

文蛤基因體之解析研究

黃慶輝、吳嘉哲
水產養殖組

文蛤 (Hard clam) 屬於軟體動物門 (Mollusca)、雙殼綱 (Bivalvia)、簾蛤目 (Veneroida)、簾蛤科 (Veneridae)、文蛤屬 (*Meretrix*)。根據漁業統計年報資料顯示，臺灣文蛤養殖產量每年約 5 萬公噸，產值每年超過新臺幣 50 億元，成為內陸養殖最重要的貝類之一，近年來本所研究持續進行相關品系之選育計畫，本計畫也希望能夠配合相關的計畫，進行文蛤基因體之解析，建立育種所需之基因背景參考資料，後續與選育至一定程度之文蛤品系的遺傳背景做定序與比較，瞭解選育效果。

為建立目前篩選文蛤之遺傳背景資料，挑選合作養殖戶場品系篩選中之健康文蛤個體，解剖取樣進行 DNA 及 RNA 萃取及品質確認，將萃取好之 DNA 製備成長片段 (PacBio) 定序用核酸庫 (library)，於 PacBio Sequel 2 上進行 SMRT 定序，第 1 次數據總產出為 161.13 Gb raw data，4.63 Gb Hi-Fi reads，本次數據觀察到大量定序片段長度較短，並且沒有達到 Hi-fi reads 的準確度，觀察 reads 的長度分布，疑似有定序提早結束的現象。第 2 次定序成功產出為 13.12 Gb Hi-Fi reads，平均讀長 14,223 也是第 1 次定序之約 2 倍 (表 1)。基因組組裝之前，根據 k-mer 分析估算了基因體大小和基因體雜合度。使用 Falcon 軟體和 PacBio 平台數據，組裝文蛤基因組片段 (contigs)，再結合 Illumina HiSeq 之定序資料，進行校正與補正後，目前資料所能組裝的最佳情況文蛤基因體的大小預估在 1.2 – 1.3 Gb 之間，目前所獲得之定序資料大約為文蛤基因體之 10X 左右，一般而言要組裝全基因組需要至少 30 – 40X，目前情況至少需要再 2 – 3 次定序。另定序完成之短片段轉錄體序列運用其他已有基因組資料之軟體動物，

以 Nr database (blast) 進行物種基因分類 (圖 1)，另運用 SwissProt 資料庫、GO、KEGG、BLASTp 等分析方式進行同源基因預測及功能性基因預測。

本研究希望能夠持續累積養殖文蛤序列資料，後續能成為未來文蛤品種培育重要的背景資料，改善臺灣文蛤養殖現況，使其更具高生產效益。

表 1 本年度研究 2 次長片段定序結果之比較

樣本	HiFiReads (#)	HiFiYield (nt)	平均長度
Xie0203_foot (1 st)	626,900	4,630,000,000	7,390
Xie0906_gill (2 nd)	935,961	13,312,315,936	14,223

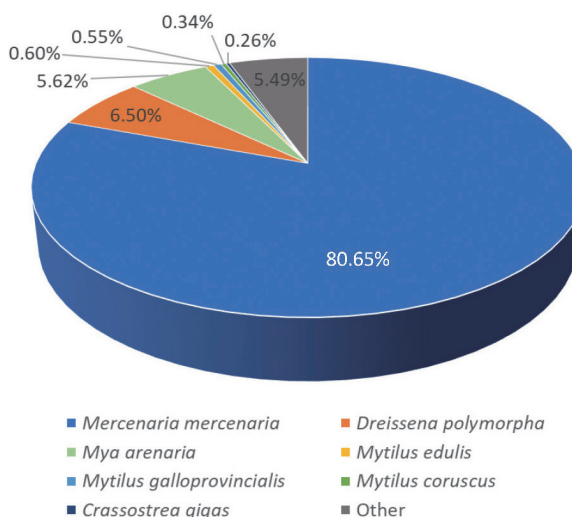


圖 1 文蛤轉錄體序列在 Nr database 下有 8 成以上 (80.65%) 之基因序列與櫻桃寶石簾蛤 (*Meretrix mercenaria*) 最為相似，其餘則為斑馬貽貝 (*Dreissena polymorpha*, 6.50%) 及大海蜆蛤 (*M. arenaria*, 5.62%) 等，顯示目前櫻桃寶石簾蛤之功能性基因資料庫資料較為完整