

重建 東亞人類的族譜

文明的發生、發展、繼承過程中，多少民族輝煌過，又迷失在歷史的煙塵中，然而基因組分析這把利剪將歷史的死結一一裁開，又把一個個人群串連起來，追根溯源，重新修編各民族生生不息的族譜。

撰文／李輝、金力

用基因組分析的方法來解讀人類群體的歷史，成就了一門新興的學科——分子人類學。這一學科所倚仗的材料就是人類個體之間的基因組差異，也就是遺傳變異。人類的基因組結構大體上是相同的，而個體差別可能在整個基因組中所佔的比例不到1%，正是這些差異造成了人與人、族與族之間外觀和生理的差別。

人類基因絕大部份都位於染色體上，每個人有兩套染色體，一套來自父親，一套來自母親，Y染色體是男性染色體，來自父系遺傳；粒線體中也有基因組，全部來自於母親。由於Y染色體和粒線體基因的遺傳來源比較單純，使得我們可以更清晰地辨認其歷史軌跡，是目前使用最廣的分子人類學材料。

根據基因組的多樣性分析，人類起源於東非，因為東非人群的多樣性積累得最高，並且位於各種遺傳材料譜系樹的根部。從東非到東亞的遷徙過程中，人類群體經歷了許多次瓶頸效應和天擇的作用，使得人類到達「目的地」以後，變得和出發地的人群非常不同（見37頁〈人群遺傳變化與瓶頸效應〉）。東非最古老的土著

是布希曼人，他們操著郭依桑語系的語言，皮膚呈灰紅色、頭髮蜷曲、個頭不高、臀部脂肪特別厚。而到達東亞的人群，已經轉變成筆直的黑髮、褐黃到淺黃的皮膚、勻稱的身材，迥異於布希曼人。

兩群先民一前一後到亞洲

由於Y染色體的特殊性質，對亞太地區的民族關係有最好的辨析度，所以我們主要以Y染色體的多樣性來說明中國各民族的形成過程。「國際Y染色體命名委員會」把全世界的Y染色體分為A-R幾個大類群。澳洲土著的Y染色體大多屬於C型，其年代非常古老，這一類型從中東的沿海地區出現，沿著印度的沿海地區擴展，一直到東南亞和東亞沿海，甚至延伸至美洲沿海，也延伸到澳洲，這批人約在六萬年前抵達，稱為「早亞洲人」。

我們相信，早亞洲人是沿著海岸線向東方擴散開來的。海岸線顯然是一條容易通行的捷徑，考古發現也已經證實了東南非的人類在10幾萬年前就開始在海邊居住，並且依靠海洋資源生存。所以這條最早的沿著印度洋海岸順時針方向的遷徙，可能始於10多萬年前的南非。

關於作者



金力（左）於美國德州大學休士頓醫學健康中心取得醫學與遺傳學博士學位，現在是中國上海復旦大學現代人類學研究中心的主任。李輝自該中心取得博士學位，目前任職於美國耶魯大學。

個體間基因組 差異的分類

短片段重複序列 (short tandem repeat, STR) 是在基因組的某一區域中一小段片段 (往往是若干個鹼基) 重複了若干次, 而且在代代相傳中, 這種重複數可能改變, 造成群體中的不同個體在這一片段可能有著不同的重複數。

單核苷酸多型性 (single nucleotide polymorphism, SNP) 是在基因組某一個特定位置上的單一鹼基發生變化, 可能是一種鹼基變為另外一種, 比如 A 變成 G; 也可能是缺失或額外插入一個鹼基; 有時候還會發生插入或缺失一個小小的片段。SNP 在基因組中非常常見, 也是形成功能性差異的主要基因組差別。

複製數變異 (copy-number variant, CNV) 是比 STR 更複雜的片段在基因組中插入、缺失或是重複若干次。CNV 的重複區段比 STR 的功能性強, 但是重複的次數卻沒有 STR 那麼多, 突變的速度則接近 SNP。研究者認為它對人類生理差異的影響更大。

亞太地區的 Y 染色體類型中還有一種 D 型, 也可能是早亞洲人帶來的, 雖然他們的攜帶者可能與 C 型的攜帶者並不同, 但是到達東亞的時間差不多。這種類型出現在一些更神秘的人群中: 亞洲的小黑人。

亞洲的小黑人也叫做尼格利陀人, 分佈在菲律賓呂宋島中部、馬來半島北部的山地等狹窄地區, 以及緬甸東南面的安達曼群島, 他們的 Y 染色體大多是年代非常古老的 D 型。另外, 分佈於日本本州島北部、北海道以及俄國庫頁島南部蝦夷人的 Y 染色體主要是 D 型, 中國青藏高原的羌族、藏族系統民族中也有一定比例的 D 型 Y 染色體, 四川和甘肅交界地區的白馬氏人, Y 染色體也全部都是 D 型。

這些 D 型 Y 染色體的人群在非常遠古的時代中有著共同的祖先, 但是我們無法得知他們怎樣擴散到這些地區。這些人群現在的外形特徵差異很大, 南方的「小黑人」往往被認為是亞洲地區的「黑種人」, 而北海道的蝦夷人曾經被認為屬於「白種人」的一支, 白馬氏人則與其他典型的「黃種人」沒有什麼差別, 或許他們都與「棕種人」更接近, 屬於「早亞洲人」的後代。

「晚亞洲人」構成了我們現代東亞和太平洋地區人群的主體, 在 Y 染色體類型上主要為 O 型。晚亞洲人可能是追逐獵物從內陸來到遠東的, 大約三、四萬年前才到達東南亞地區, 並在兩萬年前的冰河盛期才進入中國, 由南往北分佈, 整個擴張開來。在這個過程中, 他們可能與早亞洲人爭奪生存空間。最終由於技術和體力略佔優勢, 在大部份地區勝出, 早亞洲人只留下很少的比例融合入各地的人群中。

晚亞洲人從南向北擴散

晚亞洲人進入中國以後, 迅速地擴散開來, 並分化成各個民族, 我們可以用「九州」來說明當時人群散居與遷徙的區域。今天最廣為流傳的九州劃分方式是《尚

書》指出的冀州、兗州、青州、徐州、揚州、荊州、豫州、梁州、雍州, 但是兗州、青州、徐州都是山東一帶的地方, 與其他六州不成比例。最均勻的分法是把兗青徐三州合併, 加上北方的幽州和并州, 這樣的九州兼顧了虞夏時代地理區域和風土人情的差異。

最初九州的文化都是相對獨立的起源, 並長期保持各自的特色。(見 39 頁〈九州各族的文化遺跡與遺傳類型〉)。以分子人類學方法對九州各地的古代和現代人群做了詳細分析, 發現從新石器時代開始 (大約在一萬年前), 九州各地的人群漸漸在體質上發生了細微的差別, 當時幾個不同區域人群的遺傳結構, 已經有了明顯的差異。

O 型 Y 染色體有若干亞型, 在揚州良渚文化的遺骨中, 我們發現了 O1 亞型, 在荊州的大溪文化遺址中則發現了 O3a4 亞型, 并州的陶寺遺址中發現了 O3a5 亞型。這些不同的亞型在各個文化之間形成了明顯的遺傳分界。在現代民族中, O1 亞型主要分佈在侗傣民族和南島民族中, 有些台灣原住民族幾乎只有這種亞型, 這些民族都是古代百越民族的後代, 所以揚州就和百越民族、良渚文化、O1 亞型 Y 染色體聯繫在一起了。O3a4 亞型出現在現在的苗瑤族群和孟高棉族群中, 荊州是苗瑤民族的故土, 大溪文化必然是苗瑤的祖先留下的。

既然九州的居民都有同一起源, 又如何各地變成不同的民族呢? 這可能要追溯到三萬多年前, 我們的祖先在東南亞的叢林中探索的時候, 常迷失在東南亞的叢林、山嶺和谷地。遙遠的村落漸漸失去聯繫, 隨著擴散的距離越來越遠, 一個個小小的瓶頸效應發生了, 人群的遺傳結構和體質特徵開始出現差異。今天, 族群不斷的遷徙和融合, 使得大部份人有著來自不同區域的遺傳成份, 特別是部份漢族和藏緬民族從北方回到了南方, 他們的主要 Y

人群遺傳變化與瓶頸效應

一般認為，遺傳突變是因為DNA複製的錯誤而產生的。許多突變會破壞基因結構，結果這些具有突變個體就不會存活下來，這些突變也即消失了。沒有危害的突變就會在群體中積累起來，群體的歷史越久遠，這些突變的數量就越多。這就是分子人類學計算群體歷史年代的基本概念。

突變如果發生在基因結構中，可能會引起個體生理差異。這些差異在某些環境中，不會造成個體生存能力的差異，當環境改變時，具有某些突變的個體可能就無法適應，而漸漸在群體中減少。在同一區域中的個體往往具有相似的特徵，通常是天擇的結果。

一個典型的例子是好發瘧疾的熱帶地區對地中海貧血症的正向選擇。東南亞和中國南方的人群中，地中海貧血基因很常見，帶有這種疾病基因的個體對瘧疾的抗性比較強，使他們更容易生存下來，結果不管是南方的土著還是從北

方來的漢族、緬族移民，在南方瘴氣瀰漫的地區，這個疾病基因都增加了。直觀的結果使他們的長相頗相似：嘴唇變厚、鼻翼變寬，眼眶深陷、額頭突起，這都是地中海貧血的個體在胎兒期缺氧的結果。這種外形和基因組的相似性，不能體現群體之間的遺傳親緣關係，所以在分析群體親緣關係的研究中，任何可能受到天擇的基因差異，都不是良好材料。

基因組的大部份區段並沒有直接的生理功能，通常不會受到天擇的影響，可以比較自由的累積突變。在沒有天擇作用的時候，群體之間的遺傳差異是如何產生的呢？答案是隨機。在群體的歷史中，各種突變都會以一定的機率傳到後代。當群體越小，各種突變的出現頻率就越不平均，會隨著世代傳遞而波動，這種現象稱為「遺傳漂變」。當群體小到一定程度，遺傳漂變的效應非常強烈，有些突變的頻率可能突然波動成零

而消失，這種現象叫做「瓶頸效應」，這就像是一群通過一個狹小的瓶口，只有非常少的人能過去，通過群體的多樣性必然會少於原來的，之後重新發展起來的群體就與原來很不一樣了（見下圖）。這種瓶頸可能是人類遷徙路徑上的一個地理障礙（如高山、海峽或沙漠，在下圖中以瓶子標示），也可能是使得人口大量減少的戰爭、瘟疫等。人類在瓶頸中經歷了人口極少的時期，使得遺傳多樣性大量喪失，直到人口增長，才重新開始累積基因組的多樣性。

人類遷徙過程中，各個地區之間都可能出現瓶頸效應，使得各個地區人群的基因組多樣性多少不同。比如美洲大部份地區的印第安人祖先從西伯利亞通過白令陸橋到達阿拉斯加以後，人口減少到了幾十個人，後來的北美和南美的主要印第安人群體，都是從這幾十個人發展而來的，這是人類遷徙中最著名的瓶頸效應例子。



人類在從非洲散播到世界各地時，經歷了不同的瓶頸，上圖的藍色瓶子標示就是人類在遷徙時遇到的瓶頸，所以人類在散播時，各地的族群會有差異（以不同顏色的人形表示）。物種遷徙時所出現的瓶頸效應，是只有少數個體能夠通過瓶頸，這些

個體所具有的遺傳多樣性不如原來的族群，即使後來數量增加，多



樣性仍不及原來的族群，例如上圖左邊的瓶子中，只有三個紅色的

人通過瓶頸，以此三人為基礎發展出的族群，也都只會是紅色的。

染色體是O3亞型，融入了南方的本地民族中，因此我們很難再分析人群最初在東南亞向東亞擴散時，是如何群體分化的。

瓶頸效應漸次發生

不過我們發現，有些民族幾乎沒有受到漢族南下的遺傳影響，保留著某支先民進入東亞時候的遺傳結構。在東南亞和東亞之間，九州大地最南端的海南島是群體融合的死角。海南島原住民黎族和仡隆人群體中，幾乎沒有O3型的Y染色體，而只有O1和O2兩種類型。在我們的祖先進入九州時，由於冰河時期海平面降低，海南島周圍的大陸棚露在海面之上，部份先民就藉由這一條捷徑來到了揚州。他們與從緬甸進入梁州地區、攜帶著O3和O2型Y染色體的群體不同，所以我們可以推測，當時人群一路是從越南進入揚州，另一路則是從緬甸進入梁州。隔離這兩路人的地理障礙很可能是中南半島的長山山脈。

Y染色體O2型在兩路人群中都出現了，是非常古老的亞型，可以追溯到印度半島的孟高棉族群祖先門達族群。孟高棉族群可能是整個東亞人群的起源，目前分佈於印度東北地區到馬來半島北部，最北是廣西雲南貴州交界地區的俵人，該處屬於梁州。孟高棉族群在古代的分佈可能更廣泛，包括雲南和貴州大部份地區，現在這一地區常見的Y染色體O2亞型很可能是他們留下的。

人群在穿過雲貴高原來到四川盆地後，群體的遺傳和文化特徵都發生變化，在四川盆地上留下了營盤山文化、寶墩文化、三星堆文化，很可能是緬彝語支傣僳族群的祖先。大約一萬年前，人群從四川北上，抵達甘青一帶的羌塘高原，古羌人在這裡孕育成長。「蜀道難」，人群通過蜀道的過程必然產生瓶頸效應，雍州人群的遺傳類型只是梁州的一小部份，Y染色體O2型就不再出現。所以在以後的羌、

人類散佈到東亞的兩次遷徙

從非洲向東亞的早期人類遷徙並不是一次完成的，至少有兩次主要的遷徙（圖中以不同顏色標示），構成了現在遠東人群的分佈格局。在印度洋沿岸以順時針遷徙的人群，至少在六萬年前就抵達了澳洲，但是今天東亞的大部份人群的歷史顯然要晚得多。根據到達東亞和太平洋地區的先後，一般把這兩批人稱為早亞洲人（深褐色）和晚亞洲人（銘黃色）人。過去的體質人類學分類上被稱為「棕色人種」或「澳大利亞人種」的澳大利亞、新幾內亞和美拉尼西亞土著，顯然屬於早亞洲人。



電腦繪圖：姚裕評

藏、漢諸族中，幾乎沒有O2的蹤影。

約7000-8000年前，部份人群又開始沿著渭河與黃河往東，來到雍州東部和豫州，發展出粟米農業。有人群陸陸續續加入這些農業社群，最終發展成華夏族。5000-6000年前，華夏族與羌族分道揚鑣，之後幾千年華夏族發展壯大，形成了現在世界人口最多的民族漢族。奇妙的是，在父系染色體方面，漢族還是保留著華夏族原初的Y染色體遺傳結構，只有南方少數分支受到其他民族融入。這一條沿著青藏高原東部邊緣從南向北的遷徙，成為東亞人群的主幹，而在當時人類大遷徙途中，這可能只是很小的分支。

三、四萬年前跨過長山的人數並不多，否則也不會形成東西如此大的差異。只是到達長山以東的東亞沿海地區，由於物產豐饒，人口增加，新的Y染色體突變也產生了。在兩、三萬年中，百越族群的人口持續增長，擴張到整個揚州地區。良渚文化是百越發展的巔峰，留下了大量精美的玉器。有些人群一直向北進入徐州、青州乃至東北的幽州。

地理不再阻絕民族融合

商朝滅亡以後，東夷族漸漸被華夏族同化，現在很難從人群中追尋東夷族的蹤跡。但是百越的O1型在東北幽州的民族中發現，南方的O2型在朝鮮出現，這都是百越族群等南方族群通過青州的證據。

百越民族特有的O1型Y染色體普遍出現在東南亞島嶼上的馬來民族中，根據分子生物學的研究結果推測，馬來民族起源於中國南部沿海，也就是百越分佈的揚州地區，經由越南沿海散佈到了整個東南亞島嶼和大部份的太平洋地區。

進入梁州地區的孟高棉族群祖先除了繼續向北到達雍州，另一批人向東越過巫山進入雲夢大澤所在的荊州。這群人是苗瑤民族的祖先，苗瑤民族除了帶有O2和O3亞型的Y染色體，還有一種比較特殊的

九州各族的文化遺跡與遺傳類型

州名	區域	族群	代表文化	Y染色體類型
梁州	秦嶺以南至川雲貴	孟高棉族群向藏緬族群過渡的人群	李家村文化	O2, O3
雍州	陝甘青藏	藏緬族群的古羌人	仰韶文化	O3a5
荊州	湖廣	苗瑤族群	大溪文化	O3a4
揚州	北越到蘇南	百越族	大坨坑文化、良渚文化	O1, O2
青兗徐州	蘇北和山東	東夷族	青蓮崗文化、大汶口文化、龍山文化	未知
豫州	河南	華族、青州的夷族、荊州的蠻族	裴李崗文化	O3
冀州	河北	豫州、青州向幽州過渡的區域	磁山文化	O1
幽州	燕山以北	通古斯族群、古西伯利亞族群	紅山文化	O1, O2, O3
并州	山西與蒙古部份地區	匈奴、突厥、烏拉爾族群	陶寺龍山文化、細石器文化	O3a5, N

O3a4小亞型，這種類型也出現在東南亞的某些孟高棉民族中，STR的多樣性分析也證明苗瑤民族確實起源於孟高棉族群。傳說中苗族曾經往北方擴展過，但是我們並沒有發現任何證據。

冀州是南方來的兩批人群交匯處，西路和東路人群帶來的許多遺傳特徵在這裡融合，並向幽州擴散。人群的交融帶來文化的碰撞，9000多年前在遼河流域迸發出了文明的火花，玉器製作技術的發展和宗教思想的產生，是這個地區對九州文明最重要的貢獻。并州地區的居民主要源自雍州，在瓶頸效應的影響下，他們的遺傳結構繼續發生變化，原本從梁州到雍州遷徙人群非常罕見的N型，在這裡越來越多，並擴張到西伯利亞西部以及歐洲。

從虞夏時期開始，隨著文明的發展，地理的阻隔已經不能阻礙族群的交流，4000多年前的龍山文化超越了一州的區域，在各地發展出了各種分支文化，並且影響了邊遠地區，各地的文化交流融合，成為中華文明的基礎。經過4000多年的發展和交流，九州大地的人群已經你中有我、我中有你，難分彼此。 SA

延伸閱讀

Genotype, haplotype and copy-number variation in worldwide human populations. Jakobsson M, et al.(2008) in *Nature* 451(7181):998-1003.

Genetic absolute dating based on microsatellites and the origin of modern humans. Goldstein DB, Ruiz Linares A, Cavalli-Sforza LL, Feldman MW (1995) in *Proc Natl Acad Sci U S A* 92(15):6723-7.

Natives or immigrants: modern human origin in east Asia. Jin L, Su B (2000) *Nat Rev Genet* 1(2):126-33.

Y chromosome of prehistoric people along the Yantze River. Jin L, et al. (2007) *Hum Genet* (122):383-388.

Paternal genetic affinity between western Austronesians and Diac population. Jin L, et al. (2008) *BMC Evol. Biol*(8)146.

Paternal genetic structure of Hainan aborigines isolated at the entrance to East Asia. Li D, Li H, Ou C, Lu Y, Sun Y, Yang B, Qin Z, Zhou Z, Li S, Jin L (2008) in *PLoS ONE* 3(5):e2168.