



莊庭瑞
中央研究院資訊科學研究所副研究員，
合聘於資訊科技創新研究中心以及
人文社會科學研究中心。

基因資料真能個人自主嗎？

基因資料互有關聯，使用上如何兼顧群體福祉與個體權利，是道難題。

撰文／莊庭瑞

去年4月，美國沙加緬度的檢察官起訴一名72歲的退休警察，認為他是1976年起數十件殺人與強暴案的凶嫌。這名嫌犯的DNA與1980年一名受害者遺體上殘留的檢體相同，然而他並未做過基因型鑑定，茫茫人海之中，檢調單位是如何找到他呢？

原來檢調人員拿了殘留檢體的DNA樣本定序分析，送到由使用者自願提供的基因型鑑定彙集網站GEDmatch，經由親緣關係比對找到殘留檢體所有人的遠房血親。GEDmatch目前約有100萬名使用者，網站並不提供基因型鑑定服務，而是讓眾人上傳既有的基因型鑑定資料。雖然不能用這些資料反向辨識基因型所有人，但使用者可以透過在網站註冊的信箱彼此聯絡。

經由這個網站，檢調人員找到十多位嫌犯可能的遠房血親，取得他們的身分，再經由公開資料追溯他們19世紀的共同祖先，在後代數千人中，依犯罪者的背景剝繭抽絲去除不可能的人，最後剩下兩位，被逮捕的退休警察是其一。經由網站使用者的集體「協助」（大部份人不知情），懸宕40多年的重案終於找到凶嫌。

然而此舉也引發爭議，GEDmatch的創辦人說他事先不知道檢調人員使用網站的方法，而且檢調人員並沒有（也不需要）取得法院許可，就註冊進行資料搜索。事實上這個網站的目的，就是讓大眾上傳個人的基因變異資料，經由比對與群力互助來尋找失散的家人或建立家族系譜。不過上傳者可能不是當事人（例如檢調人員就上傳了檢體的資料），而且血親之間的基因變異有部份特徵相同，所以上傳的基因資料也會揭露血親的基因特徵，引發使用疑慮及倫理爭議。

舉例來說，父親上傳小孩的基因資料，是否需要母親同意？小孩又如何表示意見？許多遠房親戚根本沒見過面，他們上傳自身的基因資料，可不會過問你的意見。尤其是現在基因檢測非常簡便，許多檢測服務直接訴諸消費者，產生的資料也能讓受檢者另做使用，後續

影響變得更加廣泛。

經由基因變異特徵比對，成功緝捕連續殺人犯，在遺傳系譜學上並不算驚天動地的發現。美國哥倫比亞大學的資訊科學家厄立奇（Yaniv Erlich）等人在去年11月發表論文，使用MyHeritage網站上去除個人識別的128萬人基因變異特徵資料，在特定條件下進行「遠親搜尋」實驗，發現有60%的機會可找到同高祖父母的堂表親或更近的血親，有15%的機會可找到同曾祖父母的堂表親或更近的血親。資料彙集的人口越多，遠親搜尋效果就越好，這篇研究使用的資料人數佔美國總人口的0.4%，已屬難得。不過MyHeritage涵蓋的人口不算均勻，資料大部份是歐裔（86%），其他族裔透過這類服務尋親，命中機率應該較小。

消費者型態的個人基因資料如此普及，且被大量彙集與二次使用，讓我們對個人資料的意涵以及使用規範有了新的省思。簡單來說：個人基因資料的使用，可全然只由個人決定？目前討論個人資料，多是從個人權利出發，重視「個人自主」，亦即資料的蒐集處理需要當事人「知情同意」；而個人資料彙集處理後能否釋出另行使用，則多以「去個人識別」為依據。但在實際上，大眾往往只能接受資訊服務廠商定型化的使用條款，概括同意廠商鉅細靡遺地蒐集並處理自己的活動資料，很難稱得上個人自主。

就基因資料來說，個人自主陳意過高，在現實上很難實踐。使用GEDmatch協助偵查殺人犯，也顯示去個人識別的基因資料若彙集使用並參照外部資訊，還是可以「重新識別」。劍有兩刃，這種網站立意良善，但也可能傷人。個人資料的蒐集處理與二次使用，個人無法完全自主，在追求群體福祉與維護個體權利之間，如何規範以取得平衡，著實是艱難的議題。SA

感謝中央研究院生物多樣性研究中心呂美暉、資訊科學研究所施純傑、蔡懷寬對文稿的建議。