

計算生物學原理與應用

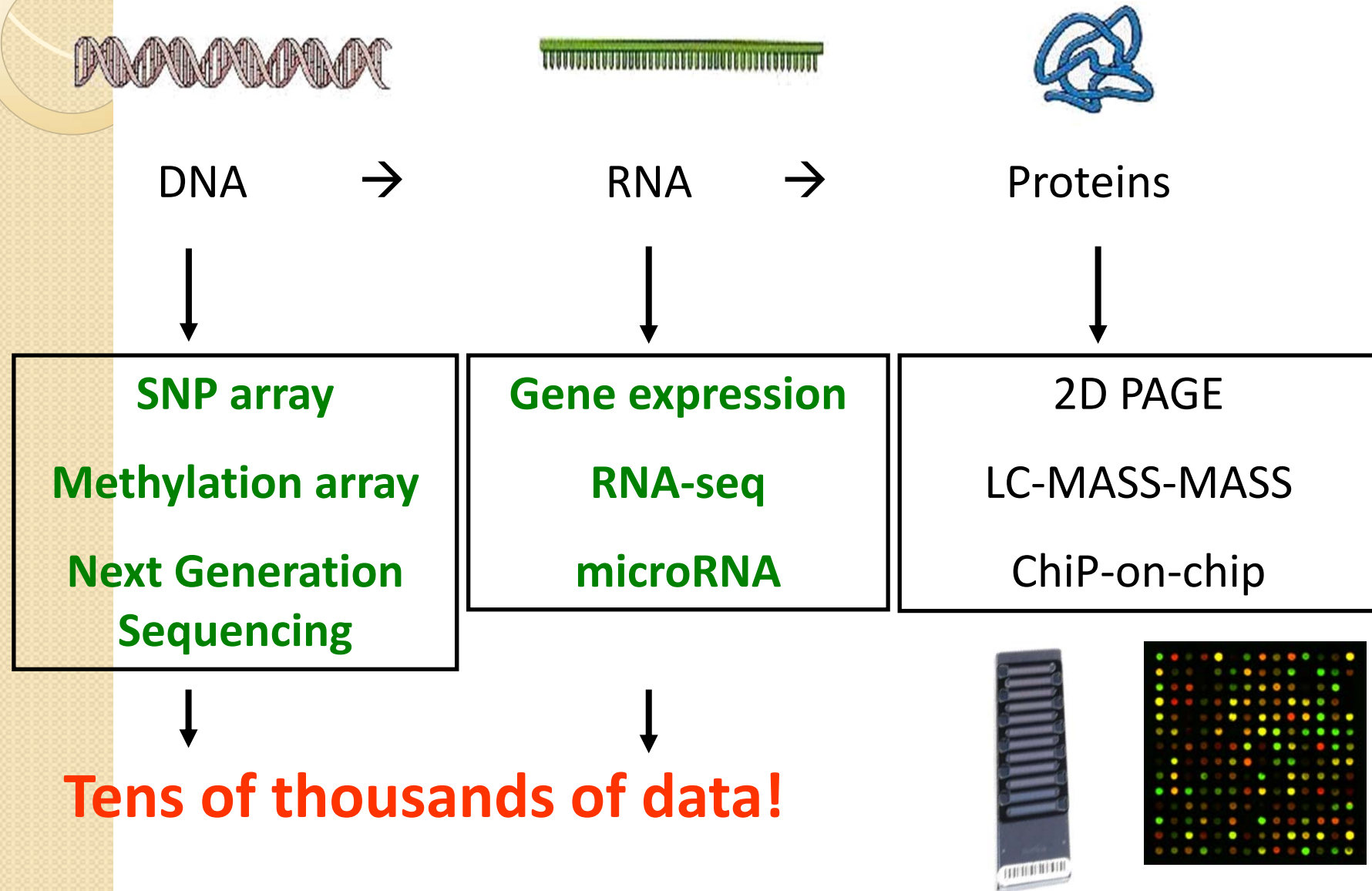


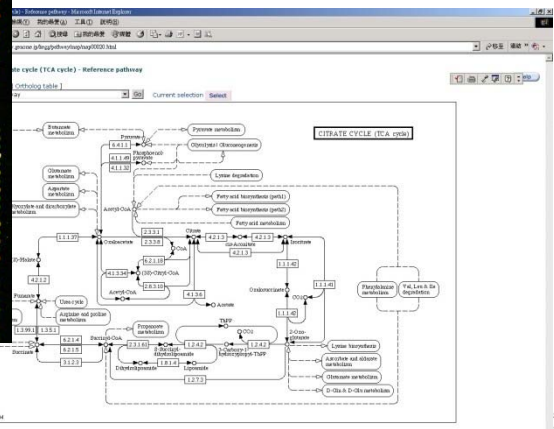


Computational Biology (wiki)

- Computational biology involves the development and application of data-analytical and theoretical methods, mathematical modeling and computational simulation techniques to the study of biological, behavioral, and social systems. The field is broadly defined and includes foundations in computer science, applied mathematics, animation, statistics, biochemistry, chemistry, biophysics, molecular biology, genetics, genomics, ecology, evolution, anatomy, neuroscience, and visualization.
- Computational biology is different from biological computation, which is a subfield of computer science and computer engineering using bioengineering and biology to build computers, but is similar to bioinformatics, which is an interdisciplinary science using computers to store and process biological data.

Applied Genomics to Study Disease Mechanisms





Tools

- BioCarta
- Pubmed
- CellSignaling
- Search

Related Links

- CellSignaling
- CellSignaling
- CellSignaling

OutLinks:

- CellSignaling
- CellSignaling
- CellSignaling

CellSignaling

bone:	3.40	8.70580
esophagus:	3.24	2.18717
thyroid:	2.98	1.10200
uncharacterized tissue:	2.70	19.213694
bone marrow:	2.20	3.41345

Go to [Gene-to-tag Mapping](#) at NCBI

sequence

[Upstream genomic sequence](#) for CDC6 cell division cycle 6 homolog (S cerevisiae)

Sequences

NM_001254	
Accession	Description
NM_001254	NA

Search PubMed using aliases AND

[illegible]

NCBI Sequence Viewer 2.0 - Microsoft Internet Explorer

File Edit View Favorites Tools Help 工 具 栏 帮助

地址: 编辑 地址 我的收藏夹 工具 帮助 说明

网上下一系列按钮

地址: <http://www.ncbi.nlm.nih.gov/blast/blast.cgi?seq1=ncbi-blast&seq2=6678719>

NCBI Nucleotide

Search Nucleotide for Go Clear

Limits Preview Index History Clipboard Details

Display GenBank Show 5 Send to

Range from begin to end Reverse complemented strand Features: SNF STS R

1. NM_008512 Reports Mus musculus low... [6678719]

Comments Features Sequence

LOCUS NM_008512 14649 bp mRNA linear ROD 12-JUN-2006

DEFINITION Mus musculus low density lipoprotein receptor-related protein 1 (Lrp1), mRNA.

ACCESSION NM_008512

VERSION NM_008512.1 GI:6678719

KEYWORDS .

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Top ten [of 40]

[Help]

	name:	2.42	8.10300
	esophagus:	3.24	218717
	thyroid:	2.98	110200
	uncharacterized tissue:	2.70	19213694
	bone marrow:	2.20	341345

SAGE (NCBI)

Go to [Gene-to-tag Mapping](#) at NCBI

Upstream Genomic Sequence

TRASER

[Upstream genomic sequence](#) for CDC6 cell division cycle 6 homolog (S. cerevisiae)

Representative mRNA Sequences

UniGene

[NM_001254](#)

LocusLink RefSeq

Accession	Description
NM_001254	NA

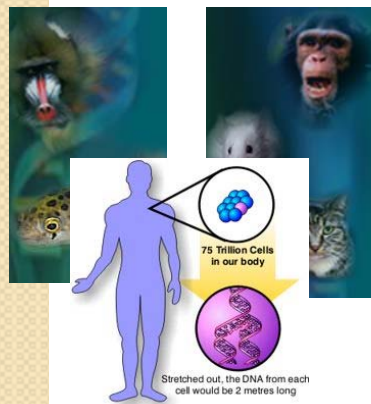
Alias PubMed Search

PubMed

Search PubMed using aliases AND

What is Computational Biology? (Bioinformatics)

Computational Biology = **Biology** + **Informatics** + Statistics



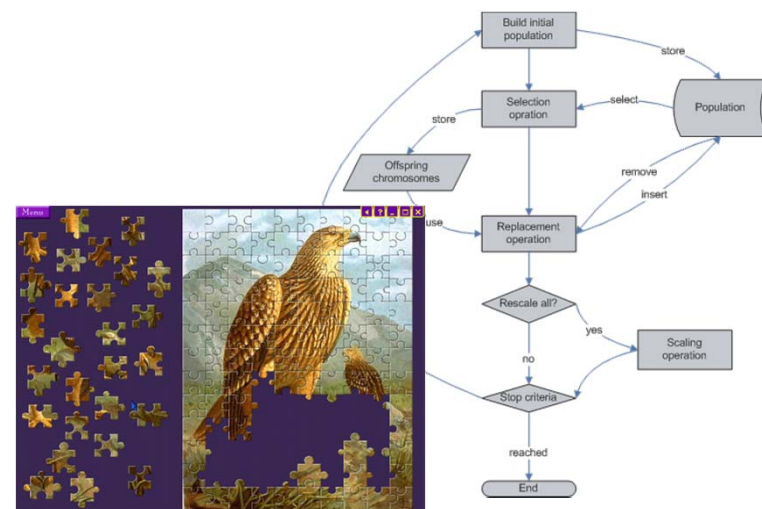
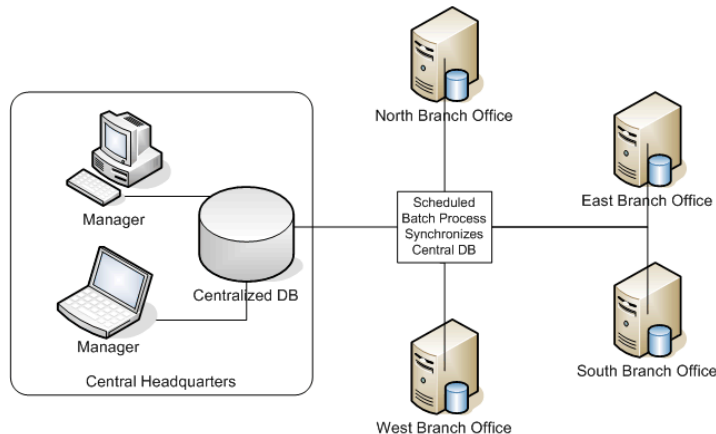
Bioinformatics

Biology

Informatics

Information

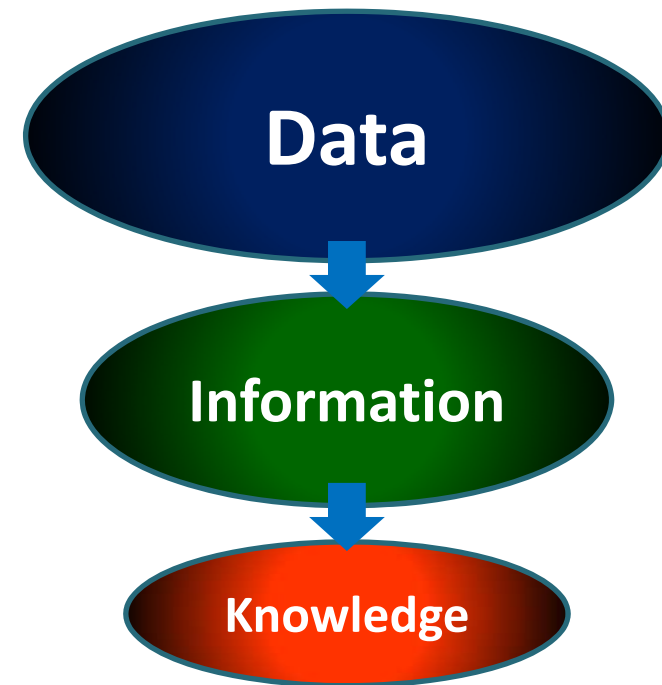
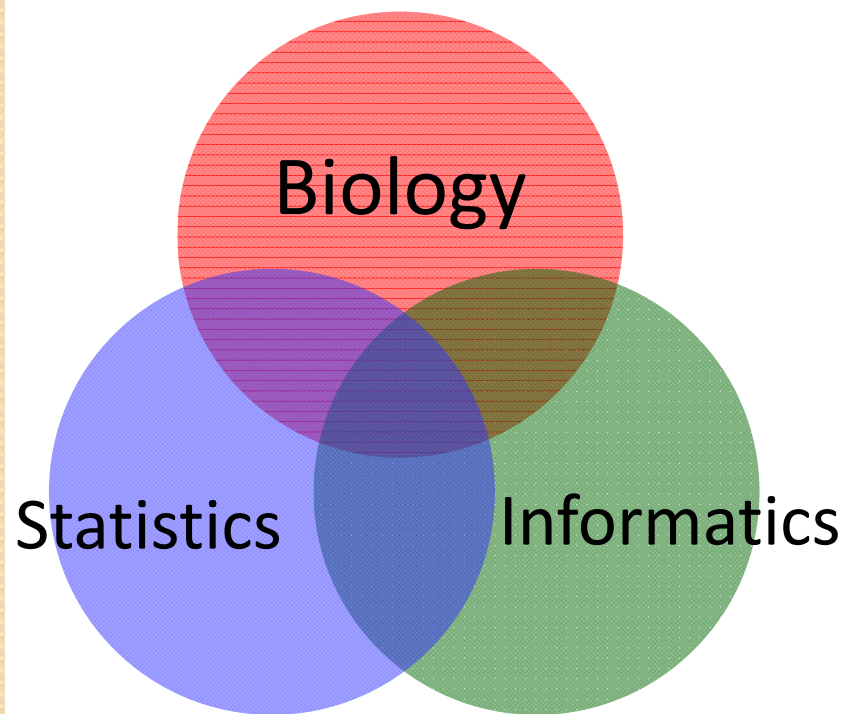
Algorithms



Research field and Goal

- **Goal:**

To transform **data** into **information**, and further integrate and analyze information to novel **knowledge**



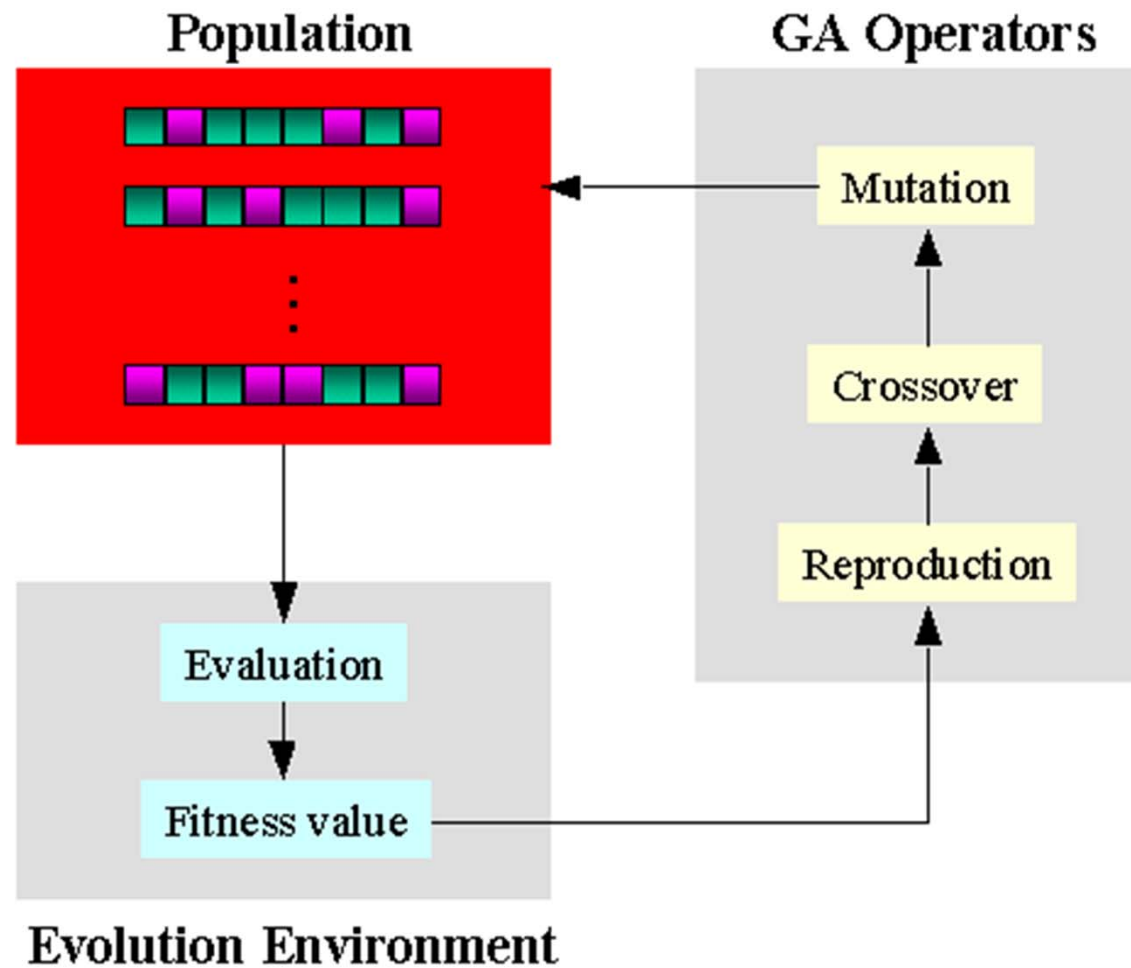


Schedule



Schedule

Genetic Algorithm



Genetic Algorithm Evolution Flow

Algorithm Comparison

0	1	2	3	4	5	6	7	8
23	17	5	90	12	44	38	84	77

↑ exchange

17	23	5	90	12	44	38	84	77
----	----	---	----	----	----	----	----	----

↑ exchange

17	5	23	90	12	44	38	84	77
----	---	----	----	----	----	----	----	----

↑ ok ↑ exchange

17	5	23	12	90	44	38	84	77
----	---	----	----	----	----	----	----	----

↑ exchange

17	5	23	12	44	90	38	84	77
----	---	----	----	----	----	----	----	----

exchange ↑

17	5	23	12	44	38	90	84	77
----	---	----	----	----	----	----	----	----

exchange ↑

17	5	23	12	44	38	84	90	77
----	---	----	----	----	----	----	----	----

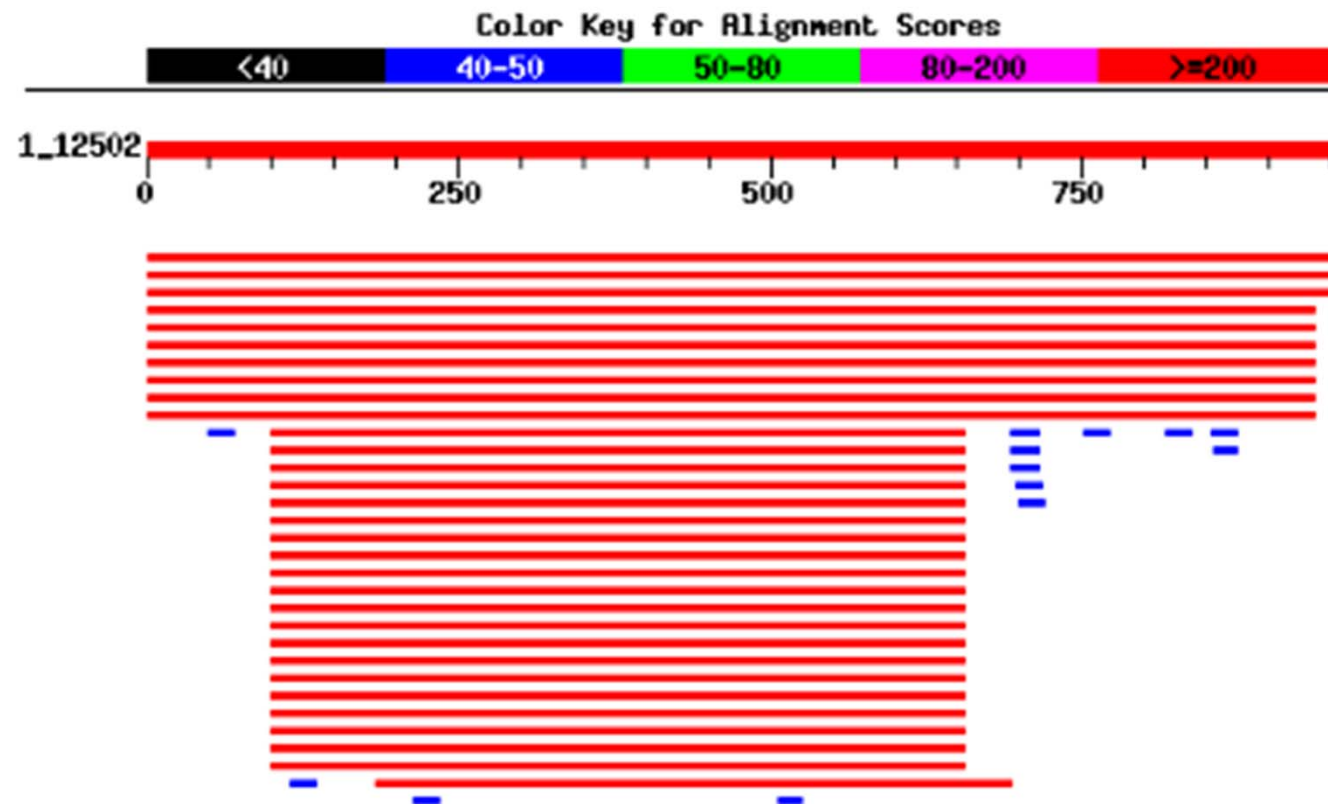
exchange ↑

17	5	23	12	44	38	84	77	90
----	---	----	----	----	----	----	----	----

The largest value 90 is at the end of the list.

Sequence Analysis

- Target sequences: AATTGGCC...GGGC



The result of a blast request. The thick red line above represents the nucleotide input, and the red and blue lines represent the results of the similarity searching in the databases.



Coronavirus -> sequence


Severe acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1, complete genome

GenBank: MN908947.3

[FASTA](#) [Graphics](#)

[Go to:](#) ☒

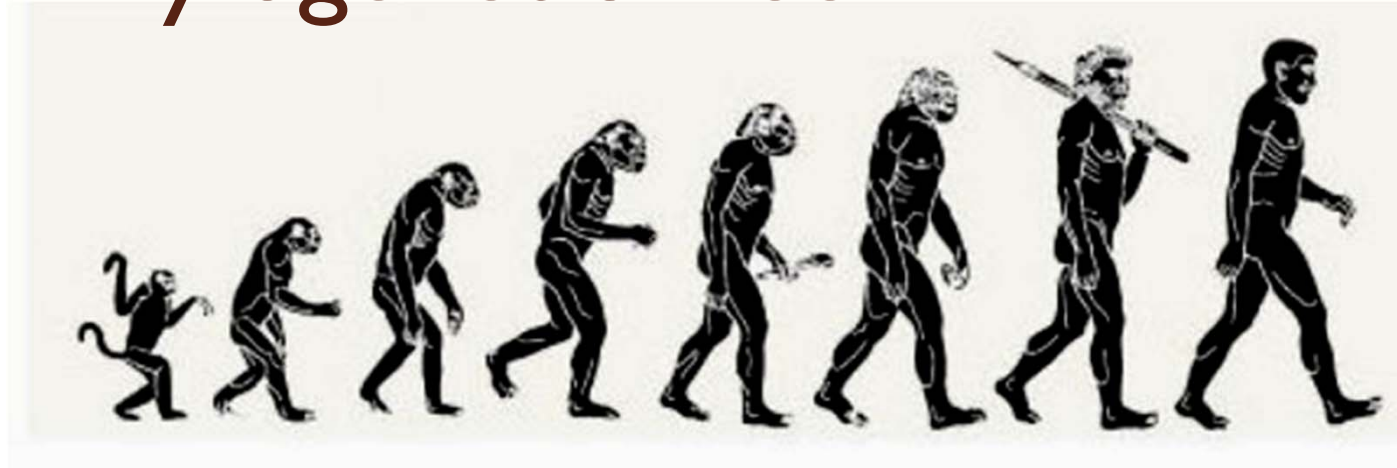
LOCUS	MN908947	29903 bp ss-RNA	linear	VRL 11-FEB-2020
DEFINITION	Severe acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1, complete genome.			
ACCESSION	MN908947			
VERSION	MN908947.3			
KEYWORDS	.			
SOURCE	Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)			
ORGANISM	Severe acute respiratory syndrome coronavirus 2 Viruses; Riboviria; Nidovirales; Coronidovirineae; Coronaviridae; Orthocoronavirinae; Betacoronavirus; Sarbecovirus.			
REFERENCE	1 (bases 1 to 29903)			
AUTHORS	Wu, F., Zhao, S., Yu, B., Chen, Y.M., Wang, W., Song, Z.G., Hu, Y., Tao, Z.W., Tian, J.H., Pei, Y.Y., Yuan, M.L., Zhang, Y.L., Dai, F.H., Liu, Y., Wang, Q.M., Zheng, J.J., Xu, L., Holmes, E.C. and Zhang, Y.Z.			



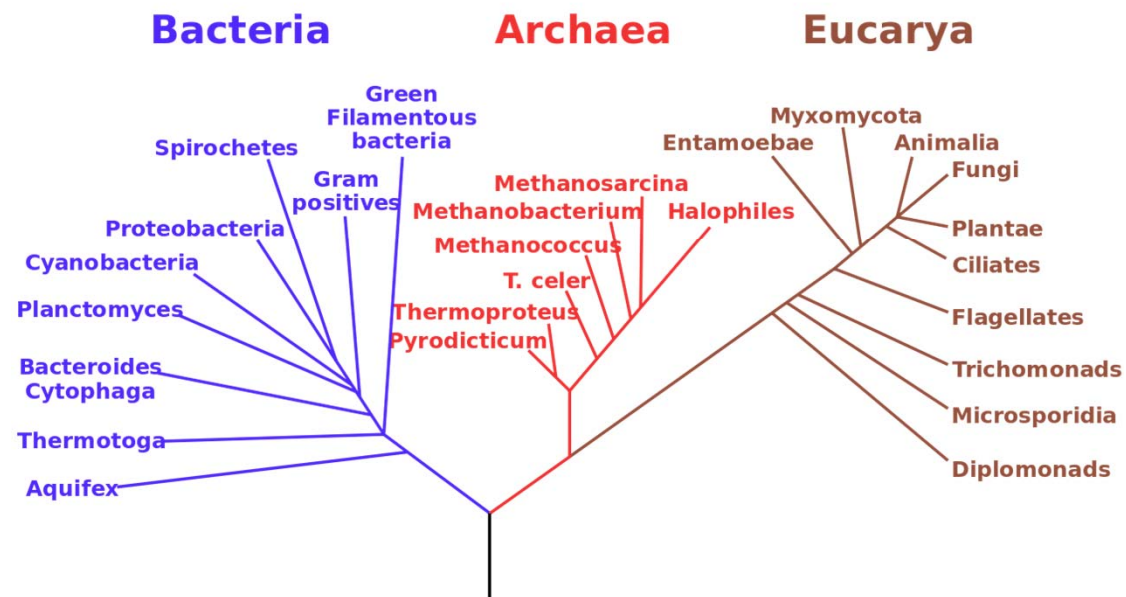
Coronavirus -> blast

Sequences producing significant alignments					Download ▾	Manage Columns ▾	Show	100 ▾	?
<input checked="" type="checkbox"/> select all 100 sequences selected					GenBank Graphics Distance tree of results				
	Description	Max Score	Total Score	Query Cover	E value	Per. Ident	Accession		
<input checked="" type="checkbox"/>	Wuhan seafood market pneumonia virus isolate Wuhan-Hu-1, complete genome	55221	55221	100%	0.0	100.00%	NC_045512.2		
<input checked="" type="checkbox"/>	Wuhan seafood market pneumonia virus isolate Wuhan-Hu-1, complete genome	55221	55221	100%	0.0	100.00%	MN908947.3		
<input checked="" type="checkbox"/>	Wuhan seafood market pneumonia virus isolate BetaCoV/Wuhan/IPBCAMS-WH-03/2019, complete genome	55208	55208	99%	0.0	100.00%	MT019531.1		
<input checked="" type="checkbox"/>	Wuhan seafood market pneumonia virus isolate WIV04, complete genome	55199	55199	99%	0.0	100.00%	MN996528.1		
<input checked="" type="checkbox"/>	Wuhan seafood market pneumonia virus isolate BetaCoV/Wuhan/IPBCAMS-WH-04/2019, complete genome	55197	55197	99%	0.0	100.00%	MT019532.1		
<input checked="" type="checkbox"/>	Wuhan seafood market pneumonia virus isolate BetaCoV/Wuhan/IPBCAMS-WH-01/2019, complete genome	55197	55197	99%	0.0	99.99%	MT019529.1		
<input checked="" type="checkbox"/>	Severe acute respiratory syndrome coronavirus 2 isolate SARS-CoV-2/Yunnan-01/human/2020/CHN, complete genome	55193	55193	100%	0.0	99.98%	MT049951.1		
<input checked="" type="checkbox"/>	Wuhan seafood market pneumonia virus isolate 2019-nCoV WHU02, complete genome	55180	55180	99%	0.0	100.00%	MN988669.1		
<input checked="" type="checkbox"/>	Wuhan seafood market pneumonia virus isolate 2019-nCoV WHU01, complete genome	55180	55180	99%	0.0	100.00%	MN988668.1		
<input checked="" type="checkbox"/>	Wuhan seafood market pneumonia virus isolate BetaCoV/Wuhan/IPBCAMS-WH-05/2020, complete genome	55179	55179	99%	0.0	100.00%	MT019533.1		
<input checked="" type="checkbox"/>	Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV/USA-CA8/2020, complete genome	55177	55177	99%	0.0	100.00%	MT106053.1		
<input checked="" type="checkbox"/>	Wuhan seafood market pneumonia virus isolate 2019-nCoV/USA-CA2/2020, complete genome	55173	55173	99%	0.0	99.99%	MN994468.1		
<input checked="" type="checkbox"/>	Wuhan seafood market pneumonia virus isolate SNU01, complete genome	55171	55171	100%	0.0	99.97%	MT039890.1		
<input checked="" type="checkbox"/>	Wuhan seafood market pneumonia virus isolate 2019-nCoV/USA-CA5/2020, complete genome	55171	55171	99%	0.0	99.99%	MT027064.1		

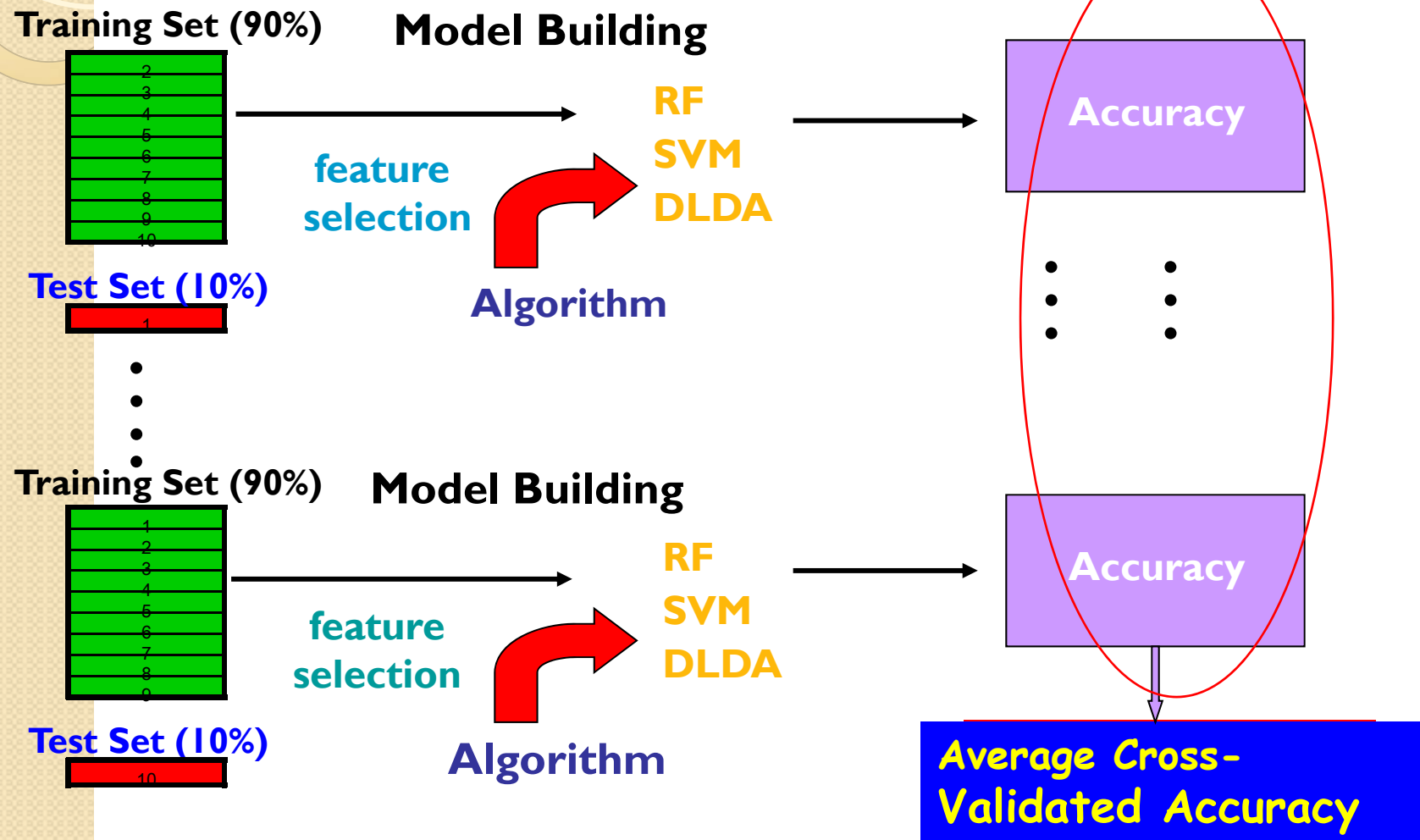
Phylogenetic Tree



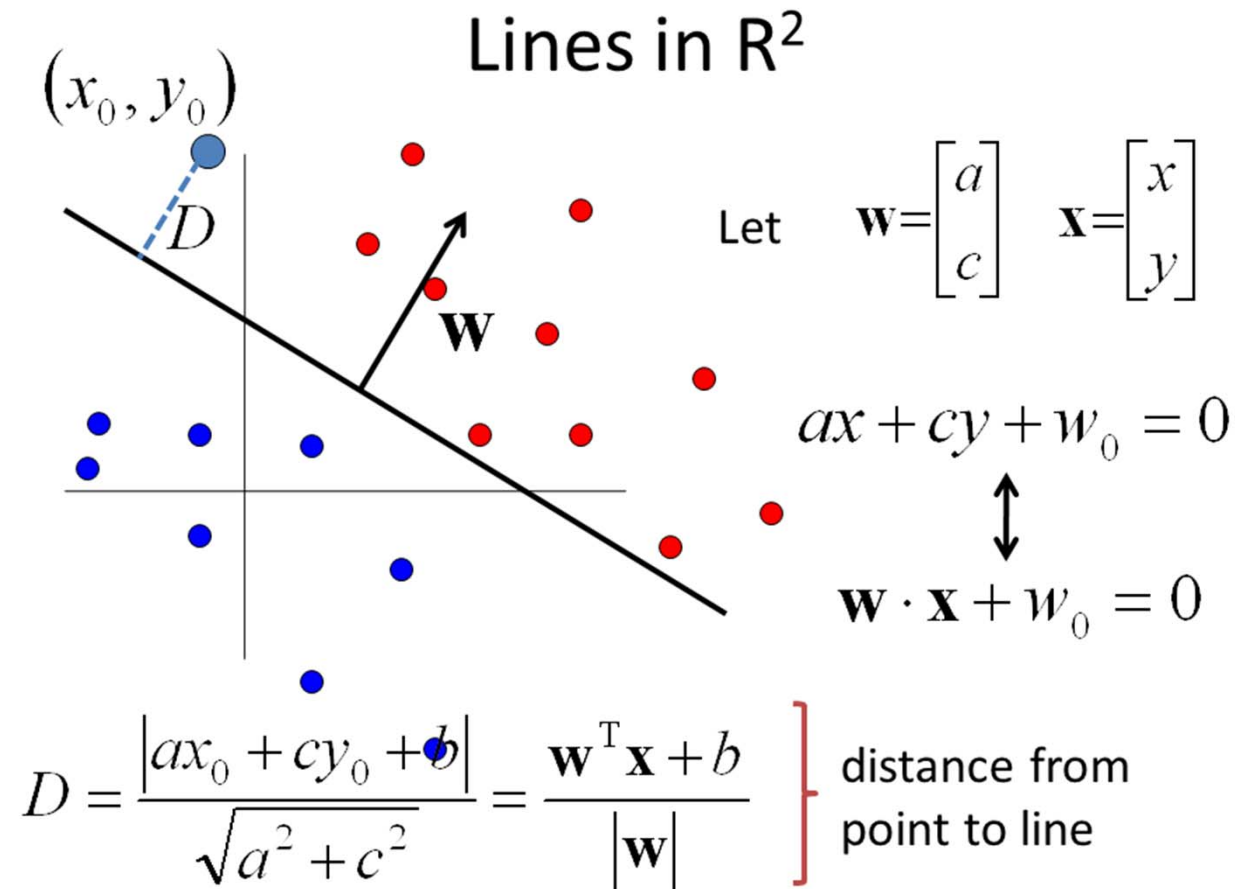
Phylogenetic Tree of Life



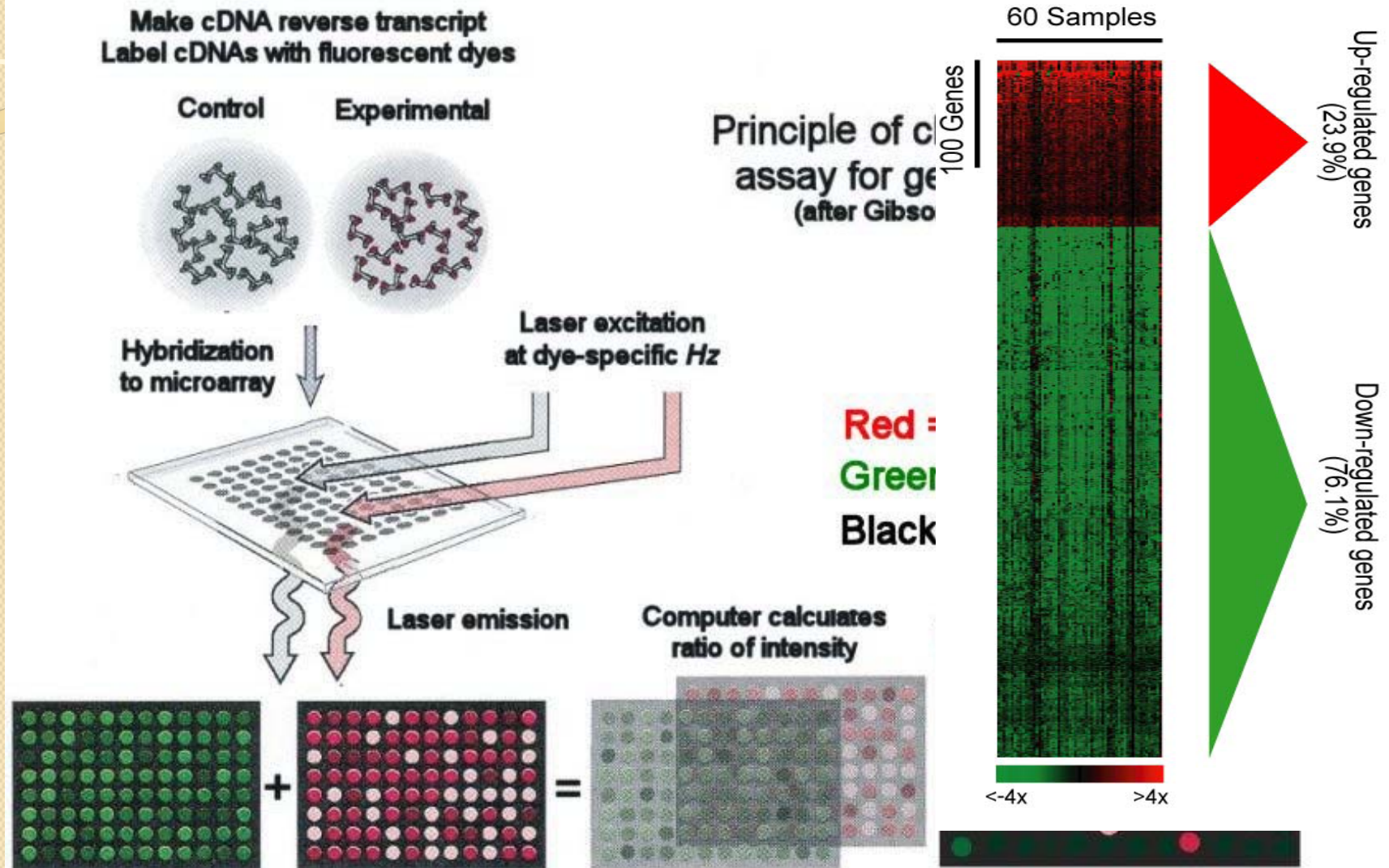
Cross-validation



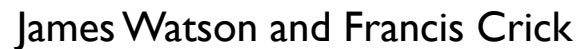
Support vector machine



Gene expression microarray



3 billion base pairs
Double-stranded helices
A, T, C, and G letters
~22,000 genes



課程資訊

- 目標：訓練學生了解各式演算法內容並具備實際上機撰寫程式能力
- 要求：程式語言撰寫及基礎生物學能力
- 上課方式：課堂前半部以上課為主，後半部以上機實作練習。課間會有作業，完成即可離開

評量方式

- 課堂討論 10%
 - 課堂實作(作業) 50%
 - 期中報告 20%
 - 期末報告 20% (含程式碼繳交)
-
- 3-4人一組，題目自由選定，但需相關於本課程相關之內容或方法。

計算生物學原理與應用 Questions?

