

大數據時代，數字信息的有效分析和研判為決策者提供了詳實的依據。如今，對於疫情的防控也離不開大數據的測算與助力。作為國家重大科技基礎設施，在國家超級計算深圳中心的超級計算機（港稱「超級電腦」），正運用超強算力篩選能抑制病毒的小分子藥物，助力研製戰疫彈「藥」，支撐各類研究抗擊疫情，為最終消滅疫情鑄造金鑰匙。

大公報記者 郭若溪

新冠肺炎疫情來勢洶洶，1月24日除夕夜，中國疾病預防控制中心對外發布，中國成功分離了第一株病毒毒種信息。從事藥學研究的中山大學藥學院羅海彬教授異常興奮，因為掌握了毒種信息就意味着藥物篩選有了方向。當晚，他立即撥通了深圳超算中心的電話，提出需要超算資源完成抵抗新冠肺炎的藥物篩選和病毒突變預測。

在隨後的1個小時內，深圳超算中心迅速協調其他用戶，在取得用戶理解後暫停部分計算，優先調度出400個節點的資源，為疫情讓出緊急通道，同時與合作夥伴快速協商後，調度了曙光昆山的DCU資源和商湯公司的GPU資源。天津超算、騰訊公司和算力地球紛紛響應，展開了一場科學與時間的賽跑。最終為羅海彬團隊開通了高性能計算帳號，免費提供所有計算資源，藥物篩選作業即刻進行。

### 兩月工夫 兩日搞掂

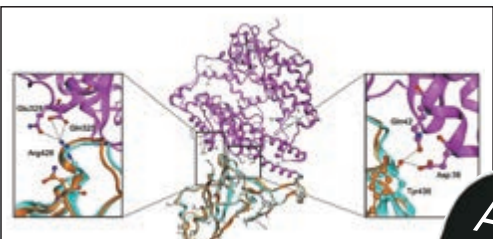
「因為我們的藥物有很多，如果用自己的電腦，可能要篩選兩個月。有了深圳超算，可能有望在兩到五天內做完這些事情，所以這是一個非常重要的一個支持。目前，初步研究成果已遞交廣東省科技廳和科技部。」羅海彬表示。

超級計算機之所以超級，在於它異常強大且高效的數據處理能力。在藥物篩選的研究過程中，就需要大量的超級計算資源幫助提高速度。「我們可以把藥物篩選的過程，看作是從成千上萬有可能的鑰匙，去找到那把正確的鑰匙來打開這把鎖。」國家超級計算深圳中心主任馮聖中用鎖和鑰匙的比喻，說明藥物篩選的原理。

那把「正確的鑰匙」，就是能夠與病毒蛋白質結合的小分子。馮聖中介紹，當小分子結合到病毒蛋白質的功能位點，病毒蛋白質就不能和它真正的底物結合，從而抑制病毒蛋白質的活性。每個小分子要通過各個角度與蛋白質結合，結合之後釋放能量越多，就代表結合越穩定，這個小分子就越有可能做成藥物。另外，運用分子動力學，也能快速分析出病毒入侵關鍵蛋白的分子結構機理，加速疫苗研製。

### 抗疫優先 免費支持

2月3日，深圳超算中心又接到了技術協助請求，深圳大學高等研究院請求下載數據集和數據預處理，用於新冠肺炎的藥物篩選工作。據了解，目前已經有近百家科研機構正在使用超算中心的資源進行疫情相關的科研工作，深圳超算優先調度出數以千萬計的機時，來服務於疫情相關的研究工作，給疫情趨勢預測、病毒基因測序、藥物篩選、病毒疫苗研製等方面研究提供強大的算力支撐。「目前，對疫情研究明確提出需求的課題組，深圳超算提供免費的支持，支持他們盡快能拿出一些好的結果來，阻止疫情的蔓延，為國家和社會做貢獻。」馮聖中說。



▲新冠病毒S蛋白質與人體ACE2分子的3維結構



▲參與開發平台的顧斐博士

#### 優點

**避免漏檢**  
相比核酸檢測主要檢測病毒序列的兩個片段，應用AI算法的全基因組檢測分析平台基本能覆蓋到病毒的全基因組，可以避免假陽性和假陰性等漏檢的問題。

#### 發現變異

由於對病毒進行全基因組檢測，可以提前發現病毒變異的情況。

#### 快速檢測

該分析平台配備了蛋白二級結構、三維結構的分析功能，可以幫助疾控人員快速檢測到病毒，並了解病毒的結構，進一步幫助他們提升分析和研究能力。

#### 缺點

**操作複雜**  
目前該平台更多還是由開發人員來使用，醫護人員操作起來比較複雜，還需要在開發人員指導下投入臨床使用。

大公報整理

# 超級電腦研製戰疫彈「藥」

## 篩選病毒抑制物 瓦解敵襲金鑰匙



### 超級電腦 抗疫應用

#### 預測疫情趨勢

- 1 通過人口基數、流動趨勢、時空分布等維度建立數學模型



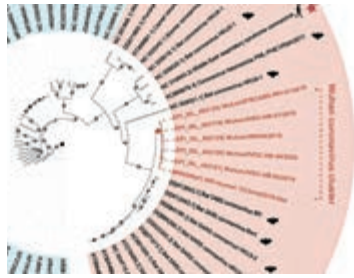
▲疫情傳播模型總體框架 網絡圖片

- 2 利用數學模型，進行超級計算的模擬仿真

- 3 預測新冠病毒的可能感染規模和傳播風險

#### 定位致病元兇

- 1 對病毒進行基因測序，通過基因相似性分析構建基因演化樹



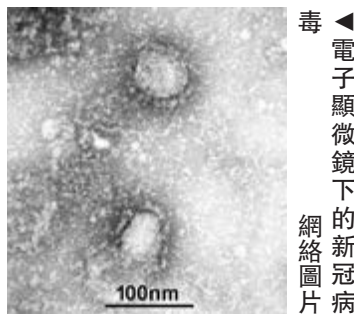
▲新冠病毒演化樹 網絡圖片

- 2 利用測序結果，輔助超級計算資源，採用生物信息學的方法來比對和查找

- 3 定位致病元兇，並揭示病毒致病機理

#### 加速藥物篩選

- 1 找到能夠讓病毒蛋白質分子失效的小分子

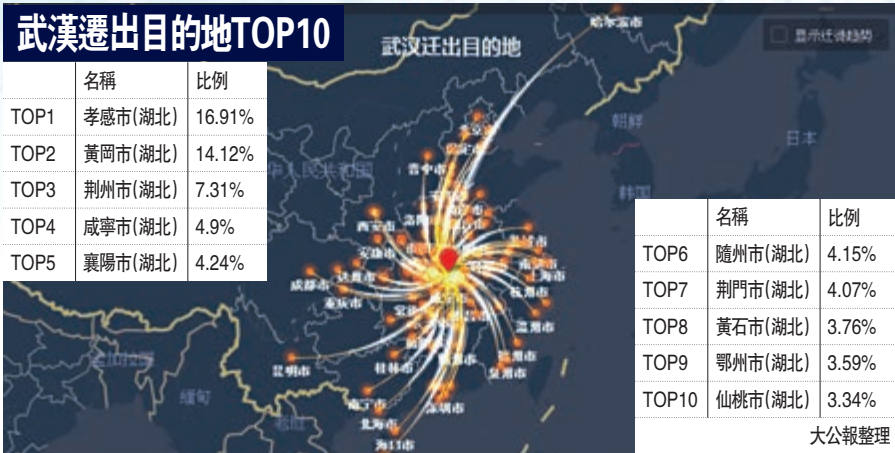


▲電子顯微鏡下的新冠病毒 網絡圖片

- 2 引入超級計算資源，將每個小分子通過各個角度與蛋白質結合

- 3 觀察能量釋放情況，加速藥物篩選

▲國家超級計算深圳中心正在使出超強算力，支撐各類研究抗擊疫情。圖為科研人員在中心內工作 網絡圖片



▲今年春運期間有500萬人離開武漢、900萬人留漢過年。武漢人口的遷徙大數據與新冠肺炎防控息息相關。據百度地圖慧眼百度遷徙數據顯示，至武漢封城當日（1月23日），從武漢向外遷徙人口前十目的地均為湖北省內城市

## 高效基因測序 鎖定致病元兇

數據匯集只是第一步，更重要的是透過數據研究疫情發展趨勢、分析病毒傳播路線、明確高危人群狀況，為下一步決策部署提供支撐。

馮聖中表示，基於疫情時空大數據，可以利用數學模型對新冠病毒的可能感染規模和傳播風險等進行預測，為評估防控措施的有效性以及醫療、衛生資源的分配提供重要的參考價值。

透過海量的數據，科學家們正通



▲國家超級計算深圳中心主理人馮聖中在接受採訪時表示，中心正在進行巡檢 受訪者供圖

過人口基數、流動趨勢、時空分布等維度，深入研究疫情趨勢分析模型，預測新增確診人數拐點並形成分析報告，支撐國家衛健委疫情防控工作，為相關機構重點管控提供數據支撐。

此外，超級計算還能加速病毒基因測序，助力病毒突變預測。對病毒進行基因測序，通過基因相似性分析構建基因演化樹，有助於定位致病元兇，並揭示病毒致病機理。「利用測序結果，採用生物信息學的方法來比對和查找，就需要超級計算資源。」馮聖中說。

在測序和序列分析初期，科學家發現新冠病毒和以前的冠狀病毒相似度並不是特別高（87%），進一步通過基因家族分析和蛋白質三維結構分析，實現了致病元兇的準確定位，超級計算在其中發揮了不可或缺的重要支撐作用。

記者 郭若溪

## 二千巡檢指標 機房穩如泰山

深圳超算中心位於西麗大學城，而當中的系統運行監控中心和中心機房是深圳超算中心內最神秘的兩個地方，是讓超算中心正常運轉的重要核心。據系統運行部部長陳開渠介紹，為保障各項作業，特別是用於藥物篩選等研究的作業安全平穩運行，今年春節期間，中心每天有16人在崗，實行7×24小時值班，另有6人遠程值班。

陳開渠告訴記者，一旦出現故障告警立馬就要處理，1分鐘之內必須響應。在系統運行監控中心值班的同時

，超算中心每天還會安排4個崗，平均兩個小時就對基礎設施和中心機房進行一次巡檢，機房是重點。「巡檢包括人員的巡檢和監控系統的輔助巡檢，前者主要是看輔機系統是否工作正常，是否有高溫狀態，是否存在安全隱患等內容，總計超過2000項巡檢指標。新春期間每天有6000多個作業在進行，確保超級計算機正常運行，是新春期間值班和日常巡檢要解決的重要內容。」陳開渠說。

記者 郭若溪

## AI分析蛋白結構 動態追蹤病毒變異

近日，浙江省疾控中心上線自動化的全基因組檢測分析平台，利用阿里達摩院研發的AI算法，將原來數小時的疑似病例基因分析縮短至半小時，並能精準檢測出病毒的變異情況。

「目前，主流檢測手段為核酸檢測方法。」達摩院算法專家顧斐博士指出，這項技術實際檢測時間相對較長，而且只能檢測到病毒基因的局部。由於病毒存在變異可能，因此核酸檢測可能會出現假陰性、假陽性的情況。

### 把關後期防疫 提升確診效率

不同於核酸檢測，新的分析平台是對樣本進行全基因組序列分析比對，能

夠有效防止漏檢。此外，平台有效縮短了全基因測序的時間。「越往後走，越需要全基因檢測，因為後期防疫的核心是防止病毒變異。」顧斐說。



▲杭州杰毅生物開發了全自動高通量測序建庫儀 網絡圖片

顧斐介紹，在整個平台中，合作夥伴杰毅生物開發了全自動高通量測序建庫儀，把常規人工需要12小時的工作縮短到2個小時。

「每次測序過程中產生的海量基因數據，則交由達摩院AI算法進行分析。」顧斐說，他們將病毒基因分析的速度由數小時縮短到半小時。同時，病毒序列拼接的速度由30分鐘-1小時縮短到15-30分鐘。

由此，AI算法可以在15-30分鐘內預測病毒蛋白二級結構，從而幫助醫護人員檢測到病毒全貌，使得變異的病毒也能精準檢測，大幅提升確診效率。

記者 俞畫