

出國報告（出國類別：進修）

美國遺傳醫學會 American Society of Human  
Genetics 年會、國家衛生研究院 National  
Institute of Health、賓州大學 University of  
Pennsylvania 參訪心得報告

服務機關：臺中榮民總醫院

姓名職稱：蕭自宏研究員

派赴國家/地區：美國/華盛頓特區與費城

出國期間：2023 年 10 月 29 日至 2023 年 11 月 12 日

報告日期：2023 年 11 月 28 日

美國遺傳醫學會 American Society of Human Genetics 年會、國家衛生研究院  
National Institute of Health、賓州大學 University of Pennsylvania 參訪心得報告

目次

摘要	6
目的	7
過程	8
心得	13
建議	14

## 摘要

在 2023 年美國遺傳醫學會年會上，發表了以 polygenic risk score 和頸動脈超音波為基礎的缺血性中風風險分層研究。該研究利用不同 polygenic risk score 探討中風風險，結合頸動脈超音波數據，旨在建立基因風險評估和頸動脈超音波監控的健康預防模式。另外，會議中包括基因定序 VUS 位點的生物資訊分析演算法、使用 TOPMed 資料插補低發生率罕見變異、Mosaic mutation 研究、單細胞定序資料的軌跡分析以及大型基因體計畫參與和價值回饋等相關議題都給我很高的學習與視野的提升。參加本次會議深化對基因遺傳最新發展的理解，也拓展了精準醫學的研究領域。

在參訪 NIH 和 Upenn 期間，我們深入了解了各實驗室的先進技術和研究方向。在 NIH，我們參觀了休閒治療中心和定序與單細胞核心實驗室，特別對於定序品質控制等細節進行深入討論。在 Upenn，我們拜訪了 Penn Neurodegeneration Genomics Center，交流其在 Alzheimer's disease 基因資料研究上的領先研究。同時，與 CHOP 的實驗室交流，掌握最新的單細胞定序技術的技術細節，也了解他們對於空間基因體領域的發展進度。這次參訪不僅技術層面有所提升，更促進了與國際頂尖機構的合作機會，為未來本院精準醫學及基因體研究提供了豐富的參考。

**關鍵字：**基因風險分數、大型基因體計畫、單細胞定序

## 目的

本次參訪同時旨在技術的精進與拓展與國際研究界的合作機會，特別是在精準醫學和遺傳疾病研究領域。藉由參訪提升本院基因體研究以及精準醫學研究的能量。此外，了解頂尖研究機構對於精準醫學與基因體學的研究方向，可作為本院未來發展的借鏡。

## 2023 年美國遺傳醫學會年會

本次在 2023 年美國遺傳醫學會年會上，團隊張貼了一張海報展示了最新的研究成果，海報題目為” Risk stratification for different subtypes of ischemic stroke based on polygenic risk score and carotid intima-media”，本次海報揭露了團隊運用了各種不同的 polygenic risk score 去探討對於各類型以及不同年齡發生中風的風險相關性，並結合了頸動脈超音波的數值進行探討，希望可以建立一個運用基因進行風險評估，對於高風險族群運用頸動脈超音波進行動態監控之精準健康預防模式。在海報的張貼過程中與很多與會的學者討論，也得到很多反饋意見，之後將用於本研究之精進。

此次年會並聆聽多場演講，得到很多啟發，相關的心得重點如下：

### 1. 對於基因定序 VUS 位點的生物資訊分析演算法

對於基因體定序的結果常常得大量的 VUS 位點(Variant of Uncertain Significance)，此次年會有多個生物資訊預測演算法用來探討這些還不清楚地位點，例如"Loss-of-function Observed/Expected Upper-bound Fraction (LOEUF) score"是一個用於評估基因變異功能影響的分數。這個分數主要用於衡量基因的突變對其功能造成損害的可能性。LOEUF 分數是基於觀察到的突變數據與期望的突變數據之比例。期望的突變數據通常是基於某種模型或參考數據集的期望突變頻率。這種比例可以提供有關特定基因突變是否比預期更可能導致基因損害的相關資訊。除此之外還有 ESM1b score，是一個擁有 650 百萬參數的蛋白質語言模型，來預測人類基因組中所有約 450 百萬可能的錯義變異效應，這個方法除了單一位點變異之外還可以使用於 in-frame indels 與 stop-gains。與會的學者運用 UK biobank 等大型基因體資料作驗證證實這兩種演算法的優異性能。

### 2. 使用 TOPMed 的全基因定序資料作為模板進行低發生率罕見變異的資料差補

過往基因資料的插補只能針對發生率 1% 以上的位點，1% 以下的位點就會有很高的差補錯誤。目前發展的最新進展是使用 TOPMed 的資料建立參考模板，結合全基因組序列數據，在 52,658 例 2 型糖尿病患者和 381,683 例對照組中進行罕見變異的插值，可以準確地識別了新的罕見變異相關性並為單基因糖尿病基因的致病性譜提供了相關資訊。過去基因體關聯研究 (GWAS) 主要集中在常見變異，由於整體基因體序列 (WGS) 數據研究的樣本數有限，以及難以準確插值罕見變異，限制了在複雜疾病中識別效應較大的罕見變異的能力。為彌補這一不足，研究使用 TOPMed 參考面板對 UK Biobank、MGB Biobank 和 GERA 等數據進行插值，並結合 All of Us Research Program 的 WGS 數據進行最大的多祖源 GWAS 草案分析，包括罕見變異 (MAF 0.005%-0.1%)。結果顯示，他們發現 HNF4A 基因的一個罕見的錯義變異 (MAF=0.007%)，對 2 型糖尿病的風險增加了 8 倍 (OR=8.2, 95% CI [4.6-14.0], p=1.08×10<sup>-13</sup>)，這是已知的 MODY (成年期發病的糖尿病) 突變，具有中等的穿透力。結果顯示目

前基因資料插補已經可以用來探討罕見變異。

### 3. Mosaic mutation 的相關研究

Mosaic mutation 指的是在體細胞中發生的基因突變，這些突變只影響個體的一部分細胞而不是所有細胞。同一個個體內的某些細胞擁有不同的基因型。這可能是由於胚胎發育過程中的細胞分裂錯誤或其他機制造成的。除了傳統的遺傳變異之外，這次 Mosaic mutation 也受到一些學者的關注。Mosaic mutation 可能造成組織功能下降、免疫系統變化、細胞增生和凋亡、DNA 損傷累積、以及癌症風險。mosaic mutation 的研究更進一步地深化對基因遺傳的理解，還可能為許多疾病的原因帶來新的可能相關性。這是一個相對新的研究領域，持續關注其潛在的應用價值。

### 4. 單細胞定序 (single cell sequencing) 資料的軌跡分析

單細胞定序 (single cell sequencing) 技術被廣泛應用於解析複雜組織和細胞群體中個別細胞的基因表達和遺傳變異。在單細胞測序數據中，軌跡分析

(Trajectory Analysis) 是一種用於研究細胞動態過程的方法。軌跡分析能夠提供不同細胞類型轉變為另一種類型之間的相關資訊。軌跡分析主要能提供偽時間 (Pseudotime)：偽時間是軌跡分析中的一個概念，表示細胞在變化過程中的相對位置。通過偽時間分析，可以將細胞按照其在變化軌跡上的相對順序進行排序，從而揭示細胞分化的時間序列。分支分析：有些細胞的表達圖譜呈現分支結構，即細胞可以選擇不同的分析路徑。軌跡分析可以幫助識別和描述這種分支結構，從而理解細胞在基因表達圖譜過程中的決策點。在單細胞定序技術的推動下，軌跡分析提供了更深入理解細胞動態過程的分析方式。

### 5 大型基因體計畫的 participant engagement 以及 return of the value

幾個國際上大型的基因體計畫一起分享他們如何進行 participant engagement 以及 return of the value。包含美國的 all of us 跟 million veterans program、英國的 UK biobank 以及商業公司 23&me。整個分享強調從研究轉換到臨床場域需要一個 confirmation test，以及精準醫學諮詢的重要性，讓研究的參與者可以因為基因體計畫的參與得到相關精準健康的照護。

在與會期間，我收取了在遺傳學以精準醫學最新的發展。包括多個大型生物基因資料庫研究，如 All of US、UK Biobank、Genomic England、Million Veterans Program、以及 23 & Me 的最新發展。也觀察到其他研究單位如何使用新穎的基因體技術例如單細胞定序、空間轉錄體學等等探索精準醫學更嶄新的領域。

這些領域的最新進展不僅擴展了我的學術視野，同時為未來的研究方向提供了寶貴的參考。在會議中，我有機會與多位台灣籍學者認識。我們討論了當前台灣精準醫學所面臨的挑戰和機遇，同時探討了未來合作的多種可能性。這次深度交流對於加強國際學術合作和交流具有重要意義。

# NIH 以及 Upenn 參訪

在 NIH，我們參訪了成人與小兒休閒治療中心、NIA CARD 實驗室及定序與單細胞核心實驗室，在 recreation therapy 部分，因為非我的專長，我抱著學習的心態聆聽他們的介紹，對於他們設計整個 program 的嚴謹跟完備程度感到敬佩。在定序與單細胞核心實驗室，我們與實驗室負責人討論。我們參訪的定序與單細胞核心實驗室是 NIH 內部定序量最大的核心實驗室，我們就整個實驗細節跟主持人討論，請教對方一些品質控制的細節，例如他們在進行大批次定序的時候會先使用小型定序儀進行小批次測試，以確定混樣上機的時候每個樣品都可以得到相同的定序量，以避免某幾個樣品定序量不足而須重新上機。這些方式可作為本院精準實驗室定序品質控制程序，因應往後的業務增加。

我們參訪了 Roy Blunt Center for Alzheimer's Disease and Related Dementias Research(CARD)。CARD 實驗室是 NIA 設立專門研究 Alzheimer disease & Parkinson disease 的新實驗室，實驗室的設計概念是 cooperation。Wet lab 跟 dye lab 的空間配置比重來到一比一。採取 google style 的開放 Design，空間的設計強調 cooperation，可以隨時隨地的討論。裡面配備了基因體學、蛋白質體學跟細胞生物學所有大型自動化儀器，包含單細胞定序儀、空間基因體相關設備正在裝機中、質譜儀、自動化神經細胞培養系統、自動化細胞共扼焦顯微鏡，以為我特別注意到他們同時配置了三台超高通量的 nanopore 第三代定序儀。整個實驗室的設計朝實驗結果數據化，進而運用數據分析進行探索。生物實驗結果大量數據化並結合醫療資訊數據化，利用資料科學驅動研究進行應該為這幾年醫學研究的嶄新型態，朝 Digital medicine 的方向前進。

我們與 CARD 裏頭使用 LLM 進行研究探索的研究人員進行對話，他們在進行一個 general CDE(common data element)的計畫，對於醫療數據裏頭同義的不同醫療字彙，使用 prompt 的技巧，驅動 LLM 去認識所有的同義字，整合成統一可分析的數據。最近很多研究學者使用 LLM 識別文字的資料的能力去整理很多原本需要人工處理的部分。

除了有充足先進的儀器設備之外，CARD 配置高比重的資料科學研究人員，並強調實驗人員與分析人員的合作，為此一實驗室無形的研究強項。

在 Upenn，我們跟 Penn Neurodegeneration Genomics Center 的研究人員進行對談，該中心是全美最大的 Alzheimer disease 基因資料中心，擁有五萬筆以上的 Alzheimer disease 全基因體定序資料，最近又獲得了 NIA 一個中心等級的大型 grant 支持。他向我們展示了他們對於失智症基因 structure variant 的最新成果，該成果目前正在 nature genetics 期刊 revise 當中，他們也提議與本院合作 Han Chinese 的相關失智基因研究。

此外 我們參觀了 2023 年諾貝爾獎得主 Drew Weissman，發明 Moderna 疫苗的 RNA 創新中心，個人化診斷中心的臨床定序實驗室，以及費城兒童醫院 (CHOP)的 Grant 實驗室。RNA 創新中心為 Dr. Drew Weissman 最新的實驗室，

還在搬遷當中，他們提及他們正與很多廠商合作努力開發其他應用，把 RNA 疫苗從預防往基因治療的方向推展。本次參訪個人化診斷中心的臨床定序實驗室，此實驗室為費城大型臨床 NGS 報告實驗室，具備有完整的實驗室規劃，更值得一提的是他們有一個非常優秀的 LIS( lab information system)系統，從樣品製備到定序實驗，到最後的基因位點判讀，都在該 LIS 系統上操作，所有相關資訊都記錄在該系統，NGS 報告有別於一般傳統檢驗檢查，其實驗過程以及後續生物資訊分析，位點判讀等步驟都相當繁瑣。如何建立一個優良的 LIS 系統以增進本院次世代定序報告的品質，該實驗室可以作為本院的參考。

我們也參訪了 CHOP 的 Grant 實驗室以及單細胞定序核心實驗室，本院的單細胞定序實驗剛起步，這次參訪請教了對方很多對於檢體單細胞的製備方式，對方的經驗解決了很多我們在面對實驗曾經遭遇過的困難。CHOP 的實驗室目前在空間基因體學上也正積極準備中，他們已經配備了多台不同的 FFPE 切片多重抗體螢光呈色的實驗設備，目前空間基因體的設備正準備裝機中。本次參訪收穫為多面向的，其中一個是技術的精進，跟 NIH 與 Upenn 專家請教我們在本院發展基因體所遇到的技術瓶頸，以他們的經驗提升本院基因體研究以及精準醫學研究的能量。此外，了解兩個頂尖研究機構對於精準醫學與基因體學的研究方向，可作為本院未來發展的借鏡。藉由此次的參訪也持續洽談與 NIH 以及 Upenn 之間的國際合作，希望能有機會促成跟國外頂尖機構的合作。



## 心得

這次參加 ASHG 與機構訪問給我很多的省思。美國頂尖研究機構在精準醫學領域的研究深度和廣度當然是世界頂尖，除了有大型尖端儀器設備支援之外，他們對於各領域專門人才之間的溝通以及合作有特別的著墨，此次參訪常常會聽到”那位跟我合作的 PI 是做什麼什麼的之類”的談話，合作這個字眼很常從美國 PI 的口中聽到。台灣因為重視論文排序的問題，比較偏向各自實驗室單打獨鬥，較缺乏合作的風氣。本次參訪可以感受到他們對於分析以及資料科學的重視，實驗室跟招募的 PI 都配置比以往更高比重的資料科學人才。除此之外，幾個大型的基因體計畫都在使用最新的基因體技術，包含 WGS、空間基因體學、單細胞定序、長片段第三代定序等，國外的藥廠投資大量資源在這些基因庫中找尋 druggable target，香港、新加坡以及韓國都已經做了上萬個 WGS 或 WES 定序，臺灣需要開始準備下一代的基因體研究。

透過這次的學習和交流，我們了解了國際上精準醫學的最新進展，更體會團隊合作的重要性，以為嶄新基因體技術將會精準醫學研究帶來突破性的發展。

## 建議

1. **增加資料科學相關研究人才**：資料科學於醫學研究扮演越來越重要的角色，應多方面招募相關人才，包含生物資訊、醫療資訊、深度學習模型、統計等專長，考慮資訊人才招聘不易，可朝院內員工篩選長期培養資料科學研究能力的方向進行。
2. **關注第三代定序於人類基因體醫學之發展**：目前本院對於第三代定序只擁有低通量儀器，故關注於菌種相關研究。國際上使用第三代定序探討人類基因體相關的甲基化、長片段變異等研究蓬勃發展，應持續關注，找尋本院合適的切入時機。
3. **強化院內外研究的合作模式**：目前本院研究屬於 PI 及醫師各自研究，每個人寫一個計劃發一篇文章，較缺乏向國外研究單位合作模式，可仿效國外的團隊合作經驗，建立由臨床醫師、實驗研究人員、資訊分析的整合性團隊，強化研究的廣度。
4. **建構本院認證臨床 NGS 報告實驗室的 LIS 系統**:可參考 Upenn 的臨床 NGS 實驗室，建構一個實驗室資訊管理系統，以增進實驗室運作品質。

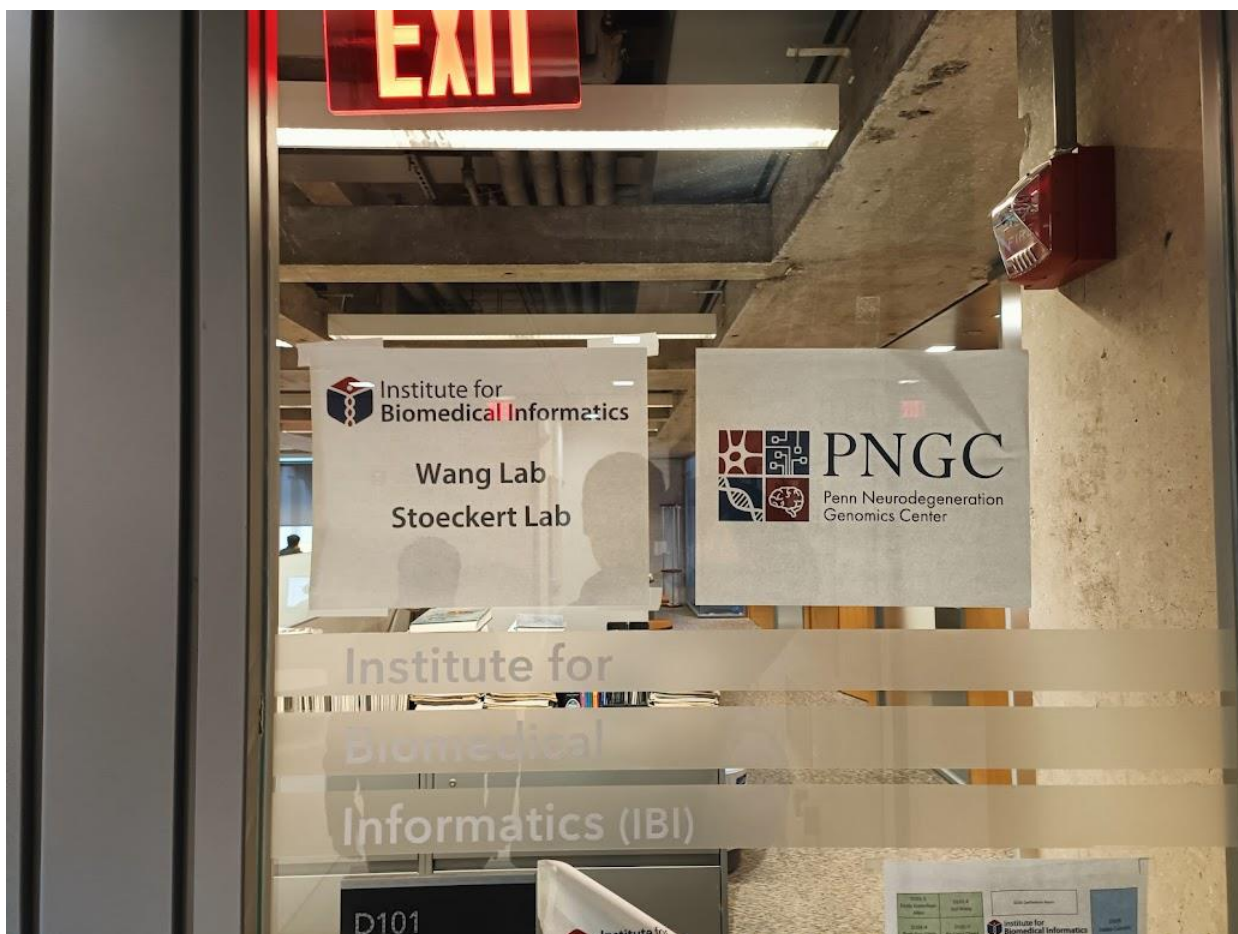


NIH 成人與小兒休閒治療中心的成人治療用健身房



NIH 的 Roy Blunt Center for Alzheimer' s Disease and Related

Dementias Research，內部不開放拍照



參訪 UPenn 的 Penn Neurodegeneration Genomics Center



2023 年諾貝爾獎得主 Drew Weissman, 發明 Moderna 疫苗的 RNA 創新中心





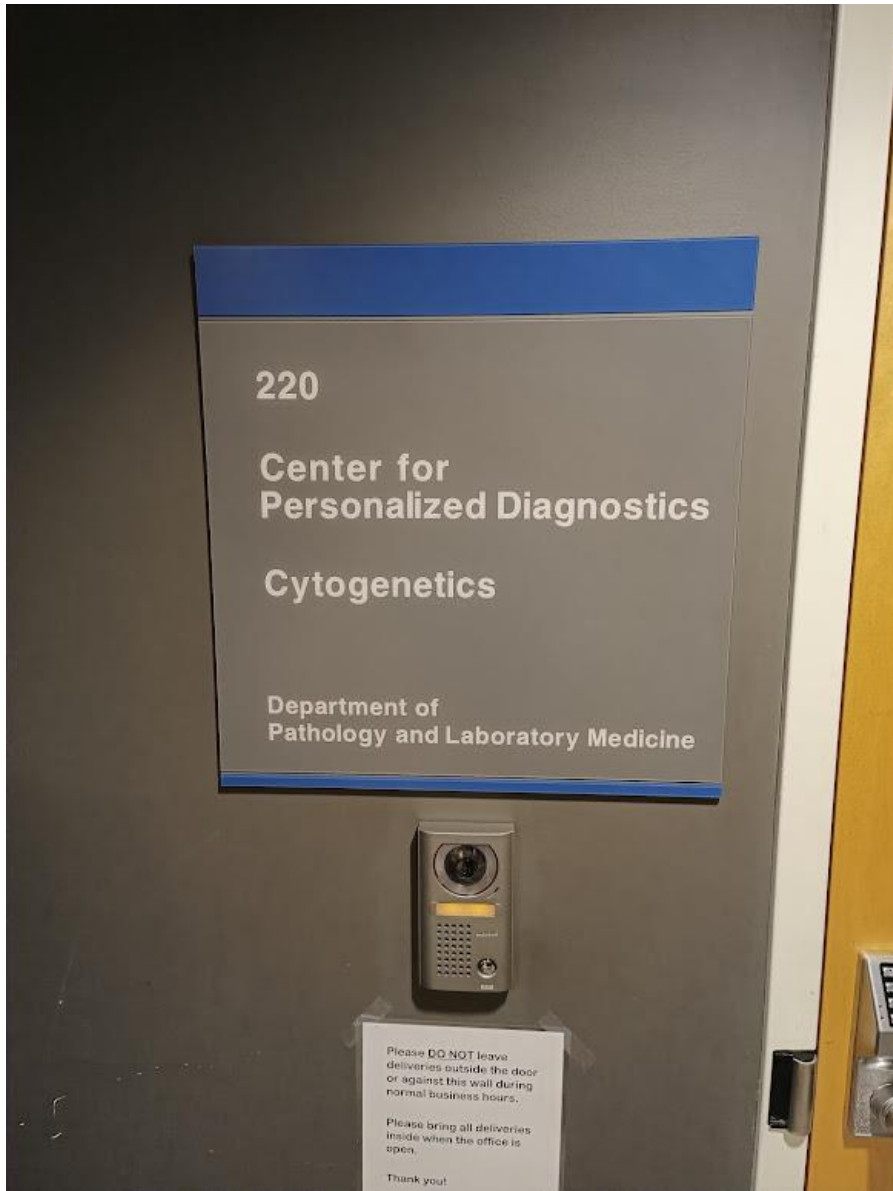
RNA 创新中心内部



RNA 創新中心內部實驗室



參訪 Upenn 病理系



Upenn 臨床 NGS 報告實驗室





Upenn 臨床 NGS 報告實驗室之大型定序設備