

TSMM研討會專題——基因檢測之困擾：突變分析

日期：2008年5月10日（星期六）

時間：13:30-17:30

地點：國泰人壽金控大樓B1

主席：簡志誠醫師（財團法人國泰綜合醫院汐止分院院長）

引言人：高國彰醫師（辜公亮和信治癌中心醫院病理科主任）、李宏謨教授（臺北醫學大學醫學檢驗生物技術學系主任）

討論者：李玲慧博士（中央研究院國家基因型鑑定中心副主任）、李銘仁醫師（國立台灣大學醫學院附設醫院基因醫學部主治醫師）、林文昌博士（中央研究院生物醫學科學研究所副研究員）、黃紀榕博士（財團法人國泰綜合醫院分子遺傳及生化研究室主持人）、曾崑元教授（馬偕紀念醫院病理科主任）（以上依姓氏筆畫順序排列）

簡志誠院長：曾理事長、高主任、李主任、還有各位先進，午安！歡迎各位來到國泰醫院剛落成的會議廳。自從人類基因體（genome）解碼後，基因檢測的進展就一日千里。基因檢測過去侷限於遺傳疾病的診斷，現在它已可用在疾病的預測。儘管用途變得更廣，但是病人、消費者及社會，都要求對基因檢測做更精確的定位。我們需要告訴他們，基因檢測的結果是真實的嗎？還是可能只是一種機率？若是後者，此機率會有風險上的考慮。所以，今天在理事長的安排下有此討論會。第一場會由李玲慧博士和黃紀榕博士向大家介紹更好、更精確、更有效率的方法，以辨識及偵測有問題的基因。第二場由李銘仁醫師與曾崑元醫師告訴大家，解讀實驗室數據時，如何應用現有的知識及不同的思考模式，以進行更合理的思考，因而能在未來的應用上更有效率並減少錯誤。相信今天的討論對未來的分子醫學臨床應用會造成影響。歡迎大家一起來接受這場知識的啟蒙。謝謝！現在請第一場的引言人—高國彰醫師發言。

第一場

高國彰醫師：首先請李玲慧博士向大家說明，如何以連鎖分析（linkage analysis）定位致病基因。



簡志誠院長致詞歡迎來賓們參加討論會。（彩圖詳見本刊網頁）

李玲慧博士：連鎖分析是一種使用統計尋找家族中致病基因所在位置的方法，其可行性建基在緊密相鄰的兩個基因座（locus）於染色體（chromosome）進行減數分裂（meiosis）時，其與另一條同源（homologous）染色體發生重組（recombination）的機率幾乎是零，因此親代兩個基因座上的對偶基因（allele）所形成的單倍基因型（haplotype）會毫無改變地傳給下一代；但距離較遠的兩個基因座發生重組的機率較高，親代兩個基因座上的對偶基因的單倍基因型，會因重組而產生新的單倍基因型傳給下一代。如果某疾病基因與某DNA標記非常接近，則某疾病基因與某DNA標記會被一起傳給罹病的子代；反之，若某疾病基因與某DNA標記距離很遠，則某疾病

基因與某DNA標記會被一起傳給罹病子代的機率就大大降低。因此利用追蹤特定DNA標記及其所形成的單倍基因型在家族中與罹病者共同分離的狀況，可以訂出疾病基因所在的位置。目前最常見的連鎖分析是全基因體連鎖分析（genome-wide linkage analysis），利用分佈於全基因體的DNA多型性（polymorphism）標記，如微衛星（microsatellite）與單一核苷酸多型性（single nucleotide polymorphism; SNP），使用Lod score分析計算連鎖的勝算之對數值，找出最高值的位置，再配合單倍基因型的演算，即可訂出疾病基因最可能座落的染色體位置。

曾嶽元教授：請您舉例說明。

李玲慧博士：我們以尋找遺傳性缺血性股骨頭壞死症（avascular necrosis of the femoral head; ANFH）的致病基因為例。我們利用微衛星進行基因型鑑定，從而計算此家族的重組率及疾病基因與微衛星之間的Lod score，再從基因型與家族資料推算出每個家族成員遺傳到的單倍基因型，最後定位到第12號染色體q13為疾病基因所在位置。然後利用對位於此區域內的基因之知識，選擇第2型膠原（collagen, type II, alpha 1; COL2A1）基因為序列分析的對象，並於其中找出家族罹病成員特有的突變（mutation）序列。此突變序列是否為致病性突變仍在進行模式動物研究階段，但以此突變序列作為家族中帶原者罹病之風險預估，已經有其實際價值。

高國彰醫師：在台灣的三個家族性股骨頭壞死（inherited form of osteonecrosis）病人中，發現家族性股骨頭壞死和COL2A1基因的突變有關，不知道其他家族的發病者或非遺傳性的病人，在COL2A1是不是也有同樣的突變產生？

李玲慧博士：我必須強調的是，這類分析是鑑定（identify）家族性疾病基因所在位置的研究方法，是從收集到的家族發病及未罹病者檢體中，



高國彰醫師引言。（彩圖詳見本刊網頁）

經過分析得到候選基因（candidate gene）位置，因此得到的突變對這三個家族是有專一性的，至於此基因突變是否也存在於其他家族或非遺傳性的病人身上，必須要經過檢驗才能知道。我們也分析過65個自發性（spontaneous）股骨頭壞死的病人，但並未發現他們的COL2A1有任何問題。換句話說，我們找到的關聯性是和特定家族的疾病有關，但是否跟有同樣疾病的其他家族，或有同樣疾病的其他病患有關？我想答案並不那麼肯定。因此若是要加入其他家族，必須要從頭再進行一次分析。

曾嶽元教授：對於股骨頭壞死的相關基因，你們鎖定的是COL2A1，不曉得文獻上是否有討論到其他膠原蛋白基因？

李玲慧博士：我並非這方面的專家，不過我們鎖定COL2A1完全是因為位置的緣故，亦即在我們定位出來的區域內，COL2A1基因是關聯性最大的，雖然一般來說，它和軟骨的形成有關，不過它的表現是在細胞外間質（extracellular matrix），可能會和骨頭的血流等有關。

高國彰醫師：如果以高密度（high density）的SNP進行你的實驗，會不會改變你現有的結論？



李玲慧博士報告主題「易感性基因與序列變異的定位與辨識」。(彩圖詳見本刊網頁)

李玲慧博士：您的問題很好！事實上我相信是不會的。如果有很大的家族族群，是可以利用增加標記密度 (marker density) 的方式，將疾病基因所在的位置鎖定到更小的範圍內，但是與疾病相關的基因仍會被包含在此範圍中。以目前這三個家族的大小來看，使用高密度的短相連重複序列多型性 (short tandem repeat polymorphism; STRP) 已經很足夠。另外要說明的是，STRP和SNP不同之處在於，STRP在對偶基因上的差異可以高達十種以上，而SNP雖然存在的密度很高 (大約每1kb即有一個SNP)，但是單一SNP在對偶基因上只有兩種變化，因此一個SNP標記的基因型為同型合子 (homozygous) 的頻率很高，鑑別率相對較差。所以STRP目前還是很常用在連鎖分析上，因為它的多型性是最高的。

高國彰醫師：你在進行連鎖分析時，會考慮到非編碼區 (non-coding region) 所造成的結果嗎？

李玲慧博士：越來越多的證據顯示，非編碼區和基因調控有關，例如microRNA和small RAN。不過這屬於更高層次的調控。而且當初我們已經先在COL2A1的編碼區 (coding region) 找到突變，我想，在這樣的情況下，大部分的實驗室大概都會認為疾病是該突變所造成的影響。目前因為高效能定序儀 (high throughput sequencer) 已問世，

我們可以將定位出來的區間做定序，雖然會增加複雜度，但不會遺失任何突變，包括在非編碼區上的變化。

高國彰醫師：接著再請李玲慧博士向大家說明，如何以相關性研究 (association study) 定位易感性基因 (disease susceptibility gene)。

李玲慧博士：相關性研究只是一種有關對偶基因及表現型 (phenotype) 同時發生的統計敘述，即患有D疾病又同時帶有對偶基因A的人，比在一般族群中沒有D疾病但帶有對偶基因A的人頻率來的高 (或來的低) 的時候，則可說對偶基因A與疾病D之間具有相關性。會發生族群相關性的因素有許多種，其中與疾病易感性基因相關的因素有 (1) 直接因果關係：對偶基因A使帶原者 (carrier) 對疾病D具有易感性；(2) 連鎖不平衡 (linkage disequilibrium; LD)：兩個基因座上的四個異型對偶基因所產生的四種單倍基因型之頻率若偏離隨機相關性太遠時，則此對偶基因間為連鎖不平衡。針對複雜疾病，族群相關性研究之目標即在於藉由標記分子與疾病易感性基因之間的LD來發現與疾病的相關性。此種研究的假設前提為兩個患有疾病D而無親屬關係的人，實際上是由很久很久以前一個共同祖先傳遞了該疾病給他們，因此除了該疾病對偶基因外，他們也傾向於共享該祖先的某個 (或某些個) 緊靠在D基因座上的對偶基因。事實上族群可說是由一個家族擴展開來的，族群與相關性程度間的關係，是由於祖先的疾病易感性基因和緊密相連結的標記之間存在著LD。假如二個「無關的」人各自遺傳來自他們共同的、遙遠的祖先的一個疾病易感性對偶基因，在經過許多世代 (generation) 和許多減數分裂的過程後，多次的染色體重組互換將使他們所共享的祖先的染色體片段變得非常小，因此只有緊緊地連接在疾病易感性基因座旁的對偶基因基因座可能同時被共享。若此二基因座上的對偶基因存在著LD，則藉由偵測緊鄰的基因座 (必須為已知的標記分子) 上的對偶基因在病患及族群中的頻

率差異，可以推算疾病與標記分子間的相關性，並且找到候選疾病易感性基因的位置。

曾嶽元教授：請您舉例說明。

李玲慧博士：我們以尋找史蒂芬強生症候群（Steven Johnson syndrome）之易感性基因座為例，說明如何利用病患、對照組與族群的對偶基因型頻率尋找與特定藥物所引發的嚴重皮膚性藥物反應相關之基因座及對偶基因型。在使用鑑定候選基因基因型的策略下，位於第6號染色體p21.3上與免疫反應密切相關的人類主要組織相容性複合體（major histocompatibility complex; MHC）之HLA-B*1502對偶基因型顯示與抗癲癇用藥—醯胺咪嗪（carbamazepine; CBZ）所引發的史蒂芬強生症候群之間有非常強烈的相關性。

聽眾：既然史蒂芬強生症候群和HLA-B*1502的關聯性極大，那麼，在給予病患CBZ之前，有無先進行基因檢測（genetic testing），檢驗其HLA是否為B*1502基因型的必要？

李玲慧博士：目前衛生署尚無此打算，只有在藥物說明書上有標識說明警告病人，若其HLA為B*1502，服用該藥物可能會引起史蒂芬強生症候群。因此可能還要收集更多病人，篩檢出HLA為B*1502的病人後，避免他們使用該藥物。若真能因此降低史蒂芬強生症候群的發生率，健保才有可能給付此篩檢的費用。另外，使用CBZ不會發生史蒂芬強生症候群的人當中，也有3%的HLA為B*1502，因此其他參與引發史蒂芬強生症候群的基因為何？還需要進一步的研究。

林文昌博士：疾病的分類對實驗很重要，若實驗初始即分類錯誤，會使後續的結果無法解釋。

高國彰醫師：你只能相信他人提供的訊息，此外，不同醫生的診斷可能也會有所不同。

曾嶽元教授：所以國外研究者開始找這類基因時，花了很多功夫在進行家庭訪問，但台灣因人事經費有限，要執行這個步驟極為困難。因此，台灣的相關研究與國外相比之下吃虧很多，因為收集的族群當中，可能有很多不是真實的。

林文昌博士：所以進行相關性研究（association study）的風險還是相當高，我們收集的檢體中，有些診斷不一定正確。

高國彰醫師：接著請黃紀榕博士向大家說明，基因差異表現與疾病的關聯性。

黃紀榕博士：細胞內約有30,000個基因，而常表現的則約有10,000個左右，這些基因許多已確定參與細胞生長、增生、分化與轉型。在整個細胞發育過程中，因應不同的發育時期或環境因素的變異，不同的細胞基因會有被激發或抑制的現象。根據研究顯示，於相類似條件下所引起的基因表現差異群組，往往具有類似或相關的基因功能。因此，特定的基因表現模式也許可成為不正常細胞調節的指標，其中也有些基因群參與整個細胞癌化的過程，這些不正常的基因表現將是癌症發生與進展的關鍵，因此單一或多個基因的差異表現是否可成為診斷特定人類癌症的標記（marker），已是許多科學家努力想要釐清



黃紀榕博士準備報告主題「基因差異表現與疾病的關聯性」。（彩圖詳見本刊網頁）

的課題。調節細胞基因的差異表現除了複雜的細胞內基因網路相互影響外，還有epigenetics及microRNA亦會造成基因的差異表現。Epigenetics是在不改變DNA序列下，利用一些化學修飾來調節基因的表現，目前已有研究顯示可改變某些癌症的基因表現。在epigenetics中最受人矚目的就是基因的甲基化（methylation），甲基化的修飾經常出現在細胞基因中，尤其是在某些癌症中更是如此，包括大腸直腸癌、乳癌及肺癌等，皆發現與不同基因的過甲基化有關。另一方面，細胞內的microRNA利用其特殊的結構影響特定基因的表現，因此若是microRNA發生變異，細胞則有致癌的可能。根據最新的研究報導，不論是肺癌或大腸直腸癌等好發的癌症，皆有基因的差異表現與microRNA有關。綜合前面所述，基因群組的差異表現與疾病息息相關，尤其是癌症更是如此。因為文獻報導顯示，單一基因的變異鮮少發展為一惡性腫瘤，相對地，致癌基因、抑癌基因或microRNA等多層面的變異為目前最為被接受的路徑。因此了解在特定因素下產生差異表現的基因群組，已視為今日對於各種癌症的相關研究刻不容緩的議題。事實上，基因差異表現的比較可應用在各不同層面，例如同一種癌症的不同期別之比較。由於這些複雜性，因此需要一高效率的平台，將與癌症相關的差異表現基因群組一併探討。

曾欽元教授：您指的高效率的平台是什麼？

黃紀榕博士：DNA微陣列（microarray）是一個可用來分析基因表現的高效率平台，迄今已有多篇文獻報導利用DNA微陣列技術發現那些與癌症或抑癌相關的基因分子群。近年我們注意到利用DNA微陣列雜交（hybridization）可發現許多基因在大腸直腸癌患者的糞便中具有差異的表現。如今大腸直腸癌已成為全球惡性腫瘤致死的原因之一，並成為主要的醫療資源負擔，然而大腸直腸癌患者的存活率與早期發現息息相關。因此能直接從糞便腸道細胞發現與大腸直腸癌相關基因的

差異表現，將可成為發展大腸直腸癌早期發現與成功治療的不二法門。多基因表現的變異調節對研究人類癌症發生與進展而言相當重要，了解差異表現的基因分子將有助於癌症患者的早期診斷及預後評估。以大腸直腸癌為例，直接定型差異表現的基因群，將有助於癌發前的篩檢、大腸直腸癌分期與治療的參考。因此基因差異表現的評估與定型，將是臨床上各種癌症的重要分子診斷依據。

林文昌博士：糞便檢體常含有極多微生物，這麼多複雜的物質在檢體中，不會造成干擾嗎？

黃紀榕博士：我們從糞便檢體中抽取的RNA，有90%以上都是屬於腸胃道的正常菌落（normal flora），為了避免干擾，在將RNA轉換成互補DNA（complementary DNA; cDNA）時，不能使用隨機六聚物（random hexamer），必須使用寡去氧胸腺苷酸（oligo(dt)），因為原核生物（prokaryote）沒有寡去氧胸腺苷酸的序列，如此一來，即可排除大部份來自原核生物的干擾。接著若是要放大（extend）某些特定的基因，聚合酶鏈鎖反應（polymerase chain reaction; PCR）的專一性（specificity）非常重要。利用這些方法即可排除干擾。



林文昌博士（右）提出問題。（彩圖詳見本刊網頁）

林文昌博士：糞便檢體中的核糖核酸酶（Ribonuclease; RNase）應該也很高。

黃紀榕博士：所以我們發展一種特別的緩衝液（buffer），可抑制RNase。該緩衝液目前已由國泰醫院提出專利申請。

曾嶽元教授：國泰醫院對大腸癌（colon cancer）的研究很多，不知道是否有將結果應用在臨床上？

簡志誠院長：目前還沒有應用在臨床上。如果病人可以接受這種沒有結果的結果，那麼現在就可以應用了。隨後麻煩李主任為第二場討論引言。

第二場

李宏謨教授：先請台灣大學的李銘仁醫師為大家說明，如何證實致病基因突變點位。

李銘仁醫師：隨著基因體科學的進步，解密基因變得簡單容易，但是在發現基因序列改變後，伴隨而來的問題是要如何解釋這個改變。在人類基因體計畫中，基因多型性變異的定義是DNA的改變並未伴隨著表現型的改變或者是該改變是無害的，而且在人群中此種改變的發生率高於0.01。而所謂「突變」則是指會致病且會使遺傳的DNA結構變化。簡單而言，就是要判斷一段DNA序列的改變是會造成表現型的改變（突變）或者是沒有造成表現型的改變（基因多型性）。前面一再被提到的單核苷酸多型性（single nucleotide polymorphism），縮寫為SNP（念法為〔snip〕），意思是「DNA序列中的單一鹼基對（base pair）變異」，也就是DNA序列中A、T、C、G的改變，即基因組的某個點位出現兩個或多個的核苷酸可能性。事實上，在目前已知的SNPs中，所佔比例最多、最常發生的單一鹼基對變異是以胸腺嘧啶（thymine; T）取代胞嘧啶



李宏謨教授引言。（彩圖詳見本刊網頁）

（cytosine; C），約佔已知總數的三分之二。目前科學界已發現了人體約400萬個SNPs，平均而言，每1kb長的DNA中，就有一個SNP存在；也就是在每個人的DNA序列中，每隔1kb單位長度，就至少會發生一個「單一鹼基對變異」。由於SNP的發生頻率非常之高，且每個人DNA上所發生的SNP皆不同，同時此SNP又可以遺傳給下一代，故SNP常被當作一種基因標記（genetic marker），以進行研究。雖然在前面提到突變與SNP的主要差異是單核苷酸多型性變異不會有表現型的改變，但是由於SNP的產生可能會造成蛋白質表現的改變，所以SNP可能會影響人類對疾病的感受性，使人特別容易或特別不容易患上某些疾病，或使得人體細胞對於治療藥物的反應性有所差異。

曾嶽元教授：請您舉例說明。

李銘仁醫師：舉例來說，引發愛滋病（acquired immunodeficiency syndrome; AIDS）的人類免疫缺陷病毒（human immunodeficiency virus; HIV）在感染人的免疫淋巴細胞時，需要淋巴細胞表面具有CCR2和CCR5。當HIV病毒陽性（positive）的感染者的免疫淋巴細胞在CCR2基因上出現SNP時，此患者會比其他感染者晚2-4年發病。另外，約9%的白種人有一種特定的CCR5缺失（deletion）變異，其CCR5基因有一段32個核苷酸長度的序列

佚失，而具有此CCR5變異的個體，會讓HIV病毒難以感染。但必須注意的是，並非所有的SNP都有臨床意義。對疾病發生和藥物治療有重大影響的SNP，只佔四百萬個SNP中很小一部分。即使產生了SNP，也不一定造成蛋白質氨基酸（amino acid）編碼改變或基因表達（expression）調控改變，或導致蛋白質結構或活性，而造成對於藥物的特殊影響。

李宏謨教授：請李醫師繼續為大家說明，如何證實致病基因突變點位。

李銘仁醫師：基因突變的定義隨著科學的進步而變得複雜困難，因為（1）傳統生物學領域中，突變是指基因序列上的任何改變，而不管其對生物體造成影響與否。（2）但是在應用醫學的領域裡，對於一個基因上的序列改變，我們首要關注的問題即是這樣的改變對生物所帶來的影響，是否會導致疾病的發生，或是根本不會造成表現型的變化。因此，在基因醫學的定義中，突變即是指一個有害的（harmful）基因序列改變，而基因多型性是無害的（harmless）序列改變。要區分基因單核苷酸多型性變異與突變之前，必須先要注意以下幾點：（1）此DNA序列的改變是否因為PCR出錯而造成。（2）DNA的分析必須要完整，切忌找到第一個核苷酸改變就終止研究並認定那

是致病的基因突變。（3）在某些特殊的狀況下，還是得靠基因功能分析（expression analysis），才能證明基因的功能是否因為此變異而受到影響。在以上三點前提之下，我們進一步可以依照以下的方法來區分DNA的變異是屬於多型性變異抑或是致病的突變：（1）根據變異的種類來區分。若是框架移動（frameshift）或者是導致產生終止密碼子（stop codon）的變化，因為會影響整個蛋白質的結構，所以多屬於會改變表現型的突變。（2）根據基因變異是否隨著性狀遺傳決定。如果得到該基因變異者皆會發病，而沒有遺傳到此基因變異者皆不會發病，那麼此變異就很有可能是致病的突變。（3）根據氨基酸被改變的重要程度區分。若此基因變異會導致蛋白質上重要的氨基酸改變，那麼此變異也較可能是會致病的突變。（4）根據變異在人類的發生率來區分。基因多型性變異的定義為在人類的發生機率高於0.01，因此當某種DNA變異的發生率低於0.01時，就是屬於很罕見的變異，就有可能是會致病的突變。在實際操作上，當我們發現一個新的基因變異時，會先檢測50個人（即100條染色體）是否也存在此基因變異。（5）根據基因的功能是否被改變來決定。如果該DNA變異發生可以被證明會改變原本基因的功能時，就是所謂的突變，而這也就是最後且最關鍵的一個條件。

曾欽元教授：您提到的基因檢測沒有想像中的那麼單純。有些人的基因發生突變不見得會產生疾病，而在有疾病的人身上也不見得找得到基因突變，所以很可能這些檢測怎麼做都不會錯，但問題是怎麼跟病人解釋這些情況？

李銘仁醫師：我認為單一基因單一疾病（one gene-one disease）的觀念必須打破，如同前面討論到找不到突變的問題，因為很多疾病並非單一基因發生突變造成的，當我們在一疾病上找不到某個基因突變時，也許在別的基因上是有突變的。例如尼曼匹克症（Niemann-Pick disease）（一種脂質代謝異常的遺傳疾病）C型的病患約有



李銘仁醫師報告主題「如何證實致病基因突變點位」。（彩圖詳見本刊網頁）

95%是NPC1 (Niemann-Pick disease type C1) 基因突變造成的，而NPC2突變佔5%。因此在一病患的NPC1上找不到突變時，不能說該病患不是尼曼匹克症，也許其突變發生在NPC2，也有可能發生在內含子 (intron)。所以跟病人解釋時，需要注意這些問題。

曾嶽元教授：因此病人有可能繳交數萬元的檢測費用之後，卻得到毫無所獲的基因檢測結果？

李銘仁醫師：所以我們會先跟病人解釋這些觀念，若他們能接受再進行檢測。

李宏謨教授：接下來請曾嶽元醫師為大家說明，如何從演化的觀點來判別突變。請曾醫師先說明基因變異在演化上的角色。

曾嶽元教授：在演化史上，蛋白質可以累積一連串的氨基酸置換，甚至是增加新的氨基酸或減少原有的氨基酸，而改變原蛋白質的特性。雖然絕大多數氨基酸的改變都是有害的，因此會被自然淘汰，但有少數是有利的。因而，改變後的氨基酸有可能使生物體更能適應變動的環境。這些有利的改變會很快地在一個族群中散佈，而取代原有的序列，此謂之固定 (fixation)。當某生物群進入全然不同的環境，而與原先之生物群隔離。如果，在漫長的時間上，再發生基因體的重組，而使得彼此不能再交互繁殖 (interbreed)，那麼新的物種就有可能出現了。譬如，大陸板塊的漂移，使得在南美的靈長類演化成新世界猴，而在非洲的靈長類則演化成舊世界猴。其相對應的基因會因為適應不同的生態環境，而有不同的序列固定。因此，兩個物種在相對應的蛋白質上的不同之處，來自於當它們的祖先分家後，彼此不再交互繁殖時，所累積下來的突變。以人類與黑猩猩 (chimpanzee) 作比較，彼此在基因體上有99%以上的相似度。這不到1%的差異包括 35×10^6 個核酸的置換 (nucleotide substitution)， 5×10^6 個佚失或插入 (insertion)，以及許多染色體的重組。



曾嶽元教授報告主題「從演化角度鑑定突變」。
(彩圖詳見本刊網頁)

人類與黑猩猩的差異絕大部份是在基因的結構 (gene structure) 和調節 (gene regulation)，而不在蛋白質的演化。因此，人類與黑猩猩有25%的核酸置換上的不同出現在CpG雙核酸上，而蛋白質則有29%完全相同，不同的蛋白質，大多也只是一兩個氨基酸的差異而已。

李宏謨教授：這些演化上的變異是否也會有致病性突變 (pathogenic mutation) ？

曾嶽元教授：由於每一個氨基酸各有其特性，因此一連串排列的氨基酸所賦予的化學特性，構成奇妙而獨特的三度空間。然而，有些氨基酸彼此的化學特性類似，因此互相取代時所造成的空間影響是有限的。這種氨基酸置換的影響，有可能讓蛋白質更能適應改變後的生理情況。化學特性不相似的氨基酸互相取代時，對蛋白質有較大的空間影響。因此，更有可能讓蛋白質適應劇烈變動的生理情況。然而，當外在生態和內在生理條件都未有太大變化時，那麼蛋白質出現太大的改變就不利於該生物體了。此外，蛋白質上的活化點，其氨基酸的側群 (side-groups) 可和其他的巨分子接觸而參與特定的化學反應。如果這幾個在活化點上的氨基酸之一出現改變的話，就會影響蛋白質對此巨分子的接觸。這些即為致病性的突變。

李宏謨教授：一般而言，突變的發生率為何？

曾嶽元教授：氨基酸的改變源自於核苷酸自發性的突變 (spontaneous mutation)。雖然真核 (eukaryotic) 細胞自發性的突變率，我們沒有確實的數據，但咸信應與組菌相似。就細菌而言，任何一個鹼基對在每個世代的突變率大約是 10^{-9} - 10^{-10} 。對於一個1,000 bp 大小的基因而言，每個世代的突變率大約是 10^{-6} 。而就細菌的基因體來說，每個世代的突變率大約是 3×10^{-3} 。人類基因體，其中25%為基因 (24%為內含子 (intron)，1%為外顯子 (exon))，絕大多數人類基因的突變為點突變 (point mutation)。更仔細的來分，58%為錯義 (missense) 突變或無義突變 (nonsense)，16%為小的佚失突變，10%為剪接 (splicing) 突變，6%為小的插入 (insertion) 突變，5%為大佚失突變，2%為大的基因體重組，而不到1%為調節性的突變。點突變和小的插入及佚失的突變為偶然發生的。大概在基因體內的任何地方都有相似的機率。發生機率特別高 (即自發性突變機率的10到100倍) 的地方，則稱為熱點 (hotspot)。

李宏謨教授：可否再說明一下，您在此處所提到的「突變」和孟德爾的「突變」，以及剛才李醫師提到的「SNP」，彼此之間的關係。

曾嶽元教授：原始的孟德爾觀點把對偶基因分為野生型 (wild-type) 和突變型 (mutant) 兩種，是太武斷些。從近代的觀點我們知道，對偶基因可以是多重的 (multiple alleles)，亦即同一基因有多種變異 (variant)，而每一種對應一種表現型。當一個基因座上有多重的對偶基因時，我們稱之為基因的多型性 (genetic polymorphism)。其中沒有任何一個對偶基因是必然的野生型。例如人類的血型基因就是如此。我們不能說血型的野生型是A型或B型。我們只能說血型基因是多型性的。因此，嚴格地說，當某變項 (不管是對偶基因、表現型、序列變異、

或者是染色體的結構變異) 在一族群中顯著出現有兩種以上時，謂之多型性。較不嚴格地說法是，當某一序列變異在一族群中出現的頻率大於1%時，或者是出現任何非病理性之序列變異時 (不管其出現的頻率)，我們都認為它是多型性的 (polymorphic)。對偶基因之間的單一核苷酸之不同，我們稱為SNP。人類基因體中，平均約1,330個鹼基即有一個SNP (因此可以說，每個人都是獨特的)。然而，當我們檢視同時存在的多個變異時，我們並不知道它們有沒有優勢或劣勢的選擇。所以，有些變異是已被固定下來的多型性；而有些則是有害的，它們只是被演化淘汰前短暫的存在而已。區分這兩種差別，就是基因檢測的困擾之一。

李宏謨教授：現在就請說明，如何從演化的觀點來評估突變。

曾嶽元教授：以人類的基因為考量基礎，大約21%的 (~8,000個) 基因是真核細胞和原核細胞所共同有的。這些基因的功能大多為維持生命所必須的，譬如代謝、複製、轉錄 (transcription) 和轉譯 (translation)。此外，大約32%的基因為所有真核細胞所共同的，這些基因是應付真核細胞特有的特徵，如特別的胞器 (organelle) 和細胞骨架。另外，大約24%的基因為多細胞動物所共有的，這和組織的分化有關。最後，約22%的 (~8,000個) 基因是脊椎動物才擁有的。這些大多和免疫及神經系統有關。因此，在演化上距離不會太遠的生物物種，彼此之間有相對應的基因。此謂之直向同源 (ortholog)。至於在同一物種中，有的基因可出現類似的副本 (duplicate)，這不叫作直向同源，而是稱為形似 (paralog)。當我們要從演化的觀點來評估突變時，我們比較分析的對象是直向同源而非形似基因。

李宏謨教授：請舉例說明直向同源和非形似基因。

曾嶽元教授：我以大家較為熟悉的球蛋白（globin）基因為例來說明直向同源和形似基因。球蛋白最原始的直向同源基因為植物的豆血紅蛋白（leghemoglobin）。它含4個外顯子。在約8億年前因外顯子融合（fusion）而衍生出含3個外顯子的球蛋白基因。此基因現在可見於八目鰻（lamprey）及盲鰻（hagfish）。此基因在5億年前再經過拷貝（duplication）和分歧化（divergence）而形成相連在一起的兩個形似基因：初始 α 基因和初始 β 基因。這種相連的基因現在可見於兩棲類及魚類，但鳥類則沒有。這兩個基因後來落在不同的染色體，並且擴張成簇（cluster），此發生於鳥類及哺乳類分家之前。約在2億年前，初始 β 基因於真獸類衍生出原 ϵ （proto- ϵ ）基因和原 β （proto- β ）基因。此兩基因後來在負鼠（opossum）演化成為 ϵ 和 β 基因。但是在胎盤哺乳類的祖先，原 ϵ 基因則在1億年前衍生出 ϵ 、 γ 、 η 三個形似基因；而原 β 基因則在4千萬年前衍生出 δ 和 β 兩個形似基因。此外，粒線體（mitochondria）基因也是一個很好的例子。除了少數真核生物如腸梨形蟲（*Giardia intestinalis*）和陰道滴蟲（*Trichomonas vaginalis*）外，每種生物的粒線體都有DNA，其所含的基因在不同生物間互為直向同源基因。

李宏謨教授：有沒有可量化直向同源基因變異的分析方法？

曾嶽元教授：Ruiz-Pesiri等人在2004年提出保留係數（Conserved Index; CI）這種計算方式來評估物種之間直向同源基因的演化變異。所謂CI即人類的氨基酸殘基（amino acid residue）在不同物種的直向同源基因所對應的蛋白質上出現的頻率。頻率愈高表示該位置上的氨基酸代表的生理角色不容改變，因此在演化過程中愈沒有變動。相反地，頻率愈低則有兩種可能性：（1）該位置上的氨基酸不具太重要的生理功能，因此被其他氨基酸代換後，也不會有太大的影響；（2）不同物種面臨演化上的獨特需求，因此個自篩選出該



學者（左起：李玲慧博士、黃紀榕博士、李銘仁醫師、林文昌博士、曾嶽元教授）熱烈討論。
（彩圖詳見本刊網頁）

物種所需的氨基酸殘基。演化上愈相近的物種，其生理環境愈相近。那麼就愈不會因演化上的獨特需求，而篩選出不同的氨基酸殘基。因此，愈重要的氨基酸殘基就有愈高的CI值；而愈不重要的氨基酸殘基就有愈低的CI值。除非，不同的物種在演化時間上竟然短到來不及在某基因上累積無害的突變。因此，到底挑選的物種到底在演化上要跨多遠，倒是值得商榷。

李玲慧博士：您在分析的時候，每個物種只取一條DNA序列當代表，但是人類以外的生物也有其SNP，您怎麼處理這種情形？

曾嶽元教授：其他動物有其SNP，但目前的資料庫並沒有足夠的數據讓我們研究這個問題。因為野生型序列最常見，所以我們只好假設，貯存在基因庫的DNA序列是野生型的。

李宏謨教授：您剛才提到，到底挑選的物種要在演化上跨多遠，值得商榷。您認為分析人類的突變時，應該跨越多廣泛才好？

曾嶽元教授：Ruiz-Pesiri等人挑選了39種物種計算CI，包括12種靈長類、22種其他哺乳類、4種其他脊椎動物和1種非脊椎動物。用散佈這麼廣

泛的物種來評估人類的突變，是否不符合實際的情形？為了檢定此一考量，於2007年，我們的研究團隊將挑選的物種限定在哺乳類，選擇15種靈長類和115種其他哺乳類，依Ruiz-Pesiri等人的方法計算係數，稱為哺乳類保留係數（mammalian conserved index; MCI）。接著，我們比較CI和MCI。首先以22個所謂的致病性（pathogenic）之粒線體基因突變，計算CI及MCI的平均值（mean）和95%信賴區間（confidence limits）。這22個粒線體基因突變是Ruiz-Pesiri等人在2004年挑選來測試的。Ruiz-Pesiri等人計算的CI之平均值和標準差（standard deviation）為 $93.4 \pm 15.7\%$ ，95%信賴區間為5.7%；我們所計算的MCI之平均值和標準差為 $95.8 \pm 15.3\%$ ，95%信賴區間為6.4%。我們將殘基之係數大於平均值者，視為演化上高度保留的氨基酸，並認為其突變為致病性的；係數小於平均值減去95%信賴區間者，視為演化上低度保留的氨基酸，其突變為非致病性的（nonpathogenic）。介於兩者之間者，則為意義未明的突變（mutation of undetermined significance）。依此標準，再去預測12個已被用質融合細胞（cybrid cell）證實為致病性的粒線體基因突變，比較CI和MCI何者有較佳的預測準確度。在12個致病性突變中，CI可預測出8個為致病性的突變，而MCI可預測出9個為致病性的突變，顯然MCI略勝一籌。因此，我認為在哺乳類的範圍內就足夠了。所以，我會用哺乳類保留係數來作計算。

簡志誠院長：從今天的座談會我們可以瞭解，基因檢測與診斷需要更多人來參與，大家從不同的角度來分析。顯然，臨床醫學和基礎醫學研究者需要更多互動。我們可以重新思考，是不是應該跨領域合作？非常謝謝大家熱心地參與討論。



生物醫學
BIOMEDICINE JOURNAL