國立臺灣大學生物資源暨農學院園藝暨景觀學系

博士論文

Department of Horticulture and Landscape Architecture
College of Bioresources and Agriculture
National Taiwan University

Doctoral Dissertation

甘藍高溫淹水耐受性之 RNA-seq 轉錄體分析

A transcriptome analysis using RNA-seq to investigate the tolerance of the cabbage (*Brassica oleracea* var. *capitata L.)* to high temperature and waterlogging stresses

洪志良

Chih-Liang Hung

指導教授：許　輔 博士，羅筱鳳 博士

Advisor: Fuu Sheu, Ph.D., Hsiao-Feng Lo, Ph.D.

中華民國 105 年 7 月

July, 2016

doi:10.6342/NTU201603060

國立臺灣大學博士學位論文
口試委員會審定書

甘藍高溫淹水耐受性之RNA-seq轉錄體分析

A transcriptome analysis using RNA-seq to investigate the tolerance of the cabbage (*Brassica oleracea* var. *capitata L.*) to high temperature and waterlogging stresses

本論文係洪志良君（D98628003）在國立臺灣大學園藝暨景觀學系、所完成之博士學位論文，於民國 105 年 7 月 13 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：許輔 羅筱鳳（簽名）

（指導教授）

doi:10.6342/NTU201603060

誌謝

本論文得以順利完成，首要感謝恩師 許輔博士及 羅筱鳳博士的指導。回想在臺大園藝系近九年的時光，兩位老師提供學生優質的設備及支援，可以專精研究，也準備自由的生活空間，讓學生在試驗空檔之餘能好好休息。同時，兩位老師的信任與支持，更使學生得以在分子生物、蛋白質分析、動物試驗、作物田間操作、及生物資訊領域，皆能涉略與精進，並學習獨立思考、執行、到最後完成的過程。師恩浩大，學生永銘於心。

論文初成，承蒙 嘉義大學園藝系洪進雄主任、臺南農業改良場王仕賢場長、宜蘭大學園藝系郭純德博士及臺大生物技術研究所林詩舜博士撥冗審閱，提供學生寶貴、精闢的建議，讓本論文更臻完善。在博士班求學期間，諸位老師在研究或生活上都給予最大的關懷，在此表達最誠摯的謝意。此外，在鑽研次世代定序技術時，有幸得到臺大農藝系 劉力瑜教授及生物產業機電系 陳倩瑜教授的指導協助，才得順利於此領域發揮，於求學時也常受到兩位老師的關懷及鼓勵，學生銘感五內。

感謝一研的所有夥伴，慧欣學姊在研究及生活上的提攜照顧；學長姊 凱婷 旭佑宜華哲哲，感謝你們的協助，那些一起奮戰的夜晚此生難忘；感謝已畢業的成員 秀豪芹籃敏婷均慈韻巍州育暄弘珊羽霖玲瑋翰琳齊辰茹萍樺及現任室友 柔於誌蓁菁倫，感謝你們的付出，讓我能專注於論文。同時，煜恆、姵華、佩玉、旻翰、晏安、佳玲、偉翔、家吟、偉順、良州、靜樺、香誼、展麒、陳薇、璧甄、至緣、彥彤、兆倫及阿嬌學姊，從臺南場的採收、調查到臺大園藝分場試驗的協助，讓我備感窩心。亦感謝技師王國卿先生及已離開加工館的淑幸姊、晶晶姊與璨哥平日辛勞，將實驗室管理的井然有序，使我能專心於研究與學習。

感謝這九年來一直當我同學的彥惠，不論是在學習上或是待人處事上，從妳身上我學到很多，由衷感激。感謝我親愛的家人的扶持與協助，特別是母親，一直作為我最大的後盾支持著我。最後，希望這份成果能獻給最愛的父親，對您只有無盡的感謝與遺憾，希望您能一同分享這份喜悅。

洪志良　謹誌於

國立臺灣大學園藝暨景觀學系

中華民國一百零五年八月十六日

doi:10.6342/NTU201603060

摘要

對臺灣夏作甘藍生產來說，颱風過後所帶來的高溫淹水是一主要逆境。此逆境會造成嚴重之生理障害進而導致產量的損失。然而，高溫淹水逆境所造成的生理障害之分子機制尚未明瞭。本論文研究目的是釐清高溫淹水逆境如何影響甘藍轉錄，同時找出甘藍中有助於增加高溫淹水耐受性之基因組。首先，利用次世代定序技術來調查 25 或 35˚C 有/無淹水處理的八週大之甘藍全轉錄體表現。其中 2,040 個基因被選作分析目標，並用其對數倍率變化來進行差異表現分析。透過階層式分群法，發現具高溫淹水逆境特異性之受 WRKY 誘導而向上調節的 ACC 氧化酶 1，其可能為降低甘藍‘夏峰一號’逆境耐受性的關鍵因子之一。而透過基因本體富集分析，結果顯示在高溫處理組的基因本體富集詞彙與高溫淹水處理組接近，但在個別處理組中仍有獨特的基因本體詞彙。為了更進一步了解逆境下甘藍之共功能網路，以 AraNet v2 預測高溫淹水逆境處理之甘藍的共表現網路模組。在預測出的 7 組共表現模組中，其中兩組與離層酸訊息傳遞及滲透壓逆境耐受性有關的模組表現向下調控，提供了甘藍‘夏峰一號’不耐高溫淹水逆境的證據。再者，利用次世代定序比較高溫淹水逆境處理之逆境耐性及逆境敏感甘藍栽培品種的轉錄體表現，預期將找到受高溫淹水逆境影響之代謝路徑及基因組。此研究植物材料使用了具高溫淹水耐受性的‘228’，及敏感性的‘Fuyudori’。逆境處理則是在生長箱中進行，取樣時間點為處理後 0、6、12、及 24 小時。藉由時程試驗，並結合階層式分群法進行初級表現測量與權重共表現基因網路分析，透過兩種不同生物資訊學方法來進行數據分析。根據此方法，256 個具有最顯著表現差異的基因被確認，並建構了 13 個與高溫淹水逆境相關之共表現模組。最後，結果顯

doi:10.6342/NTU201603060

示在‘228’中高溫淹水逆境耐受性與酚類化合物生合成高度相關，而失控的缺水逆境可能是造成‘Fuyudori’不耐受高溫淹水逆境的關鍵因子之一。這些數據顯示高溫淹水逆境對甘藍代謝及調節路徑之影響。數個具逆境特異性之模組能與次級代謝物的累積、離層酸訊息的傳遞、及熱逆境因子與熱休克蛋白的向上調控做連結。上述這些機制也許提供了甘藍在面對劇烈變化的高溫淹水逆境下適合的代謝適應力維持生理反應的靈活對策。

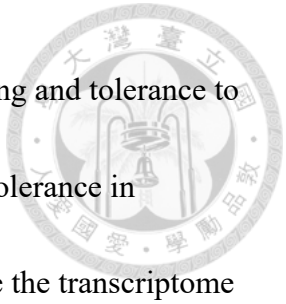
關鍵字：次世代定序、轉錄體、甘藍(*Brassica oleracea* var. *capitata L.*)、淹水、高溫、逆境、權重共表現基因網路分析。

doi:10.6342/NTU201603060

Abstract

Waterlogging at high temperature is a major stress after typhoon to the cabbage production during summer in Taiwan. This stress brings in serious physiological disorder and results in yield loss in cabbage. However, the molecular mechanisms of the physiological disorders under waterlogging stress at high temperature remain unclear. This thesis aims to identify how the waterlogging stress at high temperature (HWS) influences the cabbage transcriptome and to discover the gene sets which contribute to the tolerance of HWS in cabbage. First, RNA-seq was used to investigate the whole transcriptome of eight-week-old cabbage 'Shia Feng No. 1' treated with or without waterlogging both at 25 or 35°C. Log_2 fold change value in selected 2,040 genes was used to discriminate differentially expressed genes (DEGs). By hierarchical clustering, WRKY-induced up-regulation of ACC oxidase 1 was specifically found in HWS treatment, which to be one of the key factors that caused decreased stress tolerance in cabbage 'Shia Feng No.1'. According to gene ontology (GO) enrichment analysis, the enriched GO terms in heat treatment were close to HWS treatment; however, there were still unique GO terms enriching in each treatment. To further understand the co-functional networks in cabbages exposed to stress, AraNet v2 was used to predict co-expression network modules of HWS-treated cabbages. In the 7 predicted co-expression

doi:10.6342/NTU201603060



modules, the down-regulation of two modules related to ABA signaling and tolerance to osmotic stress in plants may provide the evidence about the HWS intolerance in cabbage. Next, next generation sequencing was employed to compare the transcriptome of stress-tolerant cultivar '228' and stress–intolerant cultivar 'Fuyudori' under HWS, which were used to find HWS-influenced metabolic pathways and gene sets. Stress treatment was performed in growth chamber at 35°C for 24 h, and sampling was performed at 0, 6, 12, and 24 h after treatment. A time-course RNA-seq analysis was performed and combined two different bioinformatic methods, primary co-expression measure with hierarchical clustering and weighted correlation network analysis (WGCNA), for analyzing the transcriptome data. 256 most significantly changed genes were identified and 13 coexpression modules associate to HWS were constructed. Finally, comparative analysis showed HWS tolerance highly linked to phenolic biosynthesis in '228', and uncontrollable water deprivation may be one of the key factors to cause HWS-affected in 'Fuyudori'. These data show how HWS influences the metabolic and regulatory pathways in cabbages. Several stress tolerance-specific gene modules were linked to the accumulation of secondary metabolites, transduction of ABA signaling, and up-regulation of heat stress factors and heat shock proteins. These

doi:10.6342/NTU201603060

may provide cabbage a flexible strategy tolerant to cope with HWS by offering appropriate metabolic adaptability under the dramatically changing environment.

Keywords: next generation sequencing; transcriptome; *Brassica oleracea* var. *capitata L.*; waterlogging; high temperature; stress; WGCNA.

doi:10.6342/NTU201603060

Content

doi:10.6342/NTU201603060

doi:10.6342/NTU201603060

doi:10.6342/NTU201603060

List of tables

List of figures

doi:10.6342/NTU201603060

Figure 9. Sample clustering and the expression profile of the 265 significantly expressed genes (FDR<0.1) (countinous). 58

doi:10.6342/NTU201603060

doi:10.6342/NTU201603060

List of Abbreviations

1-MCP	1-methylcyclopropene
ABA	abscisic acid
ABF	abscisic acid responsive elements-binding factor
ABI	abscisic acid-insensitive
ACC	1-aminocyclopropane-1-carboxylate
ACO	1-aminocyclopropane-1-carboxylate oxidase
ADH	alcohol dehydrogenase
ADK	adenosine kinase
AFP	ABI five binding protein
ANP	anaerobic polypeptide
APX	ascorbate peroxidase
ASH	ascorbic acid
ATP	adenosine triphosphate
AVG	aminoethoxyvinylglycine
bp	base pair
CAT	catalas
cDNA	complementary DNA
CDS	coding sequence
ClpB	caseinolytic peptidase B
COA	Council of Agriculture
COS	conserved orthologous sequence
DEG	differentially expressed genes
DHAR	dehydroascorbate reductase
DNA	deoxyribonucleic acid
EMBL-EBI	European Molecular Biology Laboratory- European Bioinformatics Institute
ERBP	ethylene responsive binding protein
EV	expression variability
FDA	Food and Drug Administration
FDR	false discovery rate
GA	gibberellic acid
GCL	glutamate-cysteine ligase
GM	genetically modified
GO	gene ontology
GO-BP	gene ontology-biological process

doi:10.6342/NTU201603060

GOPX	guaicol peroxidase
GPX	glutathione peroxidase
GR	glutathione reductase
GSH	glutathione
GST	glutathione-S-transferase
H	heat treatment
HRE	hypoxia-response element
HSF	heat stress factor
HSG	heat shock gene
HSP	heat shock protein
HSR	heat stress response
HWS	waterlogging stress at high temperature
IAA	indole-3-acetic acid
ISR	intron-spanning region
KRT	keratin_ type II cytoskeletal 1-like
LDH	lactate dehydrogenase
LEA	late embryogenesis abundant protein
LRT	likelihood ratio test
MDHAR	monodehydroascorbate reductase
MPK	mitogen-activated protein kinase
mTERF	mitochondrial transcription termination factor family
MYB	myeloblastosis
NADPH	nicotinamide adenine dinucleotide phosphate
NGS	next generation sequencing
ORF	open reading frame
PABP	polyadenylate-binding protein
PAL	phenylalanine ammonia lyase
PCA	principal components analysis
PDC	pyruvate decarboxylase
PE	paired-end
PEG	polyethylene glycol
PME	pectin methylesterase
PME	pectin methylesterase
PMEI	pectin methylesterase inhibitor protein
PMEI	pectin methylesterase inhibitor
qPCR	quantitative real time polymerase chain reaction

doi:10.6342/NTU201603060

R	R language
RBOHD	respiratory burst oxidase homolog D
Refseq	NCBI Reference Sequence Database
RNA	ribonucleic acid
RNase	ribonuclease
RNA-seq	RNA sequencing
ROS	reactive oxygen species
SEA	singular enrichment analysis
SNP	single nucleotide polymorphism
SOD	superoxide dismutase
SSR	simple sequence repeat
SUS	sucrose synthases
TAIR	The Arabidopsis Information Resource
TCA	tricarboxylic acid
TOM	topological overlap mapping metric
W	waterlogging treatment
WGCNA	weighted correlation network analysis
zf-HD	zinc finger homeodomain

doi:10.6342/NTU201603060

Chapter 1 Introduction

1.1 Heat and waterlogging stresses in crop

Abiotic stress is one of the important factors, which causes the limit of productivity on numerous crop growth worldwide. For crops and plants, different environment stimuli cause complex and varied responses. Meanwhile, the stress responses of crops to are dependent on multiple factors, such as the species, developmental stage, or duration of stimulus (Gan *et al.* 2004, Goulas *et al.* 2006). Immotile plants can only adapt to the environmental stimuli and maintain growth through modifying physiology and metabolism.

Heat stress.

High temperature-induced heat stress is a problem in agricultural cultivation for many crops. Raising temperature beyond the threshold of the crop for a duration, which results in the irreversible damage on crop physiology, is defined as heat stress (Wahid *et al.* 2007). The morpho-anatomical, physiological, and biochemical changes are often observed on heat stress-suffered plants. These changes always result in the damages on plant growth and development and further cause loss in economic yield (Wahid *et al.* 2007). For crop physiology, heat stress always causes instability of various compounds

doi:10.6342/NTU201603060

and cell structure such as proteins, RNAs, membranes, and cytoskeleton (Wahid *et al.* 2007). Heat stress induces the limitation of crop growth because of decreased carbohydrate accumulation, leakage of ion, production of toxic compounds and reactive oxygen species (ROS) (Almeselmani *et al.* 2006, Gill & Tuteja 2010, Wahid *et al.* 2007). Heat stress-induced reduction in leaf water status and elevation of oxidative stress, resulting in the growth inhibition and chlorosis in summer-sown mungbean (Kumar *et al.* 2011). Furthermore, the efficiency of enzymatic reactions is greatly affected, which induced metabolic disorder (Mittler *et al.* 2012, Suzuki *et al.* 2012). Nevertheless, plants also sense heat stress through these physiological changes and initiate various reactions to adapt the stress. To sense the heat stress responses (HSRs), at least four putative sensors were proposed, including an inward calcium flux channel on plasma membrane, a histone sensor in the nucleus, and two unfold sensors in the endoplasmic reticulum and cytosol (Mittler *et al.* 2012). Due to the high conservation of HSRs in plant, multiple pathways, regulatory networks, and cellular compartments involved in HSRs and different pathways may have a different optimum temperature (Mittler *et al.* 2012). In plants, metabolic pathways are highly sensitive to temperature change. Heat stress may result in disruption of coordination between multiple pathways (Apel & Hirt 2004, Kumar *et al.* 2011, Mittler *et al.* 2012). For example, ROS, such as

doi:10.6342/NTU201603060

$^1O_2$, H_2O_2, O_2^-, and $HO^\bullet$, are byproducts produced from the metabolic imbalance and able to cause oxidative damage to proteins, DNA, and lipids (Apel & Hirt 2004). ROS belongs to the HSR-initiated signals and is produced by respiratory burst oxidase homolog D (RBOHD), a NADPH oxidase (Miller *et al.* 2007). The accumulation of ROS is necessary for being rapid, cell-to-cell propagating systemic signals, and the activation of RBOHD along the signal path is also essential for signal propagation (Miller *et al.* 2009). The accumulation of ROS induced by heat stress is also an oxidative stress in plants, so the stress tolerance is related with antioxidant defense mechanism. For scavenging of ROS, plants employ two antioxidant defense systems, enzymatic (superoxide dismutase, SOD; catalase, CAT; ascorbate peroxidase, APX; glutathione reductase, GR; monodehydroascorbate reductase, MDHAR; dehydroascorbate reductase, DHAR; glutathione peroxidase, GPX; guaicol peroxidase, GOPX and glutathione-S-transferase, GST) and non-enzymatic (ascorbic acid, ASH; glutathione, GSH; phenolic compounds, alkaloids, non-protein amino acids and α-tocopherols), to regulate the cascades of uncontrolled oxidation and reduce oxidative damage in plant cells (Gill & Tuteja 2010, Hasanuzzaman *et al.* 2013). In addition to antioxidant defense, plants up-regulate heat shock genes (HSGs) which encode heat shock proteins (HSPs) to help them survive under heat stress. There are five well-

doi:10.6342/NTU201603060

characterized HSP families in plants, including HSP100 (or ClpB), HSP90, HSP70 (or DnaK), HSP60 (or GroE), and HSP 20 (or small HSP, sHSP) (Hasanuzzaman *et al.* 2013, Timperio *et al.* 2008). **Supplemental Tab. 1** shows the primary heat stress tolerance-related molecular function of five HSP families in plants. These findings suggested that the plants respond to heat stress with multiple strategies and a systemic screening is needed for studying of stress tolerance in plants.

Waterlogging.

Water involves in the most of physiological processes in plant and is one of the crucial resources in agricultural productivity. However, excess water could cause many physiological disorders and affect the growth of plant. Waterlogging is defined as the soil surface filled with free-standing water or the water saturation of soil at least 20% higher than the field capacity (Aggarwal *et al.* 2006). Soil waterlogging rapidly alters the soil physico-chemical properties such as soil redox potential, pH and O_2 level. Dependent on plant species, seasonal timing, and duration, waterlogging stress causes various levels of physiological disorders, even death. Generally, waterlogging stress induced physiological changes include decreased photosynthetic efficiency, low stomatal conductance, reduced leaf-water-potential and water-use efficiency, and

doi:10.6342/NTU201603060

modified level of phytohormones (Irfan *et al.* 2010, Ou *et al.* 2011). The primary problem of soil waterlogging in plants is O_2 restrictive conditions because the water saturates soil pores and reduces gas diffusion. Thus, conditions of hypoxia or anoxia usually result in the aerobic respiration of plant root switching to anaerobic fermentation (Parent *et al.* 2008). Actually, the activities of most enzymes involved in ethanoic fermentation, such as alcohol dehydrogenase (ADH) and pyruvate decarboxylase (PDC), are induced following hypoxia stress (Fukao & Bailey-Serres 2004). In plant, waterlogging induced oxygen deprivation causes the reduction of cellular adenosine triphosphate (ATP) supply that affects cellular metabolism and plant development. Utilizing alternative pathways/electron receptors to keep a continuous ATP supply is the common strategy for waterlogging-suffered plants (Parent *et al.* 2008). In addition, reduction in stomatal conductance, photosynthesis, and root hydraulic conductivity are also observed in plants following waterlogging. These physiological modifications then affect the membrane-transport activity and result in the reduction of cytosolic ions and carbohydrate translocation (Parent *et al.* 2008, Shabala *et al.* 2014). Moreover, the gene expression profile which is similar to above the physiological phenomena is observed in the transcriptome of waterlogging-treated cucumber (Qi *et al.* 2012). The differentially expressed genes (DEGs) of cucumber

doi:10.6342/NTU201603060

under waterlogging are mostly down-regulated and mainly involve in carbon metabolism, photosynthesis, reactive oxygen species generation/scavenging, and hormone synthesis/signaling. In addition to physiological changes, most studies have focused on the molecular mechanism of model plants such as *Arabidopsis* (Hsu *et al.* 2013). This study has shown that hypoxia strongly induced the expression of WRKY transcription factor and enhanced downstream ethylene biosynthesis. The increased ethylene further influenced other hormone biosynthesis like gibberellic acid (GA) to inhibit the hypoxia tolerance. On the other hand, hypoxia could induce the transcription factors such as *HRE1* and *HRE2* to mediate expression of hypoxia tolerance-related genes, including *ADH*, *PDCs*, and *SUSs* (Licausi *et al.* 2010, Yang *et al.* 2011). *HRE1* also was known as a negative regulator of ethylene responses to promote the hypoxia resistance. Meanwhile, the analysis of proteome and transcriptome on hypoxia stress-treated rice, tomato, and *Arabidopsis thaliana* reveals a temporary expression of more anaerobic polypeptides (ANPs) relate to starch-sucrose metabolism, glycolysis, tricarboxylic acid cycle (TCA) cycle, fermentation pathways and aerenchyma formation. Lots of hypoxia stress-induced proteins are found including heat shock factors, ethylene responsive binding proteins (ERBP), MADS-box proteins, AP2 domain, leucine zipper, zinc finger, WRKY factors (Lasanthi-Kudahettige *et al.* 2007,

doi:10.6342/NTU201603060

Loreti *et al.* 2005), NIM1-like protein2, DWARF1 protein, PAL, flavanone-3-hydroxylase mtATPase α-subunit (Ahsan *et al.* 2007) and fermentation enzymes (Irfan *et al.* 2010, Narsai *et al.* 2011). These studies mentioned above also reveals many genes, ANPs, and proteins of unknown function might regulate the survival-related mechanism in plants under waterlogging stress.

The combination of waterlogging and high temperature stress.

In general, single stress conditions such as high/low temperature, drought, waterlogging, or salinity have been studied intensively (Hirayama & Shinozaki 2010). Most studies focus on the effects of single stress to plants and simulate abiotic stress in controlled environment. However, most crops are cultured in the field and suffer a combination of different abiotic stresses (Jiang & Huang 2001; Lin *et al.* 2015; Wang *et al.* 2008). In East Asia, many crops encounter a combination of high temperature stress and waterlogging stress in summer due to typhoons. Short-term waterlogging and heat stress after typhoon is a common abiotic stress and always causes serious loss in crop productivity in summer. Although most abiotic stress had been well studied, the mechanism of different stress combination still remains unknown (Mittler 2006), especially waterlogging at high temperature. In Kentucky bluegrass, waterlogging stress

doi:10.6342/NTU201603060

at high temperature induced similar level of anaerobic metabolism-related enzymes, ADH, and lactate dehydrogenase (LDH) (Wang *et al.* 2008). Although, the leaf growth and root water-soluble carbohydrate are significantly reduced under HWS. In addition, HWS causes significantly reduction of chlorophyll content and accumulation of H_2O_2 in broccoli (Lin *et al.* 2015). However, there are few studies that focus on the HWS-related genes and the molecule mechanisms of HWS in plants.

1.2 RNA-seq in plant transcriptome

Next generation sequencing (NGS) is a new method of high throughput sequencing which combines template preparation, sequencing and imaging, and genome alignment and assembly methods. In contrast to the “first generation” technology-automate Sanger method, NGS has the advantage on producing enormous volume of data with cheaper cost (Metzker 2010). Now, NGS has many applications on sequencing of various organisms, including genomics, transcriptomics, metagenomeics, and epigenomics. Due to the development of NGS, the genome sequencing of many important crops which were used as model plants has finished in recent years. Meanwhile, NGS coordinates with new genetic mapping strategies, evolutionary analysis, optimized gene discovery, and the unitization of genetic variation to speed up the improvement of crop breeding

doi:10.6342/NTU201603060

(Morrell *et al.* 2012). More than 100 plant genome sequencing projects, which include 63% crop species, have finished and published (Michael & VanBuren 2015). Although the availability of plant reference genome is necessary for crop breeding and improvement, genome data alone is not enough for efficient crop breeding. Recently, many studies utilize NGS on transcriptome to clarify which combination of genes is conducive to crop stress tolerance.

The stress responses of plants have well-known involved in multiple regulation pathways and metabolic mechanisms. Recently, more and more studies use RNA-seq (RNA sequencing) to investigate the gene expression profile of plants under various stresses. In contract to traditional strategies on plant stress studies, NGS provides rapid gene expression screening of whole genome and multiple pathways for stress-suffered plants. In the study of polyethylene glycol (PEG)-induced osmotic stress or exogenous ABA on sorghum, the results reveal that two treatments have strong interplay among abscisic acid and 13-lipoxygenase, salicylic acid, jasmonic acid, and plant defense pathways (Dugas *et al.* 2011). They also discover more than 50 drought-responsive gene orthologs of uncharacterized function through linking with published transcriptome analyses on rice, maize, and *Arabidopsis* (Dugas *et al.* 2011). Kakumanu *et al.* utilize the transcriptome comparison to find the differential expressed genes relate to the

doi:10.6342/NTU201603060

signaling among increased ABA levels, decreased glucose levels, activation of programmed cell death/senescence through repression of a phospholipase C-mediated signaling pathway and arrest of the cell cycle in the drought-stressed maize (Kakumanu *et al.* 2012). Moreover, a study identities 728 simple sequence repeats (SSRs), 495 single nucleotide polymorphisms (SNPs), 387 conserved orthologous sequence (COS) markers, and 2088 intron-spanning region (ISR) markers from the transcriptome comparison of drought-treated and well-watered chickpeas (Hiremath *et al.* 2011). With the similar strategy of comparative transcriptome analysis, Yan *et al.* find not only heat response transcripts but also some genes involve in salt stress, organic acid metabolic and carotenoid metabolic in heat stress treated spinach leaf (Yan *et al.* 2016).

In addition to the transcriptome of stress physiology, some studies focus on the relation of stress and physiology development in plants (Mutasa-Göttgens *et al.* 2012, Min *et al.* 2014, O'Rourke *et al.* 2012). The study of global transcriptional responses in the shoot apex to vernalization and GA treatment reveals the important role of a RAV1-like AP2/B3 domain protein in vernalization and efflux transporters in Sugar beet (Mutasa-Göttgens *et al.* 2012). In high temperature sensitive-anther development stage of cotton, the results of transcriptome analysis reveal that the key regulators of the anther response to high temperature stress involve in multiple metabolic pathways,

doi:10.6342/NTU201603060

including sugar-regulated IAA biosynthesis and *GhCKI*-auxin signaling pathway (Min *et al.* 2014). Furthermore, O'Rourke *et al.* combine RNA-seq data and classic physiological experiments to investigate the biochemical and molecular mechanisms of Pi deficiency-induced cluster root development in white lupin (O'Rourke *et al.* 2012).

These studies mention above probe into the molecular mechanism of plants under stresses or specific development status with whole genomic gene expression analysis. These study strategies contribute to resolve the actual status of crops from the comprehensive perspective.

1.3 Overview of cabbage and its problem on production in Taiwan

Cabbage (*Brassica oleracea* var. *capitata L.*) is one of the important global and it is difficult to grow in the tropical and subtropical area in summer. In Taiwan, cabbage is classified in major vegetables with high annual production and favorite for people. According to the annual report of Council of Agriculture (COA) for 2014, the total planted area, total production quantity, and total value of cabbage in Taiwan are 8,428 ha., 340,853 m.t., and NT$3,728,197, respectively. Since the fact that cabbage belongs to temperate crop, the high temperature in Taiwan in summer hinders the cabbage production. There are some studies about heat tolerance of cabbages (Annamalai &

doi:10.6342/NTU201603060

Yanagihara 1999, Chauhan & Senboku 1996), cabbage breeding has been focused on heat tolerant traits for a long time. The major loss of cabbage is still resulted from the waterlogging at high temperature after typhoon. According to the statistics of crop product disasters, Typhoon Matmo alone caused approximately NT$ 315 million (approximately US$ 98 million) loss in cabbage in 2014 (COA, 2015). Although heat-tolerance cultivars of cabbage have been bred for a long time in Taiwan, waterlogging at high temperature still causes serious damages to heat-tolerant vegetable crop. Molecular mechanisms of hypoxia on most agricultural crops like cabbages are less to refer.

1.4 Aims

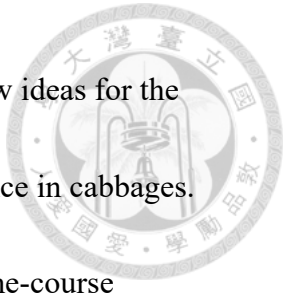
In Taiwan, the heavy losses in summer cabbage production is a serious problem due to the HWS following typhoon. Little research has focused on HWS in plants, cabbage grown in summer is an ideal material for clarifying the molecular mechanism of HWS in crop. In addition, comparison among cabbage cultivars to HWS-tolerance has been studied for two years (Chen, 2011). It is the hypothesis that cabbage cultivar with HWS-tolerance has significantly different gene expression pattern for adapting to the HWS in contrast to HWS-intolerant ones. Hence, the core aims of this HWS-suffered cabbage transcriptome sequencing project are:

doi:10.6342/NTU201603060

1. To explore the regulatory pathways involving in the stress combination of high temperature and transitory waterlogging stress from the transcriptome of cabbage in contrast to the stress of high temperature alone or waterlogging alone.
2. To verify the hypothesis that tolerance and intolerant cultivars perform different molecular mechanism to response of HWS.
3. To provide the information of HWS tolerance-related gene sets and regulatory pathways.

In order to address these aims, two separate experiments were designed and performed. First, to identify different regulatory pathways employed in cabbage among the stresses of heat alone, waterlogging alone, and the combination of both heat and waterlogging. Using RNA-seq, the transcriptome of the leaves from *B. oleracea* var. *capitata L.* 'Shia Feng No.1', a commercial cultivar with tolerance to high temperature or waterlogging at ambient temperature (25°C), but waterlogging-intolerant at 35°C was analyzed. The cabbage transcriptomes in the treatments of waterlogging, high temperature, and waterlogging at high temperature were compared with the group without waterlogging treatment at 25°C. The key gene sets which resulted in intolerance to the combination of heat and waterlogging stresses were investigated. This study

doi:10.6342/NTU201603060



provided new insights into the combination of stress and suggests new ideas for the strategies of breeding or cultivation plan to improve the stress tolerance in cabbages. Second, using comparative analysis of two cabbage cultivars with time-course experiment was used to identify major gene sets and regulatory pathways which associated with HWS tolerance trait. RNA-seq data from tolerant and intolerant cabbage cultivars treated with waterlogging at 35°C for 0, 6, 12, and 24 h were used for WGCNA of gene expression. Through building the modules of highly correlation genes from correlated network, the module eigengene or an intramodular hub gene of stress-tolerant traits can be obtained. Finally, this study is expected to integrate all the results and summarize the cabbages response strategies to waterlogging stress at high temperature.

doi:10.6342/NTU201603060

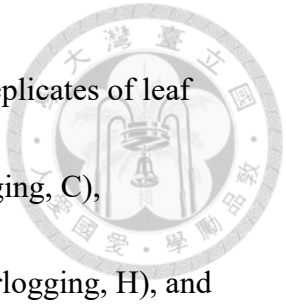
Chapter 2 Transcriptome analysis of heat tolerant cabbage under high temperature and/or waterlogging stress using RNA-seq

2.1 Materials and Methods

2.1.1 Plant materials and growth conditions

Cabbage (*B. oleracea* var. *capitata L.*) 'Shia Feng No.1' (purchased from Nong-Seng seed, Taiwan) F1 seeds were rinsed in sterile water, and allowed to soak overnight in the dark in a petri dish with a moist filter paper. Germinating seeds were planted one seedling per cell in 72-cell plug trays containing a mixture of 75% peat moss and 25% perlite. Plug trays were placed in growth chambers at 23°C/18°C with a 16-h photoperiod (300 μmol photos m^{-2} s^{-1}) for three weeks. Consequently, 3-week-old 'Shia Feng No.1' seedlings were transplanted in 8-inch pots containing a mixture of 75% coconut fiber and 25% perlite. Pots were placed in a plastic greenhouse. All plants were watered twice per day with automatic drip irrigation system for 5 weeks, applied as a foliar spay of 1/300x Foliar Nitrophoska® 20-19-19 (COMPO, Germany) with 1/1000x Tween-20 every week. After sowing for 8 weeks as describe above, 'Shia Feng No.1' plants were treated w/o waterlogging stress at 25°C or 35°C in a phytotron for 12 hours. Waterlogging stress was endured by soaking pots in a tank and maintaining 5 cm of

doi:10.6342/NTU201603060



standing water from the top of the culture medium. Four biological replicates of leaf samples (4 pots per treatment) in the control (25°C without waterlogging, C), waterlogging (25°C with waterlogging, W), heat (35°C without waterlogging, H), and heat-waterlogging (35°C with waterlogging, HWS) treatments were collected, soaked in liquid nitrogen, and stored in -80°C for future analysis.

2.1.2 RNA extraction, cDNA library preparation, and RNA-seq

Total RNA was purified from 16 leaf samples (4 treatments × 4 biological replicates) in 'Shia Feng No.1' using Plant Total RNA Miniprep Kit (GeneMark). Total RNA samples were quantified using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies), and the RNA integrity was further determined using RNA600 Nano assay with the Agilent 2100 Bioanalyzer (Agilent Technologies). After quantification and qualification, the total RNA samples of 'Shia Feng No.1'were sent to the Beijing Genome Institute (BGI; Shenzhen, China), and the four replicates of total RNA samples in each treatment were mixed equally into one for further cDNA library preparation (1 library per treatment) and then sequencing using Illumina HiSeq2000. All the cDNA libraries were constructed using selected poly(A) mRNAs following Illumina Truseq RNA-seq library preparation procedure.

doi:10.6342/NTU201603060

2.1.3 *De novo* assembly, annotation, and differential gene expression analysis of RNA-seq data

The quality control of ninety-one-base pair (bp) paired-end (PE) raw reads was performed using FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and then *de novo* assembled using CLC Genomics Workbench (https://www.qiagen-bioinformatics.com/) with the pipeline shown in **Fig. 1**. After generating the contig sequence, tBlastx was employed to annotate these contigs with The *Arabidopsis* Information Resource (TAIR) and the European Molecular Biology Laboratory- European Bioinformatics Institute (EMBL-EBI) cds databases. After removing the undefined contigs, the contigs with annotated gene ID and contig length was filtered to retain the longest contigs within the contigs with same annotated ID. This process was ensured the reads of specific genes did not disperse to different contigs, which may generate the mistake while calculating gene expression value. The TAIR ID was employed to import into AgriGO, which is a GO Analysis Toolkit and Database for Agricultural Community, and AraNet v2, a probabilistic functional gene network for *Arabidopsis thaliana*, for GO enrichment analysis and co-functional

doi:10.6342/NTU201603060

network analysis (Du *et al.* 2010, Lee *et al.* 2015). The module network was exported from AraNet v2 and visualized using Gephi v0.9.1 (https://gephi.org/).

For determining the expression value of contigs, the read count based methods combining Bowtie2, eXpress, and DESeq2 were selected for remapping the raw reads to filtered contigs (Forster *et al.* 2013, Langmead & Salzberg 2012, Love *et al.* 2014). For analyzing with likelihood ratio test (using nbinomLRT), the *p* value was indicating the significance of DEGs. Gene with a *p* value ≤ 0.1 and the $\log_2$ fold change between treatment and control ≥ 1 in at least one of all the treatments were used as the cut-off to evaluate significant differences in gene expression.

2.1.3 Validation of RNA-seq data by qPCR

A total of 4 DEGs were selected to confirm the differential expression of RNA-seq data by qPCR. This analysis was performed on a new set of 3 biological replicates for each treatment. The total RNA was isolated from leaf materials which was grinding with liquid nitrogen using GENEzol^{TM} TriRNA Pure Kit (Geneaid, Taiwan). After determining the quality and concentration of the RNA with NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies), 2 μg total RNA in 20 μL RNase-free water was used as template to synthesize the cDNA following the instruction of the

doi:10.6342/NTU201603060

RNA to cDNA EcoDry^{TM} premix (Clontech) and using the 2700 thermal cycler (Applied Biosystems). The cDNA was stored at -20℃ until next procedure.

The gene-specific primers for qPCR were designed with Primer 3 (http://bio-info.ut.ee/primer3/) **(Supplemental Tab. 2)** and selected according to the abundance of read mapping determined by Bowtie2 and Integrative Genomics Viewer (IGV) v2.3 (http://software.broadinstitute.org/software/igv/node/250). For validation of the RNA-seq results, the selected transcripts were performed in three replicates using the iQ SYBR Green Supermix protocol (Bio-Rad) on MyiQ MyiQ™ Single-Color Real-Time PCR Detection System (Bio-Rad) following the manufacturer's recommendations. qPCR was performed in a 20-μL reaction mix containing 10 μL iQ SYBR Green Supermix (2×), 8 μL distilled, deionized water, 0.5 μL forward primer (10 pmol μL^{-1}), 0.5 μL reverse primer (10 pmol μL^{-1}), and 1 μL template cDNA (75ng). The PCR conditions were as follows: 3 min of predenaturation at 95°C, 50 cycles of 15 s at 95°C, and 30 s at 55°C, followed by steps for the dissociation curve generation (60 s at 95°C, 60 s at 55°C, and 10s of 58°C). The iQ5 software (Bio-Rad) was used for data collection. Relative transcript levels for each sample were obtained using the comparative cycle threshold method (refer) and the cycle threshold value of the actin gene for each sample as a reference gene.

doi:10.6342/NTU201603060

2.2 Results

2.2.1 *De novo* assembly and annotations

To study the gene expression profile of cabbage under waterlogging stress at high temperature, RNA-seq was performed in cabbage by four different treatments. RNA samples were obtained from leaves in 8-week-old cabbage plants at 12 h post stress treatments of waterlogging (W), heat (H), and waterlogging stress at high temperature (HWS). The plants grown in normal condition were regarded as controls (N). Each treatment had three independent biological replicates, and all RNA of these replicates was mixed (with equal RNA amount in each replicate) into one RNA sample for library preparation. In total, 4 separate libraries were obtained for sequencing (**Supplemental Tab. 3A**). After sequencing and adapter trimming, approximately 215 million reads were obtained, each read was 91 nucleotides long. These clean reads were firstly merged into two fasta files according to the orientation for de novo assembly with CLC Genomics Workbench (**Fig. 1**). Finally, CLC assembly successfully generated 83,910 contigs which cover 95.4% reads and the N50 was 1,115 nt (**Supplemental Tab. 3B**).

doi:10.6342/NTU201603060

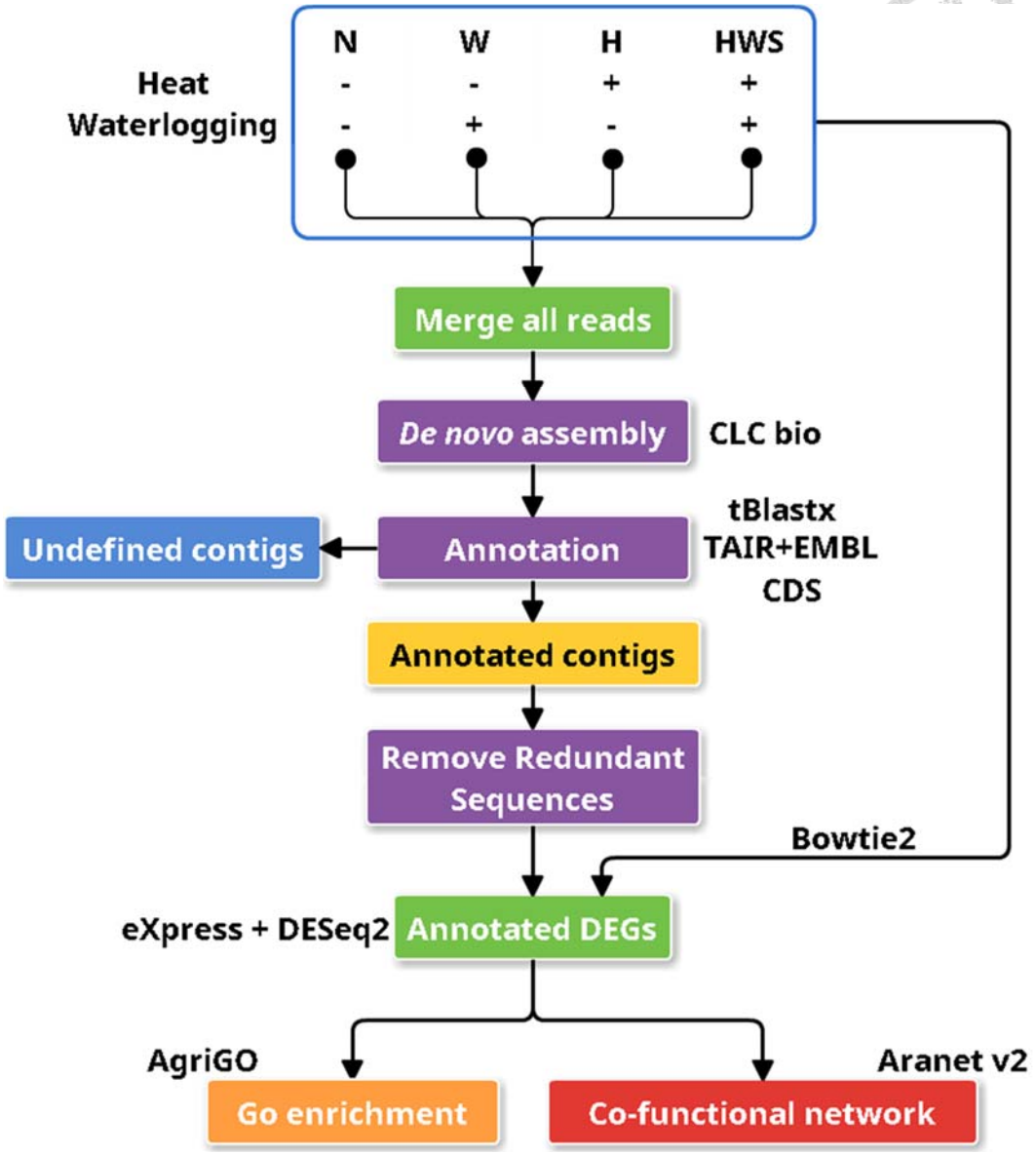


Figure 1. Workflow of whole transcriptome analysis. *B. oleracea* var. *capitata L.* 'Shia Feng No.1' in the stress treatments of waterlogging (W), high temperature (H), and waterlogging stress at high temperature (HWS). Normal (N) sample represented the plants that were cultured at 25°C without waterlogging treatment as control. Reads were generated form each sample with different treatments. First, the reads first were merged to perform *de novo* assembly using CLC. All assembled contigs were annotated with the coding sequence databases of TAIR and EMBL through tBLASTx. After annotation, reads from each treatment were mapped to annotate DEGs with Bowtie2, respectively. Expression value of the annotated DEGs was analyzed with eXpress and DESeq2. AgriGO was used to map DEGs with GO term and enriched. Aranet v2 was used to perform the enrichment analysis of MapMan bins and search for co-functional network.

doi:10.6342/NTU201603060

The *de novo* assembled contigs were analyzed using tBLASTx analysis against the *Arabidopsis thaliana* coding sequence (CDS) database (TAIR 10) and EMBL CDS database. Because cabbage and *Arabidopsis* belong to Brassicaceae, the best hit of the BLAST result against TAIR database was first selected as the annotation for the aligned contigs. For a total of 83,910 contigs, 63.33% (53,146) contigs were aligned to TAIR and EMBL databases. However, some annotated contigs were found that had high similarity and aligned to the same reference gene name. According to the sequence length and identity, longest contig as the reprehensive sequence was kept and other redundant contigs were removed. Finally, 19,548 annotated contigs were obtained as unigenes from the RNA-seq assay and these unigenes were used as cabbage transcriptome reference to perform subsequent analysis.

2.2.2 Identification of differentially expressed genes in stress treatments

For each treatment, read count of each gene was used to estimate the transcript abundance. Bowtie2 and eXpress were used for read mapping and count estimation. Then, DESeq2 was used to perform the differential analysis of read counts and quantitative analysis for each gene. For subsequent analysis, 2,040 genes (https://www.space.ntu.edu.tw/navigate/s/53232D8912634486952246360 6B47171QQY) were

doi:10.6342/NTU201603060

selected and then a cutoff of $\log_2$ fold change ≥ 1or ≤ -1 was used for identifying DEGs. DEGs were abundant in H and HW treatments, 61% and 66%, respectively. Only 9.6% DEGs were identified in W treatment. Moreover, using Venn diagram showing the distribution of these DEGs highlights that 200 up- and 788 down-regulated DEGs are commonly identified in H and HWS treatments (**Fig. 2A**). This result reveals that the gene expression profile of cabbages under waterlogging stress at high temperature was similar to the plant under heat stress. As shown in **Fig. 2B**, the expression profile of 2,040 genes among all stress treatments implied the same pattern. Most DEGs in HWS treatment had the same expression pattern as H treatment, and some DEGs were as W treatment. However, in HWS treatment, the expression pattern of 3 DEGs (BOHW43287 encoding ACC oxidase 1; BOHW57430 encoding a WRKY family transcription factor, and BOHW52477 encoding a Kinase-related protein of unknown function) were distinguished from other two treatments (**Fig. 3**). WRKY family transcription factor, especially WRKY22, is well-known as the regulator of ethylene biosynthesis during flooding (Hsu *et al.* 2013). Unlike W and H treatments, WRKY-induced up-regulation of ACC oxidase 1 could be one of the key factors that caused decreased stress tolerance in cabbage under HWS stress.

doi:10.6342/NTU201603060

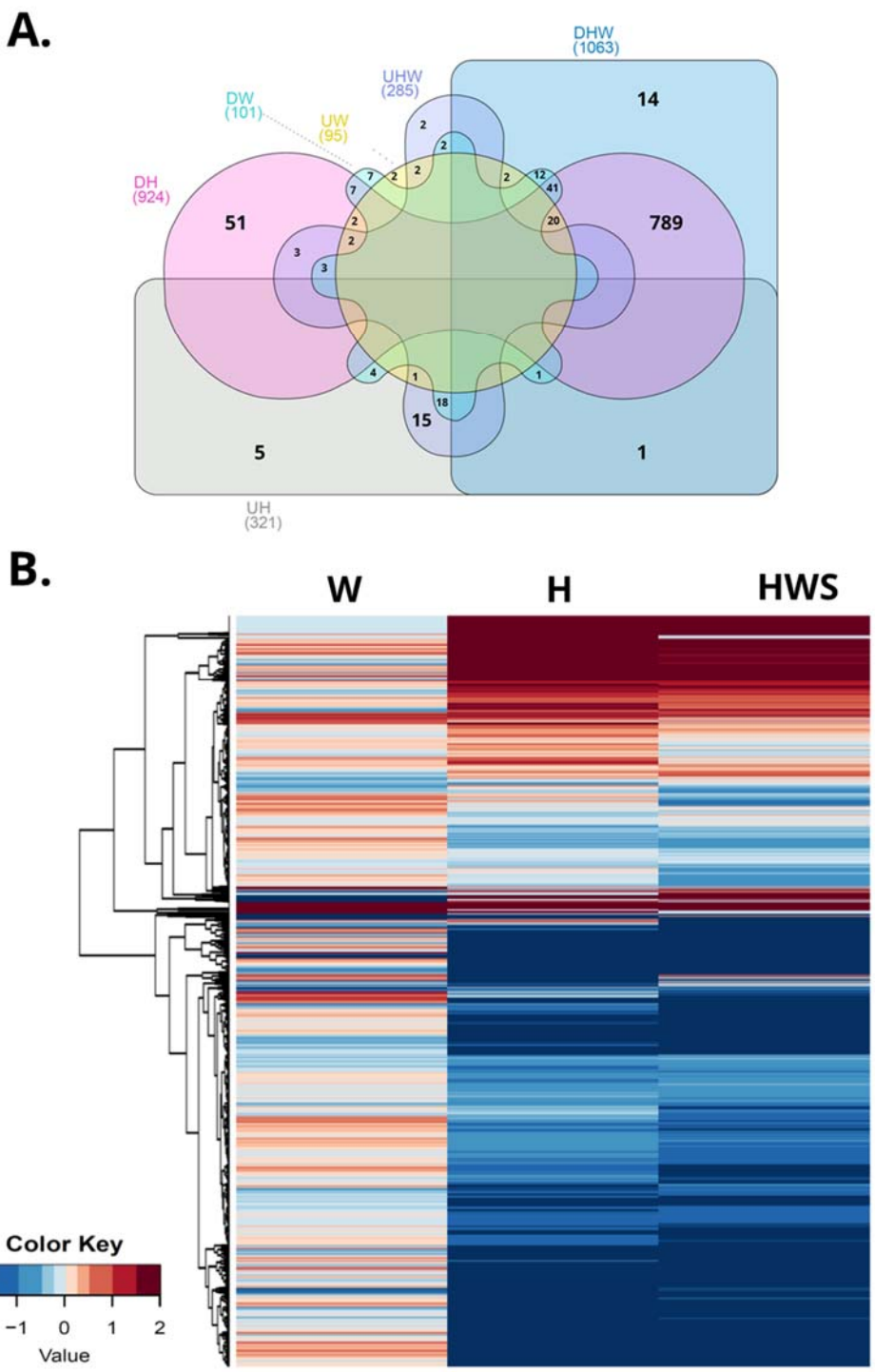


Figure 2. Analysis of differentially expressed genes. (A) Venn diagram of the 2,040 DEGs in waterlogging (W), high temperature (H), and waterlogging stress at high temperature (HWS) treatments. (B) Heat map depicting $\log_2$ fold change of 2,040 DEGs in three treatments. Deep Red and Blue colors indicate up- and down- regulated genes, respectively. *P* value < 0.1 was used as the cut-off to evaluate significant differences in gene expression.

doi:10.6342/NTU201603060

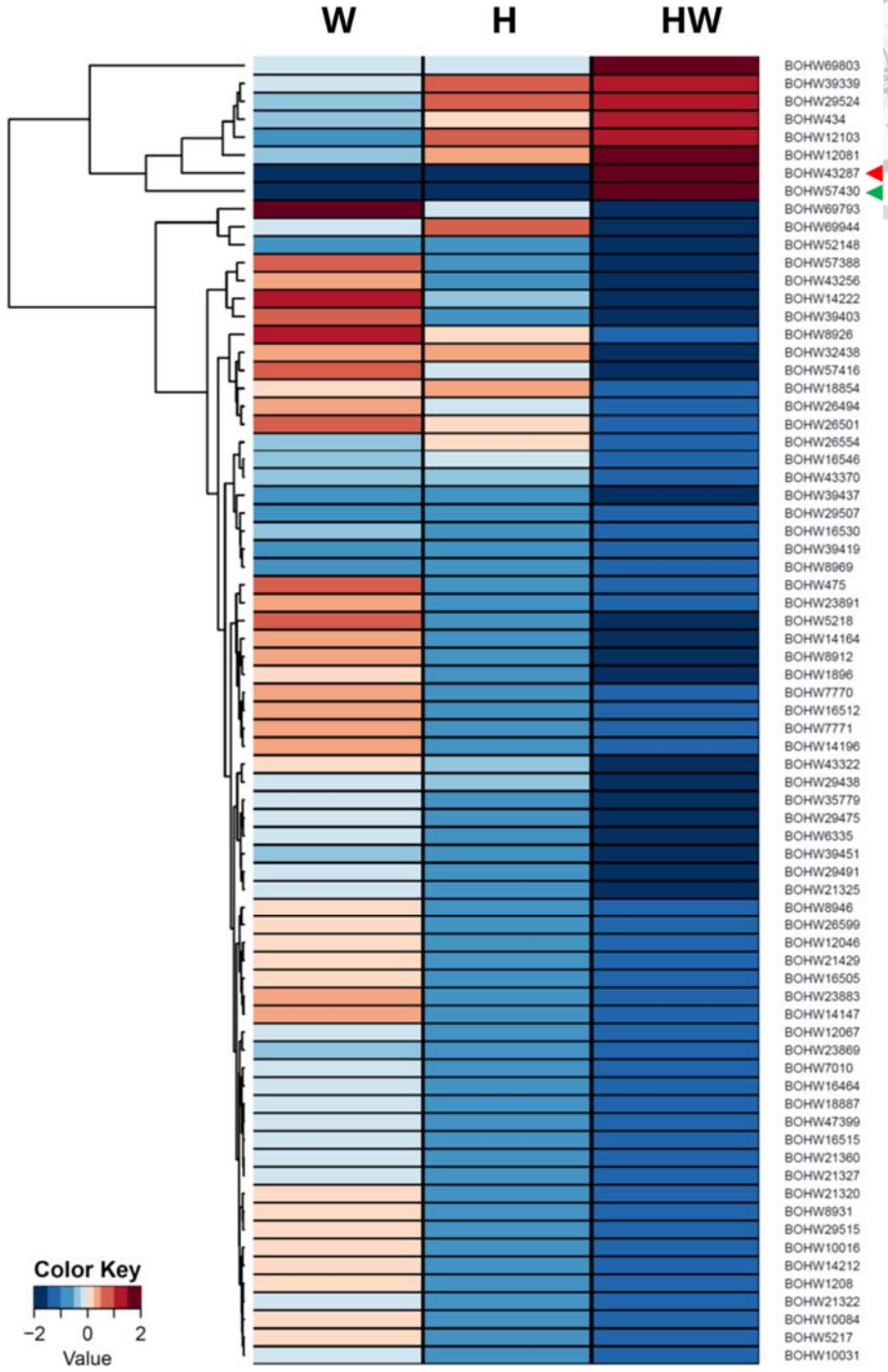


Figure 3. Heat map of DEGs in three stress treatments. Heat map depicting $\log_2$ fold change of DEGs in waterlogging (W), high temperature (H), and waterlogging stress at high temperature (HWS) treatments. Deep Red and Blue colors indicate up- and down-regulated genes, respectively. *P* value < 0.05 and the expression value of $\log_2 \geq 1$ at least one of all the treatments was used as the cut-off to evaluate significant differences in gene expression. Red and green arrow was indicated ACC oxidase 1 and WRKY family transcription factor, respectively.

doi:10.6342/NTU201603060

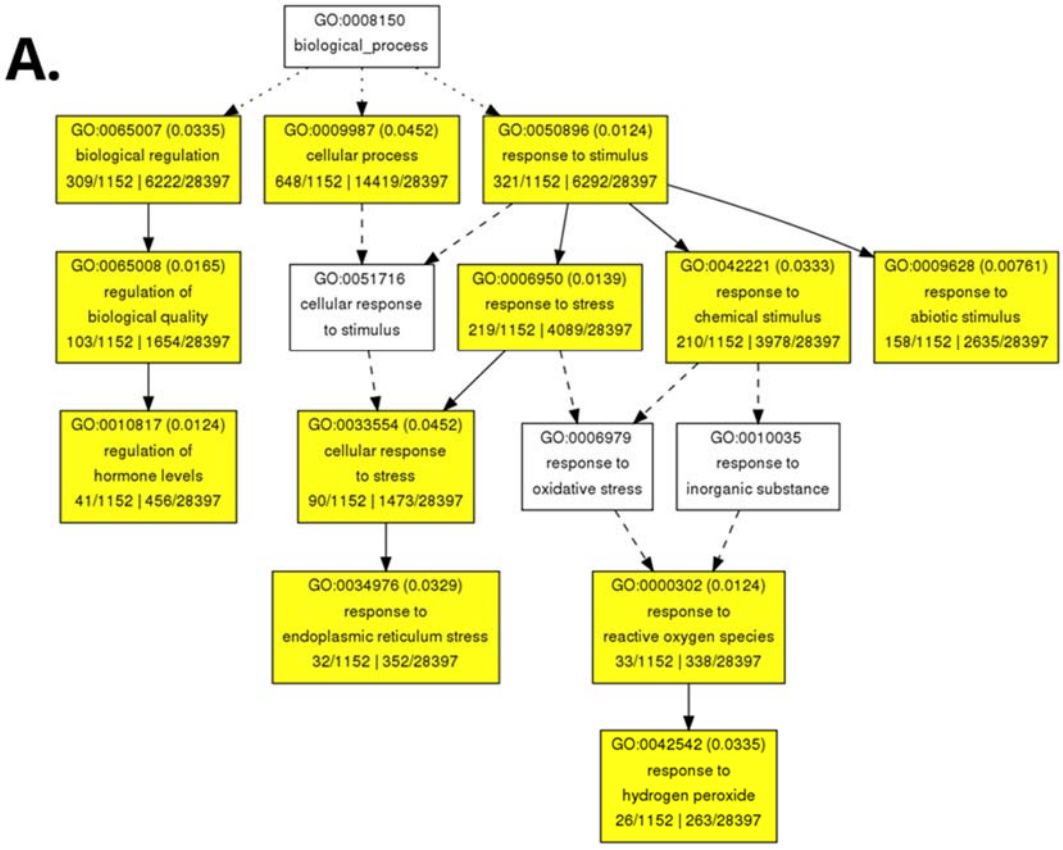
2.2.3 Singular enrichment analysis of GO for DEGs

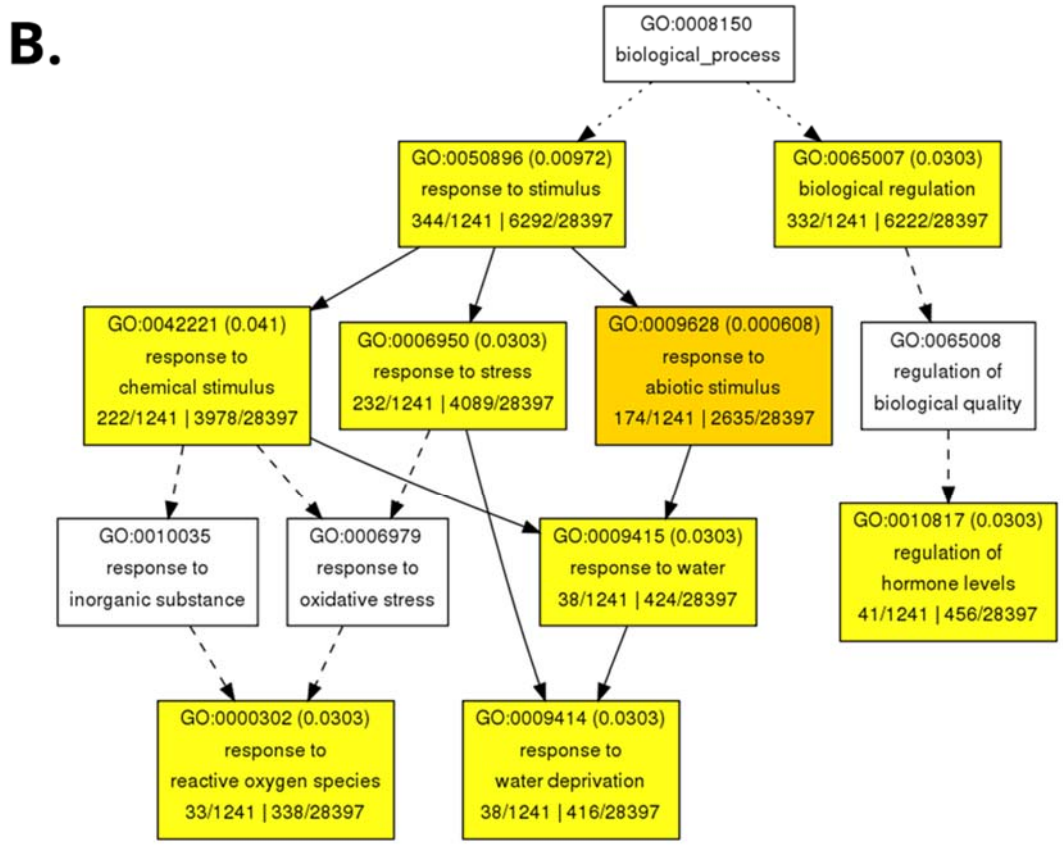
To understand the DEGs-involved biological process, AgriGO (Du *et al.* 2010) was used to assign the GO term for DEGs in each treatment and identify the enriched biological process category. Singular enrichment analysis (SEA) was performed using Fisher statistical test and false discovery rate (FDR) of 0.05 as the cutoff with complete GO. One thousand one hundred and fifty-two (92.5%) and 1,241 (92.1%) DEGs in H and HWS, respectively, were annotated to at least one of the cellular component, biological process, or molecular function GO categories. Whereas, the DEGs in W treatment were too less to enrich any categories. In both H and HWS treatments, 'Response to abiotic stimulus' (GO:0009628) was the most significantly enriched GO term in biological process (GO:0008150) (**Fig. 4**). Meanwhile, the enrichment of genes involved in 'response to stimulus' (GO:0050896), 'response to stress' (GO:0006950), 'response to reactive oxygen species (ROS)' (GO:0000302), 'response to chemical stimulus' (GO:0042221), and 'regulation of hormone levels' (GO:0010817) were found in both H (**Fig. 4A**) and HWS (**Fig. 4B**) treatments in this SEA. Additionally, SEA uniquely revealed the enrichment of genes involved in 'response to endoplasmic reticulum stress' (GO:0034976) and 'response to hydrogen peroxide' (GO:0042542) in H treatment. This result suggested that high temperature stress influenced the expression

doi:10.6342/NTU201603060

of transcription factor, kinase, receptor on the endoplasmic reticulum, and hydrogen peroxide-related signal transduction. However, more GO terms related to ‘response to water deprivation’ were found in HWS treatment, implying that the plants faced the combination of heat and water deficit stresses. Unlike individual treatment of waterlogging or high temperature, the combination of these two stresses induced unsuitable gene expression to lose the stress tolerance in cabbages.

doi:10.6342/NTU201603060





doi:10.6342/NTU201603060

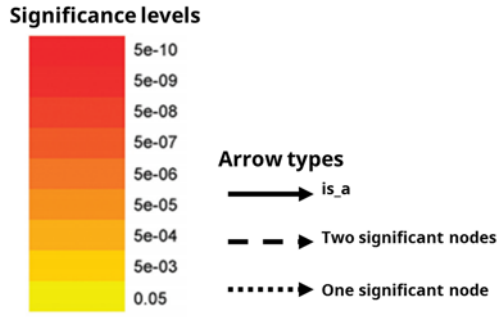


Figure 4. Singular enrichment analysis of DEGs with AgriGO. Hierarchical tree graph of heat (A) and waterlogging at high temperature (B) treatments in biological process category with overrepresented GO terms. AgriGo was utilized to map DEGs with GO term and perform singular enrichment analysis to show the enrichment level of the term. GO term number, the *p*-value in parenthesis, and GO term were shown in each box. The pair of numerals in the bottom-left of box represent number of input genes mapped with a particular GO term and total number of input genes. The pair of numerals in the bottom-right of box represent number of genes mapped with a particular GO term in *Arabidopsis* TAIR10 database and total number of *Arabidopsis* TAIR10 genes mapped with a particular GO term in *Arabidopsis* TAIR10 database. Color indicate levels of statistical significance: yellow - adjusted $p \leq 0.05$; orange - adjusted $p \leq 10e^{-5}$.

2.2.4 DEGs involved in responding to reactive oxygen species

In response to abiotic stress, ROS involved in regulation of stress-specific chemicals, compounds, and hormones as signaling molecule (Baxter *et al.* 2014). When plants face heat stress, heat stress factors (HSFs) and HSPs play a critical role in transducting ROS signal and thermotolerance (Qu *et al.* 2013). In this GO term 'response to ROS' (GO:0000302), several genes involved in heat stress response, such as Hsfs and Hsps, were up-regulated in both H and HWS treatments (**Supplemental**

doi:10.6342/NTU201603060

Tab. 4). In these two treatments, *heat shock transcription factor A7A* (*HSFA7a*) was markedly up-regulated and have been known as the key factor for heat tolerance (Larkindale and Vierling, 2008). Meanwhile, genes coding for HSPs such as *heat shock protein 18.2* (*HSP18.2*), *17.6 kDa class II heat shock protein* (*HSP17.6II*), *heat shock protein 70* (*HSP70*), and HSP20-like chaperones superfamily protein were up-regulated when suffering H and HWS stresses. In addition, two genes related to HSP binding, *ERdj3B* and *P58*^{IPK}, were down-regulated in H and HWS treatments. Most of the genes involved in metabolism, regulation of transcription, and transport showed similar expression patterns in these two treatments. However, the gene coding for *abscisic acid responsive elements-binding factor 3* (*ABF3)* was significantly down-regulated in HWS treatment but not in H treatment. *ABF3* is an important binding factor in stress-responsive ABA signaling and the regulation of stomata aperture (Kang *et al.* 2002). These results suggest that cabbage normally regulated the heat tolerance-related genes in H and HWS treatments, but the down-regulation of *ABF3* may cause intolerance to waterlogging stress at high temperature.

2.2.5 DEGs involved in responding to water deprivation

doi:10.6342/NTU201603060

It is well documented that leaf dehydration can be induced by flooding due to restricted stomatal conductance and root hydraulic conductivity (Aroca *et al.* 2012). Based on SEA results, 'response to water deprivation' (GO: 009414) was annotated in HWS treatment but not in H treatment (**Fig. 4**). Several down-regulated genes in HWS treatment had different expression patterns in W and H treatment in this GO category (**Supplemental Tab. 5**). Several transcription factors were negative regulator in response to dehydrating stress (Barak *et al.* 2014, Qiao *et al.* 2016, Tang *et al.* 2005, Yu *et al.* 2015). *ABA insensitive 5* (*ABI5*; BOHW14222), *ABI five binding protein* (*AFP1*; BOHW8926), *stress response suppressor 2* (*STRS2*; BOHW29438), and *enhanced disease resistance 1* (*EDR1*; BOHW26773) were found obviously down-regulated in HWS treatment. Interestingly, *ABI5* and *AFP1* expressed totally opposite profile between W and HW. These results showed that HWS stress induced numerous stress responses, and suppressed the stress tolerance in the cabbages. Additionally, the genes involved in drought stress tolerance, including *auxin response factor 3* (*ARF3*; BOHW5218), *autophagy-related protein 18f* (*ATG18F*; BOHW21429), and *major facilitator superfamily protein* (BOHW8946), were also down-regulated and further caused stress intolerance.

doi:10.6342/NTU201603060

2.2.6 Co-functional gene network of cabbages under HWS

The results of GO analysis showed that cabbages responded to HWS stress with multiple pathways. However, few information about the integration of DEGs under HWS was known. To further understand the signal networks in plants exposed under HWS, known plant co-functional network database, AraNet v2, was used to predict the gene co-expression network of HWS-treated cabbages. AraNet v2 is a genome wide co-functional network web tool which can perform functional predictions thought projecting non-model plant genes to *A. thaliana* gene network (Lee *et al.* 2015). In order to illustrate the signaling network of cabbage under HWS, 221 DEGs which were uniquely expressed in HWS treatment were assigned to AraNet v2 for co-expression network analysis. As shown in **Fig. 5**, 62 of 221 DEGs were found to be co-expressed and assigned in seven major network modules (nodes > 4). Based on the results of AraNet v2, several co-expression modules were related to stress responses. First of all, the DEGs related to ABA signaling pathways in module I, including *ABI5, ABF3*, *AFP1*, and *AFP4*, were co-expressed under HWS. The expression of positive (*ABI5* and *ABF3*) and negative regulators (*AFP1* and *AFP4*) in this pathway were down-regulated under HWS treatment. Meanwhile, *SEC15B*, an exocyst subunit, was also included in this module, but its role was still unclear under stress. Next, the module V and IV were

doi:10.6342/NTU201603060

the two largest modules which contained 12 and 11 genes, respectively. MapMan bin enrichment analysis showed that ‘export form nucleus/ribosome biogenesis’, ‘synthesis-degradation of auxin’, and ‘golgi/secretory pathway’ were enriched in module V (p<0.01; **Supplemental Tab. 4**). The enriched process was mainly related to ‘UXS/precursor synthesis’, ’ubiquitin E1/protein degradation’, ’E2F/DP transcription factor family/regulation of transcription’, ‘Polycomb Group (PcG)/regulation of transcription’, and ’Zn-finger (CCHC)/regulation of transcription’ in module IV (p<0.01; **Supplemental Tab. 4**). The genes in module IV were all down-regulated under HWS treatment. Whereas, there were three genes, BOHW63797 (*ILR1*), BOHW69803, and BOHW39494, up-regulated in module V but their roles in HWS were still unknown. Other two functional modules, VI and VII, contained 5 and 4 genes, respectively. Module VI contain one annotated gene and four unknown contigs. BOHW43287, which encodes *ACC oxidase 1* (*ACO1*), and BOHW12081, which encodes alactose oxidase/kelch repeat superfamily protein, were up-regulated in HW treatment. This up-regulation may reveal that the plant enhanced ethylene production and anaerobic respiration in HW stress. Additionally, module VII was consisted of kinase genes which encode *ADK1* (BOHW29524), *MPK17* (BOHW19006), and two protein kinase family proteins. Within these four kinases, *ADK1* was the only up-

doi:10.6342/NTU201603060

regulated gene, its responses to abiotic stress was unclear. *MPK17* is well known in adaptation and tolerance of osmotic stress in plants (Moustafa *et al.* 2014). However, *MPK17* was down-regulated in cabbages under HWS, and may be one of the factors which caused cabbage HWS-intolerance. On the other hand, the genes of module II and III were not significantly expressed, so they did not investigate in this study.

doi:10.6342/NTU201603060

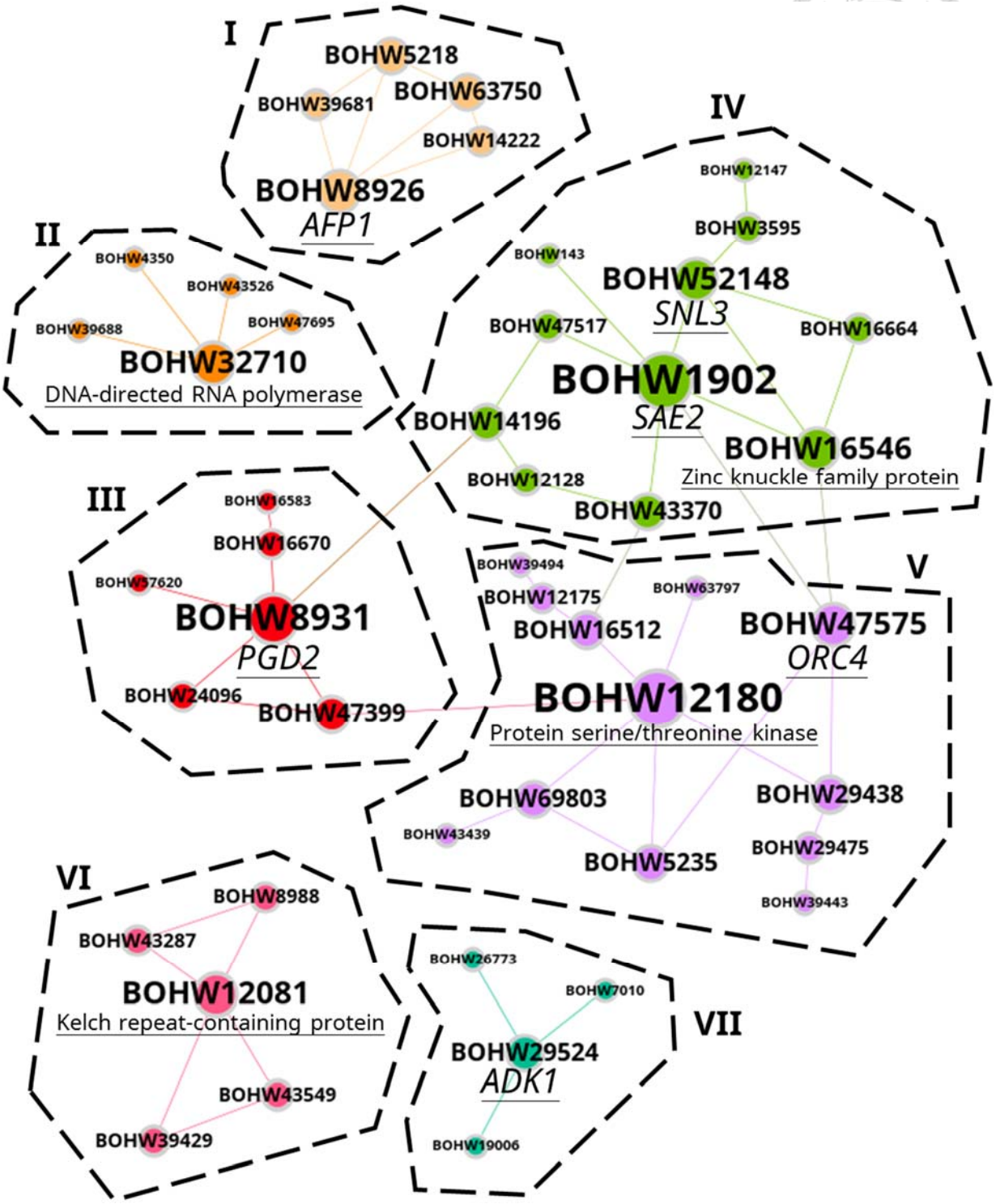


Figure 5. Co-functional network of unique DEGs in HWS-treated cabbage. The DEGs uniquely found in waterlogging stress at high temperature (HWS) treatment were analyzed with AraNet2 to obtain major network modules and labeled by Roman numerals. The gene name with underline indicated the functional annotation of the node.

doi:10.6342/NTU201603060

2.2.7 Validation of RNA-seq results by real time quantitative PCR (qPCR)

In order to confirm the reliability of the data which were obtained from RNA-seq, qPCR was performed on the DEGs found form previous analysis and different samples with the same treatments. Four genes which closely related to HWS, including *ACO1* (BOWHW43287), *ABI5* (BOHW8926), *ABF3* (BOHW5218), and *HSP18.2* (BOHW29550), were selected for qPCR analysis. To compare with the expression values obtained from RNA-seq, the relative expression values between treatments and controls using qPCR were converted to log2 fold change. The qPCR results were showed the significant correlation to RNA-seq results (R^2 =0.7191, *P*=0.0005) **(Fig. 6)**. The reliability of the gene expression patterns obtained by the RNA-seq was supported by the validation analysis of qPCR.

doi:10.6342/NTU201603060

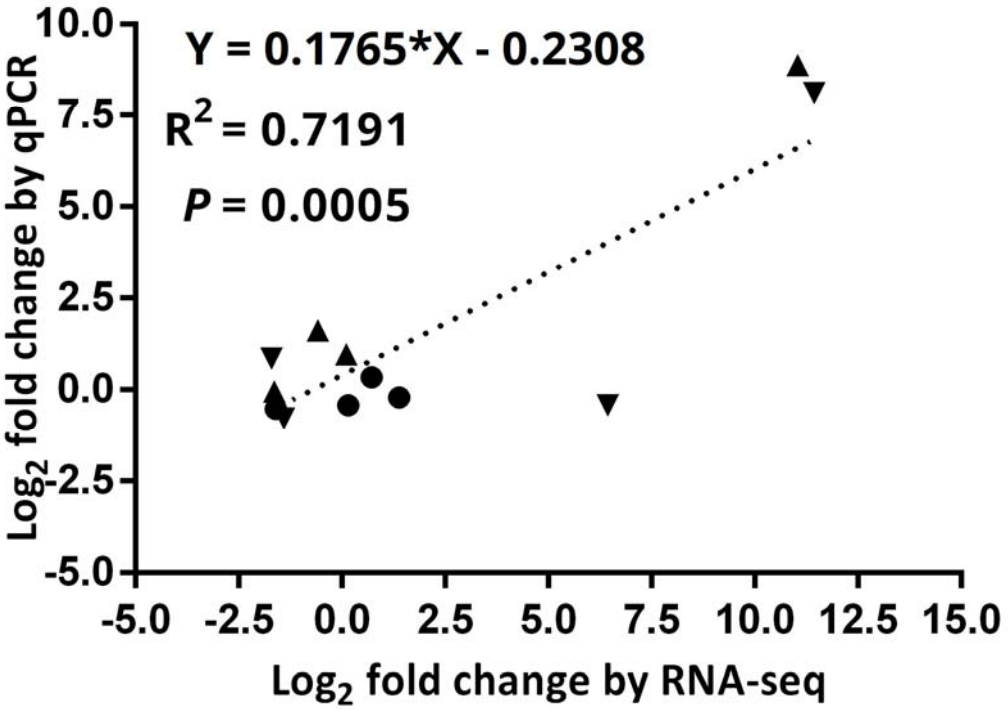


Figure 6. Correlation between RNA-seq and qPCR results for 4 DEGs under 3 treatments. The log2 fold change between waterlogging (W, dot), heat (H, triangle), and waterlogging stress at high temperature (HWS, inverted triangle) treatments is presented for both the RNA-seq results as obtain from DESeq2 and the qPCR results as obtained form 2^{-ddCt} method. The correlation coefficient and *P*-value between relative expression value is shown. Three biological replicates of qPCR were used from each sample under 3 different treatments at 12h.

doi:10.6342/NTU201603060

2.3 Discussion

Combination of waterlogging and heat stresses was common during summer in Taiwan that causes serious physiology disorders or death in cabbage, even when the the stress is temporary. The mechanism responding to waterlogging stress at high temperature in cabbage has been less studied. Most previous studies on this stress in other crops have only been focused on the activity of stress-related enzymes and amounts of secondary metabolites. Moreover, most leafy vegetables do not have the heading process in their cultivation period. The cabbage cultivar, *B. oleracea* var. *capitata L.* 'Shia Feng No.1', is well-known for tolerance to heat and waterlogging at ambient temperature. However, it is susceptible to waterlogging at high temperature. In this study, the heads of 'Shia Feng No.1' treated by waterlogging at ambient temperature, heat, and waterlogging at high temperature at initial heading period were used for RNA-seq analysis. Hence, this may contribute to clarify how the stress influenced gene expression and resulted in stress intolerance in cabbage **(Fig. 7)**. In 'Shia Feng No.1', 19,548 unigenes were successfully identified through transcriptome analysis. Among these unigenes, 2,040 had differentially responded to the stress treatments and been selected for subsequent analysis. In general, the expression profile of selected DEGs in cabbage under HWS was similar to the plant under heat stress.

doi:10.6342/NTU201603060

Most selected DEGs under waterlogging stress at ambient temperature showed similar gene expression profile with control group. However, the expression profile of several DEGs in HWS treatment was totally different between W and H treatments (**Fig. 3**). Among these DEGs, the expression of BOHW43287 and BOHW57430, which encode ACC oxidase 1 and WRKY transcription factor, were up-regulated in HWS and opposite to other two treatments. In addition, the genes in the GO category 'response to water deprivation' (GO: 009414) were uniquely found in HWS treatment and had differentially expressed pattern (**Fig. 4B**). These results demonstrated that the combination of waterlogging and high temperature stresses could cause stress intolerance in the cabbage, through with tolerated for each individual stress, due to the distinct molecular regulatory network.

The WRKY transcription factors involve in the regulation and tolerance to abiotic and biotic stresses. As shown in **Fig. 3**, the WRKY transcription factor and ACC oxidase 1 were uniquely up-regulated in HW stress but not in H and W. Nevertheless, the study focused on submergence response in *Arabidopsis* also observed the up-regulation of *WRKY22*, and suggested the submergence stress induces *WRKY22* to trigger innate immunity (Hsu *et al.* 2013). WRKY transcription factor is also a positive regulator on ethylene biosynthesis (Hsu *et al.* 2013, Li *et al.* 2012). The up-regulation of

doi:10.6342/NTU201603060

both WRKY transcription factor and ACC oxidase 1 represented that the plant needs to produce large amounts of ethylene for responding of the stress in this study. However, the activity of ACC oxidase was inhibited by low oxygen concentration (Vriezen *et al.* 1999). HWS resulted in sustaining ethylene production due to increased ACC oxidase 1 concentration. Similarly, Banga *et al.* have observed flooding-sensitive *Rumex* lacked the negative regulatory mechanism for ethylene accumulation under submergence (Banga *et al.* 1996). In this study, the up-regulation of WRKY and ACC oxidase 1 could be observed in waterlogging treatment at ambient temperature. Thus, the combination of high temperature and waterlogging stress may interfere the negative regulation of ethylene biosynthesis and then results in the loss of waterlogging tolerance in *B. oleracea* var. *capitata L.* 'Shia Feng No.1'. Furthermore, the expression levels of ABA-related genes including *ABI5*, *ABF3*, *AFP1*, and *AFP4* were down-regulated under HWS. In general, it is well-characterized about the antagonism between ABA and ethylene (Cheng *et al.* 2009). As described above, the down-regulation of ABA-related genes could result in ethylene accumulation in HWS. ABA plays a critical role in plant developmental process and adaptation during abiotic stresses especially temperature and water stress (Baron *et al.* 2012). The ability to synthesize ABA under heat stress is the key factor for plant thermotolerance and the regulator of stomatal closure that controls

doi:10.6342/NTU201603060

carbon fixation, water status and transpiration (Bita & Gerats 2013, Reynolds-Henne *et al.* 2010). During HWS, the down-regulation of ABA-related genes may cause continuous stomatal opening to bring about dramatic transportation, and finally induced water deprivation stress. The GO term enrichment analysis of DEGs in HWS treatment also contained the GO term of 'response to water deprivation' (**Fig. 4B**).

On the other hand, GO enrichment of DEGs revealed that some GO categories were enriched in both H and HWS treatments (**Fig. 4**). Most genes involved in 'response to stimulus' (GO:0050896), 'response to stress' (GO:0006950), and 'response to ROS' (GO:0000302) were enriched and had the similar expression profile in H and HWS treatments. Several DEGs in these GO categories were up-regulated and well known for heat resistance, especially Hsfs and Hsps. Hsps have multiple functions on various pathways and play the key role in several stresses such as heat, low oxygen, and freezing (Hasanuzzaman *et al.* 2013, Pucciariello *et al.* 2012, Timperio *et al.* 2008, Wahid *et al.* 2007). These results demonstrate that both H and HWS treated cabbages could perform similar molecular regulatory mechanism in response to high temperature.

doi:10.6342/NTU201603060

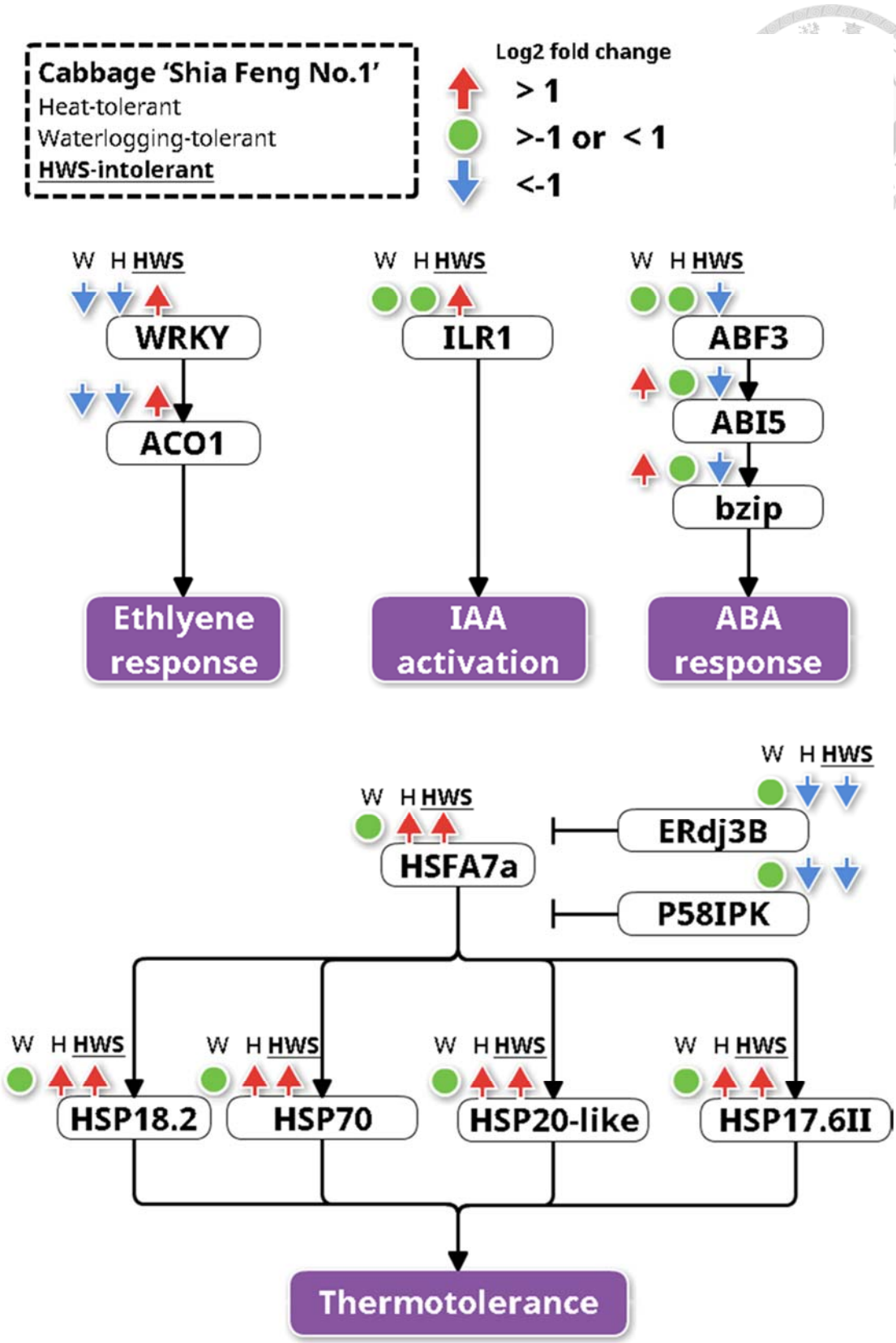


Figure 7. Summary of the complex regulatory networks in response to waterlogging (W), heat (H), and waterlogging stress at high temperature (HWS) in cabbage 'Shia Feng No. 1'.

doi:10.6342/NTU201603060

HWS induced up-regulation of ethylene response and IAA activation. The genes involving these pathways including *WRKY*, *ACO1*, and *ILR1* had opposite expression profiles in W and H. Meanwhile, ABA response-related genes were found to be down-regulated to reduce the tolerance of waterlogging stress at 12h after HWS. On the other hand, the thermotolerance-related genes such as heat shock proteins (HSP) and heat shock transcription factor (HSF) both were up-regulated in H and HWS. The negative regulators of HSF, *REdj3B* and *P58IK*, also were down-regulated in these two stresses. The combination of waterlogging and heat stress resulted in the error of hormone regulatory pathway to cause the intolerance in cabbage in HWS.

doi:10.6342/NTU201603060

Chapter 3 Comparative analysis of HWS responses in cabbages by weighted gene co-expression network analysis (WGCNA)

3.1 Materials and Methods

3.1.1 Plant materials and growth conditions

Cabbage (*B. oleracea* var. *capitata L.*) '228' (purchased from TAKII seed, Japan) and 'Fuyudori' (purchased from Kobayashi seed, Japan) F1 seeds were first soak overnight in the dark in a petri dish with a moist filter paper. After germinating, seeds were planted in 72-cell plug trays containing a mixture of 75% peat moss and 25% perlite. Plants were grown in phytotron at 23°C/18°C with a 16-h photoperiod (300 μmol photos m^{-2} s^{-1}) for three weeks. Consequently, 3-week-old seedlings were transplanted to the field in plastic greenhouse. All plants were watered twice per day with automatic drip irrigation system for 5 weeks, applied as a foliar spay of 1/300x Foliar Nitrophoska® 20-19-19 (COMPO, Germany) with 1/1000x Tween-20 every week. Eight-week-old '228' and 'Fuyudori' were transplanted to 8-inch pots to treat under waterlogging at 35 °C in a phytotron for 0, 6, 12, and 24 hours. Waterlogging stress was endured by soaking pots in a tank and maintaining 5 cm of standing water from the top of the culture medium. Leaf samples of '228' and 'Fuyudori' were

doi:10.6342/NTU201603060

collected with the design of time course experiment and 2 biological replicates for each cultivar and time point. After collecting, samples were soaked in liquid nitrogen, and stored in -80°C for future RNA extraction and RNA-seq analysis.

3.1.2 RNA extraction, cDNA library preparation, and RNA-seq

Total RNA was purified from 16 leaf samples (2 cultivars × 2 biological replicates × 4 time points) using Plant Total RNA Miniprep Kit (GeneMark). Total RNA samples were quantified using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies), and the RNA integrity was further determined using RNA600 Nano assay with the Agilent 2100 Bioanalyzer (Agilent Technologies). After quantification and qualification, total RNA samples of '228' and 'Fuyudori' were sent to Genomics BioSci & Tech company (Taipei, Taiwan), and 16 total RNA samples from each leaf sample were individually prepared and finally constructed into 16 RF strand specific cDNA libraries with 150-400 bp insert size and sequenced with Illumina NextSeq500. All the cDNA libraries were constructed using selected poly(A) mRNAs following Illumina Truseq RNA-seq library preparation procedure.

doi:10.6342/NTU201603060

3.1.3 *De novo* assembly, annotation, and differential gene expression analysis of RNA-seq data

In both cultivars, one-hundred-fifty bp strand-specific PE raw reads were used to perform one *de novo* assembly using Trinity and two genome-guided assembly using Tophat (Grabherr *et al.* 2011) and Cufflinks (Trapnell *et al.* 2012) as shown in **Fig. 8**. *De novo* assembly was performed with all raw reads in all time points for individual cultivar, and each pair-end library was inputted to Trinity without merging reads before assembly. For genome guided assembly, genome sequences of *B. oleracea* var. *oleracea* and *capitate* were employed as reference, respectively.

After assembly, 2 assembles generated from Trinity and 4 genome guided assembles were obtained and then redundant contigs were removed using Tr2aacds pipeline after merging all the assembles (http://arthropods.eugenes.org/EvidentialGene/-trassembly.html; Visser *et al.* 2015). Following the pipeline, Tr2aacds firstly predicted the CDS and amino acid sequence for each contig and removed the redundant contigs according to the coding potential. Then, removing sequence fragments was performed using substring de-replication. Finally, the primary assembled contigs for subsequent analysis were generated through clustering the highly similar sequences into loci,

doi:10.6342/NTU201603060

grouping these sequences as as ‘primary’ or ‘alternate’, and discarding the low scoring ‘drop’ sequences.

Before annotation, a process for filtering the contigs with low/zero counts was performed using Bowtie2 and eXpress. Blastx was employed to annotate the filtered contigs using TAIR and NCBI Reference Sequence Database (Refseq) plant peptide databases. Blast-Parser (http://kirill-kryukov.com/study/tools/blast-parser/) was employed for filtering the blast results to reserve the aligned gene with identity > 70%. The contigs only aligned to the genes with the identity < 70% were assigned as undefined genes. Moreover, the TAIR ID of identified contigs was imported to AraNet v2 for enrichment of GO terms and MapMan bins (Lee *et al.* 2015). On the other hand, the expression value of filtered contigs was calculated using DESeq2 and then transformed to $\log_2$ fold change value for primary expression analysis and regularized log value for WGCNA. For analyzing with likelihood ratio test (using LRT), the FDR was indicating the significance of DEGs.

3.1.4 Gene coexpression network construction and visualization

WGCNA (v1.51) package in R was employed to construct the coexpression networks of the filtered contigs (Langfelder & Horvath 2008). The regularized log value of filtered

doi:10.6342/NTU201603060

contigs was imported into WGCNA and coexpression modules were generated using the pipeline of step-by-step network construction and module detection with following parameters: softpower =6, TOM type = signed, and minModuleSize = 30). The eigengenes value was calculated for each coexpression module and employed to examine the correlation with each time point and cultivar. Eigengenes represented the first principal component in the principal components analysis (PCA) using the normalized expression values of all filter contigs in the modules and show as the average normalized gene expression for a module (Langfelder & Horvath 2008). The module networks were visualized using Gephi v0.9.1.

3.2 Results

3.2.1 Global analysis of RNA-seq data

To discover the stress-tolerance genes of cabbage under HWS, RNA-seq technique was employed to screen the transcriptome of cabbage cultivars with contrasting response (tolerance/intolerant) to HWS. In Taiwan, the HWS following typhoon in cabbages is transitory and the duration is commonly about 24-48 hours. Accordingly, a detailed time-course experiment is necessary for obtaining valuable insights on the gene response of HWS-suffered cabbage. For selecting cabbage cultivars as the materials in

doi:10.6342/NTU201603060

this study, the study on evaluating the physiological response to HWS in 25 and 24 commercial cabbage cultivars in the field was referred (Chen, 2011). Two cultivars with high and low growth score under stress, '228' and 'Fuyudori', were selected as the plant materials in this comparative study (**Fig. 8A**). RNA-seq data were generated from 16 different leaf samples (2 cultivars × 4 time points × 2 biological replicates) at different time points during HWS. In total, 16 separate cDNA libraries were obtained for RNA-seq analysis. To reach better efficient of read mapping and consistency of sequencing, the strand-specific library construction was employed and all 16 libraries were sequenced in the same flow cell. A total of 250,990,944 reads were selected for further analysis as the flow chart (**Fig. 8B**).

doi:10.6342/NTU201603060

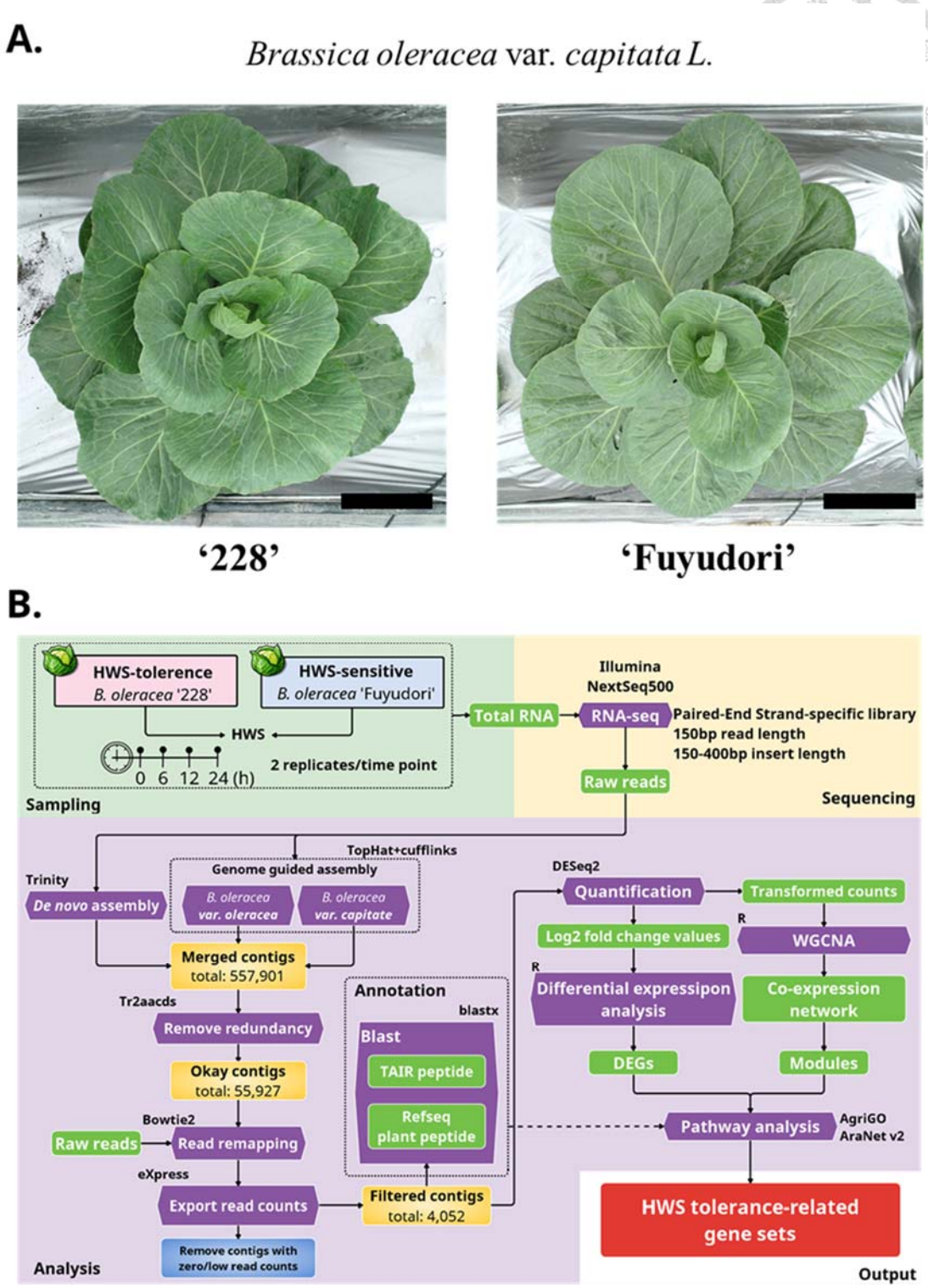


Figure 8. Appearance of plant materials and flowchart of data collection, processing and transcriptome analysis. (A)The appearance of 8-week-old cabbage '228' and 'Fuyudori' before waterlogging stress at high temperature (HWS). The black bar represented 10 cm. (B) Pipeline of differential expression and co-expression network analysis in this study.

doi:10.6342/NTU201603060

Due to the absence of a completely annotated reference sequence is, *de novo* assembly is one of the options to obtain the transcriptome sequence. However, the assembled transcripts form Illumina short reads may leverage the inherent redundancy to interfere the comprehensive transcriptome analysis (Visser *et al.* 2015). To obtain better transcriptome assembly, the strategy of combined *de novo* and genome guided assembly was used in this study. Due to very good quality of reads (over 95% of reads with a Phred score > Q30), all reads were not trimmed and directly applied to contig assembly (**Tab. 1**). For maximizing diversity of assembled transcripts, 2 de novo and 4 genome guided transcriptome assemblies was produced using Trinity and Tophat/Cufflinks. In the part of *de novo* assembly, the reads of all time points in the same cultivar were merged due to the limit of the equipment. On the other hand, genome guided transcriptome assembly used the genome sequences of *B. oleracea* var. *oleracea* and var. *capitate*, as the reference. *B. oleracea* var. *oleracea* is not a cultivated cultivar which is known as wild cabbage and have the chromosome assembly level of genome sequence (Parkin *et al.* 2014, Snogerup *et al.* 1990). Although the variety of *B. oleracea* var. *capitate* is the same as two cabbage cultivars in this study, its genome sequence is only scaffold level (Liu *et al.* 2014). Thus, assemblies with different reference genomes we expected to be more accurate and complete in different loci. The

doi:10.6342/NTU201603060

results of *de novo* assembly exhibited the longest length and largest number of contigs among other genome guided assemblies (**Tab. 2**).

Table 1. Quality statistics of the raw reads using FastQC.

Sample	Strand	Total sequences	Sequence length	>Q30 (%)*	Total (Gb)
228_0h_A	F	17,453,992	20-151	97.6	2.9
	R			96.9	2.9
228_0h_B	F	17,287,750	22-151	97.4	2.8
	R			96.5	2.8
228_6h_A	F	17,291,868	21-151	97.7	2.8
	R			97.2	2.8
228_6h_B	F	16,540,252	20-151	97.6	2.7
	R			96.8	2.7
228_12h_A	F	9,880,912	22-151	97.8	1.6
	R			97.2	1.6
228_12h_B	F	17,287,750	22-151	97.5	3.3
	R			96.2	3.3
228_24h_A	F	18,263,852	21-151	97.4	3.0
	R			96.3	3.0
228_24h_B	F	17,090,576	22-151	97.4	2.9
	R			96.7	2.9
Total		131,096,952			44.0
Fuyudori_0h_C	F	20,438,580	20-151	97.5	3.4
	R			96.3	3.4
Fuyudori_0h_D	F	19,025,956	21-151	97.3	3.2
	R			94.7	3.2
Fuyudori_6h_C	F	16,393,378	20-151	97.3	2.7
	R			96.5	2.7
Fuyudori_6h_D	F	16,540,252	24-151	96.7	1.5
	R			96.8	1.5
Fuyudori_12h_C	F	18,379,030	21-151	97.4	3.0
	R			96.7	3.0
Fuyudori_12h_D	F	17,679,090	21-151	96.9	2.9
	R			95.4	2.9
Fuyudori_24h_C	F	10,568,742	21-151	97.7	1.7
	R			97	1.7
Fuyudori_24h_D	F	16,669,726	22-151	97.5	2.7
	R			96.4	2.7
Total		16,961,844			42.2

***Percentage of reads in the library with a Phred score >30**

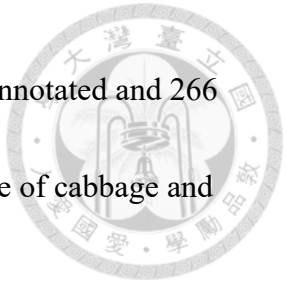
doi:10.6342/NTU201603060

Table 2. Quality statistics of assembles.

Assembly name	Total sequences	N50	Max. length (bp)	Min. length (bp)
De novo assembly				
'228'	198,949	1,344	16,485	201
'Fuyudori'	205,468	1,355	16,508	201
Genome guided assembly				
'228'_olereacea	42,331	1,306	10,585	120
'228'_capitate	34,411	1,305	8,078	146
'Fuyudori'_olereacea	42,331	1,306	10,585	120
'Fuyudori'_capitate	34,411	1,305	8,078	146

De novo and genome guided transcripts were combined into a data set with a total of 557,901 contigs, and then used as input for redundancy removal with EvidentialGene tr2aacds pipeline. The EvidentialGene pipeline is based on the coding potential prediction to select the 'best' transcripts as the best open reading frames (ORFs) assembled and the paralogs of protein sequence with high identity were removed. Totally 55,927 contigs were obtained after removing redundancy and performed the read remapping using Bowtie2 and eXpress. This reads remapping process was depended on the read abundance of contigs to filter the contigs with low/zero counts which was likely the noise of sequencing data. Further, 4,052 filtered contigs (https://www.space.ntu.edu.tw/navigate/s/0AE479853B774818B49396B3BF6A33F8QQY) were analyzed with Blastx against the *Arabidopsis* protein datatbase (TAIR10) and Refseq plant protein database. Due to the integrity of database, the contigs in this study firstly

doi:10.6342/NTU201603060



were annotated with TAIR and then Refseq database. Finally, 3,786 annotated and 266 uncharacterized transcripts was obtained from the HWS transcriptome of cabbage and as the reference for subsequent analysis (**Supplemental Fig. 1**).

3.2.2 Differentially gene expression identifies the changes in stress tolerance-associated genes

First, Bowtie2 and eXpress were employed to remap reads with reference sequence and estimate counts for each sample. For reads remapping, higher than 84% of the reads from all libraries could map to the reference sequence. Differential gene expression values and pairwise comparisons were generated with DESeq2. **Fig. 9A** and **9B** showed the results of sample-to-sample distance and PCA generated with DESeq2 function. The distance plot revealed the consistency of sample replicates and cultivars on the gene expression under HWS (**Fig. 9A**), and similar status was found in the PCA plot (**Fig. 9B**). PCA revealed that the libraries were segregated horizontally (PC1) based on the cultivar of sample sets and vertical segregation (PC2) was associated to the sampling times. Besides, as determined by the likelihood ratio test (LRT; FDR = 0.1) in DESeq2 package, 256 of the 4,052 genes had significant expression variability (EV) in two cultivars during HWS.

doi:10.6342/NTU201603060

The heatmap of 265 DEGs with significant gene variability was applied to hierarchical clustering with the "euclidean" distance and the linkage method "complete". Clear clusters were related to the cultivars during the HWS (**Fig. 9C**). Twelve clusters which represented 265 genes were clustered according to the expression trends and the dendrogram was generated by hierarchical clustering (**Fig. 10**). Then, based on the similar expression profiles, 12 clusters were combined into 7 superclusters. To identify specific biological processes enriched in specific superclusters, AraNet v2 was used to perform Gene Ontology-Biological Process (GO-BP) terms enrichment analysis. The predominant GO-BP term for HWS-tolerance cultivar '228' supercluster (supercluster I) was annotated with 'Arginine biosynthetic process'. Interestingly, there were two HWS-intolerant cultivar specific superclusters (IV and VII), exhibiting totally opposite expression trends to the tolerance one. The predominant GO-BP terms for these two superclusters were 'Response to water deprivation' and 'Chaperone mediated protein folding requiring cofactor'.

doi:10.6342/NTU201603060

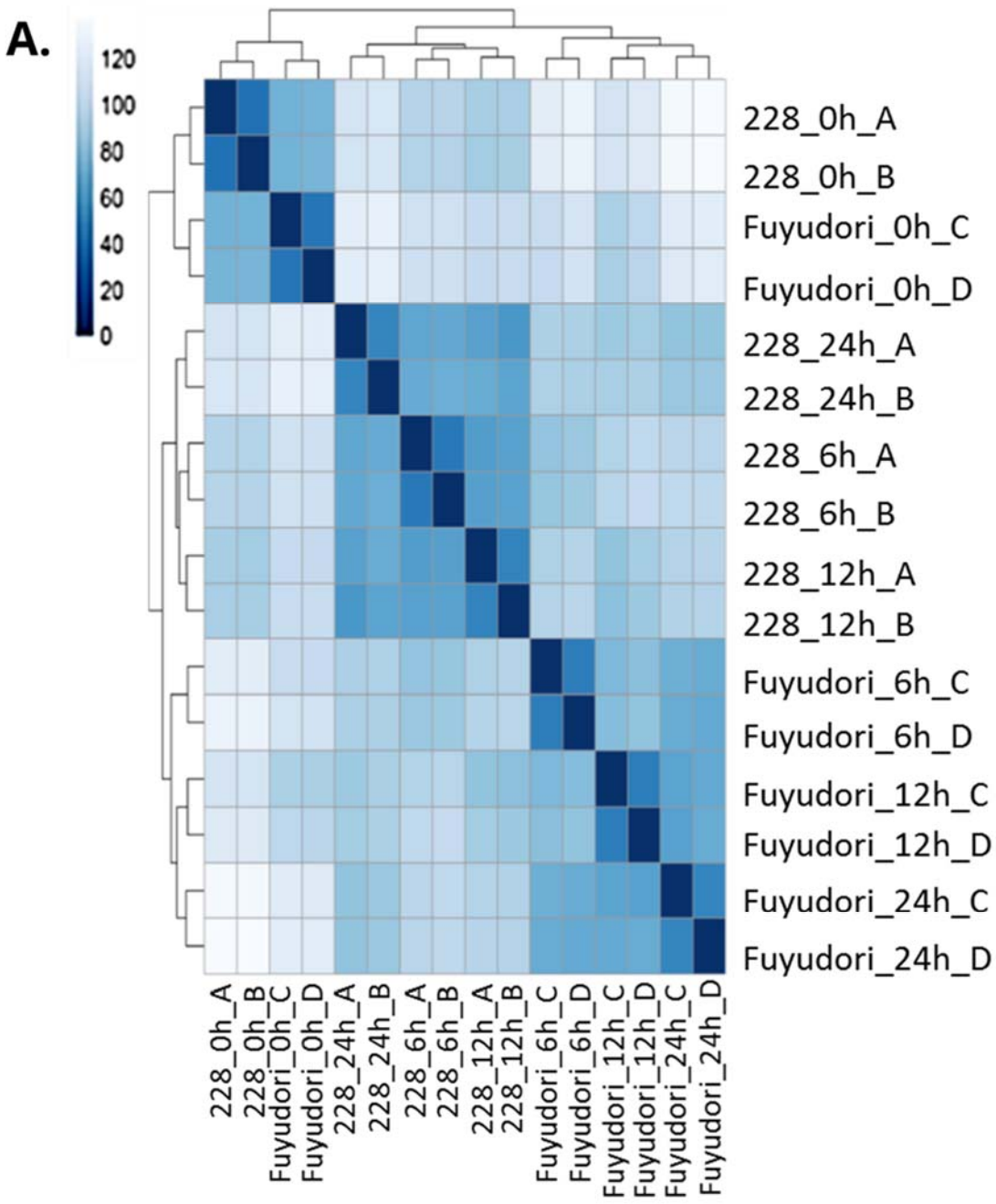


Figure 9. Sample clustering and the expression profile of the 265 significantly expressed genes (FDR<0.1). (A) Heatmap of clustered correlation matrix showed Euclidean distances in gene expression between samples. Color key was adjusted to minimal and maximal regularized log value to enhance the difference. Dendrogram showed the relationship-distance between samples.

doi:10.6342/NTU201603060

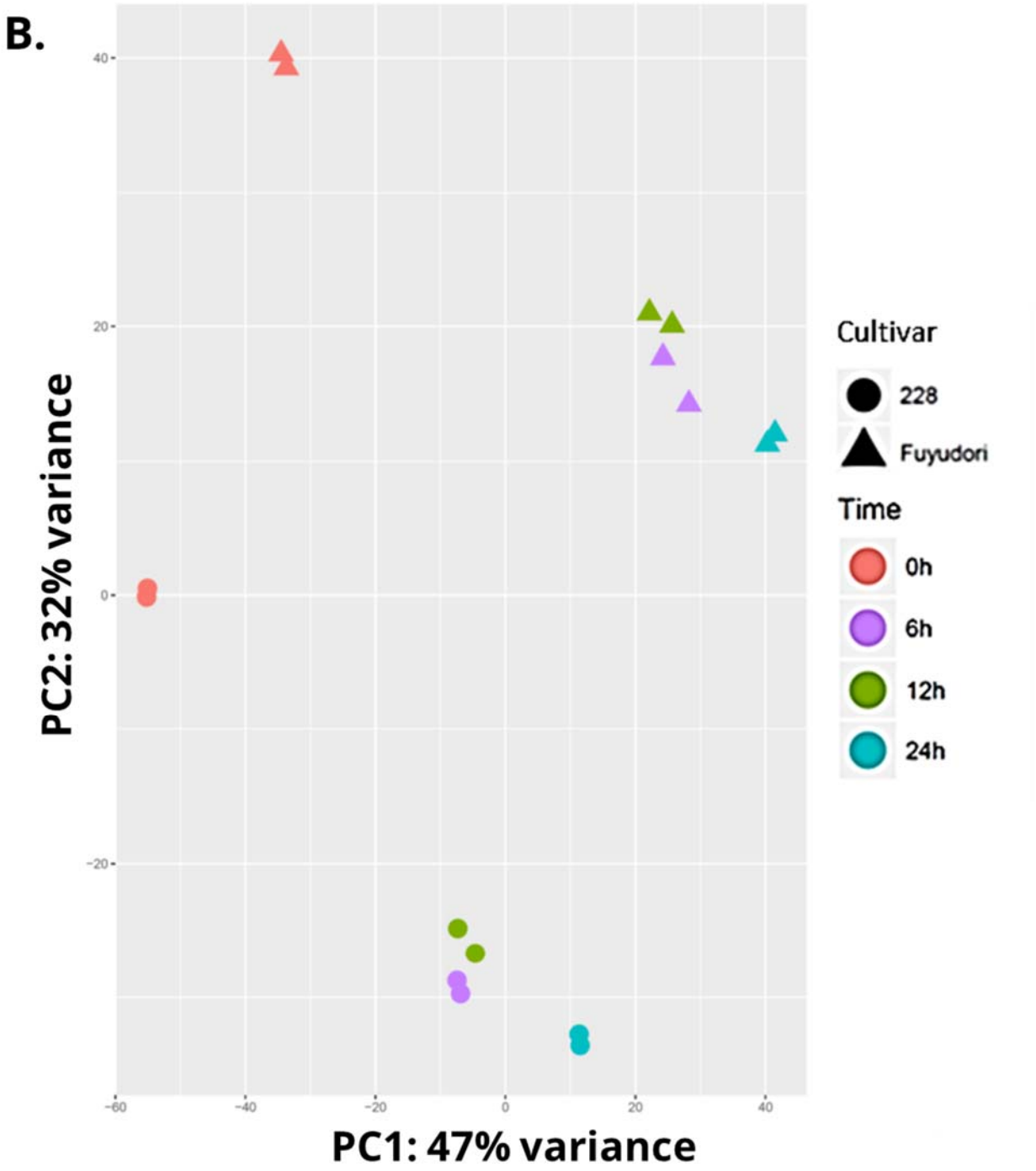


Figure 9. Sample clustering and the expression profile of the 265 significantly expressed genes (FDR<0.1) (countinous). (B) Principal component plot of the samples.

doi:10.6342/NTU201603060

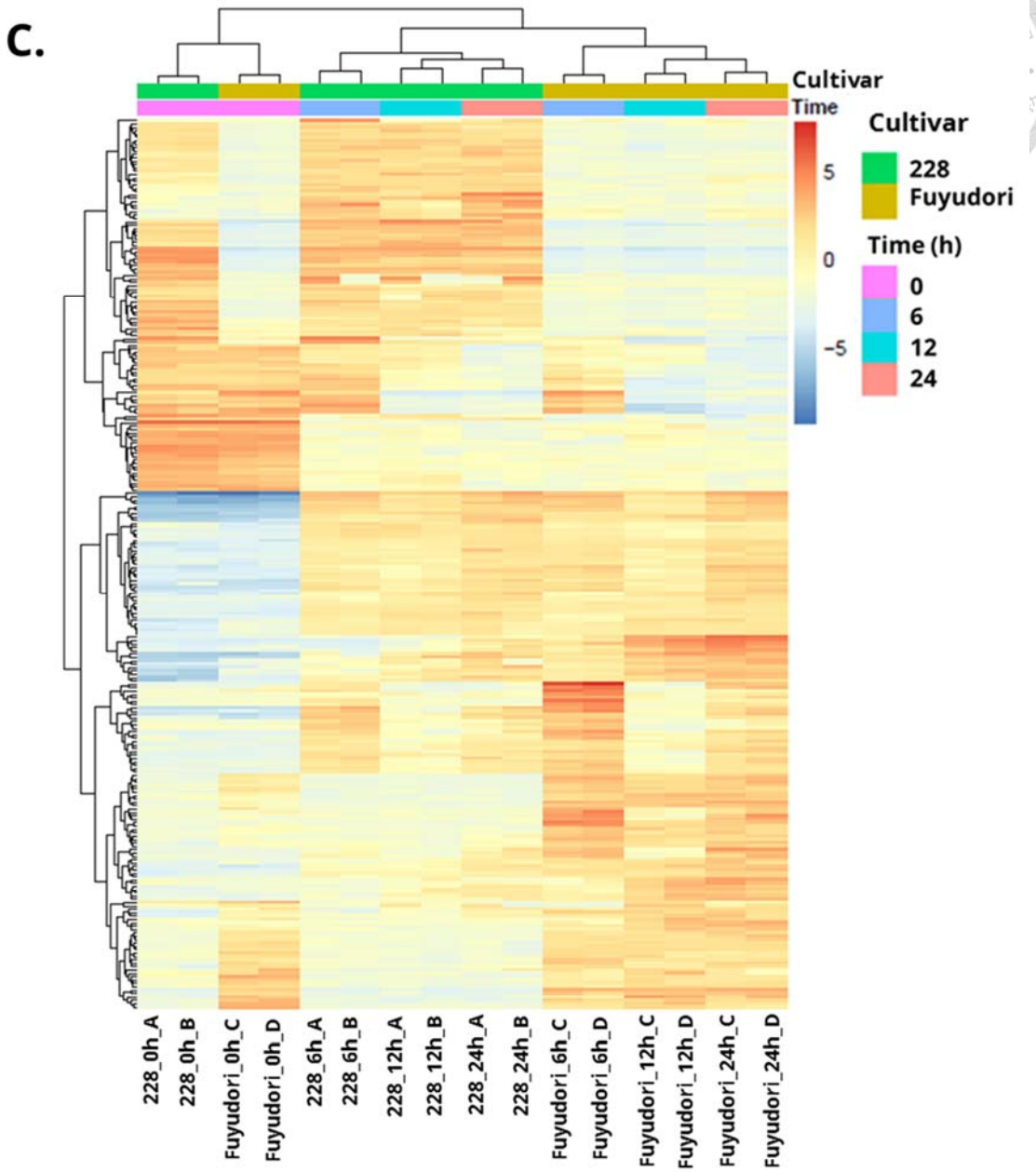


Figure 9. Sample clustering and the expression profile of the 265 significantly expressed genes (FDR<0.1) (countinous). (C) Heatmap showed the expression of the 265 significantly expressed genes in different cultivars and time points under waterlogging stress at high temperature. The data is the normalized counts from regularized log transformation.

doi:10.6342/NTU201603060

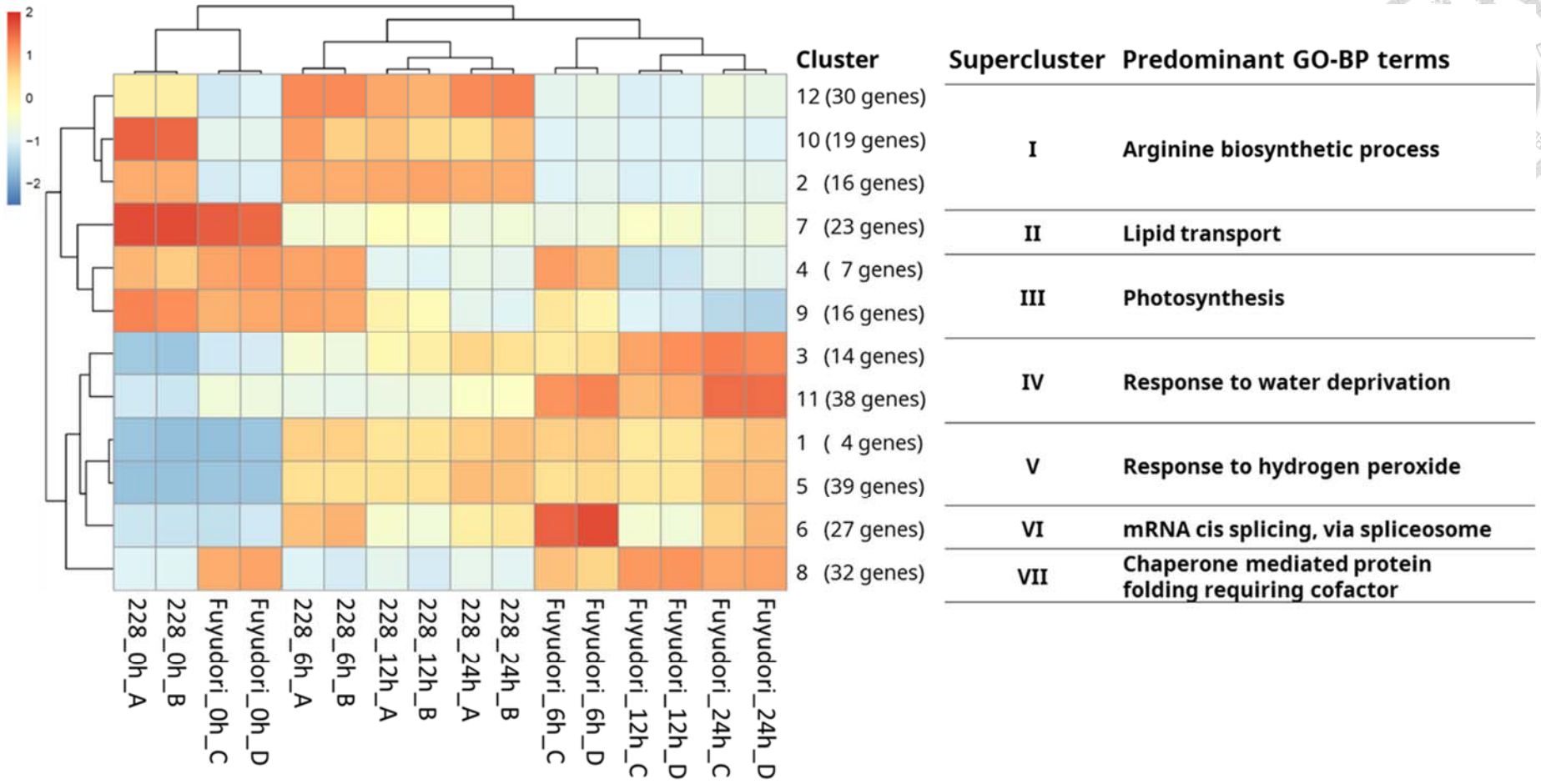


Figure 10. Heatmap of expression trends in the clusters of the 265 significantly expressed genes and the enriched GO terms of superclusters. Color key showed the mean $\log_2$ fold change value of cluster and was adjusted to minimal and maximal value to enhance the difference.

doi:10.6342/NTU201603060

The enriched metabolic pathways of DEGs responded to HWS in these two cultivars are totally different in the MapMan terms enrichment analysis with the Aranet v2 (**Supplemental Tab. 7**). The significantly enriched MapMan bins of '228'-specific supercluster I mostly belonged to lipid degradation, protein.synthesis, and secondary metabolism. Most genes of the supercluster I were continuously up-regulated during HWS in '228' (**Supplemental Fig. 3A**). These genes were non-changed or down-regulated in 'Fuyudori'. However, the up-regulated expression trends of 'Fuyudori'-specific supercluster IV which the genes were associated with protein degradation, photosynthesis, and redox were non-changed or down-regulated in '228'. Therefore, the '228'- and 'Fuyudori"-specific cluster reflect the HWS tolerance could be affected via the regulatory mechanism of these pathways.

3.2.3 Weighted gene correlation network analysis identified three cultivar-specific modules

In the differential gene expression analysis mentioned above, 265 most significantly changed genes in two cabbage cultivars under HWS were identified, but thousands of genes which may play critical role on the HWS tolerance. Gene coexpression network analysis can efficiently identify the hub genes of interest in stress

doi:10.6342/NTU201603060

responses in plants and construct the linkage between molecular mechanisms and physiological traits (Amrine *et al.* 2015, Liseron-Monfils & Ware 2015, Priest *et al.* 2014). In this study, WGCNA was employed to construct the coexpression network based on the correlation expression pattern of genes across all samples. WGCNA is a system biology method for examining the correlation network of genes across samples using gene screening approach and relies on the topological overlap mapping metric (TOM) to connect genes for constructing modules (Langfelder & Horvath 2008). As a result, 13 distinct modules composed of 4,052 genes with correlated expression trends during HWS were obtained (**Fig. 11A**). As shown in the dendrogram (**Fig. 11A**), the modules were composed of tree branches and each leaf in the branch represented one gene. Each module was assigned to a unique color label which was as the module identifier in the analysis. Meanwhile, the first principle component of a given module was assigned as the module eigengene which represented the gene expression profile of module. Each eigengene for each module were correlated with distinct cultivars and time points to calculate the corresponding correlation and *p*-value based on the cultivar- and time-specific expression profile of eigengenes (**Fig. 11B**). Remarkably, some modules are representative of genes with high expression in a single cultivar or time point such as module greenyellow, green, purple, turquoise, midnightblue, and cyan (r^2>

doi:10.6342/NTU201603060

0.75, $p < 10^{-4}$; **Fig. 11B**). In these 6 modules, 3 cultivar-specific modules with high correlation and *p*-value were identified as the cluster of genes which could be intimately associated with the HWS tolerance. For example, the green and purple modules contained 312 and 205 'Fuyudori'-specific genes, respectively. These two gene clusters also included 37 and 14 'Fuyudori'-specific genes which were identified in earlier analysis results of primary expression analysis (**Fig. 10**). WGCNA and primary expression analysis both identified two major expression patterns when comparing the stress-tolerant and stress–intolerant cultivars during the HWS.

doi:10.6342/NTU201603060

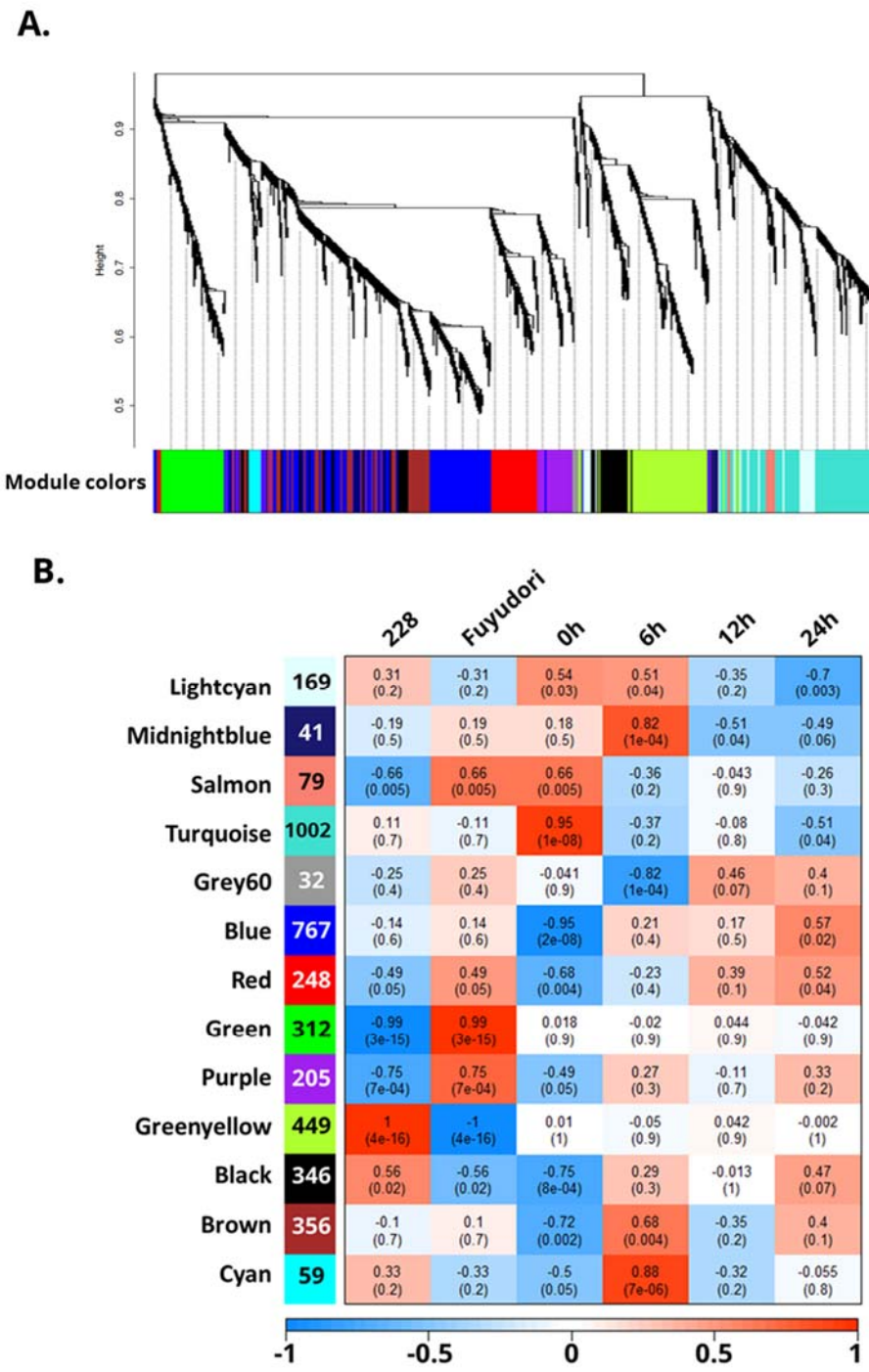


Figure 11. WGCNA of genes in two cabbage cultivars under waterlogging stress at high temperature. (A) Gene clustering and module identification using WGCNA. Each leaf in the tree represented one gene. The major branches of the tree were composed of 13 modules with different color labels. (B) Module-cultivar and module–time point association. Each row corresponded to a module which showed the gene number in the color cell. Each column corresponded to the specific cultivar or time point. The heatmap showed the correlation coefficient (r^2) between the module and cultivar/time point. The color key represented the correlation value. The *p*-value was showed in brackets.

doi:10.6342/NTU201603060

One of aims in this study is to identify the stress tolerance-specific module, such as the greenyellow, green, and purple (**Fig. 11**). **Fig. 12** showed the heatmap and pattern of eigengene expression in greenyellow, green, and purple module, respectively. The figures showed the relative normalized counts of genes from the interested modules and indicated the module of genes strongly expressed in their specific cultivars. In these three modules, the greenyellow ('228'-specific) and green ('Fuyudori'-specific) displayed totally opposite expression pattern in all time points (**Fig.12A** and **12B**). Also, the distinct expression pattern was observed between '228'-specific greenyellow module and 'Fuyudori'-specific purple module (**Fig12A** and **12C**). Whereas, the time point-specific modules such as turquoise, midnightblue, and cyan showed the same expression pattern despite of different cultivars. The expression pattern of the genes in these modules suggested that these genes related to the inherent regulatory mechanisms of cabbages when suffering HWS.

To understand the biological processes of these cultivar-specific modules, AgriGO was employed to assign the GO term to each gene and enriched them to the appropriate biological process category. As shown as **Fig. 13**, the most significantly enriched GO terms were associated to 'cellular metabolic process' (GO:0044237), 'primary metabolic process' and 'small molecule metabolic process' (GO: 0044271) for '228'-specific

doi:10.6342/NTU201603060

module greenyellow. Further, the most specific GO term of greenyellow was 'phenylpropanoid biosynthesis process' (GO:0009699) and 12 of 285 genes in greeyellow module were involved in the process. Phenylpropanoid have been well-known about its necessity for adaptation under various stresses, e.g., UV-B stress, thermal stress, and nutrient stress (Arbona *et al.* 2013, Dixon & Paiva 1995, Solecka). Unlike the HWS-intolerant 'Fuyudori', the expression trend of the greenyellow module was constantly up-regulated during the HWS (**Fig.12A**). This result suggests that high content of phenylpropanoids could increase the HWS tolerance in cabbage. In addition, the enriched GO terms in 'Fuyudori'-specific green module involved more biological processes than '228'-specific greenyellow module. Even through the GO terms related to cellular amino acid metabolism were also enriched in green module, most of the genes in this module were enriched in 'response to stimulus' (GO:0050896) (**Fig. 14A**). In this category, 'response to stress' (GO:0006950), 'response to chemical stimulus' (GO:0042221), and 'response to abiotic stimulus' (GO:0009628) all indicated that the plant suffered metabolic disturbances and attempted to up-regulate the stress-related genes for altering this biological status. Further, purple module which also belonged to 'Fuyudori'-specific modules showed high correlation with 'indole acetic acid biosynthesis process' (GO:0009684) (**Fig. 14B**). Indole acetic acid (IAA) has been

doi:10.6342/NTU201603060

reported to induce the development of adventitious roots under hypoxia stress and can influence the flooding tolerance in various plants (Arbona & Gómez-Cadenas 2008, Bai *et al.* 2010, Thirunavukkarasu *et al.* 2013). However, these genes involved in IAA biosynthesis in ‘Fuyudori’ and ‘228’ during the HWS were not different in the gene expression profile. This result does not correspond with the reports on the distinct IAA production in citrus and *Malus* species plants with different flooding tolerance.

doi:10.6342/NTU201603060

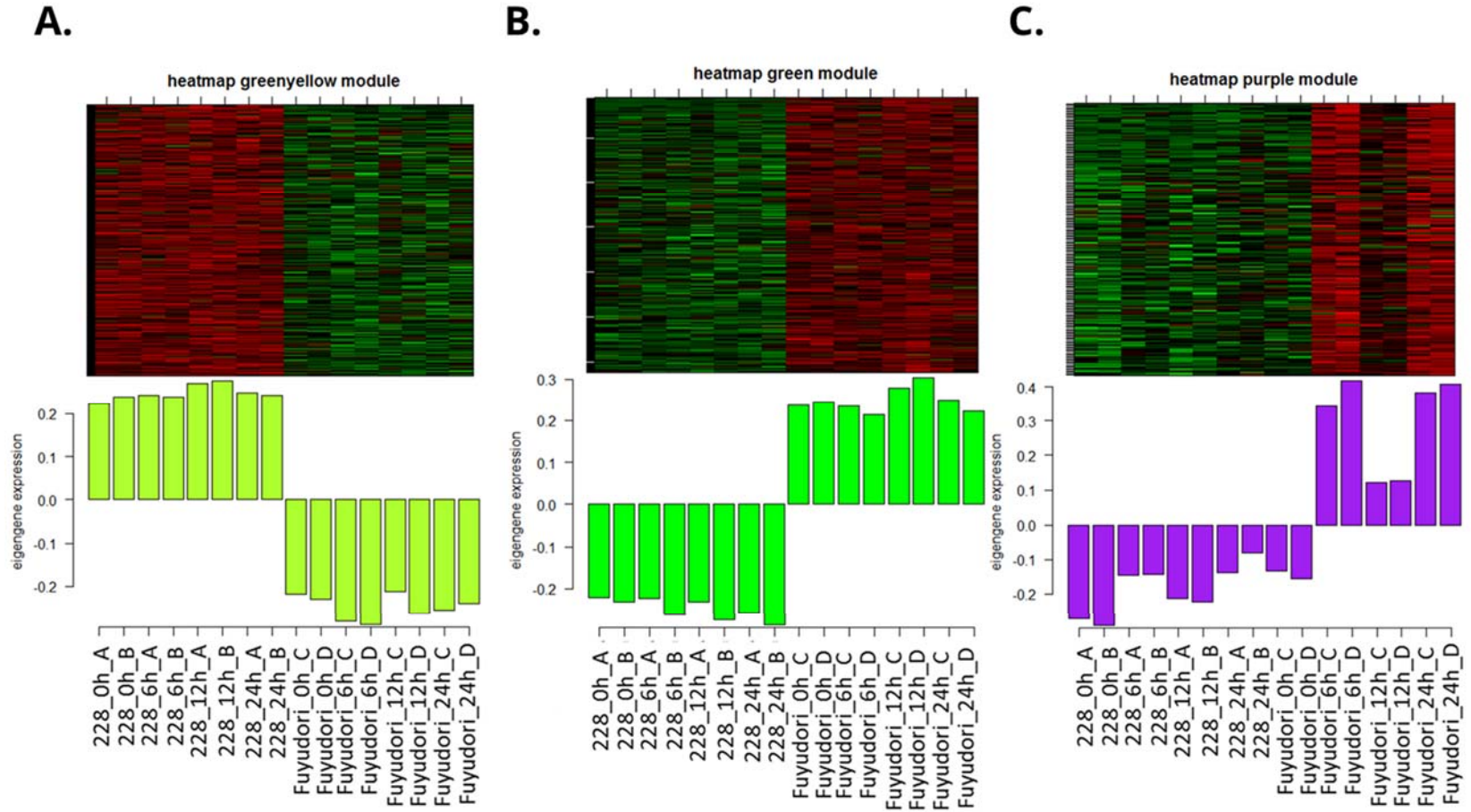


Figure 12. Heatmap and eigengene expression patterns for cultivar-specific modules. Heatmap and eigengene expression profile across each cultivar for module greenyellow (A), green (B), and purple (C). Samples were indicated below each column.

doi:10.6342/NTU201603060

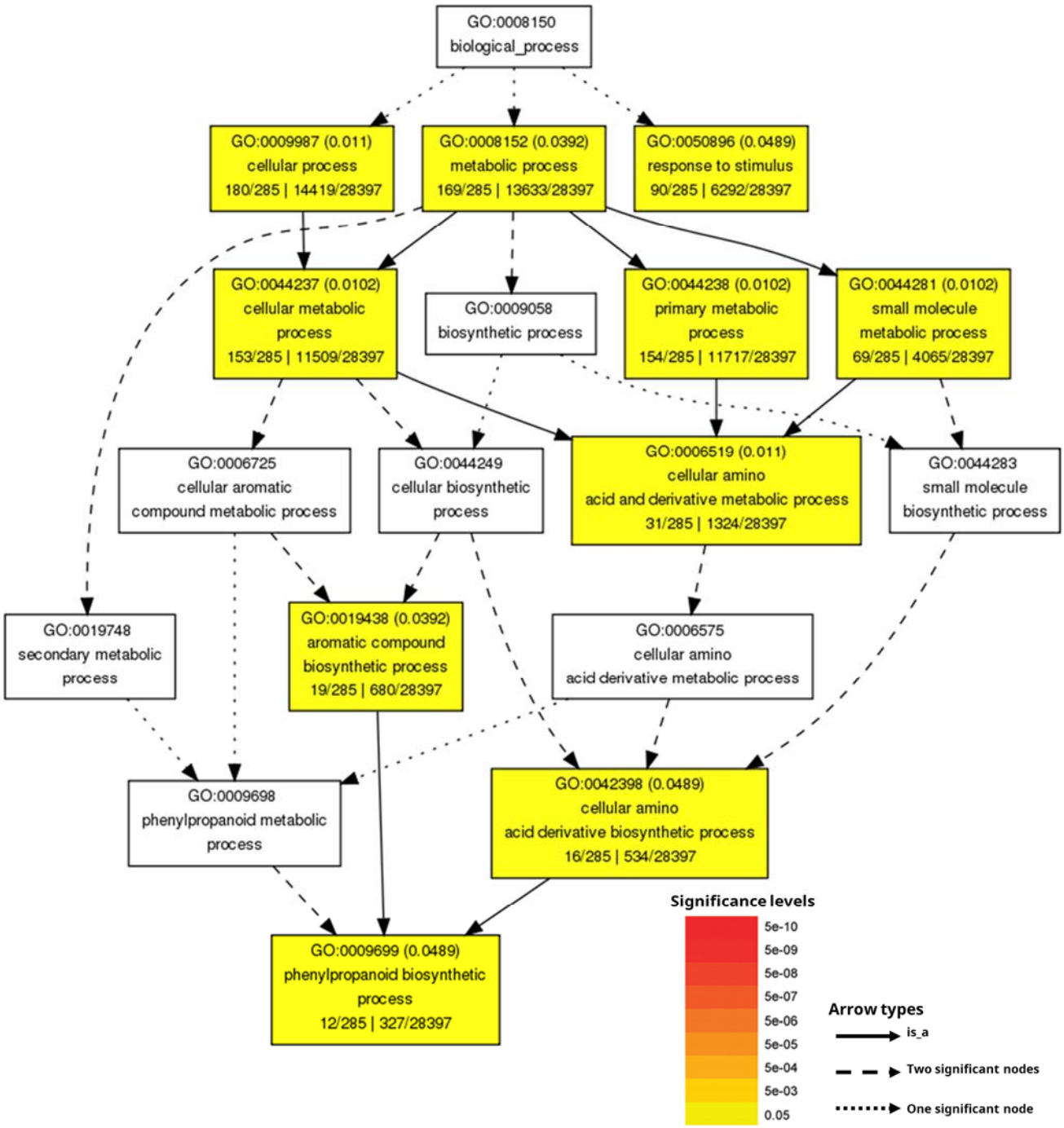
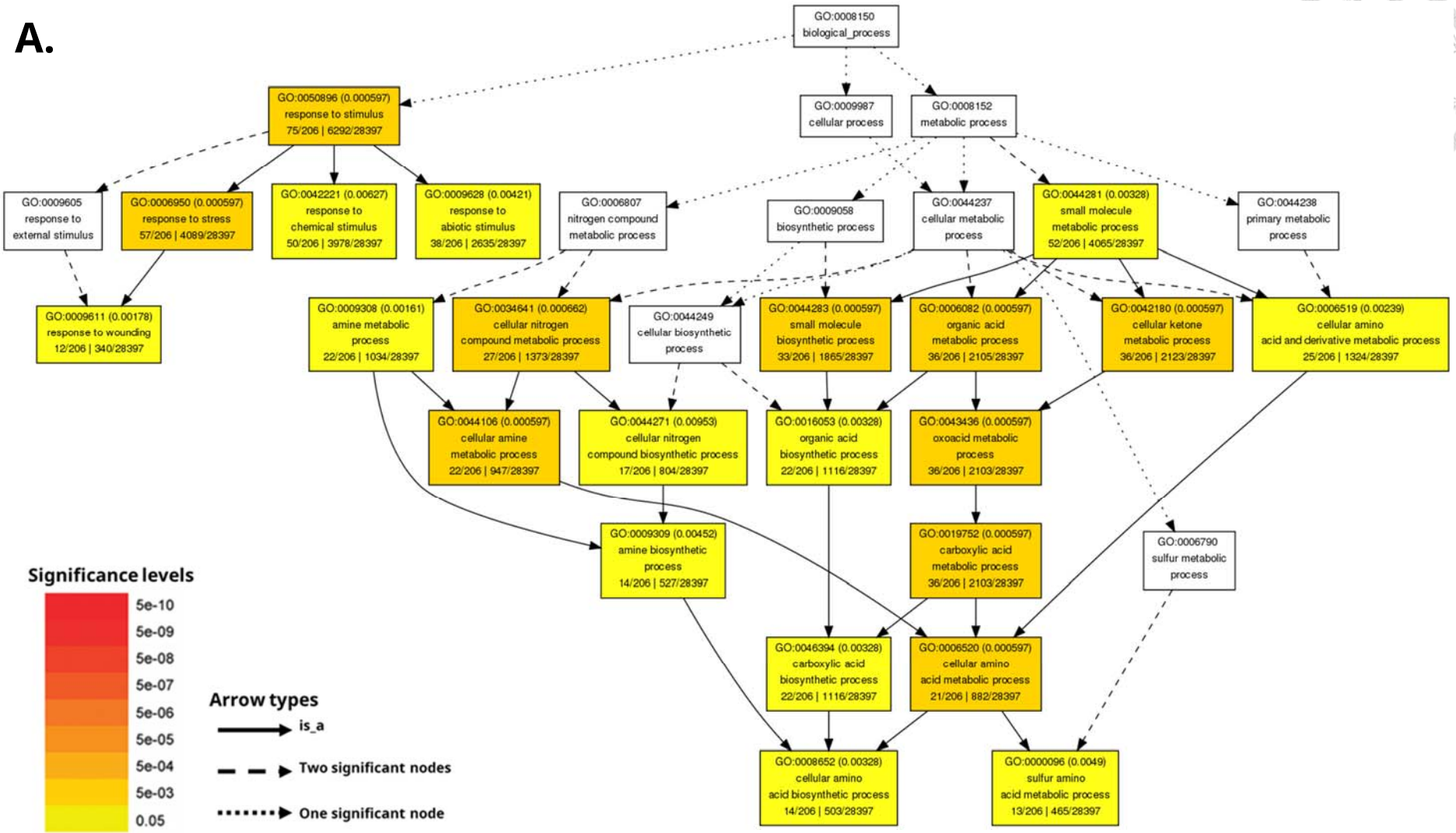


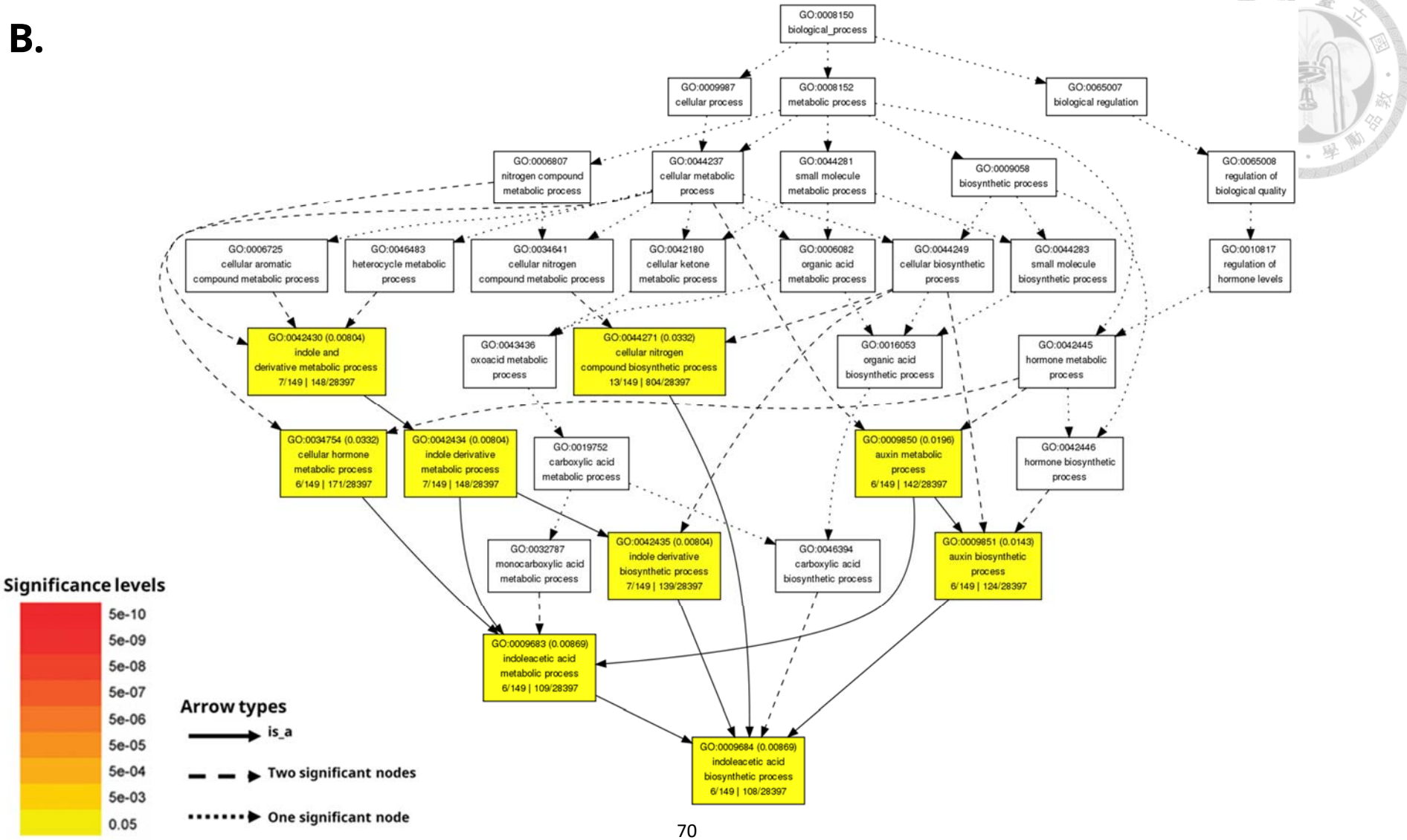
Figure 13. Singular enrichment analysis of DEGs from '228'-specific greenyellow module with AgriGO. GO term number, the *p*-value in parenthesis, and GO term were shown in each box. The pair of numerals in the bottom-left of box represented number of input genes mapped with a particular GO term and total number of input genes. The pair of numerals in the bottom-right of box represented number of genes mapped with a particular GO term in *Arabidopsis* TAIR10 database and total number of *Arabidopsis* TAIR10 genes mapped with a particular GO term in *Arabidopsis* TAIR10 database. Color indicate levels of statistical significance: yellow - adjusted $p \leq 0.05$; orange - adjusted $p \leq 10e^{-5}$.

doi:10.6342/NTU201603060



doi:10.6342/NTU201603060

B.



doi:10.6342/NTU201603060

Figure 14. Singular enrichment analysis of DEGs from 'Fuyudori'-specific green (A) and purple (B) module with AgriGO. Hierarchical tree graph of green and purple module in biological process category with overrepresented GO terms. GO term number, the *p*-value in parenthesis, and GO term were shown in each box. The pair of numerals in the bottom-left of box represented number of input genes mapped with a particular GO term and total number of input genes. The pair of numerals in the bottom-right of box represented number of genes mapped with a particular GO term in *Arabidopsis* TAIR10 database and total number of *Arabidopsis* TAIR10 genes mapped with a particular GO term in *Arabidopsis* TAIR10 database. Color indicate levels of statistical significance: yellow - adjusted $p \leq 0.05$; orange - adjusted $p \leq 10e^{-5}$.

3.2.4 Identification and analysis of coexpression network hub gene sets

WGCNA is also a tool for constructing the gene coexpression network, in which each gene is assigned as a node and the coexpression relation between genes is represented by connecting lines (edges). According to the node connectivity, the gene with comparatively high level of connectivity is assigned as hub gene which plays a critical role in the regulatory pathway (Langfelder & Horvath 2008). In this study, the hub genes identified from the network were further compared with the genes with significantly EV (FDR <0.1) from earlier analysis. In the greenyellow module network, 21 of the nodes with top level of connectivity were assigned as the hub genes of this module (**Fig. 15**, filled with green yellow). Two of the hub genes were identified as transcription factors, including a *zinc finger homeodomain (zf-HD) protein* (*Homeobox protein 28 (HB28)*/gene03215) and a *mitochondrial transcription termination factor family (mTERF) protein* (*mTERF*/gene39938). Further, *legume lectin family protein*

doi:10.6342/NTU201603060

(gene37752) and *Plant invertase/pectin methylesterase inhibitor (PMEI) superfamily protein* (gene02285) were also identified in earlier DESeq2 analysis. Pectin is one of the components composed of primary cell wall and related to plant growth, development, and defense response (Willats *et al.* 2001). Due to the demethylesterification activity of pectin methylesterase (PME), PMEs are involved in many physiological processes such as cell wall extension, cellular separation, seed germination, internode stem growth, root tip elongation, dry fruit dehiscence and soft fruit ripening (Al-Qsous *et al.* 2004, Wen *et al.* 1999). Therefore, the regulator of PMEs, pectin methylesterase inhibitor proteins (PMEIs), plays a necessary role in plant physiology processes and the function of PMEI-related proteins have been characterized in plant stress responses such as wounding, senescence, drought stress and osmotic stress (An *et al.* 2008, Greiner *et al.* 1998, 1999; Hong *et al.* 2010). As shown in **Fig. 16**, the expression trend of gene02285 encoding plant invertase/PMEI superfamily protein was largely down-regulated in HWS- intolerant 'Fuyudori' but not in '228'. Nevertheless, the most highly connected hub gene, legume lectin family protein (gene37752), was only reported on its carbohydrate binding structure and insecticidal activity in direct defense responses of plants (Vandenborre *et al.* 2011).

doi:10.6342/NTU201603060

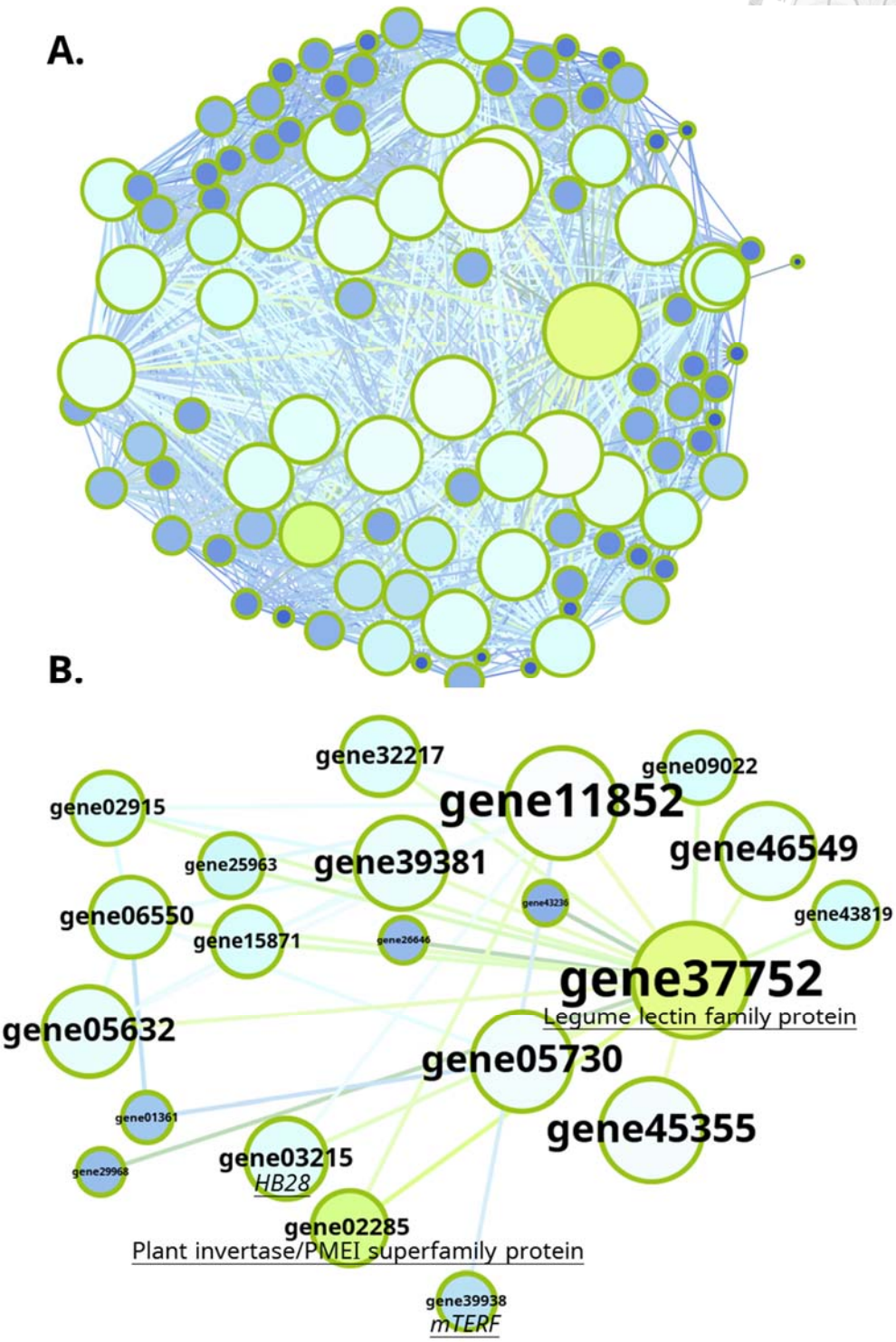


Figure 15. Correlation network of '228'-specific greenyellow module genes. (A) the coexpression network with all genes in the module. (B) the genes with highly connect number were assigned as hubs of the module. The circle size and filled color represented the number of connect line. The circle filled with greenyellow represented the gene with significantly EV (FDR <0.1).

doi:10.6342/NTU201603060

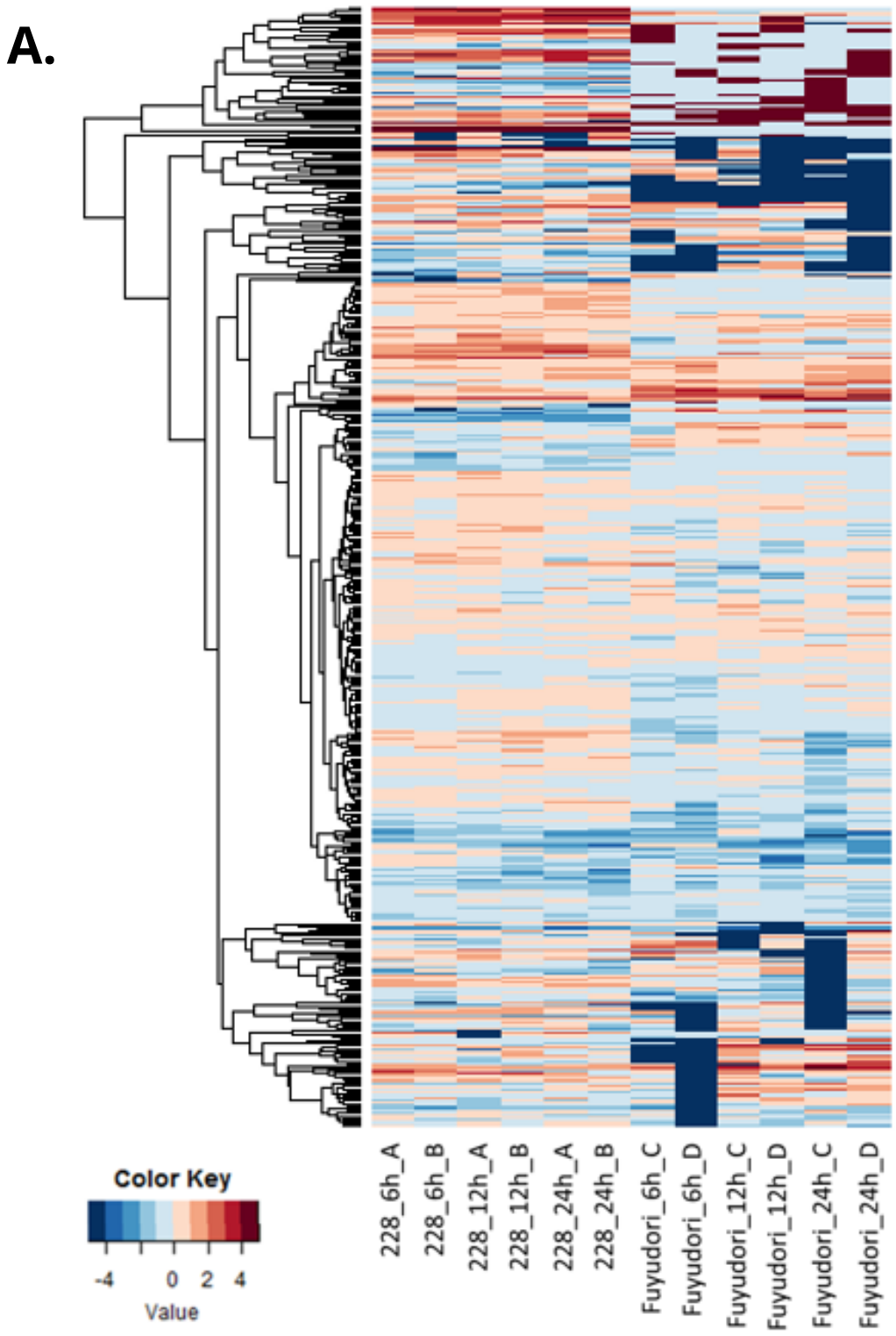


Figure 16. Heatmap showing the log_2 fold change value of '228'-specific greenyellow module genes. (A) Expression profile of all genes in the module.

doi:10.6342/NTU201603060

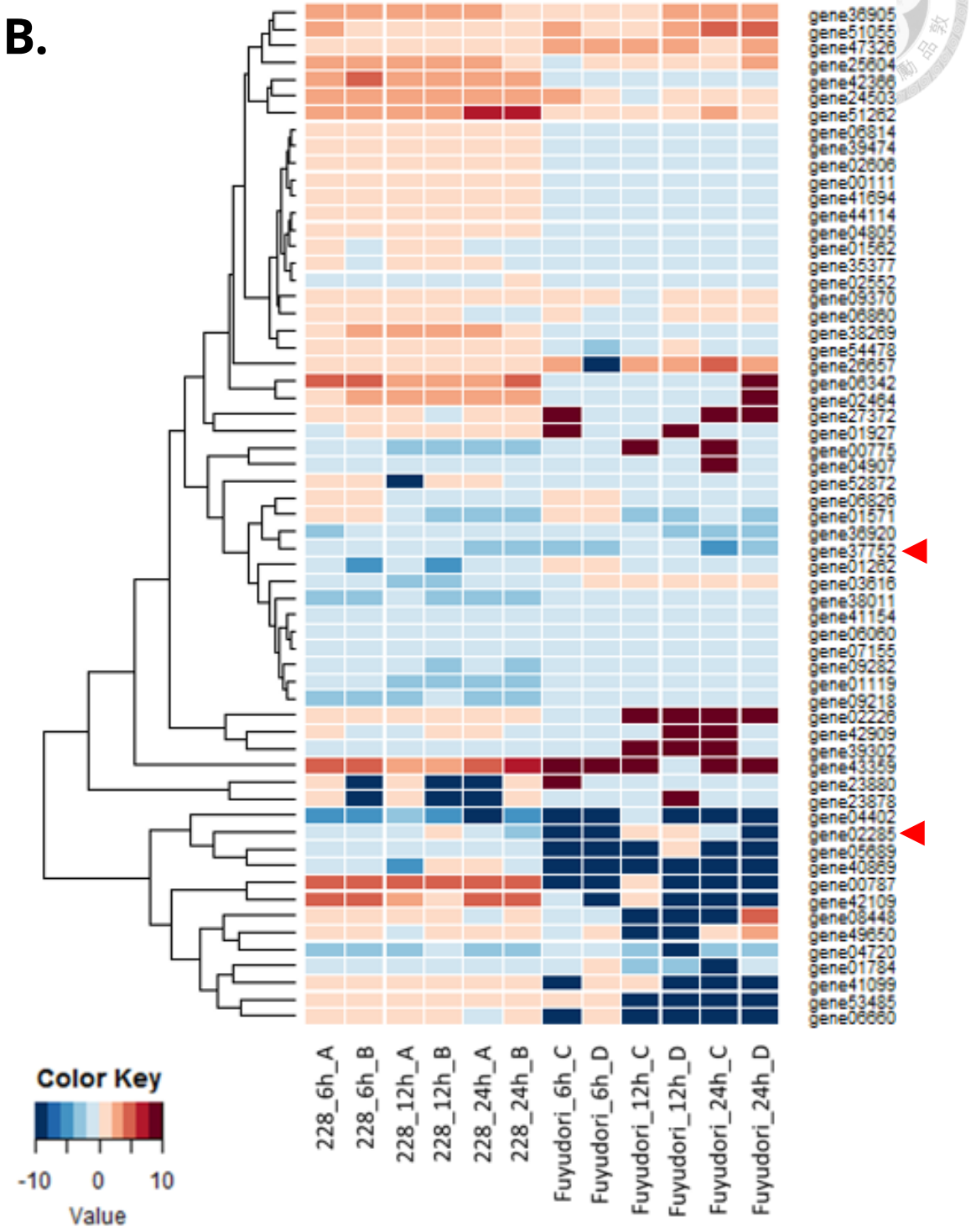
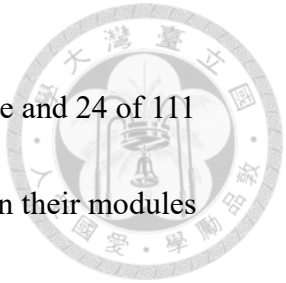


Figure 16. Heatmap showing the log$_2$ fold change value of '228'-specific greenyellow module genes (continuous). (B) Expression profile of the gene with significantly EV (FDR <0.1). Color key was adjusted to minimal and maximal log$_2$ fold change values to enhance the difference. Red triangle indicated the hub genes.

doi:10.6342/NTU201603060



In the HWS-intolerant modules, 20 of 299 genes in green module and 24 of 111 genes in purple module were individually assigned as the hub genes in their modules (**Fig. 17** and **18**). Seven hub genes in green module were also the significant EV genes (FDR <0.1), including *RING-H2 finger C2A* (*RHC2A*/gene10324), *actin 7*(gene16181), *glutamate-cysteine ligase* (*GCL*/gene13877), *polyadenylate-binding protein 2 isoform X1* (*PABP2-like isoform X1*/gene24594), *keratin_ type II cytoskeletal 1-like* (*KRT2*/gene37157), and 2 uncharacterized proteins (gene24656 and gene27811) (**Fig. 17B**, filled with green). The gene24656 was the hub gene with the highest edge number and continuously up-regulated in 'Fuyudori' but not '228' (**Fig. 19**). This may reveal that this uncharacterized protein was one of key regulators for HWS tolerance. While, physiological mechanism of these genes in plants is not clear, more information of these genes is necessary for identifying their functions in HWS. On the other hand, a third of 24 hub genes in purple module had the significant EV (**Fig. 18B**, filled with purple). A G2-like family transcription factor, *MYB-like transcription factor family protein* (gene17153), was found from the hubs with significant EV and persistently up-regulated in 'Fuyudori' during HWS but had no changes in '228' (**Fig. 20**). MYB transcription factor family has been well characterized in many studies about responses to biotic and abiotic stresses, development, cell differentiation, and secondary metabolism in various

doi:10.6342/NTU201603060

plants (Ambawat *et al.* 2013). In *Arabidopsis*, the transcription factor *AtMYB2* has been identified as a key regulatory factor of *ADH1* induction under low oxygen stress (Hoeren *et al.* 1998). Also, the tolerance to freezing, drought, and salt stress can be improved by overexpression of *OsMYB3R-2* gene in transgenic *Arabidopsis* (Ambawat *et al.* 2013, Dai *et al.* 2007). Moreover, the hub genes with similar expression pattern of gene17153 were *late embryogenesis abundant protein* (*LEA*/gene31870), *cytochrome P450 family 96 subfamily A polypeptide* (gene25393), and *E3 Ubiquitin ligase family protein parkin-like* (*parkin*/gene20919) (**Fig. 20**). Among these genes, the expression of gene31870 and gene20919 were up-regulated only at 24 h of HWS treatment. The hub genes described above all involved in the responses and regulatory pathway to osmotic stress in various plants (Cho *et al.* 2008, Hong-Bo *et al.* 2005, Lee & Kim 2011, Shinozaki & Yamaguchi-Shinozaki 2007, Umezawa *et al.* 2006). However, the up-regulation of those genes was only in stress-intolerant cultivar in this study. It suggests that intolerant cultivar suffered serious low oxygen and osmotic stresses caused by HWS earlier than the tolerant one. Finally, the hub genes with highest edge numbers encode unknown protein or protein of unknown function (**Fig. 20**). These genes very portably play the key regulatory roles for the HWS tolerance in cabbage.

doi:10.6342/NTU201603060

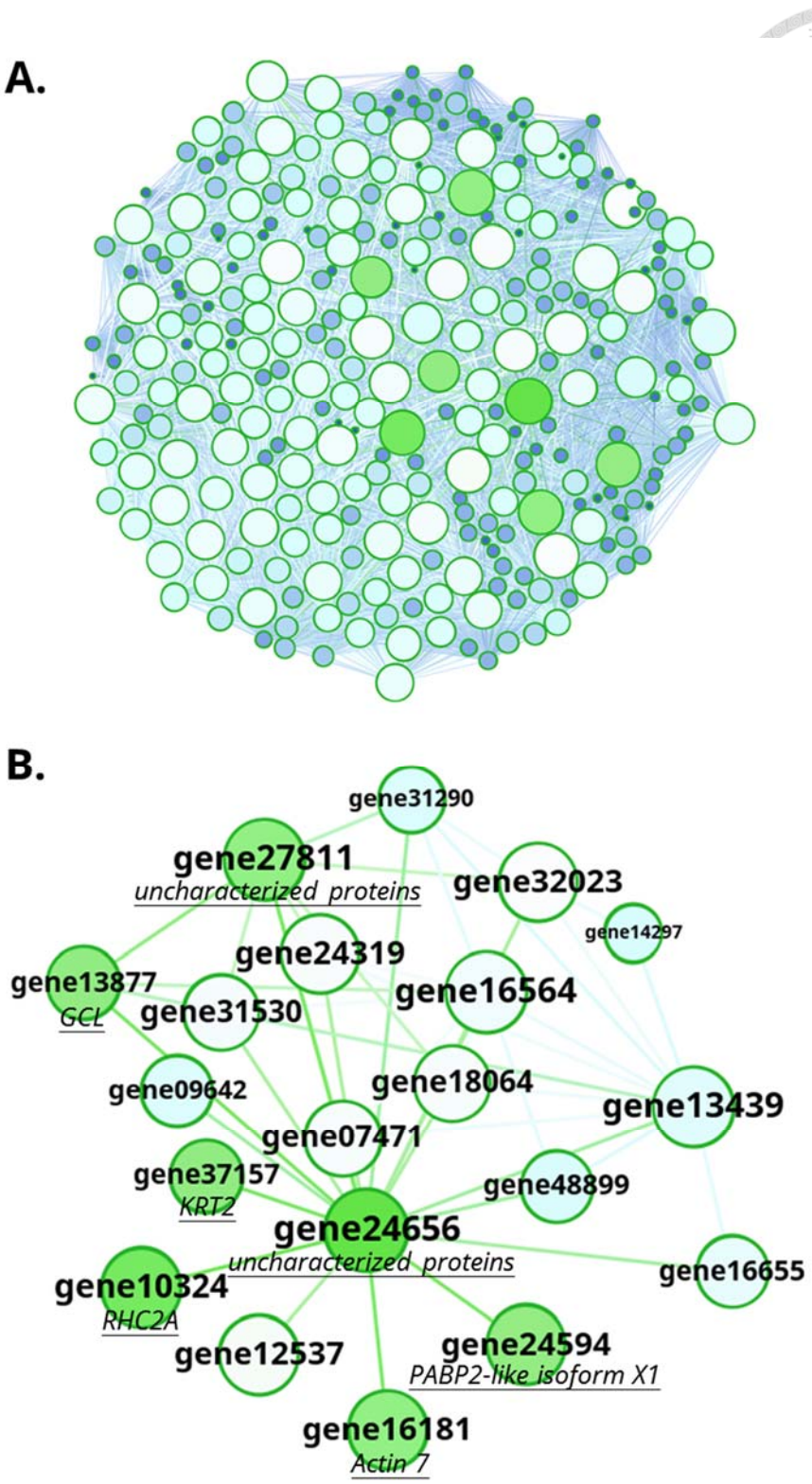


Figure 17. Correlation network of 'Fuyudori'-specific green module genes. (A)Coexpression network with all genes in the module. (B)Genes with high connect number were assigned as hubs of the module. The circle size and filled color represented the number of connect line. The circle filled with green represented the gene with significant EV (FDR <0.1).

doi:10.6342/NTU201603060

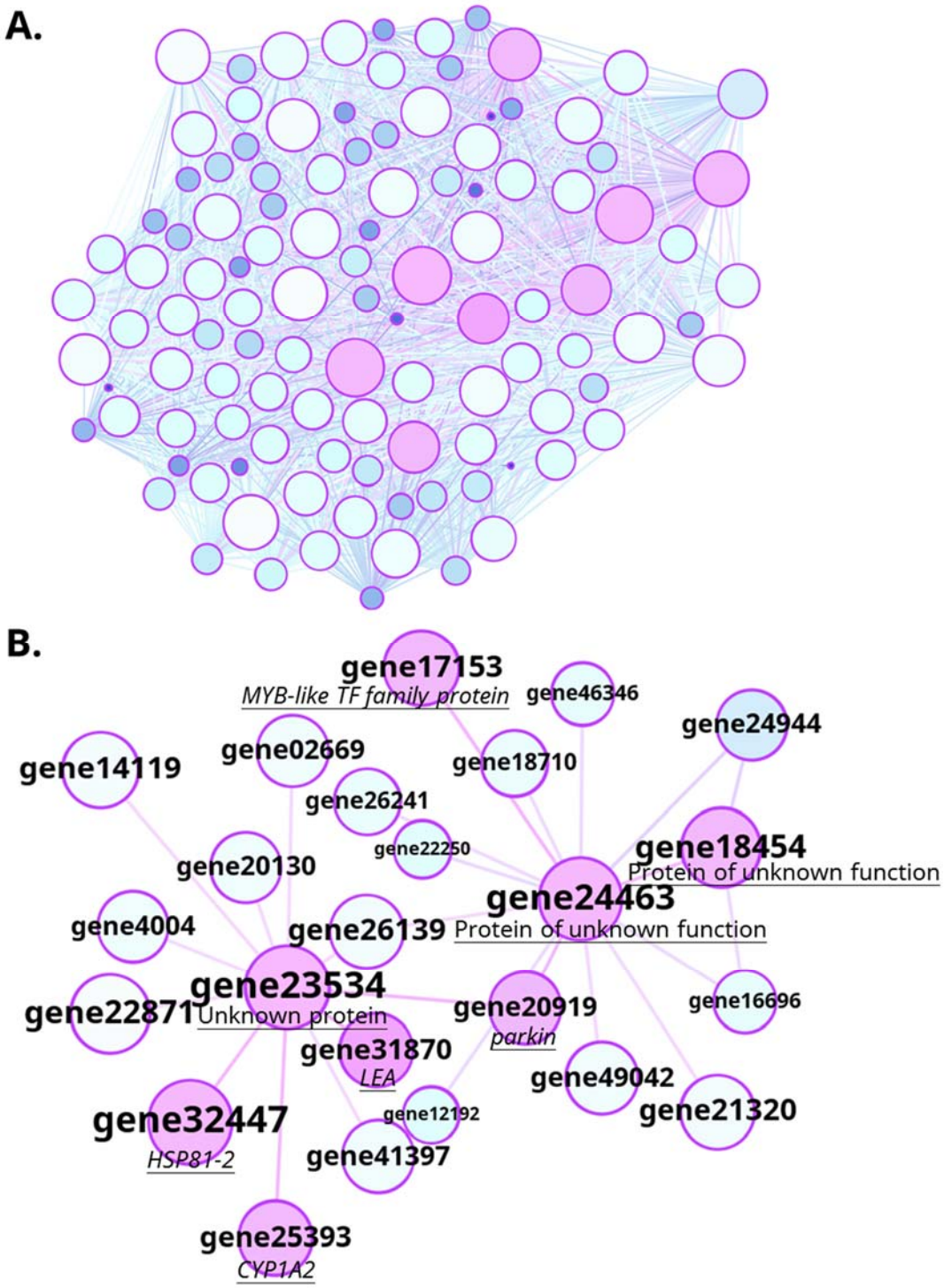


Figure 18. Correlation network of 'Fuyudori'-specific purple module genes. (A)Coexpression network with all genes in the module. (B)Genes with high connect number were assigned as hubs of the module. The circle size and filled color represented the number of connect line. The circle filled with purple represented the gene with significant EV (FDR <0.1).

doi:10.6342/NTU201603060

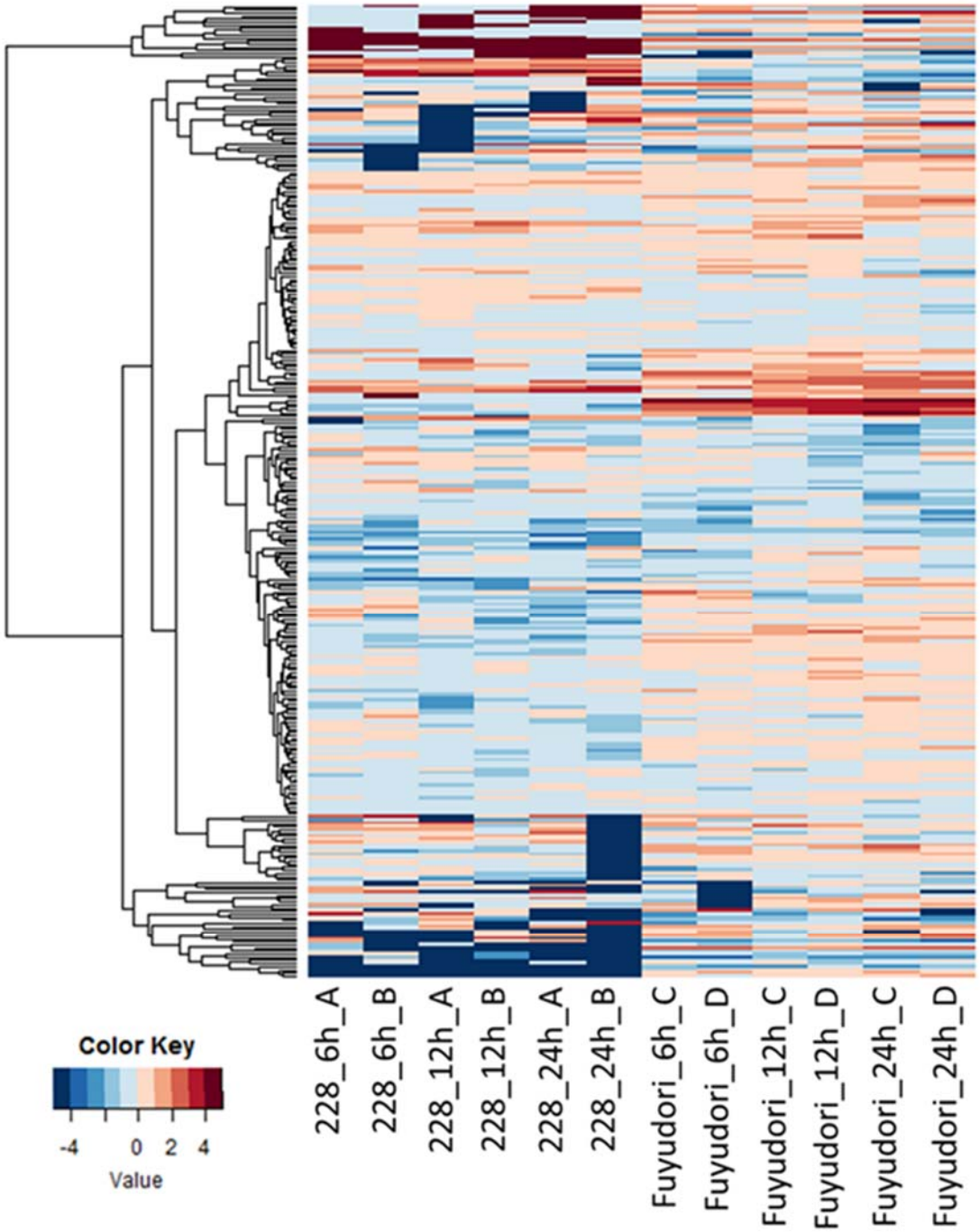


Figure 19. Heatmap showing the log_2 fold change value of 'Fuyudori'-specific green module genes. (A) Expression profile of all genes in the module. (B) Expression profile of the gene with significant EV (FDR <0.1). Color key was adjusted to minimal and maximal log_2 fold change values to enhance the difference.

doi:10.6342/NTU201603060

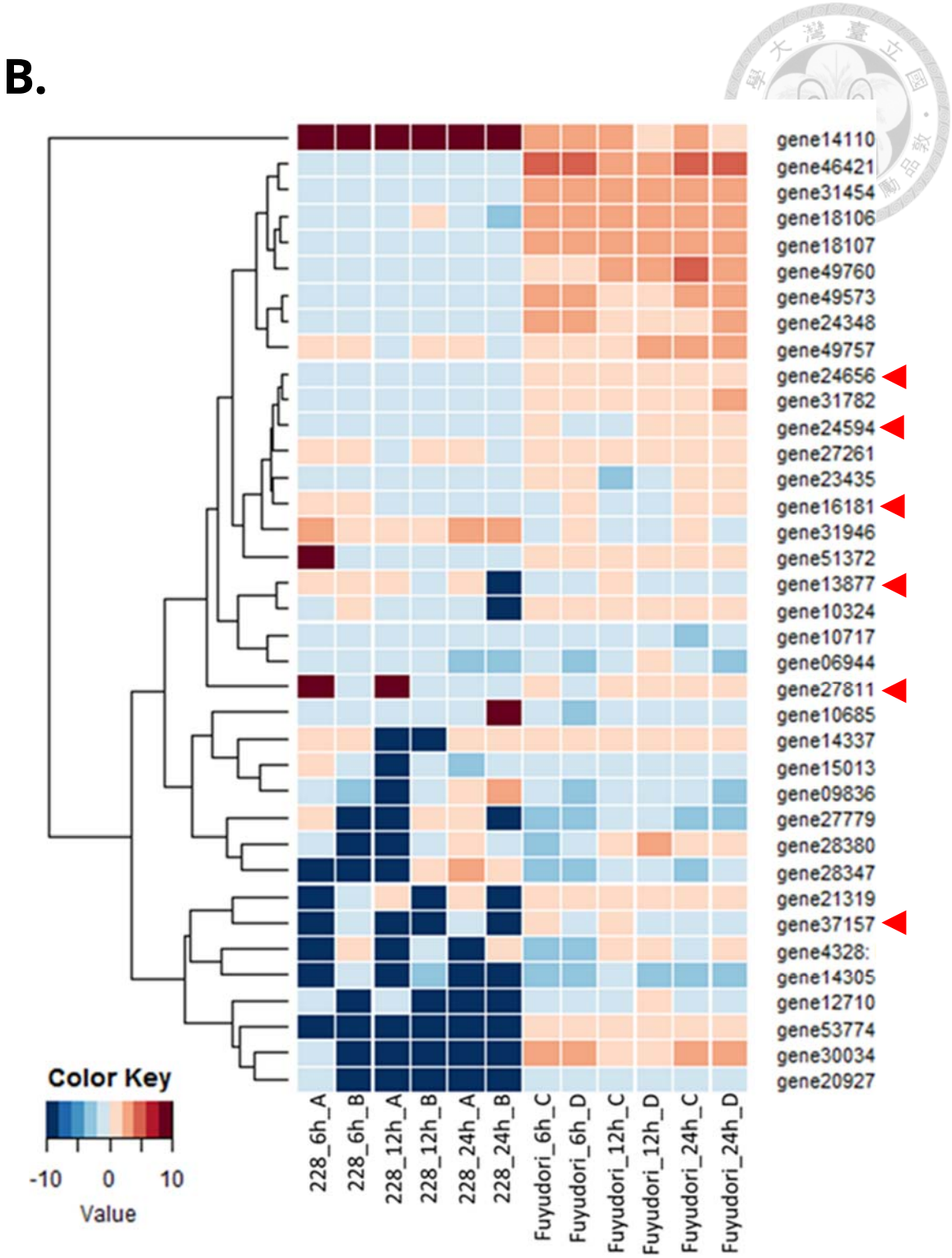


Figure 19. Heatmap showing the log$_2$ fold change value of 'Fuyudori'-specific green module genes (continuous). (B) Expression profile of the gene with significant EV (FDR <0.1). Color key was adjusted to minimal and maximal log$_2$ fold change values to enhance the difference. Red triangle indicated the hub genes.

doi:10.6342/NTU201603060

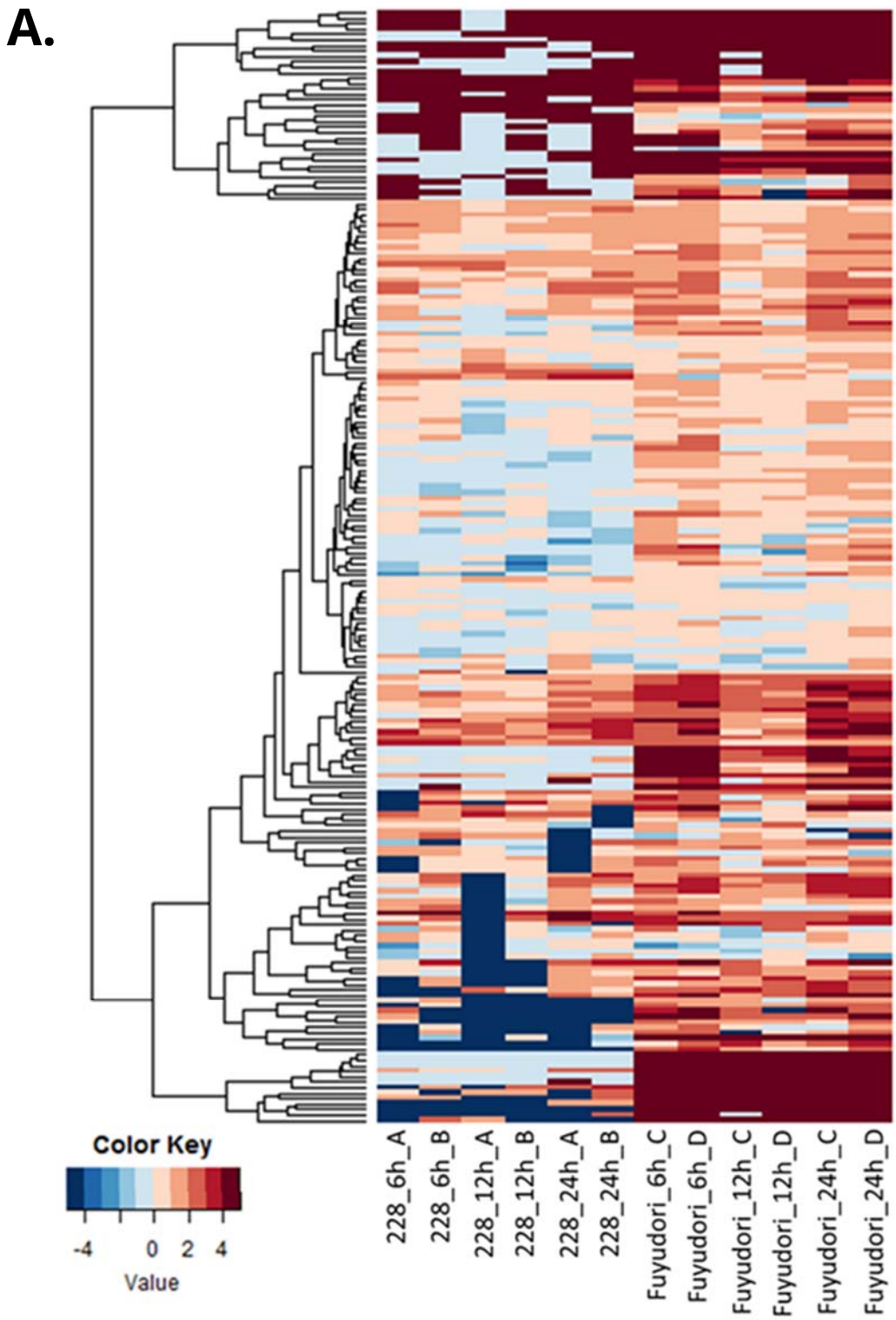


Figure 20. Heatmap showing the log_2 fold change value of 'Fuyudori'-specific purple module genes. (B) Expression profile of the gene with significant EV (FDR <0.1). Color key was adjusted to minimal and maximal log_2 fold change values to enhance the difference.

doi:10.6342/NTU201603060

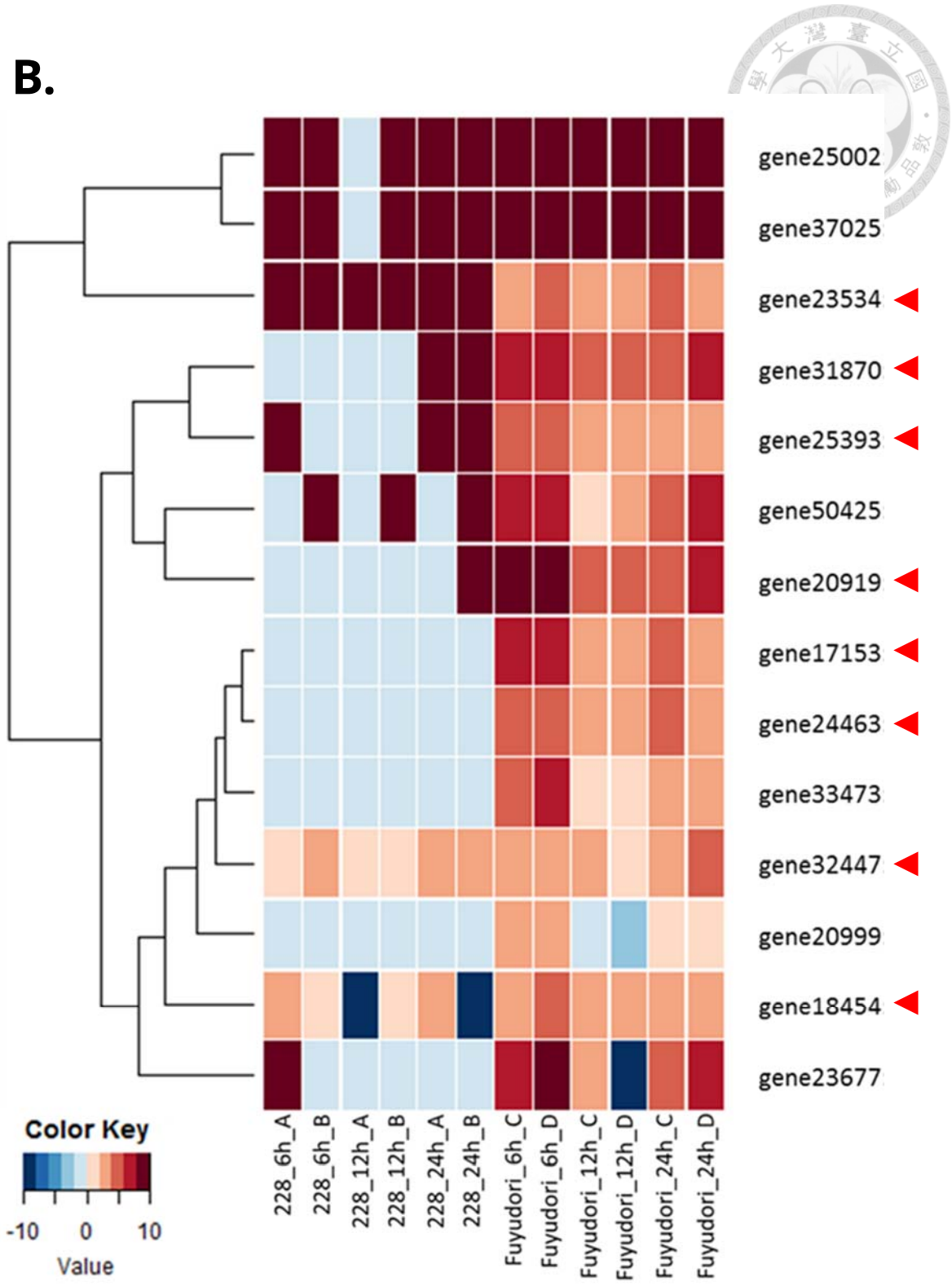


Figure 20. Heatmap showing the log_2 fold change value of 'Fuyudori'-specific purple module genes (continuous). (B) Expression profile of the gene with significant EV (FDR <0.1). Color key was adjusted to minimal and maximal log_2 fold change values to enhance the difference. Red triangle indicated the hub genes.

doi:10.6342/NTU201603060

4.3 Discussion

Cabbage is an important vegetable with a high demand in Taiwan throughout the year, but the typhoon-caused waterlogging stress in summer always results in heavy losses in the cabbage yield. The aim of this study was to discover the HWS tolerance-related gene sets in cabbage via the comparative transcriptome analysis of two cultivars with distinct HWS tolerance. Using HWS-tolerant cultivar '228' and HWS-intolerant cultivar 'Fuyudori', a time-course RNA-seq experiment was performed. Two different bioinformatic methods including primary co-expression measure with hierarchical clustering and WGCNA was combined to analyze the transcriptome data. This combination approach provided the advantage to utilize the relative expression value of genes with significant EV for examining the reliability of hub genes from the accepted modules. Meanwhile, systematic data mining using WGCNA was especially productive for discovering stress tolerance-specific modules and the critical hub genes in the mechanism of HWS tolerance.

By primary co-expression measure and enrichment analysis of the '228'-specific modules and clusters, many up-regulated genes were enriched in the terms about secondary metabolism such as phenylpropanoids and flavonoids biosynthesis (**Tab. 2 and Fig. 13**). The expression trends of these genes were totally distinct between '228'

doi:10.6342/NTU201603060

and ‘Fuyudori’, and the read counts and relative expression values of some genes in ‘228’-specific modules were not only higher than in ‘Fuyudori’ during HWS but also before HWS (0 h) (**Fig. 10** and **18**; **Supplemental Fig. 3**). Genes in HWS tolerance related to phenolic compound biosynthesis may need to continuously express in plant and especially up-regulate if stress happens. Most of the studies on plant stress responses mention the phenolic compound-related increase in amounts or/and gene expressions and consider the secondary metabolites as the indicator of stress tolerance (Akula & Ravishankar 2011, Bita & Gerats 2013, Nacif de Abreu & Mazzafera 2005, Wahid *et al.* 2007). Furthermore, the stress tolerant cultivar with higher phenolic compound content and related enzyme activity than intolerant one has been described in studies with variety comparative analysis on maize and cherry tomato (Hura *et al.* 2008, Sánchez-Rodríguez *et al.* 2011). Therefore, the appropriate application of exogenous salicylic acid for increasing the accumulation of phenolic compounds and activities of secondary metabolic enzymes to improve the stress tolerance has been demonstrated in many plants (Arfan *et al.* 2007, He *et al.* 2005, Miura & Tada 2014). Our work reported here also provided a novel evidence on the relation between phenolic metabolism and HWS tolerance from molecular level **(Fig. 21)**.

doi:10.6342/NTU201603060

On the other hand, during the time-course transcriptome analysis on cultivars with distinct tolerance, a significant enrichment of GO term 'Response to water deprivation' was noted in a 'Fuyudori' supercluster was noted (**Fig. 10**). This term was also enriched in our pervious study about the HWS treatment in cabbage 'Shia Feng No.1', a cultivar with the tolerance to heat and waterlogging stress but intolerance to HWS (**Fig. 4)**. Due to the cultivar difference between heat-intolerant 'Fuyudori' and heat-tolerant '228', the water deprivation following HWS may not be caused by high temperature independently. Under waterlogging stress, O_2 deficiency can induce cytoplasmic acidification in root cells to decrease the hydraulic conductivity of roots and then causes wilting of shoot organs in plants (Colmer & Voesenek 2009, Holbrook & Zwieniecki 2003, Tournaire-Roux *et al.* 2003). In addition, previous study about the physiological responses to HWS on the seedling of different cabbage cultivars reports that the HWS treatment on 'Fuyudori' for 48 h resulted in high stomatal conductance and membrane injury index but lower relative water content than '228' (Chen, 2011). Moreover, the down-regulation of ABA-related genes including *ABI5*, *ABF3*, *AFP1*, and *AFP4* were observed in 'Shia Feng No.1' under HWS for 12 h (**Fig. 2**). Previous studies have indicated that dehydration, high salinity and ABA treatment can induce the expression of bZIP subfamiliy members like *AREB1/ABF2*, *AREB2/ABF4* and *ABF3*. The

doi:10.6342/NTU201603060

expression of these factor enhances dehydration tolerance (Fujita *et al*. 2005, Nakashima & Yamaguchi-Shinozaki 2013). Taken together, HWS causes the down-regulation of ABA-related genes to induce the stomatal opening and further results in serious water deprivation due to incontrollable transpiration and lowered hydraulic conductivity of roots.

doi:10.6342/NTU201603060

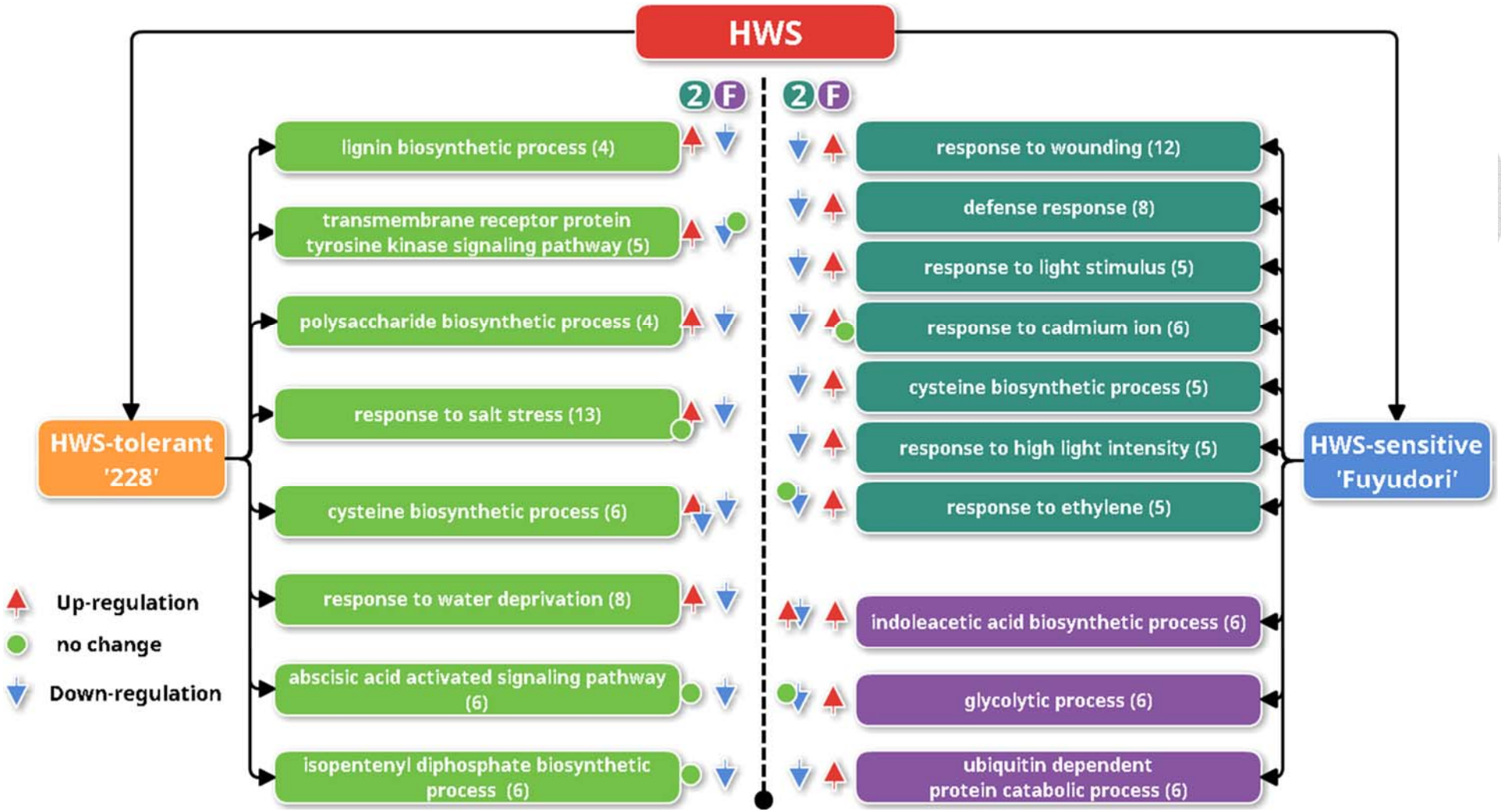


Figure 21. Overview of the complex regulatory networks in response to waterlogging stress at high temperature in stress tolerant and intolerant cabbage cultivars. GO terms were shown in each box and the number of genes involving in the GO term was shown in bracket. The arrows under letter '2' and 'F' in circle mean the expression pattern of gene in HWS-tolerant cultivar '228' and HWS-intolerance cultivar 'Fuyudori', respectively.

doi:10.6342/NTU201603060

Chapter 5 General conclusion and future directions

In conclusion, waterlogging at high temperature is a unique abiotic stress for plants and multiple regulatory pathways are involved with tolerance mechanism. Current study presents the first systematic investigation of the molecular mechanism in cabbage for the waterlogging stress at high temperature. Furthermore, the stress-tolerant cultivar simultaneously employs different regulatory pathways such as ABA, heat shock protein, and secondary metabolites for adapting to this stress (**Fig. 22**). For stress tolerance, the accumulation of secondary metabolites is important not only under stress but also in ordinary. This research also provides valuable information regarding the mechanism in combination of high temperature and waterlogging stress on plant and clarify the major factors for stress intolerance.

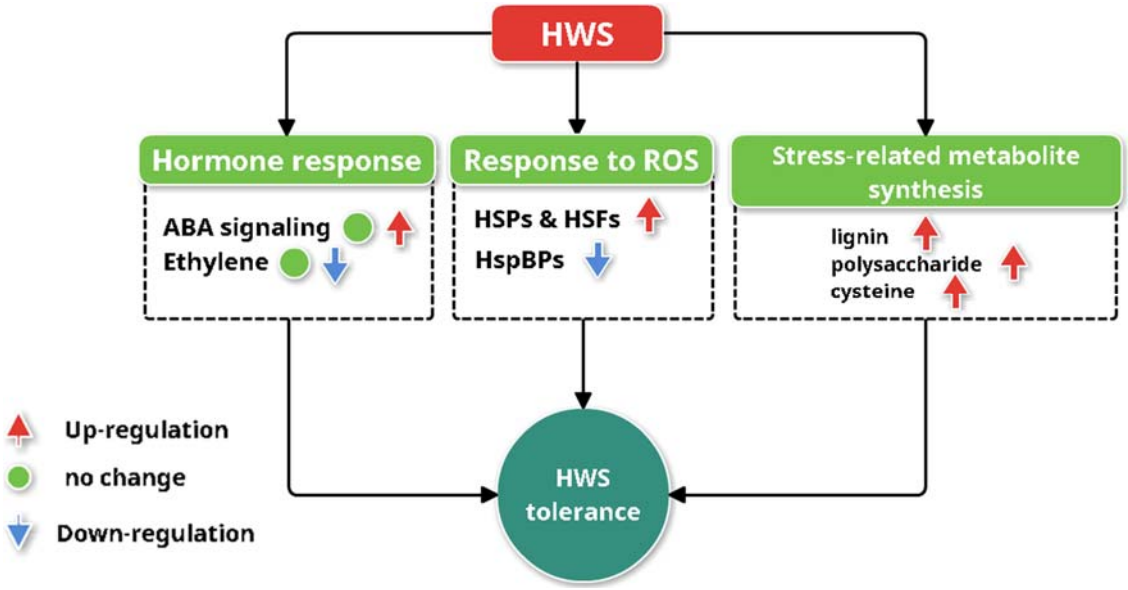
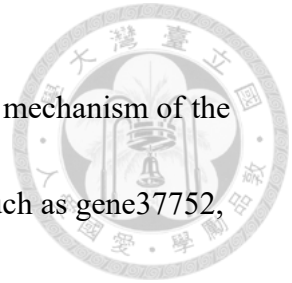


Figure 22. Summary of the stress tolerance-related pathway in HWS.

doi:10.6342/NTU201603060

However, the present work lacks the physiological evidences to link and confirm the different expression of the genes in various regulatory pathways. The regulation of hormones such as ABA and ethylene has been found that play a critical role for HWS tolerance in cabbage in this study. The quantitative analysis of these hormones during HWS is recommended and coordinated other physiological indexes such as transpiration rate, stomatal conductance, and water content of leaf to analyze the correlation with related gene expression. The combination of physiological response data and WGCNA could use to address the hub genes of hormone production in detail. Furthermore, using genome sequencing data to discover the promoters which regulate those hormone responses will be speed up the acquaintance of the HWS tolerance. Obtaining gene sequences is one of the advantages for using sequencing-based methods to analyze the gene expression pattern on non-model plants. Through traditional transgene, gene silencing, and genome edit technique, modifying the sequence and expression pattern of HWS-related genes and promotors may be used to confirm the role of these genes on regulatory mechanism of HWS tolerance. On the other hand, applying the ethylene inhibitor such as 1-methylcyclopropene (1-MCP) and aminoethoxyvinylglycine (AVG) and ABA will be one of efficient methods to verify the HWS tolerance in field.

doi:10.6342/NTU201603060



Moreover, it remains uncertain how the function and regulatory mechanism of the observed hub genes work under waterlogging at high temperature, such as gene37752, gene24656, gene27811, gene23534, gene24463, and gene18454. Although the genome sequencing of cabbage (*B. oleracea* var. *capitata L.*) has been finished, its genome annotation remains incomplete about the actual function of uncharacterized genes. Moreover, although a number of important genes have been shown in this work that they are significant differences on the gene expressions among the treatment and cultivars, the number of quantitative genes is too small to represent the profile of whole transcriptome. Deeper sequencing remains to be performed for building the complete transcriptome database of cabbage responded to waterlogging at high temperature. Future research should include the integrity of the genome to further probe the DNA sequence and structure on the genome.

Recently, genome editing technique is flourishing and very possible replaces the traditional methods used in breeding. Complete constructing the database of stress tolerance-related gene sequences and components is the trend in the future breeding. Using NGS technique to investigate gene expression pattern in crop under stresses could obtain the sequences of interested genes in the same time. In addition, traditional crop breeding dependents on the filtration of interested characters in crop but do not

doi:10.6342/NTU201603060

known which genes control the characters. Even through marker-assisted selection, it is limited for combination all the interested characters in one cultivar. In contrast to the breeding methods descried above, genome editing technique has the possibility to select and combine all the interesting genes in object cultivar. Genome-edited crops were considered not to belong genetically modified (GM) crops by Food and Drug Administration (FDA) in USA recently. This crop breeding method could be a powerful tool in crop improvement. Nevertheless, the complete sequences of interested genes are necessary for genome edited crop breeding. Using RNA-seq, clarifying the relation of physiological character and the distinct on gene sequences will be important as the future crop breeding materials.

doi:10.6342/NTU201603060

Reference

Aggarwal PK, Kalra N, Chander S, Pathak H. 2006. InfoCrop: A dynamic simulation model for the assessment of crop yields, losses due to pests, and environmental impact of agro-ecosystems in tropical environments. I. Model description. Agric. Syst. 89(1):1-25

Ahsan N, Lee D-G, Lee S-H, Lee K-W, Bahk JD, Lee B-H. 2007. A proteomic screen and identification of waterlogging-regulated proteins in tomato roots. Plant Soil. 295(1-2):37-51

Akula R, Ravishankar GA. 2011. Influence of abiotic stress signals on secondary metabolites in plants. Plant Signal. Behav. 6(11):1720-31

Almeselmani M, Deshmukh PS, Sairam RK, Kushwaha SR, Singh TP. 2006. Protective role of antioxidant enzymes under high temperature stress. Plant Sci. 171(3):382-88

Al-Qsous S, Carpentier E, Klein-Eude D, Burel C, Mareck A, *et al.* 2004. Identification and isolation of a pectin methylesterase isoform that could be involved in flax cell wall stiffening. Planta. 219(2):369-78

Ambawat S, Sharma P, Yadav NR, Yadav RC. 2013. MYB transcription factor genes as regulators for plant responses: an overview. Physiol. Mol. Biol. Plants. 19(3):307-21

Amrine KCH, Blanco-Ulate B, Cantu D. 2015. Discovery of Core Biotic Stress Responsive Genes in *Arabidopsis* by Weighted Gene Co-Expression Network Analysis. PLOS ONE. 10(3):e0118731

doi:10.6342/NTU201603060

An SH, Sohn KH, Choi HW, Hwang IS, Lee SC, Hwang BK. 2008. Pepper pectin methylesterase inhibitor protein CaPMEI1 is required for antifungal activity, basal disease resistance and abiotic stress tolerance. Planta. 228(1):61-78

Annamalai P, Yanagihara S. 1999. Identification and Characterization of a Heat-stress Induced Gene in Cabbage Encodes a Kunitz Type Protease Inhibitor. J. Plant Physiol. 155(2):226-33

Apel K, Hirt H. 2004. REACTIVE OXYGEN SPECIES: Metabolism, Oxidative Stress, and Signal Transduction. Annu. Rev. Plant Biol. 55(1):373-99

Arbona V, Gómez-Cadenas A. 2008. Hormonal Modulation of Citrus Responses to Flooding. J. Plant Growth Regul. 27(3):241-50

Arbona V, Manzi M, Ollas C de, Gómez-Cadenas A. 2013. Metabolomics as a Tool to Investigate Abiotic Stress Tolerance in Plants. Int. J. Mol. Sci. 14(3):4885-4911

Arfan M, Athar HR, Ashraf M. 2007. Does exogenous application of salicylic acid through the rooting medium modulate growth and photosynthetic capacity in two differently adapted spring wheat cultivars under salt stress? J. Plant Physiol. 164(6):685-94

Aroca R, Porcel R, Ruiz-Lozano JM. 2012. Regulation of root water uptake under abiotic stress conditions. J. Exp. Bot. 63(1):43-57

Bai T, Yin R, Li C, Ma F, Yue Z, Shu H. 2010. Comparative Analysis of Endogenous Hormones in Leaves and Roots of Two Contrasting *Malus* Species in Response to Hypoxia Stress. J. Plant Growth Regul. 30(2):119-27

Banga M, Slaa EJ, Blom C, Voesenek L. 1996. Ethylene Biosynthesis and Accumulation under Drained and Submerged Conditions (A Comparative Study of Two *Rumex* Species). Plant Physiol. 112(1):229-37

doi:10.6342/NTU201603060

Barak S, Yadav NS, Khan A. 2014. DEAD-box RNA helicases and epigenetic control of abiotic stress-responsive gene expression. Plant Signal. Behav. 9(12):e977729

Baron KN, Schroeder DF, Stasolla C. 2012. Transcriptional response of abscisic acid (ABA) metabolism and transport to cold and heat stress applied at the reproductive stage of development in *Arabidopsis* thaliana. Plant Sci. 188-189:48-59

Baxter A, Mittler R, Suzuki N. 2014. ROS as key players in plant stress signalling. J. Exp. Bot. 65(5):1229-40

Bita C, Gerats T. 2013. Plant tolerance to high temperature in a changing environment: scientific fundamentals and production of heat stress-tolerant crops. Crop Sci. Hortic. 4:273

Central Weather Bureau. http://www.cwb.gov.tw.

Chauhan YS, Senboku T. 1996. Thermostabilities of cell-membrane and photosynthesis in cabbage cultivars differing in heat tolerance. J. Plant Physiol. 149(6):729-34

Cheng W-H, Chiang M-H, Hwang S-G, Lin P-C. 2009. Antagonism between abscisic acid and ethylene in *Arabidopsis* acts in parallel with the reciprocal regulation of their metabolism and signaling pathways. Plant Mol. Biol. 71(1-2):61-80

Chen P-H. 2011, Physiological Response of Cabbage (*Brassica oleracea L.* var. *capitata L.*) to High Temperature and Waterlogging. Master Thesis of Department of Horticulture. National Taiwan University. 109 pp.

Cho SK, Ryu MY, Song C, Kwak JM, Kim WT. 2008. *Arabidopsis* PUB22 and PUB23 are homologous U-Box E3 ubiquitin ligases that play combinatory roles in response to drought stress. Plant Cell. 20(7):1899-1914

doi:10.6342/NTU201603060

Colmer TD, Voesenek LACJ. 2009. Flooding tolerance: suites of plant traits in variable environments. Funct. Plant Biol. 36(8):665-81

Dai X, Xu Y, Ma Q, Xu W, Wang T, *et al.* 2007. Overexpression of an R1R2R3 MYB Gene, OsMYB3R-2, Increases Tolerance to Freezing, Drought, and Salt Stress in Transgenic *Arabidopsis*. Plant Physiol. 143(4):1739-51

Dixon R, Paiva N. 1995. Stress-Induced Phenylpropanoid Metabolism. Plant Cell. 7(7):1085-97

Du Z, Zhou X, Ling Y, Zhang Z, Su Z. 2010. agriGO: a GO analysis toolkit for the agricultural community. Nucleic Acids Res. 38(suppl 2):W64-70

Dugas D, Monaco M, Olson A, Klein R, Kumari S, *et al.* 2011. Functional annotation of the transcriptome of Sorghum bicolor in response to osmotic stress and abscisic acid. BMC Genomics. 12(1):514

Forster SC, Finkel AM, Gould JA, Hertzog PJ. 2013. RNA-eXpress annotates novel transcript features in RNA-seq data. Bioinformatics. 29(6):810-12

Fujita Y, Fujita M, Satoh R, Maruyama K, Parvez MM, *et al.* 2005. AREB1 is a transcription activator of novel ABRE-dependent ABA signaling that enhances drought stress tolerance in *Arabidopsis*. Plant Cell. 17(12):3470-88

Fukao T, Bailey-Serres J. 2004. Plant responses to hypoxia - is survival a balancing act? Trends Plant Sci. 9(9):449-56

Gan Y, Angadi SV, Cutforth H, Potts D, Angadi VV, McDonald CL. 2004. Canola and mustard response to short periods of temperature and water stress at different developmental stages. Can. J. Plant Sci. 84(3):697-704

Gill SS, Tuteja N. 2010. Reactive oxygen species and antioxidant machinery in abiotic stress tolerance in crop plants. Plant Physiol. Biochem. 48(12):909-30

doi:10.6342/NTU201603060

Goulas E, Schubert M, Kieselbach T, Kleczkowski LA, Gardeström P, *et al.* 2006. The chloroplast lumen and stromal proteomes of *Arabidopsis thaliana* show differential sensitivity to short- and long-term exposure to low temperature. Plant J. 47(5):720-34

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, *et al.* 2011. Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. Nat. Biotechnol. 29(7):644-52

Greiner S, Krausgrill S, Rausch T. 1998. Cloning of a Tobacco Apoplasmic Invertase Inhibitor. Plant Physiol. 116(2):733-42

Greiner S, Rausch T, Sonnewald U, Herbers K. 1999. Ectopic expression of a tobacco invertase inhibitor homolog prevents cold-induced sweetening of potato tubers. Nat. Biotechnol. 17(7):708-11

Hasanuzzaman M, Nahar K, Alam MM, Roychowdhury R, Fujita M. 2013. Physiological, Biochemical, and Molecular Mechanisms of Heat Stress Tolerance in Plants. Int. J. Mol. Sci. 14(5):9643-84

He Y, Liu Y, Cao W, Huai M, Xu B, Huang B. 2005. Effects of Salicylic Acid on Heat Tolerance Associated with Antioxidant Metabolism in Kentucky Bluegrass. Crop Sci. 45(3):988-95

Hirayama T, Shinozaki K. 2010. Research on plant abiotic stress responses in the post-genome era: past, present and future. Plant J. 61(6):1041-52

Hiremath PJ, Farmer A, Cannon SB, Woodward J, Kudapa H, *et al.* 2011. Large-scale transcriptome analysis in chickpea (*Cicer arietinum L.*), an orphan legume crop of the semi-arid tropics of Asia and Africa. Plant Biotechnol. J. 9(8):922-31

doi:10.6342/NTU201603060

Hoeren FU, Dolferus R, Wu Y, Peacock WJ, Dennis ES. 1998. Evidence for a role for AtMYB2 in the induction of the *Arabidopsis* alcohol dehydrogenase gene (ADH1) by low oxygen. Genetics. 149(2):479-90

Holbrook NM, Zwieniecki MA. 2003. Plant biology: Water gate. Nature. 425(6956):361-361

Hong MJ, Kim DY, Lee TG, Jeon WB, Seo YW. 2010. Functional characterization of pectin methylesterase inhibitor (PMEI) in wheat. Genes Genet. Syst. 85(2):97-106

Hong-Bo S, Zong-Suo L, Ming-An S. 2005. LEA proteins in higher plants: Structure, function, gene expression and regulation. Colloids Surf. B Biointerfaces. 45(3-4):131-35

Hsu F-C, Chou M-Y, Chou S-J, Li Y-R, Peng H-P, Shih M-C. 2013. Submergence Confers Immunity Mediated by the WRKY22 Transcription Factor in *Arabidopsis*. Plant Cell. 25(7):2699-2713

Hura T, Hura K, Grzesiak S. 2008. Contents of Total Phenolics and Ferulic Acid, and PAL Activity during Water Potential Changes in Leaves of Maize Single-Cross Hybrids of Different Drought Tolerance. J. Agron. Crop Sci. 194(2):104-12

Irfan M, Hayat S, Hayat Q, Afroz S, Ahmad A. 2010. Physiological and biochemical changes in plants under waterlogging. Protoplasma. 241(1-4):3-17

Jiang Y, Huang B. 2001. Drought and Heat Stress Injury to Two Cool-Season Turfgrasses in Relation to Antioxidant Metabolism and Lipid Peroxidation. Crop Sci. 41(2):436

doi:10.6342/NTU201603060

Kakumanu A, Ambavaram MMR, Klumas CM, Krishnan A, Batlang U, *et al.* 2012. Effects of Drought on Gene Expression in Maize Reproductive and Leaf Meristem Tissue Revealed by RNA-Seq. Plant Physiol.

Kang J, Choi H, Im M, Kim SY. 2002. *Arabidopsis* Basic Leucine Zipper Proteins That Mediate Stress-Responsive Abscisic Acid Signaling. Plant Cell. 14(2):343-57

Kumar S, Kaur R, Kaur N, Bhandhari K, Kaushal N, *et al.* 2011. Heat-stress induced inhibition in growth and chlorosis in mungbean (*Phaseolus aureus Roxb*.) is partly mitigated by ascorbic acid application and is related to reduction in oxidative stress. Acta Physiol. Plant. 33(6):2091-2101

Langfelder P, Horvath S. 2008. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 9:559

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods. 9(4):357-59

Lasanthi-Kudahettige R, Magneschi L, Loreti E, Gonzali S, Licausi F, *et al.* 2007. Transcript profiling of the anoxic rice coleoptile. Plant Physiol. 144(1):218-31

Lee J-H, Kim WT. 2011. Regulation of abiotic stress signal transduction by E3 ubiquitin ligases in *Arabidopsis*. Mol. Cells. 31(3):201-8

Lee T, Yang S, Kim E, Ko Y, Hwang S, *et al.* 2015. AraNet v2: an improved database of co-functional gene networks for the study of *Arabidopsis thaliana* and 27 other nonmodel plant species. Nucleic Acids Res. 43(Database issue):D996-1002

Li G, Meng X, Wang R, Mao G, Han L, *et al.* 2012. Dual-Level Regulation of ACC Synthase Activity by MPK3/MPK6 Cascade and Its Downstream WRKY Transcription Factor during Ethylene Induction in *Arabidopsis*. PLOS Genet. 8(6):e1002767

doi:10.6342/NTU201603060

Licausi F, Van Dongen JT, Giuntoli B, Novi G, Santaniello A, *et al.* 2010. HRE1 and HRE2, two hypoxia-inducible ethylene response factors, affect anaerobic responses in *Arabidopsis thaliana*. Plant J. 62(2):302-15

Lin H-H, Lin K-H, Chen S-C, Shen Y-H, Lo H-F. 2015a. Proteomic analysis of broccoli (*Brassica oleracea)*. Bot. Stud. 56(1):1-11

Lin H-H, Lin K-H, Syu J-Y, Tang S-Y, Lo H-F. 2015b. Physiological and proteomic analysis in two wild tomato lines under waterlogging and high temperature stress. J. Plant Biochem. Biotechnol. 25(1):87-96

Lin KH, Chen LFO, Li SD, Lo HF. 2015c. Comparative proteomic analysis of cauliflower under high temperature and flooding stresses. Sci. Hortic. 183:118-29

Liseron-Monfils C, Ware D. 2015. Revealing gene regulation and associations through biological networks. Curr. Plant Biol. 3-4:30-39

Liu S, Liu Y, Yang X, Tong C, Edwards D, *et al.* 2014. The *Brassica oleracea* genome reveals the asymmetrical evolution of polyploid genomes. Nat. Commun. 5:3930

Loreti E, Poggi A, Novi G, Alpi A, Perata P. 2005. A Genome-Wide Analysis of the Effects of Sucrose on Gene Expression in *Arabidopsis* Seedlings under Anoxia. Plant Physiol. 137(3):1130-38

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15(12):

Metzker ML. 2010. Sequencing technologies — the next generation. Nat. Rev. Genet. 11(1):31-46

Michael TP, VanBuren R. 2015. Progress, challenges and the future of crop genomes. Curr. Opin. Plant Biol. 24:71-81

doi:10.6342/NTU201603060

Miller G, Schlauch K, Tam R, Cortes D, Torres MA, *et al.* 2009. The Plant NADPH Oxidase RBOHD Mediates Rapid Systemic Signaling in Response to Diverse Stimuli. Sci. Signal. 2(84):ra45

Miller G, Suzuki N, Rizhsky L, Hegie A, Koussevitzky S, Mittler R. 2007. Double Mutants Deficient in Cytosolic and Thylakoid Ascorbate Peroxidase Reveal a Complex Mode of Interaction between Reactive Oxygen Species, Plant Development, and Response to Abiotic Stresses. Plant Physiol. 144(4):1777-85

Min L, Li Y, Hu Q, Zhu L, Gao W, *et al.* 2014. Sugar and Auxin Signaling Pathways Respond to High-Temperature Stress during Anther Development as Revealed by Transcript Profiling Analysis in Cotton. Plant Physiol. 164(3):1293-1308

Mittler R. 2006. Abiotic stress, the field environment and stress combination. Trends Plant Sci. 11(1):15-19

Mittler R, Finka A, Goloubinoff P. 2012. How do plants feel the heat? Trends Biochem. Sci. 37(3):118-25

Miura K, Tada Y. 2014. Regulation of water, salinity, and cold stress responses by salicylic acid. Plant Physiol. 5:4

Morrell PL, Buckler ES, Ross-Ibarra J. 2012. Crop genomics: advances and applications. Nat. Rev. Genet. 13(2):85-96

Moustafa K, AbuQamar S, Jarrar M, Al-Rajab AJ, Trémouillaux-Guiller J. 2014. MAPK cascades and major abiotic stresses. Plant Cell Rep. 33(8):1217-25

Mutasa-Göttgens E, Joshi A, Holmes H, Hedden P, Göttgens B. 2012. A new RNASeq-based reference transcriptome for sugar beet and its application in transcriptome-scale analysis of vernalization and gibberellin responses. BMC Genomics. 13(1):99

doi:10.6342/NTU201603060

Nacif de Abreu I, Mazzafera P. 2005. Effect of water and temperature stress on the content of active constituents of *Hypericum brasiliense* Choisy. Plant Physiol. Biochem. 43(3):241-48

Nakashima K, Yamaguchi-Shinozaki K. 2013. ABA signaling in stress-response and seed development. Plant Cell Rep. 32(7):959-70

Narsai R, Rocha M, Geigenberger P, Whelan J, van Dongen JT. 2011. Comparative analysis between plant species of transcriptional and metabolic responses to hypoxia. New Phytol. 190(2):472-87

O'Rourke JA, Yang SS, Miller SS, Bucciarelli B, Liu J, *et al.* 2012. An RNA-Seq Transcriptome Analysis of Pi Deficient White Lupin Reveals Novel Insights Into Phosphorus Acclimation in Plants. Plant Physiol.

Ou LJ, Dai XZ, Zhang ZQ, Zou XX. 2011. Responses of pepper to waterlogging stress. Photosynthetica. 49(3):339-45

Parent C, Nicolas C, Audrey B, Crevècoeur M, Dat J. 2008. An overview of plant responses to soil waterlogging. Plant Stress. 20-27

Parkin IA, Koh C, Tang H, Robinson SJ, Kagale S, *et al.* 2014. Transcriptome and methylome profiling reveals relics of genome dominance in the mesopolyploid *Brassica oleracea*. Genome Biol. 15(6):R77

Priest HD, Fox SE, Rowley ER, Murray JR, Michael TP, Mockler TC. 2014. Analysis of Global Gene Expression in *Brachypodium distachyon* Reveals Extensive Network Plasticity in Response to Abiotic Stress. PLOS ONE. 9(1):e87499

Pucciariello C, Banti V, Perata P. 2012. ROS signaling as common element in low oxygen and heat stresses. Plant Physiol. Biochem. 59:3-10

doi:10.6342/NTU201603060

Qi X-H, Xu X-W, Lin X-J, Zhang W-J, Chen X-H. 2012. Identification of differentially expressed genes in cucumber (*Cucumis sativus L.*) root under waterlogging stress by digital gene expression profile. Genomics. 99(3):160-68

Qiao Z, Li C-L, Zhang W. 2016. WRKY1 regulates stomatal movement in drought-stressed *Arabidopsis thaliana*. Plant Mol. Biol. 1-13

Qu A-L, Ding Y-F, Jiang Q, Zhu C. 2013. Molecular mechanisms of the plant heat stress response. Biochem. Biophys. Res. Commun. 432(2):203-7

Reynolds-Henne CE, Langenegger A, Mani J, Schenk N, Zumsteg A, Feller U. 2010. Interactions between temperature, drought and stomatal opening in legumes. Environ. Exp. Bot. 68(1):37-43

Roberts A, Pachter L. 2013. Streaming fragment assignment for real-time analysis of sequencing experiments. Nat. Methods. 10(1):71-73

Sánchez-Rodríguez E, Moreno DA, Ferreres F, Rubio-Wilhelmi M del M, Ruiz JM. 2011. Differential responses of five cherry tomato varieties to water stress: Changes on phenolic metabolites and related enzymes. Phytochemistry. 72(8):723-29

Shabala S, Shabala L, Barcelo J, Poschenrieder C. 2014. Membrane transporters mediating root signalling and adaptive responses to oxygen deprivation and soil flooding. Plant Cell Environ. 37(10):2216-33

Shinozaki K, Yamaguchi-Shinozaki K. 2007. Gene networks involved in drought stress response and tolerance. J. Exp. Bot. 58(2):221-27

Snogerup S, Gustafsson M, Bothmer RV. 1990. *Brassica* sect. *Brassica* (Brassicaceae) I. Taxonomy and Variation. Willdenowia. 19(2):271-365

doi:10.6342/NTU201603060

Solecka D. Role of phenylpropanoid compounds in plant responses to different stress factors. Acta Physiol. Plant. 19(3):257-68

Suzuki N, Koussevitzky S, Mittler R, Miller G. 2012. ROS and redox signalling in the response of plants to abiotic stress. Plant Cell Environ. 35(2):259-270

Tang D, Christiansen KM, Innes RW. 2005. Regulation of Plant Disease Resistance, Stress Responses, Cell Death, and Ethylene Signaling in *Arabidopsis* by the EDR1 Protein Kinase. Plant Physiol. 138(2):1018-26

Thirunavukkarasu N, Hossain F, Mohan S, Shiriga K, Mittal S, *et al.* 2013. Genome-Wide Expression of Transcriptomes and Their Co-Expression Pattern in Subtropical Maize (*Zea mays L.*) under Waterlogging Stress. PLOS ONE. 8(8):e70433

Timperio AM, Egidi MG, Zolla L. 2008. Proteomics applied on plant abiotic stresses: Role of heat shock proteins (HSP). J. Proteomics. 71(4):391-411

Tournaire-Roux C, Sutka M, Javot H, Gout E, Gerbeau P, *et al.* 2003. Cytosolic pH regulates root water transport during anoxic stress through gating of aquaporins. Nature. 425(6956):393-97

Trapnell C, Roberts A, Goff L, Pertea G, Kim D, *et al.* 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat. Protoc. 7(3):562-78

Umezawa T, Fujita M, Fujita Y, Yamaguchi-Shinozaki K, Shinozaki K. 2006. Engineering drought tolerance in plants: discovering and tailoring genes to unlock the future. Curr. Opin. Biotechnol. 17(2):113-22

Vandenborre G, Smagghe G, Van Damme EJM. 2011. Plant lectins as defense proteins against phytophagous insects. Phytochemistry. 72(13):1538-50

doi:10.6342/NTU201603060

Visser EA, Wegrzyn JL, Steenkmap ET, Myburg AA, Naidoo S. 2015. Combined de novo and genome guided assembly and annotation of the *Pinus patula* juvenile shoot transcriptome. BMC Genomics. 16:1057

Vriezen WH, Hulzink R, Mariani C, Voesenek LACJ. 1999. 1-Aminocyclopropane-1-Carboxylate Oxidase Activity Limits Ethylene Biosynthesis in *Rumex palustrisduring* Submergence. Plant Physiol. 121(1):189-96

Wahid A, Gelani S, Ashraf M, Foolad MR. 2007. Heat tolerance in plants: An overview. Environ. Exp. Bot. 61(3):199-223

Wang K, Bian S, Jiang Y. 2008. Anaerobic metabolism in roots of Kentucky bluegrass in response to short-term waterlogging alone and in combination with high temperatures. Plant Soil. 314(1-2):221-29

Wen F, Zhu Y, Hawes MC. 1999. Effect of pectin methylesterase gene expression on pea root development. Plant Cell. 11(6):1129-40

Willats WG, McCartney L, Mackie W, Knox JP. 2001. Pectin: cell biology and prospects for functional analysis. Plant Mol. Biol. 47(1-2):9-27

Yan J, Yu L, Xuan J, Lu Y, Lu S, Zhu W. 2016. De novo transcriptome sequencing and gene expression profiling of spinach (*Spinacia oleracea L.*) leaves under heat stress. Sci. Rep. 6:

Yang C-Y, Hsu F-C, Li J-P, Wang N-N, Shih M-C. 2011. The AP2/ERF Transcription Factor AtERF73/HRE1 Modulates Ethylene Responses during Hypoxia in *Arabidopsis*. Plant Physiol. 156(1):202-12

Yu F, Wu Y, Xie Q. 2015. Precise protein post-translational modifications modulate ABI5 activity. Trends Plant Sci. 20(9):569-75

doi:10.6342/NTU201603060

Supplementary information

Supplemental Table 1. An outline of basic function of major classes of heat shock proteins in plant system for heat stress tolerance. (Hasanuzzaman *et al.* 2013)

Major classes of heat shock protein	Functions
HSP100	ATP-dependent dissociation and degradation of aggregate protein
HSP90	Co-regulator of heat stress linked signal transduction complexes and manages protein folding. It requires ATP for its function
HSP70, HSP40	Primary stabilization of newly formed proteins, ATP-dependent binding and release
HSP60, HSP10	ATP-dependent specialized folding machinery
HSP20 or small HSP (sHSP)	Formation of high molecular weight oligomeric complexes which serve as cellular matrix for stabilization of unfolded proteins. HSP100, HSP70 and HSP40 are needed for its release

Supplemental Table 2. Primers for qPCR validation.

Primer_ID	gene_ID	Sequence size	Forward Primer	Forward Primer Start	Reverse Primer	Reverse Primer Start
P1	BOHW5218	1556	TCCACTTGCTCCAGTGTCAT	1122	GTGGTTTGCAGAGACAAGGG	1234
P2	BOHW43287	1306	CCTTGCTTCATCGTCATCACA	389	TTCCTGTCCCACCATCGAAG	506
P3	BOHW8926	1859	GGAAAAGAGGCATCGTCACG	1484	AATCTGGGGTTGTCTCTCGG	1577
P4	BOHW29550	1659	TAAGCGGAGAGAGGAGCAAG	843	TCCGGCAACTTAAACCTCCT	942

doi:10.6342/NTU201603060

Supplemental Table 3. Quality statistics of the raw reads (A) and *de novo* assembly (B).

A.

Raw read	Number of nucleic acid	File size (GB)
N_F	105,479,648	5.6
N_R	105,479,648	5.6
H_F	109,663,096	5.8
H_R	109,663,096	5.8
W_F	102,522,516	5.9
W_R	102,522,516	5.9
HW_F	111,049,924	5.4
HW_R	111,049,924	5.4
Total	857,430,368	45.3

B.

Total Reads	214,357,592
Matched	204,519,631
Not matched	9,837,961
Broken paired reads	34,833,647
Reads in pairs	169,685,984
Reads usage rate (%)	95.4
Assembly contigs	83,910
Contig N50 (nt)	1,115

doi:10.6342/NTU201603060

Supplemental Table 4. MapMan terms enrichment analysis result of module IV and V. ID is the bin from MapMan. *P*-value is calculated by Hypergeometric test. N is the number of total genes in MapMan. m is the number of query gene. n is the number of genes for the MapMan term. k is the number of genes for intersection between m and n.

Rank	ID	Description	p-value	N	m	n	k
Module IV							
1	10.1.5	cell wall.precursor synthesis.UXS	0.002405	27416	11	6	1
2	29.5.11.2	protein.degradation.ubiquitin.E1	0.003206	27416	11	8	1
3	27.3.18	RNA.regulation of transcription.E2F/DP transcription factor family	0.003206	27416	11	8	1
4	27.3.65	RNA.regulation of transcription.Polycomb Group (PcG)	0.003606	27416	11	9	1
5	27.3.73	RNA.regulation of transcription.Zn-finger(CCHC)	0.006402	27416	11	16	1
6	10.1	cell wall.precursor synthesis	0.02224	27416	11	56	1
7	34.3	transport.amino acids	0.02578	27416	11	65	1
8	18	Co-factor and vitamine metabolism	0.03164	27416	11	80	1
9	31.2	cell.division	0.04172	27416	11	106	1
10	31	cell	0.04201	27416	11	830	2
11	29.5	protein.degradation	0.04386	27416	11	850	2
Module V							
1	29.2.2.1	protein.synthesis.ribosome biogenesis.export from nucleus	0.00306	27416	12	7	1
2	17.2.1	hormone metabolism.auxin.synthesis-degradation	0.006112	27416	12	14	1
3	29.3.4.2	protein.targeting.secretory pathway.golgi	0.007417	27416	12	17	1
4	29.2.2	protein.synthesis.ribosome biogenesis	0.01089	27416	12	25	1
5	9.7	mitochondrial electron transport / ATP synthesis.cytochrome c oxidase	0.01219	27416	12	28	1
6	27.1.2	RNA.processing.RNA helicase	0.01478	27416	12	34	1
7	31.2	cell.division	0.04543	27416	12	106	1
8	29.5.3	protein.degradation.cysteine protease	0.04836	27416	12	113	1
9	31	cell	0.04942	27416	12	830	2

doi:10.6342/NTU201603060

Supplemental Table 5. The expression value of DEGs involved in responding to reactive oxygen species.

Contig No.	Treatments			Functional annotation
	W	H	HW	
BOHW10022	0.346605262	-1.836224722	-2.426326166	Flavin-binding monooxygenase family protein
BOHW10032	-1.472026788	-1.354384142	-1.565597176	secE/sec61-gamma protein transport protein
BOHW10050	-0.107763318	-1.545702149	-1.35471803	DNAJ heat shock family protein
BOHW10053	0.07602255	-3.608170965	-3.791357072	sugar transporter 4
BOHW10120	-0.148368412	-1.872874191	-1.653212208	HR-like lesion-inducing protein-related
BOHW12060	0.289412791	-1.756710031	-2.303429359	glutathione peroxidase 3
BOHW12065	0.156236221	-1.020817471	-1.556309544	Plant neutral invertase family protein
BOHW12082	0.323002201	4.101662447	4.058343874	PHD finger family protein / bromo-adjacent homology (BAH) domain-containing protein
BOHW19026	0.251697727	-2.001422063	-2.448296119	receptor-like protein kinase 1
BOHW19115	-0.102724306	1.014119015	1.059391167	HSP20-like chaperones superfamily protein
BOHW21456	-0.04658194	-1.805688718	-1.295236662	homolog of mamallian P58IPK
BOHW2148	0.228122771	-1.970157312	-2.636213587	BEL1-like homeodomain 1
BOHW24060	0.344535363	-1.582572778	-1.071576608	Tetratricopeptide repeat (TPR)-like superfamily protein
BOHW26555	-0.995434618	-9	-5.022367813	ATP binding cassette subfamily B4
BOHW26575	0.105006808	2.72898995	2.539994882	Fes1A
BOHW26621	-0.177662348	1.368592326	1.387717367	senescence-associated gene 21
BOHW26679	-0.137238279	-1.302228169	-1.222392421	Lojap-related protein
BOHW29550	0.146250186	11.03936483	11.44111153	heat shock protein 18.2
BOHW3249	0.345017715	-3.58710935	-3.502444034	glycosyl hydrolase 9A1
BOHW36008	13	13	13	Glutathione S-transferase family protein
BOHW39353	0.275621295	5.674397438	5.888178266	HSP20-like chaperones superfamily protein
BOHW43451	2.444691017	5.528832303	4.528309722	EID1-like 3
BOHW47562	1.456017435	-5.807336104	-5.769092931	production of anthocyanin pigment 1
BOHW47680	0.523300852	12.34842095	12.10039488	heat shock transcription factor A7A
BOHW5218	0.715756219	-0.580041592	-1.700009705	abscisic acid responsive elements-binding factor 3
BOHW5252	-9	4.806159623	4.357552005	ferritin 3
BOHW5742	0.23932606	-1.371379038	-1.653663965	auxin response factor 2
BOHW5749	0.094970754	-1.908967596	-1.138089986	Leucine-rich repeat (LRR) family protein
BOHW57713	0.774839619	6.795105022	6.285285277	crumpled leaf
BOHW77882	13	13	13	multiprotein bridging factor 1C
BOHW7793	0.07271943	7.056991999	7.186510468	heat shock protein 70
BOHW9015	0.313508467	2.789507521	2.546016897	frataxin homolog
BOHW9067	0.250086203	9.679623852	9.475789737	17.6 kDa class II heat shock protein

doi:10.6342/NTU201603060

Supplemental Table 6. The expression value of DEGs involved in responding to water deprivation.

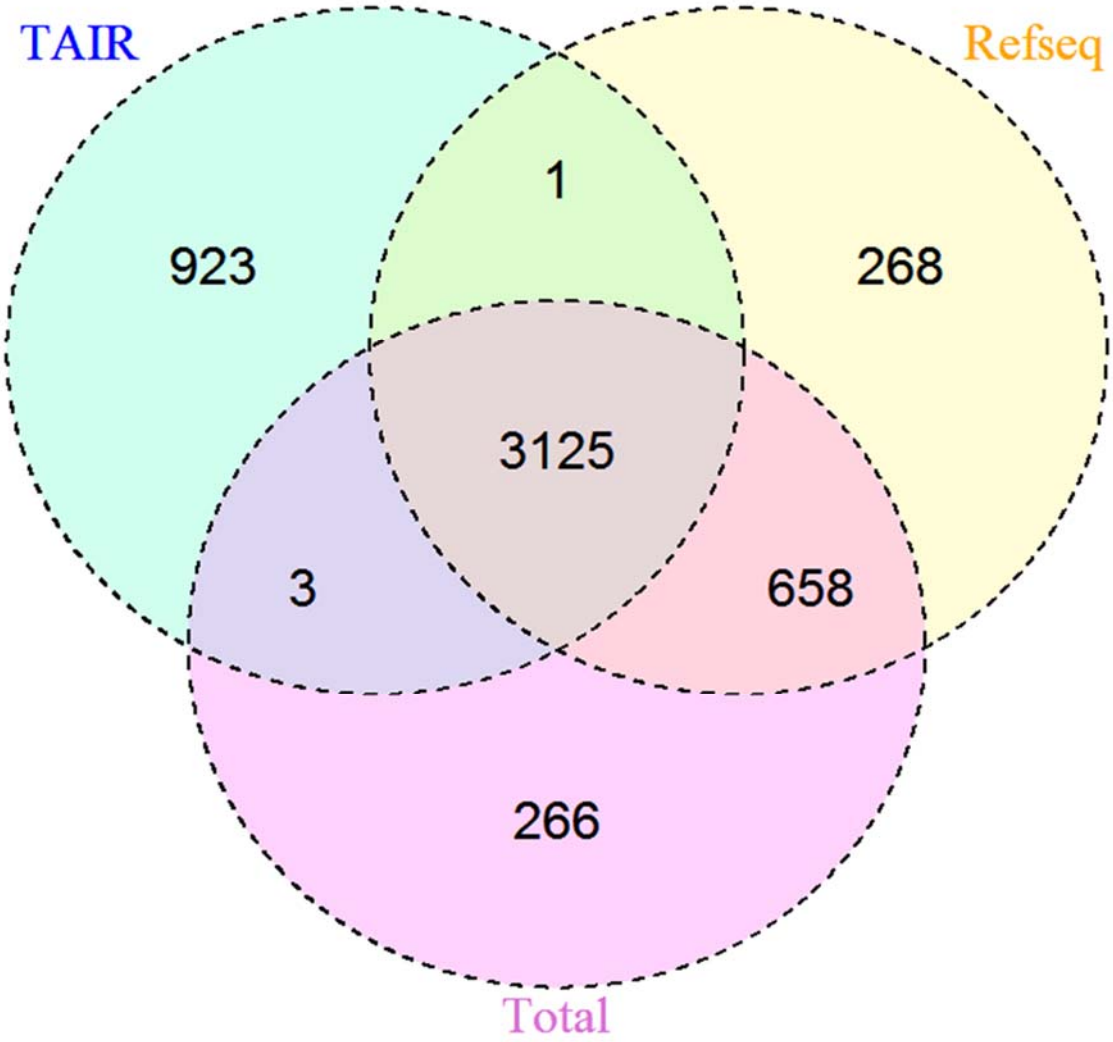
Contig No.	Treatments			Functional annotation
	W	H	HW	
BOHW10078	-0.570001948	1.419484347	1.247927513	cyclin H;1
BOHW12060	0.289412791	-1.756710031	-2.303429359	glutathione peroxidase 3
BOHW14222	1.152598358	-0.482640083	-3.459375312	ABA INSENSITIVE 5
BOHW14230	0.452911524	-1.671821133	-1.255066337	Drought-responsive family protein
BOHW16490	-0.038232243	-1.548299509	-1.662035668	Calcium-dependent lipid-binding (CaLB domain) family protein
BOHW16651	0.250516736	-1.869281517	-2.20819596	HVA22 homologue D
BOHW16670	-0.066073412	-0.895829158	-1.359346494	NAD(P)-binding Rossmann-fold superfamily protein
BOHW18895	-0.832842959	-4.949355109	-3.326149435	VQ motif-containing protein
BOHW18918	0.285801538	2.702248599	2.483845369	beta-amylase 1
BOHW19012	-0.291182975	-4.441544019	-4.336283388	histidine kinase 2
BOHW19026	0.251697727	-2.001422063	-2.448296119	receptor-like protein kinase 1
BOHW19082	-0.055785083	-6.078638126	-6.777360547	laccase 2
BOHW21429	0.057441125	-0.657924	-1.024605748	homolog of yeast autophagy 18 (ATG18) F
BOHW2143	0.064282225	-1.09783377	-1.208442863	Prenyltransferase family protein
BOHW2392	-0.251603391	-1.399480572	-1.515634266	Inorganic H pyrophosphatase family protein
BOHW26555	-0.995434618	-9	-5.022367813	ATP binding cassette subfamily B4
BOHW26621	-0.177662348	1.368592326	1.387717367	senescence-associated gene 21
BOHW26668	0.513053792	5.468251598	4.885357434	heat shock factor 4
BOHW26773	0.025732265	-0.546770421	-1.616615054	enhanced disease resistance 1
BOHW29390	0.204523894	2.143294179	1.556327261	ubiquitin-conjugating enzyme 32
BOHW29438	-0.077646498	-0.451741806	-1.576868875	stress response suppressor 2
BOHW32456	-0.210053489	-4.197282622	-3.159039449	A20/AN1-like zinc finger family protein
BOHW32509	0.39846374	-4.704242611	-7.250961939	Major facilitator superfamily protein
BOHW35733	-0.822264642	-2.546770421	-2.064074031	jasmonic acid carboxyl methyltransferase
BOHW39337	-1.068089757	-1.438245964	-2.262978099	tyrosine aminotransferase 3
BOHW39363	0.748073822	2.221197123	1.987927168	Heavy metal transport/detoxification superfamily protein
BOHW39451	-0.298217828	-0.895658415	-1.609359238	Poly (ADP-ribose) glycohydrolase (PARG)
BOHW43300	0.220170625	-4.089297655	-3.066032884	Protein kinase superfamily protein
BOHW43451	2.444691017	5.528832303	4.528309722	EID1-like 3
BOHW43506	0.091768145	1.461129884	1.699192252	Protein of unknown function (DUF3537)
BOHW47394	0.744142821	1.326279334	1.410816159	myb domain protein 60
BOHW47598	0.504900102	-3.482640083	-3.459375312	related to AP2 1
BOHW5218	0.715756219	-0.580041592	-1.700009705	abscisic acid responsive elements-binding factor 3
BOHW70388	-9	5.106074552	4.585018807	C2H2-type zinc finger family protein
BOHW77882	13	13	13	multiprotein bridging factor 1C
BOHW7805	0.1874321	-4.571184054	-5.201923428	cytochrome P450, family 79, subfamily B, polypeptide 2
BOHW8926	1.38580778	0.108837455	-1.401219292	ABI five binding protein
BOHW8946	0.138979836	-0.504516832	-1.216120544	Major facilitator superfamily protein

doi:10.6342/NTU201603060

Supplemental Table 7. Significantly enriched MapMan bins in cultivar-specific superclusters (*p*-value < 0.05).

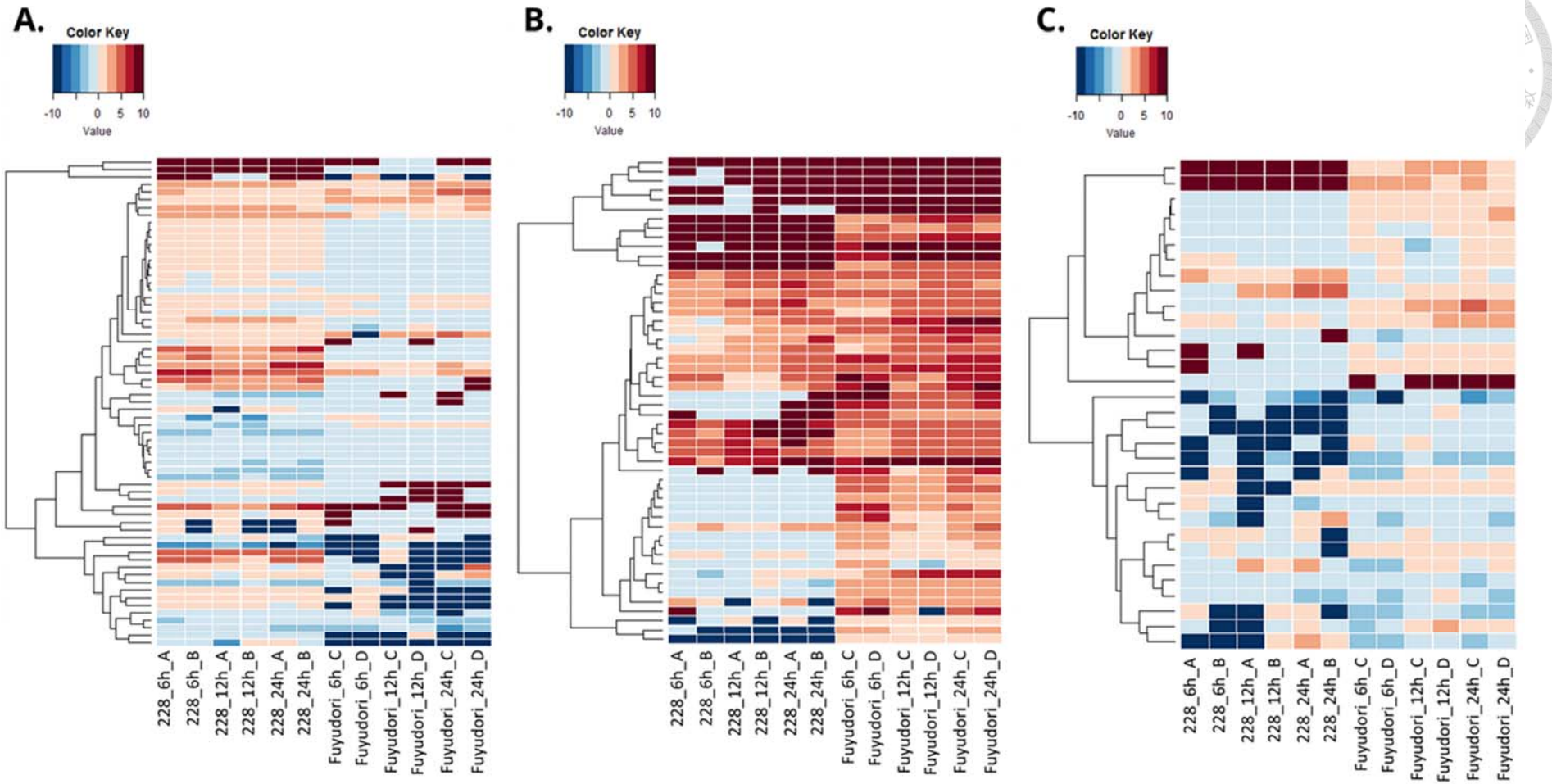
Rank	ID	Description	p-value
		Supercluster I	
1	11.9.4.9	lipid metabolism.lipid degradation.beta-oxidation.multifunctional	0.00226
2	16.2.1.7	secondary metabolism.phenylpropanoids.lignin biosynthesis.CCR1	0.005641
3	28.1.3.1	DNA.synthesis/chromatin structure.histone.H1	0.007889
4	16.8.5	secondary metabolism.flavonoids.isoflavones	0.009011
5	16.8.5.1	secondary metabolism.flavonoids.isoflavones.isoflavone reductase	0.009011
6	11.9	lipid metabolism.lipid degradation	0.01108
7	28.1.3	DNA.synthesis/chromatin structure.histone	0.01125
8	30.2.10	signalling.receptor kinases.leucine rich repeat X	0.01349
9	29.3.4.1	protein.targeting.secretory pathway.ER	0.01683
10	29.2	protein.synthesis	0.01815
11	34.1	transport.nucleotides	0.0268
12	29.2.2	protein.synthesis.ribosome biogenesis	0.0279
13	33.1	development.storage proteins	0.0312
14	11.9.2.1	lipid metabolism.lipid degradation.lipases.triacylglycerol lipase	0.03448
15	29.2.4	protein.synthesis.elongation	0.03776
16	11.9.4	lipid metabolism.lipid degradation.beta-oxidation	0.03994
17	16.2.1	secondary metabolism.phenylpropanoids.lignin biosynthesis	0.04644
18	11.9.2	lipid metabolism.lipid degradation.lipases	0.04968
		Supercluster IV	
1	33.2	development.late embryogenesis abundant	5.59E-06
2	29.1.10	protein.aa activation.methionine-tRNA ligase	0.002479
3	3.4.3	minor CHO metabolism.myo-inositol.InsP Synthases	0.003716
4	20.2.5	stress.abiotic.light	0.007419
5	29.4	protein.postranslational modification	0.01261
6	20.2	stress.abiotic	0.01751
7	3.4	minor CHO metabolism.myo-inositol	0.02088
8	11.6	lipid metabolism.lipid transfer proteins etc	0.0221
9	16.4.1	secondary metabolism.N misc.alkaloid-like	0.02694
10	16.4	secondary metabolism.N misc	0.03297
11	27.3.20	RNA.regulation of transcription.G2-like transcription factor family, GARP	0.04964
		Supercluster VII	
1	14.2	S-assimilation.APR	0.002078
2	29.5	protein.degradation	0.002451
3	1.3.13	PS.calvin cycle.rubisco interacting	0.004151
4	14	S-assimilation	0.008974
5	21.2.2	redox.ascorbate and glutathione.glutathione	0.008974
6	10.5.4	cell wall.cell wall proteins.HRGP	0.01103
7	17.5.3	hormone metabolism.ethylene.induced-regulated-responsive-activated	0.01719
8	1.3	PS.calvin cycle	0.02195
9	29.5.7	protein.degradation.metalloprotease	0.02602
10	29.5.4	protein.degradation.aspartate protease	0.02804
11	29.5.9	protein.degradation.AAA type	0.03074
12	27.1.19	RNA.processing.ribonucleases	0.03745
13	15.2	metal handling.binding, chelation and storage	0.03945
14	29.5.11.20	protein.degradation.ubiquitin.proteasom	0.04145
15	21.2	redox.ascorbate and glutathione	0.04212

doi:10.6342/NTU201603060



Supplemental Figure 1. Venn plot of the annotated contigs wih TAIR and Refseq plant peptide databases. Total represented all 4,052 contigs. TAIR and Refseq represented the contigs aligned the gene in the database and the identity of alignment results was > 70%.

doi:10.6342/NTU201603060



Supplemental Figure 2. The heatmap of gene expression in three cultivar-specific supercluster I (A), IV (B), and VII (C). Color key was adjusted to minimal and maximal $\log_2$ fold change values to enhance the difference.

doi:10.6342/NTU201603060