

基因密碼: DNA及基因體

Genetic code: DNA & Genome

藍忠昱

國立清華大學
分子與細胞生物研究所暨生命科學系

cylan@life.nthu.edu.tw

December 14, 2013

上課主題 TOPICS

1. 由基因密碼到蛋白質合成

生物巨分子：核酸(DNA & RNA)及蛋白質
生物巨分子的化學性質與反應

2. DNA之分析技術

基因重組技術 PCR DNA定序

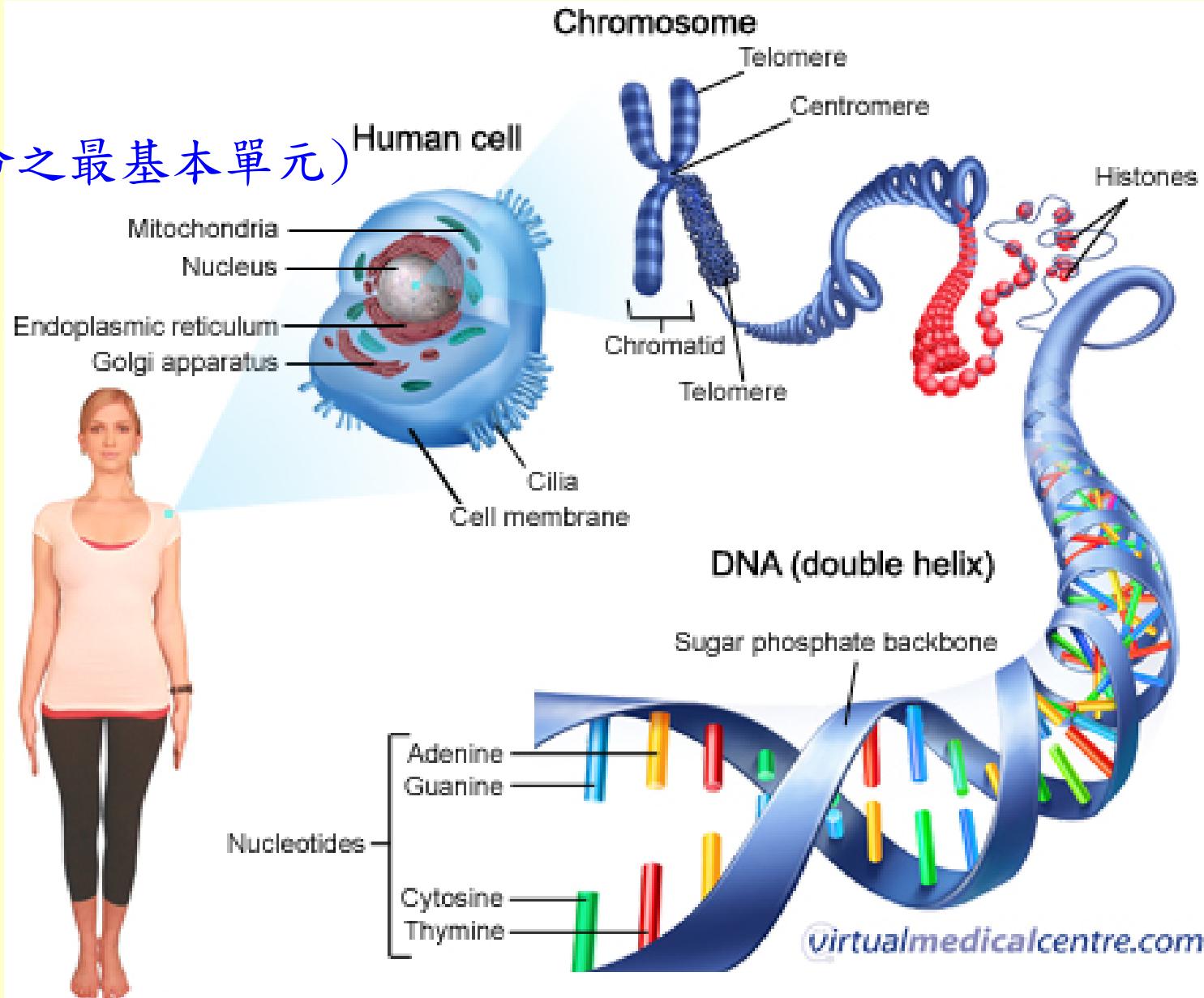
2. 基因體及其功能分析

3. DNA及基因體分析技術之應用

醫學上之應用
農業上之應用
鑑識科學及考古學上之應用

由基因密碼到蛋白質合成

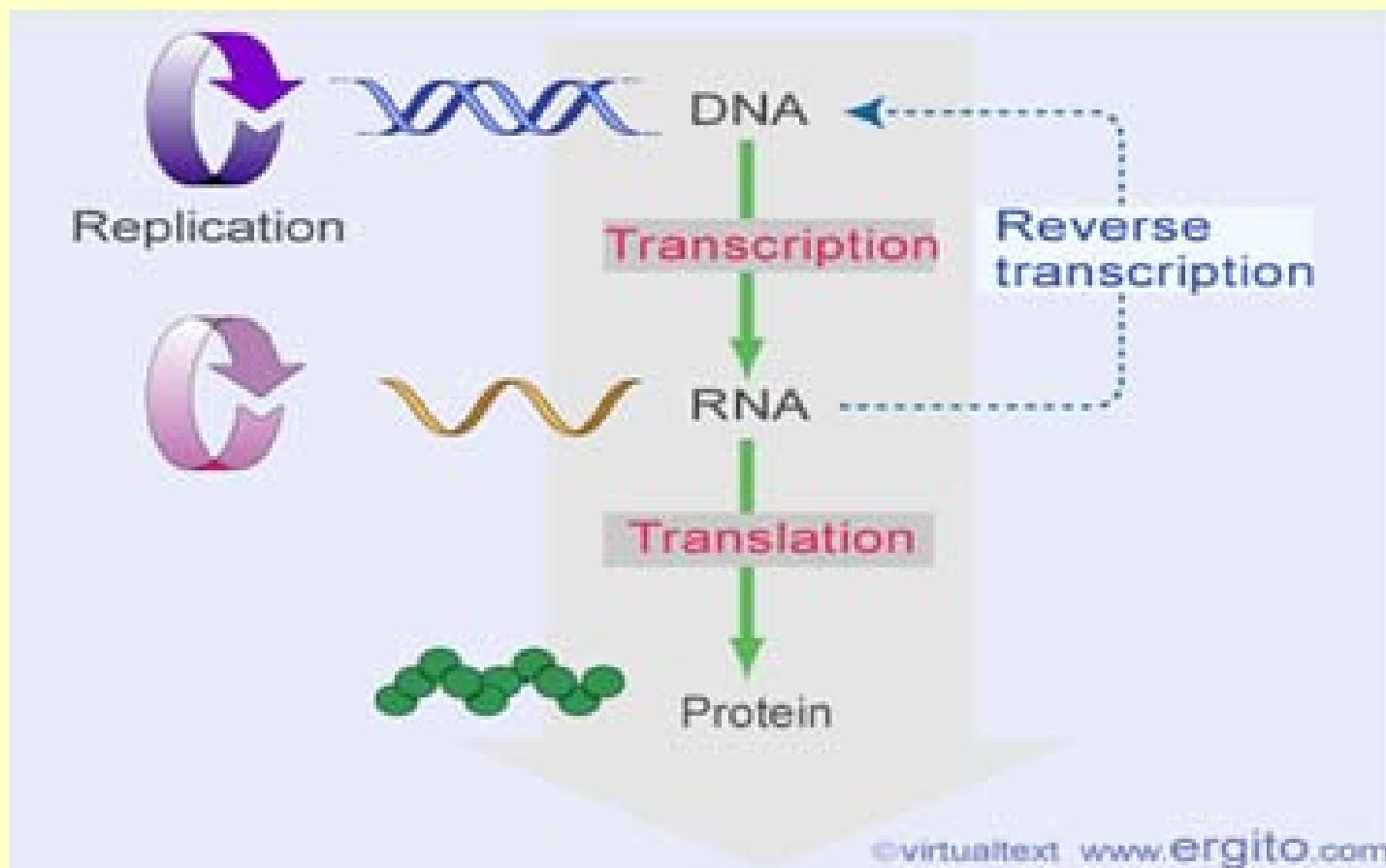
(生命之最基本單元)



[http://www.virtualmedicalcentre.com/
anatomy.asp?sid=37&title=DNA-\(Deoxyribonucleic-Acid\)](http://www.virtualmedicalcentre.com/anatomy.asp?sid=37&title=DNA-(Deoxyribonucleic-Acid))

分子生物學之中心法則 (Central dogma)

- The Central Dogma describes the Information flow from DNA → RNA → protein



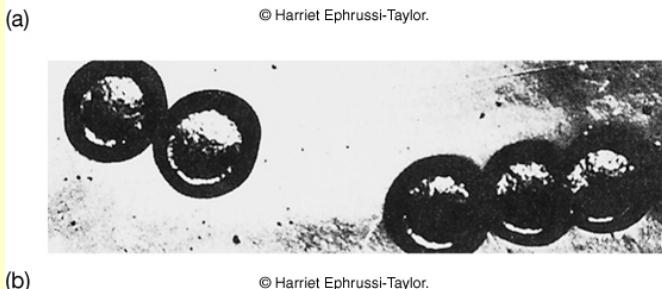
The nature of genetic material (遺傳物質)

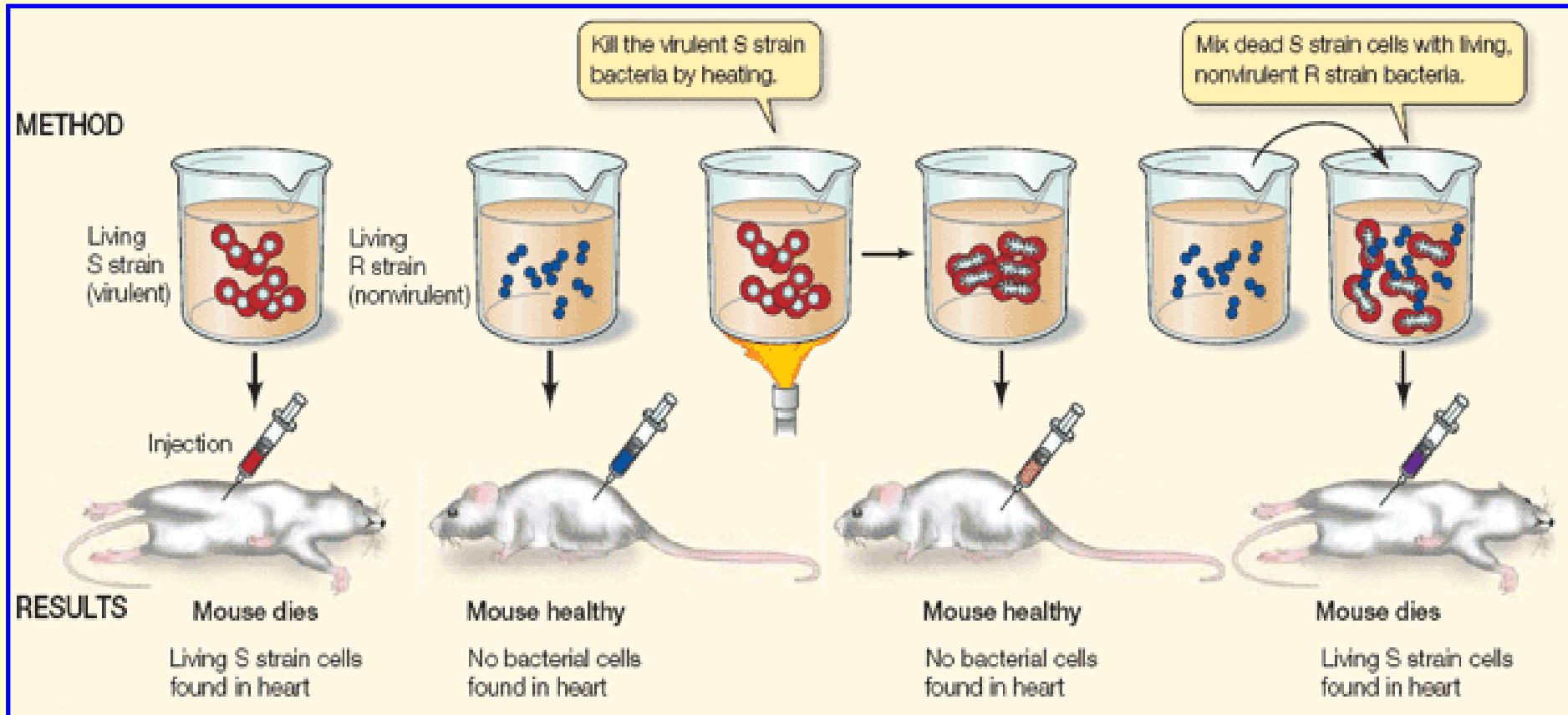
- (1) Frederick Griffith's transformation (轉形) experiments
- (2) The A. D. Hershey-Martha Chase experiment

提供重要之證據，證明：Nucleic acid (核酸) 是遺傳物質

Griffith的轉形(transformation)實驗

1. 以 *Streptococcus pneumoniae* (肺炎雙球菌)為實驗材料
2. *Streptococcus pneumoniae* 之不同變異株 (variants)
 - (a) 較大、平滑有光澤的菌落：由 **smooth (S) virulent bacteria** 構成
 - (b) 較小、斑駁的(mottled)菌落：由 **rough (R) avirulent bacteria** 構成

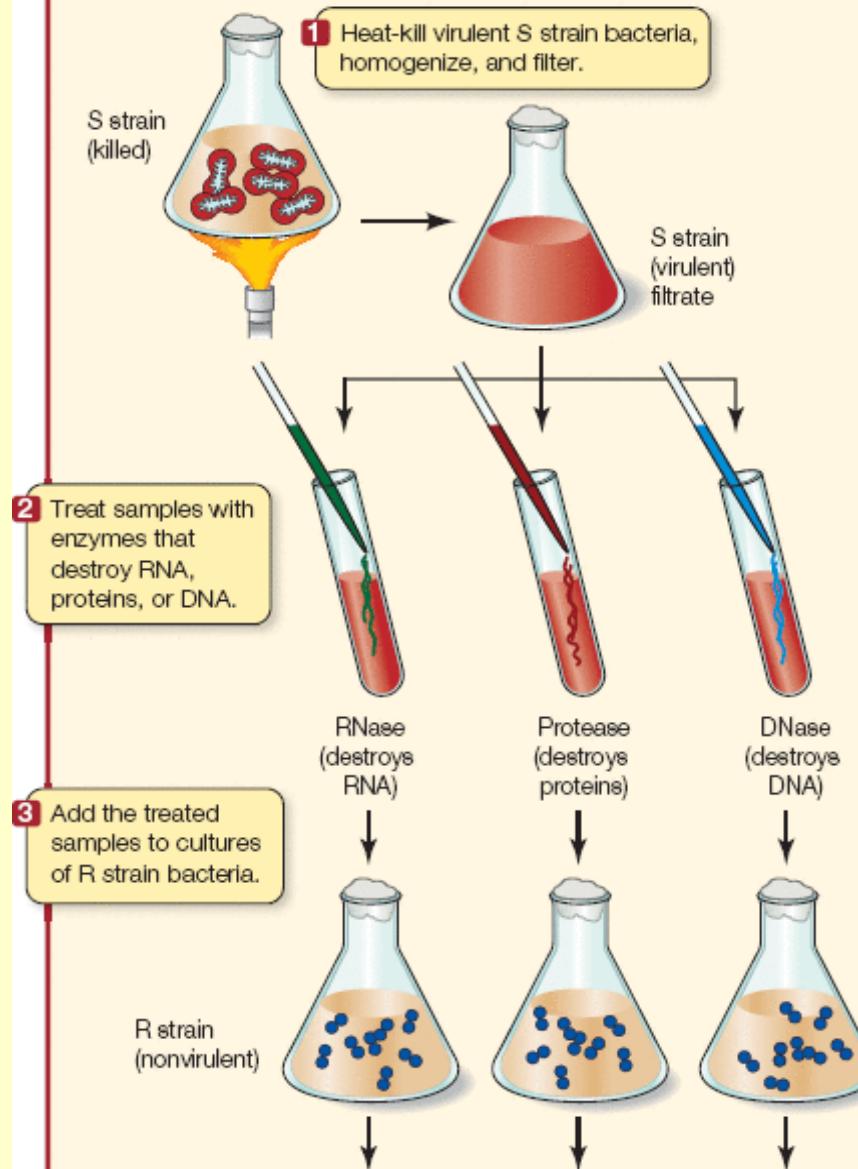




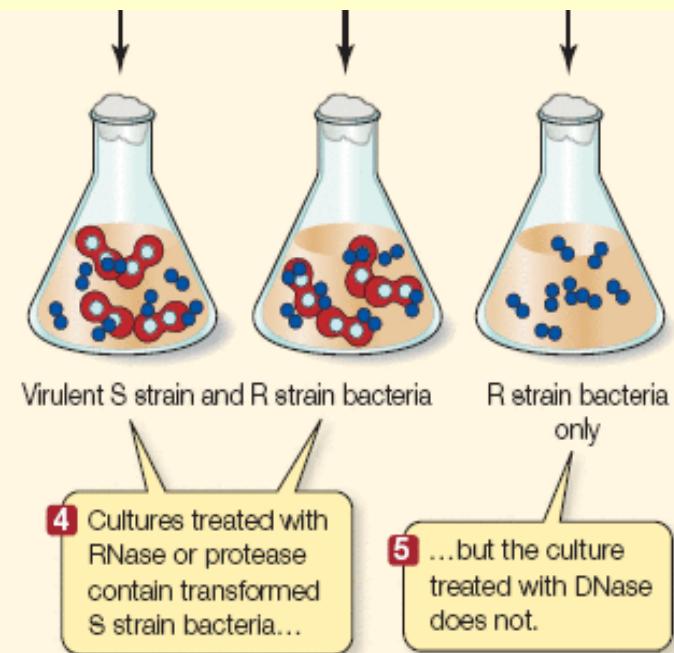
結論：有一種化學物質可由一個細菌傳至另一個細菌

[http://www.nature.com/scitable/topicpage/
isolating-hereditary-material-frederick-griffith-oswald-avery-336](http://www.nature.com/scitable/topicpage/isolating-hereditary-material-frederick-griffith-oswald-avery-336)

METHOD



RESULTS



結論：因為只有去氧核糖核酸水解酵素(DNase)可以破壞細胞間傳遞之化學物質，該物質為DNA

Hershey & Chase 在1952年進行的實驗

1. 以T2 bacteriophage (噬菌體, bacterial virus)為實驗材料

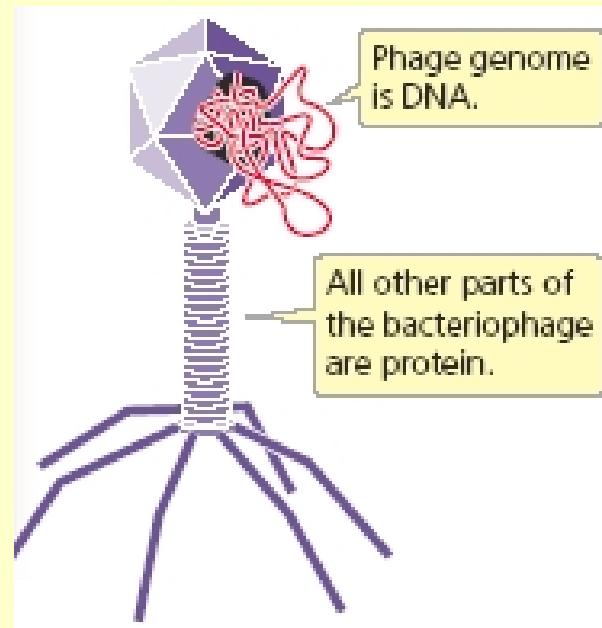
2. T2 可感染大腸桿菌(*E. coli*)

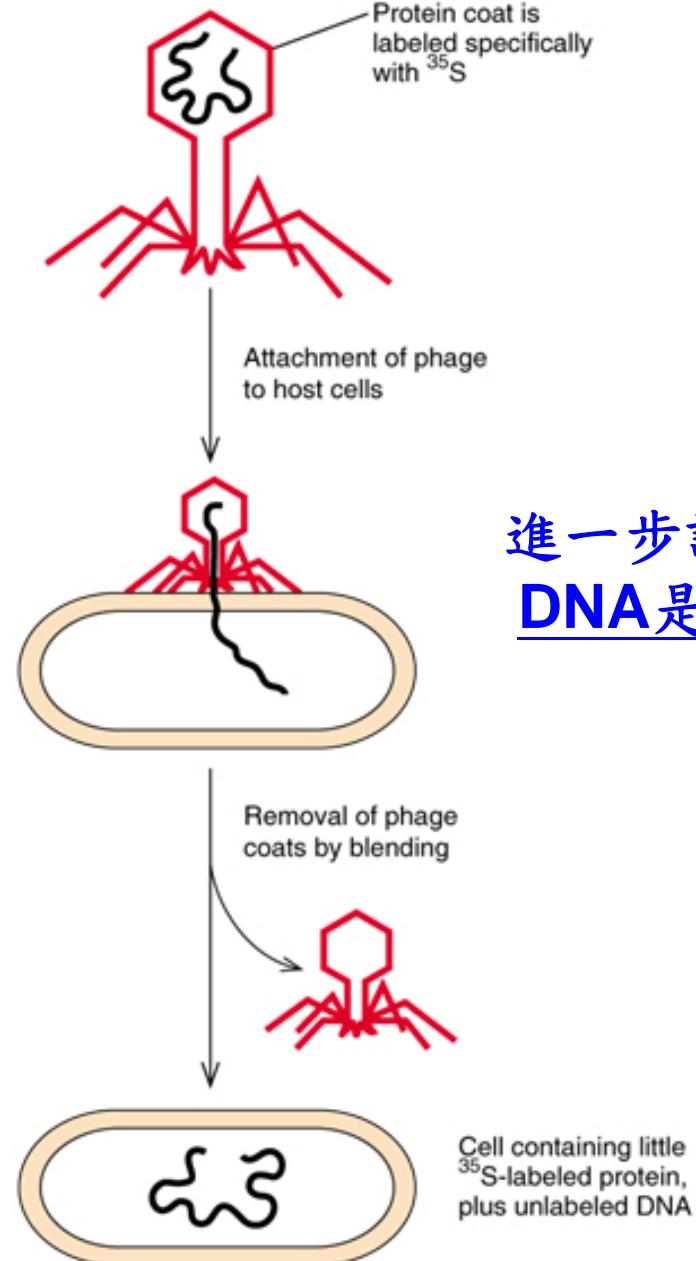
3. T2噬菌體：僅由蛋白質及DNA組成

4. 利用放射線同位素分別標定蛋白質及DNA

磷-32 (^{32}P)：標定DNA

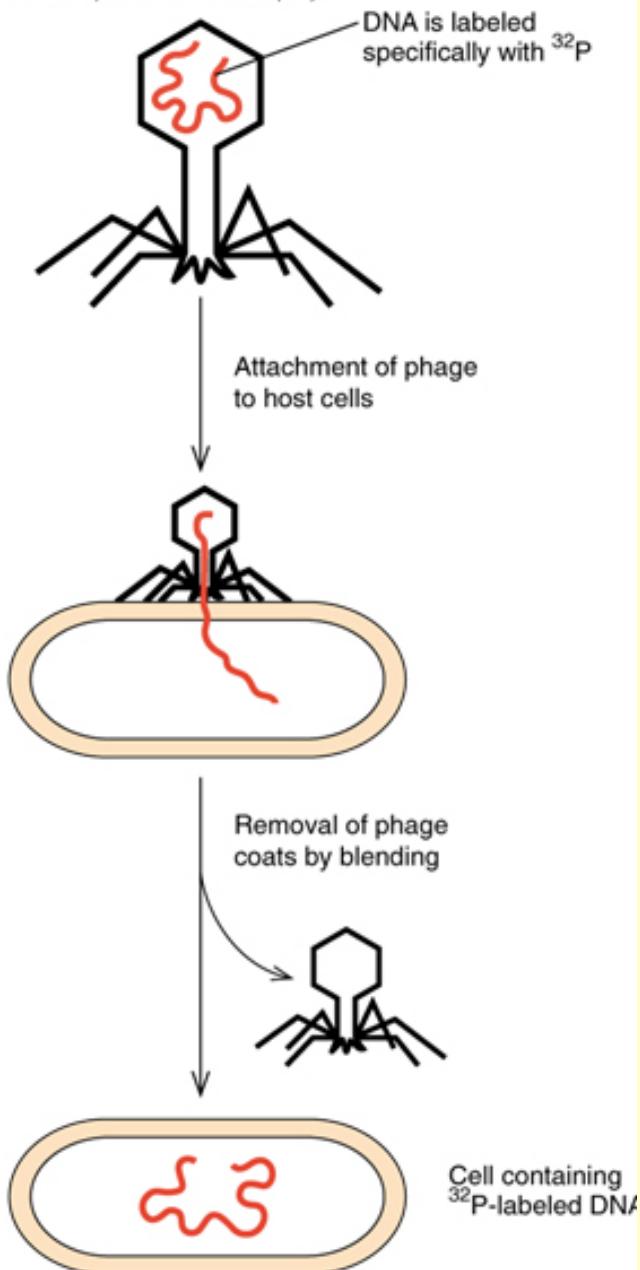
硫-35 (^{35}S)：標定蛋白質





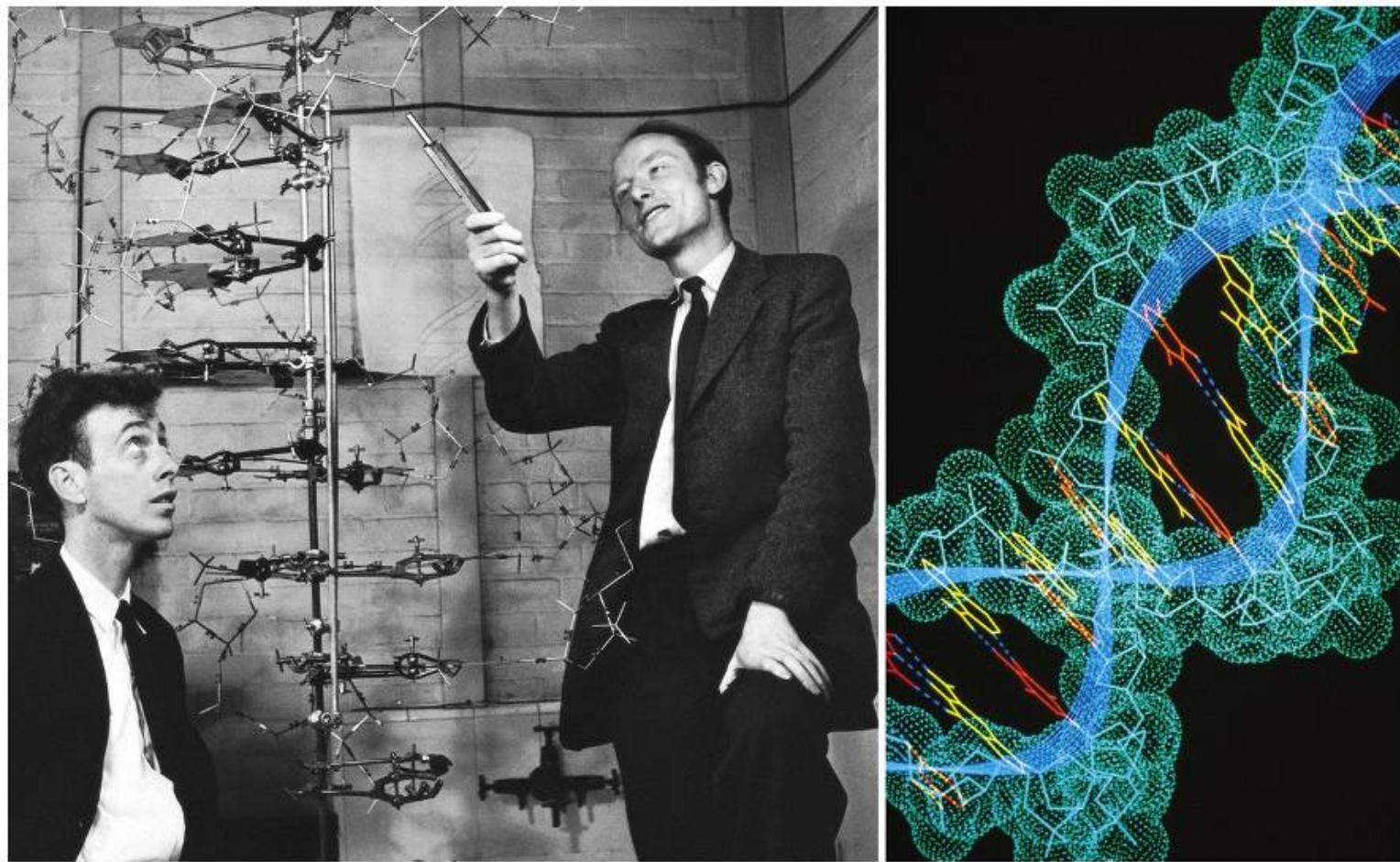
a)

進一步證明：
DNA是遺傳物質



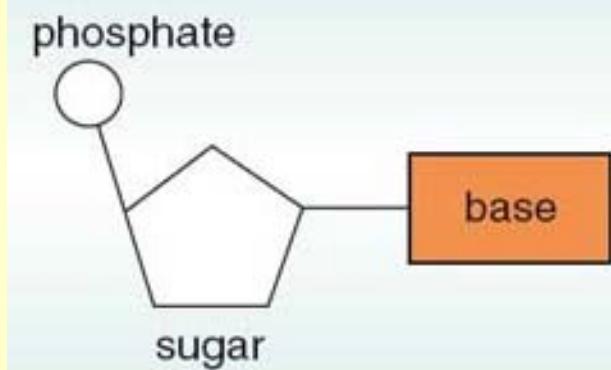
b)

DNA 結構之發現



於1953年, James Watson & Francis Crick 提出DNA 結構為
雙螺旋(double-helix)之模式

DNA之化學性質

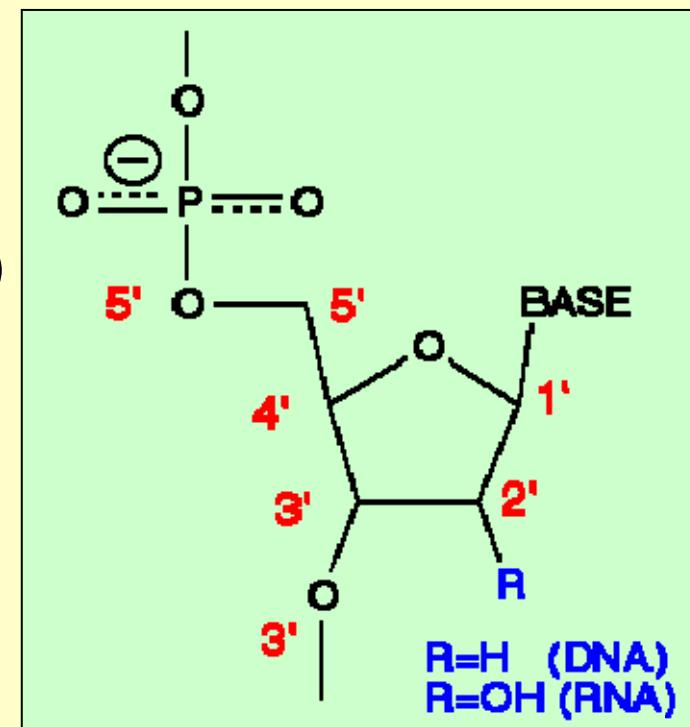


一個核苷酸，包括：

- (1) Phosphate (磷酸)
- (2) A sugar molecule (五碳醣)
- (3) A nitrogenous base (含氮鹼基)

核苷酸中之phosphate與sugar

- Sugar: 可為去氧核醣(deoxyribose, R=H)
或為核醣(ribose, R=OH)
- 在每一個核苷酸中之phosphate
與sugar都是相同



DNA密碼是由四種不同字母(含氮鹼基)A,T,C,G組成：

adenine (A), guanine (G), cytosine (C) or thymine (T)

Purine 嘌呤

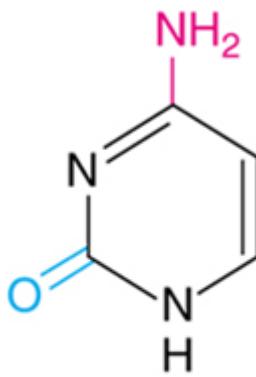


Adenine

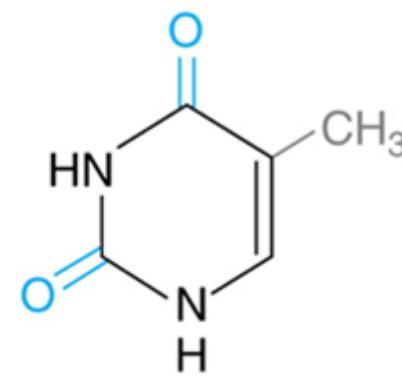


Guanine

Pyrimidine 喹啶



Cytosine



Thymine

腺嘌呤

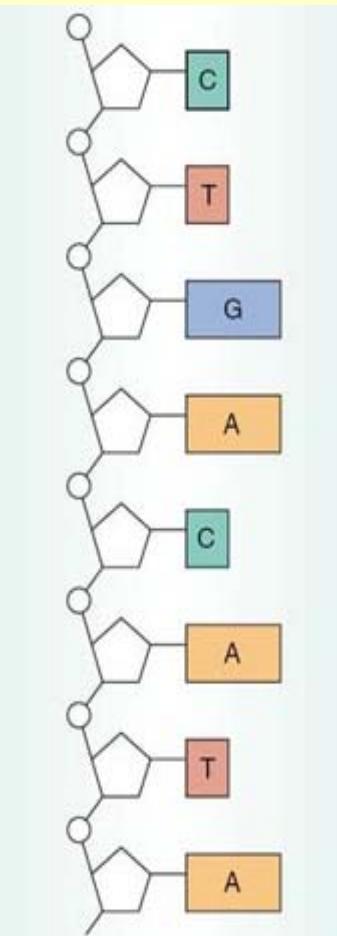
鳥糞嘌呤

胞嘧啶

胸嘧啶

DNA 雙螺旋中之每一單股:

- 是由許多核苷酸(nucleotide)串連所組成: 由一個核苷酸之phosphate與另一個核苷酸之sugar連接，形成所謂sugar-phosphate backbone(骨架)
- 核苷酸中之Bases則由骨架，朝DNA之另一股突出

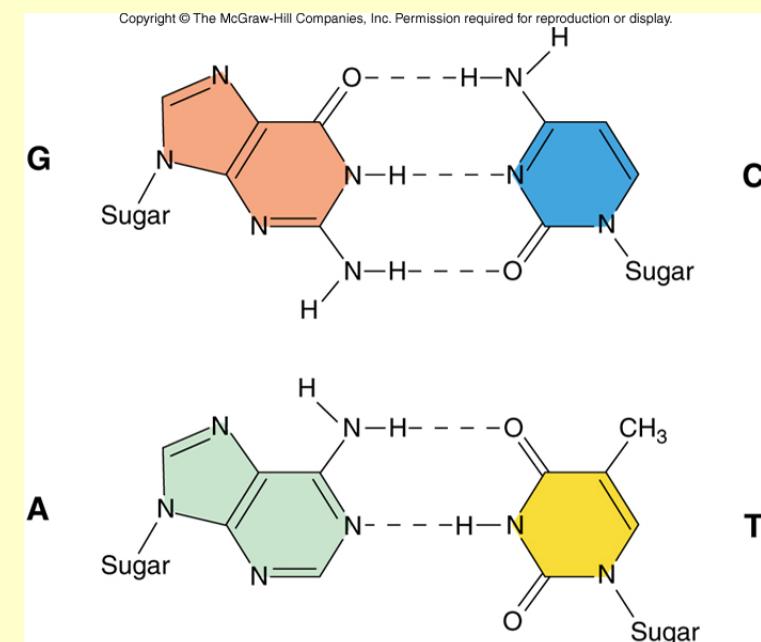


DNA 雙股螺旋結構之形成:

- Complementary base pairs 鹼基互補配對 (adenine pairs with thymine; cytosine pairs with guanine)

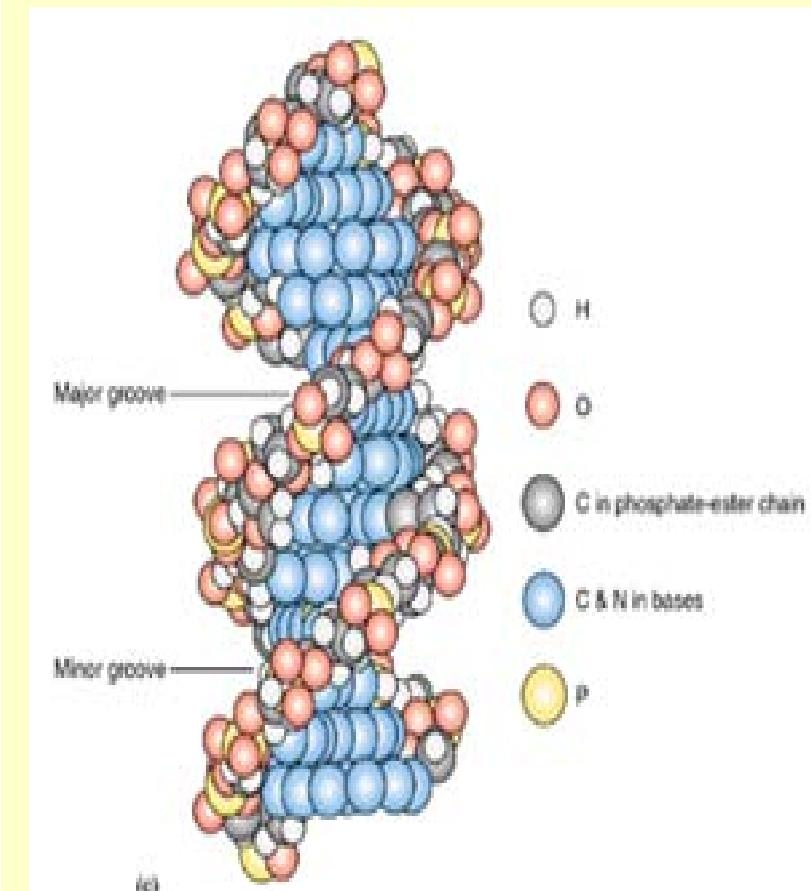
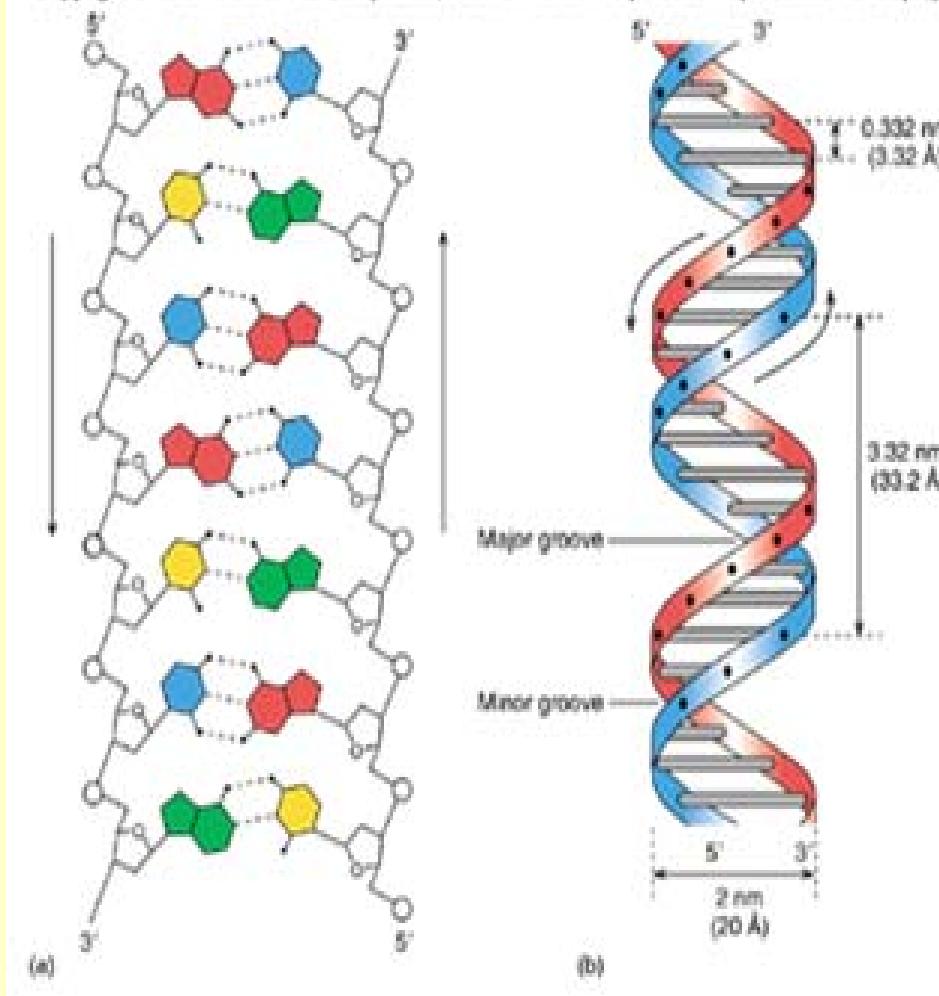
- 兩個bases之間，以氫鍵(hydrogen bonds)結合

$$A = T$$
$$G \equiv C \text{ (stronger)}$$



DNA 結構之三種模式

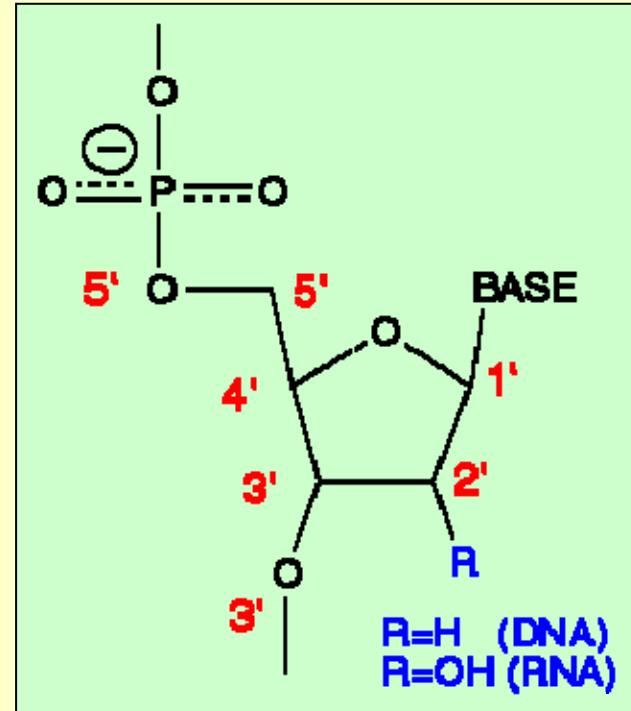
Copyright © The McGraw-Hill Companies, Inc. Permission required for reproduction or display.



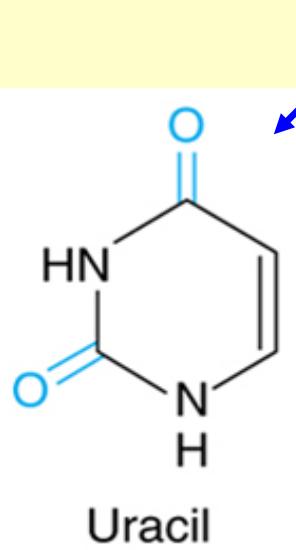
二個重點：(1) 反平行(anti-parallel)之二股DNA
(2) major groove(主要凹槽) & minor groove(次要凹槽)

RNA之化學性質

(1) RNA: ribose
(DNA: 2'-deoxyribose)



(2) RNA: U (Uracil) 取代 T (Thymine)



尿嘧啶

(3) RNA: 通常是為單股

(4) 三種主要之RNA: mRNA, tRNA & rRNA

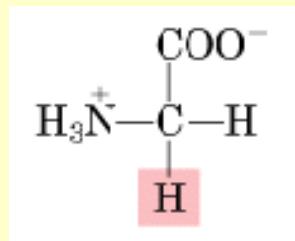
Protein (蛋白質)

(1) 蛋白質為胺基酸所組成之多肽鏈複合物 (polymers)

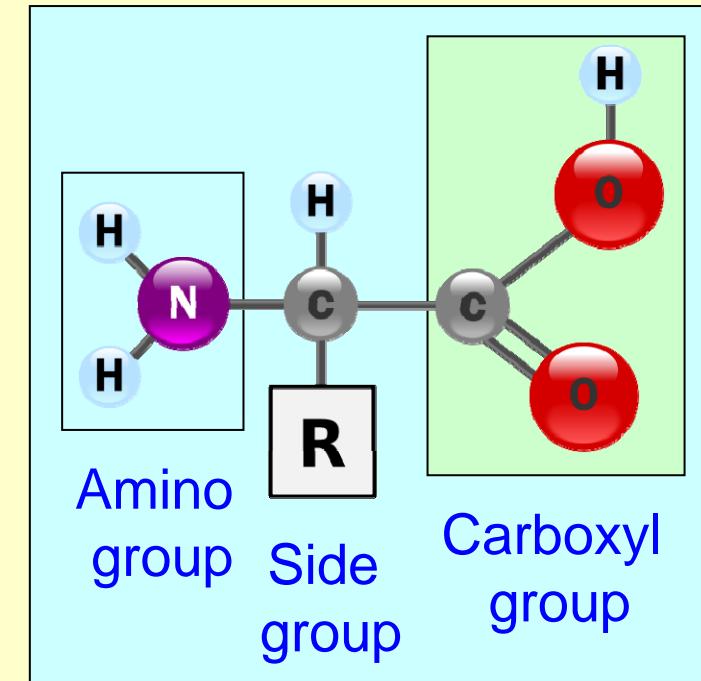
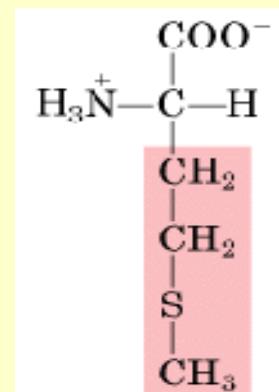
(2) 存在有20種不同之胺基酸
- 具有相同之結構
- 但有一個不同之 R group

(3) 例如：

Glycine (甘胺酸): R= -H (構造最簡單)

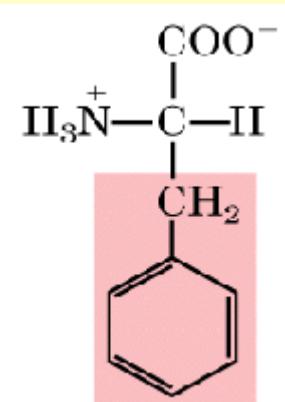


Methionine (甲硫胺酸):
R= -C-C-S-C

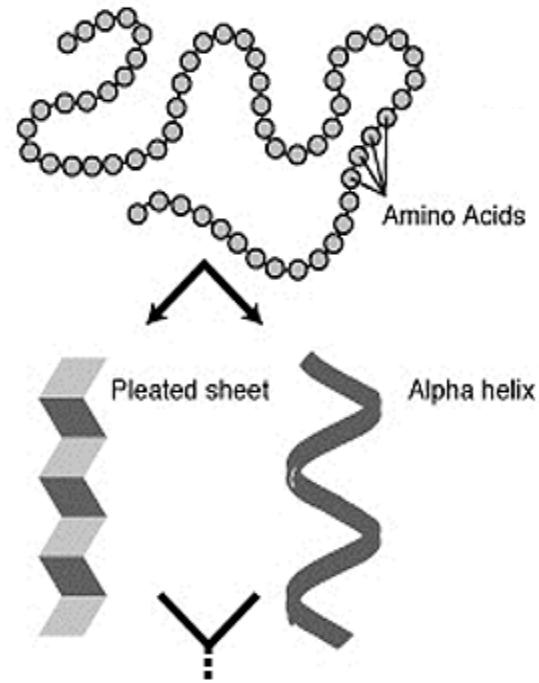


Phenylalanine (苯丙胺酸):

R= -C-[C_6H_5]



不同層次之蛋白質結構

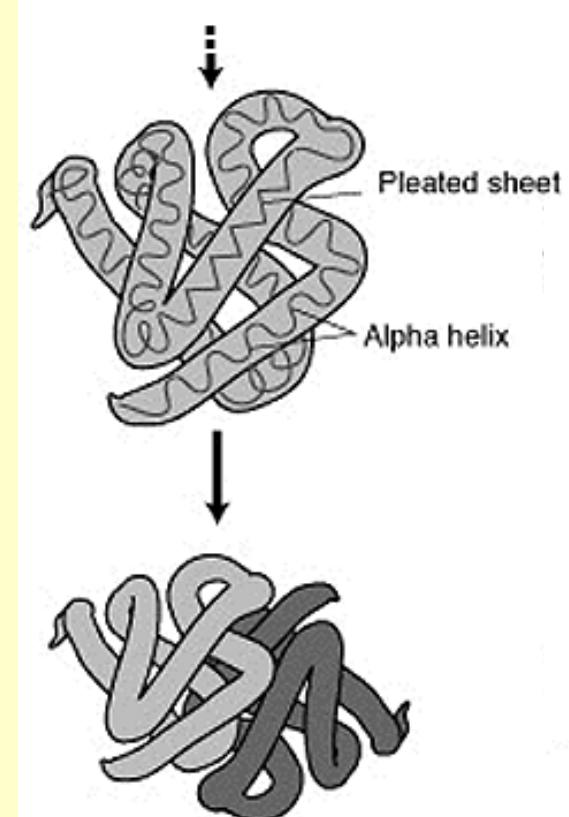


(1)一級結構:組成蛋白質多肽鏈的線性胺基酸序列

(2)二級結構:依靠不同胺基酸之間的C=O和N-H基團間的氫鍵形成的穩定結構，主要為 α 螺旋和 β 摺疊

(3)三級結構:通過多個二級結構元素在三度空間排列所形成的一個蛋白質分子的三維結構

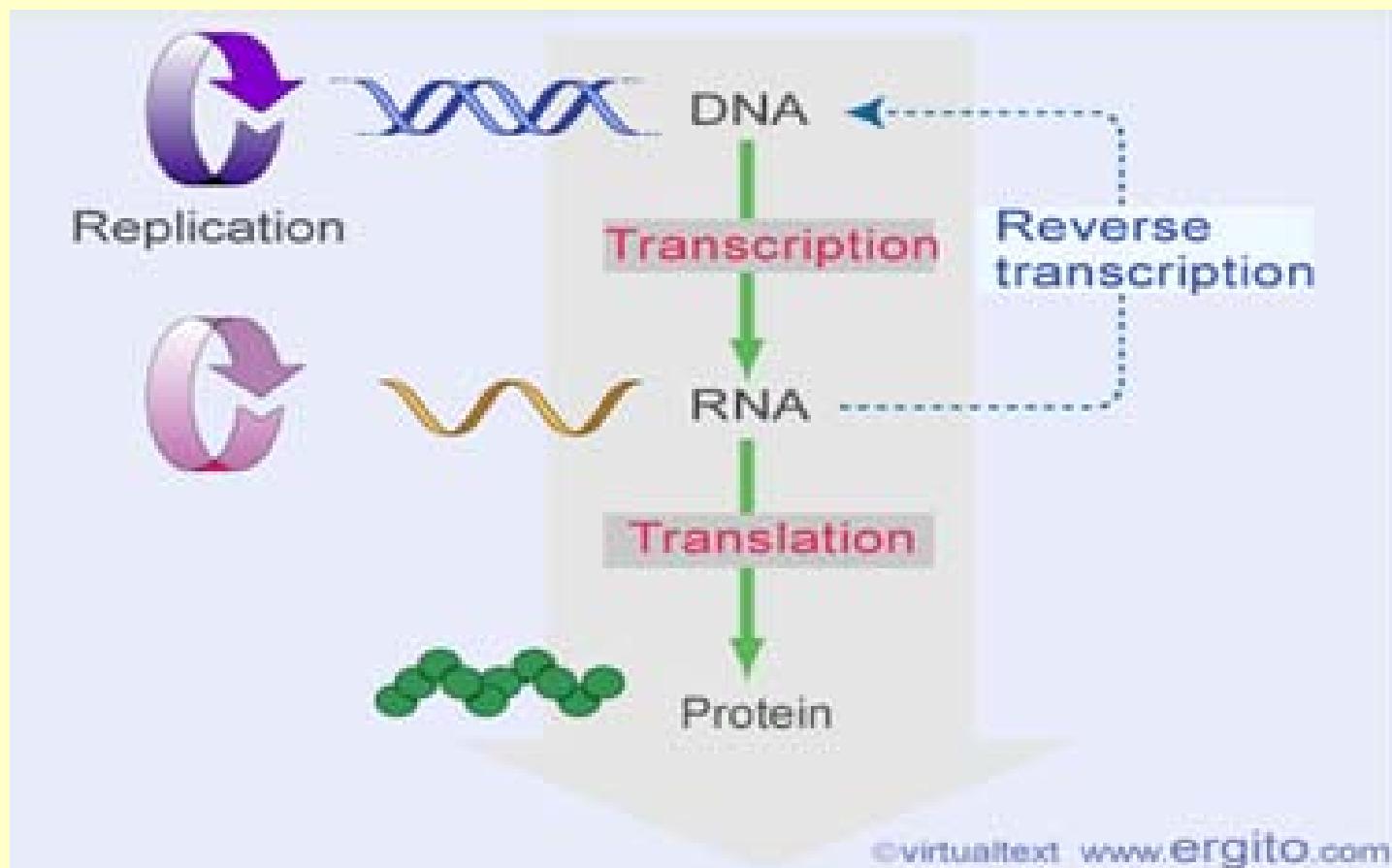
(4)四級結構:用於描述由不同多肽鏈間相互作用形成具有功能的蛋白質複合物分子



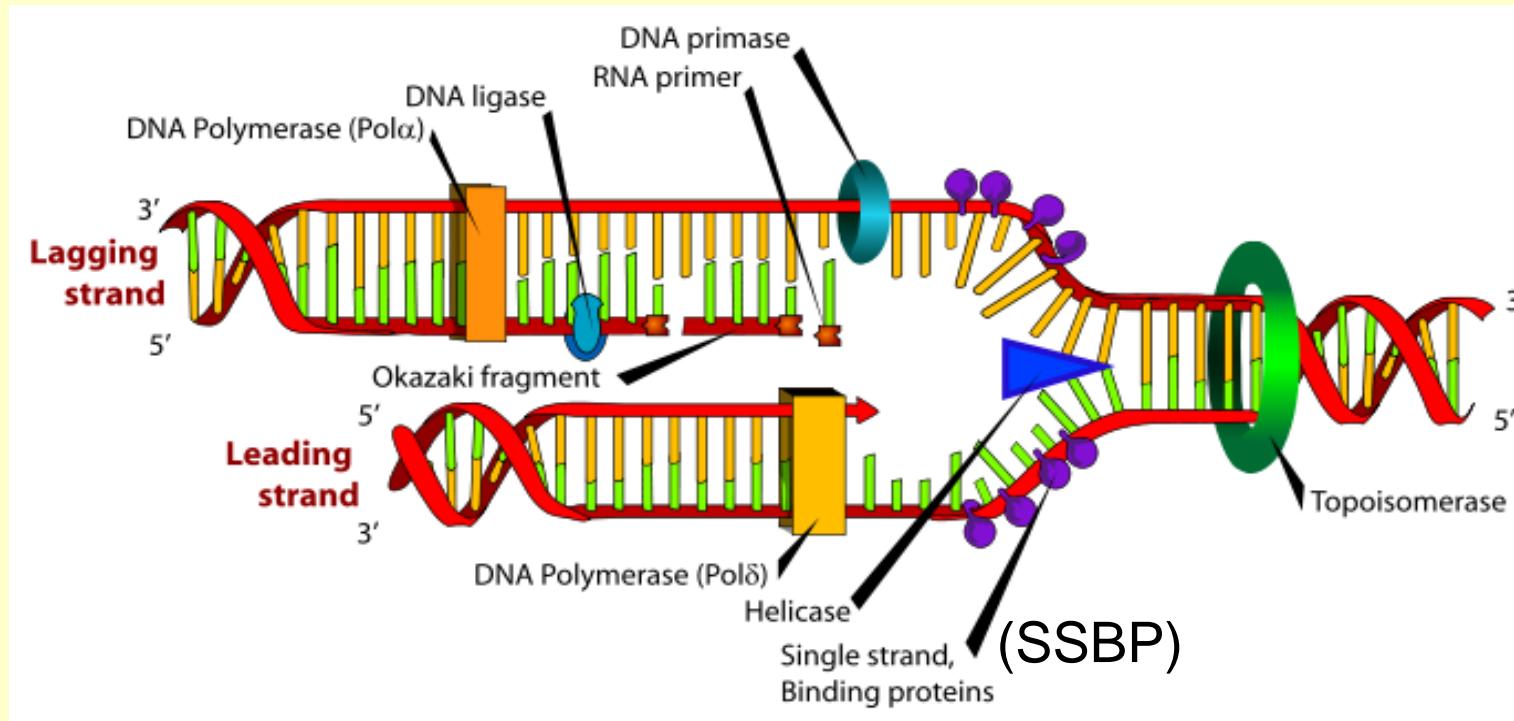
蛋白質依照其在生物體中不同功能而分類

Type:	Example:
Enzymes- Catalyze biological reactions	β -galactosidase
Transport and Storage	Hemoglobin
Movement	Actin And Myosin in muscles
Immune Protection	Immunoglobulins (antibodies)
Regulatory Function within cells	Transcription Factors
Hormones	Insulin Estrogen
Structural	Collagen 膠原蛋白

- The Central Dogma describes the Information flow from DNA → RNA → protein



DNA replication (DNA 複製): a simplified view



- 細胞中，於複製叉形成之處進行DNA複製
- DNA螺旋解開後，解旋酶(helicase)持續將複製叉之前的DNA解開，單股結合蛋白(SSBP)會結合至單股DNA上，以保持解開狀態
- 其中一股DNA會做為模板，進行延遲股("lagging" strand)之非連續式DNA合成
- 另一股DNA會做為模板，進行前導股("leading" strand)之連續式DNA合成

<http://www.stolaf.edu/people/giannini/flashanimat/molgenetics/dna-rna2.swf>

DNA ---> RNA ---> proteins

5' ...A T G G C C T G G A C T T C A... 3' **Sense strand of DNA**
3' ...T A C C G G A C C T G A A G T... 5' **Antisense strand of DNA**

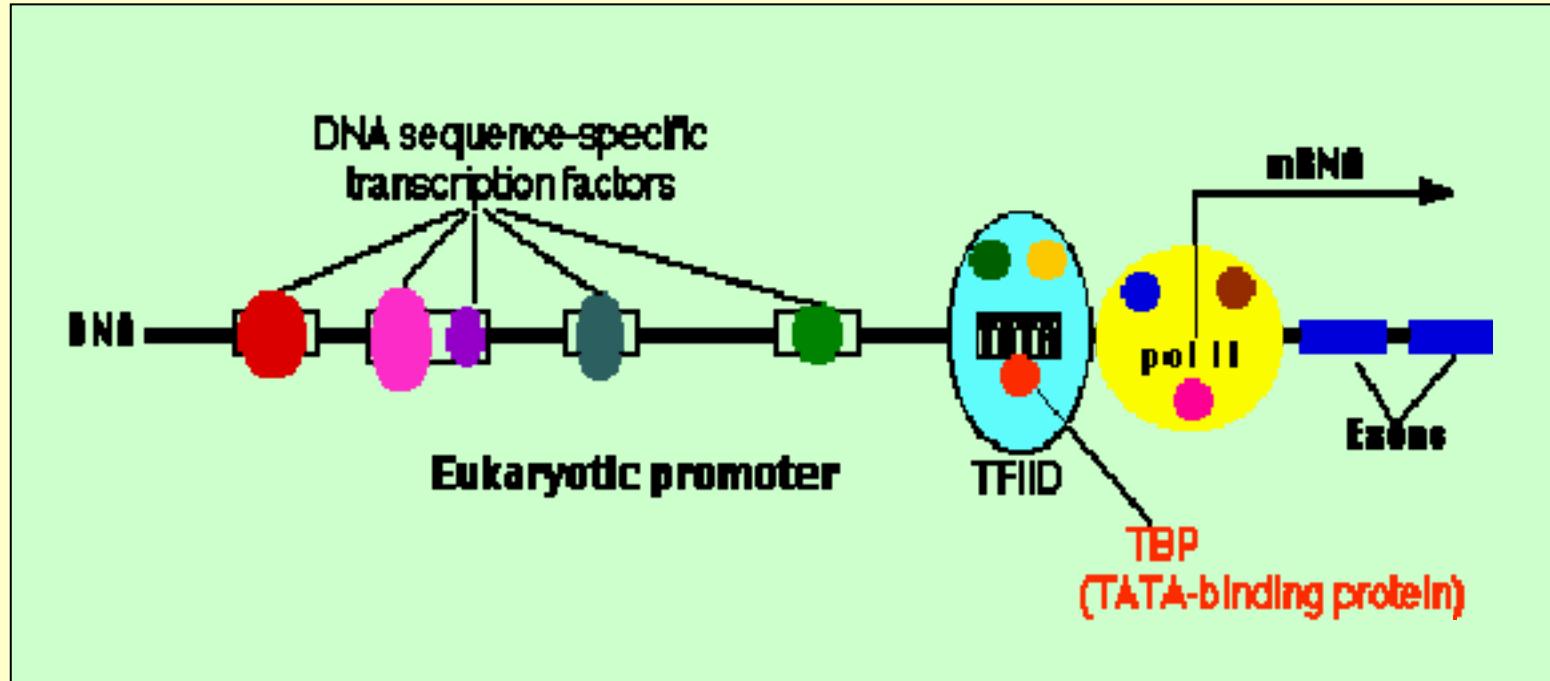
↓ **Transcription of antisense strand**

5' ...A U G G C C U G G A C U U C A... 3' **mRNA**

↓ **Translation of mRNA**

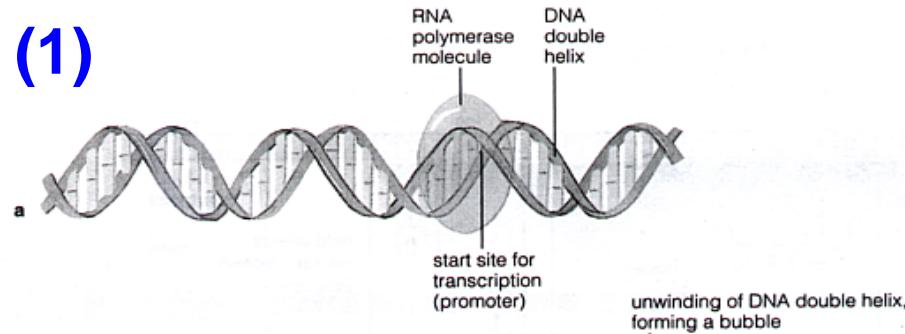
Met—Ala—Trp—Thr—Ser—**Peptide**

Transcription (轉錄): a simplified view



- 例如: 將DNA轉錄成mRNA
- DNA sequence ("promoter" 啟動子)
- Proteins: RNA polymerase
transcription factors (activators or repressors)
- Ribonucleotides

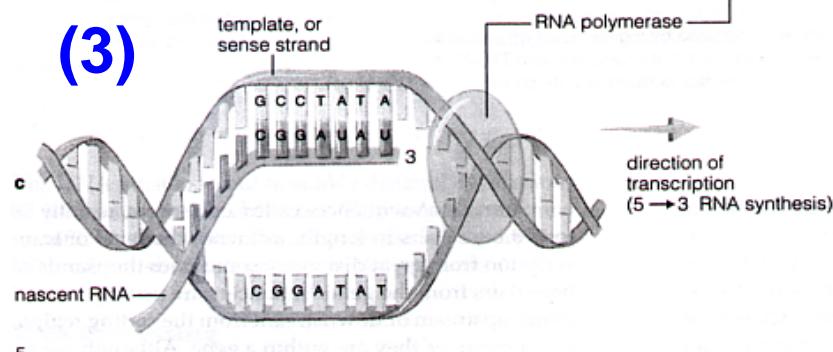
(1)



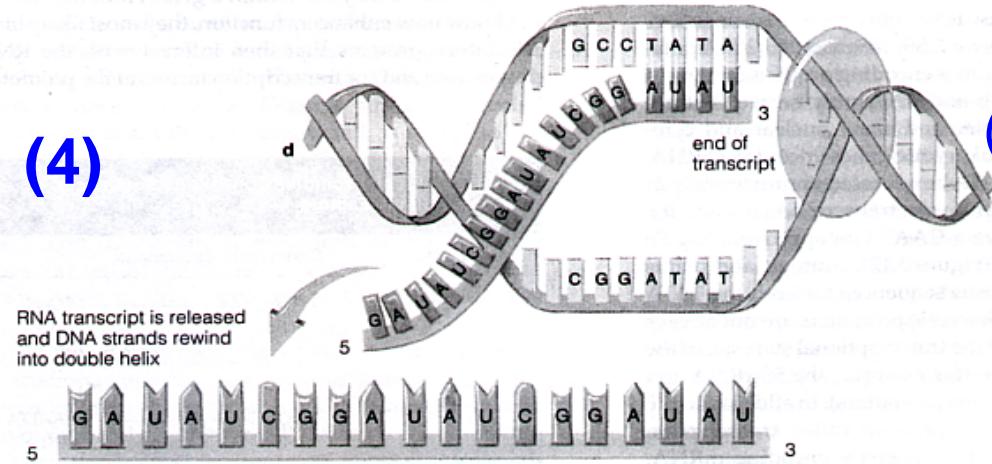
(2)



(3)



(4)



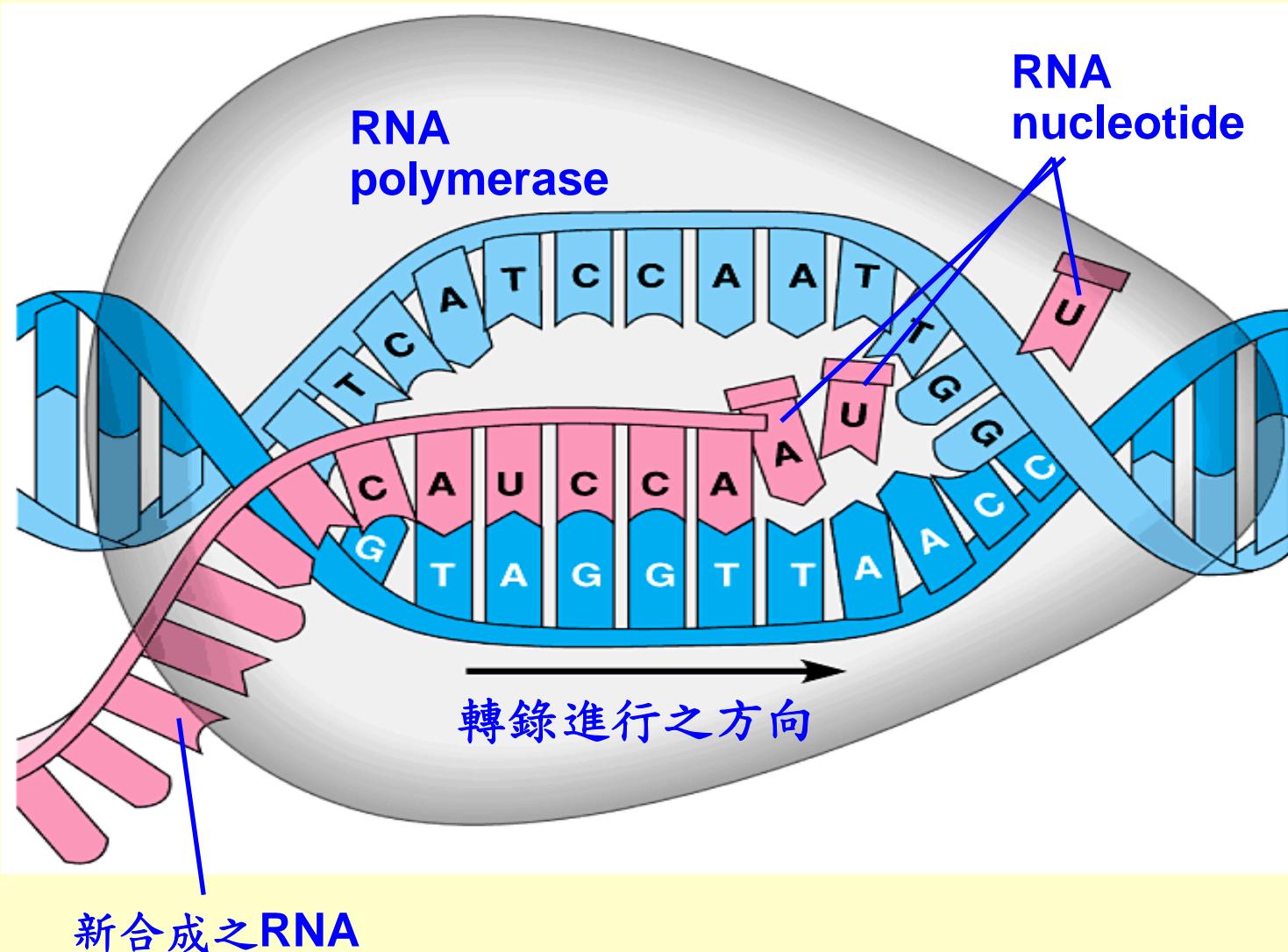
(1) Transcription initiation

(轉錄啟始): RNA polymerase binding to DNA之啟動子上

(2) 雙股DNA解開

(3) RNA polymerase 沿著DNA移動，加入核苷酸，開始生成RNA

(4) 當RNA polymerase 移動至 terminator時，合成之RNA會被釋放，而終止DNA轉錄



<http://www.stolaf.edu/people/giannini/flashanimat/molgenetics/transcription.swf>

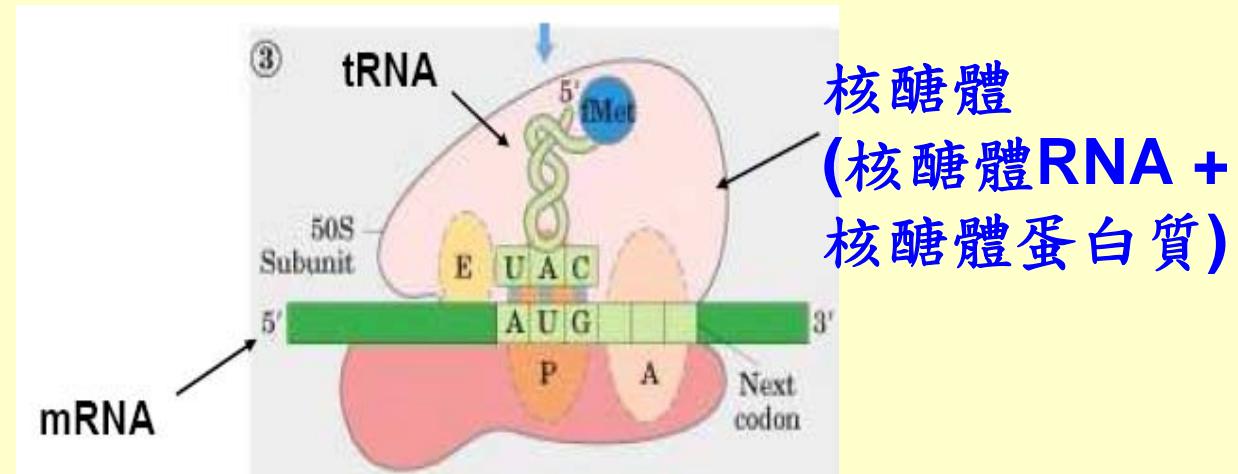
三類主要之RNA

(1) messenger RNA (mRNA, 訊息RNA)

- an intermediate between genetic information (DNA) to generate its product (protein)

(2) ribosomal RNA (rRNA, 核糖體RNA)

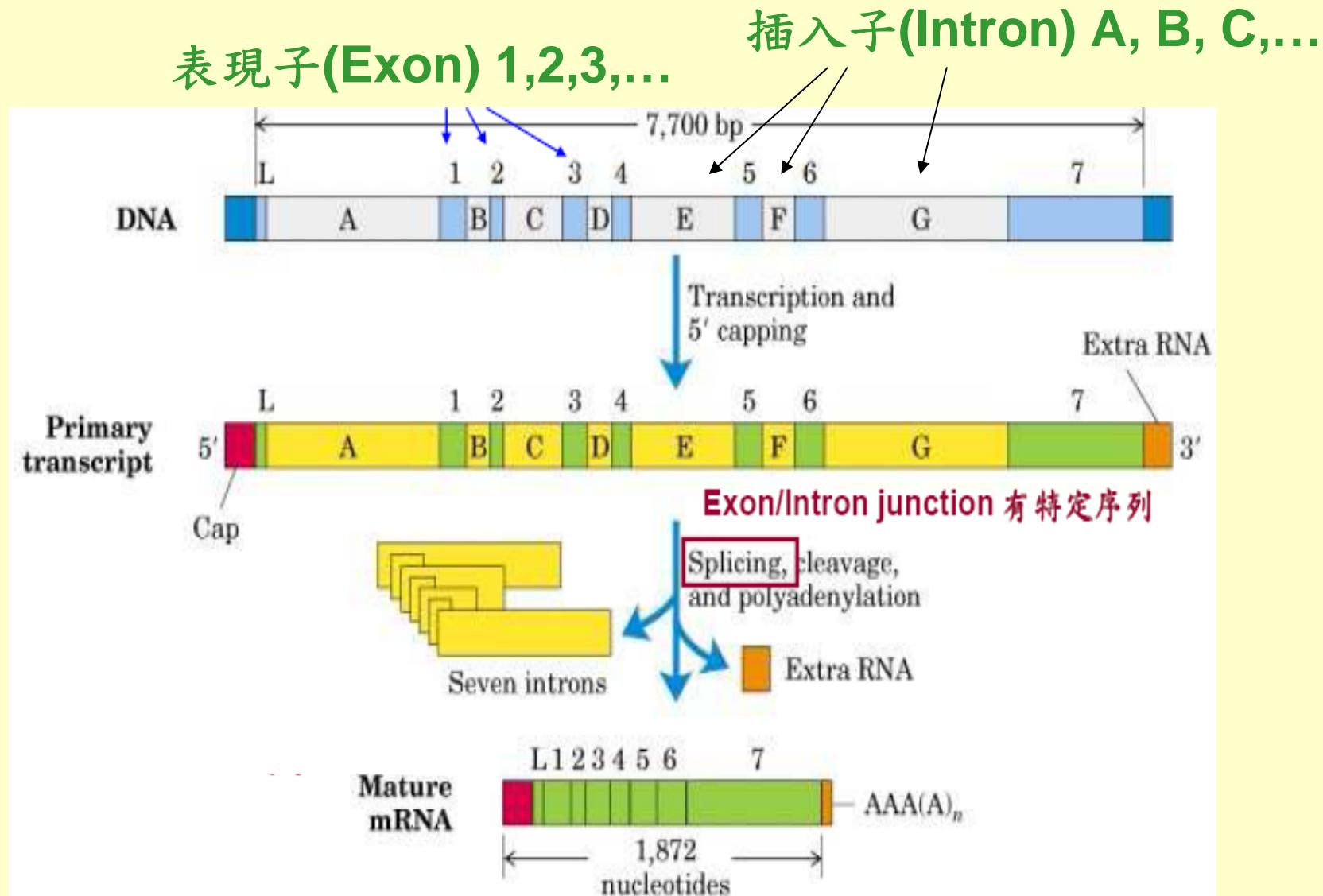
- an essential component of the ribosome where proteins are synthesized

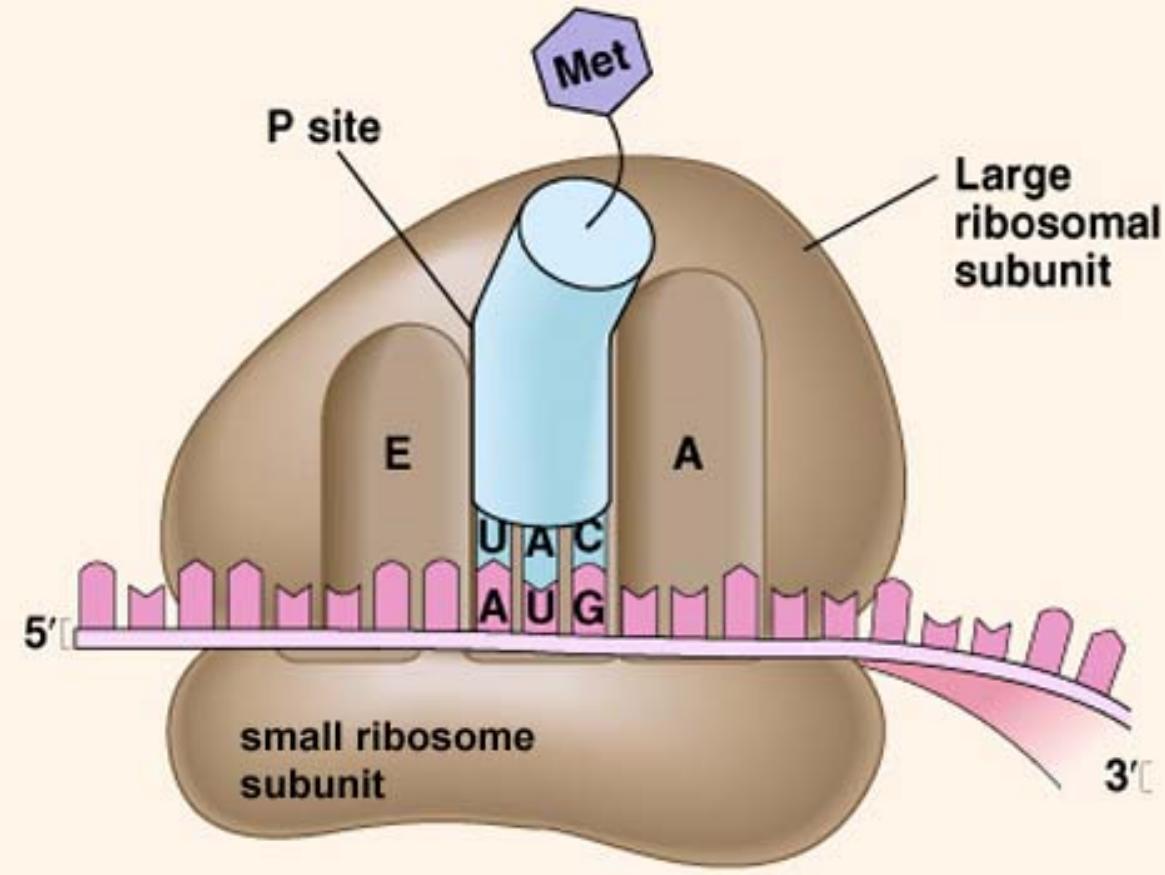


(3) transfer RNA (tRNA, 轉移RNA)

- carries amino acids to the ribosome for protein synthesis

真核細胞之mRNA 的編輯 (RNA editing)



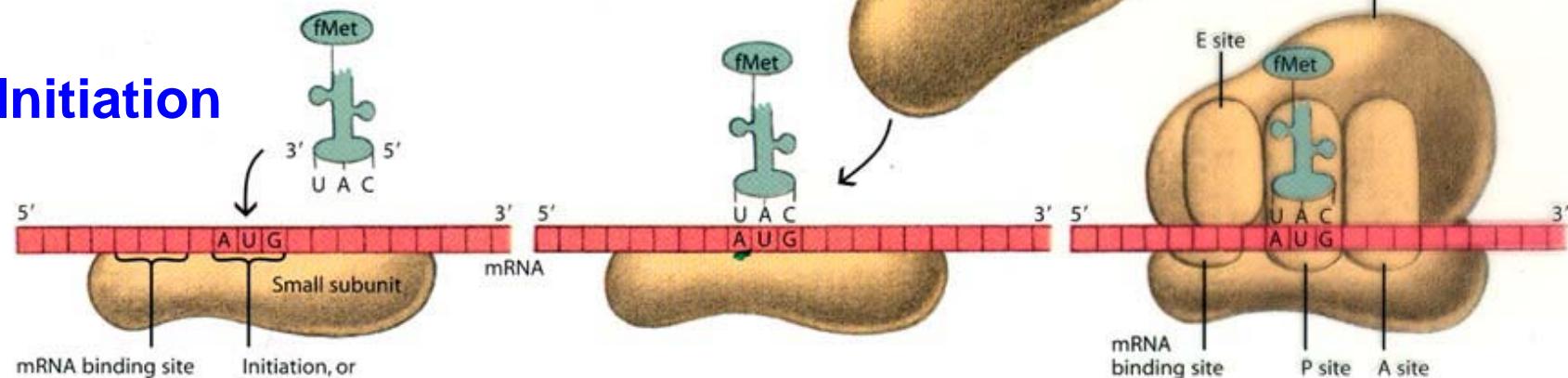


A site, P site &
E site

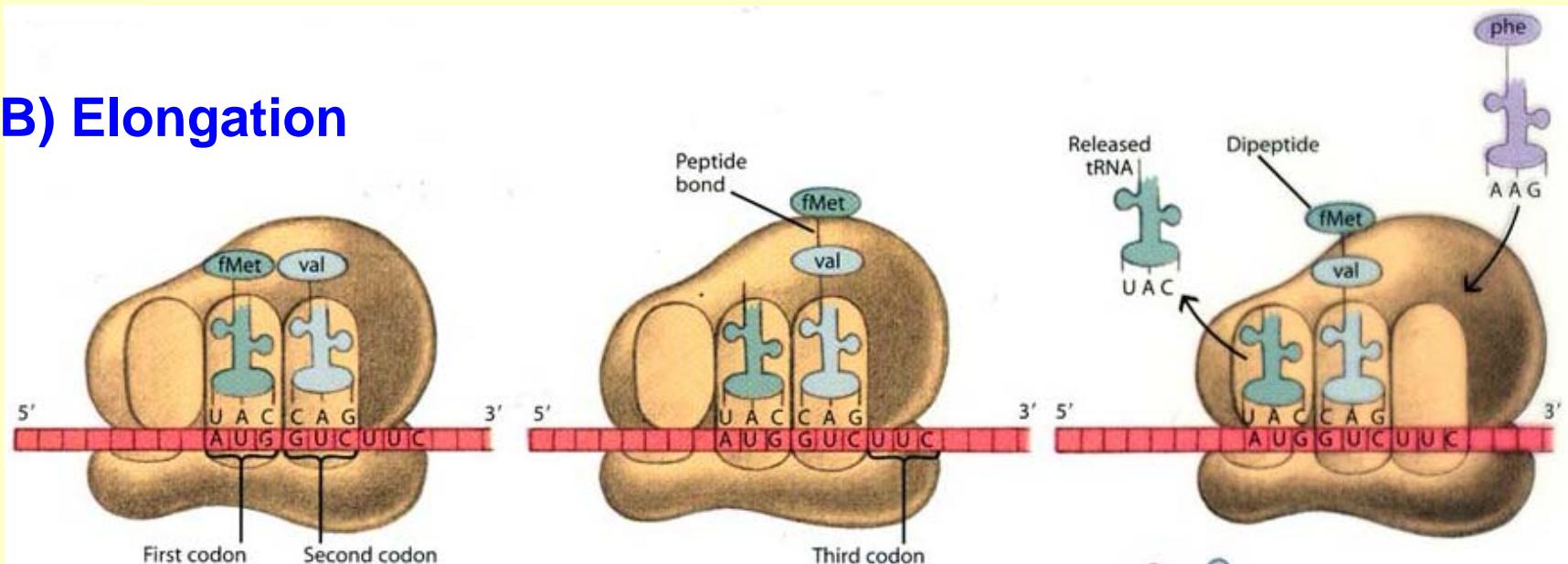
Translation (轉譯): a simplified view

- 將mRNA轉錄成蛋白質
- 需要Ribosome (核糖體) , tRNA , amino acids (胺基酸)

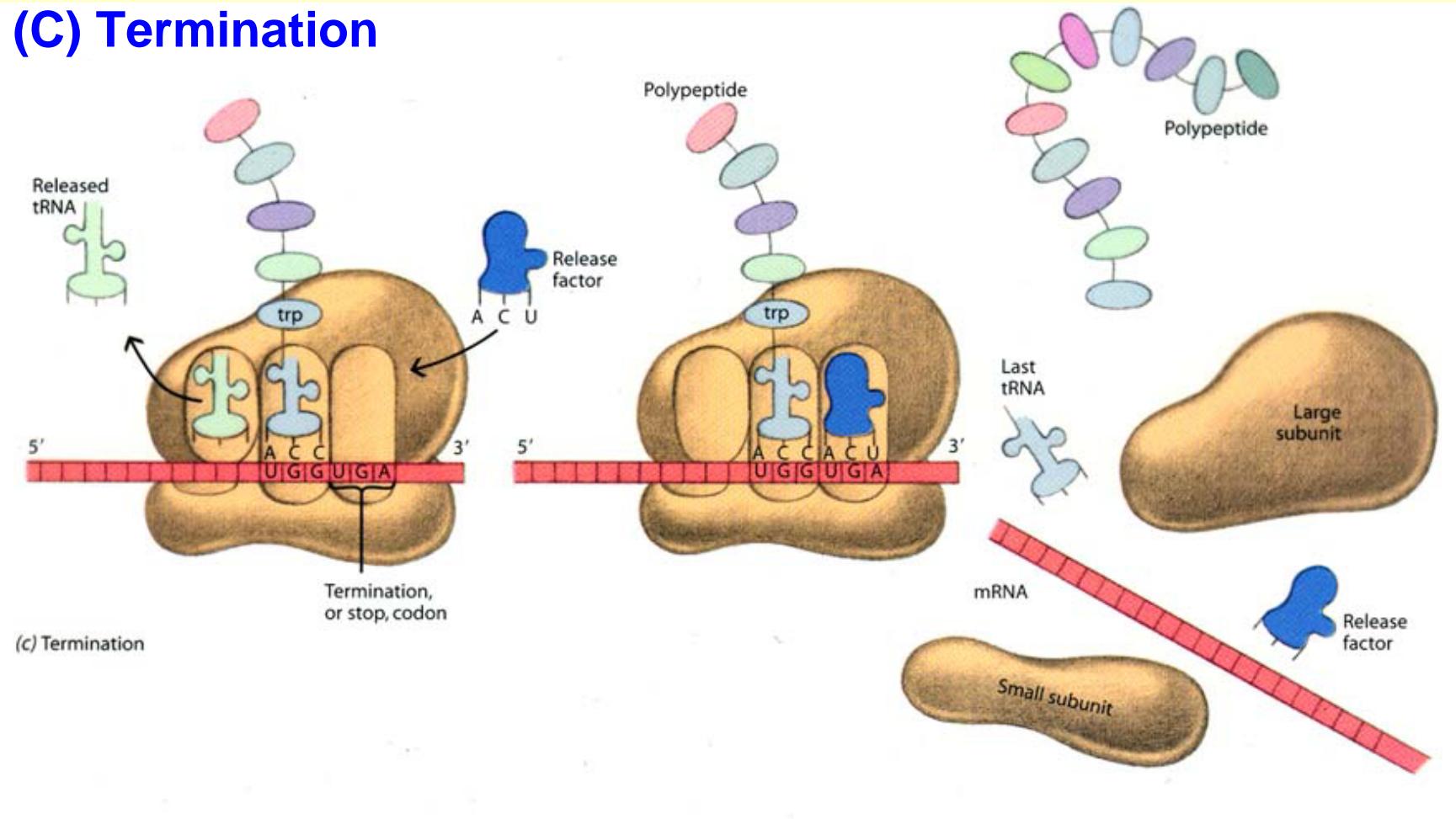
(A) Initiation

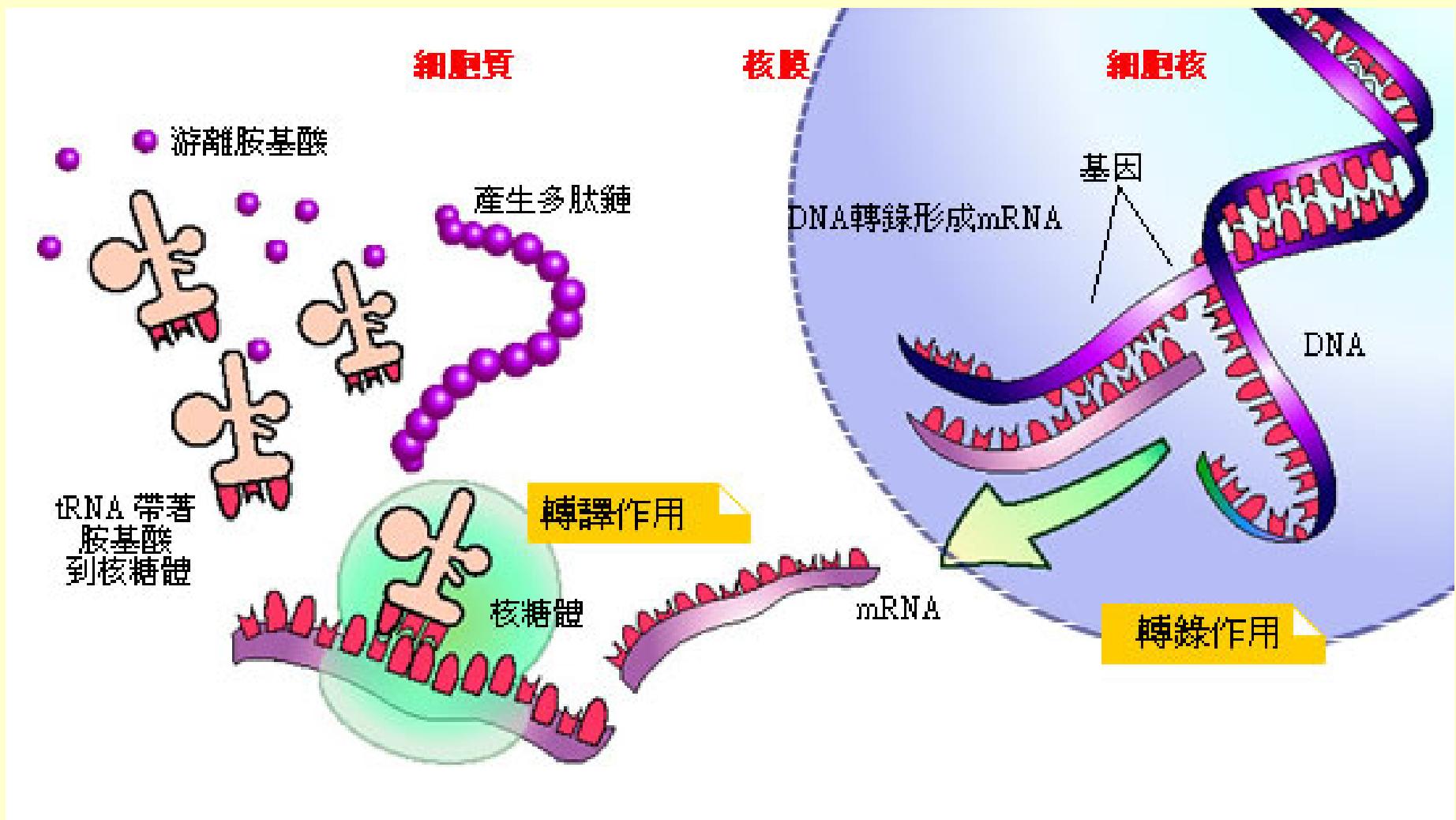


(B) Elongation



(C) Termination





細胞如何讀出DNA密碼?

美國科學家伽莫夫，首先提出了三聯體假說。他認為DNA長鏈上的連續3個核苷酸會構成一個密碼子(codon)，一個密碼子決定一種胺基酸

		Second letter					
		U	C	A	G		
First letter	U	UUU UUC UUA UUG	UCU UCC UCA UCG	UAU UAC UAA UAG	Tyr Ser Stop Stop	UGU UGC UGA UGG	Cys Stop Trp
	C	CUU CUC CUA CUG	CCU CCC CCA CCG	CAU CAC CAA CAG	His Pro Gin	CGU CGC CGA CGG	U C A G
	A	AUU AUC AUA AUG	ACU ACC ACA ACG	AAU AAC AAA AAG	Asn Thr Lys	AGU AGC AGA AGG	U C A G
	G	GUU GUC GUA GUG	GCU GCC GCA GCG	GAU GAC GAA GAG	Asp Ala Glu	GGU GGC GGA GGG	U C A G

- 密碼子使用表
(codon usage table)
- 由美國科學家霍拉那等人完成
- 20 胺基酸, 但有64 遺傳密碼子
- Some amino acids are encoded by multiple codons

密碼子使用表之另一種呈現方式

Amino Acid	SLC	DNA codons
Isoleucine	I	ATT, ATC, ATA
Leucine	L	CTT, CTC, CTA, CTG, TTA, TTG
Valine	V	GTT, GTC, GTA, GTG
Phenylalanine	F	TTT, TTC
Methionine	M	ATG
Cysteine	C	TGT, TGC
Alanine	A	GCT, GCC, GCA, GCG
Glycine	G	GGT, GGC, GGA, GGG
Proline	P	CCT, CCC, CCA, CCG
Threonine	T	ACT, ACC, ACA, ACG
Serine	S	TCT, TCC, TCA, TCG, AGT, AGC
Tyrosine	Y	TAT, TAC
Tryptophan	W	TGG
Glutamine	Q	CAA, CAG
Asparagine	N	AAT, AAC
Histidine	H	CAT, CAC
Glutamic acid	E	GAA, GAG
Aspartic acid	D	GAT, GAC
Lysine	K	AAA, AAG
Arginine	R	CGT, CGC, CGA, CGG, AGA, AGG
Stop codons	Stop	TAA, TAG, TGA

常用分子生物/DNA之分析技術

1. Recombinant DNA (Gene cloning)
2. PCR (polymerase chain reaction)
3. Genome sequencing

Recombinant DNA (重組DNA或基因操作)

Recombinant DNA: 任何以人工重組出來之DNA分子，其中常使自然界中不會同時存在之DNA序列，經重組使其同時併存

Gene manipulation (基因操作) or genetic engineering (基因工程):
任何可用於製造重組DNA之不同技術

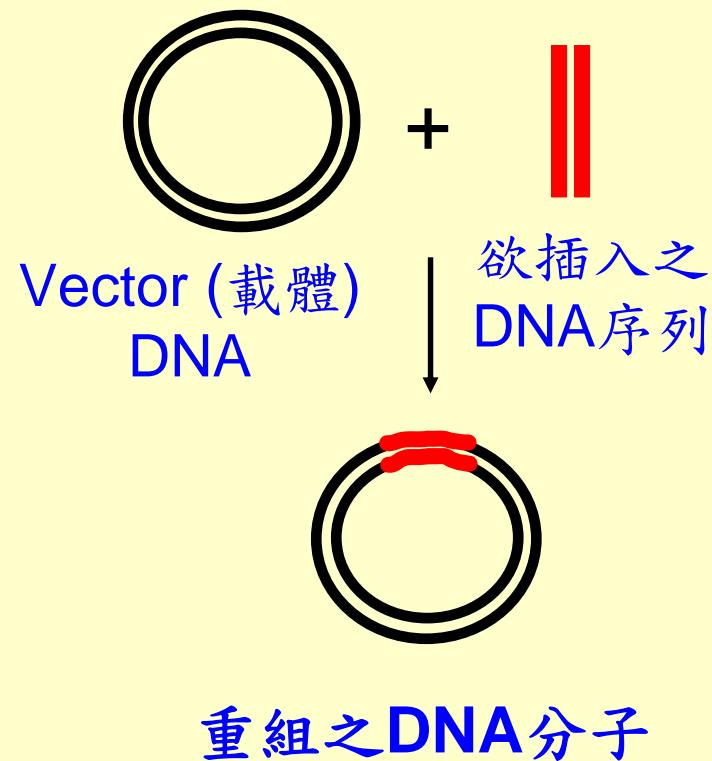
Clone: a population of identical organisms derived from a single parental organism

Cloning: 將重組DNA置於一個特殊宿主細胞(例如:大腸桿菌)使其增殖，如此可產生許多份相同之DNA序列

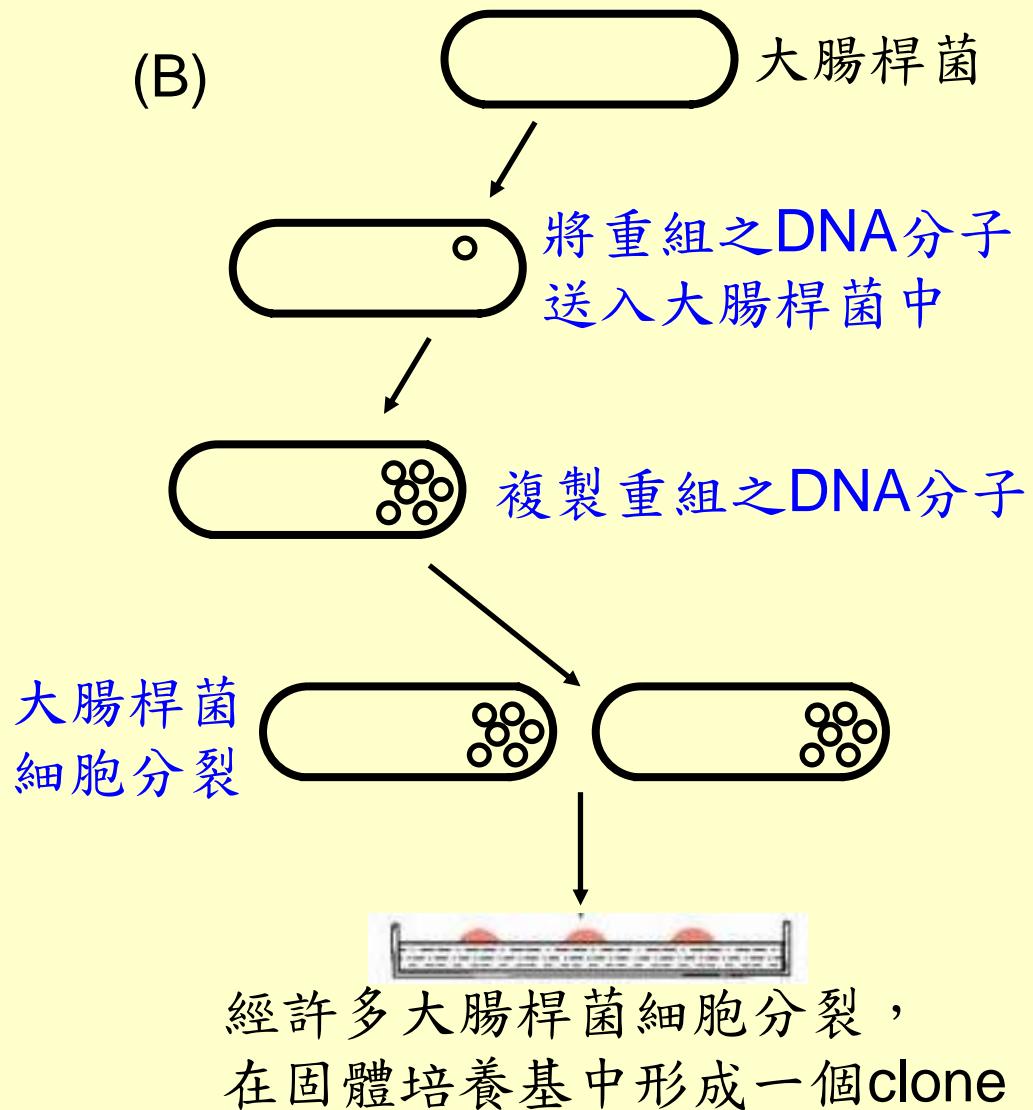
Gene manipulation(基因操作) involves the creation & cloning of recombinant DNA

基因操作之基本步驟

(A)

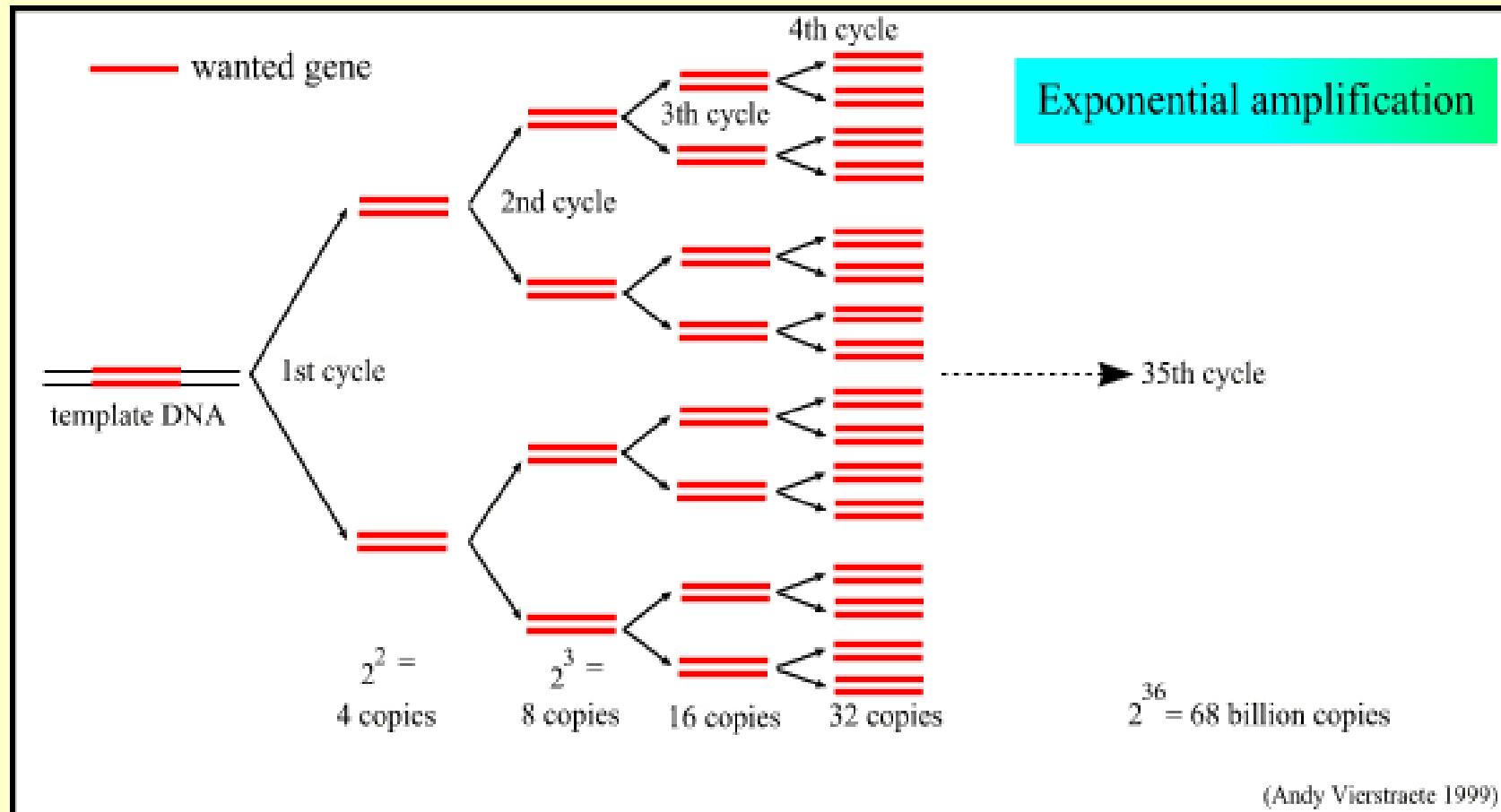


(B)



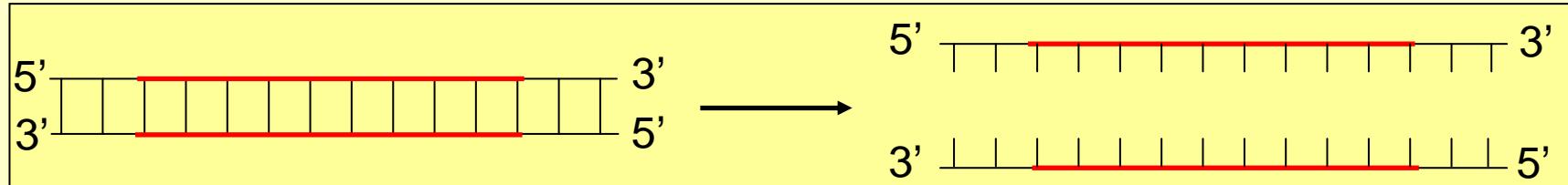
Polymerase chain reaction (PCR; DNA 聚合酶連鎖反應)

PCR可將指定的DNA基因片段之數量放大

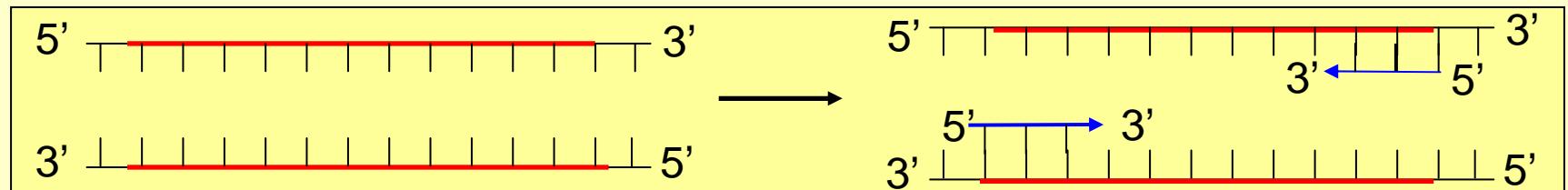


PCR之基本步驟

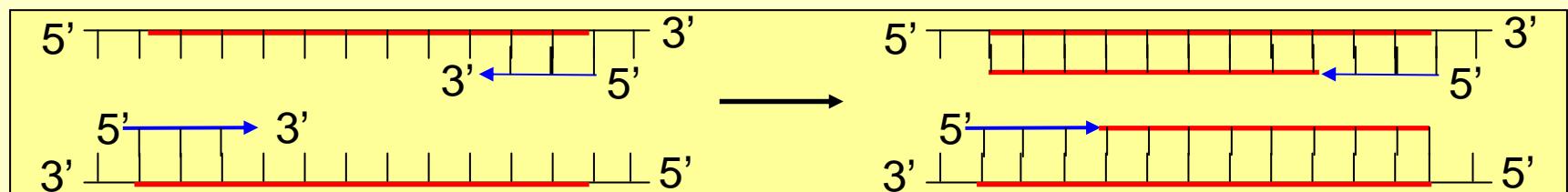
(1) 加入含有目標基因序列之DNA分子，加熱使其兩股分開



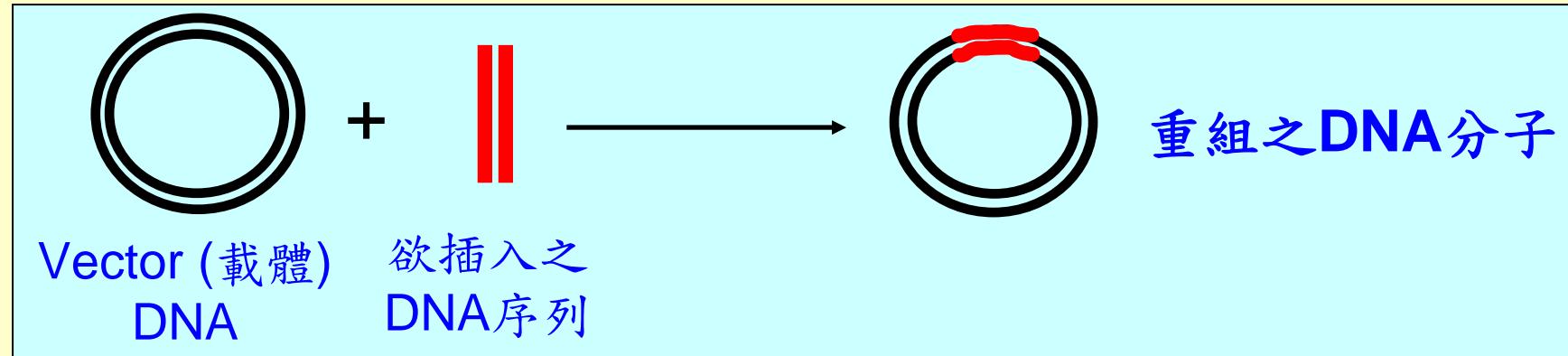
(2) 加入兩種應引子(primer)，使此二引子接合於DNA分子上，而各自界定目標序列之兩端



(3) 加入聚合酶(DNA polymerase)，由引子延伸合成另一股DNA

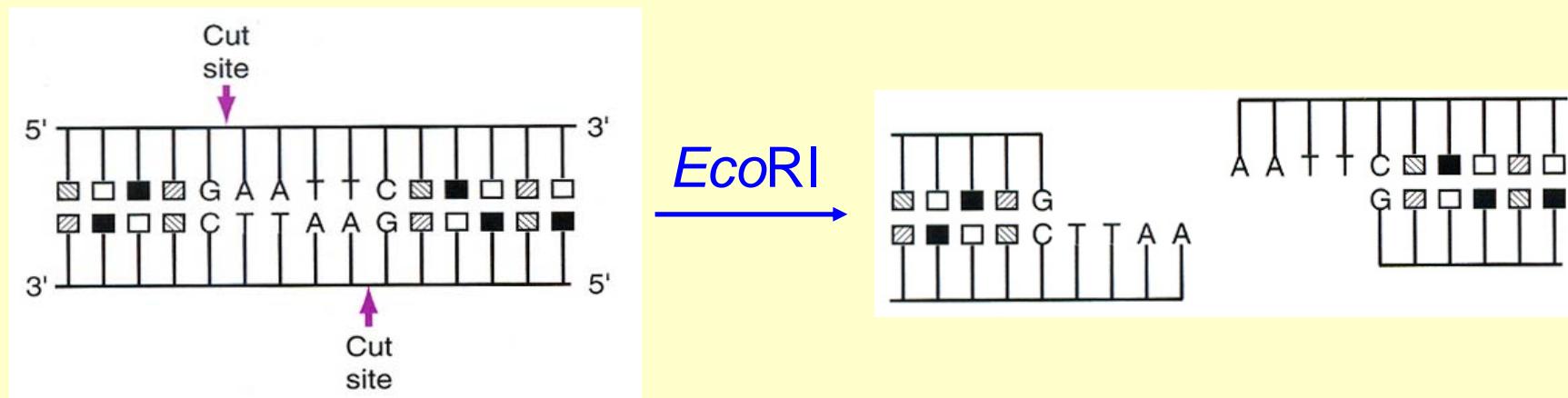


(4) 完成一次複製循環，可重複使用所有引子及酵素，經過n次複製循環，目標序列之分子數量因此倍增(2^n)



欲插入之DNA序列：可由PCR取得，或由一段DNA中”剪切“取得

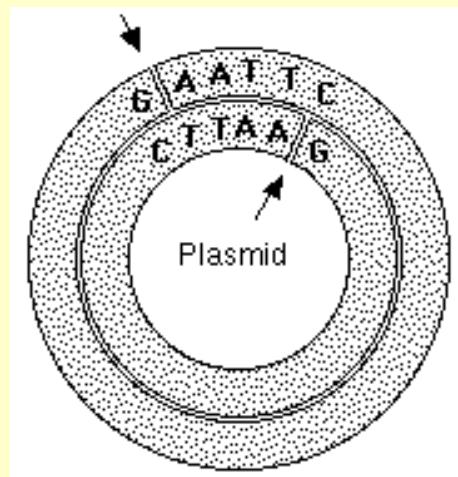
DNA之”剪切”: restriction enzymes (限制酶，又稱限制性內切酶)，切割方法是將糖類分子與磷酸之間的鍵結切斷，進而於兩條DNA鏈上各產生一個切口，且不破壞核苷酸與鹼基。例如：*EcoRI*



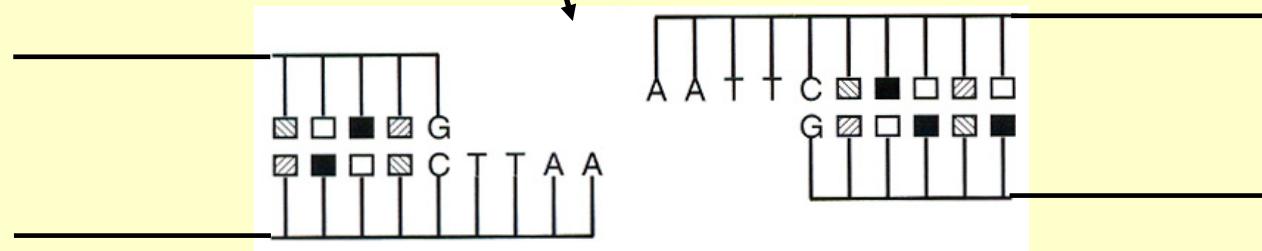
更多例子：

酵素名稱	辨識序列	切法
<i>BamHI</i>	5'GGATCC 3'CCTAGG	5'---G GATCC---3' 3'---CCTAG G---5'
<i>PstI</i>	5'CTGCAG 3'GACGTC	5'---CTGCA G---3' 3'---G ACGTC---5'
<i>SmaI</i>	5'CCCGGG 3'GGGCC	5'---CCC GGG---3' 3'---GGG CCC---5'
<i>EcoRV</i>	5'GATATC 3'CTATAG	5'---GAT ATC---3' 3'---CTA TAG---5'

EcoRI



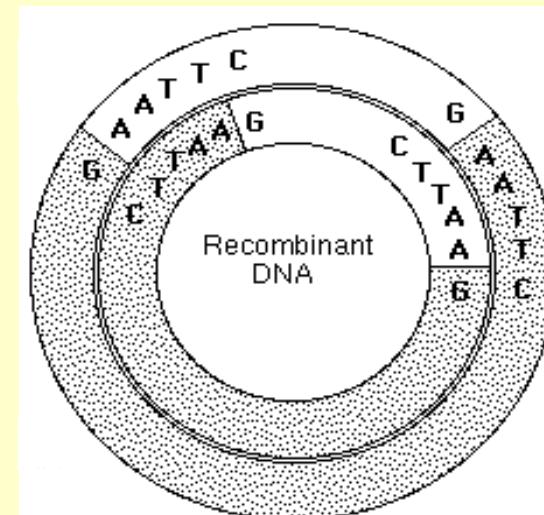
EcoRI



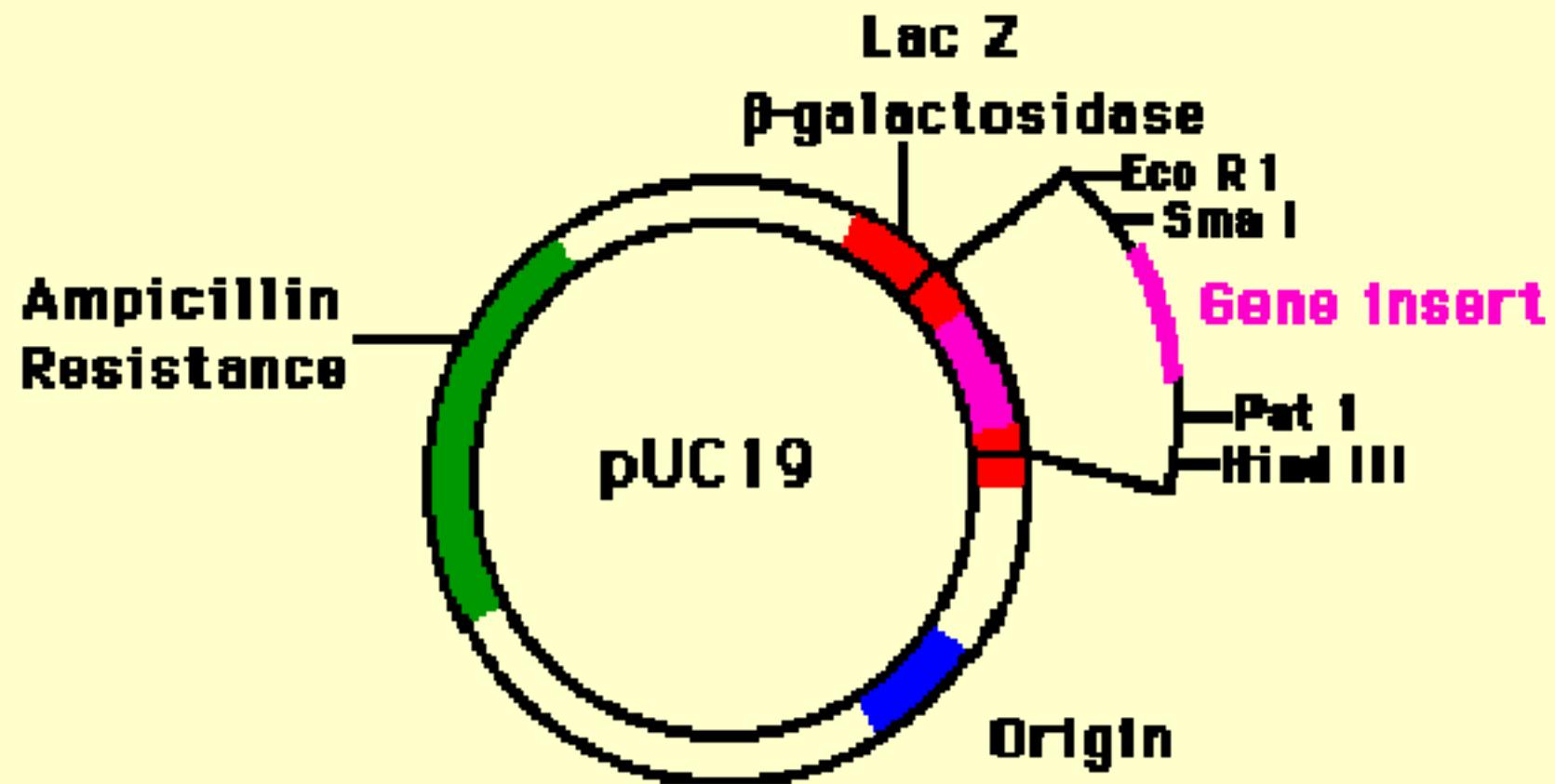
欲插入之DNA序列

Vector (載體) DNA

DNA 連接酶
(DNA ligase)

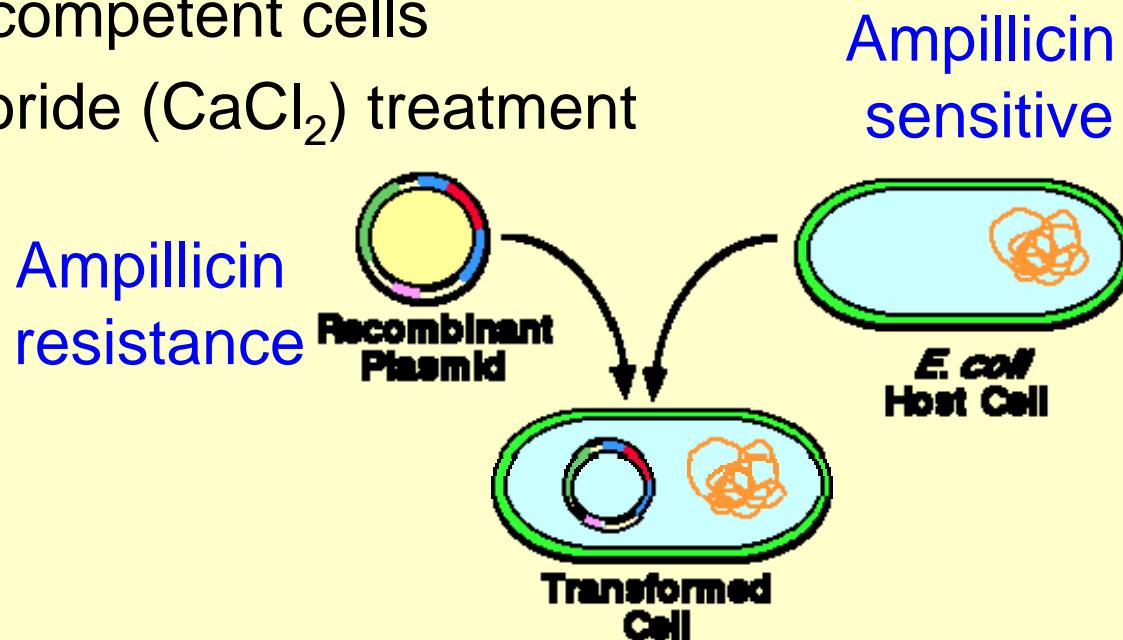


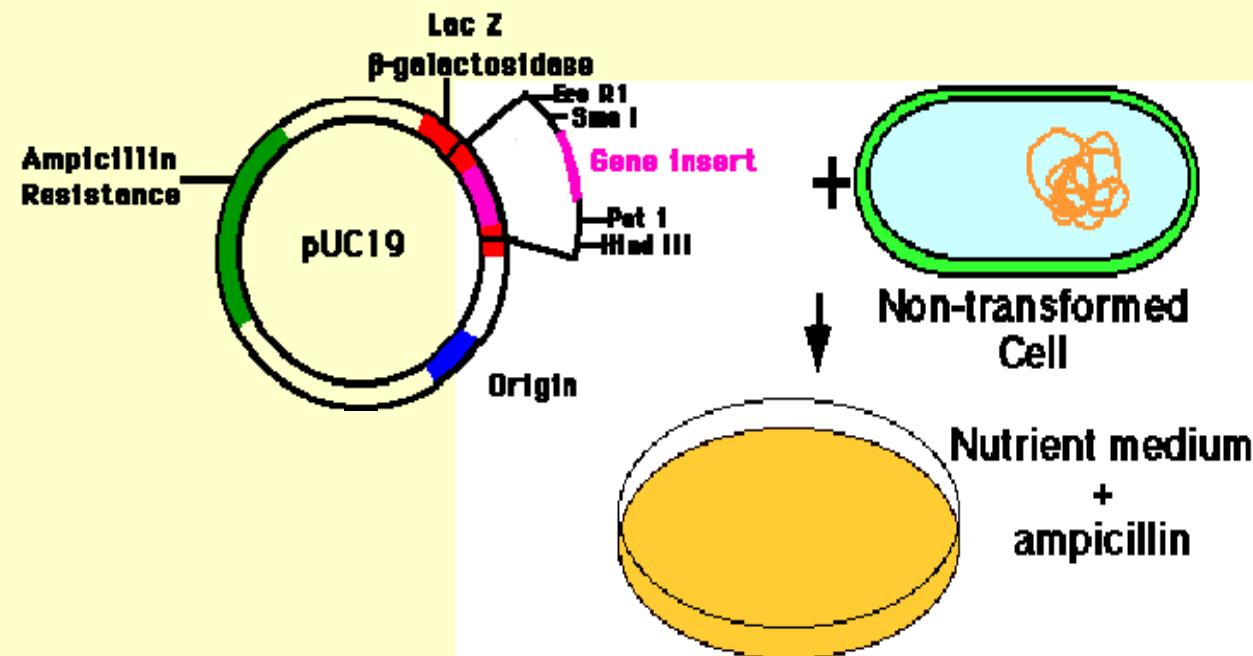
Plasmid: an example of the vector DNA



將重組之DNA分子送入大腸桿菌中

- Electroporation
- Chemical transformation
 - Competent cells (勝任細胞): make cells more permeable to DNA
 - Preparation of competent cells
 - Calcium chloride (CaCl_2) treatment





Selection for bacteria can grow in the medium containing antibiotic (e.g. ampicillin resistance)

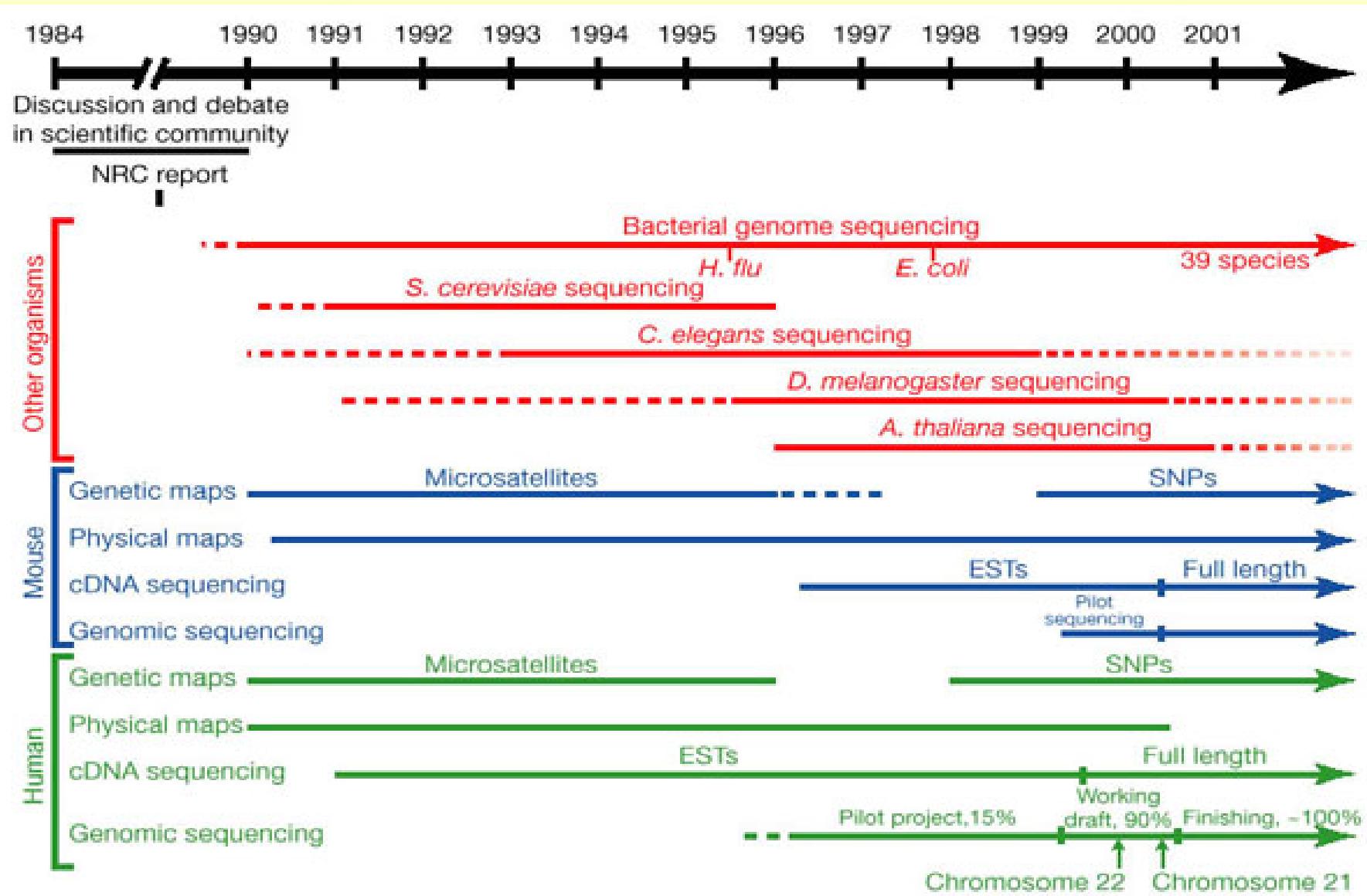
Gene (基因): 能轉譯或轉錄成rRNA, tRNA或蛋白質之DNA片段或序列

Genome (基因體或基因組): 存在於一個細胞中之整組基因(即細胞中之所有遺傳物質)

Genomics (基因體學): 研究基因體分子結構，基因體中所含之訊息，及其可轉錄生成之蛋白質等之科學

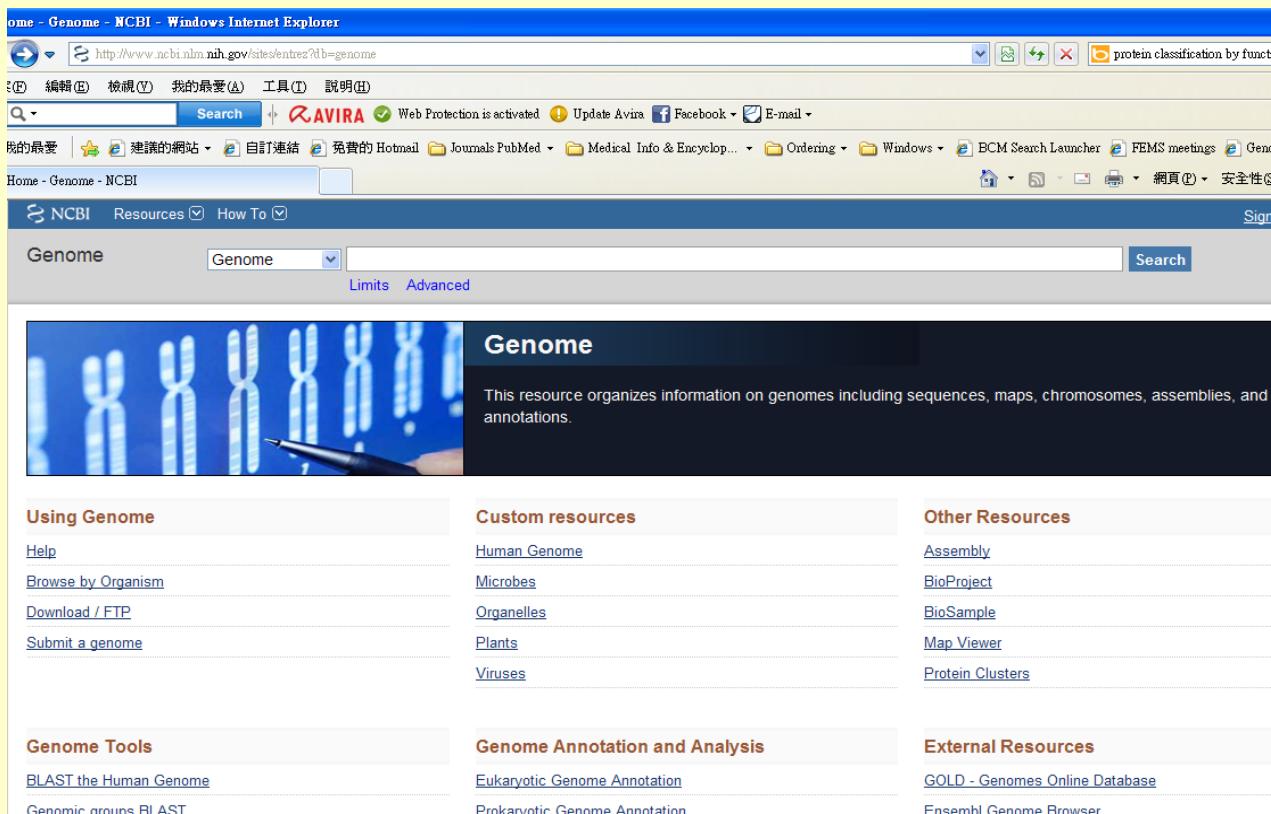
Functional Genomics (功能基因體學): 研究基因体中之基因及其它部份之功能

Genome sequencing (基因體定序)之簡史



Genome projects (基因體計劃)

- 1995: 首度完成二個細菌之基因定序
(Haemophilus influenzae, Mycoplasma genitalium)
- 之後，更多且更準確之基因定序方法被陸續發展出來



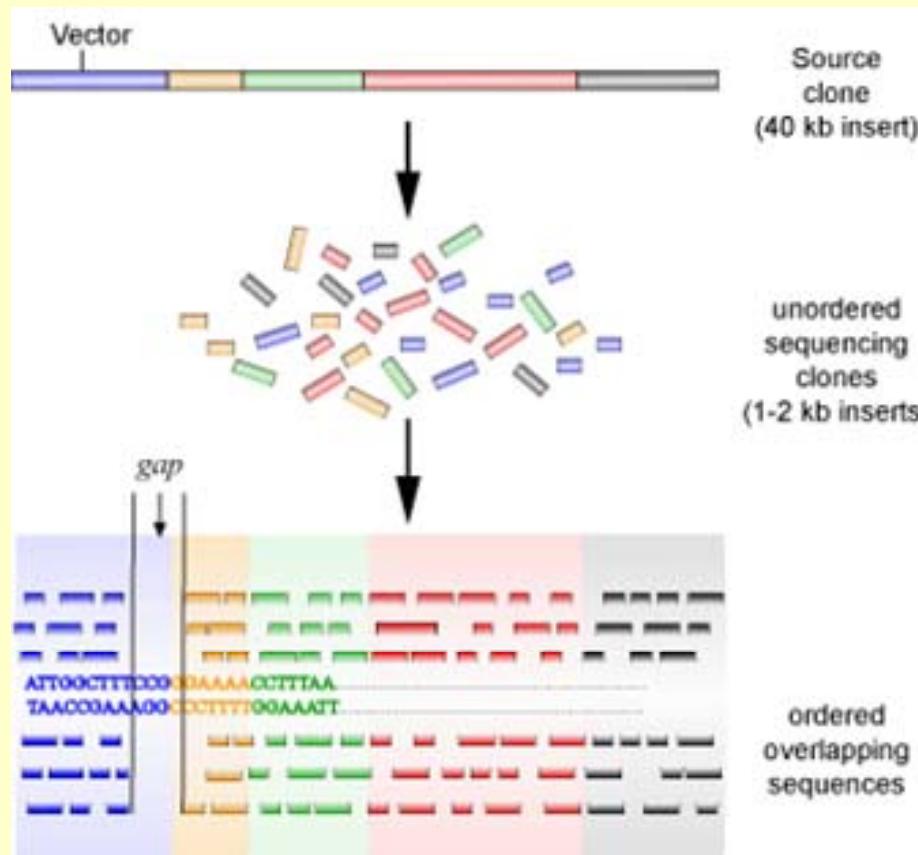
The screenshot shows the NCBI Genome homepage as it appeared in Windows Internet Explorer. The URL in the address bar is <http://www.ncbi.nlm.nih.gov/sites/entrez?db=genome>. The page features a blue header with the NCBI logo and a search bar. Below the header is a large banner with a blue background showing chromosomes and the word "Genome". To the right of the banner, text reads: "This resource organizes information on genomes including sequences, maps, chromosomes, assemblies, and annotations." The main content area is divided into several sections: "Using Genome" (Help, Browse by Organism, Download / FTP, Submit a genome), "Custom resources" (Human Genome, Microbes, Organelles, Plants, Viruses), "Other Resources" (Assembly, BioProject, BioSample, Map Viewer, Protein Clusters), "Genome Tools" (BLAST the Human Genome, Genomic groups BLAST), "Genome Annotation and Analysis" (Eukaryotic Genome Annotation, Prokaryotic Genome Annotation), and "External Resources" (GOLD - Genomes Online Database, Ensembl Genome Browser).

<http://www.ncbi.nlm.nih.gov/sites/entrez?db=genome>

微生物基因體計劃大致之進行步驟

- DNA sequencing (DNA 定序)
- Genome assembly (基因體之組合)
- Gene annotation (基因註解)
- Databases (資料庫之建立)

DNA sequencing: the shotgun method

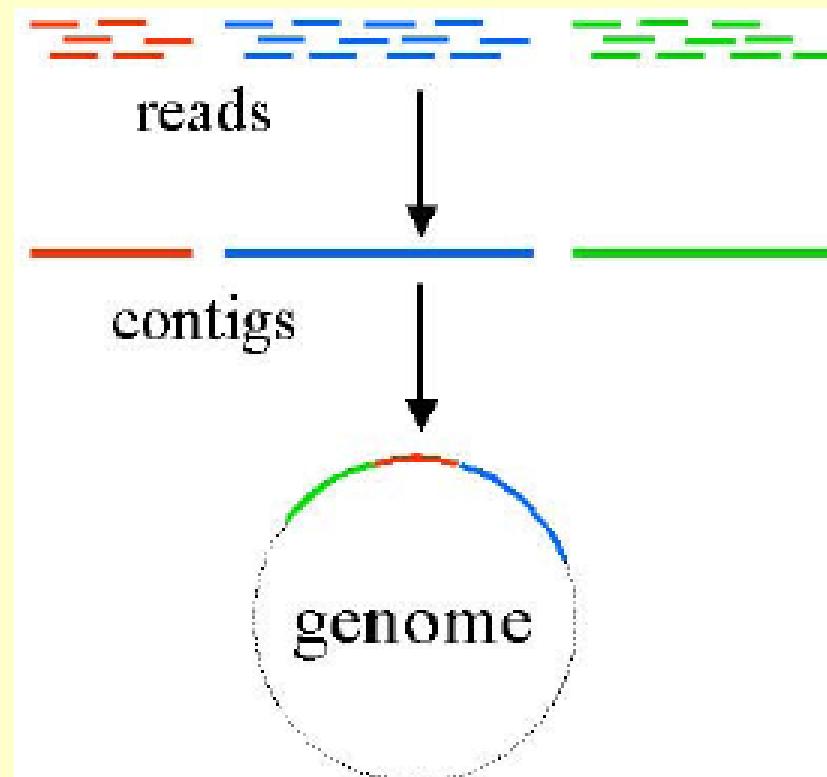


- Generate library representing whole genome
- Pick **random** clones for sequencing
- Assemble from overlap of clone sequences
- Rapid & simple

(<http://www.bioteach.ubc.ca/Bioinformatics/GenomeProjects/>)

Genome Assembly

- Sequencing reading from clones are overlapped – using sequence identities
 - to obtain large segments “Contigs”
- Resulting contigs are combined to assemble the whole genome
- By overlapping individual readings, a genome may be covered several times



(Bini, 2005)

Gene annotation

- Gene finding: determine putative open-reading-frames (ORFs)
- Translate ORFs
- Searching against databases of sequences

hit a protein of known function

hit a protein of unknown function

no match

- Assignments for molecular functions

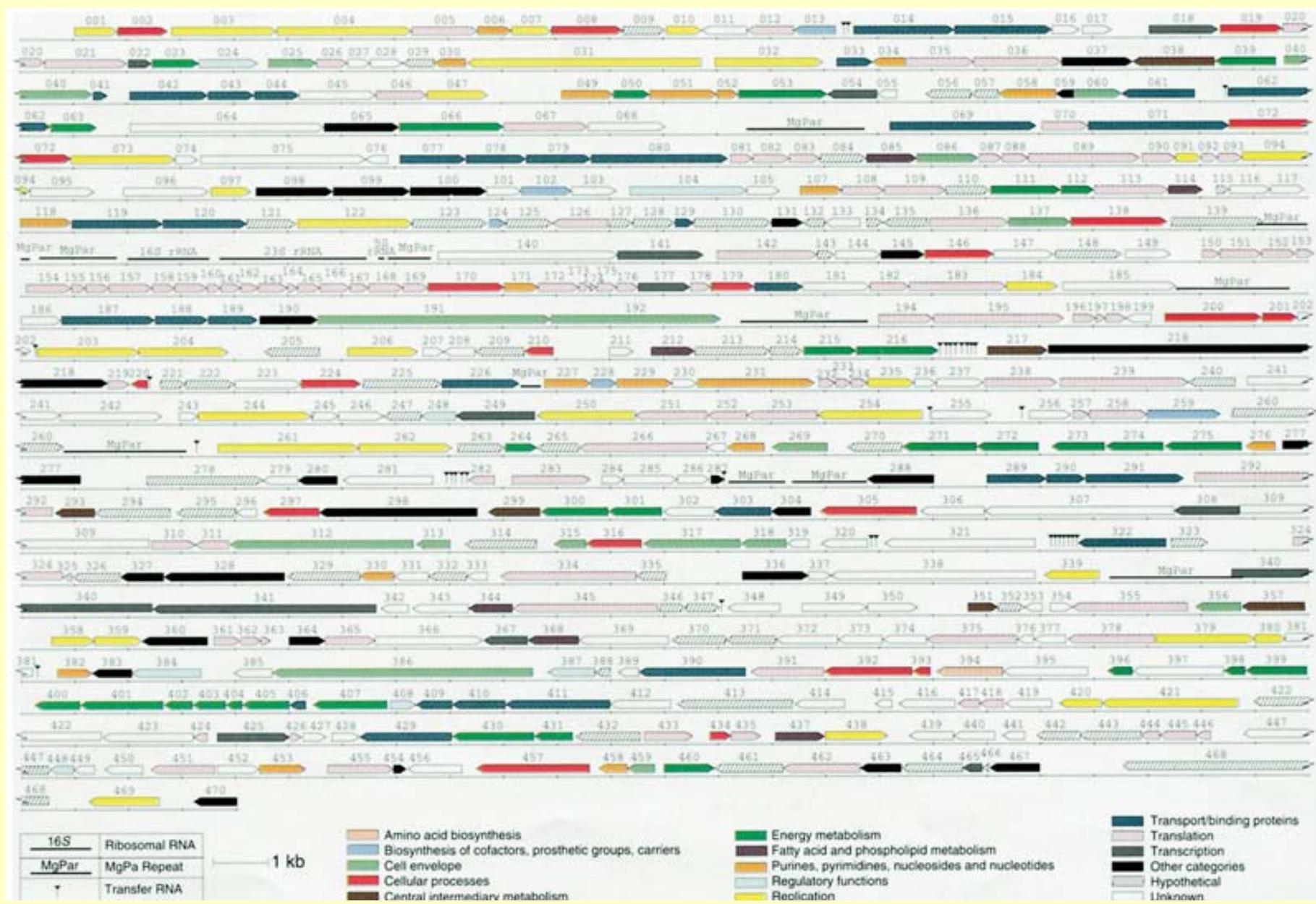
Functional assignments

Structural assignments

Genome Databases

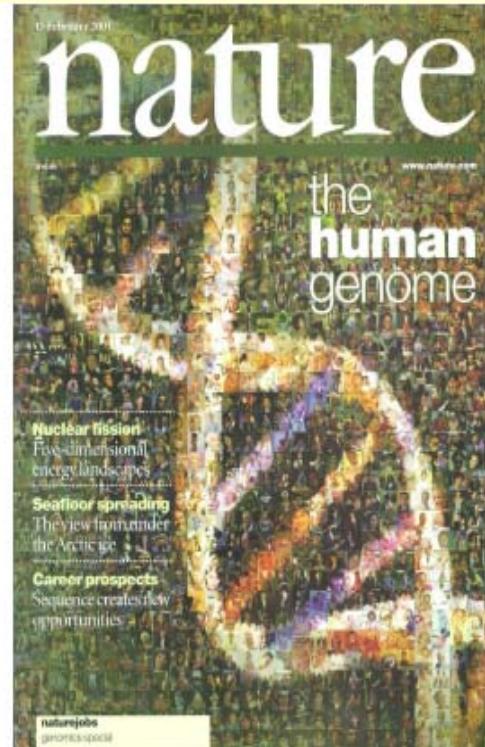
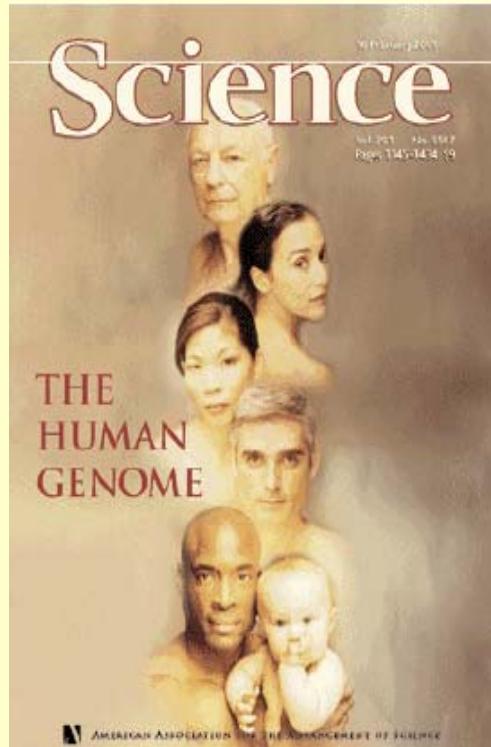
- The National Center for Biotechnology Information (NCBI)
[\(http://www.ncbi.nlm.nih.gov/genomes/\)](http://www.ncbi.nlm.nih.gov/genomes/)
- The J. Craig Venter Institute
[\(http://www.jcvi.org/\)](http://www.jcvi.org/)
- The DOE (Department of Energy) Joint Genome Institute (DGI)
[\(http://genome.jgi-psf.org/\)](http://genome.jgi-psf.org/)
- Individual genome sequencing projects
(e.g.) *Saccharomyces* Genome Database (SGD)

Mycoplasma genitalium 基因體之圖譜



Human Genome Project 人類基因體計劃

- 決定人類基因體內DNA之序列(約有18國參加)
- 人類基因體計劃正式開始於1990, 於2003完成
- 人類基因體中含有30億(3 billion)個鹼基對, 估計含3萬基因



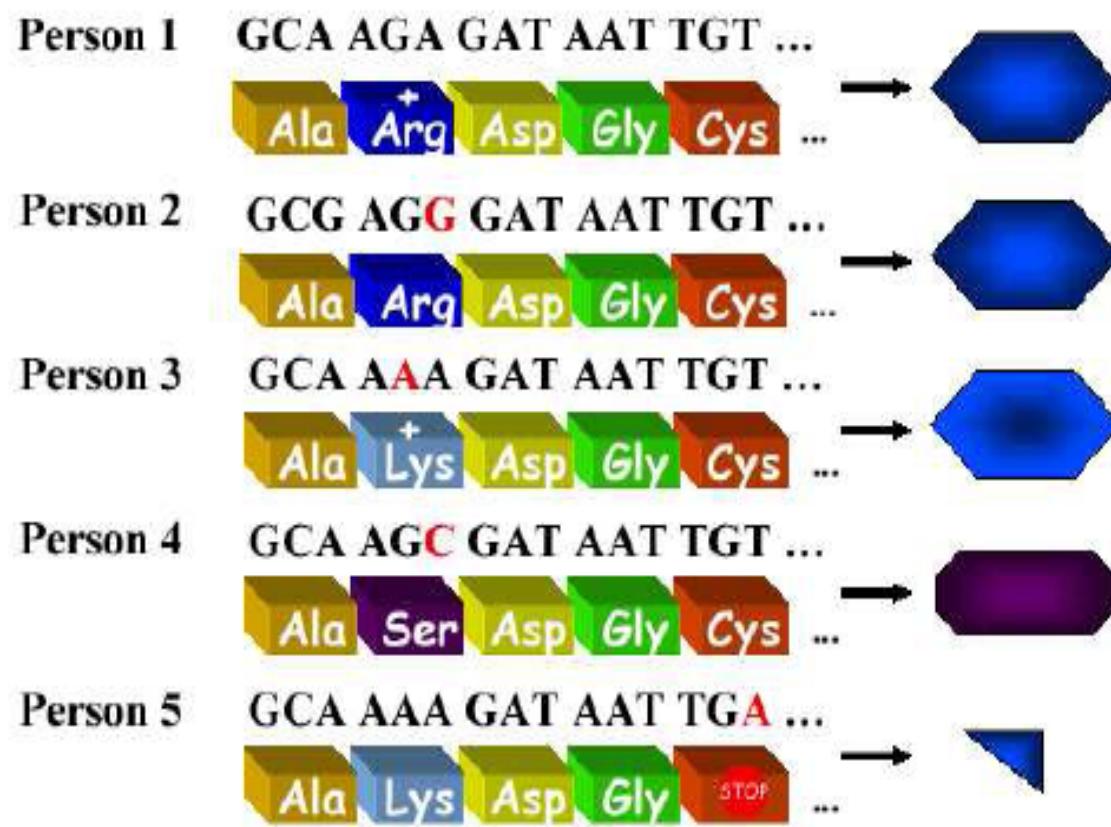
了解基因體序列之一些實際益處

- (1) 幫助我們認識不同個體DNA產生變異可能造成之後果：
因此有可能發展出新的檢驗及治療方法，並且可能於未來發展出防止數千種疾病之新方法

- (2) 除幫助了解人類之問題，知道其他非人類物種之DNA序列，
有助於了解其他物種之自然能力，有可能因此應用於醫療照護、農業、新能源產生及環境整治等

例如：由人類基因體序列找尋疾病的基因

Single-nucleotide polymorphism (SNP; 單一核酸多型性分析)

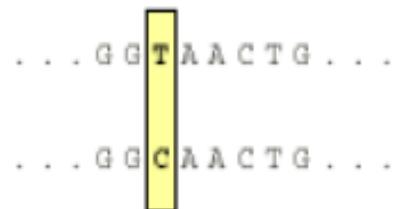


SNP可能與疾病有關，
也可應用於親源、族群
演化之研究

SNP & Pharmacogenomics (藥物基因體學)

What is an SNP?

Different people can have a different nucleotide or base at given location on a chromosome

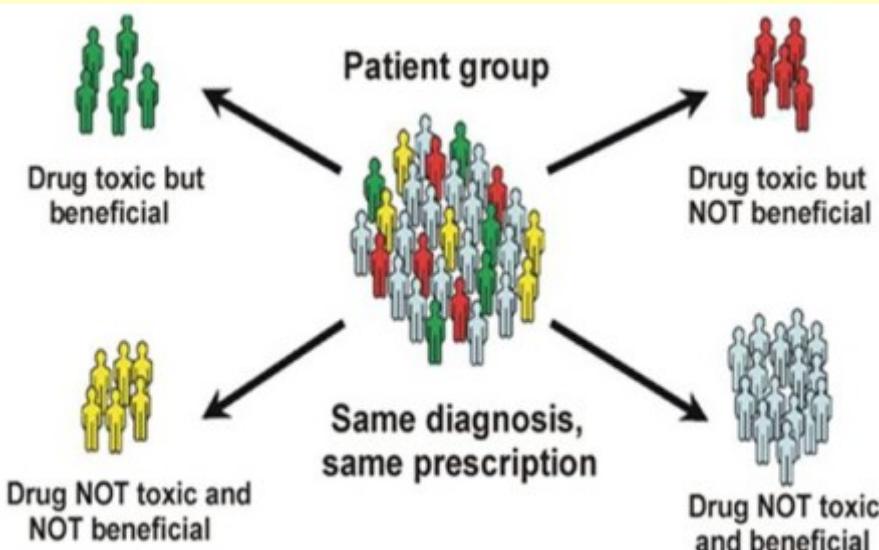


What is an SNP map?

Location of SNPs on human DNA

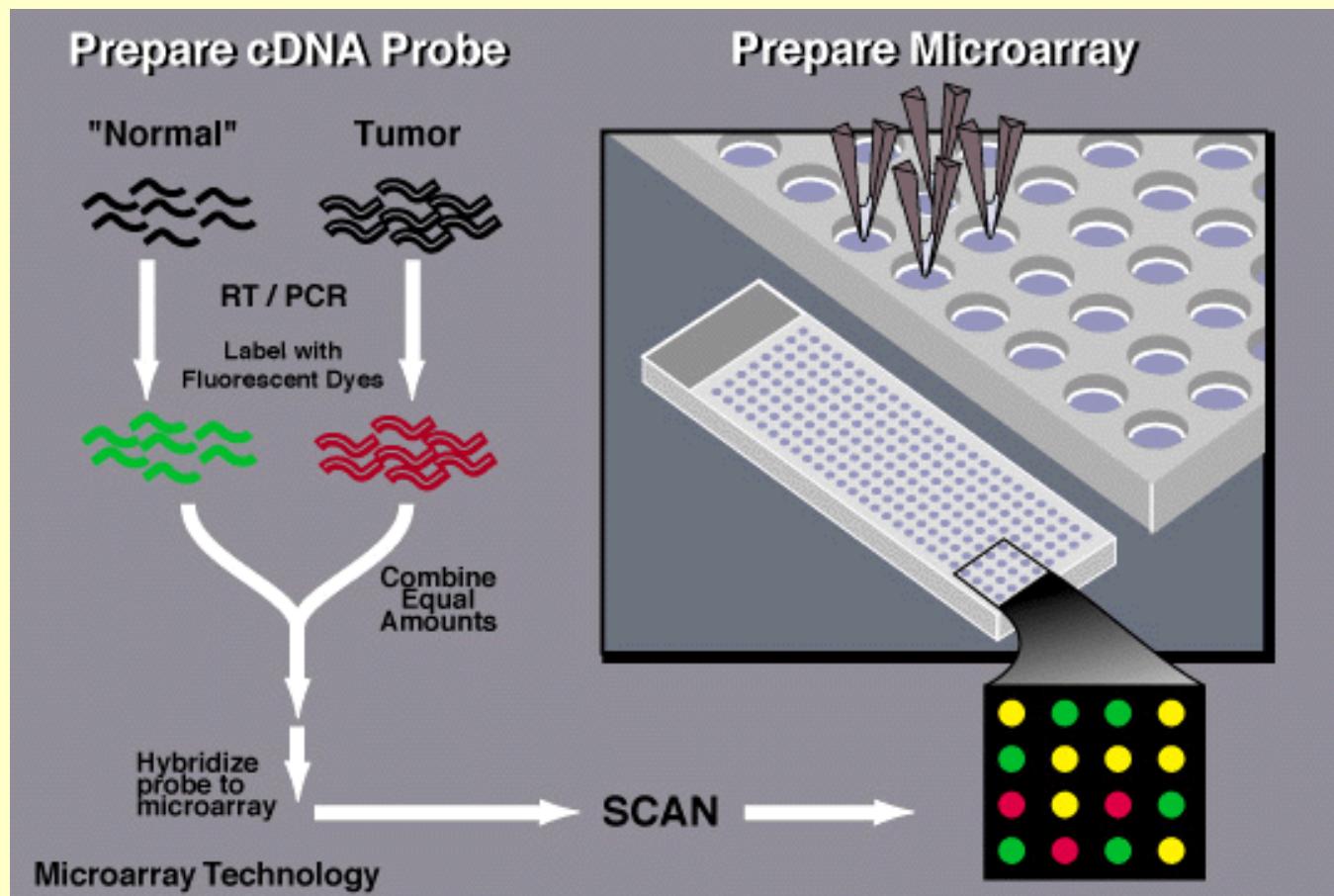


How can an SNP map be used to predict medicine response?



從基因體(遺傳性狀)的層次來了解相同藥物在不同病人中所產生的藥效或副作用上的差異性，藉此開發出有用的資訊讓醫師能「對人下藥」。除了可使有些患者因此避免嚴重甚至死亡之副作用，製藥公司也能藉此知識開發新藥或合併療法，治療對原有藥物沒有反應的病人

常用之功能基因體分析工具:DNA microarrays (基因晶片)

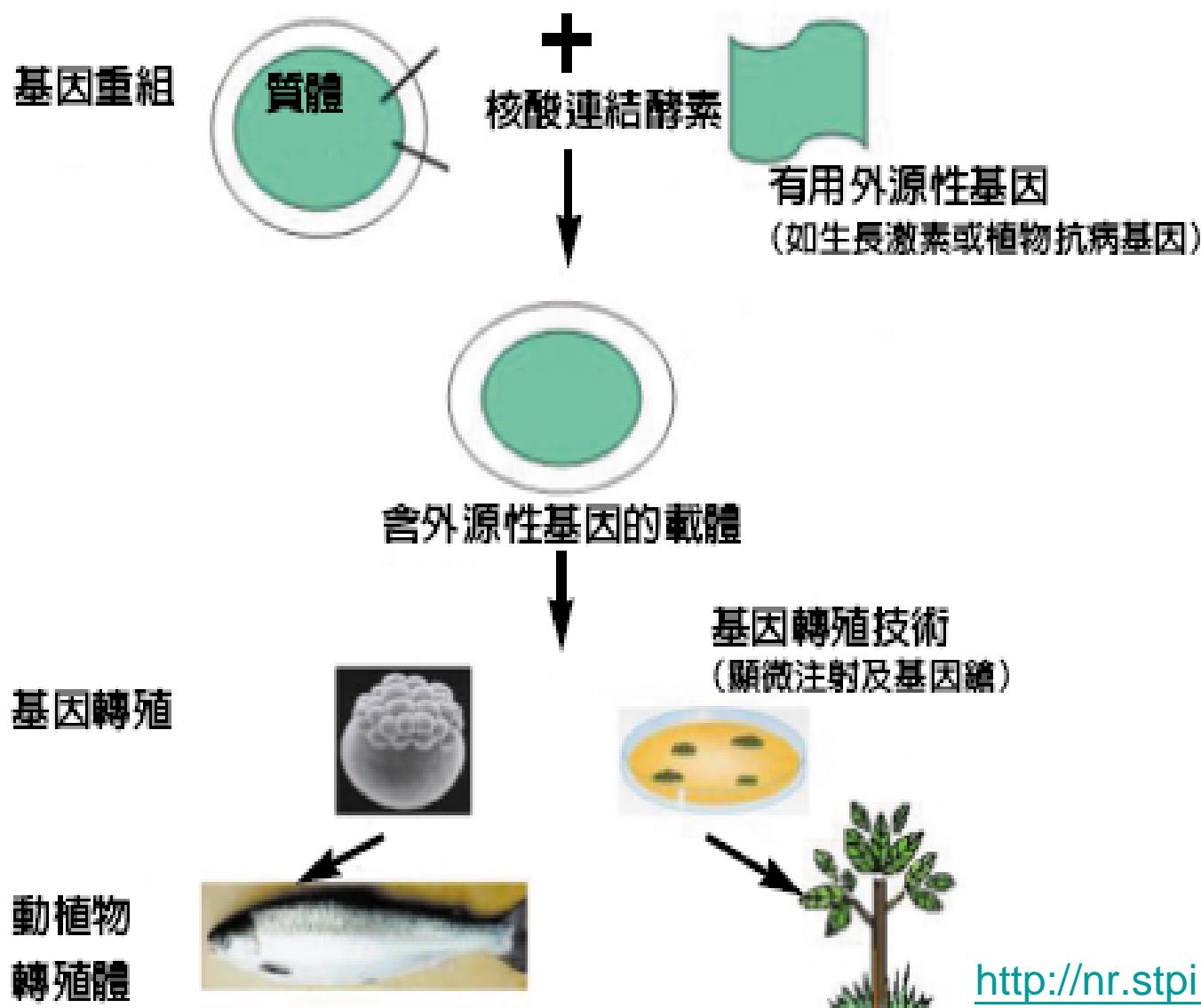


(<http://www.accessexcellence.org/RC/VL/GG/microArray.html>)

<http://www.bio.davidson.edu/Courses/genomics/chip/chip.html>

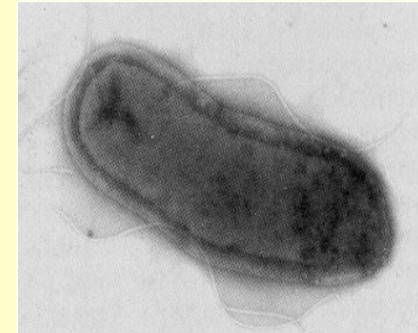
分子生物/重組DNA技術之應用

限制性核酸酵素切開質體



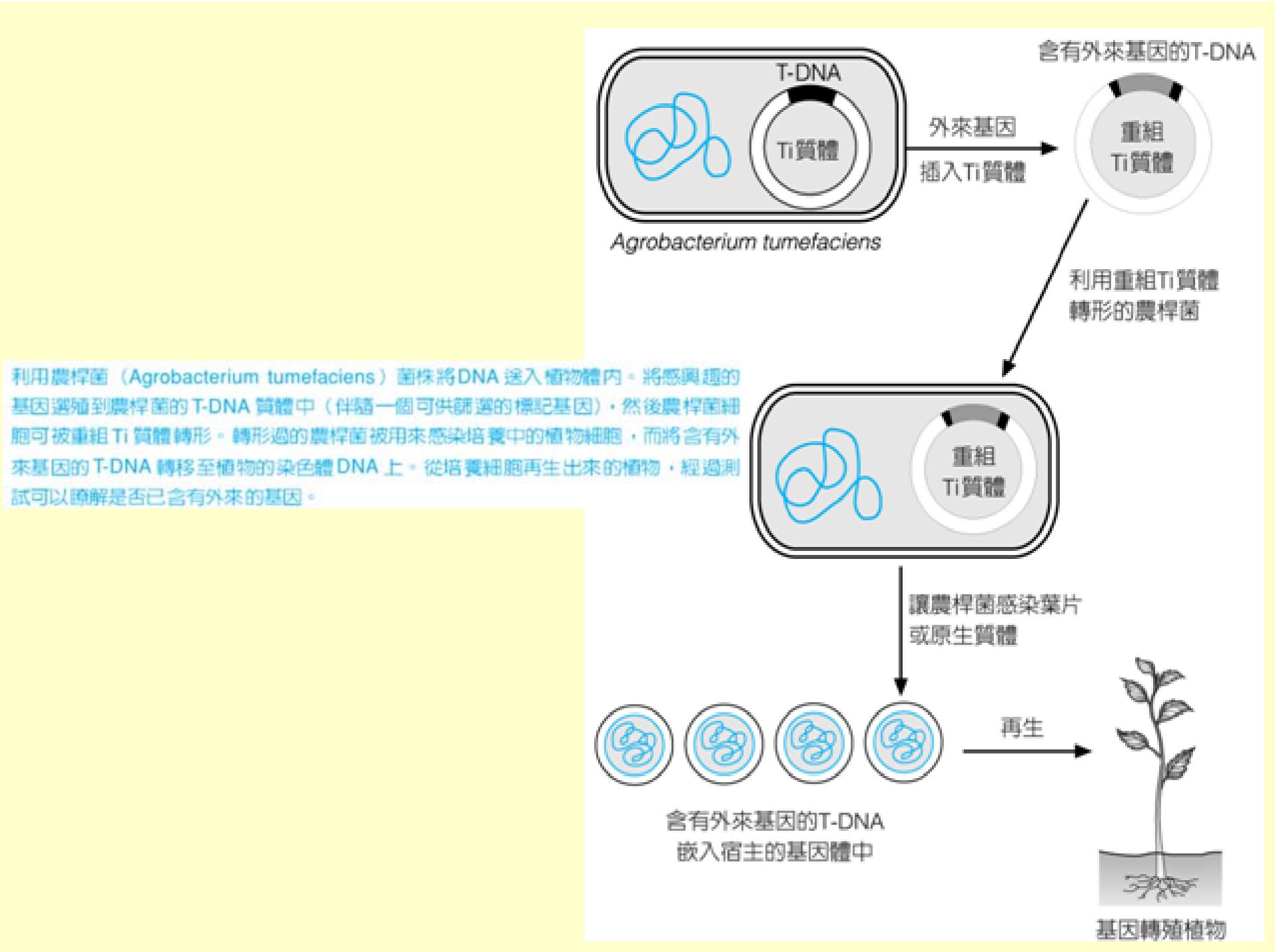
<http://nr.stpi.org.tw/ejournal/Nscm/9301/9301-03.pdf>

另一種應用於植物之基因轉植技術：
利用農桿菌(*Agrobacterium tumefaciens*)



- 侵入許多植物之根、莖等部位，造成植物腫瘤(Tumor)
- 藉由tumor-inducing (Ti) plasmid





農業上之應用

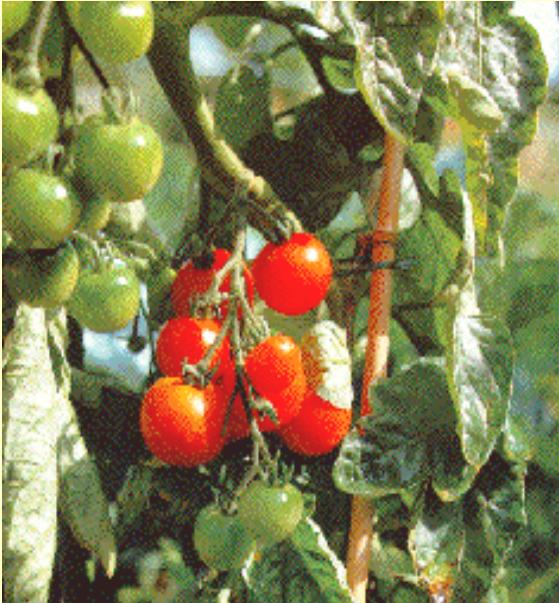
人類改良作物或飼養動物品系，求望能有更富營養、更高產量、更好肉質或更易栽種和快速收成的新品種。

過去，新品種的來源主要靠隨機取得。之後，育種交配的技術發展，可以幫助品種改良。然而此種技術複雜且不易握，遺傳性狀隨機出現於子代中，需經由篩選，才能獲得新系，費時取且困難。

基因工程(基因轉殖技術transgene technology)，對於品種改良提供一個新的契機，可以直接經由基因轉殖來改變作物或飼養動物的基因型，例如：

- (1) 抗蟲、抗病轉殖玉米、大豆及棉花
- (2) 後熟番茄（破壞番茄產生乙烯基因）

Some examples of transgenic plants

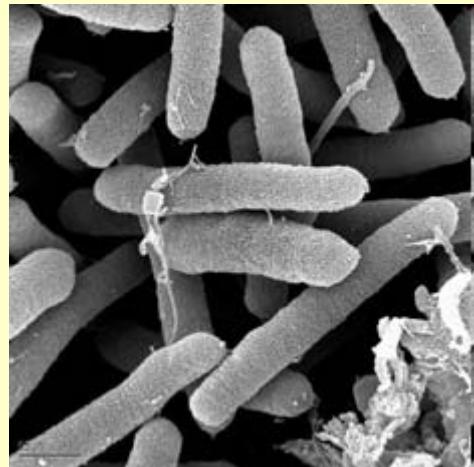


Some examples of transgenic animals



抗凍菌(ice-minus bacteria)於農業上之應用

- *Pseudomonas syringae* 為植物表面常見之Gram (-)桿狀細菌，其細胞壁外有一個冰核蛋白質(Ina protein, "Ice nucleation-active" protein)，為冰晶形成之中心



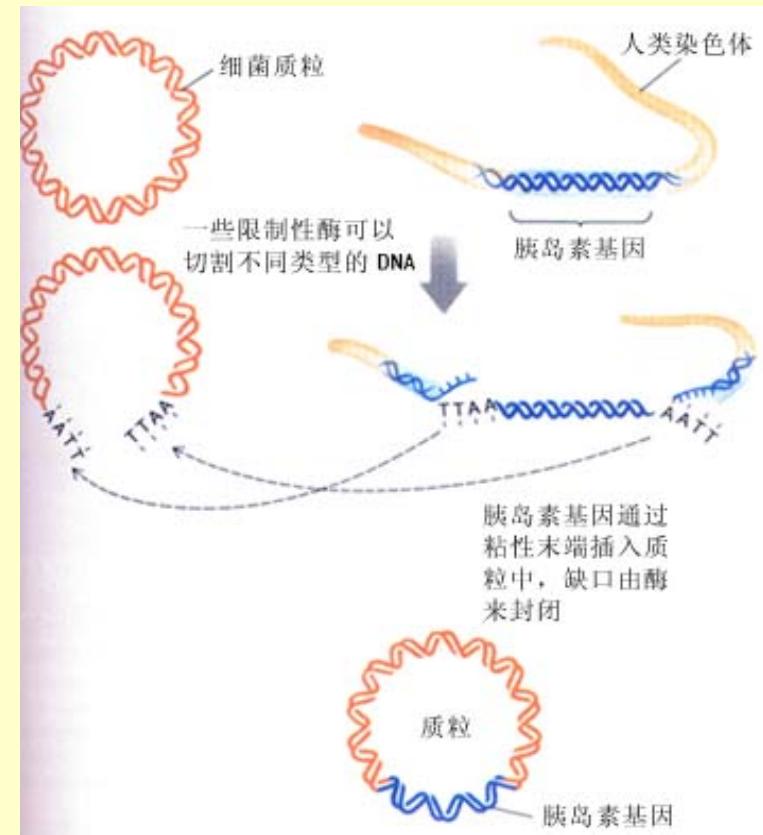
- 抗凍菌為一株*P. syringae*經遺傳工程方法去除可轉譯出Ina蛋白質之基因，因此無法生成冰核蛋白質，可降低植物表面結冰之溫度(-6 °C ~ -8 °C)，進而保護植物

醫藥上之應用

基因工程提供一個重要方法，可大量製備人類藥用蛋白質。例如：

重組胰島素(recombinant insulin)

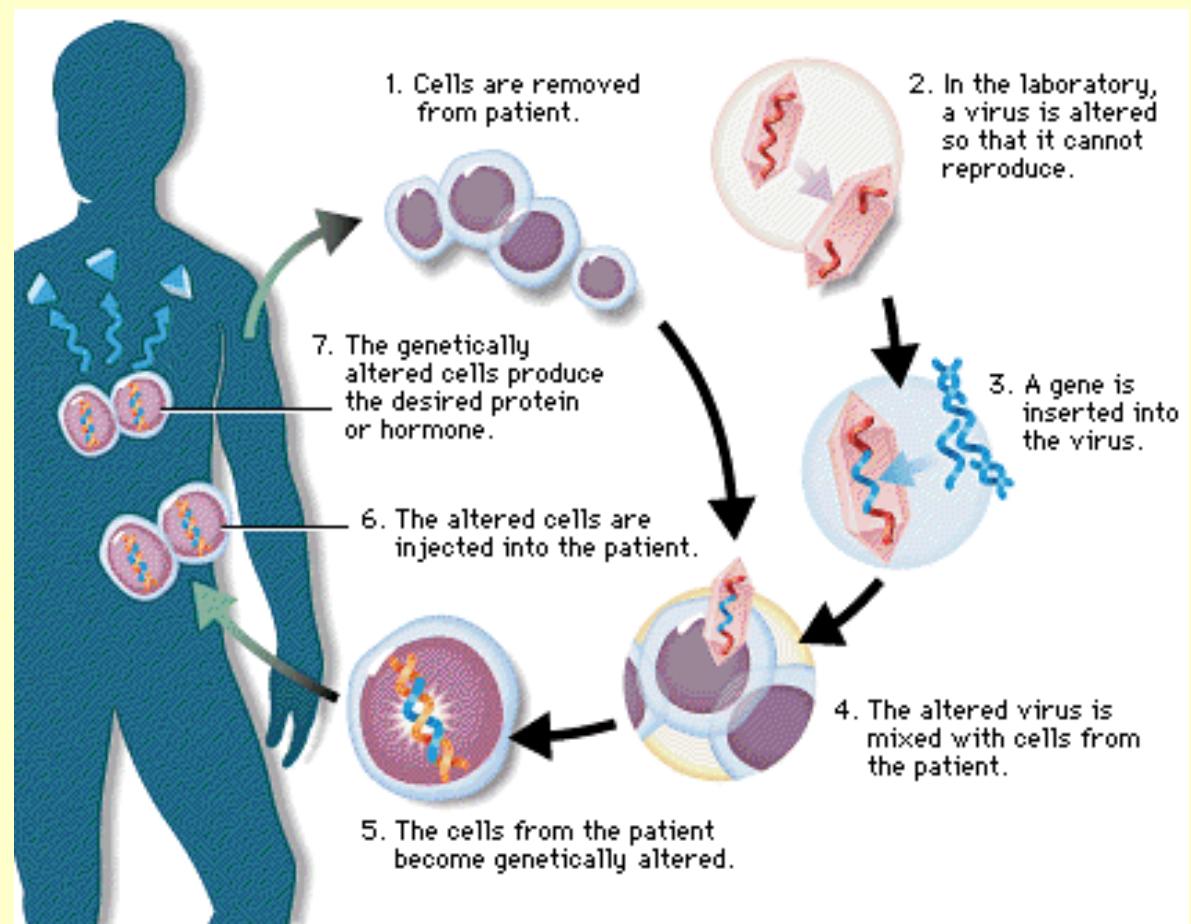
人類生長激素(human growth hormones)



許多人類疾病，肇因於身體內某種蛋白質功能缺乏或不正常，可藉由提供正常基因，使其恢復正常(如:基因療法)。

Gene therapy (基因療法)

- (1) 將正常基因送入基因機能障礙的個體，導入正常傳譯/轉錄基因的細胞或是胚胎細胞，使基因能正常作用生成正常之蛋白質或荷爾蒙。
- (2) 需發展輸送基因之最佳且安全的系統，並確定證明轉移的基因在宿主細胞能正常複製及作用。
- (3) 基因療法(包括基因操作與選擇、胚胎組織研究以及人體試驗等)，都需要考慮道德、法律等的問題。



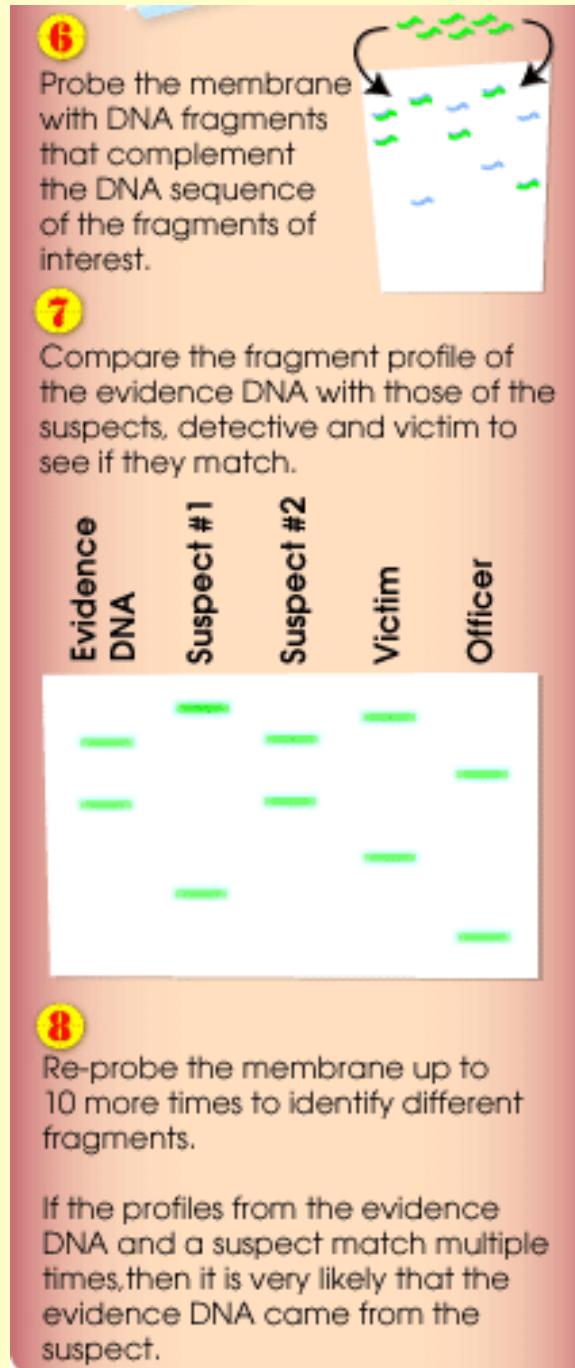
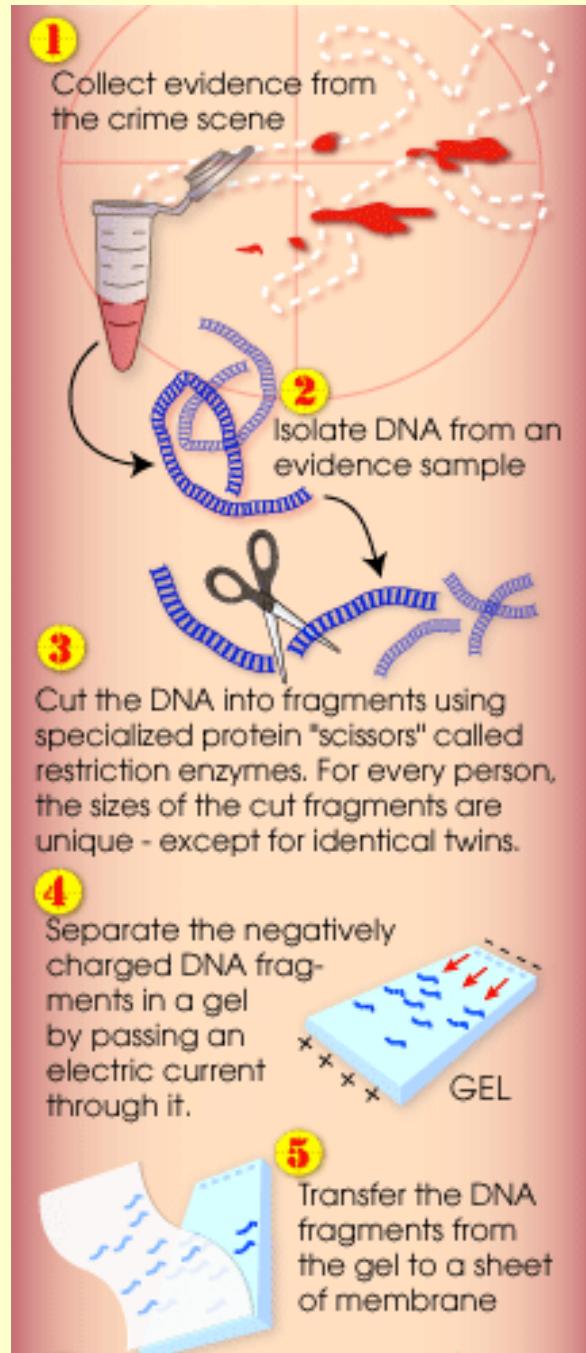
[http://www.genetherapyreview.com/
education/disease-targets/genetic/
genetic-disease-gene-therapy](http://www.genetherapyreview.com/education/disease-targets/genetic/genetic-disease-gene-therapy)

鑑識科學(Forensic science)及考古學(Archaeology)之應用

- 分子生物技術可用於鑑定犯罪現場所採集到之毛髮或血漬(含有DNA)等
- 比較現有人類與考古挖掘之人類遺骸，可幫助我們了解現代人類之演化啟源、遷徙路境等。這方面研究領域，又稱Archaeogenetics
- 例如: DNA profiling (或稱DNA typing 或Genetic fingerprinting) 之技術，常被用於此類研究



How is forensic DNA analyzed?



<http://learn.genetics.utah.edu/content/labs/gel/forensics/>

Genetic fingerprinting: an example

