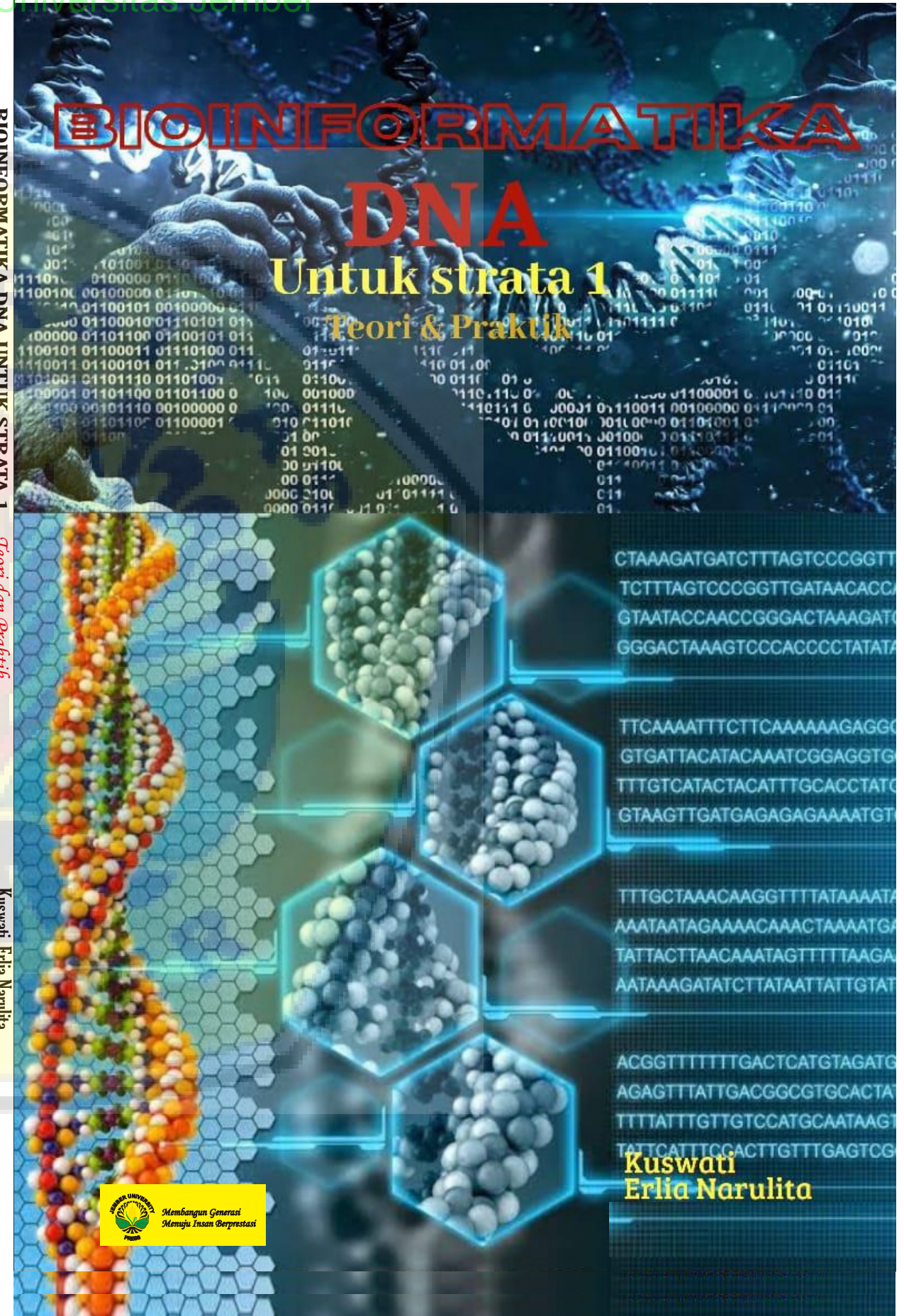




BIOINFORMATIKA DNA UNTUK STRATA 1

Teori dan Praktik

Kuswati, Erlia Narulita



Anggota APPTI No. 002.115.1.05.2020

Anggota IKAPI No. 127/JTI/2018

Jember University Press
Jl. Kalimantan 37 Jember 68121
Telp. 0331-330224, psw. 0319
E-mail: upt-penerbitan@unej.ac.id



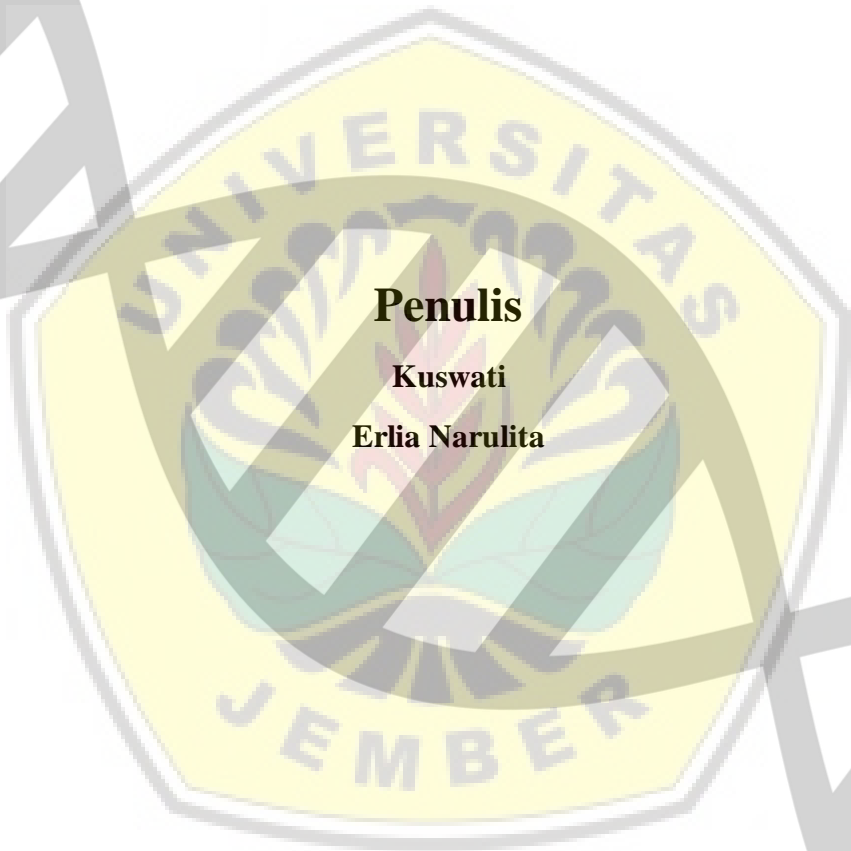
**Kuswati
Erlia Narulita**

**BIOINFORMATIKA DNA/RNA untuk
STRATA 1:
TEORI dan PRAKTIK**

Penulis

Kuswati

Erlia Narulita



Bioinformatika DNA/RNA untuk Strata 1: Teori dan Praktik

Penulis:

Kuswati
Erlia Narulita

Editor : -

Watermark : <https://pixabay.com/id/illustrations/dna-ilmu-pengetahuan-biologi-718905/>

ISBN:

978-623-7973-43-0

Penerbit:

UPT Penerbitan Universitas Jember

Redaksi:

Jl. Kalimantan 37
Jember 68121
Telp. 0331-330224, Voip. 00319
e-mail: upt-penerbitan@unej.ac.id

Distributor Tunggal:

UNEJ Press
Jl. Kalimantan 37
Jember 68121
Telp. 0331-330224, Voip. 0319
e-mail: upt-penerbitan@unej.ac.id

Hak Cipta dilindungi Undang-Undang. Dilarang memperbanyak tanpa izin tertulis dari penerbit, sebagian atau seluruhnya dalam bentuk apapun, baik cetak, photoprint, maupun microfilm.

Prakata

Puji syukur kehadirat Allah SWT atas limpahan rahmat, berkah dan hidayah penulis dapat menyusun dan menyelesaikan buku ini sebagai bahan ajar dengan tepat waktu. Sholawat dan salam selalu tercurahkan pada suri tauladan umat manusia, Nabi Muhammad SAW, yang telah menunjukkan jalan penuh nikmat dan ilmu pengetahuan. Penulis menyampaikan terima kasih yang sebesar-besarnya atas segala dukungan dari keluarga dan rekan dosen serta seluruh sivitas akademika di PS. Pendidikan Biologi Universitas Jember.

Mata kuliah Bioinformatika merupakan salah satu mata kuliah pilihan bagian dari kelompok riset *Biomolecular Engineering and Education* (BEE) PS. Pendidikan Biologi Universitas Jember. Perkembangan ilmu pengetahuan Bioinformatika sangat pesat seiring dengan berkembangnya teknologi biomolekuler yang semakin *fascinating*. Ketersediaan teknologi *sequencing* semakin hari semakin mudah dan murah, penambahan jumlah data molekuler urutan maupun struktur pada database yang terus meningkat secara eksponensial yang *publicly accessed* menyebabkan bidang ilmu bioinformatika sangat berkembang. Terdapat banyak keuntungan dalam penelitian *In silico* (bioinformatika) sehingga menjadi sangat penting membekali mahasiswa PS. Pendidikan Biologi dengan Ilmu bioinformatika DNA untuk Strata 1 yang berisi tentang teori dan praktik. Buku ajar ini terdiri atas 6 bab yaitu: Pengantar dan pengenalan database; Perbandingan urutan; Prediksi gen dan promotor; Prediksi motif dan domain; Prinsip dasar desain primer; serta Konstruksi pohon filogenetik. Keenam bab tersebut dilengkapi dengan praktik aplikasi dari teori bioinformatika. Setiap akhir Bab juga disertai dengan rangkuman dan latihan soal.

Buku ajar ini masih sangat membutuhkan pengembangan dan *update* sehingga penulis sangat menerima adanya kritik dan saran untuk perbaikan berkelanjutan sampai menghasilkan buku ajar yang terus lebih baik.

Jember, Desember 2020

Penulis

Daftar Isi

Halaman Judul	i
Kata Pengantar	ii
Daftar Isi	iii
Deskripsi Mata Kuliah	1
BAB 1. Pendahuluan	
1.1 Deskripsi dan Cakupan Bioinformatika	2
1.2 Data Molekul Biologi sebagai Basis Data Bioinformatika	4
1.3 Aplikasi Bioinformatika	7
DAFTAR PUSTAKA	8
BAB 2. Pengenalan Database Molekul Biologi	9
2.1 Klasifikasi Database Molekul Biologis	10
2.2 Struktur Data pada Database	12
RINGKASAN	15
BAHAN LATIHAN	16
DAFTAR PUSTAKA	20
LATIHAN SOAL	21
BAB 3. Penjajaran Urutan (PSA dan MSA)	22
3.1 Deskripsi PSA dan MSA	23
3.2 Aplikasi Prinsip Penjajaran Urutan pada Analisis Bioinformatika	25
3.3 Program untuk Analisis PSA dan MSA	26
RINGKASAN	27
BAHAN LATIHAN	28
DAFTAR PUSTAKA	32
LATIHAN SOAL	33
BAB 4. Prediksi Gen, Promoter, dan ORF	34
4.1 Prediksi Gen pada Organisme Prokariotik	35
4.2 Prediksi Gen pada Organisme Eukariotik	35
RINGKASAN	37
BAHAN LATIHAN	38
DAFTAR PUSTAKA	39
LATIHAN SOAL	40
BAB 5. Prediksi domain dan motif	41
5.1 Struktur Protein	42

5.2 Aplikasi Bioinformatika untuk Prediksi Protein	46
RINGKASAN	47
BAHAN LATIHAN	48
DAFTAR PUSTAKA	52
LATIHAN SOAL	53
BAB 6. Desain Primer	54
6.1 Konsep Primer	55
6.2 Program Desain Primer	57
RINGKASAN	60
BAHAN LATIHAN	60
DAFTAR PUSTAKA	64
DAFTAR PUSTAKA	66
BAB 7. Konstruksi Pohon Filogenetika Molekuler	67
7.1 Deskripsi Filogenetika	68
7.2 DNA sebagai Basis Data Analisis Filogenetik	68
7.3 Tahapan Analisis Filogenetik	69
RINGKASAN	74
BAHAN LATIHAN	75
DAFTAR PUSTAKA	76
LATIHAN SOAL	77
GLOSARIUM	79
INDEKS	81
BIOGRAFI PENULIS	82

Deskripsi Singkat Mata Kuliah

Matakuliah	: Bioinformatika Protein
Kode MK	: KPB1744
Kredit	: 2 SKS
Rumpun MK	: Mata Kuliah Pilihan Kelompok Riset (KeRis) <i>Biomoleculer Engineering and Education (BEE)</i>

Deskripsi Singkat Mata Kuliah

Mata kuliah ini dirancang untuk memberikan pengenalan dasar bioinformatika DNA/RNA yang meliputi peranan informasi urutan DNA/RNA dalam memahami proses biologi, sumber daya (basis data) dan aplikasi-aplikasi yang digunakan secara luas di bidang bioinformatika, algoritma-algoritma yang digunakan untuk menyelesaikan permasalahan di bidang bioinformatika (*in silico*), khususnya yang terkait dengan urutan DNA. Penggunaan informasi database biologi tersebut tertuang bab dalam buku ini yang meliputi: Pengenalan Database Biologi; Penjajaran Urutan: *Pairwise Sequence Alignment (PSA)* dan *Multiple Sequence Alingment (MSA)*; Prediksi Gen, Promoter dan ORF (*Open Reading Frame*); Prediksi Motif dan Domain Protein; Desain Primer; dan Konstruksi pohon filogenetik berbasis informasi data molekuler. Pembelajaran dilaksanakan dengan diskusi, ceramah, studi kasus dan dilanjutkan dengan praktik pada masing-masin teori.

Adapun Capaian Pembelajaran Program Studi (CPL) yang dibebankan pada matakuliah Bioinformatika DNA/RNA Yaitu:

- CPL 3 : Mahasiswa mampu **melakukan** kerja laboratorium (*in silico*) dan studi lapang secara mandiri dengan mengutamakan keselamatan kerja.
- CPL 6 : Mahasiswa mampu **mendemonstrasikan** keterampilan dan pengetahuan dalam menyelesaikan permasalahan sumber daya hayati, lingkungan, dan pembelajaran biologi dalam lingkup agroekosistem

Bab 1

Pendahuluan Bioinformatika DNA/RNA

1.1 Deskripsi dan Cakupan Bioinformatika

Belakangan ini banyak sekali dijumpai istilah *in silico* dalam banyak media tulisan baik berupa tulisan ilmiah akademis maupun tulisan populer. Apakah mahasiswa memahami tentang *in silico*? *In silico* diambil dari kata *silicon chips* yang merupakan komponen utama sebuah mikroprosesor (Luxon, 2003). Dalam percobaan *in silico*, alat utama yang dibutuhkan adalah komputer. Dengan pengembangan komputasi tertentu sebuah percobaan dilakukan dengan simulasi dan kalkulasi sehingga sebuah permasalahan dapat dipecahkan. Dalam eksperimen biologi, metode *in silico* bersifat komplemen dengan kedua metode yang telah ada sebelumnya yaitu *in vivo* dan *in vitro*.

Sebagai ilustrasi, Istyastono (2011) dalam percobaan penemuan obat terhadap senyawa A yang terkandung dalam tanaman X yang berpotensi sebagai antikanker karena berfungsi sebagai inhibitor pada enzim Z yang berperan dalam mekanisme kanker, maka dari informasi tersebut:

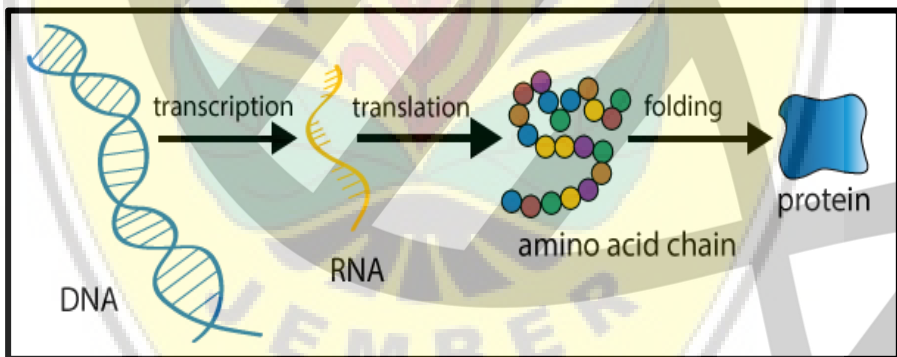
1. Secara konvensional
Akan dilakukan uji lanjutan terhadap senyawa potensial tersebut. Kemudian hal lain adalah melakukan skrining pada tanaman lain berdasarkan fakta yang ditemukan.
2. Dengan bantuan komputer
Beberapa senyawa dapat disimulasikan secara visual (menggunakan program tertentu) interaksi antara molekul senyawa aktif dan enzim, sehingga didapatkan kandidat satu senyawa yang sangat potensial. Dilakukan skrining pada tanaman lain juga dapat dilakukan dengan analisis kekerabatan berdasarkan penanda molekuler.

Bab 2

Database Bioinformatika

Sub-CPMK (Kemampuan Akhir yang Diharapkan):

- ✓ Mahasiswa mampu menjelaskan konsep dasar bioinformatika dan database biologi
- ✓ Mahasiswa mampu melaksanakan pencarian data urutan dengan prosedur yang tepat
- ✓ Mahasiswa mampu melakukan pencarian data urutan DNA organisme di daerah agroekosistem



INFO BUAT KAMU!

Bioinformatika mempelajari urutan DNA/RNA, Protein serta strukturnya.

http://sciencewithmsjones.weebly.com/uploads/3/8/6/0/38609001/central-dogma-of-dna_orig.png

Bab 3

Penjajaran Urutan (Pairwise Sequence Alignment (PSA) dan Multiple Sequence Alignment (MSA))

Sub-CPMK (Kemampuan Akhir yang Diharapkan):

- ✓ Mahasiswa mampu menjelaskan konsep dasar penjajaran urutan baik PSA maupun MSA
- ✓ Mahasiswa mampu membaca hasil penjajaran urutan

The screenshot displays the NCBI BLAST website. At the top, it says "Basic Local Alignment Search Tool" and provides a brief description of BLAST's function. Below this, there is a "Web BLAST" section with four main options: "Nucleotide BLAST" (nucleotide ↔ nucleotide), "blastx" (translated nucleotide ↔ protein), "tblastn" (protein ↔ translated nucleotide), and "Protein BLAST" (protein ↔ protein). A news alert on the right side of the page states: "A new feature was added to Primer-BLAST. We have added a new function to Primer-BLAST that helps users design primers common for a group of highly similar sequences. Tue, 20 Sep 2020 12:00:00 EST. See BLAST news..."

INFO BUAT KAMU!

BLAST merupakan sebuah program *family*, menawarkan beberapa pilihan penjajaran urutan berdasarkan *query* dan database target.

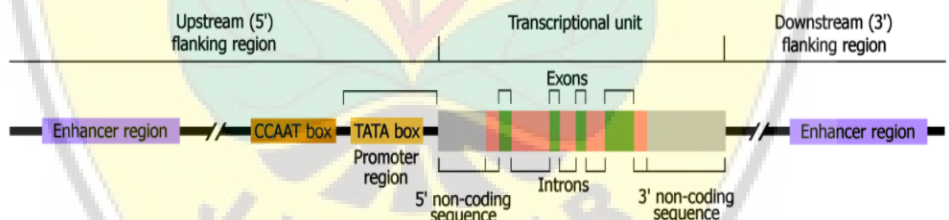
<https://blast.ncbi.nlm.nih.gov/Blast.cgi>

Bab 4

Prediksi Gen, Promoter dan ORF (Open Reading Frame)

Sub-CPMK (Kemampuan Akhir yang Diharapkan):

- ✓ Mahasiswa mampu menjelaskan konsep dasar bioinformatika dan database biologi
- ✓ Mahasiswa mampu melaksanakan pencarian data urutan dengan prosedur yang tepat
- ✓ Mahasiswa mampu melakukan pencarian data urutan DNA SDH yang terdapat di kawasan agroekosistem



INFO BUAT KAMU!

Meskipun hanya 3% daerah coding sequence pada eukariotik, namun variasi ekspresi gen jauh lebih banyak dibanding organisme prokariotik. Salah satu sebabnya adalah kompleksitas organisasi struktur gen eukariotik (Carr, M. S. 2018)

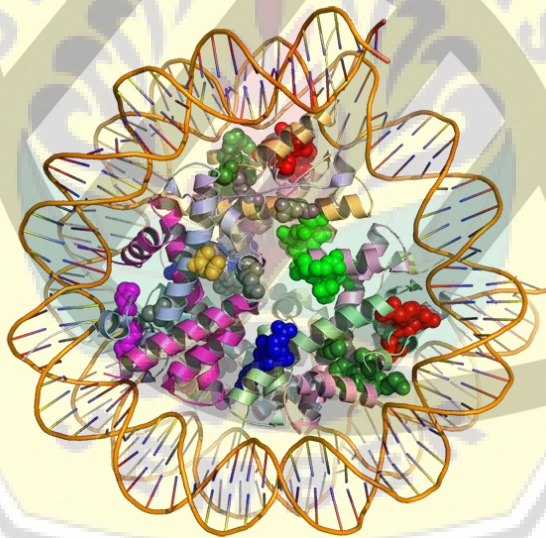
https://www.mun.ca/biology/scarr/Organization_of_a_eukaryotic_gene.html

Bab 5

Deteksi Domain dan Motif Protein

Sub-CPMK (Kemampuan Akhir yang Diharapkan):

- ✓ Mahasiswa mampu menjelaskan konsep dasar bioinformatika dan database biologi
- ✓ Mahasiswa mampu melaksanakan pencarian data urutan dengan prosedur yang tepat



INFO BUAT KAMU!

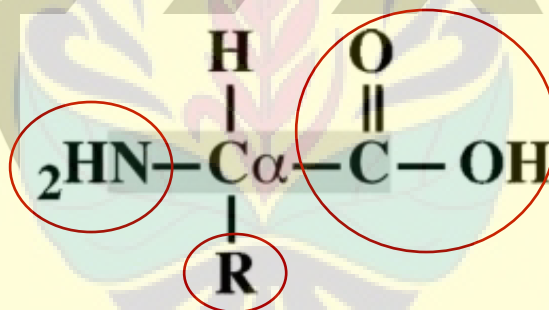
Struktur 3D protein dapat menjelaskan bagian penting dari sebuah protein meliputi motif dan domain. Informasi tersebut bermakna untuk riset selanjutnya.

Washington University in St. Louis / Li Ding, Ph.D., and Feng Chen, Ph.D.

<https://www.cancer.gov/news-events/cancer-currents-blog/2016/protein-structure-drug-targets>

5.1 Struktur Protein

Protein merupakan salah satu biomolekul yang variasi fungsinya sangat dipengaruhi oleh variasi yang terjadi di tingkat strukturnya. Jumlah struktur protein yang berhasil diketahui mencapai bilangan yang sangat besar, namun relatif sangat kecil jika dibandingkan dengan jumlah protein yang ada dalam tubuh suatu organisme. Fungsi esensial diantaranya sebagai penyusun struktural, enzim, transporter dan molekul regulator. Unit dasar penyusun protein adalah asam amino yang secara alami dapat dibuat oleh sel dengan variasi sebanyak 20. Adapun struktur dasar asam amino ditunjukkan pada Gambar 1. Asam amino memiliki gugus amino (NH_2), gugus karboksil (COOH), rantai samping (R) dimana ketiganya terikat pada atom C α . Variasi pada rantai samping (R) menentukan sifat spesifik dari asam amino seperti bersifat hidrofobik, polar, asam dan basa.



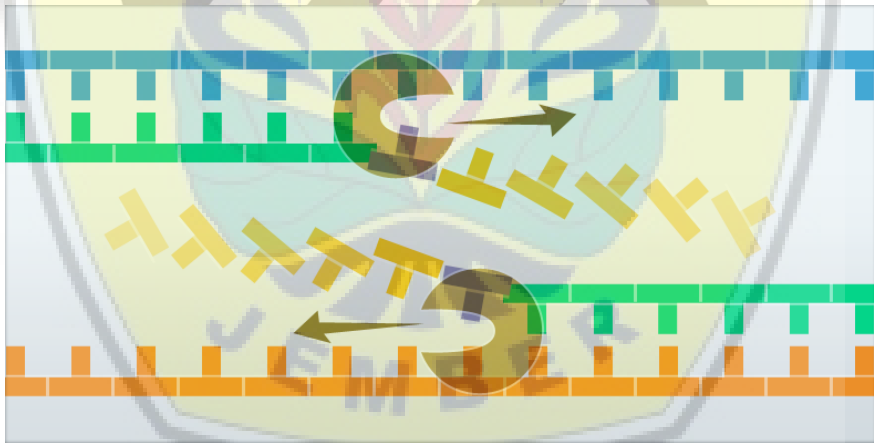
Gambar 1. Struktur Asam Amino

Asam amino dapat dikelompokkan ke dalam beberapa kategori berdasarkan sifat rantai sampingnya misalnya: ukuran molekul dan afinitas terhadap air (hidrofobik dan hidrofilik). Kelompok asam amino hidrofobik, berdasarkan susunan karbon C dibagi ke dalam kelompok asam amino alifatik dengan susunan rantai C linear; dan asam amino aromatik dengan susunan rantai karbon C siklik. Sedangkan kelompok asam amino hidrofilik dibagi menjadi kelompok bermuatan positif (basa) dan bermuatan negatif (asam). Ilustrasi pada Gambar 2 secara sederhana

Bab 6 Desain Primer

Sub-CPMK (Kemampuan Akhir yang Diharapkan):

- ✓ Mahasiswa mampu menjelaskan konsep dasar bioinformatika dan database biologi
- ✓ Mahasiswa mampu melaksanakan pencarian data urutan dengan prosedur yang tepat
- ✓ Mahasiswa mampu melakukan pencarian data urutan DNA SDH yang terdapat di kawasan agroekosistem



INFO TAHUKAH KAMU!

Primer memiliki peranan penting dalam perbanyakan gen target menggunakan teknologi PCR. Oleh sebab itu, sangat penting untuk menguji kembali spesifitas dan kualitas primer yang telah dibuat.

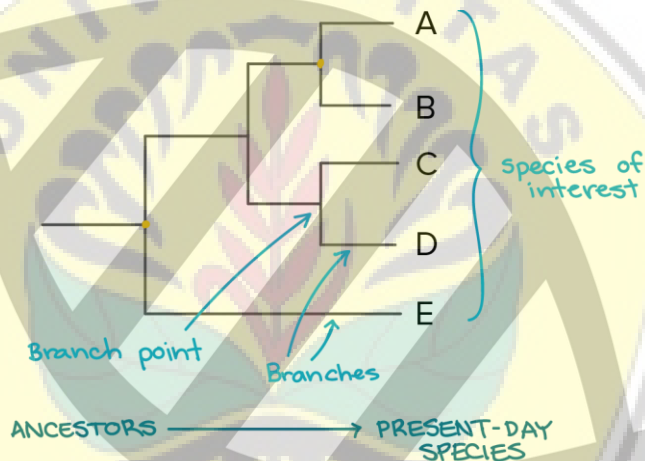
<http://www.genomecompiler.com/tips-for-efficient-primer-design/>

Bab 7

Konstruksi Pohon Filogenetika

Sub-CPMK (Kemampuan Akhir yang diharapkan)

- ✓ Mahasiswa mampu mengkonstruksi pohon filogenetik



INFO BUAT KAMU!

Pohon filogenetik merupakan diagram yang menggambarkan hubungan evolusi diantara organisme

<https://www.khanacademy.org/science/high-school-biology/hs-evolution/hs-phylogeny/a/phylogenetic-trees>

GLOSARIUM

Accession Keterangan lebih lanjut mengenai gen terpilih.

Bioinformatika Kesatuan dari ilmu biologi dan teknik informasi.

Blast Program yang menawarkan beberapa pilihan alignment berdasarkan *query* dan database target.

Database Arsip digital yang digunakan untuk menyimpan dan mengorganisasi data/informasi molekul biologi sehingga data tersebut dapat ditelusuri kembali dengan *query* tertentu.

Database primer Kumpulan data *original* berupa data urutan DNA/RNA dan protein.

Database sekunder Informasi data primer yang telah dianalisis dengan pendekatan tertentu.

DNA Polinukleotida dengan unit terkecil disebut nukleotida, tersusun atas gula pentosa, residu asam fosfat dan basa nitrogen.

Dogma sentral Keterkaitan molekul biologi DNA, RNA dan Protein.

Domain Sekuen urutan asam amino yang dilestarikan dan struktur tersier protein tertentu yang dapat berevolusi, berfungsi, dan ada secara independen dari rantai protein lainnya.

Ekson Urutan yang akan ditranslasikan.

Fenotip Karakteristik yang dapat diamati dipengaruhi oleh interaksi gen dengan lingkungannya

Filogenetika molekuler Studi tentang proses evolusi dan penggunaan filogeni data urutan DNA dan protein.

Filogenetika Studi tentang hubungan evolusi antara organisme dengan menggunakan pohon diagram sebagai representasi

Gen Sekuen nukleotida pada genome yang mengkode polipeptida tertentu

Genom Urutan sebuah gen total organisme.

Global alignment Dua urutan yang disejajarkan pada keseluruhan panjangnya.

Homologi sekuen Dua urutan diturunkan dari nenek moyang yang sama karena kesamaan urutan nukleotida/residu asam amino yang dibandingkan.

Intron Urutan yang tidak mengkodekan sifat tertentu.

KEGG Database yang memuat koleksi jalur proses tertentu dalam sel, merepresentasikan adanya interaksi molekuler yang bersifat kompleks.

Lokal alignment Dua urutan yang memiliki kesamaan pada daerah lokal tertentu.

Maximum likelihood Metode statistik pemilihan hipotesis berdasarkan nilai kemungkinan tertinggi yang sering digunakan dalam konstruksi pohon filogenetik molekuler.

Maximum parsimony Prinsip memilih solusi dengan penjelasan paling sedikit atau langkah logika. Dalam analisis filogenetik, metode parsimoni maksimum menyimpulkan pohon dengan langkah mutasi paling sedikit.

Melting Temperature Temperature atau suhu yang menyebabkan 50% untai ganda DNA terpisah.

Motif Daerah kecil dari struktur tiga dimensi protein atau urutan asam amino yang menentukan fungsi kimia atau biologis yang unik dari suatu protein.

Multiple sequence alignment (MSA) Penjajaran urutan berganda.

Ortolog sekuen Urutan sekuen homolog dari organisme atau genom berbeda yang diturunkan dari peristiwa spesiasi daripada peristiwa duplikasi gen.

Outgroup Takson atau sekelompok taksa dalam pohon filogenetik yang diketahui telah menyimpang lebih awal dari sisa taksa pada pohon filogeni dan digunakan untuk menentukan posisi akar.

Pairwise alignment (PSA) Penjajaran dua urutan DNA / residu asam amino protein.

Paralog sekuen Urutan sekuen homolog dari organisme atau genom yang sama yang berasal dari peristiwa duplikasi gen daripada peristiwa spesiasi

Polymerase Chain Reaction Teknik untuk membuat banyak salinan sekuen DNA tertentu secara *in vitro*

Primer Urutan asam nukleat pendek yang memberikan titik awal untuk sintesis DNA

Sekuen Urutan basa nukleotida pada DNA.

INDEKS

A

Accession 16, 31, 33
Amplifikasi 57, 58, 59, 62,
63, 65

B

Blast 26, 30, 31, 32, 60
Bioinformatika 3, 6, 7, 8, 11,
14, 18, 27, 49, 59

D

DNA 8, 9, 10, 15, 29, 56
Database 8, 14, 20, 27, 30
Dogma sentral 10
Domain 16, 17, 23, 47, 52,
53, 54, 55, 56, 58, 70, 82

E

Ekson 39, 40

F

Filogenetika 71, 72, 74, 75,
76, 77, 78
Fenotip 19,73

G

Gen
Genom 8, 29, 39, 71
Global alignment 28, 31

H

Homologi 41, 42

I

Intron 39, 40, 41

K

KEGG 18, 19

L

Lokal alignment 27, 70

M

Multiple sequence alignment
27, 34
Motif 41, 47, 52, 53, 54, 55,
58, 70

O

Ortolog 73

P

Pairwise alignment 27, 75
Paralog 73
Polimerase chain reaction 60,
61, 62, 63, 67
Primer 59, 60, 61, 62, 63, 64,
65, 66, 67

S

Sekuen 15, 39, 40, 41, 49, 60

BIOGRAFI PENULIS



ERLIA NARULITA. Penulis telah menempuh pendidikan dasar di SDN Jember Lor 5, sedangkan pendidikan menengah di SMPN 2 Jember dan SMAN 1 Jember. Penulis menyelesaikan studi strata satu di Pendidikan Biologi FKIP Universitas Jember lulus tahun 2002. Pendidikan strata dua di Pascasarjana Bioteknologi ITB lulus tahun 2011. Pendidikan doktor di *Department of Molecular Biotechnology Graduate School of Advance Science of Matter, Hiroshima University* (Jepang) lulus tahun 2016. Sejak tahun 2006 bekerja sebagai staf pengajar di program studi Pendidikan Biologi Jurusan PMIPA Fakultas Keguruan dan Ilmu Pendidikan Universitas Jember. Mengampu mata kuliah Bakteriofag, Fisiologi Hewan, Genetika, Biologi Sel, Bioinformatika DNA dan Protein dan Bioteknologi. Penulis merupakan peneliti dalam Divisi Bakteriologi dan Virologi, Laboratorium Kedokteran Molekuler, CDAST – Universitas Jember.



KUSWATI. Lahir di Jember, 08 Januari 1993 menempuh pendidikan dasar di SDN Sempolan 03 Jember, sekolah menengah di SMPN 1 Silo dan SMAN 2 Jember. Pendidikan Tinggi Program Sarjana ditempuh di PS. Pendidikan Biologi Universitas Jember pada tahun 2010 dan Pendidikan Magister di PS. Bioteknologi, ITB pada tahun 2016. Saat ini penulis merupakan staf pengajar di PS. Pendidikan Biologi, Jurusan Pendidikan MIPA, Fakultas Keguruan dan Ilmu Pendidikan sejak tahun 2019. Penulis memiliki fokus pada Bioteknologi tumbuhan, mengampu matakuliah Morfologi Tumbuhan, Fisiologi Tumbuhan, Biologi Dasar, Bioteknologi serta Bioinformatika DNA dan Protein. Penulis merupakan salah satu anggota dari Kelompok Riset (KeRis) PS. Pendidikan Biologi yaitu *Biological Engineering and Education* (BEE).