

# 體學概論

- 什麼是體學？
- 基因體學與轉錄體學
  - 次世代定序
- 蛋白質體學
  - 質譜法
- 體學相關應用

# 什麼是體學（omics）？

- 針對生物學中某種類別的研究對象（如DNA、RNA、蛋白質、代謝物、表型）進行統整性的系統研究。



# 各式各樣的體(-ome)、體學(-omics)

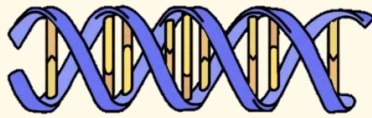
對象

對象的研究

- 基因體（組）學
  - 比較基因體學
  - 總體基因體學
  - 泛基因體學
- 轉錄體學
- 蛋白質體學
- 代謝體學
- 表型體學
- 表觀基因體學
- 脂質體學
- 食品體學
- 醣體學
- 微生物體學
- Regulomics
- Organomics
- ...

# 常見的體學研究

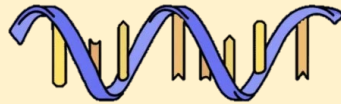
DNA  
(基因)



基因體學

轉錄

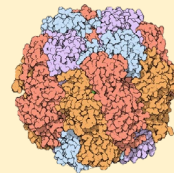
mRNA



轉錄體學

轉譯

蛋白質



蛋白質體學

參與

代謝反應/  
代謝物



代謝體學

造成

表型改變



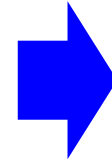
表型體學

環境等  
變因



# 研究工具與體學發展相輔相成

- 一般的體學實驗步驟：

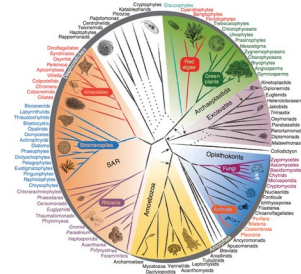
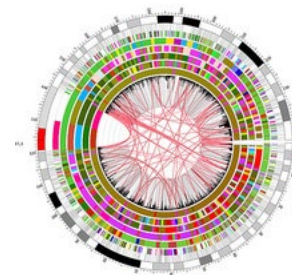
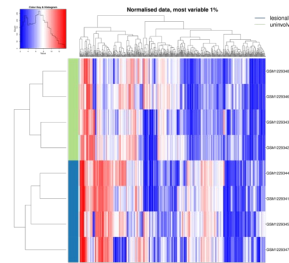


實驗設計與操作  
( 實驗組/對照組 )

樣品製備

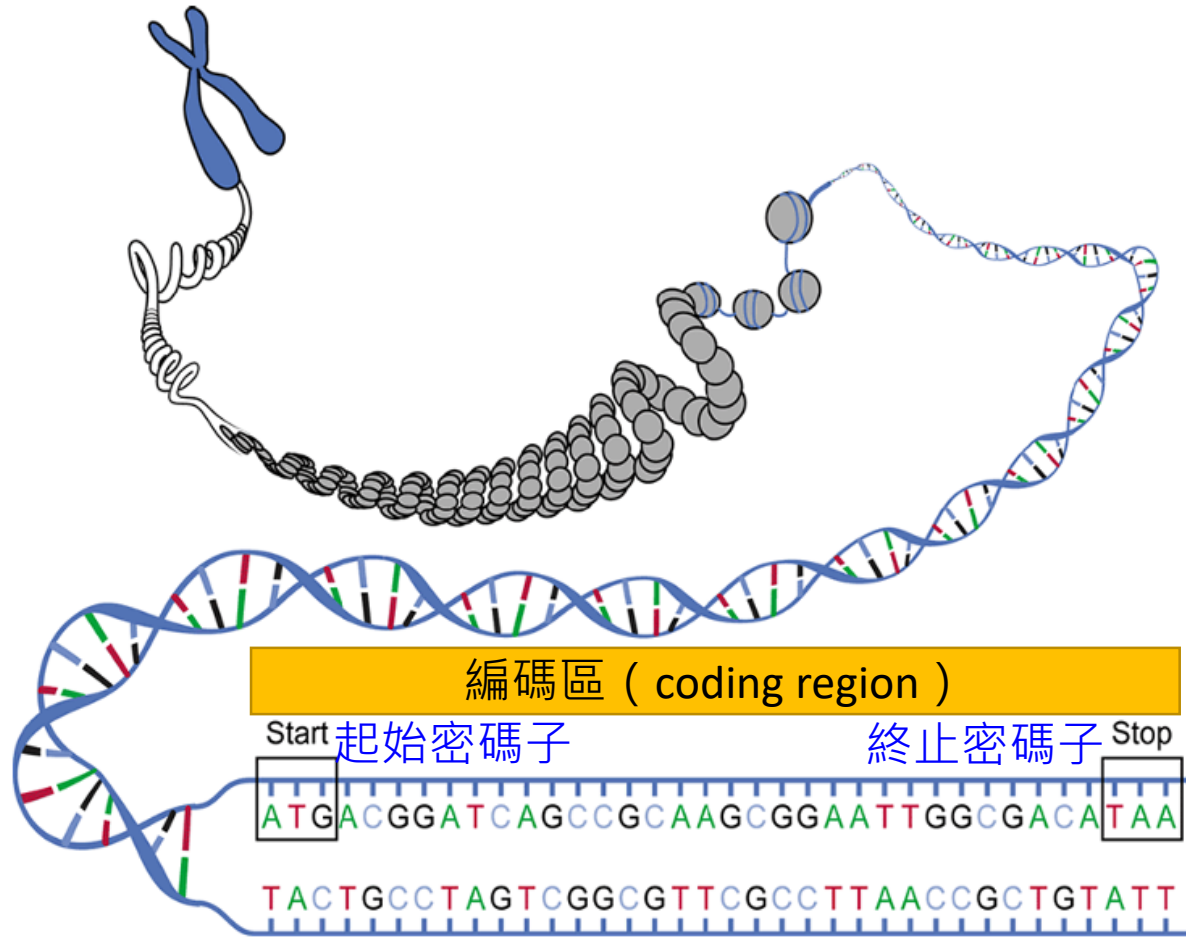
高通量  
儀器分析

大數據分析  
( 生物資訊工具 )



# 基因體學與轉錄體學

- 基因體 ( Genome: gene + -ome )





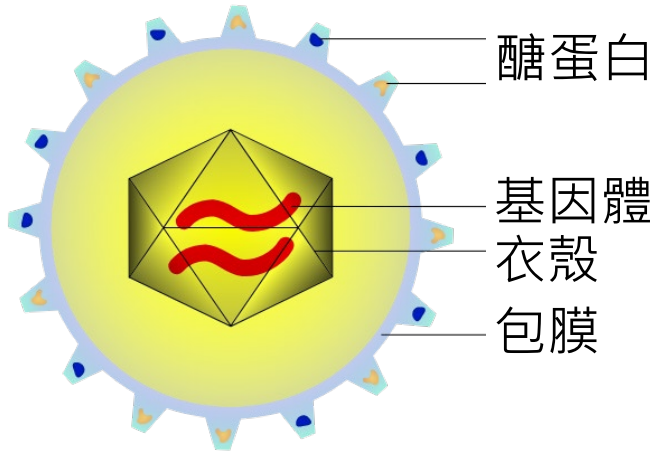
ATTCGCTCATCGATGAAATTGCGCTTAACTGCATCGACTACGTGCTACCCCTGGCGCCT  
ATATATCGCTAGCCCGATATTCATAACGGTACTGCATCGATTTCCCGACCTTAACCTTG  
GAACCTTTCCTTAACCGGTTCCATTCGAACTTGGCCATTCCGGAAATTCCGGGTAAAC  
CTTGGCCAATTGGCCCAAATTTCCCGCCTTCGGGACCTTCGACATTCGCTCATCGATGA  
AATTGCGCTTAACTGCATCGACTACGTGCTACCCCTGGCGCCTATATATCGCTAGCCCG  
ATATTCATAACGGTACTGCATCGATTTCCCGACCTTAACCTTGGAACCTTTCCTTAACC  
GGTTCATTCGAACTTGGCCATTCCGGAAATTCCGGGTAAACCTTGGCCAATTGGCCC  
AAATTTCCCGCCTTCGGGACCTTCGACATTCGCTCATCGATGAAATTGCGCTTAACTGC  
ATCGACTACGTGCTACCCCTGGCGCCTATATATCGCTAGCCCGATATTCATAACGGTAC  
TGCATCGATTTCCCGACCTTAACCTTGGAACCTTTCCTTAACCGGTTCCATTCGAACTT  
GGCCATTCCGGAAATTCCGGGTAAACCTTGGCCAATTGGCCCAAATTTCCCGCCTTCG  
GGACCTTCGACATTCGCTCATCGATGAAATTGCGCTTAACTGCATCGACTACGTGCTAC  
CCCTGGCGCCTATATATCGCTAGCCCGATATTCATAACGGTACTGCATCGATTTCCCGA  
CCTTAACCTTGGAACCTTTCCTTAACCGGTTCCATTCGAACTTGGCCATTCCGGAAATT  
CCGGGTAAACCTTGGCCAATTGGCCCAAATTTCCCGCCTTCGGGACCTTCGACGGCAT  
CCGAGTCAGCTAACCGGTTTAACCTTGGACTGACCTTAAGAGACTTAGAGCTGGAACCTC



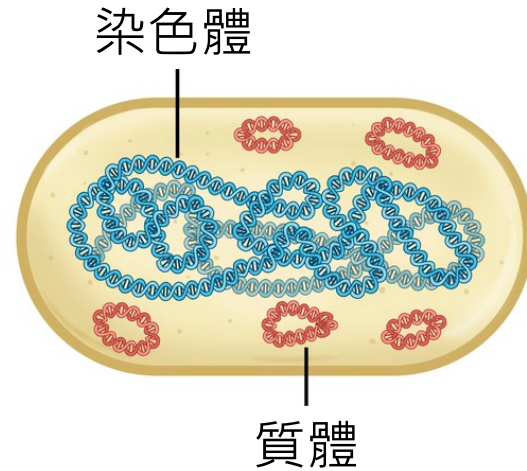
# 哪裡有基因體？

線形？環形？

## 病毒

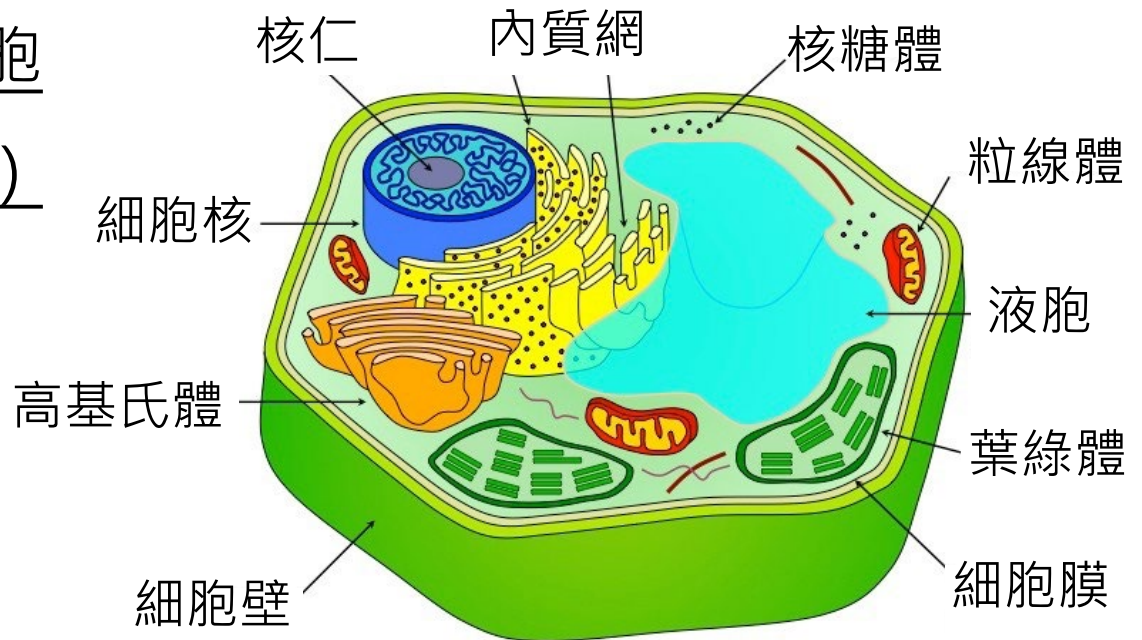


## 細菌



## 真核細胞

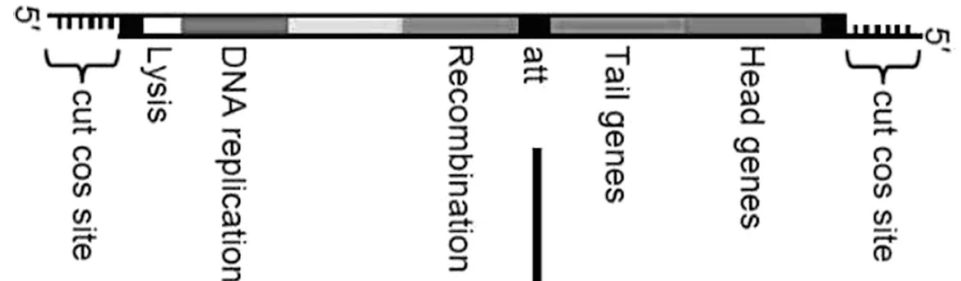
(以植物細胞為例)



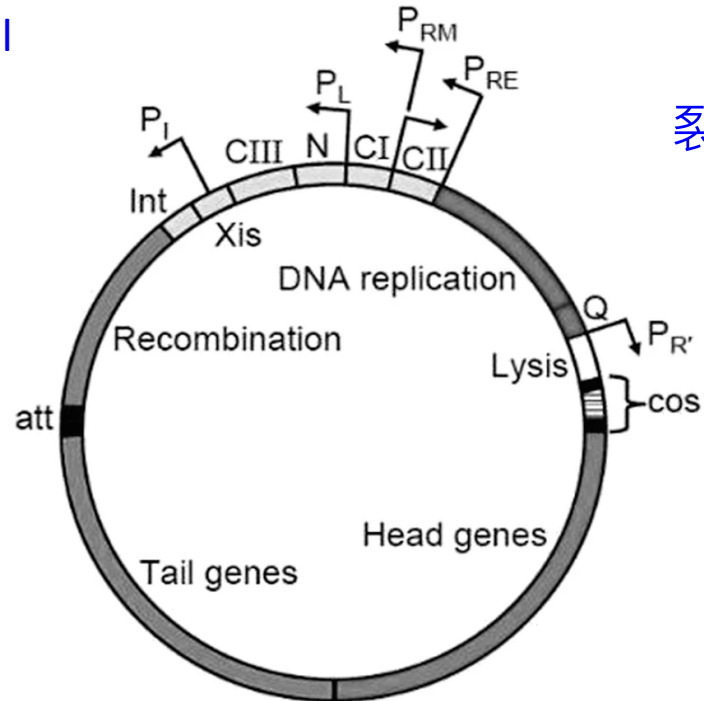
# 病毒的基因體結構

- 以λ噬菌體基因體為例

溶原期



裂解期

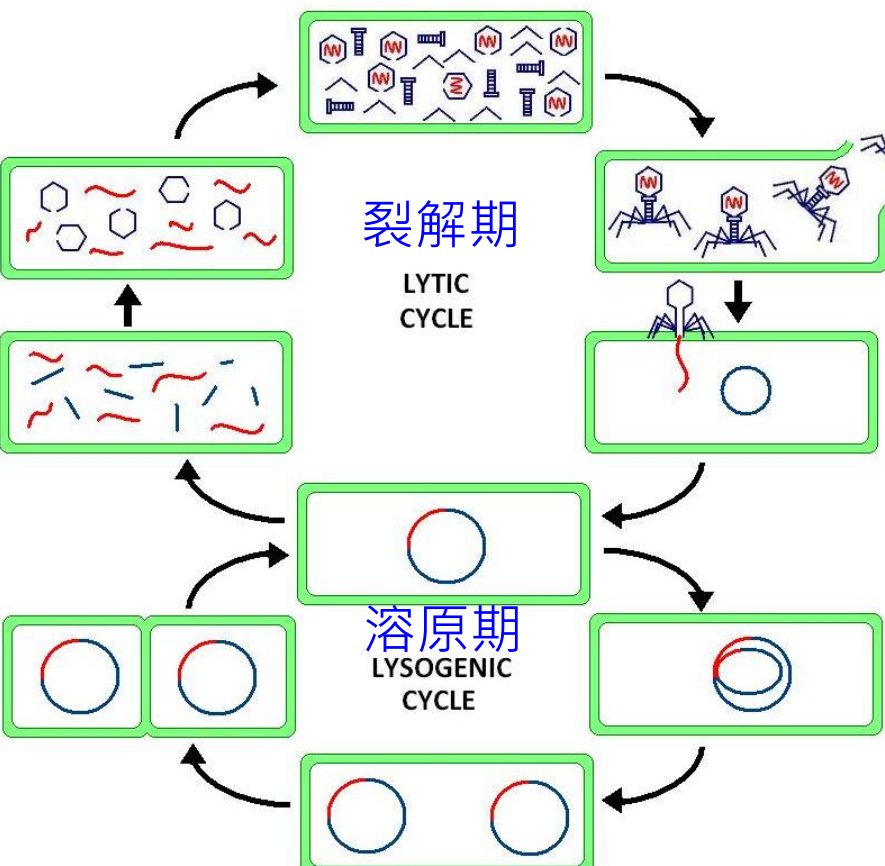


裂解期

LYTIC CYCLE

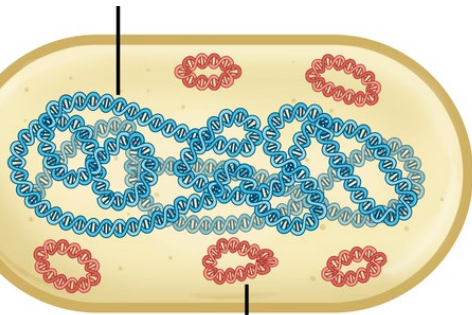
溶原期

LYSOGENIC CYCLE



# 細菌的基因體結構(I)：質體

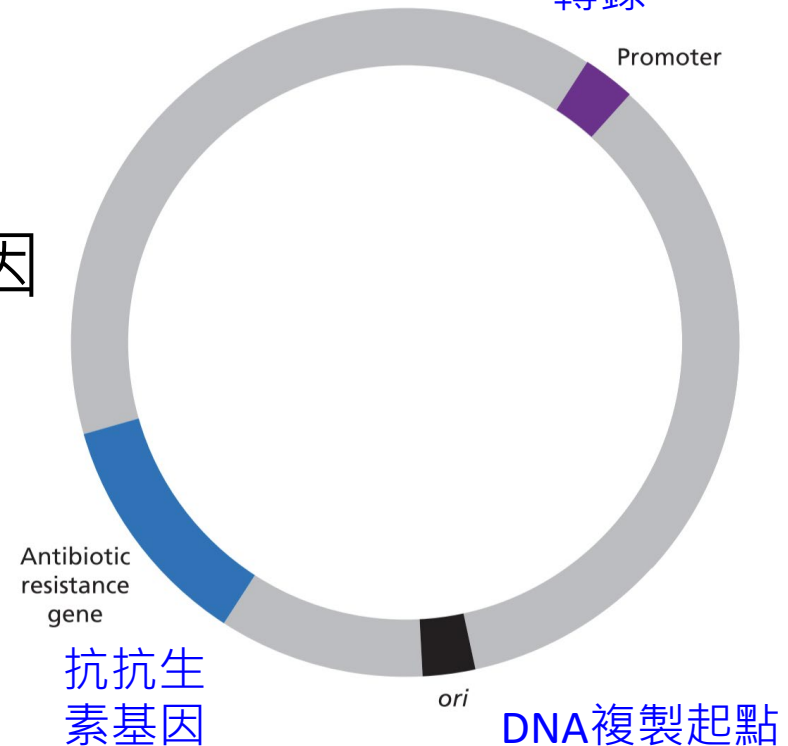
染色體



質體

## 分子選殖載體組成

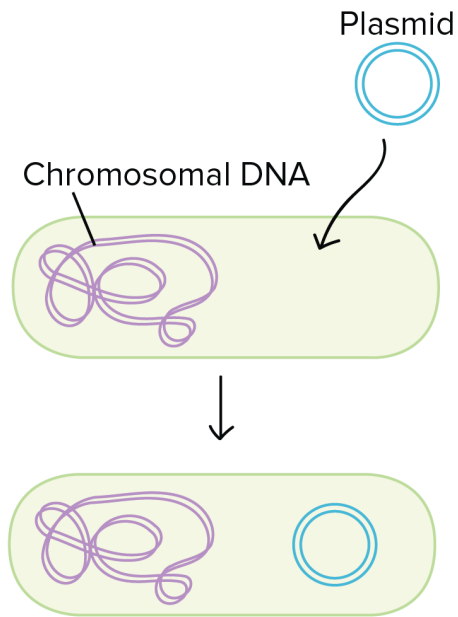
啟動子：驅  
使DNA進行  
轉錄



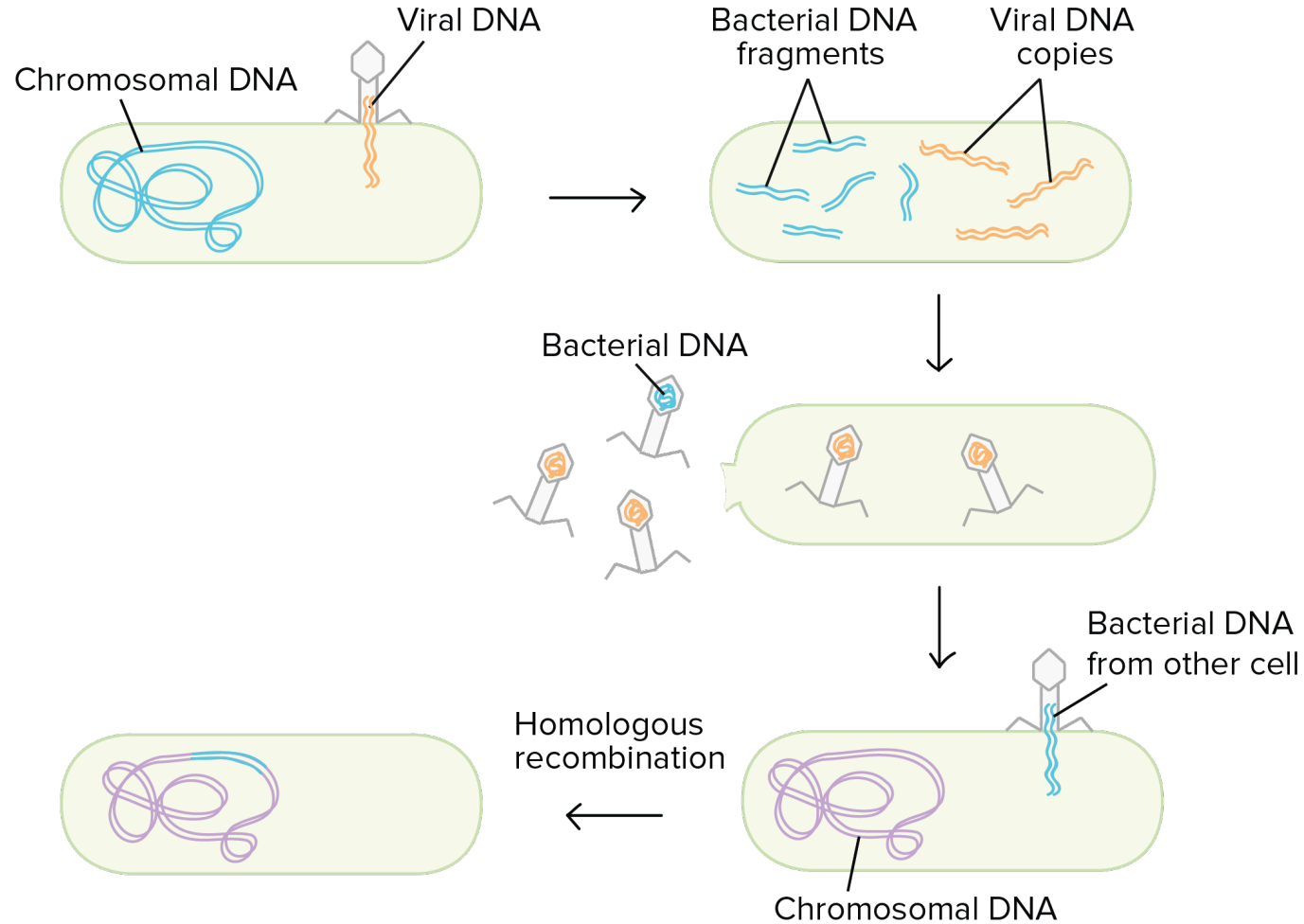
- 細菌染色體外環形DNA分子
- 獨立於染色體行自我複製
- 含有**非**細菌生存代謝必要之基因
  - 接合作用 ( conjugation )
  - 抗抗生素
  - 特殊物質合成酵素...等
- 常改造用來作為分子選殖載體

# 質體轉移途徑(I)

## 1. 轉型 ( transformation )

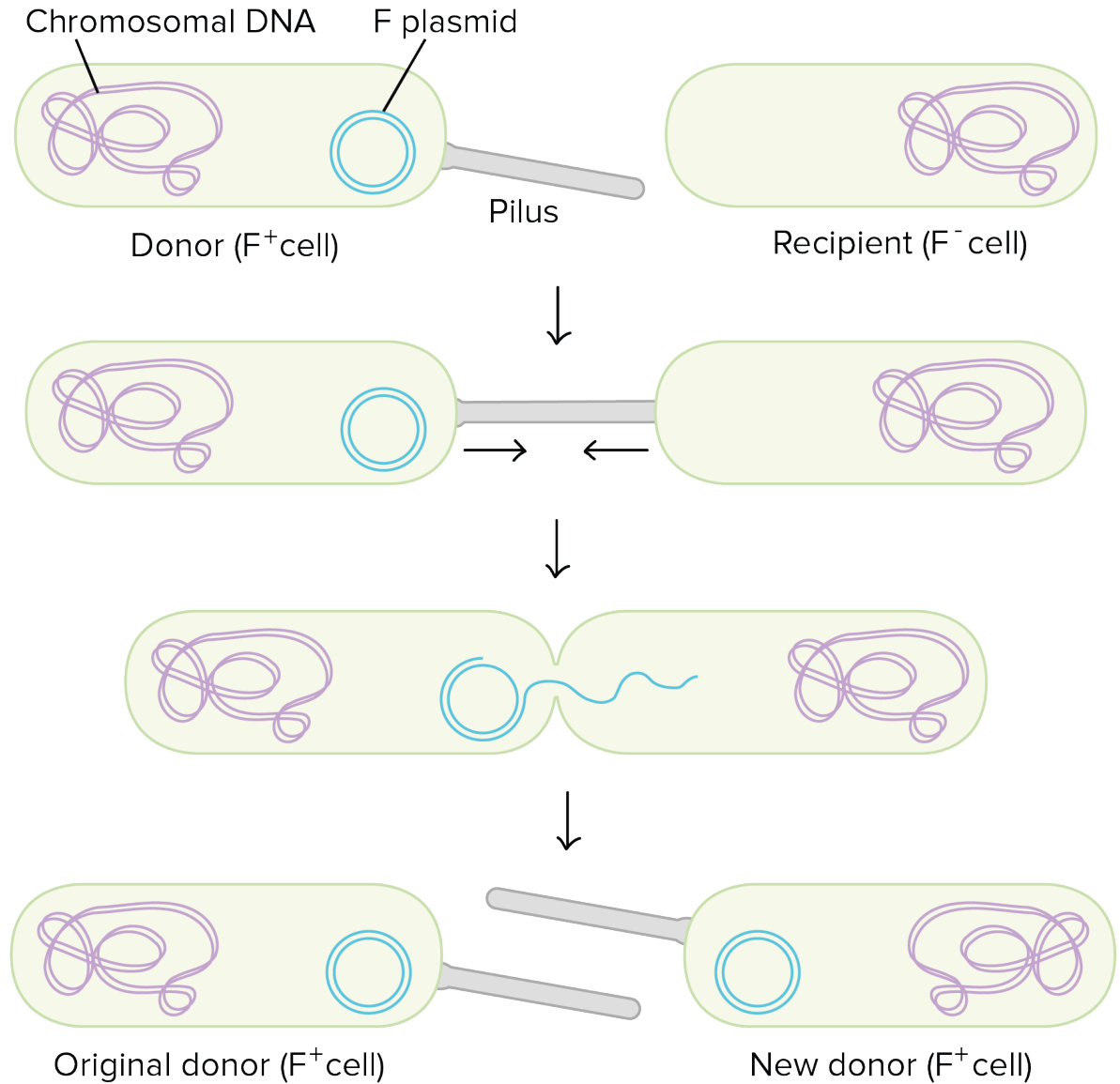


## 2. 轉導 ( transduction )



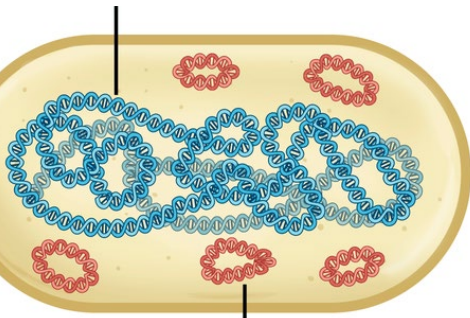
# 質體轉移途徑(II)

## 3. 接合 ( conjugation )



# 細菌的基因體結構(II)：染色體

染色體



質體



## Circular map legend

- Ring 1: ORFs plus strand  
*Escherichia coli* str. K-12 substr. MDS42 DNA
- Ring 2: ORFs minus strand  
*Escherichia coli* str. K-12 substr. MDS42 DNA
- Ring 3: GC percentage  
*Escherichia coli* str. K-12 substr. MDS42 DNA
- Ring 4: text  
*Escherichia coli* str. K-12 substr. MDS42 DNA
- Ring 5: Coordinates  
*Escherichia coli* str. K-12 substr. MDS42 DNA

開放閱讀框

( open reading frame, ORF )

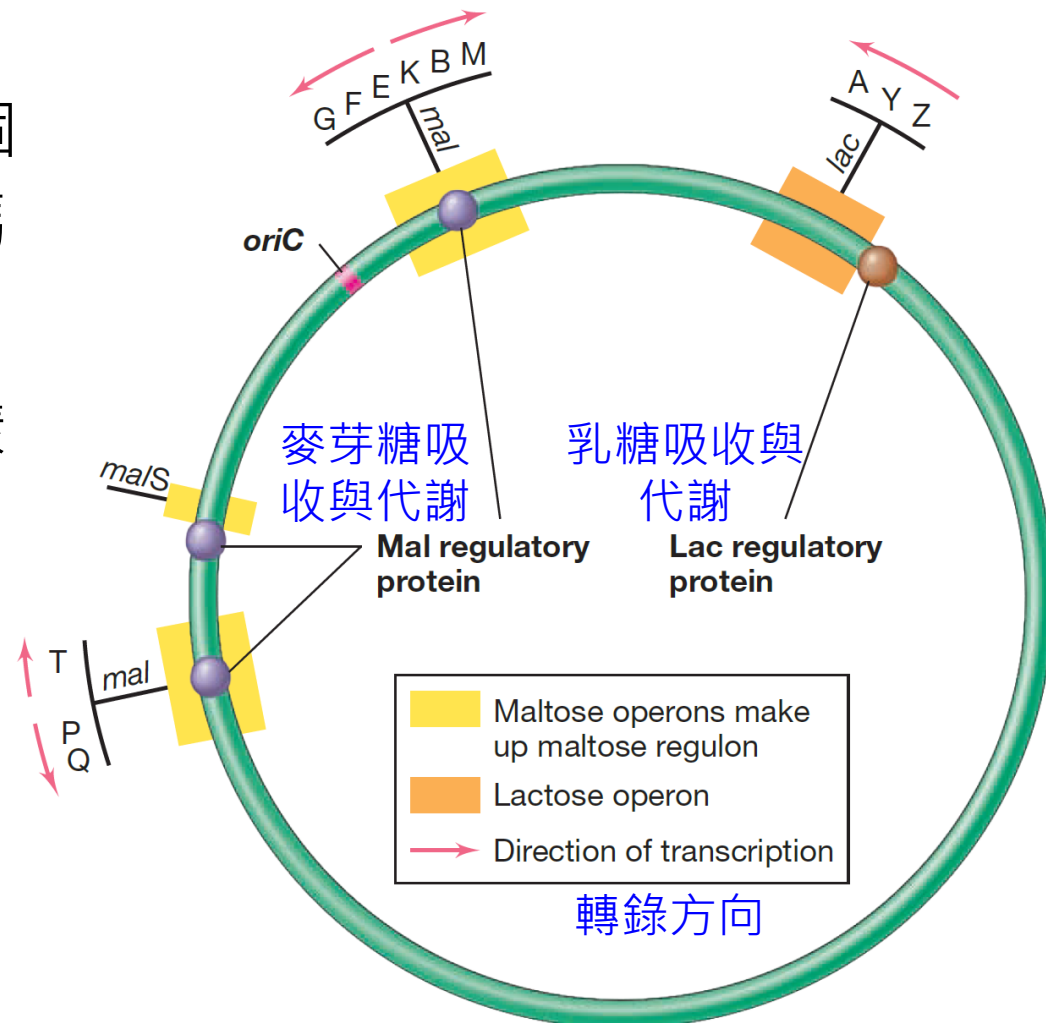
排列密集

Interactive view of the *E. coli* genome:

[http://bacmap.wishartlab.com/maps/NC\\_000913/index.html](http://bacmap.wishartlab.com/maps/NC_000913/index.html)

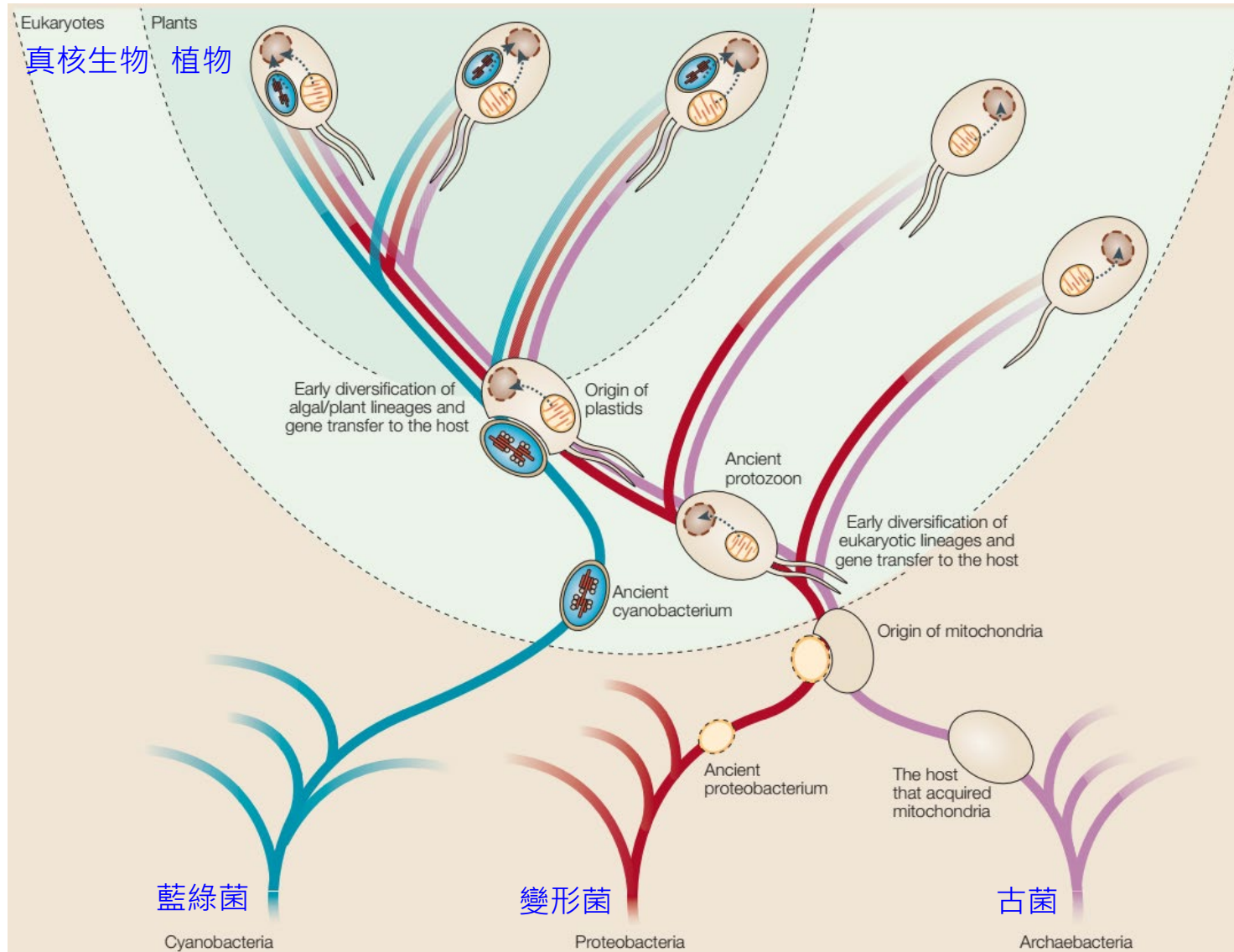
# 細菌基因組成特色：操縱組 ( operon )

- 操縱組：由單一操縱子 (operator) 所負責調控一個或一個以上的基因 ( 編碼區 ) 表現。
- 通常這一組基因負責同樣的代謝功能 ( 如右圖 ) 。



# 真核細胞基因體結構(I)：粒線體與葉綠體

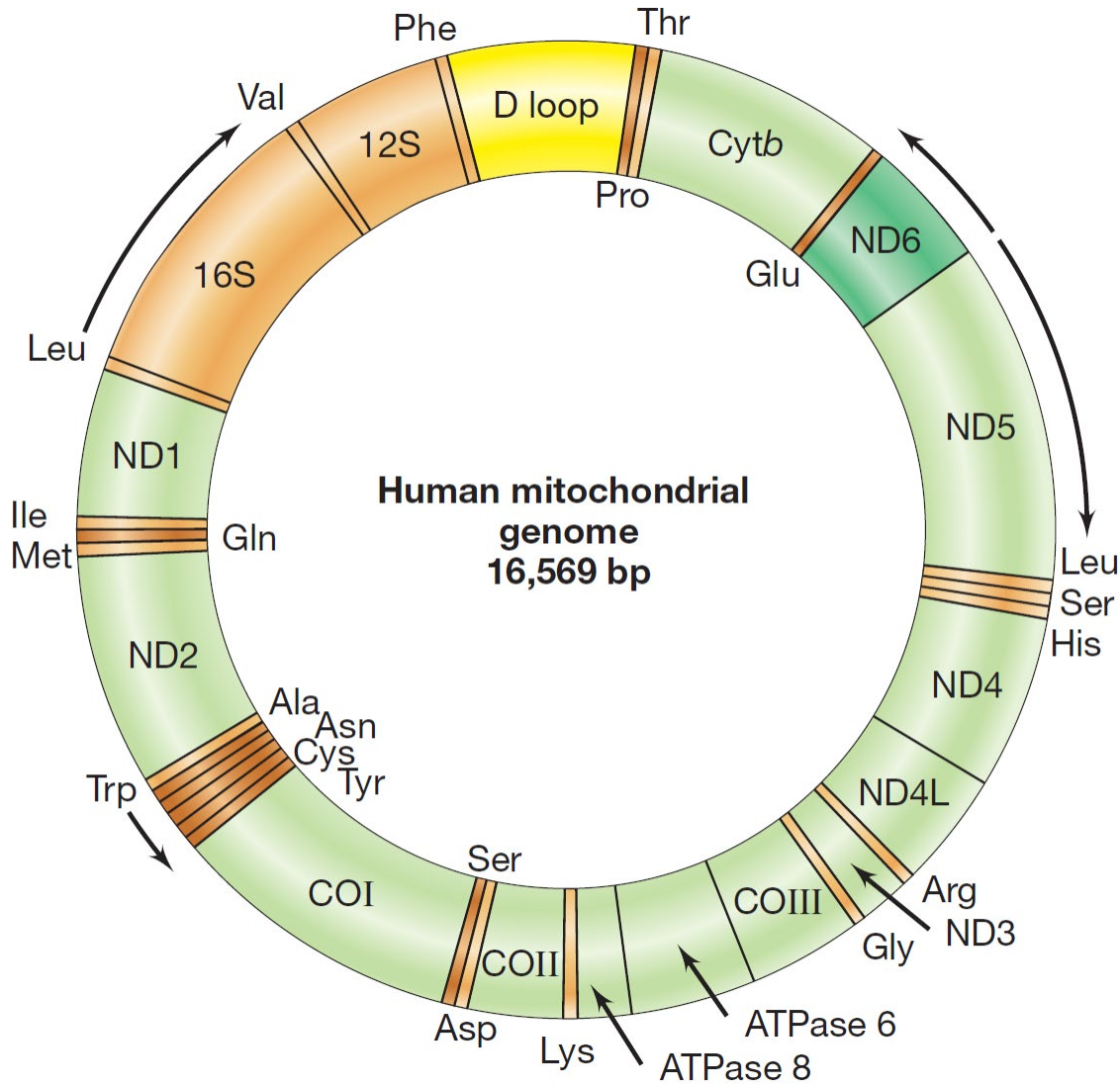
## 內共生理論



Timmis, J., Ayliffe, M., Huang, C. *et al.* (2004) Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. *Nat Rev Genet* 5, 123–135



# • 粒線體基因體



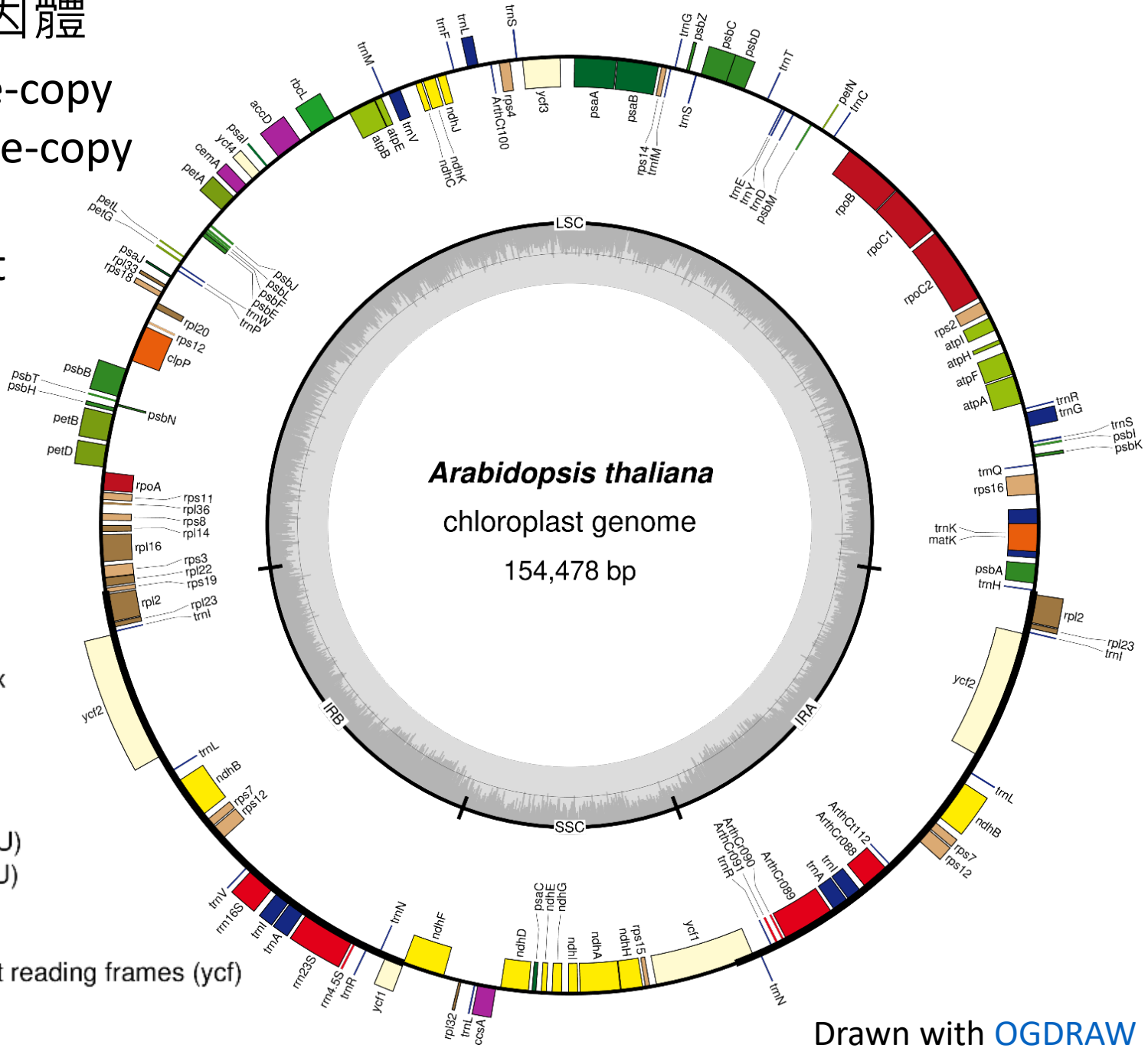
呼吸作用電子傳遞

tRNA, rRNA

DNA複製

# • 葉綠體基因體

LSC: long single-copy  
 SSC: short single-copy  
 IRA, IRB: two inverted repeat (region)

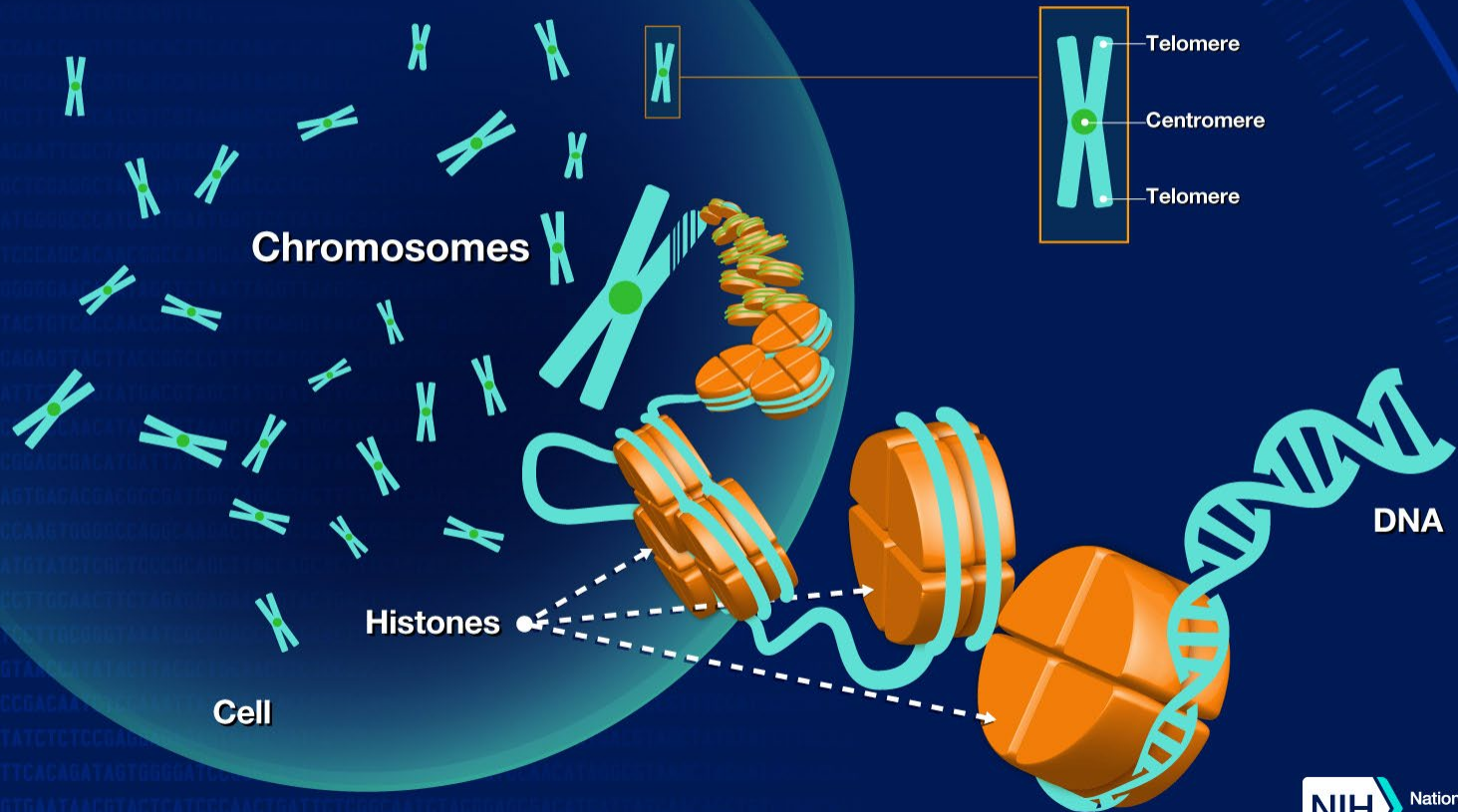


- photosystem I
- photosystem II
- cytochrome b/f complex
- ATP synthase
- NADH dehydrogenase
- RubisCO large subunit
- RNA polymerase
- ribosomal proteins (SSU)
- ribosomal proteins (LSU)
- clpP, matK
- other genes
- hypothetical chloroplast reading frames (ycf)
- transfer RNAs
- ribosomal RNAs

# 真核細胞基因體結構(II)：細胞核基因體

## Chromosome

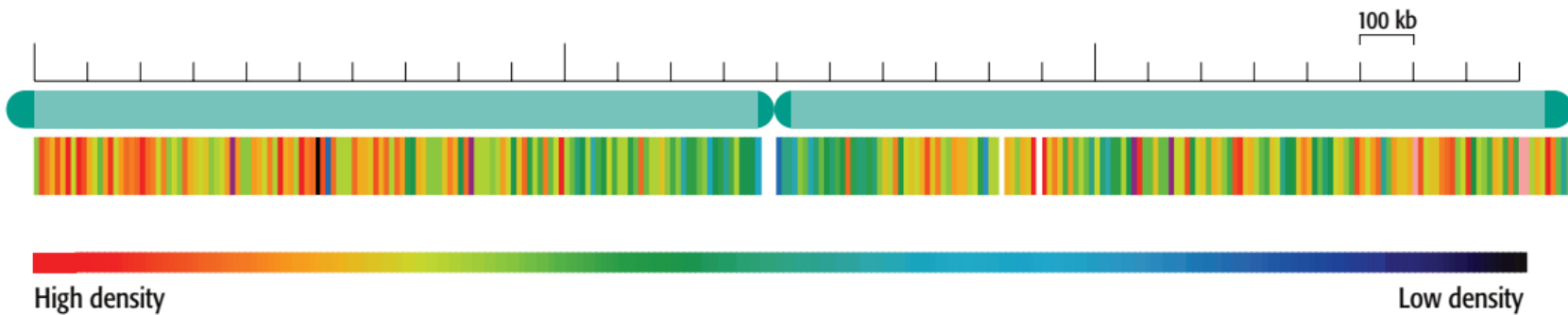
NHGRI FACT SHEETS  
genome.gov



NIH National Human Genome Research Institute

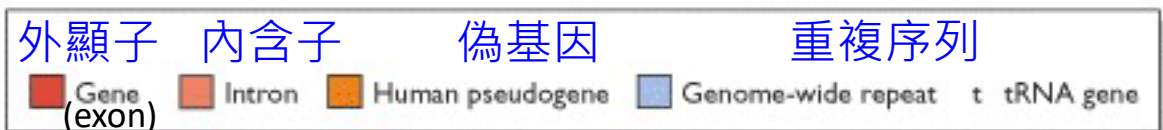
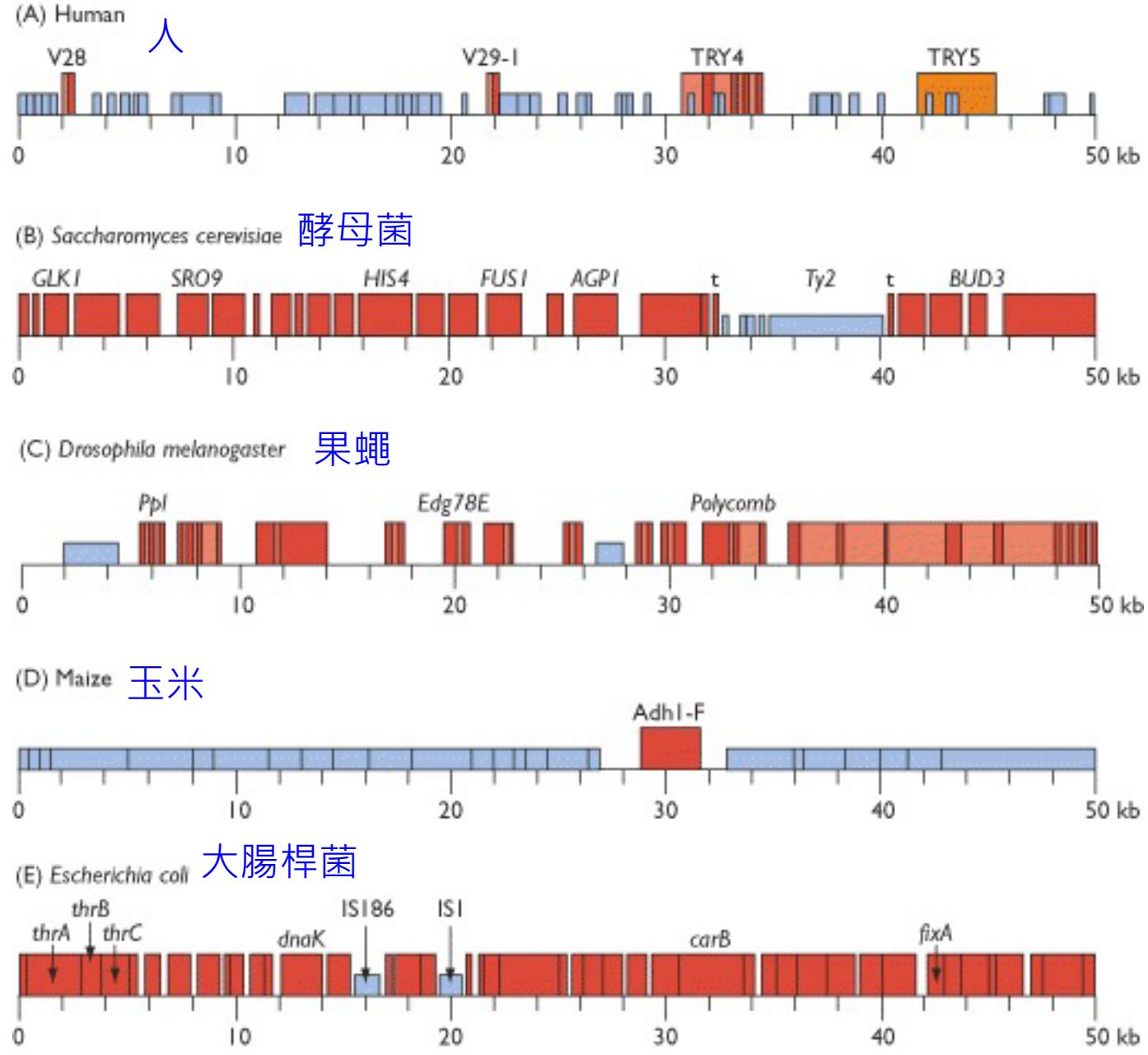
# 真核細胞核基因體特色

- 基因密度分佈不均



Gene density along the largest of the five *Arabidopsis thaliana* chromosomes

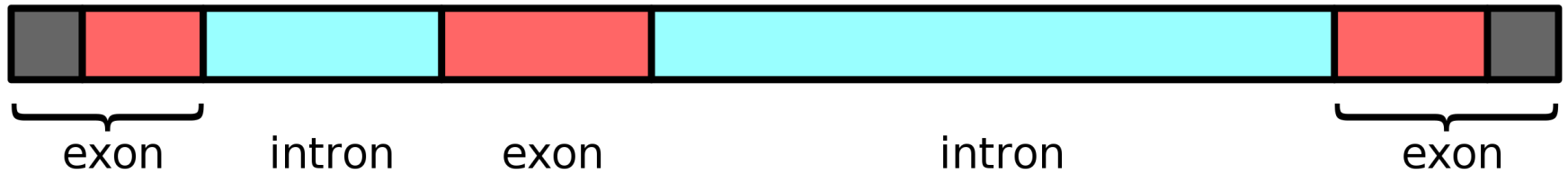
同樣長度 ( 鹼基數 )  
之基因體片段比較



<https://www.ncbi.nlm.nih.gov/books/NBK21120/>

- 外顯子？內含子？

pre-mRNA (從DNA轉錄而來)



mature mRNA

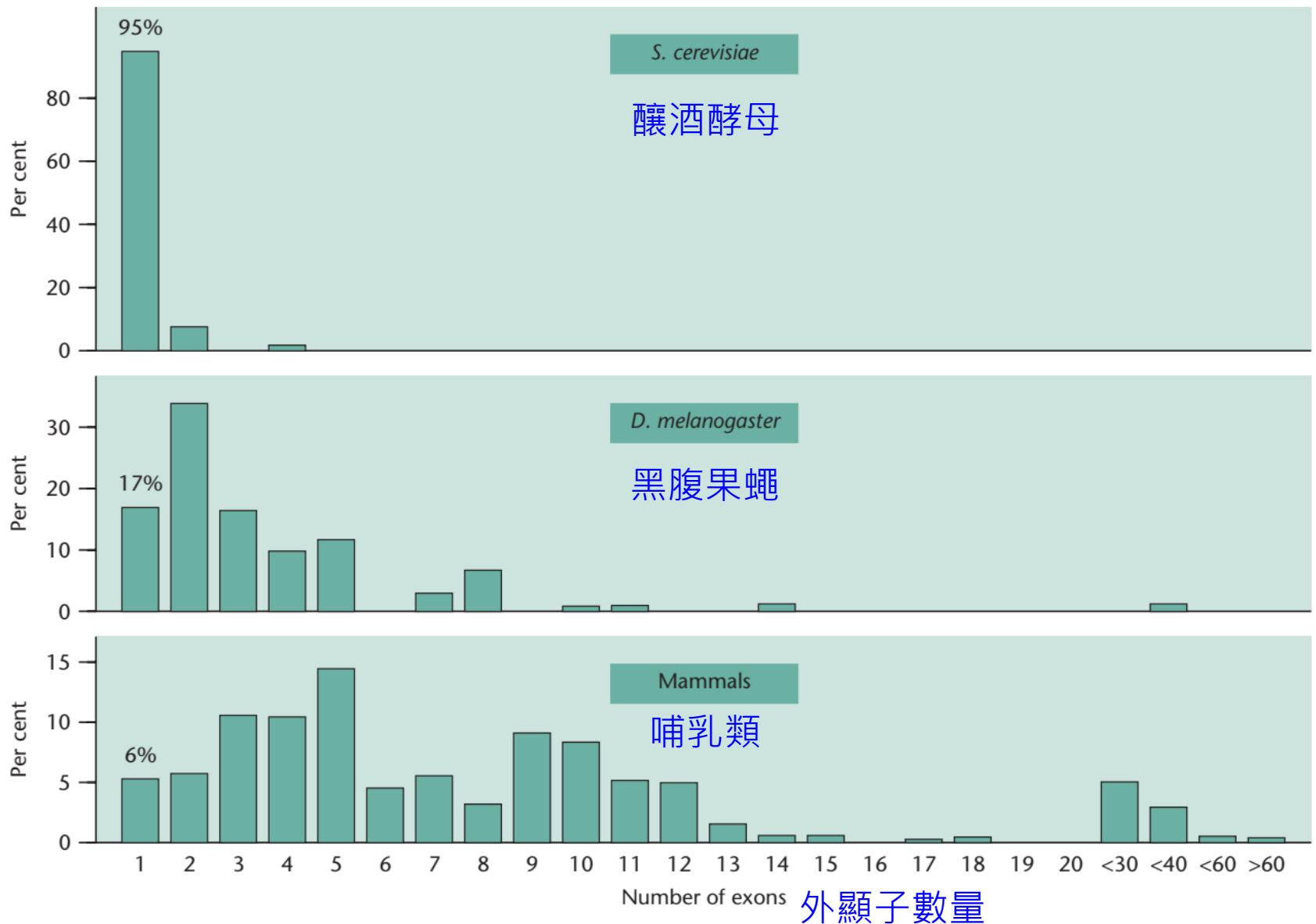


complete protein-coding sequence

■ untranslated region (UTR)    ■ coding sequence (CDS)    ■ intron

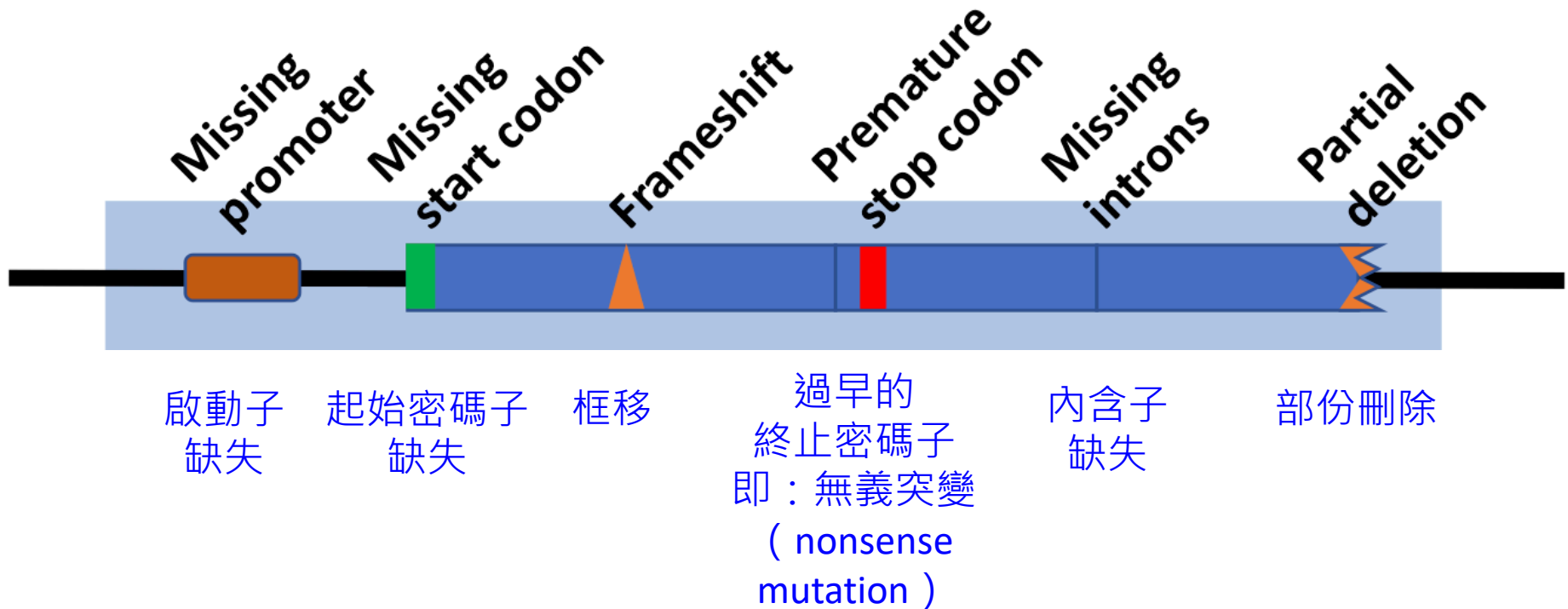
<https://en.wikipedia.org/wiki/Intron>

# 不同真核生物基因的外顯子數量分佈



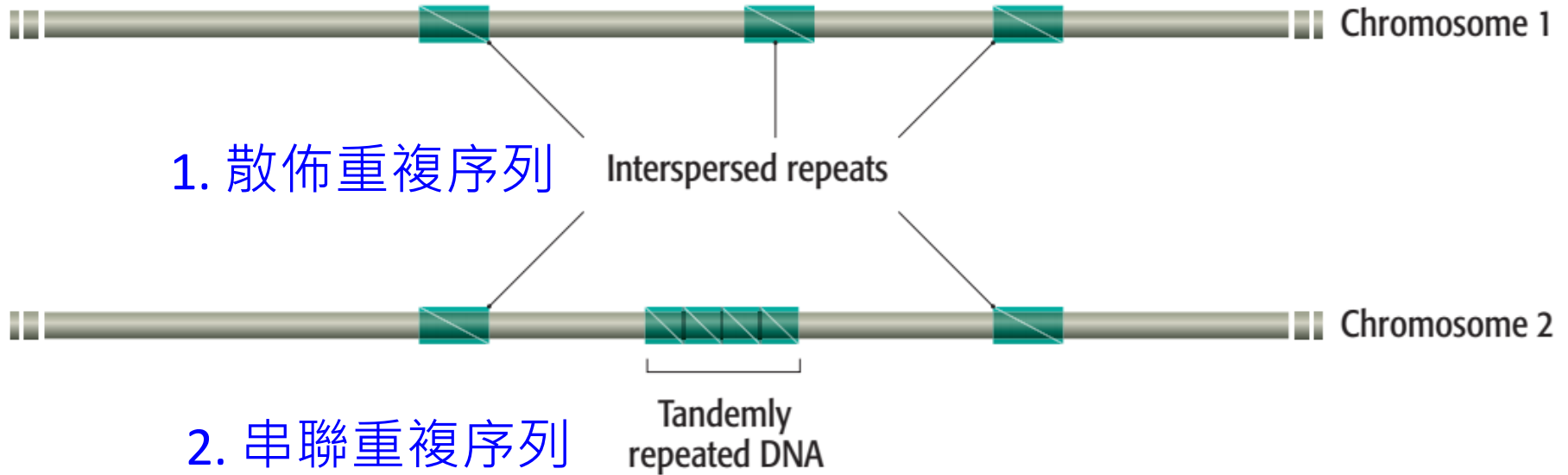
- 偽基因：無法表現或編碼的蛋白質沒有功能的基因

## 常見的偽基因缺失特徵

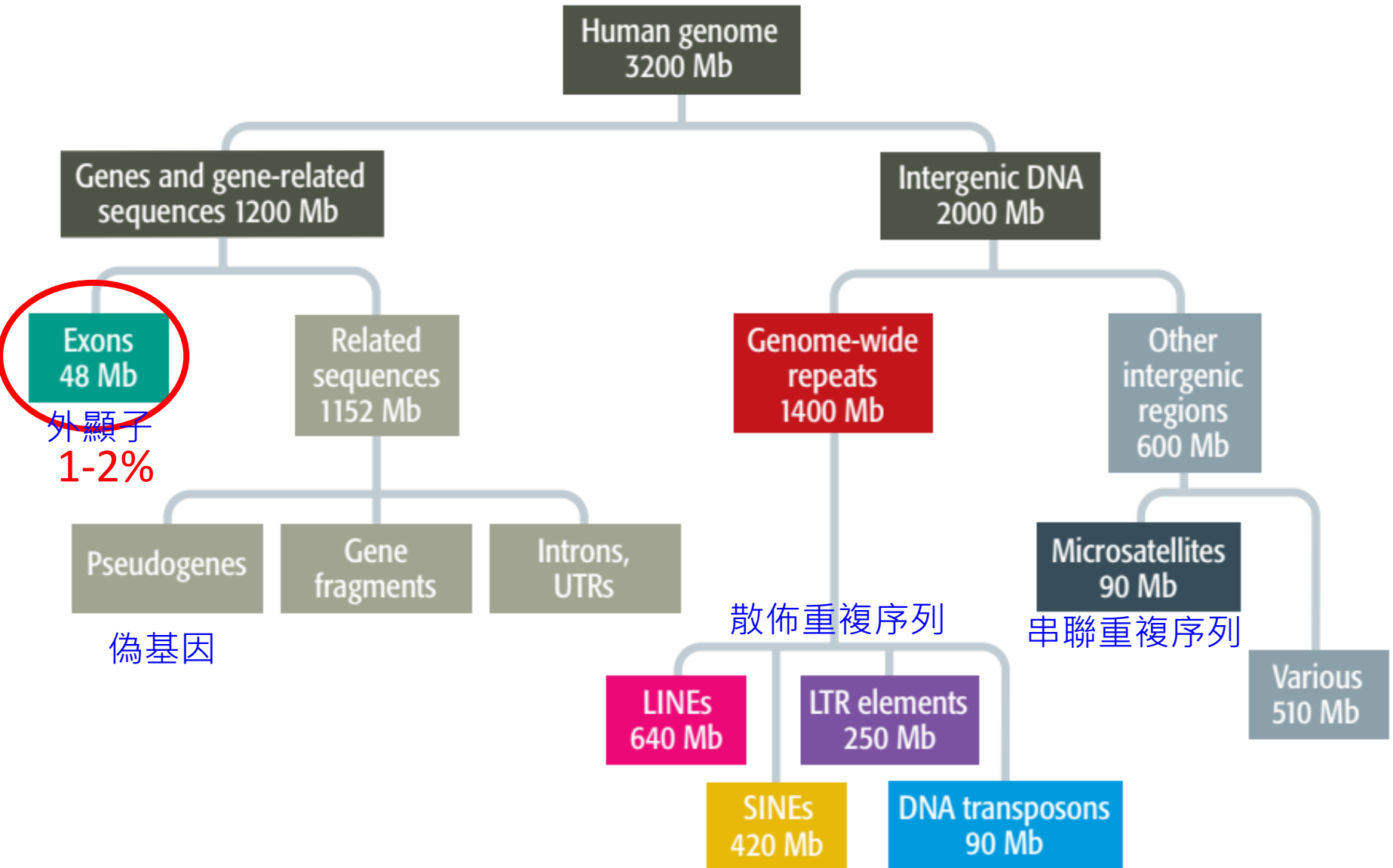




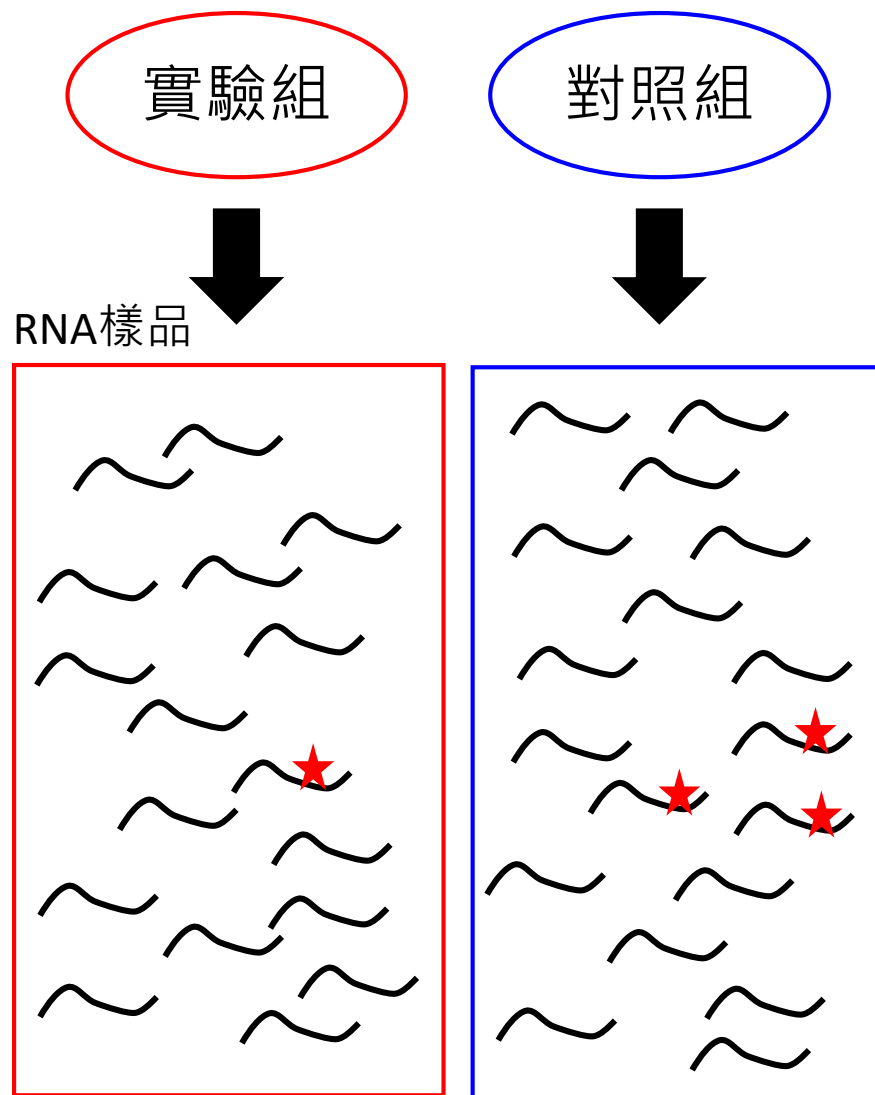
- 重複序列：分為兩種



# 人類基因體中各種組成比例



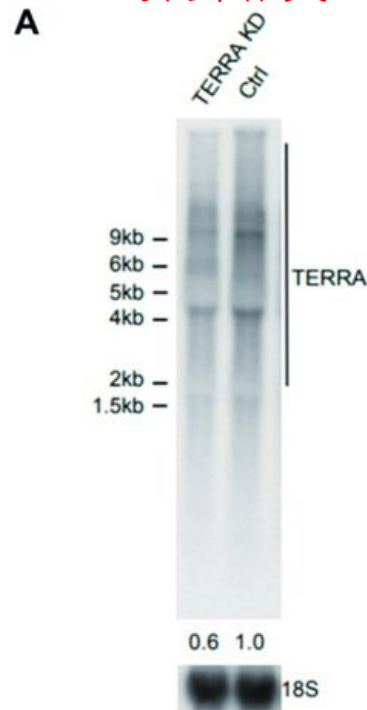
# • 轉錄體 ( Transcriptome: transcript + -ome )



專統方式分析基因表現方法  
( 僅針對少數基因 )

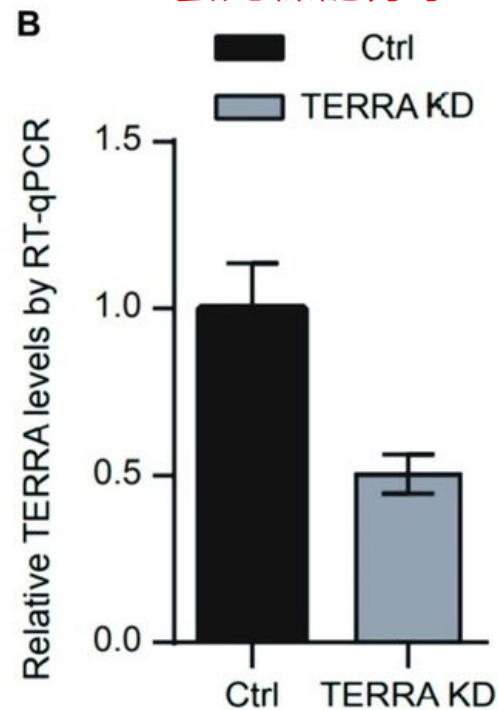
北方墨點法

★ 探針雜交



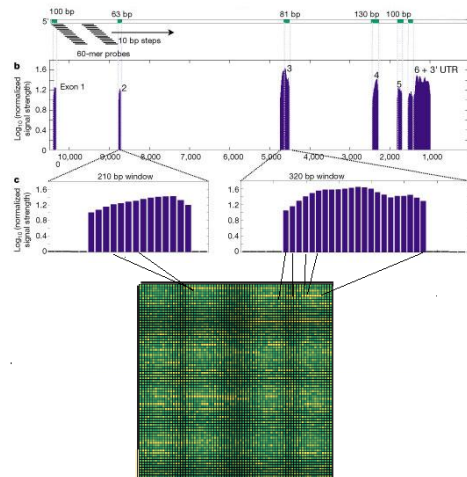
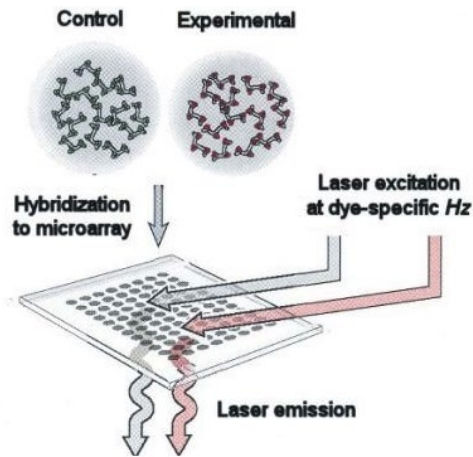
RT-qPCR

★ 螢光標記分子

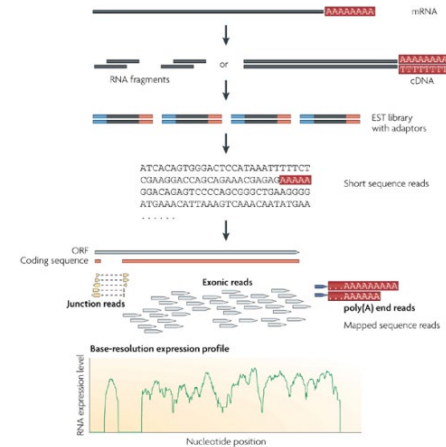


# 轉錄體分析演進

## 利用雜交技術



## 利用次世代定序



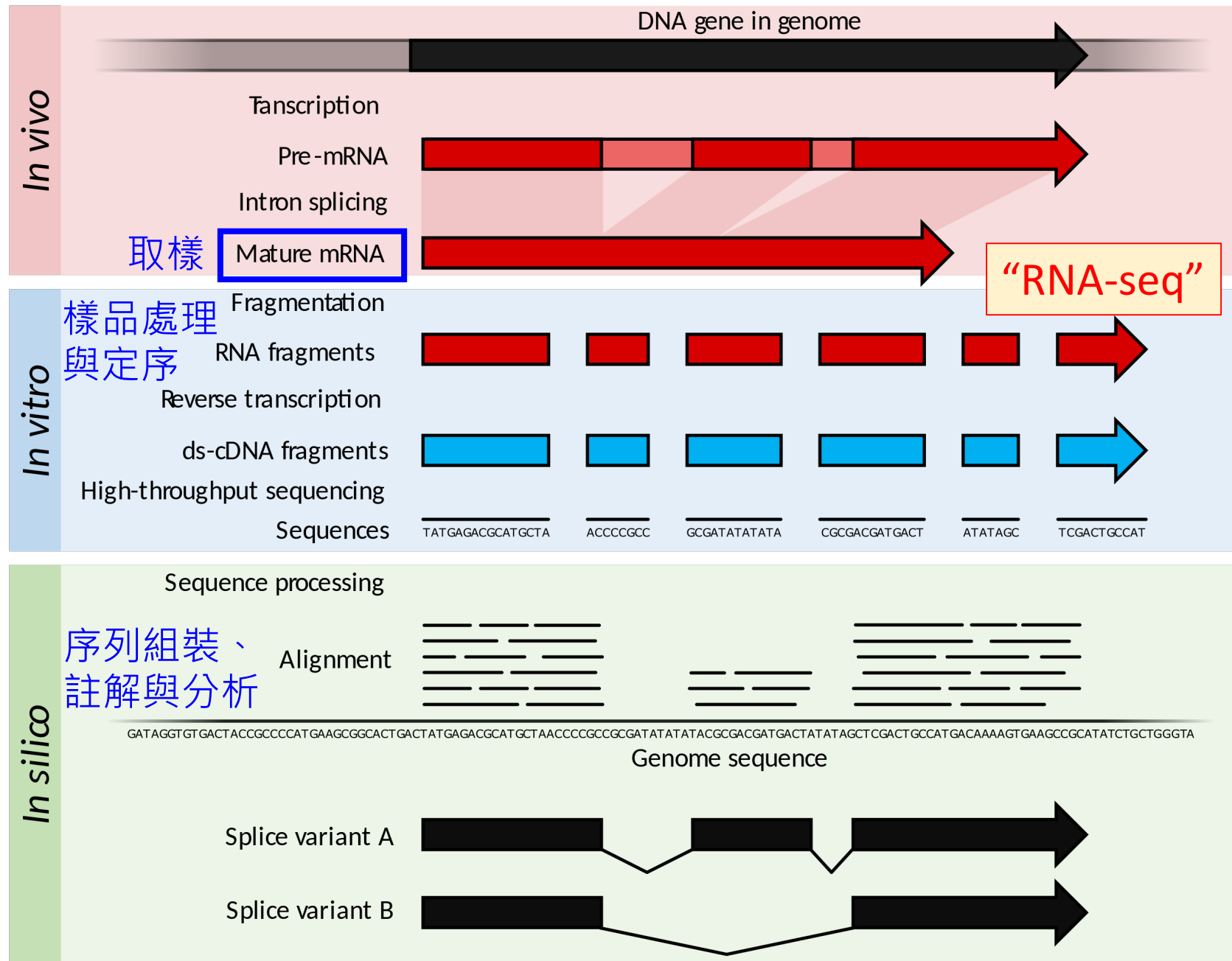
Nature Reviews | Genetics

1995 P. Brown, et. al.  
Gene expression profiling  
using spotted **cDNA  
microarray**: expression levels  
of known genes

2002 Affymetrix, **whole  
genome expression profiling**  
using tiling array: identifying  
and profiling novel genes and  
splicing variants

2008 many groups, **mRNA-seq**:  
direct sequencing of mRNAs  
using next generation  
sequencing techniques (NGS)

• 利用次世代定序進行轉錄體分析



# 兩種RNA組裝與定量方法

( 參照基因體 )

**Reference-based**

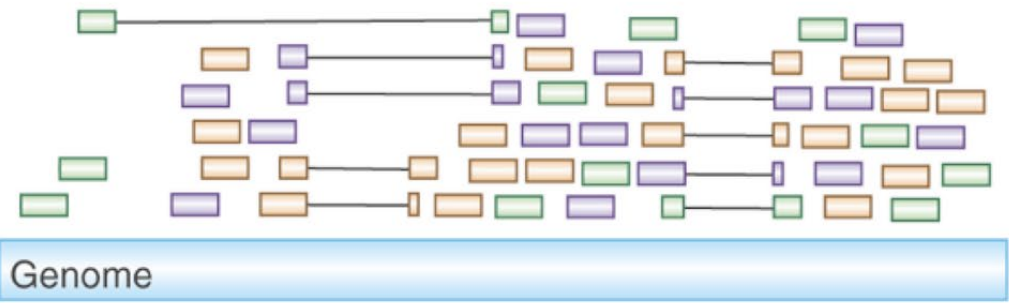
RNA-Seq reads

( 全新組裝 )

**De novo**

Align reads to genome

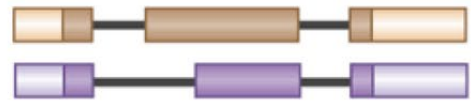
Assemble transcripts de novo



Assemble transcripts from spliced alignments



Align transcripts to genome



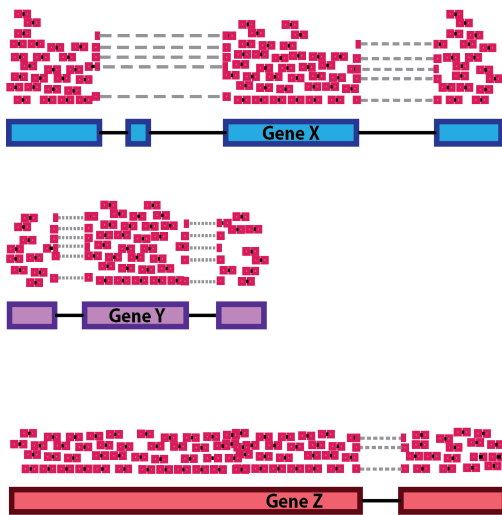
More abundant

Less abundant

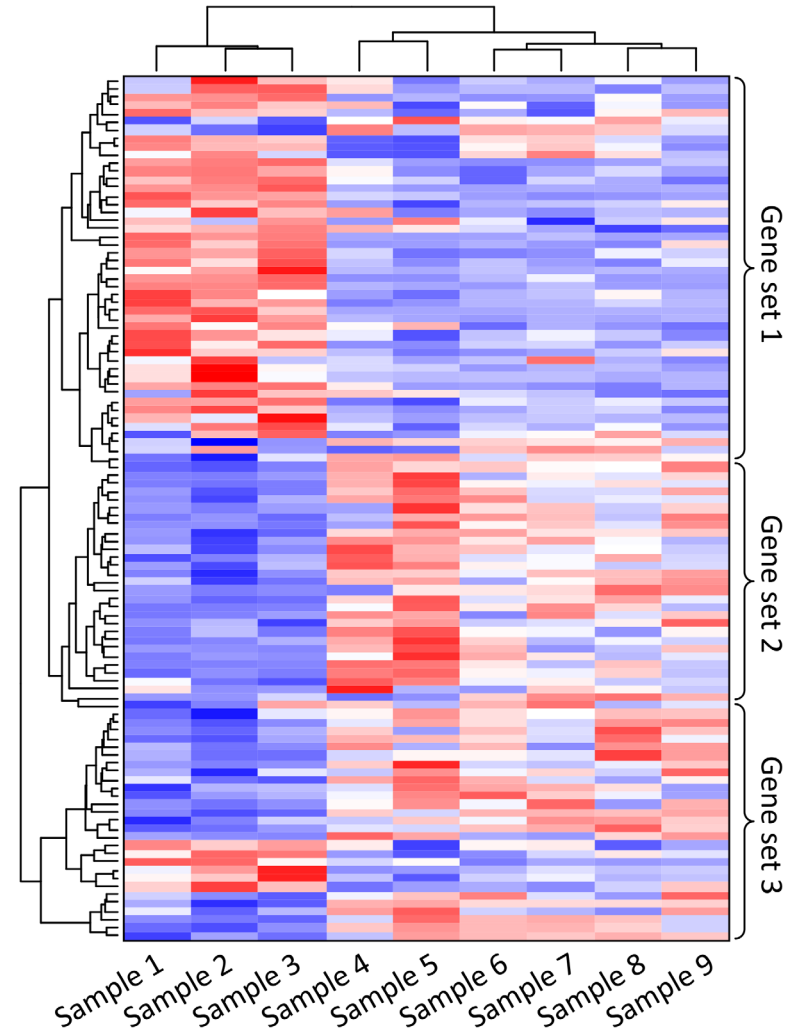
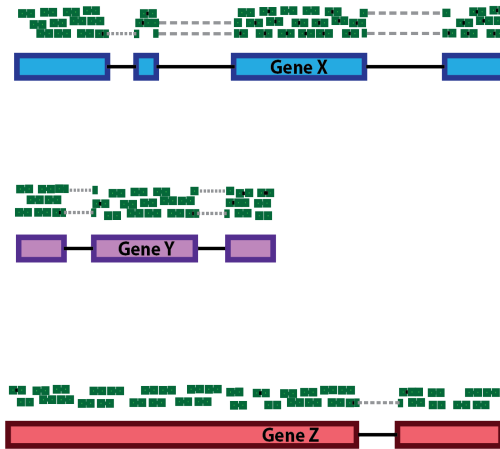
# 轉錄體定量分析與結果表示

熱圖 ( Heatmap )

Sample A Reads



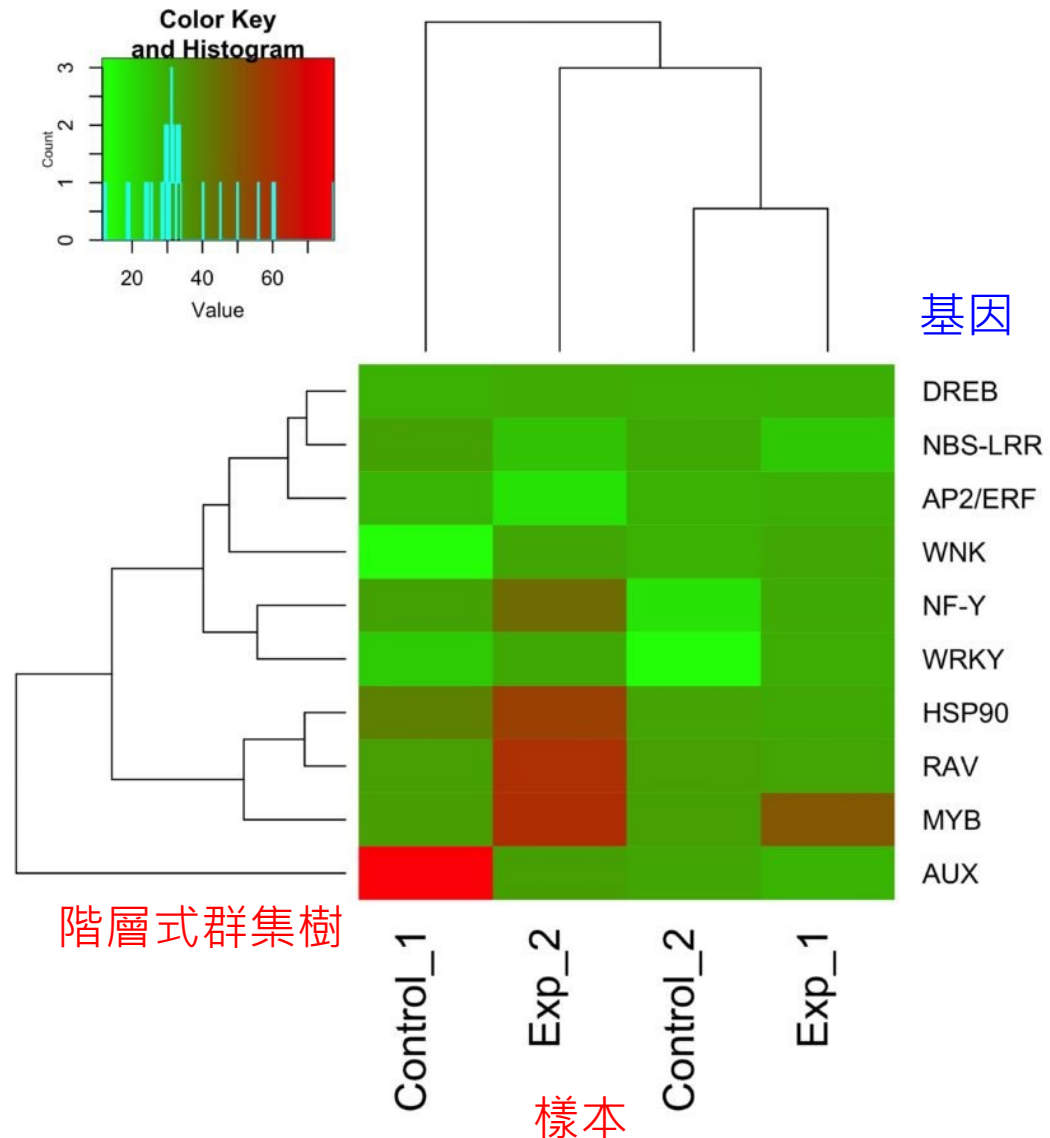
Sample B Reads



# 熱圖可用來展示群集分析的結果

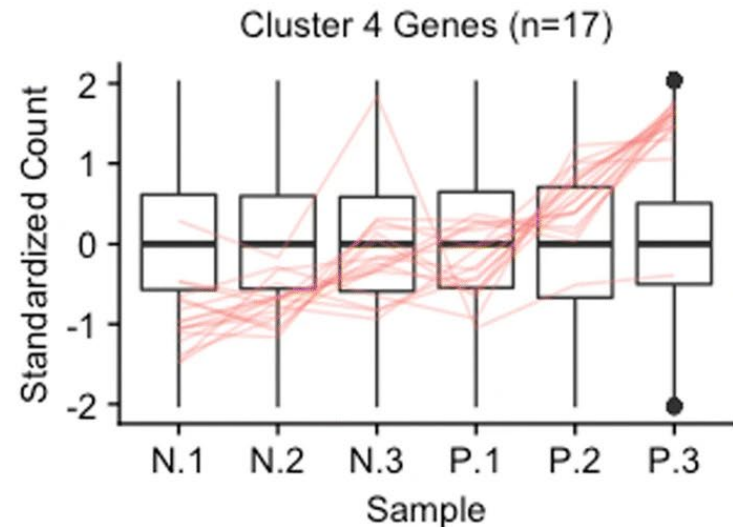
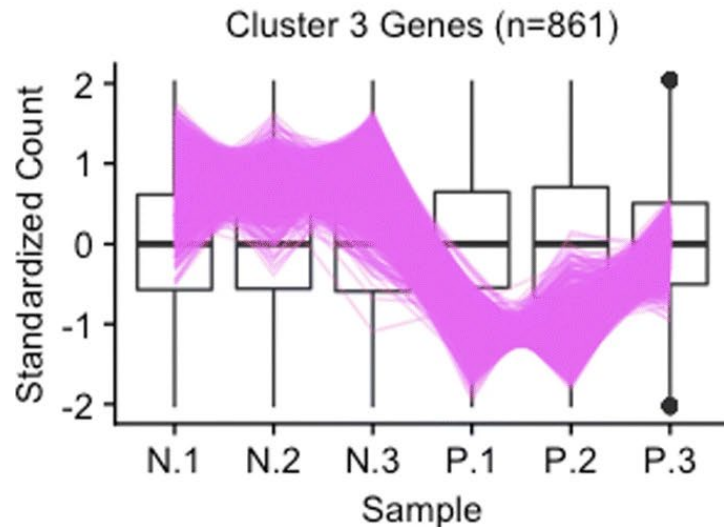
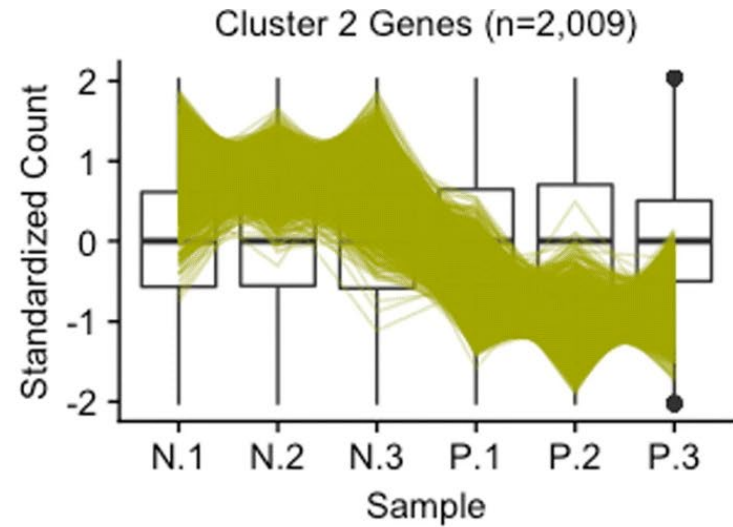
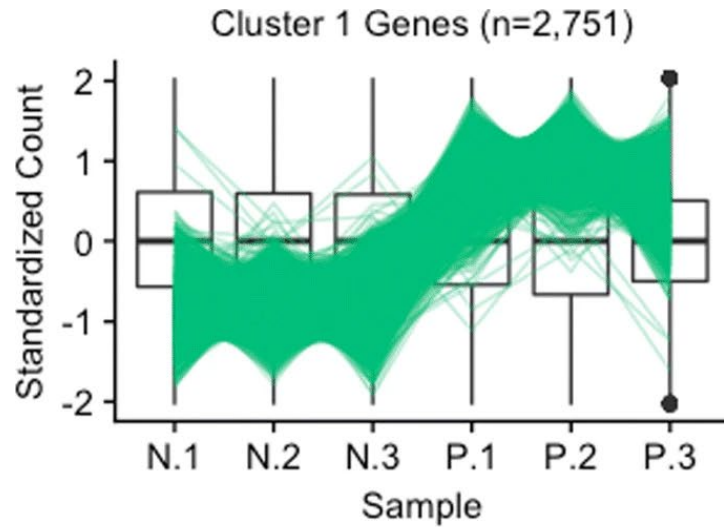
## ➤ 群集分析

- 將相似表現量的**基因**或**樣本**歸類於同一組





- 群集分析後的結果也可以使用平行坐標圖 ( Parallel Coordinates Plot ) 表示。



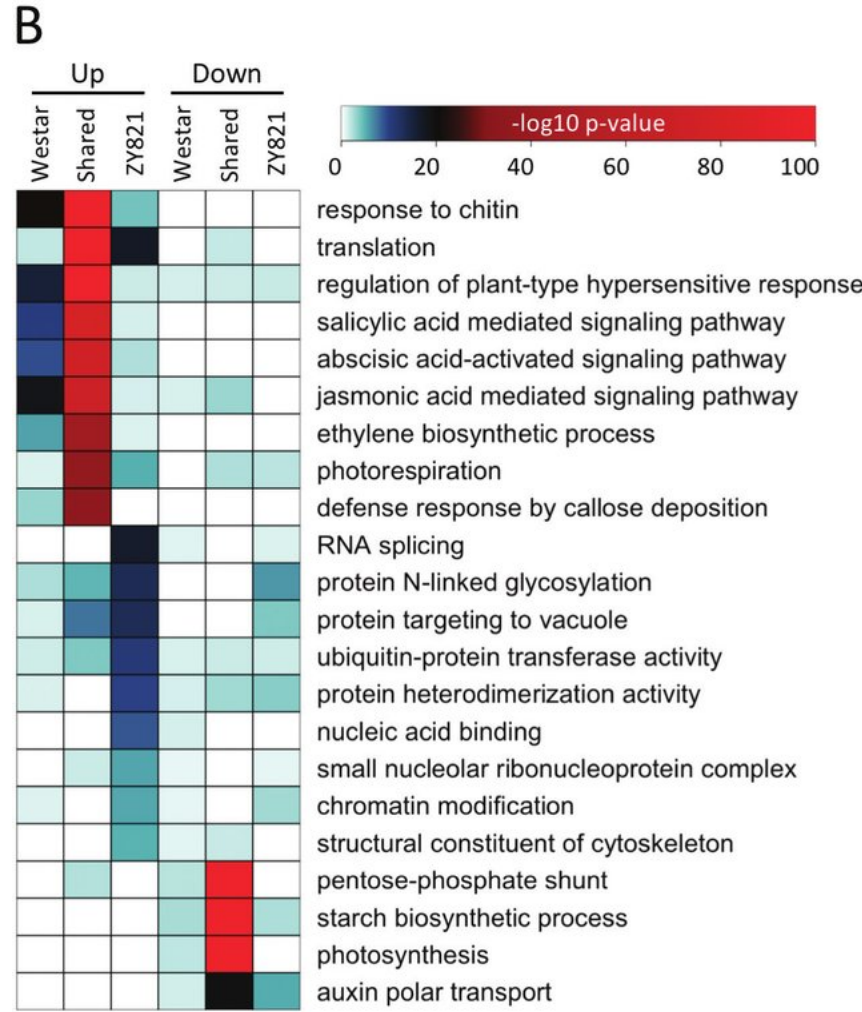
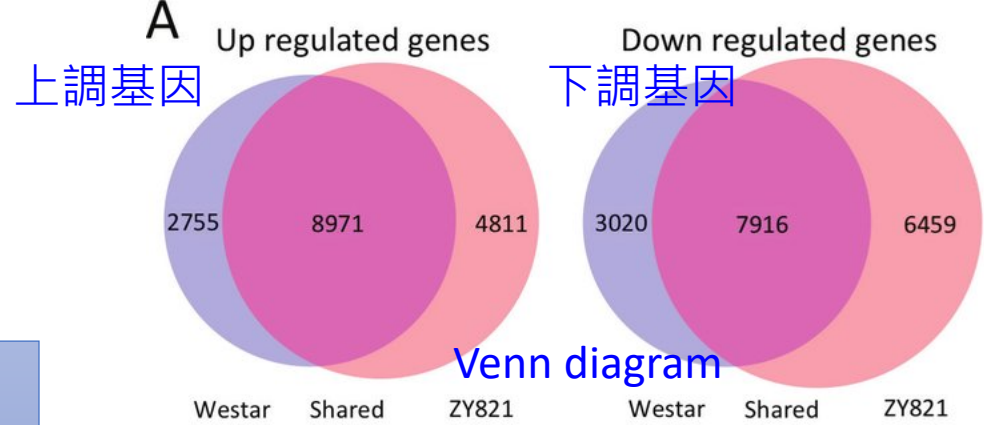
- 基因功能分析：
  - Gene ontology (GO) analysis

GO covers 3 domains:

- Cellular component 細胞組成
- Molecular function 分子功能
- Biological process 生物過程

Susceptible (Westar) and tolerant (ZY821) genotypes of *B. napus* infected with *S. sclerotiorum* 大油菜 菌核病菌

GIRARD, Ian J., et al. RNA sequencing of *Brassica napus* reveals cellular redox control of *Sclerotinia* infection. *Journal of Experimental Botany*, 2017, 68.18: 5079-5091.

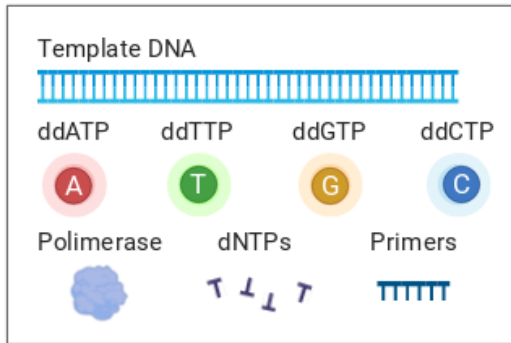


# 次世代定序

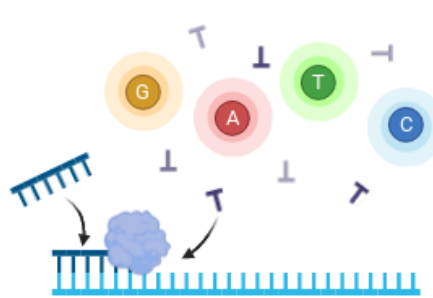
( NGS: next-generation sequencing )

- 又稱高通量定序 ( high throughput sequencing )
- 對比「第一代」Sanger定序：

## Reagents

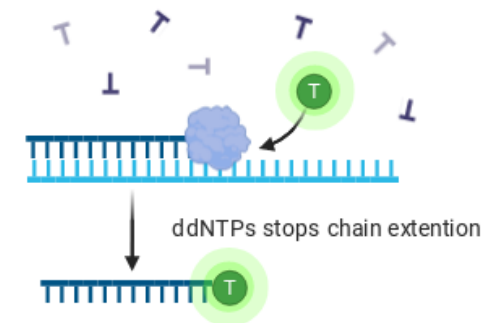


## ① Primer annealing and chain extension

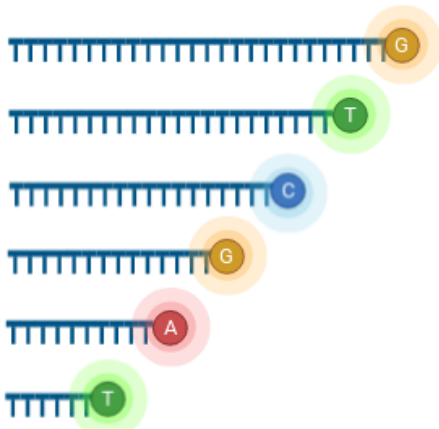


## Terminator Chemistry

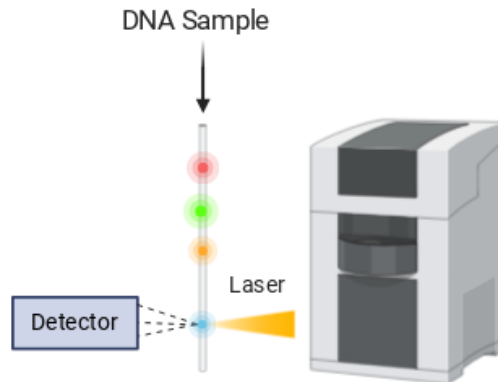
### ② ddNTP binding and chain termination



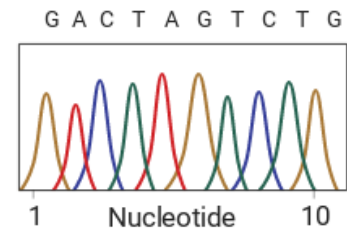
### ③ Fluorescently labelled DNA sample



### ④ Capillary gel electrophoresis and fluorescence detection



### ⑤ Sequence analysis and reconstruction



# 次世代定序：以基因體定序為例

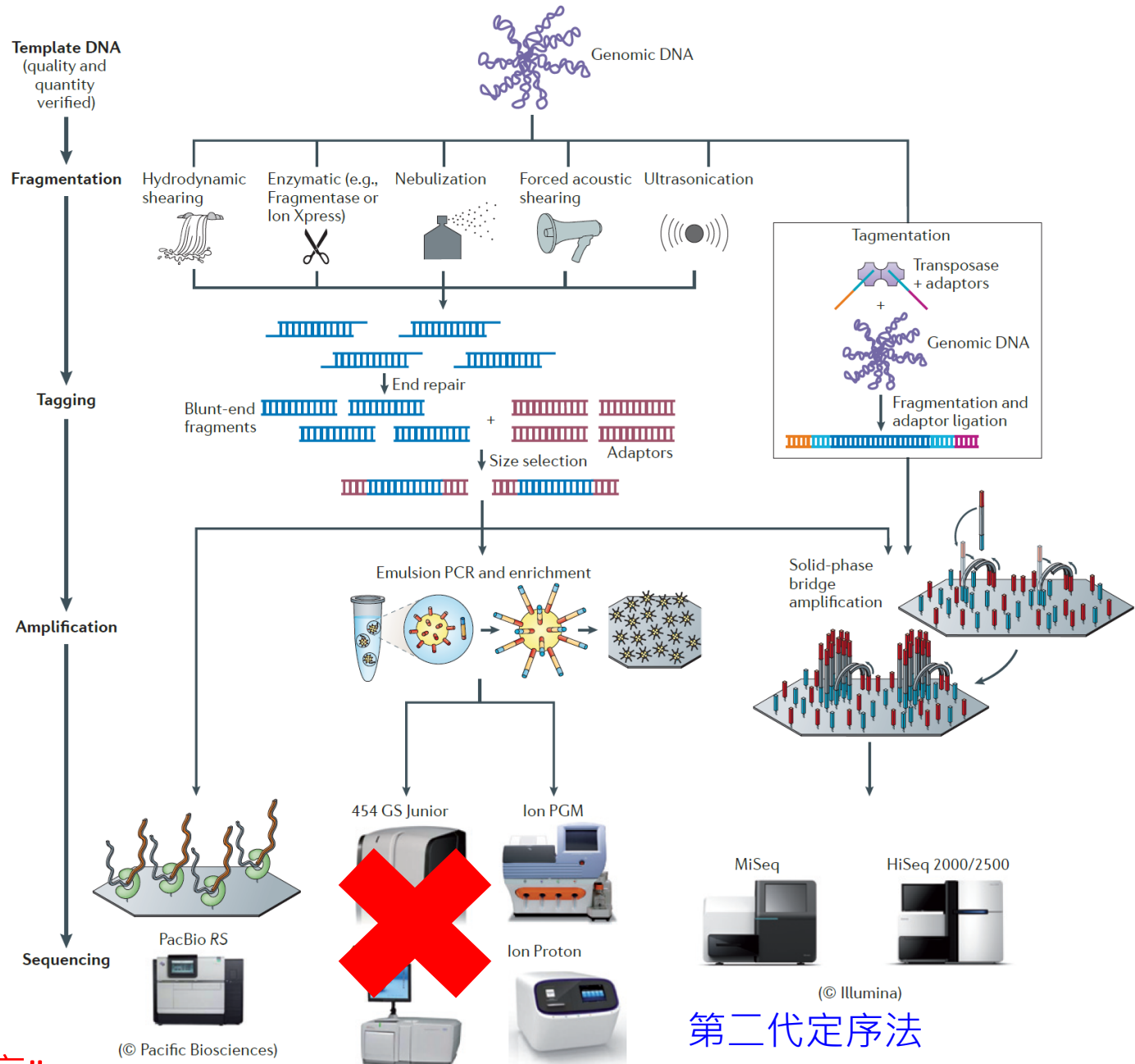
模版DNA

碎片化

標記

(DNA放大)

定序



“大量平行定序”

第三代定序法

第二代定序法

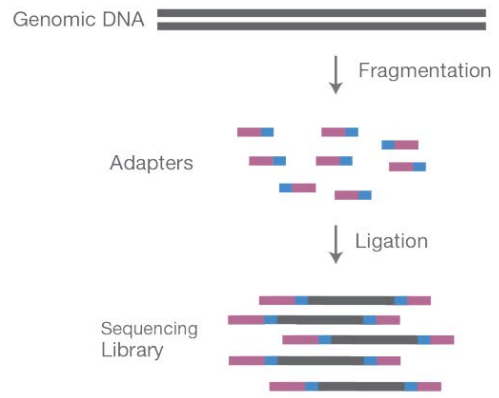
# 次世代定序特點

- 皆採用大量平行定序方式進行 ~1,000 bases per sample
  - 一個樣品中可以混雜不同序列的DNA
- 第二代定序法 > 1 Gb ( $10^9$  bases) per sample
  - 樣品仍需經過DNA擴增 ( amplification ) 的步驟
  - 可讀取的序列長度偏短 ( 約100-300 bp )
- 第三代定序法
  - 樣品**無需**經過DNA擴增的步驟
  - 可讀取的序列長度可達10-200 kb
  - 序列讀取錯誤率較第二代定序來得高



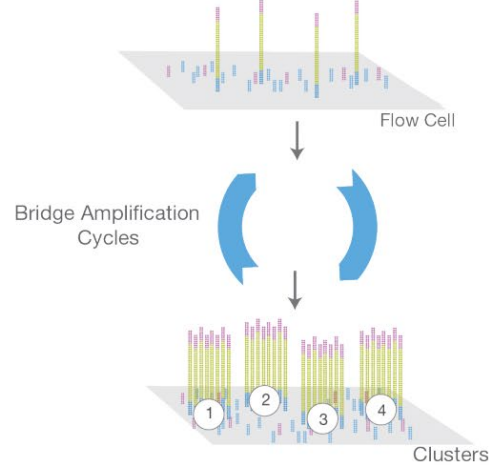
# Illumina 樣品製備 與定序原理

## A. Library Preparation 樣本庫製備



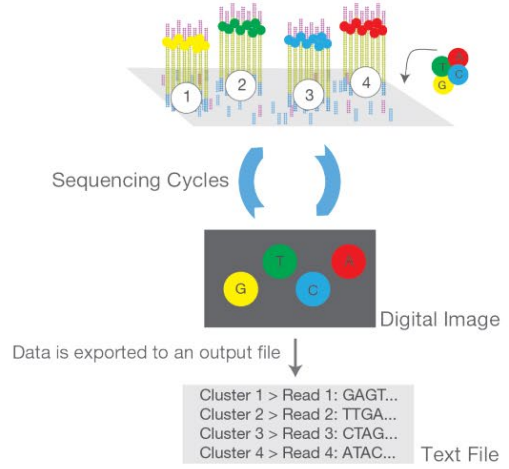
NGS library is prepared by fragmenting a gDNA sample and ligating specialized adapters to both fragment ends.

## B. Cluster Amplification 叢集擴增



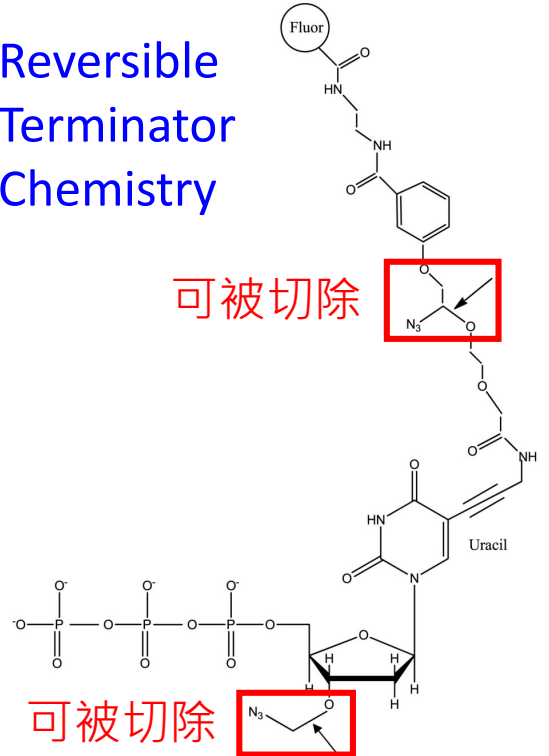
Library is loaded into a flow cell and the fragments are hybridized to the flow cell surface. Each bound fragment is amplified into a clonal cluster through bridge amplification.

## C. Sequencing 定序



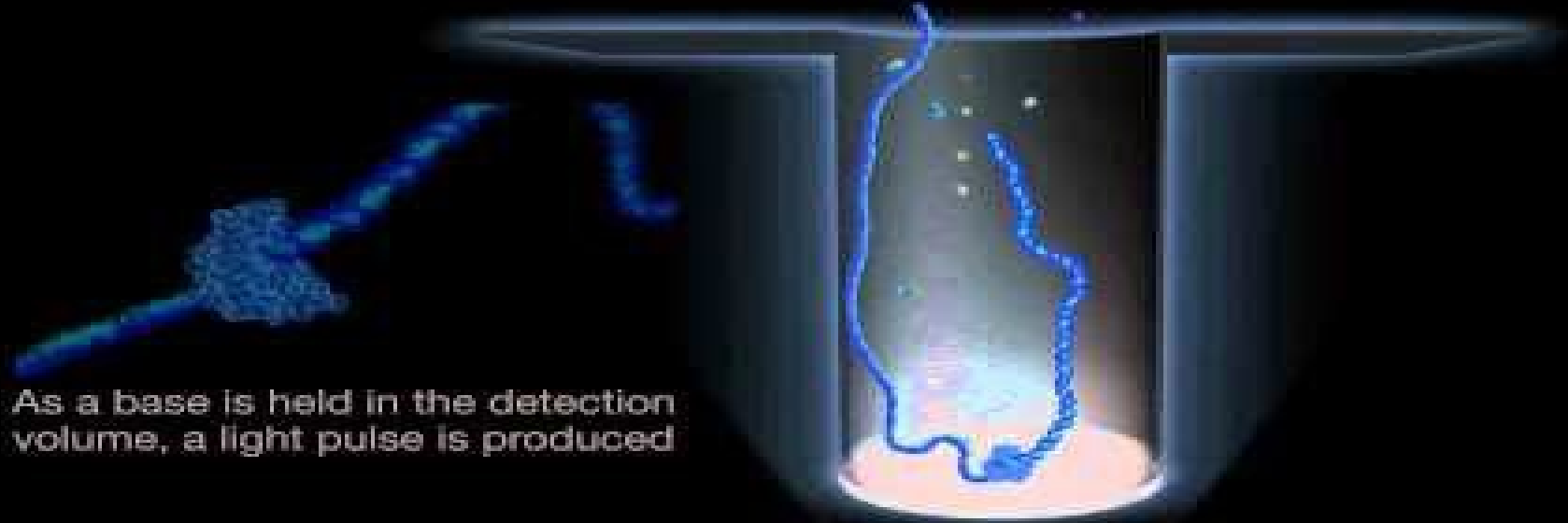
Sequencing reagents, including fluorescently labeled nucleotides, are added and the first base is incorporated. The flow cell is imaged and the emission from each cluster is recorded. The emission wavelength and intensity are used to identify the base. This cycle is repeated "n" times to create a read length of "n" bases.

## Reversible Terminator Chemistry



# 第三代定序法(I)：

PacBio SMRT (Single Molecule Real Time) sequencing (2010-)

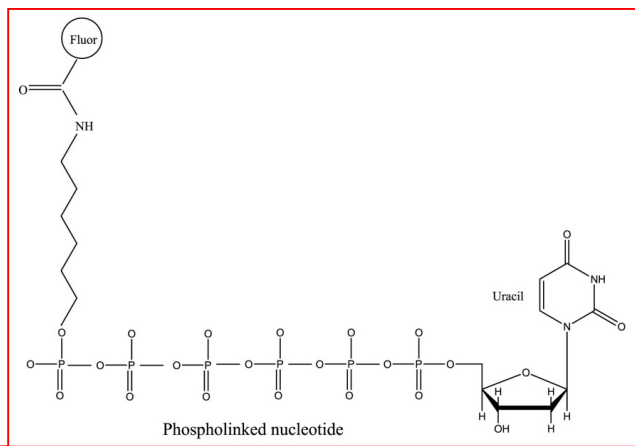


As a base is held in the detection volume, a light pulse is produced

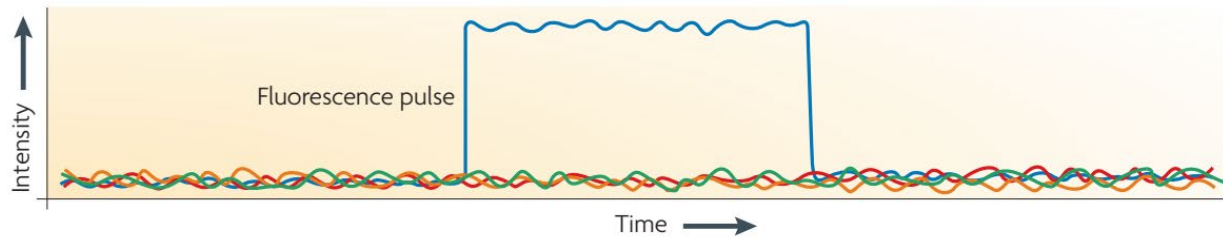
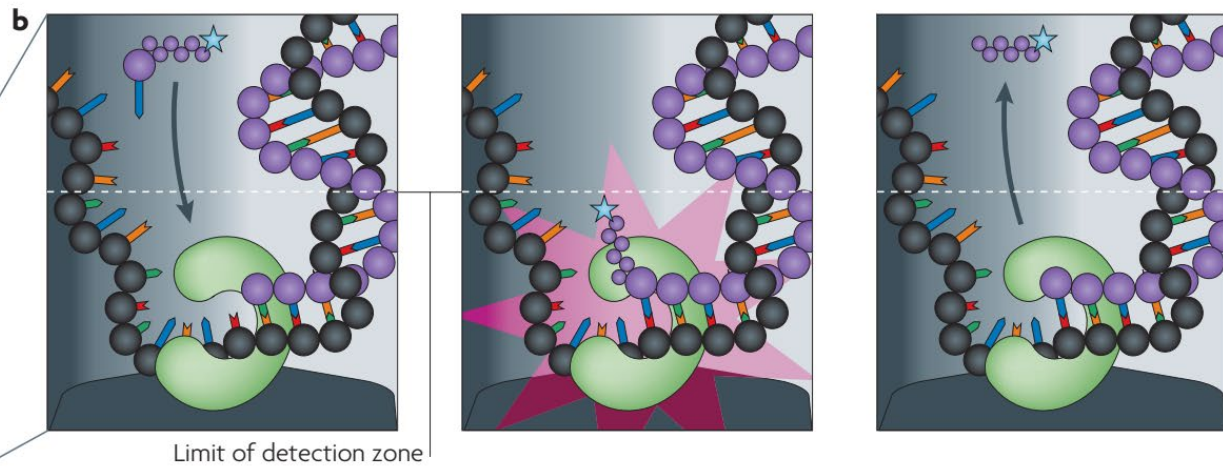
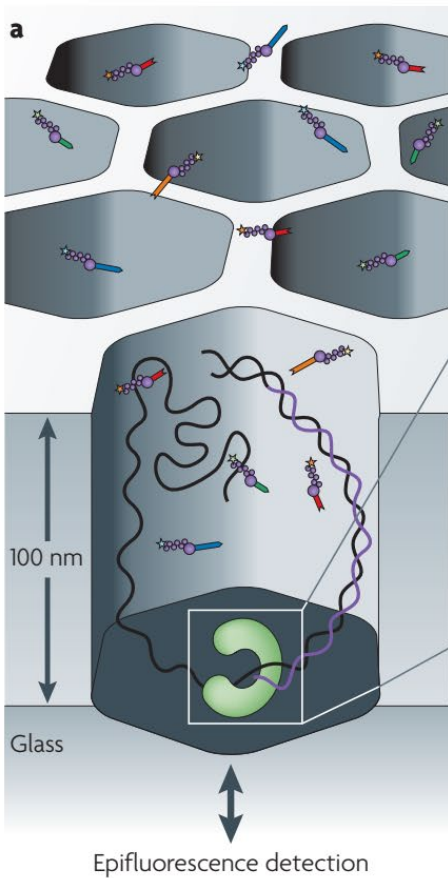
<https://youtu.be/NHCJ8PtYCFc>



# PacBio SMART 定序原理



## Pacific Biosciences — Real-time sequencing





# 第三代定序法(II)： Oxford Nanopore sequencing (2015-)

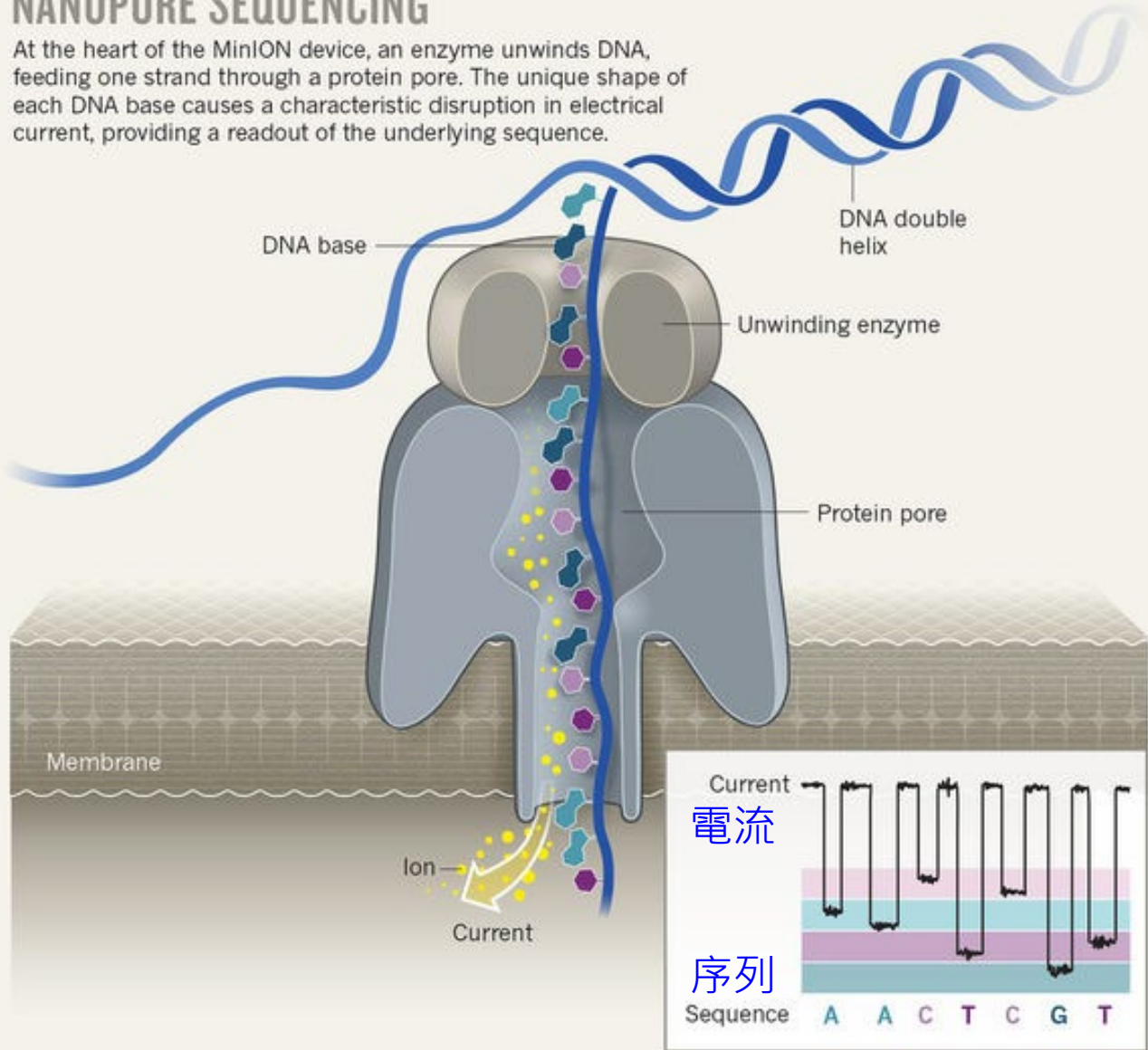


<https://youtu.be/GUb1TZvMWsw>

# Oxford Nanopore 定序原理

## NANOPORE SEQUENCING

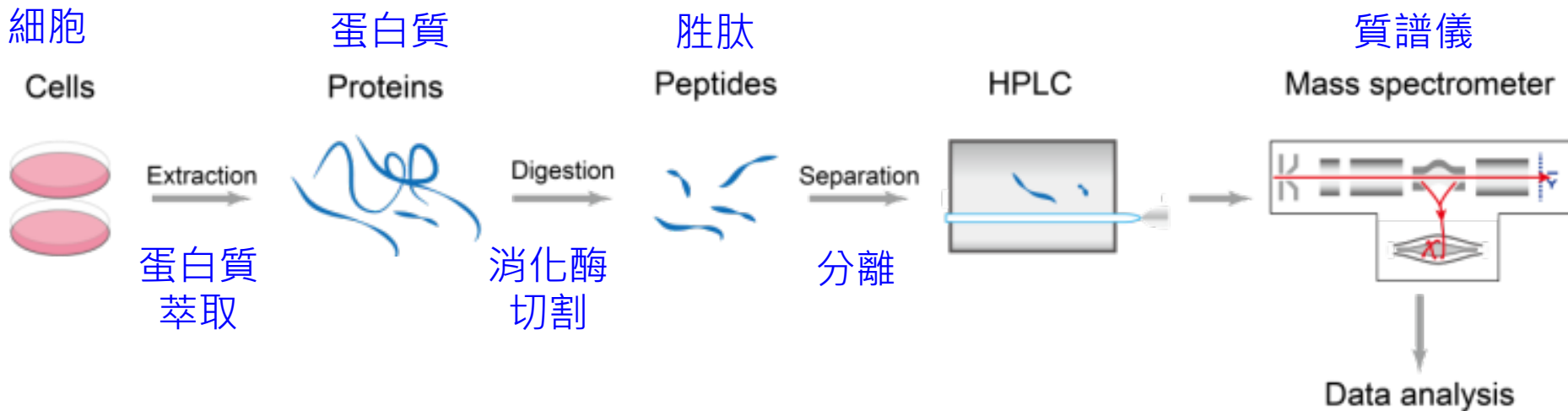
At the heart of the MinION device, an enzyme unwinds DNA, feeding one strand through a protein pore. The unique shape of each DNA base causes a characteristic disruption in electrical current, providing a readout of the underlying sequence.



Stoddart, David, et al. "Single-nucleotide discrimination in immobilized DNA oligonucleotides with a biological nanopore." PNAS 106.19 (2009): 7702-7707.

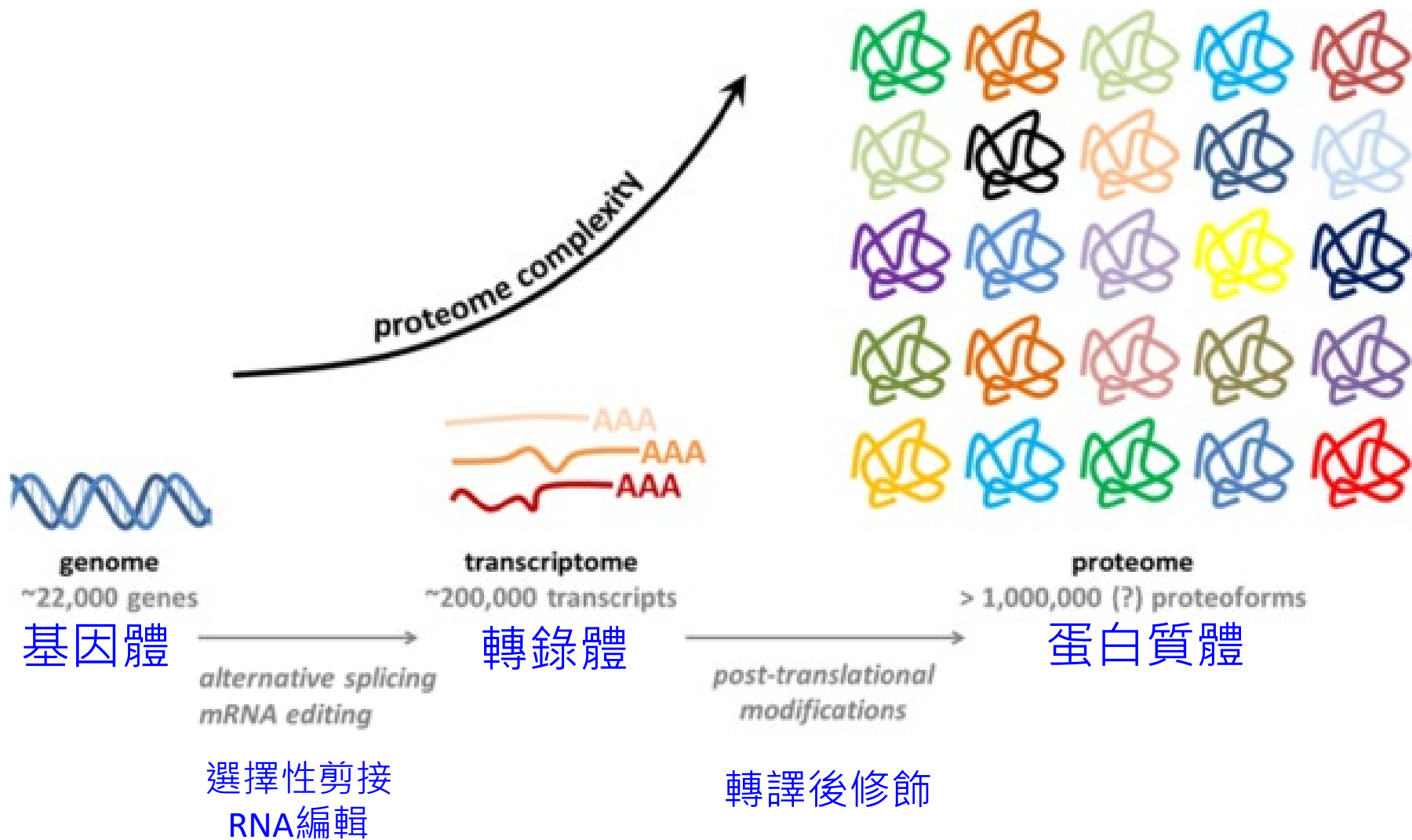
# 蛋白質體學

- 分析樣品中**所有的(?)**蛋白質組成
  - 定性：鑑定樣品中蛋白質種類
  - 定量：比較不同樣品中蛋白質表現變化
- 常見基本流程：

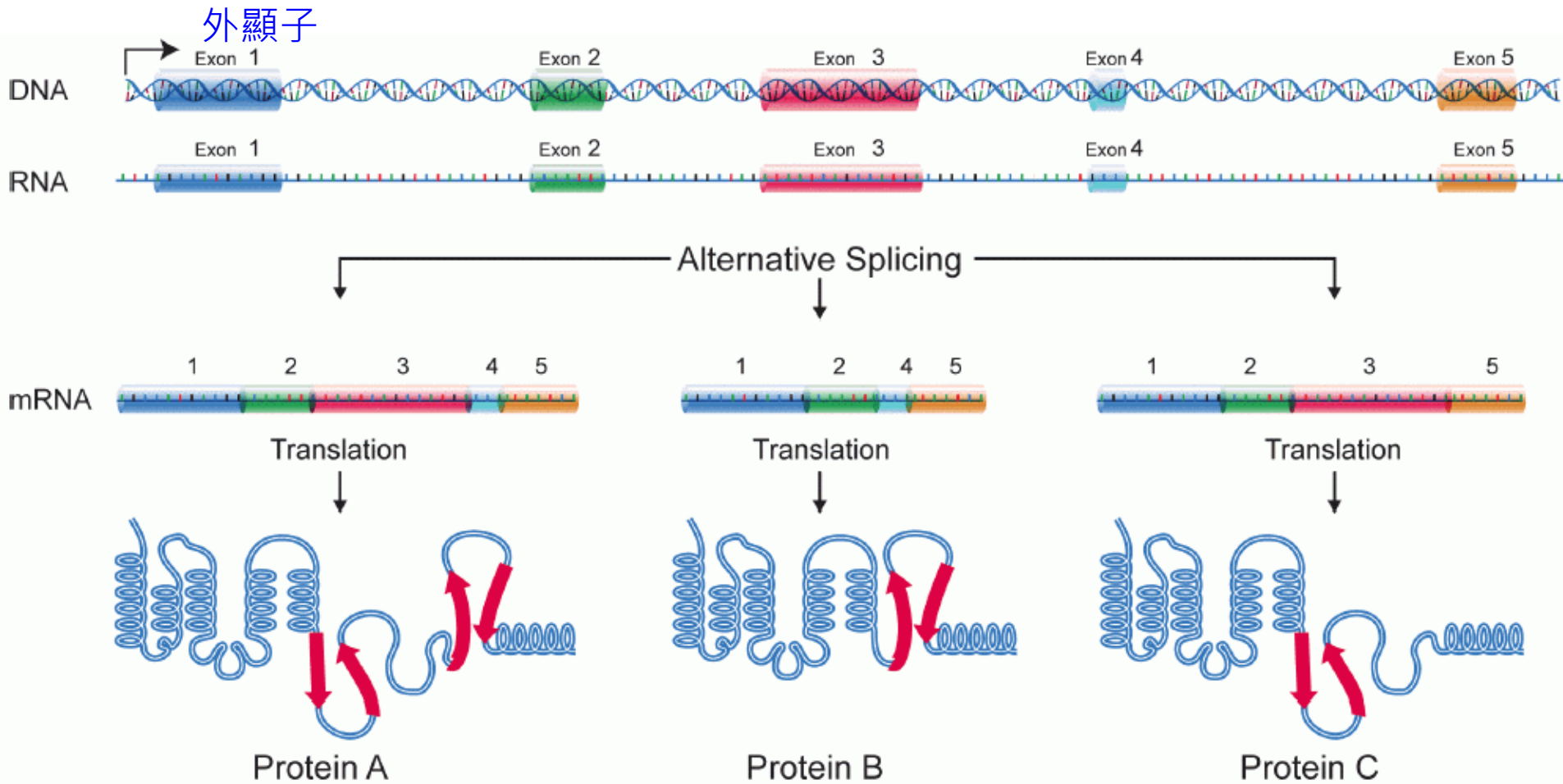


利用質量大小  
鑑定胜肽序列

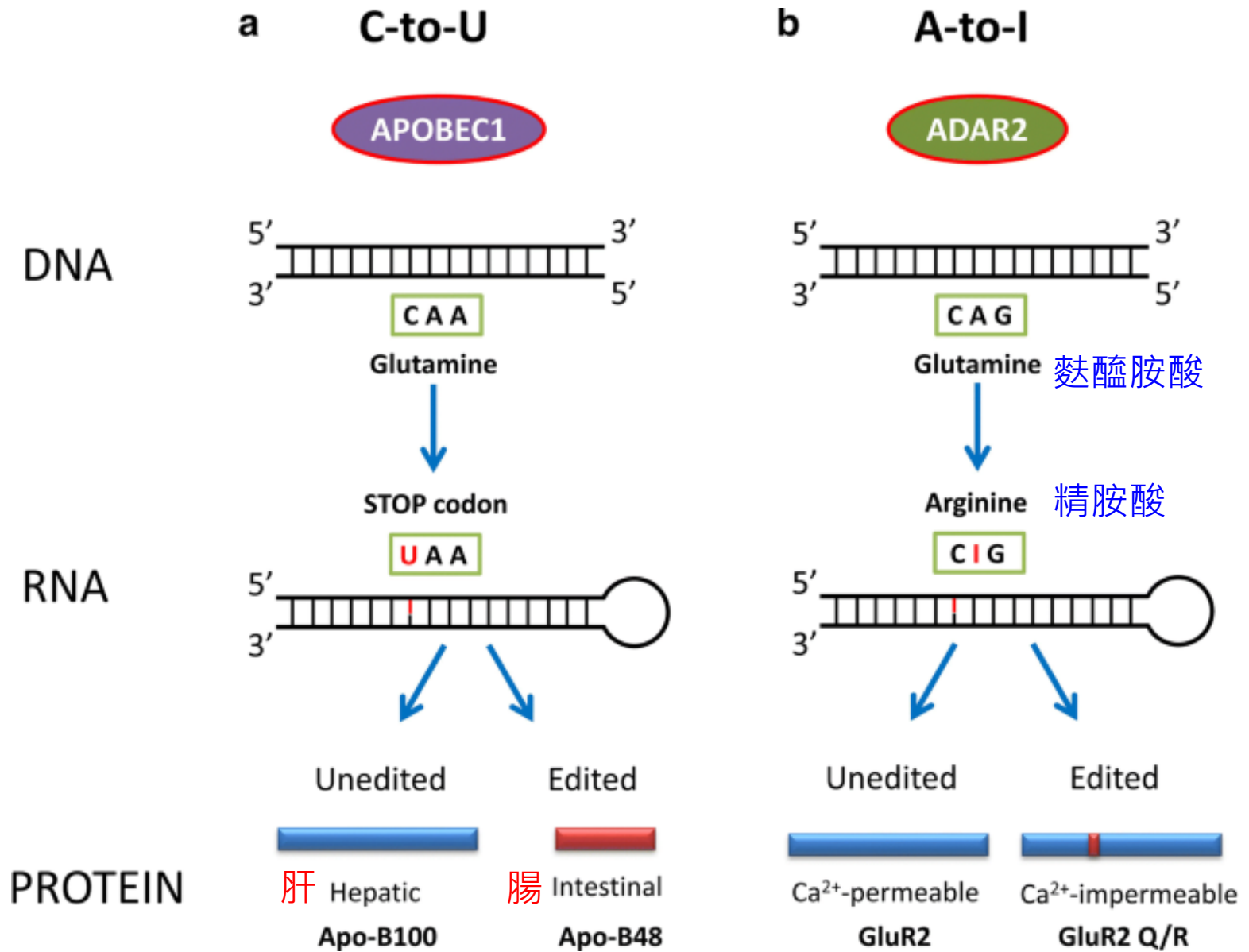
# 蛋白質體的複雜性



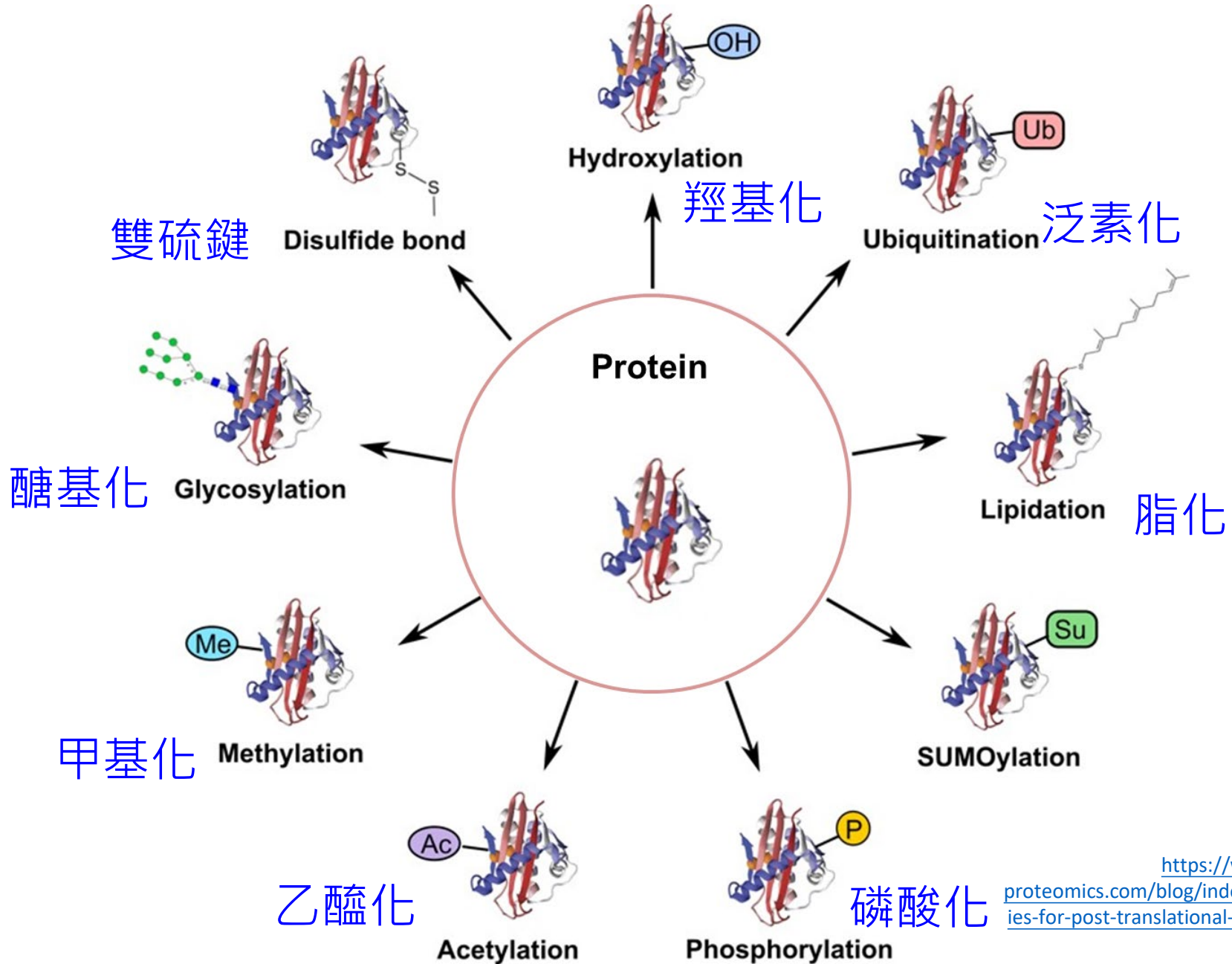
# 選擇性剪接 ( Alternative splicing )



# RNA編輯

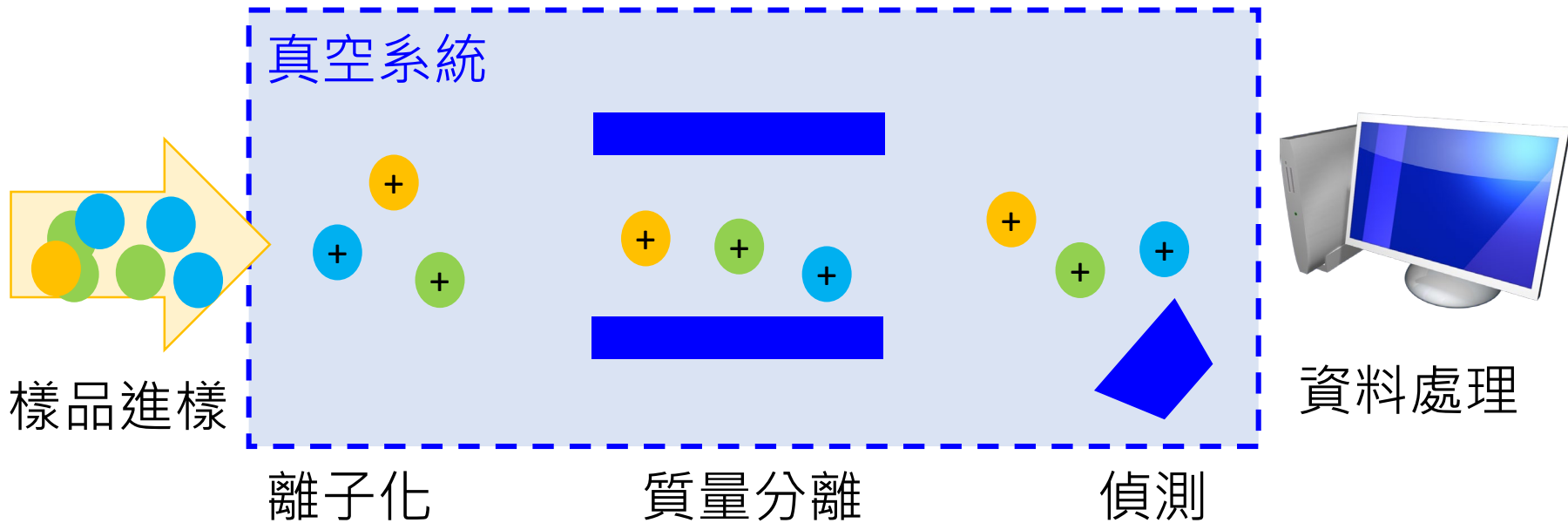


# 轉譯後蛋白質胺基酸的修飾



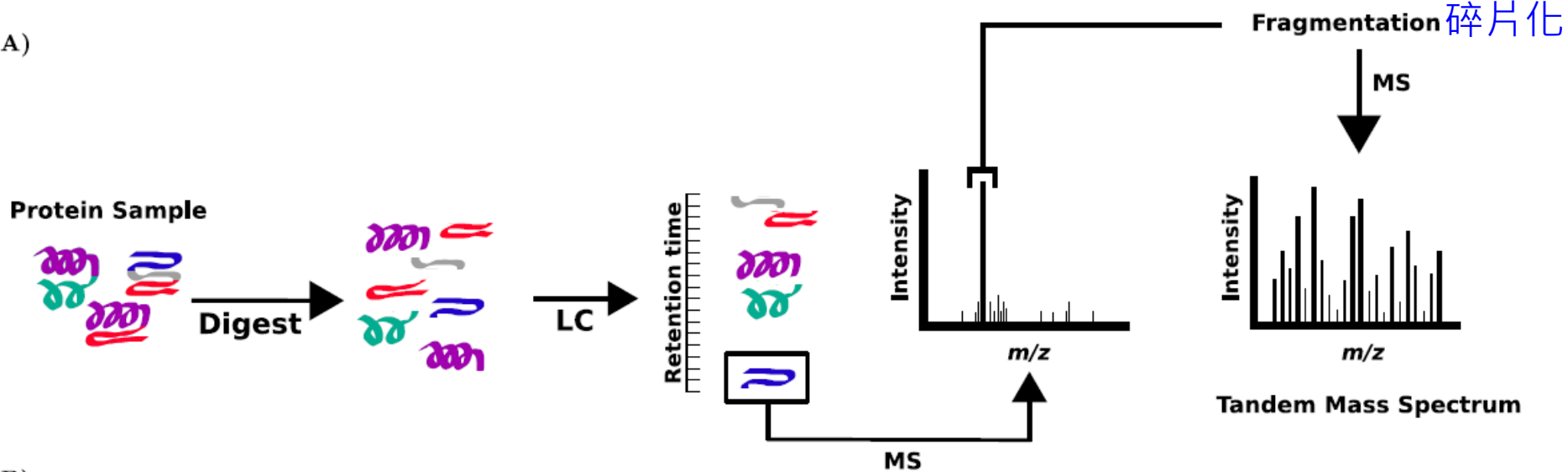


# 質譜法 ( mass spectrometry ) : 蛋白質體學關鍵技術

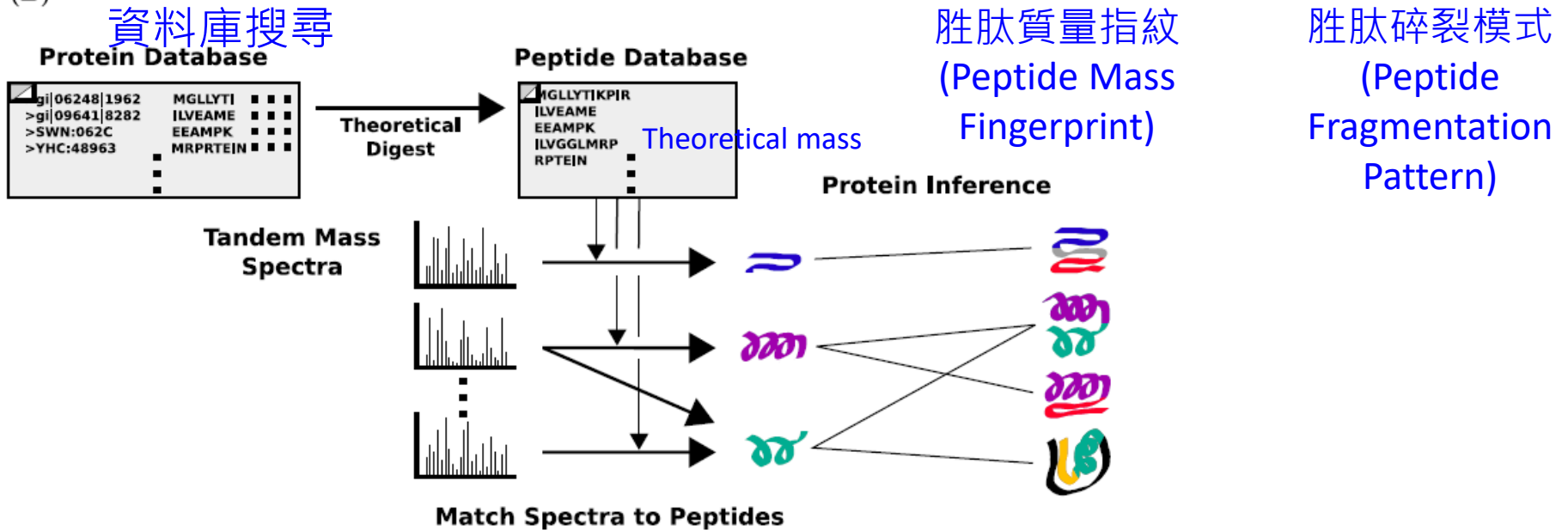


# 利用串聯質譜法進行蛋白質分析

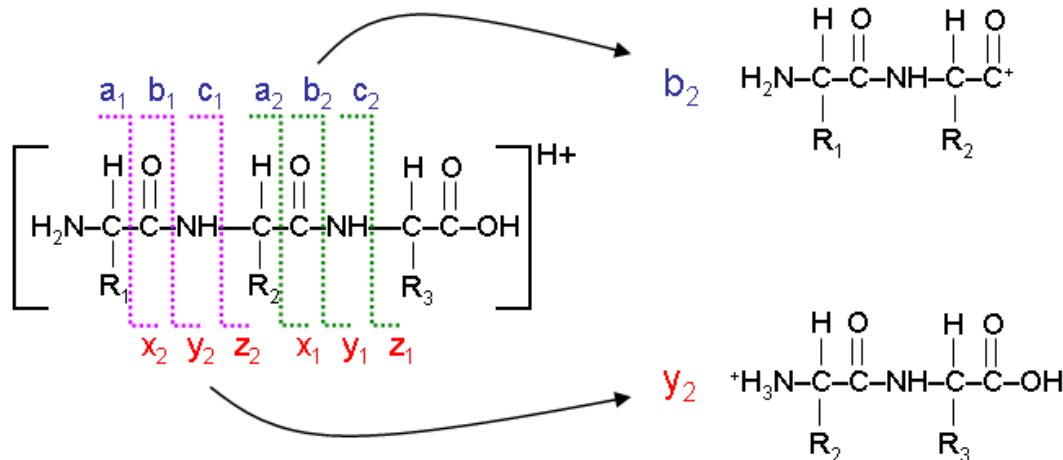
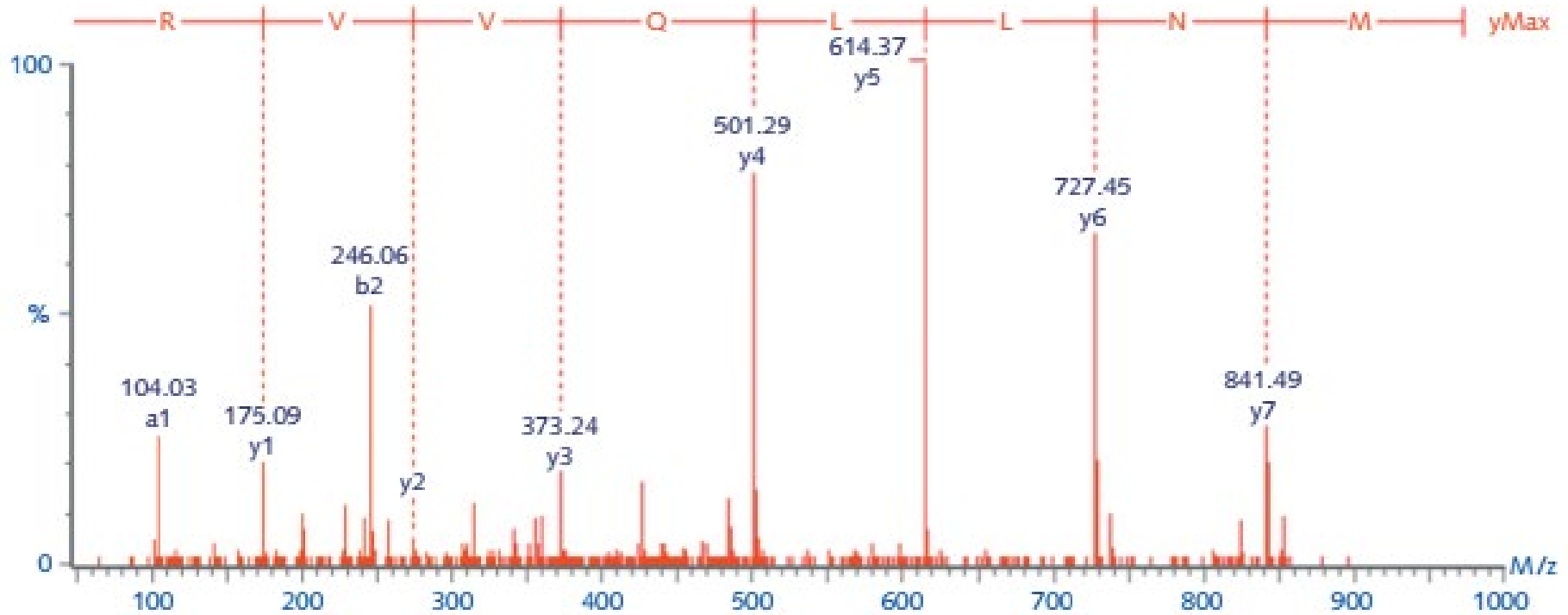
(A)



(B)



# 串聯質譜(MS/MS)定序



<http://www.chemistmatt.com/MwtHelp/PeptideFragModelling.htm>

# 資料庫搜尋與比對：Mascot

胜肽質量指紋  
( MS )

<b>Database(s)</b>	SwissProt NCBIInr contaminants cRAP MSDB	<b>Enzyme</b>	Trypsin
		<b>Allow up to</b>	1 missed cleavages
<b>Taxonomy</b>	All entries		
<b>Fixed modifications</b>	--- none selected ---	> <	Acetyl (K) Acetyl (N-term) Acetyl (Protein N-term) Amidated (C-term) Amidated (Protein C-term) Ammonia-loss (N-term C) Biotin (K) Biotin (N-term) Carbamyl (K) Carbamyl (N-term) Carboxymethyl (C)
	Display all modifications <input type="checkbox"/>		
<b>Variable modifications</b>	Carbamidomethyl (C) Deamidated (NQ) Oxidation (M)	> <	
<b>Protein mass</b>	<input type="text"/> kDa	<b>Peptide tol. ±</b>	0.3 Da
<b>Mass values</b>	<input checked="" type="radio"/> MH <sup>+</sup> <input type="radio"/> M <sub>r</sub> <input type="radio"/> M-H <sup>-</sup>	<b>Monoisotopic</b>	<input checked="" type="radio"/> Average <input type="radio"/>
<b>Data file</b>	<input type="text"/>	瀏覽...	
<b>Query</b>	NB Contents of this field are ignored if a data file is specified.		
<b>Decoy</b>	<input type="checkbox"/>	<b>Report top</b>	AUTO hits
<b>Start Search ...</b>		<b>Reset Form</b>	

# 串聯質譜(MS/MS)搜尋

<b>Your name</b>	<input type="text" value="jimmy"/>	<b>Email</b>	<input type="text" value="jimmy.tsai@mass-solutions.com"/>
<b>Search title</b>	<input type="text"/>		
<b>Database(s)</b>	<input type="text" value="Human_EST"/> <input type="text" value="Fungi_EST"/> <input type="text" value="Environmental_EST"/> <input type="text" value="SwissProt"/> <input type="text" value="NCBIInr"/>	<b>Enzyme</b>	<input type="text" value="Trypsin"/>
		<b>Allow up to</b>	<input type="text" value="1"/> missed cleavages
		<b>Quantitation</b>	<input type="text" value="None"/>
<b>Taxonomy</b>	<input type="text" value="All entries"/>		
<b>Fixed modifications</b>	<input type="text" value="--- none selected ---"/>	<input type="button" value="&gt;"/>	<input type="text" value="Acetyl (K)"/> <input type="text" value="Acetyl (N-term)"/> <input type="text" value="Acetyl (Protein N-term)"/> <input type="text" value="Amidated (C-term)"/> <input type="text" value="Amidated (Protein C-term)"/> <input type="text" value="Ammonia-loss (N-term C)"/> <input type="text" value="Biotin (K)"/> <input type="text" value="Biotin (N-term)"/> <input type="text" value="Carbamyl (K)"/> <input type="text" value="Carbamyl (N-term)"/> <input type="text" value="Carboxymethyl (C)"/>
	<input type="button" value="&lt;"/>		
<b>Variable modifications</b>	<input type="text" value="Carbamidomethyl (C)"/> <input type="text" value="Deamidated (NQ)"/> <input type="text" value="Oxidation (M)"/>	<input type="button" value="&gt;"/> <input type="button" value="&lt;"/>	
<b>Peptide tol. ±</b>	<input type="text" value="0.3"/> <input type="text" value="Da"/> # <sup>13</sup> C <input type="text" value="0"/>	<b>MS/MS tol. ±</b>	<input type="text" value="0.3"/> <input type="text" value="Da"/>
<b>Peptide charge</b>	<input type="text" value="1+, 2+ and 3+"/>	<b>Monoisotopic</b>	<input checked="" type="radio"/> Average <input type="radio"/>
<b>Data file</b>	<input type="text"/> <input type="button" value="瀏覽..."/>		
<b>Data format</b>	<input type="text" value="Micromass (.PKL)"/>	<b>Precursor</b>	<input type="text"/> m/z
<b>Instrument</b>	<input type="text" value="ESI-QUAD-TOF"/>	<b>Error tolerant</b>	<input type="checkbox"/>
<b>Decoy</b>	<input type="checkbox"/>	<b>Report top</b>	<input type="text" value="AUTO"/> hits
<input type="button" value="Start Search ..."/>		<input type="button" value="Reset Form"/>	

1. [CH60\\_HUMAN](#) Mass: 61016 Score: 1225 Matches: 31(27) Sequences: 19(17)  
 60 kDa heat shock protein, mitochondrial precursor (Hsp60) (60 kDa chaperonin) (CPN60) (Heat shock

Check to include this hit in error tolerant search

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Expect	Rank	Unique	Peptide
<input checked="" type="checkbox"/> <a href="#">11</a>	417.1822	832.3498	832.3828	-0.0329	0	45	0.016	1		K.APGFGDNR.K
<input checked="" type="checkbox"/> <a href="#">12</a>	422.7433	843.4720	843.5066	-0.0346	0	46	0.017	1	U	K.VGEVIVTK.D
<input checked="" type="checkbox"/> <a href="#">13</a>	430.7328	859.4510	859.4837	-0.0327	0	36	0.15	1	U	K.IPAMTIAK.N + Oxidation (M)
<input checked="" type="checkbox"/> <a href="#">15</a>	451.2499	900.4853	900.5280	-0.0428	0	52	0.0039	1	U	K.LSDGVAVLK.V
<input checked="" type="checkbox"/> <a href="#">16</a>	456.7806	911.5467	911.5804	-0.0337	0	59	0.00056	1	U	K.VGLQVVAVK.A
<input checked="" type="checkbox"/> <a href="#">21</a>	480.7447	959.4748	959.5036	-0.0288	0	45	0.017	1	U	R.VTDALNATR.A
<input checked="" type="checkbox"/> <a href="#">24</a>	595.7855	1189.5565	1189.6012	-0.0447	0	(57)	0.0011	1	U	K.EIGNIISDAMK.K
<input checked="" type="checkbox"/> <a href="#">25</a>	603.7720	1205.5294	1205.5962	-0.0668	0	60	0.00048	1	U	K.EIGNIISDAMK.K + Oxidation
<input checked="" type="checkbox"/> <a href="#">26</a>	608.3099	1214.6052	1214.6507	-0.0455	0	73	2.2e-05	1	U	K.NAGVEGSLIVEK.I
<input checked="" type="checkbox"/> <a href="#">27</a>	617.2857	1232.5569	1232.5885	-0.0316	0	81	4e-06	1	U	K.VGGTSDVEVNEK.K
<input checked="" type="checkbox"/> <a href="#">31</a>	672.8375	1343.6605	1343.7085	-0.0480	0	64	0.00016	1	U	R.TVIIEQSWGSPK.V
<input checked="" type="checkbox"/> <a href="#">34</a>	714.8884	1427.7623	1427.8058	-0.0435	0	(65)	0.00014	1	U	R.GVMLAVDAVIAELK.K
<input checked="" type="checkbox"/> <a href="#">35</a>	714.8938	1427.7730	1427.8058	-0.0327	0	(73)	2.1e-05	1	U	R.GVMLAVDAVIAELK.K
<input checked="" type="checkbox"/> <a href="#">36</a>	722.8849	1443.7552	1443.8007	-0.0455	0	75	1.2e-05	1	U	R.GVMLAVDAVIAELK.K + Oxidati
<input checked="" type="checkbox"/> <a href="#">37</a>	722.8934	1443.7722	1443.8007	-0.0285	0	(73)	2.2e-05	1	U	R.GVMLAVDAVIAELK.K + Oxidati
<input checked="" type="checkbox"/> <a href="#">39</a>	752.8643	1503.7141	1503.7490	-0.0349	0	90	4.3e-07	1	U	K.TLNDELEIEGGMK.F
<input checked="" type="checkbox"/> <a href="#">40</a>	760.8461	1519.6777	1519.7439	-0.0662	0	(89)	4.7e-07	1	U	K.TLNDELEIEGGMK.F + Oxidati
<input checked="" type="checkbox"/> <a href="#">45</a>	640.3281	1917.9625	1918.0636	-0.1010	0	102	2.1e-08	1	U	K.ISSIQSIVPALEIANAHR.K
<input checked="" type="checkbox"/> <a href="#">46</a>	960.0327	1918.0509	1918.0636	-0.0127	0	(87)	5.1e-07	1	U	K.ISSIQSIVPALEIANAHR.K
<input checked="" type="checkbox"/> <a href="#">48</a>	1019.5106	2037.0067	2037.0153	-0.0087	0	52	0.0015	1	U	R.IQEIEQLDVTTSEYEK.E
<input checked="" type="checkbox"/> <a href="#">51</a>	1057.0537	2112.0929	2112.1323	-0.0394	0	116	6.8e-10	1	U	R.ALMLQGVDLLADAVAVTMGPK.G
<input checked="" type="checkbox"/> <a href="#">52</a>	1065.0399	2128.0653	2128.1272	-0.0619	0	(72)	1.7e-05	1	U	R.ALMLQGVDLLADAVAVTMGPK.G +
<input checked="" type="checkbox"/> <a href="#">53</a>	1065.0623	2128.1100	2128.1272	-0.0172	0	(26)	0.63	1	U	R.ALMLQGVDLLADAVAVTMGPK.G +
<input checked="" type="checkbox"/> <a href="#">54</a>	1073.0477	2144.0809	2144.1221	-0.0412	0	(93)	1.3e-07	1	U	R.ALMLQGVDLLADAVAVTMGPK.G +

## Protein View



1. **CH60\_HUMAN** Mass: 61 60 kDa heat shock protein, mitochondrial precursor (Hsp60) (60 kDa cl  
60 kDa heat shock prot Found in search of C:\Auto MSMS output\Sample 1.pkl

Check to include this : Nominal mass ( $M_r$ ): 61016; Calculated pI value: 5.70  
NCBI BLAST search of CH60\_HUMAN against nr  
Unformatted sequence string for pasting into other applications

Query Observed Mr (e Taxonomy: Homo sapiens  
Variable modifications: Oxidation (M)  
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P  
**Sequence Coverage: 49%** 序列覆盖率

Matched peptides shown in **Bold Red**

Query	Observed	Mr (e)	Peptide Sequence
<input checked="" type="checkbox"/>	<a href="#">11</a>	417.1822	832.
<input checked="" type="checkbox"/>	<a href="#">12</a>	422.7433	843.
<input checked="" type="checkbox"/>	<a href="#">13</a>	430.7328	859.
<input checked="" type="checkbox"/>	<a href="#">15</a>	451.2499	900.
<input checked="" type="checkbox"/>	<a href="#">16</a>	456.7806	911.
<input checked="" type="checkbox"/>	<a href="#">21</a>	480.7447	959.
<input checked="" type="checkbox"/>	<a href="#">24</a>	595.7855	1189.
<input checked="" type="checkbox"/>	<a href="#">25</a>	603.7720	1205.
<input checked="" type="checkbox"/>	<a href="#">26</a>	608.3099	1214.
<input checked="" type="checkbox"/>	<a href="#">27</a>	617.2857	1232.
<input checked="" type="checkbox"/>	<a href="#">31</a>	672.8375	1343.
<input checked="" type="checkbox"/>	<a href="#">34</a>	714.8884	1427.
<input type="checkbox"/>	--	----	----

```

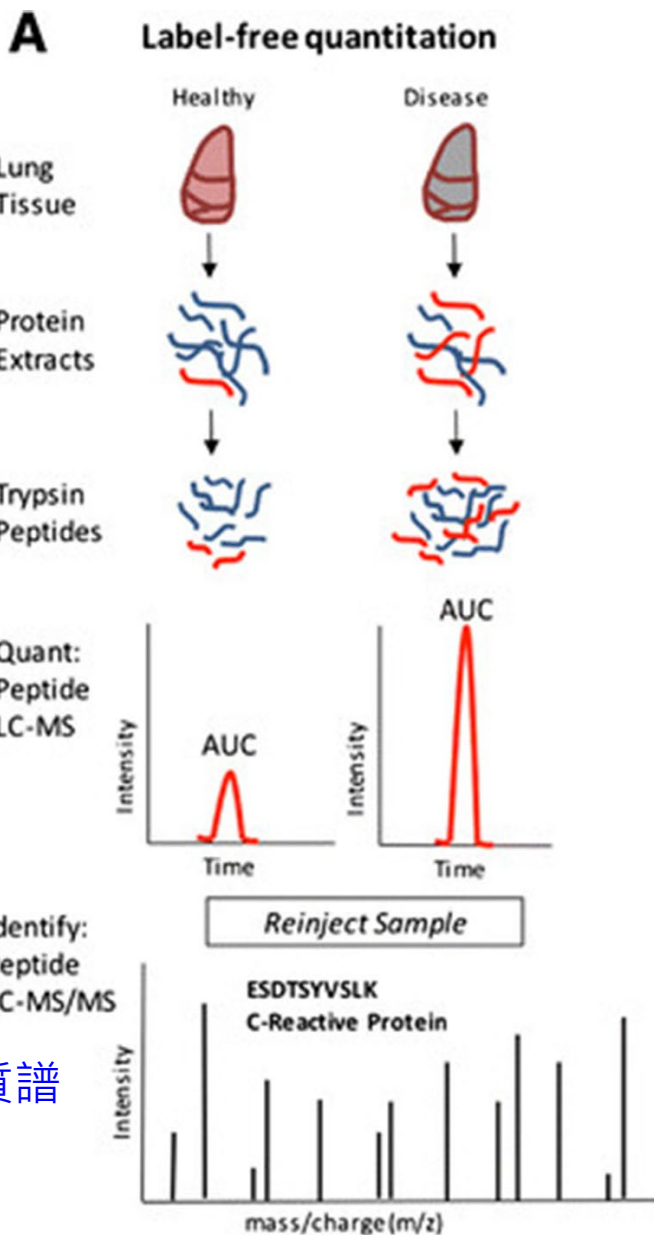
1 MLRLPTVFRQ MRPVSRVLAP HLTRAYAKDV KFGADARALM LQGVDLLADA
51 VAVTMGPKGR TVIIEQSWG PKVTKDGVTV AKSIDLKDKY KNIGAKLVQD
101 VANNTNEEAG DGTTTATVLA RSIAKEGF EK ISKGANPVEI RRGVMLAVDA
151 VIAELKKQSK PVTTP EEIAQ VATISANGDK EIGNIISDAM KKVGRKGVIT
201 VKDGKTLNDE LEIIEGMKFD RGYISPYFIN TSKGQKCEFQ DAYVLLSEKK
251 ISSIQSIVPA LEIANHRKP LVIIAEDVDG EALSTLVLNR LKVGLQVVAV
301 KAPGPGDNRK NQLKDMAIAT GGAVFGEEGL TLNLEDVQPH DLGKVGEVIV
351 TKDDAMLLKG KGDKAQIEKR IQEIIEQLDV TTSEYEKEKL NERLAKLSDG
401 VAVLKVGGTS DVEVNEKDR VTDALNATRA AVEEGIVLGG GCALLRCIPA
451 LDSLTPANED QKIGIEIIKR TLKIPAMTIA KNAGVEGSLI VEKIMQSSSE
501 VGVDAMAGDF VNMVEKGIID PTKVVRTALL DAAGVASLLT TAEVVVTEIP
551 KEEKDPGMGA MGGMGGGMGG GMF

```

# 定量蛋白質體學分析原理

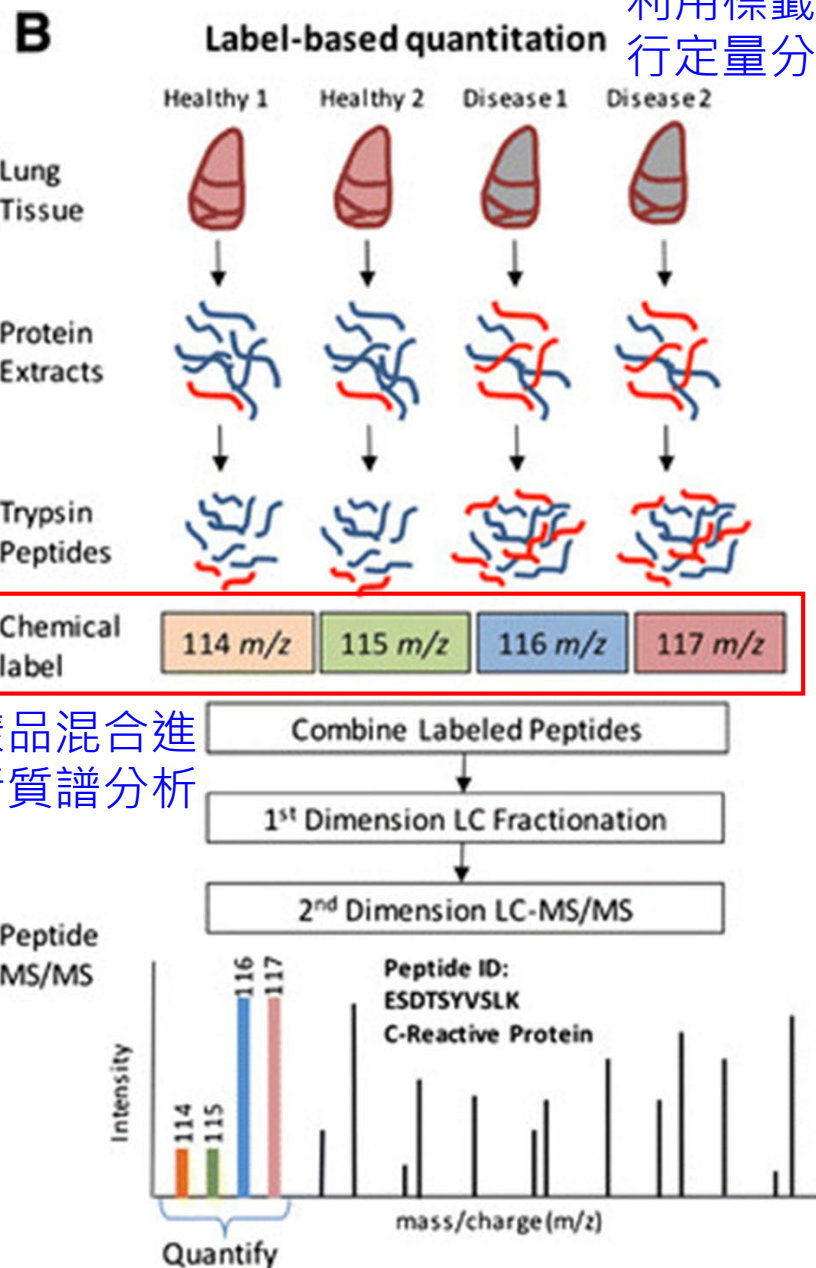
無標籤進行  
定量分析

利用標籤進  
行定量分析



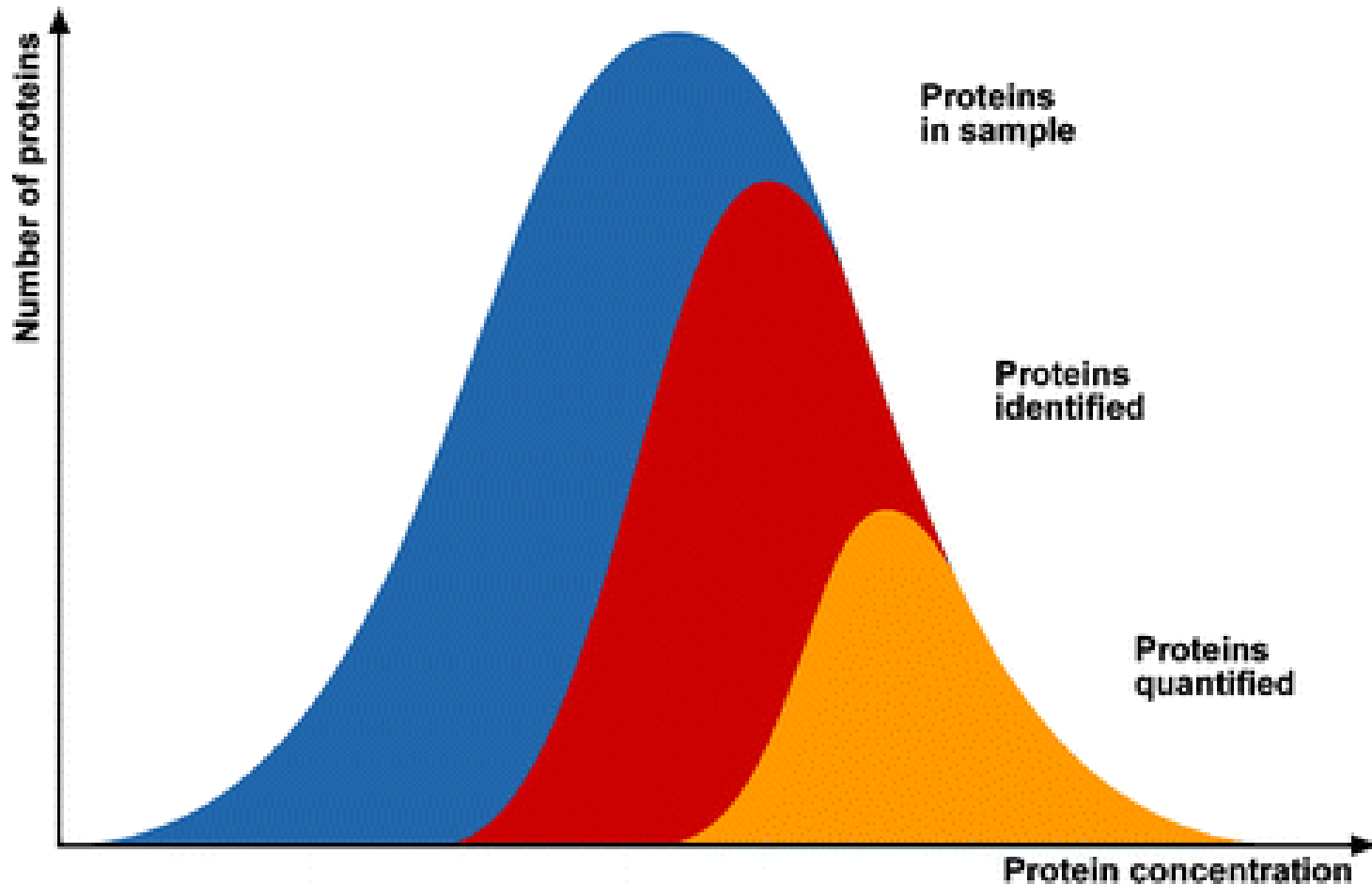
樣品分開進  
行質譜分析

串聯質譜

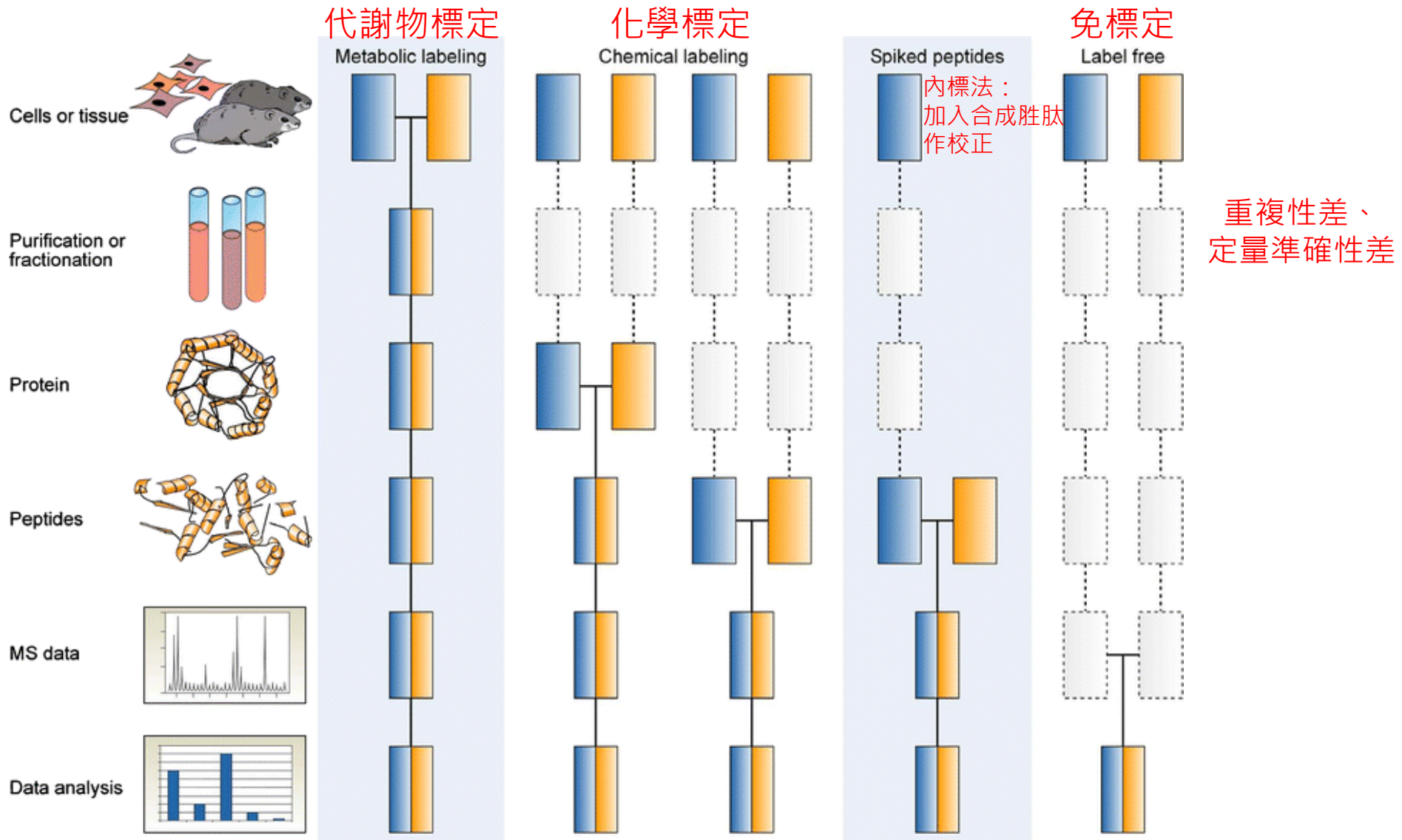




# 能被定量的蛋白質種類比 能被定性的蛋白質種類還少

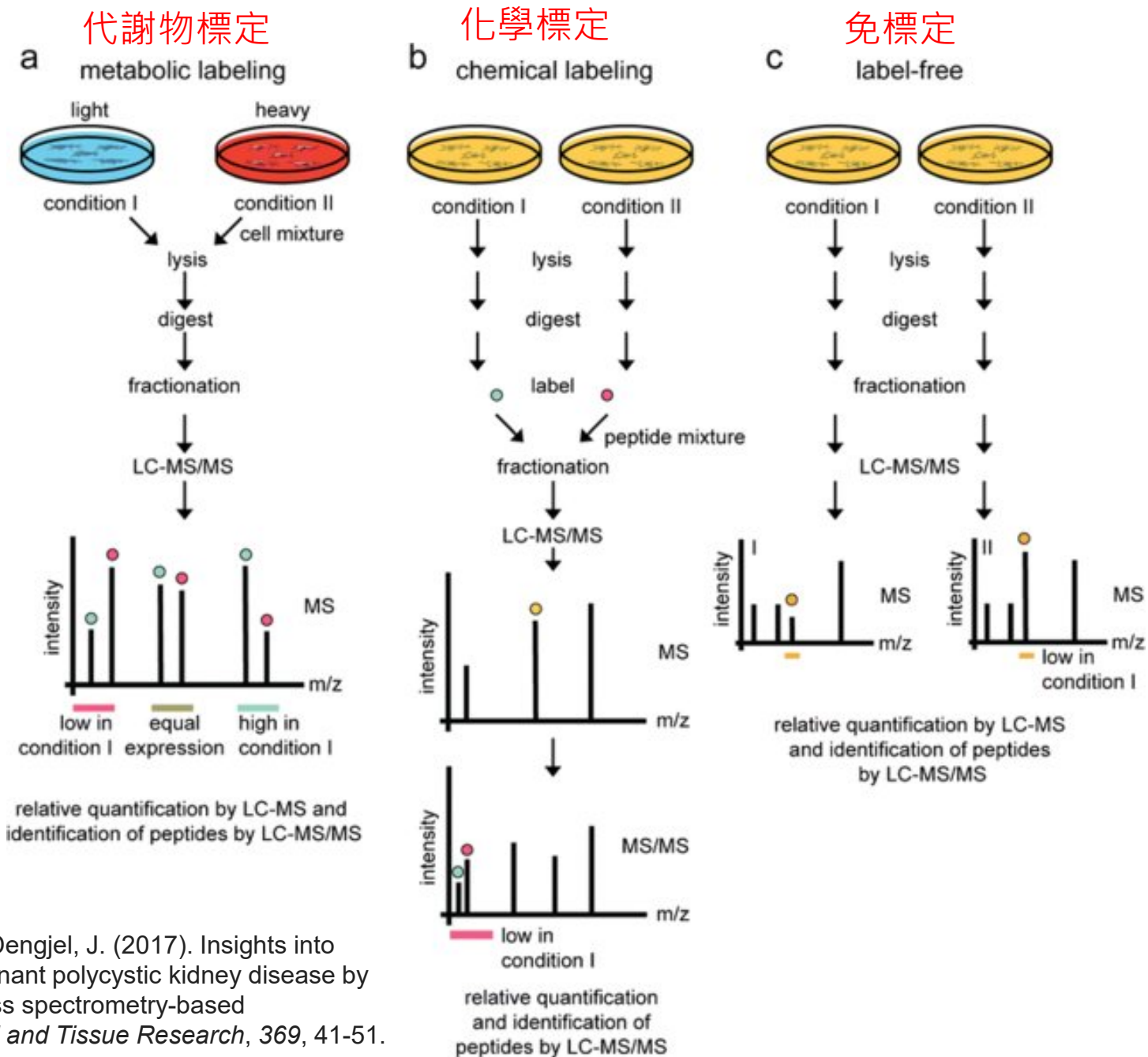


Common quantitative mass spectrometry workflows. *Boxes in blue and yellow represent two experimental conditions. Horizontal lines indicate when samples are combined. Dashed lines indicate points at which experimental variation and thus quantification errors can occur.*



Bantscheff, M., Schirle, M., Sweetman, G., Rick, J., & Kuster, B. (2007). Quantitative mass spectrometry in proteomics: a critical review. *Analytical and bioanalytical chemistry*, 389, 1017-1031.

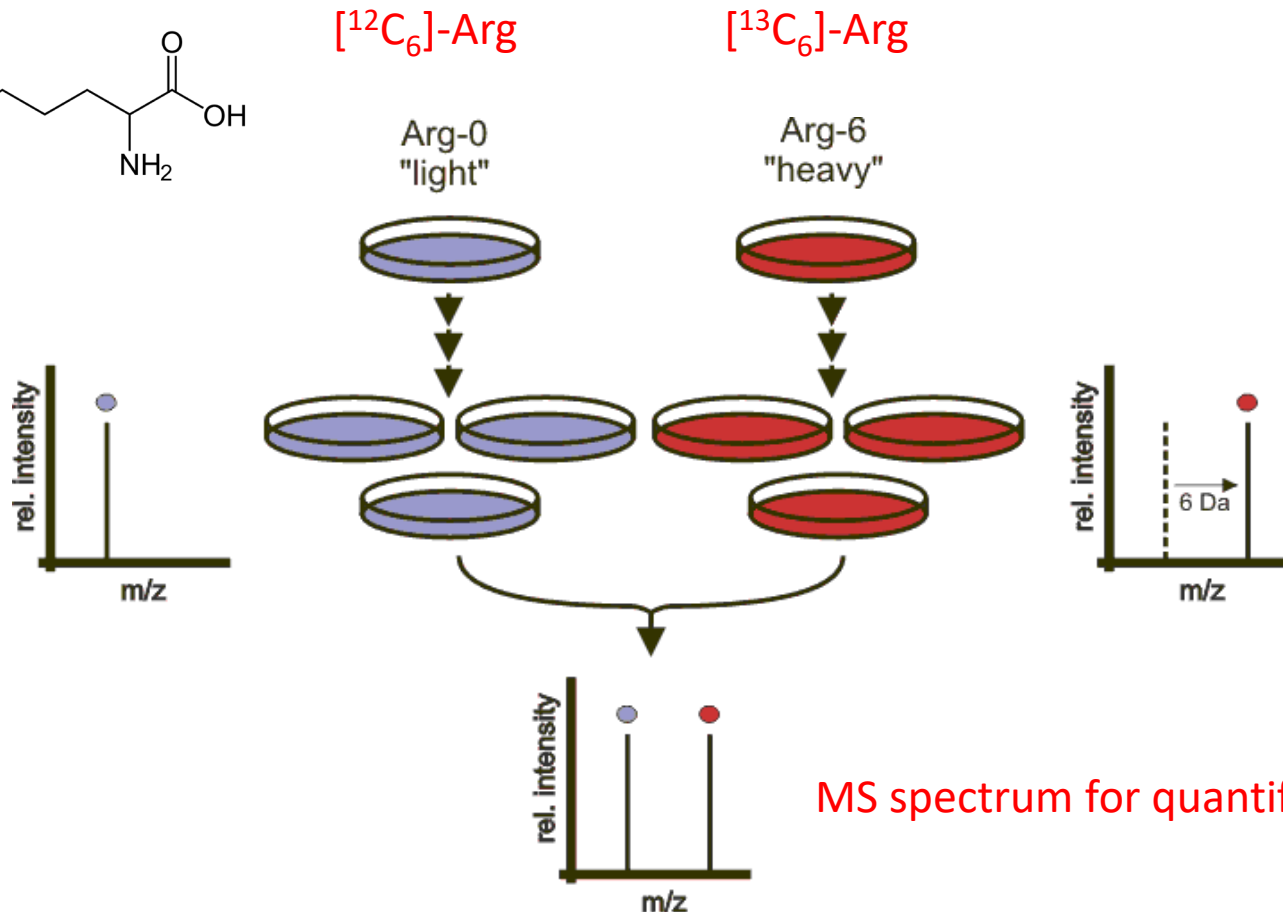
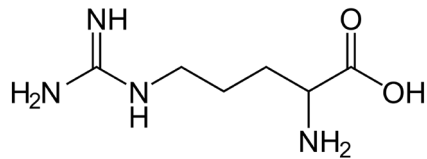
# 穩定同位素標定



Diedrich, B., & Dengjel, J. (2017). Insights into autosomal dominant polycystic kidney disease by quantitative mass spectrometry-based proteomics. *Cell and Tissue Research*, 369, 41-51.

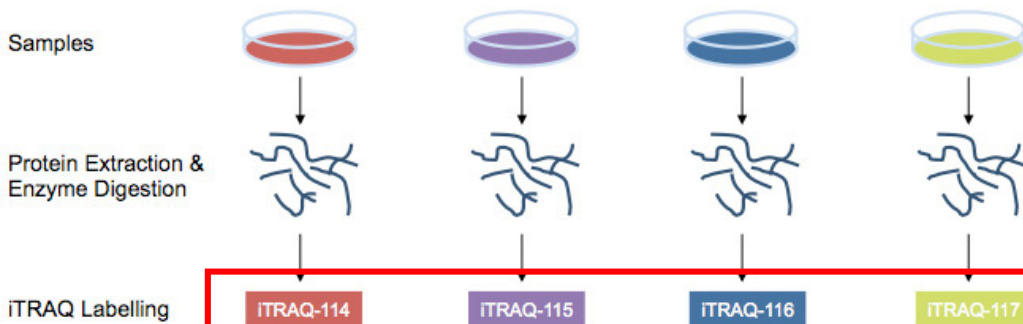
# 代謝標定：使用穩定同位素之代謝物進行標定

- 以胺基酸培養進行穩定同位素標定(stable isotope labeling by amino acids in cell culture, SILAC)
  - 只適用於體外培養細胞或細菌的標定

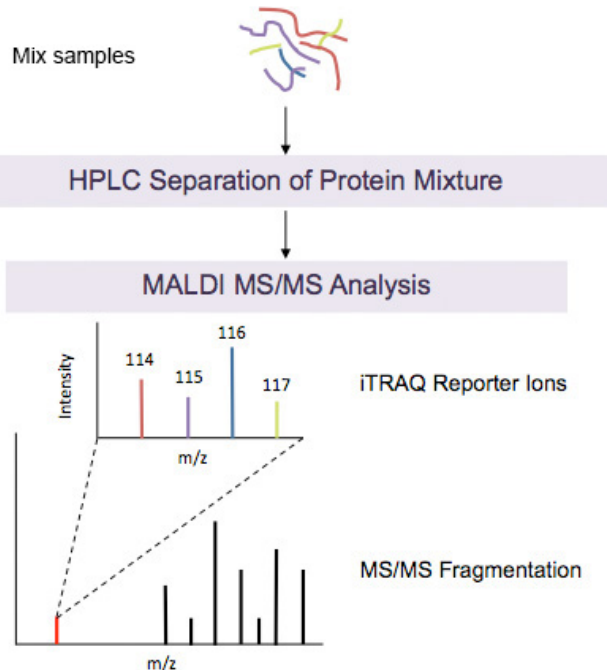


# 化學標定：

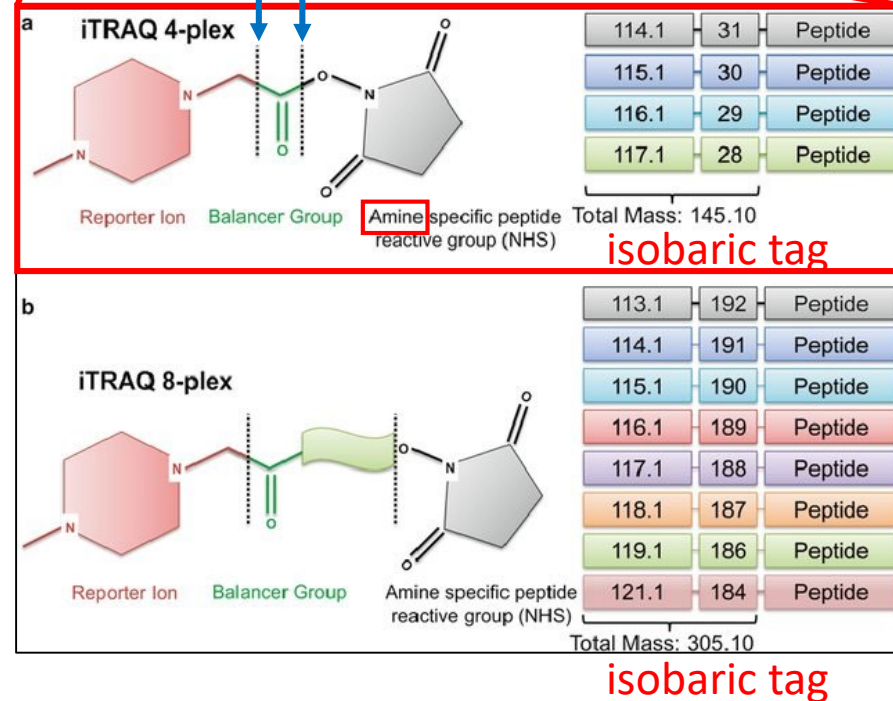
## 採用與蛋白質反應之穩定同位素標籤進行標定



- Isobaric tag for relative and absolute quantitation (iTRAQ)



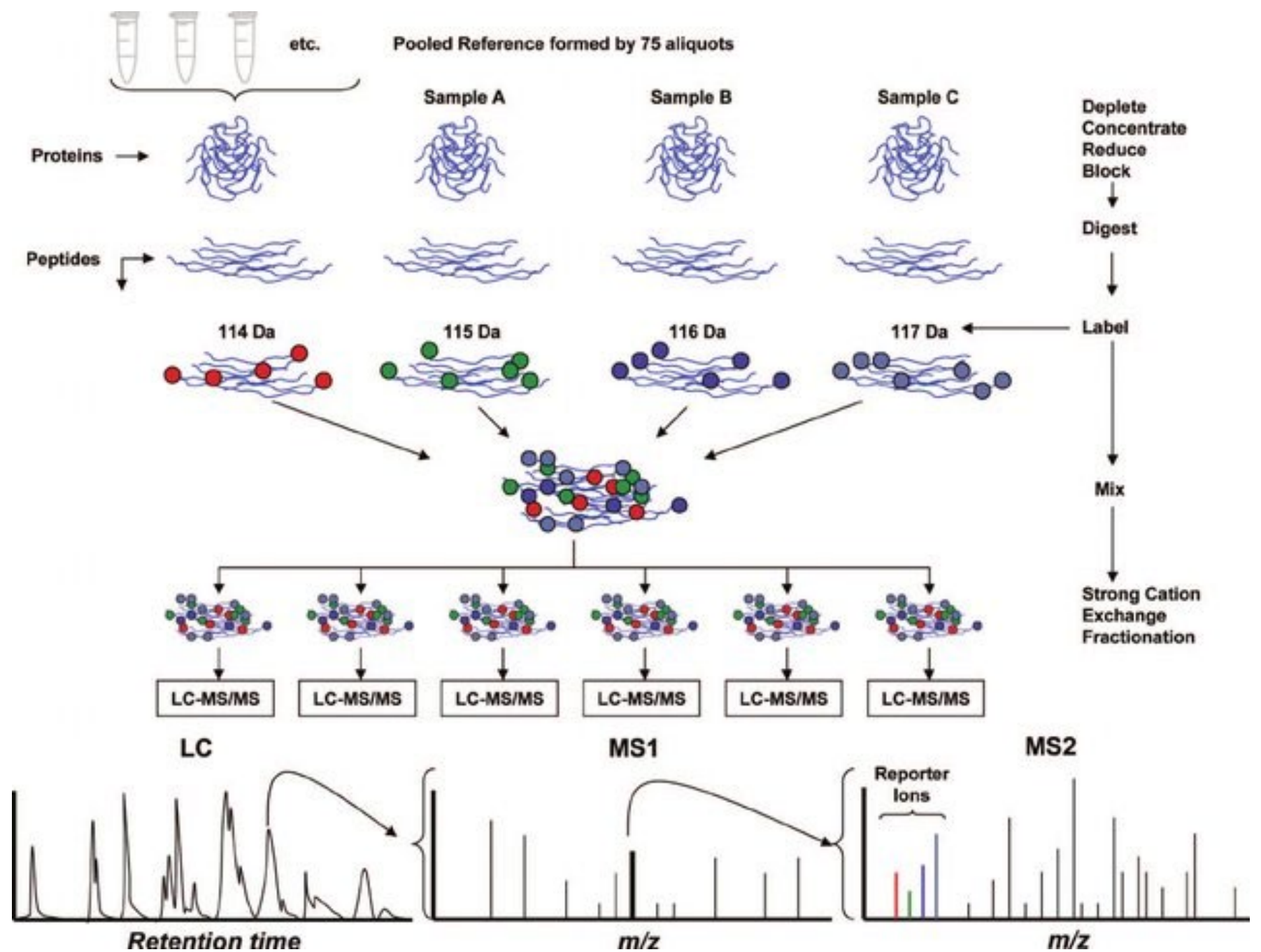
Cut during fragmentation



MS + MS/MS spectra for quantification

<https://www.creative-proteomics.com/services/itraq-based-proteomics-analysis.htm>

Aggarwal, S., & Yadav, A. K. (2016). Dissecting the iTRAQ data analysis. *Statistical analysis in proteomics*, 277-291.



# 體學相關應用

## 1. 親緣基因體學 ( Phylogenomics )

- 結合親緣關係學 ( phylogenetics ) 與基因體學的研究。
- 包含了以下研究：
  - 1) 利用基因體建立演化樹。
  - 2) 基因功能預測。
  - 3) 基因家族的演化研究。
  - 4) 基因側向轉移 ( lateral gene transfer ) 。

# 由基因體建立演化樹

**1 Input genomes**

**2 Align individual target genes**

Gene X

Genome A: [Red segment] [Pink segment] [Blue segment]

Genome B: [Red segment] [Dotted segment] [Blue segment]

Genome C: [Red segment] [Pink segment] [Blue segment]

Legend: A (Pink), B (Dotted), C (Pink), Y (Pink), Z (Blue)

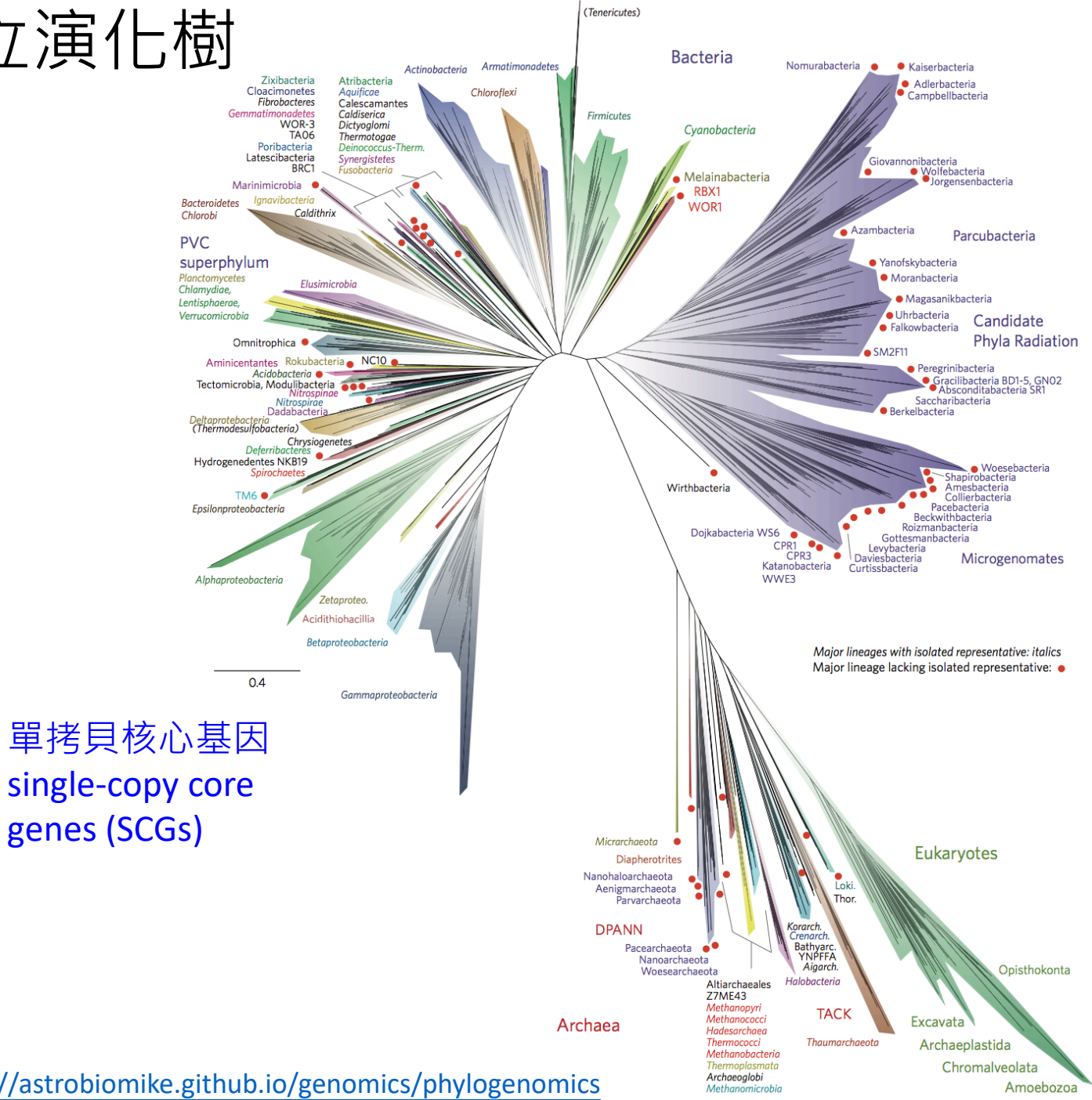
**3 Stick alignments together**

Genome A: [Red] [Pink] [Blue]

Genome B: [Red] [Dotted] [Blue]

Genome C: [Red] [Pink] [Blue]

**4 Infer evolutionary relationships**

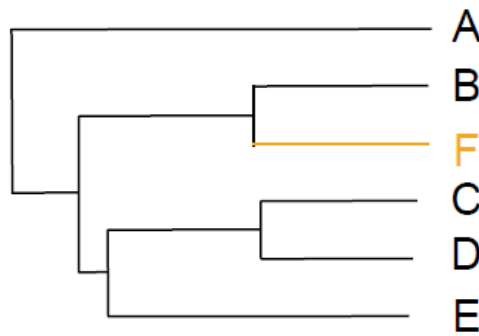


單拷貝核心基因  
single-copy core  
genes (SCGs)

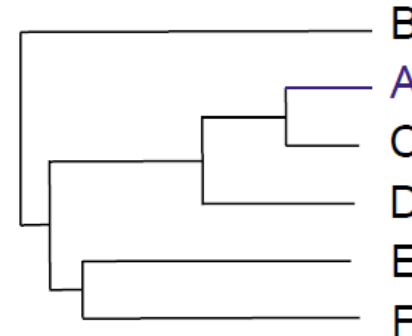


# 基因側向轉移/基因水平轉移

- 生物將遺傳物質（基因）傳遞給其他細胞而非其子代的過程，例如：接合、轉導及轉型。

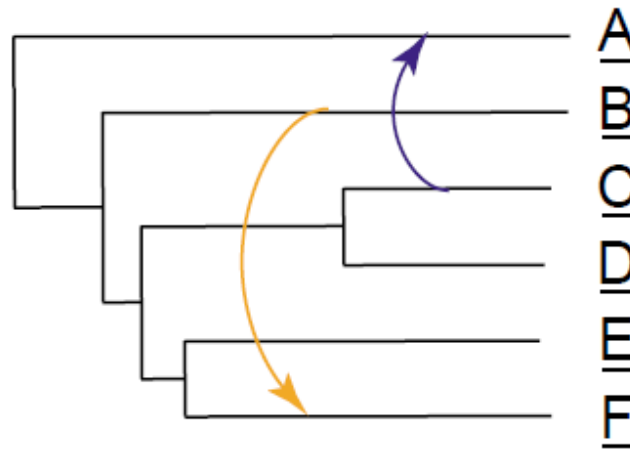


Gene 1



Gene 2

假設的基因轉移途徑

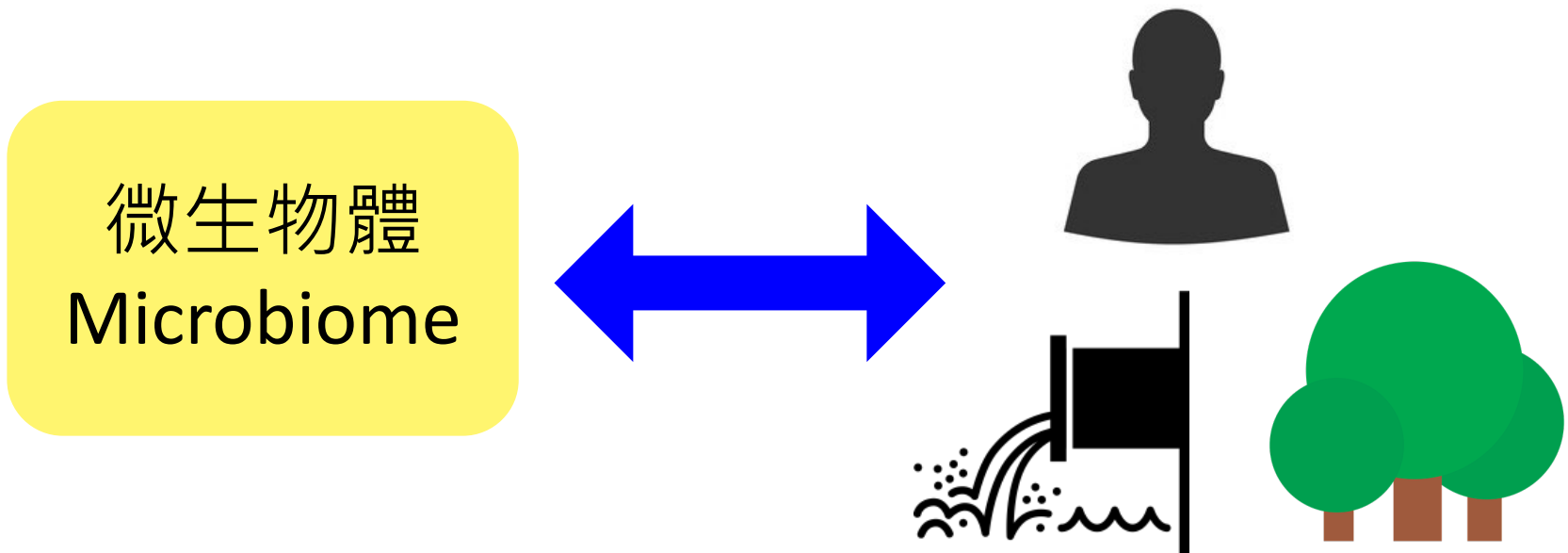


基因2可能由C轉移到A

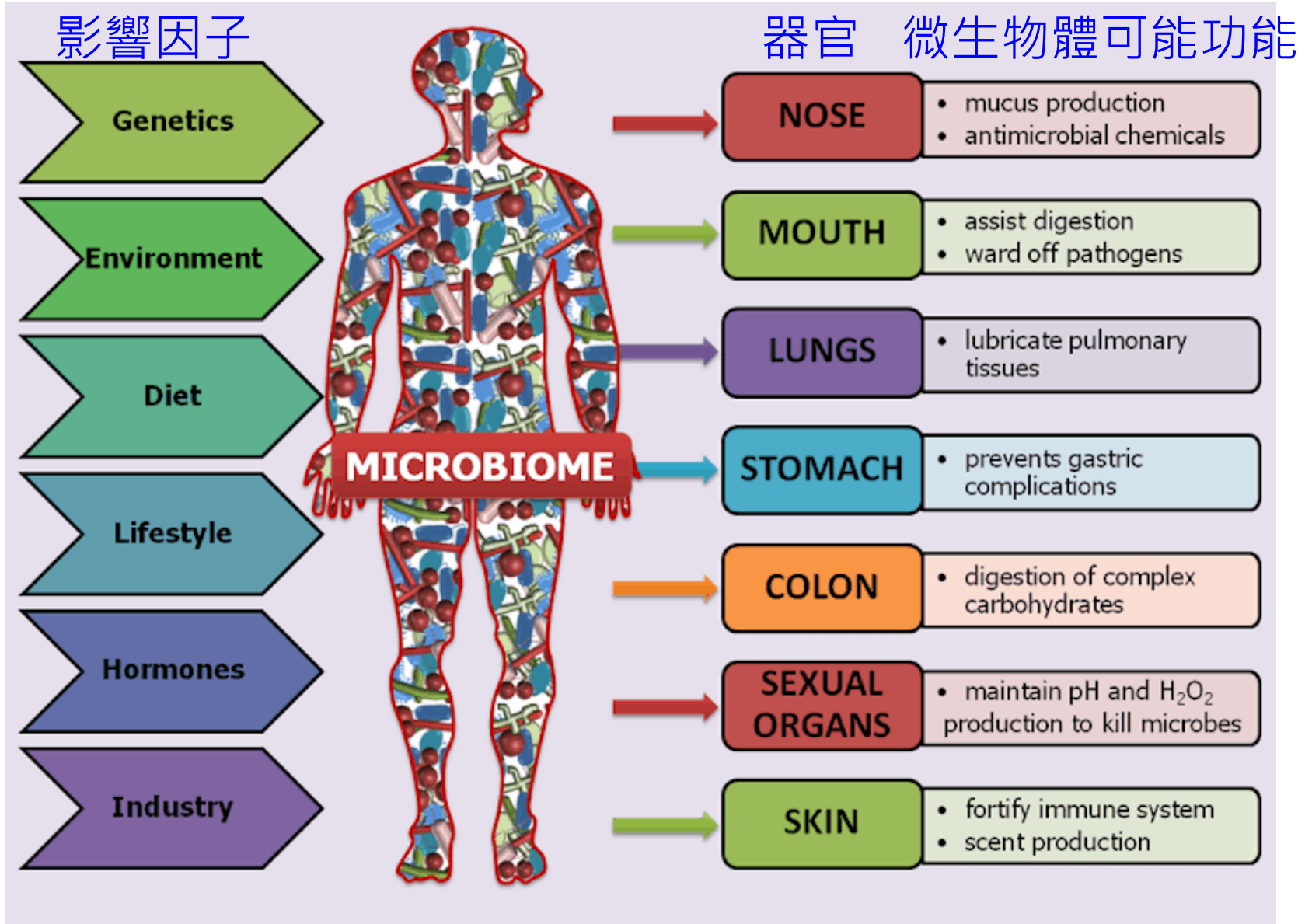
基因1可能由B轉移到F

## 2. 微生物體學 ( Microbiomics )

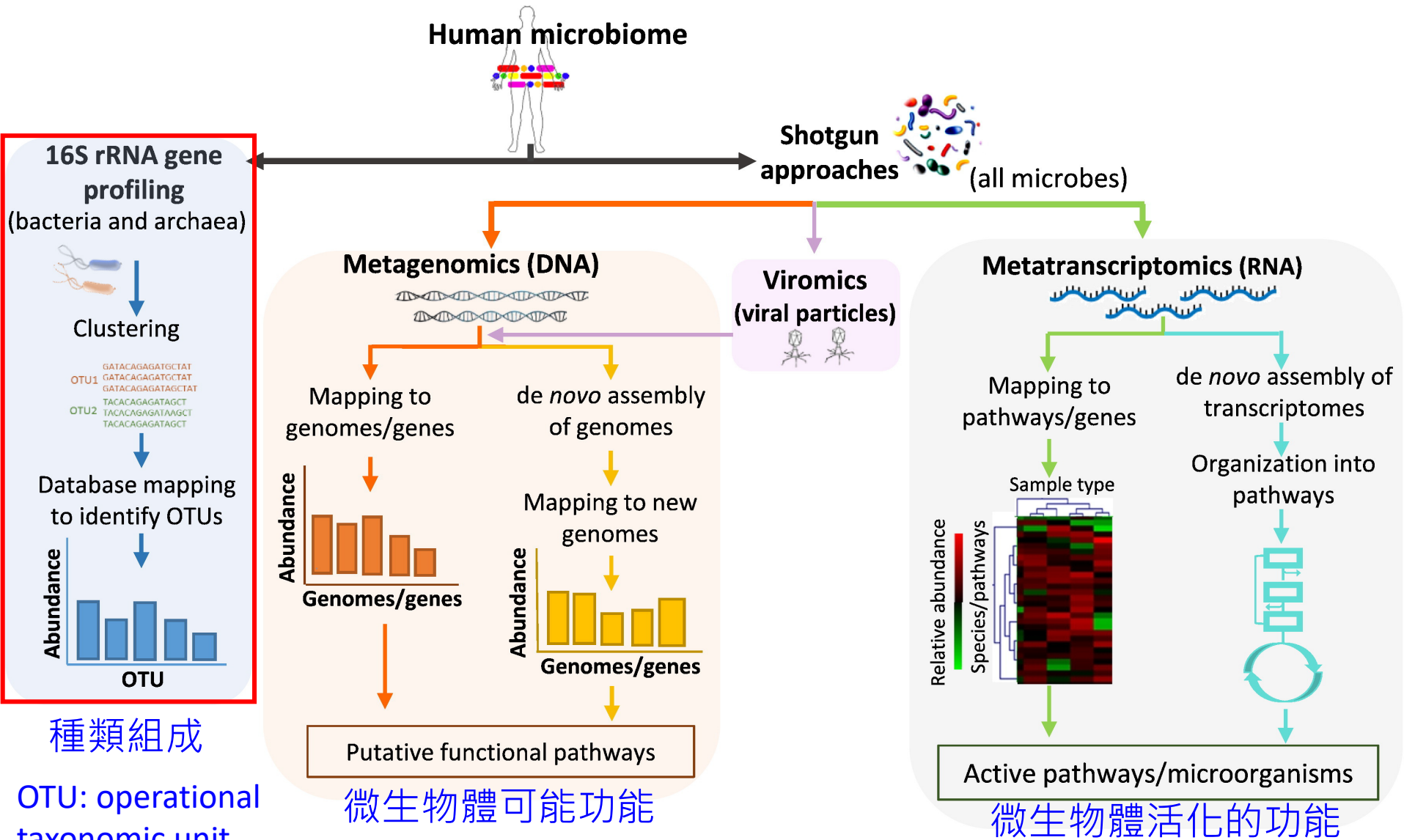
- 分析環境/宿主 ( 器官 ) 中微生物群落 ( microbial community ) 的種類組成與功能，藉以得知微生物群落與宿主/環境之間的關係。



# 人體微生物體 ( Human microbiome )



# 微生物體分析方法：以人體腸道微生物體為例



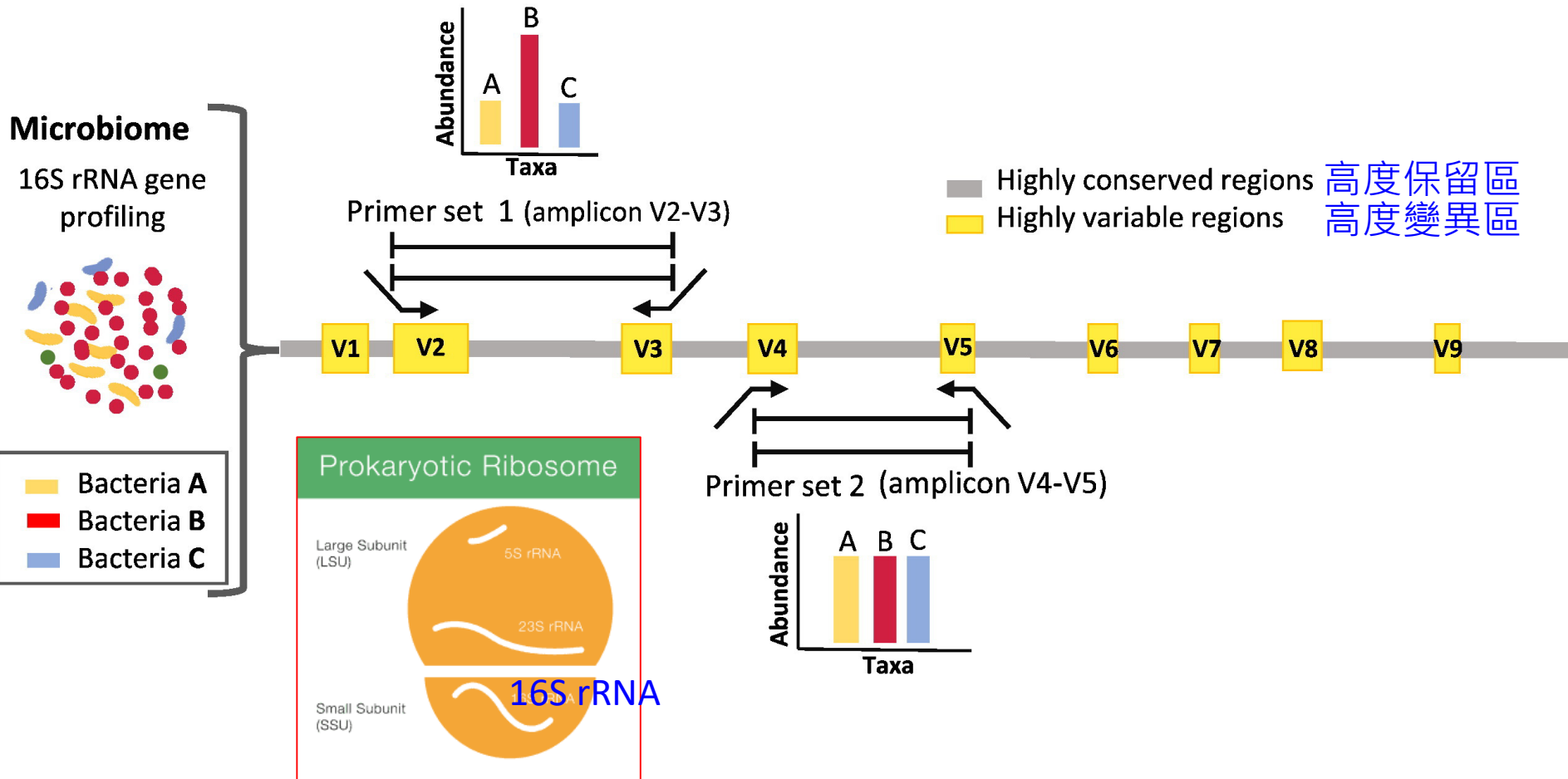
種類組成

微生物體可能功能

微生物體活化的功能

OTU: operational taxonomic unit  
分類操作單元

# 16S rRNA基因定序分析： 最常用來分析微生物體組成的方法



<https://www.sciencedirect.com/science/article/pii/S2001037015000318>

<https://help.ezbiocloud.net/16s-rna-and-16s-rna-gene/>

# 3. 多體學 ( Multi-omics )

