

國立臺灣大學生命科學院生態學與演化生物學研究所

碩士論文

Institute of Ecology and Evolutionary Biology

College of Life Science

National Taiwan University

Master's Thesis

粗首馬口鱖以高屏溪為界的南北遺傳分化機制

The mechanism of north–south genetic divergence of
Opsariichthys pachycephalus across the Kaoping River

邱宇晨

Yu-Chen Chiu

指導教授：王弘毅 博士

Advisor: Hurng-Yi Wang, Ph.D.

中華民國 114 年 7 月

July, 2025



國立臺灣大學碩學位論文
口試委員會審定書

粗首馬口鱖以高屏溪為界的南北遺傳分化機制
The mechanism of north-south genetic divergence
of *Opsariichthys pachycephalus* across the
Kaoping River

本論文係邱宇晨君（學號R11B44017）在國立臺灣大學生態學與演化生物學研究所完成之碩士學位論文，於民國114年7月17日承下列考試委員審查通過及口試及格，特此證明

口試委員：

臺灣大學生態學與演化生物學研究所(指導教授)

王弘毅 博士

王弘毅

中央研究院生物多樣性研究中心

王子元 博士

王子元

國立中山大學海洋科學系

廖德裕 博士

廖德裕

所長

何傳愷

(簽名)

誌謝



本論文致謝以下合作者的參與與合作。

1.指導教授王弘毅老師:

謝謝老師在過去提供的分析意見、樣本蒐集地以及教授的釣魚技巧，讓我學到許多以前不知道的技巧和分析。並且老師非常的有耐心，感謝老師願意指出我的錯誤與問題。

2. 合作研究者與樣本提供者:

謝謝趙淑妙老師提供的資金，使本次研究能夠使用全基因組定序，得到高品質的序列資料。

感謝廖德裕老師、王子元老師與王晨帆老師提供魚類樣本、DNA 樣本和樣本採集地點，使本研究得以涵蓋臺灣西部大部分河川，並獲得不同年份的樣本資料，有助於深入探討粗首馬口鱖之族群結構與族群歷史。

此外，亦感謝張凱翔同學提供魚類樣本及樣本採集地點，使本次研究能獲得近期河川的樣本。

3. 實驗室成員:

感謝游宗翰、戴睿紘、劉宇恩、彭翊倫、張永朋、陳冠蓁等人的協助。

其中特別感謝戴睿紘學長的幫助，謝謝學長教導辨識與測量魚類的方式，還有查看資料時的技巧與注意事項，在程式遇到問題或結果出現錯誤時也是學長提供了幫助。而且還在資料篩選以及序列分析上提供了寶貴的意見，讓此次研究的結果具有更好的邏輯與嚴謹性。謝謝戴睿紘學長三年來的教導。

另感謝游宗翰與彭翊倫提供樣本及協助進行樣本採集，並提供了一些分析建議還有程式使用方式，節約採集與分析上所花費的時間。其中游宗翰還額外提供了關於定序送樣的經驗以及基礎的族群分析方法。

最後謝謝劉宇愐、陳冠蓁與張永朋提供的意見以及協助，幫助我修正用詞與論述邏輯，使研究得以順利完成。

摘要



粗首馬口鱲 (*Opsariichthys pachycephalus*) 是屬於鮡科 (Xenocyprididae) 的初級淡水魚類，主要分布於臺灣西部河川。以高屏溪為界其粒線體基因以及同功酶基因型能分為南北兩個族群，並且過去的研究中以 *cytb* 確認到有高屏溪以北的個體被放流至南部的狀況。目前粗首馬口鱲缺少對核基因的族群研究，因此將透過全基因組定序(Whole genome sequencing)探討粗首馬口鱲南北族群的高度分化是否有天擇參與，並調查近期的河川雜交狀況。此研究蒐集了臺灣西部從瑪鋈溪至港口溪總共 48 個樣本，定序結果為 86.79% 的平均覆蓋率以及 10.79X 的平均深度。

PCA、演化樹以及 ADMIXTURE 的結果與過去的研究一致。以高屏溪為界分為兩個族群，北部族群可以再分為兩個小族群，分別為苗栗中港溪以北的北部族群與介於中港溪及高屏溪之間的中部族群。高屏溪以南的河川存在許多雜交個體，而高屏溪以北則沒有檢測到任何雜交個體；粒線體與核基因的演化樹不匹配也應證了雜交的現狀。透過族群間的固定 SNP (Fixed SNP) 以及 F3 Statistic，確認到高屏溪、枋山溪、楓港溪、四重溪以及港口溪含有雜交的基因型，高屏溪的個體則屬於對南部族群的多次回交族群，而枋山溪、楓港溪、四重溪以及港口溪則屬於 F2 或是中部族群的多次回交的族群。所以南北族群間還未演化出完全的生殖隔離機制。

在核基因與粒線體的演化樹上分別計算出 0.47% 與 4.67% 的遺傳距離 (Genetic distance)，結合南部族群不論在核基因的 π 值都是三個族群中的最低值，還有粒線體基因組在南北族群間最近共祖時間的高估 (226 萬年)。代表粒線體基因組可能受到基因漂變 (Genetic drift) 的影響而加速的分化。

此研究使用了 F_{ST} 以及族群間核苷酸多樣性(π)比率預測以高屏溪為界南北族群的候選天擇基因，檢測到 Mammalian target of rapamycin (*mTOR*)以及 Thyroid hormone receptor beta (*thrb*)，它們分別在斑馬魚(*Danio rerio*)以及食藻鱗(*Cyprinodon nevadensis amargosae*)上被證實參與溫度調節反應。另外南部族群中還檢測到了顯著的 Thyroid hormone signaling pathway (ko04919)，過去的研究顯示側線鱗數與甲狀腺素有關，而且在臺灣鬚鱯身上也發現了側線鱗與河川溫度的負相關關係。表示南北分化可能與河川溫度適應有關。

關鍵字：族群遺傳、全基因組定序、粗首馬口鱮、雜交、天擇

ABSTRACT



Opsariichthys pachycephalus, a primary freshwater fish species belonging to the family Xenocyprididae, is mainly distributed in rivers of western Taiwan. Based on mitochondrial DNA (mtDNA) and allozyme markers, populations can be divided into two group, northern and southern, using the Kaoping River as the boundary. Previous studies using the cytb gene have confirmed that individuals from north of the Kaoping River were introduced into rivers in the south. However, there is currently a lack of population studies based on nuclear genomic data.

In this study, we applied whole-genome sequencing (WGS) to investigate whether the pronounced differentiation between the northern and southern populations of *O. pachycephalus* is shaped by natural selection, and to assess recent hybridization events across rivers. A total of 48 individuals were collected from rivers across western Taiwan, ranging from the Masu River to the Gangkou River. The sequencing yielded an average coverage of 86.79% and a mean depth of 10.79×.

Principal Component Analysis (PCA), phylogenetic trees, and ADMIXTURE results were consistent with previous findings, identifying two major genetic groups divided by the Kaoping River. The northern group can be further subdivided into a northern subgroup (north of the Jhonggang River) and a central subgroup (between the Jhonggang and Kaoping Rivers). In contrast, many hybrid individuals were detected in rivers south of the Kaoping River, while no hybrids were found in the north. The incongruence between mitochondrial and nuclear phylogenies further supports the presence of hybridization.

Through analysis of fixed SNPs and the f_3 statistic, hybrid genotypes were confirmed in the Kaoping, Fangshan, Fenggang, Sichongxi, and Gangkou Rivers. Individuals from

the Kaoping River were identified as multi-generation backcrosses to the southern population, while those from the other southern rivers were either F2 hybrids or backcrosses involving the central subgroup. These findings suggest that complete reproductive isolation between the northern and southern populations has not yet evolved.

Genetic distances calculated from nuclear and mitochondrial phylogenies were 0.47% and 4.67%, respectively. Moreover, the southern group exhibited the lowest nucleotide diversity (π) among the three identified subgroups. In combination with an overestimated mitochondrial time to most recent common ancestor (TMRCA) between north and south populations (2.26 million years), these results suggest that mitochondrial divergence may have been accelerated by genetic drift.

We further used F_{ST} and inter-population nucleotide diversity ratios (π ratios) to identify candidate genes under selection between the northern and southern populations. Notably, genes such as Mammalian target of rapamycin (*mTOR*) and Thyroid hormone receptor beta (*thrb*), both previously implicated in thermal adaptation in *Danio rerio* and *Cyprinodon nevadensis amargosae*—were identified. Additionally, the southern population showed significant enrichment of the thyroid hormone signaling pathway (ko04919). Previous studies have associated lateral line scale count with thyroid hormone levels, and in *Candidia barbatus*, a negative correlation has been found between scale count and river temperature. These findings suggest that the observed north–south genetic divergence may be associated with adaptation to local thermal environments.

Keywords: Population genetics, whole-genome sequencing, *Opsariichthys pachycephalus*, hybridization, natural selection

目次



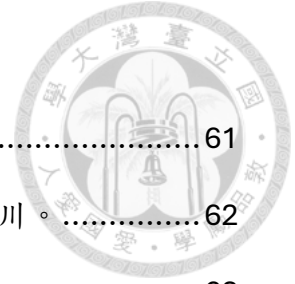
論文口試委員會審定書.....	i
誌謝.....	ii
摘要.....	iv
ABSTRACT.....	vi
目次.....	viii
表次.....	x
圖次.....	xi
第一章 前言.....	1
第二章 材料與方法.....	4
2-1 樣本來源.....	4
2-2 DNA 萃取、聚合酶連鎖反應 (PCR) 以及定序.....	4
2-3 全基因組定序 (Whole genome sequencing).....	5
2-4 粒線體全基因組組裝與分化計算.....	6
2-5 族群結構分析、親緣關係樹、分化計算.....	6
2-5-1 篩除高度連鎖點位以及族群結構分析.....	6
2-5-2 構建演化樹.....	7
2-5-3 分化計算.....	7
2-6 雜交檢測.....	7
2-6-1 族群的基因分化程度與基因多樣性.....	7
2-6-2 雜交族群檢定.....	8
2-6-3 雜交區段檢測.....	8
2-7 適應性基因檢測以及族群統計資料.....	9
2-8 族群歷史.....	10
第三章 結果.....	11
3-1 <i>cytb</i> 結果及 WGS 定序.....	11
3-2 族群結構.....	11

3-3 雜交鑑定.....	13
3-4 ABBA-BABA statistic	14
3-5 族群歷史、群分化時間及演化樹分化程度.....	14
3-6 適應性基因檢測.....	16
第四章 討論	17
4-1 南北族群的分化時間與分化程度.....	17
4-2 南北族群的雜交狀況.....	18
4-3 南部回交族群保留的基因滲流片段與免疫有關	19
4-4 南北族群的分化與溫度具有關聯性.....	19
第五章 結論	22
第六章 參考文獻	23
表	29
圖	61

表次

表一、粗首馬口鱲樣本採集地及各河川 <i>cytb</i> 定序數量	29
表二、粗首馬口鱲樣本採集地點、地圖代號及各河川全基因組定序數量	30
表三、全基因組定序樣本與粒線體譜系	31
表四、根據 SNP 計算出的族群每個點位的平均 π 值以及族群內個體數量	34
表五、透過 ADMIXTOOLS 計算的 f_3 statistic	35
表六、透過 ADMIXTOOLS 進行 ABBA-BABA 測試結果	36
表七、透過 LOTER 計算出的個體單倍體混合比例	37
表八、ABBA-BABA 中顯著富集的 GO	38
表九、ABBA-BABA 中顯著富集的 GO 中包含的基因	43
表十、依據核基因全基因組以及粒線體全基因組估算的基因距離	46
表十一、粒線體與核基因的 π 值	47
表十二、南部族群中顯著的 GO	48
表十三、南部族群中顯著的 GO 中包含的基因	52
表十四、北部族群的富集的 GO	55
表十五、北部族群中顯著的 GO 中包含的基因	58
表十六、南部族群顯著的 KEGG	60

圖次



圖一、粗首馬口鱧的照片	61
圖二、台灣地圖及河川，紅色的點位代表有採集的河川。	62
圖三、由 cytb 構建的 Neighbor-Joining tree	63
圖四、序列的對比率、覆蓋率與定序深度	64
圖五、全基因組構建的 PCA	65
圖六、CV error 值	66
圖七、全基因組構建的 ADMIXTURE	67
圖八、全基因組構建的最大似然樹	68
圖九、由全粒線體基因組構建的 Neighbor-Joining tree	69
圖十、基於全基因組所有點位計算的 F_{ST}	70
圖十一、基於全基因組所有點位計算的 π 以及 Tajima's D	71
圖十二、由每個點位估算的基因組平均 Weir and Cockerham weighted F_{ST}	72
圖十三、由族群間固定的且沒有缺失的 SNP 構建的三角散點圖	73
圖十四、由族群間固定的且沒有缺失的 SNP 構建的 100 kb 區塊的平均雜交比率	74
圖十五、SNP 構建的三角散點圖	75
圖十六、SMC++ 的預測結果	76
圖十七、GONE 的近年的有效族群預測結果	77
圖十八、PopLDdecay 的結果	78
圖十九、 F_{ST} 與 π ratio 篩選可能受到天擇區塊的示意圖	79

第一章 前言



親緣地理學 (Phylogeography) 是一門整合遺傳譜系資料與地理分布資訊，以研究物種內部與物種間演化歷史與族群結構的學問，它透過分析如線粒體 DNA 的親緣結構，來揭示不同地理族群間的遺傳差異與分化歷程[1]。淡水魚的擴散受限於河川，族群結構反映過去的冰河時期與河川匯流等地理事件[2]。臺灣根據淡水魚類物種分布將西部地區以高屏溪為界分為了南與北兩個區域[3]。反應出臺灣淡水魚類的分布與地理事件關聯性。

過去多項研究透過粒線體 DNA 發現臺灣淡水魚族群具有明顯的南北分化現象。高身小鰾鮡 (*Microphysogobio alticorpus*) 透過粒線體 D-loop 及 *cytb* (cytochrome b) 基因分析，顯示其高屏溪族群與其他地區族群呈現明顯分化[4]；埔里中華爬岩鰍 (*Sinogastromyzon puliensis*) 的 D-loop 序列亦呈現高屏溪為界的族群分化[5]。這些物種可能是在冰河時期透過陸橋遷徙進入臺灣，並隨著當時河川間的連通性擴散至各地河川。臺灣在末次冰河期的時海平面下降並露出臺灣海峽，令河川間相連[6]。進一步支持臺灣淡水魚分布與冰河時期遷徙事件的關聯。

粗首馬口鱮 (*Opsariichthys pachycephalus*) (圖一) 是臺灣的初級淡水魚類，主要分布於臺灣西部，在同功酶 (allozyme) 上以高屏溪為界分為南北兩個族群，北部族群可以進一步分為北與中兩個姊妹群[7]。透過 D-loop、側線鱗數量還有下顎顏色，以高屏溪為界被分為粗首馬口鱮以及高屏馬口鱮兩個物種[8]。過去推測粗首馬口鱮是在冰河時期，海平面下降時高屏溪以北的族群可以相互交流，僅有南部族群因高屏溪外海的海底峽谷阻隔而難以與北部族群進行基因交流，並且發現了一些可能受到天擇的同功酶[7]。南部族群可能受到了天擇與地理隔離的影響而高度分化。

近期粗首馬口鱮南部族群中，除東港溪與士文溪外，觀察到多個個體帶有北

部型的 *cytb* 基因型[9]，顯示粗首馬口鱲與高屏馬口鱲近期因人為引入發生了次級接觸（Secondary contact）。在過去的研究中淡水魚因為人為因素而造成種間雜交的案例中，人為造成的雜交事件佔到全部雜交事件的一半，包含物種引入、魚類養殖以及棲地破壞等[10]，此次的人為引入很可能造成種間的雜交。

基因組島（Genomic island）指的是族群或物種間高度分化的區段，這些區段可能包含種化或受選擇的基因。當物種間發生了次級接觸時，物種中的中性區段會被置換，而種化與受選擇的基因可以抵抗基因流並被保留下來。種化基因會造成雜交個體的適存性（Fitness）下降或是使雜交後代不孕阻止基因流[11]；受到天擇的基因則會增加連鎖不平衡（Linkage Disequilibrium）進而抑制基因滲流[12]。而基因島也有可能是基因漂變（Genetic drift）造成的高度分化區段，例如在島嶼灰狐（*Urocyon littoralis*）的族群研究中發現小族群中包含基因漂變造成的高度分化區段[13]。理論模型指出，在異域分化後發生次級接觸時，中性區域的遺傳分化會因基因流而迅速被侵蝕，受選擇影響的基因區段能維持較高的 F_{ST} [14]。當前可能出現南、北族群雜交的粗首馬口鱲，提供了一個理想系統，可用於探討自然選擇在維持遺傳分化中的角色。

過去粗首馬口鱲的研究集中在粒腺體與同功酶，從雜交以來也僅確認到了二次級接觸，而不確定南北族群是否能雜交，因此本研究使用全基因組定序來確定族群間的族群結構、雜交情況以及可能的受天擇的基因。定序瑪鍊溪至港口溪共 48 隻個體以及作為外群的 2 個長鰭馬口鱲（*Opsariichthys evolans*），將序列對齊到粗首馬口鱲的參考基因組上。確認族群間的分群以確認族群分群是接近同功酶的北、中及南分群，還是 D-loop 中的南北分群。透過雜交檢測找到個體是否有回交（Backcross）或早期雜交個體，確定兩個族群間是否有強烈的生殖隔離，並找到基因滲流（Introgression）的區段以確定雜交對南部族群的影響。最後以高度分化的基因組島來尋找可能受到正向選擇的區域，用於推斷南北部的分化是單純

由冰河時期後族群間的地理隔離造成，還是有正向選擇的參與。



第二章 材料與方法



2-1 樣本來源

本實驗中使用的樣本為 2014/7/18 至 2023/12/29 期間採集的粗首馬口鱖，包含 19 條河川，共計 124 條個體（表一），分別為瑪鍊溪（Masu Creek）、基隆河（Keelung River）、橫溪（Hengxi）、大漢溪（Dahan River）、中港溪（Jhonggang River）、後龍溪（Houlong River）、大安溪（Da-an River）、大甲溪（Dajia River）、烏溪（Wu River）、濁水溪（Jhuoshuei River）、朴子溪（Puzih River）、高屏溪（Kaoping River）、東港溪（Donggang River）、林邊溪（Linbian Rive）、士文溪（Shihwun River）、枋山溪（Fangshan River）、楓港溪（Fongkang River）、四重溪（Sichongxi River）以及港口溪（Gangkou River）。魚類樣本來源包括中研院王子元研究技師、國立中山大學的廖德裕教授提供的 DNA 或是酒精浸泡樣本，以及透過垂釣取得的個體。垂釣的樣本會透過冰塊冷凍送至實驗室，將個體置於 50 毫升離心管中並注滿 75%酒精保存於-20°C冰箱，接著每天更替 75%酒精共計 3 次以利於樣本永久保存。

2-2 DNA 萃取、聚合酶連鎖反應（PCR） 以及定序

樣本將透過鑷子以及剪刀取出肌肉組織，以 ddH₂O 洗去酒精後用餐巾紙吸乾肌肉組織內的水分，隨後置於 1.5 毫升的試管中，接著使用 Favrogen 的 Prep™ Tissue Genomic DNA Extraction Mini Kit 進行 DNA 萃取。

以 Cytb 引子（primer）對其中 109 尾個體的 DNA 樣本進行 PCR，序列分別為：

Cytb-1:

5'- YYT AAC CRR GAC YAA TGA CTT GA-3'

Cytb-2:

5'- GTT TAG AAT TCT GGC TTT GG-3'

[15]

PCR 中使用 Kapa Taq PCR kit 每個樣本的配比為 16.25 μ l ddH₂O、2.5 μ l Buffer A、2 μ l MgCl₂、2 μ l dNTP、Primer forward/reverse 0.5 μ l、0.25 μ l Kapa Taq polymerase。擴增反應的溫度設置為 Initial Denaturation 為 95°C 持續 3 分鐘，後續的 35 個循環，每個循環的設置為：

1. Denaturation 95 °C， 30 秒

2. Annealing 47 °C， 1 分鐘

3. Extension 72 °C， 1 分鐘

Final Extension 部分為 72 °C 持續 3 分鐘，完成 PCR 的擴增過程，產物保存於 4°C 冰箱並送至基龍米克斯進行定序。

2-3 全基因組定序 (Whole genome sequencing)

選用 19 條河川中的 48 尾粗首馬口鱖以及作為外群的 2 尾長鰭馬口鱖由基龍米克斯進行建庫，送至基隆米克斯以 Illumina NovaSeq X Plus 以及 Illumina NovaSeq 6000 進行雙端 150 bp 定序 (圖二、表二)。定序得到的 FASTQ 由 fastp v0.23.4[16, 17]修剪 (Trimming) 與過濾 (Filter)，指定 Trimmomatic v0.39[18]的 TruSeq3-PE.fa 為轉接子 (Adapters) 參考，修剪長度大於 20 個鹼基的 ploy-G、刪去 40%的鹼基 Phred 品質小於 20 且鹼基數量小於 80 的序列(reads)以及從 3'端至 5'端以每 4 個鹼基為一個窗口，刪去平均 Phred 品質小於 25 的窗口。

篩選完的序列透過 BWA v0.7.17[19, 20]的“bwa mem”比對到由實驗室組裝的粗首馬口鱖參考基因組[21]上，接著以 SAMtools v1.20[22]的“fixmate”功能去除錯誤的標籤；接著使用“markdup”以及“rmdup”標記並刪去重複 reads，產生的 BAM 檔會透過 gatk 4.5.0.0[23]的“AddOrReplaceReadGroups”功能添加個體名稱。然後使用 BCFtools v1.20[22, 24]的“mpileup”以及“call”生成個體的全基因組 VCF

檔案，並確保每個位點的 Mapping Quality 和 Base Quality 均不低於 20。最後以“merge”將所有 VCF 合併為一個檔案，透過“view”篩選出 SNP 的點位。

使用 VCFtools v0.1.16[25]進行篩選並保留符合條件的點位：(1) 位於染色體上的 SNP；(2) 缺失數據小於 10%的點位；(3) 次要等位基因數量 (Minor allele count) 大於等於 2；(4) 雙等位基因 (Bi-allelic) 的點位；(5) 個體的點位深度大於等於 5；(6) 個體的點位深度小等於自身平均值加上兩倍標準差。

2-4 粒線體全基因組組裝與分化計算

粒線體全基因組之組裝與註解係使用 MitoZ [26]，此工具可自動化執行動物粒線體基因組的組裝與功能註解，以未篩選的定序資料 (FASTQ 格式) 作為輸入。依據其使用說明，本研究設定以下參數：(1) 指定參考資料庫為脊索動物 (Chordata)；(2) 讀長 (read length) 設為 151 bp；(3) 使用輸入檔案中前 3 Gb 的定序資料進行分析。

組裝完成後，為統一粒線體基因組的起始位置，將特定序列“GCTAGCGTAGCTTAATTAAA”作為 5' 端標準，對所有組裝結果進行序列端點對齊。接續以 MEGA X [27] 將所得序列進行對齊 (Alignment)，然後將所有樣本依照譜系分為南北兩群，透過 p distance 模型進行 π 及遺傳距離 (genetic distance) 的統計。為了推估族群間的分化時間，使用 BEAST v2.7.7 [28]對所有粒線體樣本進行貝氏演化分析，採用 JC69 模型、1,000 萬次 Markov 鏈蒙地卡羅模擬 (MCMC)，並以 1.05% 每百萬年的演化速率[29]作為分子時鐘設定。最終以 Tracer v1.7.2[30] 估算共祖時間 (TMRCA) 作為族群分化時間的依據。

2-5 族群結構分析、親緣關係樹、分化計算

2-5-1 篩除高度連鎖點位以及族群結構分析

為了避免高度連鎖的 SNP 對族群結構的影響，於是使用 PLINK v1.90b7.1[31] 以 10 kb 鹼基為一個窗口 (Window size) 以及 1 SNP 的步長 (Step size) 計算並

標記 r^2 高於 0.2 的連鎖 SNP (`--indep-pairwise 10kb 1 0.2`)，接著以 VCFtools 篩選去除，最後透過 PLINK 推算 PCA。

ADMIXTUREv1.3.0[32]透過去除連鎖的 SNP 推算個體的祖先型組成，ADMIXTURE 分析 1 至 10 個祖先族群 (K)，由交叉驗證錯誤 (Cross validation error, CV error) 確定最佳的族群分群，並採用分群為 2 至 5 的結果，接著透過 R 的 ggplot2 包繪製結果。

2-5-2 構建演化樹

為了進一步確定族群結構，需要建構一個演化樹，使用包含外群長鰭馬口鱮的 VCF，以沒有篩除連鎖的 SNP 進行分析。透過 vcf2phylip.py[33]將未篩除連鎖的所有 SNP 串聯在一起並轉換為 phylip 格式後，親緣關係樹由 IQ-TREE v2.3.6[34] 構建，選擇 GTR+I+R+F 模型繪製最大似然樹 (Maximum Likelihood tree)。並且會透過演化樹的分支長度 (Branch length) 估算兩個族群間的基因遺傳距離，算式為使用點位數乘以分支長度並除以基因組全長。

cytb 序列以及組裝的粒線體 DNA 由 MEGA X 以 Kimura 2-parameter model 作為替換模型 (Substitution Model) 構建 Neighbor-Joining tree，並進行 1000 次拔鞋法 (Bootstrapping)，只使用沒有任何插缺變異 (indel) 的對齊序列進行建構。

2-5-3 分化計算

使用篩掉 90% 缺失點位的全點位 (All site) 檔案透過 VCFtools 進行 F_{ST} 、核苷酸多樣性 (π) 以及 Tajima's D 的計算，以每 100 kb 為一個區塊計算結果並透過 ggplot2 包繪製結果。

2-6 雜交檢測

2-6-1 族群的基因分化程度與基因多樣性

為了確定能夠作為純種的河川，需要計算高屏溪以南各河川的基因分化程度

及基因多樣性。以 VCFtools 計算每個點位成對的 (Pairwise) Weir and Cockerham weighted F_{ST} 、每個點位的核苷酸多樣性值及平均值及兩個統計量的所有點位平均值，取其中 F_{ST} 最高且 π 值最低的族群視作高屏溪以南雜交比例最低的族群。

2-6-2 雜交族群檢定

將 SNP 的 VCF 作為輸入檔案，以 ADMIXTOOLS v7.0.2 [35] 計算 f3-statistic 與 D-statistic 以檢測潛在的雜交事件與基因流現象。f3-statistic 用於判斷目標族群是否為兩個親代族群的雜交後代，設置族群關係為 (A,B;C)，其中 C 為待測族群，A 與 B 分別為推定的純種群體。當 f3 統計值為負，且 Z 分數小於等於 -3 時，表示 C 很可能為 A 與 B 的雜交後代。

D-statistic 用於評估兩族群間是否存在顯著的基因流，族群假設為 (P1,P2;P3,O)。其中 P1 與 P3 為純種參考族群，其中 P1 為南部的族群，P2 為需要檢測的族群，P3 為北部的族群，O 為外群（此處設為長鰭馬口鱮 *Opsariichthys evolans*）。當 D 值為負且 Z 分數小於等於 -3，則表示 P2 與 P3 之間存在顯著的基因流。

為更進一步評估各河川族群的雜交程度，本研究選擇純種族群間無缺失且於兩群體間固定差異的位點 (fixed SNPs) 作為資料集，計算每個個體的雜交比例與雜合性 (Heterozygosity)。上述結果將以 R 語言繪製三角形散點圖 (Triangle plot)，以視覺化個體在祖源組成空間中的位置與雜交程度。

2-6-3 雜交區段檢測

為分析基因組中可能的雜交區段，將基因組劃分為每個 100 kb 的區塊 (Block)，並計算各區塊內的平均雜交比率。考量族群間固定 SNP 的可能令存在偏差，因此進一步使用 BEAGLE v5.4 [36, 37] 以預設參數對所有樣本進行單倍型相位推定 (Haplotype phasing)，並以純種族群作為參考族群，透過 LOTER [38, 39] 執行局部祖源推斷 (Local ancestry inference)。LOTTER 的結果將用於估

算每位個體在兩條染色體上的局部祖源比例與平均雜交比率。

此外，為偵測不同河川間可能共享的基因滲流 (Introgression) 區段，首先使用 `parseVCF.py` 將原始 VCF 檔轉換為 GENO 格式，並使用 `ABBABABAwindows.py` 工具，以 100 kb 視窗與 10 kb 步長進行滑動分析，計算各區段的 `fd` (Introgression fraction) 值，該指標源自 ABBA-BABA 測試，用以量化基因滲入程度。

為確保結果合理性，排除所有 `fd` 值小於 0 或大於 1 的區段，因為這些數值在量化 P2 與 P3 之間的雜交程度時沒有意義。隨後，對每個視窗計算 `Z` 分數，並假設其為標準常態分布以轉換為 `P` 值，再以 FDR (False Discovery Rate) 進行多重檢定校正。最終篩選 $FDR < 0.05$ 的視窗，定義為具有統計顯著性的基因滲流的區段。接著，使用 R 套件 `IRanges` 將重疊或相鄰的顯著視窗整合為連續的雜交片段。確定區段位置後，透過比對至參考基因組之 GFF 註解檔，確定每個片段中所涵蓋的基因。基因會透過 R 套件 `clusterProfiler` [40] 執行功能富集分析 (enrichment analysis) 以確定基因滲流區域的主要功能。

2-7 適應性基因檢測以及族群統計資料

確認族群後，透過 `VCFtools` 以每 100 kb 為一個區塊計算 F_{ST} 、 π 以及 Tajima's D ，並計算平均值以及繪製盒狀圖。為了瞭解南北族群的連鎖不平衡衰減狀況，以 `PopLDdecay v3.43` [41] 檢測高屏溪南北之間的連鎖衰減差異。

為了鑑定高屏溪以南的候選適應性基因，本研究使用 `VCFtools` 計算每 100 kb 視窗、以 10 kb 為步長的 F_{ST} 與 π (nucleotide diversity) 值，並移除視窗內 F_{ST} 為負值的區段。南部族群包括高屏溪、東港溪、林邊溪與士文溪，並與高屏溪以北所有樣本分別配對計算 Weir and Cockerham 加權 F_{ST} 與各自族群內的 π 值。 π ratio 定義為北部族群的 π 會除以南部族群的 π 值，數值越高代表南部群體在該區段的遺傳多樣性相對越低，可能受到正向選擇作用；反之，若 π ratio 數

值偏低，則可能反映北部族群在該區段受到選擇壓力。將視窗中 F_{ST} 值排名前 5% 的區段，分別與 π ratio 前 5%及後 5%之交集視為候選適應性區段。前者代表南部群體的候選適應性區段，後者則代表北部群體的候選適應性區段。兩組候選區段將分別進行功能富集分析，以探討南北族群潛在的適應性差異。兩種交集將分別用於富集分析，確認高屏溪南北的候選適應性功能。

GO enrichment 的背景基因集為全基因組註解基因；KEGG pathway 則透過 eggNOG-mapper[35]，使用 DIAMOND 比對方式，並以脊椎動物與硬骨魚為主要參考庫（包含人類、老鼠及魚類）來構成作為背景資訊的 KEGG 資料。兩者的富集分析皆由 clusterProfiler 執行，前者透過“enrichGO”功能進行富集，後者由“enricher”功能進行富集並採納以 KO 開頭的物種間通用生理路徑作為 KO term。

2-8 族群歷史

本研究使用 GONE[42]與 SMC++[43] 兩種方法推估粗首馬口鱲的族群歷史。GONE 採用基於連鎖不平衡（Linkage Disequilibrium, LD）的方法，可用於推估近百世代內的有效族群大小變動；SMC++ 則採用順序馬可夫裔合模型（Sequentially Markovian Coalescent, SMC），適用於推估數千年至數百萬年前的族群歷史。兩種方法皆以保留單一變異型（singleton）的原始 VCF 檔案作為輸入資料，由於臺灣鬚鱲（*Candidia barbatus*）大約在鱗片 1 輪時成熟，最年長的個體為 3 輪[44]，統一設定世代間隔（generation time）為 2 年。此設定係根據近緣物種臺灣鬚鱲的生命史特徵推估，其成熟年齡為 1 歲，最大觀察年齡為 3 歲，故依下式計算世代跨度：

$$\text{世代跨度} = (\text{成熟年齡} + \text{最大壽命}) / 2 = (1 + 3) / 2 = 2 \text{ 年}$$

GONE 所需的重組率(recombination rate)設定為斑馬魚參考值 1.8 cM/Mb；SMC++ 所需的突變率（mutation rate）則使用魚類平均值 5.97×10^{-9} 變異/位點/世代[45]。此外，SMC++ 透過 split 指令推估不同族群間的分歧時間(split time)。

第三章 結果



3-1 *cytb* 結果及 WGS 定序

根據 *cytb* 在刪去訊號不佳的前後端後，保留了 1089bp 作為對齊的區段。基因構建的親緣關係樹顯示，高屏溪以北採集的樣本皆屬於北部特有譜系 (Lineage)，未檢出來自南部的粒線體譜系；相對地，南部河川中部分樣本中則檢測出北部的譜系 (Lineage A)。由於過去的 D-loop 的結果中顯示高屏溪以南不存在北部譜系的個體[8]，引此推測是由高屏溪以北往高屏溪南部的人為放流 (圖三)。此外，高屏溪以北樣本未觀察到具南部譜系 (Lineage B) 者，推論其為未受雜交影響的純種族群。

全基因組定序部分，48 隻個體共產出約 596.6 Gbp 的原始資料 (表四)，平均比對率 (mapping rate) 99.7%、定序深度為 10.79×、平均基因組覆蓋率為 86.79% (圖四)，比對率與覆蓋率不低。經過品質過濾與篩選後，最終保留 12,509,778 個高品質 SNP 用於後續分析。

3-2 族群結構

為了避免連鎖不平衡的區段影響族群結構的分析，我透過 PLINK 篩除連鎖不平衡後，共保留 665,152 個獨立的 SNP，用於後續的 PCA 與 ADMIXTURE 分析。

根據 PCA 的結果，PC1 可明確區分高屏溪以南與以北的樣本，解釋總變異量的 11.25%；PC2 則進一步將高屏溪以北的樣本區分為北部 (中港溪以北) 與中部 (中港溪至朴子溪) 兩個族群，其變異解釋率為 5.28% (圖五)。此結果顯示南北族群之間的遺傳分化程度高於北部與中部族群之間的分化。部分採自高屏溪以南河川的樣本，在 PCA 中分布於中部與南部族群之間，呈現中間型遺傳結構，推測可能為具雜交背景之個體。

綜合 PCA 結果與樣本的地理分布，本研究將所有樣本劃分為三個主要族群：北部族群（North，涵蓋瑪鍊溪至大漢溪）、中部族群（Middle，涵蓋中港溪至朴子溪）、以及南部族群（South，涵蓋高屏溪至士文溪）。

在 ADMIXTURE 分析中，交叉驗證誤差（cross-validation error, CV error）顯示最佳分群數為兩群（ $K=2$ ）（圖六），此時族群以高屏溪為界區分為北部與南部兩大類群。分為兩群以及三群時，能觀測到許多樣本具有明顯的北部祖先型信號，而當分為四群時（ $K=4$ ），混合北與南基因型的枋山溪被錯誤的分為另一個族群（圖七）。

當分為三群時，ADMIXTURE 分析結果進一步將中港溪至高屏溪之間的樣本劃分為獨立的中部族群（Middle）。其中，中港溪與後龍溪的樣本表現出部分北部祖先成分，此結果與 PCA 所呈現的分群結果一致。

根據 IQ-TREE 構建的最大似然樹（Maximum Likelihood tree）結果，中港溪與後龍溪的樣本在樹中則被歸類為北部群體的子群體（Subclades），與 PCA 與 ADMIXTURE 的分群結果不一致，顯示其基因型分布介於北部與中部族群之間，顯示該區域族群可能為北部與中部族群之間的遺傳過渡帶（圖八）。在演化樹上，雜交個體的演化枝較其他南部樣本短，對應的是 PCA 中介於兩族群間的雜交個體，顯示高屏溪以南的確存在南北族群間的基因流。此外組裝的全粒線體基因組與核基因的分析結果存在不一致的情況（圖八、圖九）。在林邊溪之中含有兩隻屬於北部譜系的個體，在 PCA、ADMIXTURE 以及核基因演化樹上皆被認定為南部的基因型，代表林邊溪曾經有發生過基因滲流。

由全部的點位估算的 F_{ST} 、 π 以及 Tajima's D 部分，南部族群相較其他族群的 F_{ST} 較高（圖十），並且就算包含了可能受到雜交影響的高屏溪 π 值依然為三個族群中最低的（圖十一 A）。另外 Tajima's D 在南部呈現出了負值（圖十一 B），然而考慮到採用了可能受基因流影響的高屏溪作為南部族群，可能是因為少量的

基因滲流導致低估了 Tajima's D。

綜合以上的結果，後續測試中的高屏溪個體會刪去雜交比例最高的個體，避免單一個體較高的雜交比例拉低整體的分化水平。由於核基因接近南部型的林邊溪中存在北部型的粒線體，代表存在的回交族群，需要對 PCA 中屬於南部族群的樣本進行進一步的測試，確定哪些河川中的個體有回交的現象。

3-3 雜交鑑定

由於被視為沒有雜交信號的林邊溪帶有北部的粒線體譜系，推測有未能被 PCA 以及 ADMIXTURE 檢測到的雜交個體。因此計算各河川族群與北部及中部族群間的 Weir and Cockerham 加權 F_{ST} 以及每個點位的 π 值（圖十二、表三），並取平均值作為比較指標，以辨識可作為雜交檢定的參考純種族群。分析結果顯示，將北部與中部族群分別與高屏溪以南的各河川進行比較時，發現士文溪的個體有最低的平均 π 值以及最高的 F_{ST} 值，推測其受外來基因滲入影響最小。因此在雜交測試中將士文溪視為高屏溪以南區域的代表性純種族群。

基於 PCA 的結果，雜交個體大多數介於中部族群與南部族群之間（圖五），因此將中部族群以及士文溪視作純種族群進行雜交檢測。以族群間沒有缺失且族群間固定的 SNP 推測樣本的雜交類別，並呈現在以三角散點圖上。結果顯示雜交比例以高屏溪、東港溪、林邊溪的順序遞減，高屏溪、東港溪與林邊溪分別為對南部族群的第二至第四次回交（Backcross）；枋山溪、楓港溪、四重溪、港口溪有較高的混合比例，其中枋山溪以及港口溪屬於 F_2 的混合族群，楓港溪、四重溪和一隻高屏溪個體則屬於對中部族群回交的個體（圖十三）。依據 100 kb 的區塊雜交比例進行平均，結果顯示高屏溪以北的族群沒有發現來自南部的片段，僅有高屏溪以南的族群存在來自中部族群的片段（圖十四）。

但是在三角散點圖中存在東港溪及林邊溪為都各自聚類（Cluter）在一起，並未發現不同的雜交程度個體（圖十五），可能是僅用士文溪導致的誤差。由於

東港溪的個體是沒有發現北部粒線體的基因滲流，而林邊溪可能在多次的回交中已經清除了大部分的中部區段，兩者被認為是回交族群可能是因為南部的基因多樣性。因此使用東港溪、林邊溪與士文溪作為純種重作三角散點圖，發現其他個體的雜交分類並沒有變化，所以東港溪與林邊溪應該也是偏向南部純種的族群。

將東港溪、林邊溪與士文溪納入南部族群，透過 ADMIXTOOLS 的 f3-statistic 結果顯示高屏溪、楓港溪、枋山溪、四重溪以及港口溪皆為高屏溪南北族群顯著的雜交族群（表五）。以 ADMIXTOOLS 進行的 ABBA-BABA 測試也顯示出這些河川存在顯著的基因流（表六）。

為了驗證雜交結果，以中部族群與士文溪作為純種族群，透過 LOTER 計算的雜交比例結果中，楓港溪、四重溪以及一個高屏溪樣本皆帶有一條近乎與中部族群一樣的單倍體（Haplotype），代表這些個體的其中一個單倍體帶有較多的南部片段，另一個單倍體的親代來自較為純的中部族群，代表個體是 F1 或 F2 族群與中部親代的回交（表七），三角散點圖並沒有誤判回交個體。

3-4 ABBA-BABA statistic

為了進行 ABBA-BABA statistic，使用了包含長期馬口鱸的資料集，共有 20,319,285 個 SNP。在 ABBA-BABA 之中以東港溪、林邊溪與士文溪作為南部族群的親代以及北部族群還有中部族群作為高屏溪以北的親代，測試高屏溪的基因滲流片段，在顯著的基因滲流片段（FDR<0.05）之中確定最長的片段為 0.97 MB，總共包含了 15.08 MB 的區域。富集分析中共有 101 個基因，顯著的 GO term 包含 T cell receptor binding（GO:0042608）、T cell selection（GO:0045058）等與免疫相關的 GO term（表八），在 KEGG 部分則沒有顯著的結果。顯著區段的部分包含了 *DOCK1* 與 *VAV2* 等與免疫相關的基因（表九）。

3-5 族群歷史、群分化時間及演化樹分化程度

為了檢測族群歷史，使用了包含單一變異型的 SNP 資料，共包含 15,545,876

個 SNP。根據 SMC++ 推估的歷史有效族群大小變化，可觀察到南部族群在末次間冰期（Last Interglacial, LIG）與末次冰盛期（Last Glacial Maximum, LGM）期間皆經歷過明顯的瓶頸效應；相較之下，高屏溪以北的族群僅在末次冰期期間出現有效族群數量的下降（圖十六）。進一步根據 SMC++ 預測的族群分化時間，高屏溪南北族群之間的分歧時間約為 0.13 百萬年前，位於倒數第二冰期（Penultimate Glacial Period）以及倒數第二間冰期（Penultimate Interglacial Period）之間，屬於海水正要上升的時段，海水上升或許與兩族群的分化有關。

GONE 的結果中顯示北部與中部族群近期屬於族群逐漸上升的狀態，並且南部族群的有效族群量非常少（圖十七），和 SMC++ 的結果一樣呈現出南部族群少於北與中部族群。南部族群預測出來的有效族群量非常少，而 PopLDdecay 中顯示出的南部族群相較北中部族群有著高連鎖（圖十八），應是南部族群連鎖較高導致 GONE 預測的有效族群量較少。

根據全粒線體序列的 BEAST 分析推估，南北兩族群的最近共同祖先（TMRCA）時間約為 2.26 百萬年前（95% HPD: 2.10–2.41 Ma），其中南部族群內部的共祖時間為 0.26 百萬年前（95% HPD: 0.21–0.31 Ma），北部族群則為 0.19 百萬年前（95% HPD: 0.16–0.22 Ma）（圖九）。然而，與基於核基因組資料之 SMC++ 分析所預測的分化時間相比，粒線體所推估的族群分化時間顯著偏高。此差異可能源自粒線體基因受限於單倍體遺傳及母系遺傳特性，易受基因漂變等過程影響，進而導致共祖時間高估。

透過核基因演化樹與粒線體演化樹計算出的平均北與中部族群與南部族群間的遺傳距離分別為 0.46% 以及 4.6%，兩者間相差 10 倍的演化距離；而北與中部族群與外群長鰭馬口鱲僅間的核基因與粒線體遺傳距離僅相差 2 倍，因此粗首馬口鱲的粒線體演化速度可能比較快。且依據核基因的演化樹，高屏溪南北的差異與外群長鰭馬口鱲的 5.56% 差距仍然很大，南北族群的核基因分化相較於粒線

體分化來說相當的低（表十）。在計算核基因與粒線體的 π 值時，南部族群都是三個族群中最低的（表十一）。代表粒線體可能受到遺傳漂變的影響造成了比核基因有更快的演化速度。



3-6 適應性基因檢測

PopLDdecay 之中預測出南部的族群有明顯低於北部與中部族群的連鎖不平衡下降速率（圖十八），代表南部相對於北部與中有較高的選擇壓力，但也可能是受到族群下降或是雜交的影響所造成的結果。

以 F_{ST} 與 π ratio 共同預測可能受到天擇的區段(圖十九)，最終得到了北部候選的適應性基因 435 個；南部候選的適應性基因 418 個。將兩組基因資料分別做富集分析，並得到顯著富集的 GO term，發現北部的基因中含有 78 個；南部的基因含有 108 個，兩者顯著的 GO term 皆為 58 個。南部族群中包含了 thyroid hormone binding (GO:0070324)、negative regulation of female receptivity (GO:0007621) 以及 cholesterol storage (GO:0010878)，透過 BLAST 到斑馬魚上發現這三個 GO term 中包含了 *thrb*、*stard4* 等基因（表十二、表十三）；中北部的族群部分則在 rDNA binding (GO:0000182)、cellular response to nutrient (GO:0031670)、response to insecticide (GO:0017085) 以及 positive regulation of translation (GO:0045727) 中找到 *mTOR* 與 *pik3cb* 等與抗寒以及缺氧有關的基因（表十四、表十五）。

在 KEGG 的富集分析中北部與中部族群並沒有任何顯著的結果，但是確認到了 mTOR signaling pathway (ko04150) 存在；南部族群中則在 Biotin metabolism (ko00780) 以及 Thyroid hormone signaling pathway (ko04919) 有顯著的結果，並沒有觀測到顯著的 mTOR signaling pathway (ko04150)（表十六）。因此南部族群的天擇可能與甲狀腺素的調控路徑有關。

第四章 討論



本研究透過多項遺傳分析方法，綜合評估粗首馬口鱖南北族群間的分化機制與生殖隔離狀況。結果顯示，兩族群間存在基因交流的現象，顯示尚未形成完全的生殖隔離，核基因遺傳距離較小，因此可能為同一物種。族群結構分析揭示粗首馬口鱖主要以高屏溪為界分為兩個族群，與過去的分子標記結果相符。進一步分析亦發現基因組島的低多樣性區域中包含了溫度調控相關的基因，顯示環境適應與自然選擇可能在族群分化與基因滲入的過程中扮演重要角色。

4-1 南北族群的分化時間與分化程度

過去的研究中主要都是以高屏溪為界分為南北兩個族群，同功異構酶 (allozyme) 以高屏溪為界分為南北族群，並且北部族群可以分為北與中兩個姊妹群[7]；D-loop 則以高屏溪為界分為南北兩個族群[8]。核基因結果顯示出與同功異構酶一致的分群 (圖八)，並且最佳的分群數為 2 (圖六)，支持了以高屏溪為界的分群的結果。

南北族群在核基因與粒線體上分別有 0.46% 以及 4.6% 的遺傳距離，顯示兩者間存在地理隔離與遺傳分化，但是相較於高屏溪以北族群與長鰭馬口鱖核基因之間的 5.5% 遺傳距離仍有差距 (表十)。南北族群間體核基因分化並未到物種間差異的程度。

不論是 GONE 或是 SMC++ 的預測結果，南部族群的有效族群量皆小於北部族群。其中 SMC++ 預測出南部族群分別在末次間冰期以及末次間冰期曾經歷過一次瓶頸效應 (圖十六、圖十七)，同時不論是核基因還是粒線體基因的 π 值南部族群皆低於北部與中部區域 (表十一)，代表南部族群應是冰河時期分出去的小族群，符合過去認為粗首馬口鱖曾經歷過創始者效應 (Founder effect) 的假設 [7]。粒線體基因有效族群大小相較於核基因小的特性，因此更易受到遺傳漂變影

響。過去的研究發現粒線體的突變率可能與有效族群量呈現負相關[46]，小族群下的粒線體基因可能出現變異率上升的現象，進而導致分化時間被高估。在北美紅雀（*Cardinalis cardinalis*）的粒線體與核基因的研究中發現，有效族群量小會使粒線體變異率增加並導致分化時間的高估[47]。在此次的研究中發現南北族群間的核基因與粒線體遺傳距離，在南北族群間差異達到 10 倍左右，但是在北部與中部族群對上長鰭馬口鱖時，核基因與粒線體基因間的遺傳距離只相差了 2 倍左右（表十）。因此粗首馬口鱖南北族群間的粒線體基因有加速演化的現象。並且 SMC++計算的核基因分化時間為 13 萬年前，而 BEAST 計算的粒線體分化時間則為 226 萬年前（圖九），然而中央山脈隆起約為 150 至 130 萬年前[48]，由於粗首馬口鱖的原生族群是分布於臺灣西部[7]，粗首馬口鱖遷徙時可能受到了中央山脈的阻隔而無法擴散到東部地區，因此粗首馬口鱖不可能在 226 萬年前遷徙至臺灣。推測此分化時間的高度差異，可能反映了南部族群因長期低有效族群量所導致的粒線體變異率上升與遺傳漂變效應，進而造成最近共祖時間被高估以及粒線體分化程度偏大的現象。

根據核基因所推算的族群分化時間約為 13 萬年前（圖十六），正處於第二次間冰期（MIS 5e）與倒數第二次冰期（MIS 6）之間的過渡時期，亦即海平面快速上升的階段 [49]。推測此時的海平面上升阻斷了原本在大陸棚上相連的河川，進而形成南、北族群間的地理隔離。

4-2 南北族群的雜交狀況

根據三角散點圖的結果，高屏溪以南的樣本中觀察到多代雜交個體，包括 F2 及回交個體，且涵蓋從第一次回交至第三次回交的類型（圖十五）。枋山溪、楓港溪、四重溪以及港口溪大多為 F2 或是對中部族群的回交個體，高屏溪也包含了一隻對中部回交的個體。由於發現了大量的 F2 與回交個體，南北族群間可能還未演化出有效的生殖隔離。

目前無法確定放流的具體原因，但可以確定該放流事件發生於2014年以前。與2017年的粒線體研究中首次發現南北族群間的次級接觸的結果一致[9]。而枋山溪、楓港溪、四重溪以及港口溪發現大量混合個體的原因，或許是因為這些河川較短（圖二），更容易產生兩個族群間的基因流。

4-3 南部回交族群保留的基因滲流片段與免疫有關

根據 ABBA-BABA 的顯著區段（FDR<0.05）測定出的顯著富集 GO term（ $q < 0.05$ ）包含了免疫相關的功能，包含 C 顯著的 GO term 包含 T cell receptor binding（GO:0042608）、T cell selection（GO:0045058）等（表八）。

其中包含的 *VAV2* 和 *DOCK1*（表九）都屬於 Chemokine signaling pathway (ko04062) 中的基因。Chemokine signaling pathway (ko04062) 的功能是調控免疫細胞的遷移與移動方向[50]。*VAV2* 在小鼠的敲除實驗上發現是與 B 的體液免疫與 B 細胞成熟有關[51]。*DOCK1* 則可以與 Rac 結合並作用於細胞的遷徙以及對凋亡細胞的吞噬[52]。因此基因滲流的區段可能與免疫具有關聯性，

4-4 南北族群的分化與溫度具有關聯性

受天擇的候選片段區域中，分別在北部與南部找到了與溫度相關的基因，北部的基因中包含 *mTOR* 及 *mapkapk2b*（表十五）。根據先前對斑馬魚的研究，飢餓會透過抑制 mTOR 訊號路徑，促進脂肪代謝與自噬作用，進而提升魚類對寒冷環境的耐受性，此機制被認為是魚類冷適應的重要策略之一[53]。*mapkapk2b* 為 *mapkapk2* 的同源基因（Paralog）。已有研究顯示，當 *mTOR* 被抑制時，BCL2 的表現量會下降 [53]。在海拉細胞（HeLa cell）的實驗中發現 *mapkapk2* 會受到 BCL2 的抑制，且 *mapkapk2* 可以正向調控自噬蛋白質 BECN 1，加強自噬作用 [54]。本研究找到了兩個相互關聯調控自噬作用的基因，推測北部族群受到了冷適應的天擇影響。

南部族群受天擇的基因中包含 *thrb*、*gsk3ba* 等基因（表十三），同時也找到

了 thyroid hormone signaling pathway 的顯著富集。有研究指出，甲狀腺素能透過調控代謝與肌肉功能，參與外溫動物的體溫調節，對於變溫動物的熱適應具有關鍵性影響[55]。棲息於高溫環境的阿馬戈薩鱗魚（*Cyprinodon nevadensis amargosae*）族群，其肝臟中 *thrb* 的 mRNA 表現量相較於低溫環境族群低[56]。在斑馬魚的三碘甲狀腺素（T3）暴露實驗中時，斑馬魚肝臟的 *thrb* 表現量低於控制組以及其他部分 *thrb* 上升的狀況[57]。另有研究發現，熱馴化（28°C）條件下的斑馬魚，其骨骼肌中三碘甲狀腺素濃度高於冷馴化（18°C）個體，然而對三碘甲狀腺素的生理反應卻較為遲鈍[58]。表示在較熱的環境之下魚類體內的三碘甲狀腺素可能會更高。另外補充三碘甲狀腺素可降低藍鯽魚的側線鱗（Lateral-line scale）數量[59]。而在臺灣鬚鱨的研究中，側線鱗數與水溫呈現負相關 [60]；並且粗首馬口鱮中，南部族群的側線鱗片數亦少於北部族群[8, 61]。所以南部的粗首馬口鱮的側線鱗相較北部少的原因，推測是熱環境下導致個體的三碘甲狀腺素上升，致使南部個體的側線鱗片減少。

過去粗首馬口鱮與高屏馬口鱮便是依據側線鱗片數量作為其中一個外觀分類標準[8]，但是根據找到的甲狀腺素路徑證據，以及國小的科展之中發現了粗首馬口鱮與高屏馬口鱮的側線鱗片數量的重疊結果[61]。粗首馬口鱮的側線鱗可能是受溫度與甲狀腺素影響而產生差異，而非完全受到基因型的影響，因此不是一個好的物種分類依據。

gsk3ba 是 *gsk3b* 的平行同源基因（paralogous gene），*gsk3b* 的功能是抑制熱休克後的 HSF1（Heat Shock Factor 1）蛋白質轉錄活性[62]。在斑馬魚的研究中發現，HSF1 的表現可以降低個體缺氧後的眼睛與大腦的細胞凋亡[63]。所以 *gsk3ba* 的演化可能與熱環境下的缺氧有關係，其可能藉由降低自身表現量，延長 HSF1 的活化時間，進而增強細胞對缺氧壓力的耐受性。

所以在富集研究中分別在南北族群中找到熱耐受以及冷耐受的高度分化基

因，而南部河川中除了主要的甲狀腺素調節外，還找到了與缺氧相關的 *gsk3ba*，
因此推測以高屏溪為界的南北的分化與溫度相關。



第五章 結論



過去對粗首馬口鱧的研究多聚焦於同功酶與粒線體，並依據側線鱗片數量將南北族群區分為不同物種。然而，本研究透過全基因組定序，確認粗首馬口鱧可分為北、中、南三個主要族群，其中南部族群與其他兩群之間的分化程度最高。南部族群亦為小族群，其粒線體的高度分化可能主要受到基因漂變的影響。以核基因的分化時間為準，南北族群的分化可能是因為海平面上升導致的。

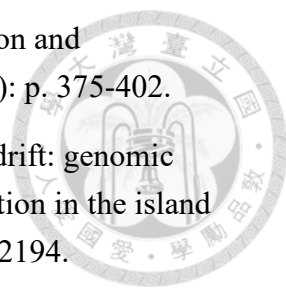
儘管族群間具有遺傳分化，但在雜交分析中未觀察到明顯的生殖隔離現象，顯示南北族群尚未演化出完全的生殖隔離機制。進一步的功能基因分析發現，南部族群呈現與甲狀腺素調控相關基因的選擇訊號，而側線鱗數、河川溫度與甲狀腺素之間存在顯著關聯。因此，側線鱗片數量可能受到環境與內分泌調控影響，不宜作為穩定的分類依據。

綜合而言，本研究結果支持南北族群應仍屬於同一物種，其間的分化可能主要受到溫度梯度所引發的自然選擇與地理隔離所驅動，為理解粗首馬口鱧的族群分化與環境適應提供了新的基因層次證據。

第六章 參考文獻



1. Avise, J.C., et al., Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics*, 1987: p. 489-522.
2. Durand, J.D., et al., Nested clade and phylogeographic analyses of the chub, *Leuciscus cephalus* (Teleostei, Cyprinidae), in Greece: implications for Balkan Peninsula biogeography. *Molecular Phylogenetics and Evolution*, 1999. 13(3): p. 566-580.
3. Tzeng, C. Distribution of the Freshwater Fishes of Taiwan. *Quarterly Journal of Taiwan Museum*, 1986. 39(2): p. 127-146.
4. Jean, C.-T., et al., Population genetic structure in the endemic cyprinid fish *Microphysogobio alticorpus* in Taiwan: evidence for a new phylogeographical area. *Biochemical Systematics and Ecology*, 2014. 57: p. 108-116.
5. Liao, T.-Y., et al., Phylogeography of the endangered species, *Sinogastromyzon puliensis* (Cypriniformes: Balitoridae), in southwestern Taiwan based on mtDNA. *Zoological Studies*, 2008. 47(4): p. 383-392.
6. Boggs, S., Sediment properties and water characteristics of the Taiwan shelf and slope. *Acta Oceanographica Taiwanica*. Taiwan., 1979. 10: p. 10-49.
7. Wang, H.Y., et al., Influence of glaciation on divergence patterns of the endemic minnow, *Zacco pachycephalus*, in Taiwan. *Molecular Ecology*, 1999. 8(11): p. 1879-1888.
8. Chen, I.-S., J.-H. Wu, and S.-P. Huang, The taxonomy and phylogeny of the cyprinid genus *Opsariichthys* *Bleeker* (Teleostei: Cyprinidae) from Taiwan, with description of a new species. *Chinese Fishes*, 2010: p. 165-183.
9. 邱柏豪，粗首鱯複合種群之粒線體基因族群遺傳研究，海洋科學系研究所。2017，碩士論文，國立中山大學海洋科學系，高雄市。
10. Scribner, K.T., K.S. Page, and M.L. Bartron, Hybridization in freshwater fishes: a review of case studies and cytonuclear methods of biological inference. *Reviews in Fish Biology and Fisheries*, 2000. 10: p. 293-323.
11. Wu, C.I., The genic view of the process of speciation. *Journal of Evolutionary Biology*, 2001. 14(6): p. 851-865.

- 
12. Nosil, P., D.J. Funk, and D. Ortiz-Barrientos, Divergent selection and heterogeneous genomic divergence. *Molecular Ecology*, 2009. 18(3): p. 375-402.
 13. Funk, W.C., et al., Adaptive divergence despite strong genetic drift: genomic analysis of the evolutionary mechanisms causing genetic differentiation in the island fox (*Urocyon littoralis*). *Molecular Ecology*, 2016. 25(10): p. 2176-2194.
 14. Yeaman, S., S. Aeschbacher, and R. Bürger, The evolution of genomic islands by increased establishment probability of linked alleles. *Molecular Ecology*, 2016. 25(11): p. 2542-2558.
 15. Wang, H.Y., et al., New insights on molecular systematics of opsariichthines based on cytochrome b sequencing. *Journal of Fish Biology*, 2007. 71: p. 18-32.
 16. Chen, S., Ultrafast one-pass FASTQ data preprocessing, quality control, and deduplication using fastp. *iMeta*, 2023. 2(2): p. e107.
 17. Chen, S., et al., fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, 2018. 34(17): p. i884-i890.
 18. Bolger, A.M., M. Lohse, and B. Usadel, Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 2014. 30(15): p. 2114-2120.
 19. Li, H., Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv*. 2013; arXiv:1303.3997.
 20. Li, H. and R. Durbin, Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*, 2009. 25(14): p. 1754-1760.
 21. Tai, J-H. Mechanisms of Genomic Divergence in the Opsariichthyini: Chromosomal Fission and Environmental Adaptation. 2025, Taipei: National Taiwan University.
 22. Danecek, P., et al., Twelve years of SAMtools and BCFtools. *GigaScience*, 2021. 10(2).
 23. Van der Auwera, G.A. and B.D. O'Connor, Genomics in the cloud: using Docker, GATK, and WDL in Terra. *O'Reilly Media* , 2020.
 24. Li, H., A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*, 2011. 27(21): p. 2987-93.
 25. Danecek, P., et al., The variant call format and VCFtools. *Bioinformatics*, 2011.

27(15): p. 2156-2158.

26. Meng, G., et al., MitoZ: a toolkit for animal mitochondrial genome assembly, annotation and visualization. *Nucleic Acids Research*, 2019. 47(11): p. e63-e63.

27. Kumar, S., et al., MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Molecular Biology and Evolution*, 2018. 35(6): p. 1547-1549.

28. Bouckaert, R., et al., BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. *PLoS Computational Biology*, 2019. 15(4): p. e1006650.

29. Dowling, T.E., et al., Evolutionary relationships of the plagopterins (Teleostei: Cyprinidae) from cytochrome b sequences. *Copeia*, 2002. 2002(3): p. 665-678.

30. Rambaut, A., et al., Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology*, 2018. 67(5): p. 901-904.

31. Chang, C.C., et al., Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience*, 2015. 4(1).

32. Alexander, D.H., J. Novembre, and K. Lange, Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, 2009. 19(9): p. 1655-64.

33. Ortiz, E., vcf2phylic v2. 0: Convert a VCF Matrix into Several Matrix Formats for Phylogenetic Analysis [Computer Software]. 2019.

34. Minh, B.Q., et al., IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Molecular Biology and Evolution*, 2020. 37(5): p. 1530-1534.

35. Patterson, N., et al., Ancient admixture in human history. *Genetics*, 2012. 192(3): p. 1065-1093.

36. Browning, B.L., et al., Fast two-stage phasing of large-scale sequence data. *The American Journal of Human Genetics*, 2021. 108(10): p. 1880-1890.

37. Browning, B.L., Y. Zhou, and S.R. Browning, A one-penny imputed genome from next-generation reference panels. *The American Journal of Human Genetics*, 2018. 103(3): p. 338-348.

38. Alves, T.D., Modélisation du déséquilibre de liaison en génomique des populations par méthodes d'optimisation. 2017, Université Grenoble Alpes.

39. Dias-Alves, T., J. Mairal, and M.G.B. Blum, Loter: A Software Package to Infer Local Ancestry for a Wide Range of Species. *Molecular Biology and Evolution*, 2018.

35(9): p. 2318-2326.

40. Wu, T., et al., clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *The Innovation*, 2021. 2(3).

41. Zhang, C., et al., PopLDdecay: a fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. *Bioinformatics*, 2018. 35(10): p. 1786-1788.

42. Santiago, E., et al., Recent Demographic History Inferred by High-Resolution Analysis of Linkage Disequilibrium. *Molecular Biology and Evolution*, 2020. 37(12): p. 3642-3653.

43. Terhorst, J., J.A. Kamm, and Y.S. Song, Robust and scalable inference of population history from hundreds of unphased whole genomes. *Nature genetics*, 2017. 49(2): p. 303-309.

44. 楊芷毓，石門水庫集水區湳仔溝溪之臺灣馬口魚族群生態學研究，漁業科學研究所。2011，碩士論文，國立臺灣大學漁業科學研究所，台北市。

45. Bergeron, L.A., et al., Evolution of the germline mutation rate across vertebrates. *Nature*, 2023. 615(7951): p. 285-291.

46. Allio, R., et al., Large Variation in the Ratio of Mitochondrial to Nuclear Mutation Rate across Animals: Implications for Genetic Diversity and the Use of Mitochondrial DNA as a Molecular Marker. *Molecular Biology and Evolution*, 2017. 34(11): p. 2762-2772.

47. Smith, B.T. and J. Klicka, Examining the role of effective population size on mitochondrial and multilocus divergence time discordance in a songbird. *PLoS One*, 2013. 8(2): p. e55161.

48. Chang, Q., et al., Rapid topographic growth of the Taiwan orogen since~ 1.3–1.5 Ma. *Science Advances*, 2023. 9(25): p. eade6415.

49. Siddall, M., et al., Sea-level fluctuations during the last glacial cycle. *Nature*, 2003. 423(6942): p. 853-858.

50. Mellado, M., et al., Chemokine signaling and functional responses: the role of receptor dimerization and TK pathway activation. *Annual Review of Immunology*, 2001. 19(1): p. 397-421.

51. Doody, G.M., et al., Signal transduction through Vav-2 participates in humoral immune responses and B cell maturation. *Nature Immunology*, 2001. 2(6): p. 542-547.

52. Koubek, E.J. and L.C. Santy, Actin up: an overview of the Rac GEF Dock1/Dock180 and its role in cytoskeleton rearrangement. *Cells*, 2022. 11(22): p. 3565.
53. Lu, D.L., et al., Fasting enhances cold resistance in fish through stimulating lipid catabolism and autophagy. *The Journal of Physiology*, 2019. 597(6): p. 1585-1603.
54. Wei, Y., et al., The stress-responsive kinases MAPKAPK2/MAPKAPK3 activate starvation-induced autophagy through Beclin 1 phosphorylation. *Elife*, 2015. 4: p. e05289.
55. Little, A., Thyroid hormone regulation of thermal acclimation in ectotherms: physiological mechanisms and ecoevolutionary implications. *Molecular and Cellular Endocrinology*, 2021. 530: p. 111285.
56. Lema, S.C., et al., Endocrine and metabolic impacts of warming aquatic habitats: differential responses between recently isolated populations of a eurythermal desert pupfish. *Conservation Physiology*, 2016. 4(1): p. cow047.
57. Parsons, A.E., et al., Expression dynamics of genes in the hypothalamic-pituitary-thyroid (HPT) cascade and their responses to 3, 3', 5-triiodo-L-thyronine (T3) highlights potential vulnerability to thyroid-disrupting chemicals in zebrafish (*Danio rerio*) embryo-larvae. *Aquatic Toxicology*, 2020. 225: p. 105547.
58. Little, A.G., et al., Thyroid hormone actions are temperature-specific and regulate thermal acclimation in zebrafish (*Danio rerio*). *BMC Biology*, 2013. 11: p. 1-15.
59. Levin, B. and M. Levina, Poly-and oligomerization of scales in the blue breem *Ballerus ballerus* (Cyprinidae) as a consequence of thyroid status regulation. *Journal of Applied Ichthyology*, 2014. 30(4): p. 809-813.
60. Chen, F., et al., Systematics and phylogeography of the Taiwanese endemic minnow *Candidia barbatus* (Pisces: Cyprinidae) based on DNA sequence, allozymic, and morphological analyses. *Zoological Journal of the Linnean Society*, 2011. 161: p. 613-632.
61. 粘閔智；邱瓏臻，粗首鱯遺傳多樣性與形態變異之研究。1996，基隆市私立聖心高級中學，基隆市。
62. Xavier, I.J., et al., Glycogen synthase kinase 3 β negatively regulates both DNA-binding and transcriptional activities of heat shock factor 1. *Journal of Biological Chemistry*, 2000. 275(37): p. 29147-29152.

63. Tucker, N.R., et al., HSF1 is essential for the resistance of zebrafish eye and brain tissues to hypoxia/reperfusion injury. *PLoS One*, 2011. 6(7): p. e22268.



表

表一、粗首馬口鱮樣本採集地及各河川 *cytb* 定序數量

Lineage A 與 Lineage B 分別代表北部譜系與南部譜系的數量。



Chinese name	Sample size	Lineage A	Lineage B
瑪鍊溪(MS)	9	9	
基隆河(KL)	6	6	
橫溪(HX)	4	4	
大漢溪(DH)	5	5	
中港溪(JG)	13	13	
後龍溪(HL)	4	4	
大安溪(DA)	4	4	
大甲溪(DJ)	4	4	
烏溪(WX)	3	3	
濁水溪(JS)	2	2	
朴子溪(PZ)	3	3	
高屏溪(KP)	12	2	10
東港溪(DG)	9	9	
林邊溪(LB)	11	3	8
士文溪(SW)	13	13	
枋山溪(FS)	11	10	1
楓港溪(FG)	9	9	
四重溪(SH)	1	1	
港口溪(GK)	1	1	

表二、粗首馬口鱖樣本採集地點、地圖代號及各河川全基因組定序數量

Symbol 對應(圖一)的採集地點。

Symbol	Chinese name	Sample size
1	瑪鍊溪(MS)	1
2	基隆河(KL)	3
3	橫溪(HX)	2
4	大漢溪(DH)	2
5	中港溪(JG)	3
6	後龍溪(HL)	2
7	大安溪(DA)	2
8	大甲溪(DJ)	3
9	濁水溪(JS)	2
10	烏溪(WX)	3
11	朴子溪(PZ)	3
12	高屏溪(KP)	5
13	東港溪(DG)	3
14	林邊溪(LB)	4
15	士文溪(SW)	3
16	枋山溪(FS)	3
17	楓港溪(FG)	2
18	四重溪(SH)	1
19	港口溪(GK)	1

表三、全基因組定序樣本與粒線體譜系

樣本	定序類型	採集地點	採集日期	總定序量 (Gbp)	粒線體譜系
OETO06	Illumina NovaSeq 6000	頭前溪	2018.08.02	17.25	<i>O. evolans</i>
OEHL01	Illumina NovaSeq 6000	後龍溪	2018.09.14	23.29	<i>O. evolans</i>
OPMS09	Illumina NovaSeq 6000	瑪鍊溪	2023.06.24	17.00	Lineage A
OPKL04	Illumina NovaSeq X Plus	基隆河	2020.08.08	8.92	Lineage A
OPKL06	Illumina NovaSeq X Plus	基隆河	2019.07.23	8.53	Lineage A
OPKL07	Illumina NovaSeq X Plus	基隆河	2019.07.18	11.42	Lineage A
OPHX03	Illumina NovaSeq X Plus	橫溪	2023.06.25	9.03	Lineage A
OPHX04	Illumina NovaSeq X Plus	橫溪	2023.06.25	10.69	Lineage A
OPDH05	Illumina NovaSeq X Plus	大漢溪	2019.07.18	9.59	Lineage A
OPDH06	Illumina NovaSeq X Plus	大漢溪	2019.07.18	11.47	Lineage A
OPJG01	Illumina NovaSeq X Plus	中港溪	2023.06.25	11.56	Lineage A
OPJG02	Illumina NovaSeq X Plus	中港溪	2023.06.25	9.61	Lineage A
OPJG12	Illumina NovaSeq 6000	中港溪	2023.06.25	16.73	Lineage A
OPHL01	Illumina NovaSeq 6000	後龍溪	2019.07.30	17.38	Lineage A
OPHL02	Illumina NovaSeq X Plus	後龍溪	2020.09.04	9.42	Lineage A
OPDA01	Illumina NovaSeq X Plus	大安溪	2015.07.26	10.68	Lineage A
OPDA02	Illumina NovaSeq X Plus	大安溪	2015.07.26	11.67	Lineage A
OPDJ01	Illumina NovaSeq X Plus	大甲溪	2015.07.27	12.58	Lineage A
OPDJ02	Illumina NovaSeq X Plus	大甲溪	2015.07.27	9.05	Lineage A
OPDJ03	Illumina NovaSeq X Plus	大甲溪	2015.07.27	9.55	Lineage A

表三、全基因組定序樣本與粒線體譜系(續)

樣本	定序類型	採集地點	採集日期	總定序量 (Gbp)	粒線體譜系
OPZS01	Illumina NovaSeq X Plus	濁水溪	2015.08.06	12.58	Lineage A
OPZS02	Illumina NovaSeq X Plus	濁水溪	2015.08.06	11.88	Lineage A
OPWU01	Illumina NovaSeq X Plus	烏溪	2023.11.14	13.03	Lineage A
OPWU02	Illumina NovaSeq X Plus	烏溪	2023.11.14	13.03	Lineage A
OPWU03	Illumina NovaSeq X Plus	烏溪	2023.11.14	14.27	Lineage A
OPPZ01	Illumina NovaSeq X Plus	朴子溪	2023.11.13	12.05	Lineage A
OPPZ02	Illumina NovaSeq X Plus	朴子溪	2023.11.13	12.03	Lineage A
OPPZ03	Illumina NovaSeq X Plus	朴子溪	2023.11.13	9.81	Lineage A
OPKP11	Illumina NovaSeq X Plus	高屏溪	2015.07.04	9.00	Lineage A
OPKP03	Illumina NovaSeq X Plus	高屏溪	2015	10.47	Lineage B
OPKP05	Illumina NovaSeq X Plus	高屏溪	2015	9.45	Lineage B
OPKP08	Illumina NovaSeq X Plus	高屏溪	2015	9.52	Lineage B
OPKP13	Illumina NovaSeq X Plus	高屏溪	2015.07.04	23.75	Lineage B
OPDG02	Illumina NovaSeq X Plus	東港溪	2023.11.26	10.88	Lineage B
OPDG03	Illumina NovaSeq X Plus	東港溪	2023.11.26	10.02	Lineage B
OPDG05	Illumina NovaSeq X Plus	東港溪	2023.11.26	12.08	Lineage B
OPLB01	Illumina NovaSeq X Plus	林邊溪	2023.11.25	10.90	Lineage B
OPLB02	Illumina NovaSeq X Plus	林邊溪	2023.11.25	10.67	Lineage A
OPLB06	Illumina NovaSeq X Plus	林邊溪	2023.11.25	8.94	Lineage A
OPLB11	Illumina NovaSeq 6000	林邊溪	2014.12.08	21.66	Lineage B

表三、全基因組定序樣本與粒線體譜系(續)



樣本	定序類型	採集地點	採集日期	總定序量 (Gbp)	粒線體譜系
OPSW01	Illumina NovaSeq X Plus	士文溪	2023.11.25	12.92	Lineage B
OPSW02	Illumina NovaSeq X Plus	士文溪	2023.11.25	10.33	Lineage B
OPSW07	Illumina NovaSeq X Plus	士文溪	2023.11.25	11.67	Lineage B
OPFS03	Illumina NovaSeq X Plus	枋山溪	2023.11.25	9.92	Lineage A
OPFS09	Illumina NovaSeq X Plus	枋山溪	2023.11.25	9.22	Lineage B
OPFS011	Illumina NovaSeq X Plus	枋山溪	2014.7.18	9.18	Lineage A
OPFG03	Illumina NovaSeq X Plus	楓港溪	2023.12.29	11.50	Lineage A
OPFG08	Illumina NovaSeq X Plus	楓港溪	2023.12.29	8.78	Lineage A
OPGK01	Illumina NovaSeq X Plus	港口溪	2014.08.03	10.96	Lineage A
OPSH01	Illumina NovaSeq X Plus	四重溪	2014.08.23	10.69	Lineage A

表四、根據 SNP 計算出的族群每個點位的平均 π 值以及族群內個體數量

Population	Mean π	Population Size
North	0.100	9
Middle	0.200	13
KP	0.206	4
DG	0.193	3
LB	0.124	4
SW	0.079	3
FS	0.155	3
FG	0.123	2
SH	0.196	1
GK	0.062	1

表五、透過 ADMIXTOOLS 計算的 f_3 statistic

若 Z score 小於 3 且 f_3 為負值，則代表目標族群為兩個親代的混合族群。此處的高屏溪（KP）沒有包含雜交最嚴重的個體。



P1	P2	Target	f_3	std. err	Z	SNPs
North, Middle	DG, LB, SW	KP	-0.22	0.00	-55.49	8,970,389
North, Middle	DG, LB, SW	FS	-0.21	0.01	-39.98	9,003,774
North, Middle	DG, LB, SW	FG	-0.13	0.01	-13.77	8,892,033
North, Middle	DG, LB, SW	SH	-0.12	0.02	-6.23	8,512,761
North, Middle	DG, LB, SW	GK	-0.16	0.02	-8.94	8,535,569

表六、透過 ADMIXTOOLS 進行 ABBA-BABA 測試結果

若 Z_{score} 小於 3 且 D 為負值，則代表目標族群為兩個親代的混合族群。此處的高屏溪（KP）沒有包含雜交最嚴重的個體。



P1	Target	P2	Outgroup	D	Z_Score	BABA	ABBA	SNPs
DG, LB, SW	KP	North, Middle	OE	-0.31	-17.08	57,507	108,860	14,420,605
DG, LB, SW	FS	North, Middle	OE	-0.62	-45.91	80,644	343,444	14,412,896
DG, LB, SW	FG	North, Middle	OE	-0.70	-80.74	98,964	553,951	14,332,526
DG, LB, SW	GK	North, Middle	OE	-0.62	-51.12	98,529	419,582	13,860,802
DG, LB, SW	SH	North, Middle	OE	-0.70	-81.73	99,695	557,005	13,830,738

表七、透過 LOTER 計算出的個體單倍體混合比例

使用士文溪與中部族群作為純種，Hybird index 代表此單倍體含有的中部片段百分比，Hybrid type 是依據(圖十五)的預測結果。

Sample name	Hybird index (Haplotype 1)	Hybird index (Haplotype 2)	Hybrid type
OPKP03	19.5	6.1	BC2
OPKP05	18.7	8.3	BC2
OPKP08	16.1	3.2	BC3
OPKP13	16.9	4.1	BC3
OPKP11	47.5	90.4	BC1
OPDG02	12.3	2.4	Parent
OPDG03	10.5	2.1	Parent
OPDG05	12.8	1.9	Parent
OPLB01	5.3	0.8	Parent
OPLB11	4.8	0.6	Parent
OPLB02	5.1	0.7	Parent
OPLB06	4.7	0.9	Parent
OPFS09	46.3	39.4	F2
OPFS11	41.1	44.4	F2
OPFS03	46.9	47.8	F2
OPFG03	67.3	91.5	BC3
OPFG08	59.7	91.7	BC2
OPSH01	67.6	95.1	BC2
OPGK01	54.5	65.9	F2

表八、ABBA-BABA 中顯著富集的 GO

Ontology	ID	Description	GeneRatio	BgRatio	pvalue	p.adjust	qvalue
BP	GO:0006689	ganglioside catabolic process	3/353	10/24014	3.50E-04	2.40E-02	2.28E-02
BP	GO:0001573	ganglioside metabolic process	5/353	34/24014	1.31E-04	1.34E-02	1.27E-02
BP	GO:0001768	establishment of T cell polarity	7/353	21/24014	1.36E-08	1.77E-05	1.68E-05
BP	GO:0001766	membrane raft polarization	7/353	22/24014	1.97E-08	1.77E-05	1.68E-05
BP	GO:0001767	establishment of lymphocyte polarity	7/353	23/24014	2.80E-08	1.77E-05	1.68E-05
BP	GO:0031580	membrane raft distribution	7/353	23/24014	2.80E-08	1.77E-05	1.68E-05
BP	GO:0051665	membrane raft localization	7/353	25/24014	5.35E-08	2.54E-05	2.41E-05
BP	GO:0031579	membrane raft organization	7/353	47/24014	5.30E-06	1.19E-03	1.12E-03
BP	GO:0070986	left/right axis specification	7/353	89/24014	3.43E-04	2.40E-02	2.28E-02
BP	GO:0030010	establishment of cell polarity	19/353	430/24014	2.33E-05	3.41E-03	3.23E-03
BP	GO:0090630	activation of GTPase activity establishment or maintenance	11/353	200/24014	1.95E-04	1.81E-02	1.71E-02
BP	GO:0045197	of epithelial cell apical/basal polarity	10/353	163/24014	1.57E-04	1.49E-02	1.41E-02
BP	GO:0035088	establishment or maintenance of apical/basal cell polarity	10/353	191/24014	5.56E-04	3.34E-02	3.16E-02
BP	GO:0061245	establishment or maintenance of bipolar cell polarity	10/353	191/24014	5.56E-04	3.34E-02	3.16E-02
BP	GO:0006659	phosphatidylserine biosynthetic process	3/353	13/24014	8.07E-04	4.58E-02	4.33E-02
BP	GO:0044351	macropinocytosis	7/353	22/24014	1.97E-08	1.77E-05	1.68E-05
BP	GO:0006907	pinocytosis	7/353	42/24014	2.42E-06	7.09E-04	6.71E-04
BP	GO:0086028	bundle of His cell to Purkinje myocyte signaling	3/353	10/24014	3.50E-04	2.40E-02	2.28E-02
BP	GO:0086043	bundle of His cell action potential	3/353	10/24014	3.50E-04	2.40E-02	2.28E-02
BP	GO:1901843	positive regulation of high voltage-gated calcium channel activity	3/353	10/24014	3.50E-04	2.40E-02	2.28E-02
BP	GO:0002277	myeloid dendritic cell activation involved in immune response	6/353	11/24014	4.20E-09	1.60E-05	1.51E-05

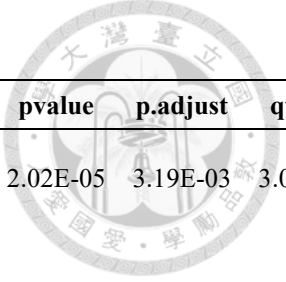
表八、ABBA-BABA 中顯著富集的 GO(續)

Ontology	ID	Description	GeneRatio	BgRatio	pvalue	p.adjust	qvalue
BP	GO:0043383	negative T cell selection	6/353	22/24014	5.91E-07	2.05E-04	1.94E-04
BP	GO:0045059	positive thymic T cell selection	6/353	22/24014	5.91E-07	2.05E-04	1.94E-04
BP	GO:0045060	negative thymic T cell selection	6/353	22/24014	5.91E-07	2.05E-04	1.94E-04
BP	GO:0007523	larval visceral muscle development	6/353	25/24014	1.35E-06	4.29E-04	4.06E-04
BP	GO:0061327	anterior Malpighian tubule development	6/353	29/24014	3.45E-06	9.38E-04	8.88E-04
BP	GO:0007522	visceral muscle development	6/353	31/24014	5.22E-06	1.19E-03	1.12E-03
BP	GO:0035022	positive regulation of Rac protein signal transduction	6/353	31/24014	5.22E-06	1.19E-03	1.12E-03
BP	GO:0045061	thymic T cell selection	6/353	32/24014	6.35E-06	1.34E-03	1.27E-03
BP	GO:0001773	myeloid dendritic cell activation	6/353	37/24014	1.53E-05	2.91E-03	2.76E-03
BP	GO:0035212	cell competition in a multicellular organism	6/353	38/24014	1.79E-05	3.19E-03	3.02E-03
BP	GO:0019049	mitigation of host antiviral defense response	6/353	39/24014	2.09E-05	3.19E-03	3.02E-03
BP	GO:0052031	modulation by symbiont of host defense response	6/353	39/24014	2.09E-05	3.19E-03	3.02E-03
BP	GO:0052173	response to defenses of other organism	6/353	43/24014	3.73E-05	4.55E-03	4.30E-03
BP	GO:0052200	response to host defenses	6/353	43/24014	3.73E-05	4.55E-03	4.30E-03
BP	GO:0075136	response to host	6/353	43/24014	3.73E-05	4.55E-03	4.30E-03
BP	GO:0043368	positive T cell selection	6/353	54/24014	1.38E-04	1.38E-02	1.31E-02
BP	GO:1903356	positive regulation of distal tip cell migration	6/353	55/24014	1.53E-04	1.49E-02	1.41E-02
BP	GO:0046633	alpha-beta T cell proliferation	6/353	58/24014	2.06E-04	1.83E-02	1.73E-02
BP	GO:1901076	positive regulation of engulfment of apoptotic cell	6/353	60/24014	2.48E-04	2.10E-02	1.99E-02
BP	GO:0035020	regulation of Rac protein signal transduction	6/353	61/24014	2.72E-04	2.20E-02	2.08E-02

表八、ABBA-BABA 中顯著富集的 GO(續)

Ontology	ID	Description	GeneRatio	BgRatio	pvalue	p.adjust	qvalue
BP	GO:0097628	distal tip cell migration	6/353	62/24014	2.97E-04	2.31E-02	2.19E-02
BP	GO:1903354	regulation of distal tip cell migration	6/353	62/24014	2.97E-04	2.31E-02	2.19E-02
BP	GO:0045058	T cell selection	6/353	66/24014	4.18E-04	2.65E-02	2.51E-02
BP	GO:0060100	positive regulation of phagocytosis, engulfment	6/353	76/24014	8.90E-04	4.70E-02	4.45E-02
BP	GO:1905155	positive regulation of membrane invagination	6/353	76/24014	8.90E-04	4.70E-02	4.45E-02
BP	GO:0050766	positive regulation of phagocytosis	9/353	151/24014	4.10E-04	2.64E-02	2.50E-02
BP	GO:2000427	positive regulation of apoptotic cell clearance	8/353	82/24014	2.79E-05	3.79E-03	3.59E-03
BP	GO:0010883	regulation of lipid storage	9/353	169/24014	9.22E-04	4.81E-02	4.55E-02
BP	GO:2000425	regulation of apoptotic cell clearance	9/353	111/24014	3.92E-05	4.55E-03	4.30E-03
BP	GO:0061158	3'-UTR-mediated mRNA destabilization	4/353	23/24014	3.26E-04	2.40E-02	2.28E-02
BP	GO:0070935	3'-UTR-mediated mRNA stabilization	4/353	28/24014	7.11E-04	4.16E-02	3.94E-02
BP	GO:2000114	regulation of establishment of cell polarity	6/353	60/24014	2.48E-04	2.10E-02	1.99E-02
BP	GO:0032878	regulation of establishment or maintenance of cell polarity	6/353	65/24014	3.85E-04	2.53E-02	2.39E-02
BP	GO:0016319	mushroom body development	9/353	155/24014	4.96E-04	3.10E-02	2.93E-02
BP	GO:0001771	immunological synapse formation	8/353	35/24014	3.35E-08	1.82E-05	1.73E-05
BP	GO:0009988	cell-cell recognition	8/353	98/24014	1.00E-04	1.06E-02	1.00E-02
BP	GO:0071364	cellular response to epidermal growth factor stimulus	6/353	75/24014	8.30E-04	4.58E-02	4.33E-02

表八、ABBA-BABA 中顯著富集的 GO(續)



Ontology	ID	Description	GeneRatio	BgRatio	pvalue	p.adjust	qvalue
BP	GO:0051701	biological process involved in interaction with host adenylate cyclase-activating G	14/353	248/24014	2.02E-05	3.19E-03	3.02E-03
BP	GO:0007189	protein-coupled receptor signaling pathway	13/353	278/24014	2.62E-04	2.17E-02	2.05E-02
BP	GO:0042448	progesterone metabolic process	5/353	25/24014	2.78E-05	3.79E-03	3.59E-03
BP	GO:1901074	regulation of engulfment of apoptotic cell	7/353	82/24014	2.06E-04	1.83E-02	1.73E-02
BP	GO:0060099	regulation of phagocytosis, engulfment	7/353	101/24014	7.37E-04	4.25E-02	4.02E-02
BP	GO:1905153	regulation of membrane invagination	7/353	103/24014	8.28E-04	4.58E-02	4.33E-02
BP	GO:0035821	modulation of process of another organism	11/353	215/24014	3.62E-04	2.42E-02	2.29E-02
BP	GO:0006672	ceramide metabolic process	9/353	148/24014	3.54E-04	2.40E-02	2.28E-02
BP	GO:0006687	glycosphingolipid metabolic process	7/353	104/24014	8.77E-04	4.70E-02	4.45E-02
BP	GO:0019048	modulation by virus of host process	9/353	91/24014	7.86E-06	1.57E-03	1.49E-03
BP	GO:0044003	modulation by symbiont of host process	9/353	102/24014	1.99E-05	3.19E-03	3.02E-03
BP	GO:0019054	modulation by virus of host cellular process	8/353	86/24014	3.94E-05	4.55E-03	4.30E-03
BP	GO:0044068	modulation by symbiont of host cellular process	8/353	96/24014	8.68E-05	9.43E-03	8.93E-03
BP	GO:0048505	regulation of timing of cell differentiation	4/353	16/24014	7.26E-05	8.13E-03	7.70E-03
BP	GO:0033077	T cell differentiation in thymus	8/353	126/24014	5.62E-04	3.34E-02	3.16E-02
CC	GO:0032045	guanyl-nucleotide exchange factor complex	6/373	50/25008	9.66E-05	3.65E-02	3.47E-02
MF	GO:0004308	exo-alpha-sialidase activity	3/341	10/23517	3.36E-04	2.87E-02	2.68E-02

表八、ABBA-BABA 中顯著富集的 GO(續)

Ontology	ID	Description	GeneRatio	BgRatio	pvalue	p.adjust	qvalue
MF	GO:0016997	alpha-sialidase activity	3/341	10/23517	3.36E-04	2.87E-02	2.68E-02
MF	GO:0005096	GTPase activator activity	18/341	479/23517	2.44E-04	2.87E-02	2.68E-02
MF	GO:0042608	T cell receptor binding	6/341	16/23517	6.30E-08	3.76E-05	3.52E-05
MF	GO:0036312	phosphatidylinositol 3-kinase regulatory subunit binding	4/341	14/23517	3.88E-05	5.78E-03	5.42E-03
MF	GO:0017124	SH3 domain binding	10/341	190/23517	4.79E-04	3.57E-02	3.35E-02
MF	GO:0004508	steroid 17-alpha- monooxygenase activity	4/341	12/23517	1.96E-05	3.90E-03	3.65E-03
MF	GO:0047442	17-alpha-hydroxyprogesterone aldolase activity	4/341	12/23517	1.96E-05	3.90E-03	3.65E-03
MF	GO:0016832	aldehyde-lyase activity	4/341	28/23517	6.75E-04	4.48E-02	4.19E-02

表九、ABBA-BABA 中顯著富集的 GO 中包含的基因

GO term count 代表此基因在多少個顯著的 GO term 中出現。

GO term		
count	Gene symbol	Gene name
56	<i>DOCK1</i>	dedicator of cytokinesis 1
8	<i>ZC3H12A</i>	zinc finger CCCH-type containing 12Ab
6	<i>dvl2</i>	dishevelled segment polarity protein 2
6	<i>kpna7</i>	karyopherin alpha 7 (importin alpha 8)
6	<i>neu3.2</i>	sialidase 3 (membrane sialidase), tandem duplicate 2
6	<i>neu3.5</i>	sialidase 3 (membrane sialidase), tandem duplicate 5
5	<i>GIT1</i>	G protein-coupled receptor kinase interacting ArfGAP 1
5	<i>MYO9A</i>	myosin IXAa
5	<i>RXRA</i>	retinoid X receptor, alpha a
5	<i>myo9ab</i>	myosin IXAb
4	<i>CBL</i>	Cbl proto-oncogene, E3 ubiquitin protein ligase
4	<i>DAB2IP</i>	DAB2 interacting protein a
4	<i>ILK</i>	integrin-linked kinase
		RAS p21 protein activator (GTPase activating protein)
4	<i>RASA1</i>	1b
4	<i>c3b.1</i>	complement component c3b, tandem duplicate 1
4	<i>c3b.2</i>	complement component c3b, tandem duplicate 2
4	<i>cyp1c1</i>	cytochrome P450, family 1, subfamily C, polypeptide 1
4	<i>cyp1c2</i>	cytochrome P450, family 1, subfamily C, polypeptide 2
4	<i>si:dkeyp-72a4.1</i>	si:dkeyp-72a4.1
3	<i>ACACA</i>	acetyl-CoA carboxylase alpha
3	<i>CACNA2D1</i>	si:dkey-266j7.2
3	<i>DLG3</i>	discs, large homolog 3 (Drosophila)
3	<i>KRIT1</i>	KRIT1 ankyrin repeat containing
3	<i>LHX2</i>	LIM homeobox 2b
3	<i>PAX6</i>	paired box 6a
3	<i>ROCK2</i>	rho-associated, coiled-coil containing protein kinase 2a
		calcium channel, voltage-dependent, alpha 2/delta subunit
3	<i>cacna2d1a</i>	1a
3	<i>pdzd11</i>	PDZ domain containing 11
	<i>si:ch211-</i>	
3	<i>153f2.3</i>	si:ch211-153f2.3

表九、ABBA-BABA 中顯著富集的 GO 中包含的基因(續)

GO term		
count	Gene symbol	Gene name
	<i>si:dkey-</i>	
3	<i>3lj3.11</i>	si:dkey-3lj3.11
3	<i>st3gal1</i>	ST3 beta-galactoside alpha-2,3-sialyltransferase 1
3	<i>st3gal1l</i>	ST3 beta-galactoside alpha-2,3-sialyltransferase 1, like
2	<i>AKT3</i>	v-akt murine thymoma viral oncogene homolog 3a
2	<i>BCR</i>	BCR activator of RhoGEF and GTPase
2	<i>MAPK14</i>	mitogen-activated protein kinase 14b
2	<i>NAGA</i>	N-acetylgalactosaminidase, alpha
2	<i>QKI</i>	QKI, KH domain containing, RNA binding b
2	<i>RBM24</i>	RNA binding motif protein 24a
2	<i>SMPD1</i>	sphingomyelin phosphodiesterase 1
2	<i>STX1B</i>	syntaxin 1B
2	<i>TBC1D22B</i>	TBC1 domain family, member 22B
2	<i>TSC1</i>	TSC complex subunit 1b
2	<i>prf1.3</i>	perforin 1.3
1	<i>ADGRL2</i>	adhesion G protein-coupled receptor L2a
1	<i>ADGRL4</i>	adhesion G protein-coupled receptor L4
1	<i>ARFGAP2</i>	ADP-ribosylation factor GTPase activating protein 2
		ATPase sarcoplasmic/endoplasmic reticulum Ca2+
1	<i>ATP2A2</i>	transporting 2b
1	<i>CERS5</i>	ceramide synthase 5
1	<i>ELMOD1</i>	ELMO/CED-12 domain containing 1
1	<i>ETNK1</i>	ethanolamine kinase 1
1	<i>HTR4</i>	5-hydroxytryptamine receptor 4
1	<i>NKAPL</i>	NFKB activating protein
1	<i>PAK6</i>	p21 protein (Cdc42/Rac) -activated kinase 6a
1	<i>PTDSS1</i>	phosphatidylserine synthase 1a
1	<i>PTPN13</i>	protein tyrosine phosphatase non-receptor type 13
1	<i>RAB38</i>	RAB38b, member of RAS oncogene family
1	<i>RICTOR</i>	RPTOR independent companion of MTOR, complex 2b
1	<i>SIVA1</i>	SIVA1, apoptosis-inducing factor
1	<i>SOX5</i>	SRY-box transcription factor 5

表九、ABBA-BABA 中顯著富集的 GO 中包含的基因(續)

GO term		
count	Gene symbol	Gene name
1	<i>SSH2</i>	slingshot protein phosphatase 2a
1	<i>STARD3</i>	StAR related lipid transfer domain containing 3
1	<i>TMEM135</i>	transmembrane protein 135
1	<i>VAV2</i>	vav 2 guanine nucleotide exchange factor
1	<i>ZNF609</i>	zinc finger protein 609a
1	<i>ghrh</i>	growth hormone releasing hormone
1	<i>lta4h</i>	leukotriene A4 hydrolase
1	<i>nectin3b</i>	nectin cell adhesion molecule 3b
1	<i>nkl.4</i>	NK-lysin tandem duplicate 4
1	<i>si:ch211-191j22.8</i>	si:ch211-191j22.8

表十、依據核基因全基因組以及粒線體全基因組估算的基因距離

核基因部分依據 PCA 分群作為依據，而粒線體則依據所屬的單系群作為分群，數值皆已乘上百分比，因此使用%標示。

Group 1	Group 2	Nucleotide genetic distance (%)	Mitochondrial genetic distance (%)
North	South	0.49	4.68
Middle	South	0.46	4.66
North, Middle	South	0.47	4.67
North	Middle	0.11	0.31
North, Middle	<i>O. evolans</i>	5.56	9.65

表十一、粒線體與核基因的 π 值

DNA 的 π 值來自於 VCFtools 每 100kb 區塊的計算結果，而粒線體 DNA 則由 MEGA X 的 p distance 模型運算。



Population	Nucleotide DNA mean π	Mitochondrial DNA mean π
North	0.00167	0.00210
Middle	0.00218	0.00274
North, Middle	0.00227	0.00276
South	0.00129	0.00134

表十二、南部族群中顯著的 GO

Ontology	ID	Description	GeneRatio	BgRatio	pvalue	p.adjust	qvalue
MF	GO:0001609	G protein-coupled adenosine receptor activity	4/362	17/23517	1.12E-04	8.54E-03	7.76E-03
BP	GO:0007621	negative regulation of female receptivity	3/363	10/24014	3.80E-04	4.45E-02	4.28E-02
MF	GO:0070324	thyroid hormone binding	3/362	15/23517	1.43E-03	4.37E-02	3.97E-02
MF	GO:0090556	phosphatidylserine floppase activity	4/362	12/23517	2.48E-05	4.31E-03	3.92E-03
MF	GO:0090554	phosphatidylcholine floppase activity	4/362	13/23517	3.54E-05	4.31E-03	3.92E-03
MF	GO:0140328	floppase activity	4/362	14/23517	4.89E-05	4.97E-03	4.51E-03
MF	GO:0008525	phosphatidylcholine transporter activity	4/362	26/23517	6.31E-04	2.40E-02	2.19E-02
BP	GO:0001973	G protein-coupled adenosine receptor signaling pathway	5/363	31/24014	9.44E-05	2.02E-02	1.94E-02
BP	GO:0035588	G protein-coupled purinergic receptor signaling pathway	5/363	36/24014	1.97E-04	2.72E-02	2.61E-02
BP	GO:1905868	regulation of 3'-UTR-mediated mRNA stabilization	5/363	11/24014	3.29E-07	1.27E-03	1.22E-03
BP	GO:0048674	collateral sprouting of injured axon	5/363	14/24014	1.37E-06	2.66E-03	2.55E-03
BP	GO:0070935	3'-UTR-mediated mRNA stabilization	5/363	28/24014	5.67E-05	1.83E-02	1.75E-02
BP	GO:0048691	positive regulation of axon extension involved in regeneration	5/363	31/24014	9.44E-05	2.02E-02	1.94E-02
BP	GO:0048687	positive regulation of sprouting of injured axon	5/363	34/24014	1.49E-04	2.38E-02	2.29E-02
MF	GO:0004709	MAP kinase kinase kinase activity	5/362	36/23517	2.14E-04	1.45E-02	1.32E-02
BP	GO:0030516	regulation of axon extension	16/363	361/24014	1.36E-04	2.38E-02	2.29E-02

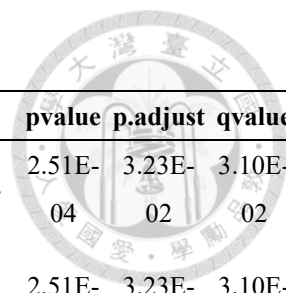
表十二、南部族群中顯著的 GO(續)

Ontology	ID	Description	GeneRatio	BgRatio	pvalue	p.adjust	qvalue
BP	GO:0061387	regulation of extent of cell growth	16/363	393/24014	3.53E-04	4.27E-02	4.10E-02
BP	GO:0048668	collateral sprouting	10/363	110/24014	6.76E-06	5.66E-03	5.43E-03
BP	GO:0031103	axon regeneration	13/363	197/24014	1.03E-05	6.52E-03	6.26E-03
BP	GO:0048679	regulation of axon regeneration	11/363	136/24014	7.31E-06	5.66E-03	5.43E-03
BP	GO:0070570	regulation of neuron projection regeneration	11/363	143/24014	1.18E-05	6.52E-03	6.26E-03
BP	GO:0031102	neuron projection regeneration	14/363	214/24014	5.21E-06	5.66E-03	5.43E-03
BP	GO:0048678	response to axon injury	14/363	254/24014	3.58E-05	1.65E-02	1.58E-02
BP	GO:0048690	regulation of axon extension involved in regeneration	6/363	44/24014	4.98E-05	1.83E-02	1.75E-02
BP	GO:0048677	axon extension involved in regeneration	6/363	45/24014	5.67E-05	1.83E-02	1.75E-02
BP	GO:0048686	regulation of sprouting of injured axon	6/363	47/24014	7.29E-05	1.94E-02	1.86E-02
BP	GO:0048682	sprouting of injured axon	6/363	53/24014	1.44E-04	2.38E-02	2.29E-02
MF	GO:0005212	structural constituent of eye lens	6/362	60/23517	3.17E-04	1.80E-02	1.63E-02
CC	GO:0038039	G protein-coupled receptor heterodimeric complex	4/363	14/25008	3.90E-05	1.59E-02	1.58E-02
MF	GO:0004965	G protein-coupled GABA receptor activity	4/362	12/23517	2.48E-05	4.31E-03	3.92E-03
BP	GO:0045665	negative regulation of neuron differentiation	19/363	466/24014	9.93E-05	2.02E-02	1.94E-02
BP	GO:0010771	negative regulation of cell morphogenesis	12/363	192/24014	3.84E-05	1.65E-02	1.58E-02

表十二、南部族群中顯著的 GO(續)

Ontology	ID	Description	GeneRatio	BgRatio	pvalue	p.adjust	qvalue
MF	GO:0048156	tau protein binding	4/362	30/23517	1.10E-03	3.53E-02	3.21E-02
MF	GO:0004879	nuclear receptor activity	7/362	96/23517	7.13E-04	2.41E-02	2.19E-02
MF	GO:0098531	ligand-activated transcription factor activity	7/362	96/23517	7.13E-04	2.41E-02	2.19E-02
MF	GO:0015079	potassium ion transmembrane transporter activity	13/362	322/23517	1.56E-03	4.53E-02	4.12E-02
MF	GO:0008187	poly-pyrimidine tract binding	7/362	58/23517	2.96E-05	4.31E-03	3.92E-03
MF	GO:0008266	poly (U) RNA binding	6/362	40/23517	3.15E-05	4.31E-03	3.92E-03
MF	GO:0003727	single-stranded RNA binding	10/362	173/23517	3.65E-04	1.85E-02	1.68E-02
CC	GO:0016363	nuclear matrix	9/363	138/25008	1.92E-04	3.92E-02	3.90E-02
BP	GO:0071360	cellular response to exogenous dsRNA	4/363	19/24014	1.66E-04	2.38E-02	2.29E-02
BP	GO:0010878	cholesterol storage	4/363	24/24014	4.30E-04	4.75E-02	4.56E-02
MF	GO:0000339	RNA cap binding	5/362	30/23517	8.73E-05	7.59E-03	6.90E-03
BP	GO:0021654	rhombomere boundary formation	5/363	29/24014	6.76E-05	1.94E-02	1.86E-02
BP	GO:0021594	rhombomere formation	5/363	30/24014	8.02E-05	1.94E-02	1.86E-02
BP	GO:0021593	rhombomere morphogenesis	5/363	39/24014	2.90E-04	3.62E-02	3.47E-02
BP	GO:0048685	negative regulation of collateral sprouting of intact axon in response to injury	4/363	17/24014	1.05E-04	2.02E-02	1.94E-02
BP	GO:0097155	fasciculation of sensory neuron axon	4/363	19/24014	1.66E-04	2.38E-02	2.29E-02

表十二、南部族群中顯著的 GO(續)



Ontology	ID	Description	GeneRatio	BgRatio	pvalue	p.adjust	qvalue
BP	GO:0048673	collateral sprouting of intact axon in response to injury	4/363	21/24014	2.51E-04	3.23E-02	3.10E-02
BP	GO:0048683	regulation of collateral sprouting of intact axon in response to injury	4/363	21/24014	2.51E-04	3.23E-02	3.10E-02
BP	GO:0003404	optic vesicle morphogenesis	4/363	24/24014	4.30E-04	4.75E-02	4.56E-02
MF	GO:0005004	GPI-linked ephrin receptor activity	4/362	22/23517	3.24E-04	1.80E-02	1.63E-02
BP	GO:0048681	negative regulation of axon regeneration	6/363	51/24014	1.16E-04	2.14E-02	2.05E-02
BP	GO:0070571	negative regulation of neuron projection regeneration	6/363	54/24014	1.60E-04	2.38E-02	2.29E-02
MF	GO:0005229	intracellular calcium activated chloride channel activity	4/362	25/23517	5.41E-04	2.20E-02	2.00E-02
MF	GO:0061778	intracellular chloride channel activity	4/362	25/23517	5.41E-04	2.20E-02	2.00E-02
MF	GO:0140303	intramembrane lipid transporter activity	6/362	63/23517	4.14E-04	1.94E-02	1.76E-02
BP	GO:0007613	memory	19/363	457/24014	7.70E-05	1.94E-02	1.86E-02

表十三、南部族群中顯著的 GO 中包含的基因

GO term count 代表此基因在多少個顯著的 GO term 中出現。

GO term count	Gene symbol	Gene name
21	<i>ephb1</i>	EPH receptor B1
21	<i>ephb2b</i>	eph receptor B2b
21	<i>ephb4b</i>	eph receptor B4b
19	<i>si:ch211-45c16.2</i>	si:ch211-45c16.2
15	<i>dusp16</i>	dual specificity phosphatase 16
8	<i>ptenb</i>	phosphatase and tensin homolog B
5	<i>abca1a</i>	ATP-binding cassette, sub-family A (ABC1), member 1A
5	<i>abca1b</i>	ATP-binding cassette, sub-family A (ABC1), member 1B
5	<i>abca4a</i>	ATP-binding cassette, sub-family A (ABC1), member 4a
5	<i>gsk3ba</i>	glycogen synthase kinase 3 beta, genome duplicate a
5	<i>ifit10</i>	interferon-induced protein with tetratricopeptide repeats 10
5	<i>jak2a</i>	Janus kinase 2a
5	<i>thrb</i>	thyroid hormone receptor beta
4	<i>gdi1</i>	GDP dissociation inhibitor 1
4	<i>mib1</i>	MIB E3 ubiquitin protein ligase 1
4	<i>trpc5a</i>	transient receptor potential cation channel, subfamily C, member 5a
	<i>adenosine receptor</i>	
3	<i>A1-like</i>	adenosine receptor A1-like
3	<i>adora3a.1</i>	adenosine A3 receptor a.1
3	<i>ano3</i>	anoctamin 3
		amyloid beta (A4) precursor protein-binding, family B, member 1
3	<i>apbb1ip</i>	interacting protein
3	<i>cnga3b</i>	cyclic nucleotide gated channel subunit alpha 3b
3	<i>khdrbs2</i>	KH domain containing, RNA binding, signal transduction associated 2
3	<i>patl2</i>	PAT1 homolog 2
3	<i>phox2bb</i>	paired like homeobox 2Bb
3	<i>pspc1</i>	paraspeckle component 1
3	<i>sa20824</i>	sa20824
3	<i>ssb</i>	small RNA binding exonuclease protection factor La
3	<i>zeb2b</i>	zinc finger E-box binding homeobox 2b
2	<i>HNF4a</i>	hepatocyte nuclear factor 4, alpha
2	<i>ada</i>	adenosine deaminase

表十三、南部族群中顯著的 GO 中包含的基因(續)

GO term count	Gene symbol	Gene name
2	<i>ano2b</i>	anoctamin 2b
2	<i>ano8a</i>	anoctamin 8a
2	<i>gpr156</i>	G protein-coupled receptor 156
2	<i>matn3a</i>	matrilin 3a
2	<i>nr1d2b</i>	nuclear receptor subfamily 1, group D, member 2b
2	<i>nr2f6a</i>	nuclear receptor subfamily 2, group F, member 6a
2	<i>nr4a2a</i>	nuclear receptor subfamily 4, group A, member 2a
2	<i>si:ch1073-446e18.3</i>	tetratricopeptide repeat domain 3
2	<i>ttc3</i>	tetratricopeptide repeat domain 3
1	<i>KCNJ15</i>	potassium inwardly rectifying channel subfamily J member 15
1	<i>SLC24A2</i>	si:cabz01080888.1
1	<i>ago4</i>	argonaute RISC component 4
1	<i>atad1b</i>	ATPase family AAA domain containing 1b
1	<i>atp1b1b</i>	ATPase Na ⁺ /K ⁺ transporting subunit beta 1b
1	<i>atxn1l</i>	ataxin 1-like
1	<i>ccnd2b</i>	cyclin D2, b
1	<i>celf5a</i>	cugbp, Elav-like family member 5a
1	<i>ciz1b</i>	cdkn1a interacting zinc finger protein 1b
1	<i>colec12</i>	collectin sub-family member 12
1	<i>crhr1</i>	corticotropin releasing hormone receptor 1
1	<i>crta1a</i>	CREB regulated transcription coactivator 1a
1	<i>crygm2d15</i>	crystallin, gamma M2d15
1	<i>crygm2f</i>	crystallin, gamma M2f
1	<i>crygm3</i>	crystallin, gamma M3
1	<i>crygs3</i>	crystallin, gamma S3
1	<i>crygs4</i>	crystallin, gamma S4
1	<i>dpys</i>	dihydropyrimidinase
1	<i>dyrk1aa</i>	dual-specificity tyrosine- (Y) -phosphorylation regulated kinase 1A, a
1	<i>ednrba</i>	endothelin receptor Ba
		eukaryotic translation initiation factor 4E family member 2 related
1	<i>EIF4E2RS1</i>	sequence 1
1	<i>EIF4G3A</i>	eukaryotic translation initiation factor 4 gamma, 3a
1	<i>eps8a</i>	EGFR pathway substrate 8a, signaling adaptor

表十三、南部族群中顯著的 GO 中包含的基因(續)

GO term count	Gene symbol	Gene name
1	<i>erc1b</i>	ELKS/RAB6-interacting/CAST family member 1b holocarboxylase synthetase (biotin- (propionyl-CoA-carboxylase (ATP-hydrolysing)) ligase)
1	<i>hlcs</i>	
1	<i>insra</i>	insulin receptor a
1	<i>kcnc2</i>	potassium voltage-gated channel, Shaw-related subfamily, member 2
1	<i>kcni3a</i>	potassium inwardly rectifying channel subfamily J member 3a
1	<i>kcni6</i>	potassium inwardly rectifying channel subfamily J member 6 potassium intermediate/small conductance calcium-activated channel, subfamily N, member 1a
1	<i>kcni1a</i>	
1	<i>mavs</i>	mitochondrial antiviral signaling protein
1	<i>ntf3</i>	neurotrophin 3
1	<i>ppig</i>	peptidylprolyl isomerase G (cyclophilin G)
1	<i>prdm1b</i>	PR domain containing 1b, with ZNF domain
1	<i>si:ch211-247n2.1</i>	S100 calcium binding protein P
1	<i>si:dkey-162b23.4</i>	si:dkey-162b23.4
1	<i>si:dkey-43k4.5</i>	si:dkey-43k4.5
1	<i>si:rp71-39b20.4</i>	si:rp71-39b20.4
1	<i>sorbs1</i>	sorbin and SH3 domain containing 1
1	<i>sptbn2</i>	spectrin, beta, non-erythrocytic 2
1	<i>stard4</i>	StAR related lipid transfer domain containing 4
1	<i>stxbp5a</i>	syntaxin binding protein 5a (tomosyn)
1	<i>ttc39b</i>	tetratricopeptide repeat domain 39B
1	<i>yeats4</i>	YEATS domain containing 4
1	<i>zgc:158403</i>	zgc:158403

表十四、北部族群的富集的 GO

Ontology	ID	Description	GeneRatio	BgRatio	pvalue	p.adjust	qvalue
BP	GO:1902075	cellular response to salt	14/308	369/24014	3.22E-04	3.41E-02	3.21E-02
BP	GO:0071103	DNA conformation change	13/308	340/24014	4.87E-04	4.39E-02	4.14E-02
BP	GO:0040029	epigenetic regulation of gene expression	16/308	481/24014	5.29E-04	4.52E-02	4.27E-02
BP	GO:1990834	response to odorant	4/308	28/24014	4.26E-04	4.14E-02	3.90E-02
BP	GO:0043488	regulation of mRNA stability	13/308	248/24014	2.10E-05	4.19E-03	3.95E-03
BP	GO:0061013	regulation of mRNA catabolic process	13/308	264/24014	4.02E-05	6.53E-03	6.16E-03
BP	GO:0043487	regulation of RNA stability	13/308	265/24014	4.18E-05	6.53E-03	6.16E-03
BP	GO:0045727	positive regulation of translation	10/308	207/24014	3.61E-04	3.70E-02	3.49E-02
BP	GO:0030261	chromosome condensation	8/308	113/24014	1.07E-04	1.53E-02	1.45E-02
BP	GO:0045739	positive regulation of DNA repair	8/308	75/24014	5.35E-06	1.37E-03	1.30E-03
BP	GO:0006282	regulation of DNA repair	8/308	132/24014	3.11E-04	3.41E-02	3.21E-02
BP	GO:2001034	positive regulation of double-strand break repair via nonhomologous end joining	7/308	16/24014	5.53E-10	9.92E-07	9.36E-07
BP	GO:2001032	regulation of double-strand break repair via nonhomologous end joining	7/308	21/24014	5.32E-09	4.77E-06	4.50E-06
BP	GO:2000781	positive regulation of double-strand break repair	7/308	36/24014	3.24E-07	2.32E-04	2.19E-04
BP	GO:2000779	regulation of double-strand break repair	7/308	82/24014	8.91E-05	1.33E-02	1.26E-02
BP	GO:0051306	mitotic sister chromatid separation	8/308	115/24014	1.21E-04	1.67E-02	1.57E-02
BP	GO:0007060	male meiosis chromosome segregation	6/308	25/24014	6.12E-07	2.75E-04	2.59E-04
BP	GO:0000712	resolution of meiotic recombination intermediates	6/308	28/24014	1.26E-06	4.69E-04	4.43E-04
BP	GO:0006265	DNA topological change	6/308	29/24014	1.57E-06	4.70E-04	4.43E-04
BP	GO:0071318	cellular response to ATP	6/308	29/24014	1.57E-06	4.70E-04	4.43E-04

表十四、北部族群的富集的 GO(續)

Ontology	ID	Description	GeneRatio	BgRatio	pvalue	p.adjust	qvalue
BP	GO:0051307	meiotic chromosome separation	6/308	41/24014	1.31E-05	3.13E-03	2.95E-03
BP	GO:0033198	response to ATP	6/308	61/24014	1.30E-04	1.73E-02	1.63E-02
BP	GO:0007076	mitotic chromosome condensation	7/308	39/24014	5.77E-07	2.75E-04	2.59E-04
BP	GO:0015074	DNA integration	20/308	160/24014	2.24E-14	8.04E-11	7.58E-11
BP	GO:0032196	transposition	10/308	60/24014	4.45E-09	4.77E-06	4.50E-06
BP	GO:0006310	DNA recombination	19/308	442/24014	5.03E-06	1.37E-03	1.30E-03
BP	GO:0006313	DNA transposition	5/308	44/24014	2.42E-04	2.90E-02	2.73E-02
BP	GO:0032197	retrotransposition	5/308	16/24014	1.31E-06	4.69E-04	4.43E-04
BP	GO:0042542	response to hydrogen peroxide	10/308	199/24014	2.64E-04	3.05E-02	2.88E-02
BP	GO:0051054	positive regulation of DNA metabolic process	13/308	331/24014	3.78E-04	3.77E-02	3.56E-02
BP	GO:0070301	cellular response to hydrogen peroxide	8/308	95/24014	3.09E-05	5.29E-03	4.99E-03
BP	GO:0006303	double-strand break repair via nonhomologous end joining	9/308	74/24014	4.45E-07	2.66E-04	2.51E-04
BP	GO:0007131	reciprocal meiotic recombination	8/308	87/24014	1.62E-05	3.43E-03	3.23E-03
BP	GO:0140527	reciprocal homologous recombination	8/308	87/24014	1.62E-05	3.43E-03	3.23E-03
BP	GO:0035825	homologous recombination	8/308	93/24014	2.65E-05	4.75E-03	4.48E-03
BP	GO:0007127	meiosis I	9/308	178/24014	5.08E-04	4.45E-02	4.20E-02
BP	GO:0031670	cellular response to nutrient	9/308	120/24014	2.52E-05	4.75E-03	4.48E-03
BP	GO:0017085	response to insecticide	5/308	51/24014	4.87E-04	4.39E-02	4.14E-02
BP	GO:0051610	serotonin uptake	3/308	10/24014	2.35E-04	2.90E-02	2.73E-02
BP	GO:1904647	response to rotenone	3/308	10/24014	2.35E-04	2.90E-02	2.73E-02
BP	GO:0051581	negative regulation of neurotransmitter uptake	3/308	11/24014	3.19E-04	3.41E-02	3.21E-02
BP	GO:0098700	neurotransmitter loading into synaptic vesicle	3/308	13/24014	5.43E-04	4.54E-02	4.28E-02
BP	GO:0006837	serotonin transport	4/308	29/24014	4.89E-04	4.39E-02	4.14E-02

表十四、北部族群的富集的 GO(續)

Ontology	ID	Description	GeneRatio	BgRatio	pvalue	p.adjust	qvalue
CC	GO:0008623	CHRA	6/294	10/25008	5.06E-10	1.88E-07	1.83E-07
CC	GO:0031010	ISWI-type complex	6/294	24/25008	2.83E-07	5.26E-05	5.10E-05
CC	GO:0060198	clathrin-sculpted vesicle	4/294	19/25008	6.31E-05	7.83E-03	7.60E-03
MF	GO:0003696	satellite DNA binding	6/325	14/23517	1.82E-08	3.30E-06	3.03E-06
MF	GO:0003916	DNA topoisomerase activity	6/325	21/23517	3.03E-07	4.12E-05	3.78E-05
MF	GO:0000400	four-way junction DNA binding	6/325	37/23517	1.08E-05	1.17E-03	1.08E-03
MF	GO:0000217	DNA secondary structure binding	6/325	59/23517	1.62E-04	8.81E-03	8.08E-03
MF	GO:0004908	interleukin-1 receptor activity	6/325	14/23517	1.82E-08	3.30E-06	3.03E-06
MF	GO:0140097	catalytic activity, acting on DNA	14/325	379/23517	8.72E-04	3.96E-02	3.63E-02
MF	GO:0003727	single-stranded RNA binding	10/325	173/23517	1.54E-04	8.81E-03	8.08E-03
MF	GO:0016853	isomerase activity	11/325	240/23517	5.44E-04	2.69E-02	2.47E-02
MF	GO:0042826	histone deacetylase binding	10/325	158/23517	7.28E-05	4.95E-03	4.54E-03
MF	GO:0000182	rDNA binding	8/325	18/23517	4.74E-11	2.58E-08	2.36E-08
MF	GO:0008504	monoamine transmembrane transporter activity	4/325	16/23517	5.72E-05	4.72E-03	4.33E-03
MF	GO:0005080	protein kinase C binding	8/325	97/23517	6.07E-05	4.72E-03	4.33E-03

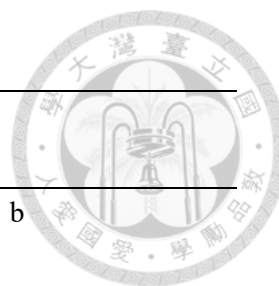
表十五、北部族群中顯著的 GO 中包含的基因

GO term count 代表此基因在多少個顯著的 GO term 中出現。

GO term count	Gene symbol	Gene name
39	<i>top2b</i>	Topoisomerase II Beta
9	<i>slc18a1</i>	solute carrier family 18 member A1
9	<i>mre11a</i>	Meiotic recombination 11
8	<i>fh</i>	Fumarate Hydratase
7	<i>mtor</i>	mammalian target of rapamycin
7	<i>fxr2</i>	FMR1 Autosomal Homolog 2
5	<i>dis3l2</i>	DIS3L2 exoribonuclease
5	<i>pum2</i>	Pumilio homolog 2
5	<i>ddx3xb</i>	DEAD-box helicase 3 X-linked b
5	<i>elavl1b</i>	Enables mRNA 3'-UTR binding activity
5	<i>pabpc1l</i>	poly (A) binding protein, cytoplasmic 1-like
4	<i>parga</i>	poly (ADP-ribose) glycohydrolase a
4	<i>ywhabb</i>	14-3-3 protein beta/alpha
4	<i>fastkd3</i>	FAST kinase domain-containing protein 3
3	<i>mapkapk2b</i>	MAP kinase-activated protein kinase 2
3	<i>slc18a3a</i>	solute carrier family 18 member A3
3	<i>zfp36l2</i>	ZFP36 Ring Finger Protein Like 2
3	<i>ptk2bb</i>	protein tyrosine kinase 2 beta, b
3	<i>ttn.2</i>	tandem duplicate 2
2	<i>rev3l</i>	Protein reversionless 3-like
2	<i>acin1b</i>	apoptotic chromatin condensation inducer 1b
2	<i>hdac4</i>	Histone Deacetylase 4
2	<i>tgfb1a</i>	transforming growth factor, beta 1a
2	<i>adam9a</i>	ADAM metalloproteinase domain 9a
2	<i>il1rl2</i>	Interleukin-1 Receptor-Like 2
2	<i>lmnb1</i>	Lamin B1
2	<i>ppp5c</i>	Protein Phosphatase 5 Catalytic Subunit
2	<i>wfs1b</i>	Wolfram syndrome 1b
2	<i>zmp:0000000936</i>	
1	<i>gatad1</i>	GATA zinc finger domain containing 1
1	<i>pik3cb</i>	phosphatidylinositol-4,5-bisphosphate 3-kinase, catalytic subunit beta

表十五、北部族群中顯著的 GO 中包含的基因(續)

GO term count	Gene symbol	Gene name
1	<i>add3b</i>	adducin 3 (gamma) b
1	<i>ajuba</i>	ajuba LIM protein
1	<i>becn1</i>	ajuba LIM protein
1	<i>cdh30</i>	cadherin 30
1	<i>cenpe</i>	Centromere protein E
1	<i>cpne8</i>	copine VIII
1	<i>dse</i>	Dermatan-sulfate epimerase
1	<i>fkbp10a</i>	FKBP prolyl isomerase 10a
1	<i>il1rap1a</i>	interleukin 1 receptor accessory protein-like 1a
1	<i>il1rap1b</i>	interleukin 1 receptor accessory protein-like 1b
1	<i>il1rapl2</i>	X-linked interleukin-1 receptor accessory protein-like 2
1	<i>kcnh1b</i>	potassium voltage-gated channel, subfamily H (eag-related) , member 1b
1	<i>kdm1a</i>	Lysine-specific histone demethylase 1A
1	<i>med1</i>	Mediator of RNA polymerase II transcription subunit 1
1	<i>mier1b</i>	mesoderm induction early response 1b, transcriptional regulator
1	<i>p3h4</i>	prolyl 3-hydroxylase family member 4 (inactive)
1	<i>per2</i>	period circadian clock 2
1	<i>ppiaa</i>	peptidylprolyl isomerase Aa (cyclophilin A)
1	<i>rbpms</i>	RNA binding protein, mRNA processing factor
1	<i>rgs7b</i>	regulator of G protein signaling 7b
1	<i>si:ch211-212k5.3</i>	
1	<i>si:ch211-57f7.7</i>	
1	<i>si:dkey-17e16.9</i>	
1	<i>syt4</i>	Synaptotagmin-4
1	<i>usp9</i>	ubiquitin specific peptidase 9



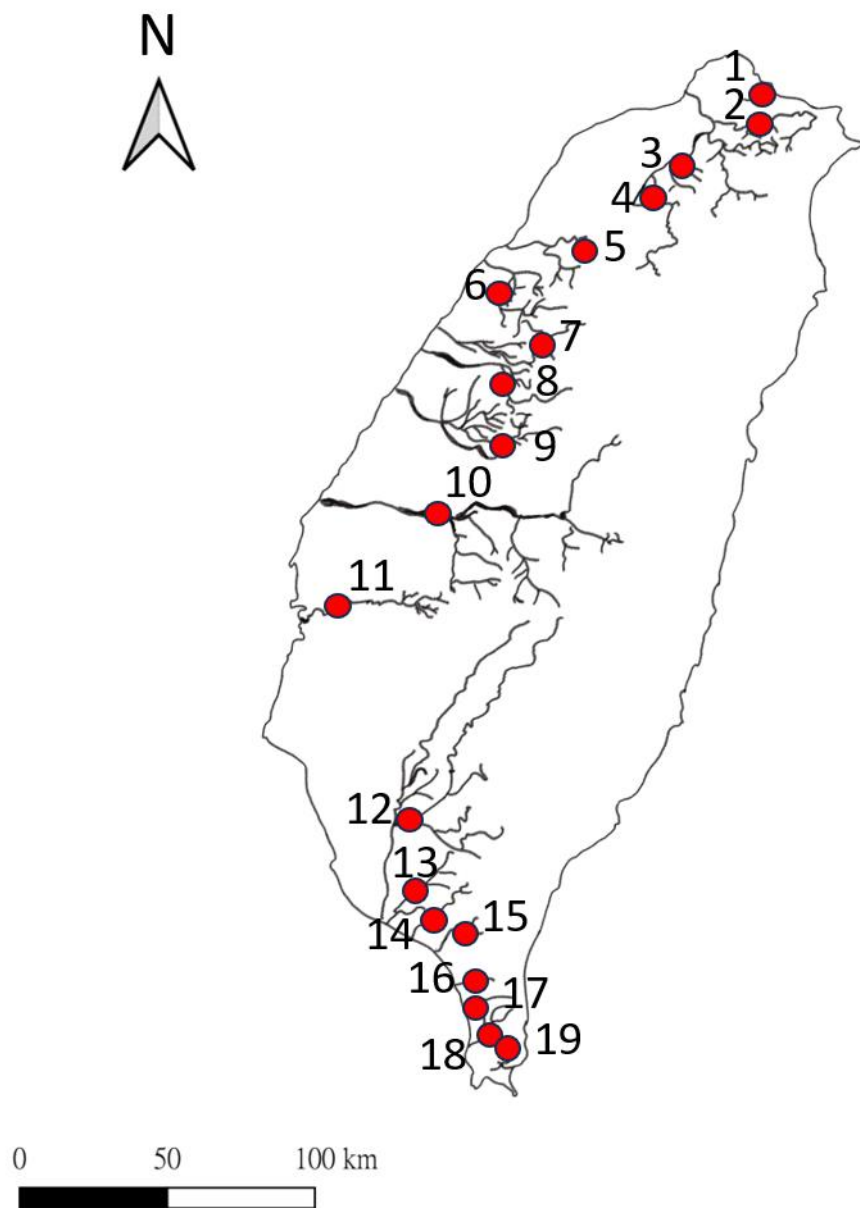
表十六、南部族群顯著的 KEGG

ID	Description	GeneRatio	BgRatio	pvalue	p.adjust	qvalue
ko00780	Biotin metabolism	5/486	10/19038	2.41E-06	5.81E-04	5.49E-04
ko04919	Thyroid hormone signaling pathway	20/486	241/19038	4.09E-06	5.81E-04	5.49E-04

圖

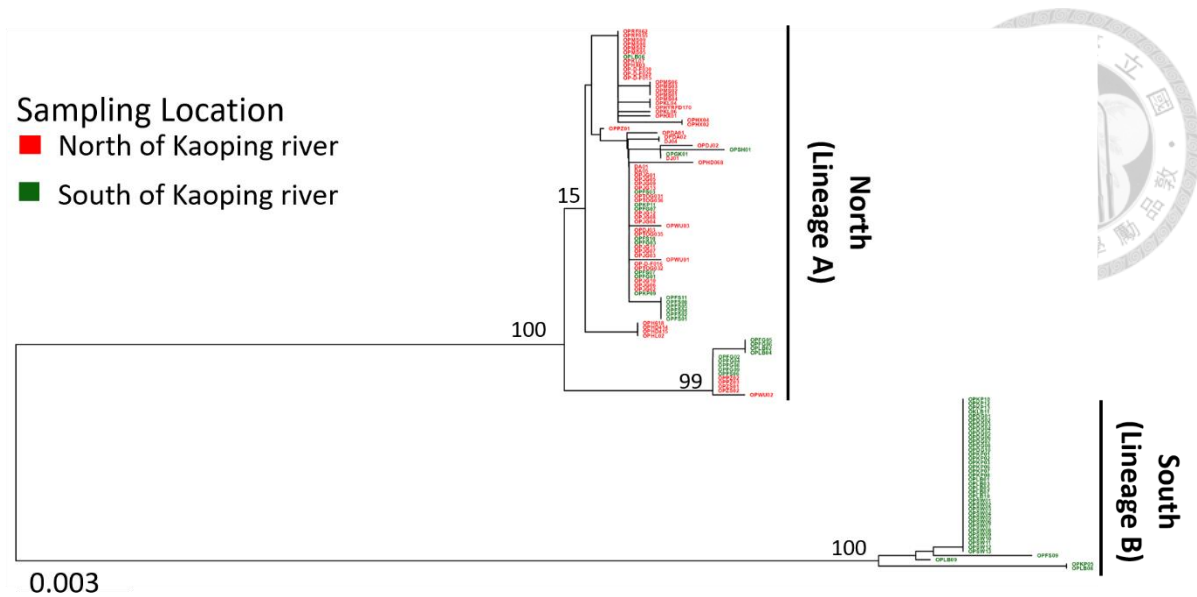


圖一、粗首馬口鱖的照片



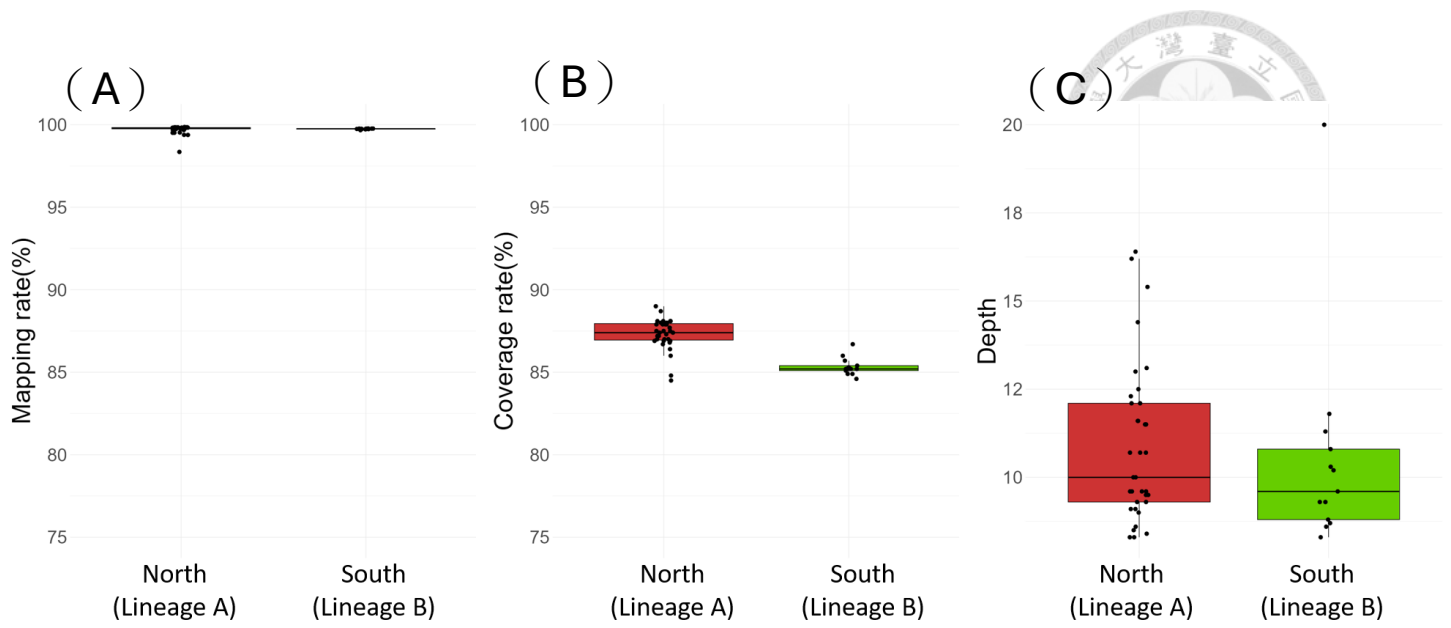
圖二、台灣地圖及河川，紅色的點位代表有採集的河川。

代號及河川名稱請參考（表二）。



圖三、由 cytb 構建的 Neighbor-Joining tree

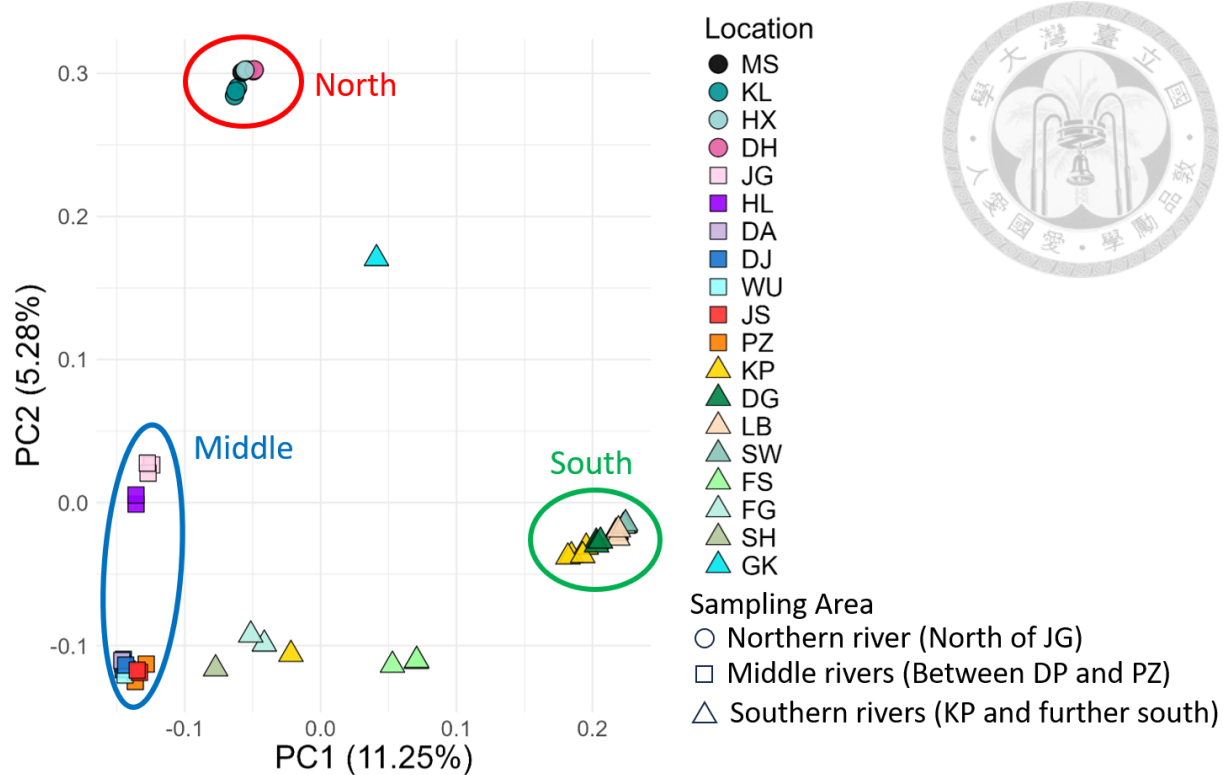
共包含 124 隻個體。紅色代表的是採自高屏溪（含）以北與高屏溪（含）以南的個體，位於演化枝上的數字是該分支在重抽樣分析中被支持的比例。



圖四、序列的對比率、覆蓋率與定序深度

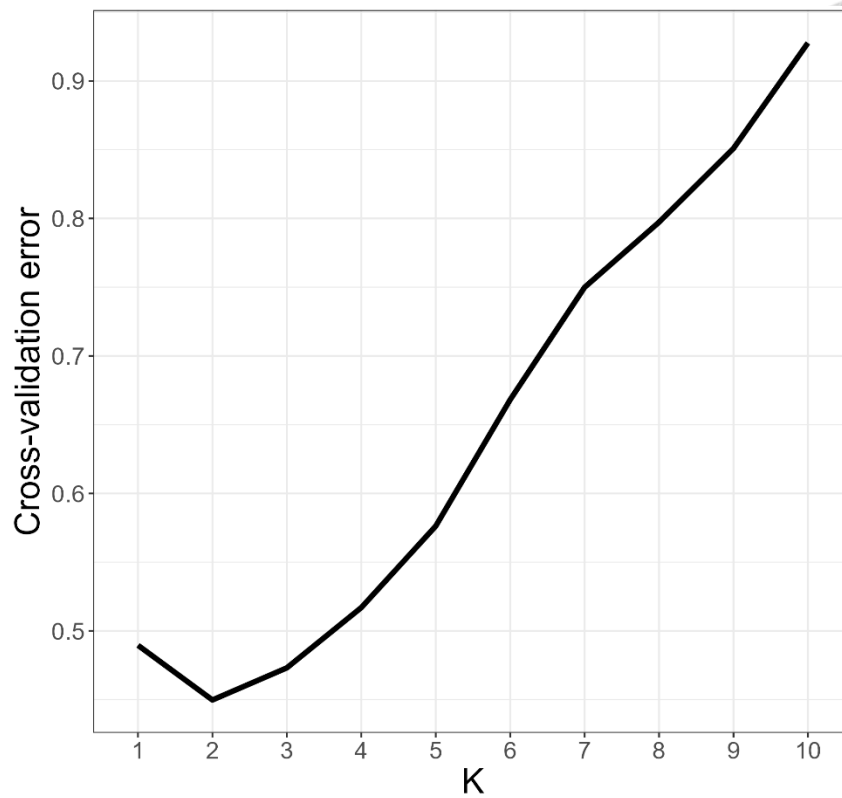
定序的結果，紅色與綠色分別代表 *cytb* 中屬於高屏溪以北與高屏溪以南的譜系。

圖 A 為對比率、圖 B 為覆蓋率、圖 C 為定序深度。



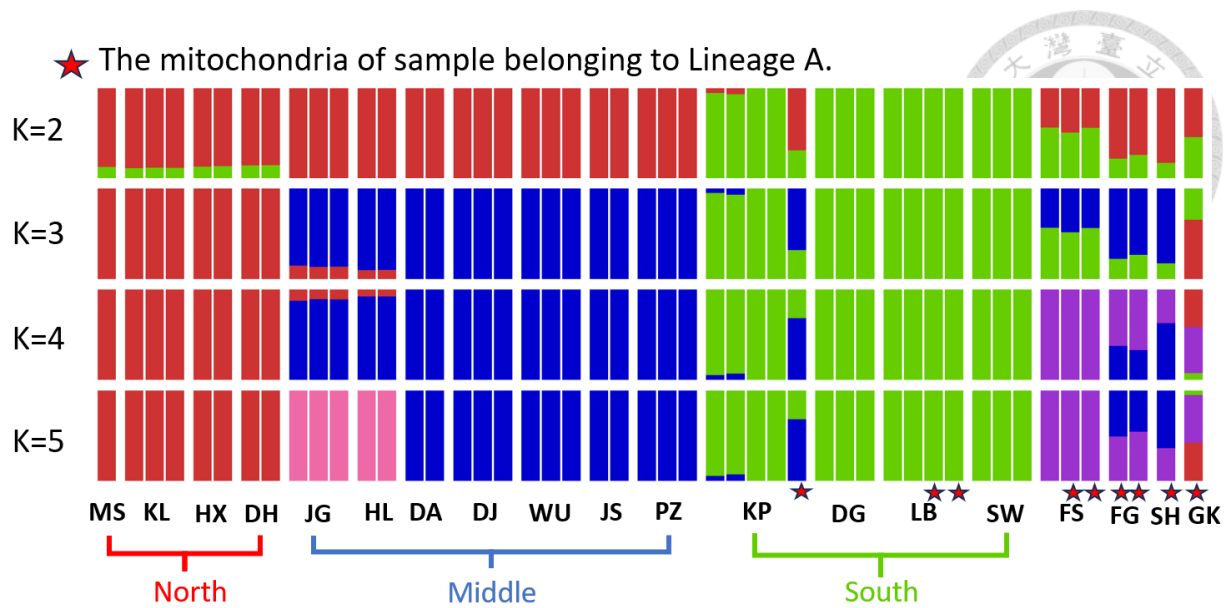
圖五、全基因組構建的 PCA

由圓圈圈起來的為目前認定的三個族群。地點標示分別為瑪鍊溪 (MS)、基隆河 (KL)、橫溪 (HX)、大漢溪 (DH)、中港溪 (JG)、後龍溪 (HL)、大安溪 (DA)、大甲溪 (DJ)、烏溪 (WU)、樸子溪 (PZ)、高屏溪 (KP)、東港溪 (DG)、林邊溪 (LB)、士文溪 (SW)、枋山溪 (FS)、楓港溪 (FG)、四重溪 (SH) 以及港口溪 (GK)。



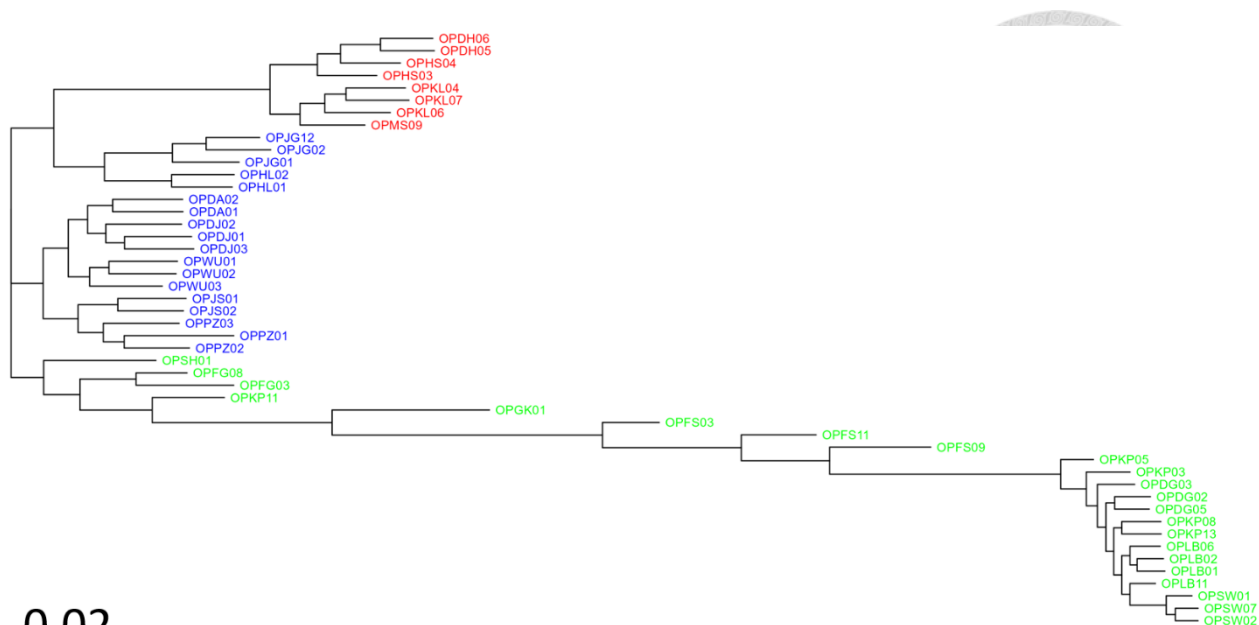
圖六、CV error 值

ADMIXTURE 預測的 CV error 值以及最佳的族群分群，圖中指出分為兩群(K=2)為最佳分群。



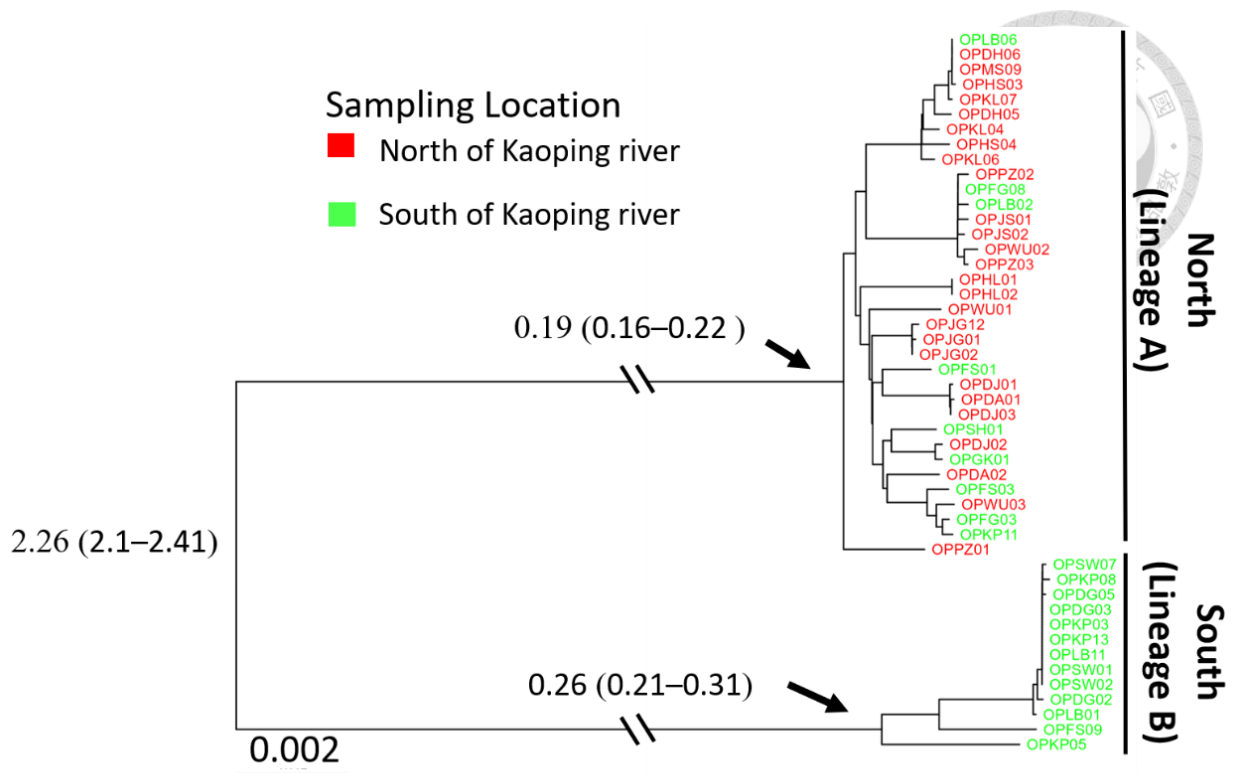
圖七、全基因組構建的 ADMIXTURE

由紅色星號標註的為屬於北部型粒線體的個體，而北（North）、中（Middle）以及南（South）各別代表三個 PCA 上的族群，其中的高屏溪不包含屬於北部粒線體的個體。地點標示分別為瑪鍊溪（MS）、基隆河（KL）、橫溪（HX）、大漢溪（DH）、中港溪（JG）、後龍溪（HL）、大安溪（DA）、大甲溪（DJ）、烏溪（WU）、樸子溪（PZ）、高屏溪（KP）、東港溪（DG）、林邊溪（LB）、士文溪（SW）、枋山溪（FS）、楓港溪（FG）、四重溪（SH）以及港口溪（GK）。



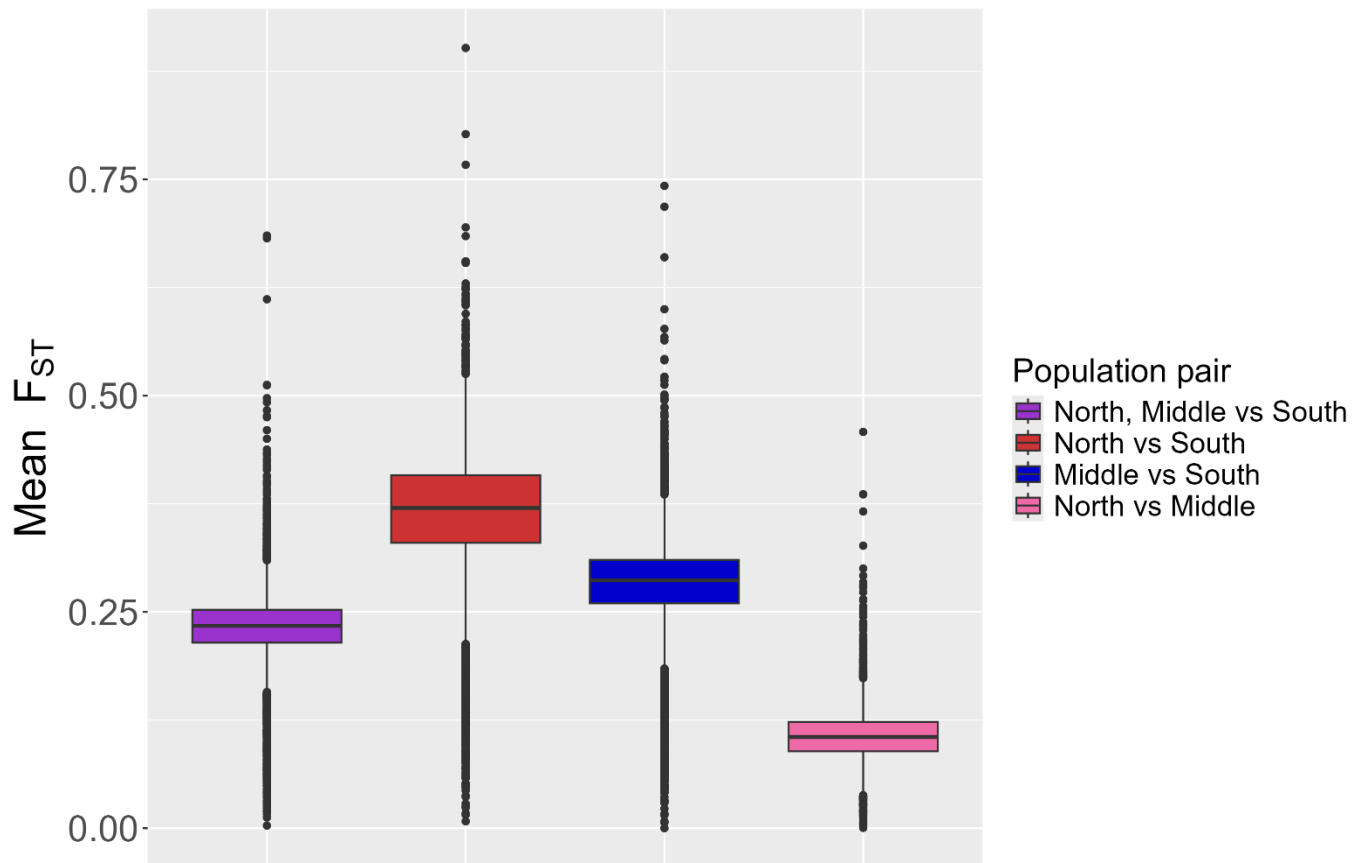
圖八、全基因組構建的最大似然樹

北（North）、中（Middle）以及南（South）各別代表三個 PCA 上的族群。紅色代表淡水河及瑪鍊溪採集的樣本；藍色代表中港溪至朴子溪採集的樣本；綠色代表高屏溪以南採集的樣本。



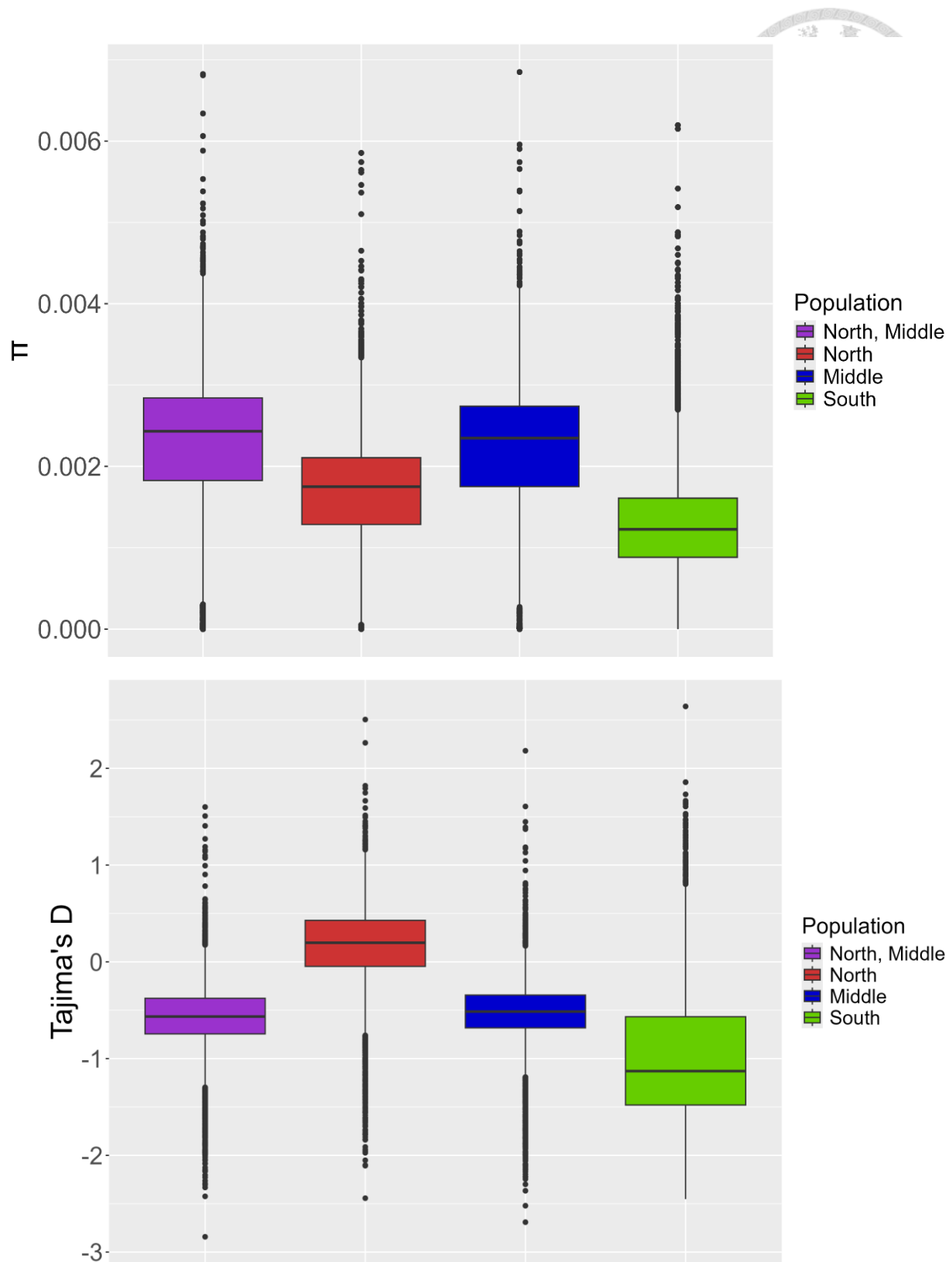
圖九、由全粒線體基因組構建的 Neighbor-Joining tree

共包含 48 隻個體。紅色代表的是採自高屏溪（含）以北與高屏溪（含）以南的個體，數字代表的是最近共祖時間，單位為百萬年。



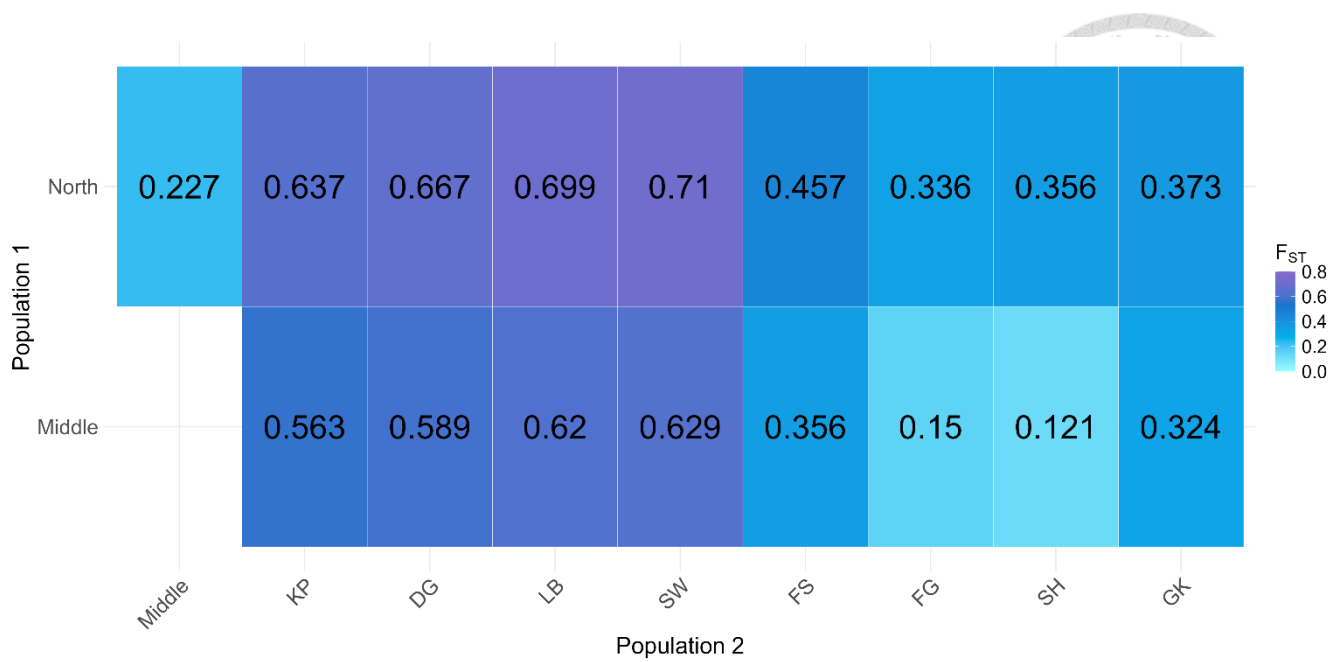
圖十、基於全基因組所有點位計算的 F_{ST}

以每 100kb 為一個區塊進行估算的結果，每一個點都唯一一個窗口。



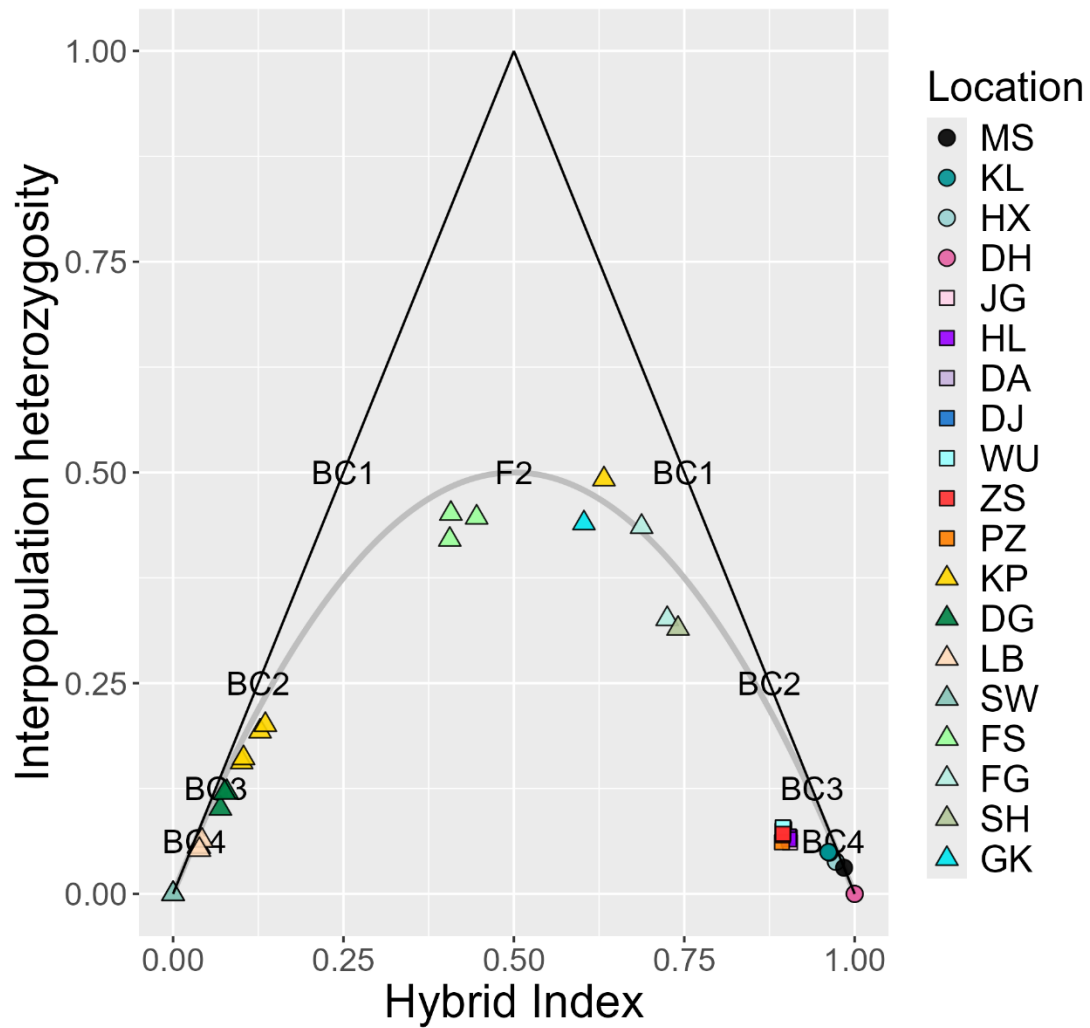
圖十一、基於全基因組所有點位計算的 π 以及 Tajima's D

以每 100kb 為一個區塊進行估算的結果，每一個點都為一個窗口。



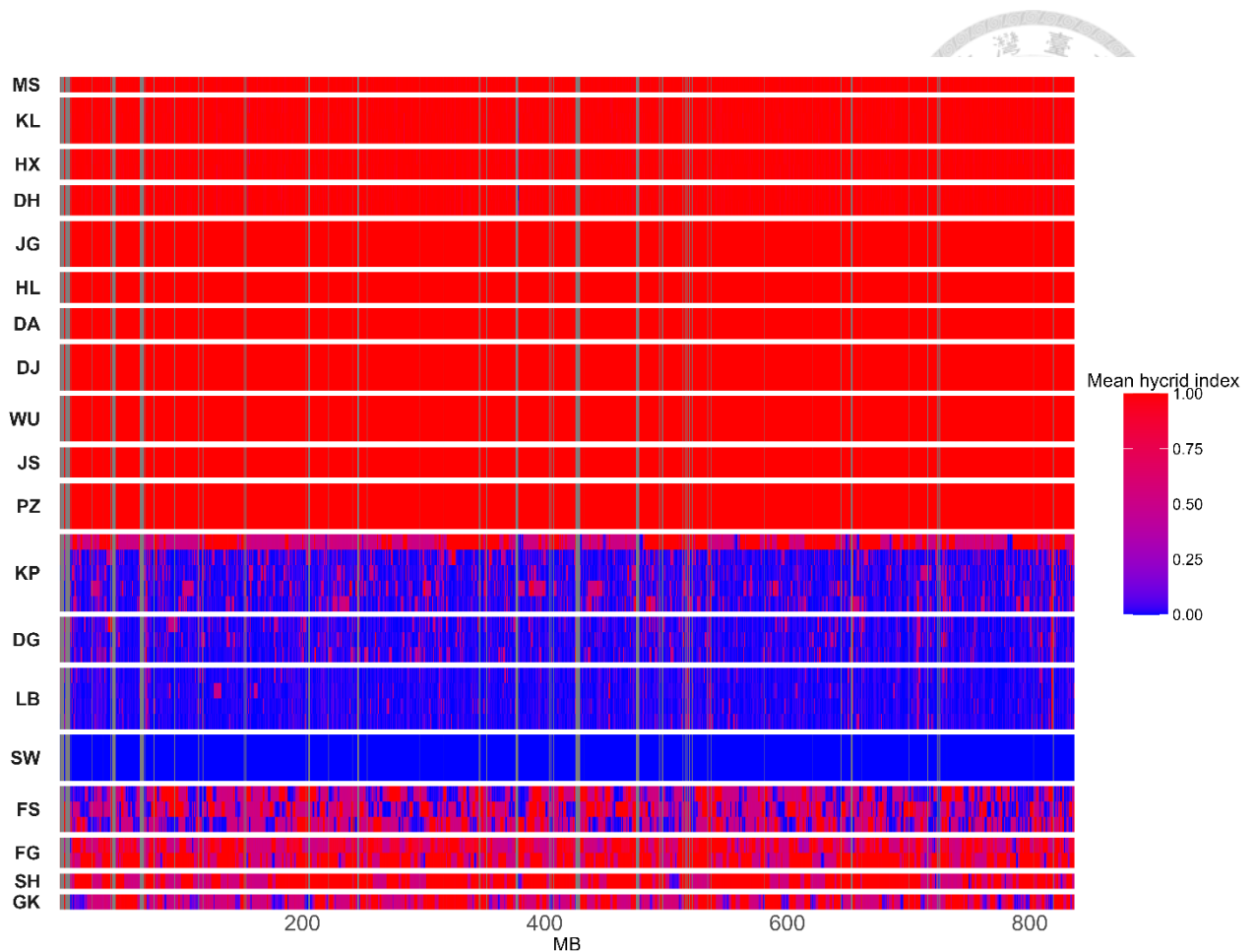
圖十二、由每個點位估算的基因組平均 Weir and Cockerham weighted F_{ST}

地點標示分別為北部族群 (North)、中部族群 (Middle)、高屏溪 (KP)、東港溪 (DG)、林邊溪 (LB)、士文溪 (SW)、枋山溪 (FS)、楓港溪 (FG)、四重溪 (SH) 以及港口溪 (GK)。



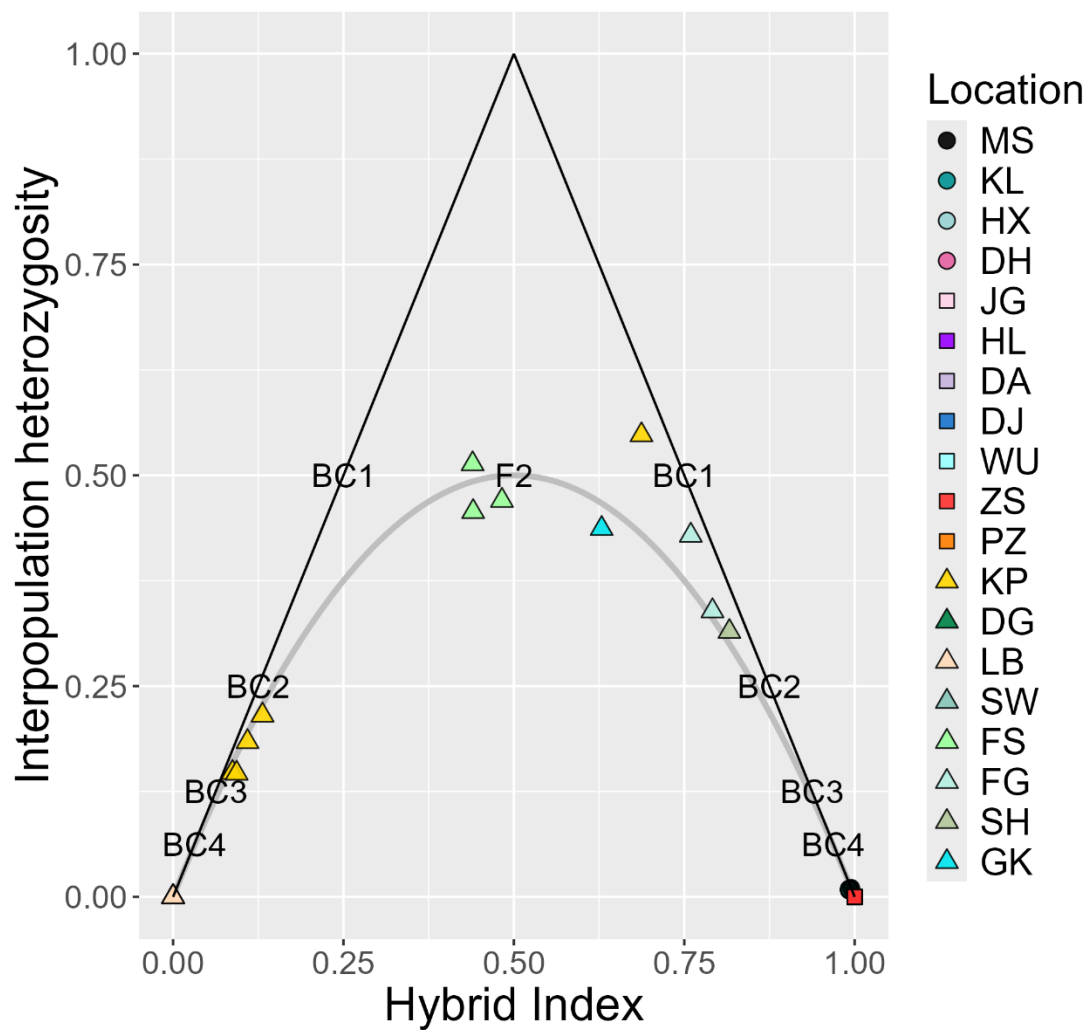
圖十三、由族群間固定的且沒有缺失的 SNP 構建的三角散點圖

以士文溪的個體及中部族群作為親代，由族群間固定的且沒有缺失的 SNP 構建的三角散點圖。BC 代表回交個體，數字代表第幾次回交；F2 為早期雜交個體。地點標示分別為瑪鍊溪 (MS)、基隆河 (KL)、橫溪 (HX)、大漢溪 (DH)、中港溪 (JG)、後龍溪 (HL)、大安溪 (DA)、大甲溪 (DJ)、烏溪 (WU)、樸子溪 (PZ)、高屏溪 (KP)、東港溪 (DG)、林邊溪 (LB)、士文溪 (SW)、枋山溪 (FS)、楓港溪 (FG)、四重溪 (SH) 以及港口溪 (GK)。



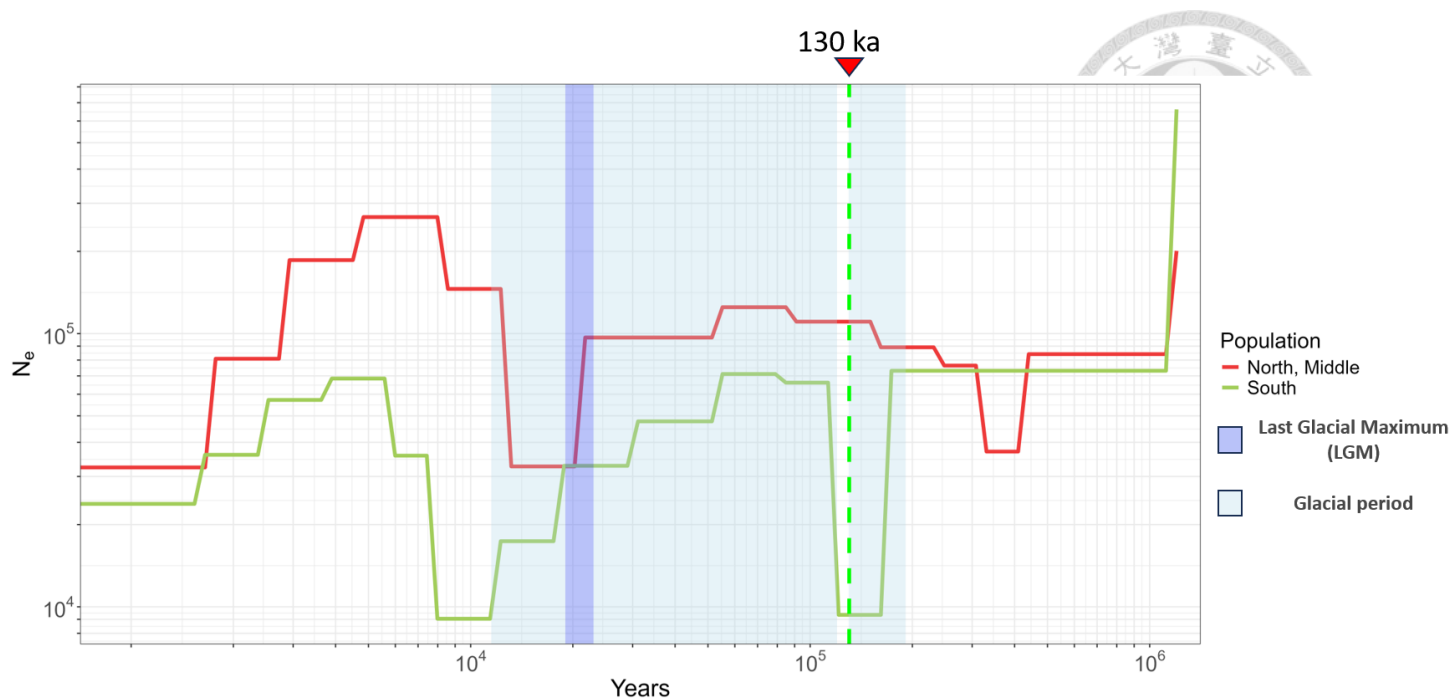
圖十四、由族群間固定的且沒有缺失的 SNP 構建的 100 kb 區塊的平均雜交比率

以士文溪的個體及中部族群作為親代，由族群間固定的且沒有缺失的 SNP 構建的 100 kb 區塊的平均雜交比率。紅色代表的是基因型來自中部族群；藍色代表基因型來自士文溪；灰色代表此處沒有固定的 SNP。BC 代表回交個體，數字代表第幾次回交；F2 為早期雜交個體。地點標示分別為瑪鍊溪(MS)、基隆河(KL)、橫溪(HX)、大漢溪(DH)、中港溪(JG)、後龍溪(HL)、大安溪(DA)、大甲溪(DJ)、烏溪(WU)、樸子溪(PZ)、高屏溪(KP)、東港溪(DG)、林邊溪(LB)、士文溪(SW)、枋山溪(FS)、楓港溪(FG)、四重溪(SH)以及港口溪(GK)。



圖十五、SNP 構建的三角散點圖

以東港溪、林邊溪與士文溪的個體及中部族群作為親代，由族群間固定的且沒有缺失的 SNP 構建的三角散點圖。BC 代表回交個體，數字代表第幾次回交；F2 為早期雜交個體。地點標示分別為瑪鍊溪（MS）、基隆河（KL）、橫溪（HX）、大漢溪（DH）、中港溪（JG）、後龍溪（HL）、大安溪（DA）、大甲溪（DJ）、烏溪（WU）、樸子溪（PZ）、高屏溪（KP）、東港溪（DG）、林邊溪（LB）、士文溪（SW）、枋山溪（FS）、楓港溪（FG）、四重溪（SH）以及港口溪（GK）。



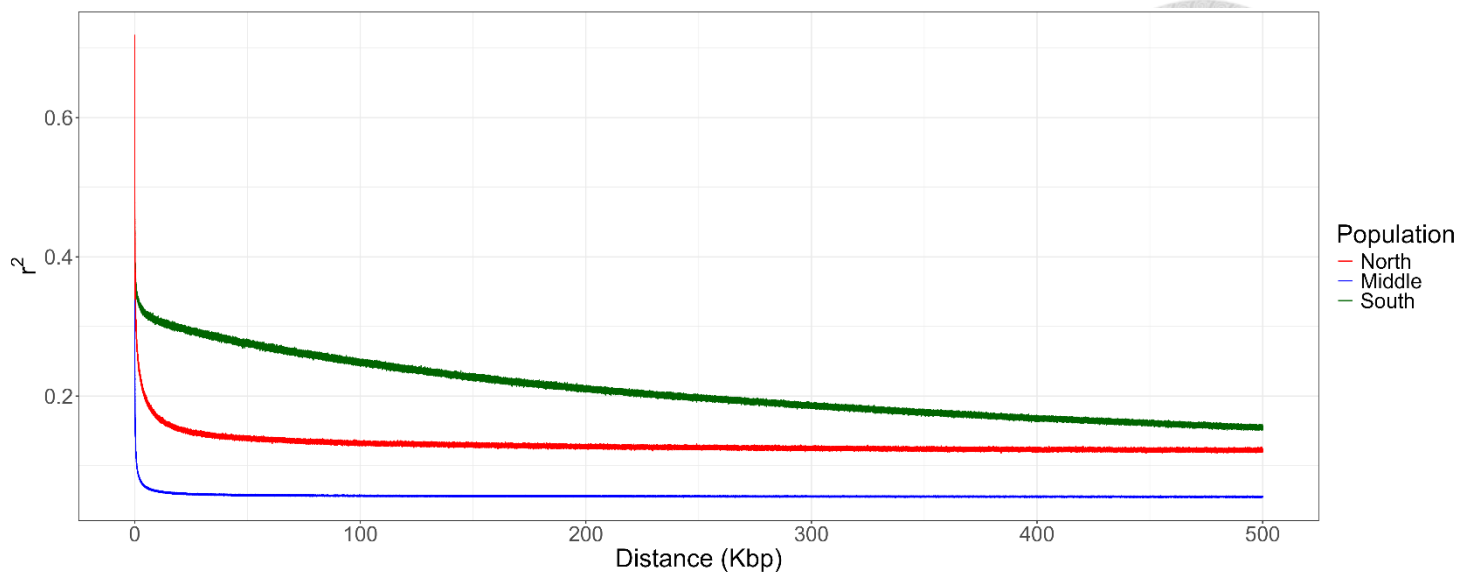
圖十六、SMC++的預測結果

設置為變異率 $5.97e-9$ 以及世代時間 2 年。淡藍色區域為冰河時期，深藍色區域為末次尖冰期 (LGM)，紅線為北部及中部族群，綠線為南部族群。綠色虛線是 SMC++ 估算出的分化時間。



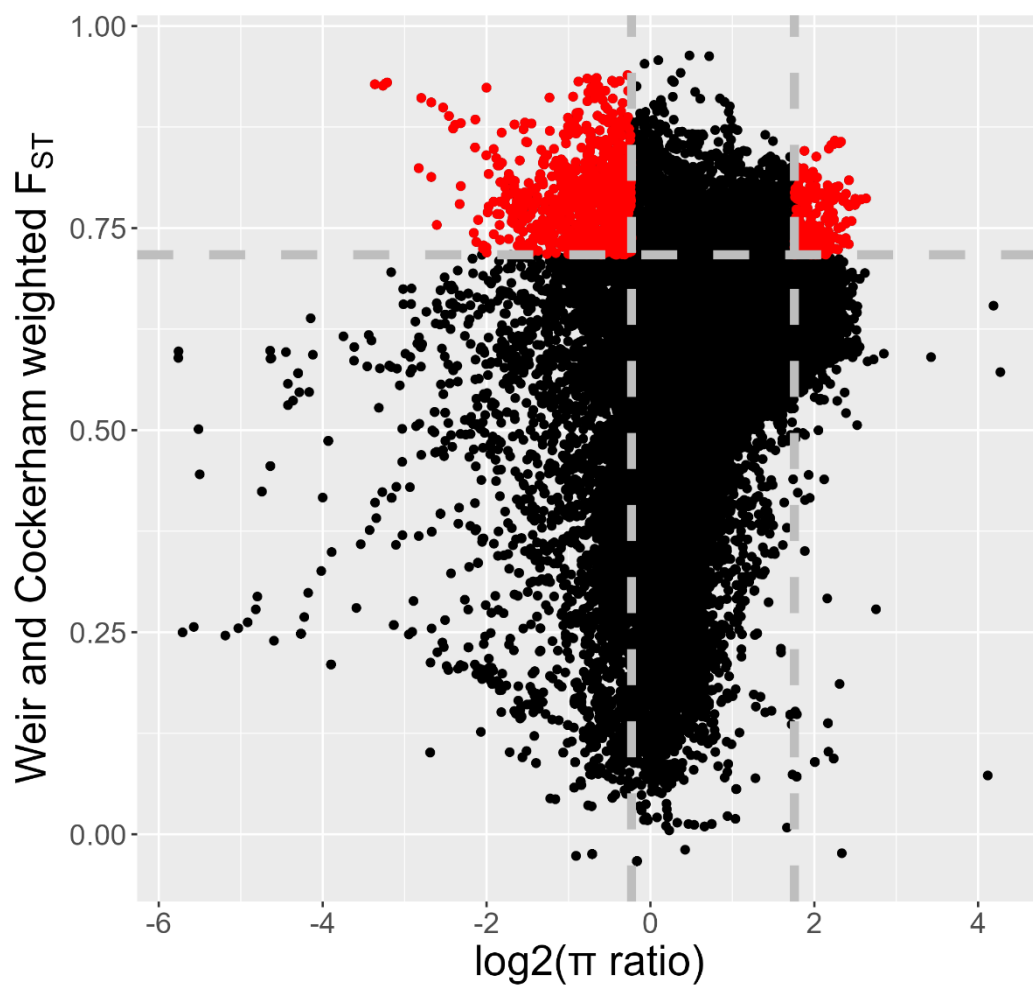
圖十七、GONE 的近年有效族群預測結果

設置圍平均重組率 1.8 cM/Mb 以及世代時間 2 年。紅線為北部及中部族群，綠線為南部族群。



圖十八、PopLDdecay 的結果

顯示出各個族群的連鎖不平衡衰減速度，其中南部族群衰減最慢，代表連鎖不平衡相較其他族群高。



圖十九、 F_{ST} 與 π ratio 篩選可能受到天擇區塊的示意圖

每一個點都為一個窗口。此處使用的是北部與中部族群和南部族群之間的 F_{ST} 與 $\log_2(\pi \text{ ratio})$ 呈現。灰色的虛線代表前後 5% 的 π ratio 以及前 5% 的 F_{ST} ，紅色的點代表可能受到天擇的區塊，右側屬於南部可能受天擇的區域，左邊則為北部與中部族群可能受天擇的區域。