

出國報告 (出國類別：研究)

高經濟型種畜禽之全基因體定序科技與 提高基因選種效益

服務機關：農業部畜產試驗所

姓名職稱：朱家德 助理研究員

派赴國家：美國

出國期間：112 年 9 月 24 日至 10 月 8 日

報告日期：112 年 12 月 5 日

摘要

本次研究以「高經濟型種畜禽之全基因體定序科技與提高基因選種效益」為主題。在美國加州大學戴維斯分校 (University of California-Davis, UC Davis) 與Dr. Huaijun Zhou 及美國農業部農業研究署 (U.S Department of Agriculture-Agricultural Research Service, USDA-ARS) Dr. Hans H. Cheng等2位遺傳育種專家，討論有關「種雞蛋蛋殼品質相關候選基因檢測平台建置」實驗計畫目前成果，Dr. Zhou與Dr. Cheng提及如能擴大試驗動物達1,000隻的性能紀錄與基因型資料，則可以更有效且準確地找到更多影響性能表現之遺傳標記。關於畜試土雞全基因體關聯性分析資料 (Genome-wide association study, GWAS) 分析與討論，Dr. Dailu Guan從旁指導生物資訊軟體的使用，包括建構虛擬伺服器Linux系統的Ubuntu，全基因體關聯分析軟體PLINK與GEMMA，對遺傳標記間連鎖效應的分析LDBlockShow，以及基因功能性分析的Ensembl等軟體。目前已完成畜試土雞全基因體關聯性資料分析，找到影響畜試土雞40週齡產蛋數的候選遺傳標記。未來，會進一步建置候選基因之基因型檢測平臺，來印證遺傳標記對40週齡產蛋數的影響程度。

目錄

壹、 目的	4
貳、 計畫行程	5
參、 計畫過程	6
一、 前言	6
二、 影響臺灣有色肉雞種蛋孵化率基因研究	7
三、 畜試土雞全基因體關聯性分析研究	8
肆、 心得與建議	10
伍、 參考文獻	11
陸、 附件	12

壹、目的

本研究旨在探討全基因體定序技術在高經濟型種畜禽選育中的應用及其對提升基因選種效益的影響。近年來，全基因體關聯性分析 (genome-wide association study, GWAS) 與全基因體定序技術 (whole genome sequencing, WGS) 在農業生物科技領域中發展迅速，尤其在種畜禽遺傳選育上展現出巨大影響潛力。此技術能夠提供全面的基因資訊，從而更精確地識別、篩選與穩固具有特定性能表現的個體，這對於改善畜禽產業的產量、品質以及抗病性等特性至關重要。

貳、計畫行程

本次赴美國舊金山加州大學戴維斯分校行程如下：

日期	星期	地區	內容
9月24日	日	臺灣-美國舊金山	<ul style="list-style-type: none"> ●長榮航空 (RB 8, EVA Air) 的航班於上午 10:15 從台灣桃園機場 (TPE) 起飛，並於上午 6:30 抵達美國舊金山機場 (SFO)。 ●乘坐巴士 (Davis Airporter) 從舊金山機場前往加州大學戴維斯分校，預計於中午 12 點抵達旅館。
9月25日	一	舊金山	與 Dr. Huaijun Zhou 及 Dr. Hans H. Cheng 討論改善臺灣土雞繁殖性狀相關的育種研究。
9月26日	二	舊金山	討論改善臺灣土雞肉質性狀相關的育種研究，並與實驗室的研究員、博士後和科學家交流。參加實驗室會議、系列講座和校園其他研討會。
9月27日	三	舊金山	討論應用全基因組關聯性分析進行基因標記篩選的育種研究：育種研究設計。與研究生、博士後及科學家討論，並參加實驗室會議、系列講座和校園其他研討會。
9月28日	四	舊金山	討論應用全基因組關聯性分析進行基因標記篩選的育種研究：使用生物資訊工具。與研究生、博士後及科學家討論，並參加實驗室會議、系列講座和校園其他研討會。
9月29日	五	舊金山	討論應用全基因組關聯性分析進行基因標記篩選的育種研究：GWAS 分析研習 I。
9月30日	六	舊金山	資料彙整與分析
10月1日	日		
10月2日	一	舊金山	討論應用全基因組關聯性分析進行基因標記篩選的育種研究：GWAS 分析研習 II。
10月3日	二	舊金山	討論應用全基因組關聯性分析進行基因標記篩選的育種研究：GWAS 分析研習 III。
10月4日	三	舊金山	討論應用全基因組關聯性分析進行基因標記篩選的育種研究：GWAS 分析研習 IV。
10月5日	四	舊金山	討論應用全基因組關聯性分析進行基因標記篩選的育種研究：GWAS 分析研習 V。
10月6日	五	舊金山	討論應用全基因組關聯性分析進行基因標記篩選的育種研究：GWAS 分析研習 VI。
10月7日	六	舊金山	<ul style="list-style-type: none"> ●乘坐巴士 (Davis Airporter) 從加利福尼亞大學戴維斯分校前往舊金山機場。 ●長榮航空 (RB 7, EVA Air) 的航班於美國舊金山機場 (SFO) 的下午 13:00 起飛。
10月8日	日	臺灣	長榮航空 (RB 7, EVA Air) 的航班於下午 5:30 抵達台灣桃園機場 (TPE)。

參、計畫過程

一、前言

美國農業部 (U.S Department of Agriculture, 以下簡稱USDA) 管轄的農業研究署 (Agricultural Research Service, 以下簡稱ARS) 以探索從田地到餐桌農產問題的解決策略為目標, 同時為全球農業挑戰提供科學依據的應對方法為宗旨。USDA-ARS的農業研究範圍橫跨15個國家, 執行超過660個研究項目, 農業研究領域成果豐碩。

2019年USDA-ARS副署長Dr. Steven M. Kappes受農業部及畜試所邀請, 來臺分享種畜禽基因選育經驗, Dr. Steven M. Kappes在研討會演講時表示, 美國乳牛產業運用多元化的基因選種技術選拔優良種畜, 可提升30 - 50%不等的性能改進成效, 為產業帶來倍數化躍進。運用分子生物技術進行經濟動物繁殖及生長性狀的選育已經是世界趨勢。全基因組關聯研究 (GWAS) 在整個基因組中檢測數千個遺傳標記, 用以識別與經濟性狀表現程度相關的特定變異。GWAS研究的優點包括深入了解表現型的基礎生物學、估算其遺傳性、計算遺傳相關性、進行臨床風險預測、為藥物開發計劃提供信息, 以及推斷風險因素和健康結果之間潛在的因果關 (Uffelmann *et al.*, 2020)。2019年, Liu等人運用Illumina 60K SNP chips基因晶片進行北京油雞肌內脂肪 (intramuscular fat, IMF) 選育研究, 其G1世代較G0世代提升9.62% IMF。Zhang等 (2021) 運用基因晶片 (50K Porcine v1 Genotyping BeadChip, Neogen) 進行杜洛克豬種IMF選育研究, 結果發現有11個單核苷酸多態性位點 (single nucleotide polymorphisms, SNPs) 可顯著影響豬肉IMF含量, 並且有12個SNPs可顯著影響背脂厚度 (Backfat thickness, BT), 可作為性狀選拔的遺傳標記。故若能運用基因組SNPs資訊與家族系譜資料結合生物資訊與數量遺傳等分析方法, 則可以提升低遺傳率性狀的選拔準確度, 而能有效增加目標經濟性狀的性能表現。

Dr. Hans H. Cheng為USDA-ARS的育種專家聯繫窗口, Dr. Hans H. Cheng帶領家禽病和腫瘤研究室 (Avian Disease and Oncology Laboratory) 研究團隊, 運用全基因體關聯性分析 (Genome-wide association study, GWAS) 與RNA seq等新的基因組策略建置家禽遺傳選種平臺, 並成功應用於家禽抵抗馬立克病 (Mark' s

disease, MD) 的基因遺傳選育。

Dr. Huaijun Zhou任教於美國加州大學戴維斯分校 (University of California-Davis, UC Davis) 動物科學系，因對動物和家禽基因組研究的成果豐碩，被美國國家科學院 (National Academy of Sciences, NAS) 授予2023年NAS食品與農業科學獎 (NAS Prize in Food and Agriculture Sciences)。Dr. Zhou團隊開發了一個高效基因選拔平臺，可以提升家禽對新城雞瘟 (Newcastle Disease, ND) 與熱緊迫 (heat stress) 的抵抗力，有助於解決食品安全和全球氣候變化等關鍵問題。

二、影響臺灣有色肉雞種蛋孵化率基因研究

筆者在UC Davis動物學系Dr. Huaijun Zhou的辦公室與Dr. Hans H. Cheng進行視訊會議，討論「種雞蛋蛋殼品質相關候選基因檢測平台建置」實驗計畫目前成果(圖1)。筆者將112年的研究進度做詳細說明，此次研究以100隻紅羽土雞母雞為試驗動物，收集自30至50週齡種蛋孵化紀錄(圖2)。運用競爭性交替基因特異性聚合酶鏈鎖反應 (kompetitive allele specific polymerase chain reaction, KASP™) 檢測技術進行候選基因之基因型與種蛋孵化率的關聯性分析。初步研究結果發現，有2個候選基因 (*OCX-32*與*PLCZI*) 對不同週齡的種蛋孵化率影響達統計顯著水準 ($P < 0.05$)。 *OCX-32*有利基因型可提升第30與第40週齡種蛋孵化率分別達29.7與22.2%，並且從30至50週齡的平均種蛋孵化率可達91.6%(圖3)； *PLCZI*有利基因型可提升第35週齡種蛋孵化率達18.0 % (圖4)。

聽完筆者簡報後，2位專家明確提種蛋孵化率屬於多基因遺傳影響的數量性狀基因，目前收集30、35、40、45及50週齡種蛋性狀與孵化資料的資料量尚不足夠，不容易找到對性狀有顯著影響的全部候選基因，能找到1至2個有影響潛力的候選基因實屬難得可貴。若能完整收集自30至50週齡各週種蛋性狀的連續性數據資訊，可以有效且全面得獲得影響種蛋孵化率的所有候選基因，對於種雞性能的改善與表現的穩固，提供顯著的貢獻。

2位專家亦提及有關家禽育種實驗所需要的族群樣本數量問題。以現有100隻紅羽土雞母雞的樣本數難有全面且整體的基因體資訊，若能擴大至1,000試驗樣本數則可有效提高基因型分析的準確程度。針對不同週齡的種蛋孵化率資料收集

部分，則需更加全面且連續性的收集性狀資料，如此即能深度了解每個週齡種蛋孵化率的差異，更有益於未來選育高孵化率種雞。對於未來種畜禽基因遺傳育種選育計畫，2人皆表示樂於合作的意願，並願意提供相關的實驗設計與研究分析等協助。

三、畜試土雞全基因體關聯性分析研究

有關「畜試土雞全基因體關聯性分析資料」分析部分，因產蛋數是影響養雞產業利潤的一個關鍵因素。為了全面且精準地了解全基因體SNPs與畜試土雞產蛋性狀表現的關聯性，將GWAS原始資料與美國生物資訊專家Dr. Dailu Guan討論交流，期藉由與美國專家的協助，找到影響產蛋數的潛力基因遺傳標記。

在UC Davis的研究室中，Dr. Dailu Guan從旁指導生物資訊軟體的使用（圖5）。本次研究交流所研習到的GWAS分析架構流程如圖6所示，包括建構虛擬伺服器Linux系統的Ubuntu，Ubuntu 提供一個整合性的界面，由於其開源的特性而受到多數生物資訊學家的喜愛，常被用作各種分析和計算工作。運用GWAS原始資料分析的基礎分析軟體PLINK，PLINK 是一套全功能的基因體關聯分析軟體，特別適合用於大規模基因型資料的處理和分析（Purcell *et al.*, 2007）。首先，需進行SNPs品質控制，剔除了次等位基因頻率小於5%的SNPs，移除了缺失率超過10%的基因型，過濾掉缺失率超過10%的樣本，以及偏離哈溫平衡（Hardy-Weinberg equilibrium, HWE）測試 ($P < 1 \times 10^{-6}$) 的SNPs。PLINK輸出了一個曼哈頓圖（Manhattan plot）和一個分位數-分位數圖（Quantile-Quantile plot, Q-Q plot），以明確區分每一條染色體上的SNPs對產蛋性狀的影響程度（圖7）。GEMMA 是用於基因體關聯研究的軟體，特別是在探討多性狀和混合模型。它可以評估不同的遺傳模型，如固定效應和隨機效應模型。線性混合模型近來因為其在遺傳關聯測試中處理族群分層和相關性方面的強大有效性，而受到相當大的關注（Zhou and Stephens, 2012）。

LDBlockShow 主要用於對遺傳標記間連鎖效應的分析，此項分析為降低未來建置基因型分析成本的重要工具（圖8）。LDBlockShow用於顯示SNPs的連鎖不平衡（linkage disequilibrium, LD）模式和單倍型（haplotype）結構，是基因組學和遺傳學研究的重要工具，能顯著減少電腦運算時間和硬體儲存的使用。此外，

LDBlockShow還能呈現更多的統計分析資訊（例如關聯性顯著程度P值）和基因組區域註釋，使得LDBlockShow成為遺傳學家進行LD和單倍型分析所需快速且方便的生物資訊工具 (Dong *et al.*, 2021)。Ensembl 是研究者獲取高質量基因組資訊的重要來源。Ensembl 提供多種動物基因組的註解資料。這不僅包括基因和蛋白質的註解，還包括變異資料、比對資料和其他功能區域的資訊 (圖9)。

筆者與Dr. Dailu Guan進行95隻畜試土雞產蛋相關性狀的GWAS分析。有關性狀資料部分，包括初產體重、初產日齡、初產蛋重、40週齡產蛋數、40週齡體重、40週齡平均蛋重以及40週齡體增重。結果發現第4號染色體上有16個候選SNPs顯著影響40週齡產蛋數，但其他測量性狀並未發現具影響潛力的SNPs存在。此發現與Wolc *et al.* (2014) 研究結果相呼應，該研究運用客製化基因晶片 (42K Illumina SNP chip) 取得4,500隻海蘭蛋雞全基因體SNPs資訊，並於收集13,000母雞產蛋性能資料後進行GWAS分析，結果發現在第4號染色體有19個SNPs可顯著影響產蛋性能表現。

筆者將在第4號染色體上的16個候選SNPs進行LDBlockShow的連鎖不平衡分析，有效取得16個候選SNPs中具代表性的9個SNPs。經查詢Ensembl的基因功能註解資料庫，發現這9個SNPs分別位於lncRNA、NAA15、NOCT、RAB33B和SLC7A11等5個候選基因的序列上，lncRNA、NAA15、RAB33B和SLC7A11主要為細胞能量代謝的調控因子，而NOCT基因在脂肪生成過程中扮演重要的角色。

藉由這次與美國3位遺傳育種專家的協助，才能讓畜式土雞全基因體分析的結果有突破性的進展，筆者原本以為當前的成果可以做為臺美雙邊研究合作計畫的發表研究報告，但經由Dr. Dailu Guan審閱過多篇GWAS研究稿件的經驗，說明因為此計畫的GWAS分析僅源自95個樣本數，尚無法達到期刊要求最少200個樣本的門檻要求，除非將找到的候選遺傳標記應用於田間試驗，進一步印證其對性能的影響程度，國際期刊才有可能會接受並採納當前的研究成果。

未來，筆者會依據美國3位遺傳育種專家的建議，進一步建置候選基因之KASP™基因型檢測平臺，並擴大田間試驗的性能資料收集，來印證9個候選遺傳標記對40週齡產蛋數的影響程度。

肆、心得與建議

在美國UC Davis學習GWAS期間，Dr. Hans H. Cheng、Dr. Huaijun Zhou及Dr. Dailu Guan給予筆者許多關於育種研究計畫的建議與幫助。Dr. Huaijun Zhou更熱心分享該研究室例行性的研究進度報告與讓筆者與北區分所趙俊炫副研究員參與，讓我們實際明瞭國際遺傳育種領域研究的最新進展，著實拓展筆者的研究視野（圖10）。

目前，國際種畜禽選育公司已跳脫單一或多個基因型的遺傳標記選育技術，而是朝向全基因組的遺傳標記篩選策略精進。因許多的經濟性狀都是屬於多基因（百個基因以上）遺傳的性狀，即使將數個影響性能表現的主效基因納入遺傳選拔，但經過3~5個世代依然會面臨性能衰退的情況，此乃因其他眾多非主效基因未經選拔的負面影響所致，所以若要永續提升且穩固種畜禽的性能表現，則全基因體選拔是箇中關鍵。

有關種畜禽的育種選育技術研究建議，美國生物資訊專家Dr. Dailu Guan提及目前世界各國種畜禽業者已將GWAS作為選種必要技術，並進一步結合系譜與性狀資料等三方面資訊，架構基因組最佳線性無偏估測（Genomic best linear unbiased prediction, gBLUP）選育系統。依據Dr. Dailu Guan與國際種畜禽業者合作的經驗分享，說明運用GWAS與gBLUP能在每個選育世代貢獻顯著的性能改進成效。Dr. Dailu Guan建議如能在臺灣完整收集到符合運用GWAS分析所需的種畜禽性能表現與系譜資料後，赴美國研習3~6個月，應可充分學習並了解此GWAS種畜禽選育技術與gBLUP種畜禽選育策略應用，期可提升臺灣種畜禽遺傳育種的研發量能與產業實質應用效益。

伍、參考文獻

1. Dong, S. S., W. M. He, J. J. Ji, C. Zhang, Y. Guo, and T. L. Yang. 2021. LDBlockShow: a fast and convenient tool for visualizing linkage disequilibrium and haplotype blocks based on variant call format files. *Brief Bioinform.* 22(4): 1-6.
2. Liu, R., M. Zheng, J. Wang, H. Cui, Q. Li, J. Liu, G. Zhao, and J. Wen. 2019. Effects of genomic selection for intramuscular fat content in breast muscle in Chinese local chickens. *Anim. Genet.* 50: 87-91.
3. Purcell, S., B. Neale, K. Todd-Brown, L. Thomas, M. A. R. Ferreira, D. Bender, J. Maller, P. Sklar, P. I. W. de Bakker, M. J. Daly, and P. C. Sham. 2007. PLINK: A Tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81(3): 559-575.
4. Uffelmann, E., Q. Q. Huang, N. S. Munung, J. de Vries, Y. Okada, A. R. Martin, H. C. Martin, T. Lappalainen, and D. Posthuma. 2020. Genome-wide association studies. *Nat. Rev. Methods Primers* 59: 1-21.
5. Wolc, A., J. Arango, T. Jankowski, I. Dunn, P. Settar, J. E. Fulton, N. P. O'Sullivan, R. Preisinger, R. L. Fernando, D. J. Garrick, and J. C. Dekkers. 2014. Genome-wide association study for egg production and quality in layer chickens. *J. Anim. Breed. Genet.* 131(3): 173-182.
6. Zhang, Z., Z. Zhang, F. O. Oyelami, H. Sun, Z. Xu, P. Ma, Q. Wang, and Y. Pan. 2021. Identification of genes related to intramuscular fat independent of backfat thickness in Duroc pigs using single-step genome-wide association. *Anim. Genet.* 52(1): 108-113.
7. Zhou, X., and M. Stephens. 2012. Genome-wide efficient mixed-model analysis for association studies. *Nat. Genet.* 44(7): 821-824.

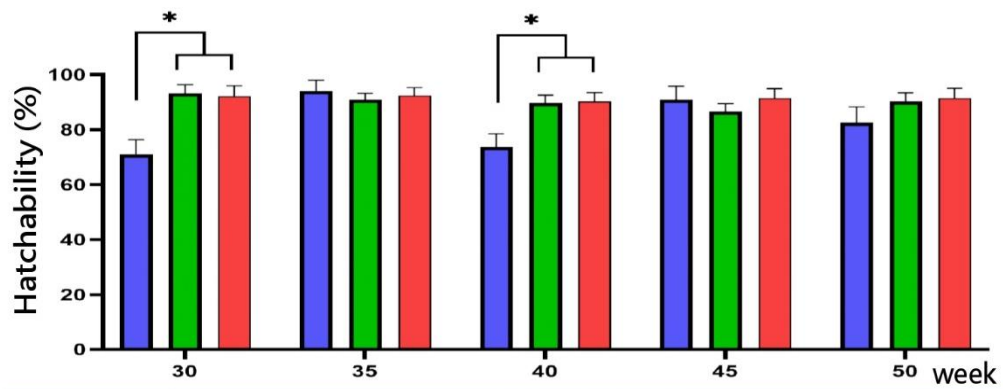
陸、附件



圖 1. 筆者 (右2) 在UC Davis動物科學系Dr. Huaijun Zhou (右1) 的辦公室與Dr. Hans H. Cheng (左) 進行種雞育種計畫視訊會議。

100 hens		Genotype		
母雞		Breed	Method	Gene
Taiwan commercial red feather native chicken		White Leghorn (WL) males and Rhode Island Red (RIR) females	PCR-RFLP	OCX-32
		White Leghorn (WL) and Dongxiang (DX) chickens	GWAS	PLCZ1, ITPR2, IAPP, ABCC9, KCNJ8, CACNA1C
Phenotype		Statistical analysis		
Trait	Content	General linear model (GLM) analysis Formula: $Y_{ij} = \mu + G_i + e_{ij}$. Y_{ij} is the trait measurement value. G_i is the single gene locus effect. e_{ij} is the residual effect.		
Hatchability	30 th , 35 th , 40 th , 45 th , 50 th week of age			

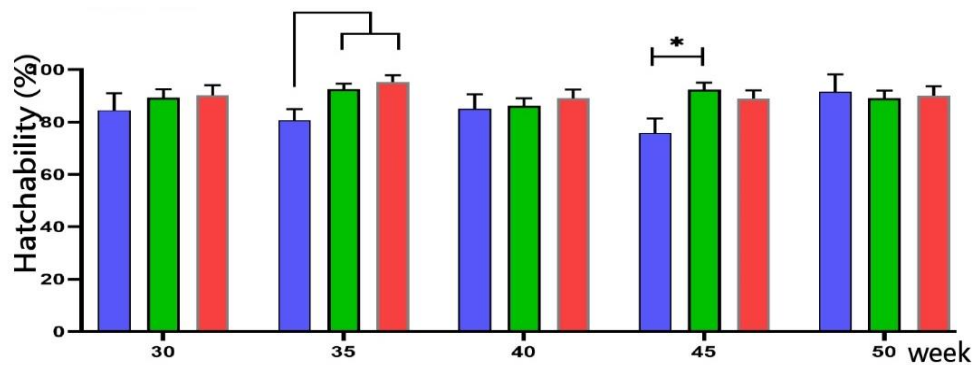
圖 2. 種雞蛋蛋殼品質相關候選基因檢測平臺建置計畫之實驗設計與統計分析。



<i>OCX-32</i> gene	Hatchability (%)				
	30 week	35 week	40 week	45 week	50 week
AA	71.1 ± 5.3 ^a	94.1 ± 3.9	73.8 ± 4.7 ^a	90.9 ± 4.9	82.6 ± 5.7
AC	93.2 ± 3.1 ^b	90.9 ± 2.3	89.7 ± 2.8 ^b	86.6 ± 2.8	90.3 ± 3.1
CC	92.2 ± 3.8 ^b	92.4 ± 2.8	90.2 ± 3.3 ^b	91.5 ± 3.4	91.5 ± 3.5

(ab means P < 0.05)

圖 3. *OCX-32* 基因對紅羽土雞特定週齡種蛋孵化率的關聯性分析。



<i>PLCZ1</i> gene	Hatchability (%)				
	30 week	35 week	40 week	45 week	50 week
AA	84.6 ± 6.4	80.7 ± 4.3 ^a	85.0 ± 5.6	75.8 ± 5.5 ^a	91.6 ± 6.6
AC	89.3 ± 3.2	92.6 ± 2.1 ^b	86.3 ± 2.8	92.3 ± 2.7 ^b	89.0 ± 3.0
CC	90.2 ± 3.9	95.3 ± 2.6 ^b	89.0 ± 3.4	88.9 ± 3.2 ^{ab}	90.0 ± 3.6

(ab means P < 0.05)

圖 4. *PLCZ1* 基因對紅羽土雞特定週齡種蛋孵化率的關聯性分析。



圖 5. 筆者 (右) 與Dr. Dailu Guan (左) 研究交流畜試土雞全基因體關聯性分析。

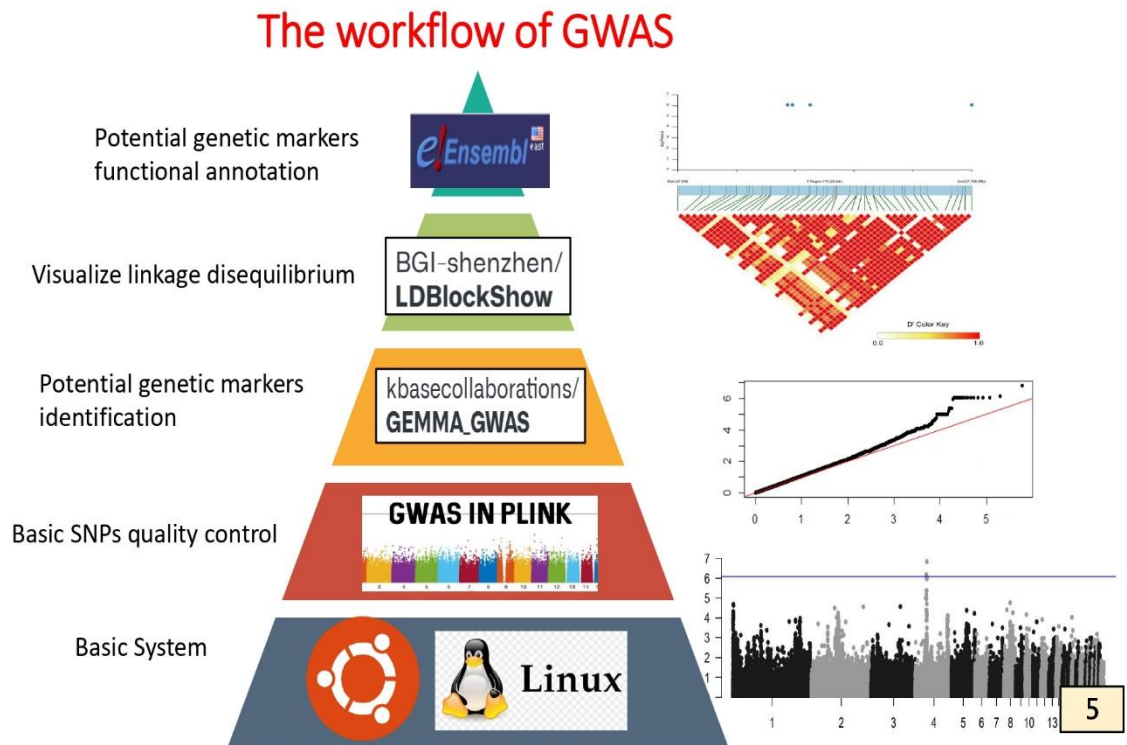


圖 6. GWAS分析流程架構。

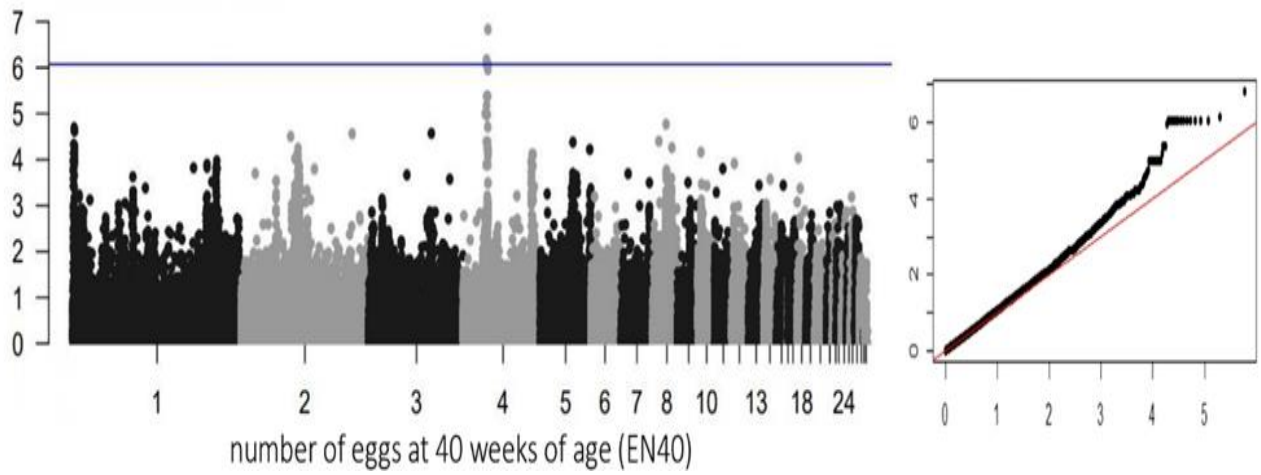


圖 7. GWAS分析Manhattan plot (左) 和 Q-Q plot (右) 的結果。

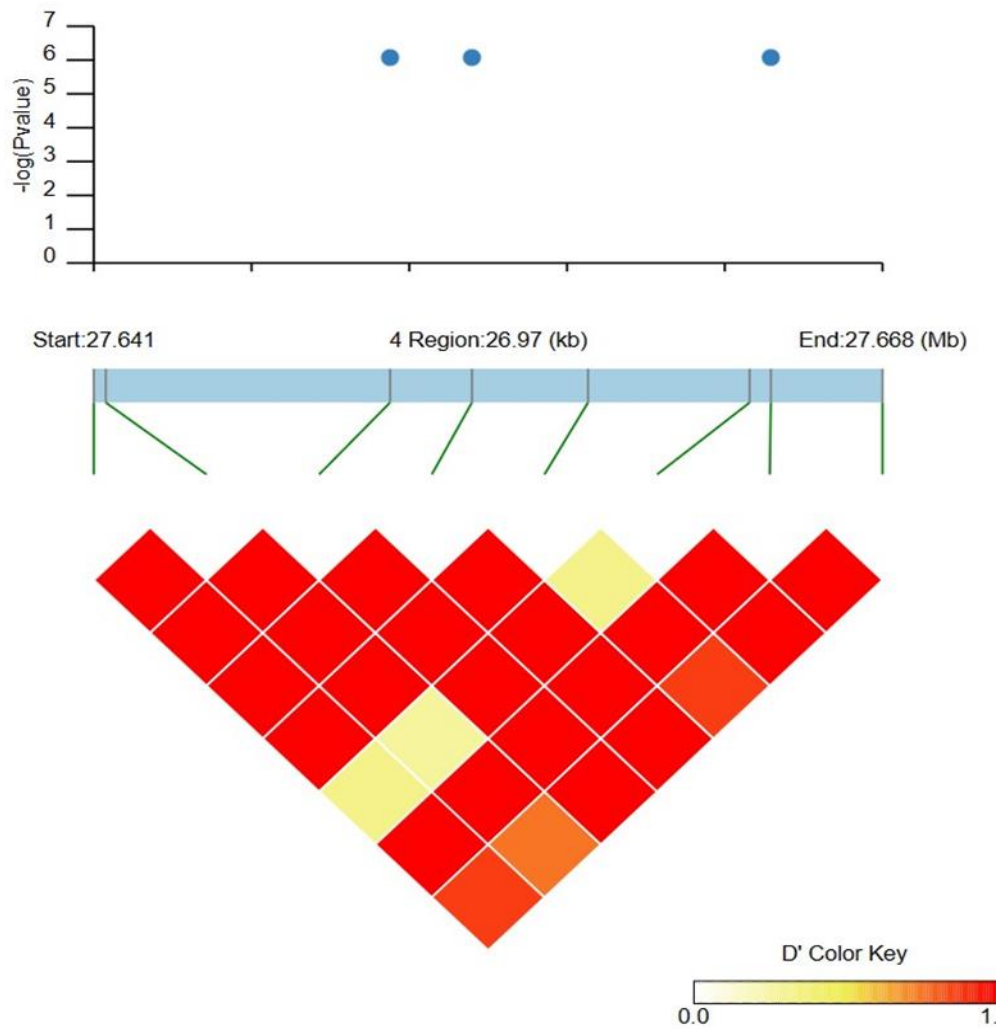
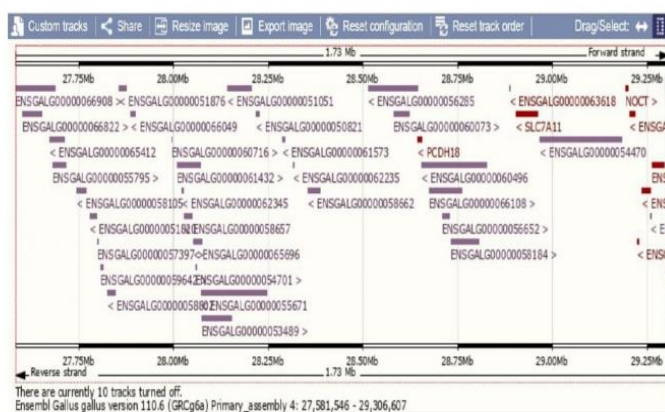


圖 8. 運用LDBlockShow分析的SNPs 連鎖不平衡結果。



Candidate genes	Functional annotation
lncRNA	Gene regulation and cellular processes
NAA15	Acetyltransferase activity
NOCT	Adipogenesis
RAB33B	GTPase activity
SLC7A11	Cystine:Glutamate antiporter activity

圖9. 運用Ensembl取得候選基因的功能註釋。



圖 10. 筆者 (左前2) 與北區分所趙俊炫副研究員 (右1) 參加Dr. Huaijun Zhou (左後1) 實驗室全體學生的實驗進度報告會議。