

APLIKASI NCBI DALAM BIOINFORMATIKA

Buku Aplikasi NCBI dalam Bioinformatika ini ditujukan untuk membantu mahasiswa dalam memahami bioinformatika dasar yang terjadi dalam sel meliputi struktur DNA, RNA, transkripsi, translasi, modifikasi pasca translasi hingga protein fungsional.

Diharapkan buku ini dapat membantu pembelajaran bioinformatika sederhana dalam pemahaman biologi sel dan molekuler. Informasi yang disajikan dibuat sederhana sehingga lebih mudah untuk dipelajari. Buku ini juga dilengkapi dengan gambar-gambar yang memudahkan pembelajaran mandiri para mahasiswa. Bagi pemula, buku ini dijadikan pegangan untuk mempelajari aplikasi virtualnya secara *online*.



PENERBIT USD
Jl. Affandi, Gejayan (Mrican) Tromol Pos 29
Yogyakarta 55002; e-mail: publisher@usd.ac.id

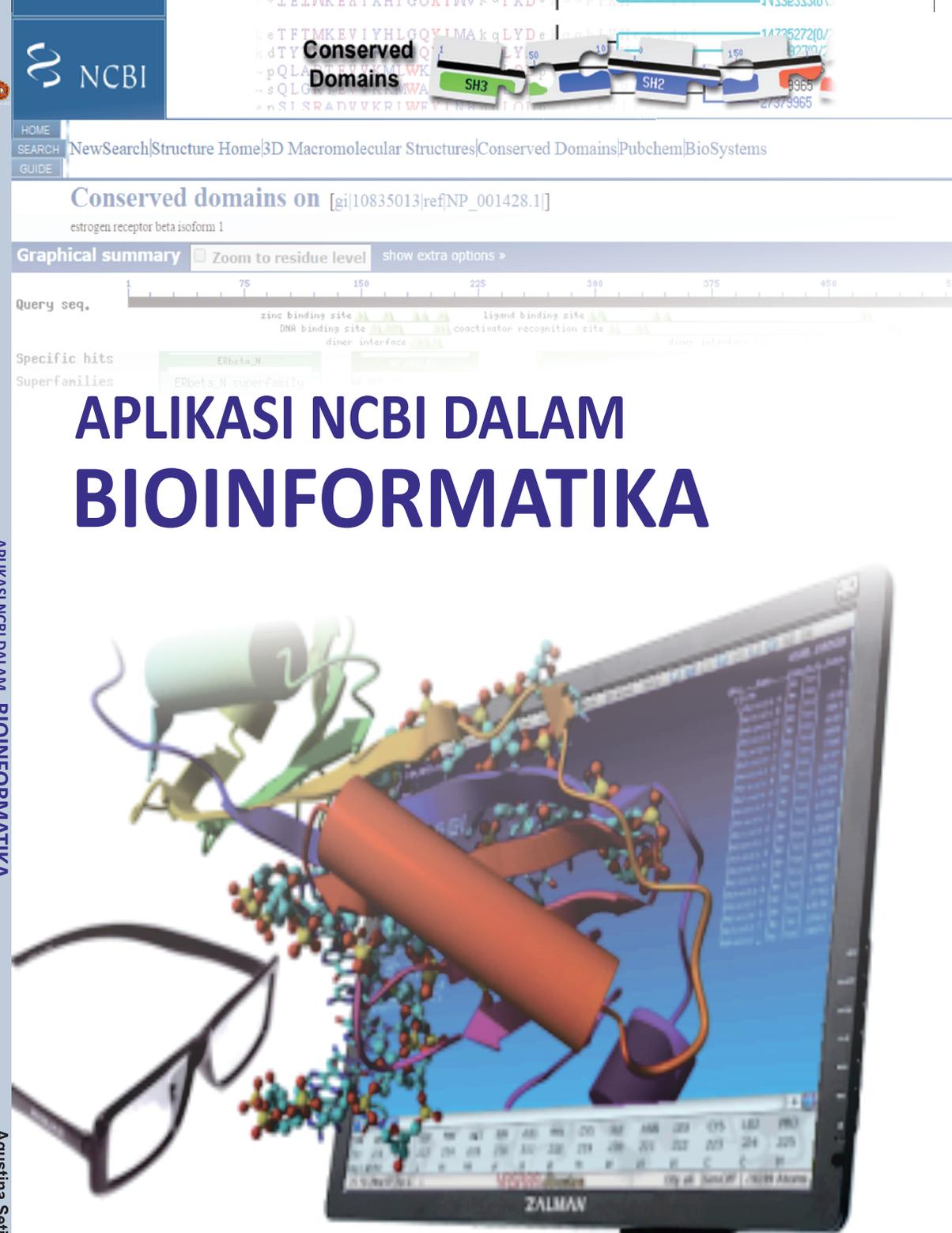


ISBN 978-602-0830-10-0



APLIKASI NCBI DALAM BIOINFORMATIKA

Agustina Setiawati



APLIKASI NCBI DALAM BIOINFORMATIKA

Agustina Setiawati

APLIKASI
N C B I
DALAM
BIOINFORMATIKA

Agustina Setiawati



Sanata Dharma University Press

APLIKASI NCBI DALAM BIOINFORMATIKA

Copyright © 2015

Agustina Setiawati. Podi Farmasi, F. Farmasi

Universitas Sanata Dharma.

Diterbitkan oleh:

Sanata Dharma University Press
Jl. Affandi (Gejayan) Mrican,
Yogyakarta 55281
Telp. (0274) 513301, 515253;
Ext.1527/1513; Fax (0274) 562383
e-mail: publisher@usd.ac.id



Universitas Sanata Dharma berlambangkan daun teratai coklat bersudut lima dengan sebuah obor hitam yang menyala merah, sebuah buku terbuka dengan tulisan "*Ad Maiorem Dei Gloriam*" dan tulisan "Universitas Sanata Dharma Yogyakarta" berwarna hitam di dalamnya. Adapun artinya sebagai berikut.

Teratai: kemuliaan dan sudut lima: Pancasila; Obor: hidup dengan semangat yang menyala-nyala; Buku yang terbuka: ilmu pengetahuan yang selalu berkembang; Teratai warna coklat: sikap dewasa yang matang; "*Ad Maiorem Dei Gloriam*": demi kemuliaan Allah yang lebih besar.

Penulis:

Agustina Setiawati

Ilustrasi Sampul & Grafis:

Skolastika

Tata Letak:

Thoms

Cetakan Pertama, September 2015

v ; 80 hlm.; 155 x 230 mm.

ISBN: 978-602-0830-10-0

EAN: 9-786020-830100



Penerbit Universitas Sanata Dharma Anggota APPTI
(Asosiasi Penerbit Perguruan Tinggi Indonesia)

Hak Cipta Dilindungi Undang-Undang.

Dilarang memperbanyak karya tulis ini dalam bentuk dan dengan cara apa pun, termasuk fotokopi, tanpa izin tertulis dari penerbit.

Puji Syukur dan Terima Kasih

**Tuhan Yesus,
pemilik genom abadi**

**Partner genomku:
Decky**

**Pewaris genomku:
Azel & Isaac**

KATA PENGANTAR

Puji Tuhan atas terselesaikannya penyusunan buku Aplikasi NCBI dalam Bioinformatika. Buku ini disusun dengan tujuan untuk membantu mahasiswa memahami bioinformatika dasar yang terjadi dalam sel meliputi struktur DNA, RNA, transkripsi, translasi, modifikasi pasca translasi hingga protein fungsional. Materi dirancang secara 'dry lab' menggunakan informasi dan basis data yang tersedia secara online tidak berbayar.

Buku diharapkan dapat membantu pembelajaran bioinformatika sederhana dalam pemahaman biologi sel dan molekuler. Informasi yang disajikan dibuat sederhana sehingga lebih mudah untuk dipelajari. Buku ini juga dilengkapi dengan gambar-gambar yang memudahkan pembelajaran mandiri para mahasiswa. Bagi pemula, buku ini dijadikan pegangan untuk mempelajari aplikasi virtualnya secara online.

Yogyakarta, 18 Agustus 2015

Agustina Setiawati, M.Sc, Apt

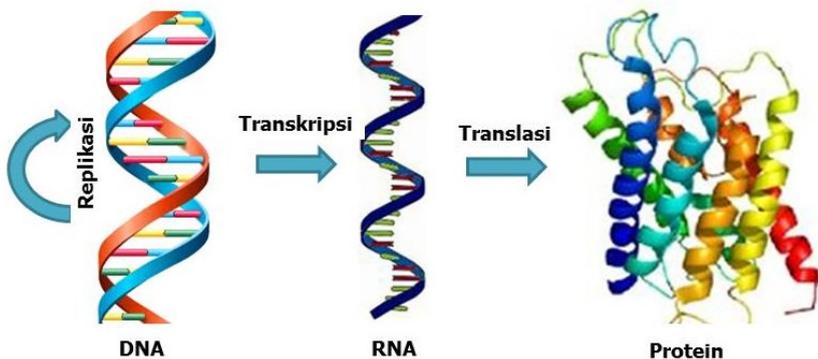
DAFTAR ISI

Halaman Judul	i
Kata Pengantar	iv
Daftar Isi	v
BAB I. PENDAHULUAN	1
BAB II. PENGENALAN NCBI	5
BAB III. ANALISIS EKSPRESI GEN	18
BAB IV. DOMAIN FUNGSIONAL DAN TRANSDUKSI SINYAL	31
BAB V. ANALISIS PENYEJAJARAN NUKLEOTIDA DAN PROTEIN	46
BAB VI. DESAIN PRIMER PADA PCR	64
DAFTAR PUSTAKA	74
LAMPIRAN	76

BAB I PENDAHULUAN

Kemajuan ilmu pengetahuan dibidang biologi molekuler telah berhasil mengungkapkan pengetahuan dasar mengenai asam nukleat baik DNA maupun RNA serta pengetahuan mengenai protein. Perkembangan ilmu biologi molekuler sangat signifikan dipengaruhi oleh perkembangan ilmu bioinformatika yang menerapkan analisis informasi biologis menggunakan bantuan komputer. Bioinformatika meliputi beberapa bidang ilmu antara lain biologi molekuler, kimia komputasi, statistika dan juga bioteknologi yang mampu membantu mengungkap rahasia proses-proses biologis dalam tingkat molekuler.

Seluruh proses dogma sentral dalam biologi sel dan molekuler dapat dipahami melalui bioinformatika (Gambar 1.1). Kode-kode genetik yang terdapat dalam DNA disalin menjadi mRNA (transkripsi) kemudian diterjemahkan menjadi protein atau polipeptida (translasi). Protein mengalami pelipatan menjadi struktur tersier atau kuarterner untuk dapat menjalankan fungsinya di dalam sel melalui proses transduksi sinyal. Selain itu, dengan bioinformatika kita dapat mengetahui kemiripan suatu gen pada organisme yang berbeda (homologi). Dengan demikian, memungkinkan kita untuk melakukan skrining atau rekayasa genetika untuk mendapatkan protein tertentu.



Gambar 1.1. Dogma sentral dalam biologi molekuler

Perkembangan ilmu bioinformatika sangat mendukung kemajuan bidang kesehatan. Aplikasi bioinformatika dapat digunakan untuk membantu teknik diagnosis penyakit, analisis asam nukleat/protein dan penemuan obat baru. Dalam diagnosis penyakit ataupun analisis asam nukleat, teknik yang sering dilakukan adalah *Polymerase Chain Reaction* (PCR). Teknik PCR adalah teknik untuk memperbanyak DNA hingga jumlahnya menjadi jutaan kali lipat jumlah semula. Reaksi yang terjadi dalam PCR didasari reaksi replikasi DNA dalam sel yaitu DNA diperbanyak oleh enzim polimerase DNA dengan bantuan primer. Primer merupakan nukleotida pendek yang digunakan sebagai awal untuk sintesis DNA baru. Primer mutlak dalam replikasi DNA atau PCR karena enzim polimerase DNA tidak bisa mensintesis DNA secara *de novo* (mulai dari nol). Primer yang digunakan dalam reaksi PCR spesifik untuk perbanyak DNA tertentu. Bioinformatika sangat berguna dalam mendesain primer yang digunakan untuk PCR. Dengan menggunakan basis data online, primer dapat didesain dengan sesuai dengan susunan DNA yang diperbanyak.

DNA hasil perbanyak reaksi PCR, selanjutnya dapat dianalisis lebih lanjut dengan menggunakan elektroforesis ataupun *southern blot*. Pada beberapa kasus kecelakaan pesawat atau penyelidikan kasus pembunuhan, PCR sangat membantu untuk identifikasi korban ataupun pelaku pembunuhan. Dengan mengisolasi DNA dari potongan jasad korban atau bercak darah yang tercecer, jumlah DNA yang didapatkan terlalu sedikit dan tidak mencukupi untuk dianalisis lebih lanjut. Oleh karena itu, supaya dapat dianalisis harus dilakukan perbanyak DNA dengan PCR.

Bioinformatika juga dapat digunakan untuk mengetahui homologi atau polimorfisme suatu protein pada berbagai organisme yang berbeda. Pengetahuan mengenai homologi itu dapat diketahui melalui teknik yang disebut sebagai penyejajaran nukleotida atau protein. Hasil analisis penyejajaran dapat diketahui tingkat kesamaan suatu protein pada

beberapa organisme yang berbeda, sebagai contohnya insulin. Insulin merupakan protein farmasetik yang digunakan untuk terapi diabetes tipe II. Penyebab dari diabetes tipe II adalah ketidakmampuan kelenjar pankreas untuk menghasilkan insulin, sehingga pemberian insulin dari luar membantu menurunkan kadar gula darah. Dengan menggunakan bioinformatika, kita dapat mengetahui insulin berbagai organisme yang mirip dengan insulin manusia.

Bagi seorang farmasis, bioinformatika membantu para ilmuwan untuk mendesain molekul obat baru. Obat memberikan efek farmakologi apabila obat tersebut dapat berikatan dengan targetnya, sebagai contoh protein dan DNA. Dengan bioinformatika, struktur protein dapat diketahui termasuk sisi aktif dan domain dari protein tersebut. Asam-asam amino pada sisi aktif protein tersebut dapat diidentifikasi kemudian dapat diprediksi interaksi sisi aktifnya dengan kandidat molekul obat. Proses pernautan protein target dan molekul obat tersebut disebut sebagai '*docking*'. *Docking* telah banyak digunakan untuk mensintesis obat baru dalam dunia kefarmasian. Prediksi aktivitas suatu obat dapat dilakukan dengan bantuan komputer sehingga mempersingkat waktu penemuan obat.

Pemahaman ilmu bioinformatika tidak lepas dari perkembangan basis data online tidak berbayar yang berisi jutaan informasi mengenai DNA, RNA dan protein dari berbagai organisme. Basis data tersebut berkembang seiring dengan semakin banyaknya penelitian genetika. Beberapa basis data online yang dapat digunakan antara lain:

1. NCBI (*National Centre for Biotechnology Information*) dalam situs www.ncbi.nlm.nih.gov
2. EBI (*European Bioinformatic Institute*) dalam situs www.ebi.ac.uk
3. DDBJ (DNA Data Bank of Japan) dalam situs www.ddbj.nig.ac.jp
4. PDB (Protein Data Bank)

Salah satu basis data yang paling populer digunakan dalam bioinformatika adalah NCBI. NCBI berisi informasi genetik berupa DNA, RNA serta protein yang mempunyai urutan sama dengan EBI atau DDBJ. Basis data tersebut berisi informasi asam nukleat (DNA dan RNA), nama organisme yang berkaitan dengan asam nukleat tersebut. Penguasaan basis data NCBI dapat membantu untuk memahami berbagai proses biologi molekuler yang terjadi dalam sel.

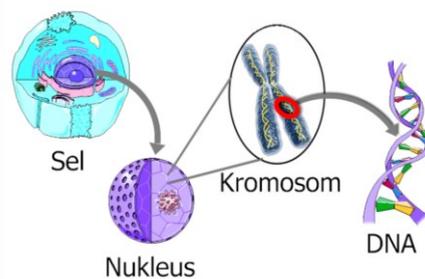
BAB II. PENGENALAN NCBI

A. SASARAN CAPAIAN

1. Mahasiswa mengenal dan menguasai NCBI dan berbagai fitur sistem informasi yang disediakan.
2. Mahasiswa mampu menjelaskan aplikasi bioinformatika dalam dunia kesehatan.

B. DASAR TEORI

Genom adalah susunan lengkap DNA, molekul yang berisi informasi genetik yang dibutuhkan untuk perkembangan dan aktivitas kehidupan suatu organisme. Dalam perkembangan penelitian mengenai genom, pemerintah Amerika Serikat mendanai suatu proyek, yaitu *Human Genome Project* (HGP). HGP merupakan proyek besar dengan tujuan untuk mengetahui seluruh susunan genetika pada manusia dan telah berhasil mengidentifikasi genom manusia yang dimulai pada tahun 1990. Susunan genetik manusia tersusun atas 3 miliar pasang basa yang menempati 23 pasang kromosom dalam nukleus sel manusia (Gambar 2.1). Pada tanggal 26 Juni 2000, diumumkan genom manusia terdapat 25.000 hingga 45.000 gen, lebih banyak dibandingkan dengan perkiraan sebelumnya.



Gambar 2.1. Letak DNA dalam sel manusia

DNA merupakan molekul pembawa informasi genetik yang strukturnya tersusun atas dua rantai yang saling berpasangan (*double helix*) yang dihubungkan dengan ikatan hidrogen. Tiap nukleotida terdiri atas gula

deoksiribosa, basa dan fosfat yang dihubungkan melalui ikatan fosfodiester. Basa pada DNA terdiri atas basa purin (adenin dan guanin) dan basa pirimidin (sitosin dan timin). Basa-basa yang terdapat dalam DNA diterjemahkan menjadi kodon pada RNA oleh proses yang disebut transkripsi, kemudian diterjemahkan menjadi protein fungsional oleh proses yang disebut translasi. Oleh karena itu, untuk memahami jalur ekspresi protein dari DNA sangat penting diketahui sebagai pengetahuan dasar ilmu biologi sel dan molekuler.

Kemajuan teknik biologi sel dan molekuler dalam mengungkapkan alur ekspresi protein yang berguna untuk mengungkap perkembangan penyakit genetik yang disebabkan oleh mutasi atau abnormalitas yang terjadi pada DNA. Pengetahuan tersebut dijadikan sebagai dasar dalam penemuan obat baru untuk pengobatan penyakit genetik tersebut. Pengobatan yang efektif dan minim efek samping dapat tercapai apabila obat dapat langsung berikatan dengan targetnya antara lain berupa DNA dan protein. Dengan demikian, seorang calon farmasis harus menguasai ilmu biologi sel dan molekuler dalam upaya penemuan obat baru.

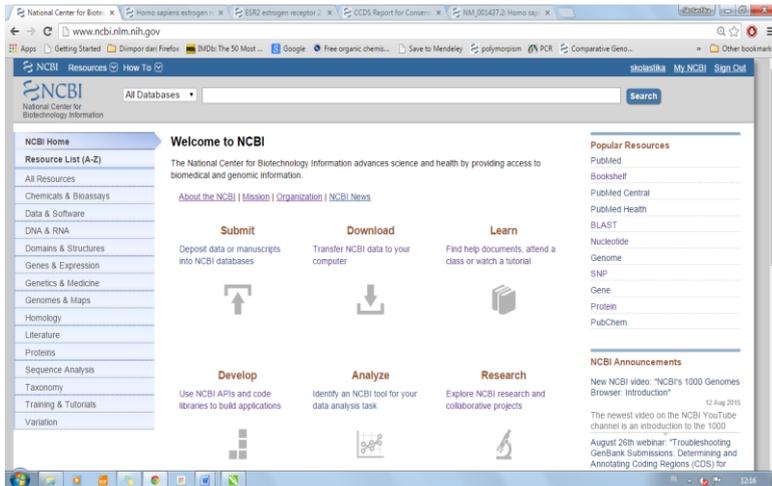
Seiring perkembangan teknologi informasi dan komputer, muncul ilmu bioinformatika yang menerapkan ilmu komputasional untuk mengelola dan menganalisis informasi biologis yang tersedia dalam basis data *online*. Basis data *online* tersebut dapat kita gunakan untuk berbagai tujuan yang mendukung ilmu biologi sel dan molekuler antara lain; mengetahui urutan DNA, RNA dan asam amino serta informasi yang terkait di dalamnya, *sequence alignment* (penyejajaran urutan nukleotida), bentuk protein dan transduksi sinyalnya. Oleh karena itu, penguasaan basis data *online* dapat diaplikasikan untuk pengembangan penemuan obat baru sesuai dengan target molekuler dalam sel tersebut.

Salah satu basis data yang dapat membantu pembelajaran biologi sel dan molekuler adalah NCBI. **NCBI (*National Centre for Biotechnology Information*)** merupakan suatu institusi yang konsen sebagai sumber informasi perkembangan biologi molekuler. NCBI membuat basis data yang dapat diakses oleh publik dan mengembangkan *software* penganalisis data genomik. Keseluruhan basis data tersebut diharapkan membantu pemahaman yang lebih proses-proses molekuler yang mempengaruhi manusia dan kesehatannya dengan lebih baik.

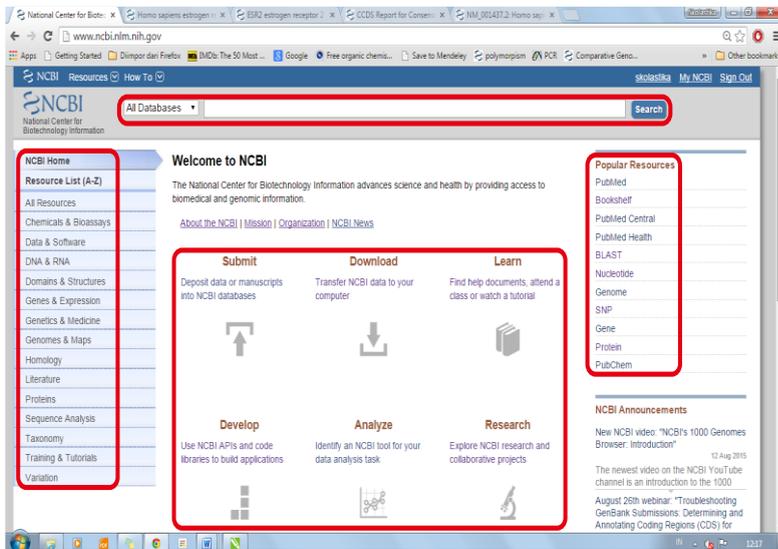
NCBI didirikan pada tanggal 4 November 1988, sebagai salah satu divisi dari *National Library of Medicine* (NLM) di *National Institutes of Health* (NIH). NIH merupakan sebuah lembaga penelitian dari Amerika Serikat yang menjadi salah satu pusat penelitian medis terkemuka di dunia. Sebagai sumber informasi biologi molekuler, misi NCBI ialah untuk mengembangkan teknologi informasi baru yang membantu memahami proses fundamental molekuler dan genetik yang mengontrol kesehatan dan penyakit.

C. PROSEDUR KERJA

1. Pengenalan situs www.ncbi.nlm.nih.gov; maka terbuka *homepage* NCBI seperti tampilan berikut:

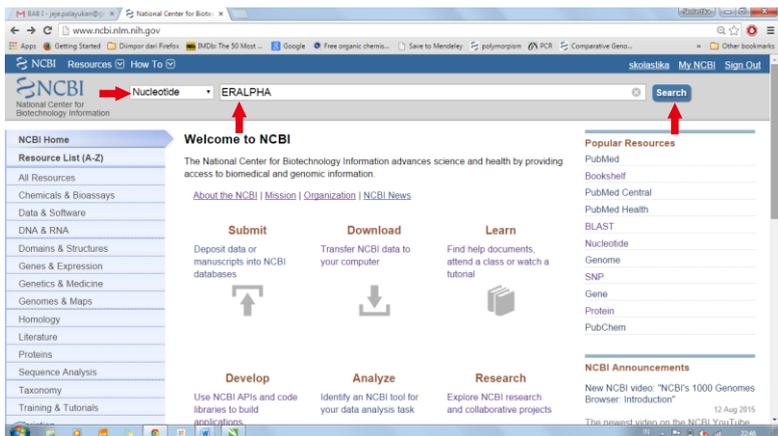


2. Pengenalan fitur yang terdapat dalam www.ncbi.nlm.nih.gov Pada bagian sebelah kiri terdapat sumber basis data yang telah dikelompokkan. Bagian tengah terdapat **fitur submit** untuk memasukkan data ke dalam basis data NCBI; **fitur download** untuk mentransfer/mendownload data NCBI ke komputer; **fitur learn** sebagai menu bantuan untuk membantu menemukan dokumen atau file yang diinginkan; **fitur develop** untuk membantu programmer dalam mengakses dan memanipulasi data NCBI pada aplikasi mereka; **fitur analyze** untuk membantu pengguna dalam analisis data; **fitur research** untuk mengetahui penelitian NCBI dan proyek kolaboratif lainnya. Di sebelah kanan terdapat sumber basis data yang populer digunakan. Pada bagian atas terdapat kotak *search engine* yang memudahkan dalam mencari dokumen/file yang dibutuhkan.

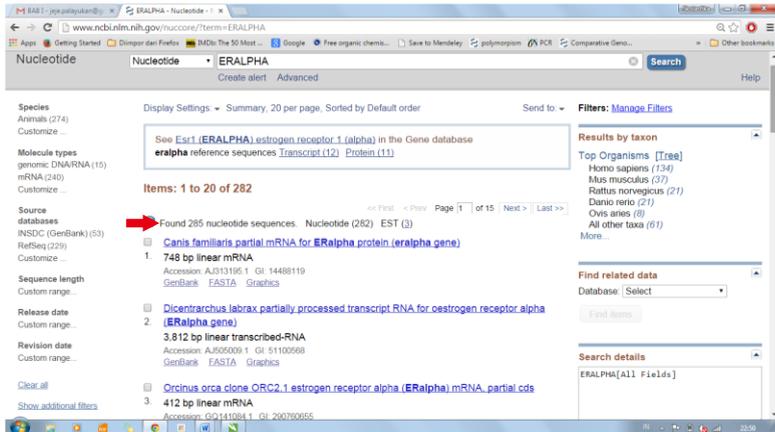


3. Pencarian Nukleotida dalam NCBI

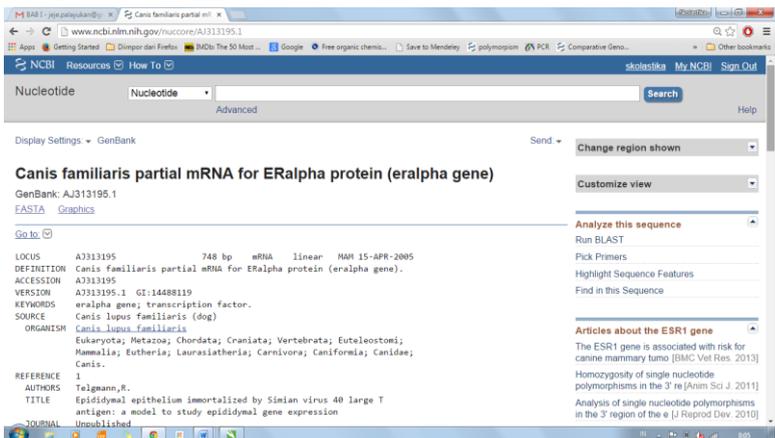
- a) Urutan nukleotida dapat dicari dengan memilih *nucleotide* pada kotak *all database* dan mengisi kotak *search* dengan nukleotida yang ingin dicari, misalnya "ERALPHA" lalu klik *search*.

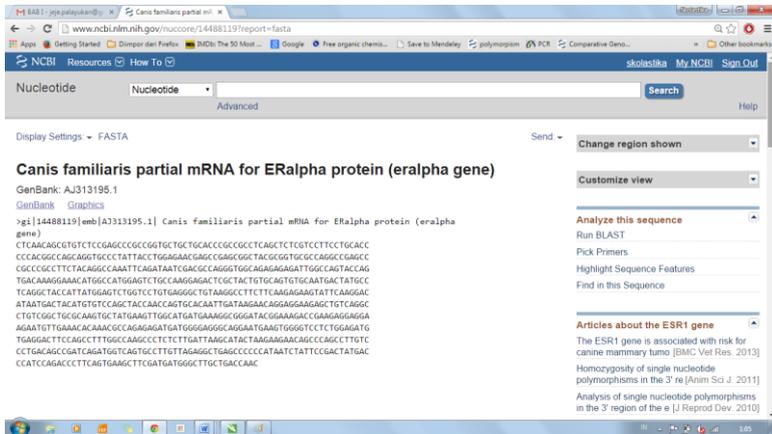


b) Beberapa hasil pencarian terkait nukleotida ERALPHA yang terdapat pada berbagai organisme ditampilkan seperti gambar berikut:

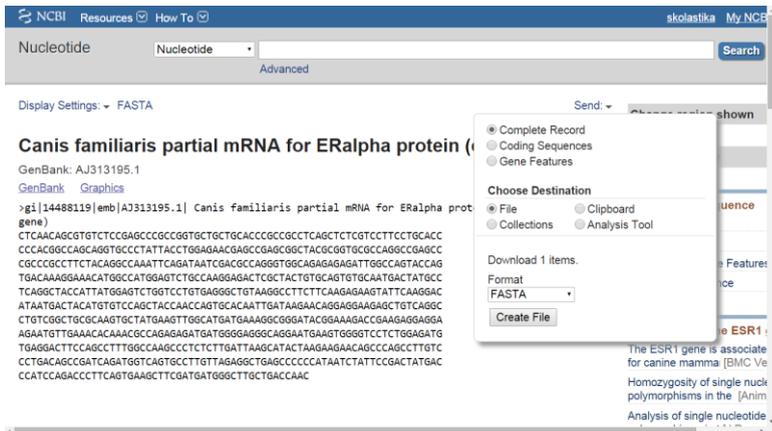


c) Klik salah satu hasil pencarian nukleotida tersebut, misalnya “Canis familiaris partial mRNA for ERalpha protein (eralpha gene)”. Informasi terkait gen tersebut muncul dalam format *genbank*.

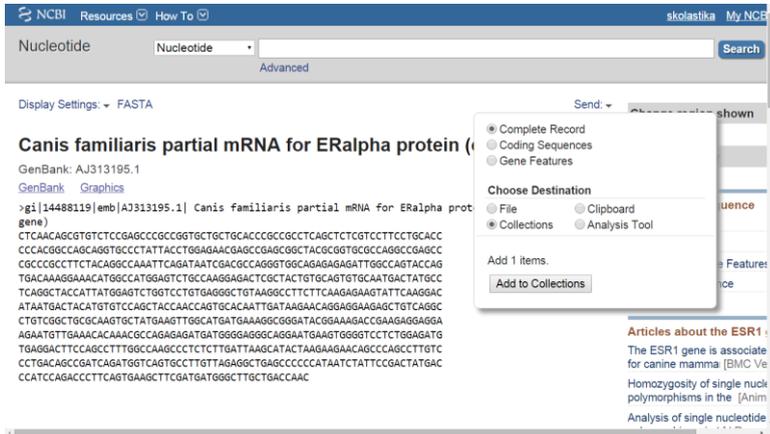




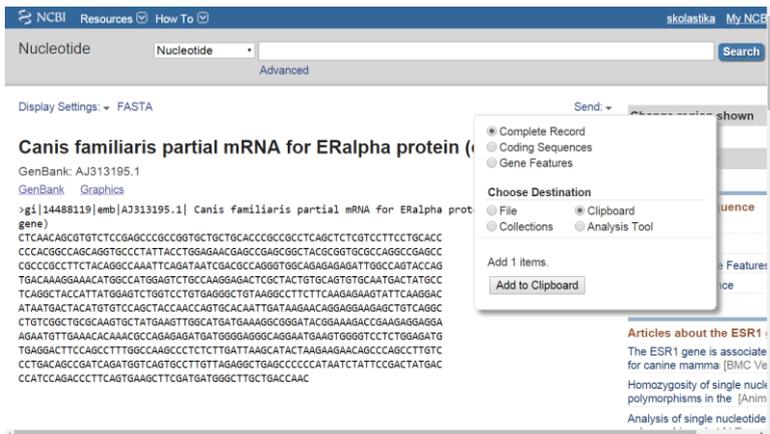
e) Ada beberapa cara untuk menyimpan data urutan nukleotida. Untuk menyimpan file dalam format FASTA, maka klik *send* → *complete record* → *File* → *FASTA* → *create file*.



f) Bila telah mempunyai akun NCBI, maka data FASTA dapat ditambahkan dalam akun dengan cara klik *send* → *complete record* → *collection* → *add to collection*.

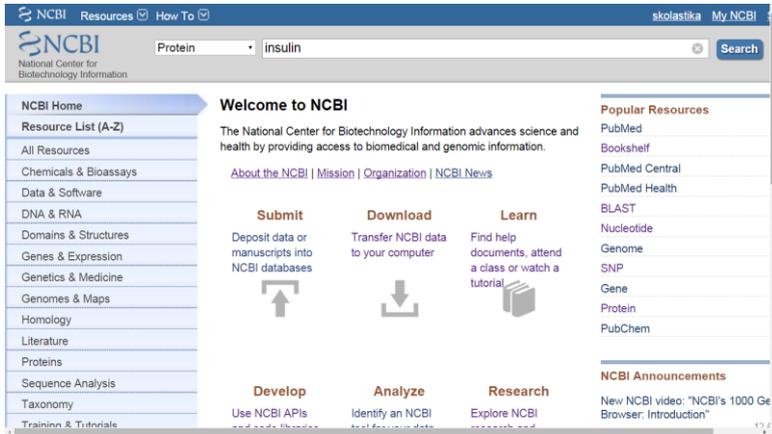


g) Selain itu, data FASTA juga dapat disimpan dengan cara klik *send* → *complete record* → *clipboard* → *add to clipboard*. Data FASTA secara otomatis tersimpan pada clipboard sehingga dapat sewaktu-waktu dibuka.

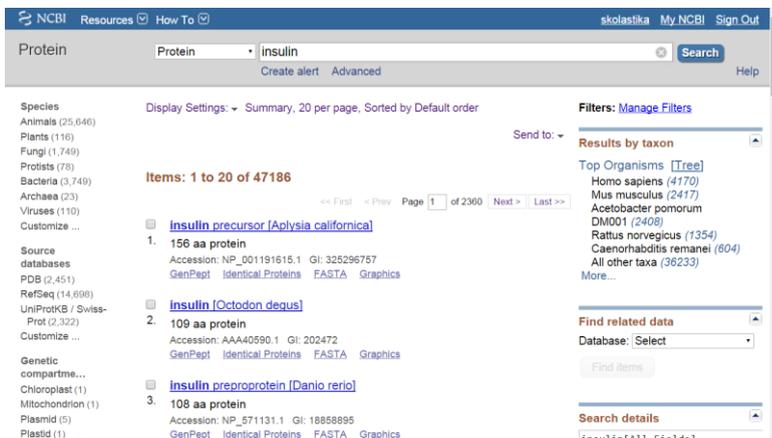


4. Pencarian Urutan Protein dalam NCBI

- a) Langkah-langkah dalam mencari urutan protein hampir sama saat mencari urutan nukleotida. Buka *website* NCBI: <http://www.ncbi.nlm.nih.gov/>
- b) Pada kotak *all database* pilih protein dan pada kotak *search* tulis nama protein yang ingin dicari. Misalnya, insulin lalu klik *search*.



c) Muncul beberapa hasil pencarian terkait protein yang dicari yang terdapat pada berbagai organisme.



d) Klik salah satu link, maka akan muncul informasi terkait gen tersebut dalam format GenPept.

NCBI Resources How To skolastika My NCBI

Protein Protein Advanced Search

Display Settings: GenPept Send to: Change region shown

insulin precursor [Aplysia californica]

NCBI Reference Sequence: NP_001191615.1

[Identical Proteins](#) [FASTA](#) [Graphics](#)

Go to:

LOCUS NP_001191615 156 aa linear INV 27-APR-2014

DEFINITION insulin precursor [Aplysia californica].

ACCESSION NP_001191615

VERSION NP_001191615.1 GI:325296757

DBSOURCE REFSEQ: accession [NM_001204686.1](#)

KEYWORDS RefSeq.

SOURCE Aplysia californica (California sea hare)

ORGANISM [Aplysia californica](#)
Eukaryota; Metazoa; Lophotrochozoa; Mollusca; Gastropoda;
Heterobranchia; Euthyneura; Euopisthobranchia; Aplysiomorpha;
Aplysioidea; Aplysidae; Aplysia.

REFERENCE 1 (residues 1 to 156)

AUTHORS Jakubowski JA, Hatcher NG, Xie F and Sweedler JV.

TITLE The first gamma-carboxyglutamate-containing neuropeptide

Analyze this sequence
Run BLAST
Identify Conserved Domains
Highlight Sequence Features
Find in this Sequence

Articles about the PIN gene
The first gamma-carboxyglutamate-containing neuropeptide [Neuroscience] Insulin prohormone processing distribution, and relative [J Neurosci]

e) Klik FASTA pada bagian atas untuk melihat urutan protein dalam format FASTA.

NCBI Resources How To skolastika My NCBI

Protein Protein Advanced Search

Display Settings: FASTA Send to: Change region shown

insulin precursor [Aplysia californica]

NCBI Reference Sequence: NP_001191615.1

[GenPept](#) [Identical Proteins](#) [Graphics](#)

```
>gi|325296757|ref|NP_001191615.1| insulin precursor [Aplysia californica]
MSKFKLLQSHSANAACLLTLLTLASNLDISLANFHSNCWYRHPHPRGLCGEDLHVIIISNLCSSLGNNRRF
LAKYMWKRDTEWVNDKLRGILLNKKFAFSYLTKREASGSITCECCFNQCRIFELAQYCRLPDHPFSSRISR
TGRSHSQAQLQEDVFS
```

Analyze this sequence
Run BLAST
Identify Conserved Domains
Highlight Sequence Features
Find in this Sequence

Articles about the PIN gene
The first gamma-carboxyglutamate-containing neuropeptide [Neuroscience] Insulin prohormone processing distribution, and relative [J Neurosci]

Reference sequence

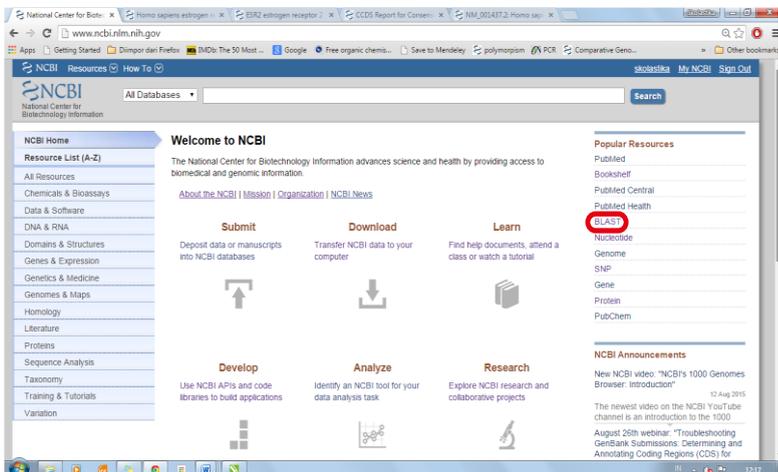
Tabel I. Asam Amino dan Kodenya (IUPAC, 1968)

Asam amino	Kode 3 huruf	Kode 1 huruf	Asam amino	Kode 3 huruf	Kode 1 huruf
Alanin	Ala	A	Leusin	Leu	L
Arginin	Arg	R	Lisin	Lys	K
Asparagin	Asn	N	Metionin	Met	M
Aspartat	Asp	D	Fenilalanin	Phe	F
Sistein	Cys	C	Prolin	Pro	P
Glutamin	Gln	Q	Serin	Ser	S
Glutamat	Glu	E	Treonin	Thr	T
Glisin	Gly	G	Triptofan	Trp	W
Histidin	His	H	Tirosin	Tyr	Y
Isoleusin	Ile	I	Valin	Val	V

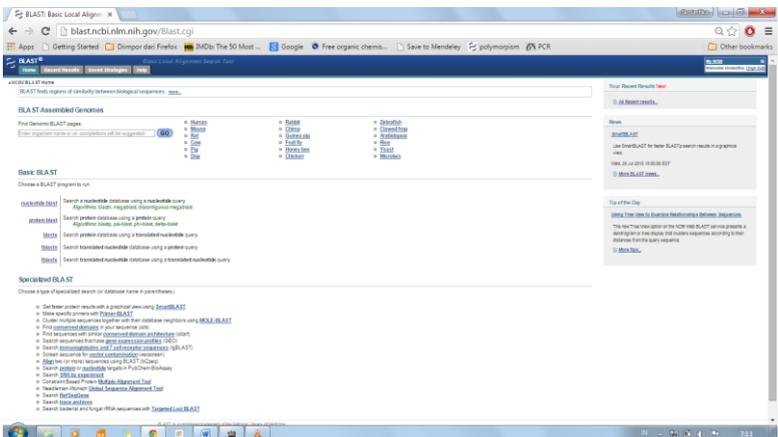
f) Data urutan protein dapat disimpan dengan cara yang sama seperti pada penyimpanan data urutan nukleotida.

5. Penyejajaran Urutan Nukleotida/Protein

BLAST (*Basic Local Alignment Search Tool*), yang terdapat pada sisi kanan *homepage* NCBI, merupakan suatu alat pencari digunakan untuk penyejajaran sekuen yang mirip dengan sekuen meragukan yang kita miliki melalui perbandingan sekuen melalui GenBank DNA.



6. Berikut merupakan tampilan basis data BLAST:



7. Ada 5 program utama dalam BLAST, yaitu :
- a) **Nucleotide blast (blastn)**: membandingkan suatu sekuen nukleotida yang kita miliki dengan database sekuen nukleotida.
 - b) **Protein blast (blastp)**: membandingkan suatu sekuen asam amino yang kita miliki dengan database sekuen protein.
 - c) **Blastx**: membandingkan produk translasi konsep 6-frame sebuah sekuen nukleotida (translated nucleotide) yang kita miliki dengan database sekuen protein.
 - d) **tblastn**: membandingkan suatu sekuen protein yang kita miliki dengan database sekuen nukleotida yang secara dinamis ditranslasi pada semua pembacaan 6 frame.
 - e) **tblastx**: membandingkan suatu translasi 6 frame dari nukleotida.

Selain itu, pada bagian bawah terdapat program khusus yang makin mempermudah dan mendukung pencarian.

D. PERTANYAAN DISKUSI

1. Mengapa *Human Genome Project* penting dalam perkembangan ilmu biologi sel dan molekuler?
2. Apakah definisi ilmu bioinformatika?
3. Mengapa mahasiswa farmasi perlu mempelajari ilmu bioinformatika?
4. Apakah yang dimaksud dengan BLAST dan aplikasinya?
5. Apakah yang dimaksud NCBI?

BAB III. ANALISIS EKSPRESI GEN: TRANSKRIPSI DAN TRANSLASI

A. SASARAN CAPAIAN

1. Mahasiswa memahami prinsip ekspresi gen yang meliputi transkripsi dan translasi serta regulasinya.
2. Mahasiswa dapat mengidentifikasi bagian gen yang berperan sebagai promoter, *coding DNA sequence* (CDS), intron dan ekson, start dan stop kodon, serta *regulating element*.
3. Mahasiswa dapat memahami terjadinya mekanisme *splicing* dan modifikasi post translasi.

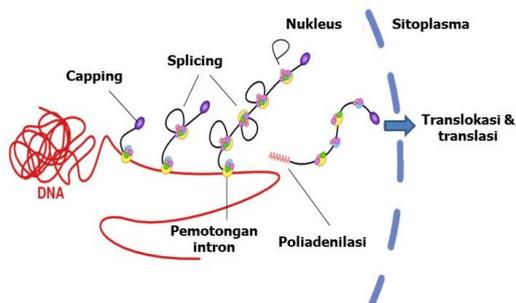
B. DASAR TEORI

Sel-sel penyusun tubuh organisme mempunyai susunan kromosom yang sama akan tetapi jenis protein yang diekspresikan sel tersebut tidak selalu sama. Variasi jenis protein tersebut disebabkan oleh perbedaan pola ketersediaan protein pada masing-masing sel, atau dengan kata lain disebabkan karena perbedaan pola ekspresi gen. Dalam sel, protein tidak selalu diekspresikan tetapi ada mekanisme pengaturan ekspresi suatu gen. Proses ekspresi gen ini ditentukan oleh sel itu sendiri. Jika sel tersebut membutuhkan suatu protein maka gen tertentu akan diekspresikan, sebaliknya jika sel tersebut tidak membutuhkan protein maka gen tersebut tidak diekspresikan.

Ekspresi protein diawali dengan proses transkripsi, yaitu proses sintesis mRNA dengan menggunakan cetakan DNA. Proses ini terjadi dalam nukleus menggunakan enzim DNA polimerase. Pada prokariotik, mRNA langsung ditranslasi menjadi polipeptida. Transkripsi dan translasi berlangsung bersamaan pada sel prokariotik sedangkan pada eukariotik keduanya berlangsung tidak bersamaan. Dalam sel eukariotik, mRNA

setelah selesai proses transkripsi mengalami proses pematangan sebelum ditranslasikan menjadi polipeptida di sitoplasma.

Proses maturasi mRNA antara meliputi proses *capping*, *splicing* dan *tailing*. *Capping* adalah proses penutupan ujung 5'P mRNA dengan menggunakan molekul 7 metil guanosa oleh enzim guanil transferase. Tujuan *capping* adalah melindungi mRNA dari degradasi dari enzim eksonuklease 5'- 3'. Selanjutnya mRNA dihilangkan intronnya dari ekson melalui proses *splicing*. Intron adalah bagian nukleotida yang tidak bisa diterjemahkan menjadi protein, sedangkan ekson sebaliknya. Proses maturasi mRNA selanjutnya adalah *tailing*, yaitu menutup ujung 3'OH dengan molekul poliadenil dengan bantuan enzim poli(A) polimerase. Tujuan *tailing* adalah untuk melindungi mRNA dari degradasi dari enzim eksonuklease 3'- 5'. Molekul mRNA yang sudah matur siap untuk diterjemahkan menjadi protein pada proses translasi.

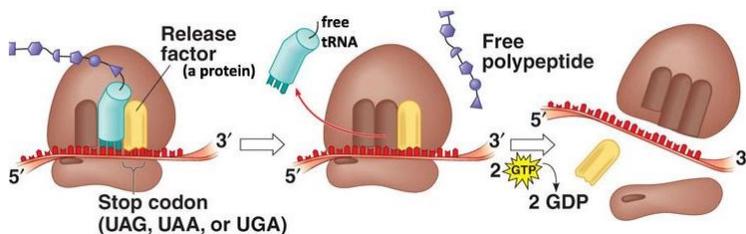


Gambar 3.1 Proses pematangan mRNA: *capping*, *splicing* dan *tailing*
(sumber: <http://bioinfo-fr.net/>)

Translasi dari proses inisiasi, yaitu berikatannya dua sub unit ribosom dengan mRNA pada daerah spesifik, yaitu 10-30 nukleotida pada daerah *upstream* dari *start* kodon. Sebelum proses translasi berlangsung, terjadi proses aktivasi yaitu proses melekatkan asam amino ke tRNA melalui penambahan gugus adenil oleh ATP dengan katalisis enzim aminoasil tRNA sintetase. Asam amino yang teraktivasi dapat membentuk ikatan ester antara gugus COOH dan gugus 3'OH tRNA yang menghasilkan aminoasil

tRNA. Aminoasil tRNA inilah yang melekat pada kodon mRNA dalam proses translasi.

Proses inisiasi memerlukan suatu kodon spesifik yaitu AUG (menyandi asam amino metionin). Setelah proses inisiasi selesai maka dilanjutkan dengan terjadinya proses perpanjangan (elongasi). Pada proses perpanjangan rantai polipeptida ini dapat dikatakan melewati siklus yang terdiri dari 3 tahap. Pertama molekul aminoasil tRNA terikat pada A-site dengan berikatan dengan kodon mRNA yang terdapat pada A-site. Kedua, gugus karboksil akhir dari rantai polipeptida terlepas dari molekul tRNA yang ada pada P-site dan membentuk ikatan peptida dengan aminoasil tRNA pada A-site. Reaksi pembentukan ikatan peptide tersebut dikatalisis oleh suatu enzim peptidil transferase. Ketiga, rantai polipeptida yang terbentuk pada A-site tersebut kemudian dipindahkan ke P-site. Selanjutnya, tRNA bebas yang semula menduduki P-site terlepas dari ribosom dan masuk sitoplasma. Proses elongasi tersebut terjadi berulang dan saat ribosom mengenali salah satu dari tiga *stop* kodon yakni UAA, UAG, dan UGA. Rantai polipeptida yang terbentuk lepas dari ribosom.



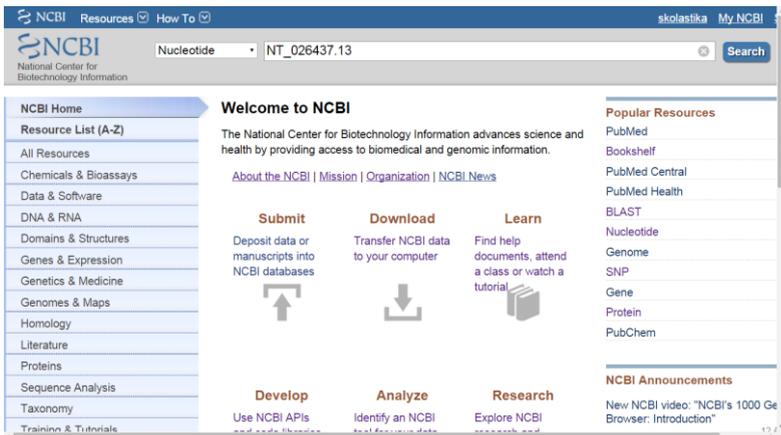
Gambar 3.2. Proses elongasi translasi
(sumber: <http://www.proteinsynthesis.org/>)

Pada percobaan kali ini mahasiswa mempelajari dogma central dalam biologi molekuler, yaitu transkripsi DNA menjadi mRNA dan translasi mRNA menjadi protein yang dengan istilah ekspresi gen. Mahasiswa harus sudah mendapatkan sekuen nukleotida dan protein pada *Homo sapiens* yang ditugaskan. Percobaan ini akan dimulai dengan melihat gambaran genom

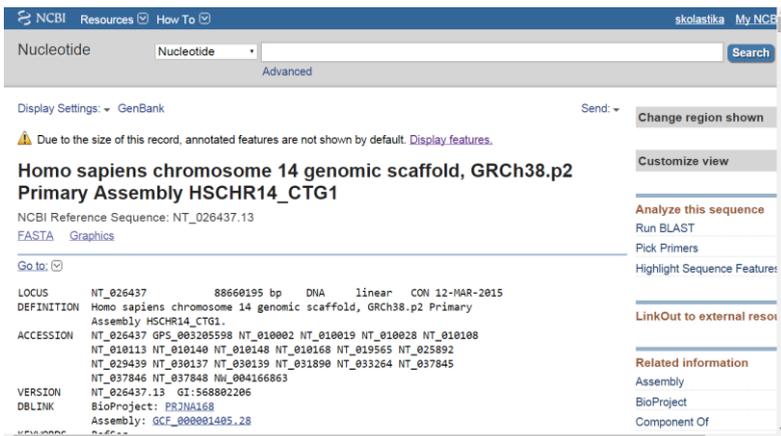
dan sekuen nukleotid terkait untuk memahami prinsip dari ekspresi gen. Selanjutnya, proses transkripsi dan translasi akan dilakukan tahap demi tahap dengan menggunakan *Open Reading Frame finder* (ORF finder).

C. PROSEDUR KERJA

1. Buka situs web <http://www.ncbi.nlm.nih.gov/> dan pilih pada kolom **Search: nucleotide, for: accession number nukleotida** yang akan dicari (**seperti tertera pada soal**) kemudian klik **Go**



2. Tampilan yang muncul seperti gambar berikut:



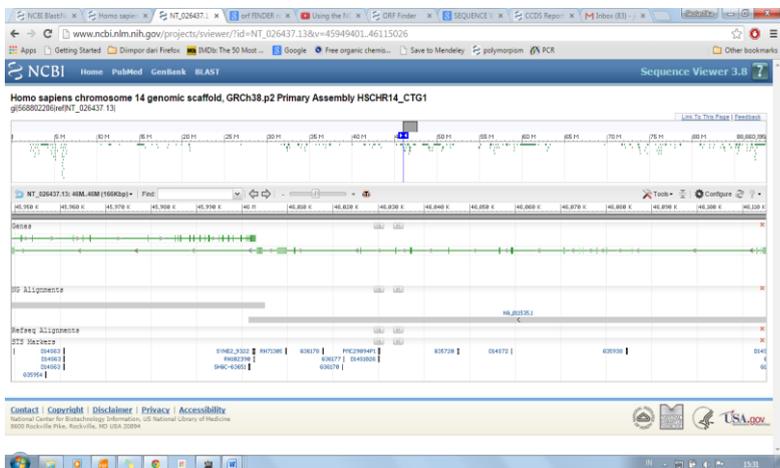
3. Klik *Graphics*, akan muncul deretan sekuen. Zoom untuk melihat informasi start kodon, dan stop kodon. Hijau adalah start kodon, dan merah adalah stop kodon.

The image displays two screenshots of a genomic browser interface, likely Ensembl, showing genomic data and gene annotations.

Top Screenshot: Shows a genomic region from 10 M to 88,660,195. The top track displays genomic coordinates. Below, the "Genes" track shows annotations for genes including *APEX1*, *HMP14*, *LGALS3*, *HIF1A*, *ESR2*, *PSENI1*, *TSHR*, *SERPINA1*, and *HSP90A*. The "NG Alignments" track shows alignment data. A "LinkOut to resources" sidebar is visible on the right.

Bottom Screenshot: Shows a zoomed-in view of a specific region from 48,121,000 to 48,121,150. The "Six-frame translations" track displays amino acid sequences with start codons (ATG) highlighted in green and stop codons (TGA, TAG, TAA) highlighted in red. The "Genes" track shows the gene *HIF1A*. The "NG Alignments", "Refseq Alignments", and "STS Markers" tracks are also visible. A "LinkOut to resources" sidebar is visible on the right.

5. Tampilan OMIM untuk mengetahui peran gen tersebut dalam fungsi biologi tubuh dan penyakit yang berhubungan dengan gen dan protein tersebut. Jika OMIM tidak terdapat pada tampilan *Map viewer*, kembali ke hasil penelusuran *accession number* dan *scroll* ke bawah. OMIM akan berada di *Related Information* yang berada di sebelah kanan. Bila tidak ada juga, maka sekuen yang anda masukkan belum memiliki data OMIM.
6. Tampilan sv menunjukkan lokasi gen di kromosom secara lebih dekat, tampilannya bisa di zoom in/out dan digeser-geser jika ingin melihat lokasi gen lain pada kromosom yang sama.

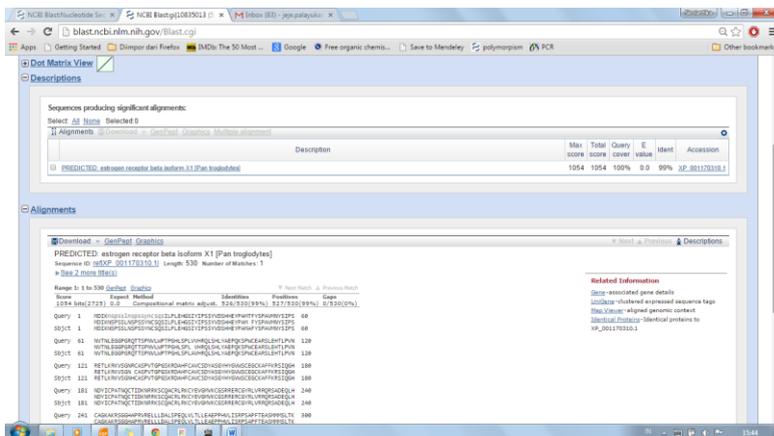


7. Pada tampilan hm atau homologene akan muncul keterangan organisme-organisme lain yang sehomolog dengan protein hasil translasi gen query, klik pada bagian yang ditunjuk dengan tanda panah, dan **analisis hasilnya untuk mengetahui persentase kemiripannya (identity)**, untuk keperluan laporan praktikum. Jika hm (*homologene*) tidak terdapat pada tampilan *Map viewer*, kembali ke hasil penelusuran *accession number* dan *scroll* ke bawah. *HomoloGene* akan berada di *Related Information* yang berada di

sebelah kanan. Bila tidak ada juga, maka sekuen yang anda masukkan belum memiliki data *HomoloGene*.



Setelah BLAST diklik, akan muncul tampilan sebagai berikut:



Cara Memperoleh CDS

1. Kembali ke hasil penelusuran *accession number* dan scroll ke bawah. Pada *related information tools* klik CCDS.

REFSeq: Data-ERBB2
COMPLETES: complete on the 5' end.

PRIMARY	REFSEQ_SPAN	PRIMARY_IDENTIFIER	PRIMARY_SPAN	CORP
1-50	AF06555.1	1-50		
51-2061	AF051427.1	1-2061		
2062-2169	AL143796.6	143227-143334		

FEATURES

Location/Qualifiers

source
1..2169
/organism="Homo sapiens"
/mol_type="RNA"
/db_xref="taxon:9606"
/chromosome="14"
/map="14q23.2"
1..2169

gene
/gene="ESR2"
/gene_synonym="ER-BETA; Erb; ESR-BETA; ESRB; ESTRB; NR3A2"
/note="estrogen receptor 2 (ER beta)"
/db_xref="GeneID:12880"
/db_xref="MIM:608136[!]"
/db_xref="HGNC:HGNC:1366[!]"
/db_xref="HUGO:ESR2[!]"
/db_xref="HUGO:ERBB2[!]"

exon
1..378
/gene_synonym="ER-BETA; Erb; ESR-BETA; ESRB; ESTRB; NR3A2"
/inference="aligner:Splice1.20.4"

misc_feature
1
/gene="ESR2"
/gene_synonym="ER-BETA; Erb; ESR-BETA; ESRB; ESTRB; NR3A2"
/note="5'-most transcription initiation site"

misc_feature
64
/gene="ESR2"
/gene_synonym="ER-BETA; Erb; ESR-BETA; ESRB; ESTRB; NR3A2"
/note="alternative transcription initiation site"

misc_feature
287
/gene="ESR2"
/gene_synonym="ER-BETA; Erb; ESR-BETA; ESRB; ESTRB; NR3A2"
/note="alternative transcription initiation site"

2. Jendela CCDS akan terbuka.

Report for CCDS9762.1 (current version)

CCDS	Status	Species	Chrom.	Gene	CCDS Release	NCBI Annotation Release	Ensembl Annotation Release	Links
9762.1	Public	Homo sapiens	14	ESR2	18	107	79	N I P R F

Public source: CCDS release 1, NCBI annotation release 35.1, Ensembl annotation release 23

Sequence IDs included in CCDS 9762.1

Original	Current	Source	Nucleotide ID	Protein ID	Status in CCDS	Seq. Status	Links
✓	✓	EBI/WTSI	ENST00000341099	ENSP00000341923	Accepted	alive	N I P R F
✓	✓	EBI/WTSI	OTTHUM0000028061	OTTHUMP00000180647	Accepted	alive	N I P R F
✓	✓	NCBI	NM_00144712	NP_0014281	Accepted	alive	N I P R F

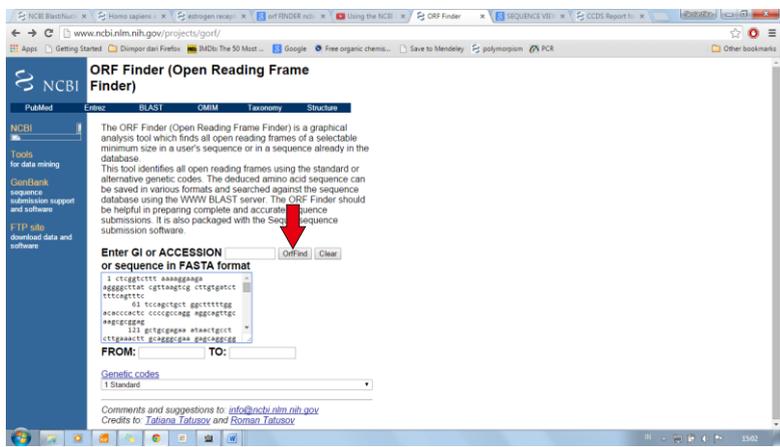
RefSeq Length Related UniProt/SwissProt Length Identity Gaps Mismatcher
NP_0014421.1 530 Q82733.4 530 100% 0 0

Chromosomal Locations for CCDS 9762.1
Assembly GRCh38.p2 [GCF_000001465.2](#)
On "+" strand of Chromosome 14 (NC_000014.9)

2. Masukkan data *origin* ke dalam kolom *ORF finder*.



```
ORF228
1 cctgctttt aaaggagga aegggattat cgttaagtc cttgtgact ttctgtttt
62 tccagctgt gcttttttg aaaccactc cccgcagcg aegcagctg aagcagcag
121 gctcagaa aaactgctt ttgaacttc gggagagaa gaaagagag agagcttgg
181 gctggagag gacaccacaa gctgagcag cctctgagc tggagagag gctggagc
241 cggagcagp gctggagag gctgcttgc ttctccaca agctggagc agggctcag
301 cggagagc cccctctag cggagaaac aactgtctc attttttag aagaaagcc
361 agctgtta tctgagag attactctg cccagagat ttgagagc tttgagagc
421 ctttggct cttctgaa ggtttctt cagtttat ctcagact cggataaaa
481 aactcaact ctgtttaa tctctctc tctcaactc gactgactc caattctcc
541 tggagagag gcttcataa caatacttc tctctgagc aagaccaca tgaatac
601 gctagact tctatgac ttgttgagc aattacagc tccagcagc tgcactaac
661 tggagagag agctgttag gaaacaa acccaactc ttgttgagc aaactgac
721 cacttttc ctttagagt ccttgagc tttcaactc tgaatagag actcaaaag
781 agctcttc gtagagag agctgagc caactctc cttgagag agagagc
841 aaaggagag ttatggag ccgtgtgac agcctgta cttgtcagc ttcagagag
901 gactcttc tctgagc ttgagctc taagagctc gactctca tggatttc
961 tggcttag gatttagc cttcttcaa agagacttc aagaccata tggattact
1021 ttctcgtc caattgag tactactg aaanagc gaaagctg ctagcctgc
1081 agctcaga agcttaga aagagagc gtagagagc gctccagc agagagct
1141 aagctcagc ttgtgagag aagagagc gctcagagc agctgactc tgcagagc
1201 gcaagagc agctgagc agctctcag gtagagagc ttgttgagc gctctgagc
1261 ccagagagc ttgtgtcac cctcttagc gctgagc ccaagctg gctagagc
1321 ccaagagc cctctcagc gctctcagc agagagc tggactgct gctgagagc
1381 aggtgtagc aactagtag ctagcagc aagatctcc gctttgagc gctagagc
1441 ttgcagagc ttgatttt agagactc ttgtagagc ttgattagc aggtttagc
1501 ttgagagc ttgagagc agagagctc actctgagc caactctg tctagagc
1561 gatggagag aatgtaga agagcttgc gaattttgc aactgactc gcaactact
1621 tcaagcttc aggtgtaa actcaactc aaagactc tctatgagc agcaactc
1681 cgtcaact caagtaga cactctgctc aagagagc agagctga cagcagc
1741 aagctgctc actgtaga agctgagc gactgagc ttgagagc ttgagagc
1801 gcaactctc ccaagagc atcagagc ctagtaact tctgagct cctctcac
1861 gtcagagc agtaaaaa aggtttaga cctctgta aactgagc caaaaatgc
1921 gctcagct agactctc actgtagc ctgagagc aactgctc agatagagc
1981 tctctcra cggatgag gtagagagc gtagagagc gtaaaagag aagagcttc
2041 aagaccac agctcagct agctgagc ctgagagc ctgagagc agagagc
2101 gtagagct gaactcagc gtagagagc ctaggcttc actctctc tgtatgctc
2161 cttantgg
//
```



NCBI ORF Finder (Open Reading Frame Finder)

The ORF Finder (Open Reading Frame Finder) is a graphical analysis tool which finds all open reading frames of a selectable minimum size in a user's sequence or in a sequence already in the database. This tool identifies all open reading frames using the standard or alternative genetic codes. The deduced amino acid sequence can be saved in various formats and searched against the sequence database using the WWW BLAST server. The ORF Finder should be helpful in preparing complete and accurate sequence submissions. It is also packaged with the Seq-Ed sequence submission software.

Enter GI or ACCESSION or sequence in FASTA format

FROM: TO:

Genetic codes:

Comments and suggestions to info@ncbi.nlm.nih.gov
Credits to [Tatiana Talavay](#) and [Domen Talavay](#)

3. Tampilan ORF akan muncul dengan menunjukkan 6 cara pembacaan CDS (*six frames*). Untuk itu, klik *frame* untuk melihat urutan asam amino yang dihasilkan. Frame yang tepat diketahui dengan mencocokkan arah pembacaan dengan urutan asam amino yang muncul pada tampilan *sequence viewer*

ORF Finder (Open Reading Frame Finder)

Anonymous

View 1 GenBank Redraw 100 Stratemes

Frame from to	Length
-1 469,2061 1593	
-2 1095,1676 582	
-1 1,342 342	
-1 616,1893 258	
-2 1749,1988 240	
-2 902,1114 213	
+3 1137,1295 159	
-2 2,145 144	
-2 1955,2074 120	
-1 892,1008 117	
-3 497,610 114	

Length: 530 aa

Accept Alternative Initiation Codons

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

Urutan Asam Amino Frame 1

Nucleotide Sequence (1593 nt):

ATGATGAAAAAACCACCATCAGCCCTAAATCTCCTCTCTACAACCTGCAGTCAATCCATCTTAC
 CCCTGGAGCACGGCTCCATATACATACCTTCCTCCTATGTAGACAGCCACCATGAAATCCAGCCATGAC
 ATTCCTATAGCCCTGCTGTGTGAATTAACAGCATCCAGCAAGTGTCACTAACTTGGAAAGTGGGCTGTGT
 CGGCACACGACCAAAAGTGTGTGTGGCCAAACCTGGGCACCTTCTCCTTATAGTGGCCATCGCC
 AGTTATCACACTGTATGCGAACCTCAAAGAGTCCCTGGTGTGAAGCAAGATCGCTAGAACAACCTT
 ACCCTGTAACAGAGAGACATGAAAAGGAAGGTAGTGGGAACGTTGCGCCAGCCCTGTACTGGTCCA
 GGTTCAAAGAGGGATGCTCACTTCGCGCTGTCTGCAGCGATTACGCATCGGGATATCACTATGGAAGT
 TGGTCTGTGAAGGATGTAAGGCCCTTTTAAAGAAAGCATCAAGGACATAATGATTATATTTGTCCAGC
 TCAAAATCAGTGTACAATCGATAAAAACGGCGCAAGAGCTGCCAGGCTGCCGACTTCGGAAGGTTTAC
 GAACTGGGAATGGTGAAGTGTGGCTCCCGAGAGAGAGAATGGTGGTACCCTCTGTCCGGAGACAGAGAA
 GTGCTGCAACGACAGGCTGCAGACTGTGCTGGAGAGCCCAAGAGAGTGGCGCCACCCGCCCGAGTGGGA
 GCTGCTGTGCAGCCCTGAGCCCGCAGCAGCTAGTGTCACTCCCTCGAGAGCTGAGCCGCACTGTG
 CTGATCAGCCGCCCCAGTGCGCCCTTACCAGGCGCTCCATGATGATGCTCCGTGACCAAGTTGGCCGACA
 AGGAGTGGTACACATGATCAGCTGGGCAAGAAGATTCGGGCTTTGTGGAGCTCAGCGCTGTATCGACCA
 AGCTGGGCTCTTGGAGAGCTGTGGATGGAAGGTTAATGATGGGGCTGATGGGCGCTCAATGGACCC
 CCGGCGAAGCTCATCTTTCAGATCTTGTCTGGACAGGGATGAGGGGAAATGCGTGAAGAAATTC
 TGGAAATCTTTGACATGCTCTGGCACTACTCAAGGTTTCGAGAGTTAAAATCCCAACCAAGAATA
 TCTCTGTCTCAGAGCTGCAGCTGCTGCTCAATTCAGATGATACCTCTGGTCCACAGCAGCCAGGATGCT
 GACAGCAGCCGGAAGCTGGCTCACTTCTGAAAGCGTGGACGATGCTTGGTTGGGTGATGCCAAGA
 CGCGCATCTCTCCAGCAGCAATCCATGCCTGGCTTAACCTCTGATGCTCTGTCCACAGTCAAGCA
 TGGGATTAACAAGGGCATGGAAATCTGCTCAACTGAAATGCAAAAATGGGTCCAGCTGTATGACCTG
 CTGCTGGAGATGCGAATGCCACGTCTTCGCGGGTGAAGTCTCTCATCAGGGGCTCGAGTGCAGCC
 CGGCAGGGACAGTAAAAGCAAGAGGGCTCCAGAAACCACAGTCTCAAGTA

Translation (530 aa):

MKNSPSSLNPSSSYNCSQSIPLLEHGSYIPSSYVDSSHHEYPAMTFYSPAVMNYSPSNVNLGEGPGR
 QTTSPMVLMPTPGHLSPLVHRQLSHLYAEPQKSPWCEARSLEHLTPVNRETLKRVSIGNRCASPVTGP
 GSKRDHFCAVKSDYASGVHYGWSGCKAFFKRSIGSHMDYECPATNQCTIDKINRRKSCQCARLRKCY
 EVGHKCSRRECGYRLVRRQSADQLHCAAGKAKRSGGHAPVRRELLDLASPEQLVLLLEAEPHVL
 ISRPSAPFTESSMMSLKLADKLVHISNAKIPGVVLSLFDQVRLLESCEMELVLMMLMRSIDHPGKLI
 FAPDLVLDREDEKCVLEILEIFDMLATTSRFRLEKLQHKLEYLCKVAMILLNSMPVLTATQDASRR
 LAHLNVAVTDLVWIAKSGISSQQSMRLANLLMLLSHVRASNKGMEHLNMMKKNVPPVYDLLLEML
 NHAHLRGCSSITGSECSPAEDSKSEGSQNPQSQ

D. PERTANYAAN DISKUSI

1. Pada tahap apakah ekspresi suatu gen dalam sel diatur? Jelaskan!
2. Apakah yang disebut sebagai promotor?
3. Apakah yang dimaksud sebagai faktor transkripsi?
Berikan contohnya!
4. Sebutkan proses maturasi mRNA dalam eukariotik berikut dengan fungsinya!
5. Apakah yang dimaksud dengan polisistronik dan monosistronik?

BAB IV. DOMAIN FUNGSIONAL DAN TRANSDUKSI SINYAL

A. SASARAN CAPAIAN

1. Mahasiswa dapat memahami karakteristik dari protein target yang meliputi struktur tiga dimensi dan plot hidropati.
2. Mahasiswa dapat mengetahui domain fungsional dari protein target dan fungsinya.
3. Mahasiswa dapat memahami peran protein target dalam transduksi sinyal di dalam sel.

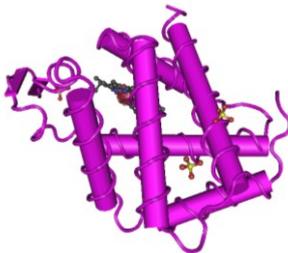
B. DASAR TEORI

Informasi genetik yang tersimpan dalam DNA disampaikan ke sel berupa ekspresinya sebagai protein. Protein merupakan makromolekul yang memegang peran penting dalam kehidupan tingkat seluler. Pemahaman struktur protein menjadi fundamental untuk memahami berbagai proses biologis yang terjadi dalam sel. Protein merupakan salah satu target spesifik aksi obat baik berupa enzim, reseptor dan hormon. Oleh karena itu, penting bagi seorang ahli biologi atau peneliti kesehatan mengetahui struktur suatu protein.

Ketika protein selesai disintesis melalui proses translasi, protein berada dalam bentuk struktur primer. Asam-asam amino penyusun struktur primer dihubungkan oleh ikatan peptida yang terbentuk antara gugus COOH dan NH₂ antar asam amino. Protein selanjutnya membentuk struktur sekunder. Dalam struktur sekunder, terbentuk ikatan-ikatan hidrogen antar asam-asam amino yang tidak bersebelahan. Struktur sekunder protein ada dua macam, yaitu α helix dan β sheet. Beberapa protein seluruhnya mempunyai struktur helix, dan beberapa protein lainnya murni β sheet. Struktur protein ada yang berupa campuran antara struktur α helix dan β sheet yang disebut struktur super sekunder atau disebut **MOTIF**. Berbagai macam bentuk motif, antara lain: $\beta\alpha\beta$ motif (yang paling umum), β hairpin

(*anti parallel strands connected by reverse turns*), $\alpha\alpha$ motif (*two successive anti-parallel α helices which are packed with their axes inclined*), β barrel (*β sheet rolling up into barrels*), *helix-loop-helix (helix turn helix)*, leucine zipper motif (*dimerization/trimerization motif*). Motif-motif tersebut bergabung membentuk suatu struktur protein yang disebut domain.

DOMAIN merupakan suatu unit penyusun protein yang berupa domain struktural ataupun domain fungsional. Protein berukuran kecil pada umumnya hanya mempunyai satu domain, sedangkan protein berukuran besar mempunyai lebih dari satu domain (multi domain). Protein eukariotik lebih cenderung multi domain dibandingkan prokariotik. Bentuk domain antara lain: α/β barrel, *four helix bundle*, α/β saddle, β/β sandwich. Domain bertanggung jawab terhadap aktivitas protein dan biasanya memiliki fungsi yang spesifik. Pembagian domain menurut fungsinya: *DNA binding domain*, *RNA binding domain*, *ligand (regulatory) domains*, *oligomerization domains*. Basis data dalam GenBank dapat dimanfaatkan untuk melacak keberadaan motif dan domain tersebut. Dari database tersebut dapat diketahui protein yang berbeda yang memiliki motif dan domain yang sama, dan mereka digolongkan dalam satu FAMILY.



Gambar 4.1. Struktur tersier hemoglobin pada *Arabidopsis thaliana*
(sumber: www.ncbi.nlm.nih.gov/structure)

Protein dapat berfungsi ketika protein berada dalam lokasi yang seharusnya dalam sel. Letak protein dalam sel sangat beragam, di membran sel atau dalam sitoplasma. Letak protein dalam sel bergantung asam-asam amino penyusun protein tersebut. Masing-masing asam amino mempunyai

skala hidropati yang berbeda-beda. Protein yang hidrofilik cenderung berada di dalam sitoplasma sedangkan protein hidrofobik cenderung berada di membran sel. Selain itu, protein yang karakteristiknya sangat khas adalah protein transmembran. Asam-asam amino yang menyusun protein transmembran terlokalisasi sesuai dengan hidrofobisitas. Asam amino hidrofobik tertanam dalam fosfolipid sedangkan asam amino hidrofilik berada di sitoplasma.

Analisis letak protein dapat dilakukan dengan analisis hidropati. Definisi hidropati adalah hidrofobisitas relatif dari suatu protein. Menurut Kyte dan Doolittle, karakteristik hidropati asam-asam amino dikategorikan menurut tabel II. Semakin besar hidropati suatu protein, maka kecenderungan protein tersebut berada dalam lingkungan yang hidrofobik seperti fosfolipid membran. Hidropati juga mempengaruhi pelipatan protein, rantai samping yang hidrofobik berada dalam struktur protein sedangkan residu yang hidrofilik berada pada permukaan. Bioinformatika sangat membantu untuk menentukan letak suatu protein berdasarkan karakteristik asam-asam amino penyusunnya. Berdasarkan basis data protein, NCBI dirancang untuk mengetahui hidropati suatu protein untuk mengetahui letak dan karakteristik protein tertentu. Selain menggunakan NCBI, program lain yang dapat digunakan untuk analisis hidropati antara lain HYDRO, ExPASy dan MPEX (Membrane Protein Explorer).

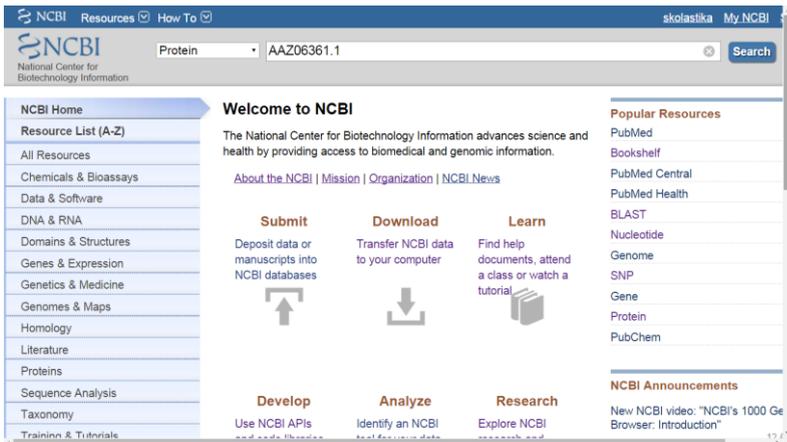
Tabel II. Indeks Hidropati Asam Amino (Kyte & Doolittle 1982)

Residue asam amino	Indeks Hidropati	Residue asam amino	Indeks Hidropati
Ile	4,5	Ser	-0,8
Val	4,2	Tyr	-1,3
Leu	3,8	Pro	-1,6
Phe	2,8	His	-3,2
Cys	2,5	Glu	-3,5
Met	1,9	Gln	-3,5
Ala	1,8	Asp	-3,5
Gly	-0,4	Asn	-3,5
Thr	-0,7	Lys	-3,9
Trp	-0,9	Arg	-4,9

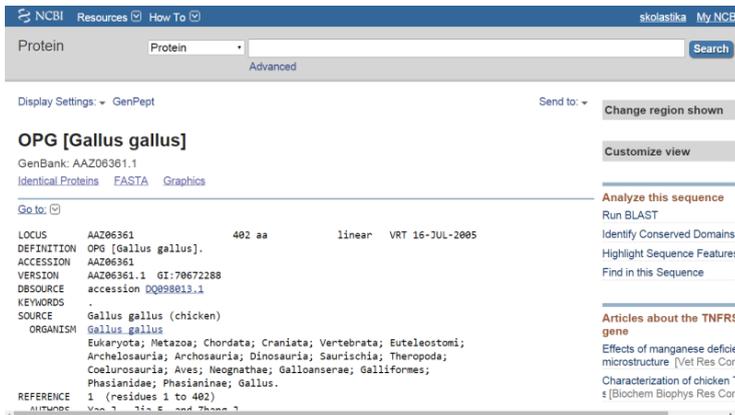
Setelah mengetahui letak suatu protein, maka hal penting yang harus diketahui ketika mempelajari biologi molekuler adalah bagaimana protein tersebut dapat menimbulkan efek fisiologis bagi sel. Protein dapat berperan untuk komunikasi sel dengan lingkungan luar dalam organisme multiseluler. Komunikasi antar sel dilakukan protein dengan berikatan dengan reseptornya pada membran atau sitoplasma untuk menimbulkan respon spesifik pada sel target. Dalam bab ini, dipelajari aplikasi bioinformatika untuk mempelajari jalur transduksi sinyal dari protein tertentu.

C. PROSEDUR KERJA

1. Buka situs web <http://www.ncbi.nlm.nih.gov/> dan pilih pada kolom **Search: protein, for: accession number protein** yang akan dicari conserved domainnya (**seperti tertera pada soal**) kemudian klik **Go**

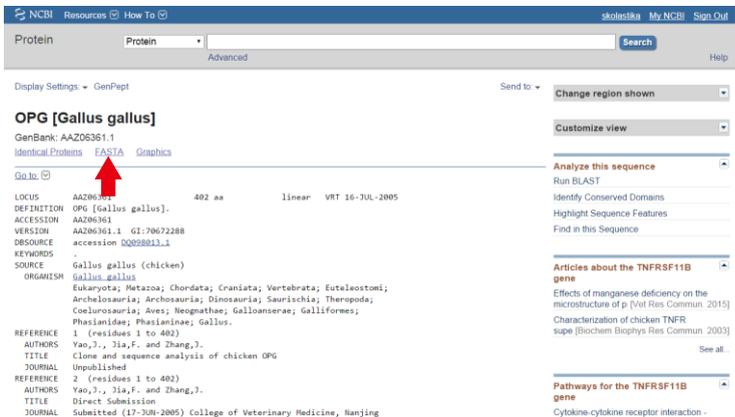


2. Tampilan yang muncul seperti gambar berikut:



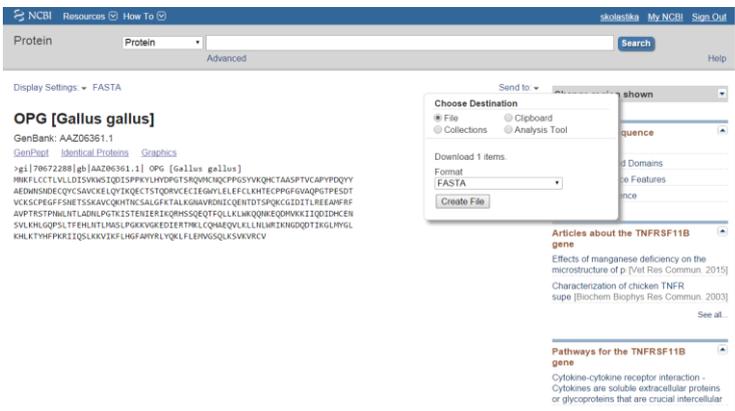
The screenshot shows the NCBI protein page for OPG [Gallus gallus]. The page includes a search bar, navigation tabs (Protein, GenPept, FASTA, Graphics), and a sidebar with options like 'Change region shown', 'Customize view', and 'Analyze this sequence'. The main content area displays protein details: GenBank: AAZ06361.1, LOCUS: AAZ06361 402 aa linear VRT 16-JUL-2005, DEFINITION: OPG [Gallus gallus], and SOURCE: Gallus gallus (chicken). A red arrow points to the 'FASTA' link in the navigation tabs.

3. Klik FASTA untuk melihat sekuen asam amino dan simpan.



This screenshot is similar to the previous one, but the 'FASTA' link in the navigation tabs is highlighted with a red arrow, indicating the next step in the process.

4. Selanjutnya klik **conserved domain** di bagian kanan.



The screenshot shows the amino acid sequence of OPG [Gallus gallus] in FASTA format. A 'Choose Destination' dialog box is open, showing options for 'File', 'Clipboard', 'Collections', and 'Analysis Tool'. The 'File' option is selected, and the 'Create File' button is visible. The sequence is: >gi|78672288|gb|AAZ06361.1| OPG [Gallus gallus] MNKFLCCFLVLDLISVWISIQDIPSPFYVHYDPGTSIQKQKICPPG5YVQKHTAASPVCAPYPPDQY AEDHNRDQVCSANCKELQYKQESTSQDVKCELEGVLELFLHTEFPFSGUAGPTPESDT VKCSPEGFNSHETSSKAVCQHTNCALSGLFKTALKGAVRNDICQENTDTSQKCGCDITLREAAHFR AVFTSTPMLNLADHLPGTKSTENIERIKQHSQEQTLQKLLWQDQKQDHWXKIQDIDKHCN SVLHLHQPSLTFEHLHTLMASSPGKGVKQEDERTFRLCQHQEQLKLLNLRHTDNGDQDITKGLMYGL NKLKTYHPFQREIQSLKXVKEFVHGFAHYLVQQLFLERPQSLQSKVXKVC

NCBI Resources How To skolatika My NCBI

Protein Protein Advanced Search

Display Settings: GenPept Send to: Change region shown

OPG [Gallus gallus]
GenBank: AAZ06361.1
[Identical Proteins](#) [FASTA](#) [Graphics](#)

Go to:

LOCUS AAZ06361 402 aa linear VRT 16-JUL-2005

DEFINITION OPG [Gallus gallus].

ACCESSION AAZ06361

VERSION AAZ06361.1 GI:79672288

DBSOURCE accession DQ988013.1

KEYWORDS

SOURCE Gallus gallus (chicken)

ORGANISM Gallus gallus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archelosauria; Archosauria; Dinosauria; Saurischia; Theropoda; Coelurosauria; Aves; Neognathae; Galliformes; Galliniformes; Phasianidae; Phasianinae; Gallus.

REFERENCE 1 (residues 1 to 402)

AUTHORS Yan J, Yan J, Liu S, and Zhou J

Analyze this sequence
Run BLAST
Identify Conserved Domains
Highlight Sequence Features
Find in this Sequence

Articles about the TNFRSF gene
Effects of manganese deficiency on the structure of the TNFRSF gene
Characterization of chicken TNFRSF11B [Biochem Biophys Res Commun]

- Jendela **conserved domain** terbuka dengan tampilan seperti berikut:

NCBI Conserved Domains

Conserved domains on [gi|70672288|gb|AAZ06361.1]

Graphical summary Zoom to residue level show extra options

Query seq. OPG [Gallus gallus]

Specific hits TNFRSF11B

Superfamilies TNFRSF superfamily TNFRSF superfamily

List of domain hits

Name	Accession	Description	Interval	E-value
TNFRSF11B	cd11041	Tumor necrosis factor receptor superfamily member 11B (TNFRSF11B), also known as	6-152	2.20e-103
TNFRSF	cd01185	Tumor necrosis factor receptor superfamily (TNFRSF); Members of TNFRF superfamily (TNFRSF)	124-188	1.21e-08
Death	cd01670	Death Domain, a protein-protein interaction domain. Death Domains (DDs) are protein-protein	324-354	4.19e-03
DD superfamily	cd14833	Death Domain Superfamily of protein-protein interaction domains. The Death Domain (DD)	232-279	1.92e-04

References:

- Marchler-Bauer A et al. (2015). "CCD: NCBI's conserved domain database." *Nucleic Acids Res.* 43:D222-6.
- Marchler-Bauer A et al. (2013). "CCD: a Conserved Domain Database for the functional annotation of proteins." *Nucleic Acids Res.* 39:D225-9.
- Marchler-Bauer A et al. (2009). "CCD: specific functional annotation with the Conserved Domain Database." *Nucleic Acids Res.* 37:D109-10.
- Marchler-Bauer A, Bryant SH (2004). "CD-Search: protein domain annotations on the fly." *Nucleic Acids Res.* 32:W337-339.

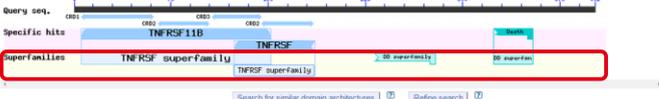
- Berdasarkan tampilan di atas, **sebutkan domain (superfamily) apa saja yang terdapat pada protein query!**
- Klik link-link yang berada di dalam kotak merah untuk melihat keterangan lebih lanjut mengenai domain tersebut.

NCBI

Conserved Domains

Conserved domains on [gi70672288|gb|AAZ06361.1]

Graphical summary

Query seq. 

Specific hits

superfamilies

TNFRSF superfamily

TNFRSF superfamily

List of domain hits

ID	Name	Accession	Description	Interval	Evalue
H	TNFRSF11B	cd10581	Tumor necrosis factor receptor superfamily member 11B (TNFRSF11B), also known as ...	6-102	2.20e-103
H	TNFRSF	cd01085	Tumor necrosis factor receptor superfamily (TNFRSF). Members of TNFR superfamily (TNFRSF) ...	124-196	1.21e-09
H	Death	cd01070	Death Domain: a protein-protein interaction domain. Death Domains (DD) are protein-protes ...	124-164	4.11e-03
H	DD super family	cd14633	Death Domain Superfamily of protein-protein interaction domains. The Death Domain (DD) ...	232-279	1.92e-04

References:

- Marchler-Bauer A et al. (2015). "CCD: NCBI's conserved domain database." *Nucleic Acids Res.* 43(Database issue):D222-6.
- Marchler-Bauer A et al. (2011). "CCD: a Conserved Domain Database for the functional annotation of proteins." *Nucleic Acids Res.* 39(Database issue):D1225-9.
- Marchler-Bauer A et al. (2009). "CCD: specific functional annotation with the Conserved Domain Database." *Nucleic Acids Res.* 37(Database issue):D1205-10.
- Marchler-Bauer A, Bryant SH (2004). "CD-Search: protein domain annotations on the fly." *Nucleic Acids Res.* 32(Database issue):W327-331.

NCBI

Conserved Protein Domain Family

TNFRSF

c12855: TNFRSF, with user query added Superfamily

Tumor necrosis factor receptor superfamily (TNFRSF)

Members of TNFR superfamily (TNFRSF) interactions with TNF superfamily (TNFSF) ligands (TNFL) control key cellular processes such as differentiation, proliferation, apoptosis, and cell growth. Dysregulation of these pathways has been shown to result in a wide range of pathological conditions, including autoimmune diseases, inflammation, cancer, and viral infection. There are 29 very diverse family members of TNFRSF reported in humans: 22 are type I transmembrane receptors (single pass with the N terminus on extracellular side of the cell membrane) and have a clear signal peptide; the remaining 7 members are either type III transmembrane receptors (single pass with the N terminus on extracellular side of the membrane but no signal sequence; TNFR13B, TNFR13C, TNFR17, and XEDAR), or attached to the membrane via a glycosylphosphatidylinositol (GPI) linker (TNFR10C), or secreted as soluble receptors (TNFR11B and TNFR6B). All TNFRs contain relatively short cysteine-rich domains (CRDs) in the ectodomain, and are involved in interaction with the TNF homology domain (THD) of their ligands. TNFRs often have multiple CRDs (between one and six), with the most frequent configuration of three or four copies; most CRDs possess three disulfide bridges, but could have between one and four. Localized or genome-wide duplication and evolution of the TNFRSF members appear to have paralleled the emergence of the adaptive immune system; teleosts (i.e. ray-finned, bony fish), which possess an immune system with B and T cells, possess primary and secondary lymphoid organs, and are capable of adaptive responses to pathogens also display several characteristics that are different from the mammalian immune system, making relevant TNFRSF orthologs and paralogs of interest to better understand immune system evolution and the immunological pathways elicited to pathogens.

Links

Taxonomy: root

PubMed: 332 links

Protein: Related Protein

Related Structure

Statistics

Accessions: c12855

PSSM Id: 277570

Names: TNFRSF

Created: 23-Apr-2015

Superfamily

cd01070 TNFRSF

cd01075 TNFRSF6B

cd10576 TNFRSF18

cd10577 TNFRSF18

cd10578 TNFRSF3

cd10579 TNFRSF6

cd10580 TNFRSF10

cd10581 TNFRSF11B

cd12855 TNFRSF

cd10576 TNFRSF18

cd10577 TNFRSF18

cd10578 TNFRSF3

cd10579 TNFRSF6

cd10580 TNFRSF10

cd10581 TNFRSF11B

Klik link yang di tunjukkan oleh panah merah. Maka akan muncul tampilan sebagai berikut:

NCBI

Conserved Protein Domain Family

stn_TNFRSF12A

pfam12191: stn_TNFRSF12A

tumor necrosis factor receptor stn_TNFRSF12A_TNFR domain

This family of proteins is found in eukaryotes. Proteins in this family are typically between 129 and 184 amino acids in length. This is the stn_TNFRSF12A_TNFR domain from the tumor necrosis factor receptor. The function of this domain is unknown.

Links

Statistics

Structure

LinkOut - more resources

Sequence Alignment

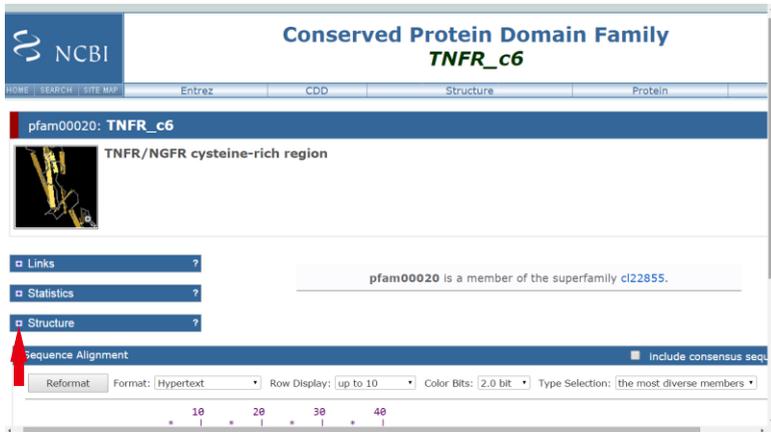
Reformat: Format: Hypertext Row Display: All 2 rows Color Bits: 2.0 bit Type Selection: top listed sequences

```

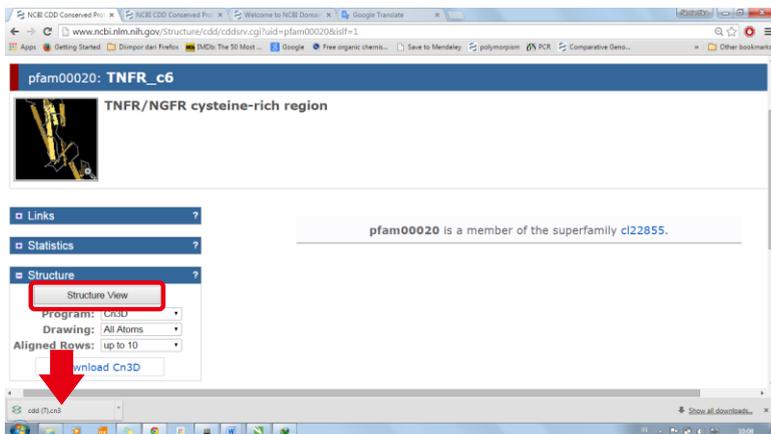
      10      20      30      40      50      60      70      80
g1 81873424 1 WAPGAPRPLQLLVISGFLVPLAragagQAPQAPMPCSSGASGLDNGKCCSCAPPHRFDICGASAPFAM-FRLW 79
g1 82088232 1 HTP---INRLITFVLLLVLLV...SSAASQGCPEGRAYSQGLGKRCSCVCKDKEKSDFCQNCPSKTPQNFPM 71
          90      100     110     120     130
g1 81873424 80 P11GALLSLALVLSVFLVLRRCRRRREFTTPEETGGCPQVAVIQ 129
g1 82088232 72 VTGFSAGGVFLVLSVFLVYVLRKRSKFTTPEETGSHSAEAL-LIH 120
  
```

Citing CCD

Marchler-Bauer A et al. (2015). "CCD: NCBI's conserved domain database." *Nucleic Acids Res.* 43(Database issue):D222-6.

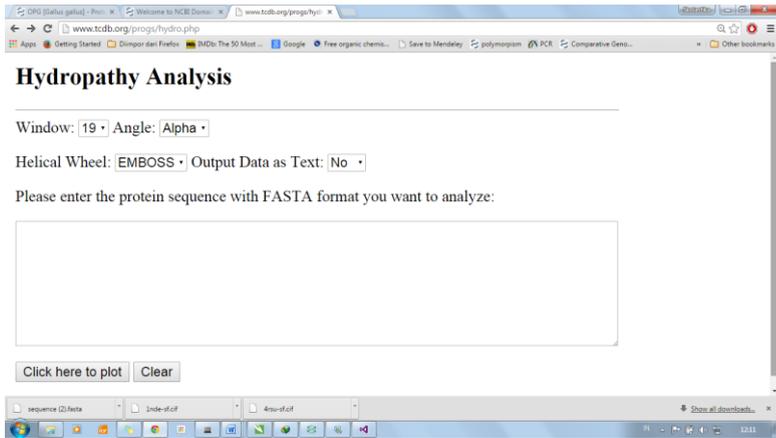


Klik tanda + yang di tunjukkan oleh panah yang terdapat di sisi kiri sebelah tulisan *Structure* lalu pilih *structure view*. Maka struktur protein akan terdownload. Buka file yang telah terdownload.

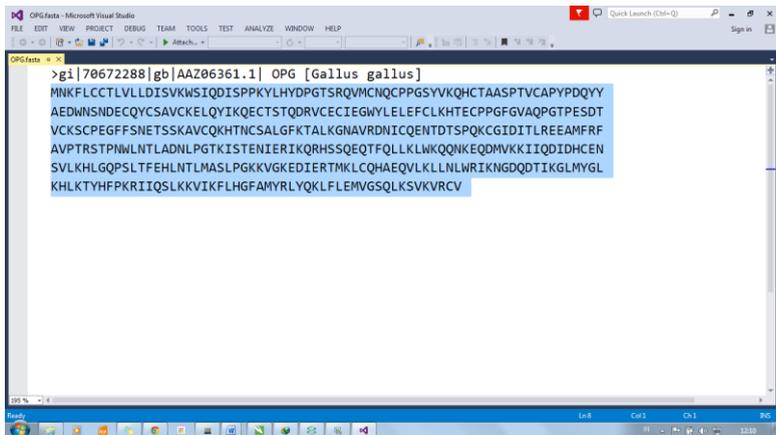


HYDROPATHY

1. Buka situs <http://www.tcdb.org/progs/hydro.php>



2. Masukkan FASTA Format sequence asam amino dari protein query, kemudian klik Click here to plot.



Hydropathy Analysis

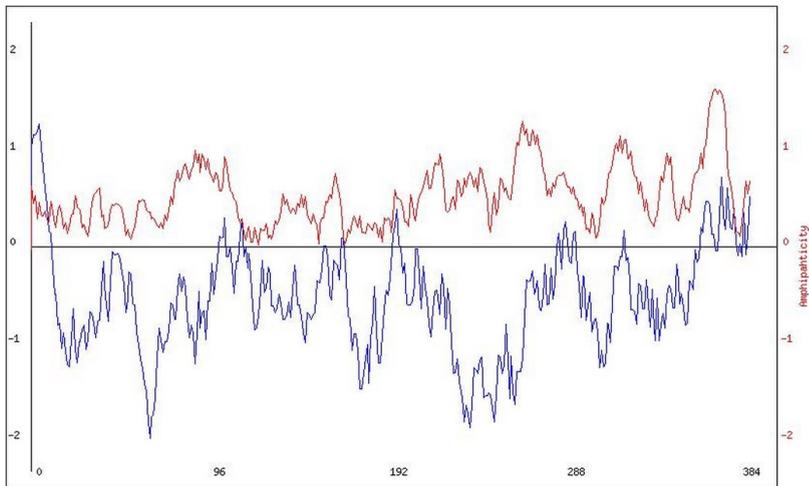
Window: 19 | Angle: Alpha

Helical Wheel: EMBOSS | Output Data as Text: No

Please enter the protein sequence with FASTA format you want to analyze:

```
MNKFCLCTLVLLDISVKHSIQDISPPKYLHYDPGTSRQVMCNQCPPGSYVKQHCTAASPTVCAPYDPQYY
AEDWNSNDECQYCSAVCKELQYIKQECTSTQDRVCECIEGWYLEFECLKHTECPPGGVAQPGTPESDT
VCKSCPFGFFSNETSSKAVCQKHTNCSALGFKTALKGNVARDNICQENTDTSQKCGIDITLREEAMFRF
AVPTRSTPNWNLTLADNLPGTKISTENIERIKQRHSSQEQTFLKLLWQKQKEQDMVKKIIQDIDHCEN
SVLKLHGQPSLTFEHLNLMASLPGKVKGKEDIERTMKLQHAQVLLKLLNLWRIKNGDQDTIKGLMVLG
KHLKTYHFKRIIQSLKKVVKLHGFMRYLYQKLFLEMVGSQKSKVKRVCV
```

[Click here to plot](#) [Clear](#)



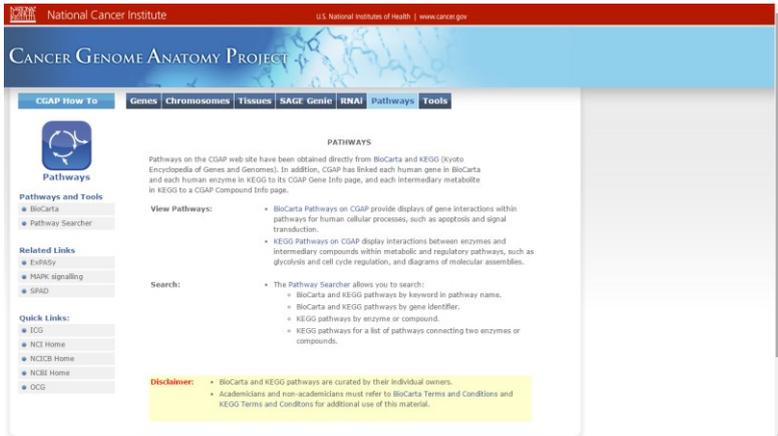
Blue Curve = Hydropathy
 Red Curve = Hydrophobic Moment (Amphipathicity)
 Blue Bars = putative transmembrane segments (TMSs)
 Putative TMSs can be clicked to plot a helical wheel diagram of the TMS

Interpretasi tampilan data yang diperoleh, mengenai sifat protein query, apakah hidrofob atau hidrofil dan bagaimana kemungkinan lokasi dari protein tersebut.

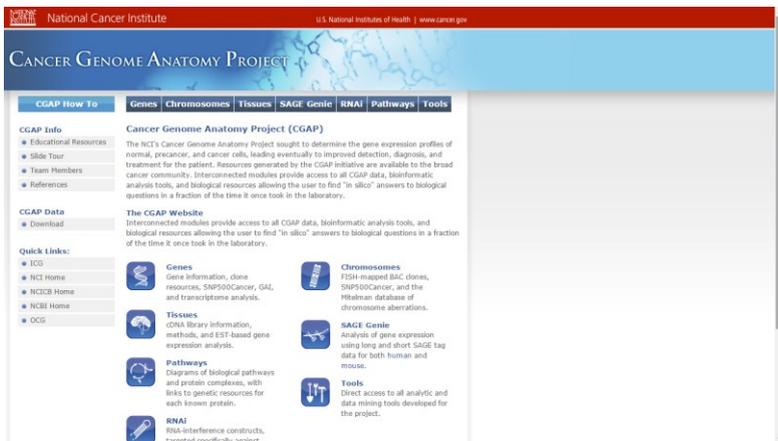
3. Analisis menggunakan situs lain seperti tertera pada petunjuk praktikum dapat dicoba sendiri.

TRANSDUKSI SINYAL

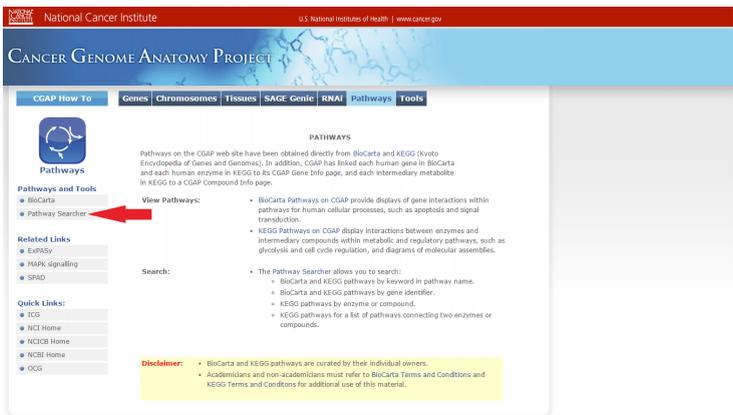
1. Buka situs <http://cgap.nci.nih.gov/>



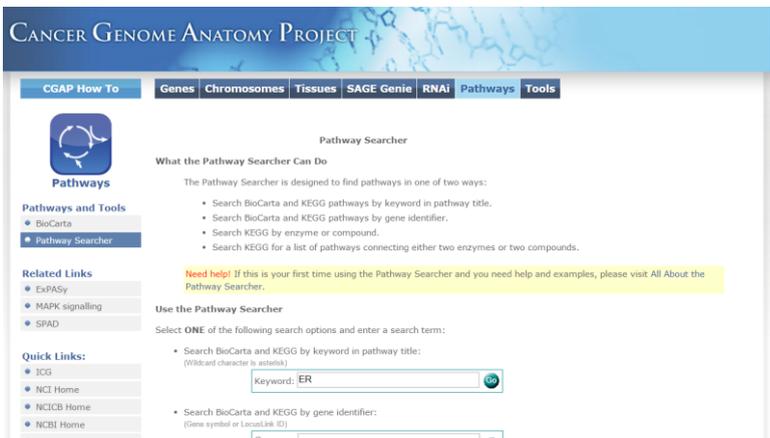
2. klik Pathways, maka muncul tampilan sebagai berikut:



3. Klik Pathway Searcher (di bagian kiri tampilan)



4. Masukkan nama protein yang akan dicari pathwaynya (protein query) pada kolom keyword dan klik Go



BAB V. ANALISIS PENYEJAJARAN NUKLEOTIDA DAN PROTEIN

A. SASARAN CAPAIAN

1. Mahasiswa mampu melakukan analisis penyejajaran urutan DNA atau protein menggunakan BLAST secara online.
2. Mahasiswa mampu memahami homologi, orthologi dan paralogi gen.
3. Mahasiswa mampu memahami polimorfisme DNA pada organisme tertentu dari sekuen nukleotide.

B. DASAR TEORI

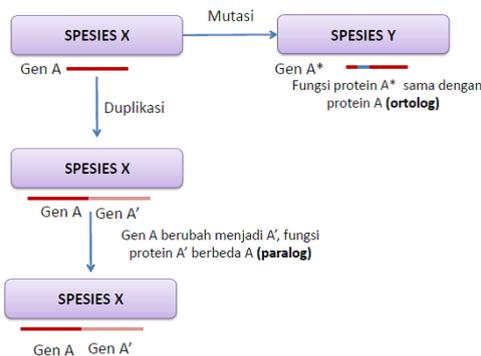
Peran biologis sebuah protein ditentukan oleh struktur protein tersebut, terutama konformasi 3 dimensi protein. Sekuen nukleotida atau protein yang mirip dapat diidentifikasi sebagai protein homolog. Homologi gen atau protein dapat didefinisikan sebagai 'kemiripan sekuen'; suatu sekuen dikatakan homolog apabila lebih dari 80% sekuen identik. Gen homolog yang ditemukan pada dua spesies yang berbeda disebut sebagai ortologi sedangkan gen homolog yang terduplikasi dalam satu spesies disebut sebagai paralogi. Spesies dengan gen ortholog kemungkinan berasal dari nenek moyang yang sama yang mengalami evolusi genetik. Gen paralog kemungkinan mempunyai struktur dan fungsi yang berbeda. Homologi suatu gen atau protein dapat ditentukan melalui analisis penyejajaran sekuen (*sequence alignment*).

Analisis penyejajaran adalah proses membandingkan urutan basa pada dua atau lebih nukleotida atau protein. Analisis bioinformatika untuk penyejajaran urutan nukleotida atau asam amino dengan menggunakan sistem sederhana secara online, yaitu BLAST. BLAST (*Basic Local Alignment Search Tool*) adalah sebuah software sederhana berdasarkan basis data untuk membandingkan informasi biologis primer

seperti basa nukleotida atau asam amino pada protein. Sebuah pencarian BLAST memungkinkan peneliti untuk membandingkan urutan QUERY dengan basis data dan mengidentifikasi urutan tersebut dengan basis data yang tersedia online. Dengan demikian, BLAST memungkinkan peneliti untuk mencari urutan nukleotida maupun protein yang mirip. Prinsip dasar yang mendasari kerja BLAST adalah sebagai berikut:



BLAST adalah program online yang paling banyak digunakan untuk membandingkan urutan nukleotida atau protein. Dengan menggunakan program *blastn* dan *blastp* yang terdapat pada situs NCBI ini kita dapat mencari protein homolog dari berbagai macam organisme. Analisis BLAST dalam NCBI juga dapat digunakan sebagai alat untuk mencari gen homolog baik ortolog maupun paralog. Gen ortolog merupakan gen-gen yang mirip serta menghasilkan protein yang berfungsi sama dalam spesies yang berbeda. Sedangkan, gen paralog merupakan gen-gen yang mirip dalam 1 spesies tetapi menghasilkan protein yang berbeda fungsi. Gen paralog dapat dijelaskan melalui proses penurunan gen dari nenek moyang sama sedangkan gen ortolog dapat dijelaskan melalui proses duplikasi gen (Gambar 5.1).



Gambar 5.1 Homologi Gen

Aplikasi BLAST dapat digunakan untuk mencari protein farmasetik untuk terapi suatu penyakit; misalnya insulin untuk terapi diabetes. Sebelum insulin dapat diproduksi dengan teknik rekayasa genetika, insulin diambil dari beberapa organisme lain yang mirip seperti babi ataupun sapi. Urutan nukleotida dan protein pada insulin babi dan sapi tidak sama persis dengan insulin manusia. Oleh karena itu, terapi diabetes militus pada manusia menggunakan insulin yang berasal dari spesies lain, dapat memicu reaksi alergi.

Selain untuk analisis homologi gen/protein, BLAST juga dapat digunakan untuk analisis polimorfisme. Polimorfisme adalah variasi dalam sekuen DNA tertentu pada individu yang berbeda. Polimorfisme disebabkan oleh mutasi genetik seperti delesi (penghilangan) dan insersi (penyisipan) juga bertanggung jawab atas variasi dalam urutan DNA. Salah satu akibat dari polimorfisme gen adalah perbedaan efek obat pada individu yang berbeda. Pada satu individu pada dosis tertentu sudah memberikan efek farmakologi sedangkan pada individu lain dibutuhkan dosis yang lebih rendah atau tinggi. Kejadian tersebut disebabkan oleh polimorfisme pada gen-gen yang menyandi enzim pemetabolisme obat.

C. PROSEDUR KERJA

Contoh soal

Nukleotida query

Sekuens nukleotida:

```
CTCGGTCTTT AAAAGGAAGA AGGGGCTTAT CGTTAAGTCG CTTGTGATCT
TTTCAGTTTC TCCAGCTGCT GGCTTTTTGG ACACCCACTC CCCC GCCAGG
AGGCAGTTGC AAGCGCGGAG GCTGCGAGAA ATA ACTGCCT CTTGAACTT
GCAGGGCGAA GAGCAGGCGG CGAGCGCTGG GCCGGGGAGG GACCACCCGA
```

GCTGCGACGG GCTCTGGGGC TGCGGGGCAG GGCTGGCGCC CGGAGCCTGA
GCTGCAGGAG GTGCGCTCGC TTTCTCAAC AGGTGGCGGC GGGGCGCGCG
CCGGGAGACC CCCCTAATG CGGGAAAAGC ACGTGTCCGC ATTTTAGAGA
AGGCAAGGCC GGTGTGTTTA TCTGCAAGCC ATTATACTTG CCCACGAATC
TTTGAGAACA TTATAATGAC CTTTGTGCCT CTTCTTGCAA GGTGTTTTCT
CAGCTGTTAT CTCAAGACAT GGATATAAAA AACTCACCAT CTAGCCTTAA
TTCTCCTTCC TCCTACAAC TCAAGTCAATC CATCTTACCC CTGGAGCACG
GCTCCATATA CATACCTTCC TCCTATGTAG ACAGCCACCA TGAATATCCA
GCCATGACAT TCTATAGCCC TGCTGTGATG AATTACAGCA TTCCAGCAA
TGTCATAAC TTGGAAGGTG GGCCTGGTCG GCAGACCACA AGCCCAAATG
TGTTGTGGCC AACACCTGGG CACCTTTCTC CTTTAGTGGT CCATCGCCAG
TTATCACATC TGTATGCGGA ACCTCAAAAAG AGTCCCTGGT GTGAAGCAAG
ATCGCTAGAA CACACCTTAC CTGTAAACAG AGAGACACTG AAAAGGAAGG
TTAGTGGGAA CCGTTGCGCC AGCCCTGTTA CTGGTCCAGG TTCAAAGAGG
GATGCTCACT TCTGCGCTGT CTGCAGCGAT TACGCATCGG GATATCACTA
TGGAGTCTGG TCGTGTGAAG GATGTAAGGC CTTTTTTAAA AGAAGCATTC
AAGGACATAA TGATTATATT TGTCCAGCTA CAAATCAGTG TACAATCGAT
AAAAACCGGC GCAAGAGCTG CCAGGCCTGC CCACTTCGGA AGTGTTACGA
AGTGGGAATG GTGAAGTGTG GCTCCCGGAG AGAGAGATGT GGGTACCGCC
TTGTGCGGAG ACAGAGAAGT GCCGACGAGC AGCTGCACTG TGCCGGCAAG
GCCAAGAGAA GTGGCGGCCA CGCGCCCCGA GTGCGGGAGC TGCTGCTGGA
CGCCCTGAGC CCGAGCAGC TAGTGCTCAC CCTCCTGGAG GCTGAGCCGC
CCCATGTGCT GATCAGCCGC CCCAGTGC GCCTTACCGA GGCCTCCATG
ATGATGTCCC TGACCAAGTT GGCCGACAAG GAGTTGGTAC ACATGATCAG
CTGGGCCAAG AAGATTCCCG GCTTTGTGGA GCTCAGCCTG TTCGACCAAG
TGCGGCTCTT GGAGAGCTGT TGGATGGAGG TGTTAATGAT GGGGCTGATG
TGGCGCTCAA TTGACCACCC CGGCAAGCTC ATCTTTGCTC CAGATCTTGT
TCTGGACAGG GATGAGGGGA AATGCGTAGA AGGAATTCTG GAAATCTTTG
ACATGCTCCT GGCAACTACT TCAAGGTTTC GAGAGTTAAA ACTCCAACAC
AAAGAATATC TCTGTGTCAA GGCCATGATC CTGCTCAATT CCAGTATGTA

CCCTCTGGTC ACAGCGACCC AGGATGCTGA CAGCAGCCGG AAGCTGGGCTC
ACTTGCTGAA CGCCGTGACC GATGCTTTGG TTTGGGTGAT TGCCAAGAGC
GGCATCTCCT CCCAGCAGCA ATCCATGCGC CTGGCTAACC TCCTGATGCT
CCTGTCCCAC GTCAGGCATG CGAGTAACAA GGGCATGGAA CATCTGCTCA
ACATGAAGTG CAAAAATGTG GTCCCAGTGT ATGACCTGCT GCTGGAGATG
CTGAATGCCC ACGTGCTTCG CGGGTGCAAG TCCTCCATCA CGGGGTCCGA
GTGCAGCCCG GCAGAGGACA GTAAAAGCAA AGAGGGCTCC CAGAACCCAC
AGTCTCAGTG ACGCCTGGCC CTGAGGTGAA CTGGCCACA GAGGTCACAG
GCTGAAGCGT GAACTCCAGT GTGTCAGGAG CCTGGGCTTC ATCTTTCTGC
TGTGTGGTCC CTCATTTGG

Asam amino query

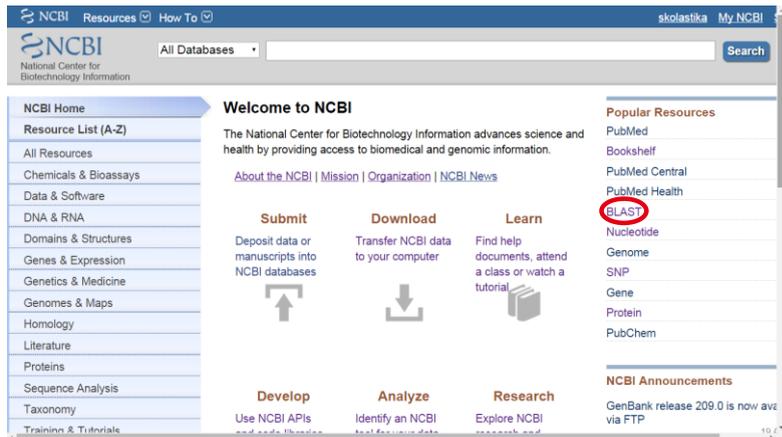
Sekuens asam amino

MDIKNSPSSL NSPSSYNCSQ SILPLEHCSI YIPSSYVDSH HEYPAMTFYS
PAVMNYSIPS NVTNLEGGPG RQTTSPNVLW PTPGHLSPV VHRQLSHLYA
EPQKSPWCEA RSLEHTLPVN RETLKRKVS G NRKASPVTPG GSKRDAHFCA
VCSYASGYH YGVWSCEGCK AFFKRSIQGH NDYICPATNQ CTIDKNRRKS
CQACRLRKY EVGMVKCGSR RERCGYRLVR RQRSADEQLH CAGKAKRSGG
HAPRVRELL DALSPQLVL TLLEAEPHV LISRPSAPFT EASMMMSLTK
LADKELVHMI SWAKKIPGFV ELSLFDQVRL LESCWMEVLM
MGLMWRSIDH PGKLIFAPDL VLDRDEGKCV EGILEIFDML LATTSRFREL
KLQHKEYLCV KAMILLNSSM YPLVTATQDA DSSRKLALL NAVTDALVWV
IAKSGISSQQ QSMRLANLLM LLSHVRHASN KGMEHLLNMK CKNVVPVYDL
LLEMLNAHVL RGCKSSITGS ECSPAEDSKS KEGSQNPQSQ

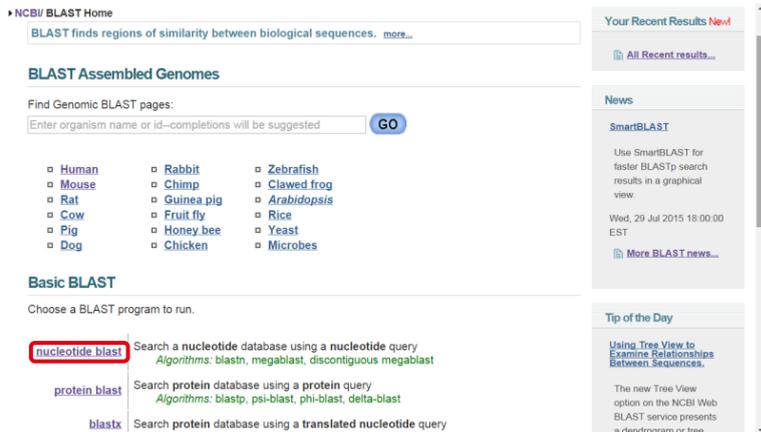
HOMOLOGI PROTEIN

Langkah Kerja dengan BLASTn

1. Buka situs www.ncbi.nlm.nih.gov



2. Pilih tool "BLAST" (klik satu kali), akan muncul tampilan pilihan program BLAST. Untuk mencari gen suatu sekuen nukleotida dari database nukleotida pilih "nucleotide blast" (blastn).



3. Setelah tampilan muncul, entri sekuen nukleotida (query) yang akan dicari; pilih setting pencarian dari database "others" (jika belum diketahui spesiesnya); pilih program "megablast"; klik "BLAST" untuk memulai proses *searching*.

NCBI BLAST® Basic Local Alignment Search Tool My NCBI Welcome skolastika. [Sign Out]

Home Recent Results Saved Strategies Help

NCBI/BLAST/blastn suite Standard Nucleotide BLAST

blastn blastp blastx tblastn tblastx

BLASTn programs search nucleotide databases using a nucleotide query. [more...](#) [Reset page](#) [Bookmark](#)

Enter Query Sequence

Enter accession number(s), gi(s), or FASTA sequence(s)

CTGGGCTTTT AAAAGGAGA AGGGGCTTAT CGTAAAGTCG CTGTGATCT TTTCAGTTTC
TCCAGCTCT GGGTTTTTGG ACACCCATC CCCCAGGG AGGGAGTGG AAGGCGGAG
GCTCGAGAA ATAACCTCT CTGAAACTT GCAGGGCGAA GAGCAGCCGG CAGCGCTGG
GCCGGGGAGG GACCAACCGA GCTCGACGG GCTCTGGGG TCGGGGCGAG GGCCTGGCC
CGAGCCTGA GCTCGAGGAG GTGGCTCGG TTCTCTCAG AGGTGGGGG GGGGCGCGG

Clear Query subrange

From To

Or, upload file Choose File No file chosen

Job Title

Enter a descriptive title for your BLAST search

Align two or more sequences

Choose Search Set

Database Human genomic + transcript Mouse genomic + transcript Others (nr etc.)
Nucleotide collection (nr/nt)

Organism Optional

Enter organism name or id—completions will be suggested Exclude +
Enter organism common name, binomial, or tax id. Only 20 top taxa will be shown

Organism Optional

Enter organism name or id—completions will be suggested Exclude +
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

Entrez Query Optional

Enter an Entrez query to limit search [YouTube](#) [Create custom database](#)

Program Selection

Optimize for

Highly similar sequences (megablast)

More dissimilar sequences (discontiguous megablast)

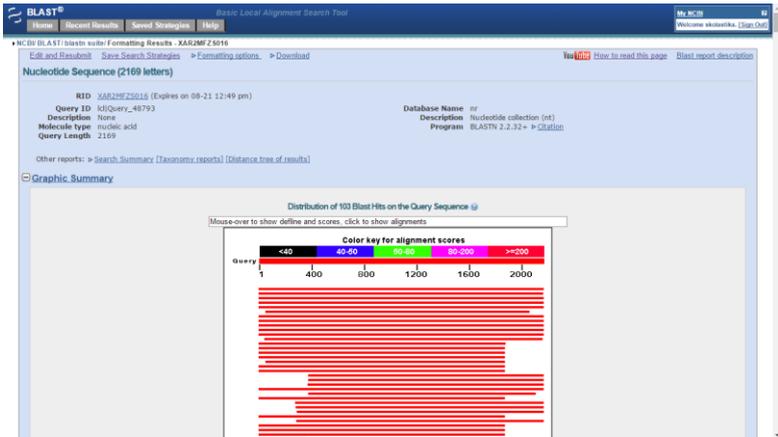
Somewhat similar sequences (blastn)

Choose a BLAST algorithm

BLAST Search database Nucleotide collection (nr/nt) using Megablast (Optimize for highly similar sequences) Show results in a new window

[Algorithm parameters](#)

4. Hasil **searching/pencarian** mendapatkan tampilan seperti berikut: Hasil blast umumnya akan menghasilkan lebih dari satu sekuen yang beresesuaian. Pilih hasil dengan skor paling tinggi.



Sequences producing significant alignments:

Select: All None Selected:0

	Description	Max score	Total score	Query cover	E value	Ident	Accession
<input checked="" type="checkbox"/>	Homo sapiens estrogen receptor 2 (ER beta) (ESR2), transcript variant a, mRNA	4006	4006	100%	0.0	100%	gi 94538323 NM_001437.2
<input type="checkbox"/>	PREDICTED: Pan paniscus estrogen receptor 2 (ER beta) (ESR2), transcript va	3916	3916	99%	0.0	99%	gi 675797822 XM_003831616.2
<input type="checkbox"/>	PREDICTED: Pan troglodytes estrogen receptor 2 (ER beta) (ESR2), transcript	3901	3901	100%	0.0	99%	gi 694956472 XM_001170310.2
<input type="checkbox"/>	PREDICTED: Gorilla gorilla gorilla estrogen receptor 2 (ER beta), transcript vari	3853	3853	100%	0.0	99%	gi 426377136 XM_004055283.1
<input type="checkbox"/>	PREDICTED: Nomascus leucogenys estrogen receptor 2 (ER beta) (ESR2), tra	3720	3720	100%	0.0	98%	gi 332237300 XM_003267794.1
<input type="checkbox"/>	Homo sapiens estrogen receptor beta mRNA, complete cds	3714	3714	92%	0.0	100%	gi 2970563 AF051427.1
<input type="checkbox"/>	PREDICTED: Pongo abelli estrogen receptor 2 (ER beta) (ESR2), transcript var	3690	3690	100%	0.0	97%	gi 395745978 XM_002824936.2
<input type="checkbox"/>	PREDICTED: Rhinopithecus rooseana estrogen receptor 2 (ER beta) (ESR2), tr	3618	3618	100%	0.0	97%	gi 724924188 XM_010383688.1
<input type="checkbox"/>	PREDICTED: Mandillus leucophaeus estrogen receptor 2 (ER beta) (ESR2), tr	3613	3613	100%	0.0	97%	gi 796127609 XM_011978290.1
<input type="checkbox"/>	PREDICTED: Colobus angolensis palliatus estrogen receptor 2 (ER beta) (ESR	3596	3596	100%	0.0	97%	gi 796314663 XM_011961967.1
<input type="checkbox"/>	PREDICTED: Chlorocebus sabaeus estrogen receptor 2 (ER beta) (ESR2), tran	3585	3585	100%	0.0	97%	gi 636127444 XM_007986933.1
<input type="checkbox"/>	Macaca mulatta estrogen receptor 2 (ER beta) (ESR2), mRNA	3530	3530	97%	0.0	97%	gi 402766370 NM_001265821.2
<input type="checkbox"/>	PREDICTED: Homo sapiens estrogen receptor 2 (ER beta) (ESR2), transcript v	3461	3461	86%	0.0	99%	gi 767979986 XM_011536945.1

- Klik "Accession" gen terpilih (hasil blastn) untuk keterangan lebih lanjut, (tanda panah menunjukkan bagian yang berisi informasi penting).

Sequences producing significant alignments:

Select: All None Selected:0

Alignments Download GenBank Graphics Distance tree of results

Description	Max score	Total score	Query cover	E value	Ident	Accession
<input type="checkbox"/> Homo sapiens estrogen receptor 2 (ER beta) (ESR2), transcript variant a, mRNA	4006	4006	100%	0.0	100%	gi 94538323 NM_001437.2
<input type="checkbox"/> PREDICTED: Pan paniscus estrogen receptor 2 (ER beta) (ESR2), transcript va	3916	3916	99%	0.0	99%	gi 675797822 XM_003831616.2
<input type="checkbox"/> PREDICTED: Pan troglodytes estrogen receptor 2 (ER beta) (ESR2), transcript	3901	3901	100%	0.0	99%	gi 684956472 XM_001170310.2
<input type="checkbox"/> PREDICTED: Gorilla gorilla gorilla estrogen receptor 2 (ER beta), transcript vari	3853	3853	100%	0.0	99%	gi 426377136 XM_004055283.1
<input type="checkbox"/> PREDICTED: Nomascus leucogenys estrogen receptor 2 (ER beta) (ESR2), tra	3720	3720	100%	0.0	98%	gi 332237300 XM_003267794.1
<input type="checkbox"/> Homo sapiens estrogen receptor beta mRNA, complete cds	3714	3714	92%	0.0	100%	gi 2970563 AF051427.1
<input type="checkbox"/> PREDICTED: Pongo abelii estrogen receptor 2 (ER beta) (ESR2), transcript var	3690	3690	100%	0.0	97%	gi 395745978 XM_002824836.2
<input type="checkbox"/> PREDICTED: Rhinopithecus roxellana estrogen receptor 2 (ER beta) (ESR2), tr	3618	3618	100%	0.0	97%	gi 724924188 XM_010383688.1
<input type="checkbox"/> PREDICTED: Mandillus leucophaeus estrogen receptor 2 (ER beta) (ESR2), tr	3613	3613	100%	0.0	97%	gi 795127609 XM_011976290.1
<input type="checkbox"/> PREDICTED: Colobus angolensis palliatus estrogen receptor 2 (ER beta) (ESR	3596	3596	100%	0.0	97%	gi 795314663 XM_011961967.1
<input type="checkbox"/> PREDICTED: Chlorocebus sabaeus estrogen receptor 2 (ER beta) (ESR2), tran	3585	3585	100%	0.0	97%	gi 635127444 XM_007986933.1
<input type="checkbox"/> Macaca mulatta estrogen receptor 2 (ER beta) (ESR2), mRNA	3530	3530	97%	0.0	97%	gi 402766370 NM_01265821.2
<input type="checkbox"/> PREDICTED: Homo sapiens estrogen receptor 2 (ER beta) (ESR2), transcript v	3461	3461	86%	0.0	99%	gi 767979986 XM_011536545.1

NCBI Resources How To skolastika

Nucleotide Nucleotide Advanced

Display Settings: GenBank Send: Change region sh

Homo sapiens estrogen receptor 2 (ER beta) (ESR2), transcript variant a, mRNA

NCBI Reference Sequence: NM_001437.2

FASTA Graphics

Go to: Custom view

LOCUS NM_001437 2169 bp mRNA linear PRI 15-MAR-2015
DEFINITION Homo sapiens estrogen receptor 2 (ER beta) (ESR2), transcript variant a, mRNA.
ACCESSION NM_001437 XM_495993
VERSION NM_001437.2 GI:94538323
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 2169)

Analyze this sequ
Run BLAST
Pick Primers
Highlight Sequence f
Find in this Sequen

Articles about the
Loss of heterozygosit
hormonal receptor (C
RSRC1 SUMOylator
SUMOylation and inh
ERB-dependent neur

```

2062-2169          AL161756.6          143227-143334
FEATURES
source             Location/Qualifiers
                  1..2169
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /chromosome="14"
                  /map="14q23.2"
gene              1..2169
                  /gene="ESR2"
                  /gene_synonym="ER-BETA; Erb; ESR-BETA; ESRB; ESTRB; NR3A2"
                  /note="estrogen receptor 2 (ER beta)"
                  /db_xref="GeneID:2100"
                  /db_xref="HGNC:HGNC:3468"
                  /db_xref="HPRD:03390"
                  /db_xref="MIM:601663"
exon              1..278
                  /gene="ESR2"
                  /gene_synonym="ER-BETA; Erb; ESR-BETA; ESRB; ESTRB; NR3A2"
                  /inference="alignment:SpIign:1.39.8"
misc_feature      1
                  /gene="ESR2"
                  /gene_synonym="ER-BETA; Erb; ESR-BETA; ESRB; ESTRB; NR3A2"
                  /note="5'-most transcription initiation site"
misc_feature      64
                  /gene="ESR2"
                  /gene_synonym="ER-BETA; Erb; ESR-BETA; ESRB; ESTRB; NR3A2"
                  /note="alternative transcription initiation site"
misc_feature      287
                  /gene="ESR2"

```

```

(HTML: 20920096)..
CDS
469..2061
/gene="ESR2"
/gene_synonym="ER-BETA; Erb; ESR-BETA; ESRB; ESTRB; NR3A2"
/note="isoform 1 is encoded by transcript variant a;
estrogen receptor beta 4; nuclear receptor subfamily 3
group A member 2"
/codon_start=1
/product="estrogen receptor beta isoform 1"
/protein_id="NP_001428.1"
/db_xref="GI:10835013"
/db_xref="CCDS:CCDS9762.1"
/db_xref="GeneID:2100"
/db_xref="HGNC:HGNC:3468"
/db_xref="HPRD:03390"
/db_xref="MIM:601663"
/translation="MDIKNSPSSLNSPSSYNCQSILPLEHGSYIPSSYVDSHHEYP
AMTFYSPAMMIVSIPSNVTILEGGPGRQTTSPNVLWPTPGHLSPLVVHRLSHLYAEP
QKSPWCEARSLEHTLPVWRETLKRVSGNRCASPVTPGPGSKRDAHFCAVCSDYASGYH
YGVWVSCGCKAFFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLRKCVEGVMKCG
SRRRCYGRVLRVRRQSADEQLHCAGKAKRSGGHAPRVRELLDALSPQVLVTLLEAE
PPHVLISRPSAPFTEASMMMSLTKLADKELVHMSIAKWIIPGFVLSLFDQVRLLESC
WHEVLLMGLMWRSIDHPGKLI FAPDLVLRDEGKCVGEILEIFDMLLATTSRFRELKL
QHKELYLVKAMILLNSHMYPLVATQDADSSRKLALHLNAVTDALVWVIKSGISSQQ
QSMRLANLLMLLSHVRHASKGMEHLNPKCKNVVYVDLLEMLNAHVLRGCKSSIT
GSECSPAEDSKSKEGSQNPQSQ"

```

Asam Amino atau Sekuen Protein

```

ORIGIN
1 ctggctttt aaaggaaag agggcttat cgtaaagtcg cttgtatct tttcagttc
61 tccactgctt ggcttttgg acaccactc cccgcagg aggcagtgc aagcgggag
121 gctcgagaa ataaactgct ctgaaactt gcaggcgaa ggcaggctg cagcgtctg
181 gcggggagg gaccaccga gctgcagcg gctcggggc tgcgggcaag gctcgtctc
241 cggagcctga gctcgagag gtcgctcgc ttctctcaac aggtgcggc gggcgcctc
301 ccggagacc cccctaatg cgggaaaag acgtgtccg attttagaa gggcagggc
361 ggtgtgttta tctgaacc attatactt cccaagaatc tttgaaaca tataatgac
421 ctttggcctt cttctgcaa ggtgttttct cagctgttat ctcaagactt gatatataaa
481 aactaccact ctatccttaa ttctctctc tctcaaacct gcagctcaatc catcttacc
541 cttggagcag gctccatata cataccttcc tctatgtag acagccacca tgaatacca
601 gccatgacat tctatgcc tctgtgtgat aattacagca tcccagcaa tgcactaac
661 ttggaaagtg ggcctggtc gcagaccaca agcccaaatg tgttgtgccc aacactggg
721 cacttttctc ctttagtgtt ccatcgccag ttatcacatc tgttgtggca acctcaaaag
781 agtccctggt gtgaagcaag atcgtagaa cacaccttac ctgtaaacag agagacactg
841 aaaaggaag tttagtgaaa ccgtgtgcc agccctgtta ctggtccaag ttcaaaagag
901 gatgctcact ctcgctcgt ctcgagcatc tacgcatcgg gatataccta tggagtctgg
961 tctgttgaag gatgtaagcc cttttttaa agaagcattc aaggacataa tgattatatt
1021 tgtccagcta caaatcagtg tacaatcgt aaaaaccgg ccaagagctg ccagcctgc
1081 cgacttcgga agtgttaca agtgggaagt gtagaagtgt gctcccggag agagagatgt
1141 gggtaaccgc ttgtcggag acagagaagt gcgacgagc agctgcactg tgcggcgaag
1201 gccaaagaaa gtggcgcca cgcgcccca gtcggggagc tgcctgtgga gccctggagc
1261 cccagcagc tagtctcac cctcctgga gctgacgac cccatgtctc gatcagccgc
1321 cccagtcgct ccttaccga ggcctccatc atgatgtccc tgaccaaagt ggcagacaag
1381 gagttggtac acatgatcag ctgggccaa aagatcccc gcttttggga gctcagcctg
1441 ttcgaccaag tgcgctctt gtagagctgt tggatggagg tttaaatgat gggcgtgat
1501 tggcctcaaa ttgaccacc cggcaagctc atcttcttc cagatctgtt tctggacagg
1561 gatgggggaa aatgctgaga aggaattctg gaaatcttgg acatgctctt ggcactact
1621 tcaaggtttc gagagttaa actccaacc aaagaatctc tctgtgtcaa ggcctatgtc
1681 ctcgtcaagt ccagatgta cctctagtc acagcagccc aggatcctga cagcagcgg
1741 aacttgctc actgtgtgta cgcctgagc gatgcttgg ttgggtgat tgcagagac
1801 ggcactctcc cccagcaga atccatcggc ctggctaac cctctgatc cctgtccca
1861 gtcaggcatg cagatacaa ggcctgtaa catctgcta acatgaagt caaaaatgtg
1921 tctccagttg atgacctgct gctggagatc ctgaaagccc agctgcttc ggggtgcaag
1981 tctccatca cgggtctga gtcagccc gcagagaca gtaaaagcaa agagggctcc
2041 cagaaccac agtctcaatc agcctggcc ctgagtgaa ctggccaca gaggtcaagc
2101 gctgaagct gaactcagt gtgtcaggag cctgggctc atctttctc tgtgtggctc
2161 ctcatctgg
//

```

tga, Stop Kodon

6. Pada *related information* klik *gene*, didapat deskripsi lebih lanjut mengenai gen tersebut.

polymorphisms with human idiopathic thin endometrium
Genet. Mol. Res. 12 (4), 5978-5985 (2013)
24338391

GeneRIF: A potential role of the ER-beta gene polymorphisms in the etiology of idiopathic atrophic endometrium.
Publication Status: Online-Only
6 (bases 1 to 2169)
Smith L, Brannan RA, Hanby AM, Shaaban AM, Verghese ET, Peter MB, Pollock S, Sathesha S, Szykiewicz M, Speirs V and Hughes TA. Differential regulation of oestrogen receptor beta isoforms by 5' untranslated regions in cancer
J. Cell. Mol. Med. 14 (8), 2172-2184 (2010)
28928096

GeneRIF: Data demonstrate that two alternative ERbeta 5'UTRs have potent and differential influences on expression acting at the level of translation.
7 (bases 1 to 2169)
Chen H, Lin RJ, Schiltz RL, Chakravarti D, Nash A, Nagy L, Privalsky ML, Nakatani Y and Evans RM. Nuclear receptor coactivator ACTR is a novel histone acetyltransferase and forms a multimeric activation complex with P/CAF and CBP/p300
Cell 90 (3), 569-580 (1997)
9267836

8 (bases 1 to 2169)
Mosselman S, Polman J and Dijkema R. ER beta: identification and characterization of a novel human estrogen receptor
FEBS Lett. 392 (1), 49-53 (1996)
8769313

Related information

- Annotated Genomic
- BioSystems
- CCDS
- Components (Core)
- Full text in PMC
- Gene**
- GeneView in dbSNP
- HomoloGene
- Map Viewer
- OMIM
- Probe
- Protein
- PubMed
- PubMed (RefSeq)
- PubMed (Weighted)
- SNP
- Taxonomy
- UniGene
- GEO Profiles

7. Jika ingin melihat lokasi gen tersebut pada kromosom, klik “map viewer”.

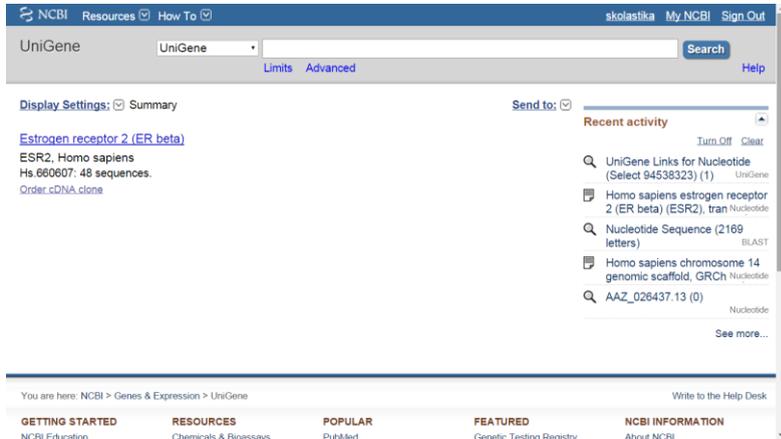
polymorphisms with human idiopathic thin endometrium
Genet. Mol. Res. 12 (4), 5978-5985 (2013)
24338391
GeneRIF: A potential role of the ER-beta gene polymorphisms in the etiology of idiopathic atrophic endometrium.
Publication Status: Online-Only
6 (bases 1 to 2169)
Smith L, Brannan RA, Hanby AM, Shaaban AM, Verghese ET, Peter MB, Pollock S, Satheesha S, Szykiewicz M, Speirs V and Hughes TA. Differential regulation of oestrogen receptor beta isoforms by 5' untranslated regions in cancer
J. Cell. Mol. Med. 14 (8), 2172-2184 (2010)
28928096
GeneRIF: Data demonstrate that two alternative ERbeta 5'UTRs have potent and differential influences on expression acting at the level of translation.
7 (bases 1 to 2169)
Chen H, Lin RJ, Schiltz RL, Chakravarti D, Nash A, Nagy L, Privalsky ML, Nakatani Y and Evans RM. Nuclear receptor coactivator ACTR is a novel histone acetyltransferase and forms a multimeric activation complex with P/CAF and CBP/p300
Cell 98 (3), 569-588 (1997)
9267036
8 (bases 1 to 2169)
Mosselman S, Polman J and Dijkema R. ER beta: identification and characterization of a novel human estrogen receptor
FEBS Lett. 392 (1), 49-53 (1996)
8769313

Related information
Annotated Genomic
BioSystems
CCDS
Components (Core)
Full text in PMC
Gene
GeneView in dbSNP
HomoloGene
Map Viewer
OMIM
Probe
Protein
PubMed
PubMed (RefSeq)
PubMed (Weighted)
SNP
Taxonomy
UniGene
GEO Profiles

8. Klik Uni gene pada gen terpilih tersebut, akan didapat homologi protein dari gen terpilih. Pilih spesies yang dikehendaki, misalnya human, klik untuk keterangan lebih lanjut. Selanjutnya dari masing-masing protein yang homolog dengan protein query bisa di klik lebih lanjut untuk mengetahui sekuens gen dan sekuens asam aminonya.

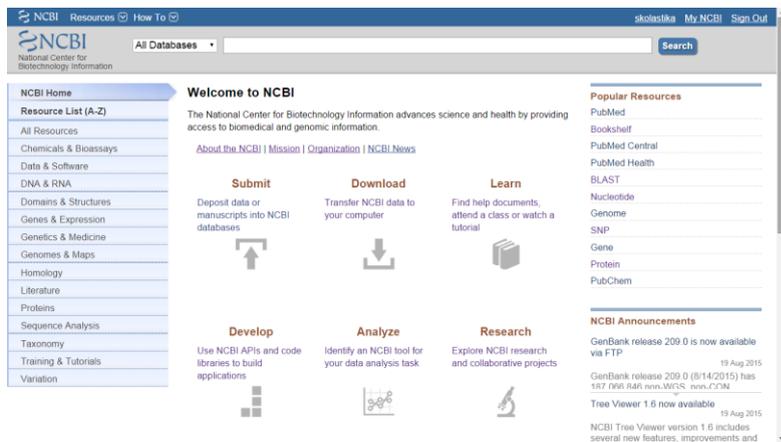
polymorphisms with human idiopathic thin endometrium
Genet. Mol. Res. 12 (4), 5978-5985 (2013)
24338391
GeneRIF: A potential role of the ER-beta gene polymorphisms in the etiology of idiopathic atrophic endometrium.
Publication Status: Online-Only
6 (bases 1 to 2169)
Smith L, Brannan RA, Hanby AM, Shaaban AM, Verghese ET, Peter MB, Pollock S, Satheesha S, Szykiewicz M, Speirs V and Hughes TA. Differential regulation of oestrogen receptor beta isoforms by 5' untranslated regions in cancer
J. Cell. Mol. Med. 14 (8), 2172-2184 (2010)
28928096
GeneRIF: Data demonstrate that two alternative ERbeta 5'UTRs have potent and differential influences on expression acting at the level of translation.
7 (bases 1 to 2169)
Chen H, Lin RJ, Schiltz RL, Chakravarti D, Nash A, Nagy L, Privalsky ML, Nakatani Y and Evans RM. Nuclear receptor coactivator ACTR is a novel histone acetyltransferase and forms a multimeric activation complex with P/CAF and CBP/p300
Cell 98 (3), 569-588 (1997)
9267036
8 (bases 1 to 2169)
Mosselman S, Polman J and Dijkema R. ER beta: identification and characterization of a novel human estrogen receptor
FEBS Lett. 392 (1), 49-53 (1996)
8769313

Related information
Annotated Genomic
BioSystems
CCDS
Components (Core)
Full text in PMC
Gene
GeneView in dbSNP
HomoloGene
Map Viewer
OMIM
Probe
Protein
PubMed
PubMed (RefSeq)
PubMed (Weighted)
SNP
Taxonomy
UniGene
GEO Profiles



Langkah Kerja dengan BLASTp

1. Buka situs www.ncbi.nlm.nih.gov



2. Pilih tool "BLAST" (klik satu kali). Untuk mencari protein homolog dari query asam amino gunakan "protein blast" (*blastp*).

NCBI/ BLAST Home

BLAST finds regions of similarity between biological sequences. [more...](#)

BLAST Assembled Genomes

Find Genomic BLAST pages:

Enter organism name or id--completions will be suggested

<input type="checkbox"/> Human	<input type="checkbox"/> Rabbit	<input type="checkbox"/> Zebrafish
<input type="checkbox"/> Mouse	<input type="checkbox"/> Chimp	<input type="checkbox"/> Clawed frog
<input type="checkbox"/> Rat	<input type="checkbox"/> Guinea pig	<input type="checkbox"/> Arabidopsis
<input type="checkbox"/> Cow	<input type="checkbox"/> Fruit fly	<input type="checkbox"/> Rice
<input type="checkbox"/> Pig	<input type="checkbox"/> Honey bee	<input type="checkbox"/> Yeast
<input type="checkbox"/> Dog	<input type="checkbox"/> Chicken	<input type="checkbox"/> Microbes

Basic BLAST

Choose a BLAST program to run.

nucleotide blast Search a nucleotide database using a nucleotide query
Algorithms: blastn, megablast, discontinuous megablast

protein blast Search protein database using a protein query
Algorithms: blastp, psi-blast, phi-blast, delta-blast

blastx Search protein database using a translated nucleotide query

Your Recent Results [New!](#)

[All Recent results...](#)

News

SmartBLAST

Use SmartBLAST for faster BLASTp search results in a graphical view.

Wed, 29 Jul 2015 16:00:00 EST

[More BLAST news...](#)

Tip of the Day

[Using Tree View to Examine Relationships Between Sequences.](#)

The new Tree View option on the NCBI Web BLAST service presents a dendrogram or tree

- Setelah tampilan muncul, entri sekuen protein (*query*) yang akan dicari; pilih seting pencarian dari database (jika membatasi hanya ingin mencari pada spesies tertentu, ketik nama organisme); pilih program "blastp"; klik "BLAST" untuk memulai proses *searching*.

BLAST® Basic Local Alignment Search Tool

Home Recent Results Saved Strategies Help

My NCBI Welcome skolastika. [Sign Out]

NCBI/ BLAST/ blastp suite Standard Protein BLAST

blastn blastp blastx tblastn tblastx

BLASTP programs search protein databases using a protein query. [more...](#) [Reset page](#) [Bookmark](#)

Enter Query Sequence

Enter accession number(s), gi(s), or FASTA sequence(s)

CAGKAKRSGG HAPRVRELLL DALSPQQLV LLEEAPPHV LISRPAFT EASPPSLTK
LADKLVMT SWAKDQFV ELSLFDQVL EESQRELVH NGLNRSIDN PGLLTPAPL
VLORDEGCV EGLEIFDHL LATTSRFREL KLQKVEYLV KANILLRSH YPLVTAQDA
DSSRLAHL NAVTALVW IAKSGISSQQ QSRRLANLL LLSHVPHASH KGNHLLHWK
CRNVPVYDL LLELNWHL RICKSSITGS ECFPAEDGSK REGSQWPOSQ

Or, upload file No file chosen

Job Title

Enter a descriptive title for your BLAST search

Align two or more sequences

Choose Search Set

Database

Organism Exclude

Enter organism name or id--completions will be suggested

Enter organism common name, binomial, or tax id. Only 20 top taxa will be shown.

Database: Protein Data Bank proteins(pdb)

Organism: Optional

Exclude: Optional

Entrez Query: Optional

Program Selection

Algorithm:

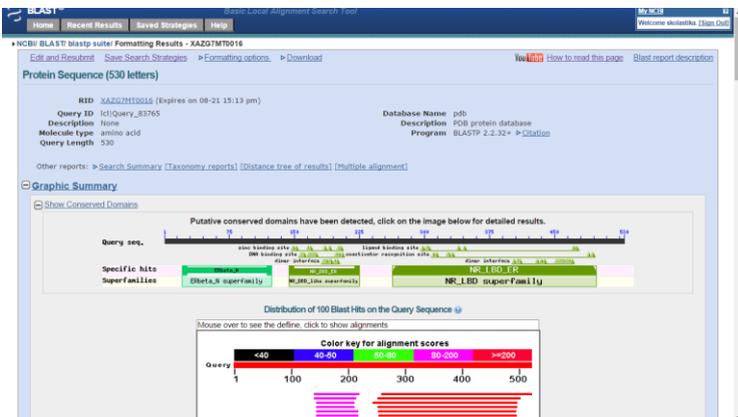
- * blastp (protein-protein BLAST)
- PSI-BLAST (Position-Specific Iterated BLAST)
- PHI-BLAST (Pattern Hit Initiated BLAST)
- DELTA-BLAST (Domain Enhanced Lookup Time Accelerated BLAST)

BLAST Search database Protein Data Bank proteins(pdb) using Blastp (protein-protein BLAST)

Algorithm parameters

Note: Parameter values that differ from the default are highlighted in yellow and marked with a sign

4. Hasil *searching* akan didapat tampilan seperti berikut:

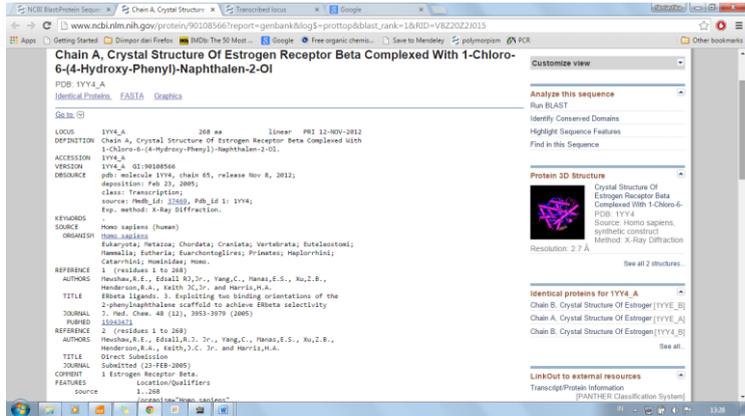


Sequences producing significant alignments:

Select: All None Selected 0

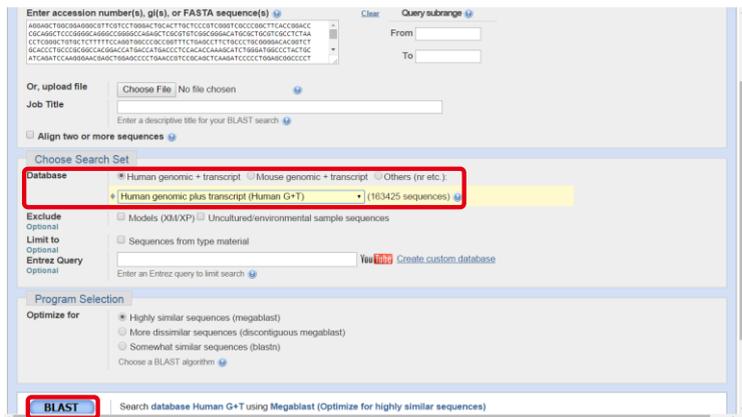
Description	Max score	Total score	Query cover	E value	Ident	Accession
Chain A, Crystal Structure Of Estrogen Receptor Beta Complexed With 1-Chloro-6-(4-Hydroxy-Phenyl)-1,3,5-Triaza-7,9-Diazepine	545	545	50%	0.0	100%	1YY4_A
Chain A, Human Estrogen Receptor Beta Ligand-Binding Domain In Complex With Partial Agonist	516	516	48%	0.0	100%	1QK0_A
Chain A, Human Estrogen Receptor Beta Ligand-Binding Domain In Complex With (R, R)-5,11-Cis-6,11-Dimethyl-2,3-Dihydro-1H-Indole-5-Carboxamide	508	508	48%	1e-178	97%	1L2J_A
Chain A, Benzopyrans Are Selective Estrogen Receptor Beta Agonists (Serbas) With Novel Activity	505	505	47%	7e-178	100%	2IG9_A
Chain A, Estrogen Receptor Beta With Selective Triazine Modulator	497	497	46%	1e-174	100%	1NDE_A
Chain A, A Second Binding Site For Hydroxytamoxifen Within The Coactivator-Binding Groove Of E	492	492	46%	8e-173	99%	2ES2_A
Chain A, Crystal Structure Of Estrogen Receptor Beta Complexed With Way-397	489	489	45%	8e-172	100%	1U85_A
Chain A, Human Estrogen Receptor Beta Ligand-Binding Domain In Complex With Compound 45	488	488	45%	2e-171	100%	2GIU_A
Chain A, Crystal Structure Of Estrogen Receptor Beta Complexed With Way-597	487	487	45%	6e-171	100%	1X78_A
Chain A, Crystal Structure Of Phosphorylated Estrogen Receptor Beta Ligand Binding Domain	484	484	45%	1e-169	99%	3OLL_A
Chain A, Crystal Structure Of Estrogen Receptor Beta Complexed With Way-555	483	483	44%	2e-169	100%	2NV7_A
Chain A, Crystal Structure Of Estrogen Receptor Beta Complexed With 3-Bromo-6-Hydroxy-2-(4-Hydroxyphenyl)-1H-Indole	482	482	44%	5e-169	99%	1ZAF_A
Chain A, Sulfonamides As Selective Estrogen Receptor Beta Agonists	482	482	45%	7e-169	99%	2YLY_A
Chain A, Structure Of Estradiol-bound Estrogen Receptor Beta Lbd In Complex With Lx01 Motif ER	477	477	48%	1e-166	92%	2J7X_A

- klik pada *accession number* yang memiliki skor paling tinggi maka akan muncul tampilan sbb:

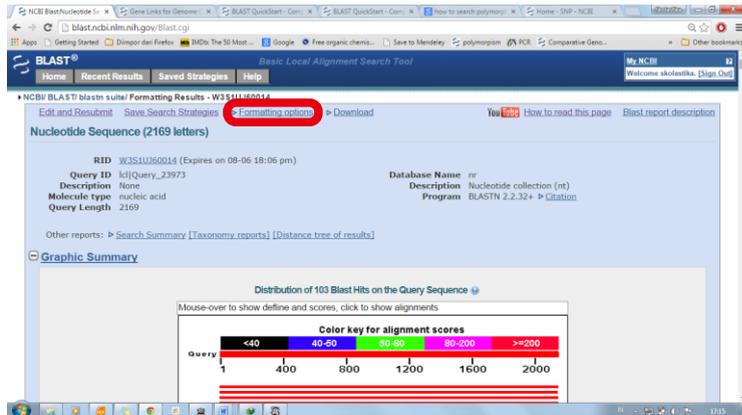


POLIMORFISME

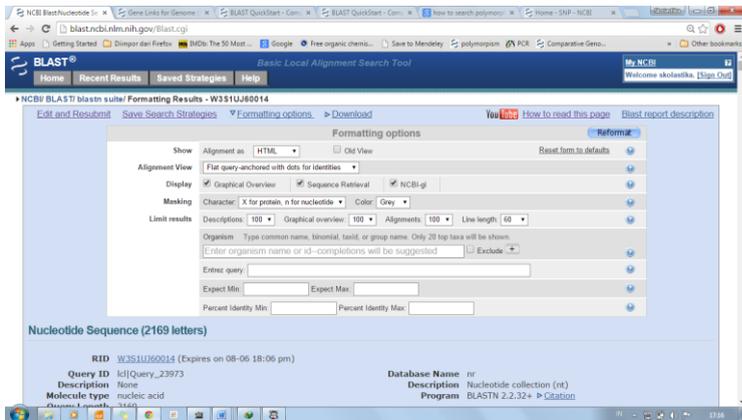
- Pada jendela BLAST pastikan *database* yang digunakan berasal dari satu organisme. Seperti pada gambar dibawah menggunakan *database human genomic+transcript*, sehingga hasil Blast yang muncul berasal dari organisme *homo sapiens*.



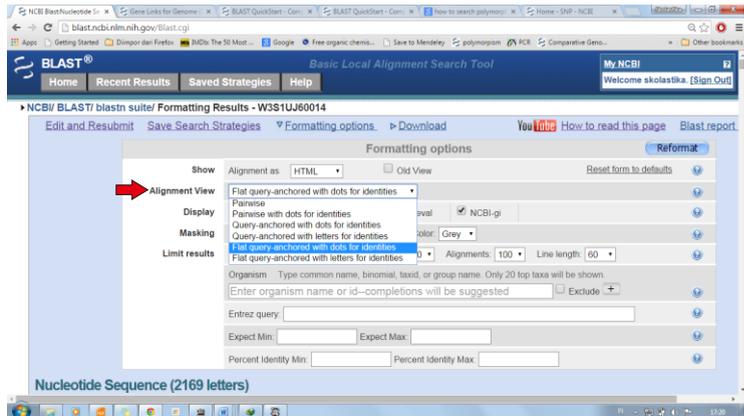
2. Pada tampilan hasil BLAST, dibagian atas terdapat tools *formatting options*. Klik tools tersebut.



3. Akan muncul tampilan seperti berikut.



4. Pada bagian *Alignment view* pilih *Flat Query-anchored with dot for identities*. Klik *Reformat* yang terletak disebelah kanan atas.



5. Scroll kebawah pada bagian *alignment*, akan terlihat polimorfisme seperti gambar berikut:



D. PERTANYAAN DISKUSI

1. Apakah yang dimaksud dengan homologi protein?
2. Apakah perbedaan homologi, ortologi dan paralogi?
3. Apakah yang disebut sebagai polimorfisme?
4. Sebutkan manfaat seorang farmasis memahami tentang polimorfisme?
5. Apakah yang disebut sebagai CDS?

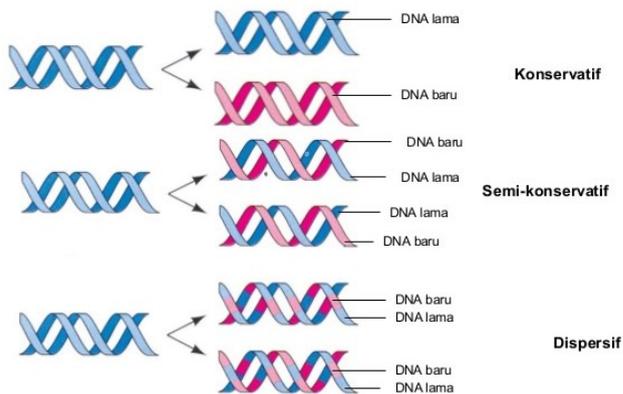
BAB VI. DESAIN PRIMER UNTUK PCR

A. SASARAN CAPAIAN

1. Mahasiswa memahami dan menjelaskan *Polymerase Chain Reaction* (PCR) dan aplikasinya dalam bidang kefarmasian.
2. Mahasiswa mampu mendesain urutan DNA primer sesuai dengan urutan gen yang diinginkan.

B. DASAR TEORI

PCR (*Polymerase Chain Reaction*), merupakan teknik perbanyakan (amplifikasi) potongan DNA secara *in vitro*. Teknik PCR ini sangat berguna untuk berbagai macam analisis DNA untuk berbagai macam kepentingan antara lain penelitian, forensik dan diagnosis penyakit. Jumlah DNA yang diisolasi dari sampel jumlahnya terbatas sehingga tidak dapat langsung dideteksi. Dalam bidang penelitian, PCR dapat digunakan untuk membandingkan urutan nukleotida pada spesies yang berbeda. Dalam bidang forensik, PCR dapat memperbanyak DNA yang digunakan untuk identifikasi pelaku atau korban pembunuhan atau kecelakaan. Selain itu, PCR dapat membantu untuk diagnosis penyakit yang terjadi akibat mutasi pada tingkat DNA.



Gambar 6.1 Teori Replikasi DNA
(sumber: <http://www.slideshare.net/ichottt28/substansi-genetika>)

Reaksi dalam PCR pada dasarnya adalah reaksi replikasi DNA secara semi-konservatif (Gambar 5.1). DNA memerlukan primer dalam sintesis strand DNA baru karena enzim DNA polimerase tidak dapat mensintesis DNA secara *de novo*. Mekanisme PCR sebagai berikut: DNA cetakan dikendorkan dengan enzim topoisomerase dan dibuka oleh enzim helikase, DNA baru disintesis oleh enzim DNA polymerase. Seperti halnya replikasi DNA, reaksi dalam PCR membutuhkan primer. Selain itu juga dibutuhkan DNA sampel yang digunakan sebagai template, Taq polymerase, deoksiribonukleotida (dNTP) dan buffer. Perbedaannya, pada PCR pembukaan *double strand* DNA dilakukan dengan manipulasi suhu reaksi. PCR tergantung kemampuan primer berhibridisasi secara spesifik pada DNA target. Secara umum, proses PCR dibagi menjadi 3 tahap yaitu:

1. Denaturasi

Denaturasi adalah pemisahan *double stranded* DNA menjadi *single stranded* DNA. Denaturasi dilakukan dengan menaikkan suhu tabung reaksi sampai dengan 95°C. Tabung reaksi ini berisi DNA yang diamplifikasi, primer, enzim Taq polymerase dan buffer.

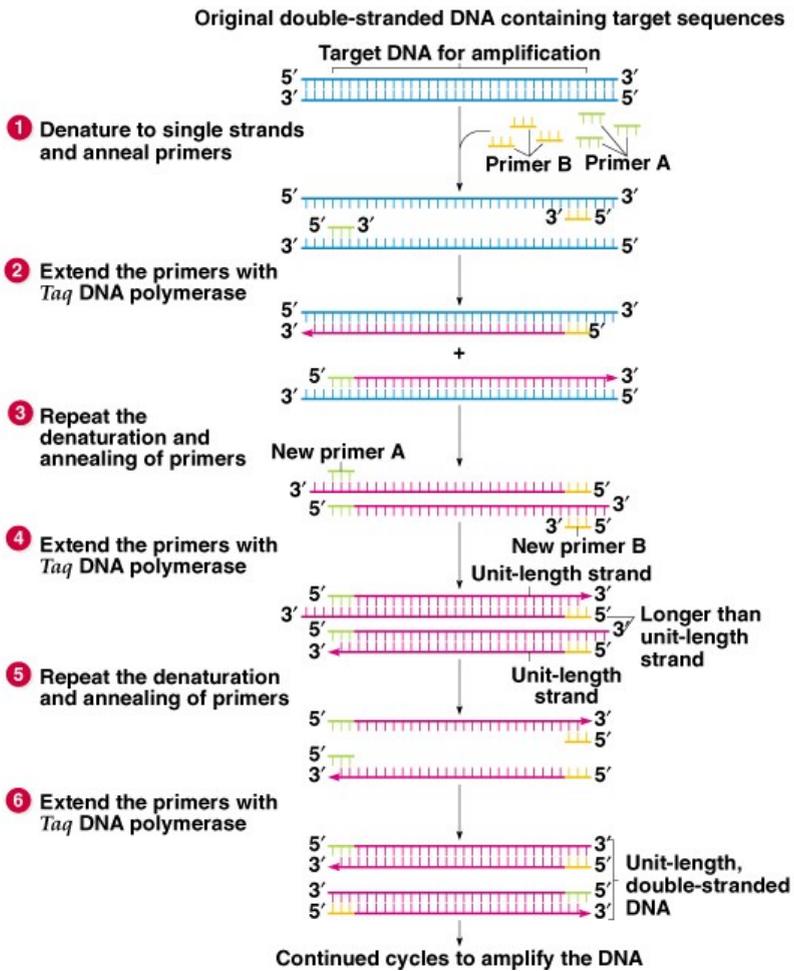
2. Penempelan primer (*annealing*/hibridisasi)

Suhu reaksi diturunkan hingga antara 37-60°C. Pada tahap ini diperlukan optimasi suhu supaya primer dapat menempel pada sekuen DNA target yang akan diamplifikasi. Hibridisasi primer pada oligonukleotida target dipengaruhi oleh kekuatan ionic (konsentrasi ion K⁺ dan Mg²⁺ dalam buffer) dan suhu. Suhu hibridisasi perlu dioptimasi supaya primer dapat menempel tepat pada DNA target. Suhu hibridisasi sangat dipengaruhi oleh desain primer yang digunakan pada PCR.

3. Polimerisasi

Pada tahap ini terjadi perpanjangan rantai DNA pada ujung 3'-OH bebas pada primer. Suhu reaksi kembali dinaikan menjadi 72°C.

Berikut adalah contoh pengaturan suhu dan siklus pada PCR



Gambar 6.2. Reaksi dalam siklus PCR
(sumber: <http://www.life.illinois.edu/ib/335/MolSyst.html>)

Siklus	Denaturasi	<i>Annealing</i>	Polimerisasi
1	94 °C, 10 menit	55 °C, 2 menit	72 °C, 2 menit
2-29	94 °C, 1 menit	55 °C, 2 menit	72 °C, 2 menit
30	94 °C, 1 menit	55 °C, 2 menit	72 °C, 2 menit

Jenis PCR ada 4 macam; antara lain: *Real time/quantitative* PCR, *Nested* PCR, *Restriction Fragment Length Polymorphism* (RFLP) PCR dan *Reverse transcript* (RT) PCR. Real time PCR merupakan PCR yang dapat digunakan untuk tujuan kuantitatif sehingga disebut kuantitatif PCR (q-PCR). Pada PCR konvensional, data yang diperoleh berupa data kualitatif. Real time PCR lebih aplikatif digunakan untuk analisis ekspresi suatu gen akibat perlakuan tertentu. Berbeda dengan PCR yang lain, nested PCR adalah modifikasi PCR yang bertujuan untuk mengurangi resiko kontaminasi akibat primer salah menempel. Pada PCR konvensional, primer menempel pada terminal DNA target. Problem yang sering muncul adalah penempelan primer pada sekuen non spesifik sehingga kemungkinan terjadi positif palsu. Nested PCR menggunakan 2 pasang primer selama proses PCR, primer kedua bertujuan mengamplifikasi DNA target yang berada diantara DNA target amplifikasi pertama. Dengan demikian, sekuen DNA yang diperbanyak lebih spesifik.

Reverse transcript-PCR (RT-PCR) merupakan PCR yang menggunakan sampel mRNA untuk diamplifikasi. Molekul mRNA kemudian ditranskripsi balik menjadi cDNA yang kemudian diamplifikasi menggunakan teknik PCR. Teknik PCR yang lain adalah *Restriction Fragment Length Polymorphism* PCR (RFLP PCR) menggunakan teknik amplifikasi PCR secara umum. Pada akhir proses PCR, DNA hasil PCR dipotong menggunakan enzim restriksi. RFLP PCR berguna untuk membedakan DNA yang mengalami mutasi pada basa tertentu. Berbagai

macam penyakit yang disebabkan oleh mutasi dapat dideteksi menggunakan RFLP PCR.

Dalam bab ini, dilakukan desain primer dengan tujuan untuk mendapatkan primer yang spesifik supaya hanya dapat memperbanyak daerah DNA yang diinginkan. Metode PCR dapat meningkatkan jumlah urutan DNA ribuan bahkan jutaan kali dari jumlah semula, sekitar $10^6 - 10^7$ kali. Primer dapat disusun secara manual. Primer akan disintesis sebagai oligonukleotida sepanjang 15 - 32 bp dengan urutan dari ujung-5' ke ujung-3' dan primer ini harus mampu mengenali urutan yang diamplifikasi. Untuk standar amplifikasi sepanjang primer mempunyai kisaran pasangan basa sekitar 20 basa panjangnya pada tiap primernya.

Beberapa syarat yang harus dipenuhi oleh primer yang baik antara lain: urutan primer mengandung urutan basa GC sebanyak 45-60%. Primer yang kandungan GC-nya kurang dari 50% perlu diperpanjang menjadi lebih dari 18 basa sehingga T_m menjadi diatas 50°C . Presentase GC, selain mempengaruhi T_m juga suhu *annealing* (Temperature annealing- T_a). Ujung-3' dari setiap primer harus G atau C, tetapi hindari susunan nukleotida G/C berturut-turut tiga pada ujung ini, misal CCG, GCG, GGC, GGG, CCC, GCC. Ujung 3' penting untuk menghindari *miss-priming*.

Urutan basa dalam primer tidak saling komplementer sehingga membentuk dimer-primers, berikatan satu sama lain, atau membentuk *hairpins*. Selain itu, untuk menghindari menyusun primer pada daerah DNA repetitif. Selain itu, faktor homologi urutan nukleotida dengan urutan DNA target sangat mempengaruhi kestabilan ikatan keduanya. Semakin tinggi persentase homologi semakin stabil ikatan yang terbentuk.

Penempelan (*annealing*) suatu primer terhadap DNA target tergantung pada ujung untai, banyaknya kandungan GC, dan konsentrasi primer itu sendiri. Optimalisasi *annealing temperature* dihitung dengan kalkulasi *melting temperature* (T_m) dari ikatan primer dan DNA template. *Annealing temperature* sangat tergantung pada primer dengan T_m yang tertentu. Cara *melting temperature* dapat menggunakan persamaan matematika sebagai berikut:

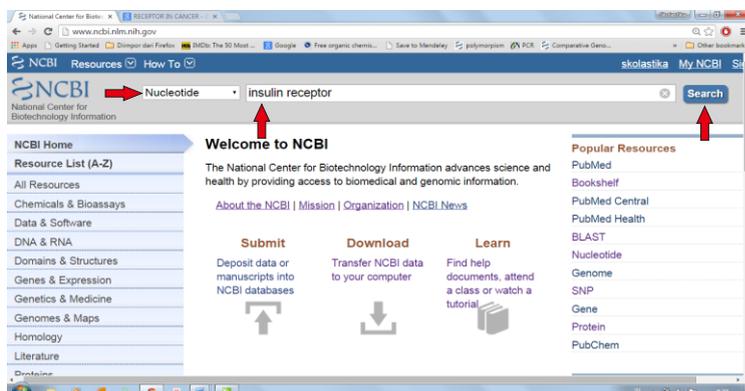
$$T_m = \{(G + C) \times 4\} + \{(A + T) \times 2\}$$

Annealing temperature umumnya 5°C dibawah T_m primer yang sebenarnya. Secara praktis, T_m ini dipengaruhi oleh komponen buffer, konsentrasi primer, dan DNA template.

Dewasa ini, desain primer dilakukan dengan menggunakan *software* baik pada computer maupun website secara online. *Software* yang digunakan untuk bidang biologi memberikan wacana tersendiri pada bidang bioinformatika. Dalam buku ini ditunjukkan salah satu cara mendesain primer yang sederhana menggunakan basis data online tidak berbayar.

C. PROSEDUR KERJA

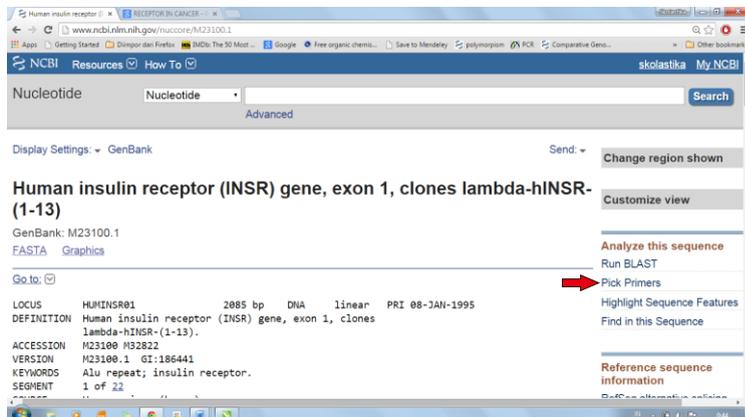
1. Buka situs web <http://www.ncbi.nlm.nih.gov/> dan pilih pada kolom **Search: nucleotide, for: nama gen** kemudian klik **search**



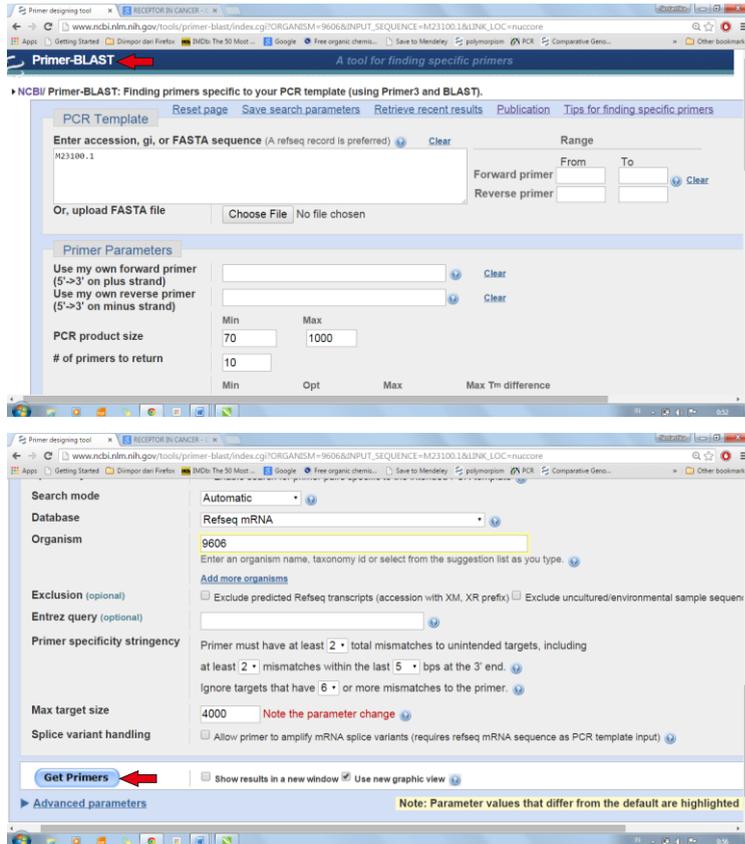
2. Hasil pencarian akan muncul. Klik salah satu gen yang diinginkan.



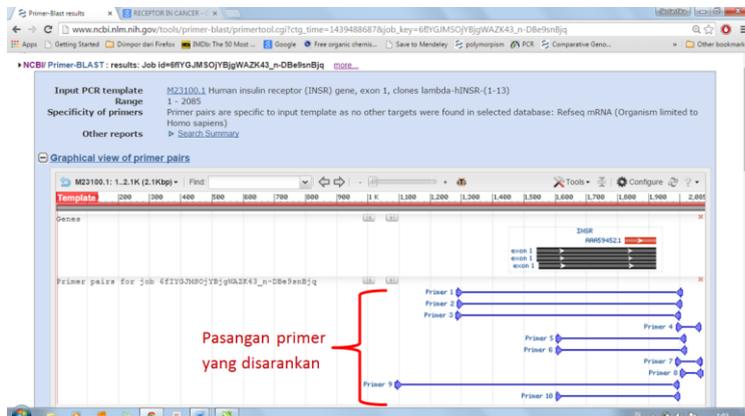
3. Pilih *pick primers* yang terdapat pada bagian kanan.



4. Selanjutnya jendela primer BLAST akan terbuka. Scroll kursor kebawah. Klik *Get primers*.



5. Berikut contoh tampilan hasil:



6. Scroll ke bawah untuk melihat urutan nukleotida pasangan primer.

Primer pair	Sequence (5'→3')	Template strand	Length	Start	Stop	Tm	GC%	Self complementarity	Self 3' complementarity
Primer pair 1									
Forward primer	GAGTCCCTTCTAGGCCAGA	Plus	20	1284	1303	60.03	60.00	6.00	2.00
Reverse primer	ACAAGCGGGGATCGATTTT	Minus	20	2007	1988	60.03	50.00	6.00	0.00
Product length: 724									
Primer pair 2									
Forward primer	TCCCTTCTAGGCCAGATCC	Plus	20	1287	1306	60.10	60.00	6.00	2.00
Reverse primer	CAAGCGGGGATCGATTTT	Minus	20	2006	1987	59.62	55.00	6.00	0.00
Product length: 720									
Primer pair 3									
Forward primer	CAGAGTCCCTTCTAGGCCA	Plus	20	1282	1301	60.33	60.00	6.00	2.00
Reverse primer	GGGGATCGATTTGGCTT	Minus	20	2001	1982	59.25	55.00	6.00	0.00
Product length: 720									
Primer pair 4									
Forward primer	AAATCGATCCCCGCTTGT	Plus	20	1988	2007	60.03	50.00	6.00	0.00
Reverse primer	CGATGAGGAGTACCGTC	Minus	19	2066	2048	60.30	63.16	4.00	2.00
Product length: 70									
Primer pair 5									
Forward primer	AGGGCCTCCCTCAGTATT	Plus	20	1604	1623	60.98	55.00	5.00	0.00
Reverse primer	GTTTCTCAGTCCACAAGCG	Minus	20	2019	2000	60.95	60.00	3.00	2.00
Product length: 416									

D. PERTANYAAN DISKUSI

1. Sebutkan aplikasi PCR dalam dunia kesehatan!
2. Mengapa dalam PCR diperlukan primer?
3. Sebutkan dan jelaskan 4 jenis PCR!
4. Mengapa mendesain primer adalah kunci keberhasilan proses PCR?
5. Sebutkan persyaratan primer yang baik!

DAFTAR PUSTAKA

- Albert, B., Bray, D., Lewis, J., Rarr, M., Roberts, K. and Watson, J. O., 1994, *Molecular Biology of The Cell*, 3rd Ed., Garland Publishing, Inc., New York.
- Chotia C., Gough J., Vogel C., Teichmann SA., 2003, Evolution of Protein Repertoire, *Science* 300, 1701.
- Ezkurdia I., Tress M.L., 2001, Protein Structural Domains: Definition and Prediction, *Current Protocols in Protein Science*. Supplement 66.
- Fatchiyah, 2014, *Prinsip Dasar Bioinformatika*, Universitas Brawijaya Press, Malang.
- Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins MR., Appel RD., Bairoch A., Protein Identification and Analysis Tools on the ExPASy Server. The Proteomics Protocols Handbook. Totowa, NJ, p 571
- IUPAC-IUB Commision on Biochemical Nomenclature (CBN), 1968, A One Letter Notation for Amino Acid Sequences, *European J.Biochem*, 5: 151- 153.
- Lodish, H., Arnold, B., Zipursky, S. L., Matsudara, P., David, B. and Darnell, J. E., 2000, *Molecular Cell Biology*, W. H. Freeman and Company, London.
- Kyte, J., and Doolittle, R. F., 1982, A simple method for displaying the hydrophathic character of a protein. *J. Mol. Biol.* 157, 105–132.
- Koonin EV. 2005. Orthologs, paralogs, and evolutionary genomics. *Annu Rev Genet.* 39:309-38. Review.

National Institutes of Health, 2015, Handbook Help Me Understand Genetics The Human Genome Project, Reprinted from Genetics Home Reference (<http://ghr.nlm.nih.gov/>)

Pearson WR., 2013. An Introduction to Sequence Similarity ("Homology") Searching. *Curr.Protoc.Bioinform.* 42:3.1.1-3.1.8.

Snider C., Jayasinghe S., Hristova K, White Sh., 2009, MPEX: A tool for exploring membrane protein. *Protein Science*, 18: 2624- 2628.

Tim pengampu Praktikum Biologi Molekuler, 2008, *Buku Petunjuk Praktikum Biologi Molekuler*, Fakultas Farmasi, Universitas Gadjah Mada, Yogyakarta.

www.ncbi.nlm.nih.gov

www.ebi.ac.uk

www.ddbj.nig.ac.jp

www.proteinsynthesis.org/

www.tcdb.org/progs/hydro.php

www.slideshare.net/ichottt28/substansi-genetika

www.life.illinois.edu/ib/335/MolSyst.html

**LAMPIRAN DAFTAR KODE PROTEIN/NUKLEOTIDA DALAM NCBI
YANG DAPAT DIGUNAKAN UNTUK PEMBELAJARAN**

1. **Protein: NP_002677.1/NM_002686.4** (phenylethanolamine N-methyltransferase). Transd: **SAM** (S-adenosylmethionine-dependent methyltransferases).
2. **Protein: NP_005949.1/ NM_005958.3** (melatonin receptor). Transd: **MTNR1A**.
3. **Nucleotide: NM_000912.3/ NP_000903.2** (opioid receptor kappa 1). Transd: **OPRK1**.
4. **Protein: NP_009163.2/NM_007232.2** (histamine H3 receptor). Transd: **HRH3**.
5. **Nucleotide: NM_004224.3** (Homo sapiens G protein-coupled receptor 50). Transd: **GPR50**.
6. **Protein: NP_006158.2/NM_006167.3** (protein Nkx-3.1 isoform 1). Transd: **NKX3-1**.
7. **Protein: NP_001287.2/NM_001296.4** (atypical chemokine receptor). Transd: **ACKR2**.
8. **Protein: NP_001020018.1/ NM_001024847.2** (TGF-beta receptor type-2). Transd: **TGFBR2**.
9. **Protein: NP_000537.3/ NM_000546.5** (cellular tumor antigen p53). Transd: **TP53**.
10. **Protein: NP_000029.2** (adenomatous polyposis coli protein). Transd: **APC**.
11. **Protein: NP_032897.2** (plasminogen activator inhibitor 1 precursor). Transd: **Serpine1**.
12. **Protein: NP_000292.1** (plasminogen isoform 1 precursor). Transd: **PLG**.
13. **Protein: NP_003591.2** (aurora kinase A). Transd: **AURKA**.

14. **Protein: NP_001091679.1/ NM_001098209.1** (catenin beta-1).
Transd: **CTNNB1**.
15. **Protein: NP_000917.3/ NM_000926.4** (progesterone receptor).
Transd: **PGR**.
16. **Protein: NP_001138303.1/ NM_001144831.1** (prohibitin-2).
Transd: **PHB2**.
17. **Protein: NP_000066.1/ NM_000075.3** (cyclin-dependent kinase 4). Transd: **CDK4**.
18. **Protein: NP_000651.3** (transforming growth factor beta-1).
Transd: **TGFB1**.
19. **Protein: NP_001175.2** (serine/threonine-protein kinase ATR).
Transd: **ATR**.
20. **Protein: NP_000042.3** (serine-protein kinase ATM). Transd: **ATM**.
21. **Protein: NP_000624.2** (apoptosis regulator Bcl-2 alpha isoform).
Transd: **BCL2**.
22. **Protein: NP_001278357.1** (apoptosis regulator BAX isoform 1).
Transd: **BAX**.
23. **Protein: NP_001218.1** (caspase-7 isoform alpha precursor).
Transd: **CASP7**.
24. **Protein: NP_004313.1** (bcl2 antagonist of cell death). Transd:
BAD.
25. **Protein: NP_009041.2** (telomerase protein component 1). Transd:
TEP1.
26. **Protein: NP_001180305.1** (telomerase reverse transcriptase
isoform 2). Tansd: **TERT**.
27. **Protein: NP_001171601.1** (epidermal growth factor). Transd:
EGF.
28. **Protein: NP_000866.1** (insulin-like growth factor 1 receptor).
Transd: **IGF1R**.
29. **Protein: NP_001092.1** (actin). Transd: **ACTB**.

30. **Protein: XP_011530308.1** (nuclear factor NF-kappa-B). Transd: **NFKB1**.
31. **Protein: NP_001231063.1** (mitogen-activated protein kinase). Transd: **MAP3K8**.
32. **Protein: NP_004351.1** (cadherin-1). Transd: **CDH1**.
33. **Protein: XP_011518955.1** (GTPase Kras). Transd: **KRAS**.
34. **Protein: NP_000068.1** (cyclin-dependent kinase inhibitor 2A). Transd: **CDKN2A**.
35. **Protein: NP_000591.1** (interleukin-6). Transd: **IL6**.
36. **Protein: NP_954592.1** (histone-arginine methyltransferase CARM1). Transd: **CARM1**.
37. **Protein: NP_000585.2** (tumor necrosis factor). Transd: **TNF**.
38. **Protein: NP_000169.1** (glutathione synthetase). Tansd: **GSS**.
39. **Protein: NP_000628.2** (glutathione reductase). Transd: **GSR**.
40. **Protein: NP_000302.1** (major prion protein). Transd: **PRNP**.
41. **Protein: NP_001019978.1** (amyloid-like protein 1). Transd: **APLP1**.
42. **Protein: NP_001138884.1** (nuclear factor erythroid 2-related factor 2). Transd: **NFE2L2**.
43. **Protein: NP_001124491.1** (glial fibrillary acidic protein isoform 2). Transd: **GFAP**.
44. **Protein: NP_001181883.1** (matrin-3 isoform a). Transd: **MATR3**.
45. **Protein: NP_001020278.1** (double-stranded RNA-specific adenosine deaminase). Transd: **ADAR**.
46. **Protein: NP_000570.1** (chemokine receptor). Transd: **CCR5**.
47. **Protein: NP_001186078.1** (lactotransferrin isoform 2). Transd: **LTF**.
48. **Protein: NP_005912.1** (mitogen-activated protein kinase kinase). Transd: **MAP3K1**.

49. **Protein: NP_001269.3** (inhibitor of nuclear factor kappa-B kinase subunit alpha). Transd: **CHUK**.
50. **Protein: NP_001020413.1** (interleukin-1 receptor-associated kinase 1). Transd: **IRAK1**.
51. **Protein: NP_001166037.1** (myeloid differentiation primary response protein MyD88). Transd: **MYD88**.
52. **Protein: NP_004611.1** (TNF receptor-associated factor 6). Transd: **TRAF6**.
53. **Protein: NP_000954.1** (prostaglandin G/H synthase 2). Transd: **PTGS2**.
54. **Protein: NP_000953.2** (prostaglandin G/H synthase 1). Transd: **PTGS1**.

BIODATA PENULIS



Agustina Setiawati adalah seorang dosen di bidang Biologi Molekuler di Fakultas Farmasi, Universitas Sanata Dharma (USD). Agustina lulus S1 Program Studi Ilmu Farmasi, Universitas Gadjah Mada (UGM) tahun 2007, selanjutnya berhasil menyelesaikan pendidikan profesi Apoteker di Perguruan Tinggi yang sama pada tahun 2008. Selanjutnya, Agustina melanjutkan studi S2 Ilmu Farmasi minat Penemuan dan Pengembangan Obat di UGM tahun 2008 dengan beasiswa ASEA UNINET. Sejak tahun 2010 tercatat aktif sebagai dosen dan penelitian yang ditekuni adalah penelitian biologi molekuler pada sel kanker. Mata kuliah yang diampu antara lain Biologi Sel dan Molekuler, Biokimia dan Bioteknologi.