

DNA 像電腦嗎？

李家同
暨南國際大學資訊工程系
rctlee@ncnu.edu.tw

我們人是一種生物，其他的動物和植物當然也是生物，生物有一個共同的特性，就是能夠繁衍下一代。繁衍下一代已經夠奇妙，最奇妙的是如何將上一代的資訊傳給下一代。舉例來說，大象必須生下的仍是一頭象，狗不能生出貓來。在曠裡的黃色野花，種子掉落在泥土中第二年長出來的仍是那種黃色野花。顏色不能變，花的樣子也不能變。

生物如何將資訊傳給下一代呢？我們是透過 DNA (deoxyribonucleic) 傳過去的，每一生物的細胞裡都有 DNA，以人為例，人一共有 23 對染色體，DNA 就在染色體裡面。

DNA 由四種鹼基組成，這四種鹼基就是 adenine (A) , guanine (G) , cytosine (C) 和 thymine (T)。DNA 可以說是這四個鹼的排列，舉例來說，以下的序列就可能是 DNA 的一段

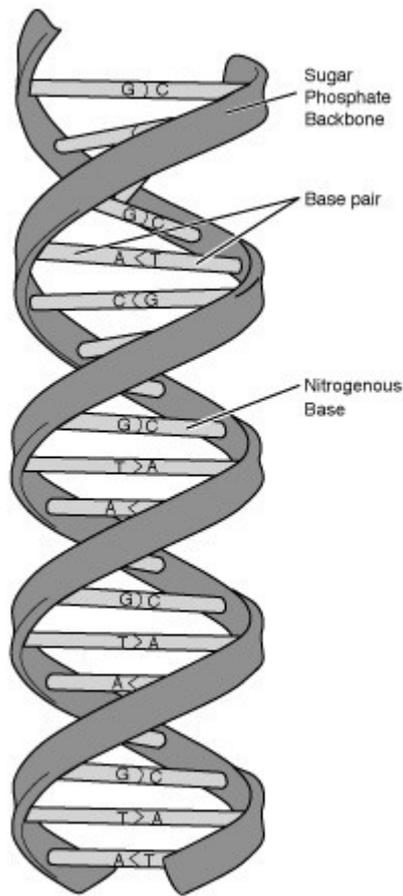
TTCATTTATCGT

而 DNA 又是一個雙螺旋 (double helix) 的東西，也就是說，DNA 有兩個序列，但這兩個序列並不互相獨立，而是互相對應的，所謂對應乃是說 A 一定對應到 T，而 G 一定對應到 C，以我們所舉的以上序列為例，他的稱應序列就是

AAGTAAATAGCA

下圖是所謂雙螺旋的示意圖。

特殊功能說細胞的奇功一樣，所擁 DNA 完全的是回原個要製種蛋以血



的功比血液就要帶氧能，是細胞不能他們的有卻是一樣這又怎麼呢？每一胞都造一白質液細

胞為例，他們就要製造一種能夠帶氧的 hameglobin 蛋白質。

細胞如何知道蛋白質怎麼造成的？答案很簡單，細胞會到 DNA 裡去找，每一個蛋白質對應 DNA 裡的一段序列，而這一段序列就是所謂基因(gene)，我們可以說基因的功能就是在告訴細胞如何製造蛋白質。

首先我要說明的是，蛋白質由二十種氨基酸(amino acid)所組成，以下是這二種氨基酸的全部名單：

Amino acid	Three-letter code	One-letter code
Alanine	ALA	A
Arginine	ARG	R
Aspartic Acid	ASP	D
Asparagine	ASN	N

Cysteine	CYS	C
Glutamic Acid	GLU	E
Glutamine	GLN	Q
Glycine	GLY	G
Histidine	HIS	H
Isoleucine	ILE	I
Leucine	LEU	L
Lysine	LYS	K
Methionine	MET	M
Phenylalanine	PHE	F
Proline	PRO	P
Serine	SER	S
Threonine	THR	T
Tryptophan	TRP	W
Tyrosine	TYR	Y
Valine	VAL	V

DNA 由四種鹼基所構成，科學家發現每一個氨基酸都對應一個由三個鹼基所構成的碼。舉例來說，TTT 就代表 Phenylalanine，TCT 代表 Serine，以下的表是這些鹼基的氨基酸的對應表，表內每三個碼就叫做一個 codon。

		Second Position of Codon					
		T	C	A	G		
F i r s t P	T	TTT Phe [F]	TCT Ser [S]	TAT Tyr [Y]	TGT Cys [C]	T	T h i r d P
		TTC Phe [F]	TCC Ser [S]	TAC Tyr [Y]	TGC Cys [C]	C	
		TTA Leu [L]	TCA Ser [S]	TAA Ter [end]	TGA Ter [end]	A	
		TTG Leu [L]	TCG Ser [S]	TAG Ter [end]	TGG Trp [W]	G	
	C	CTT Leu [L]	CCT Pro [P]	CAT His [H]	CGT Arg [R]	T	
		CTC Leu [L]	CCC Pro [P]	CAC His [H]	CGC Arg [R]	C	
		CTA Leu [L]	CCA Pro [P]	CAA Gln [Q]	CGA Arg [R]	A	
		CTG Leu [L]	CCG Pro [P]	CAG Gln [Q]	CGG Arg [R]	G	

o s i t	A	ATT Ile [I]	ACT Thr [T]	AAT Asn [N]	AGT Ser [S]	T	o s i t
		ATC Ile [I]	ACC Thr [T]	AAC Asn [N]	AGC Ser [S]	C	
		ATA Ile [I]	ACA Thr [T]	AAA Lys [K]	AGA Arg [R]	A	
		ATG Met [M]	ACG Thr [T]	AAG Lys [K]	AGG Arg [R]	G	
	G	GTT Val [V]	GCT Ala [A]	GAT Asp [D]	GGT Gly [G]	T	
		GTC Val [V]	GCC Ala [A]	GAC Asp [D]	GGC Gly [G]	C	
		GTA Val [V]	GCA Ala [A]	GAA Glu [E]	GGA Gly [G]	A	
		GTG Val [V]	GCG Ala [A]	GAG Glu [E]	GGG Gly [G]	G	

我們一共有四種鹼基，用三個鹼基，我們可以有 $4 \times 4 \times 4 = 64$ 種組合，也就是說，我們可以如此表示 64 種氨基酸。但我們只有 20 種氨基酸，所以我們的編碼是多對一 (Many to One) 的編碼，舉例來說，TTA, TTG, CTT 等等都表示 Leucine。

基因就像一個程式，這個程式每三個碼代表一個指令，這個指令使細胞裡的 RNA 去拿一個氨基酸，舉個例子來說，以下的一段就代表細胞應該去拿三個氨基酸來製造某一種蛋白質：

A T T	A G T	C G T
↑	↑	↑
Isolencine	Serine	Arginine

RNA 是細胞裡的一種東西，我們不在此介紹 RNA，我們要回答大家一定想問的問題：

(1) DNA 的長度高達 30 億個鹼基，細胞如何知道與它有關的基因從哪裡開始的？

(2) 細胞如何知道基因在何結束的？

對於第一個問題，答案是每個一基因都有一個 promotor，每一個細胞都有能力認識自己的 promotor，看見 promotor 就可以知道前面快有基因了。

基因必定開始於 ATG (對應 methionine)

對於第二個問題，請注意在以上表內有三個 codon 是不對應於任何氨基酸的，他們是 TAA，TAG 和 TGA，他們都代表 end of program，看到任何一個這種 codon，就代表基因已經結束了。

編碼學是最近所發展的學問，可是大自然在幾億年前就已發展出來了，End of Program 這一個碼更是最近才發展出來的，但是大自然早就有這種觀念了。

我們可以說我們所看到的大自然有如電腦的硬體，而 DNA 有如電腦的軟體，電腦的軟體我們看不到的，可是它卻是電腦裡最重要的

一部份，DNA 也是我們看不到的，顯然它也是大自然最重要的一部份。

學過電腦的人都同意一種說法：如果軟體不是我們自己寫的，我們就不要輕易去改它，以免造成後遺症。DNA 不是我們發展出來的，我們能夠輕率地去改變它嗎？