Pemakalah Utama 2

Proceeding Biology Education Conference
Volume 16, Nomor 1
Halaman 10 - 28

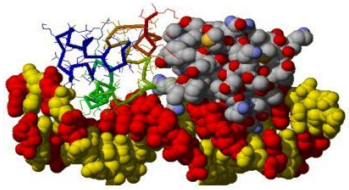
p-ISSN: 2528-5742

November 2019

PERAN BIOINFORMATIKA DALAM PENINGKATAN PENELITIAN DAN PEMBELAJARAN BIOLOGI DI ERA REVOLUSI INDUSTRI 4.0

Prof. Widodo, S.Si., M.Si., Ph.D.Med.Sc.
Guru Besar dari Universitas Brawijaya
E-mail: widodo@ub.ac.id

Peran Bioinformatika
dalam Peningkatan Penelitian dan Pembelajaran Biologi
di Era Revolusi Industri 4.0



Prof. Widodo, S.Si., M.Si., Ph.D.Med.Sc.
Guru Besar dari Universitas Brawijaya
widodo@ub.ac.id

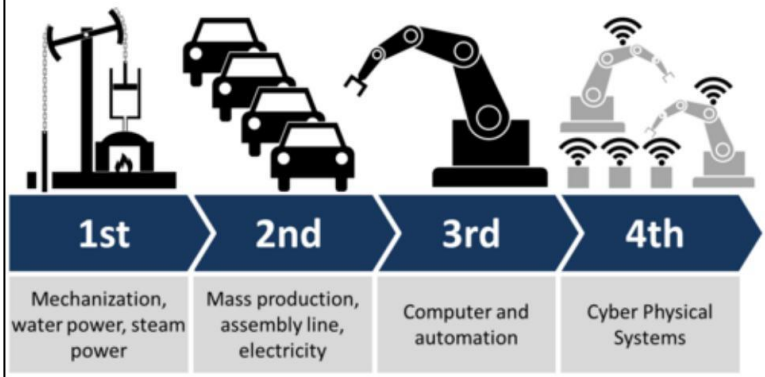
Latarbelakang

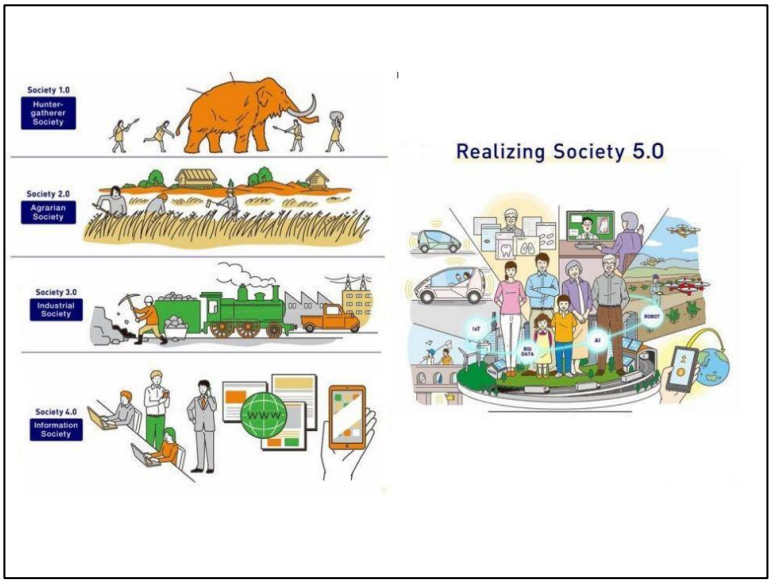
- Istilah "Industrie 4.0" berasal dari sebuah proyek dalam strategi teknologi canggih pemerintah Jerman yang mengutamakan komputerisasi pabrik.[5]
- Istilah "Industrie 4.0" diangkat kembali di Hannover Fair tahun 2011.[6]
- Pada Oktober 2012, Working Group on Industry 4.0 memaparkan rekomendasi pelaksanaan Industri 4.0 kepada pemerintah federal Jerman.
- Anggota kelompok kerja Industri 4.0 diakui sebagai bapak pendiri dan perintis Industri 4.0.
- Laporan akhir Working Group Industry 4.0 dipaparkan di Hannover Fair tanggal 8 April 2013.[7]

https://id.wikipedia.org/wiki/Industri_4.0

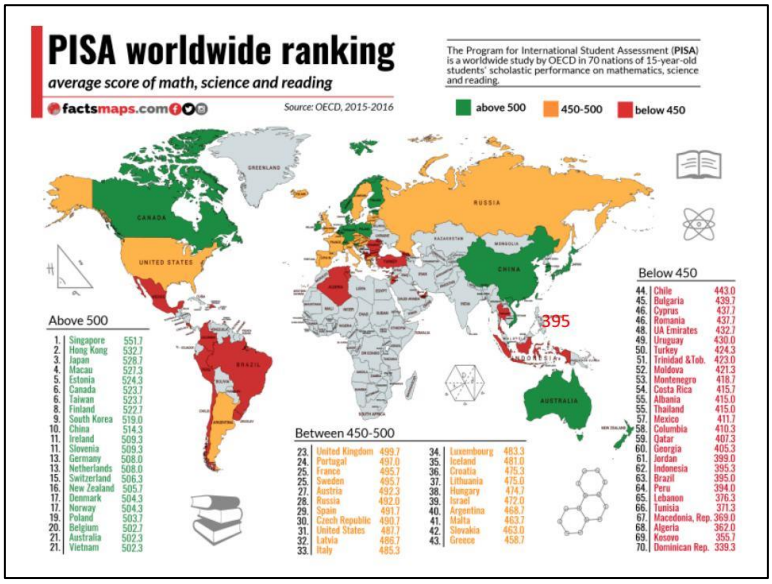
Prinsip Industri 4.0:

- **interconnection,**
- **information transparency,**
- **technical assistance,**
- **decentralized decisions.**






Watch the specia
Could this
future t
Scroll
What is Society 5.0
SCENE-01 Drones
SCENE-02 AI Home Appliances
SCENE-03 Medical Care/Nursing
SCENE-04 Smart Work
SCENE-05 Smart Management
SCENE-06 Autonomous vehicles


PISA worldwide ranking
average score of math, science and reading
factsmaps.com
Source: OECD, 2015-2016
The Program for International Student Assessment (PISA) is a worldwide study by OECD in 70 nations of 15-year-old students' scholastic performance on mathematics, science and reading.
above 500
450-500
below 450
Above 500
1. Singapore 551.7
2. Hong Kong 532.7
3. Japan 528.7
4. Macau 521.3
5. Estonia 524.3
6. Canada 523.7
6. Taiwan 523.7
8. Finland 522.7
9. South Korea 519.0
10. China 514.3
11. Ireland 509.3
11. Slovenia 509.3
13. Germany 508.0
13. Netherlands 508.0
15. Switzerland 506.3
16. New Zealand 505.7
17. Denmark 504.3
17. Norway 504.3
19. Poland 503.7
20. Belgium 502.7
21. Australia 502.3
21. Vietnam 502.3
Between 450-500
23. United Kingdom 499.7
24. Portugal 491.6
25. France 495.7
25. Sweden 495.7
27. Austria 492.3
28. Russia 492.0
29. Spain 491.7
30. Czech Republic 490.7
31. United States 487.7
32. Latvia 486.7
33. Italy 485.3
34. Luxembourg 483.3
35. Iceland 481.0
36. Croatia 475.3
37. Lithuania 475.0
38. Hungary 474.7
39. Israel 472.0
40. Argentina 468.7
41. Malta 463.7
42. Slovakia 463.0
43. Greece 458.7
Below 450
44. Chile 443.0
45. Bulgaria 439.7
46. Cyprus 437.7
46. Romania 437.7
48. UA Emirates 432.7
49. Uruguay 430.0
50. Turkey 424.3
51. Trinidad &Tob. 423.0
52. Moldova 421.3
53. Montenegro 418.7
54. Costa Rica 415.7
55. Albania 415.0
55. Thailand 415.0
57. Mexico 411.7
58. Columbia 410.3
59. Qatar 407.3
60. Georgia 405.3
61. Jordan 399.0
62. Indonesia 395.3
63. Brazil 395.0
64. Peru 394.0
65. Lebanon 376.3
66. Tunisia 371.3
67. Macedonia, Rep. 369.0
68. Algeria 362.0
69. Kosovo 355.7
70. Dominican Rep. 339.3
395


INDONESIA'S PISA SCORES COMPARED
TO OTHER ASEAN COUNTRIES
VIETNAM
2009 2012 2015
N/A 528 525
N/A 508 487
N/A 511 495
THAILAND
2009 2012 2015
425 444 421
421 441 409
419 427 415
INDONESIA
2009 2012 2015
383 382 403
402 396 397
371 375 386
SINGAPORE
2009 2012 2015
549 551 556
526 542 535
555 573 564
SCIENCE READING MATH

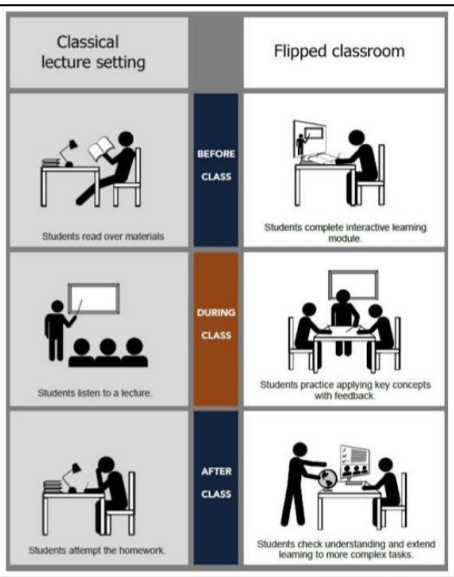
Tantangan yang dihadapi oleh Pendidikan dalam pengembangan kurikulum di era Revolusi Industri 4.0 dan society 5.0 adalah:

Menghasilkan lulusan yang memiliki kemampuan literasi baru meliputi

1. literasi data,
2. literasi teknologi
3. literasi manusia (HUMANISASI)
4. kemampuan memandang tanda- tanda akan terjadinya revolusi industri 5.0.

Capaian Pembelajaran Lulusan, CPL yang diperlukan di era industri 4.0 :

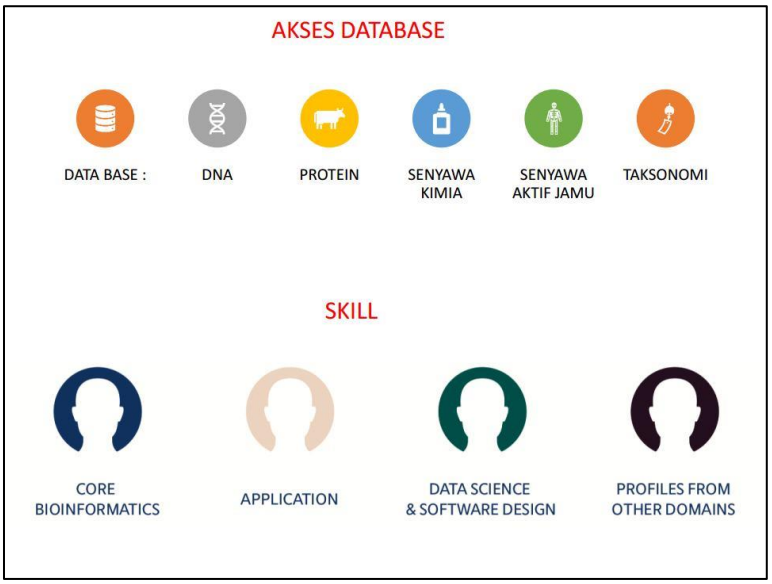
1. literasi data, kemampuan pemahaman untuk membaca, menganalisis, menggunakan data dan informasi (big data) di dunia digital;
2. literasi teknologi, kemampuan memahami cara kerja mesin, aplikasi teknologi (coding, artificial intelligance, dan engineering principle);
3. literasi manusia, kemampuan pemahaman tentang humanities, komunikasi dan desain; pemamahaman akan tanda-tanda revolusi industri 4.0; pemahaman ilmu untuk diamalkan bagi kemaslahatan bersama secara lokal, nasional, dan global.
4. kemampuan memandang tanda- tanda akan terjadinya revolusi industri 5.0. Revolusi industri 5.0 dapat dipahami sebagai pasar kolaborasi manusia dengan sistem cerdas yang berbasis pada internet of things (IoT) atau sistem fisik cyber, dengan kemampuan memanfaatkan mesin-mesin cerdas lebih efisien dengan lingkungan yang lebih bersinergi (Rada, 2017).



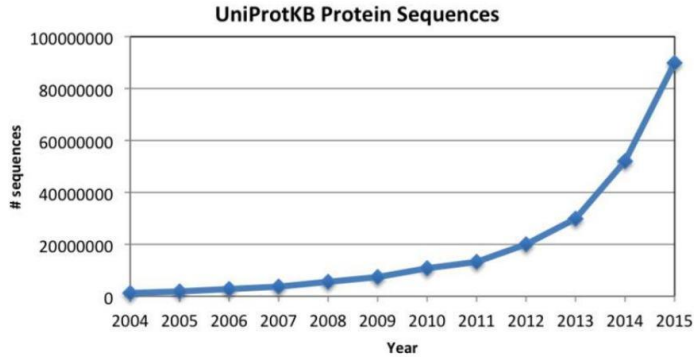
Dimana peran Bioinfromatic?

Manfaat Pendidikan Bioinformatik

1. Literasi data-BigDATA
2. Data Analyst
3. Literasi teknologi, coding & AI



The total number of protein sequences in UniProtKB (100 juta)



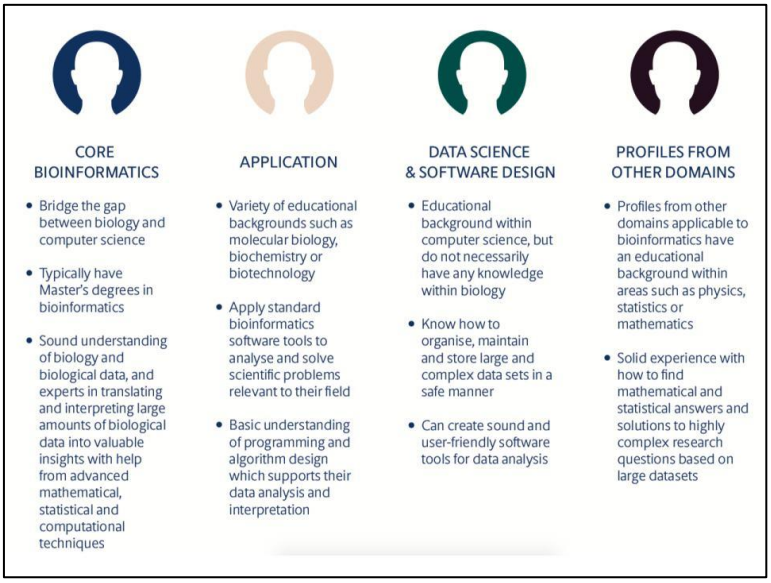
doi: 10.1007/978-1-4939-6783-4_1

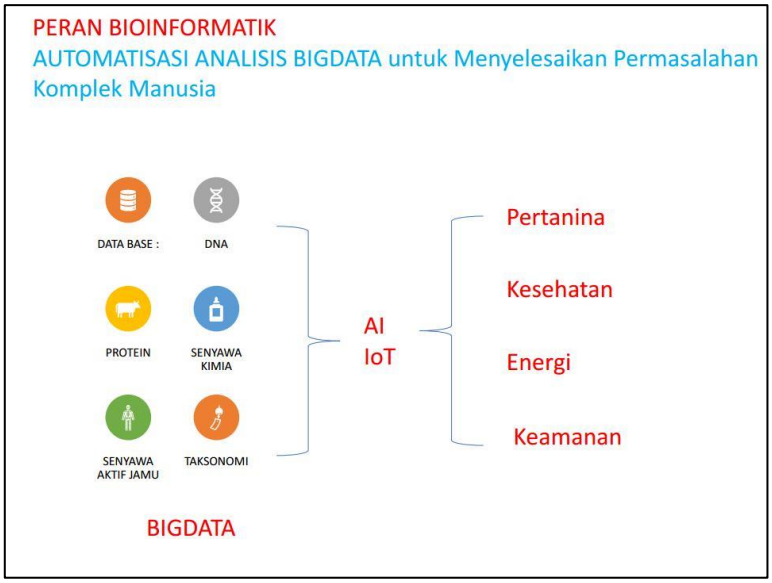
Overview of Protein Bioinformatics Databases

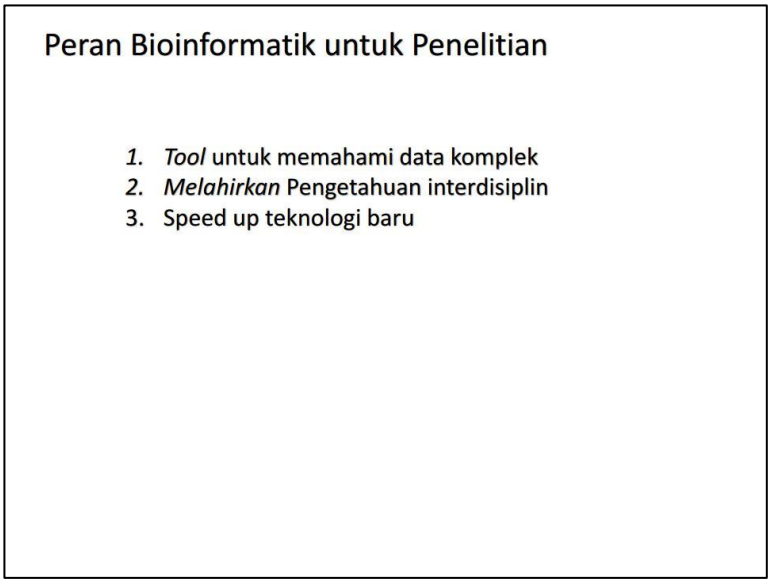
Category	DB Short Name	DB Name	URLs	Ref.
Sequence databases	CCDS	The Consensus CDS protein set database	https://www.ncbi.nlm.nih.gov/CCDS/CcdsBrowse.cgi	[9]
	DDBJ	DNA Data Bank of Japan	http://www.ddbj.nig.ac.jp/	[10]
	ENA	European Nucleotide Archive	http://www.ebi.ac.uk/ena	[11]
	GenBank	GenBank nucleotide sequence database	https://www.ncbi.nlm.nih.gov/genbank/	[12]
	RefSeq*	NCBI Reference Sequence Database	https://www.ncbi.nlm.nih.gov/refseq/	[13]
	UniGene	Database of computationally identifies transcripts from the same locus	https://www.ncbi.nlm.nih.gov/unigene/	[12]
	UniProtKB*	Universal Protein Resource (UniProt)	http://www.uniprot.org/	[14]
2D gel databases	COMPLUYEAST-2DPAGE	2-DE database at Universidad Complutense de Madrid, Spain	http://compluyeast2dpage.dacya.ucm.es/	[15]
	REPRODUCTION-2DPAGE	2-DE database at Nanjing Medical University, China	http://reprod.njmu.edu.cn/cgi-bin/2d/2d.cgi	[16]
	SWISS-2DPAGE	2-DE database at Swiss Institute of Bioinformatics, Switzerland	http://world-2dpage.expasy.org/swiss-2dpage/	[17]
	World-2DPAGE*	The World-2DPAGE database	http://world-2dpage.expasy.org/repository/	[18]
3D structure databases	DisProt	Database of Protein Disorder	http://www.disprot.org/	[19]
	MobiDB	Database of intrinsically disordered and mobile proteins	http://mobidb.bio.unipd.it/	[20]
	ModBase	Database of Comparative Protein Structure Models	http://modbase.compbio.ucsf.edu/modbase-cgi/index.cgi	[21]
	PDBe*	Protein Data Bank at Europe	http://www.ebi.ac.uk/pdbe/	[22]
	PDBj*	Protein Data Bank at Japan	http://pdbj.org/	[23]
	PDBsum	Pictorial database of 3D structures in the Protein Data Bank	http://www.ebi.ac.uk/pdbsum/	[24]
Chemistry databases	BindingDB	The Binding Datbase	http://www.bindingdb.org/	[28]
	ChEMBL*	Database of bioactive drug-like small molecules	https://www.ebi.ac.uk/chembldb	[29]
	DrugBank	Drug and Drug Target Databse	http://www.drugbank.ca/	[30]
Enzyme and pathway databases	MetaCyc/BioCyc*	MetaCyc Database of Metabolic Pathways, BioCyc Collection of Pathway/Genome Databases	http://www.biocyc.org/	[31]
	BRENDA*	BRaunschweig ENzyme DAtabase	http://www.brenda-enzymes.org	[32]
	ENZYME	Enzyme nomenclature database	http://enzyme.expasy.org/	[33]
	Reactome*	A knowledgebase of biological pathways and processes	http://www.reactome.org/	[34]
	SABIO-RK	SABIO-RK: Biochemical Reaction Kinetics Database	http://sabiork.h-its.org/	[35]
	SignaLink	A signaling pathway resource with multi-layered regulatory networks	http://signalink.org/	[36]
	UniPathway	UniPathway: a resource for the exploration of metabolic pathways	http://www.unipathway.org	[37]
Family and domain databases	Gene3D	Structural and Functional Annotation of Protein Families	http://gene3d.biochem.ucl.ac.uk/Gene3D/	[38]
	HAMAP	High-quality Automated and Manual Annotation of Proteins	http://hamap.expasy.org/	[39]
	InterPro*	Integrated resource of protein families, domains and functional sites	http://www.ebi.ac.uk/interpro/	[40]
	PANTHER	The PANTHER Classification System	http://www.pantherdb.org/	[41]
	Pfam*	The Pfam protein families database	http://pfam.xfam.org/	[42]
	PIRSF*	A whole-protein classification database	http://pir.georgetown.edu/pirwww/dbinfo/pirsf.shtml	[43]
	PRINTS	Protein Motif fingerprint database	http://www.bioinf.manchester.ac.uk/dbbrowser/PRINTS/	[44]
	ProDom	Protein domain families database	http://prodom.prabi.fr/prodom/current/html/home.php	[45]

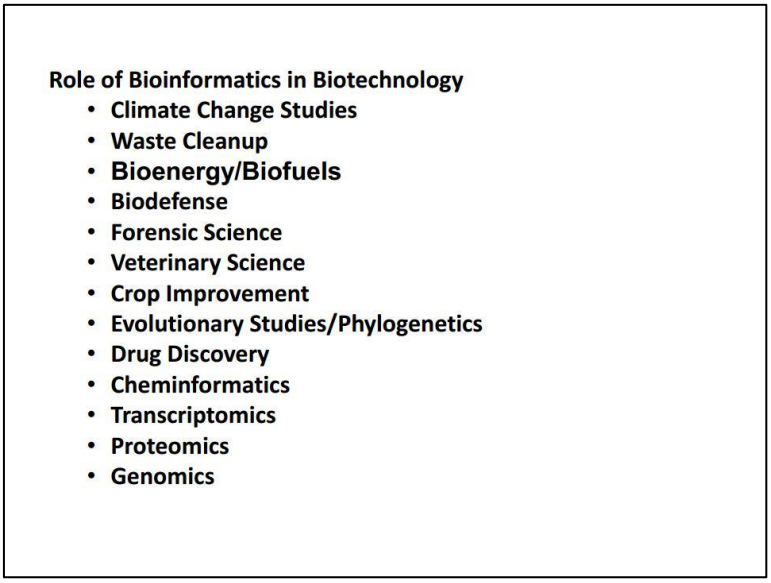
Gene expression databases	Bgee	Database for Gene Expression Evolution	http://bgee.unil.ch	[51]
	CleanEx	Database of gene expression profiles	http://cleanex.vital-it.ch/	[52]
	Genevisible	Search portal to normalized and curated expression data from Genevestigator	http://genevisible.com/search	[53]
	ExpressionAtlas*	Database of Differential and Baseline Expression	http://www.ebi.ac.uk/gxa/home	[54]
Genome annotation databases	Ensembl*	Ensembl Eukaryotic genome annotation database	http://www.ensembl.org/	[55]
	EnsemblBacteria	Ensembl Bacteria genome annotation database	http://bacteria.ensembl.org/	[56]
	EnsemblFungi	Ensembl Fungi genome annotation database	http://fungi.ensembl.org/	[56]
	EnsemblMetazoa	Ensembl Metazoa genome annotation database	http://metazoa.ensembl.org/	[56]
	EnsemblPlants	Ensembl Plants genome annotation database	http://plants.ensembl.org/	[56]
	EnsemblProtists	Ensembl Protists genome annotation database	http://protists.ensembl.org/	[56]
	Entrez Gene*	Database of Genes of Genomes in the Reference Sequence Collection	https://www.ncbi.nlm.nih.gov/gene	[57]
	KEGG	Kyoto Encyclopedia of Genes and Genomes	http://www.genome.jp/kegg/	[58]
	PATRIC	Bacterial Bioinformatics Resource Center	http://patricbrc.org/	[59]
	UCSC*	UCSC Genome Bioinformatics	http://genome.ucsc.edu	[60]
	VectorBase	Bioinformatics Resource for Invertebrate Vectors of Human Pathogens	http://www.vectorbase.org/	[61]
	WBParaSite	WormBase ParaSite	http://parasite.wormbase.org	[62]
Organism specific databases	ArachnoServer	ArachnoServer: Spider toxin database	http://www.arachnoserver.org	[63]
	CGD	Candida Genome Database	http://www.candidagenome.org/	[64]
	ConoServer	ConoServer: Cone snail toxin database	http://www.conoserver.org/	[65]
	CTD	Comparative Toxicogenomics Database	http://ctdbase.org/	[66]
	dictyBase	Central resource for Dictyostelid genomics	http://dictybase.org/	[67]
	EchoBASE	EchoBASE – an integrated post-genomic database for E. coli.	http://www.york.ac.uk/res/thomas/	[68]

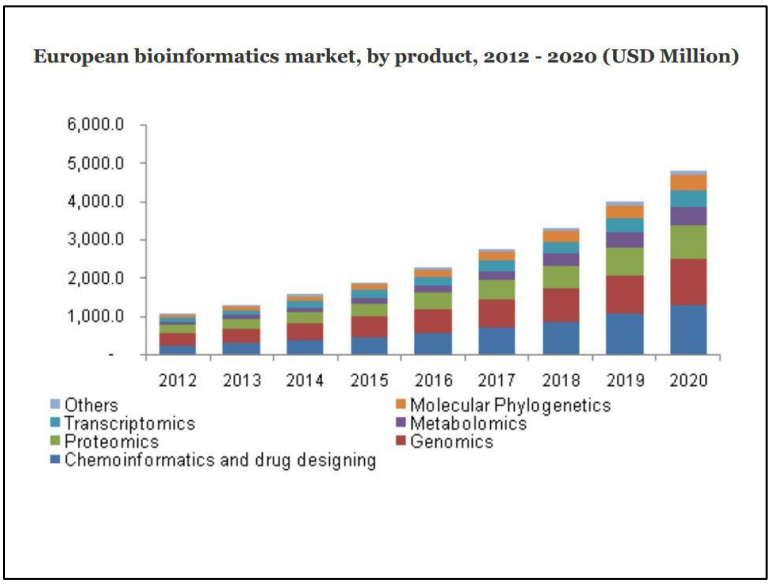
Phylogenomic databases	eggNOG	Database of orthologous groups and functional annotation	http://eggnog.embl.de/	[98]
	HOGENOM	Database of Homologous Genes from Fully Sequenced Organisms	http://pbil.univ-lyon1.fr/databases/hogenom/home.php	[99]
	HOVERGEN	Homologous Vertebrate Genes Database	http://pbil.univ-lyon1.fr/databases/hovergen.html	[100]
	InParanoid	Eukaryotic Ortholog Groups with inparalogs	http://inparanoid.sbc.su.se/	[101]
	KO	Kyoto Encyclopedia of Genes and Genomes Orthology	http://www.genome.jp/kegg/	[102]
	OMA*	The OMA orthology database	http://omabrowser.org/	[103]
	OrthoDB	Database of Orthologous Groups	http://cegg.unige.ch/orthodb6	[104]
	PhylomeDB	Database for complete catalogs of gene phylogenies (phylomes)	http://phylomedb.org/	[105]
	TreeFam	Database of animal gene trees	http://www.treefam.org	[106]
Polymorphism and mutation databases	BioMuta	Single-nucleotide variation and disease association database	https://hive.biochemistry.gwu.edu/tools/biomuta/	[107]
	dbSNP*	Database of Short Genetic Variations	https://www.ncbi.nlm.nih.gov/SNP/	[12]
	DMDM	Domain Mapping of Disease Mutations	http://bioinf.umbc.edu/dmdm/	[108]
Protein-protein interaction databases	BioGRID	The Biological General Repository for Interaction Datasets	http://thebiogrid.org	[109]
	DIP	Database of Interacting Proteins	http://dip.doe-mbi.ucla.edu/	[110]
	IntAct*	IntAct Molecular Interaction Database	http://www.ebi.ac.uk/intact/	[111]
	MINT	The Molecular INTeraction database	http://mint.bio.uniroma2.it/mint/	[112]
	STRING	Search Tool for the Retrieval of Interacting Genes/Proteins	http://string-db.org	[113]
Proteomic databases	MaxQB	The MaxQuant DataBase	http://maxqb.biochem.mpg.de/mxdb/	[114]
	PaxDb	Protein Abundance Across Organisms	http://pax-db.org	[115]



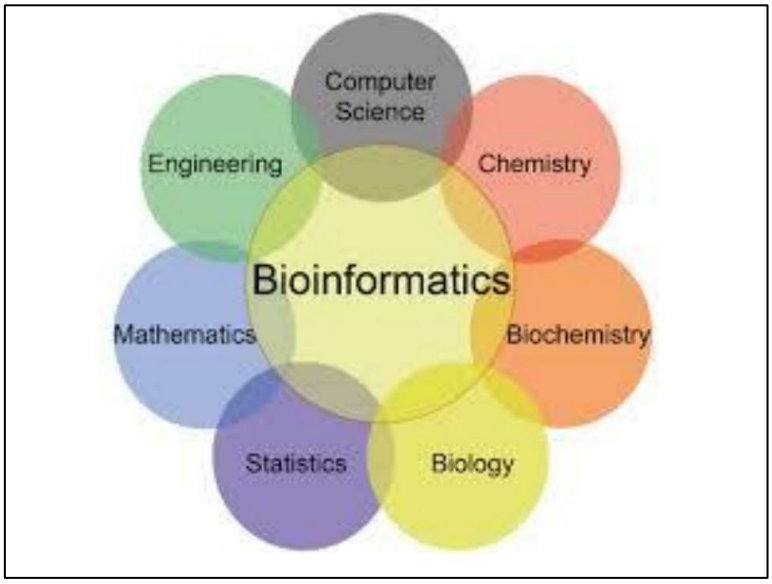

PERAN BIOINFORMATIK
AUTOMATISASI ANALISIS BIGDATA untuk Menyelesaikan Permasalahan Komplek Manusia
DATA BASE :
DNA
PROTEIN
SENYAWA KIMIA
SENYAWA AKTIF JAMU
TAKSONOMI
BIGDATA
AI
IoT
Pertanina
Kesehatan
Energi
Keamanan


Peran Bioinformatik untuk Penelitian
1. *Tool* untuk memahami data komplek
2. *Melahirkan* Pengetahuan interdisiplin
3. Speed up teknologi baru

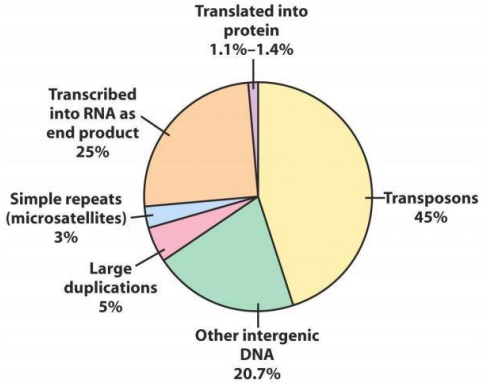

Role of Bioinformatics in Biotechnology
• Climate Change Studies
• Waste Cleanup
• Bioenergy/Biofuels
• Biodefense
• Forensic Science
• Veterinary Science
• Crop Improvement
• Evolutionary Studies/Phylogenetics
• Drug Discovery
• Cheminformatics
• Transcriptomics
• Proteomics
• Genomics



APA itu Bioinformatik ?



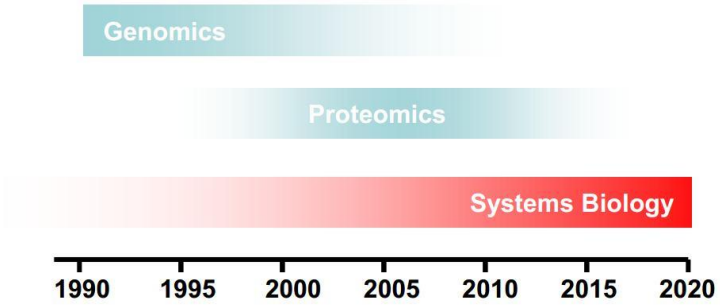
Snapshot of the human genome



- Only 1.1% to 1.4% of human genome DNA actually encodes proteins.
- More than 50% of genome consists of short, repeated sequences,
- The vast majority of which—about 45% of genome in all—come from transposons, short movable DNA sequences that are molecular parasites.

27

Genomics, Proteomics & Systems Biology



What is Systems Biology?

- Systems Biology - Systems biology is the study of an organism, viewed as an *integrated* and *interacting network* of genes, proteins and biochemical reactions. The integration of genomics, proteomics, metabolomics & modeling

- The Goal: *Predictive, Preventative and Personalized Medicine , precision agryculture*

Pre-Genomic

- Reductionist (DNA or RNA or Protein)
- Observational/ phenomenological
- Generally qualitative and non-numeric
- Hypothesis driven

Organism ⇩ Cellular ⇩ Molecular

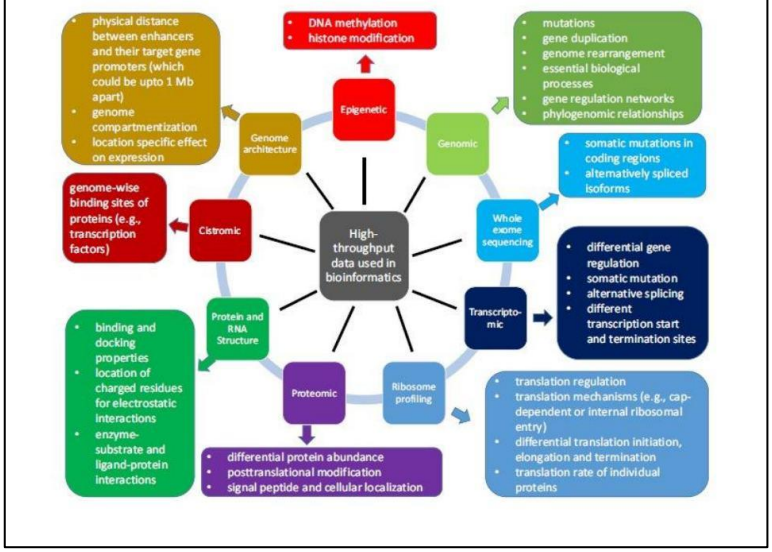
Post-Genomic

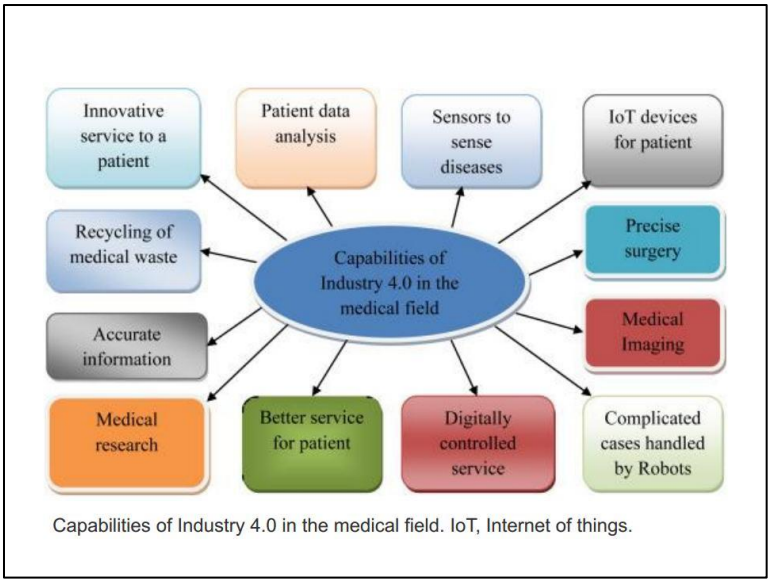
- Global/ holistic
- Systems approach (DNA and RNA and Protein)
- Quantitative and highly numeric
- Not only hypothesis driven but also data driven

Molecular ⇩ Cellular ⇩ Organism

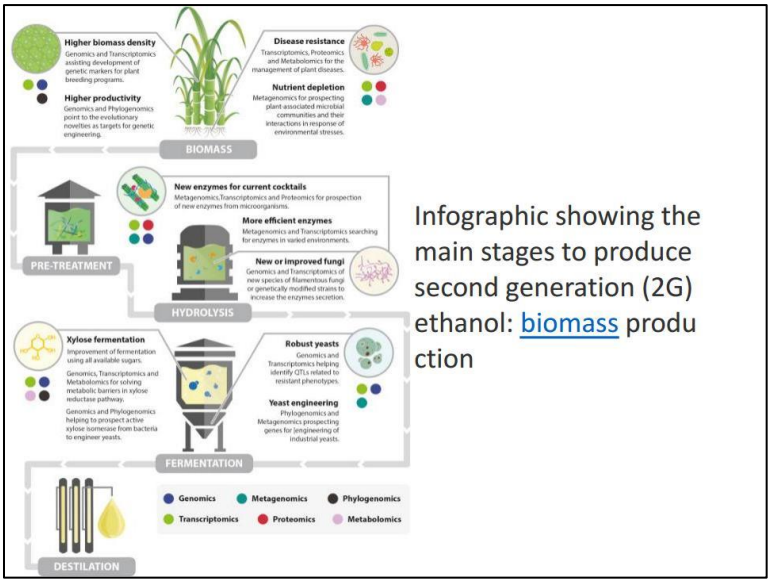
Applications of Bioinformatics

- Drug design
- Gene therapy
- Sequence analysis of proteins and DNA-Disease
- Molecular structure related to Disease
- Analysis and simulation of metabolic networks
- SNP & Disease
- Pathway Analysis (Pathomechanism)
- Diagnostic
- Prognostic
- Vaccine Design
- Molecular Mechanis

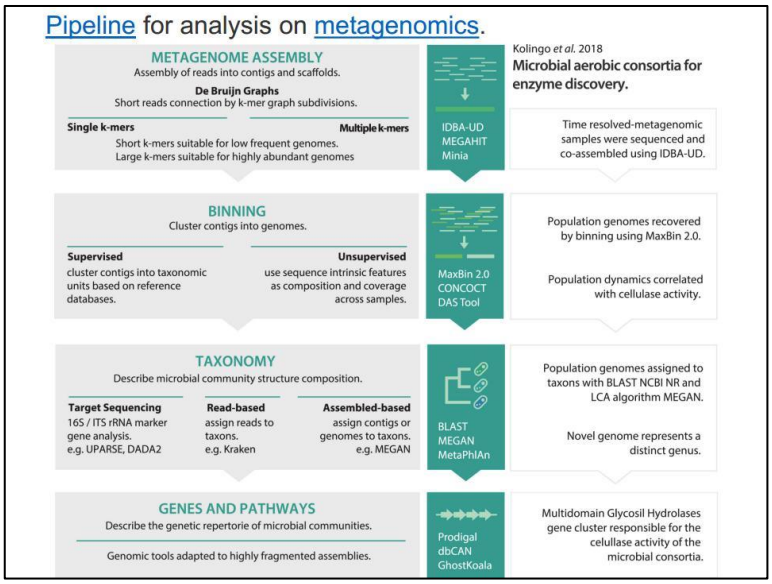


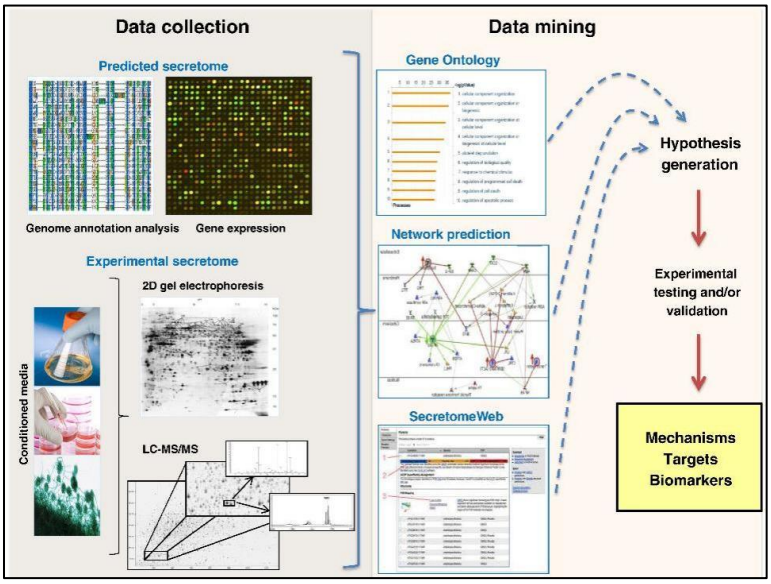


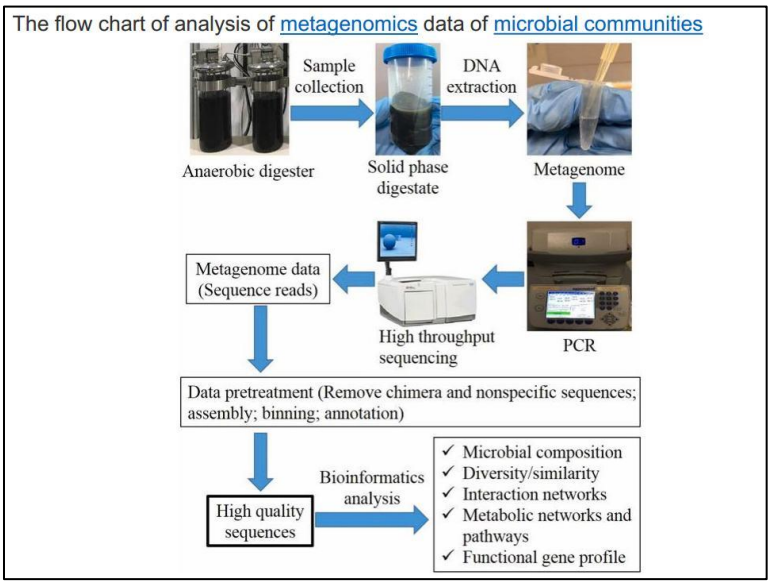
Capabilities of Industry 4.0 in the medical field. IoT, Internet of things.

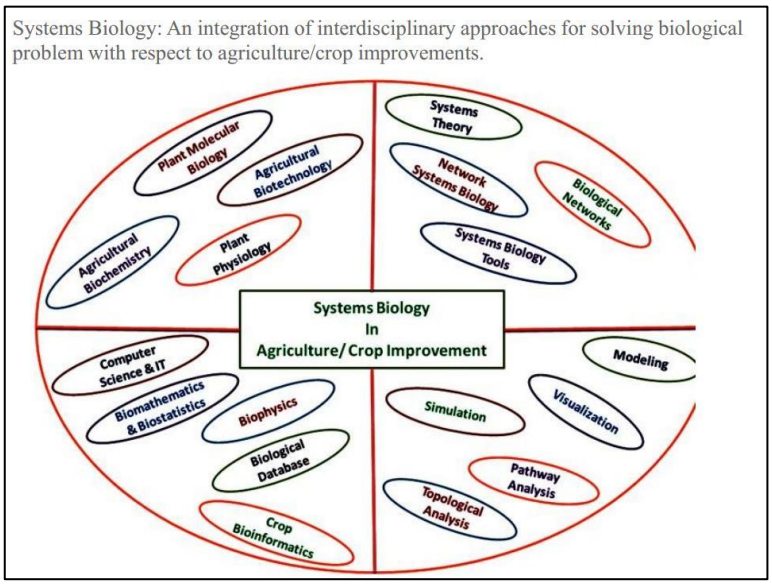


Infographic showing the main stages to produce second generation (2G) ethanol: biomass production

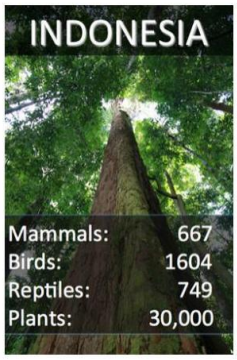



Data collection
Data mining
Predicted secretome
Genome annotation analysis
Gene expression
Experimental secretome
2D gel electrophoresis
Conditioned media
LC-MS/MS
Gene Ontology
Network prediction
SecretomeWeb
Hypothesis generation
Experimental testing and/or validation
Mechanisms
Targets
Biomarkers


The flow chart of analysis of metagenomics data of microbial communities
Sample collection
DNA extraction
Anaerobic digester
Solid phase digestate
Metagenome
Metagenome data (Sequence reads)
High throughput sequencing
PCR
Data pretreatment (Remove chimera and nonspecific sequences; assembly; binning; annotation)
Bioinformatics analysis
High quality sequences
✓ Microbial composition
✓ Diversity/similarity
✓ Interaction networks
✓ Metabolic networks and pathways
✓ Functional gene profile

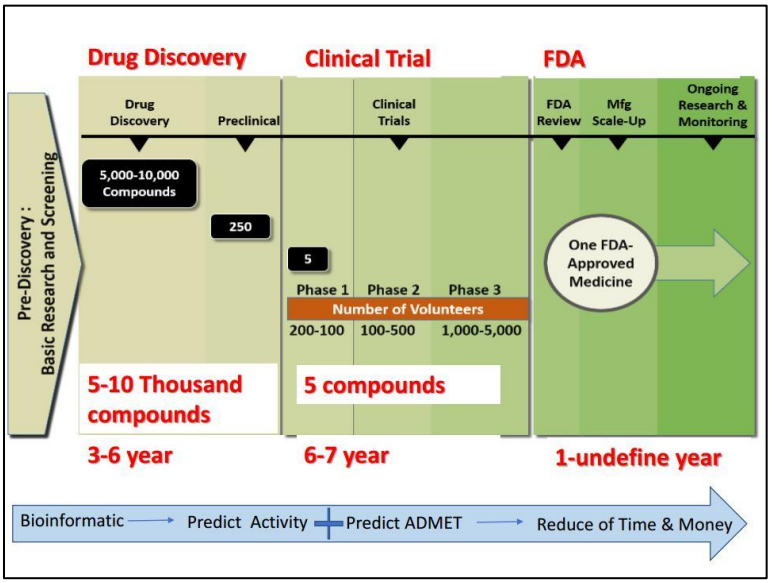

Systems Biology: An integration of interdisciplinary approaches for solving biological problem with respect to agriculture/crop improvements.
Systems Biology In Agriculture/ Crop Improvement
Plant Molecular Biology
Agricultural Biotechnology
Agricultural Biochemistry
Plant Physiology
Systems Theory
Network Systems Biology
Biological Networks
Systems Biology Tools
Computer Science & IT
Biomathematics & Biostatistics
Biophysics
Biological Database
Crop Bioinformatics
Modeling
Simulation
Visualization
Pathway Analysis
Topological Analysis

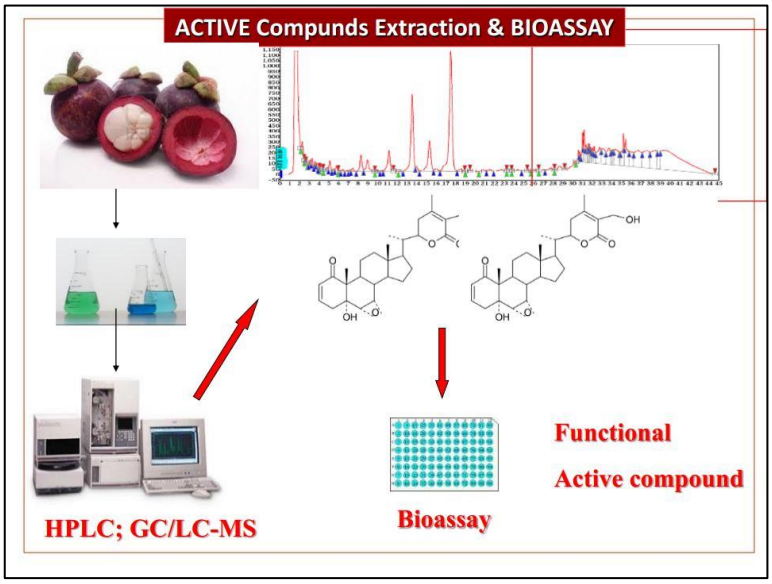
CASE STUDY

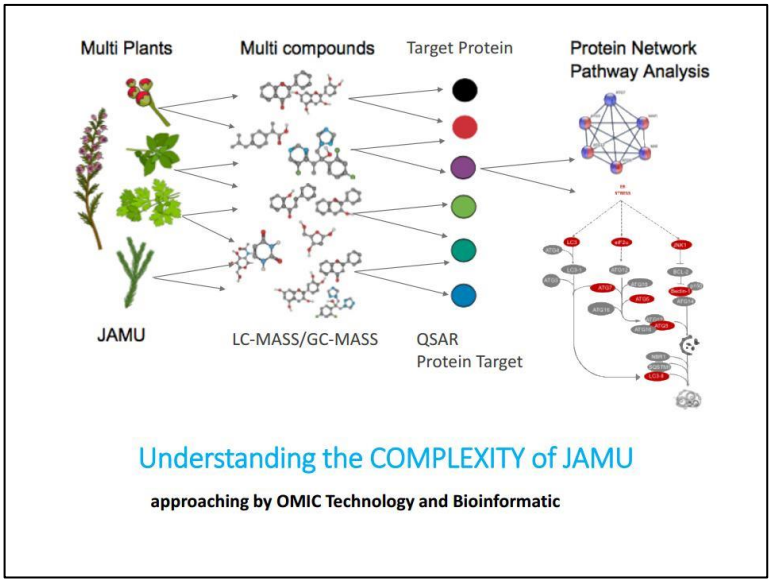


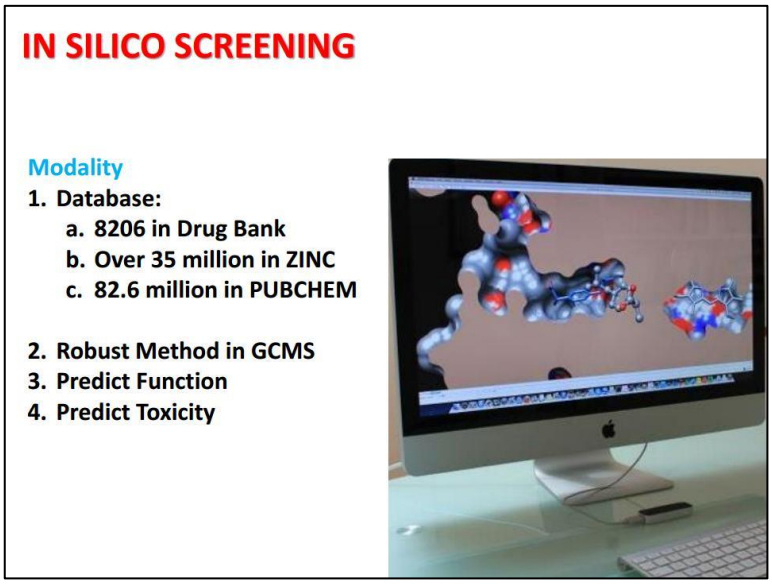
HUGE POTENCY :
RICHEST BIODIVERSITY

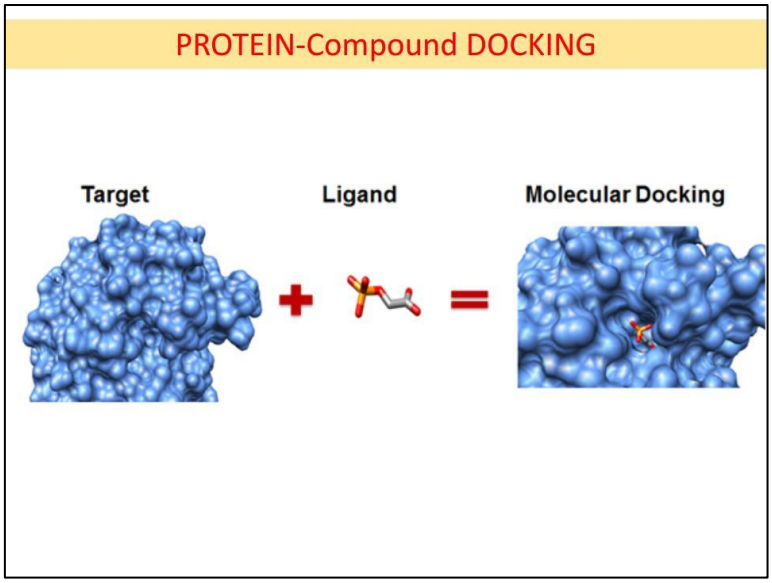
7.000 plants have Medicinal activity
(Erdelen et al.,1999)

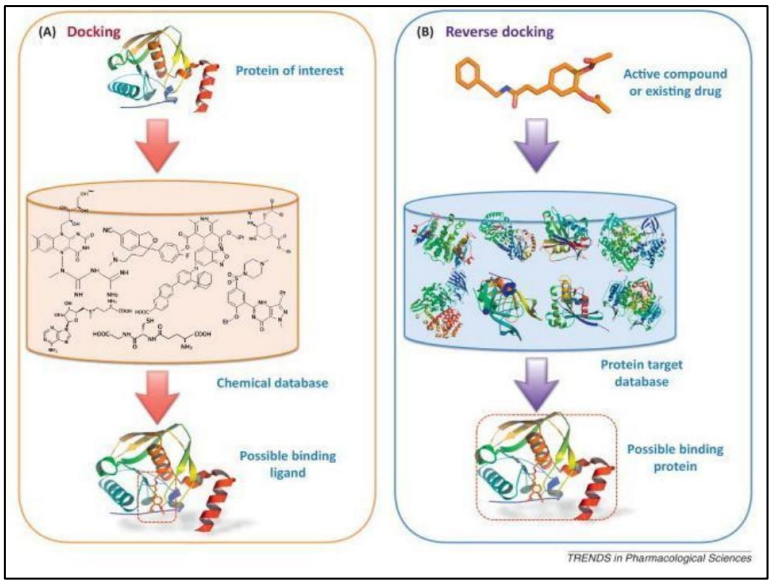


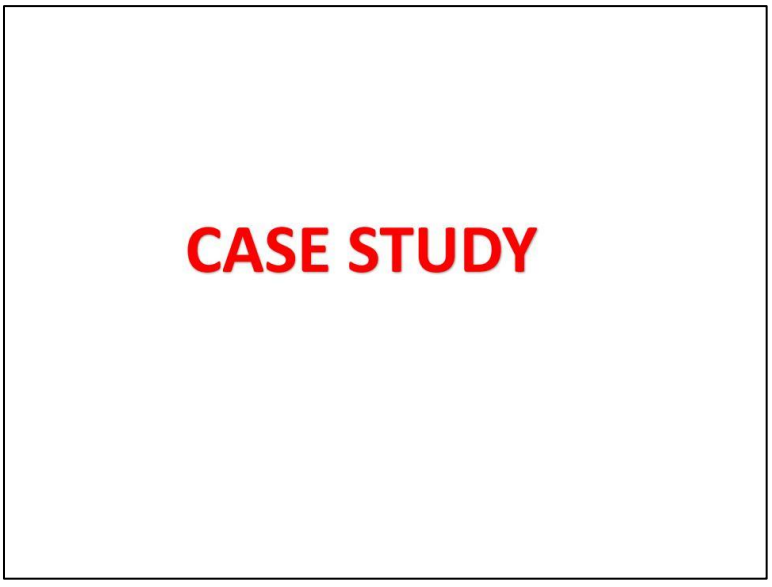

ACTIVE Compunds Extraction & BIOASSAY
HPLC; GC/LC-MS
Bioassay
Functional
Active compound

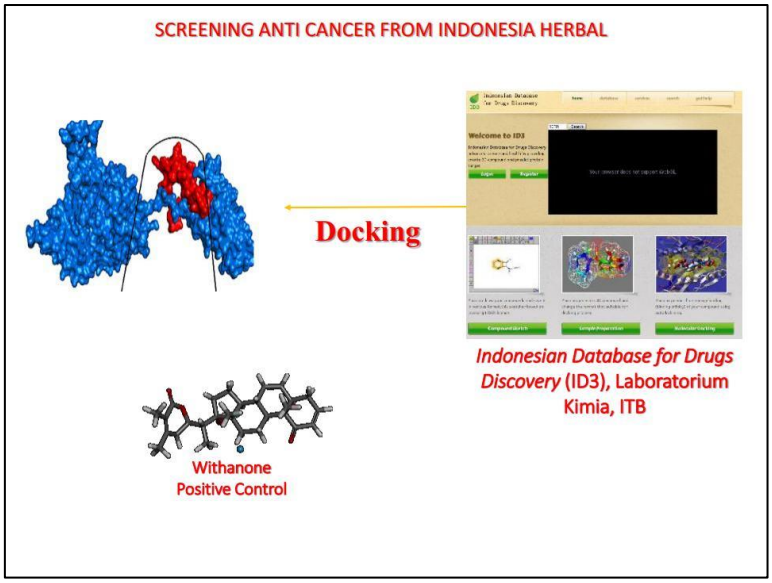

Multi Plants
Multi compounds
Target Protein
Protein Network
Pathway Analysis
JAMU
LC-MASS/GC-MASS
QSAR
Protein Target
Understanding the COMPLEXITY of JAMU
approaching by OMIC Technology and Bioinformatic

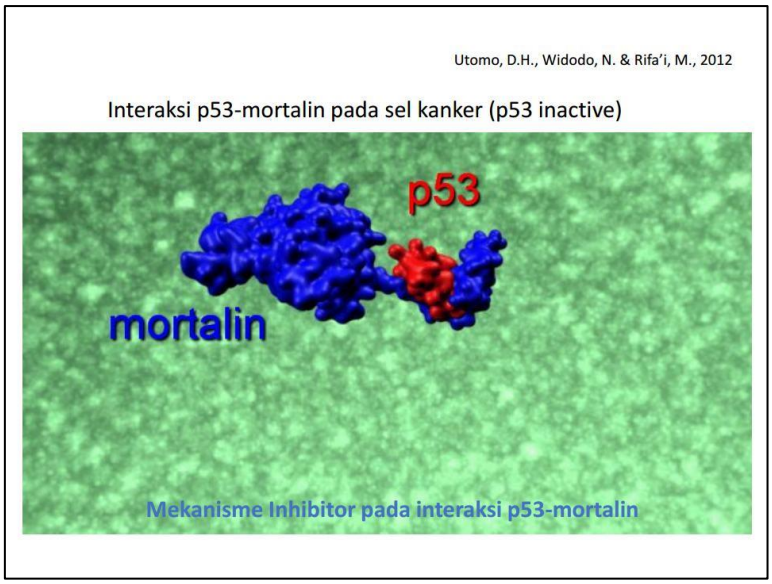

IN SILICO SCREENING
Modality
1. Database:
a. 8206 in Drug Bank
b. Over 35 million in ZINC
c. 82.6 million in PUBCHEM
2. Robust Method in GCMS
3. Predict Function
4. Predict Toxicity

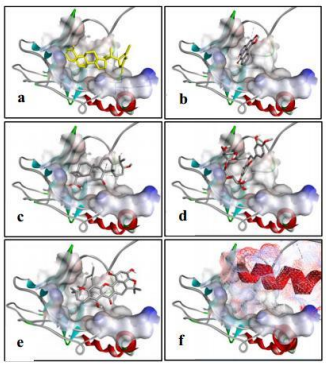

PROTEIN-Compound DOCKING
Target
Ligand
Molecular Docking


(A) Docking
Protein of interest
Chemical database
Possible binding ligand
(B) Reverse docking
Active compound or existing drug
Protein target database
Possible binding protein
TRENDS in Pharmacological Sciences


CASE STUDY







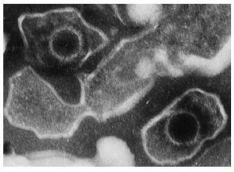
Gambar 1. Pola interkasi natara senyawa aktif (dala bentuk stik) dengan mortalin (kartun). Senyawa withanoı (**a**), Liriodenine (**b**), glycyrrhetinic acid (**c**), Hesperid (**d**), dan Artoindonesianin A (**e**) menempel di daer pengikat p53 (*p53-binding domain*) pada protein mortal (**f**), sehingga dimungkinkan dapat menghambat interak antara p53 dengan mortalin.

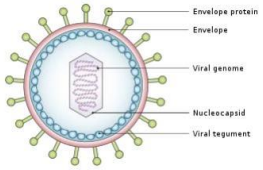
Tabel 2. Kemampuan pengikatan antara senyawa-senyawa aktif dengan protein mortalin. Empat senaywa aktif dari tanaman Indonesia memiliki binding affinity yang lebih baik di banding control positif, Withanone.

Senyawa Aktif	Binding Affinity (kcal/mol)
Hesperidin	-8.7
Artoindonesianin A	-8.6
Glycyrrhetinic Acid	-8.2
Liriodenine	-8.2
Withanone	-8.0
Gossypol	-7.8
Aloesin	-7.5
Eleagic Acid	-7.4
Rutin	-7.4
Diosbulbin B	-7.3
Betulinic Acid	-7.3
Geniposide	-7.0
Allamandin	-6.7
Resveratrol	-5.9
Anethole	-5.3
Eugenol	-5.3
Allicin	-4.0

CASE STUDY

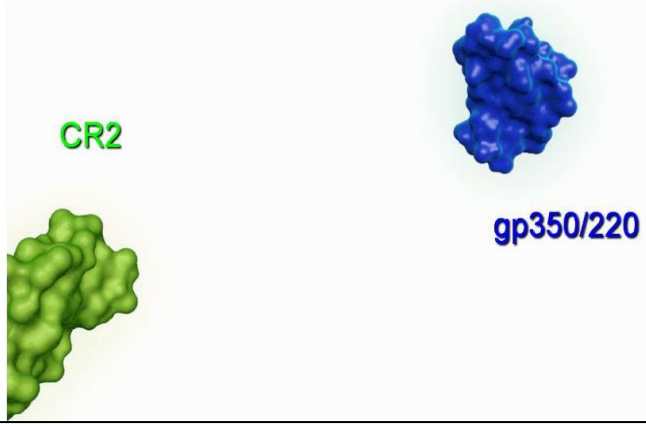
Introduction

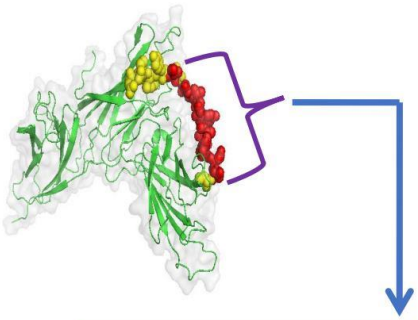




- Epstein-Barr virus (EBV) usually infects children and commonly does not show direct symptoms after infection.
- This virus infected many people, especially in the Asia region, including Indonesia.
- EBV has a very vital role in nasopharyngeal carcinoma (NPC) progression.
- One of the efforts to prevent the NPC is vaccination.

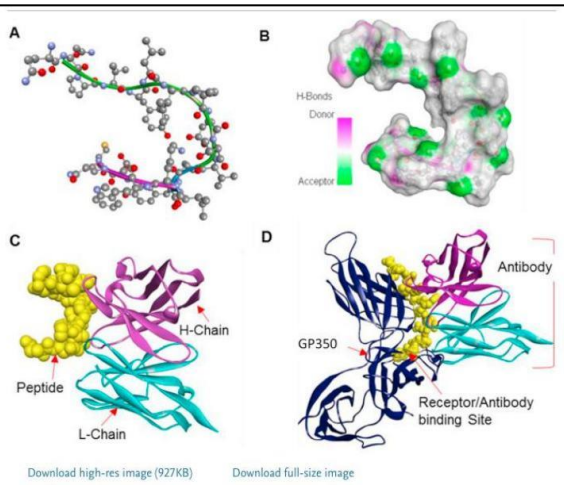
Protein gp350/220 is capsid of EBV bind into CR2, during infection process





Epitope, potential for vaccine

We have identified epitope of Gp350/220 protein that potential for the vaccine of EBV (Sitompul et al., 2012).



Download high-res image (927KB) Download full-size image

Fig. 1. The tertiary structure of EBVepitope (A) showed a coil structure with many hydrogen donors/acceptors (B). The epitope could bind 72A1 mAb in both the heavy- (purple) and light-chain (cyan) regions (C), which was similar to the binding pattern between the intact GP350/220 protein and 72A1 mAb (D). (For interpretation of the

TERIMAKASIH