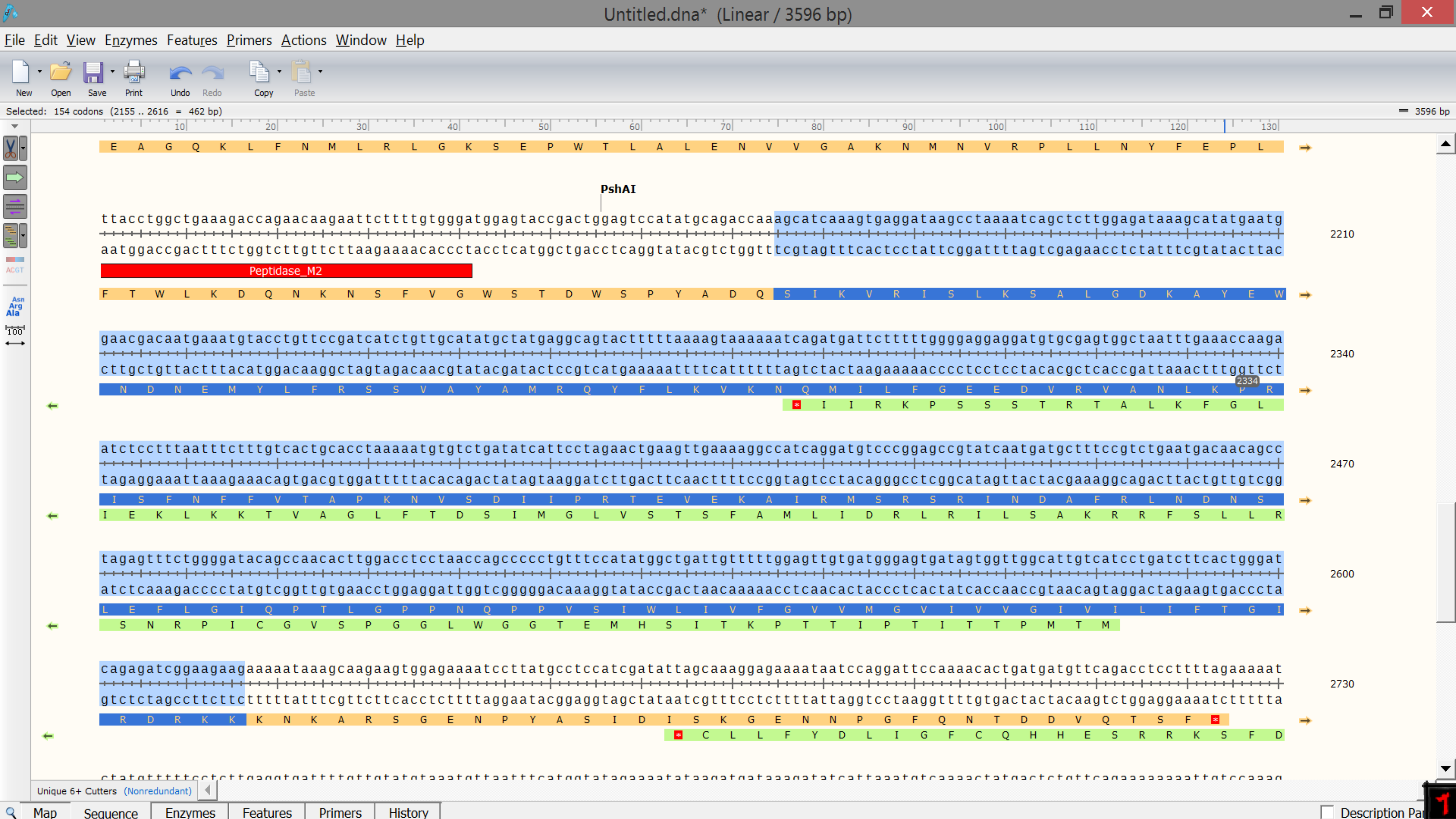
Graphical summary

List of domain hits

	Name	Accession	Description	Interval	E-value
[+]	Peptidase_M2	pfam01401	Angiotensin-converting enzyme; Members of this family are dipeptidyl carboxydipeptidases ...	19-605	0e+00
[+]	Collectrin	pfam16959	Renal amino acid transporter; Collectrin is a single-pass transmembrane protein that is ...	617-770	5.51e-93

References:

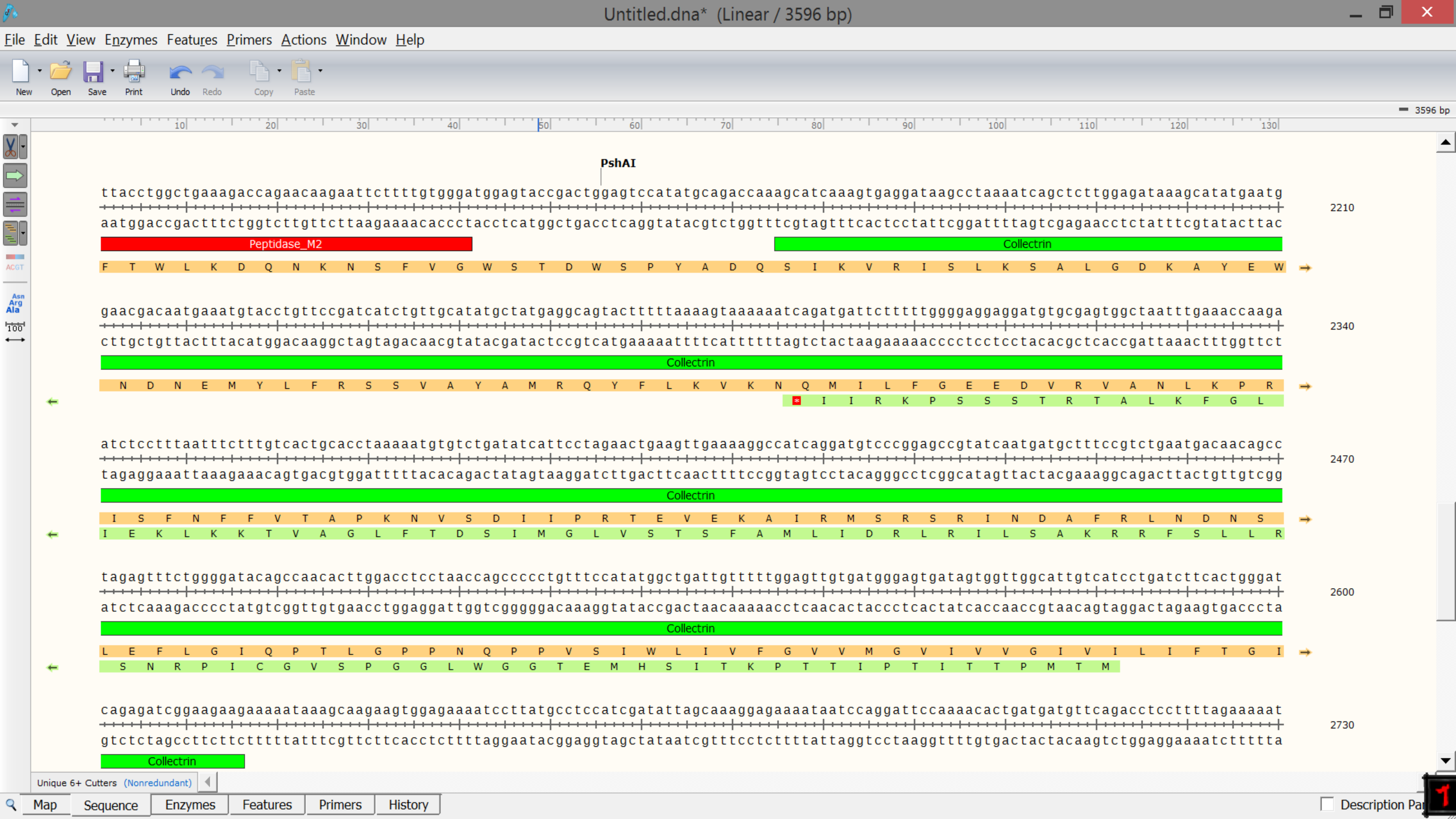
- Marchler-Bauer A et al. (2017), *"CDD/SPARCLE: functional classification of proteins via subfamily domain architectures."*, **Nucleic Acids Res.45**(D)200-3.
- Marchler-Bauer A et al. (2015), *"CDD: NCBI's conserved domain database."*, **Nucleic Acids Res.43**(D)222-6.
- Marchler-Bauer A et al. (2011), *"CDD: a Conserved Domain Database for the functional annotation of proteins."*, **Nucleic Acids Res.39**(D)225-9.
- Marchler-Bauer A, Bryant SH (2004), *"CD-Search: protein domain annotations on the fly."*, **Nucleic Acids Res.32**(W)327-331.



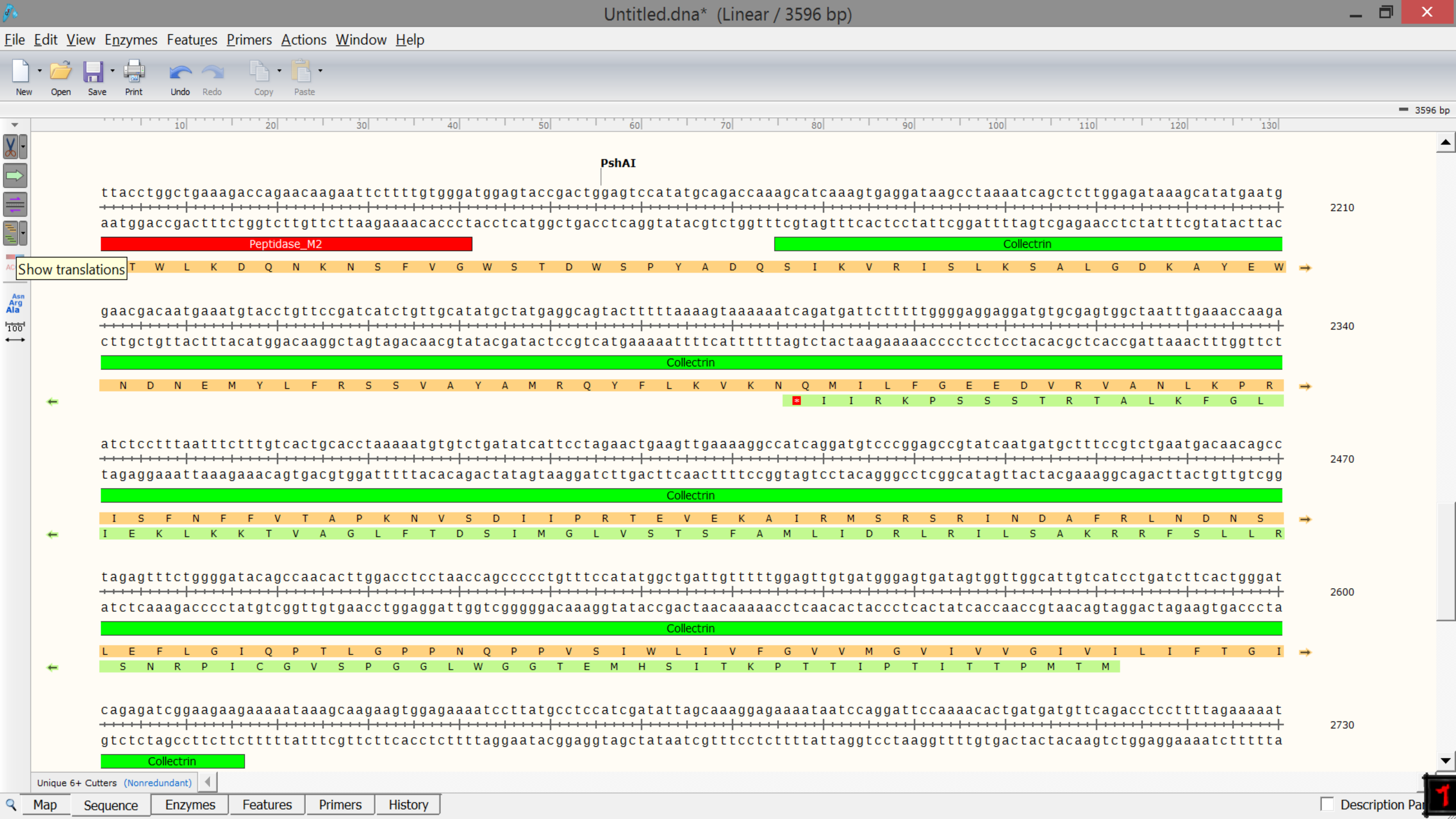
Selected: 154 codons (2155 .. 2616 = 462 bp)

3596 bp

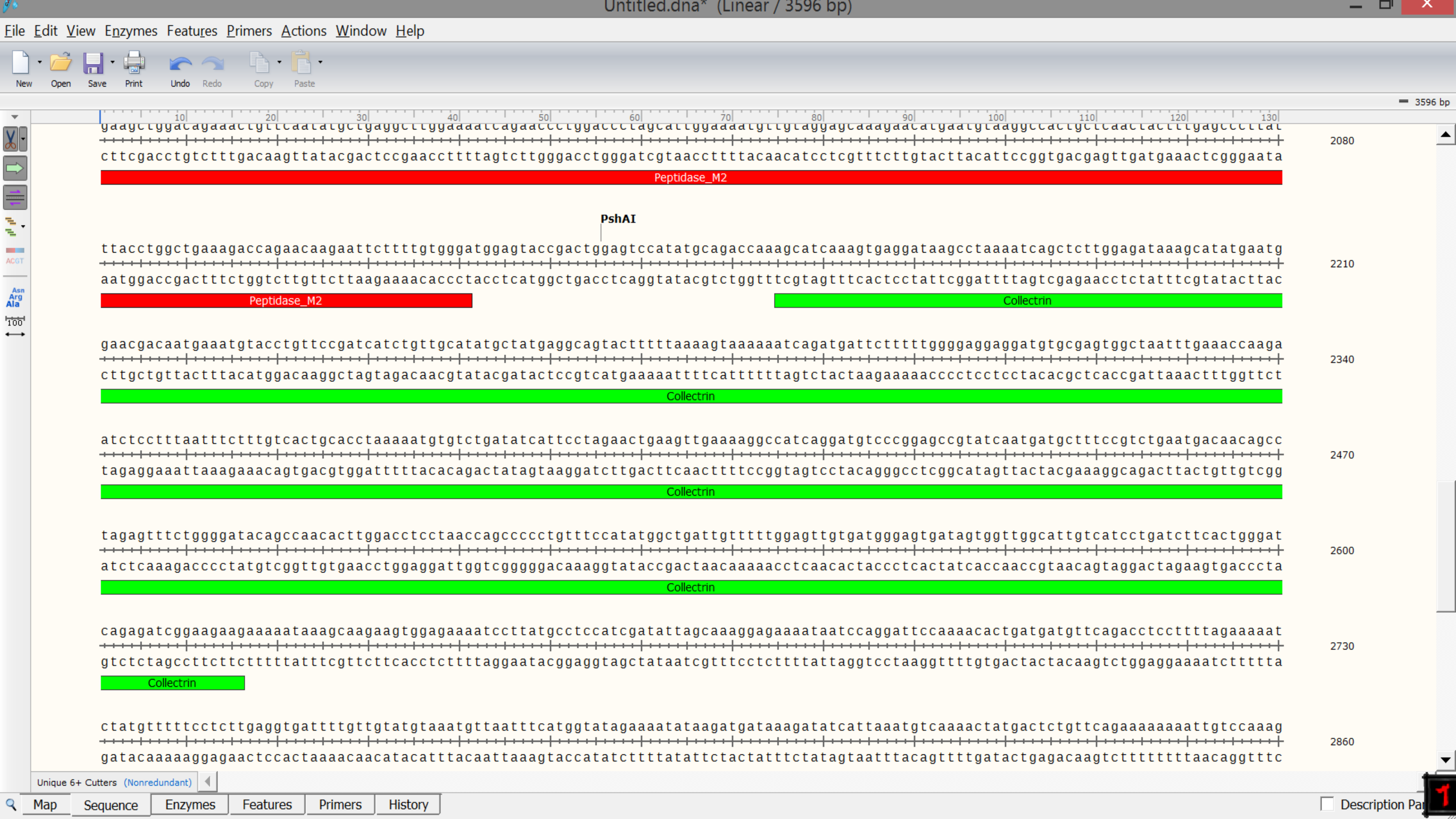
```
         10        20        30        40        50        60        70        80        90       100       110       120       130
E  A  G  Q  K  L  F  N  M  L  R  L  G  K  S  E  P  W  T  L  A  L  E  N  V  V  G  A  K  N  M  N  V  R  P  L  L  N  Y  F  E  P  L

ttacctggctgaaagaccagaacaagaattcttttgtgggatggagtaccgactggagtccatatgcagaccaaagcatcaaagtgaggataagcctaaaatcagctcttggagataaagcatatgaatg   2210
aatggaccgactttctggtcttgttcttaagaaaacaccctacctcatggctgacctcaggtatacgtctggtttcgtagtttcactcctattcggattttagtcgagaacctctatttcgtatacttac
F  T  W  L  K  D  Q  N  K  N  S  F  V  G  W  S  T  D  W  S  P  Y  A  D  Q  S  I  K  V  R  I  S  L  K  S  A  L  G  D  K  A  Y  E  W

gaacgacaatgaaatgtacctgttccgatcatctgttgcatatgctatgaggcagtactttttaaaagtaaaaaatcagatgattctttttggggaggaggatgtgcgagtggctaatttgaaaccaaga   2340
cttgctgttactttacatggacaaggctagtagacaacgtatacgatactccgtcatgaaaaatttttcattttttagtctactaagaaaaacccctcctcctacacgctcaccgattaaactttggttct
 N  D  N  E  M  Y  L  F  R  S  S  V  A  Y  A  M  R  Q  Y  F  L  K  V  K  N  Q  M  I  L  F  G  E  E  D  V  R  V  A  N  L  K  P  R
                                                                                * I  I  R  K  P  S  S  S  T  R  T  A  L  K  F  G  L

atctcctttaatttctttgtcactgcacctaaaaatgtgtctgatatcattcctagaactgaagttgaaaaggccatcaggatgtcccggagccgtatcaatgatgctttccgtctgaatgacaacagcc   2470
tagaggaaattaaagaaacagtgacgtggatttttacacagactatagtaaggatcttgacttcaacttttccggtagtcctacagggcctcggcatagttactacgaaaggcagacttactgttgtcgg
 I  S  F  N  F  F  V  T  A  P  K  N  V  S  D  I  I  P  R  T  E  V  E  K  A  I  R  M  S  R  S  R  I  N  D  A  F  R  L  N  D  N  S
I  E  K  L  K  K  T  V  A  G  L  F  T  D  S  I  M  G  L  V  S  T  S  F  A  M  L  I  D  R  L  R  I  L  S  A  K  R  R  F  S  L  L  R

tagagtttctggggatacagccaacacttggacctcctaaccagccccctgtttccatatggctgattgtttttggagttgtgatgggagtgatagtggttggcattgtcatcctgatcttcactgggat   2600
atctcaaagacccctatgtcggttgtgaacctggaggattggtcgggggacaaaggtataccgactaacaaaaacctcaacactaccctcactatcaccaaccgtaacagtaggactagaagtgaccctа
L  E  F  L  G  I  Q  P  T  L  G  P  P  N  Q  P  P  V  S  I  W  L  I  V  F  G  V  V  M  G  V  I  V  V  G  I  V  I  L  I  F  T  G  I
  S  N  R  P  I  C  G  V  S  P  G  G  L  W  G  G  T  E  M  H  S  I  T  K  P  T  T  I  P  T  I  T  T  P  M  T  M

cagagatcggaagaagaaaaataaagcaagaagtggagaaaatccttatgcctccatcgatattagcaaaggagaaaataatccaggattccaaaacactgatgatgttcagacctccttttagaaaaat   2730
gtctctagccttcttcttttatttcgttcttcacctcttttaggaatacggaggtagctataatcgtttcctcttttattaggtcctaaggttttgtgactactacaagtctggaggaaaatcttttta
 R  D  R  K  K  K  N  K  A  R  S  G  E  N  P  Y  A  S  I  D  I  S  K  G  E  N  N  P  G  F  Q  N  T  D  D  V  Q  T  S  F  *
                                                                    *  C  L  L  F  Y  D  L  I  G  F  C  Q  H  H  E  S  R  R  K  S  F  D
```



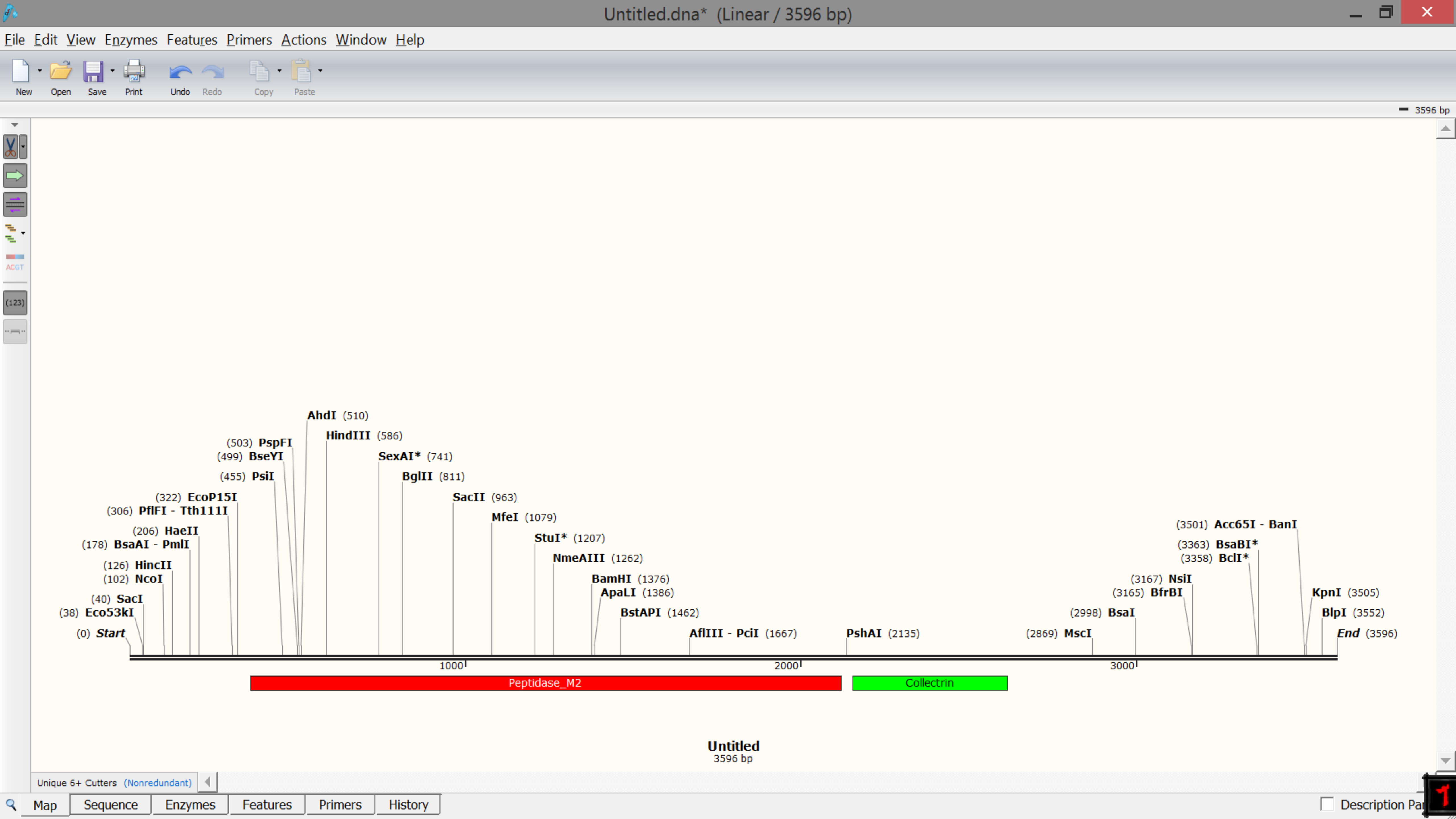
```
                                                                                      PshAI
ttacctggctgaaagaccagaacaagaattcttttgtgggatggagtaccgactggagtccatatgcagaccaaagcatcaaagtgaggataagcctaaaatcagctcttggagataaagcatatgaatg   2210
aatggaccgactttctggtcttgttcttaagaaaacaccctacctcatggctgacctcaggtatacgtctggtttcgtagtttcactcctattcggattttagtcgagaacctctatttcgtatacttac
Peptidase_M2                                                                                Collectrin
F  T  W  L  K  D  Q  N  K  N  S  F  V  G  W  S  T  D  W  S  P  Y  A  D  Q  S  I  K  V  R  I  S  L  K  S  A  L  G  D  K  A  Y  E  W  →

gaacgacaatgaaatgtacctgttccgatcatctgttgcatatgctatgaggcagtactttttaaaagtaaaaaatcagatgattctttttggggaggaggatgtgcgagtggctaatttgaaaccaaga   2340
cttgctgttactttacatggacaaggctagtagacaacgtatacgatactccgtcatgaaaaattttcatttttttagtctactaagaaaaacccctcctcctacacgctcaccgattaaactttggttct
Collectrin
 N  D  N  E  M  Y  L  F  R  S  S  V  A  Y  A  M  R  Q  Y  F  L  K  V  K  N  Q  M  I  L  F  G  E  E  D  V  R  V  A  N  L  K  P  R  →
←                                                                                        *  I  I  R  K  P  S  S  S  T  R  T  A  L  K  F  G  L

atctcctttaatttctttgtcactgcacctaaaaatgtgtctgatatcattcctagaactgaagttgaaaaggccatcaggatgtcccggagccgtatcaatgatgctttccgtctgaatgacaacagcc   2470
tagaggaaattaaagaaacagtgacgtggatttttacacagactatagtaaggatcttgacttcaacttttccggtagtcctacagggcctcggcatagttactacgaaaggcagacttactgttgtcgg
Collectrin
 I  S  F  N  F  F  V  T  A  P  K  N  V  S  D  I  I  P  R  T  E  V  E  K  A  I  R  M  S  R  S  R  I  N  D  A  F  R  L  N  D  N  S  →
← I  E  K  L  K  K  T  V  A  G  L  F  T  D  S  I  M  G  L  V  S  T  S  F  A  M  L  I  D  R  L  R  I  L  S  A  K  R  R  F  S  L  L  R

tagagtttctggggatacagccaacacttggacctcctaaccagccccctgtttccatatggctgattgtttttggagttgtgatgggagtgatagtggttggcattgtcatcctgatcttcactgggat   2600
atctcaaagacccctatgtcggttgtgaacctggaggattggtcgggggacaaaggtataccgactaacaaaaacctcaacactaccctcactatcaccaaccgtaacagtaggactagaagtgacccta
Collectrin
L  E  F  L  G  I  Q  P  T  L  G  P  P  N  Q  P  P  V  S  I  W  L  I  V  F  G  V  V  M  G  V  I  V  V  G  I  V  I  L  I  F  T  G  I  →
←  S  N  R  P  I  C  G  V  S  P  G  G  L  W  G  G  T  E  M  H  S  I  T  K  P  T  T  I  P  T  I  T  T  P  M  T  M

cagagatcggaagaagaaaaataaagcaagaagtggagaaaatccttatgcctccatcgatattagcaaaggagaaaataatccaggattccaaaacactgatgatgttcagacctccttttagaaaaat   2730
gtctctagccttcttctttttatttcgttcttcacctcttttaggaatacggaggtagctataatcgtttcctcttttattaggtcctaaggttttgtgactactacaagtctggaggaaaatcttttta
Collectrin
```

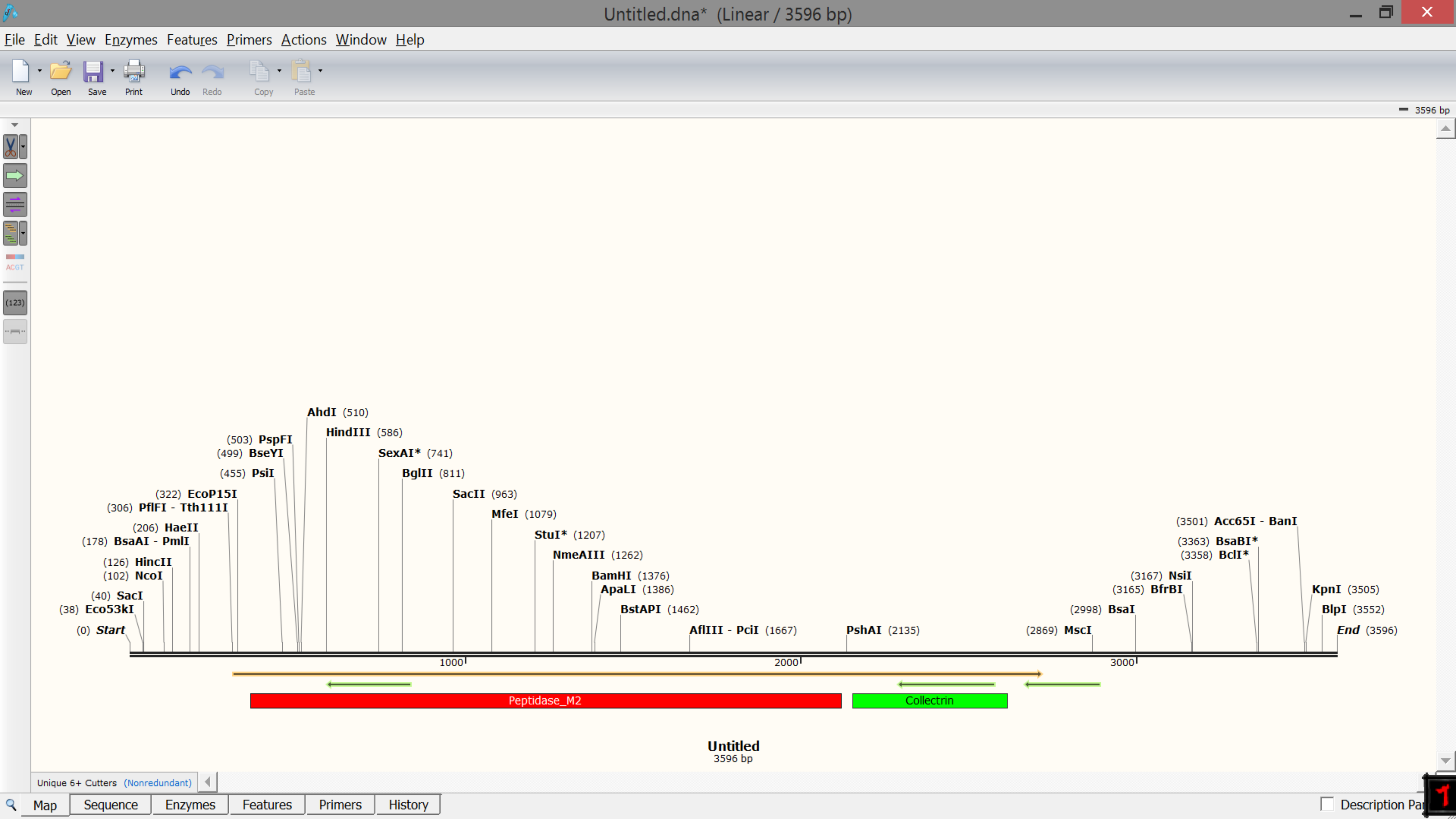


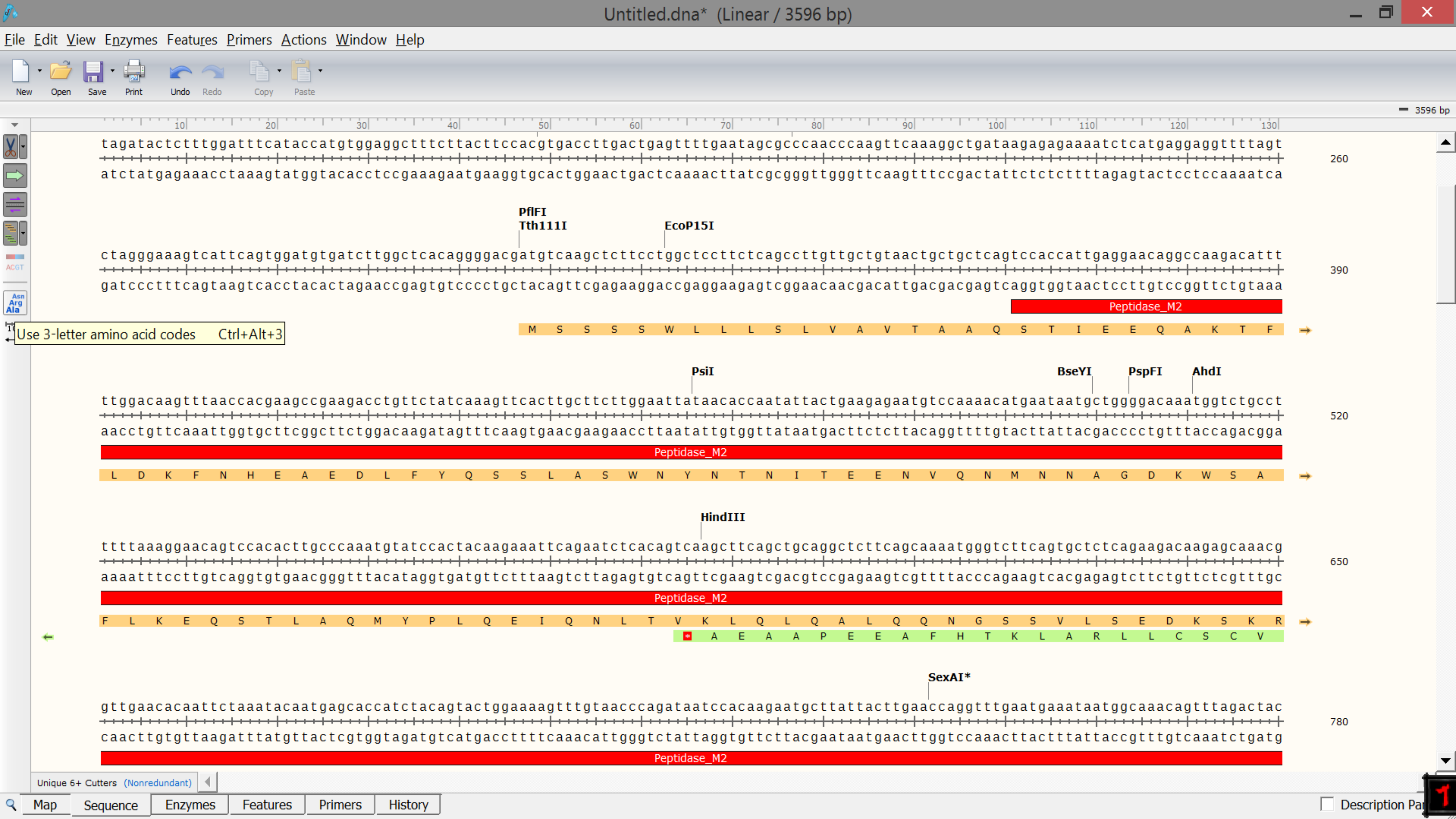
```
                                            PshAI
ttacctggctgaaagaccagaacaagaattcttttgtgggatggagtaccgactggagtccatatgcagaccaaagcatcaaagtgaggataagcctaaaatcagctcttggagataaagcatatgaatg   2210
aatggaccgactttctggtcttgttcttaagaaaacaccctacctcatggctgacctcaggtatacgtctggtttcgtagtttcactcctattcggattttagtcgagaacctctatttcgtatacttac
Peptidase_M2                                                         Collectrin
  T  W  L  K  D  Q  N  K  N  S  F  V  G  W  S  T  D  W  S  P  Y  A  D  Q  S  I  K  V  R  I  S  L  K  S  A  L  G  D  K  A  Y  E  W  →

gaacgacaatgaaatgtacctgttccgatcatctgttgcatatgctatgaggcagtactttttaaaagtaaaaaatcagatgattctttttggggaggaggatgtgcgagtggctaatttgaaaccaaga   2340
cttgctgttactttacatggacaaggctagtagacaacgtatacgatactccgtcatgaaaaattttcattttttagtctactaagaaaaacccctcctcctacacgctcaccgattaaactttggttct
Collectrin
 N  D  N  E  M  Y  L  F  R  S  S  V  A  Y  A  M  R  Q  Y  F  L  K  V  K  N  Q  M  I  L  F  G  E  E  D  V  R  V  A  N  L  K  P  R  →
←                                                                             *  I  I  R  K  P  S  S  S  T  R  T  A  L  K  F  G  L

atctcctttaatttctttgtcactgcacctaaaaatgtgtctgatatcattcctagaactgaagttgaaaaggccatcaggatgtcccggagccgtatcaatgatgctttccgtctgaatgacaacagcc   2470
tagaggaaattaaagaaacagtgacgtggatttttacacagactatagtaaggatcttgacttcaacttttccggtagtcctacagggcctcggcatagttactacgaaaggcagacttactgttgtcgg
Collectrin
 I  S  F  N  F  F  V  T  A  P  K  N  V  S  D  I  I  P  R  T  E  V  E  K  A  I  R  M  S  R  S  R  I  N  D  A  F  R  L  N  D  N  S  →
←I  E  K  L  K  K  T  V  A  G  L  F  T  D  S  I  M  G  L  V  S  T  S  F  A  M  L  I  D  R  L  R  I  L  S  A  K  R  R  F  S  L  L  R

tagagtttctggggatacagccaacacttggacctcctaaccagccccctgtttccatatggctgattgtttttggagttgtgatgggagtgatagtggttggcattgtcatcctgatcttcactgggat   2600
atctcaaagacccctatgtcggttgtgaacctggaggattggtcgggggacaaaggtataccgactaacaaaaacctcaacactaccctcactatcaccaaccgtaacagtaggactagaagtgaccctа
Collectrin
L  E  F  L  G  I  Q  P  T  L  G  P  P  N  Q  P  P  V  S  I  W  L  I  V  F  G  V  V  M  G  V  I  V  V  G  I  V  I  L  I  F  T  G  I  →
←  S  N  R  P  I  C  G  V  S  P  G  G  L  W  G  G  T  E  M  H  S  I  T  K  P  T  T  I  P  T  I  T  T  P  M  T  M

cagagatcggaagaagaaaaataaagcaagaagtggagaaaatccttatgcctccatcgatattagcaaaggagaaaataatccaggattccaaaacactgatgatgttcagacctccttttagaaaaat   2730
gtctctagccttcttctttttatttcgttcttcacctcttttaggaatacggaggtagctataatcgtttcctcttttattaggtcctaaggttttgtgactactacaagtctggaggaaaatcttttta
Collectrin
```



```
gaagctggacagaaactgttcaatatgctgaggcttggaaaatcagaaccctggaccctagcattggaaaatgttgtaggagcaaagaacatgaatgtaaggccactgctcaactactttgagcccttat   2080
cttcgacctgtctttgacaagttatacgactccgaaccttttagtcttgggacctgggatcgtaacctttttacaacatcctcgtttcttgtacttacattccggtgacgagttgatgaaactcgggaata
Peptidase_M2

PshAI
ttacctggctgaaagaccagaacaagaattcttttgtgggatggagtaccgactggagtccatatgcagaccaaagcatcaaagtgaggataagcctaaaatcagctcttggagataaagcatatgaatg   2210
aatggaccgactttctggtcttgttcttaagaaaacaccctacctcatggctgacctcaggtatacgtctggtttcgtagtttcactcctattcggattttagtcgagaacctctatttcgtatacttac
Peptidase_M2                                                            Collectrin

gaacgacaatgaaatgtacctgttccgatcatctgttgcatatgctatgaggcagtactttttaaaagtaaaaaatcagatgattctttttggggaggaggatgtgcgagtggctaatttgaaaccaaga   2340
cttgctgttactttacatggacaaggctagtagacaacgtatacgatactccgtcatgaaaaattttcattttttagtctactaagaaaaacccctcctcctacacgctcaccgattaaactttggttct
Collectrin

atctcctttaatttctttgtcactgcacctaaaaatgtgtctgatatcattcctagaactgaagttgaaaaggccatcaggatgtcccggagccgtatcaatgatgctttccgtctgaatgacaacagcc   2470
tagaggaaattaaagaaacagtgacgtggatttttacacagactatagtaaggatcttgacttcaacttttccggtagtcctacagggcctcggcatagttactacgaaaggcagacttactgttgtcgg
Collectrin

tagagtttctggggatacagccaacacttggacctcctaaccagccccctgtttccatatggctgattgtttttggagttgtgatgggagtgatagtggttggcattgtcatcctgatcttcactgggat   2600
atctcaaagacccctatgtcggttgtgaacctggaggattggtcgggggacaaaggtataccgactaacaaaaacctcaacactaccctcactatcaccaaccgtaacagtaggactagaagtgaccta
Collectrin

cagagatcggaagaagaaaaataaagcaagaagtggagaaaatccttatgcctccatcgatattagcaaaggagaaaataatccaggattccaaaacactgatgatgttcagacctccttttagaaaaat   2730
gtctctagccttcttctttttatttcgttcttcacctcttttaggaatacggaggtagctataatcgtttcctcttttattaggtcctaaggttttgtgactactacaagtctggaggaaatctttttta
Collectrin

ctatgtttttcctcttgaggtgattttgttgtatgtaaatgttaatttcatggtatagaaaatataagatgataaagatatcattaaatgtcaaaactatgactctgttcagaaaaaaaattgtccaaag   2860
gatacaaaaaggagaactccactaaaacaacatacatttacaattaaagtaccatatcttttatattctactatttctatagtaatttacagttttgatactgagacaagtcttttttttaacaggtttc
```

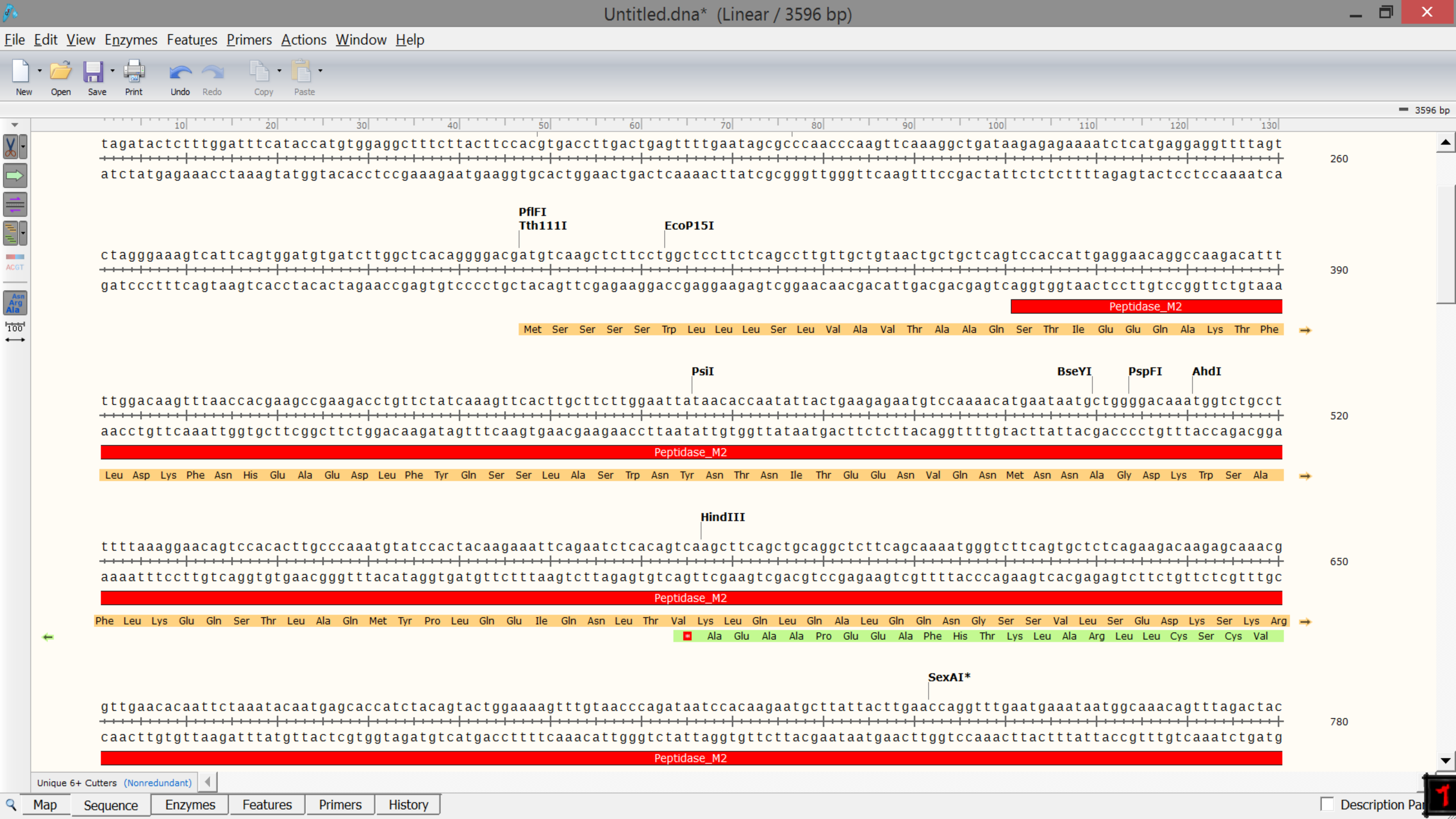
Untitled.dna* (Linear / 3596 bp)
File
Edit
View
Enzymes
Features
Primers
Actions
Window
Help
New
Open
Save
Print
Undo
Redo
Copy
Paste
3596 bp
ACGT
(123)
AhdI (510)
HindIII (586)
(503) PspFI
(499) BseYI
SexAI* (741)
(455) PsiI
BglII (811)
SacII (963)
(322) EcoP15I
(306) PflFI - Tth111I
MfeI (1079)
(206) HaeII
StuI* (1207)
(178) BsaAI - PmlI
NmeAIII (1262)
(126) HincII
BamHI (1376)
(102) NcoI
ApaLI (1386)
(40) SacI
(38) Eco53kI
BstAPI (1462)
(0) Start
AflIII - PciI (1667)
PshAI (2135)
(2869) MscI
(2998) BsaI
(3167) NsiI
(3165) BfrBI
(3501) Acc65I - BanI
(3363) BsaBI*
(3358) BclI*
KpnI (3505)
BlpI (3552)
End (3596)
1000
2000
3000
Peptidase_M2
Collectrin
Untitled
3596 bp
Unique 6+ Cutters (Nonredundant)
Map
Sequence
Enzymes
Features
Primers
History
Description Pa


Untitled.dna* (Linear / 3596 bp)
File
Edit
View
Enzymes
Features
Primers
Actions
Window
Help
New
Open
Save
Print
Undo
Redo
Copy
Paste
3596 bp
ACGT
(123)
AhdI (510)
HindIII (586)
(503) PspFI
(499) BseYI
SexAI* (741)
(455) PsiI
BglII (811)
(322) EcoP15I
SacII (963)
(306) PflFI - Tth111I
MfeI (1079)
(206) HaeII
StuI* (1207)
(178) BsaAI - PmlI
NmeAIII (1262)
(126) HincII
BamHI (1376)
(102) NcoI
ApaLI (1386)
(40) SacI
BstAPI (1462)
(38) Eco53kI
(0) Start
AflIII - PciI (1667)
PshAI (2135)
(3501) Acc65I - BanI
(3363) BsaBI*
(3358) BclI*
(3167) NsiI
(3165) BfrBI
KpnI (3505)
(2998) BsaI
BlpI (3552)
(2869) MscI
End (3596)
1000
2000
3000
Peptidase_M2
Collectrin
Untitled
3596 bp
Unique 6+ Cutters (Nonredundant)
Map
Sequence
Enzymes
Features
Primers
History
Description Pa

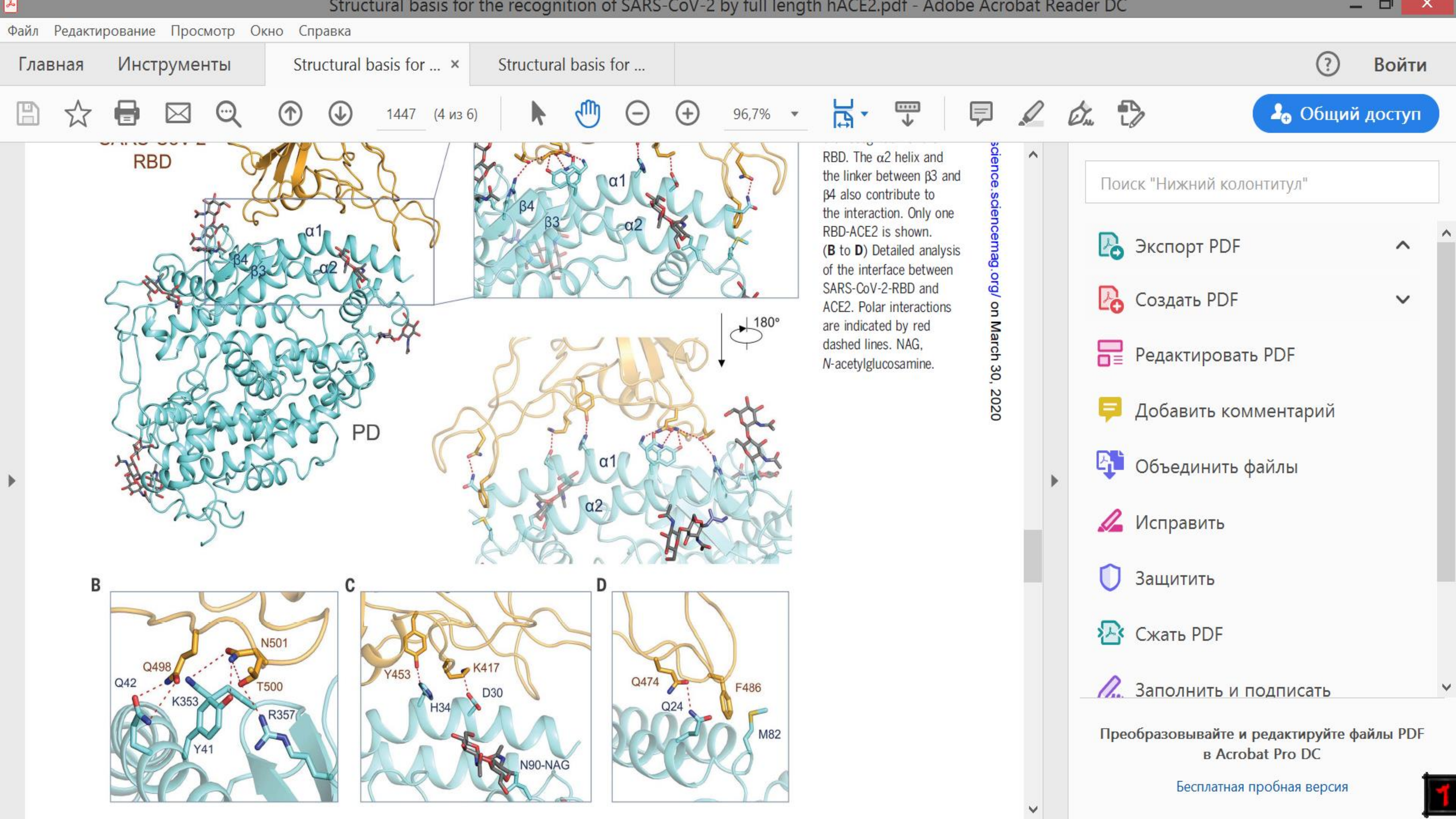


Use 3-letter amino acid codes Ctrl+Alt+3

```
tagatactctttggatttcataccatgtggaggctttcttacttccacgtgaccttgactgagtttgaatagcgcccaacccaagttcaaaggctgataagagagaaaatctcatgaggaggttttagt
atctatgagaaacctaaagtatggtacacctccgaaagaatgaaggtgcactggaactgactcaaacttatcgcgggttgggttcaagtttccgactattctctcttttagagtactcctccaaaatca    260

ctagggaaagtcattcagtggatgtgatcttggctcacaggggacgatgtcaagctcttcctggctccttctcagccttgttgctgtaactgctgctcagtccaccattgaggaacaggccaagacattt
gatccctttcagtaagtcacctacactagaaccgagtgtcccctgctacagttcgagaaggaccgaggaagagtcggaacaacgacattgacgacgagtcaggtggtaactccttgtccggttctgtaaa    390
                                              M  S  S  S  S  W  L  L  L  S  L  V  A  V  T  A  A  Q  S  T  I  E  E  Q  A  K  T  F

ttggacaagtttaaccacgaagccgaagacctgttctatcaaagttcacttgcttcttggaattataacaccaatattactgaagagaatgtccaaaacatgaataatgctggggacaaatggtctgcct
aacctgttcaaattggtgcttcggcttctggacaagatagtttcaagtgaacgaagaaccttaatattgtggttataatgacttctcttacaggttttgtacttattacgacccctgtttaccagacgga    520
 L  D  K  F  N  H  E  A  E  D  L  F  Y  Q  S  S  L  A  S  W  N  Y  N  T  N  I  T  E  E  N  V  Q  N  M  N  N  A  G  D  K  W  S  A

ttttaaaggaacagtccacacttgcccaaatgtatccactacaagaaattcagaatctcacagtcaagcttcagctgcaggctcttcagcaaaatgggtcttcagtgctctcagaagacaagagcaaacg
aaaatttccttgtcaggtgtgaacgggtttacataggtgatgttctttaagtcttagagtgtcagttcgaagtcgacgtccgagaagtcgttttacccagaagtcacgagagtcttctgttctcgtttgc    650
 F  L  K  E  Q  S  T  L  A  Q  M  Y  P  L  Q  E  I  Q  N  L  T  V  K  L  Q  L  Q  A  L  Q  Q  N  G  S  S  V  L  S  E  D  K  S  K  R
                                                                    *  A  E  A  A  P  E  E  A  F  H  T  K  L  A  R  L  L  C  S  C  V

gttgaacacaattctaaatacaatgagcaccatctacagtactggaaaagtttgtaacccagataatccacaagaatgcttattacttgaaccaggtttgaatgaaataatggcaaacagtttagactac
caacttgtgttaagatttatgttactcgtggtagatgtcatgaccttttcaaacattgggtctattaggtgttcttacgaataatgaacttggtccaaacttactttattaccgtttgtcaaatctgatg    780
```

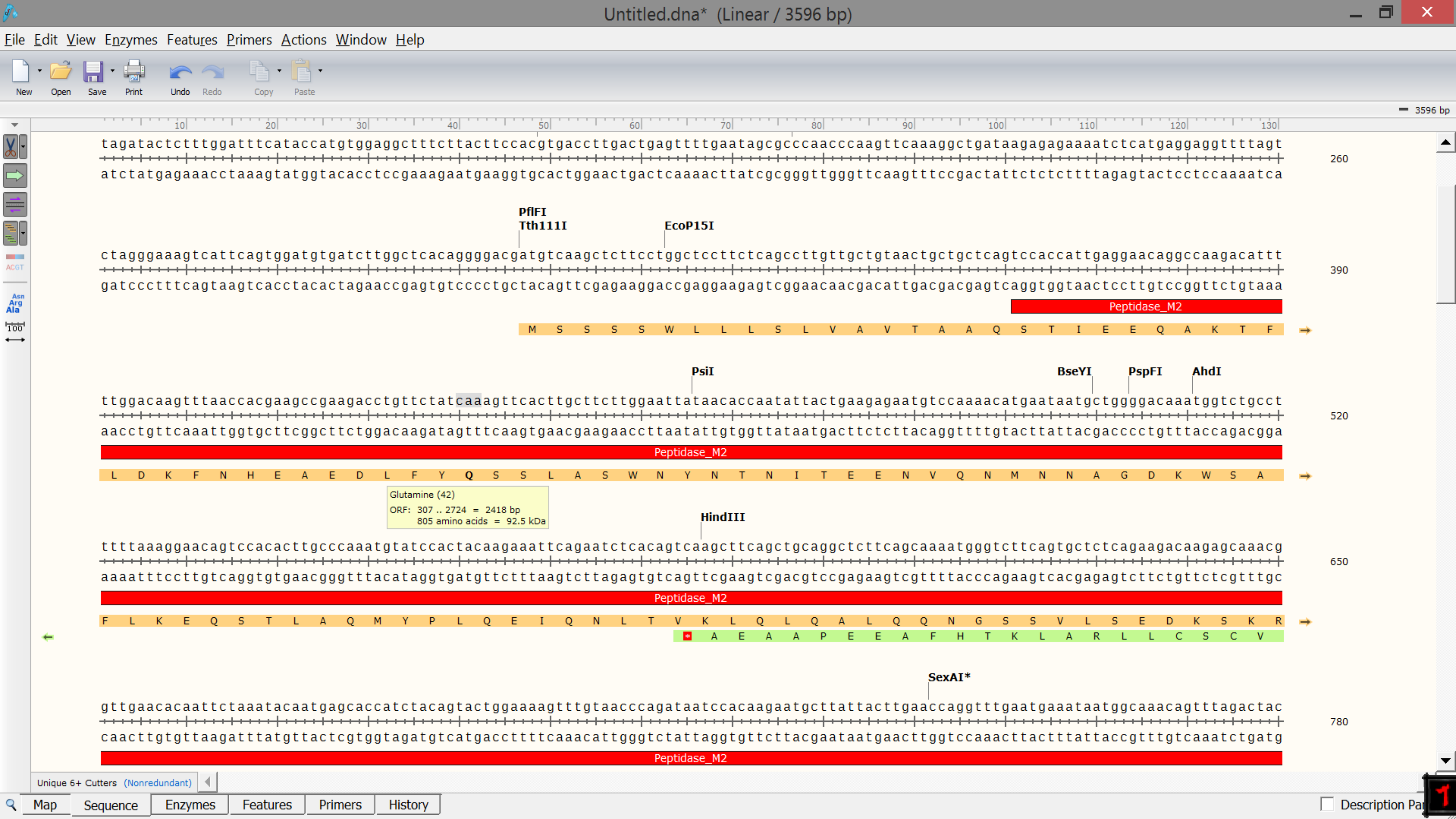


```
         10        20        30        40        50        60        70        80        90        100       110       120       130

tagatactctttggatttcataccatgtggaggctttcttacttccacgtgaccttgactgagttttgaatagcgcccaacccaagttcaaaggctgataagagagaaaatctcatgaggaggttttagt   260
atctatgagaaacctaaagtatggtacacctccgaaagaatgaaggtgcactggaactgactcaaaacttatcgcgggttgggttcaagtttccgactattctctcttttagagtactcctccaaaatca

                                             PflFI
                                             Tth111I          EcoP15I
ctagggaaagtcattcagtggatgtgatcttggctcacaggggacgatgtcaagctcttcctggctccttctcagccttgttgctgtaactgctgctcagtccaccattgaggaacaggccaagacattt   390
gatccctttcagtaagtcacctacactagaaccgagtgtcccctgctacagttcgagaaggaccgaggaagagtcggaacaacgacattgacgacgagtcaggtggtaactccttgtccggttctgtaaa
                                                                                                              Peptidase_M2
                                             Met Ser Ser Ser Ser Trp Leu Leu Leu Ser Leu Val Ala Val Thr Ala Ala Gln Ser Thr Ile Glu Glu Gln Ala Lys Thr Phe →

                                                           PsiI                                       BseYI    PspFI   AhdI
ttggacaagtttaaccacgaagccgaagacctgttctatcaaagttcacttgcttcttggaattataacaccaatattactgaagagaatgtccaaaacatgaataatgctggggacaaatggtctgcct   520
aacctgttcaaattggtgcttcggcttctggacaagatagtttcaagtgaacgaagaaccttaatattgtggttataatgacttctcttacaggttttgtacttattacgacccctgtttaccagacgga
                                                        Peptidase_M2
Leu Asp Lys Phe Asn His Glu Ala Glu Asp Leu Phe Tyr Gln Ser Ser Leu Ala Ser Trp Asn Tyr Asn Thr Asn Ile Thr Glu Glu Asn Val Gln Asn Met Asn Asn Ala Gly Asp Lys Trp Ser Ala →

                                                            HindIII
ttttaaaggaacagtccacacttgcccaaatgtatccactacaagaaattcagaatctcacagtcaagcttcagctgcaggctcttcagcaaaatgggtcttcagtgctctcagaagacaagagcaaacg   650
aaaatttccttgtcaggtgtgaacgggtttacataggtgatgttctttaagtcttagagtgtcagttcgaagtcgacgtccgagaagtcgttttacccagaagtcacgagagtcttctgttctcgtttgc
                                                        Peptidase_M2
Phe Leu Lys Glu Gln Ser Thr Leu Ala Gln Met Tyr Pro Leu Gln Glu Ile Gln Asn Leu Thr Val Lys Leu Gln Leu Gln Ala Leu Gln Gln Asn Gly Ser Ser Val Leu Ser Glu Asp Lys Ser Lys Arg →
←                                                          ■ Ala Glu Ala Ala Pro Glu Glu Ala Phe His Thr Lys Leu Ala Arg Leu Leu Cys Ser Cys Val

                                                                                    SexAI*
gttgaacacaattctaaatacaatgagcaccatctacagtactggaaaagtttgtaacccagataatccacaagaatgcttattacttgaaccaggtttgaatgaaataatggcaaacagtttagactac   780
caacttgtgttaagatttatgttactcgtggtagatgtcatgaccttttcaaacattgggtctattaggtgttcttacgaataatgaacttggtccaaacttactttattaccgtttgtcaaatctgatg
                                                        Peptidase_M2
```



RBD. The α2 helix and the linker between β3 and β4 also contribute to the interaction. Only one RBD-ACE2 is shown. (**B** to **D**) Detailed analysis of the interface between SARS-CoV-2-RBD and ACE2. Polar interactions are indicated by red dashed lines. NAG, *N*-acetylglucosamine.

Преобразовывайте и редактируйте файлы PDF в Acrobat Pro DC



```
tagatactctttggatttcataccatgtggaggctttcttacttccacgtgaccttgactgagttttgaatagcgcccaacccaagttcaaaggctgataagagagaaaatctcatgaggaggttttagt
atctatgagaaacctaaagtatggtacacctccgaaagaatgaaggtgcactggaactgactcaaaacttatcgcgggttgggttcaagtttccgactattctctcttttagagtactcctccaaaatca   260

PflFI
Tth111I  EcoP15I
ctagggaaagtcattcagtggatgtgatcttggctcacaggggacgatgtcaagctcttcctggctccttctcagccttgttgctgtaactgctgctcagtccaccattgaggaacaggccaagacattt
gatccctttcagtaagtcacctacactagaaccgagtgtcccctgctacagttcgagaaggaccgaggaagagtcggaacaacgacattgacgacgagtcaggtggtaactccttgtccggttctgtaaa   390
Peptidase_M2
M S S S S W L L L S L V A V T A A Q S T I E E Q A K T F

PsiI  BseYI  PspFI  AhdI
ttggacaagtttaaccacgaagccgaagacctgttctatcaaagttcacttgcttcttggaattataacaccaatattactgaagagaatgtccaaaacatgaataatgctggggacaaatggtctgcct
aacctgttcaaattggtgcttcggcttctggacaagatagtttcaagtgaacgaagaaccttaatattgtggttataatgacttctcttacaggttttgtacttattacgacccctgtttaccagacgga   520
Peptidase_M2
L D K F N H E A E D L F Y Q S S L A S W N Y N T N I T E E N V Q N M N N A G D K W S A
```

Glutamine (42)
ORF: 307 .. 2724 = 2418 bp
805 amino acids = 92.5 kDa

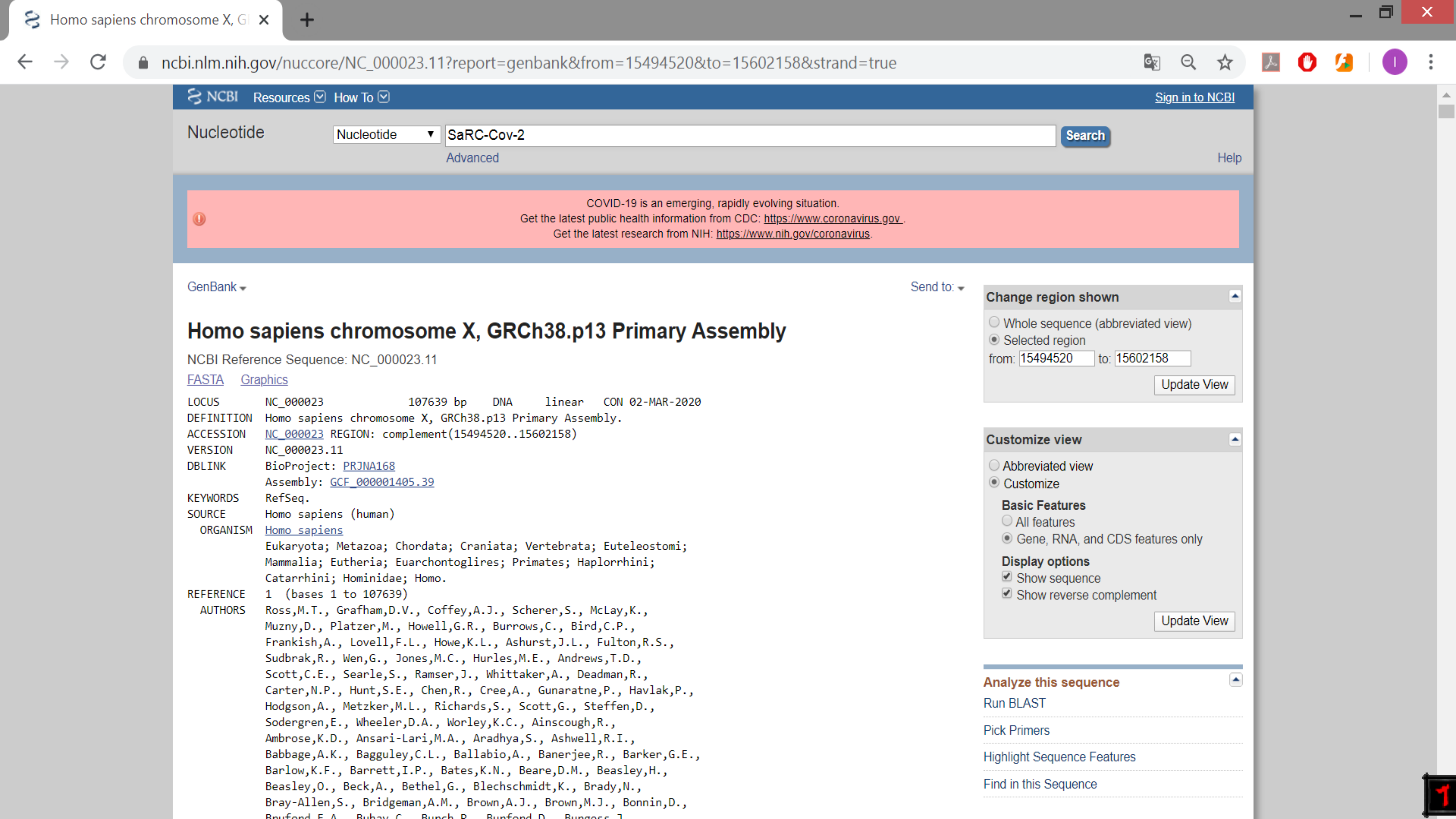
```
HindIII
ttttaaaggaacagtccacacttgcccaaatgtatccactacaagaaattcagaatctcacagtcaagcttcagctgcaggctcttcagcaaaatgggtcttcagtgctctcagaagacaagagcaaacg
aaaatttccttgtcaggtgtgaacgggtttacataggtgatgttctttaagtcttagagtgtcagttcgaagtcgacgtccgagaagtcgttttacccagaagtcacgagagtcttctgttctcgtttgc   650
Peptidase_M2
F L K E Q S T L A Q M Y P L Q E I Q N L T V K L Q L Q A L Q Q N G S S V L S E D K S K R
* A E A A P E E A F H T K L A R L L C S C V

SexAI*
gttgaacacaattctaaatacaatgagcaccatctacagtactggaaaagtttgtaacccagataatccacaagaatgcttattacttgaaccaggtttgaatgaaataatggcaaacagtttagactac
caacttgtgttaagatttatgttactcgtggtagatgtcatgaccttttcaaacattgggtctattaggtgttcttacgaataatgaacttggtccaaacttactttattaccgtttgtcaaatctgatg   780
Peptidase_M2
```

```
Y   W   R   G   D   Y   E   V   N   G   V   D   G   Y   D   Y   S   R   G   Q   L   I   E   D   V   E   H   T   F   E   E   I   K   P   L   Y   E   H   L   H   A   Y   V   R

ggcaaagttgatgaatgcctatccttcctatatcagtccaattggatgcctccctgctcatttgcttggtgatatgtggggtagattttggacaaatctgtactctttgacagttccctttggacagaaa    1170
ccgtttcaactacttacggataggaaggatatagtcaggttaacctacggagggacgagtaaacgaaccactatacaccccatctaaaacctgtttagacatgagaaactgtcaagggaaacctgtcttt
  A   K   L   M   N   A   Y   P   S   Y   I   S   P   I   G   C   L   P   A   H   L   L   G   D   M   W   G   R   F   W   T   N   L   Y   S   L   T   V   P   F   G   Q   K

ccaaacatagatgttactgatgcaatggtggaccaggcctgggatgcacagagaatattcaaggaggccgagaagttctttgtatctgttggtcttcctaatatgactcaaggattctgggaaaattcca    1300
ggtttgtatctacaatgactacgttaccacctggtccggaccctacgtgtctcttataagttcctccggctcttcaagaaacatagacaaccagaaggattatactgagttcctaagaccctttttaaggt
 P   N   I   D   V   T   D   A   M   V   D   Q   A   W   D   A   Q   R   I   F   K   E   A   E   K   F   F   V   S   V   G   L   P   N   M   T   Q   G   F   W   E   N   S

tgctaacggacccaggaaatgttcagaaagcagtctgccatcccacagcttgggacctggggaagggcgacttcaggatccttatgtgcacaaaggtgacaatggacgacttcctgacagctcatcatga    1430
acgattgcctgggtcctttacaagtctttcgtcagacggtagggtgtcgaaccctggaccccttcccgctgaagtcctaggaatacacgtgtttccactgttacctgctgaaggactgtcgagtagtact
M   L   T   D   P   G   N   V   Q   K   A   V   C   H   P   T   A   W   D   L   G   K   G   D   F   R   I   L   M   C   T   K   V   T   M   D   D   F   L   T   A   H   H   E

gatggggcatatccagtatgatatggcatatgctgcacaacctttttctgctaagaaatggagctaatgaaggattccatgaagctgttggggaaatcatgtcactttctgcagccacacctaagcattta    1560
ctaccccgtataggtcatactataccgtatacgacgtgttggaaaagacgattctttacctcgattacttcctaaggtacttcgacaacccctttagtacagtgaaagacgtcggtgtggattcgtaaat
  M   G   H   I   Q   Y   D   M   A   Y   A   A   Q   P   F   L   L   R   N   G   A   N   E   G   F   H   E   A   V   G   E   I   M   S   L   S   A   A   T   P   K   H   L
```

Lysine (353)
ORF: 307 .. 2724 = 2418 bp
805 amino acids = 92.5 kDa

Selected: active (1363 .. 1365 = 3 bp)   3596 bp

```
        10        20        30        40        50        60        70        80        90        100       110       120       130
 Y   W   R   G   D   Y   E   V   N   G   V   D   G   Y   D   Y   S   R   G   Q   L   I   E   D   V   E   H   T   F   E   E   I   K   P   L   Y   E   H   L   H   A   Y   V   R  →

                                        MfeI
ggcaaagttgatgaatgcctatccttcctatatcagtccaattggatgcctccctgctcatttgcttggtgatatgtggggtagattttggacaaatctgtactctttgacagttccctttggacagaaa   1170
ccgtttcaactacttacggataggaaggatatagtcaggttaacctacggagggacgagtaaacgaaccactatacaccccatctaaaacctgtttagacatgagaaactgtcaagggaaacctgtcttt
                                                          Peptidase_M2
  A   K   L   M   N   A   Y   P   S   Y   I   S   P   I   G   C   L   P   A   H   L   L   G   D   M   W   G   R   F   W   T   N   L   Y   S   L   T   V   P   F   G   Q   K  →

                                      StuI*                                                   NmeAIII
ccaaacatagatgttactgatgcaatggtggaccaggcctgggatgcacagagaatattcaaggaggccgagaagttctttgtatctgttggtcttcctaatatgactcaaggattctgggaaaattcca   1300
ggtttgtatctacaatgactacgttaccacctggtccggaccctacgtgtctcttataagttcctccggctcttcaagaaacatagacaaccagaaggattatactgagttcctaagacccttttaaggt
                                                          Peptidase_M2
  P   N   I   D   V   T   D   A   M   V   D   Q   A   W   D   A   Q   R   I   F   K   E   A   E   K   F   F   V   S   V   G   L   P   N   M   T   Q   G   F   W   E   N   S  →

                                                                                     BamHI     ApaLI
tgctaacggacccaggaaatgttcagaaagcagtctgccatcccacagcttgggacctggggaaggcgacttcaggatccttatgtgcacaaaggtgacaatggacgacttcctgacagctcatcatga   1430
acgattgcctgggtcctttacaagtctttcgtcagacggtagggtgtcgaaccctggaccccttcccgctgaagtcctaggaatacacgtgtttccactgttacctgctgaaggactgtcgagtagtact
                                                          Peptidase_M2
                                                             active
M   L   T   D   P   G   N   V   Q   K   A   V   C   H   P   T   A   W   D   L   G   K   G   D   F   R   I   L   M   C   T   K   V   T   M   D   D   F   L   T   A   H   H   E  →

                                  BstAPI
gatggggcatatccagtatgatatggcatatgctgcacaacctttttctgctaagaaatggagctaatgaaggattccatgaagctgttggggaaatcatgtcactttctgcagccacacctaagcattta   1560
ctaccccgtataggtcatactataccgtatacgacgtgttggaaaagacgattctttacctcgattacttcctaaggtacttcgacaacccctttagtacagtgaaagacgtcggtgtggattcgtaaat
                                                          Peptidase_M2
  M   G   H   I   Q   Y   D   M   A   Y   A   A   Q   P   F   L   L   R   N   G   A   N   E   G   F   H   E   A   V   G   E   I   M   S   L   S   A   A   T   P   K   H   L  →
```



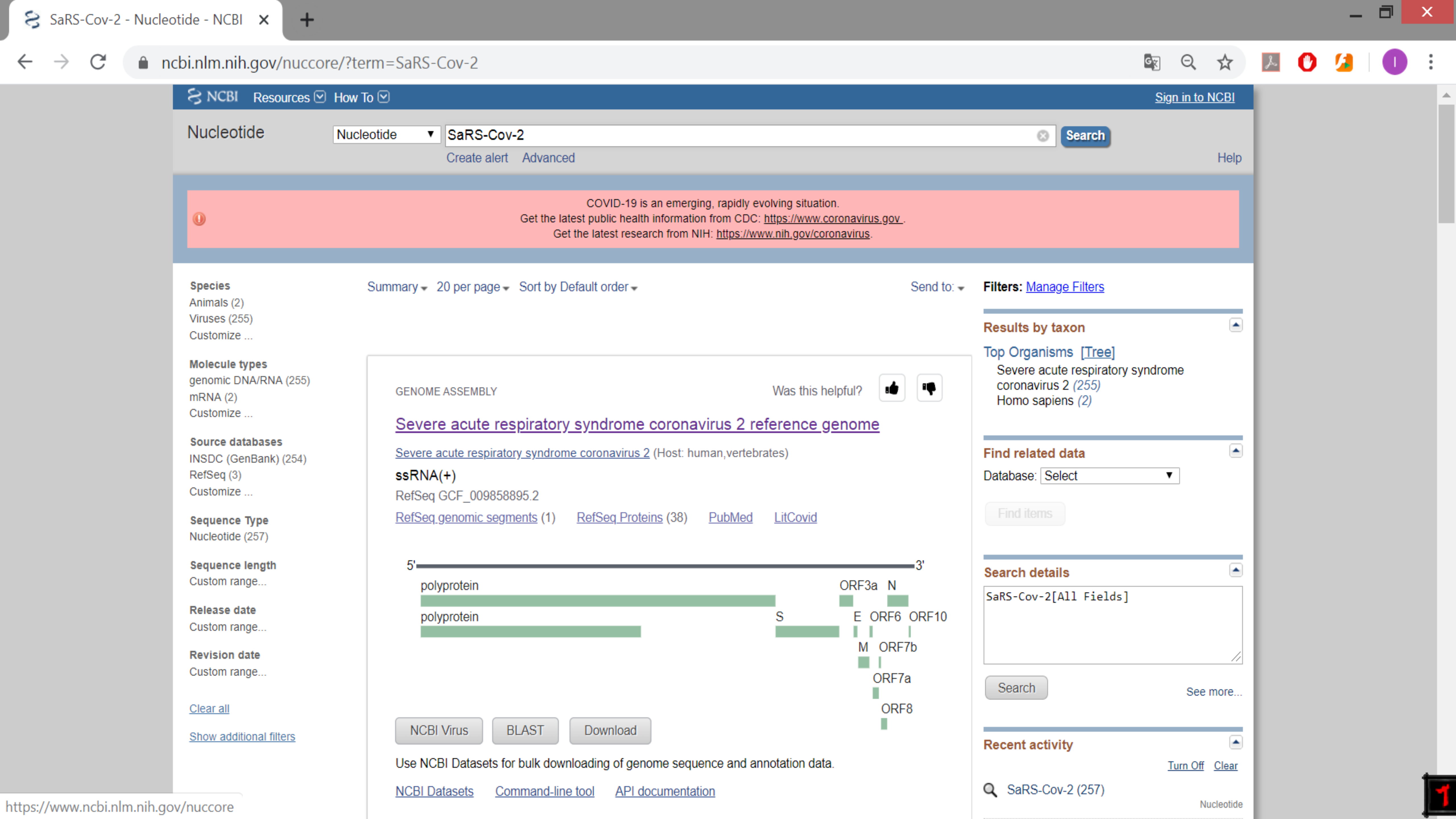
COVID-19 is an emerging, rapidly evolving situation.
Get the latest public health information from CDC: https://www.coronavirus.gov .
Get the latest research from NIH: https://www.nih.gov/coronavirus.

Homo sapiens chromosome X, GRCh38.p13 Primary Assembly

NCBI Reference Sequence: NC_000023.11

FASTA Graphics

```
LOCUS       NC_000023             107639 bp    DNA     linear   CON 02-MAR-2020
DEFINITION  Homo sapiens chromosome X, GRCh38.p13 Primary Assembly.
ACCESSION   NC_000023 REGION: complement(15494520..15602158)
VERSION     NC_000023.11
DBLINK      BioProject: PRJNA168
            Assembly: GCF_000001405.39
KEYWORDS    RefSeq.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;
            Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 107639)
  AUTHORS   Ross,M.T., Grafham,D.V., Coffey,A.J., Scherer,S., McLay,K.,
            Muzny,D., Platzer,M., Howell,G.R., Burrows,C., Bird,C.P.,
            Frankish,A., Lovell,F.L., Howe,K.L., Ashurst,J.L., Fulton,R.S.,
            Sudbrak,R., Wen,G., Jones,M.C., Hurles,M.E., Andrews,T.D.,
            Scott,C.E., Searle,S., Ramser,J., Whittaker,A., Deadman,R.,
            Carter,N.P., Hunt,S.E., Chen,R., Cree,A., Gunaratne,P., Havlak,P.,
            Hodgson,A., Metzker,M.L., Richards,S., Scott,G., Steffen,D.,
            Sodergren,E., Wheeler,D.A., Worley,K.C., Ainscough,R.,
            Ambrose,K.D., Ansari-Lari,M.A., Aradhya,S., Ashwell,R.I.,
            Babbage,A.K., Bagguley,C.L., Ballabio,A., Banerjee,R., Barker,G.E.,
            Barlow,K.F., Barrett,I.P., Bates,K.N., Beare,D.M., Beasley,H.,
            Beasley,O., Beck,A., Bethel,G., Blechschmidt,K., Brady,N.,
            Bray-Allen,S., Bridgeman,A.M., Brown,A.J., Brown,M.J., Bonnin,D.,
```



SaRS-Cov-2

COVID-19 is an emerging, rapidly evolving situation.
Get the latest public health information from CDC: https://www.coronavirus.gov .
Get the latest research from NIH: https://www.nih.gov/coronavirus.

Species
Animals (2)
Viruses (255)
Customize ...

Molecule types
genomic DNA/RNA (255)
mRNA (2)
Customize ...

Source databases
INSDC (GenBank) (254)
RefSeq (3)
Customize ...

Sequence Type
Nucleotide (257)

Sequence length
Custom range...

Release date
Custom range...

Revision date
Custom range...

Clear all

Show additional filters

GENOME ASSEMBLY

Was this helpful?

Severe acute respiratory syndrome coronavirus 2 reference genome

Severe acute respiratory syndrome coronavirus 2 (Host: human,vertebrates)

ssRNA(+)

RefSeq GCF_009858895.2

RefSeq genomic segments (1) RefSeq Proteins (38) PubMed LitCovid

Use NCBI Datasets for bulk downloading of genome sequence and annotation data.

NCBI Datasets Command-line tool API documentation

Filters: Manage Filters

Results by taxon

Top Organisms [Tree]
Severe acute respiratory syndrome coronavirus 2 (255)
Homo sapiens (2)

Find related data

Database: Select

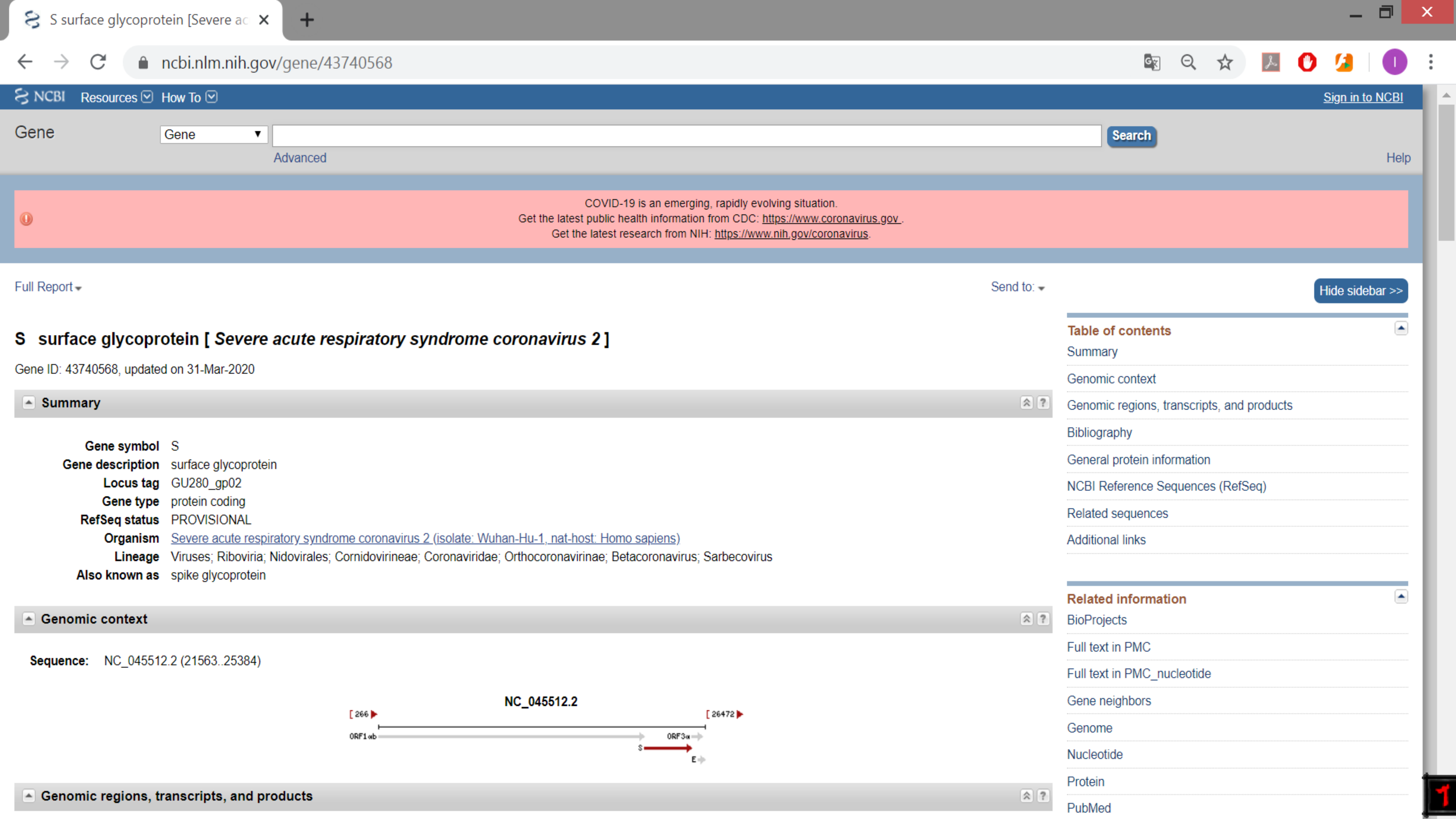
Search details

SaRS-Cov-2[All Fields]

Recent activity

SaRS-Cov-2 (257)
Nucleotide

https://www.ncbi.nlm.nih.gov/nuccore



COVID-19 is an emerging, rapidly evolving situation.
Get the latest public health information from CDC: https://www.coronavirus.gov .
Get the latest research from NIH: https://www.nih.gov/coronavirus.

S surface glycoprotein [*Severe acute respiratory syndrome coronavirus 2*]

Gene ID: 43740568, updated on 31-Mar-2020

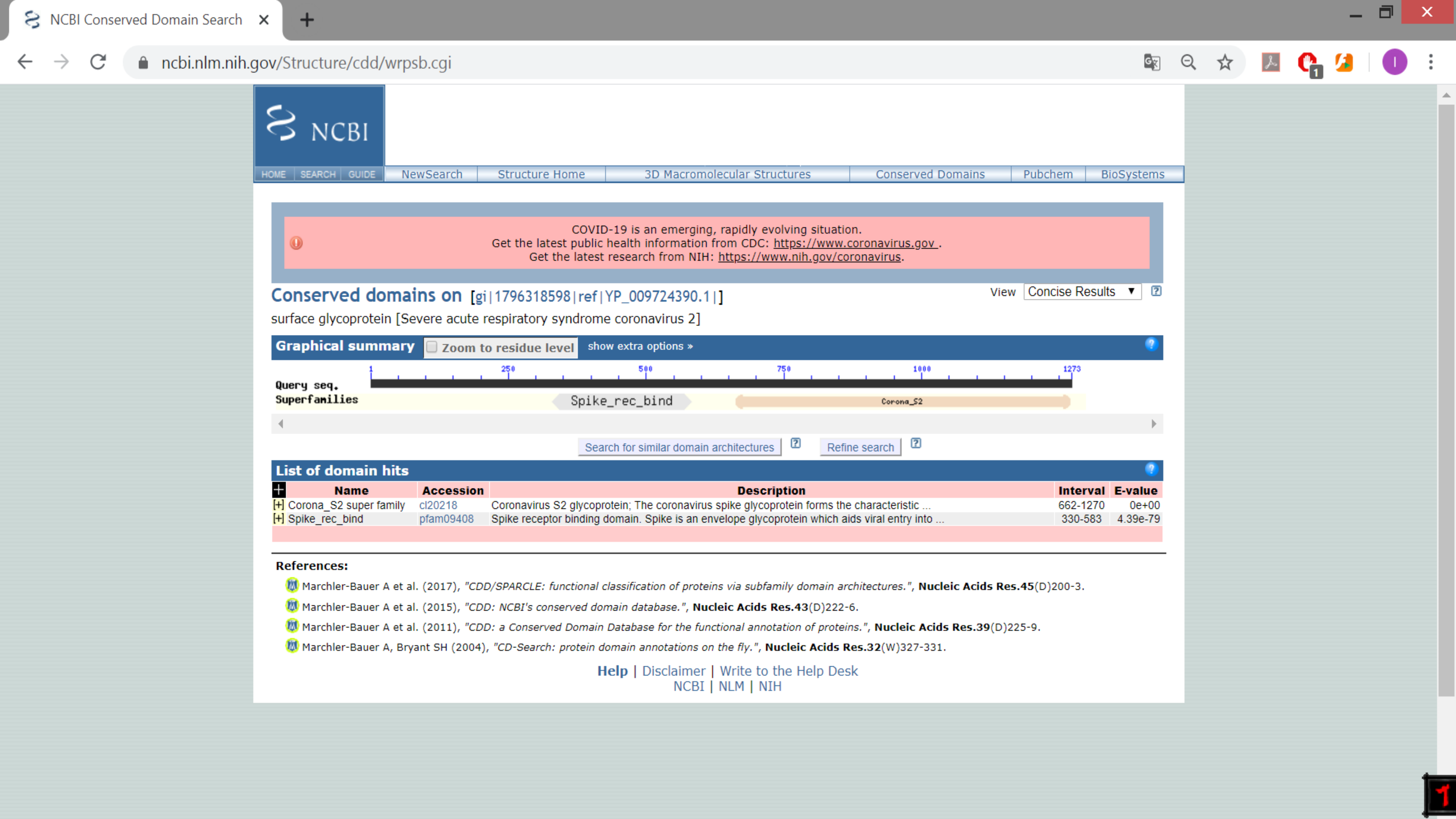
Summary

Gene symbol	S
Gene description	surface glycoprotein
Locus tag	GU280_gp02
Gene type	protein coding
RefSeq status	PROVISIONAL
Organism	Severe acute respiratory syndrome coronavirus 2 (isolate: Wuhan-Hu-1, nat-host: Homo sapiens)
Lineage	Viruses; Riboviria; Nidovirales; Cornidovirineae; Coronaviridae; Orthocoronavirinae; Betacoronavirus; Sarbecovirus
Also known as	spike glycoprotein

Genomic context

Sequence: NC_045512.2 (21563..25384)

Genomic regions, transcripts, and products



COVID-19 is an emerging, rapidly evolving situation.
Get the latest public health information from CDC: https://www.coronavirus.gov .
Get the latest research from NIH: https://www.nih.gov/coronavirus.

Conserved domains on [gi|1796318598|ref|YP_009724390.1|]

View: Concise Results

surface glycoprotein [Severe acute respiratory syndrome coronavirus 2]

Graphical summary ☐ Zoom to residue level show extra options »

List of domain hits

	Name	Accession	Description	Interval	E-value
[+]	Corona_S2 super family	cl20218	Coronavirus S2 glycoprotein; The coronavirus spike glycoprotein forms the characteristic ...	662-1270	0e+00
[+]	Spike_rec_bind	pfam09408	Spike receptor binding domain. Spike is an envelope glycoprotein which aids viral entry into ...	330-583	4.39e-79

References:

Marchler-Bauer A et al. (2017), *"CDD/SPARCLE: functional classification of proteins via subfamily domain architectures."*, **Nucleic Acids Res.45**(D)200-3.

Marchler-Bauer A et al. (2015), *"CDD: NCBI's conserved domain database."*, **Nucleic Acids Res.43**(D)222-6.

Marchler-Bauer A et al. (2011), *"CDD: a Conserved Domain Database for the functional annotation of proteins."*, **Nucleic Acids Res.39**(D)225-9.

Marchler-Bauer A, Bryant SH (2004), *"CD-Search: protein domain annotations on the fly."*, **Nucleic Acids Res.32**(W)327-331.

Help | Disclaimer | Write to the Help Desk
NCBI | NLM | NIH

1) Найдите последовательность и скажите NCBI Reference Sequence id NUMBER: для Mycobacterium virus D29 - gp10 гена.

2) Найдите домен данного белка и укажите его первую и последнюю аминокислоты.