

# 臺灣特有種馬氏扣海膽(*Sinaechinocyamus mai*)分析計畫

李士傑 (Shyh-Jye Lee)<sup>1</sup>, 蔡孟勳 (Mong-Hsun Tsai)<sup>2</sup>, 林日白 (Jih-Pai Lin)<sup>3</sup>,  
張若愚 (Lo-Yu Chang)<sup>4</sup>

<sup>1</sup> 國立臺灣大學生命科學系

<sup>2</sup> 國立臺灣大學生物科技研究所

<sup>3</sup> 國立臺灣大學地質科學系

<sup>4</sup> 國立臺灣大學醫學系

## 一、摘要(Abstract)

關鍵字：楯形目、馬氏扣海膽、微型化、支序基因體學、同源異型基因

楯形目(Clypeasteroida，又稱沙錢)之化石紀錄始於晚古新世，為現生棘皮動物門中相當年輕之成員，現今楯形目生物以拓殖至全球熱帶至溫帶海域。與正形海膽不同，楯形目物種具棘皮動物門中少見之次生性兩側對稱，為其一大特徵。目前楯形目之演化位階與基因體結構了解甚少，歸因於缺乏完整之基因體定序資料。馬氏扣海膽(*Sinaechinocyamus mai*)為棲息於臺灣西部海岸之楯形目物種，可見化石紀錄始於更新世頭嵙山層，然臺灣周遭海域之拓殖史尚不明朗。本研究希冀透過分析馬氏扣海膽之粒線體基因組，確認其演化關係與基因體結構。同時進一步以核基因體作為楯形目模式物種材料，以同源異形基因分析次生性兩側對稱之可能起源，並探討馬氏扣海膽微型化之分子機制。

With fossil record appearing in Late Palaeocene, Clypeasteroida (sand dollar) is one of the youngest class in living Echinodermata. Clypeasteroida has colonized marine benthos niche from tropical to temperate latitude around the world. The unique body plan of secondary bilateral symmetry differentiates Clypeasteroida from pentamerism body plan exhibited among regular echinoids. However, the evolution history and genomic feature of Clypeasteroida remains in dispute, largely due to the lack of published complete genome. *Sinaechinocyamus mai* (Wang, 1984) is a clypeasteroid echinoid whose tests and living individuals can be found in western coast of Taiwan. Its fossil record in Taiwan can be traced back to Toukoshan Formation (Pleistocene); while the colonization history remains elusive. This study aims to understand the phylogenetic position and mitochondrial genome structure of *S. mai*; furthermore, decipher of the first Clypeasteroida genome model *S. mai* provides insight into genome biology of Clypeasteroida as a leverage to compare the

*Hox* gene cluster with regular echinoids for secondary bilateral symmetry and the possible molecular mechanism behind the miniaturization exhibited in *S. mai*.

Keyword: Clypeasteroida; *Sinaechinocyamus mai*; miniaturization; phylogenomics; *Hox* gene

## 二、研究動機與研究問題(Specific Aims)

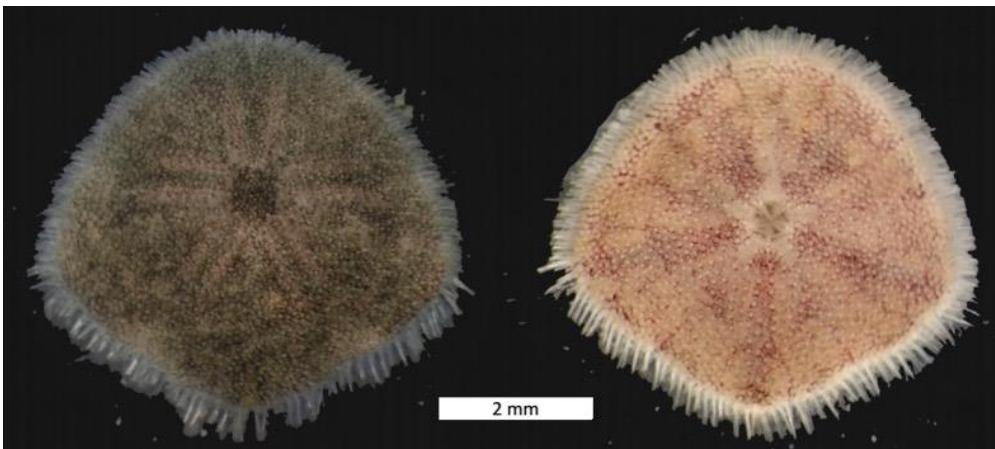
1. 以次世代定序(next generation sequencing)進行馬氏扣海膽(*Sinaechinocyamus mai*)之粒線體基因體組裝，並比較獨立樣本得到公同(consensus)序列。
2. 以馬氏扣海膽之粒線體基因體進行分子定年與正向演化(positive selection)分析。
3. 以已知生物體微型化(miniaturization)途徑為參考，探討馬氏扣海膽微型化之可能機制。
4. 對核基因體(nuclear genome)進行初步分析，以同源異型基因(Hox gene)探討楯形目之兩側對稱特徵起源。

## 三、文獻回顧與探討(Background and Previous Researches)

### 1. 馬氏扣海膽(*Sinaechinocyamus mai*)之早期研究與化石紀錄

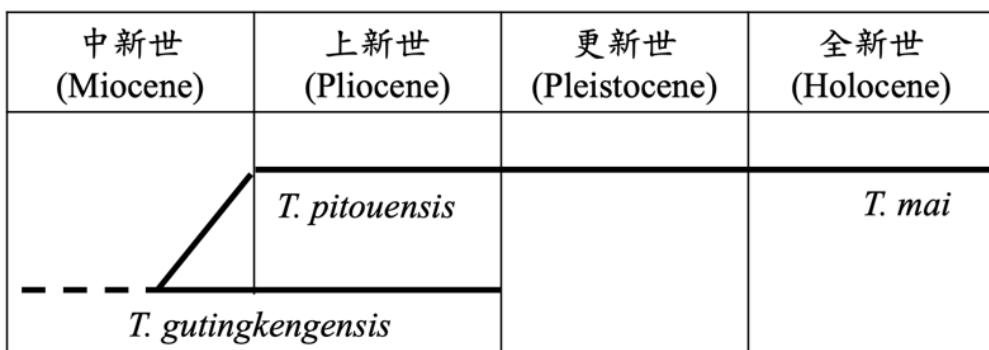
棘皮動物門(Echinodermata)與脊索動物門(Chordata)同屬後口動物(Deuterostomia)，現生物種約 7,000 種，為相當多樣之海洋無脊椎動物類群。其中海膽綱(Echinoidea)現生種約 900 種，其幼體具有兩側對稱性(bilateral symmetry)，但於變態過程中體制重新變化，大多數海膽成體為五輻對稱(pentamerism)，稱為正形海膽(regular echinoid)，而成體具有次生性兩側對稱體制者稱為歪形海膽(irregular echinoid)，體型扁平之楯形目(Clypeasteroida，又稱沙錢)與蝶團目(Spatangoida，又稱心海膽)為兩大分類群。而歪形海膽殼體趨向兩側對稱具有前後主軸與特化的帶線(fasciole)可能有助於其潛沙移動覓食。除現生物種之重要性外，海膽於古生物研究中亦扮演重要角色。目前已知最早之海膽化石紀錄始於奧陶紀(Ordovician)，相對其餘現生棘皮動物出現之地質時代較晚。現生正形海膽則始於二疊紀(Permian)，而歪形海膽則起源於侏羅紀(Jurassic)早期之正形海膽，並於新生代(Cenozoic)興盛(1, 2)。

馬氏扣海膽(*Sinaechinocyamus mai*, Wang 1984)屬楯形目樹星海膽科(Dendrasteridae)，最早由王家慶描述，其種小名紀念地質學者馬廷英。原發表文章中命名為 *Taiwanaster mai*(3)，然 Mooi (1990)(4)重新檢視其特徵，發現 *Taiwanaster* 與廖玉麟所描述之 *Sinaechinocyamus* (Liao, 1979)相同，是故 *Taiwanaster* 為次異名(junior synonym)不具有效地位，而 *Taiwanaster mai* 即改為 *Sinaechinocyamus mai*。馬氏扣海膽主要分佈於臺灣西海岸新竹至曾文溪口之潮間帶與淺海，如圖一所見，其直徑多於 1 公分以下，活體呈灰綠色而體殼白色，與多數沙錢相同以有機碎片為食並潛沙活動。



圖一、活體之馬氏扣海膽(*Sinaechinocamus mai*)

原描述者王家慶除命名馬氏扣海膽外，另發表二化石種 *T. gutingkengensis* 與 *T. pitouensis*，並藉由口面之間步帶(interambulacrum)型態推論三物種中 *T. gutingkengensis* 為基群，其分枝演化(cladogenesis)形成新種 *T. pitouensis*，而 *T. pitouensis* 再線性演化(anagenesis)形成現生種 *T. mai* (即現稱之 *S. mai*)，其三者間具有如下圖二之支序關係。



圖二、參考 Wang (1984)(3)繪製之 *Taiwanaster* 屬內三物種親緣支序關係。

然而線性演化與分支演化之種化理論近年來充滿爭議，其最大爭議點在於難以定義化石種與延續之現生種是否具有足夠差異以區分為兩物種，且區辨特徵是否具有意義亦隨研究者不同具有高度主觀性質(5)。若不論三化石種之間可能之種化模式，無庸置疑之點為馬氏扣海膽與其相關化石種於臺灣地層分佈相當連續且完整，西部麓山帶之晚中新世桂竹林層(c. 6 Mya)到更新世頭崙山層均可見，且數量相當豐富(3, 6)。

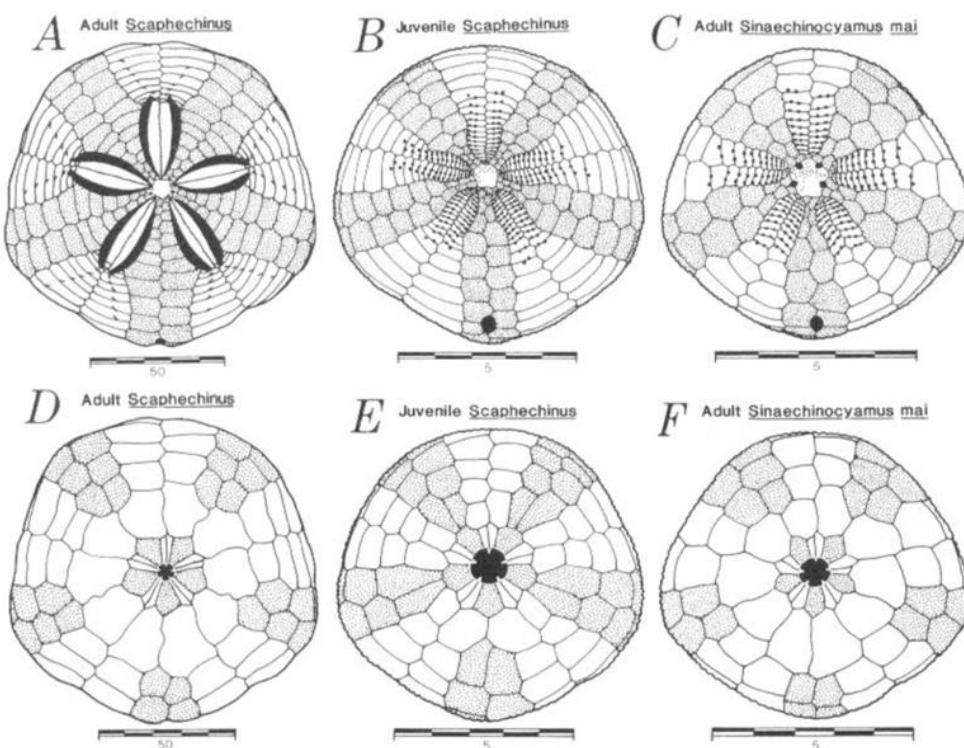
奇異掘海錢(*Scaphechinus mirabilis*)為馬氏扣海膽之現生姊妹群(sister group)，其現生種分佈於日本、韓國及中國沿海，臺灣僅於苗栗部分地區可見化石紀錄，且分佈時間僅限於更新世頭崙山層(c. 1.6 Mya)之後。奇異掘海錢與馬氏扣海膽兩者間之演化與生態地理分佈尚未有系統性研究討論之，本研究希望能透過分子演化之相關技術提供兩物種分化時間之相關資訊，並分析可能使馬氏扣海膽得以維持臺灣現生生態棲位之機制。

## 2. 馬氏扣海膽之微型化探討

馬氏扣海膽最引人注目之特徵為其微型化(miniaturization)之體型，相較近緣現生姊妹群奇異掘海錢成體直徑約 10 公分，馬氏扣海膽之成體直徑極少超過 1 公分，幾乎與奇異掘海錢之幼

體相同。個體微型化廣泛見於動物界，如哺乳類著名之島嶼侏儒化(island dwarfism)(7)，而於無脊椎動物各類群中亦不乏案例，如腕足動物(brachiopods)、軟體動物(molluscs)與棘皮動物等(8)，其係指某物種相對其與姊妹群之共同祖先(common ancestor)體型顯著縮小之趨勢，而微型化不單為體型縮小，亦有其他相關之生理、行為、生殖方式等適應。於諸多微型化案例中，Urbanek (1993)於化石紀錄所觀察之小人國效應(Lilliput effect)提供一重要演化脈絡，其顯示災變過後地層中化石紀錄之物種通常較災變前之近緣祖先種小(9)。而後續研究所提出之演化解釋指出生態環境於災變後改變，物種受有限資源篩選，較小之體型可減少需求，有利於該物種生存於資源匱乏之環境(10)。

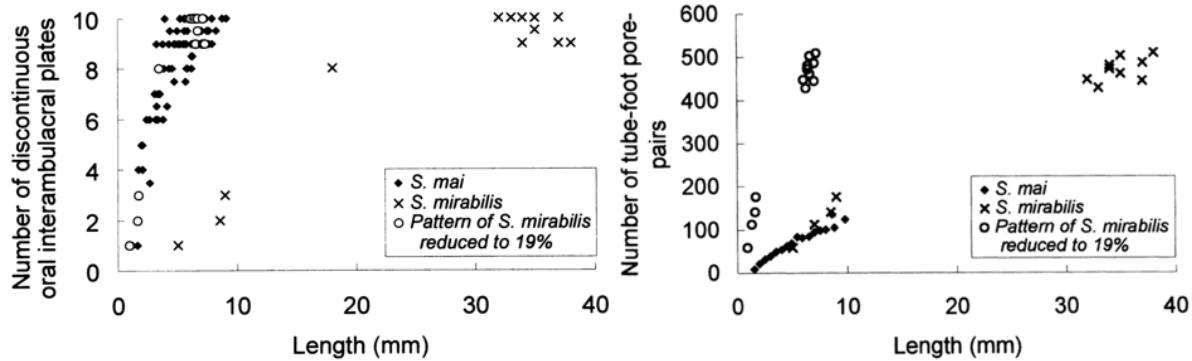
Mooi (1990)即以型態分析對馬氏扣海膽與奇異掘海錢進行比較(4)，發現除大小外，解剖形態上馬氏扣海膽成體亦相當類似奇異掘海錢之幼體，故其推測馬氏扣海膽可能表現出幼形遺留(paedomorphosis)之特徵。如下圖三所示：



圖三、奇異掘海錢(*Scaphechinus mirabilis*)成體與幼體及馬氏扣海膽(*Sinaechinocyamus mai*)成體之比較圖，A-C：反口面(aboral surface)；D-F：口面(oral surface)，陰影處為間步帶(interambulacrum)。引自 Mooi (1990)(4)。

幼形遺留之生理機制可分為兩大極端：生長速率不變，但成長期提早結束之初期形成(progenesis)與生長速率下降，但成長期長度維持不變之幼態成熟(neoteny)(8)，多數已知之幼形遺留物種均有兩者之混合特性，為了解馬氏扣海膽幼形遺留之生理機制為何者，Chen and Chao (1997)進行個體培養觀察實驗如下圖四。其發現馬氏扣海膽之成長速率約為奇異掘海錢之19%，但生長期與性成熟期並未出現顯著提早之現象，故馬氏扣海膽之幼形遺留機制應主要為幼態成熟(neoteny)，但因缺乏分子生物研究，其可能之訊息傳遞途徑與分子機制尚不明

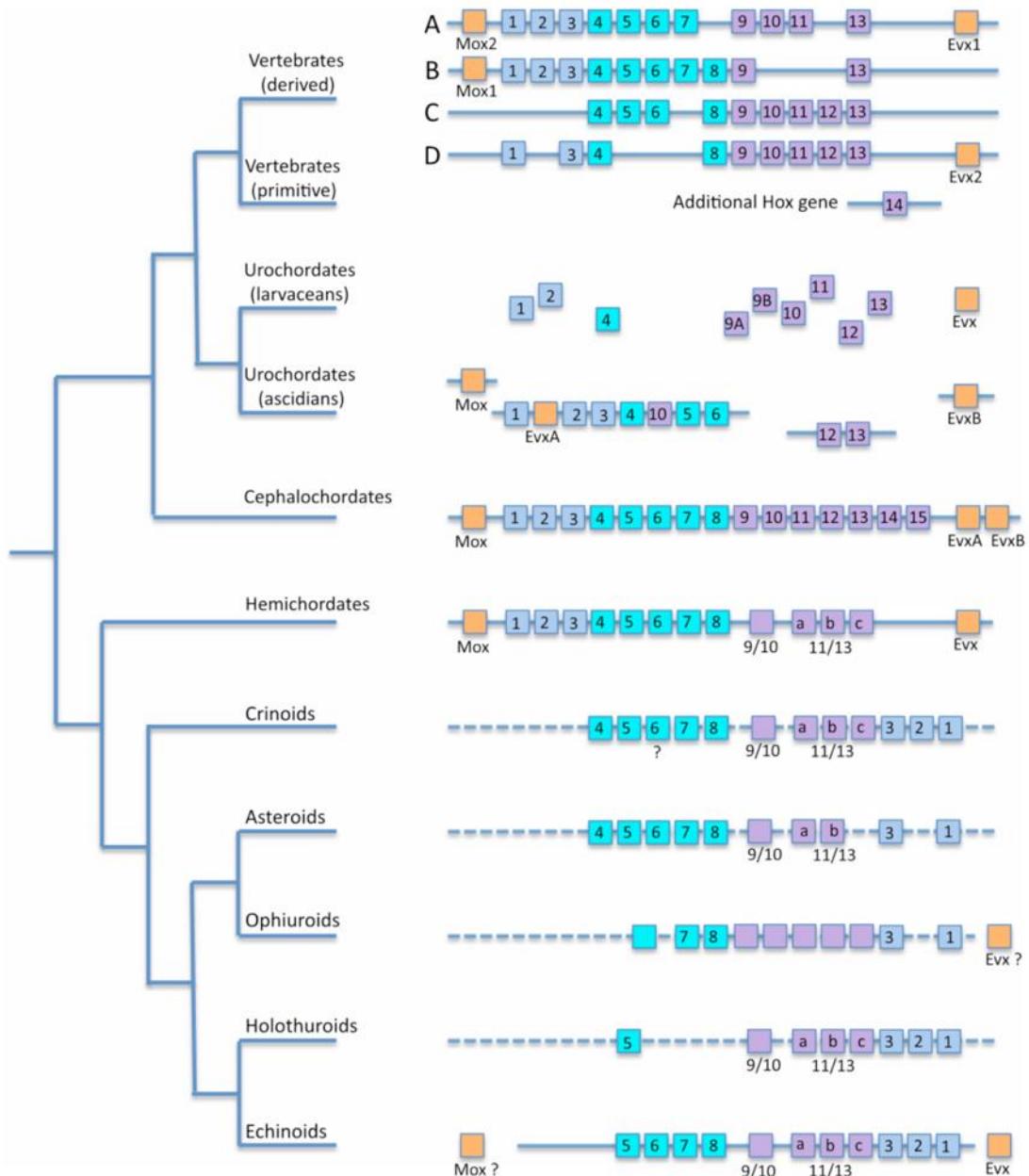
朗。近年已知部分與海膽變態(metamorphosis)及成長之關鍵調控重要基因(*Opsin4*, *Opsin5*, *Pax-Six-Eya-Dach Network* 等)(11, 12)，本研究希冀能透過基因體分析比較近緣物種，於尋找突變與正向演化(positive selection)之可能證據。



圖四、馬氏扣海膽(*S. mai*)與奇異掘海錢(*S. mirabilis*)之生長比較圖，左圖為口板(oral plate)之成長，右圖為反口板(aboral plate)之成長。引自 Chen and Chao (1997)(13)

### 3. 沙錢之體軸建立與同源異型基因(Hox gene)

同源異型基因(Hox gene)為調控生物體發育、體制(body plan)與構築兩側對稱生物前後體軸(anterior-posterior axis, A/P axis)之重要基因群，其具有特殊之共線性(collinearity)，亦即同源異型基因於染色體上之排列與體軸之排列，位於3'端之基因为發育最早期表現，亦表現於個體最前端之前端組(anterior class)；其次為居中之中端組(medial class)；最後為位於5'端最晚表現，亦表現於個體最後端之後端組(posterior class)(14, 15)。David et al. (2014)統整目前已知之後口動物(Deuterostomia)各類群之同源異型基因群結構如下圖四，可發現相較於兩側對稱類群如脊椎動物(vertebrates)、尾索動物(urochordates)、頭索動物(cephalochordates)，棘皮動物各類群同源異型基因群之共線性已喪失，且在海膽綱之模式物種紫海膽(*Strongylocentrotus purpuratus*)可發現屬於中端組之 *Hox 4* 基因消失，而失序之同源異型基因群結構可能與棘皮動物之次生性五輻對稱相關(15)。至今並無沙錢或其他歪形海膽之公開基因體序列，亦缺乏歪形海膽前後體軸次生建立之相關研究，本研究希望能透過初步分析核基因，以馬氏扣海膽為模式找出沙錢之同源異型基因群結構，探討次生性兩側對稱之起源與可能機制。



圖五、主要後口動物(Deuterostomia)各類群之同源異型基因群型態示意圖。方框表示基因，深藍色表示前端組(anterior class)；亮藍色表示中端組(medial class)；紫色表示後端組(posterior class)。引用自 David et al. (2014)(15)

#### 四、研究方法及步驟(Materials and Methods)

##### 1. 物種定序

馬氏扣海膽之活體樣本採樣自臺灣苗栗海灘之自然棲地( $120^{\circ}39' E$ ;  $24^{\circ}29' N$ )。全基因體 DNA 由 RecoverEase DNA isolation kit (Agilent Technologies) 萃取，萃取之 DNA 由 BluePippin (Sage Science) 篩選長度大於 40 kb 之片段並以 Life Technologies Qubit and Agilent Technologies Tapestation 4200 進行定量。基因庫(library)由 Chromium gel bead and library kit (10X Genomics) 以篩選過後總量 0.625 ng DNA 建立。建構之基因庫由 Illumina Novaseq 6000 sequencer 進行

151 bp 雙邊定序(paired-end reads)。

## 2. 基因體組裝(assembly)

定序所得之雙邊序列 FASTQ 檔案以 FastQC v. 0.11.8(16)進行品管篩檢(quality control)，並以 KmerGenie v. 1.7(17)用於判斷組裝時所需之 Kmer 長度。由 Minia 3 以 KmerGenie v. 1.7 決定之參數進行從頭基因體組裝(*de novo* assembly)，初步結果以紫海膽(*Strongylocentrotus purpuratus*)為參考基因體，由 Bowtie 2(18)進行回對分析(re-mapping)。

## 3. 粒線體基因體功能標註(annotation)

由 Bowtie 2(18)進行回對分析(re-mapping)並確認後之片段重疊群(contig)由 MITOS(19)與 GeSeq(20)進行自動粒線體基因體功能標註(annotation)，並以 blastX(21)手動確認。

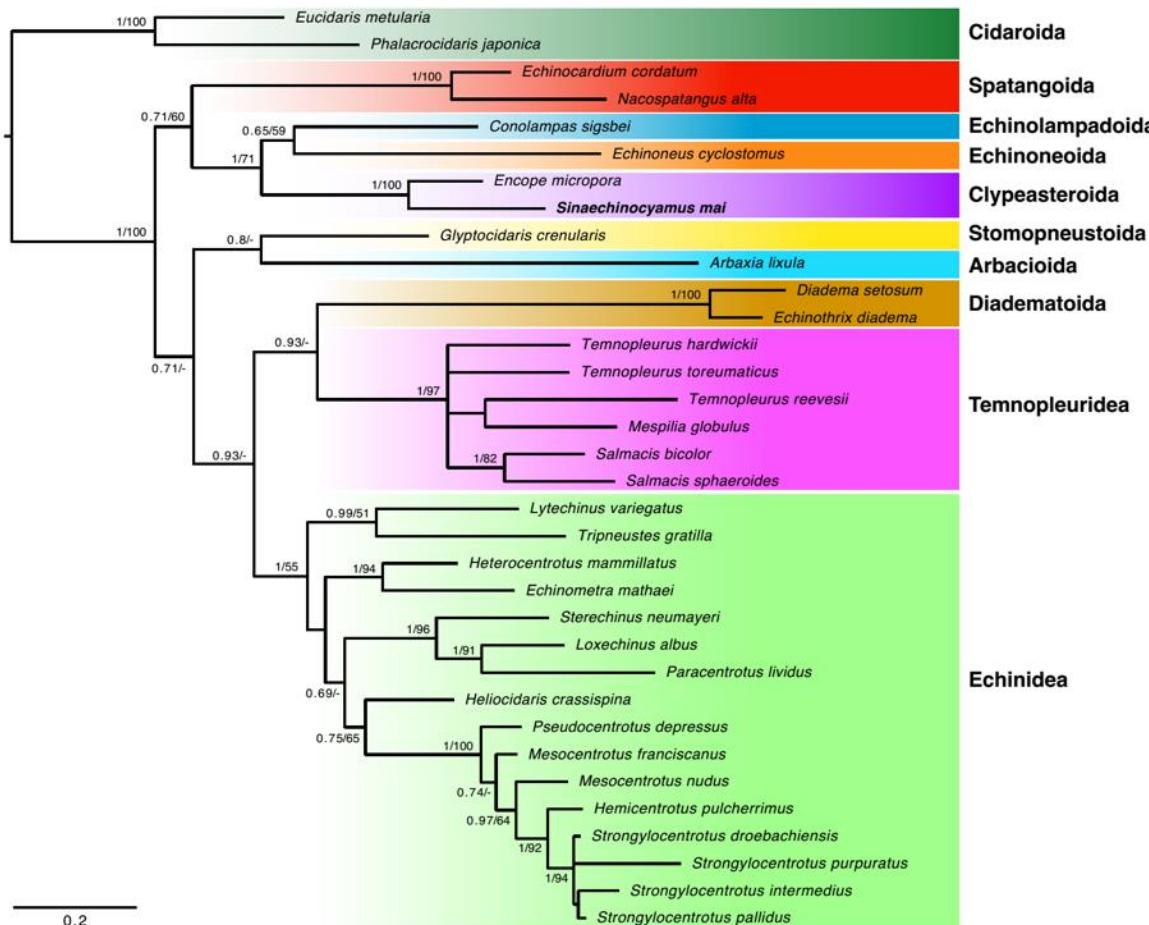
## 4. 演化分析

同源基因群以 OrthoFinder(22)進行分析，確認之同源基因以 MUSCLE(23)進行對齊(alignment)，並以 ModelGenerator(24)以赤池資訊準則(Akaike information criterion, AIC)與貝氏資訊準則(Bayesian information criterion, BIC)分析最適核苷酸與氨基酸替代模型(substitution model)。支序分析由 RAxML v. 8.2.12(25)以 ModelGenerator(24)最適模型進行最大概似法(maximum likelihood)樹形重建並重複取樣 1000 次求得拔靴值(bootstrap value)，並以 MrBayes v. 3.2.7(26)進行貝葉氏分析(Bayesian analyses)，以後驗機率(posterior probability)作為信心指標。正向演化分析以 PAML 4(27)進行。

## 五、預期結果(Expected Results)

### 1. 演化支序關係

以初步資料中之粒線體基因分析，由細胞色素 c 氧化酶次單元 1(cytochrome c oxidase subunit I, COI)與 16S rRNA 序列可得到演化支序樹形如下圖六。可發現以馬氏扣海膽為代表之楯形目(Clypeasteroida)最近姊妹群為燈海膽目(Echinolampadoida)與斜海膽目(Echinoneoida)，此結果與 Koch et al. (2018)(28)一致，指出新領總目(Neognathostomata)應為有效之單系群(monophyletic group)。未來獲得更完整之基因體組裝序列後可由 OrthoFinder 比較參考基因體(如紫海膽(*Strongylocentrotus purpuratus*))初步搜尋同源基因群，以更完善之資訊進行支序重建，但有鑑於目前樹形之節點支持度(supporting value)均高，樹形應大致同圖六之初步結果。



圖六、以細胞色素 c 氧化酶次單元 1 與 16S rRNA 序列所建構之演化關係樹。節點示貝葉氏分析(Bayesian analyses)後驗機率(posterior probability)與最大概似法(maximum likelihood)拔靴值(bootstrap value)。

## 2. 分子時鐘定年分析

以細胞色素 c 氧化酶次單元 1 所建構之樹型與其他蛋白編碼基因(protein-coding gene)可用於分子定年，從而分析馬氏扣海膽與奇異掘海錢及其餘親緣物種之分化年代，推測馬氏扣海膽與奇異掘海錢之分化年代應早於晚中新世(*c.* 6 Mya)。

## 3. 馬氏扣海膽之微型化探討

以 OrthoFinder 初步功能標註之馬氏扣海膽核基因體為材料，搜尋已知與棘皮動物成長及變態之關鍵調控重要基因(*Opsin4*, *Opsin5*, *Pax-Six-Eya-Dach Network* 等)，推測於編碼區域(coding region)或調控元件區(regulatory element)具有突變，若突變發生於編碼區域可進一步測試正向演化，以推測是否為面對演化壓力之適應。

## 4. 體軸建立與同源異型基因

以 OrthoFinder 初步功能標註之馬氏扣海膽核基因體為材料，搜尋同源異型基因群，推測馬氏扣海膽之同源異型基因群發生重新排列(rearrangement)以重建共線性與前後體軸特徵，並分析編碼區域與調控元件區是否具突變，發生於編碼區域可進一步測試正向演化以分析其演化適應。

## 六、參考文獻(References)

1. Jain S. Echinoids. Fundamentals of Invertebrate Palaeontology. New Delhi: Springer; 2017.
2. 林日白. 細說海膽化石. 地質. 2018;37(4):22-7.
3. Wang CC. New Classification Of Clypeasteroid Echinoids. Proceedings of the Geological Society of China. 1984;27:119-52.
4. Mooi R. Progenetic miniaturization in the sand dollar *Sinachinocymus*: implications for clypeastroid phylogeny. In: Jangoux M, editor. Echinoderm Research. Rotterdam: CRC Press; 1990. p. 137-43.
5. Vaux F, Trewick SA, Morgan-Richards M. Lineages, splits and divergence challenge whether the terms anagenesis and cladogenesis are necessary. Biological Journal of the Linnean Society. 2016;117(2):165-76.
6. 朱偉嘉. 臺灣的海膽化石. 地質. 2018;37(4):28-31.
7. Raia P, Meiri S. The island rule in large mammals: paleontology meets ecology. Evolution. 2006;60(8):1731-42.
8. Hanken J, Wake DB. Miniaturization of Body-Size - Organismal Consequences and Evolutionary Significance. Annual Review of Ecology and Systematics. 1993;24(1):501-19.
9. Urbanek A. Biotic crises in the history of Upper Silurian graptoloids: A Palaeobiological model. Historical Biology. 1993;7(1):29-50.
10. He WH, Shi GR, Xiao YF, Zhang KX, Yang TL, Wu HT, et al. Body-size changes of latest Permian brachiopods in varied palaeogeographic settings in South China and implications for controls on animal miniaturization in a highly stressed marine ecosystem. Palaeogeography Palaeoclimatology Palaeoecology. 2017;486:33-45.
11. Armstrong AF, Grosberg RK. The developmental transcriptomes of two sea biscuit species with differing larval types. BMC Genomics. 2018;19(1):368.
12. Byrne M, Koop D, Morris VB, Chui J, Wray GA, Cisternas P. Expression of genes and proteins of the pax-six-eya-dach network in the metamorphic sea urchin: Insights into development of the enigmatic echinoderm body plan and sensory structures. Dev Dyn. 2018;247(1):239-49.
13. Chen CP, Chao CM. Reduction of Growth Rate as the Major Process in the Miniaturization of the Sand Dollar *Sinaechinocymus mai*. Biol Bull. 1997;193(1):90-6.
14. Lacalli T. Echinoderm conundrums: Hox genes, heterochrony, and an excess of mouths. Evodevo. 2014;5(1):46.
15. David B, Mooi R. How Hox genes can shed light on the place of echinoderms among the deuterostomes. Evodevo. 2014;5:22.
16. Andrews S. FastQC: a quality control tool for high throughput sequence data 2010 [Available from: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>.

17. Chikhi R, Medvedev P. Informed and automated k-mer size selection for genome assembly. *Bioinformatics*. 2014;30(1):31-7.
18. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012;9(4):357-9.
19. Bernt M, Donath A, Juhling F, Externbrink F, Florentz C, Fritzsch G, et al. MITOS: improved de novo metazoan mitochondrial genome annotation. *Mol Phylogenet Evol*. 2013;69(2):313-9.
20. Tillich M, Lehwark P, Pellizzer T, Ulbricht-Jones ES, Fischer A, Bock R, et al. GeSeq - versatile and accurate annotation of organelle genomes. *Nucleic Acids Res*. 2017;45(W1):W6-W11.
21. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. *BMC Bioinformatics*. 2009;10:421.
22. Emms DM, Kelly S. OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol*. 2015;16:157.
23. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res*. 2004;32(5):1792-7.
24. Keane TM, Creevey CJ, Pentony MM, Naughton TJ, McLnerney JO. Assessment of methods for amino acid matrix selection and their use on empirical data shows that ad hoc assumptions for choice of matrix are not justified. *BMC Evol Biol*. 2006;6:29.
25. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*. 2014;30(9):1312-3.
26. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Hohna S, et al. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol*. 2012;61(3):539-42.
27. Yang Z. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol*. 2007;24(8):1586-91.
28. Mongiardino Koch N, Coppard SE, Lessios HA, Briggs DEG, Mooi R, Rouse GW. A phylogenomic resolution of the sea urchin tree of life. *BMC Evol Biol*. 2018;18(1):189.