



[www.esaunggul.ac.id](http://www.esaunggul.ac.id)

# PENGANTAR BIOINFORMATIKA

## IBT 431

*By Seprianto S.Pi, M.Si*



Meeting 13

# Analisis Sekuen DNA dan Protein

# Sasaran Perkuliahan

- Mahasiswa mampu menganalisis Hasil Sekuen DNA dan Protein.
- Mahasiswa dapat menentukan anotasi gen pada sekuen DNA menggunakan SnapGene
- Mahasiswa dapat menentukan struktur 3D protein dengan menggunakan **SWISS- MODEL** dan **PyMOL**

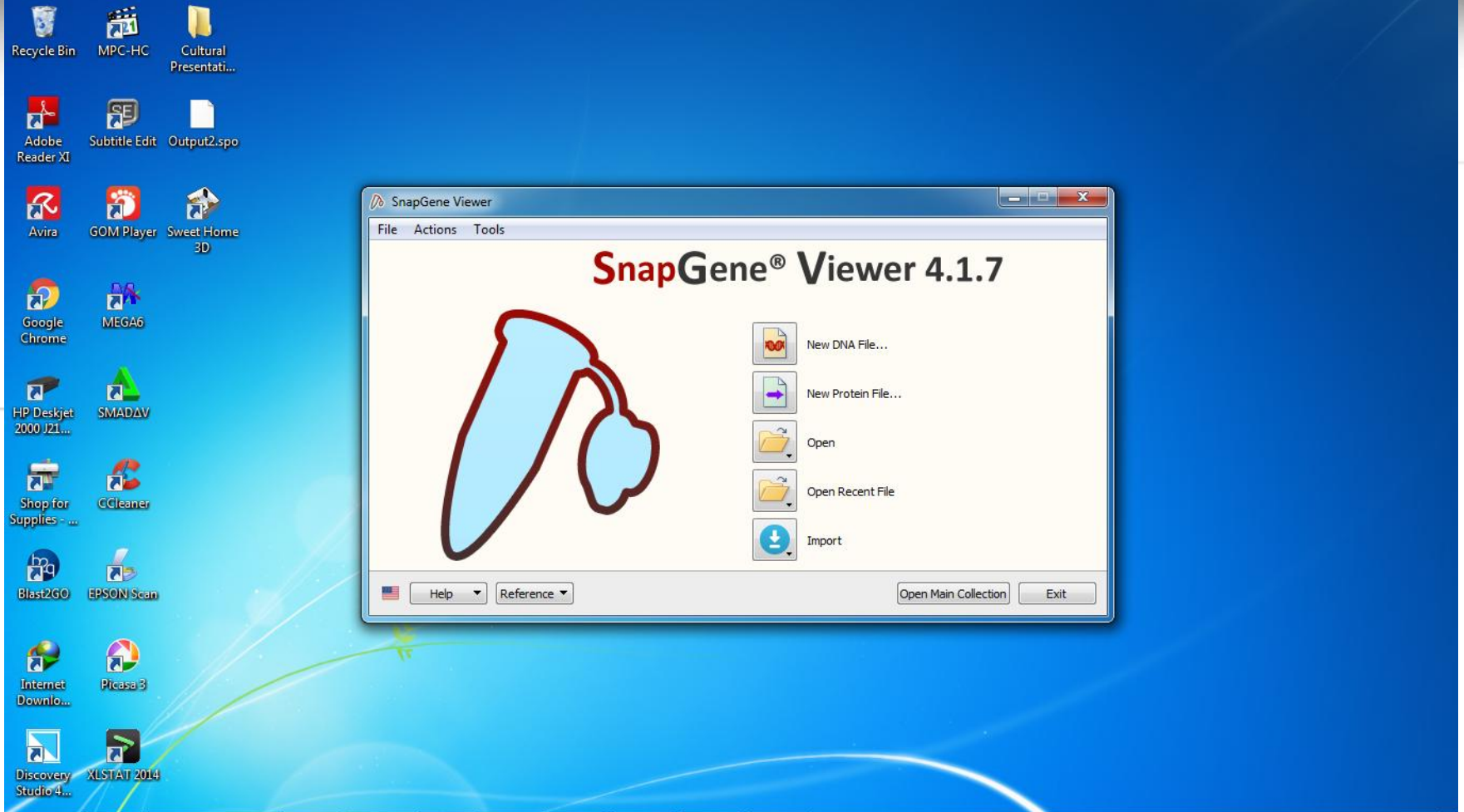
**Hands on**

**Anotasi sekuen nukleotida dan ekstrak protein dengan menggunakan SnapGene**

The screenshot displays the SnapGene Viewer product page. At the top, the SnapGene logo is on the left, and navigation links for Home, Products, Purchase, Support, Resources, Company, and Contact are in the center. A 'Manage My Account' link is on the right. Below the navigation, the 'Products' section is active, with sub-links for Feature Comparison, SnapGene, SnapGene Viewer (FREE), SnapGene Server, Screenshots, and File Compatibility. The main content area features the SnapGene Viewer logo and a question: 'Would you like to move beyond hand-drawn plasmid maps?'. To the left is a hand-drawn plasmid map of a 'Cen plasmid' (pRS315-based) with various features labeled: LEU2, pHS12-RFP (7504 bp), ~7400 bp, Pre-Cox4, ADHI, COX4, ST-10 RFP, and NotI. Restriction sites for XhoI, EcoRI, and XbaI are also marked. To the right, a 'Download' section offers four download options for SnapGene Viewer 4.1.8: Windows (32 MB), macOS (36 MB), Ubuntu Linux (36 MB), and Fedora Linux (36 MB). A link to 'Compare the features' is provided below the download options. The footer of the page includes the URL 'www.snapgene.com/contact'.

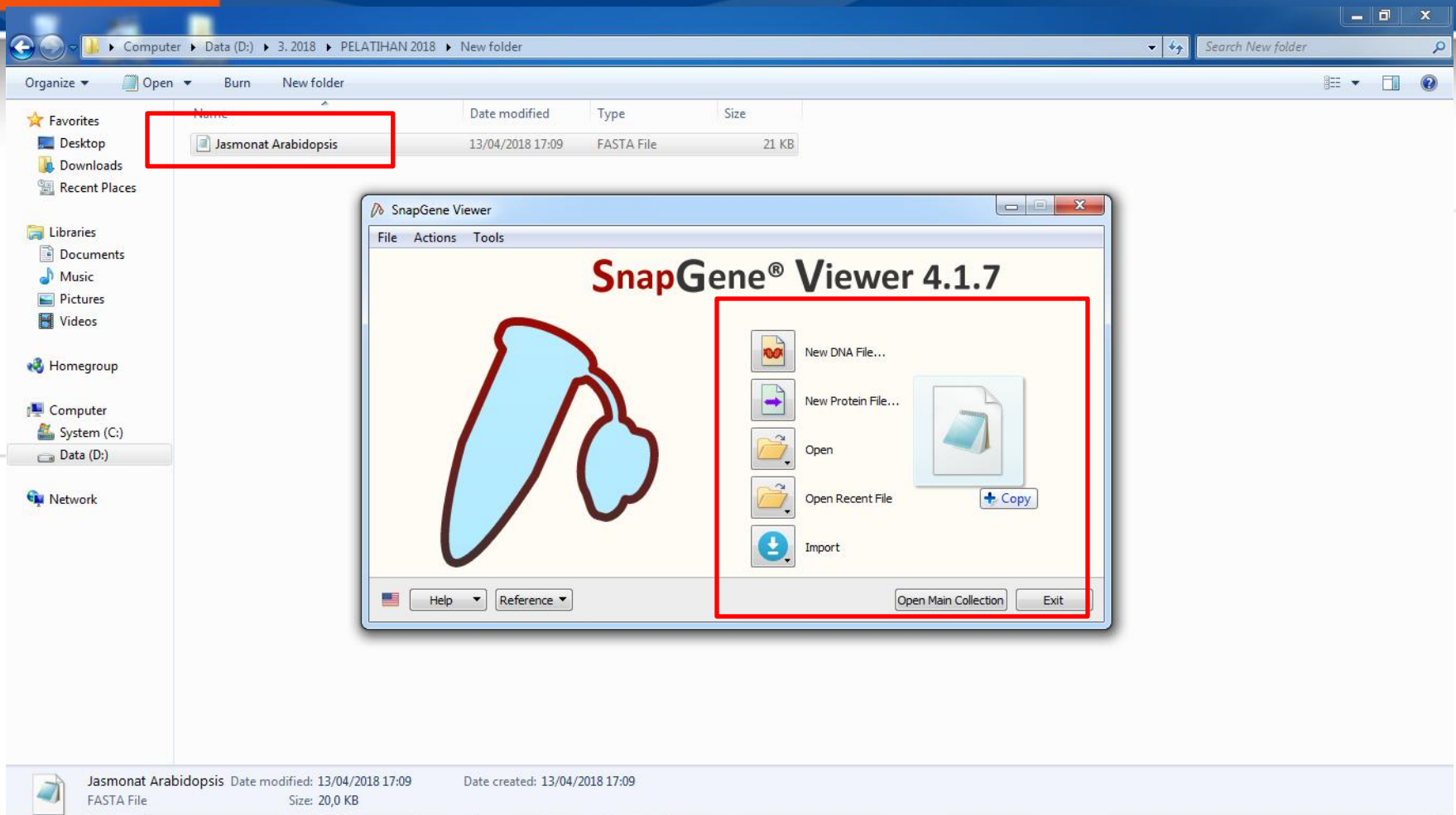
Download program **SnapGene Viewer** pada alamat website :  
[http://www.snapgene.com/products/snapgene\\_viewer/](http://www.snapgene.com/products/snapgene_viewer/)





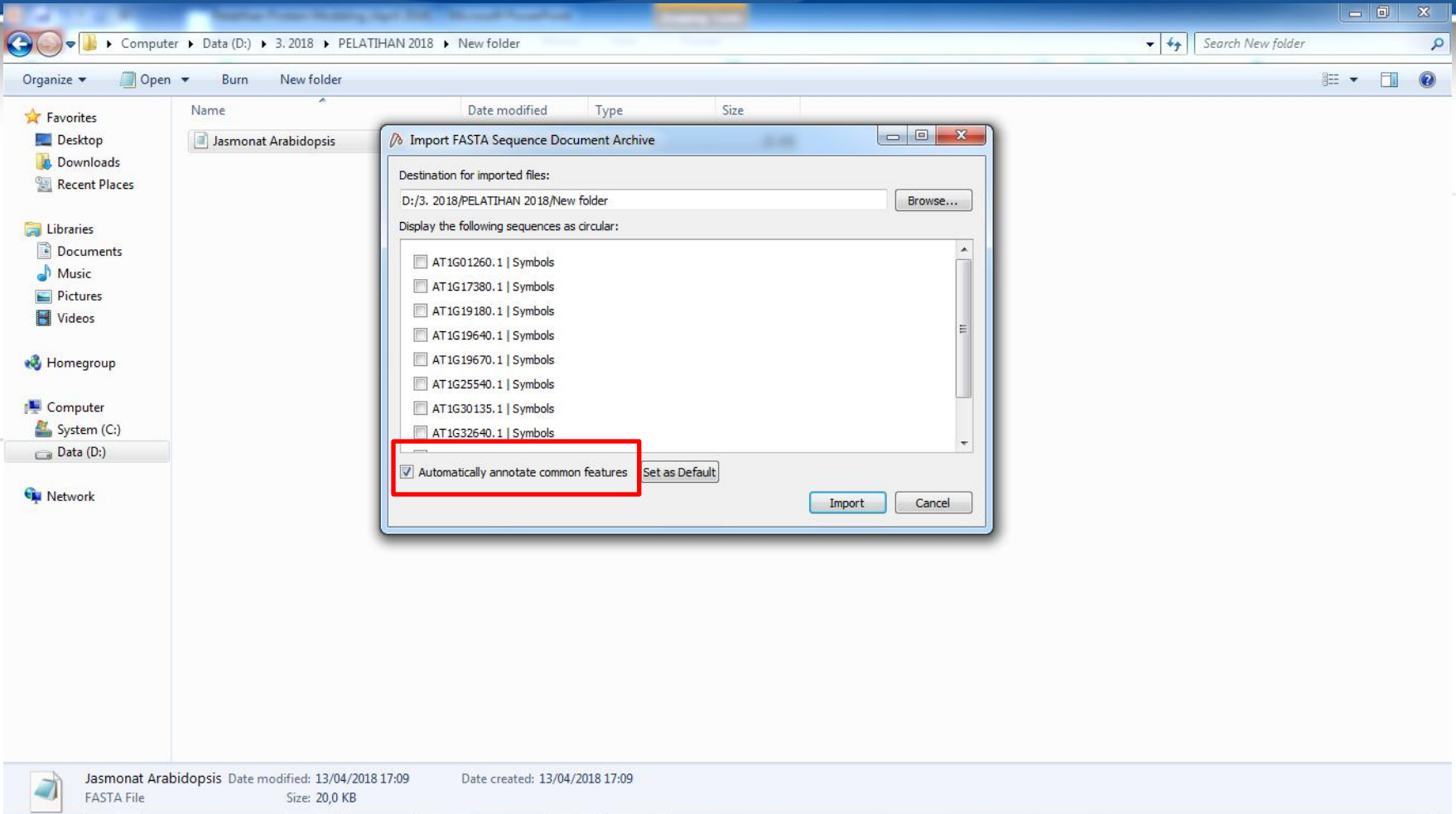
## Step 1.

Buka program **SnapGene** dengan halaman interface seperti diatas



## Step 2.

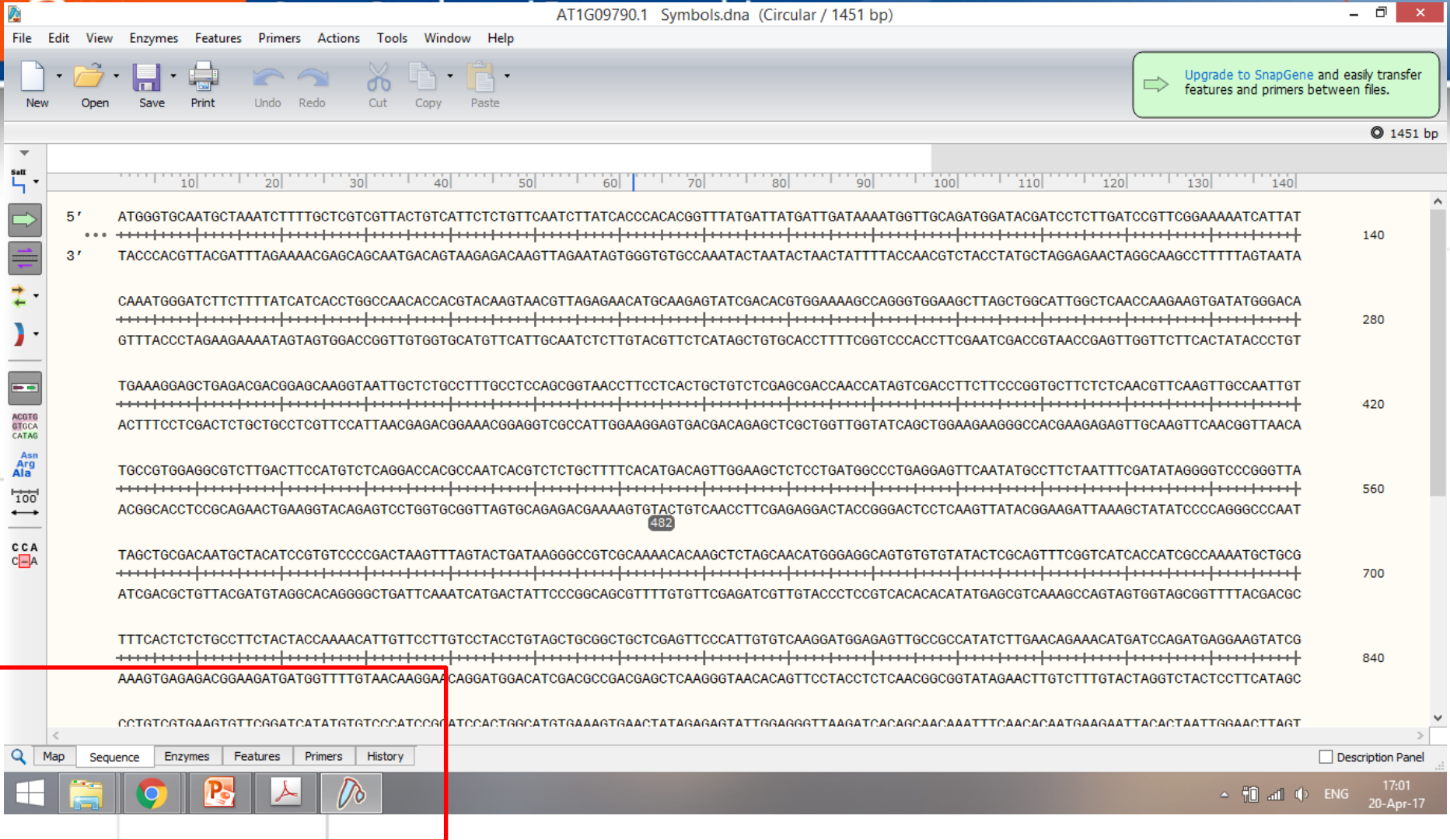
Buka file notepad yang berisi sekuen nukleotida dengan cara **drag** atau **open file** pada **SnapGene**



### Step 3.

Untuk merubah file **.fasta** menjadi file SnapGene, unggah file **.fasta** dan klik **Import**





## Step 4.

Fitur **Sequence** dapat diakses di bagian bawah dari program SnapGene  
Sekuen query ditampilkan dalam urutan 5-3' dan 3-5'

The screenshot shows the SnapGene software interface with a DNA sequence of 1451 bp. The sequence is displayed in a multi-line format with a scale at the top. The forward translation (5' to 3') is highlighted in yellow, and the reverse translation (3' to 5') is highlighted in green. A red box highlights the 'Translation' button in the left sidebar. The sequence is as follows:

```
5' ATGGGTGCAATGCTAAATCTTTTGCCTCGTCTACTGTCATTCTCTGTTCAATCTTATCACCCACACGGTTTATGATTATGATTGATAAAATGGTTGCAGATGGATACGATCCTCTTGATCCGTTCCGAAAAATCATTAT
...
3' TACCCACGTTACGATTTAGAAAAACGAGCAGCAATGACAGTAAGAGACAAGTTAGAATAGTGGGTGTGCCAAATACTAATACTAACTATTTTACCAACGCTCTACCTATGCTAGGAGAAGTGGCAAGCCTTTTGTAGTAATA
M G A M L N L L L V V T V I L C S I L S P T R F M I M I D K M V A D G Y D P L D P F G K I I I
M P A I S F R K S T T V T M R Q E I K D G V R N I I I I S L I T A S P Y S G R S G N P F I M I
CAAATGGGATCTTCTTTTATCATCACCTGGCCAACACCACGTACAAGTAACGTTAGAGAACATGCAAGAGTATCGACACGTTGAAAAGCCAGGGTGAAGCTTAGCTGGCATTGGCTCAACCAAGAAAGTGATATGGGACA
GTTTACCCTAGAAGAAAATAGTAGTGGACCGGTTGTGGTGCATGTTCAATGCAATCTCTGTACGTTCTCATAGCTGTGCACCTTTTCGGTCCCACCTTCGAATCGACCGTAACCGAGTTGGTTCTTCACTATACCTGT
K W D L L L S S P G Q H H V Q V T L E N M Q E Y R H V E K P G W K L S W H W L N Q E V I W D
L H S R R K D D G P W C W T C T V N S F M C S Y R C T S F G P H F S L Q C Q S L W S T I H S M
TGAAAGGAGCTGAGACGACGGAGCAAGGTAATTGCTCTGCCTTTGCCTCCAGCGGTAACCTTCTCACTGCTGTCTCGAGCGACCAACCATAAGTCGACCTTCTTCCCGGTGCTTCTCTCAACGTTCAAGTTGCCAATTGT
ACTTTCCTCGACTCTGCTGCCTCGTCCATTAACGAGACGAAAACGGAGGTCGCCATTGGAAGGAGTGACGACAGAGCTCGCTGGTTGGTATCAGCTGGAAGAAGGCCACGAAGAGAGTTGCAAGTTCAACGGTTAACA
M K G A E T T E Q G N C S A F A S S G N L P H C C L E R P T I V D L L P G A S L N V Q V A N C
TGCCGTGGAGGCGTCTTGACTTCCATGTCTCAGGACCACGCCAATCACGCTCTGCTTTTACATGACAGTTGGAAGCTCTCTGATGGCCCTGAGGAGTTCAATATGCCTTCTAATTTTCGATATAGGGGTCCCGGGTTA
ACGGCACCTCCGAGAACTGAAGGTACAGAGTCTGGTGCAGTTAGTGCAGAGACGAAAAGTGTACTGTCAACCTTCGAGAGGACTACGGGACTCCTCAAGTTATACGGAAGATTAAGCTATATCCCGAGGCCCAAT
C R G G V L T S M S Q D H A N H V S A F H M T V G S S P D G P E E F N M P S N F D I G V P G Y
TAGCTGCGACAATGCTACATCCGTGTCCTCCGACTAAGTTTAGTACTGATAAGGGCCGTCGCAAAACACAAGCTCTAGCAACATGGGAGGCGAGTGTGTATACTCGCAGTTTCGGTCATCACCATCGCCAAAATGCTGCG
```

## Step 5.

Klik **Translation** untuk mengakses prediksi urutan residu asam amino dari sekuen nukleotida

Sekuen asam amino forward dalam warna kuning

Sekuen asam amino reverse dalam warna hijau

The screenshot displays the SnapGene software interface for a DNA sequence file named 'AT1G09790.1 Symbols.dna (Circular / 1451 bp)'. The 'Features' menu is open, and the 'Add Translated Feature...' option is highlighted with a red box. The menu also includes options like 'Edit Features...', 'Duplicate Feature...', 'Remove Features', 'Create Feature Segment...', 'Delete Feature Segment...', 'Merge Feature Segments...', 'Add Cleavage Site...', 'Remove Cleavage Sites...', 'Feature Color...', 'Choose Alternative Codons...', 'Splice to Remove Introns...', 'Show Features Inside the Selection', 'Hide Features Inside the Selection', 'Sort Feature List...', 'Import Features from Another File...', 'Export Feature Data...', 'Detect Common Features...', 'Add to Common Features...', 'Browse Common Features', and 'Features Tutorial Video'. The main window shows a DNA sequence with various features annotated, including a CDS (Coding Sequence) highlighted in green. The sequence is displayed in both 5' to 3' and 3' to 5' orientations. The bottom of the window shows a taskbar with various application icons and a system tray with the date and time (17:04, 20-Apr-17).

## Step 6.

Lakukan anotasi dengan klik dua kali pada prediksi anotasi CDS yang diinginkan

Klik **Feature**, dan **Add Translated Features**

AT1G01260.1 Symbols.dna\* (Linear / 2377 bp)

File Edit View Enzymes Features Primers Actions Tools Window Help

New Open Save Print Undo

Selected: ORF (605 .. 2377 = 1773 bp) [44]

100 110

GACCCACAAGGTACGT

TAAAAGTCTTTCCTT

ATTTTCAGAAAGGAA

100 110

CTAGACCCAAAGA

TTGTTGCGTCATT

AACAACGCAGTAA

le Val Ala Ser Leu

le Val Ala Ser Leu

CTCTCGGATCTCG

GAGAGCCTAGAGC

Leu Ser Asp Leu

Leu Ser Asp Leu

TGGATATTGCAGA

ACCTATAACGTCT

90

Gly Tyr Cys Arg

Val Glu Arg Pro Asn Ala Ser Asn Phe Ser Trp Asn Tyr Ala Ile Phe Trp Gln Ile Ser Arg Ser Lys Ala Gly Asp Leu Val Leu Cys Trp Gly Asp Gly Tyr Cys Arg

1773 bp / 1 segment

605 2377

Split Feature...  
Merge Segments  
Delete Segment

Segment	Location	Size (bp)	Color
1	CDS	605 .. 2377	1773

/product =

Prioritize display of this feature in maps

OK Cancel

## Step 7.

Pilih tipe feature: **CDS** seperti ditampilkan diatas. Klik **Ok**

The screenshot shows a DNA sequence analysis tool window titled "AT1G01260.1 Symbols.dna\* (Linear / 2377 bp)". The interface includes a menu bar (File, Edit, View, Enzymes, Features, Primers, Actions, Tools, Window, Help) and a toolbar with icons for New, Open, Save, Print, Undo, Redo, Cut, Copy, and Paste. A scale bar at the top indicates positions from 10 to 110. The main display area shows the DNA sequence with several annotations:

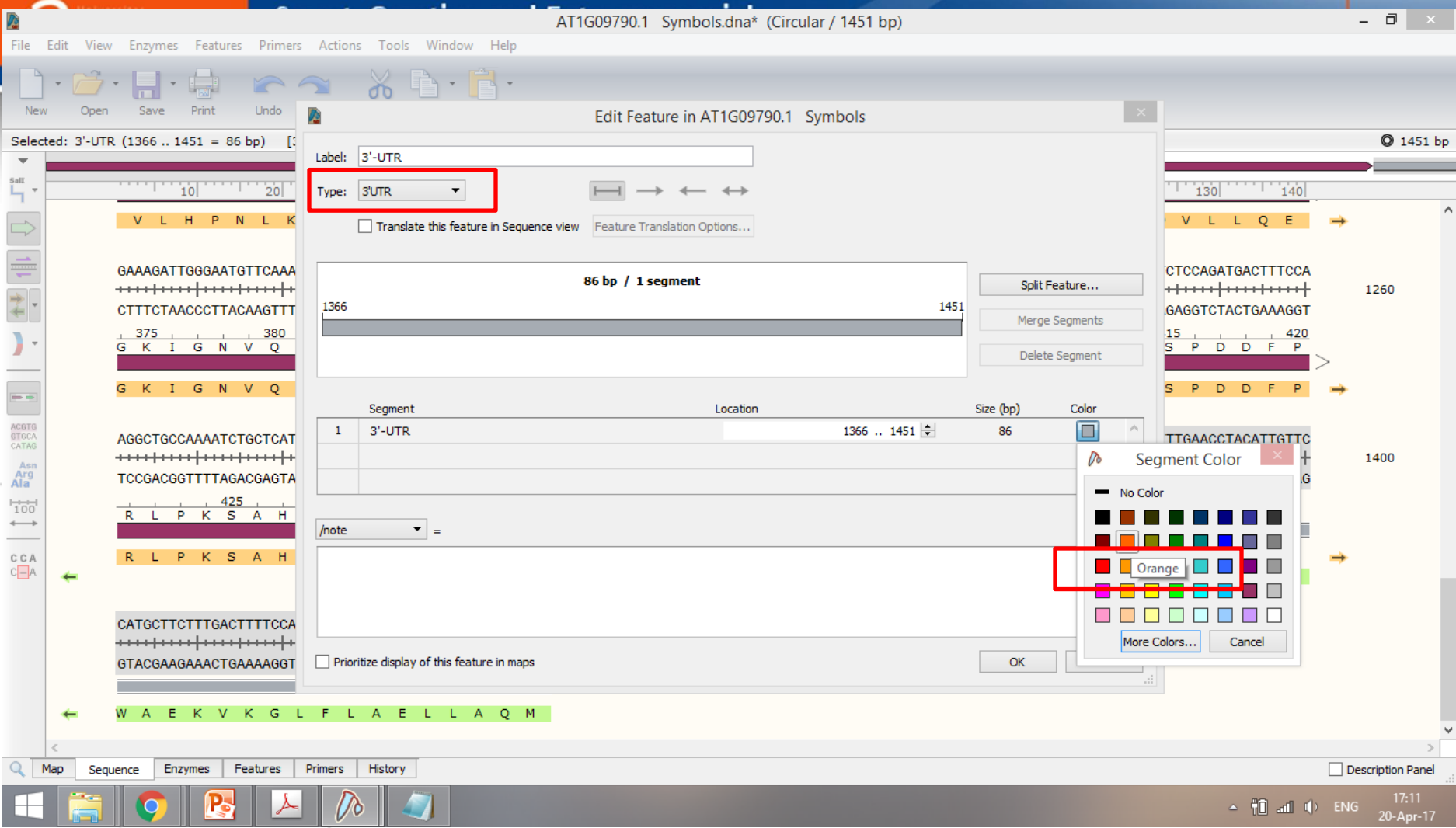
- 5' UTR:** An orange bar labeled "5'jasmonat" is positioned above the first line of the sequence: `GACCCACAAGGTACGTTTTGACAATATGAGGCCGCATATAAACGTAGACCCAGTGTGAAGATTTGGGATAGTAACTGAATTCTTAGCTAAAGACTTCTAGACCCAAAGA`.
- ORF 1:** A purple bar labeled "ORF" is positioned above the second line of the sequence: `TAAAAGTCTTTCTTTTGTGGGATCCAAAACAACAAAAGAAAAGAAAAGCAGAGATGAATATTGGTCGCCTAGTGTGGAACGAGGACGATAAAGCGATTGTTGCGTCATT`. Below this bar, the amino acid sequence is listed: Met Asn Ile Gly Arg Leu Val Trp Asn Glu Asp Asp Lys Ala Ile Val Ala Ser Leu.
- ORF 2:** A purple bar labeled "ORF" is positioned above the third line of the sequence: `ATTTTCAGAAAGGAAAAACAACCTAGGTTTGTGTTTTCTTTCTTTCTGCTCTACTTATAACCAGCGGATCACACCTTGCTCCTGCTATTTTCGTAACAACCGCAGTAA`. Below this bar, the amino acid sequence is listed: Met Asn Ile Gly Arg Leu Val Trp Asn Glu Asp Asp Lys Ala Ile Val Ala Ser Leu.
- ORF 3:** A purple bar labeled "ORF" is positioned above the fourth line of the sequence: `ACTGGGCAAACGAGCTCTCGATTACTTGTCTTTCCAACCTCTGTTTCCAATGCTAATCTCTTGATGACTCTAGGAAGCGACGAGAATCTGCAGAACAAAGCTCTCGGATCTCG`. Below this bar, the amino acid sequence is listed: Leu Gly Lys Arg Ala Leu Asp Tyr Leu Leu Ser Asn Ser Val Ser Asn Ala Asn Leu Leu Met Thr Leu Gly Ser Asp Glu Asn Leu Gln Asn Lys Leu Ser Asp Leu.
- ORF 4:** A purple bar labeled "ORF" is positioned above the fifth line of the sequence: `TGACCCGTTTGCTCGAGAGCTAATGAACGAAAGGTTGAGACAAAGGTTACGATTAGAGAACTACTGAGATCCTTCGCTGCTCTTAGACGTCTTGTTCGAGAGCCTAGAGC`. Below this bar, the amino acid sequence is listed: Leu Gly Lys Arg Ala Leu Asp Tyr Leu Leu Ser Asn Ser Val Ser Asn Ala Asn Leu Leu Met Thr Leu Gly Ser Asp Glu Asn Leu Gln Asn Lys Leu Ser Asp Leu.
- ORF 5:** A purple bar labeled "ORF" is positioned above the sixth line of the sequence: `TCGAGAGACCCAACGCTTCTAATTTCTCTTGGAACTACGCCATTTTCTGGCAGATTTCCAGGTCAAAAGGCCGGAGATTTGGTTCTCTGTTGGGGCGATGGATATTGCAGA`. Below this bar, the amino acid sequence is listed: Val Glu Arg Pro Asn Ala Ser Asn Phe Ser Trp Asn Tyr Ala Ile Phe Trp Gln Ile Ser Arg Ser Lys Ala Gly Asp Leu Val Leu Cys Trp Gly Asp Gly Tyr Cys Arg.

The software interface also includes a left sidebar with various tool icons and a bottom status bar with tabs for Map, Sequence, Enzymes, Features, Primers, and History. A "Description Panel" checkbox is visible in the bottom right corner.

## Step 8.

Tampilan hasil anotasi ORF/CDS dan 5'UTR di laman sekuen





## Step 9.

Lanjutkan anotasi untuk daerah **3'-UTR**

Warna feature dapat dipilih sesuai keinginan

File Edit View Enzymes Features Primers Actions Tools Window Help

New Open Save Print Undo Redo Cut Copy Paste

Insertion Point: 980

10 20 30 40 50 60 70

CGCAGCAGGGAAGCCGCCGCTGCGACTACATCGCTCCGTGGCCACGGCGGGCACGGCGGGCACGGCCCCGC 910

3'UTR

GTCCGCCGGCCGCCAGGTCATGGCCGGCGACCGCACCGCCACCGCCATCTCCCGCCTTTGCCGCATCTCC 980

CAGGCGGGCCGGCGGTCCAGTACCGGCCGCTGGCGGTGGCGGTAGAGGGCGGAAACGGCGTAGAGG

1 5 10 15

Met Ala Gly Asp Arg Thr Ala Thr Ala Ile Ser Arg Leu Cys Arg Ile Ser

3'UTR ORF MTGase

TTCCGCCTCGTGGCGGCGTTCCATTCTGTGCGCCGCCACCGCGCTCAGGACAGCGCGGCTGCTTACCGCGA 1050

AAGGCGGAGCACCGCCGCAAGGTAAGACAGCGGGCGGTGGCGCGAGTCTGTGCGCCGACGAATGGCGCT

20 25 30 35 40

Phe Arg Leu Val Ala Ala Phe His Ser Val Ala Ala Thr Ala Leu Arg Thr Ala Arg Leu Leu Thr Ala

ORF MTGase

ACCCCTCATGTGTCGTTTCGCTCGCATGCCCGTTTCACGGGAATCCACAACAAGGGAGTTACTGATTTTCAT 1120

TGGGGAGTACACAGCAAGCGAGCGTACGGGGCAAAGTGCCCTTAGGTGTTGTTCCCTCAATGACTAAAGTA

45 50 55 60

Asn Pro Ser Cys Val Val Arg Ser His Ala Arg Phe Thr Gly Ile His Asn Lys Gly Val Thr Asp Phe Met

ORF MTGase

GTACAAACGCCGGAGTTTACTCGCGTTTCGCCACTGTGGGTGCGCTGATATGCACCGCCGGAGTCATGCCG

**Step 10.**

Tampilan hasil anotasi **3'-UTR** di laman sekuen dengan warna berbeda

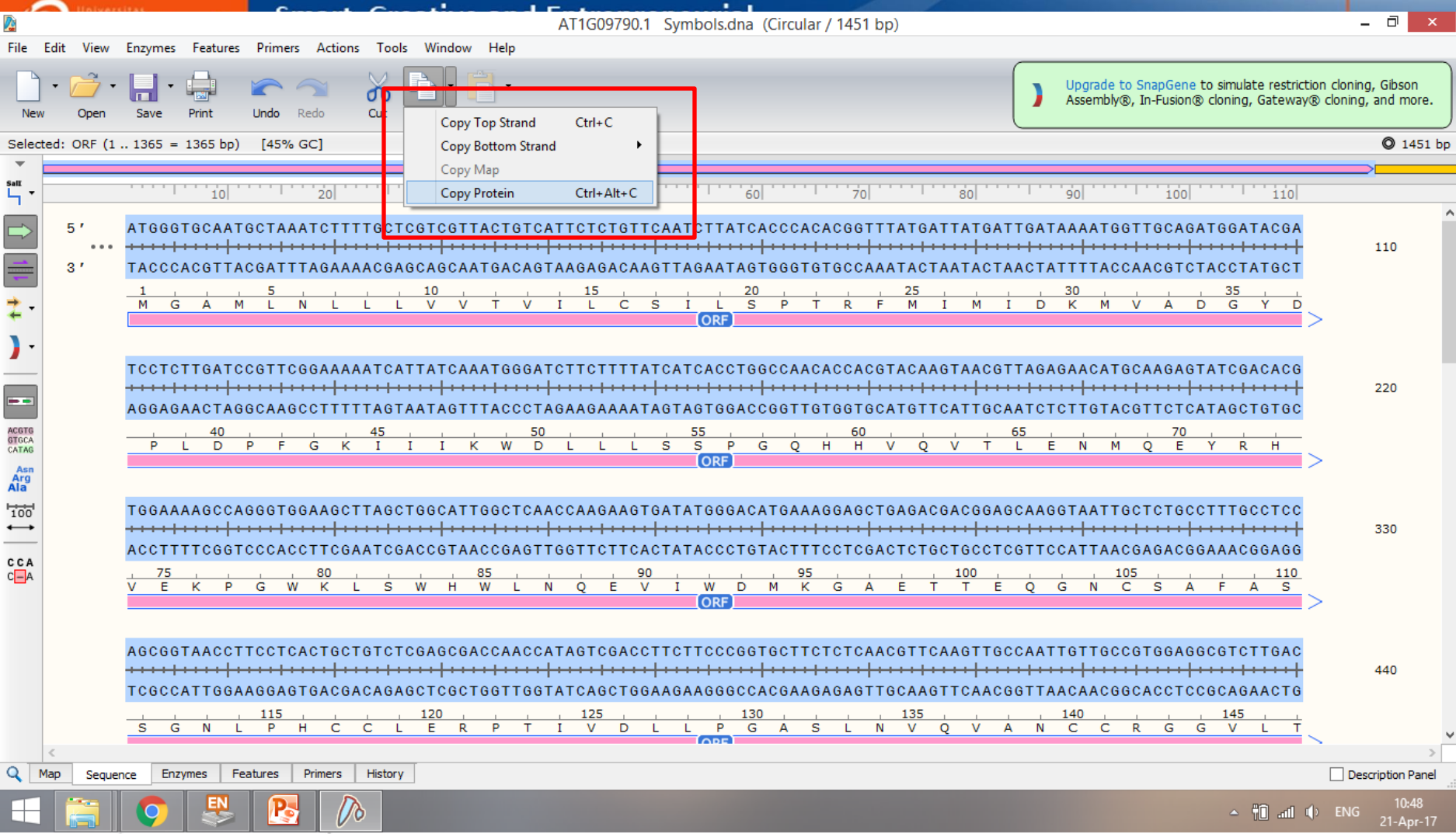


Sequences producing significant alignments:

Description	Max score	Total score	Query cover	E value	Ident	Accession
<a href="#">Arabidopsis thaliana basic helix-loop-helix (bHLH) DNA-binding superfamily protein mRNA</a>	3275	3275	100%	0.0	100%	<a href="#">gi 1063679163 NM_001197954.2</a>
<a href="#">Arabidopsis thaliana basic helix-loop-helix (bHLH) DNA-binding superfamily protein mRNA</a>	3275	3275	100%	0.0	100%	<a href="#">gi 1063679161 NM_100009.3</a>
<a href="#">Arabidopsis thaliana chromosome 1 sequence</a>	3275	3275	100%	0.0	100%	<a href="#">gi 332189094 CP002684.1</a>
<a href="#">Arabidopsis thaliana basic helix-loop-helix (bHLH) DNA-binding superfamily protein mRNA</a>	3275	3275	100%	0.0	100%	<a href="#">gi 145323701 NM_001083971.1</a>
<a href="#">Arabidopsis thaliana transcription factor MYC7E, putative (At1g01260) mRNA, complete cds</a>	3275	3275	100%	0.0	100%	<a href="#">gi 21539514 AY120752.1</a>
<a href="#">Arabidopsis thaliana At1g01260/F6F3_25 mRNA, complete cds</a>	3275	3275	100%	0.0	100%	<a href="#">gi 19310466 AY079012.1</a>
<a href="#">Arabidopsis thaliana putative transcription factor BHLH13 mRNA, complete cds</a>	3275	3275	100%	0.0	100%	<a href="#">gi 18026973 AF251698.1</a>
<a href="#">Arabidopsis thaliana chromosome I BAC F6F3 genomic sequence, complete sequence</a>	3275	3275	100%	0.0	100%	<a href="#">gi 7212005 AC023628.3</a>
<a href="#">Arabidopsis thaliana clone bHLH013 putative bHLH transcription factor mRNA, complete cds</a>	3253	3253	100%	0.0	99%	<a href="#">gi 201</a>
<a href="#">Arabidopsis thaliana Full-length cDNA Complete sequence from clone GSLTLS19ZG04 of Adult vegetative tissue of strain col-0 of Arabidopsis thaliana (t)</a>	3245	3245	100%	0.0	99%	<a href="#">gi 424</a>

## Step 12.

Nama sekuen asam amino disesuaikan dengan hasil **Blast** sebelumnya



### Step 13.

Sekuen protein ORF dapat di ekstrak dengan cara **Copy Protein**



AT1G01260.1 Symbols.dna\* (Linear / 2377 bp)

Selected: ORF (605 .. 2377 = 1773 bp) [44% GC]

5'asmonat

ACGCG  
GEOCA  
CATAG

Asn  
Arg  
Ala

100

C C A  
C C A

Map Sequence Enzymes Features Primers History

Description Panel

Untitled - Notepad

File Edit Format View Help

>Arabidopsis thaliana basic helix-loop-helix (bHLH) DNA-binding superfamily protein mRNA

MNIGRLVWNEDDKAIVASLLGKRALDYLLSNSVSNANLLMTLGSDENLQNKLSDLVERPNA  
SNFSWNYAIFWQISRSKAGDLVLCWGDGYCREPKEGEKSEIVRILSMGREEETHQTMKRKV  
LQKLHDLFGGSEEEENCALGLDRVTDTEMLLSSMYFSFPRGEGGPKCFASAKPVWSDVV  
NSGSDYCVRSFLAKSAGIQTVVLPTDLGVVELGSTSCLPESEDSILSIRSLFTSSLPPVR  
AVALPVTVAEKIDDRNTKIFGKDLHNSGFLQHHHHQQQQPPQQQHRQFREKLTVRKM  
DDRAPKRLDAYPNNGNRFMFSNPGTNNNTLLSPTWVQPENYTRPINVKEVPSDEFKFLPL  
QQSSQRLPPAQMQIDFSAASSRASENNSDGGEGGEWADAVGADESGNRRPRKRRPANG  
RAEALNHVEAERQRREKLNRFYALRSVVPNISKMDKASLLGDAVSYINELHAKLKVMEAE  
RERLGYSSNPPISLDSINVQTSGEDVTRINCPLESHPASRIFHAFEEESKVEVINSNLEV  
SQDVLHTFVVKSEELTKEKLISALSREQTNSVQRTSSGR\*

### Step 14.

Paste sekuen asam amino ke **Notepad**  
Lengkapi data dengan sekuen-sekuen lainnya

Arabidopsis Jasmonat aa - Notepad

```
>Arabidopsis thaliana basic helix-loop-helix (bHLH) DNA-binding superfamily protein mRNA
MNIIGRLVWNEDDKAIIVASLLGKRALDYLLSNSVSNANLLMTLGSDENLQNKLSDLVERPNA
SNF5WNVAIFWQISRSKAGDLVLCWGDGYCREPKEGEKSEIVRILSMGREEETHQTMKRKV
LQKLHDLFGGSEENCALGLDRVTDTEMLLSSMYFSFPRGEGGPKCFASAKPVWLSDVV
NSGSDYCVRSFLAKSAGIQTVLVLPDGLVVELGSTLPESEDSILSIRSLFTSSLPPVR
AVALPVTVAEKIDDNRTKIFGKDLHNSGFLQHHQHQQQQPPQQQHRQFREKLTVRKM
DDRPFKRLDAYPNNGNRFMSNPGTNNNTLLSPTWVQENYTRPINVKEVPTDEFKFLPL
QQSSORLLPPAQMQIDFSAASSRASENNSDGGEGGAWADAVGADESGNNRPRKRRRPPANG
RAEALNHVEAERQRREKLNRQFYALRSVVPNISKMDKASLLGDAVSYINELHAKLKVMEAE
RERLGYSSNPPISLSDSINVQTSGEDVTVRINCPLESHPASRIFHAFEEKVEVINSNLEV
SQDQTVLHTFVVKSEELTKEKLISALSREQTNSVQSRSSGR*
```

Save As

File name: Arabidopsis Jasmonat aa.fasta

Save as type: Text Documents (\*.txt)

Encoding: ANSI

Save Cancel

Step 13.  
Nama sekuen asam amino disesuaikan dengan hasil **Blast** sebelumnya

## Step 15.

Simpan dokumen dalam bentuk **.fasta**

Penting: Untuk menyimpan dalam format fasta, tambahkan **.fasta** di belakang nama file

**Selamat berjuang!**

THANK  
YOU



607132.wordpress.com

Noviani's Blog

