

IBT 432 Aplikasi Bioinformatika

Praktek: Anotasi sekuen ke genom

Riza Arief Putranto

Rencana Perkuliahan

- ~~1. Kontrak belajar dan pengenalan bioinformatika aplikatif~~
- ~~2. Database sekuen dan analisis genomika~~
3. Anotasi sekuen ke genom - Praktik
4. Analisis komparasi genomika I
5. Analisis komparasi genomika II
6. Analisis komparasi genomika III
7. Analisis komparasi genomika – Praktik
8. Protein modelling I
9. Protein modelling II
10. Protein modelling III
11. Protein modelling - Praktik
12. Visualisasi protein modelling
13. Visualisasi protein modelling - Praktik
14. Presentasi mahasiswa

Always remember this slide



Scanning....

Genomic DNA
(NC, NT, NW)

Model mRNA (XM)
(XR)



Model protein (XP)

Curated mRNA (NM)
(NR)



Curated Protein (NP)

RefSeq

GenBank
Sequences

[RNM_002467](#)

[GBC000141](#)

[GBC000917](#)

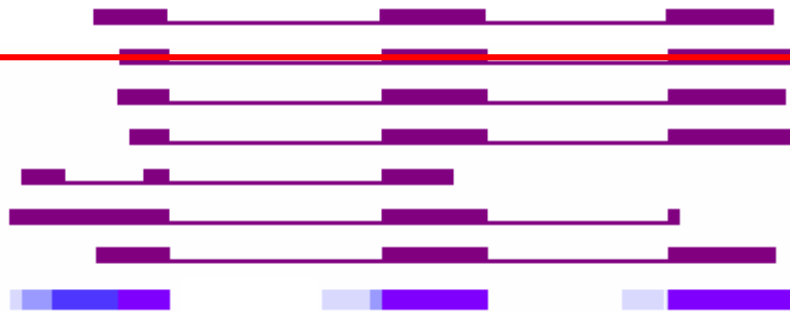
[GK02276](#)

[GM13929](#)

[GM13930](#)

[GV00568](#)

EEESTs



Remember this one also

The screenshot displays the Genome Data Viewer interface for the region **Homo sapiens: GRCh38.p11 (GCF_000001405.37) Chr 1 (NC_000001.11): 109,687,201 - 109,694,340**. The interface includes several key components:

- Current Chromosome Ideogram:** Located at the top, showing a chromosome ideogram with a red box highlighting the current region.
- Gene & Exon navigator:** A navigation bar below the ideogram showing the selected region **Region GSTM1** and transcript **NM_000561.3**, with a red box around it.
- Search results:** A sidebar on the left showing search results for **GSTM1**, including a table of genes and their locations.
- Sequence viewer:** The main central area displaying genomic tracks for **NC_000001.11: 110M..110M (7.1Kbp)**, including gene annotations, dbSNP variants, and RNA-seq coverage.
- Additional functions:** A callout box at the bottom right lists features like **Uploading data**, **Adding tracks**, and **Loading BLAST results**.

Name	Location
GSTM1	Chr1: 109,687.8K - 109,693.7K
TP53	Chr17: 7,668,402 - 7,687,550
MTHFR	Chr1: 11,785,730 - 11,806,103
GSTT1	NT_187633.1: 270.3K - 278.5K
VEGFA	Chr6: 43,770,209 - 43,786,487
GSTP1	Chr11: 67,583,595 - 67,586,653
IL6	Chr7: 22,725,889 - 22,732,002
ACE	Chr17: 63,477,061 - 63,498,380

Latar belakang praktikum

Kolaborasi INSDC merupakan basis data yang tergolong “umum”, “besar” dan banyak digunakan oleh peneliti-peneliti di seluruh dunia. Saat ini, penambahan basis data sekuen terus terjadi dan akan terus bertambah dalam tahun-tahun mendatang, seiring dengan semakin turunnya biaya untuk melakukan sekuensing genomika, transkriptomika, epigenomika, serta proteomika. Perkembangan riset yang spesifik pada spesies tertentu juga memunculkan tipe basis data spesifik dimana informasi genomik yang dianotasi berasal dari 1 jenis kingdom, famili atau bahkan genus.

Keberadaan basis-basis data tersebut juga menunjukkan bahwa riset *in silico* atau riset laboratorium kering sedang menjadi tren serta mampu menjadi alat konfirmasi dan validasi untuk hasil riset “basah” yang dilakukan di laboratorium. Riset *in silico* dapat pula digunakan untuk mempercepat prediksi dan menajamkan riset “basah”, seperti desain marka untuk seleksi pemuliaan tanaman hingga *drug design*. Hal tersebut diperkuat dengan fakta bahwa perkembangan teknologi komputasi yang pesat sudah memberikan kemudahan analisis bioinformatika skala kecil pada komputer desktop atau laptop.

Tujuan praktikum

1. Mahasiswa memahami cara pengambilan sekuen dari *database* nukleotida
2. Mahasiswa memahami runtutan penggunaan *database* nukleotida, BLAST dan bagaimana memanfaatkan informasi tersebut
3. Mahasiswa memahami cara anotasi manual dalam piranti lunak bioinformatika
4. Mahasiswa memahami makna exon (CDS) dan domain protein dalam anotasi

Sequence-to-Genome annotation

>seq

```
ATGGCCACCAACCAACTCCTCCCCCTCCTCCTCCTCGCCCTCGT
CGCCGGCTCCCATGCGGGCAGCATCGCTGTCTACTGGGGCCA
GAACGGCAATGAGGGCACCTCGCCGACACCTGCTCCAGCG
GCCTCTACGCNTACGTCATCCTCGCCTTCCTCACCACATTCGG
CAACGGCCAGACCCCCGTCCTCAACCTCGCAGGGCCACTGCGA
CCCCAACGCCGGCACCTGCACCAGCCTCTCCTCCGACATCCA
GTCCTGCCAGTCCCAGGGTGTCATCGTCCTTCTCTCCCTCGGC
GGTGGCTCCGGCAGCTACTCC
```

Google Drive: goo.gl/BbeR9o

Praktikum A. Homology search

Test de vitesse Internet | Boîte de réception (116) | Webmail - Main | Google Agenda - septem | National Center for Biotechnology Information | BLAST: Basic Local Alignment Search Tool

https://blast.ncbi.nlm.nih.gov/Blast.cgi

NIH U.S. National Library of Medicine | NCBI National Center for Biotechnology Information | putranto My NCBI Sign Out

BLAST®

Home Recent Results Saved Strategies Help

Basic Local Alignment Search Tool

BLAST finds regions of similarity between biological sequences. The program compares nucleotide or protein sequences to sequence databases and calculates the statistical significance. [Learn more](#)

NEWS
A new version (1.4.0) of the BLAST RNA-seq mapping tool, Magic-BLAST, is now available
Tue, 21 Aug 2018 16:00:00 EST [More BLAST news...](#)

Web BLAST

Nucleotide BLAST
nucleotide ► nucleotide

blastx
translated nucleotide ► protein

tblastn
protein ► translated nucleotide

Protein BLAST
protein ► protein

BLAST Genomes

Windows taskbar: 20:43, 20-Sep-18

1. Buka situs BLAST NCBI: <https://blast.ncbi.nlm.nih.gov/Blast.cgi>
2. Lakukan Nucleotide BLAST

Praktikum A. Homology search

The screenshot displays the NCBI BLAST web interface. The browser address bar shows the URL: https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome. The page title is "BLAST® >> blastn suite". The main heading is "Standard Nucleotide BLAST".

The "Enter Query Sequence" section is highlighted with a red box. It contains the following text:

```
Enter accession number(s), gi(s), or FASTA sequence(s)
ATGGCCACCAACCAACTCTCCCTCTCTCTCGCCCTCGTCCGCGCTCCATGCGGGCAGCATCGTG
TCTACTGGGGCCAGAAGGCAATGAGGGCACCTCGCCGACACCTGCTCCAGCGGCTCTACGNTACGTCAT
CCTCGCCTCTCTACACACATTCGGCAACGGCCAGACCCCGTCTCAACTCGCAGGCCACTGCGACCCCAAC
GCCGGCACCTGCACCAAGCTCTCTCCGACATCCAGTCTGCCAGTCCCAAGGGTGCATCGTCTCTCTCC
TCGGCGGTGGCTCCGGCAGCTACTCC
```

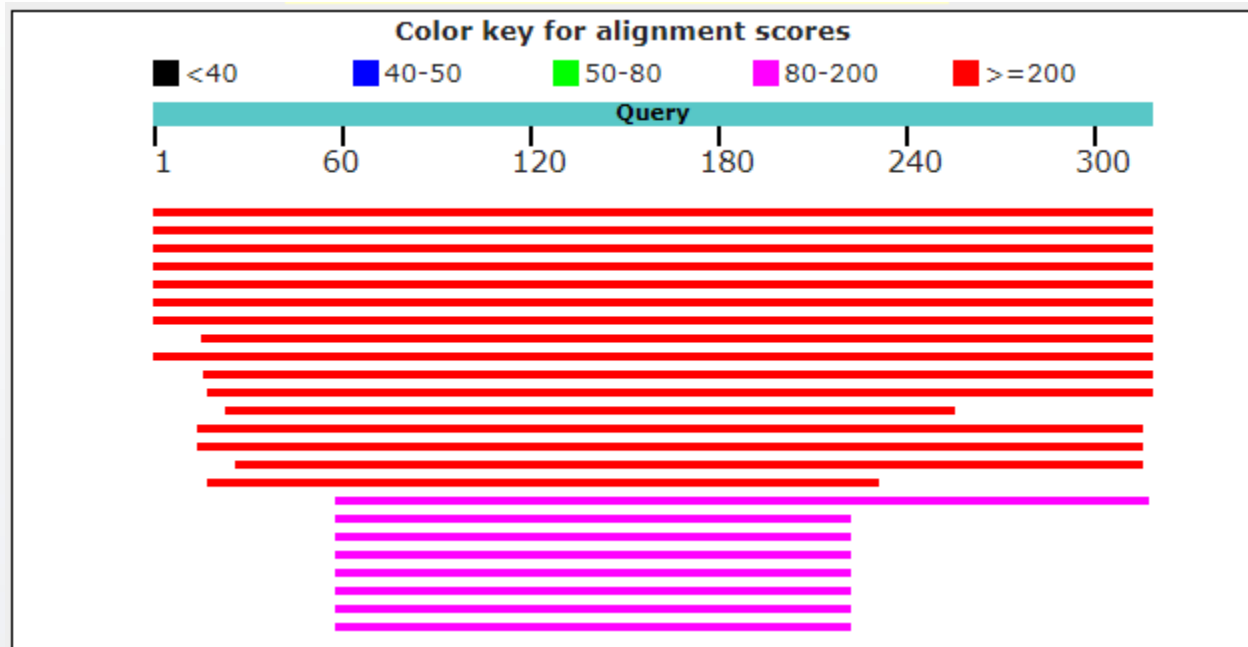
The "Choose Search Set" section is also highlighted with a red box. It shows the following options:

- Database: Human genomic + transcript Mouse genomic + transcript Others (nr etc.):
Nucleotide collection (nr/nt)
- Organism: Optional
Enter organism common name, binomial, or tax id. Only 20 top taxa will be shown
- Exclude: Optional
 Models (XM/XP) Uncultured/environmental sample sequences
- Limit to: Optional
 Sequences from type material

The "Job Title" field contains the text "seq".

1. Input sekuen ke dalam Query
2. Lakukan Nucleotide BLAST dengan parameter: Database nucleotide collection, Optimize for Highly similar sequences (megablast),

Praktikum A. Homology search



1. Apa hasil terbaik dari BLAST tersebut? Sekuen query teridentifikasi sebagai gen apa? Ambil hasil sekuen yang bukan *raw*
2. Sebutkan berapa E-value dan % homologi sekuen terbaik
3. Tuliskan no akses dari gen tersebut
4. Apakah no akses tersebut sudah tervalidasi oleh kurator? Pada tingkat mana? *Raw*, *model* atau *curated*?
5. Akses no akses tersebut dalam GenBank
6. Tuliskan no ID proteinnya dan copy paste sekuen nukleotida dan proteinnya ke dalam Notepad/Wordpad
7. Simpan sekuen dalam format .fasta

Praktikum A. Homology search

Insertion Point: 1120 1234 bp

5' ATGGTTGGTGGTAAAGCAATTAAGAGGTTGCAGATTCTTCAGACTCAACCTGTGAGGCACGAGCTCTCTCCAATTCCTCAACCCAGGGAGAAGAAGCCCATATAAG 110
3' TACCAACCACCATTTTCGTTAATTTCTCCAACGCTAAGAAGTCTGAGTTGGACACTCCGTGCTCGAGAGAGGTTAAGGGAGTTGGGTCCCTCTTTCGGGGTAATATTC

CTTTCCTCTCCTCTTACATATATTACCATGGCCACCAACCAACTCCTCCCTCTCCTCCTCCTCGCCCTAGTCGCCGGCTCCCATGCGGGCAGCATCGCTGTCTACTGGG 220
GAAAGGAGAGGAGAATGTATATAAGTGGTACCGTGGTGGTTGGTTGAGGAGGGAGAGGAGGAGGAGCGGGATCAGCGGCCGAGGGTACGCCCGTCTAGCGACAGATGACCC

GCCAGAACGGCAATGAGGGCACCCTCGCCGACACCTGCTCCAGCGGCCCTCTACGCCCTACGTCAACCTCGCCTTCTCACCACATTGCGCAACGGCCAGACCCCGTCTCTC 330
CGGTCTTGCCGTTACTCCCCTGGGAGCGGCTGTGGACGAGGTCGCCGGAGATGCGGATGCAAGTGGAGCGGAAGGAGTGGTGTAAAGCCGTTGCCGGTCTGGGGGCAAGG

AATCTCGCAGGCCACTGCGACCCAGCGCCGGTACCTGCACCAAGCCTCTCCTCCGACATCCAGTCTGCGAGTCTCAGGGCGTCAAGGTCTCCTCTCCTCTCGGGCGGTGC 440
TTAGAGCGTCCGGTGCAGCTGGGGTCGCCGCCATGGACGTGGTCGGAGAGGAGGCTGTAGGTCAGGACGGTCAGAGTCCCAGGAGTCCAGGAGGAGAGGGAGCCGCCACG

CTCCGGAAGCTACTCCCTCTCCTCCGCGACGACGCCAAAGCGTCGCCAACTACCTTTGGGACAACCTCCTTGGTGCCAGCTCCTCCTCTCGCCCTCCTCGGCATGCCG 550
GAGGCTTCGATGAGGGAGAGGAGGGCGGCTGCTGCGGGTTTCGCAAGCGGTTGATGAAACCCCTGTTGAAAGGAACACCCTCGAGGAGGAGAGCGGGGGAGCCGCTACGGC

TCCTCGATGGCATCGACTTCGACATCGAAGCAACCAAGGAGACTTCTTCGACGACCTTGCAAAGGATCTGGCGAGTTCAGCAGCCAGGGAAAGAAAGTCTACCTCACG 660
AGGAGCTACCGTAGCTGAAAGCTGTAGCTCCGTTGGTTGCCTCTGAAGAAGCTGCTGGAACGTTTCTAGACCGCGTCAAGTCTGCGGTCCCTTTCTCCAGATGGAGTGC

GCAGGCGCACAGTGCCTGACCCGATGCGCACCTCAACACGGCGCTGCAGACTGGGATTTTCGACTATGTTGGATCCAGTCTACAACAACCCATCGTGGGAGTACTC

Map Sequence Enzymes Features Primers History Description Panel

21:49
20-Sep-18

1. Lakukan anotasi manual di SnapGene untuk CDS (Exon) dan UTR
2. Berapa panjang basa CDS-nya? Berapa % GC content-nya?
3. Berapa panjang proteinnya? Ukuran molekul proteinnya?
4. Apakah ditemukan 5'-UTR?
5. Apakah ditemukan 3'-UTR?

Ingat dogma biologi molekuler!

Sekuen diatas adalah mRNA, tidak ada intron. Kita temukan sekuen DNA-nya melalui *sequence-to-genome annotation*

Praktikum B. Genome Data Viewer

The screenshot shows the NCBI Genome Data Viewer interface. At the top, there is a navigation bar with the NIH logo and the text "U.S. National Library of Medicine" and "NCBI National Center for Biotechnology Information". The main heading is "Genome Data Viewer". Below this, there is a search bar for the organism, which currently contains "Elaeis guineensis (African oil palm)". To the right of the search bar, there is a brief description: "GDV is a genome browser supporting the exploration and analysis of more than 600 eukaryotic RefSeq genome assemblies." Below the search bar, there is a phylogenetic tree showing the relationships between various plant species. The tree is rooted on the left and branches out to the right. The species shown are: rice, malo sina, African oil palm, date palm, stiff brome, Aegilops tauschii, pineapple, maize, sorghum, and foxtail millet. The "African oil palm" node is highlighted in blue. To the right of the tree, there is a panel titled "Elaeis guineensis (African oil palm) genome". This panel contains a search bar for the genome, a "Browse genome" button, and a "BLAST genome" button. Below these buttons, there are sections for "Assembly details" and "Annotation details". The "Assembly details" section lists: Name: EG5, RefSeq accession: GCF_000442705.1, GenBank accession: GCA_000442705.1, Download via FTP: RefSeq, GenBank, Submitter: Orion Genomics, Level: Chromosome, and Category: Representative genome. The "Annotation details" section lists: Annotation Release: 101 and Release date: 2017-01-11. At the bottom of the panel, there is a chromosome map showing chromosomes Chr1 through Chr14. The Windows taskbar is visible at the bottom of the screenshot, showing the Start button, taskbar icons for File Explorer, Chrome, PowerPoint, and other applications, and the system tray with the date and time (21:52, 20-Sep-18).

1. Buka situs Genome Data Viewer: <https://www.ncbi.nlm.nih.gov/genome/gdv/>
2. Select organism: *Elaeis guineensis*
3. Ada berapa kromosom tanaman ini?

Praktikum B. Genome Data Viewer

Select organism
Elaeis guineensis (African oil palm)

Genome Data Viewer
GDV is a genome browser supporting the exploration and analysis of more than 600 eukaryotic RefSeq genome assemblies.

Search in genome
XM_010926494

Genes Other

Name	Location
LOC105047536	Chr6: 39.93M - 39.93M

Examples: PG4, chr2:2403000-2408000, DNA repair

Browse genome BLAST genome

Assembly details
Name EG5
RefSeq accession GCF_000442705.1
GenBank accession GCA_000442705.1

Feedback

1. Masukkan no akses gen yang di dapat sebelumnya ke **Search in genome GDV**
2. Akses gen tersebut dalam tampilan Genome Browser

Praktikum B. Genome Data Viewer

The screenshot displays the NCBI Genome Data Viewer interface for the Elaeis guineensis genome. The main track shows the gene NC_025998.1 (40M..40M, 2.7Kbp) with its exon-intron structure. Below the gene track, RNA-seq tracks show exon coverage and intron-spanning reads. The interface includes a search bar, navigation tools, and a footer with various resources and a feedback button.

Region: LOC105047536 (Gene) | XM_010926494.2 (Transcript)

NC_025998.1: 40M..40M (2.7Kbp)

RNA-seq exon coverage, aggregate (filtered), NCBI Elaeis guineensis Annotation Release 101 - log base 2 scaled

RNA-seq intron-spanning reads, aggregate (filtered), NCBI Elaeis guineensis Annotation Release 101 - log base 2 scaled

Warning: No track data found in this range

Tracks shown: 5/138

1. Terletak di kromosom berapa gen tersebut?
2. Catat lokasi gen tersebut di dalam kromosom dari basa berapa ke berapa?
3. Ada berapa jumlah exon dan intron?

Praktikum B. Genome Data Viewer

The screenshot displays the NCBI Genome Data Viewer interface for the Elaeis guineensis genome. The main track shows the NC_025998.1 sequence with various annotations. A context menu is open over a sequence feature, with the 'Download' option selected, showing sub-options for FASTA (Visible Range), GenBank Flat File (Visible Range), and PDF file (Visible Range). The interface includes a search bar, navigation tools, and a sidebar with options like 'User Data and Track Hubs', 'BLAST', 'Add Tracks', and 'History'. The browser address bar shows the URL: https://www.ncbi.nlm.nih.gov/genome/gdv/browser/?context=genome&acc=GCF_000442705.1.

1. Klik kanan di sekuen XM tersebut
2. Download FASTA (Visible Range) – file akan tersimpan dalam format FASTA

Praktikum B. Genome Data Viewer

NC_025998.1[39926036..39928748].dna (Linear / 2713 bp)

File Edit View Enzymes Features Primers Actions Tools Window Help

New Open Save Print Undo Redo Cut Copy Paste

Insertion Point: 320

2713 bp

5' TGCTCACGAAACAAAAGAGAGGGGAAAAAATCTACACGAATGTCCTTGTGGAATATTGGATCCAAATCTGGCAGGA
3' ACGAGTGCCTTTGTTTTCTCCCTTTTTTTTTAGATGTGCTTACAGGAACACCTTATAACCTAGGTTAAGACCGTCT

AGGCGTGGCTTTGCCTCAACTCTTCAAAGTCCCAGGACAGGACGACGTCGTTTCTTGCGGAACCCGTAAATCCACGTT
TCCGCACCAGAAACGGAGTTGAGAAAGTTTCAGGGCGCCGCTGCTGCAGCAAAGAACGCCTTGGGCATTTTAGGTGCAA

TTAAGCAATCCCTCCGCGATTTTTCCACGTCAAAGGCTCGCGAGGCGGCGTGCAAGCGCTTCCACATGGTTGGTGGTAA
AATTCGTTAGGGAGGCGCTAAAAAGGTGCAGTTTCAGACGCGTCCGCCGACGTTTCGCGAAGGTGTACCAACCACATT

AGCAATTAAGAGGTTGCAGATTTCTCAGACTCAACCTGTGGTTTTTCAAAGAGAAAAAAGAAGATAATGTTATCG
TCGTTAATTTCTCCAACGCTAAGAAGTCTGAGTTGGACACCAAAAAAGTTTCTCTTTTTTTCTTCTATTACAATAGC

CCTAATTGATAGCTCAATTATTCTTATTCTGTTCAAGTTACTTGATATTCGATAAAATAATGAGACTGTGCGAAACCT
GGATTAACATATCGAGTTAATAAGGAATAAGCAAGTTCAAATGAACTATAAGCTATTTTACTCTGACACGCTTTGGGA

TATTGCATATTCTTATATTTGGTTATGCTTATACCTTGCCCACTTCATCCTTACATTAATTAATGACTGCTCTCTGGCTAA
ATAACGTATAAGAAATATAAACCAATACGAATATGGAACGGGTGAAGTAGGAATGTAATTAATGACCGAGAGAGACCGATT

TTGCCGATACGAGGGACCAAATTGACCAGGCTATTACAAGAACATAAGTTTAAACCGTGACAAATTGATAATGGTAATT

Synthetic DNA Confirmed experimentally

Laboratory Host:

Methylation:
Dam⁺ Dcm⁺ EcoKI⁺

Description:
:39926036-39928748 *Elaeis guineensis* chromosome 6, EG5, whole genome shotgun sequence

Created: Today Last Modified: Today

Accession Number: NC_025998.1 Code Number:

Sequence Author:

Comments:

References:
[Click above to import from PubMed.](#)

Map Sequence Enzymes Features Primers History

22:16
20-Sep-18

1. Buka sekuen tersimpan tersebut dalam SnapGene
2. Berapa panjang basa sekuen DNA dari gen tersebut?

Praktikum B. Genome Data Viewer

NC_025998.1[39926036..39928748].dna* (Linear / 2713 bp)

File Edit View Enzymes Features Primers Actions Tools Window Help

New Open Save Print Undo Redo Cut Copy Paste

2713 bp

Synthetic DNA Confirmed experimentally

Laboratory Host:

Methylation: Dam⁺ Dcm⁺ EcoKI⁺ Change...

Description:
:39926036-39928748 Elaeis guineensis chromosome 6, EG5, whole genome shotgun sequence

Created: Today Last Modified: Today

Accession Number: NC_025998.1 Code Number:

Sequence Author:

Comments:

References:
Click above to import from PubMed.

Find DNA sequence: aacctgtg 1 match Previous Next

Map Sequence Enzymes Features Primers History

Description Panel

22:28
20-Sep-18

1. Anotasi kembali sekuen DNA dari gen tersebut: CDS (exon 1, exon 2), intron, 5'UTR dan 3'UTR. **Perhatikan langkah-langkahnya dan catat!**

Praktikum B. Genome Data Viewer

The screenshot displays the Genome Data Viewer interface for the DNA sequence NC_025998.1. The main window shows a linear map of the sequence with the following features identified:

- 5'UTR (orange arrow pointing right)
- Exon 1 (orange box, labeled below)
- Intron 1 (black line)
- Exon 2 (orange box, labeled below)
- 3'UTR (grey box)

The sequence is 2713 bp long. The right-hand panel contains the following information:

- Synthetic DNA: Confirmed experimentally
- Laboratory Host:
- Methylation: Dam⁺ Dcm⁺ EcoK1⁺
- Description: :39926036-39928748 Elaeis guineensis chromosome 6, EG5, whole genome shotgun sequence
- Created: Today; Last Modified: Today
- Accession Number: NC_025998.1; Code Number:
- Sequence Author:
- Comments:
- References: [Click above to import from PubMed.](#)

The bottom of the window shows a taskbar with Windows, File Explorer, Chrome, PowerPoint, and a drawing application. The system tray shows the time as 22:36 on 20-Sep-18.

Anotasi selesai jika seluruh bagian sekuen telah teridentifikasi

Tugas Homology Search

>ajusi

```
AAAATATATTTAAAATAATAATATTTTGTGTAATAGCCCCCAAATTCAAATATCTAATCCGTTCTATATA  
TTTCCCAAACCTTAGTGCCACTTAGACACAATAAGAAAACCGTTAACAAGAAGAAAAAAAAAAG  
ATCGAAAATGGAATATCAAATAACTTCTTAAGTGGAGAGTTTTTCCCCGGAGAACTCTTCTTCAA  
GCTCATGGAGCTCACAAGAATCATTCTTGTGGGAAGAGAGTTTTCTTACATCAATCATTTGACCAA  
TCCTTCCTTTTATCTAGCCCTACTGATAACTACTGTGATGACTTCTTTGCATTTGAATCATCAATCAT  
AAAAGAAGAAGGAAAAGAAGCCACCGTGGCGGCCGAGGAGGAGGAGAAGTCATACAGAGG  
AGTGAGGAAACGGCCGTGGGGGAAATTCGCGGCCGAGATAAGAGACTCAACGAGGAAAGG  
GATAAGAGTGTGGCTTGGGACATTCGACACCGCGGAGGCGGCGGCTCTCGCTTATGATCAGGC  
GGCTTTCGCTTTGAAAGGCAGCCTCGCAGTACTCAATTTCCCCGCGGATGTCGTTGAAGAATCT  
CTCCGGAAGATGGAGAATGTGAATCTCAATGATGGAGAGTCTCCGGTGATAGCCTTGAAGAGA  
AAACACTCCATGAGAAACCGTCCTAGAGGAAAGAAGAATCTTCTTCTTCTTCGACGTTGACAT  
CTTCTCCTTCTTCCTCCTCCTCCTATTTCATCTTCTTCGTCTTCTTCTTCTTTGTCGTCAAGAAGTAGA  
AAACAGAGTGTTGTTATGACGCAAGAAAGTAATACAACACTTGTGGTTCTTGAGGATTTAGGTG  
CTGAATACTTAGAAGAGCTTATGAGATCATGTTCTTGATAATCTCTGCTTCTACAATTTTTATGTAA  
TTTGAGTGAGCTCTGTTTTTTTTAAGAAAAAATCCGTTGAAAAGTTTTGGAAGAGATGTGATGGC  
CGTGTATAGATGTCAACTATTTGCGTACTGTTGGAAACATGTATAGATGTGTAGTTCTTTATCTTG  
GTGCTTAATTTTAATATTTATTTTTGCCATAATTCTACAAAACCGATCCGAG
```

Tugas Homology Search

1. Apa hasil terbaik dari BLAST tersebut?
2. Sekuen query teridentifikasi sebagai gen apa? Ambil hasil sekuen yang bukan *raw*
3. Sebutkan berapa E-value dan % homologi sekuen terbaik
4. Tuliskan no aksesori dari gen tersebut
5. Apakah no aksesori tersebut sudah tervalidasi oleh kurator? Pada tingkat mana? *Raw, model* atau *curated*?
6. Akses no aksesori tersebut dalam GenBank
7. Anotasi sekuen tersebut dalam SnapGene: CDS, 5'UTR dan 3'UTR

Bukti yang harus dilampirkan – print pdf/image dari

1. Halaman hasil BLAST nucleotide
2. Halaman GenBank
3. Halaman Map dari gen hasil anotasi di SnapGene

Laporan dan Tugas disubmit oleh Kormat

Google Drive: goo.gl/BbeR9o

Folder: 3b-Laporan dan Tugas Praktikum

Maksimal tanggal Jum'at, 5 Oktober 2018

It was still the third course, don't
get dizzy yet

