


# 版納植物園基於全基因組數據解析新型冠狀病毒的演化和傳播

時間：2020-02-20    來源：科技外事處    瀏覽次數：    作者：郁文彬     打印    字體：大 中 小 【關閉】

2019年12月在湖北武漢爆發了一種新型冠狀病毒（ $\text{2019-nCoV}$ ）所致的肺炎（現稱 $\text{COVID-19}$ ），新型冠狀病毒爆發已兩個多月，確定華南海鮮市場是不是唯一的發源地，對於尋找病毒的來源，以及確定中間宿主，對疫情的控制和避免再次爆發具有至關重要的意義。中國科學院西雙版納熱帶植物園聯合華南農業大學和北京腦科中心的科研人員一起收集了全世界各領域共享到GISAID數據庫中覆蓋了四大洲12個國家的93個新型冠狀病毒樣本的基因組數據（截止2月12日），通過全基因組數據解析，追溯傳染源及擴散路徑。研究發現，收到的93個樣本包含58種單倍型，可以歸納為五組（圖1），包括3個古老超級傳播者單倍型（ $\text{H1}$ 、 $\text{H3}$ 和 $\text{H13}$ ）和2個新的超級傳播者單倍型（ $\text{H56}$ 和 $\text{H2}$ ）；華南海鮮市場的新型冠狀病毒是從其他地方傳入進來，在市場中發生快速傳播蔓延到市場之外；同時，現擴散的病例至少來自於3個途徑。新型冠狀病毒在2月12日之前發生過2次明顯的種群擴張（分別是12月8日和1月6日）。

## 華南海鮮市場的新型冠狀病毒是從其他地方傳入的

基於120個變異位點得到58種單倍型（基因類型），單倍型演化關係顯示，單倍型 $\text{H13}$ 和 $\text{H38}$ 是比較「古老的」單倍型，通過一個中間載體（ $\text{H1}$ ，可能是一個祖先單倍型，可能是來自中間宿主或者「零號病人」）與蝙蝠冠狀病毒 $\text{RaTG13}$ 關聯，並通過單倍型 $\text{H3}$ 衍生出了單倍型 $\text{H1}$ 。與華南海鮮市場有關聯的患者樣品單倍型都是 $\text{H1}$ 及其衍生的單倍型 $\text{H2}$ 、 $\text{H8}$ 、 $\text{H12}$ （圖1*a*），而一份武漢樣品單倍型 $\text{H3}$ 與華南海鮮市場無關。可見，華南海鮮市場的新型冠狀病毒是從其他地方傳入進來，在市場中發生快速傳播蔓延到市場之外。另外，根據病患發病時間記錄和種群擴張時間推斷，也印證了華南海鮮市場不是病毒發源地的推論。

對「古老的」單倍型 $\text{H13}$ 和 $\text{H38}$ 的病毒樣品溯源發現分別是來自深圳的病患（廣東首例）和美國華盛頓州的病患（美國首例）。他們的旅行記錄表明應該都是2019年12月底至2020年1月初在武漢探親期間被感染的。現有武漢樣本中沒有檢測到 $\text{H13}$ 和 $\text{H38}$ 單倍型，可能是因為現有樣品主要采自幾家定點醫院，而且樣品採集時間侷限於2019年12月24日和2020年1月5日。如果能在武漢其他醫院早期的病患檢測到這兩種單倍型，將對於尋找病毒來源非常有幫助。

## 新型冠狀病毒在2月12日之前發生過2次明顯的種群擴張

根據新型冠狀病毒基因組數據推算1月之前的種群擴張發生時間是12月8日，該結果暗示病毒可能在12月初，甚至11月下旬即已經開始有人際傳播，隨後在華南海鮮市場加快了人際傳播（圖2）。研究推算2月份之前的種群擴張時間在1月6日，這個可能與元旦假期有關聯。需要指出，這一天國家疾控中心發佈了2級應急響應。當時的預警起到了一些警示作用，公眾活動和出行都有所減少。如果當時的警示能引起大眾更廣泛的重視，那麼1月份中下旬向全國和全球蔓延的病例會有所降低。研究人員進一步確認我國其他9個省區和其他11個國家的感染病例基本都是從武漢直接或者間接輸入而來。

## 現擴散的病例至少來自於3個途徑

為了能夠細分來源，研究人員將58種單倍型分成了五組（圖1），採用標準是3個中心（古老超級傳播者）單倍型（ $\text{H1}$ 、 $\text{H3}$ 和 $\text{H13}$ ）和2個新的超級傳播者單倍型（ $\text{H56}$ 和 $\text{H2}$ ）。以此鑑別出廣東的病毒可能有三個來源，重慶和台灣的病毒有兩個來源。其中，廣東深圳一家人在早期就通過人傳人進行了傳播。有較多樣本的澳大利亞、法國、日本和美國，他們的患者感染源至少有兩個，尤其是美國包括了五個來源。非常值得關注的是 $\text{H56}$ 這個超級傳播者單倍型，它同時是澳大利亞、法國和美國，以及我國台灣患者的傳染源。其他國家患者因為樣品比較少，大多數的來源比較單一，他們除了是武漢旅遊輸入或在武漢感染外，有一些人可能是在廣東、新加坡等地被感染。

## 新型冠狀病毒基因組尚未發生重組事件

研究人員發現新型冠狀病毒基因組沒有發生重組事件，93個基因組之間有120核苷酸發生了突變（0.41%序列長度），並均勻分散在10個編碼區（ $\chi^2=1.958$ ,  $\text{df}=9$ ,  $P=0.99$ ）。120個突變的核苷酸關聯了119個氨基酸密碼子，其中79個密碼子 $\sim 65.83\%$ 改變了氨基酸類型，並有42個 $\sim 53.17\%$ 氨基酸理化性質都被改變（圖3）。這些氨基酸類型以及理化性質改變是否會影響新型冠狀病毒的活性暫不清楚，需要其他蛋白組學和結構生物學方面的專業人士進行驗證。本研究是版納植物園綜合保護中心生物多樣性研究組的科研人員利用其在系統與演化領域的專長開展的，本研究提到單倍型演化關係分析方法可以結合到傳染病學研究中，對於尋找傳染源，以及精確的傳播和擴散方向能提供非常重要的信息。

相關研究結果已提交到中國科學院科技論文預發佈平台 $\text{http://arxiv.org/abs/2002.00033}$ ，可供疾控防疫相關部門參考，並接受專家學者的交流和討論。

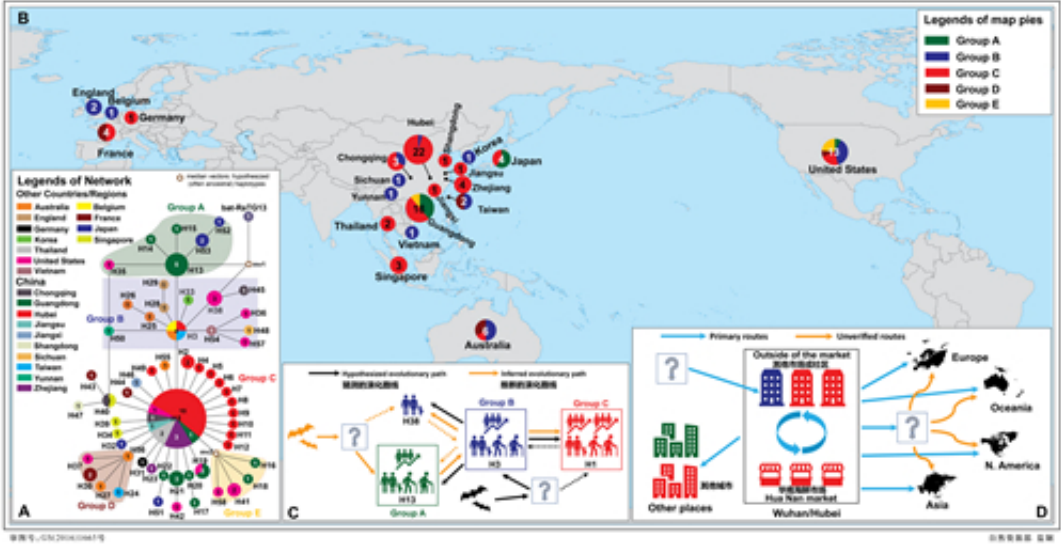


圖1. 新型冠狀病毒58種單倍型的演化關係和地理分佈格局（ $\tilde{I}$ ， $X$ ），單倍型之間的可能演化關係（ $\S$ ），以及新型冠狀病毒的可能傳播和擴散路線。圖1中 $\tilde{I}$ 和 $X$ 圓圈中的數據是樣本數量。

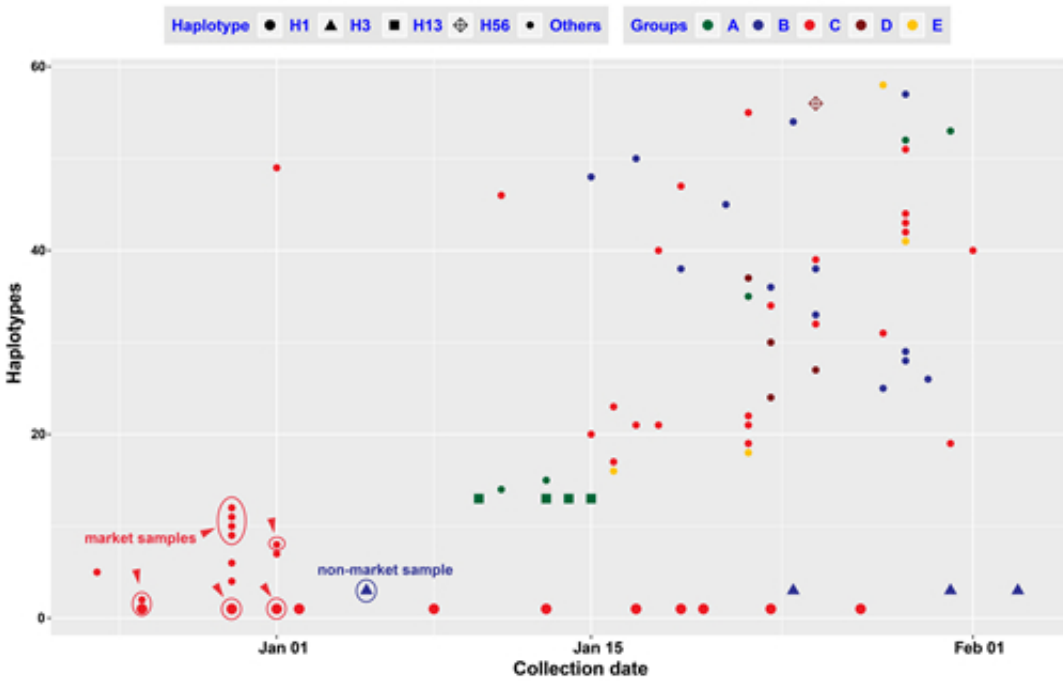


圖2. 新型冠狀病毒單倍型的樣品採集時間情況。紅色圈的樣品是確認與華南海鮮市場有關；藍色圈的樣本確認與海鮮市場無關。

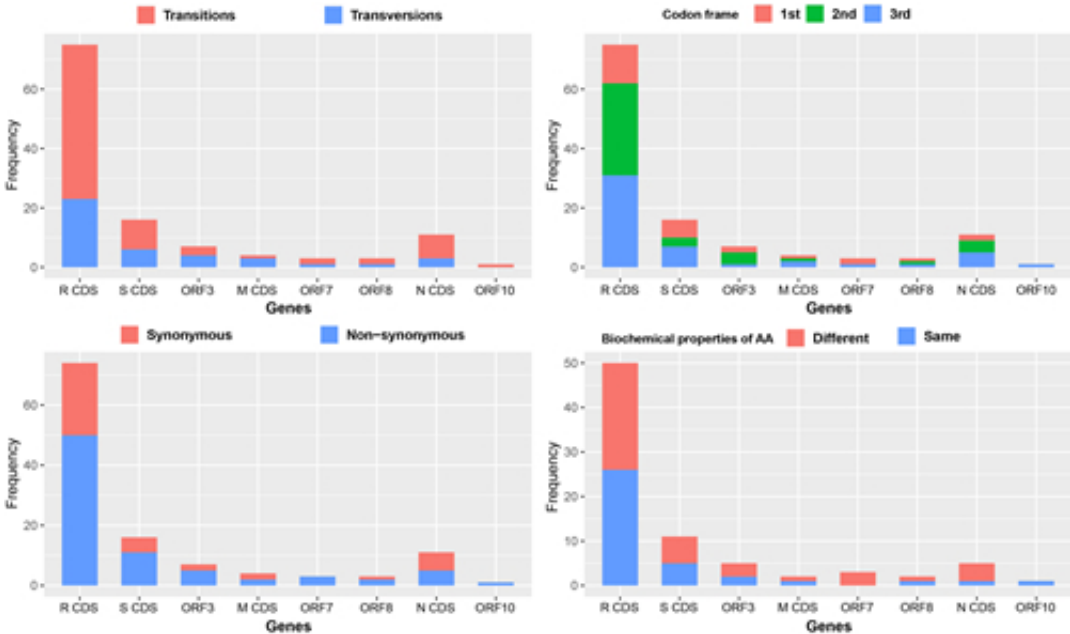


圖3. 120個變異位點在8個編碼區的分佈情況。統計類型包括替換或替代（左上角），密碼子位置1-3（右上角），同義突變或非同義突變（左下角），和氨基酸性質（右下角）。