

蕎麥葉片及種子轉錄體分析¹

陳裕星²、王捷³、陳鏗斌²、廖宜倫²

摘 要

蕎麥(*Fagopyrum esculentum* Moench)屬於蓼科蕎麥屬一年草生作物，蕎麥種子與植株中含有多種有益健康之機能性成份，同時為優良蜜源植物，因此成為世界重要雜糧作物之一。本研究分析蕎麥種子及葉片轉錄體，分別讀取143 M及247 M條序列，共組裝97,200及74,544條contigs，平均覆蓋率(coverage)分別約為1,474及3,340倍，contig序列平均長度分別為565及607 bp，全部拼接完成之序列分別為54 M及45 M bp。葉片及種子序列混合拼接時，共可獲得106,190條contig，其中屬於種子和葉片者分別為69,043及56,345條，單獨存於種子或葉片者分別為22,587及9,889條。所有組裝成功之contigs進一步與資料庫比對，註解並進行基因功能分類檢索，分別依照細胞組成(cellular component)、生理程序(biological process)及分子功能進行分類。本轉錄體資訊亦與KEGG (Kyoto Encyclopedia of Genes and Genomes)資料庫比對，完成306個代謝途徑之比對，並針對類黃酮、黃酮與黃酮醇及芸香苷代謝途徑基因加以檢視，本研究同時檢索與種子蛋白質合成相關基因，包括種子貯存蛋白、LEA protein, cupin 家族蛋白及過敏蛋白等基因，這些資訊可提供未來蕎麥栽培育種工作之參考。

關鍵字：芸香苷、過敏蛋白、蕎麥核心抗原、蕎麥糖醇

前 言

蕎麥為蓼科蕎麥屬一年生草本植物，屬內有許多種，最常見的栽培種有二種，一為普通種蕎麥(common buckwheat) *Fagopyrum esculentum* Moench，另一為韃靼種蕎麥(Tatary buckwheat) *Fagopyrum tataricum* (L.) Gaertn。蕎麥原為臺灣中部地區水田秋冬裡作作物，主要供做綠肥用^(1,2)，在臺灣主要之栽培地區以彰化二林、大城等地為主，面積約40公頃，於二期作水稻收穫後秋裏作種植，臺中區農業改良場目前已育成「臺中1號」、「臺中2號」、「臺中3號」及「臺中5號」等品種，並推廣給農民栽培，其中臺中2號為韃靼種蕎麥，臺中5號為普通種蕎麥^(1,2,3)，為目前主要栽培品種。普通種蕎麥同時為優良之蜜源植物，盛花期可長達2~3個月，每公頃並可產蜜達70~100公斤⁽⁴⁾，在農村可搭配蕎麥粉手工製麵，作為副產業與農村休閒旅遊之用。

¹行政院農業委員會臺中區農業改良場研究報告第 0839 號。

²行政院農業委員會臺中區農業改良場助理研究員。

³中研院生物多樣性中心研究助理。

本研究針對蕎麥種子及葉片轉錄體，進行基因功能分類(gene ontology)與註解(annotation)，對比葉片及種子基因表現的輪廓，並探討同一物種葉片與種子基因表現差異之情形，藉此了解對產量及品質相關重要基因的表現情形，包括澱粉合成酶、種子貯存蛋白質、過敏蛋白質的種類及表現情形。

前人研究

(一)蕎麥的營養成分

蕎麥常見的栽培種可分為普通種及韃靼種，韃靼種之種實具苦味，故又稱苦蕎；相較於苦蕎，普通種蕎麥則被稱為甜蕎，為大部份國家所栽培的種類；兩者之營養成分組成含量具差異性，例如韃靼蕎麥種子含有豐富芸香苷，普通種蕎麥則較含量低。蕎麥含約60%碳水化合物(醣)、11%蛋白質、2%脂肪、10%纖維、鎂、鈣、鉀、磷、鐵、少量維生素B1、維生素B2、菸鹼酸等⁽⁵⁾。早在唐代千金、食治等書中已有蕎麥之藥用記載，本草綱目有云：蕎麥性屬“甘味降氣、寬腸沉積、泄痢帶濁”，為漢藥材之一。

蕎麥含高水溶性蛋白質(10~13.1%)且不含麩質(gluten free)，可開發無麩質產品，也適合減肥者食用，且含各種必需胺基酸，包括必須由膳食補充的亮胺酸(leucine)，異亮胺酸(isoleucine)、離胺酸(lysine)、蛋胺酸(methionine)、苯丙胺酸(phenylalanine)、蘇胺酸(threonine)、色胺酸(tryptophen)和纈胺酸(valine)，其離胺酸(lysine)含量為所有穀類中最高者⁽⁵⁾。

蕎麥含2.2~2.7%脂肪，大多為油酸、亞油酸、棕櫚酸及亞麻油酸，其碳水化合物是所有穀類澱粉中最容易糖化者，易被人體消化吸收。其種子所含的可溶性纖維質高於其他禾穀類作物，可促進腸胃蠕動及消化，對於消除腸胃內積滯之食物有助益，適合腸胃患者食用。蕎麥含1.1~2.1%灰分，其中含量最多者為磷酸鹽類，其他如鉀、鎂、鈣、鐵之含量均高於其他禾穀類作物^(6,7)。蕎麥之葉綠素、維生素B1、B2、B6和E的含量亦顯著高於其他作物^(5,6,8,9)，並含有芸香苷(rutin)、膽鹼素、泛酸、菸鹼酸、水楊酸(salicylic acid)，4-羥基苯甲胺(4-hydroxybenzylamine)，N-水楊酸又替水楊酸(N-salicylidene-salicylamine)等多種有益人體健康成分。

(二)蕎麥的機能性成分與過敏性蛋白質

蕎麥含有非常重要的黃酮類代謝物：芸香苷(rutin)和槲皮素(querctin)。芸香苷又稱維他命P，具有防治毛細血管脆弱性出血症，對血管具有擴張及強化作用，可防止動脈硬化及高血壓，目前醫學上已確定芸香素對視網膜出血(retinal hemorrhage)、毛細管性中風(capillary apoplexy)、冠狀動脈阻塞(coronary occlusion)等疾病有顯著的療效^(9,10)。糖尿病用藥阿卡波糖(acarbose)可以抑制 β -glucosidase活性，使患者餐後血糖不劇烈波動^(11,12,13)，而蕎麥所含之芸香苷與槲皮素具有中等到高度的 β -glucosidase抑制活性，分別為阿卡波糖的0.5倍及3倍⁽¹⁴⁾，也有機會發展為輔助調節血糖產品。近年來也發現黃酮類代謝物可以作用在AP-1, NF- κ B,

C/EBP δ , MMP-2/-9, VEGF, caspases, Bcl-2, Bax等基因，而具有抗氧化、抑制腫瘤組織血管新生及促進腫瘤細胞凋亡的效果^(15,16)。

蕎麥中含有對第二型糖尿病患有功的保健成分手性肌醇D-chiral-inositol (DCI)。蕎麥的DCI與galactopyranose結合形成蕎麥糖醇(fagopyritols)，主要存在於麩皮^(17,18)。在大鼠、猴子和人體試驗中，DCI能迅速降低高血糖症(Hyperglycemia)，因為在血糖調控機制中，DCI為胰島素訊息傳導(insulin's steroidogenic actions)之媒介體D-Chiral-Inositolphosphoglycans (IPGs)的成份，在第二型糖尿病患者的組織細胞中，其細胞DCI濃度低於正常細胞，在缺乏足夠濃度受體形下使胰臟 β 細胞需分泌更多的胰島素以補償^(19,20)。有病患族群其肌肉、肝臟細胞中，胰島素訊息作用傳導之媒介體D-Chiro-Inositolphosphoglycans (IPGs)濃度較低，遠低於正常水準，因此在膳食中補充DCI此IPG受體前驅物，可增加病患對胰島素敏感性，若一連八星期每天口服一千二百毫克DCI，可幫助一些會抗拒胰島素作用的糖尿病人，令其血糖和三酸甘油酯下降^(19,20)。

此外，多囊性卵巢症候群(PCOS, Polycystic ovarian syndrome)乃生育年齡婦女最常見的內分泌疾病，據估計約有1/10生育年齡婦女有此疾病，典型的症狀包括無月經、不孕、多毛症、青春痘、體重增加等，在醫學期刊中所提到PCOS治療方式包括服用Spironolactone抗雄性素、服用Metaformin 減少肝臟葡萄糖分解，以及在膳食中補充DCI，可以改善PCOS的症狀，更曾有恢復排卵個案。美國也因蕎麥DCI對於PCOS症狀的明顯改善作用^(21,22)，於2004年有DCI膳食補充品的上市(Chiral Balance膳食補充品，即DCI)。

然而蕎麥亦含過敏性蛋白質，可透過呼吸道接觸或是食用使人產生過敏反應，症狀包括氣喘或呼吸緊迫、蕁麻疹、皮膚腫癢、腸胃不適、眼睛不適等，從接觸過敏原到發生反應的時間很短，主要為IgE免疫球蛋白所介導誘發的第一型過敏反應^(23,24)。蕎麥引發的過敏反應可使用皮膚針刺試驗、測血清特異性IgE或食物過敏原貼布測試(food atopy patch test)、嗜伊紅性球測試(eosinophili test)來檢驗⁽²⁵⁾，Nakamura等調查蕎麥過敏者，其中75%在嗜伊紅性球測試為陽性，其調查病例中，僅有14%為20歲以上成年人，男性與女性比例為1.64，其中60%為呼吸道接觸^(26,27)。Takahashi等調查橫濱市9萬名學童發現，約有0.22%的學童食用蕎麥食品後會產生過敏現象⁽²⁸⁾。Wieslander等(2000)針對中國太原市不同職業族群之調查，以皮膚針刺試驗發現1.6%的成人對蕎麥過敏，主要為透過呼吸道接觸過敏原所導致⁽²⁴⁾。Badiu等(2013)同樣以皮膚針刺調查法分析義大利的人口，在義大利南北地區對蕎麥過敏的比例介於2.2~4.5%之間⁽²⁹⁾。由對不同民族、年齡層的分析調查大致可歸納兒童較容易發生蕎麥過敏現象，此過敏症狀會隨著年紀增長而改善，東方人對蕎麥過敏的盛行率低於白種人。整體而言，對蕎麥過敏的人口比例仍低於其他食品如小麥(麩質)、甲殼類海鮮、花生、堅果等常造成過敏之食物，但如能藉由育種程序降低或減少過敏蛋白質含量，對產業發展會有很大助益。

(三)蕎麥花與種子基因表現分析及過敏性蛋白質研究

蕎麥的花具有不同長短花柱型態，以及自交不稔特性，長久以來就吸引相當多研究者矚目，探討控制花器分化發育基因的表現情形與遺傳行為，並從韃靼蕎麥中引進自交稔性基因，

其次，蕎麥的花朵中含有極高含量的芸香苷，因而為純化芸香苷的良好材料。Logacheva氏等⁽³⁰⁾利用454 pyrosequencing技術，探討2種蕎麥花朵轉錄體表現的差異，希望對蕎麥轉錄體基因表現有更深入的认识，對蕎麥(*F. esculentum ssp. ancestrale*)及韃靼蕎麥(*F. tataricum*)分別讀取267千及229千筆序列資訊(reads)，平均長度為341~349核苷酸。所讀取之序列在2個品種均組成25千個contigs，定序覆蓋率平均7.5~8.2倍，2個品種的轉錄體特徵整體而言是相似的，但是retrotransposon genes和在醣類生合成代謝相關基因的表現有差異。其中13個單套基因用來作為親緣關係的分析，以核基因的序列分析結果和用葉綠體基因的分析具有一致性，和石竹目(Caryophyllales)與菊目(asteroids)的姐妹群關係再次受到證實。

Gupta氏等⁽³¹⁾以cDNA-AFLP方法分析蕎麥與韃靼蕎麥從開花到種子成熟階段基因表現的差異，以32組引子擴增獲得509個差異表現片段，其中64%為量性，依照生物程序(biological processes)分類，33%為初級代謝與二次代謝，其餘分別為調節功能(18%)、訊息傳導(14%)、運輸(13%)、細胞組織(10%)、光合作用與能量(4%)⁽³¹⁾。蕎麥可累積高量的芸香苷，芸香苷之合成步驟為由flavonoid 3-O-glucosyltransferase在槲皮素上加上糖基(glucose)，第二個步驟為在糖基上由rhamnosyltransferase再加入鼠李糖基(rhamnose)，目前蕎麥芸香苷合成第一步驟之flavonoid 3-O-glucosyltransferase已經被選殖發表⁽³²⁾。

在過敏蛋白質的研究方面，許多具IgE結合活性不同分子量的蕎麥蛋白質已陸續分離，包括100, 50, 17 KDa⁽³³⁾，8-9 kDa⁽³⁴⁾以及24 kDa^(35,36)等。韃靼蕎麥的種子24, 35及56 KDa之過敏原蛋白質也曾經被分離研究^(37,38)，這些過敏蛋白質皆屬於cupin family，主要由種子11S及7S貯存蛋白質組成，其中Fag e1核心蛋白質為主要抗原決定位，具有最強的IgE的結合活性^(37,38,39,40)，此過敏蛋白可抗胃蛋白酶之降解，但若透過梅納反應與種子多醣體鍵結，可有效增加其溶解性並減少與IgE結合活性^(41,42)。雖然蕎麥和韃靼蕎麥的種子過敏蛋白質基因已大致被選殖分析⁽⁴³⁾，然而迄今並無蕎麥中相關種子貯存蛋白質、過敏蛋白質基因豐富度和歧異度與基因表現量的探討，本研究的轉錄體分析對蕎麥種子貯存蛋白質和過敏蛋白質將有更深入且全面的認識。

材料與方法

植物材料

蕎麥臺中5號於2013年1月以種子條播方式定植於臺中區農業改良場(彰化縣大村鄉)，定植後第6週即陸續開始開花結實，於第7週收穫充實期種皮為綠色之種子，以及與同一植株最上方之成熟葉片，採回實驗室後立即進行RNA之萃取。

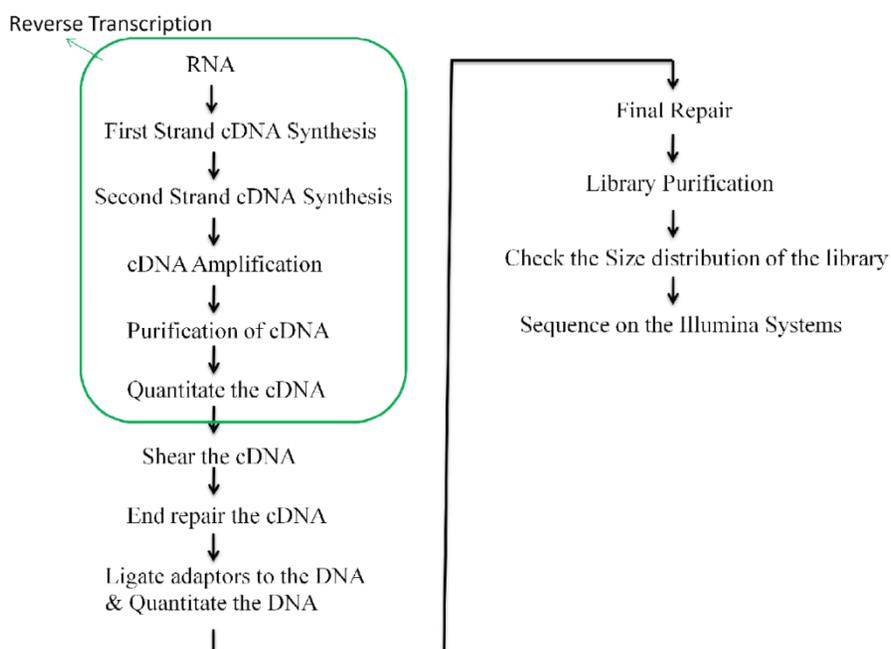
RNA萃取

種子與葉片RNA之萃取使用TANBead® Plant RNA kit，首先取50 mg種子及葉片組織各4重複，單一様品置於小試管中，加入試劑組之lysis buffer後以植物組織均質器研磨，之後取出

均質液轉到TANBead全自動磁珠操作平臺(臺灣圓點奈米技術公司)，自動萃取純化RNA，所純化之RNA暫存於-20°C冰箱保存。

轉錄體定序

種子與葉片之RNA樣品使用Agilent Bioanalyzer 2100以RNA 6000進行品質分析，以RNA完整性(RNA Integrity Number, RIN)值7.0為閾值，棄卻RIN低於7.0之樣品，本研究所採用之種子與葉片樣品RIN值分別為8.7及7.3。RNA樣品使用Nugen Ovation RNA Seq System V2進行反轉錄反應合成cDNA；cDNA以Covaris S220剪切用以製備固定長度cDNA之文庫構築，使用Nugen Encore Rapid DR system將cDNA兩端接上接合器後，再經過純化、定量與文庫長度分布確認後，隨即以Illumina HiSeq 2000上機進行雙邊各100 bp定序(如圖一工作流程)。

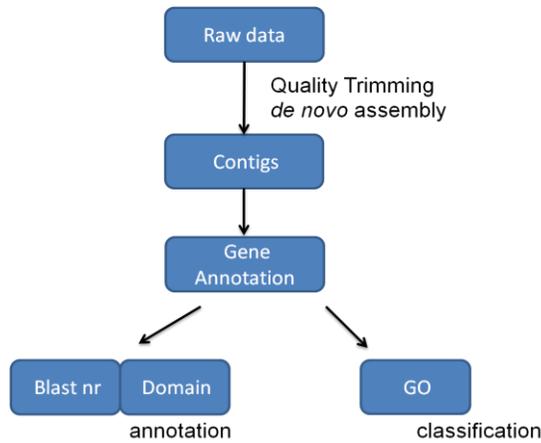


圖一、蕎麥種子與葉片轉錄體材料的製備流程。

Fig. 1. Preparation of seeds and leaves transcriptome of buckwheat.

轉錄體分析與基因分類註解

以CLC Genomics workbench套裝軟體之”trim模組”去除訊號不良之序列，並使用”de novo組裝模組”將種子與葉片之reads混合組裝(由於無法獲得蕎麥詳細的基因組資訊故以此作為本次參考基因訊息)；以blastx與NCBI nr資料庫比對，進行基因註解(annotation)⁽⁴⁴⁾(圖二)。再將此註解好之contigs當作參考基因組，分別將種子與葉片之reads進行RNA-Seq定序，並以RPKM(Reads Per Kilo-base per Million reads)來表示參考基因之表現量⁽⁴⁵⁾。將種子與葉片表達量有差異之參考基因分類(Gene Ontology)與KEGG分析其調控途徑。



圖二、轉錄體分析與基因分類註解流程。

Fig. 2. Flow chart of gene ontology classification and annotation.

結果與討論

本研究中送定序之種子(BS)與葉片(BL)樣本分別讀取143 M及247 M條序列，共組裝97,200及74,544條contigs，平均覆蓋率(coverage)分別約為1,474及3,340倍，contig序列平均長度分別為565及607 bp，全部拼接完成之序列分別為54 M及45 M bp (表一)。參考Logacheva氏等進行之蕎麥花朵轉錄體分析，共讀取266,782條序列，獲得25,435條contig，平均長度為698 bp，覆蓋率為8.2。

表一、原始資料及重疊群(contig)之特徵

Table 1. Characteristics of raw data and contigs

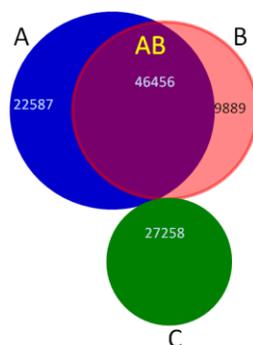
	Buckwheat seeds	Buckwheat leaf	BS-BL mixed assembly
Number of reads	143,803,530	247,195,816	390,999,346
N75	370	397	445
N50	696	810	840
N25	1,457	1,651	1,716
Minimum Length (bp)	135	68	106
Maximum Length (bp)	17,000	14,726	23,106
Average Length (bp)	565	607	640
Number of Contigs	97,200	74,544	106,190
Total bps	54,889,512	45,283,362	67,945,620

本研究所使用之種子及葉片轉錄體樣品同為蕎麥臺中5號，完成種子及葉片轉錄體分析之後，進一步混合拼接以比較種子和葉片表現基因之異同，如此可以得知哪些contigs分別屬

於種子或葉片，或是屬於兩者共有，不屬於葉片或種子則可視為錯誤拼接(圖三、C)可排除之。其次，混合拼接可將不完整的同源序列組裝為更完整之contigs。

所有讀取之序列進行混合拼接，共可獲得106,190條contig，其中屬於種子和葉片者分別為69,043及56,345條(圖三)，單獨存於種子或葉片者分別為22,587及9,889條；混合拼接後衍生27,258條(圖三，C區) contig其RPKM值為0，這些組裝後無法算出RPKM值的contig，表示其為由兩者共同assembly出的contig，不屬於種子或葉片，在此認定為錯誤之拼接，佔總contigs數量的25.6%，由原始資料也可以看出這些共同組裝的contig其長度平均也短，為 260 ± 53 bp，進行blast nr比對結果，大多也無法得到理想的結果(資料未顯示)，推測為降解的RNA。扣除錯誤拼接的contigs之後，成功組裝的contigs認定為78,932條 ($106,190 - 27258 = 78932$)

A (BS) Area: 69043, B (BL) Area: 56345, C Area: 27258,
AB Area: 46456, AC Area: 0, BC Area: 0, ABC Area: 0,
Total: 106190

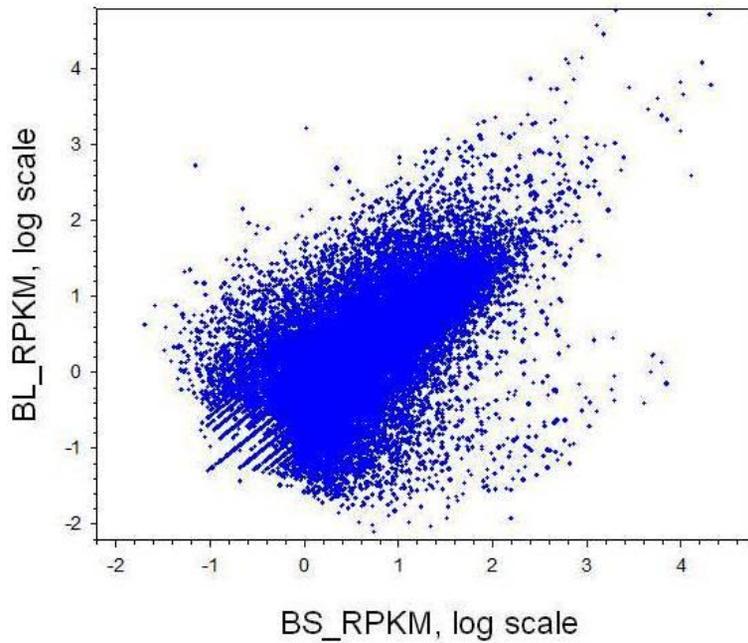


圖三、蕎麥葉片與種子個別拼接與混合拼接後所獲得 contig 之片段分類。BS：蕎麥種子；BL：蕎麥葉片。

Fig. 3. Contigs segmentation after individual and mixed assembly. BS: buckwheat seed; BL: buckwheat leaf.

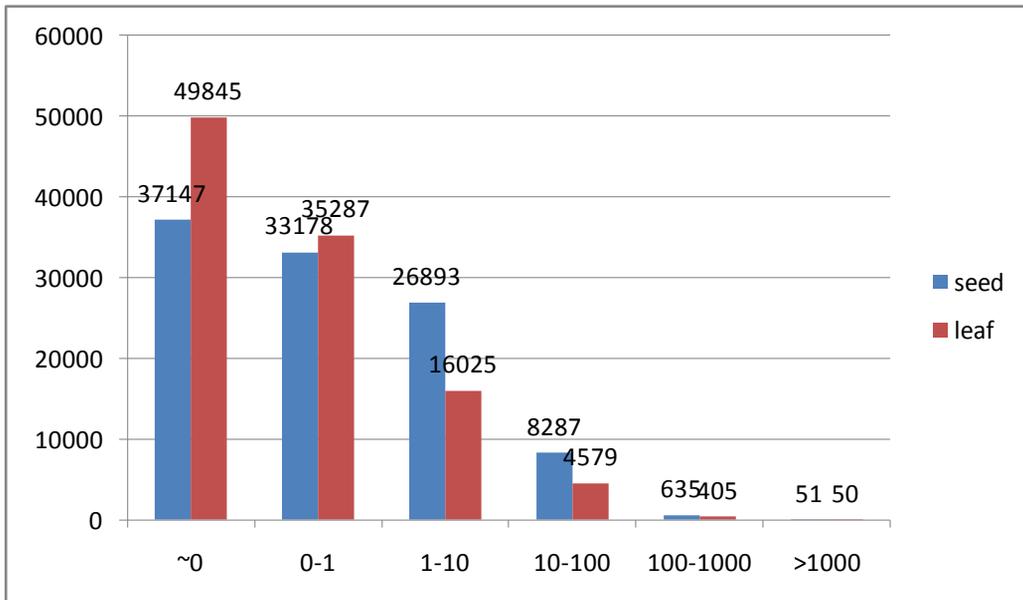
以混合拼接後各個contig在種子和葉片中標準化後的RPKM值繪製散佈圖。比較相同contig在兩種子及葉片中相對表現差異(圖四)，整體而言，同樣的contig在種子中有較高的表現量。

進一步分析各contig之RPKM值可發現，種子和葉片中各有37,147及49,845條contigs之RPKM值趨近於0，所比對到序列長度(hit length)也都極低(15~100 bp)，RPKM值介於0~1之間者各有33,178及35,287條(圖五)，這些contigs因為表現量低，或是因為錯誤組裝使得其在整體資料中相對為微量，或者是因為其本身表現即為微量，推測並非我們有興趣之基因。蕎麥引人注目的機能性成分包括芸香苷、手性肌醇(D-chiral inositol)、cupin家族的過敏蛋白質等，從文獻得知這些成分含量在蕎麥中都相當高，本研究中專注在高表現量的contig分析；在數據的處理上，篩選掉種子RPKM<10之後的contig，總共剩下8,973條contig(圖五)。此外，混合拼接之後表現量高於1的contig在種子樣品中數量也較多(圖五)，因此本研究篩選在種子中具有高表現量(RPKM>10)，且和葉片表現量相比也有高比值者優先進行分析。



圖四、種子和葉片標準化後 contig 之 RPKM 數值散佈圖，BS：蕎麥種子；BL：蕎麥葉片。

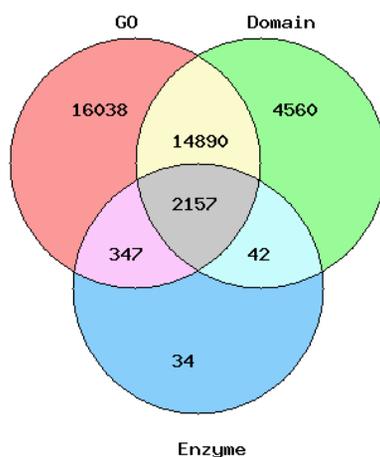
Fig. 4. Scattered plot of standardized seed and leaf RPKM value. BS: buckwheat seed; BL: buckwheat leaf.



圖五、蕎麥葉片與種子混合拼接之 contig 依照 RPKM 值分類圖。

Fig. 5. Distribution of RPKM value of mixed assembled contigs.

所有contig序列與資料庫相比對，其中33,342筆可比對出基因功能(gene ontology)，與蛋白質區塊資料庫比對有21,649筆，與酵素資料庫比對則可得到2,580筆(圖六)，而其中三者交集的序列為2,157筆，三者聯集者為37,936筆。整體而言所有比對到的筆數占全部認定為成功組裝contigs的48.1%。



圖六、蕎麥種子與葉片轉錄體註解分類。

Fig. 6. Buckwheat contigs (BS-BL) annotation segmentation.

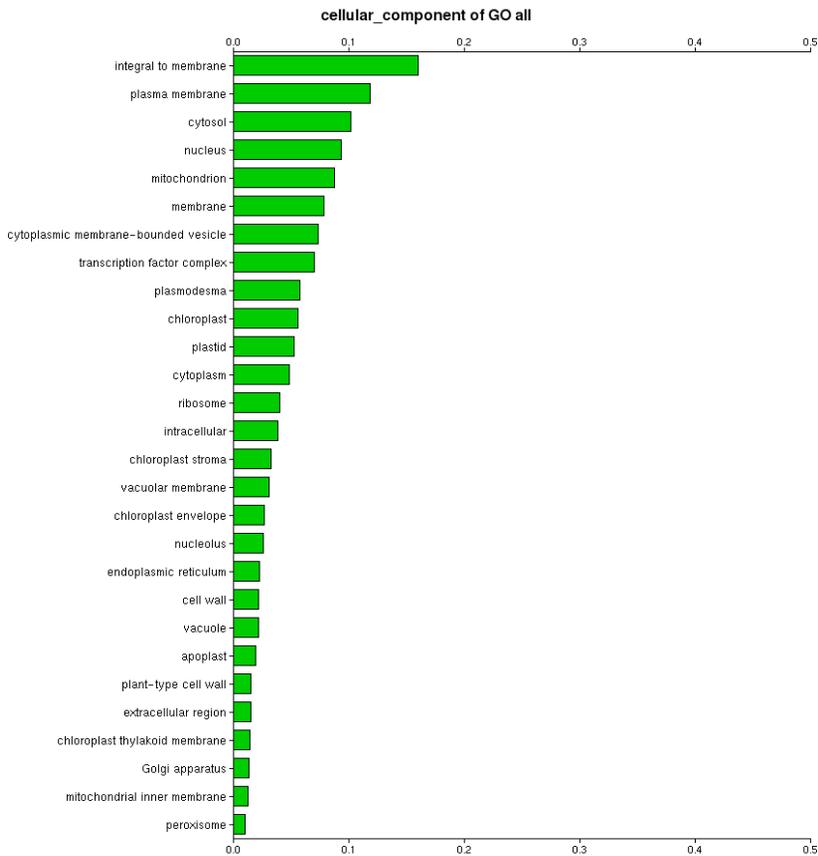
依照細胞組成加以分類，佔比重最高的前五名依序為細胞膜成分(integral to membrane)、原生質膜(plasma membrane)、細胞質(cytosol)、細胞核(nucleus)以及粒線體(mitochondria)(圖七)。前二名皆屬於細胞膜上之功能基因，可見膜功能在生物中的重要性，受限於時間因素，目前尚未和其他植物文獻報告比對，而Logacheva氏等⁽³⁰⁾所發表的蕎麥花芽轉錄體並未提出以細胞組成分類的結果。

依據 Logacheva 氏等分析蕎麥花芽轉錄體的報告⁽³⁰⁾，其使用簡化的 GOSlim註解方式和阿拉伯芥的生物資訊(TAIR)⁽⁴⁶⁾互相比較，依照生理程序分類蕎麥花芽轉錄體之五項分別為生理程序、細胞程序、代謝程序、初級代謝程序與細胞程序，可知此分類方式相當簡化。本研究比對2013年之NCBI資料庫序列並註解，生理程序排序前五名分別為氧化還原反應、絲胺酸家族胺基酸代謝、DNA依賴性的轉錄控制、蛋白質磷酸化以及電子傳遞鍊(圖八)，

當以分子功能為分類依據時，排序前五名分別為ATP binding, Zinc ion binding, DNA binding, sequence-specific DNA binding transcription factor activity, and binding(圖九)。而Logacheva氏等之報告其排序前五名分別為molecular function, binding, catalytic activity, transferase activity and protein binding⁽³⁰⁾，與本報告分類方式不同，但是並無太大之牴觸。

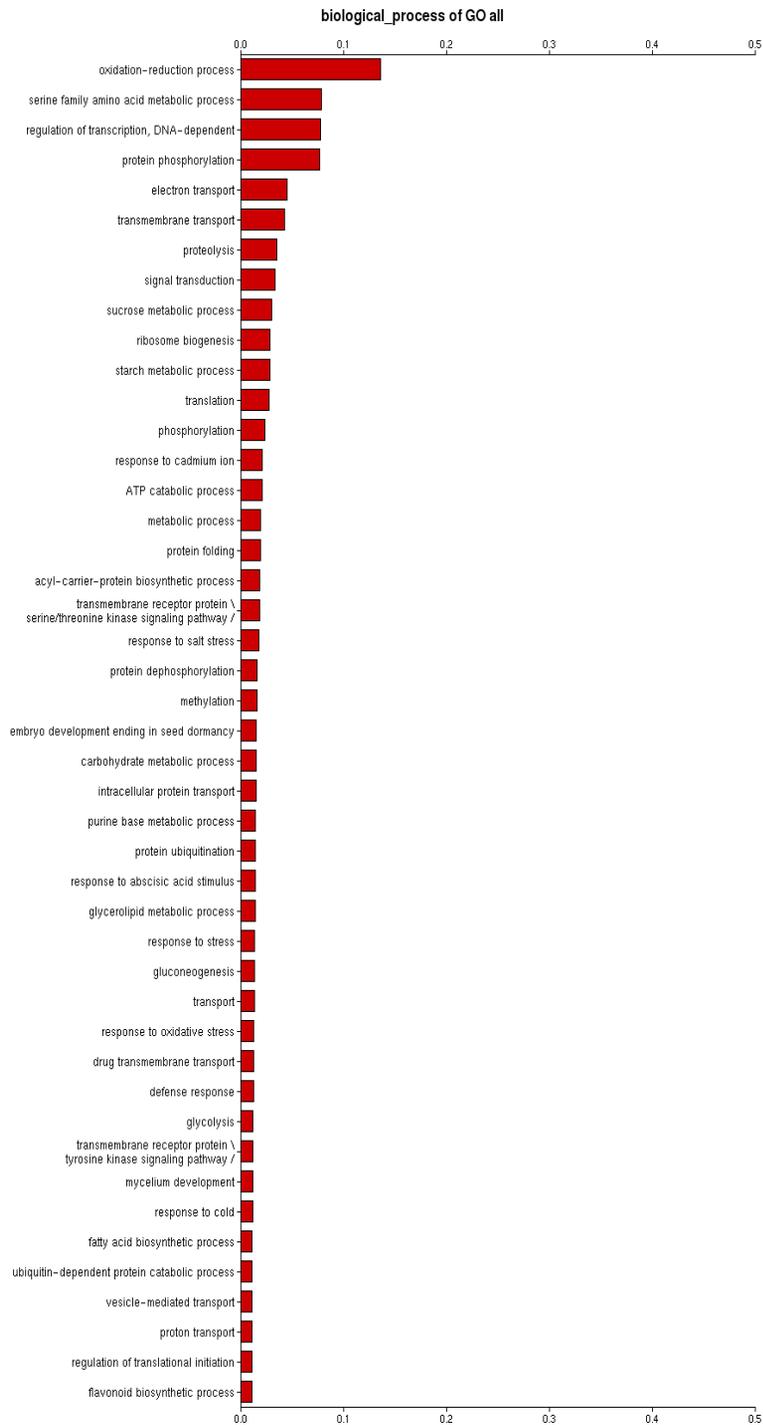
本轉錄體資訊亦與KEGG (Kyoto Encyclopedia of Genes and Genomes)^(47,48)資料庫比對，以了解所組裝contig在不同代謝途徑出現的情形，共完成包括糖類、胺基酸、核苷酸代謝、TCA循環、氧化磷酸化等306個代謝途徑之比對(圖十)，由於資訊量龐大，因此僅先針對類黃

酮、黃酮與黃酮醇(flavones and flavonol) (圖十一)及芸香苷代謝途徑基因(圖十二)加以檢視，在類黃酮生合成代謝基因方面，由p-coumaroyl-CoA等前驅物開始，至 flavone, flavonol, anthocyanin 生合成之重要節點中間物質，如 naringenin, kaempferol, apigenin, cyaniding, delphinidin, pelargonidin等，其相關生合成基因的contigs皆可比對到(資料未顯示)，進一步顯示 flavone 的生合成，也可以比對到 quercetin 的合成基因(圖十一、酵素 1.14.13.21)，以及 isoquercitrin (2.4.1.91)與 quercitrin (2.4.1.-)的基因存在，但是在下游的芸香苷合成酵素(2.4.1.159)則未比對到。依據 Lucci and Mazzafera⁽⁴⁹⁾所提供的芸香苷合成途徑，包括了由槲皮素經過UDP-glucosyl transferase形成 quercetin-3-glucoside，之後UDP-rhamnosyl transferase加上一個 rhamnose，形成芸香苷(圖十二)。在蕎麥的代謝物分析中，我們可確認蕎麥葉片含有高量的芸香苷，反而槲皮素(quercetin)為微量(詳如附件蕎麥成分分析研究報告)，因此再以UDP-glucosyl transferase 和 rhamnosyltransferase 為關鍵字檢索，仍可分別檢索到34及12筆 contig其GO功能涵蓋對應關鍵字(表二)，後續將擷取該序列進行深入分析。其他蕎麥機能性成份與有興趣之蛋白質生合成基因，其在本轉錄體之表現資料列於表二。



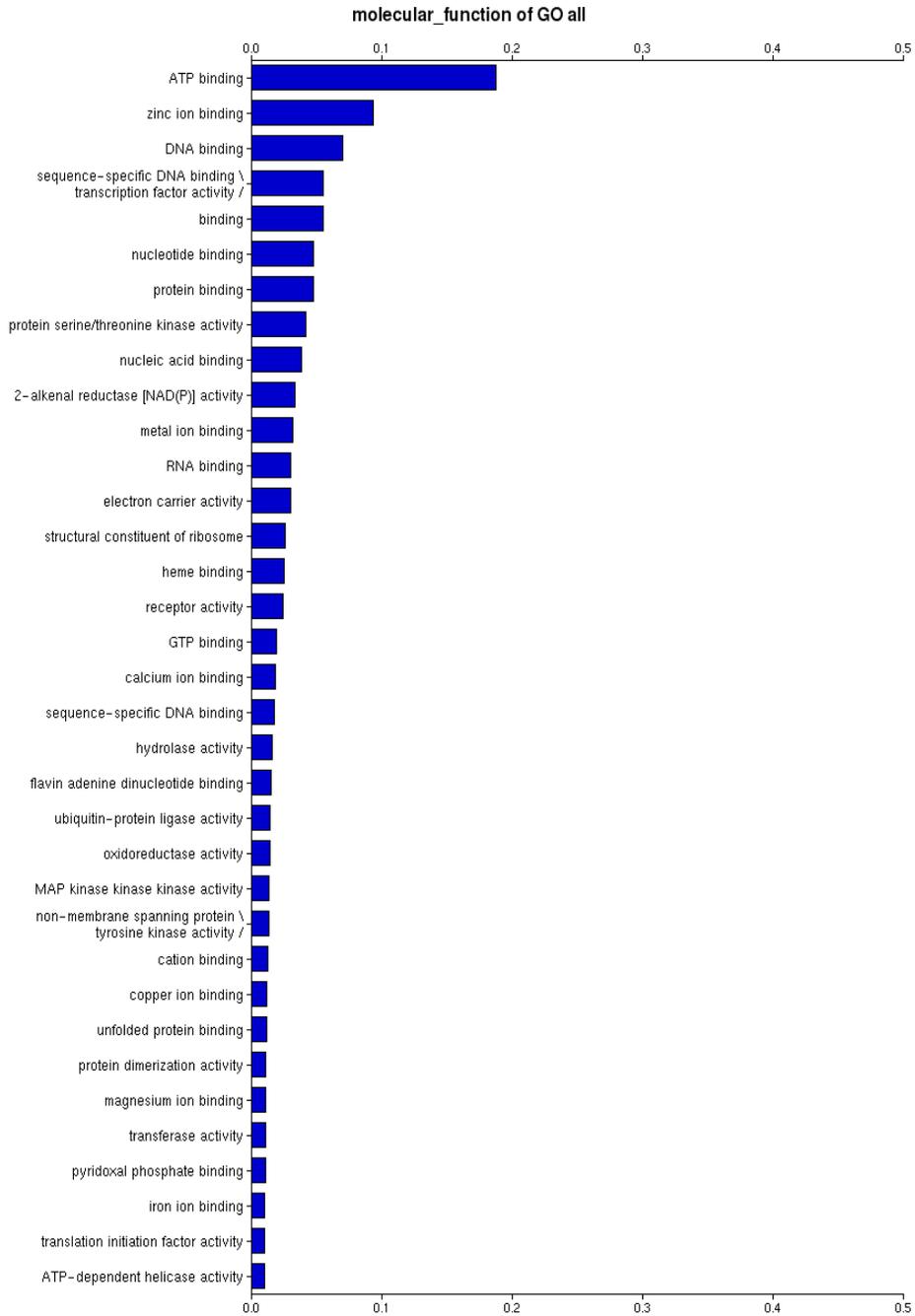
圖七、以細胞成份為標準之基因功能分類。

Fig. 7. Gene ontology classification using cellular components as criteria.



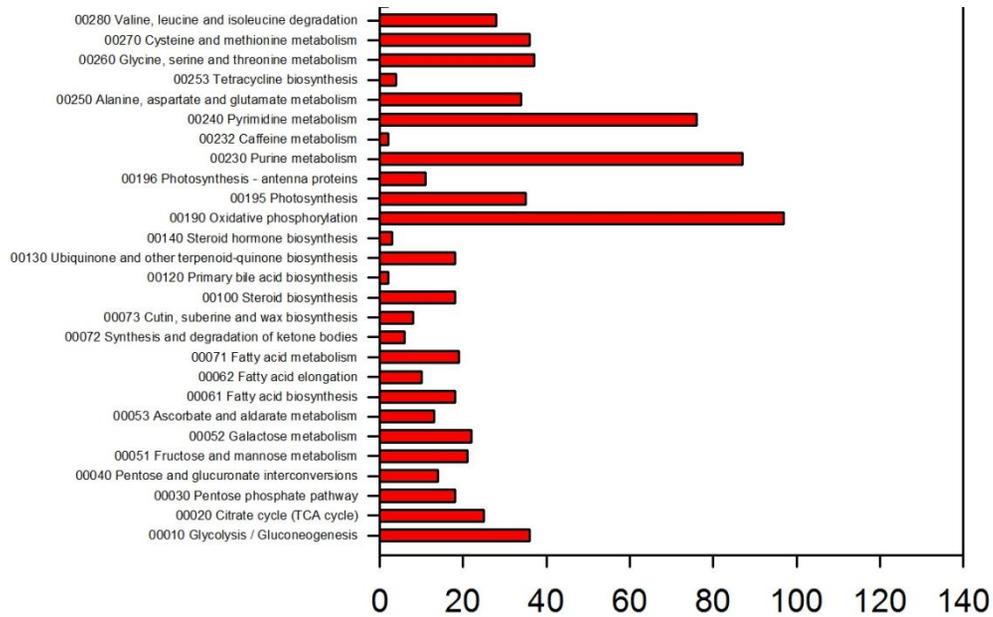
圖八、以生理程序為標準之基因功能分類。

Fig. 8. Gene ontology classification using biological processes as criteria.



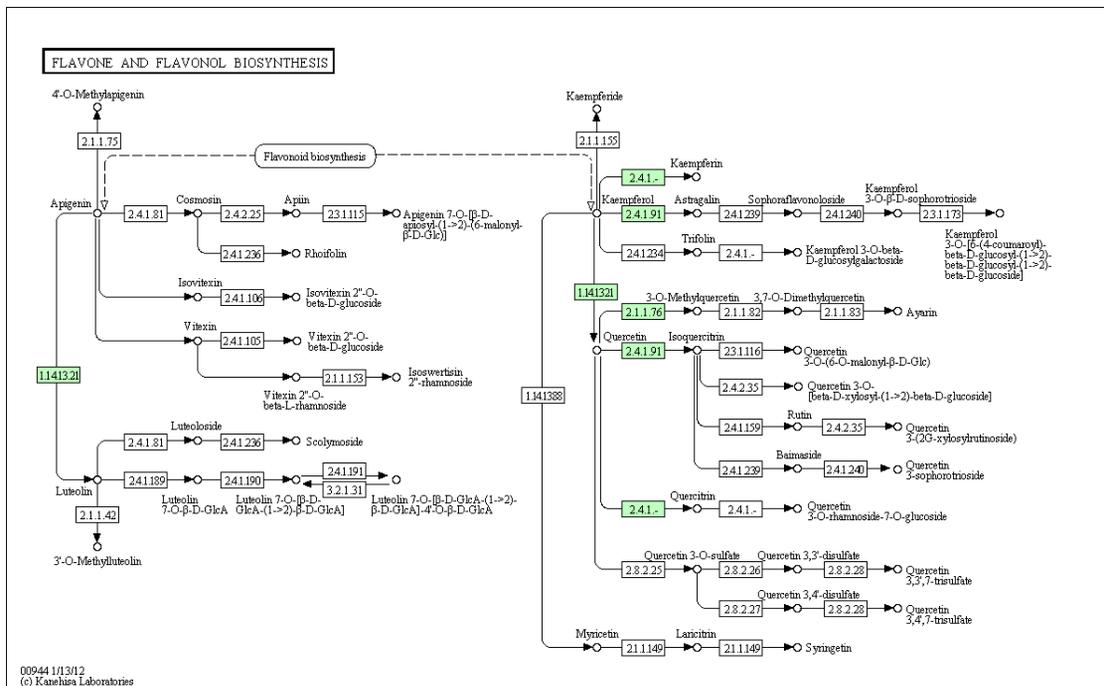
圖九、以分子功能為標準之基因功能分類。

Fig. 9. Gene ontology classification using molecular functions as criteria.



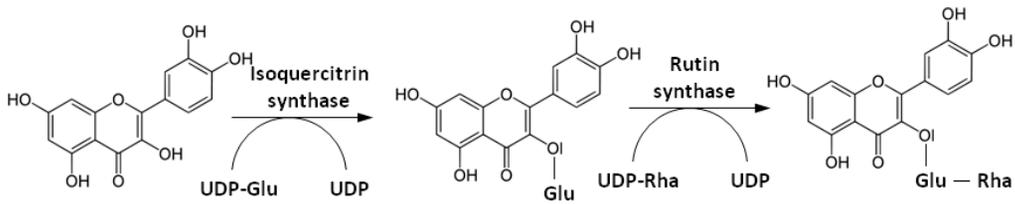
圖十、與細胞代謝資料庫 KEGG 比對之結果(僅顯示部分資料)。

Fig. 10. Nucleotide blast with KEGG database, only part of the results are shown.



圖十一、與 KEGG 資料庫比對黃酮與黃酮醇(flavones and flavonol)代謝途徑基因。

Fig. 11. Contigs blasted with KEGG Flavones and flavonol biosynthesis pathway.



圖十二、芸香苷生成代謝途徑 (引用自 Lucci and Mazzafera, 2009)。

Fig. 12. Biosynthesis pathway of rutin (Cited from Lucci and Mazzafera, 2009).

表二、於本轉錄體 contig 檢索蕎麥機能性成份與部份蛋白質合成基因之結果

Table 2. Results of contig sorting for genes might associate with buckwheat nutraceutical composition and particular protein synthesis

Keywords for sorting	No. of contigs	Descriptions or notes
Seed storage protein	15	13 S globulin seed storage protein
LEA	7	Late embryogenesis abundant protein
Trypsin inhibitor	19	8 contigs match published Tartary buckwheat trypsin inhibitor
Cupin	42	A large family of storage protein that contains buckwheat allergen
Allergen	53	Include pollen allergen, Bet v1, Ole e2, Fag e1 allergen, etc.
Fag e 1 allergen	8	Buckwheat major allergenic protein
Fagopyritol synthase	8	Contains D-chiral inositol
Glucosyltransferase	34	Need characterization
Rhamnosyltransferase	12	Need characterization

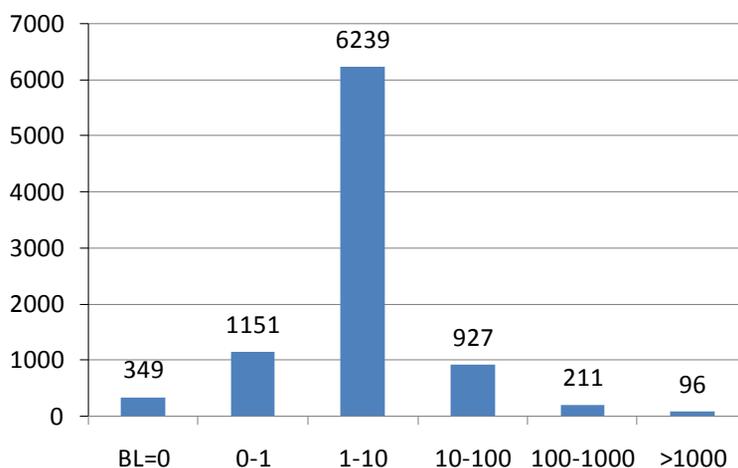
種子中高表現量 contigs 之篩選

和農業栽培最相關者為種子產量與品質，我們認為特定 contigs 在種子中表現量高 (RPKM>10)，即為與種子充實高度相關之基因。同時，比較在種子和葉片表現量比值，比值高者為該基因在種子的表現量高於葉片，如此即可得知，有多少數量的 contig 在種子中有專一性的大量表現(圖六)。由圖十三之 BS/BL RPKM 比值可知，共有 7,822 條 contig 在種子有高表現量，未來將優先探討這些高表現量 contigs 的基因功能。

結論與建議

由於次世代定序技術(Next Generation Sequencing, NGS)的快速發展，使得基因體與轉錄體的分析得以快速進行，近年來學界已經累積豐富的作物生物資訊，科研瓶頸反而出現在資料的解析。本研究雖已完成蕎麥葉片與種子轉錄體之定序與註解，然而分析工作仍在早期階段，目前僅能歸納本轉錄體大致的輪廓，尚未進入深入分析階段。蕎麥中有數項引人入勝的議題，包括植株可以合成累積超高量的芸香苷，可達 46,348 ppm，相當於 4.6 % 的乾物重，筆

者在分析本場所栽培之蕎麥臺中5號時獲得此數據，與國外所發表文獻相當。蕎麥種子含有豐富的手性肌醇，為胰島素訊息傳導受體 PIP2 的組成成分之一，可增加糖尿病第II型病患對胰島素之敏感性，也有助醫治PCOS病患之病徵。DCI與galactopyranose合成fagopyritol A2、A3、B2、B3，為麩皮上之可溶性碳水化合物，可開發為具醫療效果的膳食補充品。然而蕎麥亦具有過敏性蛋白質，雖然蕎麥所造成食物過敏現象的盛行率遠低於小麥麩質、海鮮、花生堅果等八大類常造成過敏食品，若能透過育種方式降低蕎麥過敏蛋白質含量，對蕎麥產業會有極大的幫助。本研究分析蕎麥種子和葉片轉錄體，報告中針對學術界對於蕎麥有興趣的幾項議題進行初步篩選分析，未來仍需更深入研究影響蕎麥產量與品質的因素，以期對蕎麥栽培育種及加工有更大的助益。



圖十三、蕎麥種子和葉片 BS/BL contigs RPKM 比値之分佈。

Fig. 13. Distribution of the BS/BL contigs RPKM ratio.

誌 謝

本研究由農委會科技計畫補助（計畫名稱：蕎麥及薏苡DNA生物條碼與化學指紋圖譜分析模式之建立，計畫編號：102農科-9.1.1-中-D1），轉錄體資訊在中研院生物多樣性中心趙淑妙老師指導下進行，並得莊樹諄老師提供分析建議，與研究室其他學弟妹共同討論轉錄體分析訣竅，源資生物科技公司許宏吉博士協助資料剖析，特此致謝。

參考文獻

1. 曾勝雄、陳裕星 2007 蕎麥臺中2號之育成 臺中區農業改良場研究彙報 95: 49-59。
2. 曾勝雄、陳裕星 2008 蕎麥臺中3號之育成 臺中區農業改良場研究彙報 100: 13-23。
3. 曾勝雄、陳裕星、廖宜倫 2009 蕎麥臺中5號之育成 臺中區農業改良場研究彙報 103: 1-15。

4. Cawoy, V., J. M. Kinet and A. L. Jacquemart. 2008. Morphology of nectarines and biology of nectar production in the distylous species *Fagopyrum esculentum*. *Annals of Botany* 102: 675-684.
5. Li, S. Q. and Q. H. Zhang. 2001. Advances in the development of functional foods from buckwheat. *Crit Rev Food Sci Nutr.* 41: 451-464.
6. Kusano, T., N. Nakayama, J. Doi, Y. Kiyomasa, H. Chine and K. Ikeda. 1992. Studies on nutritional aspects of tartary buckwheat. In *Proc. 5th Intl. Symp. Buckwheat.*; 453-454. August 20-26 1992, Taiyuan, China.
7. Fabjan, N., J. Rode, I. J. Kosir, Z. Wang, Z. Zhang and I. Kreft. 2003. Tartary buckwheat (*Fagopyrum tataricum* Gaertn.) as a source of dietary rutin and quercitrin. *J Agric Food Chem.* 51: 6452-6455.
8. La, Casa C., I. Villegas, C. Alarcón de la Lastra, V. Motilva and M. J. Martín Calero. 2000. Evidence for protective and antioxidant properties of rutin, a natural flavones, against ethanol induced gastric lesions. *J Ethnopharmacol.* 71: 45-53.
9. Bonafaccia, G., M. Marocchini and I. Kreft. 2003. Composition and technological properties of the flour and bran from common and tartary buckwheat. *Food Chem* 80: 9-15.
10. Landberg, R., Q. Sun, E. B. Rimm, A. Cassidy, A. Scalbert, C. S. Mantzoros, F. B. Hu and R. M. van Dam. 2011. Selected dietary flavonoids are associated with markers of inflammation and endothelial dysfunction in U.S. women. *J Nutr.* 141(4): 618-625.
11. Kahn, C. R. 1985. Pathophysiology of diabetes mellitus: an overview. In: *Joslin's diabetes mellitus.* ed. Marble, A. et al., 12th ed., Lea & Febiger, Philadelphia.
12. Lopez, M. P., M. J. Gomez-Lechon and J. V. Castell. 1991. *Diabetes* 40, 263-268.
13. Laakso, M. and S. Lehto. 1998. Epidemiology of risk factors for cardiovascular disease in diabetes and impaired glucose tolerance. *Atherosclerosis* 137, S65-S73.
14. Li, Y. Q., F. C. Zhou, F. Gao, J. S. Bian and F. Shan. 2009. Comparative evaluation of quercetin, isoquercetin and rutin as inhibitors of alpha-glucosidase. *J Agric Food Chem.* 57(24): 11463-8.
15. Lazzè, M. C., M. Savio, R. Pizzala, O. Cazzalini, P. Perucca, A. I. Scovassi, L. A. Stivala and L. Bianchi. 2004. Anthocyanins induce cell cycle perturbations and apoptosis in different human cell lines. *Carcinogenesis*, 25: 1427-1433
16. Afaq, F., D. N. Syed, A. Malik, N. Hadi, S. Sarfaraz, M. H. Kweon, N. Khan, M. A. Zaid and H. Mukhtar. 2007. Delphinidin, an anthocyanidin in pigmented fruits and vegetables, protects human HaCaT keratinocytes and mouse skin against UVB-mediated oxidative stress and apoptosis. *J Invest Dermatol.* 127: 222-232.
17. Obendorf, R. L. 1998. Buckwheat Fagopyritols. *Advances in buckwheat research. 7th Intl Symp Buckwheat.* 3: 65-71.

18. Steadman, K. J., M. S. Burgoon, R. L. Schuster, B. A. Lewis, S. E. Edwardson and R. L. Obendorf. 2000. Fagopyritols, D-chiro-inositol, and other soluble carbohydrates in buckwheat seed milling fractions. *J Agric Food Chem.* 48(7): 2843-2847.
19. Nestler, J. E., D. J. Jakubowicz, P. Reamer, R. D. Gunn and G. Allan. 1999. Ovulatory and metabolic effects of D-chiro-inositol in the polycystic ovary syndrome. *N Engl J Med.* 340: 1314-1320.
20. Nestler, J. E., D. J. Jakubowicz and M. J. Iuorno. 2000. Role of inositolphosphoglycan mediators of insulin action in the polycystic ovary syndrome. *J Pediatr Endocrinol Metab.* 5: 1295-8.
21. Galletta, M., S. Grasso, A. Vaiarelli and S. J. Roseff. 2011. Bye-bye chiro-inositol - myo-inositol: true progress in the treatment of polycystic ovary syndrome and ovulation induction. *Eur Rev Med Pharmacol Sci.* 15(10): 1212-4.
22. Wieslander, G. 1995. Allergy to buckwheat. *Current Advances in Buckwheat Research:* 951-955.
23. Wieslander, G, D. Norbäck, Z. Wang, Z. Zhang, Y. Mi, and R. Lin. 2000. Buckwheat allergy and reports on asthma and atopic disorders in Taiyuan City, Northern China. *Asian Pac J Allergy Immunol.* 18(3): 147-52.
24. Satoh, R., S. Koyano, K. Takagi, R. Nakamura and R. Teshima. 2010. Identification of an IgE-binding epitope of a major buckwheat allergen, BWp16, by SPOTs assay and mimotope screening. *Int Arch Allergy Immunol.* 153(2): 133-40.
25. Nakamura, S., M. Yamaguchi, M. Oishi and T. Hayama. 1974/1975. Studies on the Buckwheat allergose. Report 2: On the cases with the Buckwheat Allergose. *Allergie und Immunologie,* 20/21: 457-465.
26. Nakamura, S., Y. Suzuki, E. Ishikawa, T. Yakushi, H. Jing, T. Miyamoto and K. Hashizume. 2008. Reduction of in vitro allergenicity of buckwheat Fag e1 through the Maillard-type glycosylation with polysaccharides, *Food Chemistry,* 109(3), 538-545.
27. Takahashi, Y., S. Ichikawa, Y. Aihara and S. Yokota. 1998. Buckwheat allergy in 90,000 school children in Yokohama. *Arerugi.* 47(1): 26-33. [Article in Japanese]
28. Badiu, I., E. Olivieri, M. Montagni, G. Guida, S. Mietta, S. Pizzimenti, M. Caminati, M. R. Yacoub, E. Tombetti, D. Preziosi, C. Quecchia, S. Minetti, S. Facchetti, F. Fassio, I. Massaro, L. Corradi, M. C. Turi, A. Colagiovanni, L. Pascolini, F. W. Rossi, L. Losappio, L. Sansone, S. Imbesi, S. Leto Barone, G. Mistrello and Heffler E. 2013. Italian study on buckwheat allergy: prevalence and clinical features of buckwheat-sensitized patients in Italy. *Int J Immunopathol Pharmacol.* 26(3): 801-6.
29. Baillargeon, J. P., M. J. Iuorno, T. Apridonidze and J. E. Nestler. 2010. Uncoupling between insulin and release of a D-chiro-inositol-containing inositolphosphoglycan mediator of insulin action in obese women with polycystic ovary syndrome. *Metab Syndr Relat Disord.* 8(2): 127-36.

30. Logacheva, D., A. S. Kasianov, D. V. Vinogradov, T. H. Samigullin, M. S. Gelfand, V. J. Makeev and A. A. Penin. 2011. *De novo* sequencing and characterization of floral transcriptome in two species of buckwheat (*Fagopyrum*). BMC Genomics 12: 30.
31. Gupta, N., P. K. Naik and R. S. Chauhan. 2012. Differential transcript profiling through cDNA-AFLP showed complexity of rutin biosynthesis and accumulation in seeds of a nutraceutical food crop (*Fagopyrum* spp.). BMC Genomics, 13: 231
32. Suzuki, T, S-J. Kim, H. Yamauchi, A. Takigawa, Y. Honda and Y. Mukasa. 2005. Characterization of a flavonoid 3-O-glucosyltransferase and its activity during cotyledon growth in buckwheat (*Fagopyrum esculentum*). Plant Science 169: 943-948.
33. Yanagihara, Y, 1980. Buckwheat hypersensitivity. Kansensho 10: 184-188 (Japanese).
34. Yano, M., R. Nakamura, S. Hayakawa, S. Torii. 1989. Purification and properties of allergenic proteins in buckwheat seed. Agric Biol Chem 53: 2387-2392.
35. Urisu, A, Y. Kondo, Y. Morita, E. Wada, M. Tsuruta, T. Yasaki, K. Yamada, H. Kuzuya, M. Suzuki, K. Titani and K. Kurosawa, 1994. Identification of a major allergen of buckwheat seeds by immunoblotting methods. Allerg. Clin. Immunoi. 6, 151-155.
36. Nair, A. and T. Adachi. 1999. Immunodetection and Characterization of Allergenic Proteins in Common Buckwheat (*Fagopyrum esculentum*). Plant Biotechnology, 16 (3), 219- 224.
37. Wang, Z, Z. Zhang, Z. Zhao, G. Wieslander, D. Norbäck and I. Kreft. 2004. Purification and characterization of a 24 kDa protein from tartary buckwheat seeds. Biosci Biotechnol Biochem. 68(7): 1409-13.
38. Yoshimasu, M. A., J. W. Zhang, S. Hayakawa and Y. Mine. 2000. Electrophoretic and immunochemical characterization of allergenic proteins in buckwheat. Int Arch Allergy Immunol. 123(2): 130-6.
39. Yoshioka, H., T. Ohmoto, A. Urisu, Y. Mine and Adachi. 2004. Expression and epitope analysis of the major allergenic protein Fag e 1 from buckwheat. J Plant Physiol. 161(7):761-7.
40. Choi, S. Y., J. H. Sohn, Y. W. Lee, E. K. Lee, C. S. Hong and J. W. Park. 2007. Application of the 16-kDa buckwheat 2 S storage albumin protein for diagnosis of clinical reactivity. Annals of allergy, asthma & immunology 99(3), 254-260.
41. Nakamura, S., M. Yamaguchi, M. Oishi and T. Hayama. 1974/1975. Studies on the Buckwheat Allergose. Report 1: On the Cases with the Buckwheat Allergose. Allergy and Immunology 20/21, 449-456.
42. Yang, Z. H., C. Li, Y. Y. Li and Z. H. Wang. 2013. Effects of Maillard reaction on allergenicity of buckwheat allergen Fag t3 during thermal processing. J Sci Food Agric. 93(6): 1510-5.

43. Chen, P., Y. F. Guo, Q. Yan and Y. H. Li. 2011. Molecular cloning and characterization of Fag t 2: a 16-kDa major allergen from Tartary buckwheat seeds. *Allergy*. 66(10): 1393-5.
44. Chen, T. W., R. C. R. Gan, T. H. Wu, P. J. Huang, C. Y. Lee, Y. Y. M. Chen, C. C. Chen and P. Tang. 2012. FastAnnotator: an efficient transcript annotation web tool. *BMC Genomics* 13: Suppl 7, S9.
45. Wagner, G. P., K. Kin and V. J. Lynch. 2012. Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. *Theory Biosci.* 131(4): 281-5.
46. Rhee, S. Y., W. Beavis, T. Z. Berardini, G. Chen, D. Dixon, A. Doyle, M. Garcia-Hernandez, E. Huala, G. Lander, M. Montoya, N. Miller, L. A. Mueller, S. Mundodi, L. Reiser, J. Tacklind, D. C. Weems, Y. Wu, I. Xu, D. Yoo, J. Yoon and P. Zhang. 2003. The Arabidopsis Information Resource (TAIR): a model organism database providing a centralized, curated gateway to Arabidopsis biology, research materials and community. *Nucleic Acids Research*, 31: 224-228.
47. Kanehisa, M. and S. Goto. 2000. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* 28, 27-30 (2000).
48. Kanehisa, M., S. Goto, Y. Sato, M. Furumichi and M. Tanabe. 2012. KEGG for integration and interpretation of large-scale molecular datasets. *Nucleic Acids Res.* 40, D109-D114.
49. Lucci, N. and P. Mazzafera. 2009. Rutin synthase in fava d'anta: Purification and influence of stressors. *Canadian J. Plant Sci.* 89, 895-902.

The Analysis of Buckwheat Seed and Leaf Transcriptome¹

Yuhsin Chen², Chieh Wang³, Hwang-Bin Chen² and Yi-Lun Liao²

ABSTRACT

Buckwheat (*Fagopyrum esculentum* Moencho) is an annual crop belonging to the genus *Fagopyrum*, Polygonum. Buckwheat contains beneficial components and yield considerable amounts of nectar, therefore it has become one of the most important crops in the world. In this study we surveyed the transcriptome of buckwheat seeds and leaves, 143 M and 247 M sequence are read and assembled into 97,200 and 74,544 contigs, the average coverage is 1474 and 3340, respectively. Average length of contigs is 565 and 607 bp, overall sequence assembled is 54 M and 45 M bp. When seed and leaf transcriptome are combined, a total of 106,190 contigs are assembled, in which 69,043 and 56,345 contigs belong to seeds and leaf, respectively. Of all the contigs, 22,857 and 9,889 are exclusively exist in seeds and leaves, respectively. After mixed assembling, 27,258 contigs, 25.6% of all assembled contigs, are considered false assembled as the RPKM value is 0 in both seed and leaf transcriptome. The successfully assembled contigs are then blasted with databases and annotated. The gene ontology (GO) is annotated according to cellular component, biological process and molecular functions. The contigs are also blasted with KEGG (Kyoto Encyclopedia of Genes and Genomes) database and results in 306 metabolic pathways data. As the amount of information is tremendous, at the beginning stage we only inspect the metabolic pathway of flavonoids, flavones and flavonol. In the project, we are interested in seed properties include seed storage protein, LEA protein, cupin family and allergenic protein. All the above sequence are sorted and retrieved, which would be valuable for future breeding.

Keywords: Rutin, Allergenic protein, Fag e1, Fagopyritol

¹ Contribution No. 0839 from Taichung DARES, COA.

² Assistant Researcher of Taichung DARES, COA.

³ Research assistant of Institute of Biodiversity, Academia Sinica.