

出國報告(出國類別：研究)

科技部補助前往西班牙高等科學研
究委員會研習計畫案

Genomic and genetic mapping approaches to reveal trait-associated loci in peach:
Genetic architecture of fruit metabolic composition

服務機關：農委會林業試驗所

姓名職稱：吳家禎助理研究員

派赴國家：西班牙

出國期間：107年9月6日至107年11月6日

報告日期：107年12月15日

關鍵詞：基因體關聯性分析(Genome-wide association study)、分子標誌輔助育種(Marker assistant breeding)、單一核甘酸多型性(SNP)、族群遺傳(population genetics)

摘要

EEAD (Estación Experimental de Aula Dei)隸屬於西班牙高等科學研究委員會 (Consejo Superior de Investigaciones Científicas, CSIC)，主要研究的領域圍繞在植物的遺傳、生理、分子、土壤和水、育種以及農業生產，是重要的農林業研究重鎮。本次研習時間為期2個月，達到充分建立關係，以及擬定未來可合作的領域，研習期間都待在Dr. Gogorcena帶領的研究團隊內，該實驗室進行多種櫻屬植物(*Prunus*)的遺傳育種研究，本次研習的物種主要為水蜜桃(*Prunus persica*)，從育種品種的收集、培育，分子標誌的使用，再到全基因組關聯性分析暨生物資訊學，都有所學習。本次研習協助並且和該實驗室討論如何從單一核甘酸多型性(single nucleotide polymorphism, SNP)分子標誌，結合過去3年的性狀資料，比較不同的統計模型進行全基因組關聯性分析，找出具有顯著性的基因位點，有助於未來實驗以及分子育種的研究工作，本次研習充分實踐SNP的關聯性分析，並且協助該實驗室完成相關分析工作，完成的實驗數據將交由該實驗室博士生接續完成，也是一項極為良好的台西交流項目。

目錄

一、目的	4
二、研習過程	5
三、心得	8
四、建議	23

一、目的：

EEAD(Estación Experimental de Aula Dei)隸屬於西班牙高等科學研究委員會(Consejo Superior de Investigaciones Científicas, CSIC (英譯：Spanish National Research Council))，位於西班牙阿拉貢省(Aragon)的首府薩拉戈薩(Zaragoza)，EEAD主要分為四大主軸，圍繞在植物的遺傳生理與分子、土壤和水、育種以及農業生產，是重要的農業研究中心。本次加入Dr. Gogorcena帶領的研究團隊，對於木本植物的水蜜桃或是櫻屬(*Prunus*)植物，針對其育種、選種與鑑定有深入的研究。當地已有完善的水蜜桃育種試驗園，進行長期調查，同時強調以DNA分子標誌，進行選拔、遺傳評估以及鑑定的工作，本次研習學習目標為：

- 1.如何建立有效的種源圃。
- 2.收集的種質資源的遺傳規模結構大小如何界定，方法為何。
- 3.在缺乏大量序列資訊下，如何應用最經濟的方式，包含生物資訊學的使用，建立有效的分子育種體系。

另外，透過本次的研習，將來也能夠建立和EEAD的關係，建立合作網絡，增進國際學術交流關係。

二、 研習過程

本次科技部計畫由CSIC轄下各研究機構提供研習機會，以增進渠對西班牙科研機構之了解，汲取其研究方式，並協同雙方研究人員就未來兩國可能合作之主題及方向進行交流，促進雙方團隊實質合作研究。計畫規定的時間為期2個月(自2018年9月6日至11月6日)，本次研習的地點在西班牙阿拉貢省的薩拉戈薩EEAD研究中心，研習過程及主要研習時間表如下：

日期	地點	行程與紀要
9月3日	台灣桃園-香港-卡達	桃園機場出發
9月4日	卡達-馬德里	抵達馬德里國際機場
9月5日	馬德里-薩拉戈薩	與Dr. Gogorcena碰面，前往租屋處，辦理租屋事宜
9月6日	薩拉戈薩	整理住家與安頓，聯繫科技部法國辦事處
9月7日	薩拉戈薩	前往EEAD辦公室與實驗室，行政辦公室處理報到手續，學習如何搭乘院內接駁車，並且討論將要著手的主題：利用SSR與SNP重新計算94個peach品系的genetic structure.
9月10日	薩拉戈薩	進行約50分鐘的對內簡報，介紹台灣以及自己所做的研究工作內容，並認識實驗室環境與其他夥伴Maria(Spanish)、Chesco(Spanish) and Najila(Tunisia)。
9月11日	薩拉戈薩	認識EEAD整個環境與相關設施、圖書館資源等，也認識在EEAD的中國四川中科院博士生其進行聯合國水保相關研究
9月12日	薩拉戈薩	和Dr. Gogorcena討論論文內容，探討SNP、SSR的雜交應用與遺傳結構分析
9月13日	薩拉戈薩	前往阿拉貢政府農業研究中心，其和Dr. Gogorcena合作的團隊，他們目前是使用ABI3130進行SSR的genotyping，目前Dr. Gogorcena有25組multiplex PCR對於prunus植物進行分析鑑定

9月14日	薩拉戈薩	前往EEAD種原收集圃，由Falando和maria一起看Aula Del的收集園圃，這邊有桃子、櫻桃、蘋果、橄欖、梅子、開心果等，從扦插、嫁接、園子的灌溉系統都進行解說介紹
9月17日	薩拉戈薩	資料分析(使用R)以及整理水蜜桃SNP資料
9月18日	薩拉戈薩	使用Structure試著完成SNP試算
9月19日	薩拉戈薩	整理GBS與NGS定序資料，成功轉換格式，繪製phylogenetic tree，Structure軟體的SNP格式也輸入成功，開始運算
9月20日	Biescas	前往Biescas，訪問實驗室合作的葡萄園區，了解葡萄11種品系，請Dr. Gogorcena進行品種鑑定，拿到實驗室去進行品質的測試。
9月21日	薩拉戈薩	和馬德里的老師(Dr. Sandiego)視訊，以及討論遺傳分析方式
9月24日	薩拉戈薩	整理資料以及討論Structure，試算運算permutation
9月25日	薩拉戈薩	文獻閱讀
9月26日	薩拉戈薩	文獻閱讀與研究人員討論關於Structure運算速度問題
9月27日	薩拉戈薩	準備歐洲研究人員之夜活動
9月28日	薩拉戈薩	參加歐洲研究人員之夜
10月1日	薩拉戈薩	繪製SSR的相關資料給Dr. Gogorcena
10月2日	薩拉戈薩	和Dr. Gogorcena討論SNP data，進行篩選SNP工作
10月3日	薩拉戈薩	和當地高中有合作的實習參訪，帶領學生使用DNA偵測鑑定的活動
10月4日	薩拉戈薩	和當地高中有合作的實習參訪，帶領學生使用DNA偵測鑑定的活動
10月5日	薩拉戈薩	和Dr. Gogorcena討論博士論文與資料整理
10月8日	薩拉戈薩	研究室會議，Chesco論文報告。計算Q和K矩陣
10月9日	薩拉戈薩	完成SNP的filtering
10月10日	當地國定假日	當地國定假日

10月11日	當地國定假日	當地國定假日
10月12日	當地國定假日	當地國定假日
10月13日	當地國定假日	當地國定假日
10月14日	當地國定假日	當地國定假日
10月15日	薩拉戈薩	實驗室會議，Najila、Chesco完整簡報，以及完成Data filtering
10月16日	薩拉戈薩	實驗室會議以及帶領教導新的國際交流生Jabier生活須知與房屋陪同前往
10月17日	薩拉戈薩	使用TASSEL input檔案
10月18日	薩拉戈薩	使用MLM或GLM，以及資料整理，帶Jabier去看他的公寓
10月19日	薩拉戈薩	使用TASSEL、文獻閱讀
10月22日	薩拉戈薩	調整水蜜桃資料SNP的chromosome位置等
10月23日	薩拉戈薩	重新計算TASSEL
10月24日	薩拉戈薩	和Dr. Gogorcena討論TASSEL結果以及significant的使用
10月25日	薩拉戈薩	和Dr. Gogorcena討論雜交資料
10月26日	薩拉戈薩	Chesco碩士班口試，和薩拉戈薩大學(Universidad de Zaragoza)Dr. Pilar Catalan討論雜交研究實驗，以及後續合作項目
10月29日	薩拉戈薩	和Dr. Sandeigo討論TASSEL以及運算6種TASSEL in Peach
10月30日	薩拉戈薩	回復Pilar關於協助採集植物樣本，以及未來有機會合作撰寫文章之類的
10月31日	薩拉戈薩	行政作業，領取在地生活津貼
11月1日	當地國定假日	當地國定假日
11月2日	薩拉戈薩	完成交接檔案給Dr. Gogorcena以及餐聚

11月5日	薩拉戈薩	給大家小紀念品，檔案交代清楚，進行交接，和圖書館館長會面，贈送林試所套書
11月6日	薩拉戈薩	研習簡報暨討論會議，研習結束

三、心得

本次研習主要都於進行兩單位間的合作認識，對於彼此領域的交流學習，筆者前往該地主要學習觀摩在地育種工作，透過在當地生活，互動，交流，進而更了解當地的相關研究工作，以渠成為雙方交流合作時彼此瞭解的優勢。相關心得與學習分述如下：

育種的基本介紹與操作

人類可以進行植物育種，是因為植物具有三個進化的要素：遺傳、變異和選擇。遺傳是進化的基礎，變異是進化的動力，變異提供了可以選擇的材料，如果沒有遺傳，有利的變異就無法傳遞給後代，也不能使有利的變異得到穩定，而選擇也就決定生物進化的方向。所謂常規的育種方式，在林木來說，少則 8-10 年為一次育種循環，多則 20-30 年一次循環，十分耗時費工，簡單介紹幾個重要方法如下：

選擇育種，又分為混合選擇法(**bulk selection**)和單株選擇法(**individual selection**)。混合選擇法是從一個原始的群體或是品種中選拔出符合育種目標的優良單株，然後再收集它們的繁殖材料(種子或繁殖器官)，混合後再次栽種於同一地點進行觀察，待性狀穩定後，再與原始群體或對照品種進行比較，又可分為一次混合選擇法(**single bulk selection**)及多次混合選擇法(**multiple bulk selection**)。自交授粉植物，後代不分離，性狀容易穩定，故選擇一次混合選擇法；而大多數的木本植物屬於異花授粉，後代多產生性狀分離情形，則必須採用多次混合選擇法。

混合選擇法的優點是方法簡便易行、時間短、異花自由授粉，不易有近親繁殖導致生活力衰退的情形。缺點主要就是混合選擇主要是依據單株種子混合繁殖，不能鑑別該後代遺傳性的真實優劣，有時可能因為在優良環境下，表現型優異而實際上遺傳性並不優良的個體也入選，而降低主要的選擇效果；另外選擇的效果會越來越不顯著(遺傳增益低)。

單株選擇法，也是透過原始族群中，優良的單株選拔後，將其繁殖材料分別收集(不進行混合)，分別保存、分別繁殖成不同家系後，統一栽種並各家系進行隔離，再根據各個家系性狀的表現(以後代家系為單位)，評斷原始族群中入選之單株的優劣，進行選留及淘汰的工作。該選擇法的優點是：經過多代的

連續選擇，可以消除環境影響，選出真正屬於遺傳性變異的優良類型；多次單株選擇可以定向的累積遺傳變異，選出超越原始群體內的優良單株。缺點在於技術較複雜，一花授粉需要隔離試驗成本提高，多代的近親交配，可能造成近親衰退(inbreeding depression)。

雜交育種，意即透過人工控制後粉，選定親本後，按照一定的程序目標性的進行雜交育種工作，透過該育種方式產生後代再次重複的培育後與選拔，使基因重組後，達到純系同型合子的新品種，有時這樣的育種方式多為近親雜交，因此，雜交育種的目的是使同型結合的基因位點在族群中不斷的增加，而異型結合的基因位點減少，在數代雜交選育後，得到穩定遺傳的新品種類型。該項育種方式也是大多數的育種方式使用，先挑選出優良的父母本後，進行單交(single cross)，這時又可以分為正交、反交，從而可以了解性狀傾向於父本還是母本。該方法類型甚多，端視情況與試驗設計而使用。

分子標誌

隨著分子生物學與電腦科技的進步，科學家也開始藉由分子標誌，縮短育種的時間。而分子標誌，一般被認為是 DNA 標誌(DNA marker)，應用的面向極為廣泛，分子標誌在不同個體或是族群間所呈現出不同的變異(指多型性(polymorphism))，成為實驗者判別以及實驗應用的依據，而若找出適當的分子標誌和目標性狀有緊密連結的關聯性，便可以透過分子標誌選育品種，將可以為育種者縮短育種時間，或是計算其與性狀的關聯性，近而透過 DNA 分子標誌，預測性狀。

當分子標誌輔助育種，以及 QTL 關聯性定位的研究趨緩後，早期礙於定序成本高昂，以致於全基因組定序選拔的方案發展緩慢。但是隨著定序成本不斷降低，全基因組選拔就開始被科學界的關注，在全基因組選拔計畫中，最常使用在分子標誌輔助育種的 2 種分子標誌為微衛星體(microsatellite, simple sequence repeat, SSR)以及單一核苷酸位點多型性(single nucleotide polymorphism, SNP)，這兩種分子標誌同樣在分子標誌輔助育種中扮演重要角色。另一方面，SSR 以及 SNP 的應用，也助於深入研究生物的遺傳、演化、近親交配、基因漂流、以及馴化與野生物種的保育。同時在育種計畫中，這些分子標誌扮演著推算親屬關係(kinship)以及雙親關係的資訊。親屬關係是指出去量測在親屬間共有的對偶基因出現的頻率，這樣的對偶基因會有兩種出現的可能性，第一個是由共祖傳遞下來，因此出現相同的對偶基因；另一個可能性是不同的共祖，但是卻是來自相同的基因座。所以，家系譜的重建或是親緣關係的推定，涉及到該個體應該被分配到父親或是母親，也就是追溯該對偶基因的來源：不是從父親來的就是母親來的。微衛星體標誌很適合用來做為親子分析，主要是因為這樣的分子標誌具有高度的信息(multi-allelic，這個字是表示一個基因可能具有超過 2 種以上的多種形式)、為共顯性標誌(就是異型結合子可

以和同型結合子有所區別)，以及具有可信度。而近日來，僅具有 2 種基因型式的 SNP 已經成為大家研究的主流，主要在於基因型定序平台費用降低、育種者可以很容易的取得大量個體 SNP 的資訊，使得全基因組選拔工作，會以 SNP 做為主要鑑定基因型的分子標誌系統。

EEAD 果樹與葡萄基因體學研究室(Genomics of Fruit Trees and Grapevine)介紹

EEAD(圖1)主要分為四大研究主題：遺傳與植物生產(Genetics and Plant Production)、植物養分(Plant Nutrition)、果樹學(Pomology)與水土研究(Soil and Water)，筆者這次前往的實驗室是由Yolanda Gogorcena Aoiz所領導的「果樹與葡萄基因體學研究室」，這個實驗室從2002年開始成立，以木本植物（薔薇科和葡萄科）的育種為主，並進行三個主要主題：i) 選擇適應地中海條件不同區域的水果品種和砧木; ii) 開發用於植物育種的多學科技術; iii) 保護果樹和葡萄的遺傳資源研究。近年來的研究重點是基因組學和多學科技術的整合使用，像是：鑑定候選基因和開發植物育種的分子標記，致力於水桃，李屬植物和蘋果的遺傳資源保護和分子鑑定，以及阿拉貢種質庫中葡萄種質鑑定分子圖譜的標準化。另外，也展開*Prunus*，*Malus*和*Vitis*屬的基因組學研究，目前實驗室也收集了相關的基因資料庫，擬定作業流程，對於上述相關物種提供鑑定服務，像是：品種鑑定、親子鑑定或基因組學功能分析。



圖1、EEAD大樓，裡面包含辦公室、實驗室、圖書館等。

農民合作與品種分子鑑定

筆者前往該研究室參與了所有的活動，像是與農民合作結合地理資訊系統(GIS, global information system)進行葡萄園栽培與分子圖譜鑑定工作，以本次訪問的例子來說，我們參訪一個生產葡萄酒的科技農夫(Pilar Gracia)，對於葡萄品種種苗有疑問，便找EEAD協助，並且在新植葡萄園搭配GIS系統標註相關訊息與繪圖。她的葡萄園位於西班牙與法國交界的庇里牛斯(Pyrenean)山腳下的Biescas，距離Zaragoza北方約150公里的位置，市區海拔約800公尺，該園自2012年開始種植，有11種品系，總共約3500株，位於海拔1200公尺位置，占地0.6公頃(圖2)，目標是要選拔這個環境是否合適那一個品種，並且進行分子品種鑑定，因為，種苗商購得的苗木，有被懷疑是標示錯誤，後來在基因型分析的結果中，原本標示不同的兩個品系，但是基因型相同，因此，被分類至正確的品系，同時也把這邊種植的品種拿到實驗室去進行品質的測試，像是總糖含量等，訪問過程中，剛好遇到測試採收的日期，我們把11個品系採集少量，進行後續分析，像是糖量的測定、pH值、baume、alcohol degree、成熟度等，本次成熟度皆測為7，最好是5，指示劑會從粉紅色變色到紅棕色，如果5就是最好的採收時間(圖3)。

過程中也稍微學習詢問當地種植的條件以及品系差異，舉例來說：有些品系如果有病菌感染，他們就不會採收，也不會進行品質檢測(只採收好的果實)，有些品系果實性狀結果很緊密(Noir品系)，會造成果時間不通風，容易產生真菌感染，有些品系則不會。整個園子不施行灌溉，只靠老天降雨，春天只施肥一次，有機肥，讓樹高生長，但是不要發生徒長枝，突出走道，以防競爭與妨礙通行。每年這片葡萄園可以生產約2000瓶白酒和1000瓶紅酒。Pilar Gracia也持有其他的葡萄園(圖4)，並且和周邊種有葡萄的農民(由於年紀較大，管理不易)進行簽約採收與管理，Pilar Gracia也說到，進行科技化管理與保護，在部分的葡萄園周邊都架設太陽能電力防護圍欄，防止動物與人的危害，提高產量。



圖2、以科學經營的葡萄栽植地，每年可以生產上千瓶葡萄酒。



圖3、滴定檢測葡萄最是採收時間預測。



圖4、傳統的葡萄園種植方式，不搭網不架支柱。

EEAD種源收集園

EEAD 的育種收集園(collection plantation)從 EEAD 成立 70 多年就已經設立，並且陸續擴增，在參訪種原收集前，在 EEAD 的周邊可以看到種源收集園附近有一間卡爾特修道院(Charterhouse)(圖 5)，也是西班牙的重要國畫師戈雅(Francisco Goya)早期重要藝術創作地點，修道院鐘樓是戈雅做畫的對象，由於該修道院為著名的藝術觀光之地，也為 EEAD 意外增添不少藝術的浪漫成分。

EEAD 從成立以來，就一直進行 Prunus 的育種工作，從水蜜桃、櫻桃、梅子、蘋果、橄欖、開心果等都有相當的成果，育種工作中最重要的一項環節就是收集不同地區的品種，也就是種源，EEAD 多年來努力收集種源，以 peach 來說就維持經營了 70 年。參訪的一開始，研究室同仁從砧木(rootstock)的培育開始講起，要培育壯大且良好的根系，就是從嫁接苗開始進行。砧木的培育都是選取一年生的枝條，從頂端扣除上頭約 50-60 公分，直徑約 1 公分，長度約 30 公分的木質化枝條，做為扦插材料(圖 6)，待扦插發根後，種植一年，就可以當作砧木進行嫁接(圖 7)。由於水蜜桃種源圃的配置圖，算是研究機密，所以，不能給我，在種源配置上，同仁解釋道：總共有 56 個品系，每個品系有 3 株，依照結果實的時序開始排列(圖 8)，因此這片種源圃，也已經更新多次，如果需要營養系苗木，就從這邊採取木質化枝條，進行嫁接培育複製，品種收集的目的是要維持遺傳多樣性，並且確保下一代的培育可以具備面對下一個未知環境對於植物的衝擊。因此，這邊收集的種源，並非全部都符合商業需求與價

值，因此，賣相不佳或是無法抗逆境的品系這邊都是雙手歡迎(圖 9)。因此，該地研究人員也表示如果我們需要進行育種工作，應該盡可能的收集完整的品系，進而進行遺傳結構分析，看看是否族群中有亞族群或是遺傳多樣性的存在，確保保存資源是有科學數據佐證。



圖5、鄰接EEAD種源收集園的卡爾特修道院(Charterhouse)與其鐘樓，為EEAD增添不少藝術氣息。下方看到的是即將種植的試驗田地，已經進行整地。



圖6、水蜜桃插條，待成苗後，一年後可以嫁接。



圖7、直插水蜜桃插穗入土。



圖8、按照果實採收季節依序種植的水蜜桃種源圃。



圖9、水蜜桃種植情況，並非所有品種都是可口的水蜜桃，這也是保持多樣性而收集品種的目的。

歐洲研究者之夜(European Researcher Night)

實驗室也和當地研究機構合作，一同參加歐洲研究者之夜(圖10)，希望將艱澀的難懂的科學，免費的讓當地的民眾參與和認識，本實驗室舉辦4場的水蜜桃DNA萃取課程，讓小學生可以透過活動親身參與(圖11)，EEAD中也有其他研

究室一同參與，像是蛋白質科學實驗室，讓民眾了解生活周遭的蛋白質有哪些，有哪些相關的科學反應，表示EEAD也是一個會與民眾接觸與互動的單位。除此之外，周邊的大學各科系也設有攤位展出，讓科學變得平易近人，筆者也走訪相關的活動展場，像是盲人手持導航器，指引盲人、還有消防博物館，說明火的相關議題(圖12)。

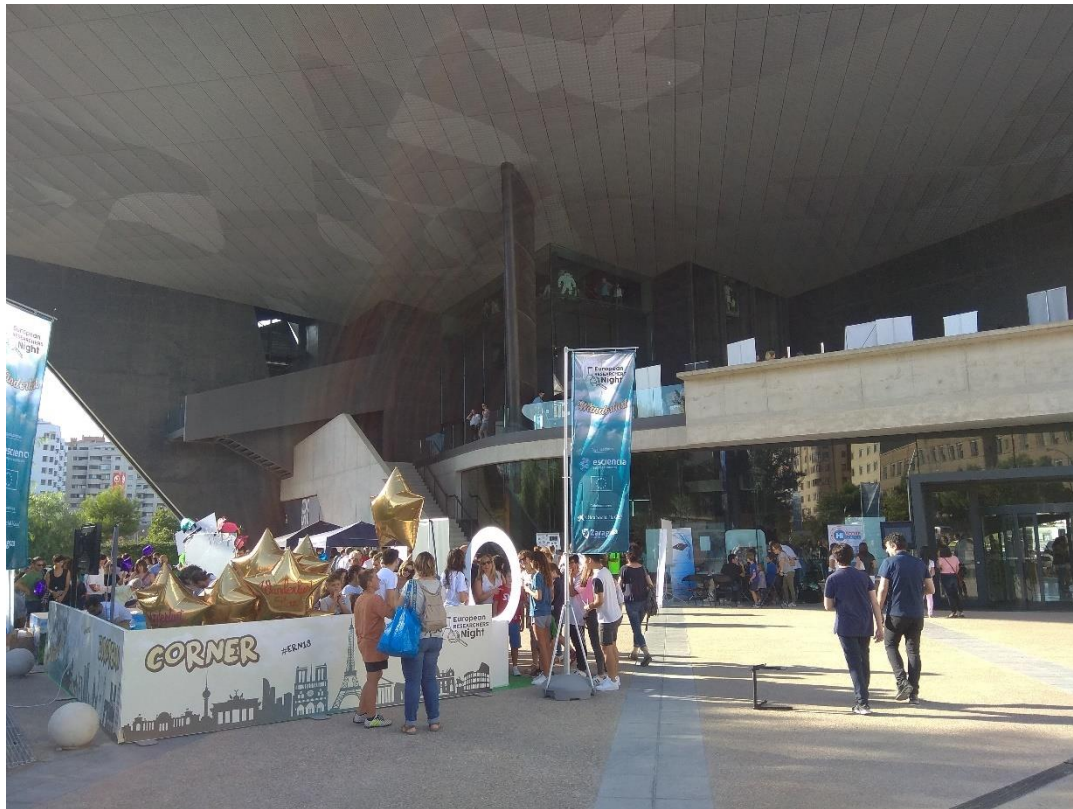


圖10、歐洲研究者之夜現場門口。



圖11、帶領學生進行水蜜桃DNA萃取體驗活動。



圖12、消防博物館中歐洲研究者之夜的活動，讓民眾免費參加，了解科學是甚麼一回事。

使用SNP進行全基因選拔分析

筆者前往EEAD的研究室進行訪問，在到達後，除了日常生活起居的部分之外，在研究中，最重要的就是必須要協助Dr. Gogorcena相關的研究工作，該實驗室在2010年時，使用Illumina Peach RosBREED Peach 9K晶片，以其水蜜桃種源收集團中的94個樣本進行基因型定序，得到約9000個SNP的基因型資料，從這些資料中，筆者必須完成其遺傳結構以及遺傳多樣性的分析，並且將資料和之前的SSR資料相互比較。另一方面，這94個樣本中，該研究室於2009年開始，進行連續3年的性狀調查，透過這些性狀資料，也必須結合SNP資料進行全基因體的關聯性分析，由於本分析結果尚未發表，下方僅能以陳述性的方式進行呈現。

從該研究室早期的報告發表顯示出這些水蜜桃品種的馴化是一個複雜的過程，最主要是由於這些收集的樣本其基因型的無性繁殖和當地野生桃的有性繁殖數代後所致。品種馴化和繁殖通常會導致多樣性喪失，從而導致瓶頸和遺傳漂變的效應。瓶頸後的遺傳多樣性取決於野生族群和耕種出去新的品系組群的

比例，以及瓶頸效應所持續的時間。在許多水果品種中，馴化發生的時間相對較晚，所以持續時間較短。從SSR的結果顯示這94個樣本中，從當地(西班牙族群)栽培品種的多樣性略低於現代栽培品種(其他國際水蜜桃品種)，這可能是因為西班牙收集的樣本是局限所致。當地栽培品種可能多代從自己的族群選出傳播的種群，而現代栽培品種是通過雜交兩個個體並選擇後代而獲得的。而引種也是可以增加或解決當地西班牙品種的遺傳變異性，以適應不同的生育條件特別是在EEAD的埃布羅谷品種(Ebro valley)與美國品種的收集，這顯示出這兩種具有共同的基因庫。因此也吻合了兩國之間水蜜桃種質資源積極交換的結果。

本次透過SNP的資料結果顯示的數據，和SSR分析的結果相當，但是更為細緻詳細，而在分析的過程中，筆者也有許多的學習，像是如果篩選SNP，從一開始拿到SNP的原始資料檔案開始，首先必須檢查是否SNP資料在每個樣本都有達到標準，也就是看是否Gentrain score是否大於0.4，所有的基因型在每個樣本中是否齊全，這邊可以視SNP資料的完整性(no call)，進行篩選：舉例來說，有一個SNP位點，在100個定序樣本中，有5個樣本是缺少資料(也就是no call)，則這個SNP位點genotyping calling的比率就是95% (95/100)，一般都是設定在no call大於5%的SNP就會進行刪除，但是也可以依照不同的實驗需求，定序品質進行調整。上面的no call也視為缺值(missing data)。在本次的定序資料中，會出現非核苷酸的符號，因此，在分析的時候也需要格外注意，將其刪除，可能的原因推測可能是定序廠商分析時，沒有設定好所致，由於分子標誌需要看同一個位點上是否有多型性，所以，如果SNP位點屬於mono-morphic也必須刪除，再來的是異型結合子的數量要大於10%，最後是要移除相同變異形式的SNP排序(similar pattern)，這樣才不會對於同樣的SNP在不同樣本間呈現相同的多型性，在進行關聯性分析時重複計算。

得到SNP資料後，可以透過遺傳分析軟體進行遺傳多樣性的分析，得到遺傳距離，進行親緣關係樹的繪製(phylogenetic tree)，而遺傳結構的分析則是經由Structure進行分析，在進行軟體分析，最重要的是數據的格式化，必須吻合軟體的要求，而Structure的分析，一般來說，除了瞭解樣本的遺傳結構外，評估是否有明顯的亞族群(sub-population)，另外，也是為了得到其中的Q矩陣(Q matrix)，使用於全基因組關聯性分析(Genome-wide association study)。

本次的全基因組關聯性分析使用TASSEL(Trait Analysis by aSSociation, Evolution and Linkage)軟體分析，這套軟體也是學術界中進行關聯性分析研究中，使用率很高的一套分析軟體，同樣必須將輸入資料格式符合TASSEL要求的格式進行輸入，在此簡單介紹筆者在當地進行TASSEL分析時的經驗與流程，提供後續如果國內相關研究人員進行參考。

在進行TASSEL之前，必須具備3種資料：1. 基因型資料(Genotyping data with filtering)、2. Q矩陣資料(Genetic structure data from STRUCTURE software)、3. 表現型資料(Phenotyping data(Traits))，在此性狀資料也可以因不同的試驗設計，分不同處理(年份、地點)，可以整合或是分開處理皆可，但是得

到的結果也會不同，看試驗者如何解釋。首先必須先讀取資料，將輸入的格式，另存為輸出的.hmp檔案(圖13)(該檔案可以使用excel開啟)；接著選擇File，以Make best Guess開啟檔案(圖14)。當檔案讀取之後，先點選基因型資料，進行數據化(Numerical genotyping data)(圖15)，基因型以及表現型資料可以使用impute，進行補缺值的模擬預測(圖16)。

接著進行關聯性分析，一開始先使用General liner model(GLM)進行分析，GLM只需要基因型與表現型資料在操作上，必須使用Intersect Join的選項，將這2個檔案反白合在一起(圖17)，即可以進行GLM的分析(圖18)，計算出來之後，就可以選擇不同的性狀，看每個SNP對於不同性狀的關聯性，藉由p值是否有顯著性(這邊的顯著性，還需要經過邦費羅尼校正(Bonferroni correction))，評估該SNP的效應。接著可以加入Q矩陣當作遺傳結構進行校正，這時Q矩陣先必須經過filter，將其中的1個Q值刪除，再進行分析，也就是所謂的(GLM+Q correction)，這時如同上述，也是使用Intersect Join選項，將3個檔案合併分析，可以比較經過Q矩陣校正的GLM之間的差異。一般進行關聯性分析，也會進行MLM(Mixed liner Model)，MLM需要額外的K矩陣(家系矩陣，kinship matrix)，K矩陣可以透過TASSEL，以基因型資料進行計算。在當地實驗室，透過和Gogocerna博士的指導以及學生的互動，了解到更多目前在水蜜桃分子育種的研究現況。

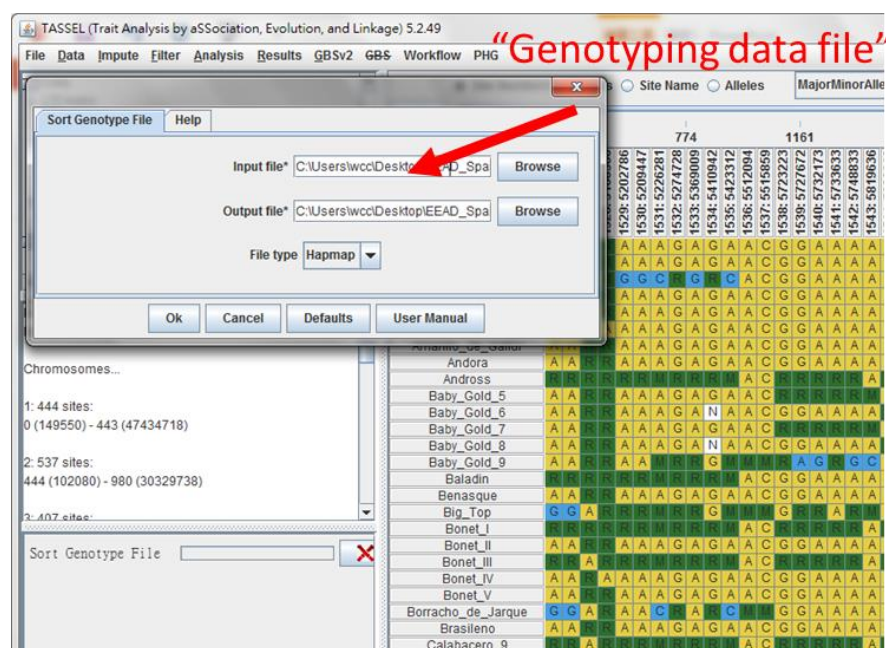


圖13、轉換基因型資料，讀取進入TASSEL程式之中。

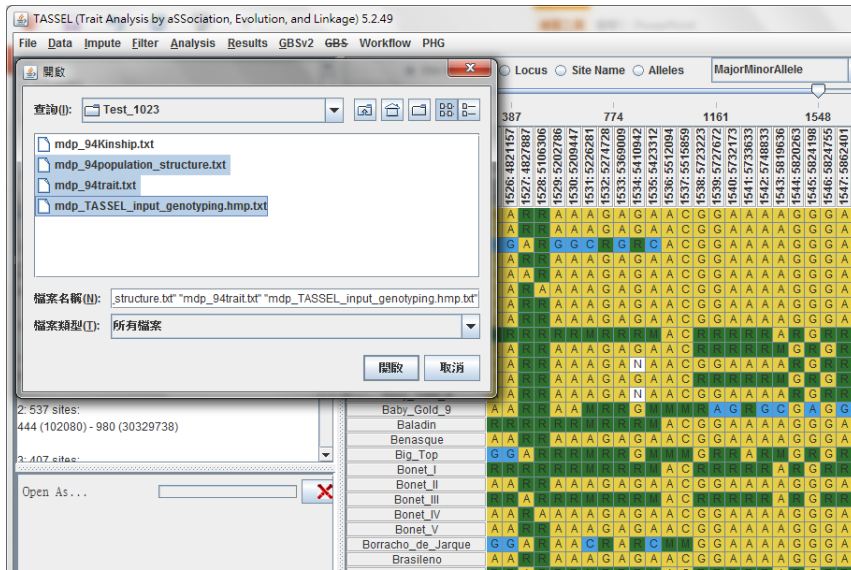


圖14、將分析必要的檔案，如基因型資料(.hmp)、表現型資料(.txt)與遺傳結構的Q矩陣輸入。

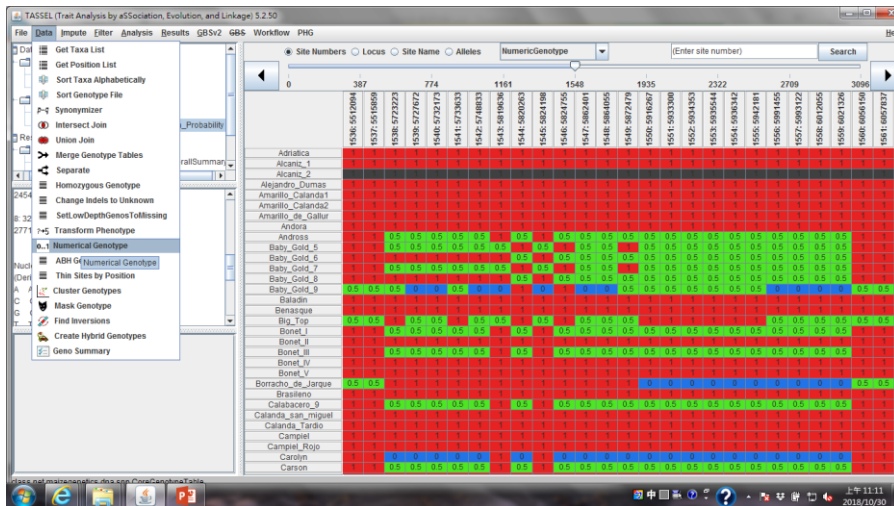


圖15、數據化資料，SNP的3種基因型以0、0.5、1進行表示。

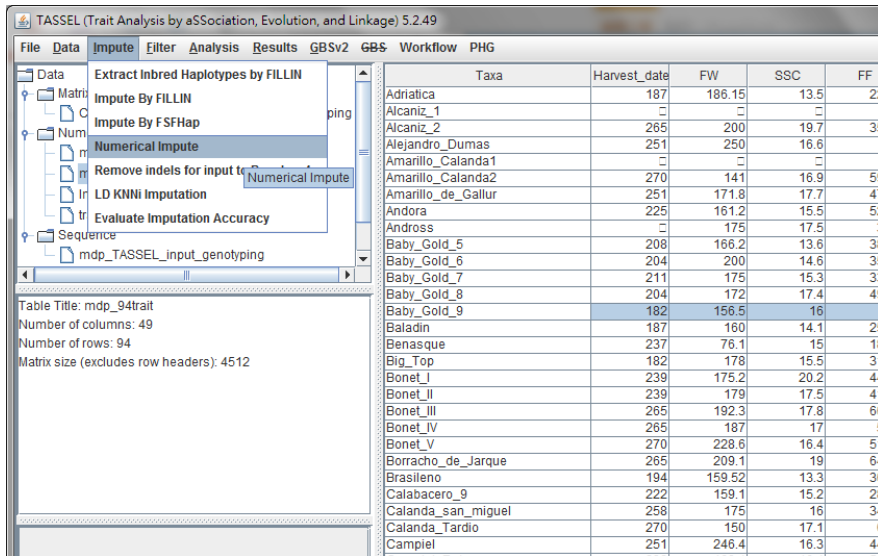


圖16、進行補缺值的運算，補缺值有很多選項，試需要而定，上圖為表現型資料補缺值的示意圖。

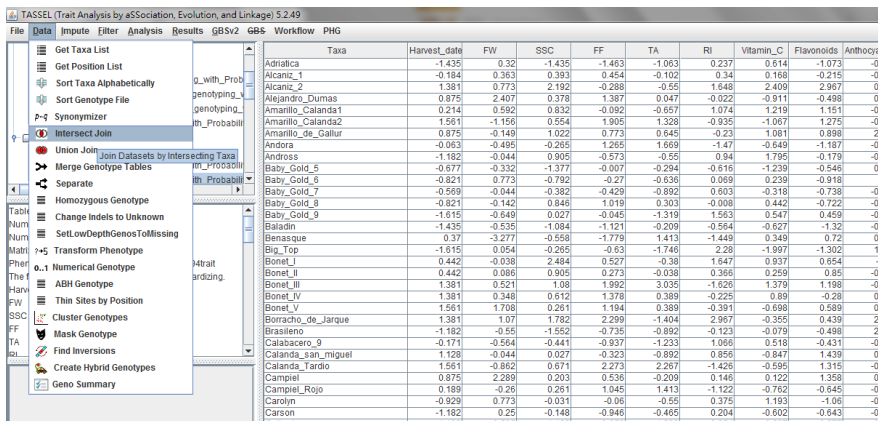


圖17、使用Data中的intersect join將想要分析的檔案組合再一起，接著就可以進行Analysis選項。

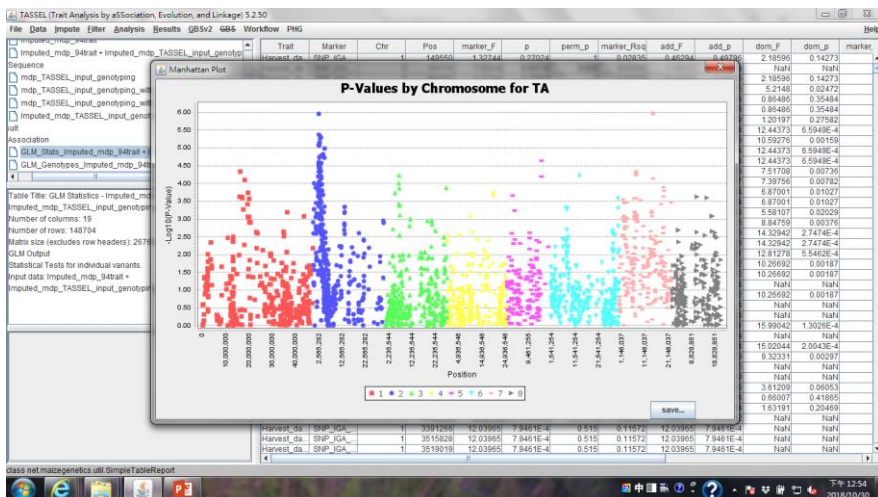


圖18、GLM的分析，以其中一個性狀的manhattan plot，下方的數據資料可以藉由export檔案輸出，進行分析。

四、建議

筆者本次參訪學習分子育種相關研究試驗，除了對當地機構與人脈建立具有良好關係之外，對於相較台西雙方研究有一些初步的體認與比較，對於本次研習也提供相關對於未來國人有前往西班牙交流時的一些建議，分述如下：

一、西班牙風俗民情與相關環境：

西班牙位處南歐，氣候較為溫和，夏季7-9月較為炎熱、酷熱，10月過後，天氣漸冷，然而各地氣候迥異，北邊靠近法國的地區水氣豐沛，綠意盎然，南部地中海地區較為乾燥，10月中旬之後，若受北方冷氣團南下影響，低溫也會將進零度。而筆者訪問期間，氣候異常，9月出現多年來的9月最高溫，南部安達魯西亞地區也強降雨，水患成災，建議未來前往西班牙之國人在出訪時段上，可以提前避開季節交替時期，可以減低衣物攜帶行李的重量與空間。

在西班牙的飲食，最值得注意的是西班牙的吃飯用餐時間和全世界所有的國家都不相同，EEAD早上大約10-11點是咖啡時間，到下午2點才是正式用餐時間的開始，一直工作到晚上6點左右；直到晚上9點才是晚餐時間，而商店禮拜天全部都是沒有營業，因此，在外頭飲食、採買等作息，必須特別注意，星期天只有少數的中國、印度、中東的店家有所營業。在購物方面，西班牙的消費以農產品來說，絕大部分的生鮮食品是比台灣便宜，且品質優良，對於在西班牙生活，在經濟上的開銷相對與台灣相當。

另外，西班牙算是屬於早期大航海時代的殖民強國，但後期因為打仗而衰弱，目前以西班牙為官方語言的國家也多位於中南美洲，而西班牙是最靠近北非的先進國家，因為地理與歷史的因素，西班牙和摩洛哥、突尼西亞、阿爾及利亞等國家交流頻繁，因此，許多店家的服務人員可能是來自摩洛哥等地，抑或當地會有較多的中南美洲學生、居民，因此對於相關國家的歷史、文化應該也要有所了解，可以幫助融入交流關係，對於不同文化種族、或是信仰的人，都應保持開放與接納的態度。

二、EEAD 的研究人員較聚焦於研究與教學且計畫具延續性

在和 Gogocerna 博士交流後，了解到他們僅專注於學術研究、產業發展、社會服務，而非一味政策導向，每一年以該研究是來說，4 年一次一個計畫，不會被砍掉，她都是計畫結束的前一年，再提後續 4 年的計畫，專心執行，而不求計畫數量多，而是追求計畫的目的與品質，不似國內實驗研究單位有太多會議以及過多的管考；另一個較大的差異是 EEAD 的研究人員，可以指導碩博士生，這些碩博士生有求學的動機，也成為研究室的研

發主力人員，最大的差異是，他們的育種研究工作具有延續性，像是水蜜桃收集圃就持續維持 70 多年，經費是 EEAD 維持提供的經常性預算，這對於農業林業的作物育種有很大的省思，一般農林業都是長期，且不見得比得上電子高科技產業的產值，但是民以食為天，最基礎的工作，往往需要靠政府支援，這樣才容易累積農林業的產業能力。

三、收納國際交流生的能力

在 Gogorcena 實驗室只有短暫的 2 個月時間，就已經接納 4 名國際博士生前來交流訪問，其中筆者是台灣目前第一個至 EEAD 訪問的研究人員，而其他 3 名，分別來自突尼西亞以及伊朗，主要進行化學分析以及植物轉錄組研究，預計訪問的時間分別為 3 個月和 6 個月，讓筆者感到訝異的是，收納國際人才的能力，據了解，EEAD 並沒有資助這些國際交流學生，都是透過像是 Research Gate 之類的網站，找到該實驗室的研究領域，進而聯繫是否有交流機會。其實，這樣的研究模式也可以試著引入台灣的相關學研機構，由於台灣的社會發展以及學術能力，在國際上也是在研究競爭力的水平平均之上，尤其是公部門的學術研究機關，是否有機會應用公務門的房舍與研究能量，有計畫性依照法規的方式，收納吸引國際年輕學者，已有目標導向的研究主題，進行交流，透過機關徵選，釋放出相關的名額，舉例來說，若是機關擁有宿舍，則可以開出提供在台住宿或是適當的生活津貼，完成研習交流，必須達成相當的文章或是研究的貢獻，以台灣目前的學術發展，勢必可以吸引很多國際學生(最好是博士生，這樣就是即戰力，可以隨時進實驗室進行操作)，短期可以刺激接待方文化、語言、研究的能力，長期可以培養實驗室雙方的合作關係，可說是透過小小的經費，可能會有得到極大報酬的投資。

四、西班牙是適合建立交流關係的國家

西班牙是一個適合與台灣建立學術交流的國家，一般大眾對於學術交流的首選網網是美國，不可諱言的是，美國的確學術發展領先國際，且語言上也是使用學術使用的第一語言：英語；但是在西班牙，除了要使用西班牙語較為方便，但是反過來思考，透過英語和西班牙人溝通，比較不會有語言弱勢的差異，畢竟英語對於西班牙人來說也是外語，因此，在溝通上，會有平等的感覺，另外一方面，西班牙也很重視與歐洲各國的學術往來，在許多國際領域上，也多有貢獻與長處；而在西班牙生活，對於台灣人來說，兩國經濟能力相近，在生活消費的負擔上也是比較輕鬆，有助提升台灣相關學者前往的動機。