DISCOVERIES OF PATHWAY AND REGULATION OF BRANCHED-CHAIN AMINO ACID CATABOLISM IN *ARABIDOPSIS THALIANA* REVEALED THROUGH TRANSCRIPT AND GENETIC STUDIES

By

Cheng Peng

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Plant Biology - Doctor of Philosophy

2015

ABSTRACT

DISCOVERIES OF PATHWAY AND REGULATION OF BRANCHED-CHAIN AMINO ACID CATABOLISM IN *ARABIDOPSIS THALIANA* REVEALED THROUGH TRANSCRIPT AND GENETIC STUDIES

By

Cheng Peng

The branched-chain amino acids (BCAAs) leucine, isoleucine and valine are among nine essential amino acids that humans and other animals must obtain from their diets, and can be nutritionally limiting in plant foods. Rapid development of transcript profiling technologies has enabled research on plant metabolism that offers potential to improve crop nutritional quality by allowing researchers to apply correlative analysis in hypothesis generation followed by experimental validation. Despite genetic evidence of its importance in regulating seed amino acid levels, the full BCAA catabolic network is not completely understood in plants, and limited information is available regarding its regulation.

In this study, a combination of transcript and mutant analyses was performed to study the pathway and regulation of BCAA catabolism in *Arabidopsis thaliana*. Transcript coexpression analyses revealed positive correlations among BCAA catabolic genes in stress, development, diurnal/circadian and light datasets. BCAA catabolism genes show coordinated oscillation in diurnal and circadian treatments, and their expression patterns are altered in clock and phytochrome B mutants, providing evidence for the regulation of BCAA catabolism by the circadian clock and light. Functional divergence is suggested by transcript profile comparison between four pairs of BCAA catabolic enzyme paralogs, and the paralogs do not increase their transcript levels upon the loss of their duplicated copies in the dark. In addition, mutants defective in putative branched-chain ketoacid dehydrogenase subunits accumulate higher levels of BCAAs in mature seeds, providing genetic evidence for their function in BCAA catabolism.

BCAA catabolism genes are highly expressed during the night on a diel cycle and during prolonged darkness, and mutants undergo senescence early and over-accumulate leaf BCAAs during prolonged darkness. These results extend the previous evidence that BCAAs can be catabolized and serve as respiratory substrates at multiple steps. Furthermore, comparison of amino acid profiles between mature seeds and dark-treated leaves revealed differences in amino acid accumulation when BCAA catabolism is perturbed. Together, these results demonstrate the consequences of blocking BCAA catabolism during both normal growth conditions and under energy-limited conditions.

ACKNOWLEDGEMENTS

It has been a very pleasant and rewarding six years studying plant biology at Michigan State University. When I first came to the US in the summer of 2009, I was naïve and didn't know what to expect. Six years later, I've become a young scientist equipped with knowledge and skills, and better prepared to solve problems in research and in ordinary life. Within the years, I've not only deepened my knowledge in science, but also picked up skills in a variety of aspects in life that will be beneficial in the future. And I know I could not have accomplished these and become a better me without help from other people along the way.

First and foremost, I want to thank my advisor, Dr. Robert Last, for his guidance and understanding throughout the years. I have been privileged to join his lab and work with him. Rob is a role model as an enthusiastic scientist who is always curious about science and dedicated at work, while at the same time enjoying life. He allowed creative freedom for me to pursue my way to finish my projects, and was always there when I needed suggestions and guidance. He also taught me to work independently and to think critically. I have learned so much from him in the past six years, and he will always be someone that I look up to.

Next, I want to thank my committee members - Drs. Dan Jones, Robin Buell and Sheng Yang He. They provided honest and constructive suggestions to my research that better refined my projects, encouraged me when I met obstacles, and were always willing to offer useful resources. They greatly facilitated my research.

I joined the Last lab in the summer of 2010, and since then, my lab mates have been very generous and helpful. Everybody in the lab is an expert on what they are doing, and they were always willing to share their experience and patiently taught me new skills. I appreciate all their

honest, critical and constructive suggestions for my projects during the past few years.

I was lucky to join the collaboration between the Last lab and the Shiu lab. As my main collaborator, Sahra Uygun provided numerous advice and help in my research. I am proud to have such a smart and dedicated collaborator and great friend. I also want to thank Dr. Shin-Han Shiu, who is not in my committee, but knows so much about my projects, and has always been extremely supportive and encouraging for my own projects as well as the projects we are collaborating on.

The highly integrated plant research community at MSU is one of the finest and largest in the country, and a lot of great plant scientists were trained here. I am so proud to be part of the Department of Plant Biology and MSU-DOE Plant Research Laboratory. I greatly appreciate their generous financial supports, their great effort in creating a collaborating and friendly research environment, and the kind and helpful faculty members and staffs. I also want to thank the Research Technology Support Facility at MSU. The Genomics core helped me prepare RNA-Seq libraries and sequenced my samples. Lijun Chen at the Mass Spectrometry core was always friendly and patient, and was extremely helpful in solving problems regarding to the LC-MS/MS.

My friends at East Lansing have become my second family over the years. They are the ones that I can discuss scientific issues with, share the highs and lows in life with, exercise with, attend events with, or just simply eat meals with. My friends have always been extremely helpful and supportive. My life is so much more colorful and entertaining because of them, and I cannot imagine survive graduate school without them.

At last, I want to thank my parents. They gave me the freedom to pursue my dream thousand miles away, while building a warm and supportive harbor back home. Their unconditional love and support are the driven force for my achievements today.

PREFACE

This dissertation research was conducted by me under the supervision of Professor Robert L. Last in the Department of Plant Biology, at Michigan State University, between July 2010 to April 2015, except where acknowledgements and references of previous work are stated.

Part of this work has been included in a manuscript submitted to Plant Physiology: **Peng C, Uygun S, Shiu, S-H, Last, R L** (2015) The Impact of the Branched-Chain Ketoacid Dehydrogenase Complex on Amino Acid Homeostasis in Arabidopsis.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

KEY TO ABBREVIATIONS

Ala	Alanine
amiRNA	Artificial micro RNA
Arg	Arginine
Asn	Asparagine
Asp	Aspartate
BCAA	Branched-chain amino acid
bZIP	Basic leucine zipper
Col-0	Columbia-0
Cys	Cysteine
GFP	Green fluorescent protein
Glu	Glutamate
Gln	Glutamine
Gly	Glycine
His	Histidine
Ile	Isoleucine
LC-MS/MS	Liquid chromatography - tandem mass spectrometry
Leu	Leucine
Lys	Lysine
Met	Methionine
mRNA	Messenger RNA
MS	Mass spectrometry

PCR	Polymerase chain reaction
Phe	Phenylalanine
Pro	Proline
PSII	Photosystem II
qRT-PCR	Quantitative reverse transcriptase PCR
RNA	Ribonucelic acid
RNA-Seq	RNA sequencing
Ser	Serine
TF	Transcription factor
Thr	Threonine
Trp	Tryptophan
Tyr	Tyrosine
UTR	Untranslated region
Val	Valine
Y1H	Yeast one hybrid

Chapter 1 Literature review

Plant branched-chain amino acid metabolism and utilization of transcript coexpression analysis in plant metabolism

1.1 Overview of plant branched-chain amino acid catabolism

1.1.1 The biological functions of BCAAs

The branched-chain amino acids (BCAAs, Figure 1.1) Leucine (Leu), Isoleucine (Ile) and Valine (Val) are among nine amino acids essential for humans and other animals because they cannot be synthesized *de novo* (Harper et al., 1984). Plants synthesize BCAAs and are the main source of these essential nutrients in the diets of humans and agriculturally important animals. However, the BCAA contents in plant foods are insufficient to meet dietary requirements, making genetic improvement in increased BCAA levels in plants a desiring target for metabolic engineering (Angelovici et al., 2013). Unfortunately, few attempts in optimization of nutritional values in plants have been successful, largely due to other unexpected deleterious traits. For example, Zhu and Galili's attempts to increase seed free lysine content led to an undesirable low germination rate in *Arabidopsis thaliana* (Zhu and Galili, 2003; Zhu and Galili, 2004). This emphasizes the idea that regulation of metabolism and the connections between metabolic pathways are far more complex than previously thought. Thus, a deeper understanding in the architecture of the metabolic networks should lead us to a better-optimized engineering in plant metabolism.

Besides serving as building blocks for proteins, BCAAs, together with BCAA-derived metabolites, have additional physiological functions. In plants, it was demonstrated that BCAAs and a wide variety of BCAA-derived products such as glucosinolates, fatty acids, acyl sugars, and volatiles contribute to normal growth, development and defense (Walters and Steffens, 1990; Pérez et al., 2002; Mikkelsen and Halkier, 2003; Taylor et al., 2004; Ishizaki et al., 2005; Matich and Rowan, 2007; Slocombe et al., 2008; Araujo et al., 2010; Gonda et al., 2010; Ding et al.,

2012; Kochevenko et al., 2012; Zhang et al., 2014). The biological functions of BCAAs have been extensively studied in animals and humans. BCAAs, mainly Leu, were shown to serve as a regulator in a number of cell signaling pathways monitoring food intake, stimulating translation initiation, promoting fatty acid oxidation, and modulating hormone (mainly insulin) levels (Shimomura et al., 2004; Blomstrand et al., 2006; Layman and Walker, 2006; Norton and Layman, 2006). Recent studies also established the link between BCAAs and diseases such as certain types of diabetes and cancer - because of their roles as signaling pathway regulators (Adeva et al., 2012; McCormack et al., 2013; O'Connell, 2013; Wubetu et al., 2014), and the potential of using BCAAs as biomarkers for the progression of such diseases is being evaluated (Tom and Nair, 2006).

1.1.2 Free BCAA levels are strictly regulated

BCAA homeostasis is well balanced in living organisms, and disruption in BCAA catabolism leads to a variety of consequences. In humans and animals, excessive BCAAs are efficiently degraded, perhaps consistent with a role of Leu as regulator in signaling pathways. The importance of a well-balanced BCAA catabolism was demonstrated by multiple severe metabolic disorders associated with enzymes involved in this pathway (Chuang et al., 2006). For an example, impaired branched-chain ketoacid dehydrogenase (BCKDH) activity leads to the maple syrup urine disease, resulting in symptoms including the buildup of BCAAs and their toxic by-products (ketoacids) in the blood and urine, and in worst case scenario, brain damage and death of patients (Harris et al., 1990). In plants, maintaining BCAA homeostasis seems much more complex, because plants both make and degrade these amino acids. Allosteric regulation is important for BCAA biosynthesis (Figure 1.2) (Halgand et al., 2002; Garcia and Mourad, 2004;

de Kraker et al., 2007; Curien et al., 2008), while no evidence was demonstrated for allosteric regulation on BCAA catabolism (Binder, 2010). The most consistent phenotype of plant mutants with impaired BCAA catabolic enzymes is the over-accumulation of free BCAAs and CoA intermediates in their dry seeds or mature fruits (Gu et al., 2010; Maloney et al., 2010; Lu et al., 2011; Angelovici et al., 2013). In addition, some BCAA catabolic mutants show altered levels of primary and specialized metabolites including amino acids that are biosynthetically unrelated, tricarboxylic acid (TCA) cycle intermediates and glucosinolates (Gu et al., 2010; Lu et al., 2011; Kochevenko et al., 2012). Certain BCAA catabolic mutants also exhibit aberrant reproductive growth such as abnormal flower and silique development (Ding et al., 2012), and early senescence in prolonged darkness (Araujo et al., 2010). These pleiotropic phenotypes suggest that disruption of BCAA catabolism impacts the function or regulation of other primary and specialized metabolic pathways during normal growth and development, and under energy-limited conditions in plants.

1.1.3 Plant BCAA biosynthesis and herbicide development targeting this pathway

Considerable effort has been devoted to studying plant BCAA biosynthesis (Figure 1.2), mostly due to the popular demand for developing low-use-rate herbicides of low animal toxicity two to three decades ago. The three BCAAs are synthesized in plant plastids - including chloroplasts - in parallel pathways sharing a set of four enzymes (as highlighted in grey rectangles in Figure 1.2) by using different substrates. The first common enzyme is acetolactate synthase (ALS, also known as acetohydroxy acid synthase, or AHAS), which catalyzes the conversion of two molecules of pyruvate into acetolactate in the formation of Val and Leu, or one molecule of pyruvate and one molecule of 2-oxobutanoate into 2-aceto-2-hydroxybutyrate in

Ile biosynthesis (Mourad and King, 1995; Singh, 1999; Binder, 2010). The discovery of a variety of structurally diverse herbicides inhibiting ALS activity was a key milestone in the history of weed control (see summary in Table 1.1) (Whitcomb, 1999; Tan et al., 2006; Duggleby et al., 2008). These herbicides are specific to plants, microbes, algae and fungi, but not animals. They also exhibit high selectivity, can work at extremely low concentrations, and lead to stunted growth, malformation and reduced seed production of sensitive plants (Duggleby et al., 2008). A series of non-transgenic crops that are resistant to herbicides were developed and commercialized by chemical mutagenesis or selection of tolerant varieties (Tranel and Wright, 2009). However, herbicide resistance in weeds rapidly developed after the commercial release of the first ALS-inhibiting herbicide - chlorosulfuron (commercial name Glean, DuPont Co.) - and to date at least 151 weed species were reported to have evolved resistance to a variety of ALS-inhibiting herbicides (Heap, 2015). To overcome this issue, mutations causing resistance were intensively identified and characterized genetically and biochemically (Haughn et al., 1988; Haughn and Somerville, 1990; Sathasivan et al., 1990; Hattori et al., 1992; Ott et al., 1996; Duggleby et al., 2008), and the crystal structures of the Arabidopsis ALS enzyme with or without herbicide binding were obtained (McCourt et al., 2006). This greatly facilitated our understanding on the molecular basis for weed resistance development, and enabled new visions for design of novel inhibitors from a more structure-based rationale (McCourt et al., 2006).

1.1.4 Plant BCAA catabolism

Compared to the well-studied BCAA biosynthetic pathways, our understanding of plant BCAA catabolic pathways is far from complete. The genes proposed for *A. thaliana* BCAA catabolism were assigned based on sequence similarity with their animal counterparts (Figure 1.3)

(Binder, 2010). Enzymes in earlier steps of the catabolic pathways, particularly the Leu degradative pathway, have been partially identified and characterized in recent years (Fujiki et al., 2000; Daschner et al., 2001; Fujiki et al., 2001; Lutziger and Oliver, 2001; Zolman et al., 2001; Che et al., 2002; Diebold et al., 2002; Lange et al., 2004; Taylor et al., 2004; Schuster and Binder, 2005; Binder, 2010; Gu et al., 2010; Maloney et al., 2010; Lu et al., 2011; Ding et al., 2012; Angelovici et al., 2013). However, putative enzymes in later steps of BCAA degradation, especially the breakdown of Ile and Val, still require further evaluation. The catabolism of animal BCAAs, which has been extensively analyzed because of their roles in genetic disease, occurs in mitochondria (Harper et al., 1984; Harris et al., 1990; Chuang et al., 2006). However, there are ongoing controversies about whether the localization of plant BCAA catabolism is in the mitochondrion, peroxisome or both (Zolman et al., 2001; Lange et al., 2004; Taylor et al., 2004; Lucas et al., 2007; Binder, 2010), and the majority of the genes discovered to date appear to have mitochondrial targeting signals (Binder, 2010).

1.1.4.1 Shared early steps in BCAA catabolism

The degradation of all three BCAAs shares one enzyme (branched-chain aminotransferase, or BCAT) and one enzyme complex (BCKDH) in the first two steps (Figure 1.3) (Binder, 2010). There are seven *BCAT* genes in the *A. thaliana* Columbia-0 genome (Table 1.2): six are transcribed and one with no transcript detected in a variety of tissues tested (Diebold et al., 2002). Based on a subcellular localization in chloroplasts (where plant BCAA biosynthesis occurs), BCAT3 and BCAT5 are annotated as being involved in BCAA biosynthesis (Diebold et al., 2002; Schuster and Binder, 2005; Knill et al., 2008), while BCAT1 (Schuster and Binder, 2005) and BCAT2 (Angelovici et al., 2013) are proposed to play roles in BCAA catabolism

based upon their mitochondrial localization. *BCAT1* seems to be transcribed at low level in almost all tissues except roots, however, *BCAT2* responds within hours to carbohydrate deficit (Schuster and Binder, 2005; Angelovici et al., 2013) and exhibits mRNA expression patterns strikingly similar to validated BCAA catabolism genes under various stresses and hormone treatments (Matsui et al., 2008; Mentzen et al., 2008; Urano et al., 2009; Angelovici et al., 2013). Angelovici and coworkers characterized T-DNA mutants defective in *BCAT1* and *BCAT2*, and revealed that *bcat2* mutants exhibit moderate but statistically significant increases in seed free BCAAs, while *bcat1* mutants show no increases (Angelovici et al., 2013). These results suggest the hypothesis that *A. thaliana BCAT1* serves a housekeeping role in BCAA catabolism while *BCAT2* responds to environmental stimuli. In cultivated tomato (*Solanum lycopersicum*), two BCATs (*SlBCAT1* and *SlBCAT2*) are targeted to the mitochondrion, and antisense-mediated reduction of *SlBCAT1* led to BCAA increases in mature fruits, consistent with a role in BCAA catabolism (Maloney et al., 2010).

The shared second step is catalyzed by the multi-subunit BCKDH complex (Figure 1.3) (Binder, 2010), which is well characterized in animals. This high molecular weight complex of up to 9 MDa converts branched-chain ketoacids (which are intermediates in both BCAA biosynthesis and degradation) into acyl-CoA esters (Figure 1.4). Despite the biochemical evidence demonstrating the existence of the enzyme activity in isolated *A. thaliana* mitochondria (Taylor et al., 2004), the functions of the genes encoding subunits of this complex still need to be evaluated. It is similar to mitochondrial and chloroplastic pyruvate dehydrogenases and mitochondrial α-ketoglutarate dehydrogenase, and shares enzyme subunits with the Glycine (Gly) decarboxylase complex (Oliver, 1994; Mooney et al., 2002). The biochemically characterized mammalian BCKDH is comprised of multiple copies of three proteins (Figure 1.4): the α-

ketoacid dehydrogenase/carboxylase E1 (E1α and E1β), dihydrolipoyl acyltransferase E2, and dihydrolipoyl dehydrogenase E3 (also known as the mitochondrial lipoamide dehydrogenase, or mtLPD) (Mooney et al., 2002; Lynch et al., 2003). The large size of this complex has hindered detailed *in vitro* characterization in plants, and identification of the plant BCKDH complex subunits is based upon sequence annotation rather than functional analysis. Pairs of paralogous genes are annotated as encoding *A. thaliana* E1α, E1β and E3 subunits, and one gene is annotated for E2 (Table 1.3). These assignments are not based on enzymology, but rather are from: 1) sequence similarity with proteins identified from other organisms, especially mammals, and 2) mitochondrial localization identified by tandem mass spectrometry (Fujiki et al., 2000; Mooney et al., 2000; Taylor et al., 2004).

1.1.4.2 Later steps in BCAA catabolism

A. thaliana isovaleryl-CoA dehydrogenase (IVD) - the sole documented acyl-CoA dehydrogenase family member in mitochondria - is hypothesized to catalyze the third step in the degradation of all three BCAAs (Figure 1.3) (Daschner et al., 2001; Binder, 2010). Daschner and colleagues first isolated *IVD1* cDNA, and demonstrated the strong substrate specificity toward isovaleryl-CoA (a Leu degradation intermediate) and isobutyryl-CoA (a Val degradation intermediate) in *A. thaliana*, indicating that it functions in the degradation of both Leu and Val (Daschner et al., 2001). In contrast to *A. thaliana*, two enzymes were found in potato: one with the same substrate specificity as *A. thaliana* IVD, while the other enzyme exhibits high catalytic efficiency and low K_m with 2-methylbutyryl-CoA - an Ile degradation intermediate (Faivre-Nitschke et al., 2001; Goetzman et al., 2005). However, no homolog of the latter gene has been found in other plant species, calling into question whether IVD has a role in Ile

catabolism across plant species (Goetzman et al., 2005). Two different *A. thaliana* mutants with IVD defects exhibit pleiotropic phenotypes with increases in 12 free seed amino acids, including the three BCAAs, (Gu et al., 2010). A recent study also indicated a link between IVD and the mitochondrial electron transport chain: during prolonged darkness, *ivd1-2* mutant shows early senescence at a level intermediate between electron-transfer flavoprotein (ETF) and electron-transfer flavoprotein: ubiquinone oxidoreductase (ETFQO) complex mutants and wild-type plants (Araujo et al., 2010).

BCAA catabolism separates into parallel branches after IVD (Figure 1.3). Methylcrotonyl-CoA carboxylase (MCCase) catalyzes the next step in Leu degradation, and functions as a biotin containing heterodimer, formed by α (MCCA) and β (MCCB) subunits (Alban et al., 1993; Anderson et al., 1998). *A. thaliana* mutants deficient in either MCCA (*mcca1-1*) or MCCB (*mccb1-1*) showed higher accumulation of all BCAAs, and amino acids that are biosynthetically unrelated - such as Arginine (Arg) and Histidine (His) - in seeds compared to wild type (Lu et al., 2011).

Hydroxymethylglutaryl-CoA lyase (HML) is hypothesized to catalyze the last step in Leu degradation (Figure 1.3). The ability of *A. thaliana* HML to hydrolyze hydroxymethylglutaryl-CoA to acetyl-CoA was tested and confirmed *in vitro* (Lu et al., 2011). The *A. thaliana hml1* mutants also exhibited over-accumulation of all BCAAs, and several amino acids that are biosynthetically diverse, including Arg, His and Tryptophan (Trp) (Lu et al., 2011).

Although most genes in Leu degradation have been discovered, we still lack information of the genes participating in Val and Ile degradation. The group of Bonnie Bartel at Rice University characterized a gene called *CHY1* (stands for *β-HYDROXYISOBUTYRYL-CoA HYDROLASE 1*) and demonstrated β-hydroxyisobutyryl-CoA hydrolase activity for the encoded

protein, suggesting that *CHY1* participate at a later step downstream of IVD in Val degradation (Zolman et al., 2001). However, CHY1 protein appears to have a peroxisomal targeting signal (Zolman et al., 2001), which is different from the mitochondrial location of other BCAA catabolism genes characterized so far. Further study is needed to identify new pathway genes and clarify the subcellular localization of Val and Ile degradation in plants.

In summary, most BCAA catabolic mutants identified and characterized to date exhibit over accumulation of free seed BCAAs (Gu et al., 2010; Lu et al., 2011; Angelovici et al., 2013), making it a signature phenotype and a reference for future studies. However, several unexpected phenotypes were also found in some mutants, including early senescence in prolonged darkness and accumulation of amino acids that are biosynthetically diverse (Araujo et al., 2010; Gu et al., 2010; Lu et al., 2011). However, whether these unexpected phenotypes apply to all BCAA catabolic mutants still needs to be examined. Regardless of the specificity of the phenotypes, existing data to date show that defects in amino acid catabolism lead to increases in biosynthetically unrelated amino acids, suggesting that plant amino acid metabolic networks are more interconnected than that was previously thought. These results reinforce the idea that there are gaps in our knowledge of the regulation of plant amino acid metabolism.

1.1.5 Regulation of BCAA biosynthesis and catabolism

The accumulation of free BCAAs is co-regulated in wild-type *A. thaliana* seeds and tomato fruits (Schauer et al., 2006; Lu et al., 2008). This presumably is due in part to the fact that they share four common biosynthetic enzymes and three catabolic steps (Figure 1.2, Figure 1.3). Recent studies also revealed that BCAA levels oscillate in diurnal and circadian conditions (Espinoza et al., 2010) and increase during prolonged darkness in rosette leaves of wild-type

Arabidopsis plants (Ishizaki et al., 2005), suggesting a coordinate regulation of the parallel pathways of BCAA biosynthesis and/or catabolism.

The well-characterized BCAA biosynthetic pathways have three enzymes subject to allosteric regulation: threonine deaminase (TD), ALS, and isopropylmalate synthase (IPMS) (Figure 1.2) (Binder, 2010). TD catalyzes the first step in Ile biosynthesis, and Ile allosterically inhibits its activity by affecting the quaternary structure (Halgand et al., 2002). It is reported that Val impedes the inhibition by competing with Ile for binding to TD (Halgand et al., 2002). ALS leads to the biosynthesis of all three BCAAs and its activity is inhibited by all three BCAAs, with synergistic positive effects documented between Leu and Ile, as well as Leu and Val (Curien et al., 2008). IPMS initiates Leu biosynthesis from 2-oxoisovalerate - the last intermediate in Val biosynthesis, hence it is feedback inhibited by Leu (de Kraker et al., 2007). Aside from allosteric regulation, additional regulatory mechanisms have not been found to date for BCAA biosynthesis. Transcripts of BCAA biosynthesis genes have been analyzed in plants under a variety of conditions (Less and Galili, 2008; Espinoza et al., 2010). However, strong common patterns regarding spatiotemporal regulation or responses to internal and external stimuli were not found.

Although we know less about the genes participating in BCAA catabolism, results of recent studies suggest that there is co-regulation of steady state levels of the mRNAs of the BCAA catabolism genes identified so far, due to their hypothesized role in the dark. It was reported that (Ishizaki et al., 2005; Ishizaki et al., 2006; Araujo et al., 2010), in addition to catalyzing the third step in the degradation of BCAAs, IVD helps plants survive under energy-limited conditions by serving as a source of electrons for the mitochondrial electron transport chain (miETC) via the electron-transfer flavoprotein α and β subunits (ETFα and ETFβ) and the

electron-transfer flavoprotein ubiquinone oxidoreductase (ETFQO) (Figure 1.3). Evidence for this role is that, in prolonged darkness: 1) the *ivd1-2* mutant becomes senescent faster than wild type, and 2) mutants defective in ETFβ and ETFQO accumulate more BCAAs and IVD substrate isovaleryl-CoA. In addition, the expression of several other genes proposed or validated to be involved in BCAA catabolism - including *BCAT2, BCKDH E2, IVD, MCCA* and *MCCB* - is rapidly induced following transition from light to dark, and is inhibited by sucrose (Fujiki et al., 2000; Daschner et al., 2001; Schuster and Binder, 2005). These observations suggest that *IVD1* and other BCAA catabolism gene transcripts might be co-regulated in the dark, and the encoded enzymes possibly contribute to plant fitness under energy-limited conditions.

In another study, Mentzen and coworkers at Iowa State University performed a global coexpression analysis using microarray data from 70 experiments and pointed out a coexpression supermodule capable of maintaining cellular energy balance via catabolism, and it included not only BCAA degradation, but also catabolism of other amino acids, carbohydrate catabolism, lipid breakdown and cell wall degradation (Mentzen et al., 2008). In this study, they also demonstrated that a basic leucine zipper (bZIP) transcription factor (TF) candidate bZIP1 (At5g49450) was included in the supermodule, and was hypothesized to participate in the regulation of these catabolic processes, but no experimental validation was performed (Mentzen et al., 2008).

The Gad Galili group at the Weizmann Institute of Science utilized publicly available microarray data to seek for co-regulated transcripts and study interactions within and between the Aspartate (Asp) -family and aromatic amino acid networks (Less et al., 2010). Two distinct gene modules regulated by two oppositely expressed subsets of genes in Asp-derived amino acid metabolism were revealed. *BCAT2* belongs to one of the two highly coexpressed modules,

together with genes encoding the bifunctional lysine-ketoglutarate reductase/saccharopine dehydrogenase (LKR/SDH), methionine γ lyase (MGL) and threonine aldolase (THA) at committed steps towards the degradation of Lysine, Methionine and Threonine, respectively. Other genes known and proposed for BCAA catabolism were not included in this study, and thus expression correlation cannot be evaluated from this work. Further examination of whether other BCAA catabolism genes exhibit coexpression with the highly co-regulated amino acid catabolic genes *LKR/SDH*, *MGL* and *THA1* should aid in our understanding of the complex and delicate regulation of plant amino acid metabolic networks.

In summary, published transcript analyses suggest a common expression pattern for BCAA catabolism genes and a potential regulation of BCAA catabolism at the transcript level. However, no comprehensive evaluation using all known and proposed BCAA catabolism genes is published. Such an in-depth analysis would aid in functional characterization of new pathway genes and to reveal additional physiological role(s) of plant BCAA catabolism.

1.2 Transcript coexpression analysis in functional studies in plant metabolism

1.2.1 Overview of two major transcript profiling techniques

Tremendous effort has been devoted to the development of novel technologies for reliable and high-throughput quantification of the transcripts in biological samples. Hybridization- and sequencing-based approaches are the two main types for such transcript profiling techniques, and both approaches were invented two decades ago (Schena et al., 1995; Velculescu et al., 1995; Wang et al., 2009). Because of lower workload and less cost, hybridization-based methods became favorable at first and were rapidly developed. However, as faster and cheaper sequencing technologies became available in the past few years, more and more researchers often prefer to use sequencing-based transcript profiling approaches in their studies now. Hybridization-based approaches include custom-made microarrays and commercial high-density oligo photo-lithographically produced microarrays such as the Affymetrix GeneChip arrays (Woo et al., 2004). They require prior knowledge of the DNA sequence for probe design, and involve incubation of fluorescently labeled cDNA with microarray chips. In contrast, no prior knowledge is required for sequencing-based approaches, because the cDNA sequence is directly determined. Being the most popular sequencing-based approach, 'RNA sequencing' (RNA-Seq) takes advantage of the recently developed next generation sequencing technology to accurately quantify the transcriptome without prior knowledge of transcriptome space (Brautigam and Gowik, 2010). Several platforms have been developed to date, with Illumina HiSeq sequencing system being one of the widely used platforms due to its reliability and efficiency (Liu et al., 2012). Currently, GeneChip and RNA-Seq are the two major approaches for high-throughput transcript profiling (Mantione et al., 2014).

The commercialization of the custom-designed microarray GeneChips made DNA microarray technologies relative inexpensive and capable of measuring transcripts in a high-throughput manner (Table 1.4). The Affymetrix ATH1 GeneChip was designed for *A. thaliana* (Hennig et al., 2003; Redman et al., 2004). Tremendous amount of data were generated and shared among scientists, and platforms and applications were created for data storage and comparison (Craigon et al., 2004; Zimmermann et al., 2004). However, several disadvantages are associated with this technique: 1) the probe design on chips depends on prior knowledge of DNA sequence, thus its application is restricted to organisms with well-explored genomes; 2) low transcriptome coverage per chip, due to the limited number of probe spots; 3) limited detection range, because of its high background during hybridization at the low end of the range and signal saturation during detection for high steady-state level transcripts; and 4) low reproducibility, therefore sophisticated normalization is required for comparison across different experiments (Wang et al., 2009; Mantione et al., 2014).

The recently developed RNA-Seq technique has advantages compared to hybridization-based methods (Table 1.4). It is compatible with organisms without a sequenced genome or extensive transcriptome information, and can detect more features including exon/intron boundaries, splice variants, novel transcripts, fusion genes, and so on (Marioni et al., 2008; Maher et al., 2009; Arnvig et al., 2011; Howard et al., 2013). RNA-Seq also provides accurate detection of transcripts within a large dynamic range. This approach is highly reproducible, and therefore no sophisticated normalization is required for comparison across experiments (Marioni et al., 2008; Gonzalez-Ballester et al., 2010; Zhao et al., 2014). However, several considerations still need to be addressed: 1) certain biases towards longer transcripts and the end or body of transcripts could be introduced by the fragmentation methods used during library preparation

(Wang et al., 2009); 2) the short reads acquired by standard Illumina RNA-Seq methods complicate analysis procedures; and 3) the large size of data files (~5Gb per sample) impedes public data sharing and storage (Wang et al., 2009; Zhao et al., 2014).

As the cost of sequencing continues to fall, RNA-Seq clearly is more advantageous and is expected to become increasingly commonly used for quantitative and qualitative assessment of the transcriptome among researchers (Gonzalez-Ballester et al., 2010; Nookaew et al., 2012; Howard et al., 2013). However, GeneChips are still quite cost-efficient, and the large amount of public data that has been generated so far makes it extremely valuable until the amount of data generated by RNA-Seq catches up. In the future, GeneChips might be restricted to specialized applications, such as clinical diagnosis (Mantione et al., 2014).

1.2.2 Utilizing transcript coexpression in hypothesis generation and testing

Technological advances have greatly facilitated the capture of huge amounts of transcriptome data. To utilize these data and further our understanding of plant biology, correlative approaches have been adopted mainly by searching for transcript coordination (Tohge and Fernie, 2012), with the underlying assumption that correlated genes are likely to be functionally related and involved in similar or identical processes (Saito et al., 2008). Several web-based coexpression applications have been developed for *A. thaliana*, for example, ATTED-II (Obayashi et al., 2009; Obayashi et al., 2011), AraNet (Hwang et al., 2011; Lee et al., 2015) and Expression Angler of the Bio-Array Resource (Toufighi et al., 2005). In addition, custom-based coexpression analysis protocols have also rapidly evolved, allowing for a more in depth analysis of specific pathways (Reich et al., 2006; Teknomo, 2006; Aoki et al., 2007; Langfelder and Horvath, 2008). Although correlation does not prove that the co-regulated genes work in the

same biological process, it is a good way to generate hypotheses for experimental testing (Fukushima et al., 2009; Stitt, 2013).

The construction of coexpression networks using validated genes as baits has facilitated the identification of novel genes in plant metabolic pathways where transcriptional regulation are important, such as cell wall biosynthesis (Brown et al., 2005; Persson et al., 2005; Mutwil et al., 2009) and specialized metabolism (Rischer et al., 2006; Hirai et al., 2007; Saito et al., 2008; Yonekura-Sakakibara et al., 2008; Fukushima et al., 2009; De Luca et al., 2012; Kliebenstein, 2012; Higashi and Saito, 2013). In addition, genes that participate in signaling pathways or respond to specific environmental perturbations have also been identified via this approach: for example, genes responding to cold (Hannah et al., 2005), or involved in jasmonate signaling (McGrath et al., 2005). Moreover, pathway regulators can also be identified through this approach. For example, two TFs - MYB28 and MYB29 - regulating aliphatic glucosinolate biosynthesis were identified and experimentally evaluated in *A. thaliana* (Hirai et al., 2007).

Although correlation-based gene functional identification has been successful in advancing our understanding of specialized metabolism, fewer examples were demonstrated for primary metabolism, perhaps due to its complex pathway topology and overlapping functionality (Stitt, 2013). Successful examples include the demonstration of transcriptional regulation of the starch degradation pathway (Smith et al., 2004), and the identification of novel transporters in C4 photosynthesis (Furumoto et al., 2011; Pick et al., 2011) and in photorespiration (Eisenhut et al., 2013; Pick et al., 2013).

In addition to characterizing individual pathways, coexpression analysis also expedites elucidation of the extensive coordination and communication between metabolic pathways (Fukushima et al., 2009). For example, studies on diverse biological processes such as seed

germination (Bassel et al., 2011) and dark-induced senescence (Araujo et al., 2011) revealed that pathways sharing similar physiological functions can be highly correlated.

1.2.3 Considerations

Several considerations need to be taken into account when performing transcript coexpression analysis in functional studies. First, transcript levels are subject to spatial and temporal regulation (Brady et al., 2007; Matas et al., 2011; Petricka et al., 2012; Rogers et al., 2012; Moussaieff et al., 2013), making conditional transcript co-regulations hard to be revealed without using relevant datasets (Usadel et al., 2009). Second, coexpression analysis alone does not guarantee the finding of all genes in the same pathway, in part due to the complex regulation and redundancy in plant metabolism (Stitt, 2013). Third, transcript changes might not lead to alterations in protein and metabolite levels. It is known that the levels of transcripts, proteins and metabolites are poorly correlated, partially because of the diverse turnover rates that range from minutes to days (Gibon et al., 2006; Caldana et al., 2011; Baerenfaller et al., 2012; Buescher et al., 2012) and the complicated regulatory mechanisms that act at multiple steps on both synthesis and degradation of transcripts, proteins and metabolites (Bailey-Serres, 1999; Piques et al., 2009; Li et al., 2012; Liu et al., 2012).

1.3 Aims of this research

The goal of this dissertation research is to achieve a better understanding of the genes and enzymes participating in BCAA catabolism and their regulation at the transcript level in *A. thaliana*. The first aim was to study the transcriptional regulation of BCAA catabolism genes. Transcript coexpression analyses were performed to look for coexpression among proposed and experimentally validated BCAA catabolism genes, and to identify conditions where coexpression might occur. Expression profiling comparisons were done to evaluate proposed genes and potential pathway regulators, and to examine transcriptional divergence among enzyme paralogs. Promoter mutagenesis was performed with the goal to locate the *cis*-element(s) responsible for the diurnal oscillation of BCAA catabolism gene transcripts. These results are described in Chapter 2. Another aim was to evaluate and explore the function(s) of putative and validated BCAA catabolic enzymes within and beyond BCAA catabolism. Mutants were characterized under normal growth conditions and prolonged darkness, and the results are described in Chapter 3. Taken together, these results provide insights into the regulation and physiological roles of BCAA catabolism, and emphasize the complex inter-connections within the plant metabolic networks.

APPENDIX

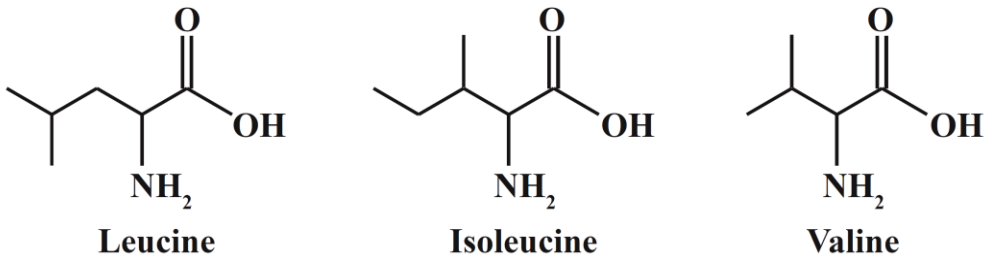


Figure 1.1 Structures of the three branched-chain amino acids.

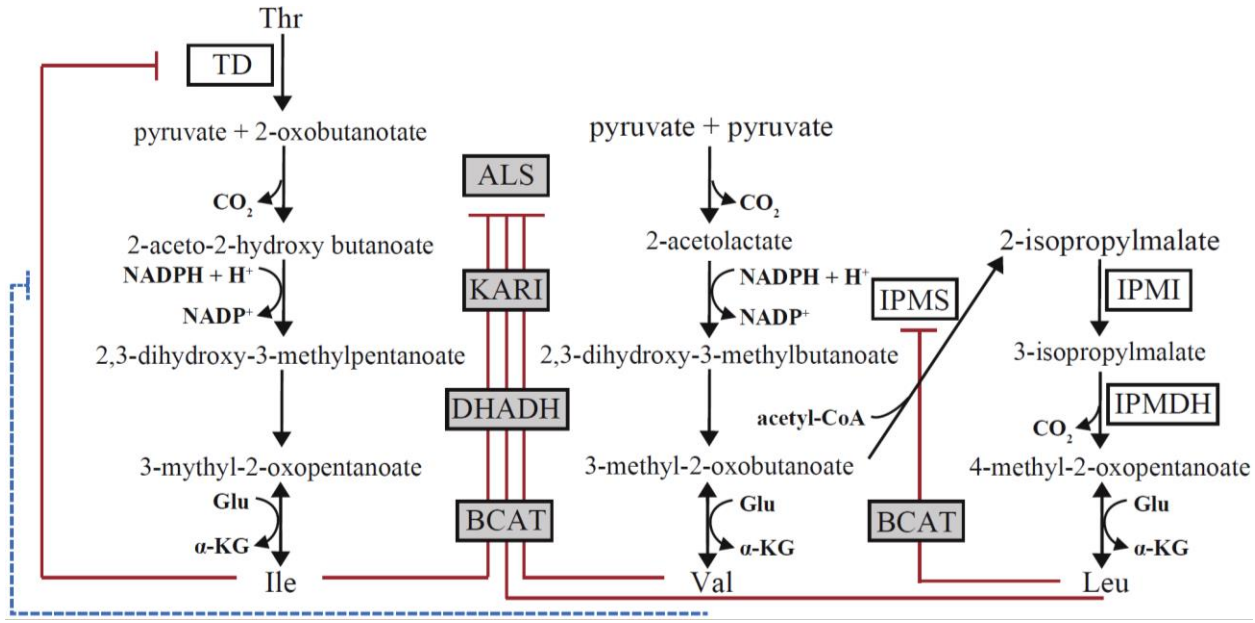


Figure 1.2 Known enzymes and allosteric regulation of BCAA biosynthesis. Enzyme names are abbreviated in rectangles with shared ones highlighted in grey. Allosteric inhibition is indicated by red lines. The restoring effect of Val on TD inhibition by Ile is shown as a dotted blue line. Four shared enzymes are: ALS, acetolactate synthase; KARI, ketoacid reductoisomerase; DHADH, dihydroxyacid dehydratase; and BCAT, branched-chain aminotransferase. TD, threonine deaminase; IPMS, isopropylmalate synthase; IPMI, isopropylmalate isomerase; IPMDH, isopropylmalate dehydrogenase; α-KG, α-ketoglutarate. Modified from Binder, 2010.

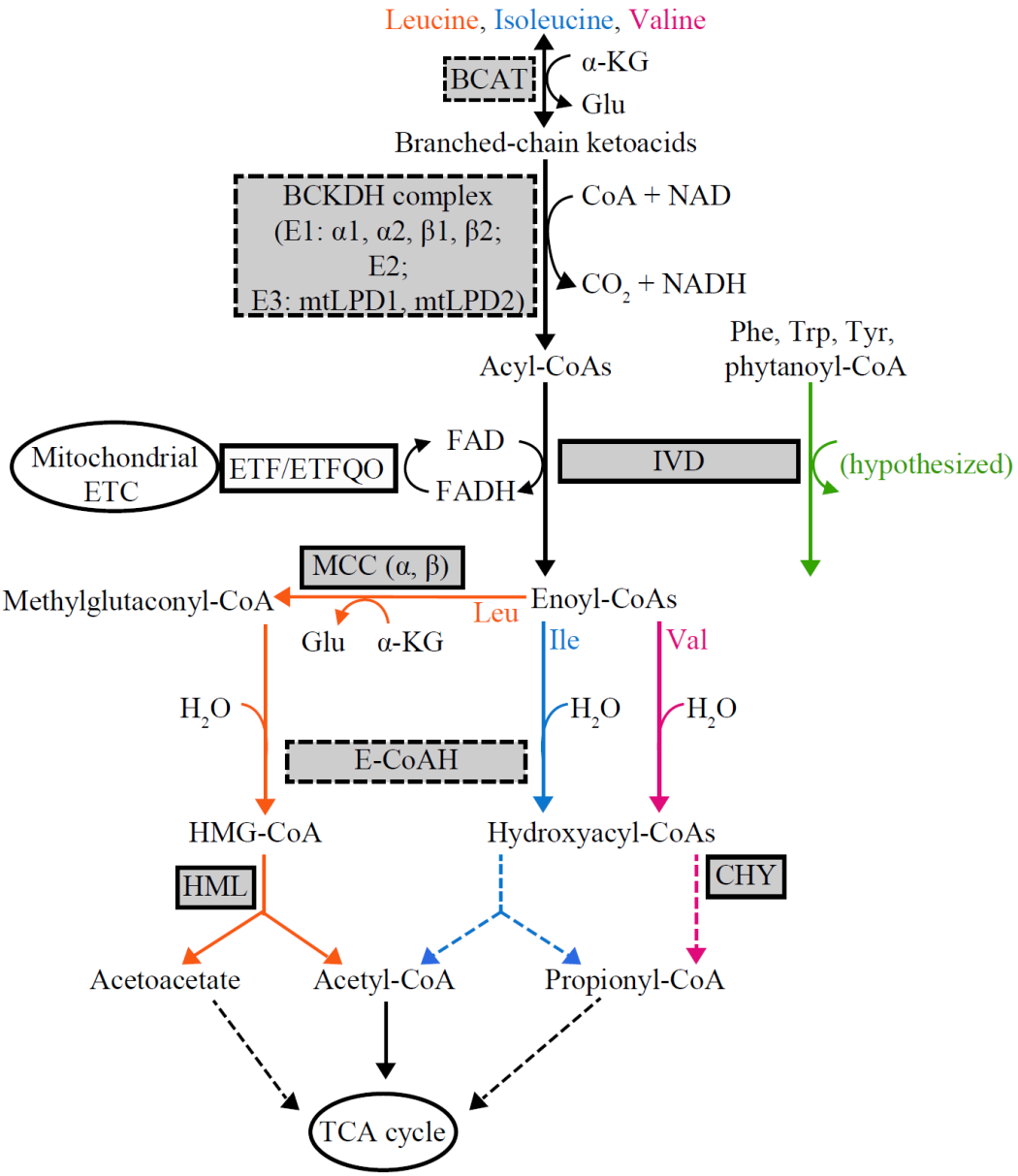


Figure 1.3 Proposed *A. thaliana* BCAA catabolic pathway. Enzyme names are abbreviated in rectangles with BCAA catabolic enzymes highlighted in grey. Validated BCAA catabolic enzyme activities are surrounded by solid lines, and putative BCAA catabolic enzyme activities by dashed lines. Metabolic processes directly or indirectly connected to BCAA catabolism in the mitochondrion are represented in ellipses. Reactions with one step are shown with solid arrows, and those with multiple steps with dashed arrows. Enzyme activities specific to Leu, Ile and Val degradation are highlighted in orange, blue and pink, respectively. Hypothesized IVD enzyme

Figure 1.3 (cont'd)

activity towards non-BCAA catabolic intermediates is indicated with a green arrow. BCAT, branched-chain aminotransferase; α-KG, α-ketoglutarate; BCKDH, branched-chain ketoacid dehydrogenase; mtLPD, mitochondrial lipoamide dehydrogenase; IVD, isovaleryl-CoA dehydrogenase; ETF, electron transfer flavoprotein; ETFQO, electron-transfer flavoprotein: ubiquinone oxidoreductase; ETC, electron transport chain; MCC, 3-methylcrotonyl-CoA carboxylase; E-CoAH, enoyl-CoA hydratase; HMG-CoA, 3-hydroxy-3-methyl-glutaryl-CoA; HML, 3-hydroxyl-3-methylglutaryl-CoA lyase; CHY, β-hydroxyisobutyryl-CoA hydrolase; TCA, tricarboxylic acid.

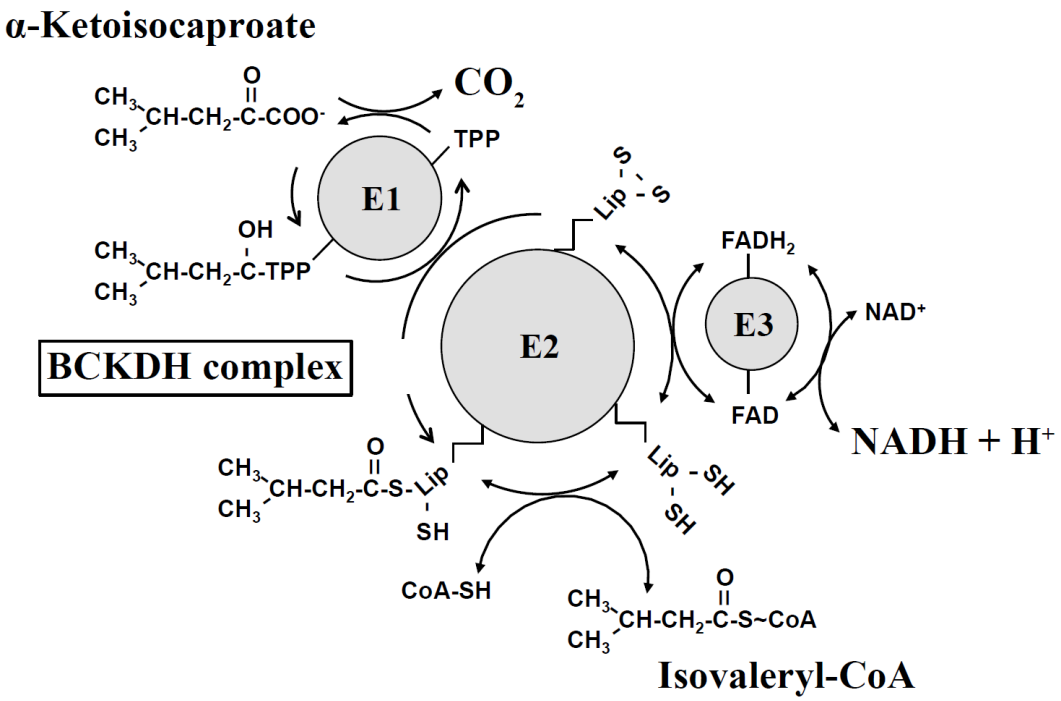


Figure 1.4 Schematic representation of the reaction mechanism of the BCKDH complex. The BCKDH complex is composed of three components (E1, E2 and E3) and α-ketoisovalerate was arbitrarily chosen as the substrate in this figure. TPP, thiamine pyrophosphate; Lip, lipoyl domains. This model is based on the mammalian BCKDH complex. Modified from Lynch et al., 2003.

Table 1.1 Summary of acetolactate synthase inhibiting herbicides.

Chemical family	Herbicide common name	Trade name	Crop use	Weed Spectrum
Imidazolinone	Imazamox Imazapic Imazapyr Imazaquin Imazethapyr	Raptor Cadre, Plateau Arsenal, Chopper, Stalker Image, Scepter Pursuit	Soybeans, other legumes, forests	Mostly broadleaf species, but also some grass species
Sulfonylurea	Chlorimuron-ethyl Chlorsulfuron Halosulfuron-methyl Metsulfuron-methyl Nicosulfuron	Classic Glean, Corsair Permit, Manage Ally, Escort, Manor Accent	Corn, soybeans, small grains, turfgrass, non-cropland	Very broad spectrum control of both grasses and broadleaves
Triazolopyrimidine	Cloransulam Flumetsulam Diclosulam Florasulam Penoxsulam	Firstrate Python Strongarm Orion, GoldSky, FirstStep, and Spitfire Granite	Corn, soybeans	Active on broadleaf species; little activity on grasses
Pyrimidinylthio (or oxy)-benzoate	Bispyribac-sodium Pyrithiobac-sodium	Nominee Staple	Rice, turfgrass, cotton	Broad spectrum control of grasses, broadleaf weeds and sedges
Sulfonylaminocarbonyltriazolinone	Flucarbazone-sodium Propoxycarbazone-sodium	Everest, Sierra Attribut, Olympus	Wheat, rye, triticale	Certain grasses and broadleaf species

Table 1.2 Summary of *A. thaliana* branched-chain aminotransferases (BCATs).

Name	AGI	Designated metabolic pathway	Subcellular localization	Detection approach*
BCAT1	AT1G10060	BCAA catabolism	mitochondria	GFP
BCAT2	AT1G10070	BCAA catabolism	mitochondria	GFP
BCAT3	AT3G49680	BCAA biosynthesis & Met-derived glucosinolate chain elongation	plastid	GFP & MS
BCAT4	AT3G19710	Met-derived glucosinolate chain elongation	cytosol	GFP
BCAT5	AT5G65780	BCAA biosynthesis	plastid, mitochondria**	GFP & MS
BCAT6	AT1G50110	/	cytosol	GFP
BCAT7***	AT1G50090	/	/	/

* GFP, green fluorescent protein; MS, mass spectrometry.

** The mitochondrial subcellular localization of BCAT5 was only shown in a proteomic study, which is questionable and likely due to contamination as was mentioned by the authors (Taylor et al., 2004).

*** *BCAT7* transcripts were not detected in a variety of tissues (Diebold et al., 2002).

Table 1.3 List of putative BCKDH enzyme subunits in *A. thaliana.*

Enzyme	AGI	Annotation
E1α1	AT1G21400	α subunit of branched-chain ketoacid dehydrogenase E1, putative
E1α2	AT5G09300	α subunit of branched-chain ketoacid dehydrogenase E1, putative
E1β1	AT1G55510	β subunit of branched-chain ketoacid dehydrogenase E1, putative
E1β2	AT3G13450	β subunit of branched-chain ketoacid dehydrogenase E1, putative
E2	AT3G06850	branched-chain ketoacid dehydrogenase E2, putative
mtLPD1	AT1G48030	branched-chain ketoacid dehydrogenase E3, putative
mtLPD2	AT3G17240	branched-chain ketoacid dehydrogenase E3, putative

Table 1.4 Comparison between GeneChip and RNA-Seq.

	Microarray	**RNA-Seq**
Pros	High-throughput	High-throughput
	Relatively inexpensive	Does not require prior knowledge
	Tons of existing data	Sensitive, high resolution and large detection range
	Mature analysis procedures	Detection of more features
Cons	Reliance on prior knowledge	Biases due to fragmentation during library preparation
	Limited detection range	Short reads
	Low reproducibility	Large data files
	Limited coverage	More expensive, especially for larger genomes

LITERATURE CITED

LITERATURE CITED

Adeva MM, Calvino J, Souto G, Donapetry C (2012) Insulin resistance and the metabolism of branched-chain amino acids in humans. Amino Acids **43:** 171-181

Alban C, Baldet P, Axiotis S, Douce R (1993) Purification and characterization of 3-methylcrotonyl-coenzyme A carboxylase from higher plant mitochondria. Plant Physiol **102:** 957-965

Anderson MD, Che P, Song J, Nikolau BJ, Wurtele ES (1998) 3-Methylcrotonyl-coenzyme A carboxylase is a component of the mitochondrial leucine catabolic pathway in plants. Plant Physiol **118:** 1127-1138

Angelovici R, Lipka AE, Deason N, Gonzalez-Jorge S, Lin H, Cepela J, Buell R, Gore MA, Dellapenna D (2013) Genome-wide analysis of branched-chain amino acid levels in *Arabidopsis* seeds. Plant Cell **25:** 4827-4843

Aoki K, Ogata Y, Shibata D (2007) Approaches for extracting practical information from gene co-expression networks in plant biology. Plant Cell Physiol **48:** 381-390

Araujo WL, Ishizaki K, Nunes-Nesi A, Larson TR, Tohge T, Krahnert I, Witt S, Obata T, Schauer N, Graham IA, Leaver CJ, Fernie AR (2010) Identification of the 2-hydroxyglutarate and isovaleryl-CoA dehydrogenases as alternative electron donors linking lysine catabolism to the electron transport chain of *Arabidopsis* mitochondria. Plant Cell **22:** 1549-1563

Araujo WL, Tohge T, Ishizaki K, Leaver CJ, Fernie AR (2011) Protein degradation - an alternative respiratory substrate for stressed plants. Trends Plant Sci **16:** 489-498

Arnvig KB, Comas I, Thomson NR, Houghton J, Boshoff HI, Croucher NJ, Rose G, Perkins TT, Parkhill J, Dougan G, Young DB (2011) Sequence-based analysis uncovers an abundance of non-coding RNA in the total transcriptome of *Mycobacterium tuberculosis*. PLoS Pathog **7:** e1002342

Baerenfaller K, Massonnet C, Walsh S, Baginsky S, Buhlmann P, Hennig L, Hirsch-Hoffmann M, Howell KA, Kahlau S, Radziejwoski A, Russenberger D, Rutishauser D, Small I, Stekhoven D, Sulpice R, Svozil J, Wuyts N, Stitt M, Hilson P, Granier C, Gruissem W (2012) Systems-based analysis of Arabidopsis leaf growth reveals adaptation to water deficit. Mol Syst Biol **8:** 606

Bailey-Serres J (1999) Selective translation of cytoplasmic mRNAs in plants. Trends Plant Sci **4:** 142-148

Bassel GW, Lan H, Glaab E, Gibbs DJ, Gerjets T, Krasnogor N, Bonner AJ, Holdsworth MJ, Provart NJ (2011) Genome-wide network model capturing seed germination

reveals coordinated regulation of plant cellular phase transitions. Proc Natl Acad Sci U S A **108:** 9709-9714

Binder S (2010) Branched-chain amino acid metabolism in *Arabidopsis thaliana*. Arabidopsis Book **8:** e0137

Blomstrand E, Eliasson J, Karlsson HK, Kohnke R (2006) Branched-chain amino acids activate key enzymes in protein synthesis after physical exercise. J Nutr **136:** 269s-273s

Brady SM, Orlando DA, Lee JY, Wang JY, Koch J, Dinneny JR, Mace D, Ohler U, Benfey PN (2007) A high-resolution root spatiotemporal map reveals dominant expression patterns. Science **318:** 801-806

Brautigam A, Gowik U (2010) What can next generation sequencing do for you? Next generation sequencing as a valuable tool in plant research. Plant Biol (Stuttg) **12:** 831-841

Brown DM, Zeef LA, Ellis J, Goodacre R, Turner SR (2005) Identification of novel genes in Arabidopsis involved in secondary cell wall formation using expression profiling and reverse genetics. Plant Cell **17:** 2281-2295

Buescher JM, Liebermeister W, Jules M, Uhr M, Muntel J, Botella E, Hessling B, Kleijn RJ, Le Chat L, Lecointe F, Mader U, Nicolas P, Piersma S, Rugheimer F, Becher D, Bessieres P, Bidnenko E, Denham EL, Dervyn E, Devine KM, Doherty G, Drulhe S, Felicori L, Fogg MJ, Goelzer A, Hansen A, Harwood CR, Hecker M, Hubner S, Hultschig C, Jarmer H, Klipp E, Leduc A, Lewis P, Molina F, Noirot P, Peres S, Pigeonneau N, Pohl S, Rasmussen S, Rinn B, Schaffer M, Schnidder J, Schwikowski B, Van Dijl JM, Veiga P, Walsh S, Wilkinson AJ, Stelling J, Aymerich S, Sauer U (2012) Global network reorganization during dynamic adaptations of *Bacillus subtilis* metabolism. Science **335:** 1099-1103

Caldana C, Degenkolbe T, Cuadros-Inostroza A, Klie S, Sulpice R, Leisse A, Steinhauser D, Fernie AR, Willmitzer L, Hannah MA (2011) High-density kinetic analysis of the metabolomic and transcriptomic response of Arabidopsis to eight environmental conditions. Plant J **67:** 869-884

Che P, Wurtele ES, Nikolau BJ (2002) Metabolic and environmental regulation of 3-methylcrotonyl-coenzyme A carboxylase expression in Arabidopsis. Plant Physiol **129:** 625-637

Chuang DT, Chuang JL, Wynn RM (2006) Lessons from genetic disorders of branched-chain amino acid metabolism. J Nutr **136:** 243S-249S

Craigon DJ, James N, Okyere J, Higgins J, Jotham J, May S (2004) NASCArrays: a repository for microarray data generated by NASC's transcriptomics service. Nucleic Acids Res **32:** D575-577

Curien G, Biou V, Mas-Droux C, Robert-Genthon M, Ferrer JL, Dumas R (2008) Amino acid biosynthesis: new architectures in allosteric enzymes. Plant Physiol Biochem **46:** 325-339

Daschner K, Couee I, Binder S (2001) The mitochondrial isovaleryl-coenzyme A dehydrogenase of Arabidopsis oxidizes intermediates of leucine and valine catabolism. Plant Physiol **126:** 601-612

de Kraker JW, Luck K, Textor S, Tokuhisa JG, Gershenzon J (2007) Two Arabidopsis genes (*IPMS1* and *IPMS2*) encode isopropylmalate synthase, the branchpoint step in the biosynthesis of leucine. Plant Physiol **143:** 970-986

De Luca V, Salim V, Atsumi SM, Yu F (2012) Mining the biodiversity of plants: a revolution in the making. Science **336:** 1658-1661

Diebold R, Schuster J, Daschner K, Binder S (2002) The branched-chain amino acid transaminase gene family in Arabidopsis encodes plastid and mitochondrial proteins. Plant Physiol **129:** 540-550

Ding G, Che P, Ilarslan H, Wurtele ES, Nikolau BJ (2012) Genetic dissection of methylcrotonyl CoA carboxylase indicates a complex role for mitochondrial leucine catabolism during seed development and germination. Plant J **70:** 562-577

Duggleby RG, McCourt JA, Guddat LW (2008) Structure and mechanism of inhibition of plant acetohydroxyacid synthase. Plant Physiol Biochem **46:** 309-324

Eisenhut M, Planchais S, Cabassa C, Guivarc'h A, Justin AM, Taconnat L, Renou JP, Linka M, Gagneul D, Timm S, Bauwe H, Carol P, Weber AP (2013) Arabidopsis A BOUT DE SOUFFLE is a putative mitochondrial transporter involved in photorespiratory metabolism and is required for meristem growth at ambient CO_2 levels. Plant J **73:** 836-849

Espinoza C, Degenkolbe T, Caldana C, Zuther E, Leisse A, Willmitzer L, Hincha DK, Hannah MA (2010) Interaction with diurnal and circadian regulation results in dynamic metabolic and transcriptional changes during cold acclimation in Arabidopsis. PLoS One **5:** e14101

Faivre-Nitschke SE, Couée I, Vermel M, Grienenberger JM, Gualberto JM (2001) Purification, characterization and cloning of isovaleryl-CoA dehydrogenase from higher plant mitochondria. Eur J Biochem **268:** 1332-1339

Fujiki Y, Ito M, Nishida I, Watanabe A (2001) Leucine and its keto acid enhance the coordinated expression of genes for branched-chain amino acid catabolism in *Arabidopsis* under sugar starvation. FEBS Lett **499:** 161-165

Fujiki Y, Sato T, Ito M, Watanabe A (2000) Isolation and characterization of cDNA clones for the E1β and E2 subunits of the branched-chain α-ketoacid dehydrogenase complex in *Arabidopsis*. J Biol Chem **275:** 6007-6013

Fukushima A, Kusano M, Redestig H, Arita M, Saito K (2009) Integrated omics approaches in plant systems biology. Curr Opin Chem Biol **13:** 532-538

Furumoto T, Yamaguchi T, Ohshima-Ichie Y, Nakamura M, Tsuchida-Iwata Y, Shimamura M, Ohnishi J, Hata S, Gowik U, Westhoff P, Brautigam A, Weber AP, Izui K (2011) A plastidial sodium-dependent pyruvate transporter. Nature **476:** 472-475

Garcia EL, Mourad GS (2004) A site-directed mutagenesis interrogation of the carboxy-terminal end of *Arabidopsis thaliana* threonine dehydratase/deaminase reveals a synergistic interaction between two effector-binding sites and contributes to the development of a novel selectable marker. Plant Mol Biol **55:** 121-134

Gibon Y, Usadel B, Blaesing OE, Kamlage B, Hoehne M, Trethewey R, Stitt M (2006) Integration of metabolite with transcript and enzyme activity profiling during diurnal cycles in *Arabidopsis* rosettes. Genome Biol **7:** R76

Goetzman ES, Mohsen A-WA, Prasad K, Vockley J (2005) Convergent evolution of a 2-methylbutyryl-CoA dehydrogenase from isovaleryl-CoA dehydrogenase in *Solanum tuberosum*. J Biol Chem **280:** 4873-4879

Gonda I, Bar E, Portnoy V, Lev S, Burger J, Schaffer AA, Tadmor Ya, Gepstein S, Giovannoni JJ, Katzir N, Lewinsohn E (2010) Branched-chain and aromatic amino acid catabolism into aroma volatiles in *Cucumis melo* L. fruit. J Exp Bot **61:** 1111-1123

Gonzalez-Ballester D, Casero D, Cokus S, Pellegrini M, Merchant SS, Grossman AR (2010) RNA-seq analysis of sulfur-deprived *Chlamydomonas* cells reveals aspects of acclimation critical for cell survival. Plant Cell **22:** 2058-2084

Gu LP, Jones AD, Last RL (2010) Broad connections in the Arabidopsis seed metabolic network revealed by metabolite profiling of an amino acid catabolism mutant. Plant J **61:** 579-590

Halgand F, Wessel PM, Laprevote O, Dumas R (2002) Biochemical and mass spectrometric evidence for quaternary structure modifications of plant threonine deaminase induced by isoleucine. Biochemistry **41:** 13767-13773

Hannah MA, Heyer AG, Hincha DK (2005) A global survey of gene regulation during cold acclimation in *Arabidopsis thaliana*. PLoS Genet **1:** e26

Harper AE, Miller RH, Block KP (1984) Branched-chain amino acid metabolism. Annu Rev Nutr **4:** 409-454

Harris RA, Zhang B, Goodwin GW, Kuntz MJ, Shimomura Y, Rougraff P, Dexter P, Zhao Y, Gibson R, Crabb DW (1990) Regulation of the branched-chain α-ketoacid dehydrogenase and elucidation of a molecular basis for maple syrup urine disease. Adv Enzyme Regul **30:** 245-263

Hattori J, Rutledge R, Labbé H, Brown D, Sunohara G, Miki B (1992) Multiple resistance to sulfonylureas and imidazolinones conferred by an acetohydroxyacid synthase gene with separate mutations for selective resistance. Mol Gen Genet **232:** 167-173

Haughn GW, Smith J, Mazur B, Somerville C (1988) Transformation with a mutant *Arabidopsis* acetolactate synthase gene renders tobacco resistant to sulfonylurea herbicides. Mol Gen Genet **211:** 266-271

Haughn GW, Somerville CR (1990) A mutation causing imidazolinone resistance maps to the *Csr1* locus of *Arabidopsis thaliana*. Plant Physiol **92:** 1081-1085

Heap IM (2015) The International Survey of Herbicide Resistant Weeds. Online. Internet. Wednesday, February 4, 2015. Available http://www.weedscience.org

Hennig L, Menges M, Murray JA, Gruissem W (2003) *Arabidopsis* transcript profiling on Affymetrix GeneChip arrays. Plant Mol Biol **53:** 457-465

Higashi Y, Saito K (2013) Network analysis for gene discovery in plant-specialized metabolism. Plant Cell Environ **36:** 1597-1606

Hirai MY, Sugiyama K, Sawada Y, Tohge T, Obayashi T, Suzuki A, Araki R, Sakurai N, Suzuki H, Aoki K, Goda H, Nishizawa OI, Shibata D, Saito K (2007) Omics-based identification of *Arabidopsis* Myb transcription factors regulating aliphatic glucosinolate biosynthesis. Proc Natl Acad Sci U S A **104:** 6478-6483

Howard BE, Hu Q, Babaoglu AC, Chandra M, Borghi M, Tan X, He L, Winter-Sederoff H, Gassmann W, Veronese P, Heber S (2013) High-throughput RNA sequencing of pseudomonas-infected Arabidopsis reveals hidden transcriptome complexity and novel splice variants. PLoS One **8:** e74183

Hwang S, Rhee SY, Marcotte EM, Lee I (2011) Systematic prediction of gene function in *Arabidopsis thaliana* using a probabilistic functional gene network. Nat Protoc **6:** 1429-1442

Ishizaki K, Larson TR, Schauer N, Fernie AR, Graham IA, Leaver CJ (2005) The critical role of *Arabidopsis* electron-transfer flavoprotein:ubiquinone oxidoreductase during dark-induced starvation. Plant Cell **17:** 2587-2600

Ishizaki K, Schauer N, Larson TR, Graham IA, Fernie AR, Leaver CJ (2006) The mitochondrial electron transfer flavoprotein complex is essential for survival of Arabidopsis in extended darkness. Plant J **47:** 751-760

Kliebenstein DJ (2012) Plant defense compounds: systems approaches to metabolic analysis. Annu Rev Phytopathol **50:** 155-173

Knill T, Schuster J, Reichelt M, Gershenzon J, Binder S (2008) Arabidopsis branched-chain aminotransferase 3 functions in both amino acid and glucosinolate biosynthesis. Plant Physiol **146:** 1028-1039

Kochevenko A, Araujo WL, Maloney GS, Tieman DM, Do PT, Taylor MG, Klee HJ, Fernie AR (2012) Catabolism of branched chain amino acids supports respiration but not volatile synthesis in tomato fruits. Mol Plant **5:** 366-375

Lange PR, Eastmond PJ, Madagan K, Graham IA (2004) An Arabidopsis mutant disrupted in valine catabolism is also compromised in peroxisomal fatty acid β-oxidation. FEBS Lett **571:** 147-153

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

Layman DK, Walker DA (2006) Potential importance of leucine in treatment of obesity and the metabolic syndrome. J Nutr **136:** 319s-323s

Lee T, Yang S, Kim E, Ko Y, Hwang S, Shin J, Shim JE, Shim H, Kim H, Kim C, Lee I (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:** D996-d1002

Less H, Angelovici R, Tzin V, Galili G (2010) Principal transcriptional regulation and genome-wide system interactions of the Asp-family and aromatic amino acid networks of amino acid metabolism in plants. Amino Acids **39:** 1023-1028

Less H, Galili G (2008) Principal transcriptional programs regulating plant amino acid metabolism in response to abiotic stresses. Plant Physiol **147:** 316-330

Li L, Nelson CJ, Solheim C, Whelan J, Millar AH (2012) Determining degradation and synthesis rates of arabidopsis proteins using the kinetics of progressive 15N labeling of two-dimensional gel-separated protein spots. Mol Cell Proteomics **11:** M111.010025

Liu L, Li Y, Li S, Hu N, He Y, Pong R, Lin D, Lu L, Law M (2012) Comparison of next-generation sequencing systems. J Biomed Biotechnol **2012:** 251364

Liu MJ, Wu SH, Chen HM, Wu SH (2012) Widespread translational control contributes to the regulation of Arabidopsis photomorphogenesis. Mol Syst Biol **8:** 566

Lu Y, Savage LJ, Ajjawi I, Imre KM, Yoder DW, Benning C, Dellapenna D, Ohlrogge JB, Osteryoung KW, Weber AP, Wilkerson CG, Last RL (2008) New connections across pathways and cellular processes: industrialized mutant screening reveals novel associations between diverse phenotypes in Arabidopsis. Plant Physiol **146:** 1482-1500

Lu Y, Savage LJ, Larson MD, Wilkerson CG, Last RL (2011) Chloroplast 2010: a database for large-scale phenotypic screening of Arabidopsis mutants. Plant Physiol **155:** 1589-1600

Lucas KA, Filley JR, Erb JM, Graybill ER, Hawes JW (2007) Peroxisomal metabolism of propionic acid and isobutyric acid in plants. J Biol Chem **282:** 24980-24989

Lutziger I, Oliver DJ (2001) Characterization of two cDNAs encoding mitochondrial lipoamide dehydrogenase from Arabidopsis. Plant Physiol **127:** 615-623

Lynch CJ, Halle B, Fujii H, Vary TC, Wallin R, Damuni Z, Hutson SM (2003) Potential role of leucine metabolism in the leucine-signaling pathway involving mTOR. Am J Physiol Endocrinol Metab **285:** E854-863

Maher CA, Kumar-Sinha C, Cao X, Kalyana-Sundaram S, Han B, Jing X, Sam L, Barrette T, Palanisamy N, Chinnaiyan AM (2009) Transcriptome sequencing to detect gene fusions in cancer. Nature **458:** 97-101

Maloney GS, Kochevenko A, Tieman DM, Tohge T, Krieger U, Zamir D, Taylor MG, Fernie AR, Klee HJ (2010) Characterization of the branched-chain amino acid aminotransferase enzyme family in tomato. Plant Physiol **153:** 925-936

Mantione KJ, Kream RM, Kuzelova H, Ptacek R, Raboch J, Samuel JM, Stefano GB (2014) Comparing bioinformatic gene expression profiling methods: microarray and RNA-Seq. Med Sci Monit Basic Res **20:** 138-142

Marioni JC, Mason CE, Mane SM, Stephens M, Gilad Y (2008) RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. Genome Res **18:** 1509-1517

Matas AJ, Yeats TH, Buda GJ, Zheng Y, Chatterjee S, Tohge T, Ponnala L, Adato A, Aharoni A, Stark R, Fernie AR, Fei Z, Giovannoni JJ, Rose JK (2011) Tissue- and cell-type specific transcriptome profiling of expanding tomato fruit provides insights into metabolic and regulatory specialization and cuticle formation. Plant Cell **23:** 3893-3910

Matich A, Rowan D (2007) Pathway analysis of branched-chain ester biosynthesis in apple using deuterium labeling and enantioselective gas chromatography-mass spectrometry. J Agr Food Chem **55:** 2727-2735

Matsui A, Ishida J, Morosawa T, Mochizuki Y, Kaminuma E, Endo TA, Okamoto M, Nambara E, Nakajima M, Kawashima M, Satou M, Kim JM, Kobayashi N, Toyoda T, Shinozaki K, Seki M (2008) *Arabidopsis* transcriptome analysis under drought, cold, high-salinity and ABA treatment conditions using a tiling array. Plant Cell Physiol **49:** 1135-1149

McCormack SE, Shaham O, McCarthy MA, Deik AA, Wang TJ, Gerszten RE, Clish CB, Mootha VK, Grinspoon SK, Fleischman A (2013) Circulating branched-chain amino acid concentrations are associated with obesity and future insulin resistance in children and adolescents. Pediatr Obes **8:** 52-61

McCourt JA, Pang SS, King-Scott J, Guddat LW, Duggleby RG (2006) Herbicide-binding sites revealed in the structure of plant acetohydroxyacid synthase. Proc Natl Acad Sci U S A **103:** 569-573

McGrath KC, Dombrecht B, Manners JM, Schenk PM, Edgar CI, Maclean DJ, Scheible WR, Udvardi MK, Kazan K (2005) Repressor- and activator-type ethylene response factors functioning in jasmonate signaling and disease resistance identified via a genome-wide screen of Arabidopsis transcription factor gene expression. Plant Physiol **139:** 949-959

Mentzen WI, Peng J, Ransom N, Nikolau BJ, Wurtele ES (2008) Articulation of three core metabolic processes in Arabidopsis: fatty acid biosynthesis, leucine catabolism and starch metabolism. BMC Plant Biol **8:** 76

Mikkelsen MD, Halkier BA (2003) Metabolic engineering of valine- and isoleucine-derived glucosinolates in Arabidopsis expressing CYP79D2 from Cassava. Plant Physiol **131:** 773-779

Mooney BP, Henzl MT, Miernyk JA, Randall DD (2000) The dihydrolipoyl acyltransferase (BCE2) subunit of the plant branched-chain α-ketoacid dehydrogenase complex forms a 24-mer core with octagonal symmetry. Protein Sci **9:** 1334-1339

Mooney BP, Miernyk JA, Randall DD (2002) The complex fate of α-ketoacids. Annu Rev Plant Biol **53:** 357-375

Mourad G, King J (1995) L-O-Methylthreonine-resistant mutant of Arabidopsis defective in isoleucine feedback regulation. Plant Physiol **107:** 43-52

Moussaieff A, Rogachev I, Brodsky L, Malitsky S, Toal TW, Belcher H, Yativ M, Brady SM, Benfey PN, Aharoni A (2013) High-resolution metabolic mapping of cell types in plant roots. Proc Natl Acad Sci U S A **110:** E1232-1241

Mutwil M, Ruprecht C, Giorgi FM, Bringmann M, Usadel B, Persson S (2009) Transcriptional wiring of cell wall-related genes in *Arabidopsis*. Mol Plant **2:** 1015-1024

Nookaew I, Papini M, Pornputtapong N, Scalcinati G, Fagerberg L, Uhlen M, Nielsen J (2012) A comprehensive comparison of RNA-Seq-based transcriptome analysis from reads to differential gene expression and cross-comparison with microarrays: a case study in *Saccharomyces cerevisiae*. Nucleic Acids Res **40:** 10084-10097

Norton LE, Layman DK (2006) Leucine regulates translation initiation of protein synthesis in skeletal muscle after exercise. J Nutr **136:** 533s-537s

O'Connell TM (2013) The complex role of branched chain amino acids in diabetes and cancer. Metabolites **3:** 931-945

Obayashi T, Hayashi S, Saeki M, Ohta H, Kinoshita K (2009) ATTED-II provides coexpressed gene networks for Arabidopsis. Nucleic Acids Res **37:** D987-991

Obayashi T, Nishida K, Kasahara K, Kinoshita K (2011) ATTED-II updates: condition-specific gene coexpression to extend coexpression analyses and applications to a broad range of flowering plants. Plant Cell Physiol **52:** 213-219

Oliver DJ (1994) The glycine decarboxylase complex from plant mitochondria. Annu Rev Plant Physiol Plant Mol Biol **45:** 323-337

Ott K-H, Kwagh J-G, Stockton GW, Sidorov V, Kakefuda G (1996) Rational molecular design and genetic engineering of herbicide resistant crops by structure modeling and site-directed mutagenesis of acetohydroxyacid synthase. J Mol Biol **263:** 359-368

Pérez AG, Olías R, Luaces P, Sanz C (2002) Biosynthesis of strawberry aroma compounds through amino acid metabolism. J Agr Food Chem **50:** 4037-4042

Persson S, Wei H, Milne J, Page GP, Somerville CR (2005) Identification of genes required for cellulose synthesis by regression analysis of public microarray data sets. Proc Natl Acad Sci U S A **102:** 8633-8638

Petricka JJ, Schauer MA, Megraw M, Breakfield NW, Thompson JW, Georgiev S, Soderblom EJ, Ohler U, Moseley MA, Grossniklaus U, Benfey PN (2012) The protein expression landscape of the Arabidopsis root. Proc Natl Acad Sci U S A **109:** 6811-6818

Pick TR, Brautigam A, Schluter U, Denton AK, Colmsee C, Scholz U, Fahnenstich H, Pieruschka R, Rascher U, Sonnewald U, Weber AP (2011) Systems analysis of a maize leaf developmental gradient redefines the current C4 model and provides candidates for regulation. Plant Cell **23:** 4208-4220

Pick TR, Brautigam A, Schulz MA, Obata T, Fernie AR, Weber AP (2013) PLGG1, a plastidic glycolate glycerate transporter, is required for photorespiration and defines a unique class of metabolite transporters. Proc Natl Acad Sci U S A **110:** 3185-3190

Piques M, Schulze WX, Hohne M, Usadel B, Gibon Y, Rohwer J, Stitt M (2009) Ribosome and transcript copy numbers, polysome occupancy and enzyme dynamics in *Arabidopsis*. Mol Syst Biol **5:** 314

Redman JC, Haas BJ, Tanimoto G, Town CD (2004) Development and evaluation of an *Arabidopsis* whole genome Affymetrix probe array. Plant J **38:** 545-561

Reich M, Liefeld T, Gould J, Lerner J, Tamayo P, Mesirov JP (2006) GenePattern 2.0. Nat Genet **38:** 500-501

Rischer H, Oresic M, Seppanen-Laakso T, Katajamaa M, Lammertyn F, Ardiles-Diaz W, Van Montagu MC, Inze D, Oksman-Caldentey KM, Goossens A (2006) Gene-to-metabolite networks for terpenoid indole alkaloid biosynthesis in *Catharanthus roseus* cells. Proc Natl Acad Sci U S A **103:** 5614-5619

Rogers ED, Jackson T, Moussaieff A, Aharoni A, Benfey PN (2012) Cell type-specific transcriptional profiling: implications for metabolite profiling. Plant J **70:** 5-17

Saito K, Hirai MY, Yonekura-Sakakibara K (2008) Decoding genes with coexpression networks and metabolomics - 'majority report by precogs'. Trends Plant Sci **13:** 36-43

Sathasivan K, Haughn GW, Murai N (1990) Nucleotide sequence of a mutant acetolactate synthase gene from an imidazolinone-resistant *Arabidopsis thaliana* var. Columbia. Nucleic Acids Res **18:** 2188

Schauer N, Semel Y, Roessner U, Gur A, Balbo I, Carrari F, Pleban T, Perez-Melis A, Bruedigam C, Kopka J, Willmitzer L, Zamir D, Fernie AR (2006) Comprehensive metabolic profiling and phenotyping of interspecific introgression lines for tomato improvement. Nat Biotechnol **24:** 447-454

Schena M, Shalon D, Davis RW, Brown PO (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science **270:** 467-470

Schuster J, Binder S (2005) The mitochondrial branched-chain aminotransferase (AtBCAT-1) is capable to initiate degradation of leucine, isoleucine and valine in almost all tissues in *Arabidopsis thaliana*. Plant Mol Biol **57:** 241-254

Shimomura Y, Murakami T, Nakai N, Nagasaki M, Harris RA (2004) Exercise promotes BCAA catabolism: effects of BCAA supplementation on skeletal muscle during exercise. J Nutr **134:** 1583s-1587s

Singh BK (1999) Biosynthesis of valine, leucine, and isoleucine. Plant amino acids: biochemistry and biotechnology**:** 227-247

Slocombe SP, Schauvinhold I, McQuinn RP, Besser K, Welsby NA, Harper A, Aziz N, Li Y, Larson TR, Giovannoni J, Dixon RA, Broun P (2008) Transcriptomic and reverse genetic analyses of branched-chain fatty acid and acyl sugar production in *Solanum pennellii* and *Nicotiana benthamiana*. Plant Physiol **148:** 1830-1846

Smith SM, Fulton DC, Chia T, Thorneycroft D, Chapple A, Dunstan H, Hylton C, Zeeman SC, Smith AM (2004) Diurnal changes in the transcriptome encoding enzymes of starch metabolism provide evidence for both transcriptional and posttranscriptional regulation of starch metabolism in Arabidopsis leaves. Plant Physiol **136:** 2687-2699

Stitt M (2013) Systems-integration of plant metabolism: means, motive and opportunity. Curr Opin Plant Biol **16:** 381-388

Tan S, Evans R, Singh B (2006) Herbicidal inhibitors of amino acid biosynthesis and herbicide-tolerant crops. Amino Acids **30:** 195-204

Taylor NL, Heazlewood JL, Day DA, Millar AH (2004) Lipoic acid-dependent oxidative catabolism of α-keto acids in mitochondria provides evidence for branched-chain amino acid catabolism in Arabidopsis. Plant Physiol **134:** 838-848

Teknomo K (2006) K-means clustering tutorial. Medicine **100:** 3

Tohge T, Fernie AR (2012) Co-expression and co-responses: within and beyond transcription. Front Plant Sci **3:** 248

Tom A, Nair KS (2006) Assessment of branched-chain amino acid status and potential for biomarkers. J Nutr **136:** 324s-330s

Toufighi K, Brady SM, Austin R, Ly E, Provart NJ (2005) The Botany Array Resource: e-Northerns, Expression Angling, and promoter analyses. Plant J **43:** 153-163

Tranel PJ, Wright TR (2009) Resistance of weeds to ALS-inhibiting herbicides: what have we learned? Weed Sci **50:** 700-712

Urano K, Maruyama K, Ogata Y, Morishita Y, Takeda M, Sakurai N, Suzuki H, Saito K, Shibata D, Kobayashi M, Yamaguchi-Shinozaki K, Shinozaki K (2009) Characterization of the ABA-regulated global responses to dehydration in Arabidopsis by metabolomics. Plant J **57:** 1065-1078

Usadel B, Obayashi T, Mutwil M, Giorgi FM, Bassel GW, Tanimoto M, Chow A, Steinhauser D, Persson S, Provart NJ (2009) Co-expression tools for plant biology: opportunities for hypothesis generation and caveats. Plant Cell Environ **32:** 1633-1651

Velculescu VE, Zhang L, Vogelstein B, Kinzler KW (1995) Serial analysis of gene expression. Science **270:** 484-487

Walters DS, Steffens JC (1990) Branched chain amino acid metabolism in the biosynthesis of *Lycopersicon pennellii* glucose esters. Plant Physiol **93:** 1544-1551

Wang Z, Gerstein M, Snyder M (2009) RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet **10:** 57-63

Whitcomb CE (1999) An introduction to ALS-inhibiting herbicides. Toxicol Ind Health **15:** 231-239

Woo Y, Affourtit J, Daigle S, Viale A, Johnson K, Naggert J, Churchill G (2004) A comparison of cDNA, oligonucleotide, and Affymetrix GeneChip gene expression microarray platforms. J Biomol Tech **15:** 276-284

Wubetu GY, Utsunomiya T, Ishikawa D, Ikemoto T, Yamada S, Morine Y, Iwahashi S, Saito Y, Arakawa Y, Imura S, Arimochi H, Shimada M (2014) Branched chain amino acid suppressed insulin-initiated proliferation of human cancer cells through induction of autophagy. Anticancer Res **34:** 4789-4796

Yonekura-Sakakibara K, Tohge T, Matsuda F, Nakabayashi R, Takayama H, Niida R, Watanabe-Takahashi A, Inoue E, Saito K (2008) Comprehensive flavonol profiling and transcriptome coexpression analysis leading to decoding gene-metabolite correlations in *Arabidopsis*. Plant Cell **20:** 2160-2176

Zhang C, Pang Q, Jiang L, Wang S, Yan X, Chen S, He Y (2014) Dihydroxyacid dehydratase is important for gametophyte development and disruption causes increased susceptibility to salinity stress in *Arabidopsis*. J Exp Bot

Zhao S, Fung-Leung WP, Bittner A, Ngo K, Liu X (2014) Comparison of RNA-Seq and microarray in transcriptome profiling of activated T cells. PLoS One **9:** e78644

Zhu X, Galili G (2003) Increased lysine synthesis coupled with a knockout of its catabolism synergistically boosts lysine content and also transregulates the metabolism of other amino acids in Arabidopsis seeds. Plant Cell **15:** 845-853

Zhu X, Galili G (2004) Lysine metabolism is concurrently regulated by synthesis and catabolism in both reproductive and vegetative tissues. Plant Physiol **135:** 129-136

Zimmermann P, Hirsch-Hoffmann M, Hennig L, Gruissem W (2004) GENEVESTIGATOR. Arabidopsis microarray database and analysis toolbox. Plant Physiol **136:** 2621-2632

Zolman BK, Monroe-Augustus M, Thompson B, Hawes JW, Krukenberg KA, Matsuda SP, Bartel B (2001) *chy1*, an *Arabidopsis* mutant with impaired β-oxidation, is defective in a peroxisomal β-hydroxyisobutyryl-CoA hydrolase. J Biol Chem **276:** 31037-31046

Chapter 2

Transcript analyses of BCAA catabolism genes support the participation of putative BCKDH complex subunits in BCAA catabolism[1]

[1] BCAA catabolism gene transcript coexpression analyses in four microarray datasets presented in this chapter were from collaboration with Sahra Uygun. Sahra collected data from four microarray datasets, calculated pairwise PCC values, determined thresholds for calling significant coexpression in each dataset, and provided R codes for drawing heat maps. I interpreted the transcript coexpression results, designed, performed and analyzed all other experiments, and wrote the first draft.

2.1 Abstract

Rapid development of transcript profiling technologies facilitates the utilization of correlative analysis in hypothesis generation, while experimental validation becomes the costly and time-consuming step. Although successful examples have been demonstrated using this approach to study plant specialized metabolism, limited cases were shown for primary metabolism. Branched-chain amino acids (BCAAs) are essential amino acids for humans and other animals because they cannot be synthesized *de novo*. Plants synthesize BCAAs and are the main source of these essential nutrients for humans and agriculturally important animals in their diets. However, the BCAA contents in plant foods are insufficient to meet dietary needs. Therefore, an optimized BCAA metabolism is desired. However, the full BCAA catabolic network is still not completely understood in plants, and limited information is available regarding its regulation. In this chapter, transcript analyses were performed to provide evidence supporting the participation of putative branched-chain ketoacid dehydrogenase subunits. The results revealed positive correlations between transcripts of proposed and validated BCAA catabolism genes in plants subjected to various stress conditions and light quality and quantities, and revealed regulation under diurnal/circadian conditions. These gene transcripts show coordinated oscillation in diurnal and circadian treatments, and exhibit altered expression patterns in mutants defective in components of the circadian clock and photoreceptors, suggesting that the clock and light regulate BCAA catabolism. In addition, BCAA catabolic enzyme gene transcripts are elevated and remain high during prolonged darkness, supporting the hypothesis that BCAA catabolic enzymes serve physiological roles in the dark. Moreover, possible functional divergence is suggested by transcript profile differences between four pairs of

BCAA catabolism gene paralogs, and is supported by no transcript compensation in single mutants of these paralogs. Together, these results provide evidence that putative catabolic enzymes participate in BCAA catabolism, and suggest transcriptional regulation of BCAA catabolism.

2.2 Introduction

With the rapid development of transcript profiling technologies, scientists are able to obtain a better view of the quantity and quality of the entire transcriptome to date. Correlative approaches have been developed to utilize large-scale transcript data in advancing our understanding in plant biology, with the underlying hypothesis being genes correlated tend to participate in the same process or are functionally related (Saito et al., 2008). One of such correlative approaches – transcript coexpression analysis with validated genes of a particular biological process as baits - has been successfully used for identification of novel pathway genes and regulators in cell wall biosynthesis and secondary metabolism (Brown et al., 2005; Persson et al., 2005; Hirai et al., 2007; De Luca et al., 2012). However, fewer examples were demonstrated in primary metabolism - possibly due to the complicated and highly inter-connected pathway topology and overlapping functionality (Stitt, 2013).

The branched-chain amino acids (BCAAs) Leu, Ile and Val are among nine amino acids essential for humans and other animals because they cannot be synthesized *de novo* (Harper et al., 1984). Plants synthesize BCAAs and are the main source of these essential nutrients in the diets of humans and agriculturally important animals. Strong correlations between the levels of free BCAAs were found in wild-type *Arabidopsis thaliana* seeds and tomato (*Solanum lycopersicum*) fruits (Schauer et al., 2006; Lu et al., 2008), which suggests co-regulation of biosynthesis and/or degradation. This presumably is due - at least in part - to the fact that they share four common biosynthetic enzymes and three catabolic steps (Figure 1.2, Figure 1.3). In addition, recent studies revealed that free BCAA levels oscillate in diurnal and circadian conditions (Espinoza et al., 2010) and increase during prolonged darkness in rosette leaves of *A.*

thaliana wild-type plants (Ishizaki et al., 2005). These results further suggest a coordinated regulation of the parallel pathways between BCAA biosynthesis and/or catabolism.

The BCAA biosynthetic pathway and its regulation have been investigated in *A. thaliana* and other plants for the past two decades, in large part because of the commercial importance of herbicides that inhibit acetohydroxy acid synthase, which is the committing enzyme of BCAA biosynthesis (Singh and Shaner, 1995; Aubert et al., 1997; Singh, 1999; McCourt et al., 2006; Tan et al., 2006; Binder, 2010; Chen et al., 2010; Yu et al., 2010). Three enzymes - threonine deaminase, acetolactate synthase and isopropylmalate synthase - locating at committed steps towards the biosynthesis of individual BCAAs are subject to allosteric regulation and feedback inhibited by the synthesized amino acids (Figure 1.2) (Binder, 2010). Aside from allosteric regulation, additional regulatory mechanisms - such as transcriptional regulation - have not been found to date for BCAA biosynthesis (Less and Galili, 2008; Espinoza et al., 2010).

Despite long-term interest in the desirability of optimizing the content of these essential amino acids in plants, the genes and proteins that constitute the full BCAA catabolic network are not completely characterized in *A. thaliana* or any other plant, and there is much to learn about the genetic and biochemical regulation of this process (Figure 1.3). Evidence suggesting a common expression pattern for validated and proposed BCAA catabolism genes was demonstrated in several studies. Dramatic increases in the transcript levels of *A. thaliana BRANCHED-CHAIN AMINOTRANSFERASE 2* (*BCAT2*), *BRANCHED-CHAIN KETOACID DEHYDROGENASE E2* (*E2*), *ISOVALERYL-COA DEHYDROGENASE 1* (*IVD1*), *METHYLCROTONYL-COA CARBOXYLASE A1* (*MCCA1*) and *METHYLCROTONYL-COA CARBOXYLASE B1* (*MCCB1*) were observed following transition from light to dark, and the induction could be inhibited by sucrose (Fujiki et al., 2000; Daschner et al., 2001; Schuster and

Binder, 2005). Increased *IVD1* transcripts were also demonstrated during prolonged darkness (Araujo et al., 2010). Moreover, the Eve Wurtele group at Iowa State University reported that several BCAA catabolism genes were in a coexpression supermodule capable of maintaining cellular energy balance via catabolism of a variety of compounds, including amino acids, carbohydrates, lipids and cell wall components (Mentzen et al., 2008). The Gad Galili group at the Weizmann Institute of Science demonstrated *BCAT2* coexpressed with genes encoding the committed enzymes in Lys, Met and Thr degradation (Less and Galili, 2008). These results suggest a potential co-regulation of BCAA catabolism gene transcripts. However, no comprehensive evaluation using all known and proposed BCAA catabolism genes is published. Such an in-depth analysis would aid in functional characterization of new pathway genes and to reveal additional physiological role(s) of plant BCAA catabolism.

Although genetic and biochemical evidence exist for the participation of *A. thaliana* enzymes BCAT2, IVD, MCCA, MCCB and HML in BCAA catabolism (Gu et al., 2010; Lu et al., 2011; Ding et al., 2012; Angelovici et al., 2013), much less is known about the genes and encoded proteins for the branched-chain ketoacid dehydrogenase (BCKDH) complex. Published biochemical evidence demonstrated the BCKDH complex enzyme activity in isolated *A. thaliana* mitochondria (Taylor et al., 2004). The better characterized mammalian BCKDH is comprised of multiple copies of three proteins: the α-ketoacid dehydrogenase/carboxylase E1 (E1α and E1β), dihydrolipoyl acyltransferase E2, and dihydrolipoyl dehydrogenase E3 (also known as the mitochondrial lipoamide dehydrogenase, mtLPD) (Mooney et al., 2002). This multi-mega-Dalton molecular mass complex catalyzes the second step of BCAA degradation, converting branched-chain ketoacids (intermediates in BCAA biosynthesis and catabolism) into acyl-CoA esters (Figure 1.2, Figure 1.3). The complex is homologous to mitochondrial and chloroplastic

pyruvate dehydrogenase complexes, and mitochondrial α-ketoglutarate dehydrogenase complex, and shares enzyme subunits with the Gly decarboxylase complex (Oliver, 1994; Mooney et al., 2002).

The large size of the BCKDH complex has hindered a detailed *in vitro* characterization in plants, and identification of the plant BCKDH complex subunits is based upon sequence annotation rather than functional analysis. Pairs of paralogous genes are annotated as encoding *A. thaliana* E1α, E1β and E3 subunits, and one gene for E2. These assignments are based on: 1) sequence similarity with proteins identified from other organisms, especially mammals, and 2) mitochondrial localization evidence using tandem mass spectrometry, rather than enzyme activities (Fujiki et al., 2000; Mooney et al., 2000; Taylor et al., 2004).

This chapter describes results from transcript profiling analyses used to evaluate genes encoding putative subunits of the *A. thaliana* BCKDH complex, and to study the regulation of BCAA catabolism at the transcript level. The steady-state mRNA levels for eight experimentally validated or proposed BCAA catabolism genes are co-regulated in a variety of conditions. The expression of these eight genes is altered in mutants defective in components of the circadian clock and photoreceptors. In addition, gene paralogs in this pathway exhibit different expression patterns and do not compensate for the loss of one paralog by elevating the transcript level of the other, suggesting possible functional divergence. These results provide evidence supporting the participation of the putative BCKDH complex subunits E1α1, E1β1, E1β2 and E2 in BCAA catabolism. Moreover, these data suggest a role for BCAA catabolism in maintaining amino acid homeostasis under day/night cycles and prolonged darkness.

2.3 Results

2.3.1 Known and hypothesized *A. thaliana* BCAA catabolism genes form a coexpression module

Transcripts of the functionally validated *A. thaliana* BCAA catabolism genes *BCAT2*, *IVD1*, *MCCA1* and *MCCB1* were reported to increase rapidly following the transition from light to dark (Daschner et al., 2001; Fujiki et al., 2001; Schuster and Binder, 2005; Araujo et al., 2010; Angelovici et al., 2013). These results led us to hypothesize that candidate genes encoding additional enzymes in BCAA catabolism could be identified by their coexpression with these known genes. Coexpression analysis was performed with 13 genes encoding proteins proposed (based on sequence similarity) or experimentally validated to participate in BCAA catabolism (the gene list is shown in Table 2.1). First, to determine the extent of coexpression among these genes, pairwise Pearson's Correlation Coefficients (PCCs) were calculated, and average-linkage hierarchical clustering was performed with four datasets: development, abiotic and biotic stress, and light datasets from AtGenExpress (Schmid et al., 2005; Kilian et al., 2007), and the diurnal/circadian dataset from the DIURNAL database (Mockler et al., 2007) (Figure 2.1, Table 2.2).

The degrees of coexpression among the five validated BCAA catabolism 'bait' genes (*BCAT2*, *IVD1*, *MCCA1*, *MCCB1* and *HML1*) were evaluated to establish the foundation for the coexpression analyses. In the stress dataset, all five bait genes form a coexpression module with significant expression correlations (Figure 2.1, Figure 2.2). In the other three datasets (development, diurnal/circadian and light), four out of the five bait genes form coexpression modules. These modules (one from each dataset) contain three common bait members - *BCAT2*,

IVD1 and *MCCA1* - with *MCCB1* in only the diurnal/circadian and the light modules, and *HML1* only in the development (Figure 2.1, Figure 2.2). Because of the significant expression correlations among bait genes, these results indicate the feasibility of this approach for identifying candidate genes in the BCAA catabolic pathway. Next, coexpression of the eight proposed BCAA catabolism genes with the five bait genes was evaluated (dotted rectangles in Figure 2.1, Figure 2.2). The stress coexpression module contains the largest number of members: five proposed genes in addition to the five bait genes. Furthermore, four genes proposed to encode proteins of BCAA catabolism and four validated bait genes are in the light and the diurnal/circadian coexpression modules. In contrast, one of the proposed genes is in the development module that contains four bait genes. In summary, highly interconnected modules containing validated and proposed BCAA catabolic enzyme genes emerged from three datasets: stress, diurnal/circadian, and light. These modules have eight members in common: the validated bait genes *BCAT2*, *IVD1*, *MCCA1* and *MCCB1*, and proposed BCKDH genes *E1A1*, *E1B1*, *E1B2* and *E2* (Figure 2.1, Figure 2.2, Table 2.2).

2.3.2 BCAA catabolic gene transcripts oscillate in diurnal/circadian conditions

The observation that four out of five bait genes - and four out of eight proposed BCAA catabolic enzyme genes - are strongly coexpressed in the diurnal/circadian dataset (Figure 2.1, Figure 2.2), suggests that these genes are coordinately regulated by light and the circadian clock. Further examination of transcript profiles confirmed common oscillation patterns among these eight validated or proposed BCAA catabolism genes under most diurnal/circadian experiments (Figure 2.3). It is notable that these gene transcript levels are reduced during the day, and increased during the night in the rosette leaves of 4-week-old, short-day-grown (8h light/16h

dark) plants (Figure 2.4A). This common oscillation pattern is consistent with the previous observation that free BCAA levels fluctuate in day/night cycles and peak towards the end of the day (Gibon et al., 2006; Espinoza et al., 2010), suggesting that up-regulation of the catabolism genes contributes to the decreased free BCAAs at night. In addition, the eight transcripts also oscillate in a similar fashion in constant light after entrainment in light/dark or hot/cold cycles (Figure 2.3, Figure 2.4B), consistent with the hypothesis that BCAA catabolism is subject to regulation by the circadian clock. These results further demonstrated the coexpression of BCAA catabolic enzyme gene transcripts in diurnal/circadian conditions, especially on a day/night cycle.

2.3.3 Altered BCAA catabolic gene transcript levels in mutants defective in components of the circadian clock and photoreceptors

The expression profiles of *BCAT2*, *E1A1*, *E1B1*, *E1B2*, *E2*, *IVD1*, *MCCA1* and *MCCB1* - identified as highly coexpressed in the diurnal/circadian dataset (Figure 2.1, Figure 2.2) - were examined in mutants defective in various components of the circadian clock and photoreceptors for further evaluation of the participation of the clock and light in regulating BCAA catabolism. These gene transcripts exhibited highly correlated oscillation during short day in *LHYOX* (overexpressor of *LATE ELONGATED HYPOCOTY*, *LHY*) and *phyB-9* (mutant defective in PHYTOCHROME B, PHYB) (Figure 2.5 A and B). In addition, expression profile comparison between the two clock mutants and their corresponding wild type lines revealed elevated transcript levels of these eight genes later at night during short day, with the maximum increased to two- to three-fold relative to that of the wild type (*BCAT2* is shown in Figure 2.5C as an example). Together, these results are consistent with the hypothesis that the circadian clock and

light regulate BCAA catabolism directly or indirectly via transcription factor LHY and red/far-red photoreceptor PhyB.

2.3.4 Regulatory regions for *IVD1* transcript oscillation during short day revealed by promoter mutagenesis experiments

Because of the highly co-regulated transcript levels of *BCAT2*, *E1A1*, *E1B1*, *E1B2*, *E2*, *IVD1*, *MCCA1* and *MCCB1*, a common *cis*-element was hypothesized in the promoter of these genes to regulate their oscillation on day/night cycles. *IVD1* was chosen as a representative for promoter mutagenesis assays to locate the regulatory *cis*-element(s). *IVD1* promoter fragments with or without 5' untranslated regions (UTRs) were inserted into a construct that has a luciferase (LUC) reporter next to the inserted region (Figure 2.6A). These constructs were transformed into *A. thaliana* Col-0 plants, and bioluminescence was monitored for primary transformants (see Methods and Materials). A total of 11 different combinations of various *IVD1* promoter fragments with or without 5' UTR were analyzed (Figure 2.6B) and compared to the *IVD1* expression pattern in Col-0 during short day growth conditions (Figure 2.4A). Representative results are shown in Figure 2.6C. The *IVD1* promoter fragment containing 100bp upstream from the transcription start site (TSS) and the 5' UTR was necessary to drive *LUC* expression in a similar oscillation pattern with *IVD1* in wild type under short day. Constructs containing either the 5'UTR alone or any promoter fragment without the 5'UTR failed to mimic the oscillation of *IVD1* transcripts. Together, these results suggest that *IVD1* transcript oscillation on a day/night cycle is regulated by a combined effect of the *IVD1* -100bp promoter region and its 5'UTR.

2.3.5 BCAA catabolism gene transcripts increase during prolonged darkness

Published studies revealed that transcripts of the known BCAA catabolism genes *BCAT2, IVD1*, *MCCA1* and *MCCB1* increase following transition from light to dark in plants grown under light-dark conditions and during prolonged darkness (Daschner et al., 2001; Fujiki et al., 2001; Schuster and Binder, 2005; Araujo et al., 2010). I used RNA-Seq analysis to ask whether the transcripts for the proposed BCKDH subunits are also upregulated in plants subjected to prolonged darkness. These experiments were performed using rosette leaves from 5-week-old, short-day-grown (8h light/16h dark) Col-0 wild-type plants moved to constant darkness for 6h, 24h, 48h and 72h (Figure 2.7). The transcripts of the eight genes found in the coexpression modules derived from stress, light and diurnal/circadian datasets (Figure 2.1, Figure 2.2) were increased nine- to 400-fold within the first 6h of prolonged darkness, and remained high until the last time point (gene names highlighted in blue in Figure 2.7). These results are consistent with the hypothesis that BCAA catabolic enzymes - including BCKDH subunits E1A1, E1B1, E1B2 and E2 - have one or more physiological roles in the dark.

2.3.6 Whole genome transcript coexpression analyses in prolonged darkness

Because eight BCAA catabolism genes - *BCAT2*, *E1A1*, *E1B1*, *E1B2*, *E2*, *IVD1*, *MCCA1* and *MCCB1* - showed elevated transcripts in prolonged darkness (Figure 2.7), the pairwise PCCs were calculated to determine the extent of the coexpression among these genes in prolonged darkness. All gene pairs were significantly coexpressed (Table 2.3), with PCC values ranging from 0.918 (between *E1B2* and *E2*) to 0.995 (between *E1B2* and *IVD1*). To further identify genes that coexpress with the eight BCAA catabolism genes, pairwise PCCs were calculated for all genes detected in Col-0 wild-type plants in replicates of at least one out of the six time points

in my RNA-Seq dataset. A total of 3244 genes were found to be significantly coexpressed with the eight genes (with PCC r>0.854, which is the 95th percentile of the null distribution of 500,000 random gene pairs in my RNA-Seq dataset, see details in material and methods). To further reduce the number of coexpressed genes for gene ontology (GO) enrichment analyses, a more stringent threshold (r>0.900) was used, and 2338 genes were found to have PCCs above 0.900 with the eight BCAA catabolism genes. Out of the 2338 genes, 2084 genes coexpressed with at least two out of the eight genes in prolonged darkness, and the 2084 genes were used for GO enrichment analyses. Two GO enrichment analysis tools were used: the Biological Network Gene Ontology tool (BiNGO, http://www.psb.ugent.be/cbd/papers/BiNGO/Home.html) (Maere et al., 2005), and the Database for Annotation, Visualization and Integrated Discovery (DAVID, http://david.abcc.ncifcrf.gov/home.jsp) (Huang da et al., 2009). A plant-specific, simplified ontology file, GOSlim-Plants, was selected for BiNGO to present a broad overview of enriched GO categories with focus on biological process (GO_BP), and 17 GO_BP slim terms were found to be enriched (Table 2.4). A more detailed GO enrichment analysis was performed using all GO_BP terms with the DAVID functional classification tool. 285 GO_BP terms were significantly enriched (corrected p-value<0.05), 94 clusters of functionally related GO terms were generated, and the top five clusters contain GO_BP terms related to RNA splicing, processing and metabolic processes, protein and macromolecule catabolic processes, cellular metabolism and regulation, fatty acid and lipid catabolism, and protein and macromolecule transport (Table 2.5).

2.3.7 Distinct expression patterns among BCAA catabolic enzyme paralogs

There are four pairs of paralogous genes encoding known or proposed enzymes in the first two steps of BCAA catabolism (Table 2.1). However, it is not known whether all paralogs participate in BCAA catabolic process because BCAT activities are involved both in BCAA biosynthesis and degradation (Diebold et al., 2002), and some BCKDH subunits are similar to those of pyruvate dehydrogenases in mitochondria and chloroplasts, and mitochondrial α-ketoglutarate dehydrogenase and Gly decarboxylase (Oliver, 1994; Mooney et al., 2002). It was hypothesized that genes encoding these BCAA degradation enzyme paralogs would show coexpression with the documented catabolic enzymes because of their significant expression correlations revealed from our previous coexpression analyses (Figure 2.1, Figure 2.2). Indeed, four out of eight paralogs - *BCAT2* and BCKDH subunits *E1A1*, *E1B1* and *E1B2* - are members of the coexpression modules identified from the light and the diurnal/circadian datasets (Figure 2.1, Figure 2.2). This observation provided evidence that these proteins have roles in BCAA catabolism.

Expression divergence is often observed among paralogous gene pairs that can be due to modification of regulatory elements during evolution (Shariati and De Strooper, 2013), and likely results in function diversification (Stern and Orgogozo, 2008). To evaluate the degree of expression divergence, I compared the expression profiles of the four paralogous pairs. The flowering and seed development datasets were included in this analysis because the known BCAA catabolism genes *BCAT2*, *IVD1* and *MCCA1* and *HML1* have increased transcripts during flowering and seed development (http://bar.utoronto.ca/efp/cgi-bin/efpWeb.cgi). In addition, data for the short-day growth (8h light/16h dark) experiments were also selected because of the highly coordinated transcript oscillations described above (Figure 2.4A). Moreover, RNA-Seq data

during prolonged darkness were also chosen because of the transcript increases of eight out of 13 proposed or validated BCAA catabolism genes shown above (Figure 2.7).

All four pairs of paralogs were found to have divergent expression profiles in these datasets that followed two general patterns (Figure 2.8). First, expression of paralogs *BCAT1/2*, *E1A1/2* and *E1B1/2* was anti-correlated during the late stages of seed development (Figure 2.8, left panel, PCC value r=-0.98, -0.55, -0.87, respectively, for the last four time points), suggesting more important roles in seed amino acid regulation for the isoforms encoded by the more highly-expressed transcripts. Varying degrees of anti-correlation were also found for *BCAT1/2*, *E1A1/2* and *mtLPD1/2* under short day (Figure 2.8, middle panel, r=-0.41, -0.82, -0.32, respectively), and between *E1A1* and *E1A2* during prolonged darkness (Figure 2.8, right panel, r=-0.94). The second pattern of expression divergence is illustrated by *mtLPD1/2* during development (Figure 2.8, left panel), *E1A1/2*, *E1B1/2* and *mtLPD1/2* under short day growth conditions (Figure 2.8, middle panel), and *BCAT1*/*BCAT2*, *E1B1*/*E1B2* and *mtLPD1*/*mtLPD2* during prolonged darkness (Figure 2.8, right panel), where one copy is expressed at a higher level than the other (Ganko et al., 2007; Zou et al., 2009). Among these pairs, the expression profiles of *E1B1* and *E1B2* are most similar than the other pairs under short day and during prolonged darkness (r=0.66 and 0.99, respectively).

In summary, examples of divergent expression profiles of paralogs can be distinguished in the development, short day diurnal cycling, or prolonged darkness datasets, indicating divergence in mRNA regulation. The differences in expression patterns suggest a possible functional divergence between the paralogs. In addition, this expression divergence under short day also explains why *BCAT1*, *E1A2*, *mtLPD1* and *mtLPD2* - genes proposed for BCAA

catabolism - are not found in the coexpression module in the diurnal/circadian dataset (Figure 2.1, Figure 2.2).

2.3.8 BCAA catabolism gene paralogs do not compensate at transcript level in prolonged darkness

Deleting a duplicate gene often leads to less severe phenotype comparing to removing a singleton gene (Gu et al., 2003; Hanada et al., 2011), partially because paralogs sometimes compensate the loss of one copy by elevating the expression of the other (Kafri et al., 2005). To look for potential compensation at the transcript level, homozygous T-DNA insertion lines of these paralogs were obtained. Quantitative PCR analysis revealed that *e1a1-1, e1a1-2, e1a2-1* and *e1b2-1* are null alleles (Figure 2.9). Because the *mtlpd2-2* mutant showed approximately 50% reduction of the *mtLPD2* transcripts relative to wild type, it was not included in further evaluation. Unfortunately, no homozygous T-DNA lines could be identified for *E1B1* and *mtLPD1*. The mRNA of *BCAT1*, *BCAT2*, *E1A1*, *E1A2* and *E1B1* were examined in 5-week-old *bcat1-1*, *bcat2-1*, *e1a1-1*, *e1a2-1* and *e1b2-1* single mutants treated in prolonged darkness for three days. For all transcripts tested, no statistical significance ($p<0.05$ determined by Student's *t*-test) was detected between single mutants and corresponding wild types (Figure 2.10). These results revealed that BCAA catabolism gene paralogs *BCAT1*/*BCAT2* and *E1A1*/*E1A2* do not compensate for each other, and *E1B1* does not compensate for the loss of *E1B2* at transcript level in prolonged darkness, further suggesting functional divergence between these paralogs in prolonged darkness.

2.4 Discussion

Advances in genomics technologies have created new opportunities to identify candidate genes and proteins for complex physiological and biochemical processes. This has created a situation where it is relatively inexpensive and straightforward to mine data and create hypotheses, while testing theories is often the costly and time consuming step. Analysis of large multi-subunit enzymes is especially complicated since it requires expression of multiple subunits, assembly of the component parts and establishment of an *in vitro* activity assay that faithfully represents the *in vivo* process. While *A. thaliana* genes were annotated for the component subunits, establishing *in vitro* functions of the proteins in a complex is problematic. The plant megadalton, multi-subunit BCKDH enzyme is an example of such a complex. In this chapter I describe results from transcript analyses that provide evidence of the participation of putative BCKDH complex subunits in BCAA catabolism, and suggest transcriptional regulation of BCAA catabolism.

2.4.1 Considerations in using transcript coexpression analysis in gene functional characterization

There are important considerations in using gene expression analysis to develop or test hypotheses regarding gene function. First, it is essential to have a set of validated 'bait' genes for the analysis; in this study five bait genes of BCAA catabolism were employed (Table 2.1). Second, success of the approach requires that appropriate transcript profiling data be available. In this study I explored multiple microarray datasets (Figure 2.1, Table 2.2); these include a developmental dataset, with a variety of tissues ('development'), plants grown under differing light and temperature regimes ('diurnal/circadian' and 'light'), as well as abiotic and biotic stress

conditions ('stress'). The broad choice of types of datasets was important in this study because each yielded coexpression modules of different sizes and structures (Figure 2.2). For example, analysis of the stress dataset produced a module with the largest number of nodes (10), while the diurnal/circadian dataset yielded a smaller (eight) but more highly connected module (Figure 2.2). In contrast, the development dataset was least useful, with a weakly connected module of four 'bait' genes that have no significant expression correlation with the proposed BCAA catabolic enzyme genes (Figure 2.2). Taken together, our results demonstrate the importance of using validated bait genes and different types of expression data to identify candidate genes.

Coexpression analysis can be confounded in a number of ways. One scenario is when paralogous genes exist, with one copy that actively responds to environmental perturbations, while the other generally expresses at a much lower level (Keith et al., 1991; Duarte et al., 2006; Ganko et al., 2007; Szekely et al., 2008; Zou et al., 2009). For example, I found that the E1α subunit gene *E1A1* is coexpressed with the other core BCAA catabolic genes, while the paralog *E1A2* is not (Figure 2.1, Figure 2.2). This result suggests possible functional divergence, and functional characterization is needed for further evaluation, which is shown in Chapter 3.

Using coexpression analysis to develop hypotheses regarding function also can be confounded if the protein product participates in multiple different processes. For example, mtLPD is proposed to function as a subunit of the mitochondrial pyruvate dehydrogenase, α-ketoglutarate dehydrogenase, Gly decarboxylase and BCKDH complexes (Lutziger and Oliver, 2001). Needs for this subunit in physiologically diverse processes might be responsible for the observed lack of co-regulation with other known and proposed genes of BCAA degradation (Figure 2.1, Figure 2.2). In addition, the single-copy and experimentally validated hydroxymethylglutaryl-CoA lyase gene *HML1* displayed coexpression with other BCAA

catabolic genes in the stress and the development datasets (Figure 2.1, Figure 2.2). I hypothesize that lack of membership of *HML1* mRNA in the light and the diurnal/circadian coexpression modules reflects that this enzyme participates in other metabolic processes (Ashmarina et al., 1994; Ashmarina et al., 1999). Alternatively, perhaps this and other non-co-regulated BCAA catabolic enzymes are subject to post-transcriptional regulation.

2.4.2 Regulation of BCAA catabolism by the circadian clock and light

Due to their sessile nature, plants have evolved with complex regulatory mechanisms helping them with adaptation to the constantly changing environment. Light and temperature are two major environmental factors for plant growth and development (Went, 1953; Kami et al., 2010). The plant circadian clock can perceive such changes in light and temperature, and regulate relevant metabolic processes accordingly (Harmer et al., 2000). Moreover, once the clock is entrained by light and temperature, plants can predict cycling environmental changes and actively re-adjust themselves. As important primary metabolites, free BCAA levels were found to oscillate during diurnal and circadian conditions (Espinoza et al., 2010). However, no information was available regarding to the regulation of BCAA metabolism gene transcripts or enzymes in diurnal and circadian conditions that contributes to the metabolite fluctuation. In this analysis, transcripts of eight BCAA catabolism genes (*BCAT2*, *E1A1*, *E1B1*, *E1B2*, *E2*, *IVD1*, *MCCA1* and *MCCB1*) were found to coordinately oscillate in diurnal and circadian conditions (Figure 2.1, Figure 2.3, Figure 2.4), and these genes are directly or indirectly regulated by components of the circadian clock and photoreceptors (Figure 2.5). These data provides first hand evidence that is consistent with the hypothesis that BCAA catabolism is regulated by the clock and light at transcript level. In addition, results shown in this chapter revealed that both the

-100bp promoter region and 5' UTR of *IVD1* are required for rhythmic oscillation of the *IVD1* transcript during short day (Figure 2.6). This suggests that multiple regulators might be involved in maintaining rhythmic BCAA catabolism gene expression during day/night cycles.

2.4.3 Induction of BCAA catabolism gene expression by dark

Prolonged darkness is often used by scientists to study plant response upon carbon starvation. Several metabolic processes were found to be upregulated and essential for plant survival in such conditions, including autophagy (Hanaoka et al., 2002; Thompson and Vierstra, 2005) and the mitochondrial electron transport (Ishizaki et al., 2005; Ishizaki et al., 2006; Schertl and Braun, 2014). In this study, elevated transcript levels of BCAA catabolism genes *BCAT2*, *E1A1*, *E1B1*, *E1B2*, *E2*, *IVD1*, *MCCA1* and *MCCB1* were demonstrated in the first 6h in prolonged darkness, and the transcripts remained high until the last time point (3d) (Figure 2.7). These results suggest that BCAA catabolism serves physiological role(s) in the dark. Testing of this hypothesis and related results are shown in Chapter 3.

In an attempt to locate *cis*-element(s) regulating the dark inducibility of *IVD1*, I used the same 2-week-old transgenic lines containing LUC reporter driven by various regions of *IVD1* promoter with or without its 5' UTR – the ones created for investigation of the *cis*-element(s) regulating the diel oscillation of *IVD1* (Figure LUC 2.6A) – to look for induction in bioluminescence during a 2-day dark treatment. However, no transgenic lines tested showed consistent luminescence induction by prolonged darkness. To eliminate the possibility that ATP depletion was the cause, 2-week-old seedlings were transferred to soil, grown for an additional two weeks, and tested for luminescence induction by prolonged darkness again. However, no consistent induction was observed, suggesting that *IVD1* promoter with 5' UTR is not sufficient

for its elevated transcript level in prolonged darkness. Besides promoter and 5' UTR, introns were shown to stimulate gene expression in both monocots and dicots (Callis et al., 1987; Luehrsen and Walbot, 1991; Norris et al., 1993; Xu et al., 1994). For an example, either of the first two introns of the *A. thaliana* phosphoribosylanthranilate transferase gene *PAT1* – which encodes an enzyme in tryptophan biosynthesis - enhances mRNA accumulation without affecting the rate of transcription(Rose and Last, 1997; Rose, 2002, 2004, 2008). Further investigation should examine the effect of *IVD1* introns on gene expression to test this hypothesis.

2.5 Material and methods

2.5.1 Plant materials and growth conditions

A. thaliana ecotype Columbia CS60000, T-DNA lines *bcat1-1* (SALK_138630), *bcat2-1* (SALK_037854), *e1a1-1* (SALK_071680), *e1a1-2* (WiscDsLox470G12), *e1a2-1* (SAIL_113_D07), *e1b2-1* (SALK_098054) and *mtlpd2-2* (SALK_027039) were obtained from the Arabidopsis Biological Resource Center (ABRC). Homozygous mutant lines that were obtained or validated were deposited at ABRC: *bcat1-1* stock CS68922, *bcat2-1* CS68923, *e1a1-1* CS68924, *e1a1-2* CS68925, *e1a2-1* CS68926, *e1b2-1* CS68927, *mtlpd2-2* CS68928, double mutants *e1a1-1*; *e1a2-1* CS68929 and *e1a1-2*; *e1a2-1* CS68930. The *ivd1-2*, *mcca1-1*, *mccb1-1*, and *hml1-2* mutants were described previously (Gu et al., 2010; Lu et al., 2011). All plants, including double mutants, were genotyped and confirmed to be homozygous with primers in Table 2.3 prior to further analyses. Plants were grown in soil in chambers at 21°C with fluorescent lamps (100 μmol m^{-2} s^{-1}) under different photoperiods (16h for LD, 8h for SD, and continuous darkness). Seeds for amino acid assay were harvested from mature plants grown under LD. Leaves for transcript analysis and amino acid assay were harvested from plants grown under SD for 5 weeks and subjected to various lengths of prolonged darkness after the end of the night.

2.5.2 Transcript coexpression analysis

Coexpression analysis of the previously documented and hypothesized BCAA catabolic enzyme genes was performed with the following microarray datasets: stress (AtGenExpress, abiotic and biotic stress treatments in roots and shoots with time points from 0.5 to 24h) (Kilian

et al., 2007), development (AtGenExpress, atlas of developmental stages consisting of a variety of tissues) (Schmid et al., 2005), light (AtGenExpress, light treatments of different wavelengths, fluence and durations) (Kilian et al., 2007) and diurnal/circadian (Diurnal, combinations of light and temperature conditions) (Mockler et al., 2007). The development, light and diurnal/circadian data were downloaded in a Robust Multi-array Average normalized form from: http://www.weigelworld.org/resources/microarray/AtGenExpress/ and ftp://www.mocklerlab.org/diurnal/. For the stress dataset, the degrees of differential expression (in the form of fold change) was obtained from an earlier study (Zou et al., 2009). For all datasets, pairwise Pearson's Correlation Coefficient (PCC) values were calculated among the genes of interest using SciPy library in Python (http://www.scipy.org/) (Jones et al., 2001). To identify gene pairs with significantly higher than randomly expected PCC values, 500,000 gene pairs were selected randomly from each dataset to calculate PCCs and establish a null PCC distribution. The 95th percentile of each null PCC distribution was used as the threshold for calling two genes as significantly positively correlated with a 5% false positive rate (arrows in Figure 2.1; Figure 2.3; Table 2.2). Coexpression modules were defined by clustered genes with significant correlations in each dataset. Heat maps were generated with levelplot within the R lattice package (Sarkar and Sarkar, 2007).

2.5.3 *IVD1* promoter mutagenesis experiments

IVD1 promoter fragments (-1500bp, -750bp, -500bp, -350bp, -200bp and -100bp, with or without 5'UTR, where +1 denotes the *IVD1* transcriptional start site) were isolated from Col-0 genomic DNA (primers see Table 2.3) and cloned into pZPXomegaLUC^{+} (Schultz et al., 2001). All constructs were sequenced prior to further analyses to check for mutations and/or unwanted

DNA fragments introduced during cloning. Constructs were transformed into *A. thaliana* wild type (Col-0) by floral dip (Clough and Bent, 1998). Two-week-old primary transformants (T1s) were selected by gentamycin resistance and tested for bioluminescence under short day using a Berthold LB960XS3 luminometer (Michael and McClung, 2002; Liu et al., 2013). Average luciferase intensity at every hour for each construct during the 2-day short day growth was calculated based on ~24 T1s tested at the same time.

2.5.4 RNA-Seq Analysis

Leaf tissues from Col-0, *ivd1-2* and *hml1-2* treated for 0h, 6h, 24h, 48h and 72h in prolonged darkness and 72h under SD were harvested for RNA-seq. For each genotype under the same treatment, two replicates were grown and harvested a month apart from each other. Therefore, in total 36 samples were sequenced.

Total RNA from 9th to 12th rosette leaves was extracted using RNeasy plant mini kit (Qiagen, Germantown, MD) and on-column digestion performed with RNase-free DNase Set (Qiagen, Germantown, MD). RNA quality was assessed using the Agilent 2100 Bioanalyzer with the RNA 6000 Pico Chip (Agilent Technologies, Santa Clara, CA). Library construction and sequencing were conducted by the Michigan State University Research Technology Support Facility using Illumina Tru-Seq Stranded kit and following the manufacturer's protocols. Six samples were multiplexed and sequenced in one lane using the Illumina HiSeq 2500 sequencer, and 50 nucleotide single end reads were generated. 16 to 60 million reads were obtained per sample.

For read processing and assembly, the sequencing adapters were removed using the following parameters in Trimmomatic version 0.30 (Bolger et al., 2014): ILLUMINACLIP:

TruSeq3-SE, SLIDINGWINDOW: 4:15, and MINLEN:35. Processed reads were filtered by fastq_quality_filter in the FASTX-toolkit version 0.0.13.2 (http://hannonlab.cshl.edu/fastx_toolkit/index.html), satisfying the criterion that > 85% bases must have Q-score ≥ 20. The processed and filtered reads were mapped to the *A. thaliana* reference genome (TAIR10) using TopHat version 1.4.1 (Trapnell et al., 2009), sorted by SAMtools version 0.1.19 (Li et al., 2009), and analyzed with Cuffdiff version 2.1.1 (Trapnell et al., 2013). The transcript levels represented as Fragments Per Kilobase of exon model per Million mapped reads (FPKM) were visualized with CummeRbund version 2.6.1 (Goff et al., 2012) in R version 3.0.3 (Statistical Package, 2012). The RNA-Seq dataset was deposited in the National Center for Biotechnology Information Gene Expression Omnibus under GEO accession number GSE67956.

Raw read counts from each sample were normalized to Counts Per Million reads (CPM), and an expression matrix was generated for wild-type plants at all six time points each with two replicates. Pairwise Pearson's Correlation Coefficient (PCC) values were calculated using the SciPy library in Python (http://www.scipy.org/) (Jones et al., 2001). The 95th percentile of the null PCC distribution from 500,000 random gene pairs was calculated and used as the threshold for calling two genes as significantly positively correlated with a 5% false positive rate (indicated in Table 2.3). Tables containing all genes that have PCCs with the eight highly coexpressed BCAA catabolism genes (*BCAT2*, *E1A1*, *E1B1*, *E1B2*, *E2*, *IVD1*, *MCCA1* and *MCCB1*) above the thresholds (either 95th percentile of the random PCC distribution or r=0.9000) were generated, and only genes showing significant coexpression with at least two out of the eight BCAA catabolism genes were included for gene ontology enrichment analyses. The BiNGO plugin for Cytoscape version 3.0.3 was used for analyzing GO enrichment

(http://www.psb.ugent.be/cbd/papers/BiNGO/Home.html) (Maere et al., 2005), and DAVID version 6.7 was used for functional annotation clustering (http://david.abcc.ncifcrf.gov/home.jsp) (Huang da et al., 2009). Parameters used were indicated in table legends of Table 2.4 and Table 2.5.

2.5.5 RNA extraction and quantitative reverse transcription-PCR analysis

Total RNA from 9th to 12th rosette leaves was extracted using RNeasy plant mini kit (Qiagen, Germantown, MD) and digested with RNase-free DNase Set (Qiagen, Germantown, MD) on-column. First-strand cDNA was synthesized from 2 μg of total RNA with M-MLV reverse transcriptase (Invitrogen, Carlsbad, CA) and oligo(dT)$_{12\text{-}18}$ primer (Invitrogen, Carlsbad, CA) (Kotewicz et al., 1985). Gene-specific primers were designed to span two or more exons as listed in Table 2.3. Quantitative PCR analyses were performed on a 7500 Fast Real-Time PCR System with Power or Fast SYBR green PCR master mix (Applied Biosystems, Waltham, MA), with *ACT2* (At3G18780) transcript level used as an internal control (Czechowski et al., 2005).

2.6 Acknowledgments

I would like to thank Sahra Uygun for help with coexpression analysis, Kathleen Imre and David Hall for isolating homozygous T-DNA mutants, and Linsey Newton for providing vectors and guidance for luciferase assays.

APPENDICES

APPENDIX A Yeast one hybrid screen for TFs interacting with *IVD1* promoter and 5'UTR

My previous analyses demonstrated that genes proposed or experimentally validated to participate in BCAA catabolism - including *BCAT2*, *E1A1*, *E1B1*, *E1B2*, *E2*, *IVD1*, *MCCA1* and *MCCB2* - coexpress in the stress, diurnal/circadian and light datasets, have common oscillation patterns during diurnal and circadian conditions, and show elevated transcript levels in prolonged darkness. These results suggest a common regulatory mechanism for these eight genes. I hypothesized that one or a group of TFs bound to the promoters of these eight genes and regulated their expression.

To find the promoter-binding TFs, *IVD1* was selected as a representative, and 880bp upstream of the *IVD1* transcriptional start site (TSS, +1) and its 5'UTR were included for yeast one hybrid (Y1H) screen against an *A. thaliana* TF library containing ~2000 TFs. Three overlapping fragments were cloned into pLacZi-GW vectors (Gendron et al., 2012): fragment #1 corresponds to -180bp to +200bp relative to the TSS of *IVD1*, fragment #2 -530bp to -150bp, and fragment #3 -880bp to -500bp. The three constructs were sent to the Pruneda-Paz lab at University of California, San Diego for the Y1H screen (Pruneda-Paz et al., 2014).

TFs presented in Table 2.7 are ones with high confidence as concluded by the Pruneda-Paz lab. A total of 72 TFs were found to interacting with fragment #1, 45 with #2 and 35 with #3. Combined with the results from the luciferase assay that the -100bp *IVD1* promoter region with 5'UTR was able to drive its diurnal oscillation (Figure 2.6), TFs found to bind to fragment #1 should be considered for future evaluation.

APPENDIX B Figures and tables

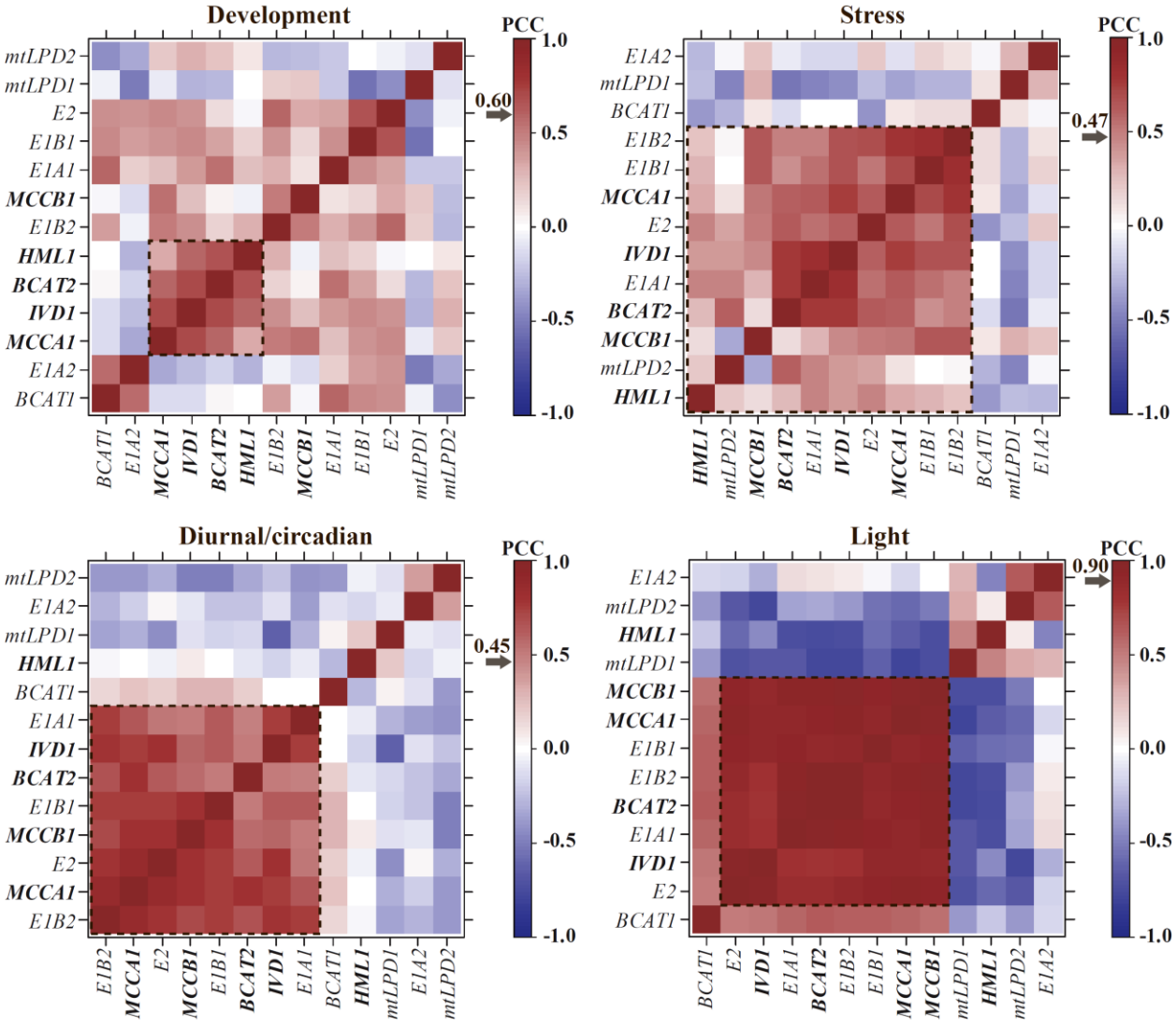


Figure 2.1 Transcript coexpression analysis of known or proposed BCAA catabolism gene transcripts in wild type (Col-0). Four datasets were tested: development (top left), stress (top right), diurnal/circadian (bottom left), and light (bottom right). Values represent the Pearson's Correlation Coefficient (PCC) for each gene pair. Arrows indicate the PCC value representing the 95th percentile of each null PCC distribution for individual datasets. Transcript names for validated genes are shown in bold and italic text, and those for proposed genes in regular italic

Figure 2.1 (cont'd)

text. Dotted rectangles represent coexpression modules in each dataset, defined by clustered genes with significant expression correlations. Heat maps were generated using microarray data from AtGenExpress for stress, development and light datasets (Schmid et al., 2005; Kilian et al., 2007), and the Diurnal database for diurnal/circadian dataset (Mockler et al., 2007).

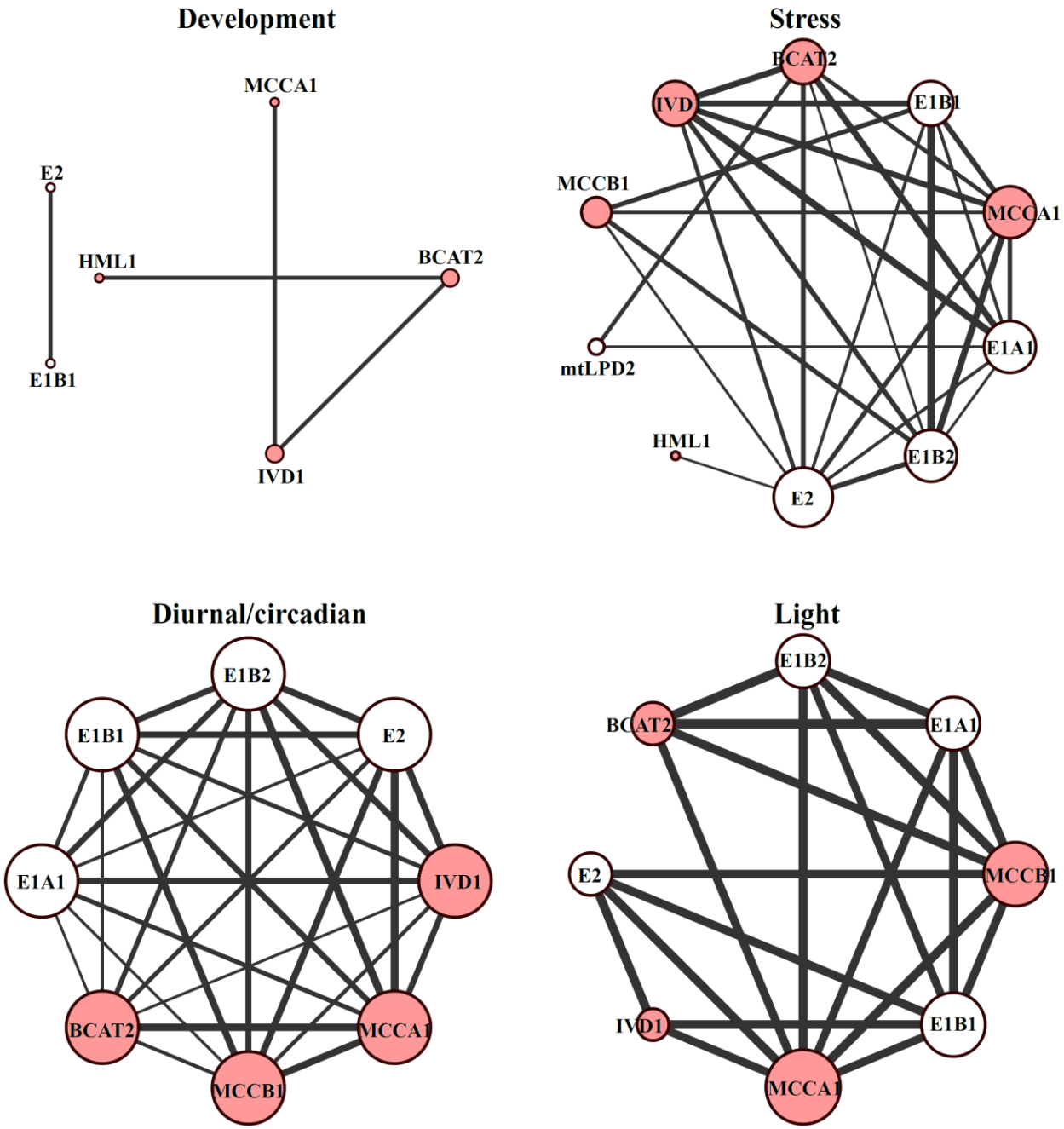


Figure 2.2 Graphical representation of transcript correlation modules among the four datasets. Graphs were constructed using development (top left), stress (top right), diurnal/circadian (bottom left), and light (bottom right) datasets. Nodes represent transcripts and edges indicate pairwise PCC values exceeding the threshold in each database. Transcript names of validated genes are in bold, and proposed genes in regular font. The size of nodes corresponds

Figure 2.2 (cont'd)

to the connectivity of the transcript (smaller node indicates fewer significant correlations with other transcripts, and vice versa). The thickness of edges correlates with their PCC values. See Table 2.2 for PCC values and the corresponding threshold in each dataset. The network graphs were created using Cytoscape.

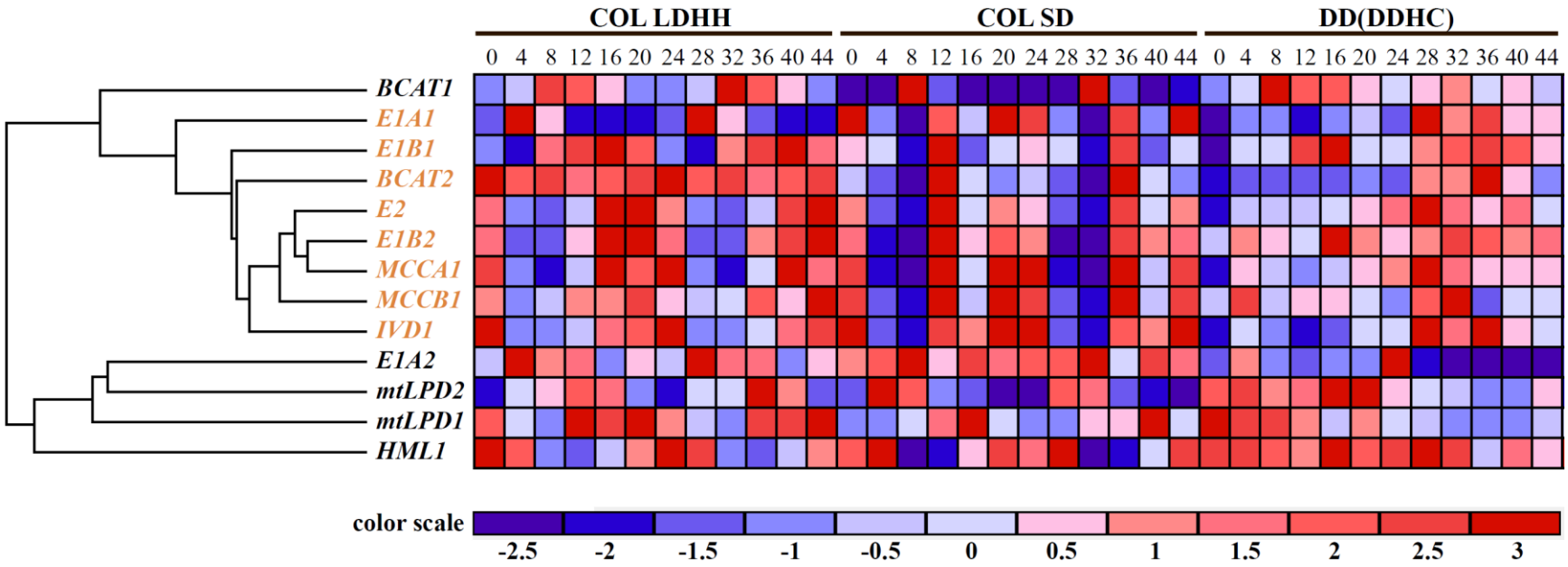
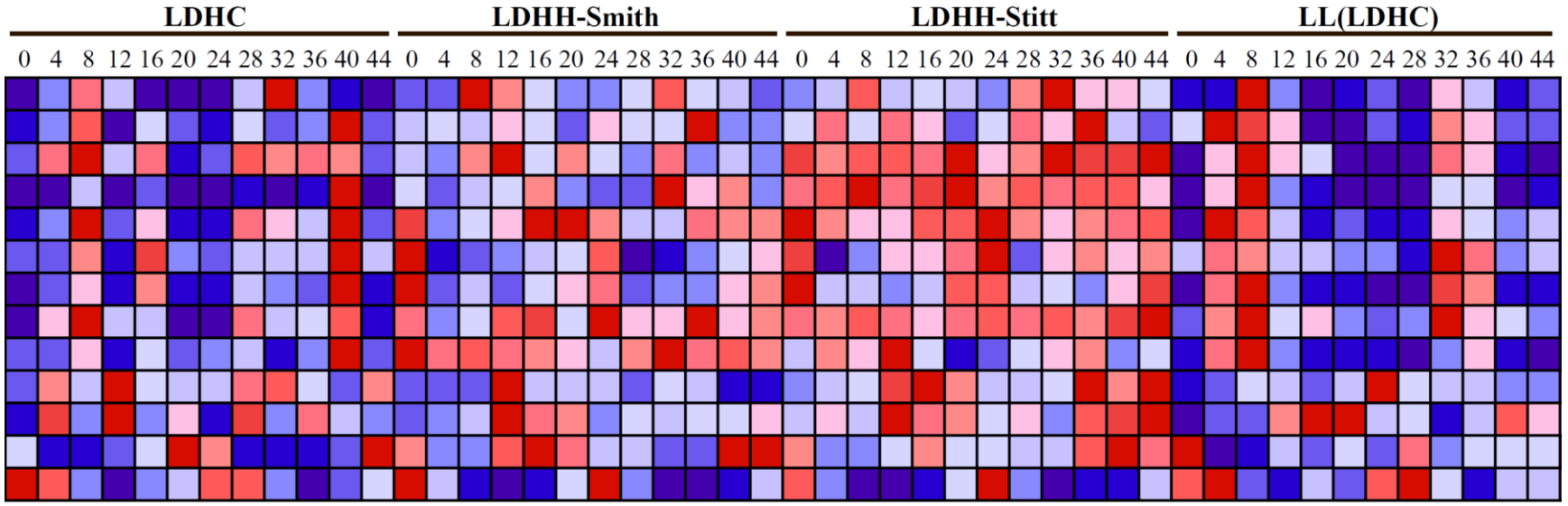


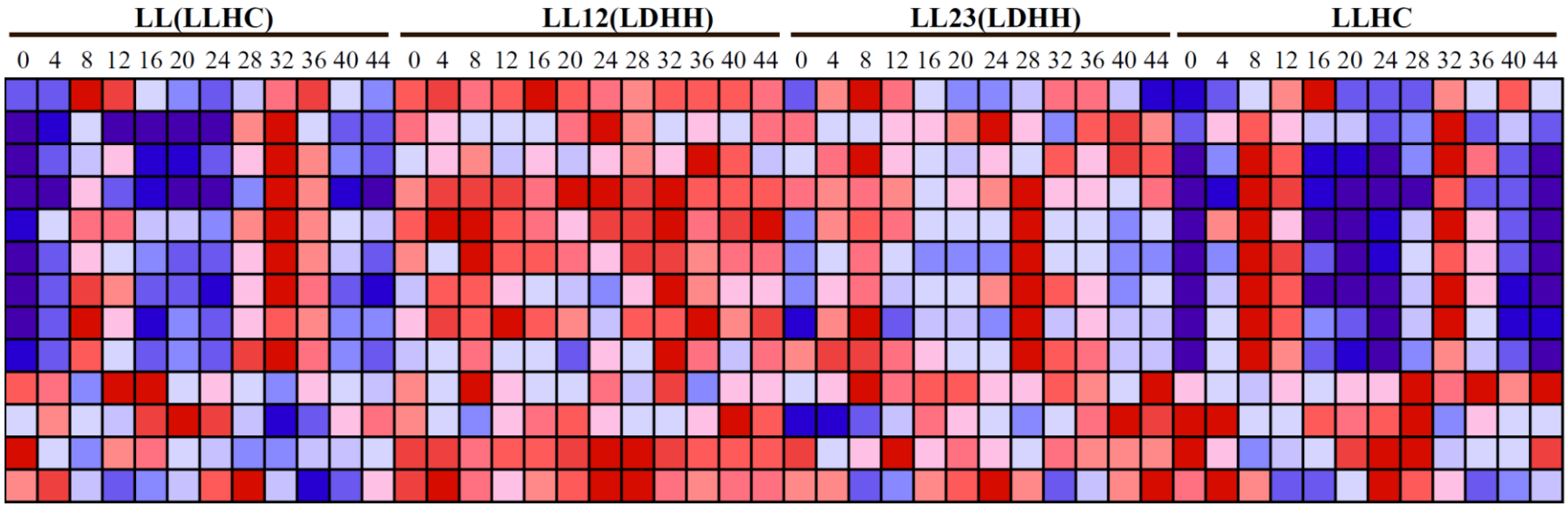
Figure 2.3 Heat map of known and proposed BCAA catabolism gene expression profiles under diurnal/circadian conditions. Microarray data were obtained from the Diurnal database (Mockler et al., 2007), and normalized to the maximum expression levels of each gene in every treatment. Pearson's Correlation Coefficient and average linkage were used for gene clustering. Gene names in orange text represent members of the highly coexpressed module identified from the diurnal/circadian dataset. Refer to the Diurnal database website (http://diurnal.mocklerlab.org/) for detailed information on each condition. COL, Col-0; LDHH, 12h light/12h dark and 24h hot; SD, 8h light/16h dark; DD(DDHC), entrained on 24h dark and 12h hot/12h cold, and subjected to 24h dark and 24h hot; LDHC, 12h light/12h dark and 12h hot/12h cold; LDHH-Smith, 12h light/12h dark and 24h hot (Smith et al., 2004); LDHH-Stitt, 12h light/12h dark and 24h hot (Bläsing et al., 2005); LL(LDHC), entrained on LDHC and subjected to 24h light and 24h hot; LL(LLHC),

Figure 2.3 (cont'd)



entrained on 24h light and 12h hot/12h cold, and subjected to 24h light and 24h hot; LL12(LDHH), entrained on LDHH and subjected to 24h light and 24h hot; LL23(LDHH), entrained on LDHH and subjected to 24h light and 24h hot; LLHC, 24h light and 12h hot/12h cold.

Figure 2.3 (cont'd)



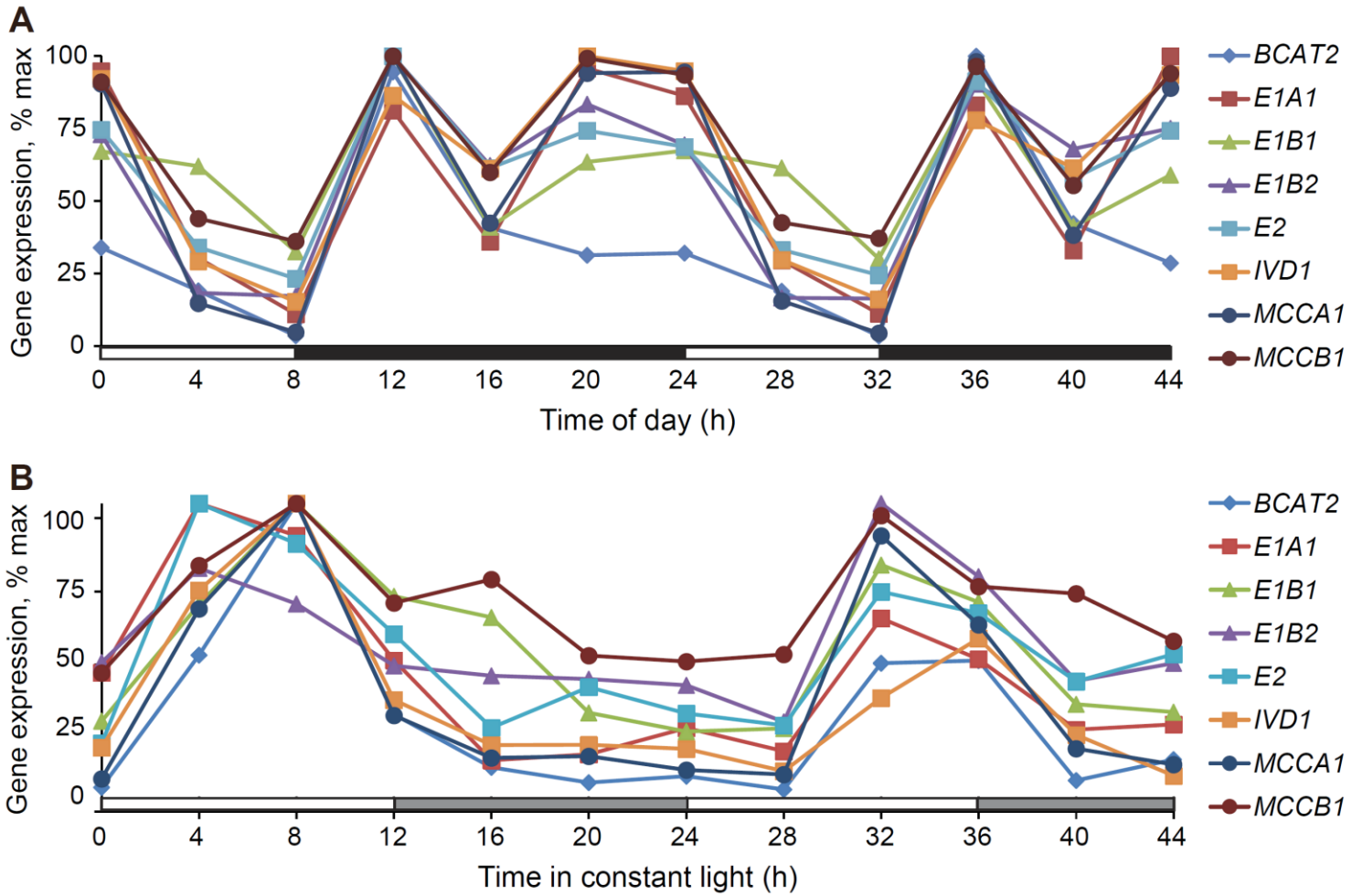


Figure 2.4 Expression profiles of highly coexpressed BCAA catabolism genes under short day (A) and constant light (LL(LDHC), B). Rosette leaves from 4-week-old Col-0 plants were used in microarray experiments from the Diurnal database (http://diurnal.mocklerlab.org) (Mockler et al., 2007). Each gene was normalized to its maximal expression. White bars on the *x* axis

Figure 2.4 (cont'd)

represent the time in the light (A) or subjective day (B), black bars represent the time in the dark, and grey bars represent the time in the subjective night. See Figure 2.3 for details on conditions.

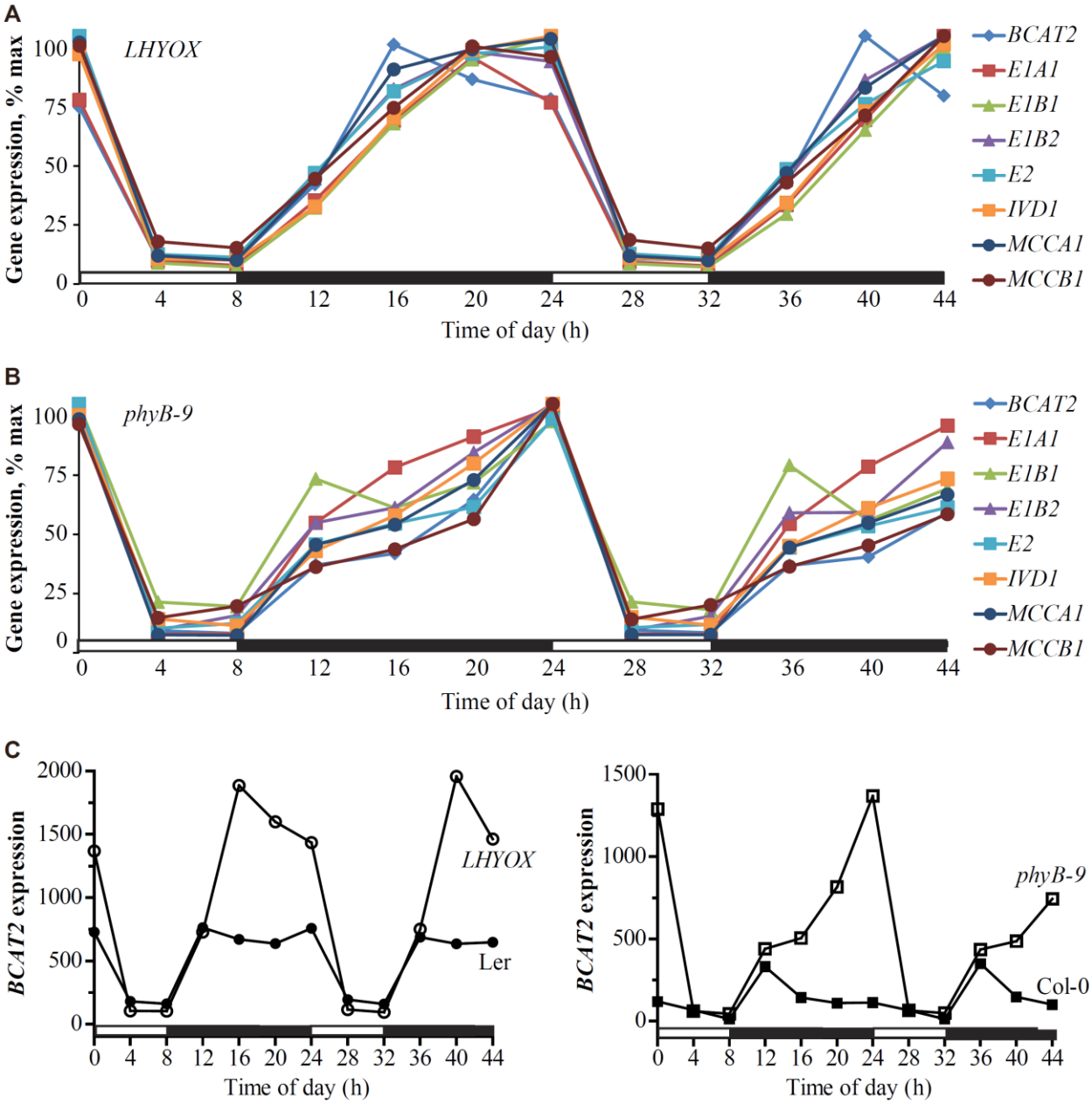


Figure 2.5 Analyses on BCAA catabolism gene transcripts in *A. thaliana* circadian clock and photoreceptor mutants. BCAA catabolism gene transcripts coordinately oscillate in *LHYOX* (A) and *phyB-9* (B) under short day. Rosette leaves from 4-week-old Col-0 plants were used in microarray experiments from the Diurnal database (http://diurnal.mocklerlab.org) (Mockler et al., 2007). Each gene was normalized to its maximal expression. White bars on the *x*

Figure 2.5 (cont'd)

axis represent the time in the light, and black bars represent the time in the dark. (C) Increased BCAA catabolism gene transcripts in *LHYOX* (left) and *phyB-9* (right) compared to corresponding wild type at night during short day. *BCAT2* is shown as an example, and the same changes were also found for *E1A1*, *E1B1*, *E1B2*, *E2*, *IVD1*, *MCCA1* and *MCCB1*. White bars on the *x* axis represent the time in the light, and black bars represent the time in the dark. Microarray data obtained from the Diurnal database (http://diurnal.mocklerlab.org) (Mockler et al., 2007).

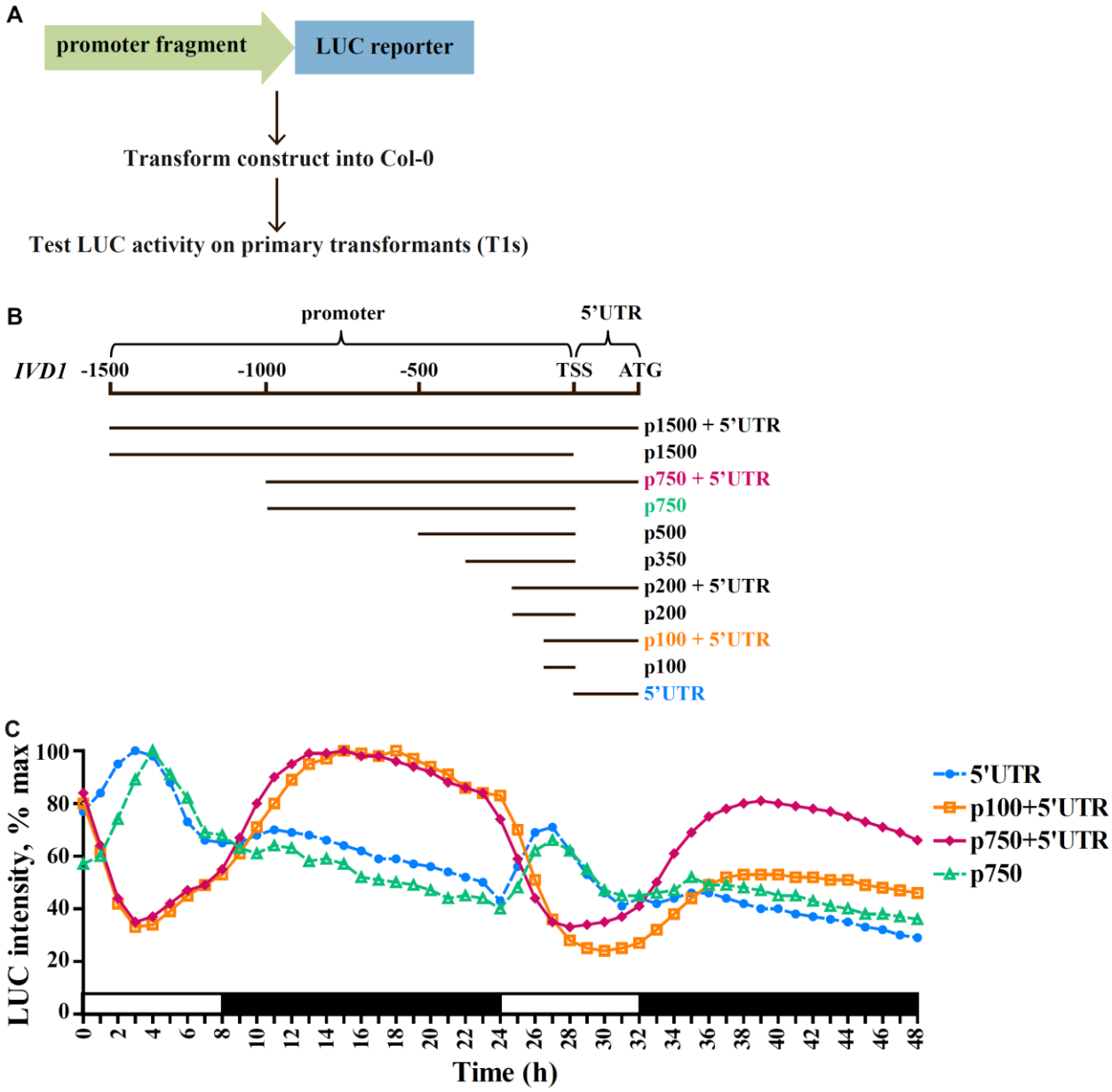


Figure 2.6 *IVD1* promoter mutagenesis experiments. (A) Simplified working flow. Constructs containing luciferase reporter gene (*LUC*) driven by different *IVD1* promoter fragments with or without 5'UTR were transformed into *A. thaliana* wildtype Col-0 plants by agrobacterium-mediated transformation. Primary transformants (T1s) were selected against antibiotics and LUC activity were monitored by luminometer. (B) Fragmentation scheme for *IVD1* promoter with or without 5'UTR. All fragments were monitored for bioluminescence. Fragment name texts

Figure 2.6 (cont'd)

highlighted in colors were shown in C as representatives. (C) LUC intensity for selected *IVD1* promoter/5'UTR fragments. Bioluminescence was recorded every hour for two days under short day (8h light/16h dark). Average intensity of >20 T1s were normalized to the maximum intensity of each fragment. Among the four fragments shown, p100 + 5'UTR and p750 + 5'UTR best mimic the *IVD1* transcript oscillation in Col-0 (shown in Figure 2.4A). TSS, transcription start site; ATG, translation start site.

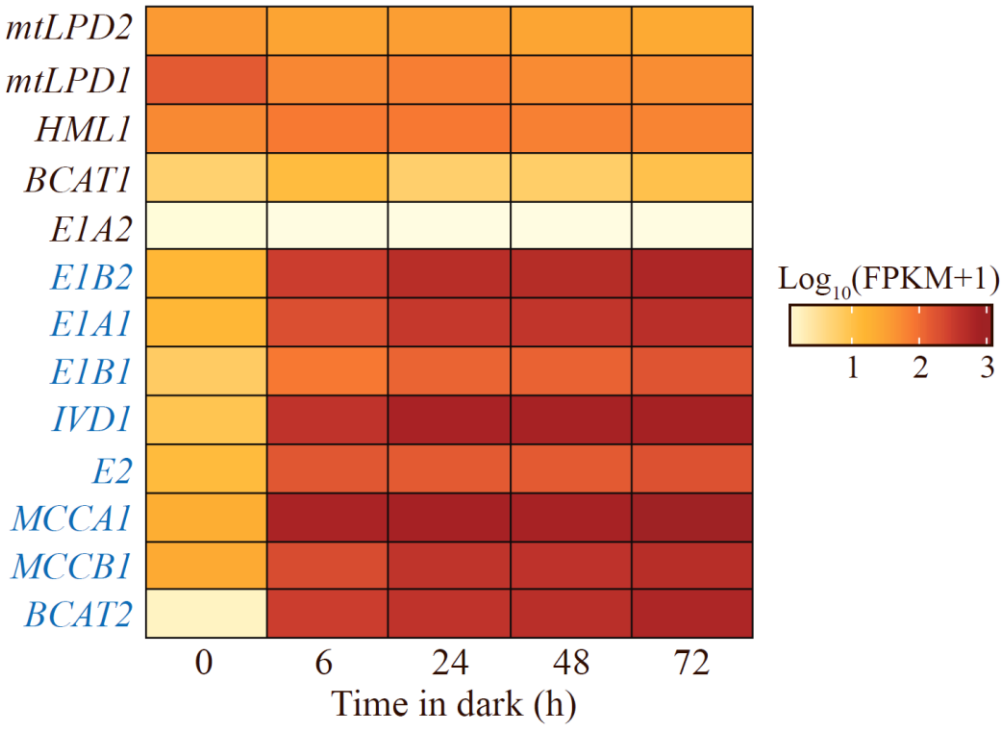


Figure 2.7 Heat map of log_{10} (FPKM+1) transcript levels of known and proposed BCAA catabolism genes in prolonged darkness in Col-0. Values are the mean of two independent biological replicates, which were grown and sequenced separately. FPKM, fragments per kilobase of transcript per million mapped reads. Genes with names in blue text showed statistically significantly increased transcript levels during prolonged darkness compared with time zero ($p<0.05$, determined by exact test implemented in edgeR with Benjamini-Hochberg correction at FDR<0.05) (Chen et al., 2011).

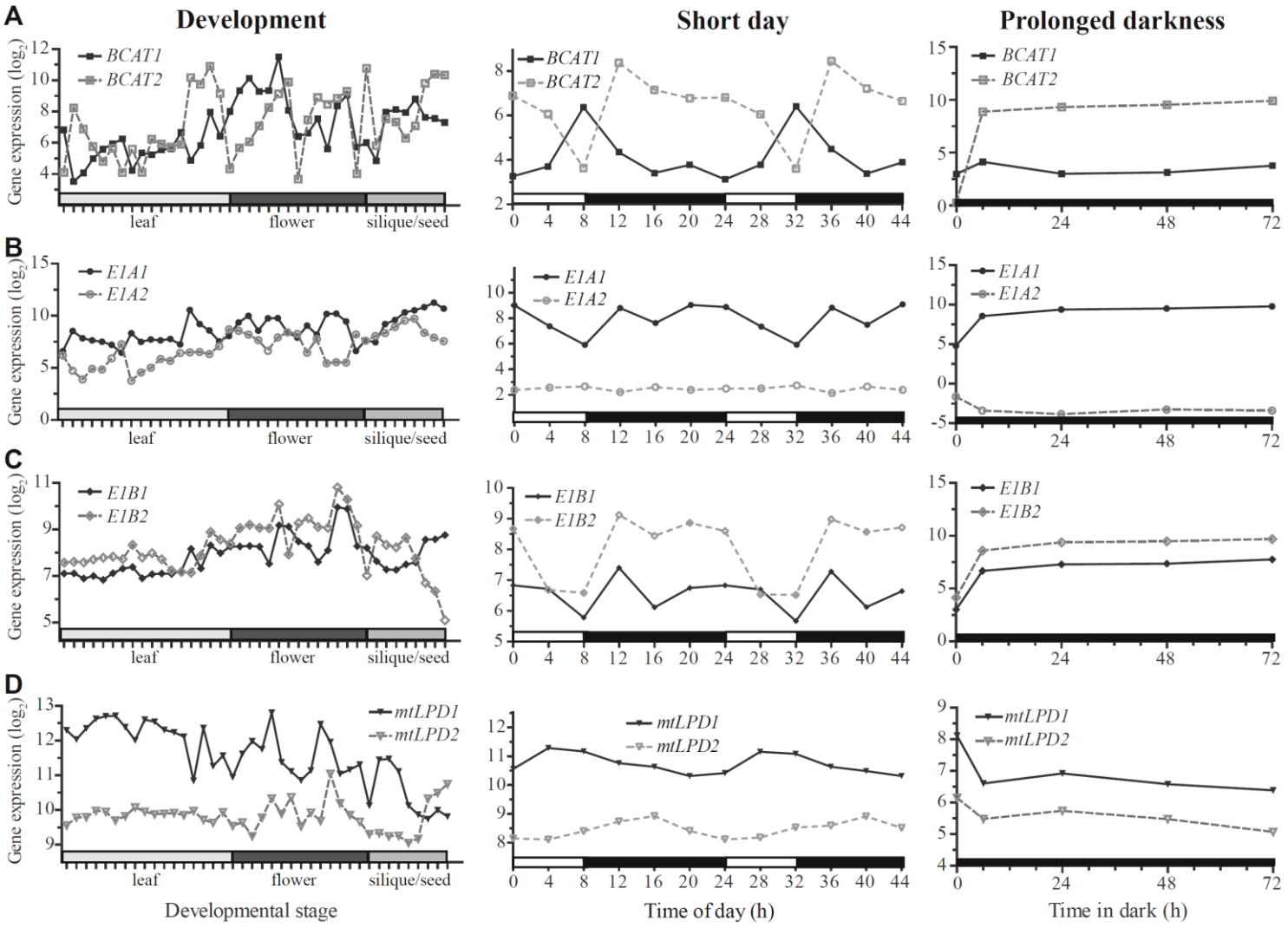


Figure 2.8 Expression analysis of paralogous genes known or proposed to be involved in BCAA catabolism. Microarray data were obtained from (left column) the AtGenExpress development dataset (Schmid et al., 2005) and (middle column) the Diurnal

Figure 2.8 (cont'd)

database (Mockler et al., 2007), respectively; data for prolonged darkness were obtained from RNA-seq experiments (right column). White bars on the *x* axis in short day represent the time in the light, and black bars represent the time in the dark. Gene expression values were converted to $\log_2$ and shown on the *y* axis. Gene pairs indicated in each row are: A, *BCAT1 vs. BCAT2*; B, *E1A1 vs. E1A2*; C, *E1B1 vs. E1B2*; and D, *mtLPD1 vs. mtLPD2*.

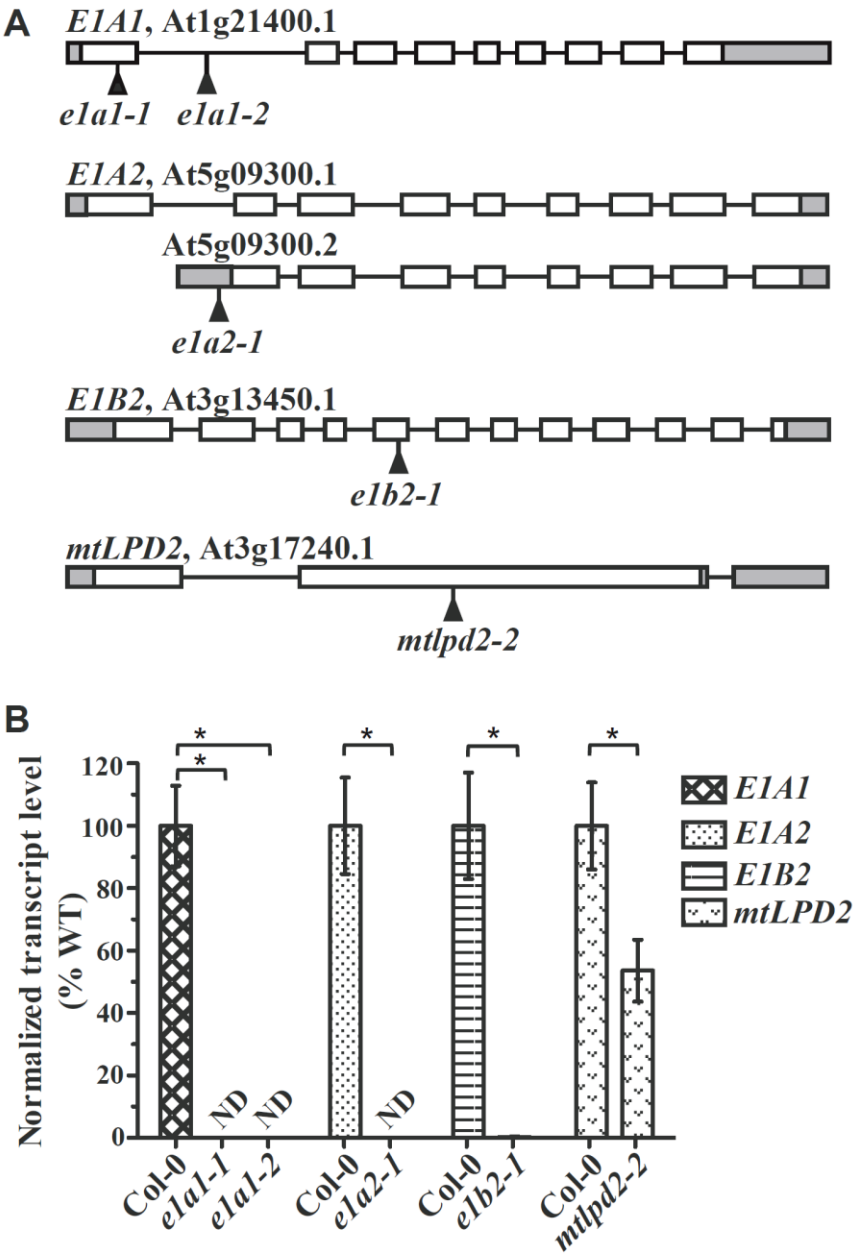


Figure 2.9 Characterization of BCKDH complex subunit mutants. (A) Schematic representations of the T-DNA insertion sites of newly characterized BCAA catabolic mutants. Exons are shown as white rectangles, UTRs as grey rectangles, and introns as solid lines. The sites of T-DNA insertion confirmed in this study are indicated by black triangles. (B) Decreased BCAA catabolic enzyme transcript accumulation in *e1a1-1*, *e1a1-2*, *e1a2-1*, *e1b2-1*, and *mtlpd2-2* mutants. Values represent mean ± SE from four biological replicates. An asterisk indicates a

Figure 2.9 (cont'd)

significant difference determined by the Student's *t*-test ($p<0.05$). ND, not detectable. The normalized *E1B2* transcript level in *e1b2-1* is less than 0.5% relative to Col-0.

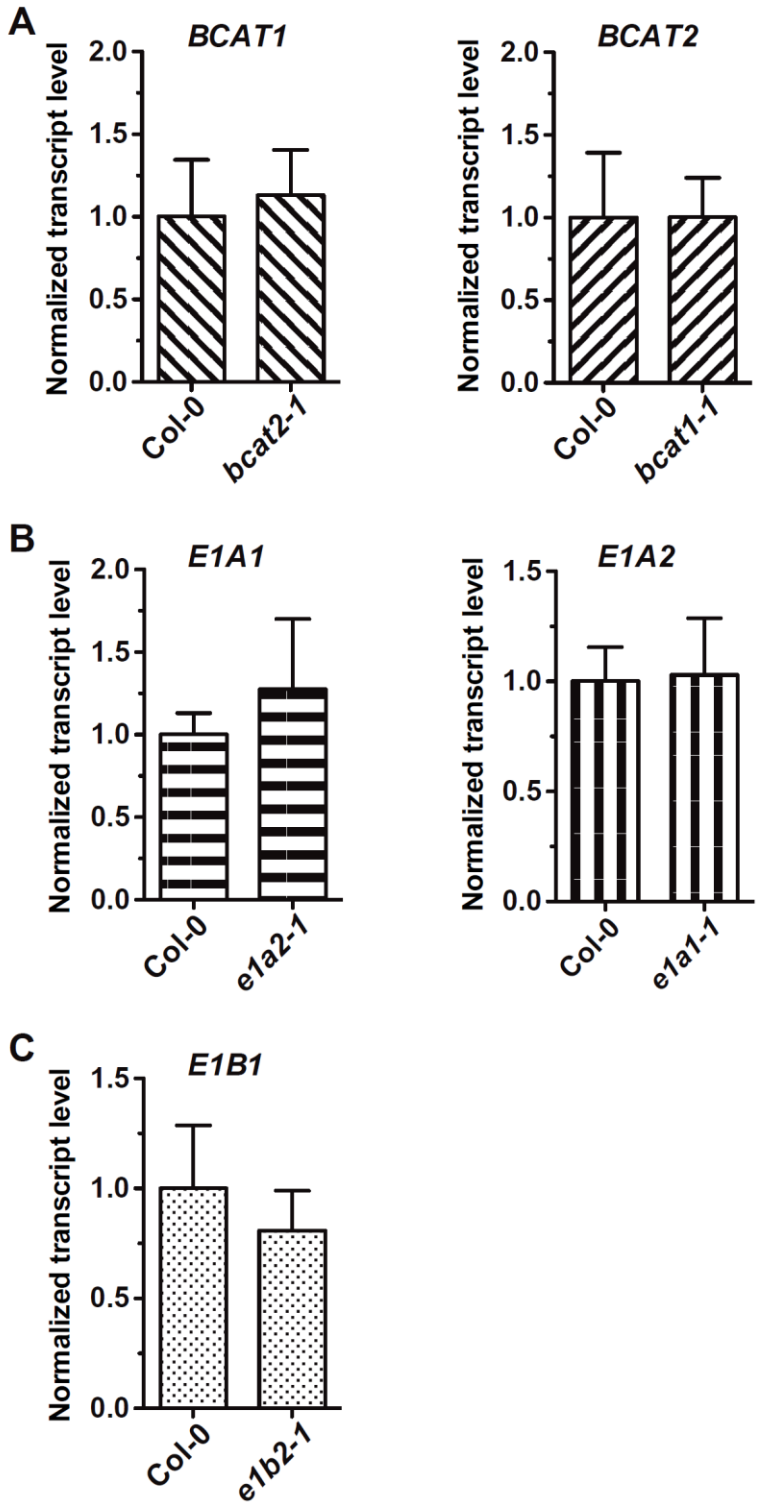


Figure 2.10 Single mutant transcript analysis for compensation between gene paralogs. Values represent mean ± SE from four biological replicates. Transcripts tested are indicated on

Figure 2.10 (cont'd)

top of each figure, and mutants (and corresponding wild type) at the bottom. Transcript levels of tested genes were normalized to *ACT2* transcript levels, and shown as fold changes relative to the WT. Transcripts were measured from rosette leaves of 5-week-old plants treated in prolonged darkness for 3 days.

Table 2.1 List of genes encoding experimentally validated and computationally annotated BCAA catabolic enzymes.

Gene	AGI	Annotation	Paralogous gene	Duplication type*
BCAT1	AT1G10060	Branched-chain aminotransferase 1, putative	*BCAT2*	Tandem
BCAT2	AT1G10070	Branched-chain aminotransferase 2	*BCAT1*	Tandem
E1A1	AT1G21400	α subunit of branched-chain ketoacid dehydrogenase E1, putative	*E1A2*	Before α-WGD
E1A2	AT5G09300	α subunit of branched-chain ketoacid dehydrogenase E1, putative	*E1A1*	Before α-WGD
E1B1	AT1G55510	β subunit of branched-chain ketoacid dehydrogenase E1, putative	*E1B2*	α-WGD
E1B2	AT3G13450	β subunit of branched-chain ketoacid dehydrogenase E1, putative	*E1B1*	α-WGD
E2	AT3G06850	Branched-chain ketoacid dehydrogenase E2, putative	None	/
mtLPD1	AT1G48030	Branched-chain ketoacid dehydrogenase E3, putative	*mtLPD2*	α-WGD
mtLPD2	AT3G17240	Branched-chain ketoacid dehydrogenase E3, putative	*mtLPD1*	α-WGD
IVD1	AT3G45300	Isovaleryl-CoA dehydrogenase	None	/
MCCA1	AT1G03090	α subunit of 3-methylcrotonyl-CoA carboxylase	None	/
MCCB1	AT4G34030	β subunit of 3-methylcrotonyl-CoA carboxylase	None	/
HML1	AT2G26800	Hydroxymethylglutaryl-CoA lyase	None	/

* Duplication type was derived from the physical location of the gene paralogs on chromosomes and published data (Bowers et al., 2003). α-WGD: most recent whole genome duplication event.

Table 2.2 Pairwise Pearson's Correlation Coefficients (PCCs) for transcripts of BCAA catabolism genes*, **.

Development

95th percentile of random PCC distribution = 0.60

	BCAT1	BCAT2	E1A1	E1A2	E1B1	E1B2	E2	mtLPD1	mtLPD2	IVD1	MCCA1	MCCB1	HML1
BCAT1	1.000	0.031	0.586	0.576	0.454	0.379	0.433	-0.049	-0.426	-0.124	-0.123	-0.034	-0.006
BCAT2	0.031	1.000	0.550	-0.173	0.356	0.177	0.138	-0.280	0.220	0.704	0.580	0.044	0.666
E1A1	0.586	0.550	1.000	0.196	0.436	0.315	0.359	-0.216	-0.216	0.387	0.241	0.113	0.251
E1A2	0.576	-0.173	0.196	1.000	0.370	-0.049	0.408	-0.516	-0.335	-0.241	-0.325	-0.133	-0.289
E1B1	0.454	0.356	0.436	0.370	1.000	0.410	0.673	-0.544	0.008	0.443	0.411	0.140	0.163
E1B2	0.379	0.177	0.315	-0.049	0.410	1.000	0.600	0.191	-0.261	0.440	0.511	0.510	0.307
E2	0.433	0.138	0.359	0.408	0.673	0.600	1.000	-0.420	-0.058	0.407	0.458	0.330	-0.029
mtLPD1	-0.049	-0.280	-0.216	-0.516	-0.544	0.191	-0.420	1.000	-0.105	-0.282	-0.075	0.206	0.018
mtLPD2	-0.426	0.220	-0.216	-0.335	0.008	-0.261	-0.058	-0.105	1.000	0.301	0.201	-0.257	0.084
IVD1	-0.124	0.704	0.387	-0.241	0.443	0.440	0.407	-0.282	0.301	1.000	0.706	0.251	0.599
MCCA1	-0.123	0.580	0.241	-0.325	0.411	0.511	0.458	-0.075	0.201	0.706	1.000	0.548	0.322
MCCB1	-0.034	0.044	0.113	-0.133	0.140	0.510	0.330	0.206	-0.257	0.251	0.548	1.000	-0.041
HML1	-0.006	0.666	0.251	-0.289	0.163	0.307	-0.029	0.018	0.084	0.599	0.322	-0.041	1.000

* PCCs above the 95th percentile threshold for each dataset are highlighted in red, and the values are indicated below.

** Genes that have PCCs above threshold with one or more other genes are highlighted in yellow.

Table 2.2 (cont'd)

Stress

95th percentile of random PCC distribution = 0.47

	BCAT1	BCAT2	E1A1	E1A2	E1B1	E1B2	E2	mtLPD1	mtLPD2	IVD1	MCCA1	MCCB1	HML1
BCAT1	1.000	-0.133	0.005	-0.027	0.120	0.134	-0.417	0.102	-0.284	0.011	0.094	0.081	-0.385
BCAT2	-0.133	1.000	0.777	-0.064	0.425	0.487	0.631	-0.538	0.607	0.767	0.608	0.125	0.275
E1A1	0.005	0.777	1.000	-0.149	0.545	0.497	0.543	-0.462	0.477	0.839	0.622	0.282	0.464
E1A2	-0.027	-0.064	-0.149	1.000	0.173	0.110	0.200	0.293	-0.032	-0.142	-0.113	0.236	-0.270
E1B1	0.120	0.425	0.545	0.173	1.000	0.825	0.535	-0.298	0.009	0.668	0.701	0.653	0.288
E1B2	0.134	0.487	0.497	0.110	0.825	1.000	0.685	-0.287	0.022	0.669	0.787	0.656	0.239
E2	-0.417	0.631	0.543	0.200	0.535	0.685	1.000	-0.250	0.357	0.605	0.634	0.504	0.480
mtLPD1	0.102	-0.538	-0.462	0.293	-0.298	-0.287	-0.250	1.000	-0.478	-0.432	-0.340	0.300	-0.249
mtLPD2	-0.284	0.607	0.477	-0.032	0.009	0.022	0.357	-0.478	1.000	0.393	0.113	-0.330	0.202
IVD1	0.011	0.767	0.839	-0.142	0.668	0.669	0.605	-0.432	0.393	1.000	0.742	0.449	0.380
MCCA1	0.094	0.608	0.622	-0.113	0.701	0.787	0.634	-0.340	0.113	0.742	1.000	0.524	0.328
MCCB1	0.081	0.125	0.282	0.236	0.653	0.656	0.504	0.300	-0.330	0.449	0.524	1.000	0.125
HML1	-0.385	0.275	0.464	-0.270	0.288	0.239	0.480	-0.249	0.202	0.380	0.328	0.125	1.000

Table 2.2 (cont'd)

Diurnal/circadian

95th percentile of random PCC distribution = 0.45

	BCAT1	BCAT2	E1A1	E1A2	E1B1	E1B2	E2	mtLPD1	mtLPD2	IVD1	MCCA1	MCCB1	HML1
BCAT1	1.000	0.195	-0.017	-0.109	0.287	0.155	0.186	0.059	-0.392	-0.002	0.227	0.298	-0.278
BCAT2	0.195	1.000	0.497	-0.222	0.515	0.670	0.633	-0.152	-0.334	0.506	0.816	0.565	-0.092
E1A1	-0.017	0.497	1.000	-0.363	0.629	0.740	0.533	-0.286	-0.409	0.750	0.644	0.511	-0.068
E1A2	-0.109	-0.222	-0.363	1.000	-0.240	-0.284	0.031	-0.070	0.379	-0.115	-0.190	-0.090	-0.155
E1B1	0.287	0.515	0.629	-0.240	1.000	0.756	0.745	-0.173	-0.487	0.635	0.746	0.811	-0.005
E1B2	0.155	0.670	0.740	-0.284	0.756	1.000	0.797	-0.354	-0.399	0.794	0.874	0.717	-0.039
E2	0.186	0.633	0.533	0.031	0.745	0.797	1.000	-0.435	-0.320	0.816	0.862	0.810	-0.046
mtLPD1	0.059	-0.152	-0.286	-0.070	-0.173	-0.354	-0.435	1.000	-0.107	-0.627	-0.320	-0.110	0.204
mtLPD2	-0.392	-0.334	-0.409	0.379	-0.487	-0.399	-0.320	-0.107	1.000	-0.239	-0.385	-0.482	-0.050
IVD1	-0.002	0.506	0.750	-0.115	0.635	0.794	0.816	-0.627	-0.239	1.000	0.748	0.591	-0.161
MCCA1	0.227	0.816	0.644	-0.190	0.746	0.874	0.862	-0.320	-0.385	0.748	1.000	0.820	-0.006
MCCB1	0.298	0.565	0.511	-0.090	0.811	0.717	0.810	-0.110	-0.482	0.591	0.820	1.000	0.071
HML1	-0.278	-0.092	-0.068	-0.155	-0.005	-0.039	-0.046	0.204	-0.050	-0.161	-0.006	0.071	1.000

Table 2.2 (cont'd)

Light

95th percentile of random PCC distribution = 0.90

	BCAT1	BCAT2	E1A1	E1A2	E1B1	E1B2	E2	mtLPD1	mtLPD2	IVD1	MCCA1	MCCB1	HML1
BCAT1	1.000	0.622	0.589	-0.146	0.612	0.618	0.510	-0.388	-0.398	0.520	0.589	0.555	-0.210
BCAT2	0.622	1.000	0.972	0.101	0.894	0.983	0.843	-0.762	-0.328	0.785	0.931	0.958	-0.752
E1A1	0.589	0.972	1.000	0.134	0.927	0.952	0.855	-0.667	-0.349	0.816	0.908	0.958	-0.722
E1A2	-0.146	0.101	0.134	1.000	-0.021	0.072	-0.172	0.292	0.621	-0.314	-0.144	0.017	-0.462
E1B1	0.612	0.894	0.927	-0.021	1.000	0.908	0.935	-0.630	-0.544	0.912	0.906	0.923	-0.567
E1B2	0.618	0.983	0.952	0.072	0.908	1.000	0.885	-0.780	-0.397	0.818	0.958	0.969	-0.738
E2	0.510	0.843	0.855	-0.172	0.935	0.885	1.000	-0.710	-0.680	0.964	0.941	0.927	-0.587
mtLPD1	-0.388	-0.762	-0.667	0.292	-0.630	-0.780	-0.710	1.000	0.328	-0.664	-0.796	-0.738	0.472
mtLPD2	-0.398	-0.328	-0.349	0.621	-0.544	-0.397	-0.680	0.328	1.000	-0.776	-0.586	-0.517	0.068
IVD1	0.520	0.785	0.816	-0.314	0.912	0.818	0.964	-0.664	-0.776	1.000	0.907	0.886	-0.452
MCCA1	0.589	0.931	0.908	-0.144	0.906	0.958	0.941	-0.796	-0.586	0.907	1.000	0.973	-0.645
MCCB1	0.555	0.958	0.958	0.017	0.923	0.969	0.927	-0.738	-0.517	0.886	0.973	1.000	-0.722
HML1	-0.210	-0.752	-0.722	-0.462	-0.567	-0.738	-0.587	0.472	0.068	-0.452	-0.645	-0.722	1.000

Table 2.3 Pairwise Pearson's Correlation Coefficients (PCCs) among eight BCAA catabolism genes with elevated transcript levels in prolonged darkness.

95th percentile of random PCC distribution = 0.854

	BCAT2	E1A1	E1B1	E1B2	E2	IVD1	MCCA1	MCCB1
BCAT2	1.000	0.983	0.980	0.967	0.942	0.963	0.959	0.974
E1A1	0.983	1.000	0.985	0.976	0.922	0.975	0.957	0.985
E1B1	0.980	0.985	1.000	0.993	0.933	0.986	0.972	0.988
E1B2	0.967	0.976	0.993	1.000	0.918	0.995	0.964	0.991
E2	0.942	0.922	0.933	0.918	1.000	0.936	0.986	0.919
IVD1	0.963	0.975	0.986	0.995	0.936	1.000	0.971	0.989
MCCA1	0.959	0.957	0.972	0.964	0.986	0.971	1.000	0.960
MCCB1	0.974	0.985	0.988	0.991	0.919	0.989	0.960	1.000

Table 2.4 Gene ontology enrichment analysis by BiNGO*.

GO-ID	Description	p-value	corr p-value	count
9056	catabolic process	1.46E-12	3.40E-11	102
9628	response to abiotic stimulus	9.74E-07	8.24E-06	132
7275	multicellular organismal development	1.38E-04	7.14E-04	163
9605	response to external stimulus	2.05E-04	9.53E-04	33
9791	post-embryonic development	7.29E-04	3.08E-03	92
6464	protein modification process	1.41E-03	5.45E-03	121
6139	nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	1.58E-03	5.87E-03	109
3	reproduction	1.65E-03	5.92E-03	94
9987	cellular process	5.57E-03	1.73E-02	601
9719	response to endogenous stimulus	6.42E-03	1.93E-02	82
9653	anatomical structure morphogenesis	7.14E-03	2.08E-02	54
16043	cellular component organization	7.55E-03	2.13E-02	90
9606	tropism	8.25E-03	2.26E-02	9
9790	embryonic development	1.30E-02	3.46E-02	45
6950	response to stress	1.43E-02	3.68E-02	163
6810	transport	1.65E-02	4.15E-02	134
6629	lipid metabolic process	1.88E-02	4.61E-02	57

* 2084 genes that coexpressed with at least two out of the eight BCAA catabolism genes - *BCAT2*, *E1A1*, *E1B1*, *E1B2*, *E2*, *IVD1*, *MCCA1*, and *MCCB1* - were used in this analysis with the following parameters:

Selected ontology file: GOSlim-Plants, with focus on biological process

Selected statistical test: hypergeometric test

Selected correction: Benjamini & Hochberg False Discovery Rate correction

Selected significance level: 0.05

Testing option: use whole annotation as reference set

Table 2.5 Functional annotation clustering by DAVID*.

Annotation Cluster 1

Enrichment Score: 9.09

GO-ID	Description	Count	p-value	corr p-value	FDR
16071	mRNA metabolic process	49	2.30E-17	3.73E-14	3.85E-14
6397	mRNA processing	42	1.25E-14	1.02E-11	2.10E-11
8380	RNA splicing	36	3.16E-13	1.71E-10	5.30E-10
16070	RNA metabolic process	95	4.73E-09	3.84E-07	7.93E-06
6396	RNA processing	68	6.65E-08	4.32E-06	1.12E-04
398	nuclear mRNA splicing, via spliceosome	15	6.69E-06	2.79E-04	1.12E-02
375	RNA splicing, via transesterification reactions	15	3.16E-05	1.05E-03	5.30E-02
377	RNA splicing, via transesterification reactions with bulged adenosine as nucleophile	15	3.16E-05	1.05E-03	5.30E-02

Annotation Cluster 2

Enrichment Score: 8.31

GO-ID	Description	Count	p-value	corr p-value	FDR
44248	cellular catabolic process	135	1.89E-12	7.68E-10	3.17E-09
44265	cellular macromolecule catabolic process	94	3.83E-11	1.25E-08	6.43E-08
51603	proteolysis involved in cellular protein catabolic process	87	5.64E-10	9.16E-08	9.45E-07
44257	cellular protein catabolic process	87	9.71E-10	1.43E-07	1.63E-06
30163	protein catabolic process	88	1.62E-09	2.20E-07	2.73E-06
19941	modification-dependent protein catabolic process	85	1.84E-09	2.30E-07	3.09E-06
43632	modification-dependent macromolecule catabolic process	85	1.84E-09	2.30E-07	3.09E-06
9057	macromolecule catabolic process	97	9.50E-09	7.35E-07	1.59E-05
9056	catabolic process	157	2.74E-08	2.02E-06	4.60E-05
6508	proteolysis	123	2.44E-06	1.16E-04	4.09E-03
6511	ubiquitin-dependent protein catabolic process	40	2.87E-04	7.48E-03	4.80E-01

Table 2.5 (cont'd)

Annotation Cluster 3

Enrichment Score: 5.76

GO-ID	Description	Count	p-value	corr p-value	FDR
50789	regulation of biological process	386	6.86E-11	1.86E-08	1.15E-07
19222	regulation of metabolic process	268	8.37E-11	1.94E-08	1.40E-07
65007	biological regulation	420	9.08E-11	1.84E-08	1.52E-07
80090	regulation of primary metabolic process	249	1.58E-10	2.85E-08	2.64E-07
60255	regulation of macromolecule metabolic process	252	3.38E-09	3.66E-07	5.66E-06
31323	regulation of cellular metabolic process	243	3.61E-09	3.67E-07	6.06E-06
50794	regulation of cellular process	352	4.54E-09	4.10E-07	7.62E-06
19219	regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	225	4.98E-08	3.52E-06	8.35E-05
51171	regulation of nitrogen compound metabolic process	226	5.02E-08	3.40E-06	8.42E-05
45449	regulation of transcription	221	9.74E-08	6.09E-06	1.63E-04
9889	regulation of biosynthetic process	228	9.79E-08	5.89E-06	1.64E-04
31326	regulation of cellular biosynthetic process	228	9.79E-08	5.89E-06	1.64E-04
10556	regulation of macromolecule biosynthetic process	225	1.11E-07	6.42E-06	1.85E-04
10468	regulation of gene expression	236	1.26E-07	7.04E-06	2.11E-04
6139	nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	260	2.95E-07	1.60E-05	4.94E-04
6807	nitrogen compound metabolic process	305	7.34E-07	3.85E-05	1.23E-03
34641	cellular nitrogen compound metabolic process	295	1.50E-06	7.41E-05	2.52E-03
51252	regulation of RNA metabolic process	126	8.68E-06	3.52E-04	1.46E-02
6355	regulation of transcription, DNA-dependent	123	2.60E-05	8.80E-04	4.36E-02
6350	transcription	136	4.64E-04	1.12E-02	7.76E-01
44260	cellular macromolecule metabolic process	493	5.27E-04	1.25E-02	8.81E-01
10467	gene expression	226	7.22E-01	9.91E-01	1.00E+02
34645	cellular macromolecule biosynthetic process	174	1.00E+00	1.00E+00	1.00E+02
9059	macromolecule biosynthetic process	174	1.00E+00	1.00E+00	1.00E+02
9058	biosynthetic process	267	1.00E+00	1.00E+00	1.00E+02

Table 2.5 (cont'd)

44249	cellular biosynthetic process	247	1.00E+00	1.00E+00	1.00E+02

Annotation Cluster 4

Enrichment Score: 4.85

GO-ID	Description	Count	p-value	corr p-value	FDR
46395	carboxylic acid catabolic process	26	2.24E-09	2.60E-07	3.76E-06
16054	organic acid catabolic process	26	2.24E-09	2.60E-07	3.76E-06
9062	fatty acid catabolic process	13	3.79E-06	1.71E-04	6.36E-03
6635	fatty acid beta-oxidation	12	4.01E-06	1.76E-04	6.72E-03
19395	fatty acid oxidation	12	1.27E-05	4.79E-04	2.13E-02
34440	lipid oxidation	12	1.27E-05	4.79E-04	2.13E-02
30258	lipid modification	14	1.54E-05	5.70E-04	2.59E-02
44242	cellular lipid catabolic process	15	9.50E-05	2.71E-03	1.59E-01
6631	fatty acid metabolic process	24	7.52E-02	4.37E-01	7.31E+01
16042	lipid catabolic process	20	2.54E-01	7.89E-01	9.93E+01

Annotation Cluster 5

Enrichment Score: 4.79

GO-ID	Description	Count	p-value	corr p-value	FDR
8104	protein localization	79	4.29E-09	4.10E-07	7.20E-06
15031	protein transport	77	4.70E-09	4.02E-07	7.88E-06
45184	establishment of protein localization	77	4.70E-09	4.02E-07	7.88E-06
16192	vesicle-mediated transport	48	1.35E-06	6.88E-05	2.27E-03
33036	macromolecule localization	89	6.15E-06	2.63E-04	1.03E-02
46907	intracellular transport	57	1.16E-05	4.47E-04	1.94E-02
51641	cellular localization	67	1.59E-05	5.72E-04	2.66E-02
51649	establishment of localization in cell	62	2.50E-05	8.63E-04	4.19E-02
6886	intracellular protein transport	42	4.38E-05	1.39E-03	7.34E-02
70727	cellular macromolecule localization	45	4.87E-05	1.46E-03	8.17E-02
34613	cellular protein localization	43	5.10E-05	1.51E-03	8.55E-02
6810	transport	182	5.64E-02	3.76E-01	6.22E+01

Table 2.5 (cont'd)

51234	establishment of localization	182	6.28E-02	3.97E-01	6.63E+01
51179	localization	186	7.18E-02	4.31E-01	7.13E+01

* 2084 genes that coexpressed with at least two out of the eight BCAA catabolism genes - *BCAT2*, *E1A1*, *E1B1*, *E1B2*, *E2*, *IVD1*, *MCCA1*, and *MCCB1* - were used in GO enrichment analysis focusing on biological processes related GO terms.

Selected ontology file: GOTERM_BP_ALL

Selected correction: Benjamini-Hochberg method

285 GO terms were enriched, 94 clusters of similar GO terms were generated by DAVID functional annotation clustering, and the top five clusters were shown in this table.

Table 2.6 Primers for genotyping, qPCR and *IVD1* promoter mutagenesis experiments.

Primer	Sequence	Note
LP_SALK_138630	TGAACCTGTATGTGGAGGAGG	Genotyping primer for *bcat1-1*
RP_SALK_138630	TTCAAAAGCTTTTGATGGGTG	Genotyping primer for *bcat1-1*
LP SALK_037854	CAAATTCAACGATTTGCCAAG	Genotyping primer for *bcat2-1*
RP SALK_037854	TTTTACCCAACGTTTGTTTGC	Genotyping primer for *bcat2-1*
LP_SALK_071680C	ACCTTTACCATGACTTGTGCG	Genotyping primer for *e1a1-1*
RP_SALK_071680C	AGTTGGAGATGGATACGGATG	Genotyping primer for *e1a1-1*
LP SALK_098054	GATGTTGGATTTGGTGGTGTC	Genotyping primer for *e1b2-1*
RP SALK_098054	TGGAACCTATATACCTCTGCCTC	Genotyping primer for *e1b2-1*
LP SALK_027039	TTGTTCTCGTTGCATATGCTG	Genotyping primer for *mtlpd2-2*
RP SALK_027039	CATCTTCTTCGGCTTTGTGAG	Genotyping primer for *mtlpd2-2*
LP SALK_137966	AATATCTTGCTCATGGCCATG	Genotyping primer for *mcca1-1*
RP SALK_137966	TGCAGCCTTTCTTAATGCTTC	Genotyping primer for *mcca1-1*
LP SALK_117349	CATATTTTTAGCAGGACCGCC	Genotyping primer for *mccb1-1*
RP SALK_117349	AGCACAGGATACTGCCATCAC	Genotyping primer for *mccb1-1*
LP SALK_145226	TTCCTTTGCACCTGCAGATAC	Genotyping primer for *hml1-2*
RP SALK_145226	GAAGTTGGTCCAAGAGATGGC	Genotyping primer for *hml1-2*
LBa1	TGGTTCACGTAGTGGGCCATCG	Universal left genotyping primer for SALK lines
P1 At1g21400	GGTTTGCTAGATCCAAAACCC	Genotyping primer for *e1a1-2*
P2 At1g21400	AGAACCCGGTAACATGGAATC	Genotyping primer for *e1a1-2*
p745	AACGTCCGCAATGTGTTATTAAGTTGTC	Universal left genotyping primer for WiscDsLox lines
LP_SAIL_113_D07	TTTTTACAGACGAAGGCCTTG	Genotyping primer for *e1a2-1*
RP_SAIL_113_D07	CTCTTCACCGATTGCAGTAGC	Genotyping primer for *e1a2-1*
LB3	TAGCATCTGAATTTCATAACCAATCTCGATACAC	Universal left genotyping primer for SAIL lines

Table 2.6 (cont'd)

GABI_756G02-LP	LPAATCTGCAAAGCAACCACAAC	Genotyping primer for *ivd1-2*
GABI_756G02-RP	RPACCTGCAGAGGAATATGGAGG	Genotyping primer for *ivd1-2*
LB-GABI-KAT-o8409	ATATTGACCATCATACTCATTGC	Insertion specific genotyping left primer for *ivd1-2*
AT1G21400L3	CGTATTTGAGTCCCTTCGGTA	qPCR primer for *E1A1*
AT1G21400R3	TTTCATCTCCGATGTGTAACC	qPCR primer for *E1A1*
AT5G09300L2	CACGAATACGCCAACAATCA	qPCR primer for *E1A2*
AT5G09300R2	TCATCAAGAACACGGTAGCA	qPCR primer for *E1A2*
At1g55510L	ATCCTCGGTCTTATGTCTTT	qPCR primer for *E1B1*
At1g55510R	CCAATGCCAAATCCAACAAT	qPCR primer for *E1B1*
At3g13450L	AGGTTCCGACATAACTCTTG	qPCR primer for *E1B2*
At3g13450R	CACTGAGGTCTCAACGATTT	qPCR primer for *E1B2*
LPD2_At3g17240L	CTCGGTGGTACTTGTCTTAA	qPCR primer for *mtLPD2*
LPD2_At3g17240R	CAACCGAAGAGACCTTAACA	qPCR primer for *mtLPD2*
IVD_1500P_F	CACCTGAGGATGATAATGAGAAG	forward primer for amplifying *IVD1*-1500 promoter with CACC at 5'
IVD_750P_F	CACCTGACACATTGTATCGCAT	forward primer for amplifying *IVD1*-750 promoter with CACC at 5'
IVD_500P_F	CACCTCAATGAGTCAATGTTAAAC	forward primer for amplifying *IVD1*-500 promoter with CACC at 5'
IVD_350P_F	CACCTTAACGTCGTTGTACATGAA	forward primer for amplifying *IVD1*-350 promoter with CACC at 5'
IVD_200P_F	CACCTCCTAAACACTAATGTGTTC	forward primer for amplifying *IVD1*-200 promoter with CACC at 5'
IVD_100P_F	CACCATGGTATAATAGAGCAGTGT	forward primer for amplifying *IVD1*-100 promoter with CACC at 5'
IVD_0P_F	CACCGACTCATTGCTCATATCTTC	Forward primer for amplifying *IVD1* 5'UTR with CACC at 5'
IVD_P_R	ATTGGTCCATCTAATCTAGTTCCG	reverse primer for amplifying *IVD1* promoter
IVD_5UTR_R	ATCTTCGTTATTACCGGTAAG	PCR primer for *IVD1* 5'UTR region. Located at 3' end of 5'UTR of *IVD1*

Table 2.7 List of TFs interacting with *IVD1* promoter and 5'UTR by Y1H assay.

Fragment	AGI	Transcription factor family
#1	AT1G04370	AP2-EREBP
#1	AT1G14580	C2H2
#1	AT1G21000	PLATZ
#1	AT1G22190	AP2-EREBP
#1	AT1G25330	bHLH
#1	AT1G46480	HB
#1	AT1G50640	AP2-EREBP
#1	AT1G51140	bHLH
#1	AT1G64620	C2C2-DOF
#1	AT1G66140	C2H2
#1	AT1G66600	WRKY
#1	AT1G68550	AP2-EREBP
#1	AT1G72570	AP2-EREBP
#1	AT1G73730	EIL
#1	AT1G76420	NAC
#1	AT1G76900	TUB
#1	AT1G78080	AP2-EREBP
#1	AT2G02820	MYB
#1	AT2G20880	AP2-EREBP
#1	AT2G29580	C3H
#1	AT2G31230	AP2-EREBP
#1	AT2G31370	bZIP
#1	AT2G35530	bZIP
#1	AT2G37060	CCAAT/CCAAT-HAP3
#1	AT2G38340	AP2-EREBP
#1	AT2G45410	LOB/AS2

Table 2.7 (cont'd)

#1	AT2G45420	LOB/AS2
#1	AT2G46870	ABI3-VP1
#1	AT3G02990	HSF
#1	AT3G04280	Orphans
#1	AT3G07670	SET/PcG
#1	AT3G10030	TRIHELIX
#1	AT3G10470	C2H2
#1	AT3G10500	NAC
#1	AT3G15540	AUX-IAA
#1	AT3G16770	AP2-EREBP
#1	AT3G17609	bZIP
#1	AT3G20310	AP2-EREBP
#1	AT3G21880	C2C2-CO-like
#1	AT3G23030	AUX-IAA
#1	AT3G23220	AP2-EREBP
#1	AT3G30210	MYB
#1	AT3G53200	MYB
#1	AT3G54320	AP2-EREBP
#1	AT3G59060	bHLH
#1	AT3G61250	MYB
#1	AT3G61630	AP2-EREBP
#1	AT3G62100	AUX-IAA
#1	AT4G00270	GeBP
#1	AT4G01720	WRKY
#1	AT4G14770	CPP
#1	AT4G18020	ARR-B/G2-like
#1	AT4G21040	C2C2-DOF
#1	AT4G31550	WRKY

Table 2.7 (cont'd)

#1	AT4G35040	bZIP
#1	AT4G35280	C2H2
#1	AT4G37730	bZIP
#1	AT4G37750	AP2-EREBP
#1	AT4G39100	PHD
#1	AT4G39780	AP2-EREBP
#1	AT5G05410	AP2-EREBP
#1	AT5G10140	MADS
#1	AT5G15840	C2C2-CO-like
#1	AT5G18450	AP2-EREBP
#1	AT5G40220	MADS
#1	AT5G53950	NAC
#1	AT5G56860	C2C2-GATA
#1	AT5G57390	AP2-EREBP
#1	AT5G61590	AP2-EREBP
#1	AT5G65130	AP2-EREBP
#1	AT5G65310	HB
#1	AT5G66730	C2H2
#2	AT1G04370	AP2-EREBP
#2	AT1G08540	SIGMA70-like
#2	AT1G22190	AP2-EREBP
#2	AT1G28160	AP2-EREBP
#2	AT1G47270	TUB
#2	AT1G51140	bHLH
#2	AT1G72010	TCP
#2	AT1G72570	AP2-EREBP
#2	AT1G73730	EIL
#2	AT1G78080	AP2-EREBP

Table 2.7 (cont'd)

#2	AT2G20180	bHLH
#2	AT2G20880	AP2-EREBP
#2	AT2G22540	MADS
#2	AT2G24430	NAC
#2	AT2G27230	bHLH
#2	AT2G31460	REM(B3)
#2	AT2G37430	C2H2
#2	AT2G39030	GNAT
#2	AT3G04280	Orphans
#2	AT3G15500	NAC
#2	AT3G16770	AP2-EREBP
#2	AT3G23220	AP2-EREBP
#2	AT3G23240	AP2-EREBP
#2	AT3G52540	OFP
#2	AT3G57600	AP2-EREBP
#2	AT3G61630	AP2-EREBP
#2	AT3G61740	PHD
#2	AT4G00238	GeBP
#2	AT4G00270	GeBP
#2	AT4G00390	GeBP
#2	AT4G33280	ABI3-VP1
#2	AT4G36900	AP2-EREBP
#2	AT4G37750	AP2-EREBP
#2	AT4G38960	Orphans
#2	AT4G39160	MYB-related
#2	AT5G05410	AP2-EREBP
#2	AT5G06160	C2H2
#2	AT5G18450	AP2-EREBP

Table 2.7 (cont'd)

#2	AT5G20240	MADS
#2	AT5G25790	CPP
#2	AT5G27130	MADS
#2	AT5G47390	MYB-related
#2	AT5G57390	AP2-EREBP
#2	AT5G61590	AP2-EREBP
#2	AT5G65130	AP2-EREBP
#3	AT1G04370	AP2-EREBP
#3	AT1G07640	C2C2-DOF
#3	AT1G12260	NAC
#3	AT1G14580	C2H2
#3	AT1G15580	AUX-IAA
#3	AT1G25580	NAC
#3	AT1G26610	C2H2
#3	AT1G49480	ABI3-VP1
#3	AT1G53160	SBP
#3	AT1G64620	C2C2-DOF
#3	AT1G68480	C2H2
#3	AT1G72570	AP2-EREBP
#3	AT1G75490	AP2-EREBP
#3	AT1G76420	NAC
#3	AT2G23660	LOB/AS2
#3	AT2G47190	MYB
#3	AT3G02990	HSF
#3	AT3G06490	MYB
#3	AT3G10000	TRIHELIX
#3	AT3G10030	TRIHELIX
#3	AT3G10470	C2H2

Table 2.7 (cont'd)

#3	AT3G16770	AP2-EREBP
#3	AT3G23210	bHLH
#3	AT3G27810	MYB
#3	AT3G57920	SBP
#3	AT3G59060	bHLH
#3	AT3G62100	AUX-IAA
#3	AT4G35040	bZIP
#3	AT4G35280	C2H2
#3	AT5G04390	C2H2
#3	AT5G15480	C2H2
#3	AT5G18240	G2-like
#3	AT5G59340	HB
#3	AT5G63790	NAC
#3	AT5G67450	C2H2

LITERATURE CITED

LITERATURE CITED

Angelovici R, Fait A, Fernie AR, Galili G (2011) A seed high-lysine trait is negatively associated with the TCA cycle and slows down Arabidopsis seed germination. New Phytol **189:** 148-159

Angelovici R, Lipka AE, Deason N, Gonzalez-Jorge S, Lin H, Cepela J, Buell R, Gore MA, Dellapenna D (2013) Genome-wide analysis of branched-chain amino acid levels in *Arabidopsis* seeds. Plant Cell **25:** 4827-4843

Araujo WL, Ishizaki K, Nunes-Nesi A, Larson TR, Tohge T, Krahnert I, Witt S, Obata T, Schauer N, Graham IA, Leaver CJ, Fernie AR (2010) Identification of the 2-hydroxyglutarate and isovaleryl-CoA dehydrogenases as alternative electron donors linking lysine catabolism to the electron transport chain of *Arabidopsis* mitochondria. Plant Cell **22:** 1549-1563

Ashmarina LI, Pshezhetsky AV, Branda SS, Isaya G, Mitchell GA (1999) 3-Hydroxy-3-methylglutaryl coenzyme A lyase: targeting and processing in peroxisomes and mitochondria. J Lipid Res **40:** 70-75

Ashmarina LI, Rusnak N, Miziorko HM, Mitchell GA (1994) 3-Hydroxy-3-methylglutaryl-CoA lyase is present in mouse and human liver peroxisomes. J Biol Chem **269:** 31929-31932

Aubert S, Bligny R, Day DA, Whelan J, Douce R (1997) Induction of alternative oxidase synthesis by herbicides inhibiting branched-chain amino acid synthesis. Plant J **11:** 649-657

Binder S (2010) Branched-chain amino acid metabolism in *Arabidopsis thaliana*. Arabidopsis Book **8:** e0137

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics **30:** 2114-2120

Bowers JE, Chapman BA, Rong J, Paterson AH (2003) Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. Nature **422:** 433-438

Brown DM, Zeef LA, Ellis J, Goodacre R, Turner SR (2005) Identification of novel genes in Arabidopsis involved in secondary cell wall formation using expression profiling and reverse genetics. Plant Cell **17:** 2281-2295

Callis J, Fromm M, Walbot V (1987) Introns increase gene expression in cultured maize cells. Genes Dev **1:** 1183-1200

Chen H, Saksa K, Zhao F, Qiu J, Xiong L (2010) Genetic analysis of pathway regulation for enhancing branched-chain amino acid biosynthesis in plants. Plant J **63:** 573-583

Chen Y, McCarthy D, Robinson M, Smyth GK (2011) edgeR: differential expression analysis of digital gene expression data User's Guide. http://www.bioconductor.org/packages/release/bioc/vignettes/edgeR/inst/doc/edgeRUsersGuide.pdf

Clough SJ, Bent AF (1998) Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. Plant J **16:** 735-743

Czechowski T, Stitt M, Altmann T, Udvardi MK, Scheible W-R (2005) Genome-wide identification and testing of superior reference genes for transcript normalization in Arabidopsis. Plant Physiol **139:** 5-17

Daschner K, Couee I, Binder S (2001) The mitochondrial isovaleryl-coenzyme A dehydrogenase of Arabidopsis oxidizes intermediates of leucine and valine catabolism. Plant Physiol **126:** 601-612

De Luca V, Salim V, Atsumi SM, Yu F (2012) Mining the biodiversity of plants: a revolution in the making. Science **336:** 1658-1661

Diebold R, Schuster J, Daschner K, Binder S (2002) The branched-chain amino acid transaminase gene family in Arabidopsis encodes plastid and mitochondrial proteins. Plant Physiol **129:** 540-550

Ding G, Che P, Ilarslan H, Wurtele ES, Nikolau BJ (2012) Genetic dissection of methylcrotonyl CoA carboxylase indicates a complex role for mitochondrial leucine catabolism during seed development and germination. Plant J **70:** 562-577

Duarte JM, Cui L, Wall PK, Zhang Q, Zhang X, Leebens-Mack J, Ma H, Altman N, dePamphilis CW (2006) Expression pattern shifts following duplication indicative of subfunctionalization and neofunctionalization in regulatory genes of *Arabidopsis*. Mol Biol Evol **23:** 469-478

Espinoza C, Degenkolbe T, Caldana C, Zuther E, Leisse A, Willmitzer L, Hincha DK, Hannah MA (2010) Interaction with diurnal and circadian regulation results in dynamic metabolic and transcriptional changes during cold acclimation in Arabidopsis. PLoS One **5:** e14101

Fujiki Y, Ito M, Nishida I, Watanabe A (2001) Leucine and its keto acid enhance the coordinated expression of genes for branched-chain amino acid catabolism in *Arabidopsis* under sugar starvation. FEBS Lett **499:** 161-165

Fujiki Y, Sato T, Ito M, Watanabe A (2000) Isolation and characterization of cDNA clones for the E1β and E2 subunits of the branched-chain α-ketoacid dehydrogenase complex in *Arabidopsis*. J Biol Chem **275:** 6007-6013

Ganko EW, Meyers BC, Vision TJ (2007) Divergence in expression between duplicated genes in Arabidopsis. Mol Biol Evol **24:** 2298-2309

Gendron JM, Pruneda-Paz JL, Doherty CJ, Gross AM, Kang SE, Kay SA (2012) *Arabidopsis* circadian clock protein, TOC1, is a DNA-binding transcription factor. Proc Natl Acad Sci U S A **109:** 3167-3172

Gibon Y, Usadel B, Blaesing OE, Kamlage B, Hoehne M, Trethewey R, Stitt M (2006) Integration of metabolite with transcript and enzyme activity profiling during diurnal cycles in *Arabidopsis* rosettes. Genome Biol **7:** R76

Goff LA, Trapnell C, Kelley D (2012) CummeRbund: visualization and exploration of Cufflinks high-throughput sequencing data. R Package Version 2.2.0. http://www.bioconductor.org/packages/release/bioc/vignettes/cummeRbund/inst/doc/cummeRbund-manual.pdf

Gu LP, Jones AD, Last RL (2010) Broad connections in the Arabidopsis seed metabolic network revealed by metabolite profiling of an amino acid catabolism mutant. Plant J **61:** 579-590

Gu Z, Steinmetz LM, Gu X, Scharfe C, Davis RW, Li W-H (2003) Role of duplicate genes in genetic robustness against null mutations. Nature **421:** 63-66

Hanada K, Sawada Y, Kuromori T, Klausnitzer R, Saito K, Toyoda T, Shinozaki K, Li WH, Hirai MY (2011) Functional compensation of primary and secondary metabolites by duplicate genes in *Arabidopsis thaliana*. Mol Biol Evol **28:** 377-382

Hanaoka H, Noda T, Shirano Y, Kato T, Hayashi H, Shibata D, Tabata S, Ohsumi Y (2002) Leaf senescence and starvation-induced chlorosis are accelerated by the disruption of an Arabidopsis autophagy gene. Plant Physiol **129:** 1181-1193

Harmer SL, Hogenesch JB, Straume M, Chang H-S, Han B, Zhu T, Wang X, Kreps JA, Kay SA (2000) Orchestrated transcription of key pathways in *Arabidopsis* by the circadian clock. Science **290:** 2110-2113

Harper AE, Miller RH, Block KP (1984) Branched-chain amino acid metabolism. Annu Rev Nutr **4:** 409-454

Hirai MY, Sugiyama K, Sawada Y, Tohge T, Obayashi T, Suzuki A, Araki R, Sakurai N, Suzuki H, Aoki K, Goda H, Nishizawa OI, Shibata D, Saito K (2007) Omics-based identification of *Arabidopsis* Myb transcription factors regulating aliphatic glucosinolate biosynthesis. Proc Natl Acad Sci U S A **104:** 6478-6483

Huang da W, Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc **4**: 44-57

Ishizaki K, Larson TR, Schauer N, Fernie AR, Graham IA, Leaver CJ (2005) The critical role of *Arabidopsis* electron-transfer flavoprotein:ubiquinone oxidoreductase during dark-induced starvation. Plant Cell **17:** 2587-2600

Ishizaki K, Schauer N, Larson TR, Graham IA, Fernie AR, Leaver CJ (2006) The mitochondrial electron transfer flavoprotein complex is essential for survival of Arabidopsis in extended darkness. Plant J **47:** 751-760

Jones E, Oliphant T, Peterson P (2001) SciPy: Open source scientific tools for Python. http://www.scipy.org/

Kafri R, Bar-Even A, Pilpel Y (2005) Transcription control reprogramming in genetic backup circuits. Nat Genet **37:** 295-299

Kami C, Lorrain S, Hornitschek P, Fankhauser C (2010) Chapter two - Light-regulated plant growth and development. *In* CPT Marja, ed, Current Topics in Developmental Biology, Vol Volume 91. Academic Press, pp 29-66

Keith B, Dong XN, Ausubel FM, Fink GR (1991) Differential induction of 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase genes in *Arabidopsis thaliana* by wounding and pathogenic attack. Proc Natl Acad Sci U S A **88:** 8821-8825

Kilian J, Whitehead D, Horak J, Wanke D, Weinl S, Batistic O, D'Angelo C, Bornberg-Bauer E, Kudla J, Harter K (2007) The AtGenExpress global stress expression data set: protocols, evaluation and model data analysis of UV-B light, drought and cold stress responses. Plant J **50:** 347-363

Kotewicz ML, D'Alessio JM, Driftmier KM, Blodgett KP, Gerard GF (1985) Cloning and overexpression of Moloney murine leukemia virus reverse transcriptase in *Escherichia coli*. Gene **35:** 249-258

Less H, Galili G (2008) Principal transcriptional programs regulating plant amino acid metabolism in response to abiotic stresses. Plant Physiol **147:** 316-330

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Genome Project Data P (2009) The sequence alignment/map format and SAMtools. Bioinformatics **25:** 2078-2079

Liu T, Carlsson J, Takeuchi T, Newton L, Farre EM (2013) Direct regulation of abiotic responses by the Arabidopsis circadian clock component PRR7. Plant J **76:** 101-114

Lu Y, Savage LJ, Ajjawi I, Imre KM, Yoder DW, Benning C, Dellapenna D, Ohlrogge JB, Osteryoung KW, Weber AP, Wilkerson CG, Last RL (2008) New connections across pathways and cellular processes: industrialized mutant screening reveals novel associations between diverse phenotypes in Arabidopsis. Plant Physiol **146:** 1482-1500

Lu Y, Savage LJ, Larson MD, Wilkerson CG, Last RL (2011) Chloroplast 2010: a database for large-scale phenotypic screening of Arabidopsis mutants. Plant Physiol **155:** 1589-1600

Luehrsen KR, Walbot V (1991) Intron enhancement of gene-expressionand the splicing efficiency of introns in maize cells. Mol Gen Genet **225:** 81-93

Lutziger I, Oliver DJ (2001) Characterization of two cDNAs encoding mitochondrial lipoamide dehydrogenase from Arabidopsis. Plant Physiol **127:** 615-623

Maere S, Heymans K, Kuiper M (2005) BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. Bioinformatics **21**: 3448-344

McCourt JA, Pang SS, King-Scott J, Guddat LW, Duggleby RG (2006) Herbicide-binding sites revealed in the structure of plant acetohydroxyacid synthase. Proc Natl Acad Sci U S A **103:** 569-573

Mentzen WI, Peng J, Ransom N, Nikolau BJ, Wurtele ES (2008) Articulation of three core metabolic processes in Arabidopsis: fatty acid biosynthesis, leucine catabolism and starch metabolism. BMC Plant Biol **8:** 76

Michael TP, McClung CR (2002) Phase-specific circadian clock regulatory elements in Arabidopsis. Plant Physiol **130:** 627-638

Miyashita Y, Good AG (2008) NAD(H)-dependent glutamate dehydrogenase is essential for the survival of *Arabidopsis thaliana* during dark-induced carbon starvation. J Exp Bot **59:** 667-680

Mockler TC, Michael TP, Priest HD, Shen R, Sullivan CM, Givan SA, McEntee C, Kay SA, Chory J (2007) The DIURNAL project: diurnal and circadian expression profiling, model-based pattern matching, and promoter analysis. Cold Spring Harb Symp Quant Biol **72:** 353-363

Mooney BP, Henzl MT, Miernyk JA, Randall DD (2000) The dihydrolipoyl acyltransferase (BCE2) subunit of the plant branched-chain α-ketoacid dehydrogenase complex forms a 24-mer core with octagonal symmetry. Protein Sci **9:** 1334-1339

Mooney BP, Miernyk JA, Randall DD (2002) The complex fate of α-ketoacids. Annu Rev Plant Biol **53:** 357-375

Mueller LA, Zhang P, Rhee SY (2003) AraCyc: a biochemical pathway database for Arabidopsis. Plant Physiol **132:** 453-460

Norris SR, Meyer SE, Callis J (1993) The intron of *Arabidopsis thaliana* polyubiquitin genes is conserved in location and is a quantitative determinant of chimeric gene expression. Plant Mol Biol **21:** 895-906

Oliver DJ (1994) The glycine decarboxylase complex from plant mitochondria. Annu Rev Plant Physiol Plant Mol Biol **45:** 323-337

Persson S, Wei H, Milne J, Page GP, Somerville CR (2005) Identification of genes required for cellulose synthesis by regression analysis of public microarray data sets. Proc Natl Acad Sci U S A **102:** 8633-8638

Pruneda-Paz JL, Breton G, Nagel DH, Kang SE, Bonaldi K, Doherty CJ, Ravelo S, Galli M, Ecker JR, Kay SA (2014) A genome-scale resource for the functional characterization of *Arabidopsis* transcription factors. Cell Rep **8:** 622-632

Rose AB (2002) Requirements for intron-mediated enhancement of gene expression in *Arabidopsis*. RNA **8:** 1444-1453

Rose AB (2004) The effect of intron location on intron-mediated enhancement of gene expression in *Arabidopsis*. Plant J **40:** 744-751

Rose AB (2008) Intron-mediated regulation of gene expression. Curr Top Microbiol Immunol **326:** 277-290

Rose AB, Last RL (1997) Introns act post-transcriptionally to increase expression of the *Arabidopsis thaliana* tryptophan pathway gene *PAT1*. Plant J **11:** 455-464

Saito K, Hirai MY, Yonekura-Sakakibara K (2008) Decoding genes with coexpression networks and metabolomics - 'majority report by precogs'. Trends Plant Sci **13:** 36-43

Sarkar D, Sarkar MD (2007) The lattice Package. http://cran.r-project.org/web/packages/lattice/lattice.pdf

Schauer N, Semel Y, Roessner U, Gur A, Balbo I, Carrari F, Pleban T, Perez-Melis A, Bruedigam C, Kopka J, Willmitzer L, Zamir D, Fernie AR (2006) Comprehensive metabolic profiling and phenotyping of interspecific introgression lines for tomato improvement. Nat Biotechnol **24:** 447-454

Schertl P, Braun H-P (2014) Respiratory electron transfer pathways in plant mitochondria. Front Plant Sci **5**

Schmid M, Davison TS, Henz SR, Pape UJ, Demar M, Vingron M, Scholkopf B, Weigel D, Lohmann JU (2005) A gene expression map of *Arabidopsis thaliana* development. Nat Genet **37:** 501-506

Schultz TF, Kiyosue T, Yanovsky M, Wada M, Kay SA (2001) A role for LKP2 in the circadian clock of Arabidopsis. Plant Cell **13:** 2659-2670

Schuster J, Binder S (2005) The mitochondrial branched-chain aminotransferase (AtBCAT-1) is capable to initiate degradation of leucine, isoleucine and valine in almost all tissues in *Arabidopsis thaliana*. Plant Mol Biol **57:** 241-254

Shariati SA, De Strooper B (2013) Redundancy and divergence in the amyloid precursor protein family. FEBS Lett **587:** 2036-2045

Singh BK (1999) Biosynthesis of valine, leucine, and isoleucine. *In* BK Singh, ed, Plant Amino Acids: Biochemistry and Biotechnology. Marcel Dekker, New York, pp 227-247

Singh BK, Shaner DL (1995) Biosynthesis of branched chain amino acids: from test tube to field. Plant Cell **7:** 935-944

Statistical Package R (2012) R: A language and environment for statistical computing. *In* R Foundation for Statistical Computing, Vienna, Austria

Stern DL, Orgogozo V (2008) The loci of evolution: How predictable is genetic evolution? Evolution **62:** 2155-2177

Stitt M (2013) Systems-integration of plant metabolism: means, motive and opportunity. Curr Opin Plant Biol **16:** 381-388

Szekely G, Abraham E, Cseplo A, Rigo G, Zsigmond L, Csiszar J, Ayaydin F, Strizhov N, Jasik J, Schmelzer E, Koncz C, Szabados L (2008) Duplicated *P5CS* genes of Arabidopsis play distinct roles in stress regulation and developmental control of proline biosynthesis. Plant J **53:** 11-28

Tan S, Evans R, Singh B (2006) Herbicidal inhibitors of amino acid biosynthesis and herbicide-tolerant crops. Amino Acids **30:** 195-204

Taylor NL, Heazlewood JL, Day DA, Millar AH (2004) Lipoic acid-dependent oxidative catabolism of α-keto acids in mitochondria provides evidence for branched-chain amino acid catabolism in Arabidopsis. Plant Physiol **134:** 838-848

Thompson AR, Vierstra RD (2005) Autophagic recycling: lessons from yeast help define the process in plants. Curr Opin Plant Biol **8:** 165-173

Trapnell C, Hendrickson DG, Sauvageau M, Goff L, Rinn JL, Pachter L (2013) Differential analysis of gene regulation at transcript resolution with RNA-seq. Nat Biotechnol **31:** 46-53

Trapnell C, Pachter L, Salzberg SL (2009) TopHat: discovering splice junctions with RNA-Seq. Bioinformatics **25:** 1105-1111

Went F (1953) The effect of temperature on plant growth. Annu Rev Plant Physiol **4:** 347-362

Xu Y, Yu H, Hall TC (1994) Rice triosephosphate isomerase gene 5' sequence directs β-Glucuronidase activity in transgenic tobacco but requires an intron for expression in Rice. Plant Physiol **106:** 459-467

Yu Q, Han H, Vila-Aiub MM, Powles SB (2010) AHAS herbicide resistance endowing mutations: effect on AHAS functionality and plant growth. J Exp Bot **61:** 3925-3934

Zou C, Lehti-Shiu MD, Thomashow M, Shiu SH (2009) Evolution of stress-regulated gene expression in duplicate genes of *Arabidopsis thaliana*. PLoS Genet **5:** e1000581

Chapter 3

Characterization of mutants defective in BCAA catabolism

3.1 Abstract

The branched-chain amino acids (BCAAs) leucine, isoleucine and valine are among nine essential amino acids that must be obtained from the diet of humans and other animals, and can be nutritionally limiting in plant foods. Despite genetic evidence of its importance in regulating seed amino acid levels, the full BCAA catabolic network is not completely understood in plants, and limited information is available regarding its regulation. In this study, mutants defective in putative branched-chain ketoacid dehydrogenase complex subunits were demonstrated to accumulate higher levels of BCAAs in mature seeds, providing genetic evidence for their function in BCAA catabolism. In addition, prolonged dark treatment caused the mutants to undergo senescence early and over-accumulate leaf BCAAs compared with the isogenic wild-type plants. These results extend the previous evidence that BCAAs can be catabolized and serve as respiratory substrates at multiple steps. Moreover, comparison of amino acid profiles between mature seeds and dark-treated leaves revealed differences in amino acid accumulation when BCAA catabolism is perturbed. Together, these results demonstrate the consequences of blocking BCAA catabolism during both normal growth conditions and under energy-limited conditions.

3.2 Introduction

The branched-chain amino acids (BCAAs) Leu, Ile and Val are among nine amino acids essential for human and other animals because they cannot be synthesized *de novo* (Harper et al., 1984). Plants synthesize BCAAs and are the main source of these essential nutrients in the diets of humans and agriculturally important animals. In addition to their nutritional value, BCAAs and BCAA-derived metabolites such as glucosinolates, fatty acids and acyl sugars, contribute to plant growth, development, defense and flavor (Mikkelsen and Halkier, 2003; Taylor et al., 2004; Ishizaki et al., 2005; Slocombe et al., 2008; Araujo et al., 2010; Ding et al., 2012; Kochevenko et al., 2012). Despite long-term interest in optimizing the content of these essential amino acids in plants, the genes and proteins that constitute the full BCAA catabolic network are not completely characterized in *A. thaliana* or any other plant, and there is much to learn about the genetic and biochemical regulation of this process.

Recent observations using mutants blocked in BCAA catabolism (Figure 1.3) indicate that the regulation of seed amino acid metabolism has unexpected, and potentially important, features. Previous work revealed that mutants defective in the *A. thaliana* mitochondrial enzymes branched-chain aminotransferase 2 (BCAT2), isovaleryl-CoA dehydrogenase (IVD), α (MCCA) and β (MCCB) subunits of 3-methylcrotonyl-CoA carboxylase, and hydroxymethylglutaryl-CoA lyase (HML) exhibit increases in all three free BCAAs in mature seeds (Gu et al., 2010; Lu et al., 2011; Angelovici et al., 2013). While it is logical that defects in the enzymes early in the catabolic pathway (the *bcat2* and *ivd1* mutants) would cause accumulation of Leu, Ile and Val, it was not expected that mutants blocked in three enzymes specific to Leu degradation (*mcca1*, *mccb1* and *hml1*) would also accumulate Ile and Val. Even more surprising is that

biosynthetically unrelated amino acids - including His and Arg - accumulate to higher levels in seeds of *ivd1*, *mcca1-1*, *mccb1-1* and *hml1* mutants compared to the wild type (Gu et al., 2010; Lu et al., 2011). This suggests that the *A. thaliana* amino acid networks are more interconnected than previously thought, and reveals that there are important gaps in our knowledge of the regulation of amino acid metabolism.

Recent studies in *A. thaliana* revealed that BCAA catabolism plays physiological roles beyond maintaining free amino acid homeostasis (Ishizaki et al., 2005; Araujo et al., 2010). In addition to catalyzing the third step in the degradation of BCAAs, IVD helps plant survive under energy-limited conditions by serving as a source of electrons for the mitochondrial electron transport chain via the electron-transfer flavoprotein α and β subunits (ETFα and ETFβ) and the electron-transfer flavoprotein ubiquinone oxidoreductase (ETFQO) (Figure 1.3). Two lines of evidence for this role are that the *ivd1-2* mutant becomes senescent faster than wild type in prolonged darkness, and mutants defective in ETFβ and ETFQO accumulate more free BCAAs and the IVD substrate isovaleryl-CoA (Ishizaki et al., 2005; Ishizaki et al., 2006; Araujo et al., 2010). In addition, the transcripts of the functionally validated BCAA catabolism genes *BCAT2*, *IVD1*, *MCCA1* and *MCCB1* rapidly increase following transition from light to dark, and this increase is inhibited by sucrose (Fujiki et al., 2000; Che et al., 2002; Binder, 2010; Angelovici et al., 2013). These observations suggest that IVD and other BCAA catabolic enzymes contribute to plant fitness under energy-limited conditions.

Although genetic and biochemical evidence exists for the participation of *A. thaliana* enzymes BCAT2, IVD, MCCA, MCCB and HML in BCAA catabolism (Gu et al., 2010; Lu et al., 2011; Ding et al., 2012; Angelovici et al., 2013), much less is known about the genes and encoded proteins for the branched-chain ketoacid dehydrogenase (BCKDH) complex. The *A.*

thaliana BCKDH complex is a megadalton enzyme complex comprised of multiple subunits (See Chapter 1 and the Introduction of Chapter 2 for descriptions of the BCKDH complex and its subunits). BCKDH enzyme activity was detected in isolated *A. thaliana* mitochondria (Taylor et al., 2004). However, the size of this complex has hindered its detailed *in vitro* characterization in plants. The current identification of its subunits is based upon sequence annotation rather than functional analysis (Fujiki et al., 2000; Mooney et al., 2000; Taylor et al., 2004).

In my dissertation research, a functional genomics analysis of genes annotated as encoding subunits of the BCKDH complex was performed. A variety of transcript analyses were utilized to provide evidence for the participation of the putative BCKDH complex subunits in BCAA catabolism, and the results are described in Chapter 2. This chapter reports results of characterization of BCAA catabolic mutants, including different tissues and following various treatments. Mutants in highly coexpressed BCAA catabolism genes described in Chapter 2 accumulate higher levels of BCAAs in seeds and have enhanced senescence and increased amino acid accumulation in leaves of plants subjected to prolonged darkness. These data provide experimental evidence for the participation of the putative BCKDH subunit genes *E1A1*, *E1A2*, *E1B1*, *E1B2* and *E2* in BCAA catabolism, reinforcing the importance of BCAA catabolism in regulating amino acid homeostasis under day-night cycles and prolonged darkness, and during seed development. These results are consistent with the hypothesis that the *A. thaliana* BCAA catabolic network interacts with energy metabolism at multiple steps.

3.3 Results

3.3.1 Double mutants of BCKDH E1α subunits accumulate increased seed free BCAAs

The coexpression analyses described in Chapter 2 (Figure 2.1, Figure 2.2) led to the hypothesis that the putative BCKDH E1α1, E1β1, E1β2 and E2 subunits, which are co-regulated with the known BCAA catabolic enzymes, function in BCAA degradation. Because *ivd1-2*, *mcca1-1*, *mccb1-1*, *hml1* (Gu et al., 2010; Lu et al., 2011) and *bcat2* mutants (Angelovici et al., 2013) have increased free BCAAs in mature seeds, I hypothesized that mutants defective in *bona fide* BCKDH subunits would share this phenotype.

To test this idea, the free BCAA content of mature dry seed was examined in homozygous mutants of proposed BCAA catabolic enzyme genes described in Chapter 2 (Figure 2.9). The positive control mutants *bcat1-1*, *bcat2-1*, *ivd1-2*, *mcca1-1*, *mccb1-1*, and *hml1-2* (Figure 3.1, Table 3.1) displayed BCAA changes similar to those previously reported (Gu et al., 2010; Lu et al., 2011; Angelovici et al., 2013). Some of the single mutants defective in the putative BCKDH subunits yielded small but significant ($p<0.05$) changes in free BCAAs. For example, both of the *e1a1* mutants and *e1b2-1* had modest increases (1.5 to 5.8 fold relative to the wild type) in Ile and Val, or all three BCAAs (Figure 3.1, Table 3.1). I tested the hypothesis that the existence of paralogs was responsible for the modest increases in the single mutants by examining free seed BCAA levels in *e1a1-1*; *e1a2-1* and *e1a1-2*; *e1a2-1* double mutants, defective in both genes annotated as encoding E1α subunits. Consistent with the hypothesis, both double mutants had large and statistically significant ($p<0.05$) increases in seed free BCAAs: 45-, 22- and nine-fold increases for Leu, Ile and Val, respectively (Figure 3.1, Table 3.1). These data

support the hypothesis that E1α1 and E1α2 contribute to the degradation of BCAAs during seed development.

Because no null *e1b1* mutants were available, constructs encoding artificial microRNAs (amiRNAs) targeting *E1B1* were transformed into the *e1b2-1* mutant background in an attempt to reduce expression of both paralogous E1β subunit genes (Figure 3.2). True breeding homozygous T3 transgenic lines from two independent primary transformants had *E1B1* transcript reduction to ~30% of the wild type level (Figure 3.2A). The free amino acid content from the seeds of homozygous T3 lines exhibited moderate but significant ($p<0.05$) BCAA increases compared with wild type (Figure 3.2B). Thus, modest reduction of the *E1B1* transcript in the *e1b2-1* mutant background results in more seed free BCAAs. This supports the hypothesis that both E1β enzyme paralogs participate in BCAA catabolism during seed development. Genetic analysis of the role of E2 and mtLPDs was not possible since no null alleles were available for *E2*, *mtLPD1* or *mtLPD2*.

3.3.2 BCAA catabolic mutants exhibit early senescence under prolonged darkness

The increased seed free BCAAs in the mutant and amiRNA lines deficient in *E1A1*, *E1A2*, *E1B1* and *E1B2* support the hypothesis that the tested genes encode proteins of BCAA catabolism. Moreover, elevated transcript levels during prolonged darkness were observed for the eight highly coexpressed known or proposed BCAA catabolism genes identified in Chapter 2 (Figure 2.7), supporting the hypothesis that BCAA catabolic enzymes - including BCKDH subunits E1A1, E1B1, E1B2 and E2 - have one or more physiological roles in the dark.

To explore this idea, BCAA catabolic mutant leaf morphology was monitored during prolonged darkness for 15 days. Prior to dark treatment, all of the lines being assessed were

green and showed normal rosette leaf morphology (Figure 3.3A, Figure 3.4). As previously described for *ivd1-2* (Araujo et al., 2010), I found that mutants defective in the other validated Leu catabolic enzyme genes - *MCCA1*, *MCCB1* and *HML1* - exhibit enhanced dark-induced senescence phenotypes (Figure 3.3A, Figure 3.4). These results demonstrate that early leaf senescence is observed in mutants blocked from the middle to the end of the Leu catabolism pathway. I asked whether the putative BCKDH subunit mutants had abnormal dark induced senescence. The two mutants defective in the putative E1A1 subunit gene that is upregulated in the dark (*e1a1-1* and *e1a1-2*) showed early dark-induced senescence (Figure 3.3A, Figure 3.4). Double mutants that combine these mutations with the loss of function allele for the paralogous gene (*e1a1-1; e1a2-1* and *e1a1-2; e1a2-1* double mutants) exhibited senescence phenotypes similar to the *e1a1* single mutants. These results are not surprising because - in contrast to the dark inducibility of *E1A1* transcript – *E1A2* mRNA is not induced in the dark (Figure 2.7). In fact the *e1a2-1* single mutant did not exhibit enhanced dark-induced senescence compared to the wild type control, even after 15 days. Two other single mutants, *bcat2-1* and *e1b2-1*, also did not exhibit enhanced senescence. To assess viability following the 15-day dark treatment, plants were transferred back to 8h light/16h dark photoperiod and examined for new growth after one week. All genotypes recovered from this extended dark treatment except for *ivd1-2*, which had dry and yellow leaves at the end of the prolonged dark treatment (Figure 3.3A, Figure 3.4). These results demonstrate that, in addition to *IVD1*, the disruption of *E1A1*, *MCCA1*, *MCCB1* or *HML1* also results in early senescence in prolonged darkness.

To complement the analysis of gross morphological changes, the maximum photochemical efficiency of Photosystem II (F_v/F_m) was used to quantify leaf senescence caused by prolonged darkness (Oh et al., 1997). Mutant chlorophyll fluorescence was indistinguishable

from that of wild type prior to dark treatment (Figure 3.3B, day 0). In contrast, on day nine the mutants that showed early senescence also exhibited statistically significantly lower F_v/F_m values relative to the corresponding wild type (Figure 3.3B).

3.3.3 Enhanced leaf free Leu, Ile and Val accumulation during prolonged darkness

The *ivd1-2* mutant was demonstrated to have increased leaf free BCAAs during prolonged darkness (Araujo et al., 2010). To test whether other mutants defective in BCAA catabolism also show enhanced BCAA accumulation, free rosette leaf amino acid content was analyzed in wild type and catabolic mutants subjected to prolonged darkness. All three BCAAs increased in dark-treated wild-type plants at all three time points, and further increases due to blocked BCAA catabolism were seen for some combinations of mutants and amino acids (Figure 3.5, Table 3.2). The effects of all mutations, except for *e1a2* and *e1b2*, were strongest and most consistent for Leu (eight to 14 fold increases compared with wild type at day nine), with the least significant increases in Ile. Of the mutants, *e1a2-1* and *e1b2-1* displayed the least significant increases compared to wild type and this is especially obvious for Leu and Val. The small impact of these mutations is reminiscent of the results with free seed amino acid changes (Figure 3.1, Table 3.1). The *E1B1*-silenced *e1b2-1* plants were also evaluated for leaf amino acid content after six days in prolonged darkness. The homozygous T3 lines from two independent primary transformants showed stronger BCAA increases compared with the *e1b2-1* mutant and type (Figure 3.2C). In summary, the BCAA increases in the mutants support the hypothesized roles of BCKDH E1α1, E1β1 and E1β2 subunits in BCAA catabolism - especially in Leu degradation - during prolonged darkness.

Taken together, analysis of the mutants and amiRNA lines demonstrated that disruption of BCAT2, E1α1, E1β1, E1β2, IVD, MCCA, MCCB, HML enzymes led to early senescence and/or increased leaf free BCAAs in prolonged darkness. These results further support the hypothesis that the full BCAA catabolic pathway plays a role in plant survival under carbon-limited conditions. These data also add to the evidence that *E1A1*, *E1B1* and *E1B2* encode subunits of the BCKDH complex.

3.3.4 Beyond Leu, Ile and Val: Blocking branched-chain ketoacid dehydrogenase causes broad changes in leaf and seed amino acids

In addition to accumulating seed free BCAAs, *ivd1-2*, *mcca1-1*, *mccb1-2* and *hml1-2* mutants were previously shown to have increased levels of biosynthetically seemingly unrelated amino acids, with increased Arg and His common to all (Gu et al., 2010; Lu et al., 2011). I asked whether this inter-pathway phenomenon was observed in mutants annotated as defective in BCKDH subunit genes (Figure 3.6, Table 3.1, Table 3.2). The most obvious sign of cross pathway seed amino acid accumulation was seen in the *e1a1-1; e1a2-1* and *e1a1-2; e1a2-1* double mutants, which had two to seven fold and statistically significant ($p<0.001$) increases in Arg, His, Met and Ser (Figure 3.6A, Table 3.1). The *e1b2-1* mutant also had more modest, but quite widespread increases in multiple seed amino acids, despite the presence of the intact paralogous *E1B1* gene. As previously reported for mutants blocked later in the pathway (Gu et al., 2010; Lu et al., 2011), high seed His was a hallmark of all the putative BCKDH subunit mutants, with the exception of *e1a2-1*. The positive control *ivd1-2* mutant showed the expected highly pleiotropic amino acid changes in both seeds and leaves by the end of the dark treatment: it exhibited significant increases for 16 out of 19 amino acids detected in seeds as was previously

reported (Gu et al., 2010; Lu et al., 2011) (Figure 3.6A, Table 3.1), and 17 out of 19 amino acids detected by the end of the nine-day dark treatment (Figure 3.6B, Table 3.2). The *ivd1-2* mutant displayed severe dehydration and senescence when tissue was extracted after 9d in the dark. Thus, the data from 6d dark treated *ivd1-2* mutant plants - which is similar to the 9d data in Figure 3.6B (Table 3.3) - is likely to be more reliable. With the possible exception of increased Asp and Met, dark-induced leaf amino acid changes were less consistent across the mutants in BCKDH and in other enzymes of BCAA catabolism.

3.4 Discussion

In the current research, I took a systems approach to hypothesizing and testing functions of putative BCKDH subunits through a combination of sequence similarity, transcript coexpression analyses, and mutant characterization. I provided evidence showing the participation of putative BCKDH subunits in BCAA catabolism, and demonstrated examples of genetic redundancy and functional divergence between enzyme paralogs. Moreover, my data revealed that BCAA catabolic mutants defective at multiple steps in the pathway exhibit early senescence with high leaf free BCAAs in prolonged darkness, supporting a previous hypothesis that BCAA degradation generates alternative sources of energy in plants under energy-limited conditions (Ishizaki et al., 2005; Araujo et al., 2010).

3.4.1 Evidence that BCAA catabolism and energy metabolism interact at multiple steps

In this study, I demonstrated that eight out of the 13 proposed or validated BCAA catabolism genes are coexpressed and share common transcript oscillation patterns in diurnal and circadian treatments (Figure 2.1, Figure 2.2). These findings are consistent with the observed fluctuation of free BCAA levels on diel cycles (Gibon et al., 2006; Espinoza et al., 2010), demonstrating the physiological importance of BCAA catabolism at night. BCAAs are proposed to provide their downstream catabolic products - acetoacetate, acetyl-CoA and propionyl-CoA - to the tricarboxylic acid (TCA) cycle for energy generation (Figure 1.3) (Anderson et al., 1998). In addition, Mentzen and coworkers performed a global coexpression analysis with microarray data from 70 experiments, and pointed out a coexpression supermodule capable of maintaining cellular energy balance via catabolism (Mentzen et al., 2008). This supermodule contained genes

encoding enzymes in the catabolism of amino acids (including BCAAs), carbohydrates, lipids and cell wall components. Despite the proposed role of BCAA catabolism in energy generation, no vegetative or reproductive phenotype was observed for BCAA catabolic mutants grown in photoperiods with 16h, 12h or 8h light (Lu et al., 2011, and day 0 in Figure 3.3 and Figure 3.4 of this study), in contrast to the reported aberrant reproductive development of the *mcca1* and *mccb1* mutants (Ding et al., 2012). My results suggest a limited role of BCAA catabolism in providing TCA cycle substrates in day/night cycling conditions (Lu et al., 2011 and this study).

While diurnal regulation is more physiologically relevant, the prolonged darkness assay is experimentally convenient for studying the intricate regulation of energy metabolism and interacting metabolic processes in plants. Success using such an approach was previously demonstrated in identifying and characterizing mutants defective in genes participating energy related processes including autophagy, mitochondrial electron transport chain and starch metabolism (Gibon et al., 2004; Ishizaki et al., 2005; Ishizaki et al., 2006; Liu and Bassham, 2012). During energy-limited conditions, autophagy was demonstrated to promote organelle degradation including chloroplasts, partially contributing to the increase of free amino acid pools including free BCAAs in vegetative tissues (Hanaoka et al., 2002; Wada et al., 2009; Izumi et al., 2010). Consistent with this idea, in the current study we also observed dramatic free amino acid increases including BCAAs, aromatic amino acids, Lys, His, Asn and Arg in the wild type within three days in prolonged darkness (Table 3.2). In addition, Lys metabolism was previously demonstrated to interact with plant energy metabolism and the high seed Lys KD genotype showed reduced levels of the TCA cycle intermediates (Angelovici et al., 2011).

My results are consistent with the notion that multiple steps of the BCAA degradation pathway provide alternative sources of energy under long-term dark treatment conditions. Prior

to the current study, IVD was the only enzyme in BCAA catabolism that was shown to play a role in plant survival in energy-limited conditions, because *ivd1-2* exhibited enhanced senescence relative to the wild type in prolonged darkness (Araujo et al., 2010). The authors hypothesized that the early senescence of *ivd1-2* resulted from deficiency in supplying electrons to the mitochondrial electron transport chain and/or providing the BCAA catabolic products to the TCA cycle (Ishizaki et al., 2005; Araujo et al., 2010) (Figure 1.3). In my study, BCAA catabolic mutants defective in enzymes both upstream (*e1a1* single mutants and *e1a1; e1a2* double mutants) and downstream (*mcca1-1*, *mccb1-1* and *hml1-2* mutants) of IVD displayed enhanced senescence in prolonged darkness (Figure 3.3, Figure 3.4), supporting the hypothesized role of BCAA catabolism in providing TCA cycle substrates in energy-limited conditions. Interestingly, *ivd1-2* exhibited dehydrated and yellowed rosette leaves more rapidly than the other mutants (Figure 3.4), and was the only mutant that did not recover after the 15d dark treatment (Figure 3.3). This is consistent with the hypothesis that IVD influences energy homeostasis in multiple ways, not only by providing BCAA catabolic CoA intermediates to the mitochondrial electron transport chain, but also by catabolizing additional substrates such as phytanoyl-CoA and aromatic amino acids (Araujo et al., 2010). The observation that *e1a1* single mutants and *e1a1; e1a2* double mutants, which are defective in subunits of the upstream BCKDH complex, showed enhanced senescence at a weaker level than *ivd1-2* (Figure 3.3, Figure 3.4) supports the hypothesized broader substrate range for IVD. Taken together, my results provide genetic evidence supporting the hypothesized interaction between BCAA catabolism and energy metabolism at multiple steps, and are consistent for a broader substrate range for IVD in plants.

3.5 Material and methods

3.5.1 Free amino acid analysis by LC-MS/MS

Free amino acids from dry seeds or the 10th and 11th rosette leaves were extracted and analyzed by modifying a previously described method (Gu et al., 2007; Lu et al., 2008; Gu et al., 2012). 1 μM of five heavy amino acids were added to the extraction buffer for a more accurate quantification: Leu-d_{10} for Leu and Ile, His-d_3 for His, Trp-d_5 for Trp, Val-d_8 for Val and Phe-d_8 for all other amino acids. Details on the detection of selected ion monitoring pairs were as described previously (Angelovici et al., 2013). The amino acid quantities were normalized to the fresh weight of the harvested samples. All heavy amino acids were purchased from Cambridge Isotope Laboratories (Tewksbury, MA).

3.5.2 Generation of *E1B1*-silenced *e1b2-1* mutant

To clone the artificial micro RNA (amiRNA) constructs, two amiRNAs targeting *E1B1* were designed using the WMD3-Web MicroRNA Designer (http://wmd3.weigelworld.org/cgi-bin/webapp.cgi): *amiE1B1-1* (5'-TATGCGATTACATTAGTCCTT-3') and *amiE1B1-2* (5'-TAACTACAGATAGTACGCCTA-3'). The precursor amiRNAs were cloned by overlapping PCR following protocols provided by WMD3-Web MicroRNA Designer (http://wmd3.weigelworld.org/downloads/Cloning_of_artificial_microRNAs.pdf), amplified by Gateway-compatible primers (Table 3.4) and inserted into pEarley100 (Earley et al., 2006). The resultant constructs were transformed into the *e1b2-1* single mutant. All progeny lines were tested for the presence of the amiRNA and T-DNA prior to further analyses.

3.5.3 Determination of the PSII photochemical efficiency

All chlorophyll fluorescence experiments were performed at the Center for Advanced Algal and Plant Phenotyping (CAAPP) at Michigan State University (http://www.prl.msu.edu/caapp) with previously described setups for the growth chambers (Attaran et al., 2014). Plants were grown under SD for five weeks and then subjected to prolonged darkness for 14 days. Chlorophyll fluorescence parameters were measured each day at the end of the subjective night from day 0 to 14 in prolonged darkness. Experiments were repeated twice and representative results were shown.

3.6 Acknowledgments

I would like to thank Dr. Ruthie Angelovici for providing heavy amino acid standards and MS method for amino acid detection, Lijun Chen for help with fine-tuning the MS method and assistance in using the LC-MS/MS, Dr. Ronghui Pan for providing vectors and protocols for amiRNA mutant generation, and Linda Savage and David Hall for help with PSII maximum photochemical efficiency measurements.

APPENDIX

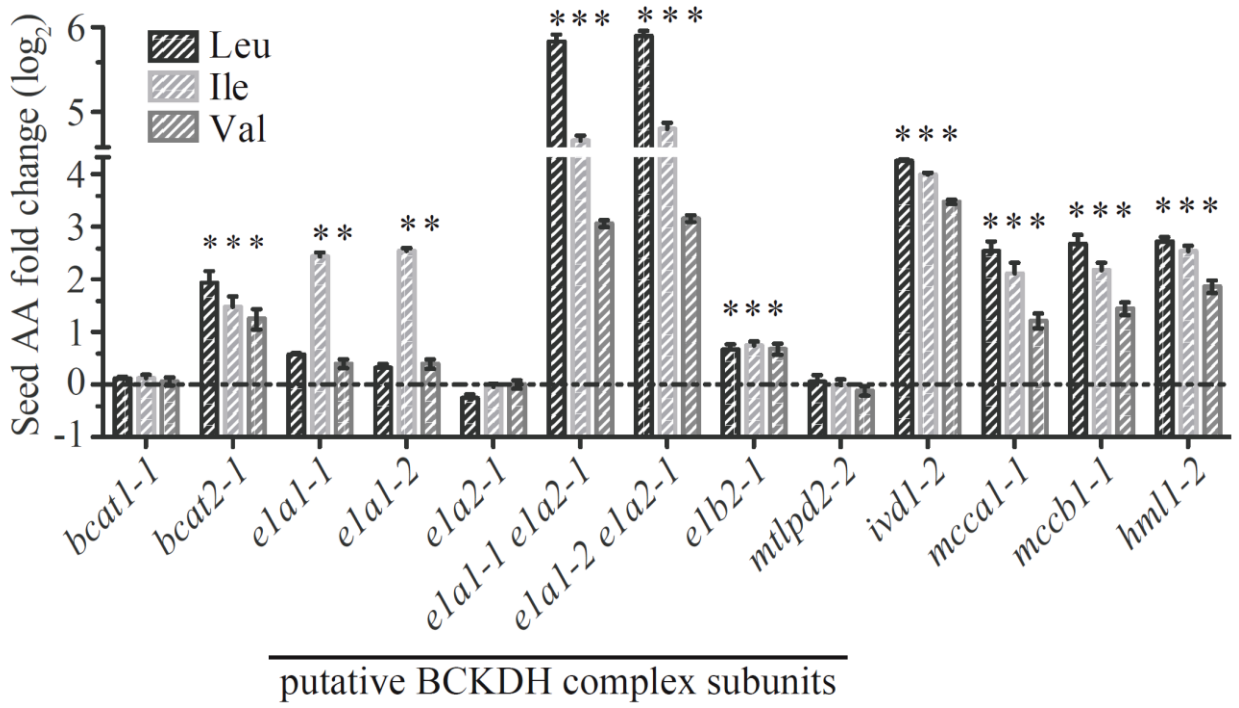


Figure 3.1 Changes in free BCAA content in dry seeds of mutants relative to wild type. The bars show the fold change of individual amino acids in each mutant compared to the wild type (horizontal dashed line) grown at the same time. Four or more biological replicates were measured for each genotype. An asterisk indicates a significant difference from the wild type, determined by the Student's *t*-test ($p<0.05$). Error bars represent means ± SE. The experiments were done at least three times with similar results obtained, and representative results are shown.

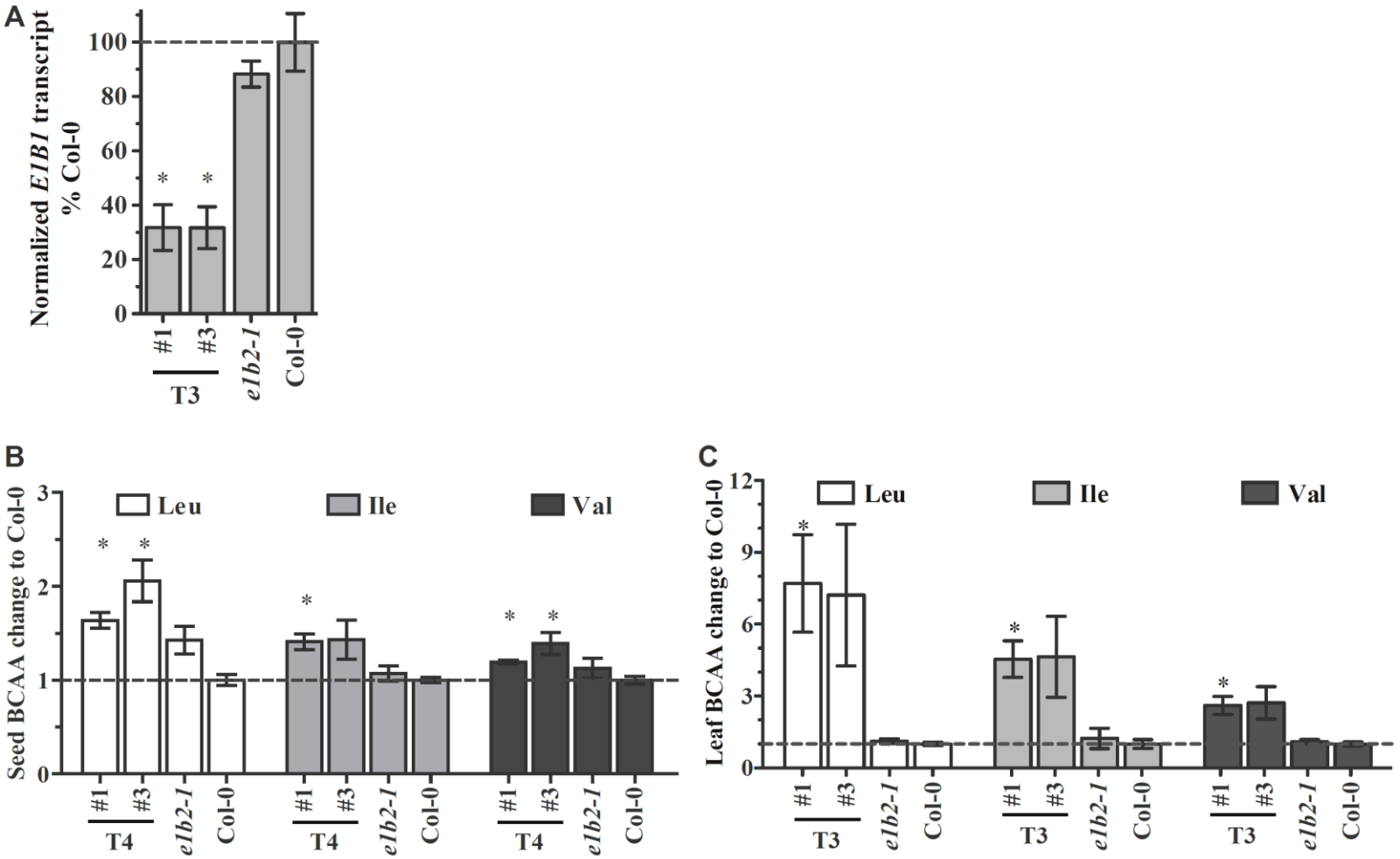


Figure 3.2 Transcript and mutant analyses of *E1B1*-silenced *e1b2-1* lines. (A) Relative *E1B1* transcript abundance in two homozygous *E1B1*-silenced *e1b2-1* T3 lines (progeny seed pools of independent primary transformants #1 and #3) by qPCR. The *y* axis values represent the normalized *E1B1* transcript levels relative to Col-0 (shown as the horizontal dashed line, n=5, mean ± SE).

Figure 3.2 (cont'd)

The *E1B1* transcript levels were normalized to *ACT2* transcript levels. An asterisk indicates a significant difference from the wild type, determined by the Student's *t*-test ($p<0.05$). Leaf tissues were harvested by the end of night on a 16h light/8h dark photoperiod. (B) Relative levels of seed free BCAAs in homozygous *E1B1*-silenced *e1b2-1* T4 seeds (progeny seed pools of #1 and #3). The *y* axis values represent the amino acid levels relative to Col-0 (shown as the dashed line, n=4, mean ± SE). An asterisk indicates a significant difference from the wild type, determined by the Student's *t*-test ($p<0.05$). (C) Relative levels of leaf free BCAAs in homozygous *E1B1*-silenced *e1b2-1* T3 lines (progeny seed pools of #1 and #3) after 6 days in prolonged darkness. The *y* axis values represent the amino acid levels relative to Col-0 (shown as the dashed line, n=5, mean ± SE). An asterisk indicates a significant difference from the wild type, determined by the Student's *t*-test ($p<0.05$).

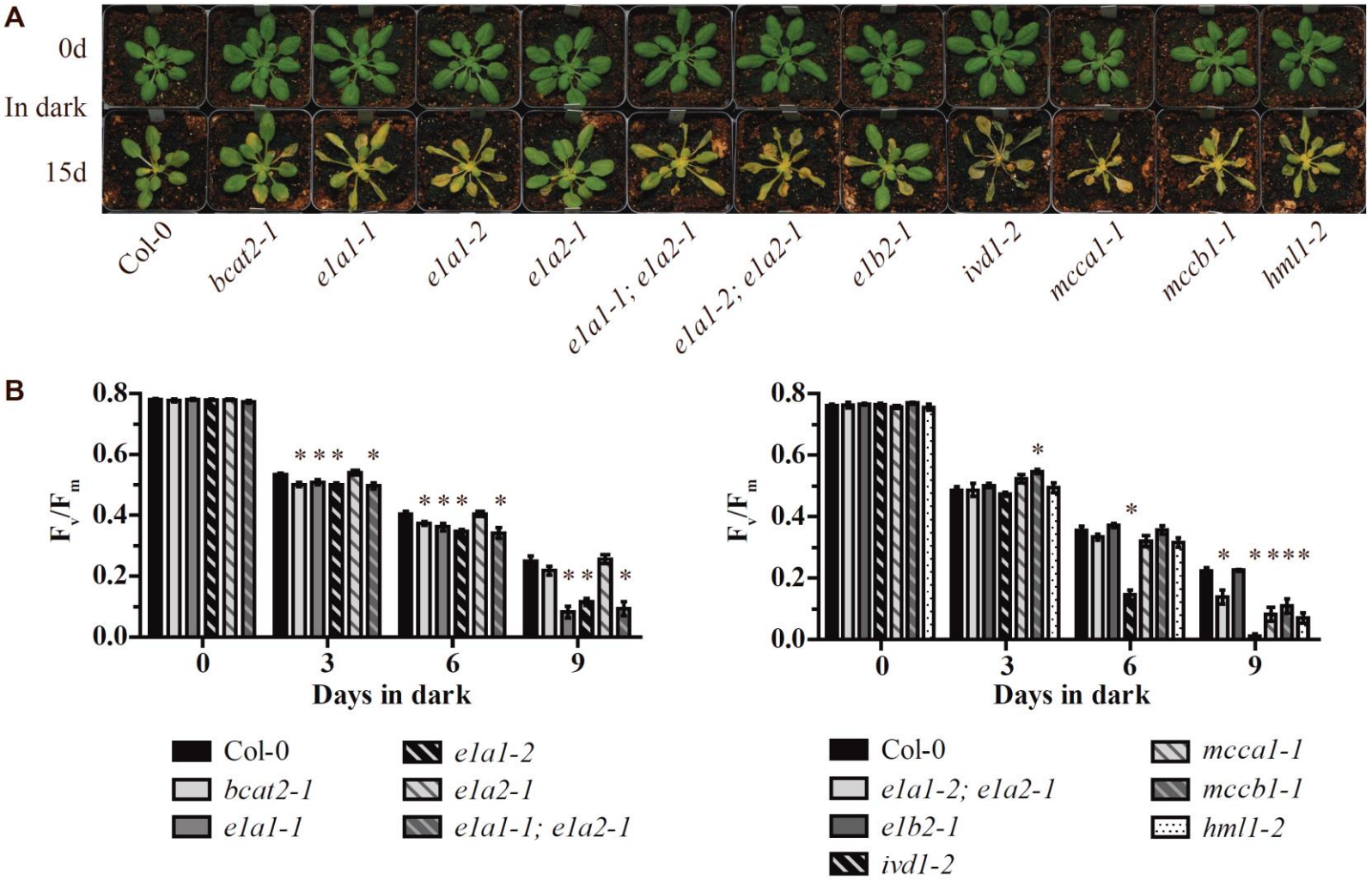


Figure 3.3 Phenotypes of BCAA mutants subjected to prolonged darkness. (A) Photographs of 5-week-old, short-day-grown *A. thaliana* plants taken prior to (0d) and after 15 days of prolonged darkness. The leaves of *e1a1-1*, *e1a1-2*, both *e1a1; e1a2* double mutants, *ivd1-2*, *mcca1-1*, *mccb1-1*, and *hml1-2* were visibly yellowed and dehydrated following 15 days of prolonged darkness

Figure 3.3 (cont'd)

compared to the wild type. The experiments were done at least three times with similar results, and representative results are shown. (B) Analysis of the maximum photochemical efficiency of PSII (F_v/F_m) to quantify the kinetics of leaf senescence. Plants were grown under the same conditions as in (A). Values are means ± SE of three to five biological replicates. An asterisk indicates a significant difference from the wild type, determined by the Student's *t*-test ($p<0.05$). The two bar graphs represent data from plants grown in separate flats. Each experiment was done twice with similar results, and representative results are shown.

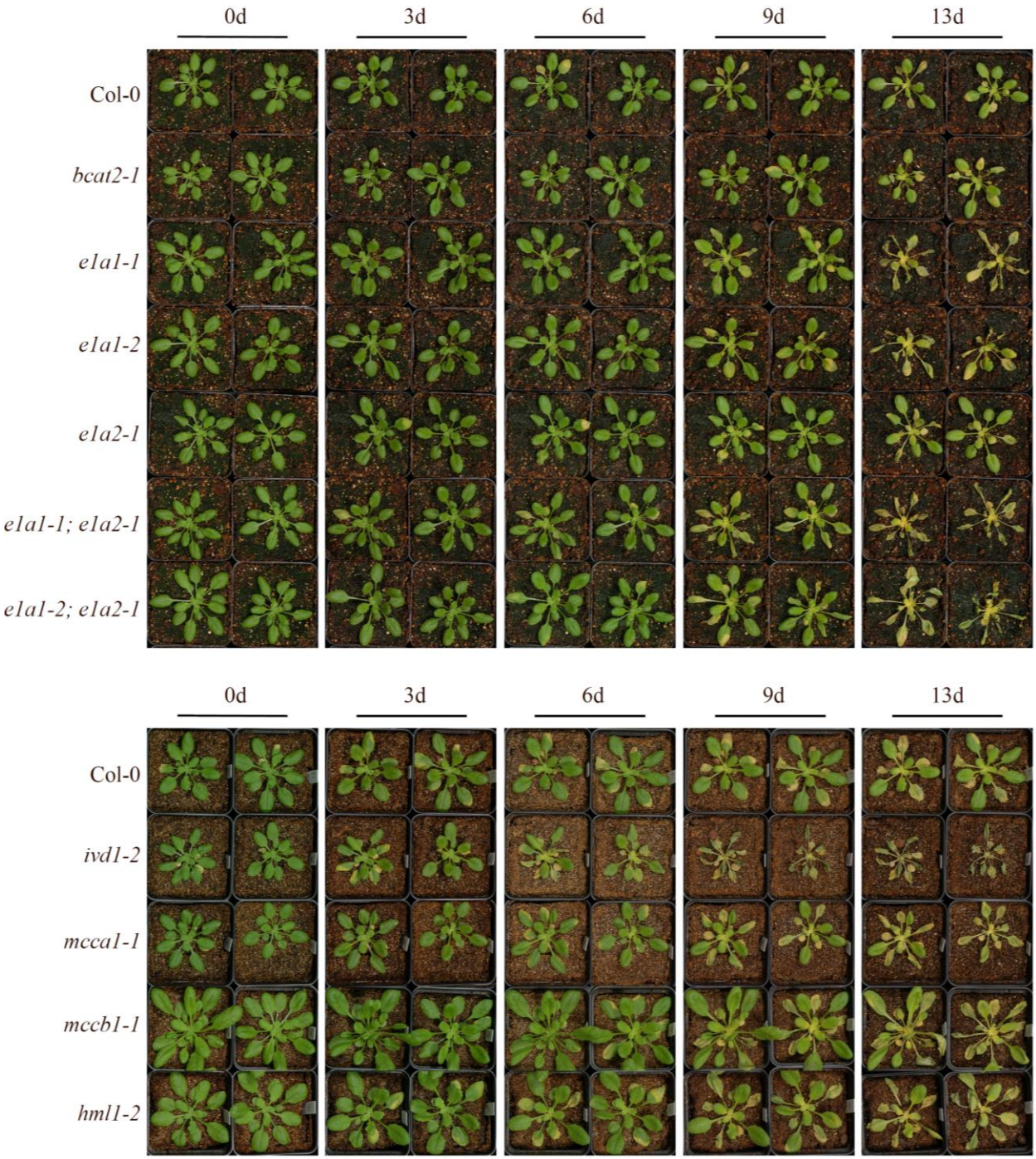


Figure 3.4 Phenotypes of BCAA mutants subjected to prolonged darkness - early time points. Photographs of 5-week-old, short-day-grown Arabidopsis plants taken prior to (0d) and after 3, 6, 9 and 13 days of prolonged darkness. The leaves of *e1a1-1*, *e1a1-2*, both *e1a1; e1a2* double mutants, *ivd1-2*, *mcca1-1*, *mccb1-1*, and *hml1-2* were visibly yellowed and dehydrated

Figure 3.4 (cont'd)

following 13 days of prolonged darkness compared to the wild type. The *ivd1-2* mutant started showing visible senescence symptoms at day 6. The experiments were done at least three times with similar results, and representative results are shown. Plants in the upper and lower panels were grown and assayed in individual experiments. Two replicates were shown for each mutant at every time point.

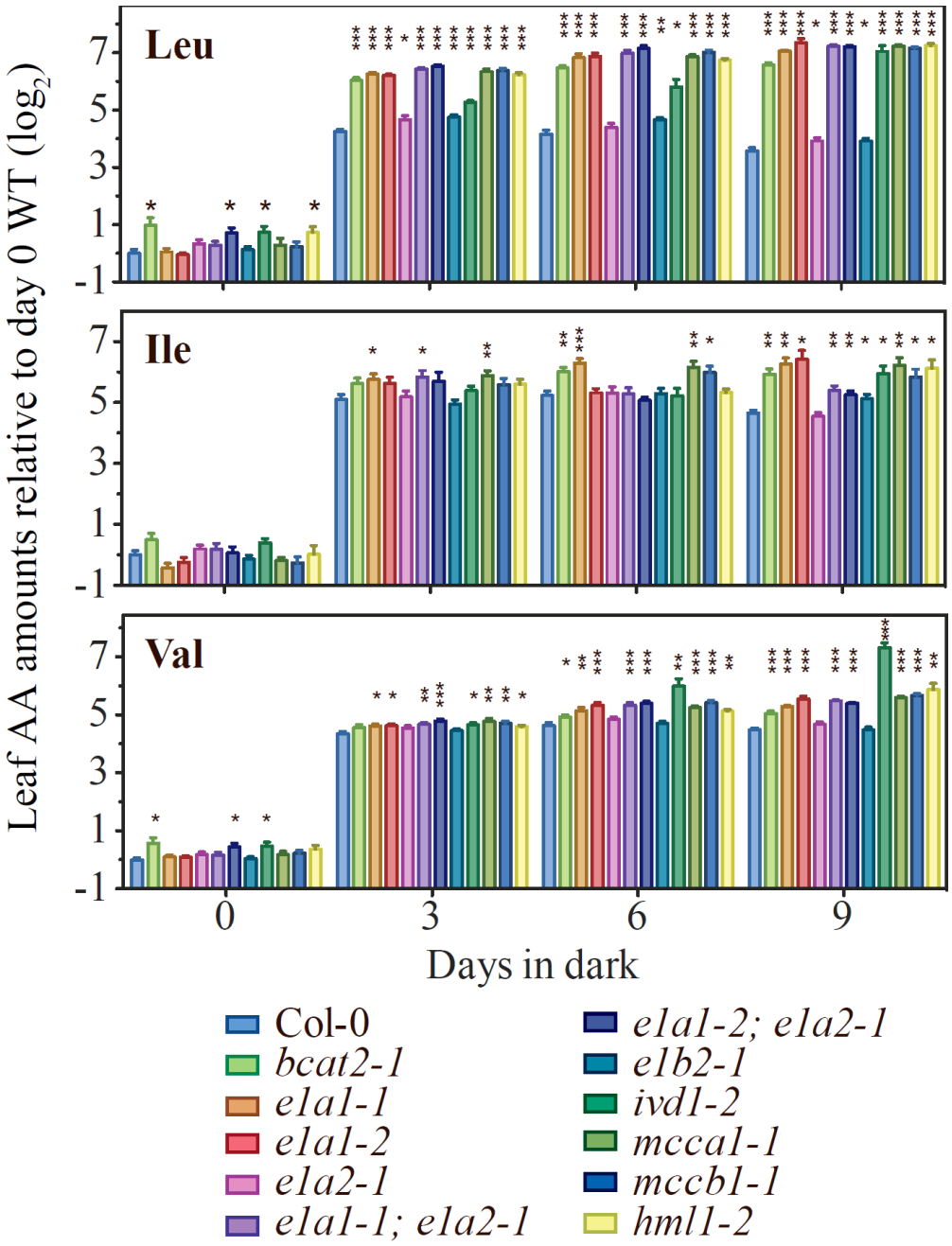


Figure 3.5 Relative levels of leaf free BCAAs in mutants during prolonged darkness. The *y* axis values represent the $\log_2$ transformed amino acid levels normalized to the wild type (Col-0) at day 0. Values are means ± SE of five biological replicates. * indicates significant difference from the wild type at the same time point, determined by the Student's *t*-test $p<0.05$, ** $p<0.01$, *** $p<0.001$.

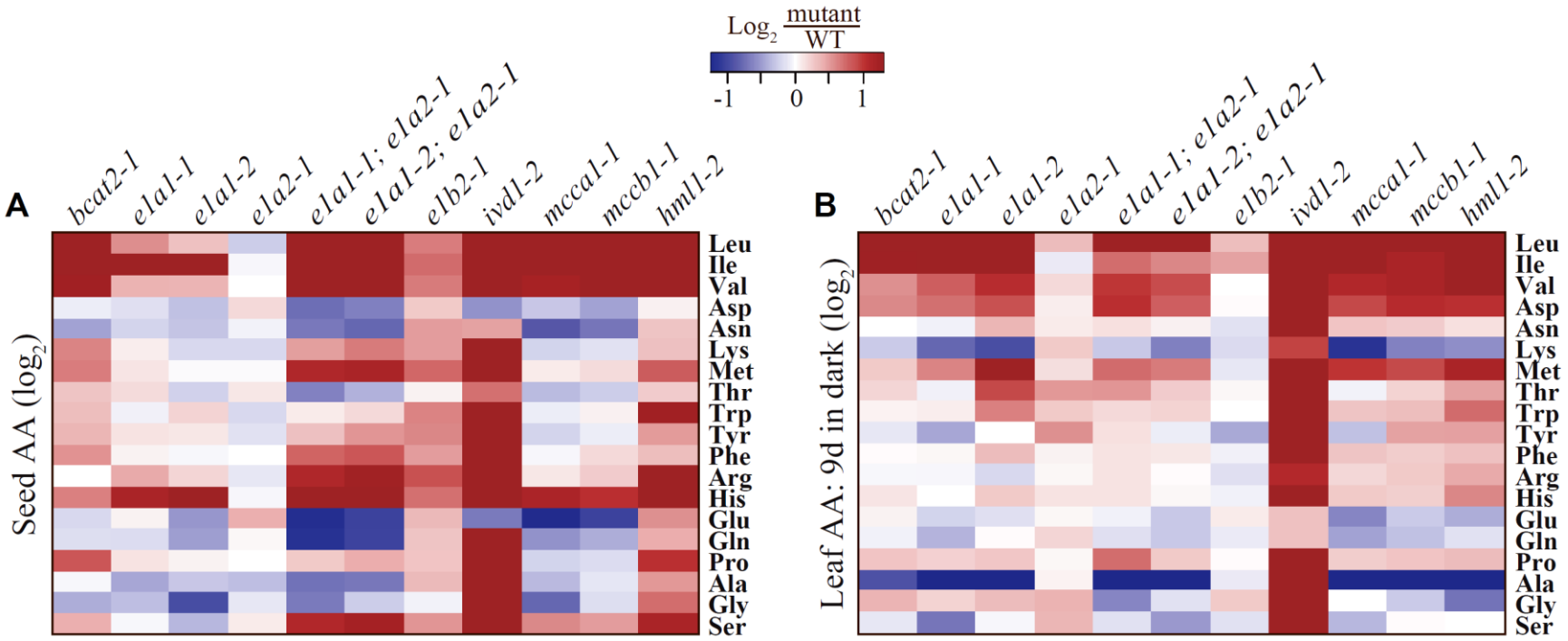


Figure 3.6 Heat map showing the effect of disrupting known or proposed BCAA catabolism genes on amino acid homeostasis. Comparisons of amino acid contents between (A) dry seeds and (B) leaves of 5-week-old plants kept in the dark for 9d. Values represent mean fold change (log_2) in mutants compared to the wild type. Four biological replicates were used for analyzing seed amino acid content, and five for leaf. Maximum color intensities correspond to -1.3 and +1.3, which are equivalent to fold change of 0.4 and 2.5, respectively.

Table 3.1 Mutant seed free amino acid profiles relative to the wild type (Col-0).

Amino acid	*bcat1-1*			*bcat2-1*			*e1a1-1*			*e1a1-2*			*e1a2-1*		
	Mean*	**SE**	***t*-test****	**Mean**	**SE**	***t*-test**	**Mean**	**SE**	***t*-test**	**Mean**	**SE**	***t*-test**	**Mean**	**SE**	***t*-test**
Ala	0.80	0.05	0.23	0.98	0.17	0.92	0.73	0.11	0.30	0.82	0.13	0.48	0.79	0.23	0.53
Arg	1.19	0.08	0.37	1.01	0.16	0.95	1.36	0.18	0.18	1.16	0.07	0.13	0.92	0.06	0.40
Asn	0.97	0.07	0.91	0.72	0.09	0.29	0.86	0.05	0.49	0.81	0.05	0.37	0.97	0.09	0.87
Asp	1.78	0.32	0.35	0.94	0.09	0.76	0.90	0.11	0.45	0.80	0.09	0.10	1.14	0.16	0.46
Gln	1.15	0.04	0.47	0.89	0.14	0.78	0.89	0.20	0.64	0.71	0.05	0.02	1.03	0.09	0.83
Glu	1.17	0.07	0.50	0.87	0.11	0.51	1.04	0.30	0.91	0.69	0.09	0.03	1.34	0.21	0.21
Gly	1.56	0.12	0.13	0.75	0.20	0.50	0.80	0.07	0.27	0.51	0.07	0.03	0.92	0.12	0.70
His	1.08	0.02	0.48	1.56	0.18	0.05	2.28	0.44	0.10	2.44	0.19	0.00	0.97	0.10	0.81
Ile	1.09	0.05	0.57	2.79	0.41	0.02	5.41	0.31	0.00	5.83	0.22	0.00	0.97	0.04	0.79
Leu	1.09	0.03	0.35	3.83	0.63	0.01	1.49	0.03	0.05	1.25	0.06	0.21	0.84	0.04	0.38
Lys	0.99	0.03	0.95	1.53	0.34	0.24	1.06	0.08	0.64	0.87	0.07	0.36	0.87	0.09	0.38
Met	1.03	0.04	0.87	1.58	0.51	0.37	1.09	0.09	0.56	0.98	0.10	0.92	0.98	0.06	0.90
Phe	1.07	0.05	0.63	1.46	0.27	0.19	1.04	0.03	0.57	0.98	0.03	0.75	1.00	0.06	0.99
Pro	1.06	0.21	0.95	1.80	0.57	0.28	1.10	0.07	0.40	1.04	0.07	0.71	1.00	0.10	0.98
Ser	1.01	0.06	0.95	1.34	0.24	0.25	0.97	0.16	0.89	0.77	0.04	0.07	1.08	0.07	0.52
Thr	0.61	0.08	0.30	1.24	0.24	0.58	1.13	0.11	0.34	0.85	0.04	0.02	1.08	0.12	0.56
Trp	1.23	0.07	0.20	1.27	0.17	0.39	0.95	0.06	0.53	1.16	0.11	0.23	0.87	0.04	0.11
Tyr	1.04	0.05	0.78	1.31	0.37	0.55	1.10	0.06	0.45	1.08	0.12	0.63	0.90	0.07	0.47
Val	1.04	0.06	0.82	2.39	0.31	0.02	1.32	0.07	0.03	1.31	0.08	0.02	1.00	0.05	0.96

* The mean represents the fold change between the averages of five biological replicates of mutants and Col-0 grown at the same time.

** '*t*-test' indicates the value at which the mean measurement deviates from expectation (for example '0.01' means significant at <0.01).

Table 3.1 (cont'd)

e1a1-1;e1a2-1			*e1a1-2;e1a2-1*			*e1b2-1*			*ivd1-2*			*mcca1-1*		
Mean	**SE**	***t*-test**	**Mean**	**SE**	***t*-test**	**Mean**	**SE**	***t*-test**	**Mean**	**SE**	***t*-test**	**Mean**	**SE**	***t*-test**
0.60	0.08	0.15	0.62	0.06	0.16	1.28	0.07	0.02	3.18	0.14	0.00	0.78	0.05	0.07
2.18	0.20	0.00	2.39	0.16	0.00	1.84	0.23	0.03	10.67	0.33	0.00	1.09	0.07	0.32
0.62	0.04	0.12	0.58	0.03	0.09	1.42	0.12	0.04	1.39	0.05	0.01	0.54	0.03	0.02
0.59	0.04	0.00	0.64	0.03	0.00	1.21	0.08	0.14	0.68	0.05	0.06	0.82	0.08	0.27
0.44	0.05	0.00	0.49	0.04	0.00	1.24	0.11	0.16	2.60	0.10	0.00	0.68	0.03	0.11
0.43	0.05	0.00	0.48	0.06	0.00	1.30	0.11	0.10	0.63	0.03	0.01	0.42	0.02	0.01
0.63	0.06	0.07	0.84	0.10	0.38	0.96	0.08	0.89	4.00	0.12	0.00	0.58	0.13	0.43
6.78	0.65	0.00	6.93	0.50	0.00	1.64	0.12	0.02	13.35	0.69	0.00	2.23	0.32	0.03
25.39	1.06	0.00	27.92	1.42	0.00	1.67	0.10	0.00	15.96	0.32	0.00	4.33	0.66	0.01
56.96	3.32	0.00	59.78	2.45	0.00	1.59	0.11	0.01	19.11	0.37	0.00	5.81	0.78	0.01
1.40	0.14	0.02	1.60	0.16	0.01	1.42	0.13	0.04	4.87	0.14	0.00	0.85	0.05	0.06
2.18	0.14	0.00	2.27	0.12	0.00	1.70	0.15	0.01	5.22	0.21	0.00	1.08	0.11	0.55
1.71	0.08	0.00	1.80	0.07	0.00	1.42	0.14	0.05	3.12	0.11	0.00	0.97	0.05	0.78
1.21	0.06	0.07	1.35	0.10	0.02	1.23	0.11	0.18	2.51	0.22	0.13	0.85	0.22	0.60
2.17	0.17	0.00	2.34	0.18	0.00	1.44	0.10	0.02	5.05	0.22	0.00	1.51	0.14	0.03
0.64	0.09	0.00	0.76	0.07	0.01	1.04	0.19	0.89	1.65	0.08	0.04	0.79	0.11	0.23
1.08	0.09	0.41	1.14	0.07	0.16	1.54	0.08	0.00	7.16	0.27	0.00	0.94	0.12	0.65
1.26	0.08	0.08	1.45	0.09	0.01	1.52	0.13	0.02	3.95	0.14	0.00	0.86	0.08	0.22
8.35	0.40	0.00	8.93	0.41	0.00	1.60	0.11	0.01	11.12	0.33	0.00	2.32	0.23	0.01

Table 3.1 (cont'd)

mccb1-1			*hml1-2*		
Mean	**SE**	***t*-test**	**Mean**	**SE**	***t*-test**
0.92	0.10	0.53	1.43	0.18	0.09
1.21	0.18	0.32	3.00	0.38	0.01
0.61	0.09	0.03	1.24	0.12	0.16
0.72	0.07	0.12	1.06	0.14	0.76
0.75	0.09	0.16	1.35	0.04	0.09
0.49	0.07	0.00	1.47	0.10	0.01
0.89	0.04	0.82	1.67	0.30	0.27
2.06	0.15	0.00	3.30	0.20	0.00
4.52	0.45	0.00	5.82	0.43	0.00
6.40	0.80	0.01	6.59	0.41	0.00
0.91	0.16	0.61	1.26	0.06	0.02
1.13	0.15	0.46	1.76	0.23	0.04
1.14	0.13	0.43	1.27	0.11	0.12
0.89	0.12	0.58	2.04	0.47	0.11
1.42	0.14	0.05	2.26	0.22	0.01
0.83	0.05	0.26	1.21	0.22	0.45
1.05	0.08	0.58	2.38	0.24	0.01
0.94	0.05	0.46	1.42	0.17	0.09
2.73	0.22	0.00	3.64	0.31	0.00

Table 3.2 Mutant leaf free amino acid levels in prolonged darkness (µmol/mg FW).

Col-0

Amino acid	0d in dark		3d in dark		6d in dark		9d in dark	
	Mean*	SE	Mean	SE	Mean	SE	Mean	SE
Ala	1.37	0.09	2.48	0.13	2.54	0.24	3.17	0.24
Arg	0.03	0.01	2.36	0.14	3.70	0.34	4.14	0.42
Asn	0.12	0.01	1.37	0.06	2.68	0.20	4.18	0.34
Asp	1.55	0.19	0.59	0.07	0.97	0.08	0.83	0.04
Gln	0.88	0.05	0.78	0.04	1.06	0.11	1.53	0.10
Glu	6.20	0.31	6.65	0.43	7.95	0.63	9.05	0.40
Gly	0.27	0.03	0.30	0.04	0.35	0.05	0.36	0.05
His	0.08	0.00	1.42	0.07	2.14	0.13	2.38	0.10
Ile	0.04	0.00	1.44	0.16	1.57	0.16	1.06	0.07
Leu	0.07	0.01	1.35	0.06	1.26	0.13	0.84	0.07
Lys	0.11	0.01	1.35	0.08	2.02	0.19	1.94	0.15
Met	0.02	0.00	0.07	0.01	0.12	0.01	0.15	0.01
Phe	0.12	0.00	2.69	0.12	4.11	0.24	4.29	0.22
Pro	0.17	0.01	0.16	0.02	0.09	0.01	0.09	0.01
Ser	0.41	0.02	0.91	0.07	1.75	0.21	1.98	0.14
Thr	0.37	0.04	0.82	0.10	0.77	0.14	0.57	0.04
Trp	0.02	0.00	0.84	0.03	1.33	0.09	1.54	0.07
Tyr	0.03	0.00	1.00	0.06	1.21	0.13	1.46	0.15
Val	0.16	0.01	3.17	0.17	3.86	0.29	3.48	0.14

* The mean represents the average of five biological replicates (µmol/mg fresh weight).

** Significance by Student's *t*-test, compared to wild types at the same time point: $p<0.05$, **$p<0.01$**, ***$p<0.001$***.

Table 3.2 (cont'd)

bcat2-1

Amino acid	0d in dark			3d in dark			6d in dark			9d in dark		
	Mean	**SE**	***t*-test**[**]	**Mean**	**SE**	***t*-test**	**Mean**	**SE**	***t*-test**	**Mean**	**SE**	***t*-test**
Ala	1.56	0.20	0.412	2.22	0.20	0.283	1.46	0.13	**0.002**	1.66	0.19	***0.000***
Arg	0.05	0.01	0.123	2.51	0.15	0.472	3.89	0.17	0.624	4.04	0.38	0.854
Asn	0.12	0.02	0.641	1.45	0.11	0.496	2.64	0.19	0.872	4.22	0.44	0.943
Asp	1.41	0.16	0.587	1.00	0.10	**0.004**	1.36	0.08	**0.004**	1.26	0.11	**0.003**
Gln	0.97	0.12	0.496	0.78	0.07	0.987	0.88	0.08	0.194	1.46	0.17	0.746
Glu	5.77	0.36	0.373	5.95	0.46	0.282	6.97	0.45	0.222	9.37	0.81	0.725
Gly	0.24	0.05	0.608	0.27	0.04	0.622	0.17	0.03	**0.006**	0.47	0.09	0.291
His	0.12	0.01	0.010	1.46	0.11	0.793	2.08	0.13	0.727	2.60	0.20	0.330
Ile	0.06	0.01	0.108	2.06	0.27	0.069	2.70	0.28	**0.003**	2.55	0.33	**0.001**
Leu	0.14	0.03	0.036	4.59	0.32	***0.000***	6.24	0.30	***0.000***	6.66	0.42	***0.000***
Lys	0.12	0.01	0.153	1.41	0.16	0.714	1.60	0.11	0.069	1.61	0.15	0.134
Met	0.02	0.00	0.991	0.09	0.01	0.167	0.14	0.01	0.153	0.18	0.02	0.210
Phe	0.14	0.01	0.061	2.77	0.22	0.728	3.95	0.21	0.613	4.37	0.26	0.823
Pro	0.18	0.01	0.643	0.23	0.02	0.022	0.10	0.01	0.651	0.11	0.01	0.137
Ser	0.54	0.06	0.086	1.16	0.17	0.222	1.48	0.15	0.311	1.82	0.24	0.586
Thr	0.42	0.06	0.431	0.99	0.14	0.361	0.82	0.11	0.818	0.66	0.08	0.363
Trp	0.03	0.00	**0.008**	0.81	0.02	0.580	1.16	0.07	0.150	1.62	0.10	0.543
Tyr	0.05	0.01	0.056	0.98	0.07	0.858	1.14	0.10	0.679	1.34	0.23	0.676
Val	0.23	0.03	0.042	3.66	0.29	0.162	4.73	0.26	0.037	5.16	0.29	***0.000***

Table 3.2 (cont'd)

e1a1-1

Amino acid	0d in dark			3d in dark			6d in dark			9d in dark		
	Mean	**SE**	***t*-test**	**Mean**	**SE**	***t*-test**	**Mean**	**SE**	***t*-test**	**Mean**	**SE**	***t*-test**
Ala	1.31	0.07	0.607	1.74	0.08	***0.000***	1.02	0.11	***0.000***	0.86	0.08	***0.000***
Arg	0.04	0.01	0.369	2.36	0.22	0.983	3.86	0.40	0.756	4.01	0.20	0.784
Asn	0.11	0.01	0.475	1.43	0.10	0.593	2.92	0.33	0.551	4.00	0.27	0.683
Asp	1.28	0.12	0.260	0.99	0.07	***0.001***	1.58	0.16	**0.005**	1.38	0.07	***0.000***
Gln	0.90	0.07	0.773	0.72	0.06	0.427	0.90	0.10	0.291	1.17	0.06	**0.009**
Glu	5.37	0.35	0.092	5.43	0.13	0.021	6.19	0.55	0.050	7.71	0.56	0.069
Gly	0.23	0.04	0.478	0.23	0.02	0.150	0.22	0.04	0.055	0.42	0.05	0.399
His	0.09	0.00	0.013	1.32	0.07	0.328	2.04	0.21	0.668	2.35	0.09	0.827
Ile	0.03	0.00	0.060	2.26	0.32	0.038	3.29	0.35	***0.001***	3.21	0.48	**0.001**
Leu	0.07	0.01	0.779	5.38	0.21	***0.000***	8.00	0.77	***0.000***	9.41	0.18	***0.000***
Lys	0.11	0.01	0.873	1.31	0.09	0.753	1.39	0.19	0.029	1.13	0.11	***0.000***
Met	0.02	0.00	0.767	0.09	0.00	0.047	0.16	0.02	0.055	0.22	0.02	***0.001***
Phe	0.12	0.00	0.765	2.66	0.10	0.896	4.03	0.34	0.845	4.43	0.12	0.588
Pro	0.17	0.01	0.935	0.17	0.02	0.749	0.09	0.01	0.635	0.11	0.01	0.163
Ser	0.33	0.03	0.044	1.02	0.08	0.330	1.46	0.26	0.397	1.22	0.06	***0.000***
Thr	0.27	0.02	0.043	0.83	0.04	0.938	0.75	0.12	0.892	0.54	0.05	0.662
Trp	0.02	0.00	0.526	0.84	0.04	0.976	1.29	0.08	0.700	1.64	0.09	0.416
Tyr	0.03	0.00	0.392	0.97	0.06	0.772	1.08	0.18	0.549	1.06	0.16	0.084
Val	0.17	0.01	0.277	3.82	0.16	0.013	5.48	0.46	**0.009**	6.08	0.15	***0.000***

Table 3.2 (cont'd)

e1a1-2

Amino acid	0d in dark			3d in dark			6d in dark			9d in dark		
	Mean	**SE**	***t*-test**	**Mean**	**SE**	***t*-test**	**Mean**	**SE**	***t*-test**	**Mean**	**SE**	***t*-test**
Ala	1.39	0.10	0.868	2.01	0.15	0.030	1.34	0.17	***0.001***	0.76	0.10	***0.000***
Arg	0.03	0.01	0.790	2.29	0.23	0.799	4.30	0.23	0.164	3.61	0.48	0.420
Asn	0.12	0.01	0.909	1.68	0.12	0.038	3.50	0.36	0.067	5.43	0.80	0.173
Asp	1.47	0.15	0.753	1.08	0.10	***0.001***	1.61	0.23	0.021	1.53	0.16	**0.002**
Gln	0.99	0.08	0.244	0.86	0.06	0.301	1.12	0.17	0.752	1.55	0.16	0.888
Glu	5.96	0.36	0.617	6.13	0.20	0.291	6.68	0.97	0.290	8.14	0.87	0.362
Gly	0.23	0.02	0.286	0.25	0.04	0.464	0.20	0.04	0.027	0.46	0.09	0.331
His	0.09	0.01	0.126	1.37	0.07	0.602	2.22	0.19	0.729	2.89	0.29	0.128
Ile	0.04	0.00	0.271	2.06	0.33	0.116	1.68	0.16	0.654	3.58	0.81	0.015
Leu	0.07	0.00	0.785	5.18	0.14	***0.000***	8.22	0.66	***0.000***	11.45	1.25	***0.000***
Lys	0.11	0.01	0.707	1.40	0.07	0.632	1.57	0.18	0.101	1.01	0.10	***0.000***
Met	0.03	0.00	0.137	0.09	0.01	0.079	0.18	0.02	**0.004**	0.38	0.08	0.014
Phe	0.12	0.00	0.699	2.66	0.12	0.892	4.36	0.34	0.557	5.45	0.60	0.100
Pro	0.17	0.01	0.856	0.19	0.02	0.325	0.10	0.01	0.500	0.11	0.02	0.270
Ser	0.37	0.03	0.305	1.04	0.07	0.230	1.34	0.25	0.227	1.91	0.38	0.870
Thr	0.35	0.03	0.804	0.86	0.05	0.716	0.61	0.10	0.373	1.08	0.32	0.140
Trp	0.02	0.00	0.393	0.80	0.03	0.510	1.44	0.07	0.338	2.41	0.34	0.037
Tyr	0.03	0.00	0.124	0.97	0.08	0.788	1.27	0.14	0.750	1.44	0.30	0.973
Val	0.17	0.01	0.280	3.84	0.17	0.013	6.24	0.43	***0.000***	7.26	0.49	***0.000***

Table 3.2 (cont'd)

e1a2-1

Amino acid	0d in dark			3d in dark			6d in dark			9d in dark		
	Mean	**SE**	***t*-test**	**Mean**	**SE**	***t*-test**	**Mean**	**SE**	***t*-test**	**Mean**	**SE**	***t*-test**
Ala	1.30	0.11	0.610	3.63	0.40	0.022	3.07	0.32	0.253	3.31	0.28	0.712
Arg	0.04	0.01	0.618	2.55	0.30	0.580	3.90	0.22	0.533	4.29	0.26	0.774
Asn	0.12	0.01	0.966	1.84	0.18	0.029	3.55	0.31	0.032	4.49	0.39	0.558
Asp	1.46	0.16	0.706	1.03	0.12	**0.006**	1.25	0.14	0.059	0.88	0.07	0.538
Gln	0.89	0.08	0.901	1.15	0.14	0.030	1.30	0.15	0.145	1.77	0.25	0.384
Glu	5.80	0.33	0.390	7.92	0.48	0.066	8.70	0.87	0.415	9.36	0.69	0.700
Gly	0.20	0.04	0.182	0.33	0.03	0.569	0.26	0.03	0.106	0.47	0.05	0.101
His	0.09	0.00	0.012	1.57	0.08	0.203	2.32	0.14	0.271	2.61	0.16	0.241
Ile	0.05	0.00	0.319	1.53	0.21	0.751	1.67	0.26	0.594	0.98	0.08	0.491
Leu	0.09	0.01	0.125	1.78	0.18	0.043	1.48	0.16	0.239	1.07	0.08	0.046
Lys	0.11	0.01	0.567	1.77	0.18	0.053	2.26	0.20	0.325	2.37	0.14	0.052
Met	0.02	0.00	0.363	0.09	0.01	0.159	0.15	0.01	0.037	0.17	0.02	0.374
Phe	0.13	0.00	0.103	3.02	0.20	0.176	4.41	0.29	0.299	4.46	0.22	0.584
Pro	0.17	0.01	0.872	0.23	0.03	0.105	0.12	0.01	0.156	0.10	0.01	0.803
Ser	0.50	0.06	0.162	1.81	0.24	**0.005**	2.35	0.28	0.097	2.58	0.29	0.082
Thr	0.38	0.03	0.826	1.53	0.23	0.017	1.07	0.15	0.143	0.82	0.12	0.083
Trp	0.03	0.00	0.083	0.91	0.04	0.153	1.42	0.07	0.616	1.87	0.18	0.111
Tyr	0.03	0.00	0.820	1.17	0.13	0.234	1.41	0.08	0.312	2.14	0.23	0.023
Val	0.18	0.01	0.188	3.64	0.21	0.102	4.47	0.24	0.106	3.97	0.27	0.128

Table 3.2 (cont'd)

e1a1-1;e1a2-1

Amino acid	0d in dark			3d in dark			6d in dark			9d in dark		
	Mean	**SE**	***t*-test**	**Mean**	**SE**	***t*-test**	**Mean**	**SE**	***t*-test**	**Mean**	**SE**	***t*-test**
Ala	1.33	0.13	0.809	2.67	0.21	0.448	1.26	0.08	***0.000***	1.28	0.16	***0.000***
Arg	0.04	0.01	0.570	2.31	0.24	0.853	4.35	0.23	0.066	4.57	0.19	0.372
Asn	0.13	0.02	0.555	1.78	0.15	0.023	3.49	0.31	0.051	4.67	0.28	0.279
Asp	1.67	0.20	0.671	1.34	0.06	***0.000***	1.74	0.14	***0.000***	1.71	0.08	***0.000***
Gln	1.00	0.07	0.191	0.98	0.08	0.040	1.13	0.14	0.520	1.37	0.07	0.209
Glu	6.50	0.29	0.491	7.00	0.25	0.497	6.67	0.54	0.212	8.62	0.52	0.517
Gly	0.21	0.04	0.205	0.29	0.04	0.932	0.23	0.02	0.029	0.23	0.03	0.045
His	0.09	0.01	0.060	1.51	0.06	0.388	2.25	0.16	0.420	2.63	0.12	0.128
Ile	0.05	0.01	0.475	2.40	0.37	0.036	1.63	0.24	0.567	1.76	0.20	**0.007**
Leu	0.09	0.01	0.175	6.08	0.19	***0.000***	8.94	0.67	***0.000***	10.57	0.37	***0.000***
Lys	0.11	0.01	0.532	1.47	0.11	0.374	1.78	0.13	0.454	1.60	0.11	0.084
Met	0.03	0.00	0.527	0.11	0.01	**0.002**	0.18	0.01	**0.001**	0.25	0.02	***0.000***
Phe	0.13	0.01	0.507	2.86	0.12	0.309	4.39	0.27	0.253	4.74	0.16	0.113
Pro	0.18	0.01	0.593	0.27	0.03	**0.004**	0.10	0.01	0.491	0.15	0.01	**0.003**
Ser	0.52	0.09	0.263	1.46	0.12	**0.002**	1.77	0.22	0.786	1.79	0.12	0.313
Thr	0.46	0.08	0.335	1.48	0.18	**0.006**	0.85	0.10	0.509	0.83	0.08	0.014
Trp	0.02	0.00	0.104	0.80	0.02	0.342	1.26	0.07	0.347	1.76	0.12	0.136
Tyr	0.03	0.00	0.722	1.04	0.06	0.657	1.52	0.09	0.085	1.63	0.13	0.398
Val	0.17	0.01	0.213	3.96	0.18	**0.005**	6.25	0.42	***0.000***	6.95	0.16	***0.000***

Table 3.2 (cont'd)

e1a1-2;e1a2-1

Amino acid	0d in dark			3d in dark			6d in dark			9d in dark		
	Mean	**SE**	***t*-test**	**Mean**	**SE**	***t*-test**	**Mean**	**SE**	***t*-test**	**Mean**	**SE**	***t*-test**
Ala	1.57	0.07	0.099	2.82	0.30	0.318	1.41	0.13	**0.001**	1.24	0.14	***0.000***
Arg	0.06	0.01	0.021	2.66	0.33	0.423	4.68	0.33	0.015	4.22	0.15	0.869
Asn	0.12	0.01	0.679	1.96	0.20	0.021	3.79	0.34	0.011	4.40	0.38	0.671
Asp	1.67	0.19	0.682	1.55	0.12	***0.000***	1.99	0.19	***0.000***	1.44	0.10	***0.000***
Gln	1.00	0.05	0.098	1.18	0.17	0.053	1.28	0.15	0.168	1.26	0.10	0.069
Glu	6.39	0.27	0.645	7.25	0.36	0.304	7.41	0.63	0.730	7.47	0.56	0.034
Gly	0.20	0.02	0.087	0.30	0.05	0.977	0.30	0.02	0.348	0.32	0.06	0.688
His	0.11	0.01	**0.003**	1.52	0.08	0.377	2.40	0.18	0.102	2.43	0.06	0.626
Ile	0.04	0.01	0.810	2.17	0.50	0.205	1.41	0.09	0.518	1.60	0.14	**0.006**
Leu	0.12	0.01	0.016	6.48	0.26	***0.000***	9.99	0.69	***0.000***	10.43	0.27	***0.000***
Lys	0.13	0.01	0.148	1.75	0.10	**0.007**	1.92	0.14	0.953	1.24	0.08	**0.001**
Met	0.03	0.00	0.202	0.11	0.01	***0.001***	0.18	0.02	**0.008**	0.23	0.02	**0.001**
Phe	0.13	0.01	0.426	2.97	0.10	0.096	4.53	0.28	0.135	4.64	0.12	0.183
Pro	0.19	0.02	0.249	0.27	0.03	**0.006**	0.11	0.01	0.193	0.11	0.01	0.160
Ser	0.58	0.06	0.018	1.59	0.15	**0.002**	2.05	0.26	0.298	1.38	0.16	0.010
Thr	0.50	0.07	0.137	1.47	0.20	0.015	1.06	0.14	0.140	0.69	0.07	0.189
Trp	0.03	0.00	0.030	0.83	0.03	0.919	1.36	0.07	0.867	1.80	0.11	0.068
Tyr	0.03	0.00	0.207	1.19	0.11	0.139	1.84	0.14	**0.001**	1.37	0.16	0.706
Val	0.21	0.02	0.016	4.31	0.18	***0.000***	6.55	0.34	***0.000***	6.54	0.20	***0.000***

Table 3.2 (cont'd)

e1b2-1

Amino acid	0d in dark			3d in dark			6d in dark			9d in dark		
	Mean	**SE**	***t*-test**	**Mean**	**SE**	***t*-test**	**Mean**	**SE**	***t*-test**	**Mean**	**SE**	***t*-test**
Ala	1.25	0.09	0.352	2.60	0.15	0.554	2.36	0.15	0.580	2.94	0.25	0.503
Arg	0.05	0.01	0.078	2.30	0.16	0.514	3.76	0.09	0.907	3.73	0.31	0.441
Asn	0.10	0.01	0.116	1.36	0.13	0.468	2.45	0.19	0.248	3.76	0.33	0.392
Asp	1.34	0.14	0.376	0.73	0.05	0.066	1.01	0.04	0.559	0.85	0.07	0.833
Gln	0.83	0.08	0.629	0.77	0.07	0.543	0.75	0.04	0.018	1.43	0.16	0.619
Glu	5.77	0.37	0.382	6.78	0.34	0.460	6.65	0.33	0.100	9.77	0.72	0.394
Gly	0.24	0.04	0.600	0.26	0.04	0.486	0.18	0.03	**0.005**	0.43	0.07	0.357
His	0.09	0.00	0.097	1.33	0.06	0.680	1.88	0.07	0.112	2.27	0.16	0.571
Ile	0.04	0.00	0.484	1.29	0.14	0.930	1.64	0.21	0.833	1.46	0.15	0.032
Leu	0.08	0.01	0.461	1.89	0.12	***0.001***	1.78	0.08	**0.003**	1.06	0.08	0.037
Lys	0.11	0.01	0.786	1.23	0.10	0.719	1.65	0.09	0.088	1.71	0.18	0.340
Met	0.02	0.00	0.883	0.07	0.00	0.561	0.10	0.01	0.192	0.13	0.01	0.512
Phe	0.12	0.01	0.722	2.55	0.12	0.638	3.77	0.15	0.292	4.08	0.26	0.545
Pro	0.18	0.01	0.436	0.18	0.02	0.677	0.09	0.01	0.698	0.09	0.01	0.910
Ser	0.35	0.04	0.189	0.93	0.09	0.356	1.25	0.11	0.051	1.79	0.23	0.485
Thr	0.37	0.04	0.952	0.79	0.11	0.710	0.57	0.07	0.265	0.58	0.07	0.878
Trp	0.02	0.00	0.096	0.77	0.03	0.843	1.20	0.06	0.187	1.56	0.08	0.873
Tyr	0.03	0.00	0.111	0.93	0.08	0.907	1.29	0.06	0.573	1.08	0.11	0.051
Val	0.16	0.01	0.658	3.42	0.13	0.136	4.05	0.21	0.609	3.48	0.23	0.999

Table 3.2 (cont'd)

ivd1-2

Amino acid	0d in dark			3d in dark			6d in dark			9d in dark		
	Mean	**SE**	***t*-test**	**Mean**	**SE**	***t*-test**	**Mean**	**SE**	***t*-test**	**Mean**	**SE**	***t*-test**
Ala	1.58	0.14	0.210	0.88	0.07	***0.000***	3.49	0.58	0.012	10.33	2.26	0.011
Arg	0.05	0.01	0.182	2.51	0.21	0.563	5.72	0.52	**0.003**	8.92	0.94	***0.001***
Asn	0.13	0.01	0.594	2.18	0.18	**0.001**	5.00	0.59	**0.006**	12.03	1.22	***0.000***
Asp	1.66	0.15	0.674	1.75	0.18	***0.000***	2.21	0.26	**0.001**	3.91	0.55	***0.000***
Gln	1.05	0.07	0.071	0.73	0.05	0.364	1.02	0.18	0.911	1.91	0.15	0.051
Glu	6.87	0.26	0.115	5.15	0.27	**0.010**	5.67	1.10	0.181	11.30	1.46	0.165
Gly	0.24	0.03	0.485	0.16	0.02	0.012	0.43	0.10	0.379	1.68	0.54	0.038
His	0.11	0.01	**0.003**	1.48	0.12	0.700	5.40	1.04	0.011	14.19	1.46	***0.000***
Ile	0.06	0.01	0.074	1.77	0.18	0.195	1.55	0.30	0.959	2.59	0.48	0.013
Leu	0.12	0.02	0.033	2.72	0.11	***0.000***	3.92	0.81	0.014	9.30	1.33	***0.000***
Lys	0.13	0.01	0.021	1.63	0.14	0.117	1.96	0.30	0.822	3.75	0.64	0.021
Met	0.03	0.00	0.130	0.07	0.00	0.963	0.70	0.19	0.013	1.93	0.28	***0.000***
Phe	0.14	0.01	0.043	2.73	0.19	0.852	10.10	1.93	0.012	26.23	2.54	***0.000***
Pro	0.21	0.01	0.013	0.19	0.01	0.261	0.63	0.16	0.010	2.13	0.35	***0.000***
Ser	0.54	0.05	0.046	1.46	0.21	0.030	3.67	0.71	0.016	11.20	1.48	***0.000***
Thr	0.50	0.07	0.126	0.86	0.08	0.781	1.13	0.21	0.118	3.22	0.56	**0.001**
Trp	0.03	0.00	0.018	0.82	0.03	0.748	3.22	0.62	0.016	10.02	1.15	***0.000***
Tyr	0.04	0.01	0.103	1.13	0.08	0.199	2.93	0.51	**0.004**	7.68	1.27	***0.001***
Val	0.22	0.02	0.022	3.93	0.20	0.011	9.90	1.86	**0.010**	24.92	3.02	***0.000***

Table 3.2 (cont'd)

mcca1-1

Amino acid	0d in dark			3d in dark			6d in dark			9d in dark		
	Mean	**SE**	***t*-test**	**Mean**	**SE**	***t*-test**	**Mean**	**SE**	***t*-test**	**Mean**	**SE**	***t*-test**
Ala	1.31	0.07	0.584	3.07	0.29	0.082	1.36	0.13	**0.001**	0.80	0.11	***0.000***
Arg	0.05	0.01	0.172	2.58	0.24	0.440	4.00	0.16	0.399	4.75	0.22	0.225
Asn	0.12	0.01	0.843	1.74	0.15	0.041	3.13	0.14	0.127	5.14	0.48	0.121
Asp	1.58	0.16	0.904	1.48	0.17	***0.000***	1.73	0.13	***0.000***	1.58	0.11	***0.000***
Gln	0.93	0.04	0.453	0.76	0.06	0.761	0.68	0.04	**0.008**	1.10	0.08	**0.004**
Glu	5.73	0.20	0.224	5.56	0.31	0.056	5.02	0.34	**0.002**	5.92	0.38	***0.000***
Gly	0.19	0.03	0.067	0.25	0.03	0.390	0.30	0.03	0.255	0.36	0.07	0.955
His	0.09	0.01	0.088	1.52	0.14	0.552	2.16	0.07	0.868	2.88	0.10	**0.002**
Ile	0.04	0.00	0.323	2.48	0.28	**0.007**	2.99	0.42	**0.007**	3.13	0.59	**0.006**
Leu	0.09	0.02	0.382	5.71	0.39	***0.000***	8.20	0.33	***0.000***	10.61	0.32	***0.000***
Lys	0.11	0.01	0.800	1.26	0.13	0.561	1.24	0.07	**0.003**	0.87	0.07	***0.000***
Met	0.03	0.00	0.330	0.10	0.01	0.015	0.18	0.01	**0.003**	0.30	0.02	***0.000***
Phe	0.12	0.00	0.861	2.83	0.18	0.512	4.35	0.18	0.336	5.33	0.25	**0.006**
Pro	0.18	0.02	0.615	0.19	0.02	0.341	0.07	0.01	0.156	0.11	0.01	0.115
Ser	0.38	0.03	0.346	1.33	0.17	0.042	1.64	0.16	0.801	1.54	0.13	0.035
Thr	0.36	0.05	0.967	0.83	0.08	0.947	0.50	0.08	0.129	0.55	0.06	0.763
Trp	0.02	0.00	0.783	0.81	0.03	0.533	1.38	0.08	0.804	1.91	0.10	**0.009**
Tyr	0.03	0.00	0.959	1.01	0.07	0.890	1.21	0.10	0.902	1.17	0.15	0.179
Val	0.18	0.02	0.258	4.26	0.31	**0.008**	5.92	0.24	***0.000***	7.50	0.27	***0.000***

Table 3.2 (cont'd)

mccb1-1

Amino acid	0d in dark			3d in dark			6d in dark			9d in dark		
	Mean	**SE**	***t*-test**	**Mean**	**SE**	***t*-test**	**Mean**	**SE**	***t*-test**	**Mean**	**SE**	***t*-test**
Ala	1.40	0.10	0.851	3.05	0.27	0.073	1.62	0.10	**0.005**	1.04	0.07	***0.000***
Arg	0.04	0.00	0.154	2.73	0.15	0.090	4.68	0.34	0.032	5.02	0.17	0.076
Asn	0.15	0.01	0.095	1.70	0.08	**0.006**	3.35	0.23	0.027	5.04	0.30	0.075
Asp	1.66	0.18	0.702	1.39	0.13	***0.000***	1.82	0.10	***0.000***	1.75	0.09	***0.000***
Gln	1.16	0.08	0.011	0.76	0.05	0.689	0.86	0.07	0.165	1.22	0.07	0.024
Glu	6.68	0.29	0.275	5.49	0.21	0.031	5.72	0.46	0.014	7.50	0.46	0.021
Gly	0.22	0.04	0.345	0.23	0.03	0.209	0.27	0.04	0.126	0.30	0.05	0.366
His	0.11	0.01	***0.000***	1.50	0.09	0.497	2.35	0.13	0.213	2.82	0.13	0.013
Ile	0.03	0.01	0.299	2.02	0.30	0.112	2.67	0.40	0.048	2.39	0.48	0.021
Leu	0.08	0.01	0.320	5.83	0.32	***0.000***	9.06	0.54	***0.000***	10.06	0.26	***0.000***
Lys	0.12	0.01	0.315	1.33	0.09	0.916	1.59	0.09	0.093	1.25	0.05	**0.001**
Met	0.03	0.00	0.017	0.10	0.01	***0.000***	0.19	0.02	**0.001**	0.28	0.01	***0.000***
Phe	0.13	0.01	0.450	2.89	0.13	0.270	4.59	0.26	0.117	5.13	0.13	**0.005**
Pro	0.23	0.02	**0.007**	0.20	0.02	0.230	0.08	0.00	0.302	0.11	0.01	0.124
Ser	0.42	0.03	0.847	1.26	0.09	**0.007**	1.91	0.21	0.542	2.02	0.16	0.862
Thr	0.37	0.04	0.973	0.82	0.05	0.992	0.72	0.13	0.904	0.67	0.09	0.339
Trp	0.02	0.00	0.012	0.87	0.04	0.499	1.43	0.09	0.635	1.95	0.10	**0.005**
Tyr	0.03	0.00	0.495	1.14	0.07	0.137	1.57	0.16	0.086	2.02	0.31	0.124
Val	0.18	0.01	0.091	4.07	0.25	**0.008**	6.66	0.35	***0.000***	7.89	0.35	***0.000***

Table 3.2 (cont'd)

hml1-2

Amino acid	0d in dark			3d in dark			6d in dark			9d in dark		
	Mean	**SE**	***t*-test**	**Mean**	**SE**	***t*-test**	**Mean**	**SE**	***t*-test**	**Mean**	**SE**	***t*-test**
Ala	1.47	0.09	0.409	2.65	0.17	0.412	1.30	0.09	***0.001***	0.75	0.10	***0.000***
Arg	0.07	0.01	***0.001***	2.55	0.13	0.346	3.85	0.20	0.848	5.63	0.56	0.049
Asn	0.12	0.01	0.569	1.31	0.10	0.657	2.38	0.22	0.174	4.65	0.34	0.346
Asp	1.46	0.11	0.679	1.38	0.08	***0.000***	1.69	0.08	***0.000***	1.70	0.29	0.015
Gln	0.83	0.07	0.639	0.59	0.07	0.027	0.47	0.03	***0.000***	1.39	0.22	0.590
Glu	5.20	0.30	0.032	4.62	0.22	**0.001**	4.21	0.18	***0.000***	6.75	0.69	0.015
Gly	0.19	0.03	0.076	0.19	0.03	0.047	0.14	0.03	**0.002**	0.22	0.04	0.032
His	0.10	0.01	0.076	1.38	0.07	0.693	1.90	0.09	0.161	3.61	0.67	0.105
Ile	0.04	0.01	0.962	2.03	0.26	0.074	1.68	0.15	0.614	2.93	0.62	0.016
Leu	0.12	0.02	0.031	5.31	0.23	***0.000***	7.51	0.32	***0.000***	10.67	0.59	***0.000***
Lys	0.13	0.01	0.100	1.09	0.09	0.056	1.19	0.08	**0.001**	1.31	0.15	**0.007**
Met	0.03	0.01	0.329	0.09	0.00	0.011	0.15	0.01	**0.008**	0.33	0.04	**0.002**
Phe	0.13	0.01	0.330	2.62	0.10	0.701	3.89	0.16	0.464	5.36	0.36	0.025
Pro	0.18	0.01	0.649	0.14	0.01	0.243	0.07	0.00	0.047	0.12	0.01	0.142
Ser	0.45	0.03	0.368	1.12	0.10	0.101	1.32	0.09	0.089	1.99	0.27	0.984
Thr	0.36	0.03	0.830	0.84	0.14	0.913	0.37	0.03	0.026	0.79	0.14	0.171
Trp	0.03	0.00	0.034	0.85	0.05	0.823	1.37	0.07	0.820	2.59	0.42	0.037
Tyr	0.04	0.00	0.171	0.96	0.13	0.815	1.47	0.11	0.190	2.04	0.23	0.056
Val	0.20	0.02	0.051	3.74	0.14	0.016	5.43	0.25	**0.001**	9.09	1.48	**0.005**

Table 3.3 Leaf amino acid profiles of *ivd1-2* relative to wild type (Col-0) at 6d and 9d in prolonged darkness.

Amino acid	6d in dark			9d in dark		
	Mean*	SE	*t*-test	Mean	SE	*t*-test
Ala	1.37	0.23	0.012	3.26	0.71	0.011
Arg	1.55	0.14	0.003	2.15	0.23	0.001
Asn	1.86	0.22	0.006	2.88	0.29	0.000
Asp	2.27	0.27	0.001	4.70	0.66	0.000
Gln	0.96	0.17	0.911	1.25	0.10	0.051
Glu	0.71	0.14	0.181	1.25	0.16	0.165
Gly	1.20	0.29	0.379	4.71	1.52	0.038
His	2.52	0.48	0.011	5.97	0.62	0.000
Ile	0.99	0.19	0.959	2.46	0.46	0.013
Leu	3.11	0.64	0.014	11.10	1.59	0.000
Lys	0.97	0.15	0.822	1.93	0.33	0.021
Met	5.92	1.58	0.013	13.17	1.92	0.000
Phe	2.45	0.47	0.012	6.11	0.59	0.000
Pro	6.80	1.77	0.010	23.07	3.79	0.000
Ser	2.09	0.40	0.016	5.66	0.75	0.000
Thr	1.46	0.27	0.118	5.66	0.99	0.001
Trp	2.41	0.46	0.016	6.49	0.75	0.000
Tyr	2.42	0.42	0.004	5.28	0.87	0.001
Val	2.57	0.48	0.010	7.16	0.87	0.000

* The mean represents the fold change between the averages of five biological replicates of *ivd1-2* and Col-0 at the same time point.

Table 3.4 Primers for amiRNA generation.

Primer	Sequence	Note
amiRNA_A_cacc	CACCCTGCAAGGCGATTAAGTTGGGTAAC	amiRNA_a primer with cacc at 5', for amiRNA inserts cloning into pENTR-d/topo
amiRNA_B	GCGGATAACAATTTCACACAGGAAACAG	amiRNA_B, for amiRNA inserts cloning into pENTR-d/topo
amiE1B1-1_I	gaTATGCGATTACATTAGTCCTTtctctcttttgtattcc	I miR-s primer designed for amiE1B1-1
amiE1B1-1_II	gaAAGGACTAATGTAATCGCATAtcaaagagaatcaatga	II miR-a primer designed for amiE1B1-1
amiE1B1-1_III	gaAAAGACTAATGTATTCGCATTtcacaggtcgtgatatg	III miR*s primer designed for amiE1B1-1
amiE1B1-1_IV	gaAATGCGAATACATTAGTCTTTtctacatatatattcct	IV miR*a primer designed for amiE1B1-1
amiE1B1-2_I	gaTAACTACAGATAGTACGCCTAtctctcttttgtattcc	I miR-s primer designed for amiE1B1-2
amiE1B1-2_II	gaTAGGCGTACTATCTGTAGTTAtcaaagagaatcaatga	II miR-a primer designed for amiE1B1-2
amiE1B1-2_III	gaTAAGCGTACTATCAGTAGTTTtcacaggtcgtgatatg	III miR*s primer designed for amiE1B1-2
amiE1B1-2_IV	gaAAACTACTGATAGTACGCTTAtctacatatatattcct	IV miR*a primer designed for amiE1B1-2
pEARLEY-100_For	CATCGTGGAAAAAGAAGACGT	Forward primer for amplifying/genotyping insert on pEarleyGate100 vector
pEARLEY-100_Rev	AGGATCTGAGCTACACATGCT	Reverse primer for amplifying/genotyping insert on pEarleyGate100 vector

LITERATURE CITED

LITERATURE CITED

Anderson MD, Che P, Song J, Nikolau BJ, Wurtele ES (1998) 3-Methylcrotonyl-coenzyme A carboxylase is a component of the mitochondrial leucine catabolic pathway in plants. Plant Physiol **118:** 1127-1138

Angelovici R, Fait A, Fernie AR, Galili G (2011) A seed high-lysine trait is negatively associated with the TCA cycle and slows down Arabidopsis seed germination. New Phytol **189:** 148-159

Angelovici R, Lipka AE, Deason N, Gonzalez-Jorge S, Lin H, Cepela J, Buell R, Gore MA, Dellapenna D (2013) Genome-wide analysis of branched-chain amino acid levels in *Arabidopsis* seeds. Plant Cell **25:** 4827-4843

Araujo WL, Ishizaki K, Nunes-Nesi A, Larson TR, Tohge T, Krahnert I, Witt S, Obata T, Schauer N, Graham IA, Leaver CJ, Fernie AR (2010) Identification of the 2-hydroxyglutarate and isovaleryl-CoA dehydrogenases as alternative electron donors linking lysine catabolism to the electron transport chain of *Arabidopsis* mitochondria. Plant Cell **22:** 1549-1563

Attaran E, Major IT, Cruz JA, Rosa BA, Koo AJ, Chen J, Kramer DM, He SY, Howe GA (2014) Temporal dynamics of growth and photosynthesis suppression in response to jasmonate signaling. Plant Physiol **165:** 1302-1314

Binder S (2010) Branched-chain amino acid metabolism in *Arabidopsis thaliana*. Arabidopsis Book **8:** e0137

Che P, Wurtele ES, Nikolau BJ (2002) Metabolic and environmental regulation of 3-methylcrotonyl-coenzyme A carboxylase expression in Arabidopsis. Plant Physiol **129:** 625-637

Ding G, Che P, Ilarslan H, Wurtele ES, Nikolau BJ (2012) Genetic dissection of methylcrotonyl CoA carboxylase indicates a complex role for mitochondrial leucine catabolism during seed development and germination. Plant J **70:** 562-577

Earley KW, Haag JR, Pontes O, Opper K, Juehne T, Song K, Pikaard CS (2006) Gateway-compatible vectors for plant functional genomics and proteomics. Plant J **45:** 616-629

Espinoza C, Degenkolbe T, Caldana C, Zuther E, Leisse A, Willmitzer L, Hincha DK, Hannah MA (2010) Interaction with diurnal and circadian regulation results in dynamic metabolic and transcriptional changes during cold acclimation in Arabidopsis. PLoS One **5:** e14101

Fujiki Y, Sato T, Ito M, Watanabe A (2000) Isolation and characterization of cDNA clones for the E1β and E2 subunits of the branched-chain α-ketoacid dehydrogenase complex in *Arabidopsis*. J Biol Chem **275:** 6007-6013

Gibon Y, Blaesing OE, Hannemann J, Carillo P, Hohne M, Hendriks JH, Palacios N, Cross J, Selbig J, Stitt M (2004) A Robot-based platform to measure multiple enzyme activities in Arabidopsis using a set of cycling assays: comparison of changes of enzyme activities and transcript levels during diurnal cycles and in prolonged darkness. Plant Cell **16:** 3304-3325

Gibon Y, Usadel B, Blaesing OE, Kamlage B, Hoehne M, Trethewey R, Stitt M (2006) Integration of metabolite with transcript and enzyme activity profiling during diurnal cycles in *Arabidopsis* rosettes. Genome Biol **7:** R76

Gu L, Jones AD, Last R (2012) Rapid LC–MS/MS profiling of protein amino acids and metabolically related compounds for large-scale assessment of metabolic phenotypes. *In* MA Alterman, P Hunziker, eds, Amino Acid Analysis, Vol 828. Humana Press, pp 1-11

Gu L, Jones AD, Last RL (2007) LC-MS/MS assay for protein amino acids and metabolically related compounds for large-scale screening of metabolic phenotypes. Anal Chem **79:** 8067-8075

Gu LP, Jones AD, Last RL (2010) Broad connections in the Arabidopsis seed metabolic network revealed by metabolite profiling of an amino acid catabolism mutant. Plant J **61:** 579-590

Hanaoka H, Noda T, Shirano Y, Kato T, Hayashi H, Shibata D, Tabata S, Ohsumi Y (2002) Leaf senescence and starvation-induced chlorosis are accelerated by the disruption of an Arabidopsis autophagy gene. Plant Physiol **129:** 1181-1193

Harper AE, Miller RH, Block KP (1984) Branched-chain amino acid metabolism. Annu Rev Nutr **4:** 409-454

Ishizaki K, Larson TR, Schauer N, Fernie AR, Graham IA, Leaver CJ (2005) The critical role of *Arabidopsis* electron-transfer flavoprotein:ubiquinone oxidoreductase during dark-induced starvation. Plant Cell **17:** 2587-2600

Ishizaki K, Schauer N, Larson TR, Graham IA, Fernie AR, Leaver CJ (2006) The mitochondrial electron transfer flavoprotein complex is essential for survival of Arabidopsis in extended darkness. Plant J **47:** 751-760

Izumi M, Wada S, Makino A, Ishida H (2010) The autophagic degradation of chloroplasts via rubisco-containing bodies is specifically linked to leaf carbon status but not nitrogen status in Arabidopsis. Plant Physiol **154:** 1196-1209

Kochevenko A, Araujo WL, Maloney GS, Tieman DM, Do PT, Taylor MG, Klee HJ, Fernie AR (2012) Catabolism of branched chain amino acids supports respiration but not volatile synthesis in tomato fruits. Mol Plant **5:** 366-375

Liu Y, Bassham DC (2012) Autophagy: pathways for self-eating in plant cells. Annu Rev Plant Biol **63:** 215-237

Lu Y, Savage LJ, Ajjawi I, Imre KM, Yoder DW, Benning C, Dellapenna D, Ohlrogge JB, Osteryoung KW, Weber AP, Wilkerson CG, Last RL (2008) New connections across pathways and cellular processes: industrialized mutant screening reveals novel associations between diverse phenotypes in Arabidopsis. Plant Physiol **146:** 1482-1500

Lu Y, Savage LJ, Larson MD, Wilkerson CG, Last RL (2011) Chloroplast 2010: a database for large-scale phenotypic screening of Arabidopsis mutants. Plant Physiol **155:** 1589-1600

Mentzen WI, Peng J, Ransom N, Nikolau BJ, Wurtele ES (2008) Articulation of three core metabolic processes in Arabidopsis: fatty acid biosynthesis, leucine catabolism and starch metabolism. BMC Plant Biol **8:** 76

Mikkelsen MD, Halkier BA (2003) Metabolic engineering of valine- and isoleucine-derived glucosinolates in Arabidopsis expressing CYP79D2 from Cassava. Plant Physiol **131:** 773-779

Mooney BP, Henzl MT, Miernyk JA, Randall DD (2000) The dihydrolipoyl acyltransferase (BCE2) subunit of the plant branched-chain α-ketoacid dehydrogenase complex forms a 24-mer core with octagonal symmetry. Protein Sci **9:** 1334-1339

Oh SA, Park JH, Lee GI, Paek KH, Park SK, Nam HG (1997) Identification of three genetic loci controlling leaf senescence in *Arabidopsis thaliana*. Plant J **12:** 527-535

Slocombe SP, Schauvinhold I, McQuinn RP, Besser K, Welsby NA, Harper A, Aziz N, Li Y, Larson TR, Giovannoni J, Dixon RA, Broun P (2008) Transcriptomic and reverse genetic analyses of branched-chain fatty acid and acyl sugar production in *Solanum pennellii* and *Nicotiana benthamiana*. Plant Physiol **148:** 1830-1846

Taylor NL, Heazlewood JL, Day DA, Millar AH (2004) Lipoic acid-dependent oxidative catabolism of α-keto acids in mitochondria provides evidence for branched-chain amino acid catabolism in Arabidopsis. Plant Physiol **134:** 838-848

Wada S, Ishida H, Izumi M, Yoshimoto K, Ohsumi Y, Mae T, Makino A (2009) Autophagy plays a role in chloroplast degradation during senescence in individually darkened leaves. Plant Physiol **149:** 885-893

Chapter 4

Conclusions and future perspectives

4.1 Conclusions

The goal of this dissertation research was to further our understanding of the architecture of BCAA catabolism and its regulation in *A. thaliana*. A systems approach combining transcript and functional analyses was used in this dissertation research, to study the subunits encoding the large and biochemically challenging BCKDH complex, and to explore the role of BCAA catabolism within and beyond amino acid metabolism. I demonstrated that eight BCAA catabolic enzyme genes - *BCAT2*, *E1A1*, *E1B1*, *E1B2*, *E2*, *IVD1*, *MCCA1* and *MCCB1* - are highly coexpressed in stress, light, and diurnal/circadian conditions, and their expression is subjected to the regulation of the circadian clock and light. Promoter mutagenesis experiments demonstrated that both the -100bp promoter region and 5'UTR are necessary for the diurnal oscillation of the *IVD1* transcript. In addition, mutant characterization provides evidence that putative BCKDH enzyme subunits E1α1, E1α2, E1β1 and E1β2 participate in BCAA catabolism during normal growth conditions. Mutants deficient in BCAA catabolism undergo early senescence and accumulate leaf free BCAAs during prolonged darkness, supporting the hypothesized connection of BCAA catabolism with the plant energy metabolism.

Taken together, my dissertation research provides insights into the regulation and physiological roles of BCAA catabolism, and explored the physiological role(s) of BCAA catabolism within and beyond amino acid metabolism during normal growth conditions and under energy-limited conditions.

4.2 Future perspectives

4.2.1 Exploration of the physiological role(s) of BCAA catabolism in stress conditions

In my dissertation research, I performed transcript coexpression analyses, and found that 10 out of the 13 genes proposed or experimentally validated to be involved in BCAA catabolism are positively coexpressed in the stress dataset (Figure 2.1, Figure 2.2). Because little information is known regarding to changes in BCAA contents and the physiological function(s) of BCAA catabolism upon abiotic and biotic stress treatments, it would be interesting to further explore the expression profiles of the 10 significantly coexpressed genes to find out the treatments (stimuli) that contribute to the transcript coexpression, and to test if the gene expression responses reflect the physiological function(s). To further study the role of BCAA catabolism in the identified stress treatments, BCAA catabolic mutants could be treated with the identified stimuli, both morphological phenotypes and changes in BCAA contents could be monitored and compared to the wild type. These experiments would further our understanding in the physiological role(s) of BCAA catabolism upon particular stress treatments, and help reveal the complex inter-connections within the plant metabolic networks.

4.2.2 Follow-up of the TFs interacting with the *IVD1* promoter and 5'UTR by Y1H

In my dissertation research, I demonstrated that the *IVD1* -100bp promoter region with 5'UTR is necessary for *IVD1* transcript diurnal oscillation (Figure 2.6). Y1H screen presented in Chapter 2 Appendix 1 demonstrated that a total of 72 TFs were able to bind with high confidence to fragment #1, which contains 180bp *IVD1* promoter region upstream of TSS and its full 5'UTR (Table 2.7). In an earlier coexpression analysis that was collaborated with Sahra Uygun, we used

three algorithms - k-mean clustering, pair-wise Pearson Correlation Coefficient and mutual ranking from the ATTED-II database (http://atted.jp/) (Obayashi et al., 2009; Obayashi et al., 2011), to compute a list of genes that coexpress with the experimentally validated BCAA catabolic enzyme genes *BCAT2*, *IVD1*, *MCCA1*, *MCCB1* and *HML1*. A scoring system was developed to present the times that a gene showed coexpression with any one of the five bait genes using individual algorithm. A total of 378 genes were found to coexpress with the five bait genes for three or more than three times. Among these 72 TFs identified from Y1H, 10 TFs (Table 4.1) were included in the coexpression gene list. Electrophoretic mobility shift assay could be performed to obtain a second line of evidence demonstrating the direct binding between the TF candidates and the *IVD1* promoter. To further evaluate the *in vivo* role of the 10 TFs in regulating BCAA catabolism gene transcripts, homozygous T-DNA insertion lines and overexpression lines of these TF candidates should be obtained. BCAA catabolism gene transcripts and leaf BCAA contents could be examined, to compare the changes between the TF overexpression or mutant lines and the wild type under day-night cycles. In addition, my results also showed that the clock component LHY and photoreceptor PhyB participate in the regulation of BCAA catabolism genes in short day. It would be interesting to see if the transcript levels of the identified TF candidates are altered in the *LHYOX* (*LHY* overexpressor) and *phyB-9* lines, and correlate the transcript changes of these TFs to BCAA catabolism genes. Taken together, these experiments could lead to the identification and characterization of TFs regulating BCAA catabolism genes, and would provide insights on how the circadian clock and light regulate primary metabolism in plants.

4.2.3 Evaluation of the effect of *IVD1* introns on *IVD1* transcript accumulation in prolonged darkness

My analyses suggest that the *IVD1* promoter and its 5'UTR were not sufficient for the elevated *IVD1* transcript in prolonged darkness. One possibility could be that one or more introns are required for transcript accumulation (Callis et al., 1987; Luehrsen and Walbot, 1991; Norris et al., 1993; Xu et al., 1994; Rose and Last, 1997; Rose, 2008). To test this hypothesis, translational fusions of the *IVD1* promoter with 5'UTR and different combinations of its transit peptide and/or introns, together with the *LUC* reporter gene could be generated and transformed into *A. thaliana* Col-0 wild type. Bioluminescence from primary transformants could be examined in prolonged darkness to evaluate the effect of introns on *IVD1* transcript accumulation. These experiments would provide insights into the molecular basis of the regulation of BCAA catabolism.

4.2.4 Exploration of the interaction between the amino acid metabolic network and energy metabolism

I identified 53 out of 481 amino acid metabolism genes that showed elevated transcript levels in prolonged darkness from my RNA-Seq experiments, including genes in BCAA catabolism, Lys catabolism, Gly biosynthesis and so on (Table A.2). This suggests that more amino acid metabolic pathways serve function(s) in the dark. To further explore their physiological function in prolonged darkness, homozygous T-DNA mutants defective in the genes shown in Table A.2 could be obtained, and morphological phenotypes, amino acid contents and TCA intermediates should be monitored and measured during prolonged darkness. These experiments would provide insights into the physiological roles of amino acid metabolic

pathways beyond maintaining amino acid homeostasis, and further our understanding of the connection between amino acid metabolism and energy metabolism in plants.

4.3 Practical implications

The presented dissertation research has a variety of implications for our knowledge of plant metabolism, and provides insights into plant metabolic engineering of productivity and nutritional quality. In this dissertation research, amino acid increases - including all three BCAAs and multiple essential amino acids that are biosynthetically unrelated - were observed in seeds of BCAA catabolic mutants (Figure 3.6A). These results suggest uncharacterized inter-connections within the plant amino acid metabolic networks. Previously published studies are consistent with this idea, and showed inter-connections regulating levels of six of the nine free essential amino acids (Ile, Leu, Lys, Met, Thr and Val) in seeds of a variety of plants (Karchi et al., 1994; Zhu and Galili, 2003; Jander et al., 2004; Joshi et al., 2006; Lee et al., 2008; Gu et al., 2010; Lu et al., 2011; Angelovici et al., 2013). Together, these results and my dissertation research indicate that modifying catabolism has potential for improving the nutritional quality of crop seeds and vegetative tissues.

However, a variety of pleiotropic effects beyond amino acid metabolism were demonstrated, such as defects in vegetative and reproductive development or seed viability (Zhu and Galili, 2004; Lee et al., 2008; Ding et al., 2012), and early senescence in prolonged darkness in this study (Figure 3.3, Figure 3.4). While prolonged darkness is a rather artificial condition, the extent of the interaction between BCAA catabolism and energy metabolism during day-night cycles should be investigated in the future.

It is not surprising that primary metabolic processes - which are relatively ‘old’, compared to specialized metabolism - are broadly connected and difficult to perturb without resultant undesirable changes (Milo and Last, 2012). While these pleiotropic phenotypes create

difficulties for metabolic engineering of productivity and nutritional quality, knowing the basis for these syndromes should lead to a deeper understanding of the architecture and regulation of the plant metabolic networks.

APPENDIX

Table 4.1 List of TFs that both interact with *IVD1* -180/+200bp region and coexpress with BCAA catabolism genes

AGI	TF family
AT1G21000	PLATZ
AT1G64620	C2C2-DOF
AT1G76900	TUB
AT1G78080	AP2-EREBP
AT3G10030	TRIHELIX
AT3G16770	AP2-EREBP
AT3G53200	MYB
AT4G18020	ARR-B/G2-like
AT4G39100	PHD
AT5G61590	AP2-EREBP

LITERATURE CITED

LITERATURE CITED

Angelovici R, Lipka AE, Deason N, Gonzalez-Jorge S, Lin H, Cepela J, Buell R, Gore MA, Dellapenna D (2013) Genome-wide analysis of branched-chain amino acid levels in *Arabidopsis* seeds. Plant Cell **25:** 4827-4843

Callis J, Fromm M, Walbot V (1987) Introns increase gene expression in cultured maize cells. Genes Dev **1:** 1183-1200

Ding G, Che P, Ilarslan H, Wurtele ES, Nikolau BJ (2012) Genetic dissection of methylcrotonyl CoA carboxylase indicates a complex role for mitochondrial leucine catabolism during seed development and germination. Plant J **70:** 562-577

Gu LP, Jones AD, Last RL (2010) Broad connections in the Arabidopsis seed metabolic network revealed by metabolite profiling of an amino acid catabolism mutant. Plant J **61:** 579-590

Jander G, Norris SR, Joshi V, Fraga M, Rugg A, Yu S, Li L, Last RL (2004) Application of a high-throughput HPLC-MS/MS assay to *Arabidopsis* mutant screening; evidence that threonine aldolase plays a role in seed nutritional quality. Plant J **39:** 465-475

Joshi V, Laubengayer KM, Schauer N, Fernie AR, Jander G (2006) Two *Arabidopsis* threonine aldolases are nonredundant and compete with threonine deaminase for a common substrate pool. Plant Cell **18:** 3564-3575

Karchi H, Shaul O, Galili G (1994) Lysine synthesis and catabolism are coordinately regulated during tobacco seed development. Proc Natl Acad Sci U S A **91:** 2577-2581

Lee M, Huang T, Toro-Ramos T, Fraga M, Last RL, Jander G (2008) Reduced activity of *Arabidopsis thaliana* HMT2, a methionine biosynthetic enzyme, increases seed methionine content. Plant J **54:** 310-320

Lu Y, Savage LJ, Larson MD, Wilkerson CG, Last RL (2011) Chloroplast 2010: a database for large-scale phenotypic screening of Arabidopsis mutants. Plant Physiol **155:** 1589-1600

Luehrsen KR, Walbot V (1991) Intron enhancement of gene-expressionand the splicing efficiency of introns in maize cells. Mol Gen Genet **225:** 81-93

Milo R, Last RL (2012) Achieving diversity in the face of constraints: lessons from metabolism. Science **336:** 1663-1667

Norris SR, Meyer SE, Callis J (1993) The intron of *Arabidopsis thaliana* polyubiquitin genes is conserved in location and is a quantitative determinant of chimeric gene expression. Plant Mol Biol **21:** 895-906

Obayashi T, Hayashi S, Saeki M, Ohta H, Kinoshita K (2009) ATTED-II provides coexpressed gene networks for Arabidopsis. Nucleic Acids Res **37:** D987-991

Obayashi T, Nishida K, Kasahara K, Kinoshita K (2011) ATTED-II updates: condition-specific gene coexpression to extend coexpression analyses and applications to a broad range of flowering plants. Plant Cell Physiol **52:** 213-219

Rose AB (2008) Intron-mediated regulation of gene expression. Curr Top Microbiol Immunol **326:** 277-290

Rose AB, Last RL (1997) Introns act post-transcriptionally to increase expression of the *Arabidopsis thaliana* tryptophan pathway gene *PAT1*. Plant J **11:** 455-464

Xu Y, Yu H, Hall TC (1994) Rice triosephosphate isomerase gene 5' sequence directs β-Glucuronidase activity in transgenic tobacco but requires an intron for expression in Rice. Plant Physiol **106:** 459-467

Zhu X, Galili G (2003) Increased lysine synthesis coupled with a knockout of its catabolism synergistically boosts lysine content and also transregulates the metabolism of other amino acids in Arabidopsis seeds. Plant Cell **15:** 845-853

Zhu X, Galili G (2004) Lysine metabolism is concurrently regulated by synthesis and catabolism in both reproductive and vegetative tissues. Plant Physiol **135:** 129-136