

出國報告（出國類別：進修）

2024 執行「補助任務導向型團隊赴國外研
習」龍門計畫【運用單細胞和多尺度基因組
數據進行深度學習預測以提供個人化癌症治
療指南】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：美國匹茲堡

出國期間：2024 年 7 月 27 日至 2024 年 8 月 30 日

報告日期：2024 年 9 月 30 日

2024 執行「補助任務導向型團隊赴國外研習」龍門計畫【運用單細胞和多尺度
基因組數據進行深度學習預測以提供個人化癌症治療指南】心得報告

目次

摘要	6
目的	7
過程	8
心得	13
建議	14

摘要

本次出國研究專注於基因體研究在癌症治療中的生物標記研發，特別是利用空間轉錄體學技術（spatial transcriptomics, ST）來深入研究腫瘤細胞的空間分布及其對治療反應的影響。本次出國為執行國科會資助的「運用單細胞和多尺度基因組數據進行深度學習預測以提供個人化癌症治療指南」龍門計畫，與美國匹茲堡大學合作執行。計畫的核心目標是通過收集多種癌症類型的空間轉錄體學數據，識別與治療效果相關的空間特徵，並利用圖卷積網絡（Graph Convolutional Network, GCN）等深度學習技術，開發能夠預測患者治療效果的精準模型。這些模型可根據腫瘤微環境中的細胞交互模式，找尋單細胞層級的治療生物標記，以作為後續治療效果評估的臨床應用。在本次研究中，團隊使用了多種癌症的公開數據集，通過數據處理與分析技術，識別了與抗藥性和治療失敗相關的空間基因特徵，並探索不同癌症類型中共通與特異的治療反應機制。

，本研究的國際合作部分也取得了重要進展，我受邀至 Johns Hopkins 醫學院進行學術交流，與攝護腺癌基因體學的領域專家 Jun Lou 教授探討了最新的研究成果，並確立了雙邊合作計畫。

本次出國為本院未來空間基因體學研究提供了關鍵的技術支持，還促進了跨國學術合作，並為未來的研究方向提供方向。研究成果有望推動癌症生物標記的發展，尤其是在抗藥性及腫瘤異質性分析領域的應用。

關鍵字：深度學習、空間生物學、基因體學

目的

本次研究旨在進行空間基因體學之研究，並結合深度學習，進而找尋可能的癌症生物標記，本次研究也拓展了與國際研究界的合作機會，特別是在空間領域。藉由本次研究提升本院空間基因體以及精準癌症研究的實力。此外，參與頂尖研究機構的研究對本院醫學與基因體學能量有所提升。

精準醫療在當代癌症治療中日益受到重視，尤其是在治療的個體化和針對性上。然而，癌症治療的效果經常受到諸多因素影響，這些因素包括腫瘤的基因表達、細胞間的相互作用、腫瘤微環境等。近年來，空間轉錄體學（spatial transcriptomics, ST）技術的迅速發展，為我們提供了一個全新的角度來理解腫瘤細胞在空間中的分布，從而更深入地分析其在治療過程中的行為。今年我與陽明交通大學的柯泰銘副教授一同合作「補助任務導向型團隊赴國外研習」龍門計畫，計畫名稱為補助【運用單細胞和多尺度基因組數據進行深度學習預測以提供個人化癌症治療指南】，獲得國科會兩年期補助，計畫內容為一位博士班學生至匹茲堡大學移地研究兩年，主持人與共同主持人於國外單位移地研究兩個月。

透過此龍門計畫，我得以參與匹茲堡大學的前沿研究，專注於運用深度學習技術，結合空間轉錄體學數據，探索腫瘤細胞的空間特徵如何影響治療反應。我們利用這些資訊開發精準的預測模型，以期提升癌症治療的效果與個人化治療方案的準確性。本次出國為一個月的移地研究，與國外團隊以及國外進修的博士班學生深度探討研究主題，進行雙方合作研究計畫的推展。

確認研究目標

本次移地研究的第一件事就是確認研究的主要目標。本次研究室收集最新的空間轉錄體相關公開資料，通過分析收集到的多種癌症的空間轉錄體學數據，尋找與治療效果相關的空間特徵，並基於這些特徵開發預測模型。在多次的團隊討論之後，我們確認了以下幾個核心研究方向與假說：

- 1. 癌症類型與治療專一性：運用收集到的資料確認是否存影響治療效果的共同空間特徵，還是這些特徵僅在特定癌症類型或治療方法中出現？後續的研究設計將針對兩種假說，一是運用資料分析找尋所有治療的療效相關的單細胞層級的生物標記，第二是針對各種不同的癌症治療方式找尋專屬的生物標記。
- 2. 抗藥性的空間基因體特徵預測模型的開發：在抗藥性患者中，是否可以通過空間轉錄體學數據發現新的抗藥性標誌？這些空間特徵是否可以提升治療效果的預測能力？本此研究將使用圖卷積網絡深度學習模型來建模。

建立研究方法

1. 資料收集與數據處理

本次研究的數據來源主要是公開的空間轉錄組學數據集和單細胞 RNA 測序數據。這些數據集來自多種癌症類型，包括卵巢癌、胰腺癌、肝細胞癌等。通過多種平台（如 10x Visium 和 GeoMx）獲取空間轉錄體學數據，並對其進行過濾和預處理。包含品質控制：過濾低品質的細胞和測序點，剔除基因數低於 200 的點，並根據線粒體基因比例等質量指標篩選數據。標準化數據：使用 SCTransform 演算法進行數據標準化，以減少技術偏差。數據整合與降維分析：使用 Seurat 軟體進行數據整合，並通過主成分分析（PCA）進行降維映

射，幫助識別主要的細胞類型和基因表達模式。

2. 空間特徵的識別: 我們採用了多種分析方法識別與治療結果相關的空間特徵，包含細胞類型的注釋：使用 GPTcelltype 進行細胞類型的自動注釋，並根據不同的基因標誌對細胞進行分類。腫瘤細胞的鄰近分析：通過 Cottazm 工具分析腫瘤細胞與周圍細胞的相互作用，特別是免疫細胞與腫瘤細胞之間的空間關係。空間異質性分析：對不同區域的基因表達差異進行分析，並使用 GSEA 進行通路活性分析，以探索潛在的抗藥性機制。

3. 預測模型的開發: 使用演算法分析之識別特徵，研究團隊進一步將規劃使用一種基於圖卷積網絡（Graph Convolutional Network, GCN）的預測模型。此模型能夠根據腫瘤微環境中的空間特徵，預測不同患者的治療效果。GCN 模型的核心是通過節點之間的關聯和特徵，來學習腫瘤內部細胞之間的交互模式，並將其應用於預測不同區域對治療的反應。每個節點代表腫瘤內的一個細胞群，特徵包括該細胞群的基因表達、細胞類型以及與其他細胞的空間關係。通過多層卷積操作，將鄰近節點的特徵傳遞給目標節點，從而捕捉到細胞之間的複雜交互。最終的輸出層負責將每個腫瘤區域分類為“良好反應”或“差反應”，幫助預測整體治療效果。

未來預期研究成果

空間分佈與基因表達的相關性

初步成果發現，在治療效果較差的患者中，腫瘤細胞在微環境中的分佈呈現出特定的模式。特別是在免疫細胞與腫瘤細胞的交界處，往往存在著基因表達的顯著變化。這些變化與治療後腫瘤的生存狀態有密切關聯。

共同與特異的空間特徵

在多種癌症類型中，我們初步觀察到一些共同的空間特徵。這些特徵包括腫瘤細胞的高度聚集區域，以及免疫抑制性細胞的分佈模式。這些區域通常與抗藥性和治療失敗相關。然而，不同癌症類型間仍存在明顯的特異性。例如，在胰腺癌中，腫瘤細胞的空間異質性更為顯著，這可能與該癌症的高度侵襲性有關。

未來研究挑戰

數據整合: 雖然空間轉錄體學技術為我們提供了巨量高解析度的數據，但不同平台產生的數據在解析度和值域分布上存在差異。這使得數據整合和比較變得複雜。規劃開發新的算法來處理不同平台之間的數據偏差。空間特徵的解讀:

空間特徵的複雜性是本次研究中的一大挑戰。不同癌症的腫瘤內部結構差異很大，這使得我們在解讀空間特徵時需要針對每種癌症進行個別分析。同時，腫瘤微環境中的多重細胞類型和它們之間的複雜相互作用，使得特徵的識別和解釋變得困難。計算資源的需求: 處理大規模的空間轉錄體學和單細胞 RNA 定序數據需要大量的計算資源。本次研究依賴匹茲堡大學高性能算力，這對於之後本院發展相關領域是一個挑戰。

未來將繼續擴展資料集，涵蓋更多的癌症類型。我們計畫將研究範圍擴展至免

疫治療等，以進一步驗證我們的預測模型。本龍門計畫為兩年期，將持續推動計畫進行。

Johns Hopkins 醫學院泌尿科學術交流

本次匹茲堡大學進修期間也受到 Jun Lou 教授的邀請，受邀到 **Johns Hopkins** 醫學院泌尿科進行學術交流，Jun Lou 教授是攝護腺癌 AR-v7 的發現者，也是 Alan W. Partin 講座教授。我與 Jun Lou 教授認識多年，這次受邀與他們學科進行交流。在 Jun Lou 教授的安排之下，一整天與該學科六位研究人員一對一的交流，並在中午給了一個全科演講，跟他們交流本院目前精準醫學在攝護腺癌上的研究成果。

在此次的學術交流中，我有幸參與了多位泌尿學研究專家的討論與演講，從中學習到了許多關於前列腺癌與膀胱癌基因體學的最新研究進展，以及如何將遺傳變異與臨床數據結合進行疾病風險預測。更進一步地，因為本次參訪我們與 Jun Lou 教授開啟了一個合作計畫的契機，他很歡迎我們醫院派員至 Johns Hopkins 泌尿科與他們一起進行相關攝護腺癌基因體研究，回國之後我們將持續保持聯絡，開始雙邊合作計畫。

心得

本次研究在與國外團隊合作之下，使我增進了對空間轉錄體學技術的理解，還使我掌握了大量數據分析技術和模型開發方法。這次經歷對團隊未來的研究工作具有重要啟發。我希望能夠將這些技術帶回的本院的團隊，增進本院對於資料分析的能量。

在這次的研究過程中，我再次體會到跨領域合作的重要性。空間轉錄體學的研究涉及多種學科的知識，包括生物資訊學、腫瘤學、免疫學和計算科學。單靠個人力量難以完成這麼複雜的研究項目，因此在匹茲堡大學的這段時間，我有機會與來自不同背景的專家學者共同合作，讓我更深刻地理解如何將不同學科的觀點和技術整合到一個具體的研究問題中。

最為重要的是，這次的經驗讓我更加確認腫瘤內部的異質性為未來重要的研究方向。透過空間轉錄體學技術加上深度學習演算法的剖析，我們有能力更精細地剖析腫瘤內部的異質性。我期待未來能將這些技術引入本院的基因體研究，進一步促進本院精準醫療的發展。

建議

1. 建構本院核心算力：深度學習對於資料科學與醫學研究扮演越來越重要的角色，算力是深度學習的基礎需求，建議本院建立基礎算力以供未來發展使用。
2. 強化深度學習在基因體學的研發能量：建議建構深度學習相關專業研究能量，深入理解數學模型，進而驅動如 transformer 或者 embedding 等深度學習的前沿發展與本院的基因資料做結合，並延伸與醫療資訊整合，將可強化本院之研發能量。
3. 關注基因體學多模態 multi-modal 資料分析的發展：本次 AACR 已有多個頂尖團隊發表了空間生物學的 multi-modal analysis 的成果。除空間生物學之外，LLM 等領域也在討論多模態的資料處理。本院還在發展基礎的空間基因體學，多重螢光染色技術也還在建立當中。須持續專注多模態在空間生物學未來更進一步的發展
4. 關注 embedding 與檢索增強生成 (Retrieval-Augmented Generation, RAG) 技術於精準醫學的相關發展：本院 TPMI 資料庫須結合很多文本資料，例如 SOAP 或者出院摘要，可以使用 RAG 進行相關應用，或者使用 embedding 進行分析研究，將針對此技術找尋研究方向。

Wednesday, August 21st, 2024

10:15AM-11:00AM Dr. Shawn Lupold, Professor, Urology Research
Marburg 123 Research Office

11:15AM-11:45AM Dr. Eugene Shenderov, Assistant Professor, Urologic
Oncology
Marburg 123 Research Office

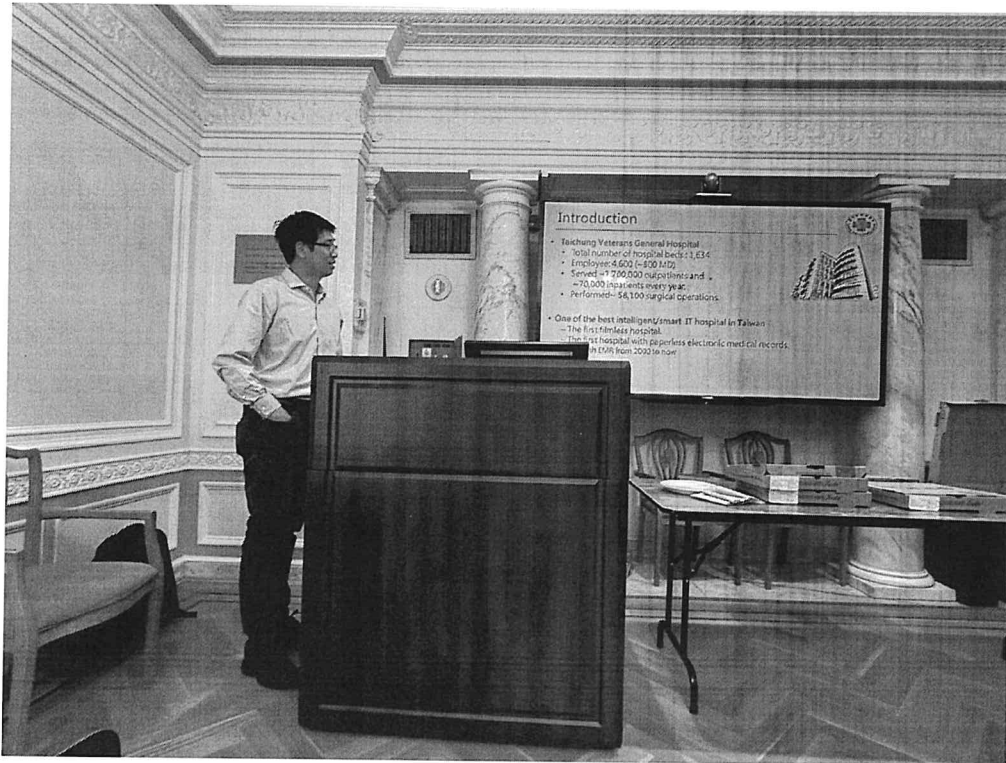
12:00PM-1:00PM Lunch Seminar, Title Talk: ***Combined Effects of Rare Pathogenic
and Common Germline Variants on the Risk of Prostate Cancer in Han Chinese***
Marburg Conference Room

2:30PM-3:00PM Dr. Jonathan Coulter, Assistant Professor, Urology Research
Marburg 123 Research Office

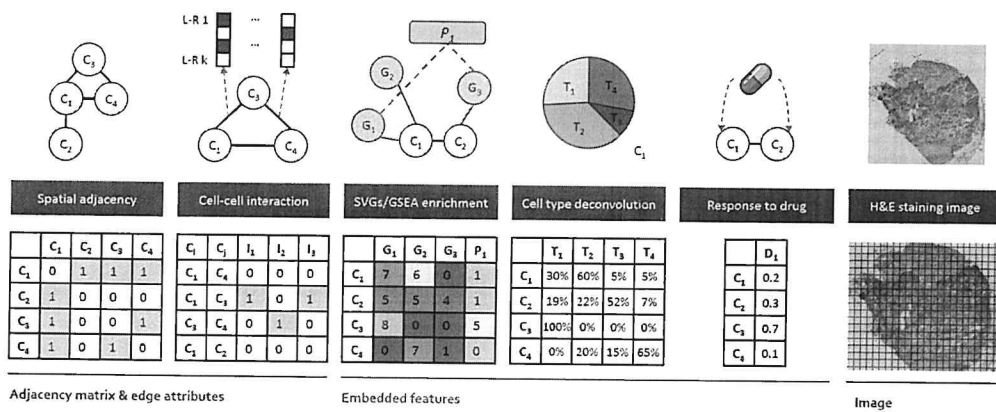
3:15PM-3:45PM Dr. Sarah Amend, Associate Professor, Urology Research
Marburg 123 Research Office

4:00PM-4:45PM Dr. David McConkey, Professor, Urology Director Greenberg
Bladder Cancer Institute
Marburg 123 Research Office

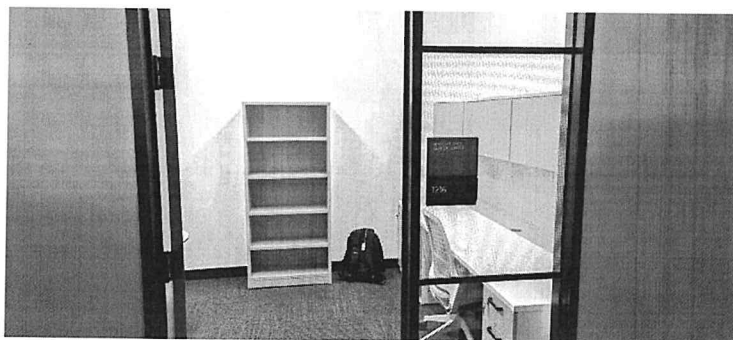
圖一: Johns Hopkins 參訪議程



圖二: Johns Hopkins 受邀演講



圖三: 空間基因體與 GNN 研究設計



圖四: 匹茲堡大學研究室



圖五:Johns Hopkins 討論室