



基因密碼與蛋白質合成

王佳文

前言

即將於2007年7月在新加坡舉行的第18屆世界家庭醫師組織世界會議(18th Wonca World Conference)主題為Genomics and Family Medicine 距會議舉行日期至今僅短短一年多，各位會員在基因體學上應多所著墨，積極參與大會。

二十一世紀是生命科學的世紀，醫學進步很快，分子生物學的進步更是日新月異，譬如Ezetimibe最初以為其降膽固醇的作用，是促進HDL receptor SR-B1或小腸上皮膽固醇的輸送蛋白ABCG5/G8及抑制Caveolin-1等所致。最終在2004年確定機制為抑制小腸絨毛上刷狀邊緣的膽固醇吸收輸送蛋白NPC1L1。

最近喧騰一時的羅倫佐的油，常不知其為何物。其實引起羅倫佐致病的是腎上腺白質退化症(adrenoleukodystrophy, ALD)，依2002年第四版Nelson Essentials of Pediatrics所載，認為其是因very long chain fatty acid CoA synthetase 的缺陷所引起，但在2004年由 Darnell等人所著第

五版的Molecular Cell Biology 則認為是在過氧化物酶體膜(peroxisomal membrane)上輸入very long chain fatty acid 的輸送蛋白(ABC transporter protein, ABCD1)缺陷導致。最近該基因也已選殖確認(ABCD1, ATP-binding cassette, subfamily D, Member 1 [*Homo Sapiens*], Chromosome Xq28, Gene ID:215.)。

另外發現mRNA轉譯前可在mRNA 3'端非轉譯區(3'untranslated region)用MicroRNA(miRNA)與其雜交加以抑制或用RNA interference(RNAi)與其完全互補配對加以切離，以達壓抑轉譯的目的。這是近幾年才發現，前一版的書中並未見，所以我們除了多讀期刊外，也必需讀最新版的書。以上例子是基因突變與轉譯調控導致產生異常蛋白質，致影響正常生理功能。

基因組(GENOME)

人體的基因組(genome)含有三十億(3 billion, 3×10^9)個核苷酸對，其中屬於獨特序列(unique copy DNA，意即整個基因體上只有一套copy)的部分約佔45%，另外55%則是反覆序列(repetitive

王佳文內兒科診所 負責醫師

關鍵字：Genome, gene, DNA, RNA



sequence)，後者含10-15% 銜接重複序列(tandem repeat sequence)，其DNA亦稱衛星DNA(satellite DNA)，法醫即用其重複序列多樣性的STR(short tandem repeat)來鑑定親子關係，另含45%散布重複序列，它是一種轉位元(transposable element)，亦稱跳躍單元(mobile element)。Francis Crick更指它為“selfish DNA”能藉著DNA重組的機轉，轉至他處。著名的佔基因組10-15%的Alu家族即在其中。人類的體細胞有46條染色體，由不同的DNA分子組成22對體染色體和兩條性染色體(X,Y或X,X)及約1%的粒線體DNA，而後者幾乎99.9%都來自母親卵子，與父親毫無關聯，且複製時是以D LOOP而非如細菌之 θ 型態複製。每一個人類細胞中皆帶有許多的粒線體，每個粒線體基因組共含16,569個鹼基對，為一條雙鏈環狀的DNA分子，含有37個基因，其中13個是編碼蛋白質基因，兩條rRNAs和22種tRNAs分子，粒線體大多數(67中之54)參與電子傳遞鏈的蛋白質都是核基因所編碼的粒線體蛋白質。

核酸的結構

DNA及RNA屬於核苷酸(nucleic acids)，由鹼基、五碳糖與磷酸根組成。RNA的五碳糖是核糖(ribose)，DNA的是去氧核糖(deoxyribose)。去氧核糖在2'位置比核糖少一個氧原子。核苷酸的鹼基共有5種：腺嘌呤(adenine, A)、胸腺嘧啶(thymine, T)、鳥嘌呤(guanine, G)、胞嘧啶(cytosine, C)及尿嘧啶(uracil, U)。

G)、胞嘧啶(cytosine, C)及尿嘧啶(uracil, U)。

DNA由A.T.C.G四種不同鹼基組成；RNA由A.U.C.G組成。核苷酸藉著磷酸二酯鍵(phosphodiesterase bond)將一個核苷酸的5'端與另一個的3'端結合在一起。真核細胞DNA的構型有A.B.Z.H四種，前三種皆以兩股互相繞成雙螺旋(double helix)，一股鹼基與另一股鹼基利用氫鍵以相反平行的方向(antiparallel)互相配對，G與C以3氫鍵結合而A與T以2氫鍵結合。

1953年，Watson及Crick提出在生理狀況的低鹽高水下存在的最主要DNA的B型結構，在此結構裡，兩股以較穩定的右手螺旋(right-handed helix)纏繞著，每前進3.4 nm繞一圈，每圈內含10個鹼基對，因此兩個鹼基對的距離為0.34 nm。

依Chargaff等人分析真核生物DNA的鹼基組成時發現：1. A=T, G=C, 2. A+G=T+C, 3. A+T \neq G+C, 此即所謂的Chargaff's rule。

1990美國與各國合作開始成立人類基因體研究計劃(Human Genomic Project, HPG)以來，經不斷的努力，於2000年6月宣佈完成人類遺傳基因體的草圖(由美國當時能原部長Francis Collins及私人Celera Genomics公司的Craig Venter聯合公佈)，估計其中蛋白質的密碼只佔基因組的1.1-1.5%，約只有30,000至40,000個蛋白質表現基因。蛋白質的合成需要三種RNAs: mRNA, rRNA, tRNA。mRNA的核



苷酸序列隱含所要合成的蛋白質密碼。rRNA 與蛋白質結成核糖體，執行蛋白質的合成。tRNA負責將mRNA的序列轉譯成蛋白質的胺基酸序列。

基因的要素

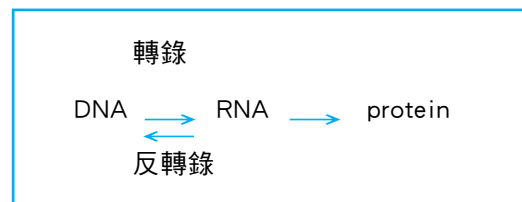
一個基因(gene)包含合成一個具有功能的RNA分子或胜肽鏈所需的全部序列。它可分成轉錄區和轉錄調控區。一段含有100個密碼子以上的DNA，始於start codon終於stop codon稱為開放轉錄區(open reading frame, ORF)。真核生物的轉錄區包括外顯子(exon)和內含子(intron)。外顯子具有所要合成的RNA或胜肽鏈的密碼，亦稱編碼區(coding region)。內含子在轉錄後的RNA處理會被切除。轉錄調控區是轉錄因子與DNA作用的區域，後者稱辨認序列(recognition sequence)。調控區依位置之不同，又分成啟動子(promoter)，增強子(enhancer)和靜默子(silencer)等。一般將轉錄起點(Inr, initiator)定為+1，蛋白質基因最主要的啟動子是TATA box 和Inr，他們亦合稱為核心啟動子(core promoter)，TATA box 位於轉錄起點上游25 bp至30 bp，與TATA box作用的蛋白質叫 TBP(TATA box binding protein)，該部位是RNA polymerase II 及其他轉錄蛋白質結合的位點，與轉錄起始的準確定位有關，位於Inr上游-70至-80 bp含有CAAT序列，約在-80至-110 bp有GC box，合稱為啟動相鄰元，它們是基因有效轉錄所必需的序列，

主要控制轉錄起始頻率，不參與起始位點的確定。增強子位於轉錄起點遠處，可能在其上游或下游數千鹼基或數萬鹼基處，不管從5'→3'或3'→5'均能提高啟動子的活性。譬如steroid hormone (glucocorticoid等)通過細胞膜與steroid receptor的配體結合區(ligand binding domain, LBD)結合，受體即進入細胞核，以其DNA結合區(DNA-binding domain, DBD)和染色體上類似enhancer的激素反應元(hormone response element, HRE)相結合，而活化基因的表現。

中心信條與其修正

1950年代 Francis Crick 對細胞內整個訊息傳遞的過程提出所謂的“中心信條(Central Dogma)”：即以DNA為模板轉錄(transcription)做出RNA，再以RNA為模板轉譯(translation)做出蛋白質。以上步驟皆為單方向，無法以蛋白質做出RNA，也無法以RNA為模板做出DNA。

1970年，Temin 提出reverse transcriptase的存在，它能以RNA為模板合成DNA，再以新合成的DNA為模板合成mRNA，然後轉譯成蛋白質，故中心信條應加以修正(圖一)。



圖一：中心信條與其修正

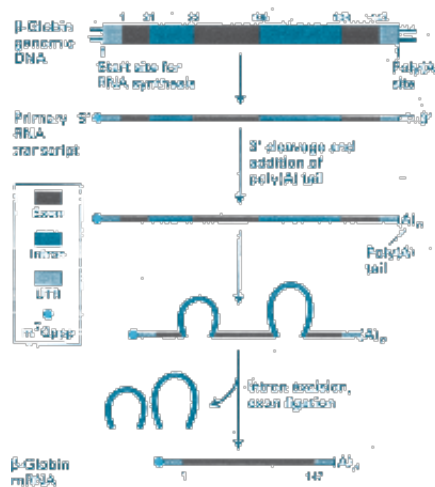


蛋白質合成與基因密碼

真核細胞經RNA polymerase II及轉錄因子做出的pre-mRNA需經四個步驟處理：

1. 5'加帽(5'capping)-加7-methylguanylate (圖二)。
2. 切除部分3'端(cleavage)。
3. 加多腺苷酸(polyadenylation)-poly-A尾巴。
4. 剪接(splicing)(圖三)。5'capping與3'tailing可協助mRNA由細胞核運至細胞質。此外，有些還經過編輯(editing)。剪接係將內含子剪掉，並連接外顯子，藉由小核RNA(sn RNA, U1, U2, U4, U5, U6)與蛋白質的協助，在內含子5'端GU處及3'AG處(尚需3'端上游20-25 bp處支點A的協助)剪接成簡單(simple transcript)或複雜不同剪接方式的mRNAs(complex transcript 如alternative splicing)。

生物的蛋白質合成都是由DNA上連續排列的遺傳密碼(genetic code)所決定，DNA的遺傳信息轉錄到mRNA上，通過翻譯讀寫mRNA上的遺傳密碼，將DNA上遺傳信息轉變為蛋白質的胺基酸排列順序。mRNA分子中每三個相鄰鹼基代表一種胺基酸或終止信號，這3個相鄰鹼基稱為密碼子(codon)。mRNA上有4種鹼基，按排列可形成64(4³)種密碼子，其中3個為終止密碼子(stop codons)，又稱“無意義密碼(nonsense codons)”，即UAA(ochre)、UAG(amber)及



圖二：β-globin基因形成 mRNA的程序(5' m7Gppp Capping, poly-A tailing及splicing) (源自Molecular Cell Biology 5th ed)



圖三：真核 mRNA 5' cap 的結構(7 methylated 5' -5' triphosphate linkage)(源自Molecular Cell Biology 5th ed.)

UGA(opal)。AUG為起始密碼(其在胜肽鏈尚可做為methionine的密碼)，其它60個代表相對應胺基酸。在生物體中，組成蛋白質的胺基酸共有20種，但決定胺基酸



的密碼有61個，顯然每一種胺基酸必有一種或一種以上遺傳密碼與之對應，此稱為遺傳密碼的退化性(*degeneracy of genetic code*)，例如GUU、GUC、GUA、GUG皆決定valine，若突變使密碼GUU改變成GUC，則仍然可轉譯為valine，不影響蛋白質的活性(稱為靜默突變，*silent mutation*)，此意味著tRNA的反密碼子(*anticodon*)的第一個鹼基(從5'端算起)與mRNA的密碼子的第三個鹼基並非一對一的關係，此即所謂的搖擺位子(*wobble position*)，例如arginine的兩個codon AGA及AGG皆與同一個第一位子含U的反密碼子(UCU)配對，glycine的GGU, GGC, GGA也與同一個反密碼子(CCI)配對(I, *inosine nucleotide*)。

若mRNA上遺傳密碼發生改變，使得所對應的胺基酸改變成另一種時，此稱為誤意突變(*missense mutation*)，譬如sickle anemia Hb S其β chain的第6個密碼子GAG(*glutamic acid*)突變成GUG(*valine*)，因而影響紅血球的形狀及生理功能。所有胜肽鏈的合成皆從甲硫胺酸(*methionine, Met*)開始，細菌在mRNA起始密碼子AUG上游約10個鹼基處具有一段多嘧呤序列(5'AGGAGG3')稱為Shine-Dalgarno sequence能與核糖體小亞單元互補結合，藉此能正確找到起始密碼AUG開始轉譯。真核細胞則係由小亞單元(40S)與啟動因子(*initiation factors, IF*)、Met-tRNA_i從5'冠帽往3'端掃描，碰到第一個

AUG(*Kozak consensus sequence, 5'-ACCAUGG-*)後，大亞單元(60S)才加進來合成80S起始複體(*initiation complex*)。

每一個核糖體有三個tRNA結合位置，分別叫做A、P、E位。開始合成時，P及A位分別與mRNA的起始密碼子及下一個密碼子對齊。因此tRNA將第一個胺基酸帶進P位，第二個帶進A位。經酵素催化，兩個胺基酸即以胜肽鍵結合，附著於A位的tRNA上。然後核糖體在mRNA位移一個密碼子。第一個已去掉胺基酸的tRNA從P位移至E位，旋即離開核糖體。第二個攜帶胜肽的tRNA從A位移至空出的P位。於是A位又能接納下一個tRNA。如此週而復始，直至碰到stop codon，完成peptide的合成。

結論

自從2000年人類基因體被宣佈解碼以來，吾人了解蛋白質基因只有30,000至40,000個，可能比低等動物還少，但是人類的蛋白質基因，並非一個基因只產生一個蛋白質，為適應環境，在不同生長發育過程、營養狀態及外在環境(如溫差、輻射)下，同一個基因可用*alternative splicing*(約60%人類基因被發現具有此種多樣剪接)製造不同*isoforms*(如NFAT1 NFAT2等)，再加上基因突變所產生的各種異常蛋白質，種類繁多，我們都要積極去掌控了解，庶幾對人類的健康促進與疾病防治多所助益。



參考資料

1. Altmann SW, Davis HR Jr, Zhu LJ, et al. Niemann-Pick C1 Like 1 protein is critical for intestinal cholesterol absorption. *Science* 2004; 303(5661):1149-50.
2. Cerutti, H: RNA interference: traveling in the cell and gaining functions? *Trends Genet.* 2003; 19:39-46.
3. Lodish HF, Berk AJ, Matsudaira PT, et al: *Molecular Cell Biology*. 5th ed. New York: Macmillan Magazines Ltd, 2004:113.
4. Murray RK, Granner DK, Mayes PA et al: *Harper's Illustrated Biochemistry*. 26th ed. New York: McGraw-Hill, 2003.
5. Behrman RE, Kliegman RM: 2002 *Nelson Essentials of Pediatrics*. 4th ed.
6. Lian JS et al: Adrenoleukodystrophy: clinical analysis of 9 Taiwanese children. *Acta paedr Taiwan.* 2004; 45:272-7.

下 期 預 告

家庭醫學	進食困難病人的營養評估與餵食管的選擇	黃彥皓、陳鼎達
家庭醫學	臨床上常見的症狀高膽紅素血症-吉爾伯特症狀群	邱千芳、蘇世斌、林勤益
社區醫學	大眾傳播媒體對健康之影響	林竹君、黃政仁
家庭醫學	風濕病之另類療法：溫泉療法	王保琳、陳偉立、張煥禎