Seed Biology: A Proteomics View

Von der Naturwissenschaftlichen Fakultät der
Gottfried Wilhelm Leibniz Universität Hannover

zur Erlangung des Grades
Doktorin der Naturwissenschaften
(Dr. rer. nat.)

genehmigte Dissertation
von
M. Sc. Christin Lorenz, geb. Haase
geboren am 12.01.1987 in Schwerin

2015

Referent: Prof. Dr. rer. nat. Hans-Peter Braun

Korreferent: Prof. Dr. rer. nat. Udo-Klaus Schmitz

Korreferent: Dr. rer. nat. habil. Hardy Rolletschek

Tag der Promotion: 17.12.2015

The following publications contributed to this thesis

1. **Lorenz C.**, Rolletschek H., Sunderhaus S., Braun H.P. (2014) *B. napus* seed endosperm – metabolism and signaling in a dead end tissue. Journal of Proteomics 108: 382-426.

2. **Lorenz C.**, Borisjuk L., Rolletschek H., Heinzel N., Tohge T., Fernie A.R., Braun H.P., Hildebrandt T.M. (*in preparation*) Amino acid metabolism and the role of ETHE1 in *Arabidopsis thaliana* seeds.

3. Höfler S., **Lorenz C.**, Busch T., Brinkkötter M., Tohge T., Fernie A.R., Braun H.P., Hildebrandt T.M. (*in preparation*) Dealing with the sulfur part of cysteine: four enzymatic steps degrade L-cysteine to pyruvate and thiosulfate in *Arabidopsis* mitochondria preventing toxic effects of persulfides.

4. Nietzel T., Dudkina N.V., **Haase C.**, Denolf P., Semchonok D., Boekema E.J., Braun H.P., Sunderhaus S. (2013) The native structure and composition of the cruciferin complex in *Brassica napus*. Journal of Biological Chemistry (288): 2238-2245.

5. Mwangi J.W., Rode C., Colditz F., **Haase C.**, Braun H.P., Winkelmann T. (2013) Proteomic and histological analyses of endosperm development in *Cyclamen persicum* as a basis for optimization of somatic embryogenesis. Plant Science 201(20): 52-65.

6. Noah A.M., Niemenak N., Sunderhaus S., **Haase C.**, Omokolo D.N., Winkelmann T., Braun H.P. (2013) Comparative proteomic analysis of early somatic and zygotic embryogenesis in *Theobroma cacao* L.. Journal of Proteomics 78: 123-133.

7. Fromm S., Göing J., **Lorenz C.**, Peterhänsel C., Braun H.P. (2016) Depletion of the "gamma-type carbonic anhydrase-like" subunits of complex I specifically affects central mitochondrial metabolism in *Arabidopsis thaliana*. Biochimica et Biophysica Acta 1857: 60-71.

8. Petereit J., Katayama K., **Lorenz C.**, Schertl P., Kitsche A., Wada H., Frentzen M., Braun H.P., Eubel H. (*in preparation*) Cardiolipin deficiency has pleiotropic effects on the plant mitochondrial electron chain.

Zusammenfassung

Die Samenentwicklung von Pflanzen wird durch den einzigartigen Mechanismus der doppelten Befruchtung initiiert. Die daraus entstehenden Samengewebe, Endosperm und Embryo, durchlaufen während ihrer Entwicklung charakteristische Stadien von der Gewebedifferenzierung bis hin zur Einlagerung von Speicherstoffen. Dieser Prozess ist durch vielfältige Veränderungen auf zellulärer und molekularer Ebene gekennzeichnet. In der vorliegenden Arbeit wurden die physiologischen und molekularen Mechanismen während der Samenentwicklung und die Rolle der verschiedenen Samengewebe untersucht (Kapitel 2.1). Im Rahmen dieser Untersuchung wurde eine proteomische Strategie entwickelt, welche (i) die Etablierung einer Endosperm-Referenzkarte, (ii) die Charakterisierung der Endosperm-entwicklung und (iii) den Vergleich des Endosperm- und Embryoproteoms umfasste. In dieser ersten Untersuchung zur Endosperm-entwicklung von *Brassica napus* konnte gezeigt werden, dass das Endosperm neben seiner Funktion als Nährstoffspeicher ein eigenständiges und stoffwechselaktives Gewebe innerhalb des Samens repräsentiert. Die Identifizierung spezifischer Stoffwechselwege und regulatorischer Elemente lieferte einen Hinweis auf molekulare Interaktionen zwischen dem Endosperm und dem Embryo. Der etablierte proteomische Versuchsaufbau eignet sich für die Analyse von Samen unabhängig von der Spezies und lieferte neue Erkenntnisse zur Endospermfunktion in *Cyclamen persicum* (Kapitel 2.3) und zu spezifischen Eigenschaften zygotischer und somatischer Embryonen von *Theobroma cacao* (Kapitel 2.4). Sobald die Differenzierung der Samengewebe abgeschlossen ist, beginnt die massive Einlagerung von Speicherstoffen wie Lipiden und Proteinen. Da die Akkumulierung von Speicherproteinen von zentraler Bedeutung für die Keimung ist, wurde die native Struktur des Cruciferins aus *B. napus* Samen charakterisiert (Kapitel 2.2). Der Cruciferin-Komplex ist durch eine oktamere, fassförmige Struktur gekennzeichnet. Diese Struktur ermöglicht eine effiziente Speicherung von Aminosäuren auf kleinstem Raum. Sowohl Entwicklungsprozesse als auch die Speicherstoffsynthese erfordern große Mengen an Reduktionsäquivalenten und Energie. Es wird vermutet, dass Samen alternative Möglichkeiten zur Energiebereitstellung, wie den Abbau von Aminosäuren, besitzen. Erste detaillierte Analysen zur Funktion einer mitochondrialen Schwefeldioxygenase ETHE1 unterstützen die Hypothese, dass der Aminosäurestoffwechsel in Samen eine zentrale Rolle spielt und an der Bereitstellung von Stickstoff und Energie beteiligt ist (Kapitel 2.5, Kapitel 2.6). Die in dieser Arbeit präsentierten neuen Ergebnisse zur Biologie der Samenentwicklung, insbesondere zur Speicherstoffsynthese und zum Aminosäure-stoffwechsel, könnten künftig genutzt werden, um die Samenqualität und den Ertrag von Nutzpflanzen zu verbessern.

Schlagworte: Proteomik, Samen, Endosperm, Embryo, Speicherstoffeinlagerung

Abstract

Plant seed development starts with the unique event of double fertilization. The resulting seed tissues endosperm and embryo undergo characteristic developmental sequences with dramatic changes on cellular and molecular level spreading from tissue differentiation and growth to storage deposition during seed filling. In the course of this thesis special emphasis was placed on the physiological and molecular characterization of developmental mechanisms and their impact on different seed compartments (chapter 2.1). For this purpose a proteomic strategy including (i) establishment of an endosperm proteome reference map, (ii) characterization of endosperm development and (iii) comparison of endosperm and embryo proteomes was designed. This first study of endosperm development in *Brassica napus* highlighted that beside its nutritive function the endosperm represents a self-competent and metabolically active tissue. Furthermore, it has been shown that several pathways and regulatory elements govern the biochemical interplay between the embryo and the endosperm. The established proteomic workflow can be applied to seeds of different species. Here, it revealed new insights into endosperm function of *Cyclamen persicum* (chapter 2.3) and characteristics of zygotic and somatic embryos of *Theobroma cacao* (chapter 2.4). Once the tissue differentiation is terminated storage compounds such as lipids and proteins are massively accumulated in the embryo of dicotyledonous species such as oilseeds. Since efficient protein storage is essential for germination success the native structure of *B. napus* cruciferin was characterized (chapter 2.2). It revealed insights into structure and packing of seed storage proteins. The cruciferin complex has an octameric barrel-like structure optimized for maximum amino acid storage capacity with minimal space requirements. Both, seed development and storage activity require high amounts of energy and reduction equivalents. It is proposed that in seeds alternative routes like amino acid degradation contribute to energy provision for efficient energy metabolism. First detailed investigations on the function of the mitochondrial sulfur dioxygenase ETHE1 in seeds support the hypothesis that amino acid metabolism is highly important in seeds and may serve in nitrogen and energy provision during seed development (chapter 2.5, chapter 2.6). The novel insights into seed biology presented in this thesis, in particular for storage product synthesis and amino acid metabolism, are potential targets for strategies aiming to improving seed quality and yield of crops in the future.

Keywords: proteomics, seed, endosperm, embryo, seed filling

Contents

Abbreviations 1

Chapter 1: Seed biology and plant proteomics

1.1 Seed development of dicotyledons: An overview 3

1.1.1 Embryogenesis 4
1.1.2 Endosperm development 4
1.1.3 Physiological and molecular determinants controlling seed development 5

1.2 Seed proteomics: Methods and applications 8

1.2.1 General concept of seed harvesting and sample preparation 8
1.2.2 Gel-based versus gel-free proteomics 9

1.3 Insights into central metabolic pathways essential for proper seed development 14

1.3.1 Seed photosynthesis 14
1.3.2 Mitochondrial metabolism in seeds 15
1.3.3 Carbohydrate metabolism in developing seeds 16
1.3.4 Storage lipid biosynthesis 17
1.3.5 Seed storage protein accumulation 18

1.4 Objectives of the thesis 21

1.5 References 22

Chapter 2: Publications and Manuscripts

2.1 *B. napus* seed endosperm – metabolism and signaling in a dead end tissue 33
Journal of Proteomics 108: 382-426

2.2 The native structure and composition of the cruciferin complex in *Brassica napus* 79
Journal of Biological Chemistry 288: 2238-2245

2.3 Proteomic and histological analyses of endosperm development in *Cyclamen persicum* as a basis for optimization of somatic embryogenesis 89
Plant Science 201/202: 52-65

2.4 Comparative proteomic analysis of early somatic and zygotic embryogenesis in *Theobroma cacao* L. 105
Journal of Proteomics 78: 123-133

2.5 Dealing with the sulfur part of cysteine: four enzymatic steps degrade L-cysteine to pyruvate and thiosulfate in *Arabidopsis* mitochondria preventing toxic effects of persulfides 117

2.6 Amino acid metabolism and the role of ETHE1 in *Arabidopsis thaliana* seeds 147

Affix 181

Curriculum Vitae

Danksagung

Abbreviations

2D	two dimensional
2D-GE	two dimensional gel electrophoresis
ACCase	acetyl-CoA carboxylase
ACP	acyl carrier protein
ATP	adenosine triphosphate
Chla	chlorophyll a
Chlb	chlorophyll b
CLE8	CLAVATA3/embryo surrounding region-related 8
CO_2	carbon dioxide
CoA	coenzyme A
DAP	days after pollination
DIGE	differential gel electrophoresis
ER	endoplasmatic reticulum
ESR	embryo surrounding region
ETC	endosperm transport cells
ETHE1	ETHYLMALONIC ENCEPHALOPATHY PROTEIN1
FA	fatty acid
IEF	isoelectric focussing
IPG	immobilized pH gradient
KAS	3-ketoacyl-ACP synthase
KOD	KISS OF DEATH
LC-MS/MS	liquid chromatography tandem mass spectrometry
MDH	NAD-dependent malic enzyme
ME	NADP-dependent malic enzyme
MQ	MaxQuant
MS	mass spectrometry
NADPH	nicotinamide adenine dinucleotide phosphate
O_2	oxygen
OPPP	oxidative pentose phosphate pathway
OXPHOS	oxidative phosphorylation
PAGE	polyacrylamide gel electrophoresis
PD	Proteome Discoverer

PEPC	phosphoenolpyruvate carboxylase
PSII	photosystem II
PSV	protein storage vacuole
rER	rough endoplasmatic reticulum
RNA	ribonucleic acid
RuBisCO	ribulose-1,5-bisphosphate carboxylase/oxygenase
SDS	sodium dodecyl sulfate
SRM	selective reaction monitoring
SSP	seed storage protein
TAG	triacylglycerol
TCA	tricarboxylic acid
t-SIM	targeted-selected ion monitoring
VLCFA	very long chain fatty acid

Chapter 1: Seed biology and plant proteomics

Why studying seed biology is of great importance? Seeds contain the entire repertoire of compounds needed for plant germination. They adapt to multitude (sometimes adverse) environmental conditions and plants developed unique mechanisms for their dispersal. Most of the seeds contain large and specific quantities of storage compounds which are important and sometimes essential components of animal and human diets. Thus, the biological and economic importance of seeds is evident. This chapter provides insights into processes involved in seed development and a global overview of the theoretical background of the thesis. In addition, recent findings in the field of seed biology and plant proteomics are presented and discussed.

1.1 Seed development of dicotyledons: An overview

Seed development starts with the unique event of double fertilization in which one sperm cell fuses with the egg cell to form the embryo while a second sperm cell unites with the central cell resulting in the triploid endosperm. The seed compartments undergo unique developmental changes underpinned by dramatic changes on cellular and molecular level (Figure 1). These differentiations are controlled by maternal tissues first and by filial organs in the later stages.

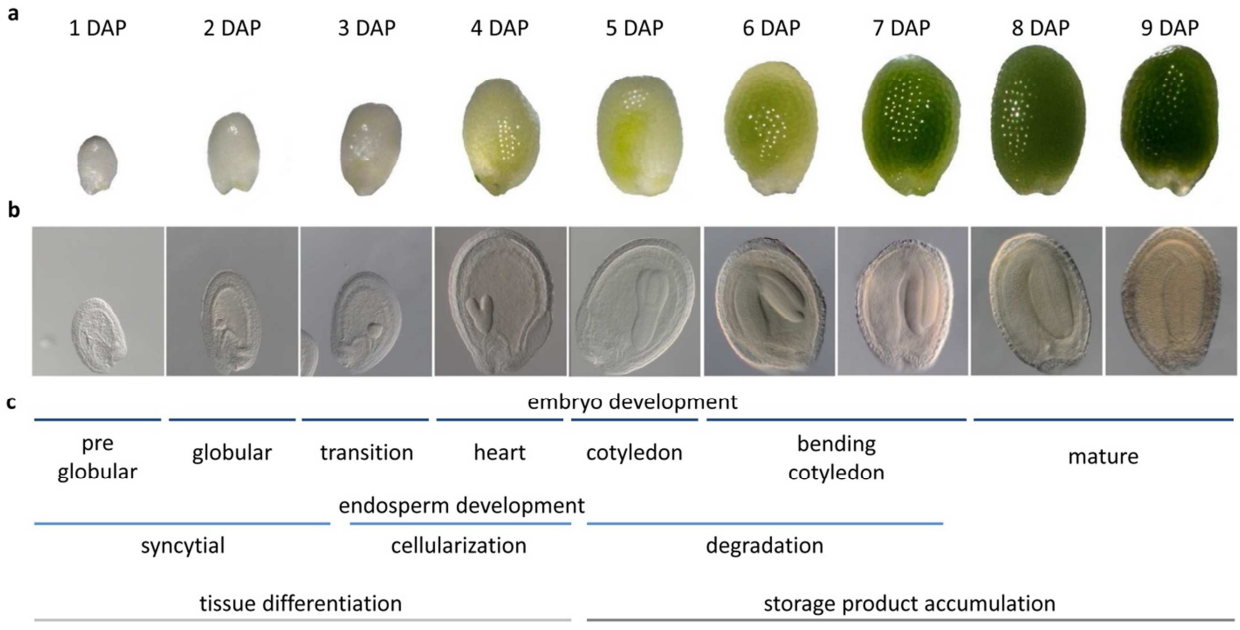


Figure 1: Embryo and endosperm development in the model plant *Arabidopsis thaliana*. Seeds were harvested from 1 to 9 DAP (**a**) and subsequently analysed by light microscopy (**b**). The endosperm and embryo follow distinct differentiation processes (**c**). **DAP**, days after pollination.

1.1.1 Embryogenesis

Embryogenesis comprises a stepwise differentiation of the fertilized zygote including cell division, cell elongation, greening and photosynthetic activity as well as accumulation of seed storage products. Overall, two distinct phases can be defined: (i) morphogenesis and (ii) maturation (Goldberg et al. 1994). The initiation of embryo formation and development is controlled by pronounced gene activity (Rodoeva and Weijers 2014). After fertilization of the egg cell a pro-embryo containing an apical and a basal cell merges defining the polarity of the embryo. The apical cell gives rise to the embryo, while the basal cell forms the suspensor, a connective structure between filial and maternal tissues. At later stages the suspensor degenerates and the metabolic connection to the mother plant is interrupted (Kawashima and Goldberg 2010). Through further cell divisions, elongation and differentiation processes the embryo acquires the building plan of the plant (De Smet et al. 2010, Lau et al. 2012). The morphogenetic processes during embryogenesis depend on regulatory networks, particularly on cell cycle regulation which ensures the formation of different cell types and tissues (Dante et al. 2014, Wendrich and Weijers 2013). During maturation the synthesis and massive accumulation of storage products is prominent. At a very late maturation stage, the embryos metabolic activity is terminated and the seed becomes tolerant to desiccation. The previously accumulated storage products can be degraded during germination to provide nutrients to the growing seedling, before photosynthetic activity is established (Baud et al. 2002).

1.1.2 Endosperm development

The endosperm is essential for embryo development during seed formation in plants. It embeds the embryo and is itself surrounded by the seed coat. In dicotyledonous species, e.g. *A. thaliana* and *B. napus*, the endosperm is of transient nature and only the aleurone layer remains until seed maturity (Novack et al. 2010). The endosperm development is characterized by a stepwise differentiation including transition from a syncytial to a cellular phase (Olsen 2004, Berger 2003, Brown et al. 2003, Brown et al. 1999). The syncytium is produced by several cycles of mitosis of the fertilized central cell in the absence of cytokinesis, thus leading to a multi nuclear cell (Olsen 2004). In the course of seed development the endosperm is divided into three mitotic domains: (1) the micropylar, (2) the peripheral and (3) the chalazal region in which nuclei divide simultaneously (Boisnard-Lorig et al. 2001). The syncytium becomes cellularized by the formation of radial microtubule systems originating from the micropylar region (Olsen 2004, Sørensen 2002, Brown et al.

1999). The process of cellularization has a decisive impact on embryo development (Hehenberger et al. 2012) and conducts the initiation of embryo growth (Brown et al. 1999). When endosperm development is terminated, programmed cell death leads to endosperm cell degradation which is probably controlled by ethylene and interactions between embryo and endosperm (Berger et al. 2006, Young and Gallie 2000, Berger 1999). The remaining aleuronic layer holds key functions in the deposition of storage compounds and enzymes, as well as desiccation and germination of seeds (Yan et al. 2015, Holdsworth et al. 2008, Bethke et al. 2007, Penfield et al. 2004).

1.1.3 Physiological and molecular determinants controlling seed development

The interplay of different seed compartments on various levels is crucial for coordinated seed development. However, regulatory mechanisms and communication between compartments are not yet fully elucidated. Recent studies provided further insights into signal transduction and transport processes.

In dicotyledonous seeds the endosperm develops first, while embryo development is paused. Due to its location the endosperm is able to coordinate maternal and filial contributions (Novack et al. 2010). Studies of mutants carrying endosperm specific defects have given insights into development and function of this compartment. Many mutations of endosperm related genes affect functions in cell differentiation, cellularization and cell growth (Hehenberger et al. 2012, Lu et al. 2012, Cavel et al. 2011, Lee et al. 2012, Sabelli and Larkins 2009, Ohto et al. 2005, Garcia 2003). Seed growth is directly linked to seed yield (Adamski et al. 2009) and understanding the mechanisms controlling the seed size therefore is essential (Orozco-Arroyo et al. 2015, Boisnard-Lorig et al. 2001). The seed coat acts as a physical barrier restricting embryo expansion, particularly that of the cotyledons (Fang et al. 2012, Haughn and Chaudhury 2005). Elements influencing endosperm development comprise cell cycle control, genomic imprinting, hormones and environmental conditions (Bauer and Fischer 2011, Wolff et al. 2011, Huh et al. 2007, Berger 2004, Gehring et al. 2004). The endosperm has distinct functional regions to coordinate developmental processes. For cereals it is known that the endosperm forms transfer cells (ETC) and embryo surrounding region (ESR) which mediate nutrient shuffling to the embryo and might be involved in signaling processes (Thiel 2014, Olsen 2004). So far, ESRs have not been characterized in dicotyledons, but it is very likely that also species with a transient endosperm are building

metabolic active regions supporting embryo growth. However, it remains to be investigated if putative dicotyledonous ESR is functionally equivalent to the one identified in cereals. The general and specific metabolic functions of the endosperm are just beginning to emerge (Lafon-Placette and Köhler 2014, Belmonte et al. 2013, Huang et al. 2009). A recent analysis has shown that the molecular architecture of *B. napus* endosperm comprises the entire set of central metabolic pathways (Lorenz et al. 2014, chapter 2.1). Besides its nutritive function, the endosperm fulfils key functions in metabolism and signaling. Its enzymatic machinery turns the endosperm into a self-sustaining and metabolically active tissue, which receives major assimilates from the mother plant, but which is also able to produce metabolic intermediates by itself. Shifts in metabolism are reflecting the dramatic changes on cellular level. It is particularly unknown to which extent endosperm metabolism can modulate the incoming stream of assimilates thereby shaping differentiation, growth and storage compound formation of the embryo. Belmonte et al. (2013) presented developmental profiles of *A. thaliana* gene activity and characterized functional differentiation of subregions from fertilization to maturation. Coexpression analyses of identified genes suggest that processes defining seed size and regulating storage product accumulation are coordinated across several regions within the seed. Investigations indicate a high number of transcription factors and regulatory components in seeds (Lorenz et al. 2014, Le et al. 2010), but only a minor number of transcription factors have been characterized so far. It appears very likely that their involvement in various signaling cascades makes a fundamental contribution to programming and regulation of seed development.

Seed development as a whole not only depends on the progress of each of the individual compartments (seed coat, embryo and endosperm), it also relies on the cross-talk between all compartments ensuring adjustment of the developmental processes in response to developmental cues. The endosperm is a connective tissue within the seed and provides the environment meeting the needs of the growing embryo. Communication routes in plant cells can either run through the symplast or the apoplast. Apoplastic transport is supported by membrane-localized transporters, while symplastic transport involves plasmodesmata. These micro channels are the connecting elements of cytoplasma and endoplasmic reticulum (ER) of neighboring cells and allow the exchange of hormones, metabolites, RNAs and small proteins (Xu and Jackson 2010). It is proposed that communication routes change several times during development (Kim et al. 2005, Lee and Yeung 2010). In seeds secreted peptides are highly abundant in specific cells and compartments. Recent findings suggest that their specific

accumulation might regulate signal cascades (Ingram and Gutierrez-Marcos 2015). For example, a peptide derived from the *Arabidopsis* CLAVATA3/embryo surrounding region-related 8 (CLE8) influences embryo cell division, endosperm proliferation and the timing of endosperm differentiation (Fiume et al. 2011). A different peptide (KISS OF DEATH, KOD) expressed in embryos is involved in the degradation of the suspensor cell (Blanvillain et al. 2011). It is suggested that peptides might be utilized to control developmental processes in plants in general (Costa et al. 2014), although their exact role in regulatory processes have not yet been fully elucidated. Furthermore, a tight redox control in the endosperm is of highest significance to ensure non-oxidizing conditions. Such conditions were recently proposed to be of key importance for germ cell fate and sexual reproduction (Kelliher and Walbot 2012). In vegetative tissues redox signals derived by photosynthetic activity regulate Calvin cycle, ATP generation, NADPH export, starch metabolism, carbon fixation, lipid synthesis and amino acid synthesis (Considine and Foyer 2014). The components of the redox-signaling pathway mainly respond to physiological and environmental inputs. The redox status in seeds is expected to be regulated using the glutathione system (Cairns et al. 2006).

Remarkable progress is currently made in understanding the interactions of different seed compartments and signaling processes, but many gaps are yet to be filled. Future challenges include the integration of different research fields such as microscopic analysis, metabolomics, transcriptomics and proteomics. Such a system biology approach is a promising avenue to unravel communication systems and to identify checkpoints of developmental control in seeds.

1.2 Seed proteomics: Methods and applications

The term proteome was introduced in 1996 by Wilkins et al. and defined as "the total protein complement of a genome". However, a proteome is highly complex with respect to the chemical properties and the variable range of its constituents. In addition the proteome is dynamic and therefore its composition depends on several factors. Proteins are the functional and structural units within a cell and continually undergo changes due to biosynthesis and degradation, metabolic activity, interaction with other proteins to form complexes and post translational modifications. Protein function itself is regulated by external and internal cues. Thus, the investigated proteome reflects the current biological state and can be defined as the molecular phenotype of any given biological sample. Proteomic analyses are suitable for whole organisms, organs, specific tissues, cell types or organelles. Among all "omics"-techniques, proteomics represents a key technology to investigate dynamic and complex biological systems such as seeds.

1.2.1 General concept of seed harvesting and sample preparation

Harvest and analysis of developing seeds is not trivial. For comparative analysis it is necessary to ensure developmental and physiological uniformity of the seed samples. To investigate seeds of a specific developmental stage, flowers are labeled on the day of pollination and are harvested at time points specific for the developmental stages in question. In addition to flower tagging and time-dependent harvesting, developmental stages of seeds can be checked by light microscopy of the growing embryo (Goldberg et al. 1994). In most seed proteomic studies whole seeds have been used since seeds of model plants such as *Arabidopsis* are comparatively small and separation of the seed in its different tissues (for example endosperm and embryo) is difficult. Schiebold et al. (2011) were able to show that laser micro dissection is a powerful method to collect different tissues from a seed for comparative analysis. However, to study the biological function of different seed compartments species producing large seeds are favorable. The endosperm of *B. napus* seeds can be easily isolated during its development using a micro syringe and provides an excellent source for proteomic approaches aimed at investigating the impact of a transient tissue on development and storage product accumulation (Lorenz et al. 2014, chapter 2.1). Sample preparation and protein extraction are critical steps to guarantee consistent and high quality proteomic data. Protein extraction from seed cells is a challenging task since the cells contain

compounds such as polyphenols, polysaccharides, starch, lipids and several secondary metabolites interfering protein separation methods such as gel electrophoresis and mass spectrometry (MS). In addition, storage proteins are present at high amounts while other proteins are comparatively low abundant in seeds. For total soluble proteins of seed samples an extraction method established by Colditz et al. (2004) has proven suitable and results in high-quality protein extracts. This method is based on protein solubilization in the phenol phase and several rounds of centrifugation followed by ammonium acetate precipitation to reduce interfering molecules in the sample. The major challenge is to analyse as many proteins as possible in a given sample. Proteins have diverse molecular characteristics such as molecular weight, isoelectric point, post translational modifications and hydrophobicity. The complexity of proteins and their dynamic range of concentration within seed samples can for example be reduced by sub-fractionation of different organs and tissues, e.g. endosperm and embryo (Lorenz et al. 2014). In order to analyse low abundant proteins during seed filling it is advantageous to deplete high abundant storage proteins from the sample (Krishnan et al. 2009). Many protocols have been established to analyse plant organellar proteomes such as chloroplasts (Kubis et al. 2008), mitochondria (Keech et al. 2005) and peroxisomes (Eubel et al. 2008). Nevertheless, to apply those protocols to seed samples they need to be adjusted for low amounts of seed material.

1.2.2 Gel-based versus gel-free proteomics

Gel-based proteomics represents the most popular approach of global separation and quantification of proteins and became a standard tool in many biological studies. Two dimensional gel electrophoresis (2D-GE) resolves approximately 1000 proteins per gel and enables simultaneous characterization of different biochemical properties such as isoforms and molecular mass. The analysis of quantitative changes of proteins is based on spot volumes followed by protein identification via mass spectrometry. However, gel based methods also have limits in terms of analysing high molecular weight proteins, alkaline proteins, hydrophobic proteins and low abundant proteins. In the recent years gel-free approaches such as shotgun mass spectrometry were established. Prior to analysis proteins are digested and the resulting complex peptide mixtures are separated by liquid chromatography and subsequently analysed by tandem MS (LC-MS/MS), which offers high throughput analysis of proteomes and provides information on protein quantities. Hence, considering gel-based and/or gel-free methods mainly relies on the biological question addressed. The continuous improvement of

proteomic methods largely contributed to the characterization of seed proteomes and elucidation of complex interaction networks (reviewed in Wang et al. 2014 and Hajduch et al. 2011). In this thesis gel-based as well as shotgun (gel-free) proteomics were implemented to unravel physiological transitions and regulatory mechanisms required for proper seed development and reserve deposition (Figure 2). In the future, this information could be translated from model plants to crops to increase seed quality and yield.

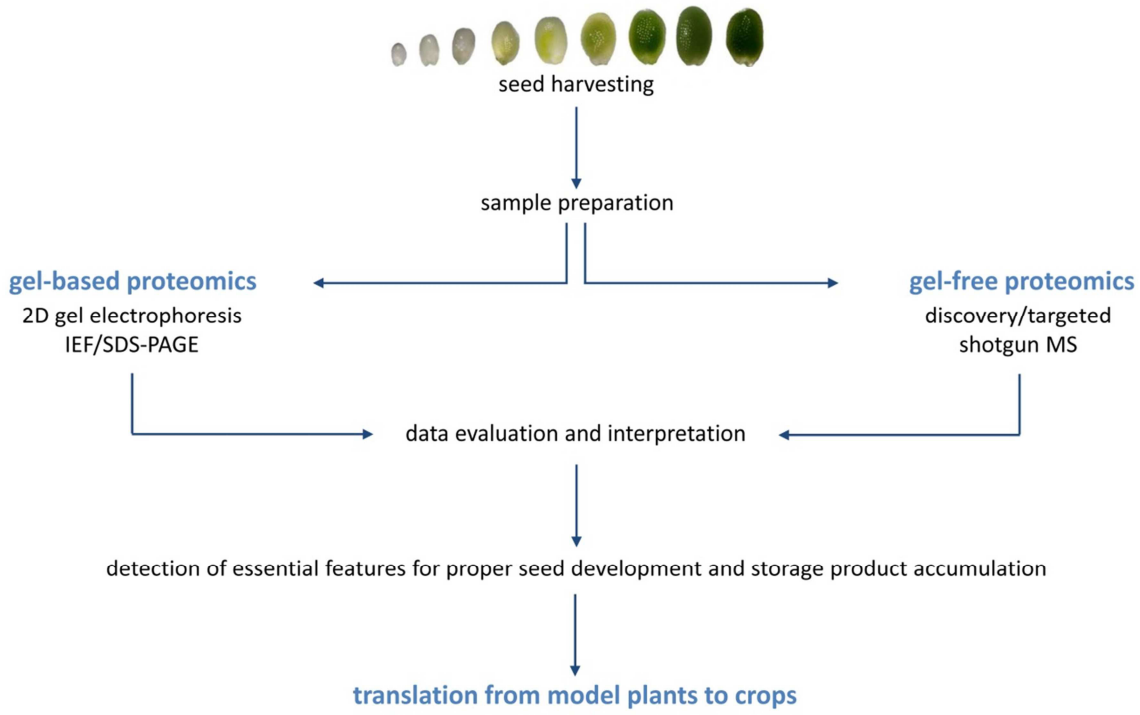


Figure 2: Proteomics workflow towards understanding of developmental processes and storage product accumulation in seeds. After isolation, seed proteins can be either analysed by gel-based or gel-free proteomics. The combination of both techniques delivers comprehensive information on essential features of seed formation and is able to identify the key proteins involved in this process, thereby revealing the biological mechanisms controlling tissue transition to seed filling. This information can be translated from model plants to crops to increase seed quality and yield. **2D**, two dimensional; **IEF/SDS-PAGE**, isoelectric focussing/sodium dodecylsulfate polyacrylamide gel electrophoresis; **MS**, mass spectrometry.

Comparative proteomic analysis using two dimensional gel electrophoresis

A proteomic workflow including two dimensional gel electrophoresis with isoelectric focussing (IEF) in immobilized pH gradients (IPG) followed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE), comparative analysis using gel analysis software and subsequent protein identification via LC-MS/MS has been established to analyse seed samples (Figure 3). Two dimensional gel electrophoresis delivers high resolution separation of different proteins including information on molecular masses, isoforms and post

translational modifications. The large number of simultaneously separated proteins characterizes the current biological state of different seed tissues or developmental stages. A known disadvantage of 2D gel electrophoresis is the low level of reproducibility due to differences in the gel runs that impede direct comparison of spot pattern of different 2D gels. Such gel to gel variations can be reduced by using differential gel electrophoresis (DIGE) that involves fluorescent labeling before protein separation, thereby allowing multiplexing of different seed samples (Hajduch et al. 2007). A positive side effect of this approach is an increased sensitivity (when compared to classical, non-fluorescent protein dyes such as coomassie). However, the multiplexing capacity is limited to the number of available fluorescent dyes. Several commercial and freely available software solutions can be used to analyse multiplexing and non-multiplexing IEF/SDS gels (e.g. Delta2D, Decyder, PD-Quest, Image Master 2D, Melanie, Same Spots). In the course of this thesis, the Delta2D software for gel image analysis and comparative proteomic approaches as described in Berth et al. (2007) was used. This software is able to match and relatively quantify protein spot volume from a large set of different samples. It has been shown that the advanced statistics and visualization methods included in Delta2D allow profiling of spot volumes and relative quantification of individual spots to predict metabolic pathways and functional dynamics in seed proteomes of different species. Physiological clustering of regulated proteins revealed new insights into endosperm function of *Brassica napus* (Lorenz et al. 2014, chapter 2.1) and *Cyclamen persicum* (Mwangi et al. 2013, chapter 2.3). The proteomic analysis of zygotic and somatic embryos of *Theobroma cacao* has been shown that carbohydrate metabolism especially glycolysis is of special importance for embryogenesis (Noah et al. 2013, chapter 2.4). These gel based proteomic studies therefore contributed in identifying quantitative changes of proteins that are particular important for developmental and metabolic processes in seeds.

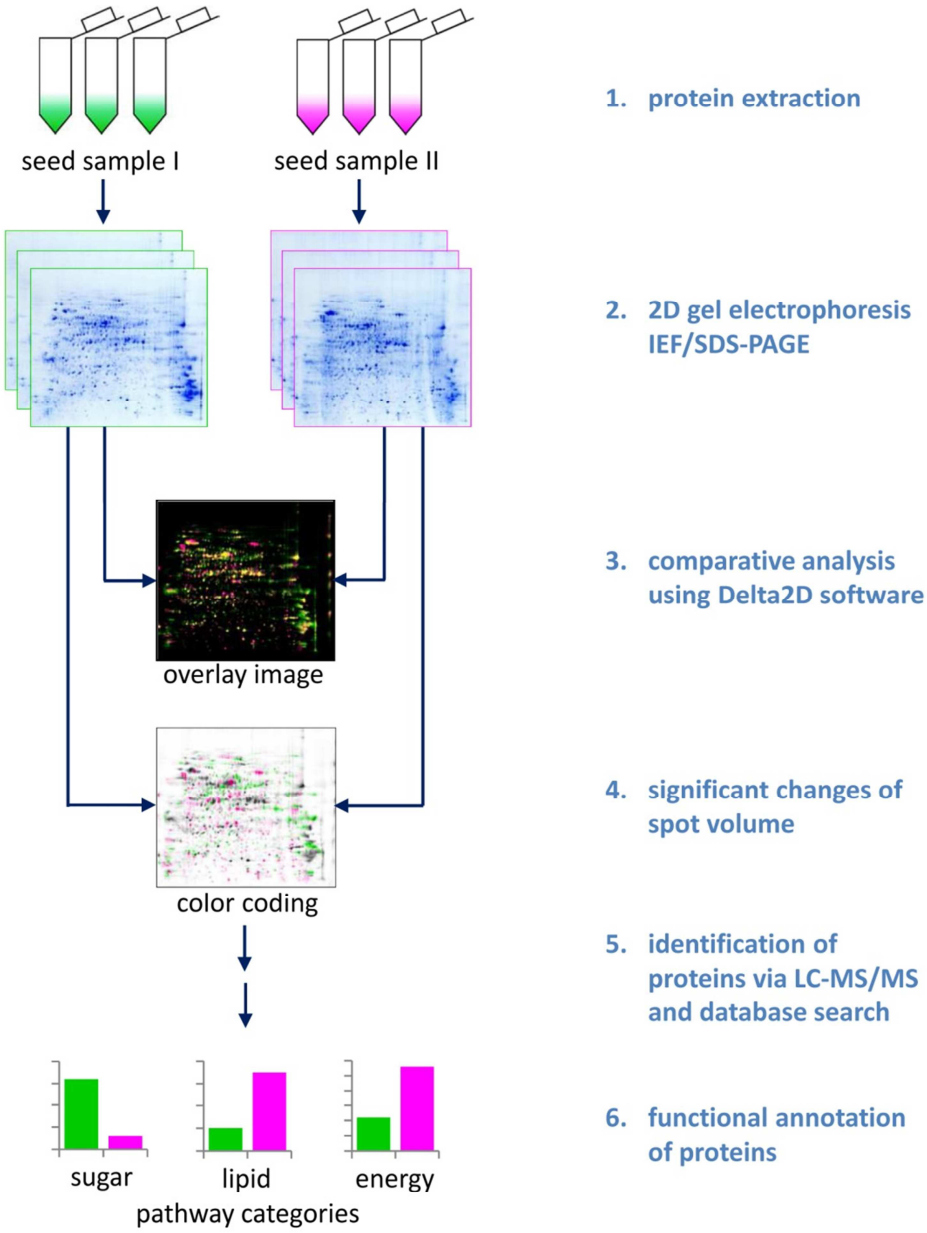


Figure 3: Workflow for gel based comparative proteomics. The total soluble proteins are extracted from each seed sample (**1**) and separated by two dimensional (**2D**) IEF/SDS-PAGE (**2**). For statistical analyses at least three replicates per sample are required. The Delta2D software (Decodon) was used for the comparative quantitative analysis (**3**). First a digital overlay image of the samples is produced. Spots shown in green are of higher volume in seed sample 1 compared to seed sample 2, magenta spots are of reduced volume and yellow spots are of equal volumes. Statistically significant changes across all samples are marked in a color coding image (**4**). The proteins within the differentially abundant spots can be analysed via liquid chromatography (**LC**) followed by tandem mass spectrometry (**MS/MS**) and database search (**5**). Finally, functions of identified proteins are assigned to metabolic pathways (**6**). **IEF,** isoelectric focussing; **SDS**, sodium dodecylsulfate; **PAGE**, polyacrylamide gel electrophoresis.

Shotgun mass spectrometry based proteomics

Supported by the ongoing development of mass spectrometers with respect to mass accuracy and sensitivity the characterization of whole proteomes in a high throughput manner became feasible without the need for prior separation of proteins (Nilsson et al. 2010). A shotgun MS based proteomic workflow includes protein extraction and digestion with a protease e.g. trypsin. The resulting complex peptide mixture is separated by liquid chromatography and spectra of the mass-to-charge ratio of precursor peptide ions and fragment peptide ions are acquired by MS scans. The huge sample throughput and high sensitivity are major advantages of LC-MS/MS methods. In addition to the pure identification of proteins (discovery MS), quantitative shotgun MS methods were successfully established in the last years. It has been shown that label free shot gun proteomics can be applied to detect and relatively quantify peptides and therefore proteins in seed samples using high resolution instruments. Applying this method revealed new insights into the role of seed photosynthesis and amino acid metabolism during development (Lorenz et al. 2015, chapter 2.6). Label-free quantification methods include either peak integration of precursor ions or spectral counting of fragment ions (Cox et al. 2014, Zhu et al. 2010). Peak integration seems to produce more accurate results, whereas spectral counting is more sensitive (Old et al. 2005). Both methods are promising to give broad insights into differentially abundant proteins within two or more proteomes. However, bioinformatics and statistics are crucial steps also in quantitative shotgun MS proteomic approaches and specialized software tools have been developed such as MaxQuant (MQ), which integrates peptide peaks and matches them from multiple samples (Cox et al. 2008) and Proteome Discoverer (PD, Thermo Fisher Scientific) for spectral counting. More recently targeted proteomic approaches such as selective reaction monitoring (SRM) and targeted-selected ion monitoring (t-SIM) were developed to identify and quantify specific proteins of interest. Picotti et al. (2013) have shown that SRM mass spectrometry is a powerful tool and represents an alternative application to immune detection by the use of antibodies. This technique also termed 'mass-western' allows quantification of low abundant proteins in complex samples.

1.3 Insights into central metabolic pathways essential for proper seed development

During seed formation, changes in cellular structure coincide with distinct switches in metabolic activities. At early stages a metabolic cross-talk between the endosperm and embryo ensures cell differentiation and proper embryo growth. Later, the embryo massively accumulates storage compounds within the cotyledons which become a highly specialized storage tissue (Figure 4). Nutrients such as sugars and amino acids delivered from the mother plant are precursors for storage product biosynthesis in seeds (Melkus et al. 2009). In oilseeds like *A. thaliana* and *B. napus* mainly lipids and proteins are deposited during the seed filling phase. Only limited information on carbon storage in the cotyledons of oilseeds is available. At early stages, starch and hexoses (glucose and fructose) accumulate transiently, but their amount gradually decreases accompanied with a rapid increase of oil and lipid storage (Borisjuk et al. 2005, Baud et al. 2002). During germination, storage compounds are subsequently degraded and provided to the growing seedling. The efficiency of the storage compound mobilization depends on their amount accumulated during maturation and the action of distinct metabolic switches towards germination (Gallardo et al. 2008, Fait et al. 2006). The results obtained in the course of this thesis indicate that the activity of central metabolic pathways such as photosynthesis, mitochondrial metabolism and carbohydrate metabolism are essential for proper tissue formation and storage product synthesis during seed filling.

1.3.1 Seed photosynthesis

Seeds of many different species are green during development, e.g. *B. napus* and *A. thaliana* contain photoheterotrophic plastids. The permeability of gases into seeds is rather low and plastidial activity during maturation phase contributes to oxygen allocation and reassimilation of CO_2 (Borisjuk and Rolletschek 2009, Rolletschek et al. 2005, Borisjuk et al. 2005, Ruuska et al. 2004). Seed plastids hold special structures and have adapted their metabolism to cope with reduced light levels available for photosynthetic reactions (Borisjuk et al. 2013, Borisjuk et al. 2004). They have an elongated shape and contain high amounts of grana stacks and large starch grains (Asokanthan et al. 1997). The electron transport activity in seed plastids is different as shown by an enhanced chlorophyll a (Chla) to chlorophyll b (Chlb) ratio and high abundance of photosystem II (PSII) proteins, indicative of a cyclic electron transport via PSII. Highest photosynthetic activity is observed during storage product synthesis, suggesting a

correlation of metabolic processes and photosynthesis (Fait et al. 2006, Ruuska et al. 2002). In contrast to chloroplasts of leaves, seed plastids import carbon, mainly glucose-6-phosphate, phosphoenolpyruvate and pyruvate via specific transporters (Eastmond and Rawsthorne 1998). The enhanced activity of PSII in seeds contributes to O_2 supply needed for energy provision by redox equivalents and ATP (Borisjuk et al. 2005). At the same time, seed capacity of CO_2 fixation is rather low (Asokanthan et al. 1997). Hence, the metabolism of photoheterotrophic plastids in seeds is changed into the direction of fatty acid biosynthesis (Andriotis et al. 2010, Goffman et al. 2005). Therefore it is proposed that seed photosynthesis effects embryogenesis (Hsu et al. 2010) and consequently germination (Allorent et al. 2015).

1.3.2 Mitochondrial metabolism in seeds

Mitochondria are essential organelles with pivotal roles in many cellular processes such as tricarboxylic acid (TCA) cycle activity, oxidative phosphorylation and energy production (Millar et al. 2011). In addition to the previously described functions, seed mitochondria have special roles in stress response and desiccation tolerance during germination (Macherel et al. 2007). As seeds have very compact tissues, the capacity of gas exchange is rather low, requiring an adjustment of mitochondrial metabolism (Borisjuk et al. 2005). Mitochondrial activities in seeds are well adapted to promote storage product accumulation (Vigeolas et al. 2003). It is assumed that embryos of *B. napus* seeds hold an unconventional mitochondrial metabolism since the cyclic flux around the TCA cycle is absent and substrate oxidation is reduced and hardly contributes to ATP production. Furthermore, isocitrate dehydrogenase activity is reversible and malic enzyme activity is enhanced (Schwender et al. 2006). Recently it was postulated that amino acids are not only building blocks for storage proteins within seeds, but also serve as alternative substrates for mitochondrial metabolism during stress situations (Galili et al. 2014, Fait et al. 2006). In vegetative tissues, amino acid catabolism is induced by carbon starvation situations e.g. periods of extended darkness (Hildebrandt et al. 2015, Araujo et al. 2011 and 2010). First studies denote a potential role of amino acid catabolism to the energy status in seeds (Krüßel et al. 2014, Credali et al. 2013, Angelovici et al. 2011, Gu et al. 2010, Angelovici et al. 2009, Weigelt et al. 2008, Zhu and Galili 2003). It is likely that also in seeds together with photosynthesis amino acid degradation presumably contributes to energy provision. It has been shown that in *Arabidopsis* leaves a mitochondrial sulfur dioxygenase, is part of a sulfur catabolic pathway that catalyses the oxidation of sulfide or persulfides derived from amino acids to thiosulfate and sulfate and has key functions in

amino acid catabolism (Höfler et al. 2015, chapter 2.5, Krüßel et al. 2014). In seeds, ETHYLMALONIC ENCEPHALOPATHY PROTEIN1 (ETHE1) knock out leads to alterations in endosperm formation and finally causes seed abortion (Holdorf et al. 2012). It has been shown that a sulfur dioxygenase activity of 1 % present in an ETHE1 knock down mutant (*ethe1-1*) is sufficient for embryo survival, but development is severely delayed (Krüßel et al. 2014), which underlines the importance of this enzyme for seed development. A first detailed investigation on the function of the mitochondrial sulfur dioxygenase ETHE1 in seeds supports the hypothesis that amino acid metabolism is pronounced and has a potential role in nitrogen and energy provision during seed development (Lorenz et al. 2015, chapter 2.6).

1.3.3 Carbohydrate metabolism of developing seeds

Proteomic investigations of seed metabolism in different species have shown that carbohydrate metabolism is of central importance and has pivotal roles in development including seed filling, stress response and signal transduction (Lorenz et al. 2014, chapter 2.1; Mwangi et al. 2013, chapter 2.3; Noah et al. 2013, chapter 2.4). Sugars are degraded within a seed which is initially realized by invertases and sucrose synthases (Hill et al. 2003). Invertases also have a regulatory role in tissue formation by enhancing cell division whereas sucrose synthase activity affects biosynthesis of cellulose, proteins, lipids and starch in seeds (Wang and Ruan 2013, Xu et al. 2012, Pugh et al. 2010, Fallahi et al. 2008, Ruan et al. 2008, Chourey et al. 1998). At very early stages of seed development, vacuolar invertases are able to convert imported sucrose into hexoses in the endosperm and provide them to the growing embryo. The accumulation of hexoses subsequently drives water and nutrient uptake of the seed. Afterwards the cellularization of the endosperm causes shrinkage of the central vacuole and the ratio of hexoses and sucrose is switched (Hehenberger et al. 2012, Moreley-Smith et al. 2008). During later stages of development sucrose accumulates in the seeds. The switch from high hexose to high sucrose level is related to differentiation, cell expansion and storage activity in the embryo (Rolland et al. 2006, Weber et al. 2005). Thus, the embryo is able to develop into a self-competent sink tissue within the seed. Hexoses can be metabolized in the cytosol and in plastids either by glycolysis or oxidative pentose phosphate pathway (OPPP) to provide precursors for fatty acid biosynthesis (Schwender et al. 2003). Beyond these two pathways, Schwender et al. (2004) discovered that green seeds are able to contribute to fatty acid (FA) precursor formation by the activity of a ribulose-1,5-bisphosphate

carboxylase/oxygenase (RuBisCO) -bypass. In this unique pathway, RuBisCO is able to act without the Calvin cycle in a so far unknown reaction to improve carbon efficiency in seeds. Acetyl-CoA, a metabolite taking part in various metabolic reactions, is needed for proper seed development and storage product deposition (Ke et al. 2000). It is proposed that several enzymes like NADP-dependent malic enzyme (ME) (Shearer et al. 2004), plastidial acetyl-CoA synthase (Lin and Oliver 2008), cytosolic phosphoenolpyruvate carboxylase (PEPC) (Sebei et al. 2006), NAD-dependent malic enzyme (MDH) (Junker et al. 2007) and ATP citrate lyase (Rawsthorne 2002) contribute to fatty acid biosynthesis in seeds. However, the regulation of carbon flux into fatty acid production by the proposed pathways is complex and their specific activity *in vivo* needs to be further investigated. In addition to providing carbohydrates for metabolism and storage product accumulation, sugars hold special functions in signal transduction (Hanson and Smeekens 2009). Sugar signal molecules like glucose, fructose, and trehalose-6-phosphate may modulate development. Defective sucrose metabolism and signaling trigger stress responses and reproductive failure (O'Hara et al. 2013, Ruan et al. 2008). It is proposed that low glucose levels inhibit cell division (Wang and Ruan 2013, Weber et al. 2005) and induce over-production of reactive oxygen species due to reduced hexokinase activity (Kim et al. 2006). Investigations on mutants indicate a considerable role of glycolytic enzymes on plant development, growth and storage product biosynthesis (Dorion et al. 2012, Chen and Thelen 2010, Muñoz-Bertomeu et al. 2010 and 2009, Andre and Benning 2007, Lee et al. 2002, Plaxton 1996). It is postulated that the occurrence of a nucleotide sugar recycling pathway is essential for vegetative and reproductive growth in *A. thaliana*. Due to cellularization of the endosperm, nucleotide sugars are important cell wall monomers that represent substrates for the turnover of cell wall polymers at later stages of seed development (Geserick and Tenhaken 2013).

1.3.4 Storage lipid biosynthesis

In seeds storage lipids are accumulated in form of triacylglycerols (TAGs) and fatty acids (FAs) (Gallardo et al. 2008, Murphy and Cummins 1989). Predominant FAs in seeds are palmitate, stearate, oleate, linoleate and alpha-linoleate (Voelker and Kinney 2001). Oilseeds like *B. napus* are enriched with very long chain FAs (VLCFA) (Sharafia et al. 2015). FA biosynthetic pathways in plants are well characterized (Baud and Lepiniec 2009, Rawsthorne 2002) and strongly dependent on the supply of energy, reductants and carbon sources (see chapter 1.3.3). Sucrose delivered from the mother plant is sequentially metabolized by

pathways in the cytosol and the plastids to produce fatty acids (Hills 2004). The first committed step is catalyzed by an acetyl-CoA carboxylase (ACCase) which carboxylates acetyl-CoA to form malonyl-CoA. The assembly of FAs occurs at the acyl carrier protein (ACP). The resulting acyl-ACP can either be hydrolyzed by acyl-ACP thioesterases or further elongated by specific condensing enzymes, the 3-ketoacyl-ACP synthases (KAS). The biosynthesis of FAs is mainly regulated by interactions of transcriptional regulation and optimization of enzyme activity (Baud and Lepiniec 2009, Ohlrogge and Jaworski 1997). FAs are transported into the ER where they may subsequently be desaturated (Baud and Lepiniec 2009), prior to their esterification to a glycerol backbone and being transferred to dedicated storage organelles so called 'oil bodies' (Huang 1992). The biosynthesis of seed storage lipids is a potential target for metabolic engineering in an attempt increasing their content of healthy FAs, improve oil stability and generally increase oil content (Thelen and Ohlrogge 2002).

1.3.5 Seed storage protein accumulation

Seed storage proteins (SSPs) accumulate during the seed filling phase and are a major source of nitrogen and amino acids. The composition of storage proteins varies between different species. In the dicotyledons *Arabidopsis* and *B. napus* the mature seeds contain globulins (cruciferin), albumins (napin) and oleosins (oil body protein) (Höglund et al. 1992, Huang 1992) with cruciferins accounting for the bulk (approximately 60 %) of the total seed protein content (Crouch and Sussex 1981). Globulins are co-translationally synthesized into the rough endoplasmic reticulum (rER). The mature protein consists of an α- and a β-polypeptide chain linked by a disulfide bond. Both chains are derived from the same precursor molecule and the proglobulin is newly assembled into trimers by cleavage of the ER signal peptide followed by formation a disulfide bond between the N- and C- terminus of the polypeptide (Ereken-Tumer et al. 1982, Chrispeels et al. 1982, Sengupta et al. 1981). The formation of trimers is essential for a directed transport from the ER into protein storage vacuoles (PSVs) (Chrispeels et al. 1982). After import the trimers are disassembled, which is realized by a vacuolar processing enzyme, and organized into hexamers to form the mature protein (Dickinson et al. 1989, Jung et al. 1998, Shimada et al. 2003). In contrast to previous reports on globulin complexes, the native *B. napus* cruciferin complex exhibits a unique octameric barrel-like structure. This structure represents a very compact building block optimized for maximal storage of amino acids (Nietzel et al. 2013, chapter 2.2).

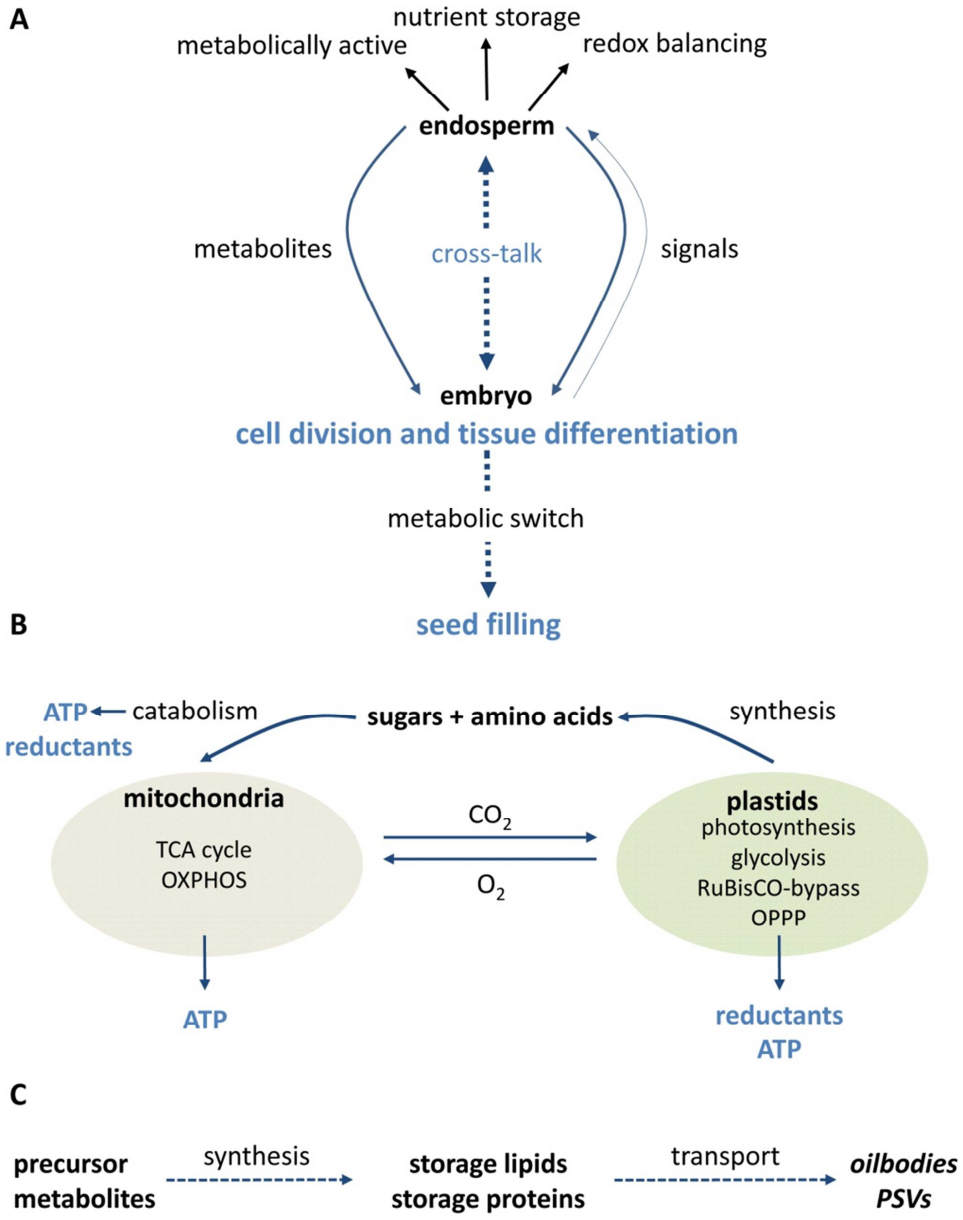


Figure 4: Essential features of developing seeds needed for proper tissue differentiation and storage product synthesis. At early stages a metabolic cross-talk between the endosperm and embryo ensures cell differentiation and proper embryo growth (**A**). Later the metabolism of the seed is switched to storage product accumulation. This process is marked by an increase of central metabolic pathways in cytosol, mitochondria and plastids (**B**). The provision of sufficient amounts of precursor molecules, reductants and energy highly influence the composition of stored lipids (*oilbodies*) and proteins (*PSVs*) (**C**). **ATP**, adenosine triphosphate; **CO_2**, carbon dioxide; **O_2,** oxygen; **OPPP**, oxidative pentose phosphate pathway; **OXPHOS**, oxidative phosphorylation; **PSV**, protein storage vacuole; **RuBisCO**, ribulose-1,5-bisphosphate carboxylase/oxygenase; **TCA**, tricarboxylic acid.

1.4 Objectives of the thesis

The overall aim of this thesis was to increase our understanding of seed biology and create a deeper knowledge of the physiological role of differentially regulated proteins during tissue differentiation and storage product accumulation. A detailed characterization of regulated metabolic pathways and the identification of developmental and tissue specific proteins using proteome analyses allowed identifying molecular events of key importance for seed biology. The new insights will be of relevance for improving seed quality and yield in the future.

This thesis was focused on the following objectives:

(i) Establishment of a comparative proteomic workflow for investigating seed development and storage product accumulation in different species

(ii) Detailed analysis of endosperm function and metabolism to characterize its contribution to seed development

(iii) Characterization of the native cruciferin complex, a highly abundant storage protein complex in *Brassica napus*

(iv) Investigation of photosynthesis and amino acid metabolism in developing seeds

1.5 References

Adamski N., Anastasiou E., Eriksson S., O'Neill C.M., Lenhard M. (2009) Local maternal control of seed size by KLUH/CYP78A5-dependent growth signaling. Proceedings of the National Academics Science of the USA 106: 20115-20120.

Allorent G., Osorio S, Vu J.L., Falconet D., Jouhet J., Kuntz M., Fernie A.R., Lerbs-Mache S., Macherel D., Courtois F., Finazzi G. (2015) Adjustments of embryonic photosynthetic activity modulate seed fitness in Arabidopsis thaliana. New Phytologist 205(2): 707-719.

Andre C., Benning C. (2007) Arabidopsis seedlings deficient in a plastidic pyruvate kinase are unable to utilize seed storage compounds for germination and establishment. Plant Physiology 145 (4): 1670-1680.

Andriotis V.M., Kruger N.J., Pike M.J., Smith A.M (2010) Plastidial glycolysis in developing Arabidopsis embryos. New Phytologist 185(3): 649-662.

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

Angelovici R., Fait A., Zhu X., Szymanski J., Feldmesser E., Fernie A.R., Galili G. (2009) Deciphering transcriptional and metabolic networks associated with lysine metabolism during Arabidopsis seed development. Plant Physiology 151(4): 2058-2072.

Araujo W.L., Ishizaki K., Nunes-Nesi A., Larson T.R., Tohge T., Krahnert I., Bauwe H., Hagemann M., Fernie A.R. (2010) Photorespiration: players, partners and origin. Trends in Plant Science 15: 330-336.

Araujo W.L., Tohge T., Ishizaki K., Leaver C.J., Fernie A.R. (2011) Protein degradation-an alternative respiratory substrate for stressed plants. Trends in Plant Science 16:489-498.

Asokanthan P.S., Johnson R.W., Griffith M., Krol M. (1997) The photosynthetic potential of canola embryos. Physiologia Plantarum 101: 353-360.

Baud S., Boutin J-P., Miquel M., Lepiniec L., Rochat C. (2002) An integrated overview of seed development in Arabidopsis thaliana ecotype Ws. Plant Physiology Biochemistry 40: 151-160.

Baud S., Lepiniec L. (2009) Regulation of de novo fatty acid synthesis in maturing oilseeds of Arabidopsis. Plant Physiology and Biochemistry 47(6): 448-455.

Bauer M. J., Fischer R. L. (2011) Genome demethylation and imprinting in the endosperm. Current Opinion in Plant Biology 14(2): 162-167.

Belmonte M.F., Kirkbride R.C., Stone S.L., Pelletier J.M., Bui A.Q., Yeung E.C., Hashimoto M., Fei J., Harada C.M., Munoz M.D., Le B.H., Drews G.N., Brady S.M., Goldberg R.B., Harada J.J. (2013) Comprehensive developmental profiles of gene activity in regions and subregions of the Arabidopsis seed. Proceedings of the National Academy of Sciences 110: 435-444.

Berger F. (1999) Endosperm development. Current Opinion in Plant Biology 2: 28-32.

Berger F. (2003) Endosperm: The crossroad of seed metabolism. Current Opinion in Plant Biology 6 (1): 42-50.

Berger F. (2004) Imprinting-a green variation. Science 303: 483-485.

Berger F., Grini P. E., Schnittger A. (2006) Endosperm: an integrator of seed growth and development. Current Opinion in Plant Biology 9: 664-670.

Berth M., Moser F. M., Kolbe M., Bernhardt J. (2007) The state of the art in the analysis of two-dimensional gel electrophoresis images. Applied Microbiology and Biotechnology 76(6): 1223-1243

Bethke, P.C., Libourel, I.G., Aoyama, N., Chung, Y.Y., Still, D.W. and Jones, R.L. (2007) The Arabidopsis aleurone layer responds to nitric oxide, gibberellin, and abscisic acid and is sufficient and necessary for seed dormancy. Plant Physiology 143: 1173-1188

Blanvillain R., Young B., Cai Y.M., Hecht V., Varoquaux F., Delorme V., Lancelin J.M., Delseny M., Gallois P. (2011) The Arabidopsis peptide kiss of death is an inducer of programmed cell death. The EMBO Journal 30: 1173-1183.

Boisnard-Lorig C., Colon-Carmona A., Bauch M., Hodge S., Doerner P., Bancharel E., Dumas C., Haseloff J., Berger F. (2001) Dynamic analyses of the expression of the HISTONE::YFP fusion protein in arabidopsis show that syncytial endosperm is divided in mitotic domains. Plant Cell. 13(3): 495-509.

Borisjuk L., Neuberger T., Schwender J., Heinzel N., Sunderhaus S., Fuchs J., Hay J.O., Tschiersch H., Braun H.P., Denolf P., Lambert B., Jakob P.M., Rolletschek H. (2013) Seed architecture shapes embryo metabolism in oilseed rape. Plant Cell 25: 1625-1640.

Borisjuk L., Nguyen T.H., Neuberger T., Rutten T., Tschiersch H., Claus B., Feussner I., Webb A.G., Jakob P., Weber H., Wobus U., Rolletschek H. (2005) Gradients of lipid storage, photosynthesis and plastid differentiation in developing soybean seeds. New Phytologist 167(3): 761-776.

Borisjuk L., Rolletschek H. (2009) The oxygen status of the developing seed. New Phytologist 182(1): 17-30.

Borisjuk L., Rolletschek H., Radchuk R., Weschke W., Wobus U., Weber H. (2004) Seed development and differentiation: a role for metabolic regulation. Plant Biology 6(4): 375-386.

Brown R. C., Lemmon B. E., Nguyen H. (2003) Events during the first four rounds of mitosis establish three developmental domains in the syncytial endosperm of Arabidopsis thaliana. Protoplasma 222(3-4): 167-174.

Brown R. C., Lemmon B. E., Nguyen H., Olsen O. A. (1999) Development of endosperm in Arabidopsis thaliana. Plant Reproduction 12(1): 32-42.

Cairns N.G., Pasternak M., Wachter A., Cobbett C.S., Meyer A.J. (2006) Maturation of Arabidopsis seeds is dependent on glutathione biosynthesis within the embryo. Plant Physiology 141: 446-455.

Cavel E., Pillot M., Pontier D., Lahmy S., Bies-Etheve N., Vega D., Grimanelli D., Lagrange T. (2011) A Plant-Specific Transcription Factor IIB-Related Protein, pBRP2, Is Involved in Endosperm Growth Control. PLoS ONE 6 (2): e17216.

Chen M., Thelen J. J. (2010) The plastid isoform of triose phosphate isomerase is required for the postgerminative transition from heterotrophic to autotrophic growth in Arabidopsis. Plant Cell 22(1): 77-90.

Chourey P.S., Taliercio E.W., Carlson S.J., Ruan Y.-L. (1998) Genetic evidence that the two isozymes of sucrose synthase present in developing maize endosperm are critical, one for cell wall integrity and the other for starch biosynthesis. Molecular Genomics and Genetics 259: 88-96.

Chrispeels M. J., Higgins T. J., Spencer D. (1982) Assembly of storage protein oligomers in the endoplasmic reticulum and processing of the polypeptides in the protein bodies of developing pea cotyledons. The Journal of Cell Biology 93: 306-313.

Colditz F., Nyamsuren O., Niehaus K., Eubel H., Braun H.P., Krajinski F. (2004) Proteomic approach: identification of Medicago truncatula proteins differentially expressed after infection with the pathogenic oomycete Aphanomyces euteiches. Plant Molecular Biology 55: 109-120.

Considine M.J., Foyer C.H. (2014) Redox Regulation of Plant development. ANTIOXIDANT & REDOX SIGNALING 21(9): 1305-1326.

Costa L.M., Marshall E., Tesfaye M., Silverstein K.A., Mori M., Umetsu Y., Otterbach S.L., Papareddy R., Dickinson H.G., Boutiller K., VandenBosch K.A., Ohki S., Gutierrez-Marcos J.F (2014) Central cell-derived peptides regulate early embryo patterning in flowering plants. Science 344(6180): 168-172.

Cox J., Mann M. (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nature Biotechnology, 26: 1367-1372.

Cox, J., Hein M.Y., Luber C.A., Paron I., Nagaraj N. and Mann M. (2014) Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. Mol Cell Proteomics 13(9): 2513-2526.

Credali A., García-Calderón M., Dam S., Perry J., Díaz-Quintana A., Parniske M., Wang T.L., Stougaard J., Vega J.M., Márquez A.J. (2013) The K+-dependent asparaginase, NSE1, is crucial for plant growth and seed production in Lotus japonicus. Plant Cell Physiology 54(1): 107-118.

Crouch M. L., Sussex I. M. (1981) Development and storage-protein synthesis in Brassica napus L. embryos in vivo and in vitro. Planta 153: 64-74.

Dante R.A., Larkins B.A., Sabelli P.A. (2014) Cell cycle control and seed development. Frontiers in Plant Science 5:493.

De Smet I., Lau S., Mayer U., Jürgens G. (2010) Embryogenesis-the humble beginnings of plant life. Plant Journal 61(6): 959-970.

Dickinson C. D., Hussein E. H., Nielsen N. C. (1989) Role of posttranslational cleavage in glycinin assembly. Plant Cell 1: 459-469.

Dorion S., Clendenning A., Jeukens J., Salas J. J., Parveen N., Haner A.A., Law R. D., Force E. M., Rivoal J. (2012) A large decrease of cytosolic triosephosphate isomerase in transgenic potato roots affects the distribution of carbon in primary metabolism. Planta 236(4): 1177-1190.

Eastmond P., Rawsthorne S. (1998) Comparison of the metabolic properties of plastids isolated from developing leaves or embryos of Brassica napus. Journal of Experimental Botany 49: 1105-1111.

Ereken-Tumer N., Richter J. D., Nielsen N. C. (1982) Structural characterization of the glycinin precursors. J. Biol. Chem. 257: 4016-4018.

Eubel H., Meyer E.H., Taylor N.L., Bussell J.D., O'Toole N., Heazlewood J.L., Castleden I., Small I.D., Smith S.M., Millar A.H. (2008) Novel proteins, putative membrane transporters, and an integrated metabolic network are revealed by quantitative proteomic analysis of Arabidopsis cell culture peroxisomes. Plant Physiology 148(4): 1809-1829.

Fait A., Angelovici R., Less H., Ohad I., Urbanczyk-Wochniak E., Fernie. A.R., Galili G. (2006) Arabidopsis seed development and germination is associated with temporally distinct metabolic switches. Plant Physiology (142): 839-854.

Fallahi H., Scofield G.N., Badger M.R., Chow W.S., Furbank R.T., Ruan Y-L. (2008) Localization of sucrose synthase in developing seed and siliques of Arabidopsis thaliana reveals diverse roles for SUS during development. Journal of Experimental Botany 59: 3283-3295.

Fang W., Wang Z., Cui R. Li J., Li Y. (2012) Maternal control of seed size by EOD3/CYP78A6 in Arabidopsis thaliana. Plant Journal 70: 929-939.

Fiume E., Monfared M., Jun J., Fletcher J.C (2011) CLE polypeptide signaling gene expression in Arabidopsis embryos. Plant Signaling & Behavior 6: 443-444.
Galili G., Avin-Wittenberg T., Angelovici R., Fernie A.R. (2014) The role of photosynthesis and amino acid metabolism in the energy status during seed development. Frontiers in Plant Science 5: 447

Gallardo K., Thompson R.D., Burstin J. (2008) Reserve accumulation in legume seeds. Comptes Rendus Biologies 331: 755-762.

Garcia D. (2003) Arabidopsis haiku Mutants Reveal New Controls of Seed Size by Endosperm. Plant Physiology 131(4): 1661-1670.

Gehring M., Choi Y., Fischer R.L. (2004). Imprinting and seed development. Plant Cell 16: 203-213.

Geserick C., Tenhaken R. (2013). UDP-sugar pyrophosphorylase is essential for arabinose and xylose recycling, and is required during vegetative and reproductive growth in Arabidopsis. The Plant Journal 74 (2): 239-247.

Goffman F.D., Alonso A.P., Schwender J., Shachar-Hill Y., Ohlrogge J.B. (2005) Light enables a very high efficiency of carbon storage in developing embryos of rapeseed. Plant Physiology 138(4): 2269-2279.

Goldberg R.B., De Paiva G., Yadegari R. (1994) Plant embryogenesis: zygote to seed. Science 266: 605-614.

Gu L., Jones A.D., Last R.L. (2010) Broad connections in the Arabidopsis seed metabolic network revealed by metabolite profiling of an amino acid catabolism mutant. Plant Journal 61: 579-590.

Hajduch M., Casteel J.E., Tang S., Hearne L.B., Knapp S., Thelen J.J. (2007) Proteomic analysis of near-isogenic sunflower varieties differing in seed oil traits. Journal of Proteome Research 6(8): 3232-3241.

Hajduch M., Matusova R., Houston N.L., Thelen J.J. (2011) Comparative proteomics of seed maturation in oilseeds reveals differences in intermediary metabolism. Proteomics 11(9):1619-1629.

Hanson J., Smeekens S. (2009) Sugar perception and signaling-an update. Current Opinion in Plant Biology 12(5): 562-567.

Haughn G., Chaudhury A. (2005) Genetic analysis of seed coat development in Arabidopsis. Trends in Plant Science 10: 472-477.

Hehenberger E., Kradolfer D., Köhler C. (2012) Endosperm cellularization defines an important developmental transition for embryo development. Plant Cell Physiology 53(1): 16-27.

Hildebrandt T.M., Nesi A.N., Araújo W.L., Braun H.P. (2015) Amino acid catabolism in plants. Molecular Plant pii: S1674-2052(15)00366-4.

Hill L. M., Morley-Smith E. R., Rawsthorne S. (2003) Metabolism of sugars in the endosperm of developing seeds in oilseed rape. Plant Physiology 131(1): 228-236.

Hills M. (2004) Control of storage-product synthesis in seeds. Current opinion in Plant Biology 7: 302-308.

Höglund A.S., Rödin J., Larsson E., Rask L. (1992) Distribution of napin and cruciferin in developing rapeseed embryo. Plant Physiology 98: 509-515.

Holdorf M.M., Owen H.A., Lieber S.R., Yuan L., Adams N., Dabney-Smith C., Makaroff C.A. (2012) Arabidopsis ETHE1 encodes a sulfur dioxygenase that is essential for embryo and endosperm development. Plant Physiology 160(1): 226-236.

Holdsworth, M.J., Bentsink, L. and Soppe, W.J. (2008) Molecular networks regulating Arabidopsis seed maturation, after-ripening, dormancy and germination. New Phytologist 179: 33-54.

Hsu SC1, Belmonte MF, Harada JJ, Inoue K. (2010) Indispensable Roles of Plastids in Arabidopsis thaliana Embryogenesis. Current Genomics 11(5): 338-349.

Huang A.H.C. (1992) Oil bodies and oleosins in seeds. Annual Review of Plant Physiology and Plant Molecular Biology, 43: 177-200.

Huang Y., Chen L., Wang L., Vijayan K., Phan S., Liu Z., Wan L., Ross A., XiangD., Datla R., Pan Y., Zou J. (2009) Probing the endosperm gene expression landscape in Brassica napus. BMC Genomics 10:256.

Huh J. H., Bauer M. J., Hsieh T-F.; Fischer R. (2007) Endosperm gene imprinting and seed development. Current Opinion in Genetics & Development 17 (6): 480-485.

Ingram G., Gutierrez-Marcos J. (2015) Peptide signalling during angiosperm seed development. Journal of Experimental Botany 66(17): 5151-5159.

Jung R., Scott M. P., Nam Y. W., Beaman T. W., Bassüner R., Saalbach I., Müntz K., Nielsen N. C. (1998) The role of proteolysis in the processing and assembly of 11S seed globulins. Plant Cell 10: 343-357.

Junker B.H., Lonien J., Heady L.E., Rogers A., Schwender J. (2007) Parallel determination of enzyme activities and in vivo fluxes in Brassica napus embryos grown on organic or inorganic nitrogen source. Phytochemistry 68: 2232-2242.

Kawashima T., Goldberg R.B. (2010) The suspensor: not just suspending the embryo. Trends in Plant Science 15(1): 23-30.

Ke J., Behal R. H., Back S. L., Nikolau B. J., Wurtele E. S., Oliver D. J. (2000) The role of pyruvate dehydrogenase and acetyl-coenzyme A synthetase in fatty acid synthesis in developing Arabidopsis seeds. Plant Physiology 123(2): 497-508.

Keech O., Dizengremel P., Gardestrom P. (2005) Preparation of leaf mitochondria from Arabidopsis thaliana. Physiologica Plantarum 124: 403-409.

Kelliher T., Walbot V. (2012) Hypoxia triggers meiotic fate acquisition in maize. Science 337(6092): 345-348.

Kim I., Kobayashi K., Cho E., Zambryski P.C. (2005) Subdomains for transport via plasmodesmata corresponding to the apical-basal axis are established during Arabidopsis embryogenesis. Proceedings of the National Academy of Sciences of the USA (102): 11945-11950.

Kim M., Lim J.H., Chang S.A., Park K., Kim G.T., Kim W.T., Pai H.S. (2006). Mitochondria-associated hexokinases play a role in the control of programmed cell death in Nicotiana benthamiana. Plant Cell 18: 2341-2355.

Krishnan H.B., Oehrle N.W., Natarajan S.S. (2009) A rapid and simple procedure for the depletion of abundant storage proteins from legume seeds to advance proteome analysis: a case study using Glycine max. Proteomics 9(11): 3174-3188.

Krüßel L., Junemann J., Wirtz M., Birke H., Thornton J.D., Browning L.W., Poschet G., Hell R., Balk J., Braun H.P., Hildebrandt T.M. (2014) The mitochondrial sulfur dioxygenase ETHYLMALONIC ENCEPHALOPATHY PROTEIN1 is required for amino acid catabolism during carbohydrate starvation and embryo development in Arabidopsis. Plant Physiology 165(1): 92-104.

Kubis S.E., Lilley K.S., Jarvis P. (2008) Isolation and preparation of chloroplasts from Arabidopsis thaliana plants. Methods in Molecular Biology 425: 171-186.

Lafon-Placette C., Köhler C. (2014) Embryo and endosperm, partners in seed development. Current Opinion in Plant Biology 17: 64-69.

Lau S., Slane D., Herud O., Kong J., Jurgens G. (2012) Early embryogenesis in flowering plants: setting up the basic body pattern. Annu Rev Plant Biol 63:483-506.

Le B. H., Cheng C., Bui A. Q., Wagmaister J. A., Henry K. F., Pelletier J., Kwong L., Belmonte M., Kirkbride R., Horvath S., Drews G. N., FischerR. L., Okamuro K. K., Harada J. J., Goldberg R. B. (2010) Global analysis of gene activity during Arabidopsis seed development and identification of seed-specific transcription factors. Proceedings of the National Academy of Sciences of the USA 107(18): 8063-8070.

Lee H., Guo Y., Ohta M., Xiong L., Stevenson B., Zhu J.-K. (2002) LOS2, a genetic locus required for cold-responsive gene transcription encodes a bi-functional enolase. EMBO Journal 21(11): 2692-2702.

Lee Y.I., Yeung E.C. (2010) The osmotic property and fluoreszent tracer movement of developing orchid embryos of Phaius tankervilliae (Aiton) BI. Sexual Plant Production 23: 337-341.

Lee, K.J.D., Dekkers, B.J.W., Steinbrecher, T., Walsh, C.T., Bacic, A., Bentsink, L. Leubner-Metzger G., Knox J.P. (2012) Distinct cell wall architectures in seed endosperms in representatives of the Brassicaceae and Solanaceae. Plant Physiology 160: 1551-1566.

Lin M., Oliver D.J. (2008) The role of acetyl-Coenzyme A synthetase in Arabidopsis. Plant Physiology 147: 1822-1829.

Lorenz, C., Rolletschek, H., Sunderhaus, S. and Braun, H.P. (2014) B. napus seed endosperm – metabolism and signaling in a dead end tissue. Journal of Proteomics 108: 382-426.

Lu J., Zhang C., Baulcombe D. C., Chen Z. J. (2012) Maternal siRNAs as regulators of parental genome imbalance and gene expression in endosperm of Arabidopsis seeds. Proceedings of the National Academy of Sciences of the USA 109(14): 5529-5534.

Macherel D., Benamar A., Avelange-Macherel M-H. Tolleter D. (2007) Function and stress tolerance of seed mitochondria. 129(1): 233-241.

Melkus G., Rolletschek H., Radchuk R., Fuchs J., Rutten T., Wobus U., Altmann T., Jakob P., Borisjuk L. (2009) The metabolic role of the legume endosperm: A noninvasive imaging study. Plant Physiology 151: 1139-1154.

Millar A.H., Whelan J., Soole K.L., Day D.A. (2011) Organization and regulation of mitochondrial respiration in plants. Annual Review of Plant Biology 62:79-104.

Moreley-Smith E.R., Pike M.J., Findlay K., Köckenberger W., Hill L.M., Smith A. M., Rawsthorne S. (2008) The Transport of Sugars to Developing Embryos Is Not via the Bulk Endosperm in Oilseed Rape Seeds. Plant Physiology 147: 2121-2130.

Muñoz-Bertomeu J., Cascales-Miñana B., Irles-Segura A., Mateu I., Nunes-Nesi A., Fernie A.R., Segura J., Ros R. (2010) The plastidial glyceraldehyde-3-phosphate dehydrogenase is critical for viable pollen development in Arabidopsis. Plant Physiology 152(4): 1830-1841.

Muñoz-Bertomeu J., Cascales-Miñana B., Mulet J. M., Baroja-Fernández E., Pozueta-Romero J., Kuhn J. M., Segura J., Ros R. (2009) Plastidial glyceraldehyde-3-phosphate dehydrogenase deficiency leads to altered root development and affects the sugar and amino acid balance in Arabidopsis. Plant Physiology 151(2): 541-558.

Murphy D.J., Cummins I. (1989) Biosynthesis of seed storage products during embryogenesis in rapeseed, Brassica napus. Plant Physiology 135: 63-69.

Mwangi J.W., Rode C., Colditz F., Haase C., Braun H.P., Winkelmann T. (2013) Proteomic and histological analyses of endosperm development in Cyclamen persicum as a basis for optimization of somatic embryogenesis. Plant Science 201(20): 52-65.

Nilsson T., Mann M., Aebersold R., Yates J.R. 3rd Bairoch A., Bergeron J.J. (2010) Mass spectrometry in high-throughput proteomics: ready for the big time. Nature Methods 7(9): 681-685.

Nietzel T., Dudkina N.V., Haase C., Denolf P., Semchonok D., Boekema E.J., Braun H.P., Sunderhaus, S. (2013) The native structure and composition of the cruciferin complex in Brassica napus. Journal of Biological Chemistry 288: 2238-2245.

Noah, A.M., Niemenak, N., Sunderhaus, S., Haase, C., Omokolo1, D.N., Winkelmann, T. and Braun, H.P. (2013) Comparative proteomic analysis of early somatic and zygotic embryogenesis in Theobroma cacao L. Journal of Proteomics 78: 123-133.

Novack M.K., Unguru A., Bjerkan K.N., Grini P.E., Schnittger A. (2010) Reproductive cross-talk: seed development in flowering plants. Biochemical Society Transactions, 38: 604-612.

O'Hara L.E., Paul M.J, Wingler A. (2013) How do sugars regulate plant growth and development? New insight into the role of trehalose-6-phosphate. Molecular Plant 6(2): 261-274.

Ohlrogge J.B., Jaworski J.G. (1997) REGULATION OF FATTY ACID SYNTHESIS. Annual Review of Plant Physiology and Plant Molecular Biology 48: 109-136.

Ohto M. A., Fischer R. L., Goldberg R. B., Nakamura K., Harada J. J. (2005) Control of seed mass by APETALA2. Proceedings of the National Academy of Sciences 102: 3123-3128.

Olsen O.-A. (2004) Nuclear Endosperm Development in Cereals and Arabidopsis thaliana. The Plant Cell 16: 214-222.

Old W.M., Meyer-Arendt K., Aveline-Wolf L., Pierce K.G., Mendoza A., Sevinsky J.R., Resing K.A., Ahn N.G. (2005) Comparison of label-free methods for quantifying human proteins by shotgun proteomics. Molecular Cell Proteomics 4(10): 1487-1502.

Orozco-Arroyo G., Paolo D., Ezquer I., Colombo L. (2015) Networks controlling seed size in Arabidopsis. Plant Reproduction 28(1): 17-32.

Penfield, S., Rylott, E.L., Gilday, A.D., Graham, S., Larson, T.R. and Graham, I.A. (2004) Reserve mobilization in the Arabidopsis endosperm fuels hypocotyl elongation in the dark, is independent of abscisic acid, and requires PHOSPHOENOLPYRUVATE CARBOXYKINASE1. Plant Cell 16: 2705-2718.

Picotti P., Bodenmiller B., Aebersold R. (2013) Proteomics meets the scientific method. Nature Methods 10(1): 24-27.

Plaxton W.C. (1996) THE ORGANIZATION AND REGULATION OF PLANT GLYCOLYSIS. Annual Review of Plant Physiology and Plant Molecular Biology 47: 85-214.

Pugh D.A., Offler C.E., Talbot M.J., Ruan Y-L. (2010) Evidence for the role of transfer cells in the evolutionary increase of seed and fiber biomass yield in cotton. Molecular Plant 3: 1075-1086.

Rawsthorne S. (2002) Carbon flux and fatty acid synthesis in plants. Progress in Lipid Research 41: 182-196.

Rodoeva R., Weijers D. (2014) A Roadmap to embryo identity in plants. Trends in Plant Science 19 (11): 709-716.

Rolland F., Baena-Gonzalez E., Jen S. (2006) SUGAR SENSING AND SIGNALING IN PLANTS: Conserved and Novel Mechanisms. Annual Review of Plant Biology 57(1): 675-709.

Rolletschek H., Radchuk R., Klukas C., Schreiber F., Wobus U., Borisjuk L.(2005) Evidence of a key role for photosynthetic oxygen release in oil storage in developing soybean seeds. New Phytologist 167(3): 777-786.

Ruan Y-L., Llewellyn D.J., Liu Q., Xu S.M., Wu L.M., Wang L., Furbank R.T. (2008) Expression of sucrose synthase in the developing endosperm is essential for early seed development in cotton. Functional Plant Biology 35: 382-393.

Ruuska S. A., Schwender J., Ohlrogge J. B. (2004) The Capacity of Green Oilseeds to Utilize Photosynthesis to Drive Biosynthetic Processes. Plant Physiology 136(1): 2700-2709.

Ruuska S.A., Girke T., Benning C., Ohlrogge J.B. (2002) Contrapuntal networks of gene expression during Arabidopsis seed filling. Plant Cell 14(6): 1191-1206.

Sabelli P.A., Larkins B.A. (2009) The development of endosperm in grasses. Plant Physiology 149(1): 14-26.

Schiebold S., Tschiersch H., Borisjuk L., Heinzel N., Radchuk R., Rolletschek H. (2011) A novel procedure for the quantitative analysis of metabolites, storage products and transcripts of laser microdissected seed tissues of Brassica napus. Plant Methods 7:19.

Schwender J., Goffman F., Ohlrogge J.B., Shachar-Hill Y. (2004) RubisCO without the Calvin cycle improves the carbon efficiency of developing green seeds. Nature 432: 779-782.

Schwender J., Ohlrogge J.B., Shachar-Hill Y. (2003) A flux model of glycolysis and the oxidative pentosephosphate pathway in developing Brassica napus embryos. The Journal of Biological Chemistry 278(32): 29442-29453.

Schwender J., Shachar-Hill Y., Ohlrogge J. B. (2006) Mitochondrial Metabolism in Developing Embryos of Brassica napus. Journal of Biological Chemistry 281(45): 34040‑34047.

Sebei K., Ouerghi Z., Kallel H., Boukhchina S. (2006) Evolution of phosphoenolpyruvate carboxylase activity and lipid content during seed maturation of two spring rapeseed cultivars (Brassica napus L.) Comptes Rendus Biologies 329: 719-725.

Sengupta C., Deluca V., Bailey D. S., Verma D. P. (1981) Post-translational processing of 7S and 11S components of soybean storage proteins. Plant Molecular Biology 1:19-34.

Sharafia Y., Majidia M.M., Golib S.A.H., Rashidia F. (2015) Oil Content and Fatty Acids Composition in Brassica Species. INTERNATIONAL JOURNAL OF FOOD PROPERTIES 18(10): 2145-2154.

Shearer H.L., Turpin D.H., Dennis D.T. (2004) Characterization of NADP-dependent malic enzyme from developing castor oil seed endosperm. Archieves Biochemistry Biophysics 429: 134-144.

Shimada T., Yamada K., Kataoka M., Nakaune S., Koumoto Y., Kuroyanagi M., Tabata S., Kato T., Shinozaki K., Seki M., Kobayashi M., Kondo M., Nishimura M., Hara-Nishimura I. (2003) Vacuolar processing enzymes are essential for proper processing of seed storage proteins in Arabidopsis thaliana. Journal of Biological Chemistry 278: 32292-32299.

Sørensen, M. B. (2002) Cellularization in the endosperm of Arabidopsis thaliana is coupled to mitosis and shares multiple components with cytokinesis. Development 129(24): 5567‑5576.

Thelen J.J., Ohlrogge B. (2002) Metabolic Engineering of Fatty Acid Biosynthesis in Plants. Metabolic Engineering 4: 12-21.

Thiel J. (2014) Development of endosperm transfer cells in barley. Frontiers in Plant Science 5:108 Proceedings National Academic Science of the USA 107(18): 8063-8070.

Vigeolas H., van Dongen J.T., Waldeck P., Huhn D., Geigenberger P. (2003) Lipid storage metabolism is limited by the prevailing low oxygen concentrations within developing seeds of oilseed rape. Plant Physiology 133(4): 2048-2060.

Voelker T.A., Kinney A.J. (2001) Variations in the biosynthesis of seed storage lipids. Annual Review of Plant Physiology and Molecular Biology, 52: 261-335.

Wang L., Ruan Y-L. (2013) Regulation of cell division and expansion by sugar and auxin signaling. Frontiers in Plant Science 4(163): 1-9.

Wang W.Q., Liu S.J., Song S.Q., Møller I.M. (2014) Proteomics of seed development, desiccation tolerance, germination and vigor. Plant Physiology and Biochemistry 86: 1-15.

Weber H., Borisjuk L., Wobus U. (2005) Molecular physiology of legume seed development. Annual Review of Plant Biology 56: 253-279.

Weigelt K., Küster H., Radchuk R., Müller M., Weichert H., Fait A., Fernie A.R., Saalbach I., Weber H. (2008) Increasing amino acid supply in pea embryos reveals specific interactions of N and C metabolism, and highlights the importance of mitochondrial metabolism. Plant Journal 55(6): 909-1026.

Wendrich J.R., Weijers D. (2013) The Arabidopsis embryo as a miniature morphogenesis model. New Phytologist 199: 14-25.

Wilkins M.R., Sanchez J.C., Gooley A.A., Appel R.D., Humphery-Smith I., Hochstrasser D.F., Williams K.L. (1996) Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it. Biotechnology & Genetic Engineering Reviews 13: 19-50.

Wolff P., Weinhofer I., Seguin J., Roszak P., Beisel C., Donoghue M.T., Spillane C., Nordborg M., Rehmsmeier M., Köhler C. (2011) High-resolution analysis of parent-of-origin allelic expression in the Arabidopsis Endosperm. PLoS Genetics 7 (6): e1002126.

Xu S.M., Brill E., Llewellyn D.J., Furbank R.T., Ruan Y-L. (2012) Overexpression of a potato sucrose synthase gene in cotton accelerates leaf expansion, reduces seed abortion, and enhances fiber production. Molecular Plant 5: 430-441.

Xu X.M., Jackson D. (2010) Lights at the end of a tunnel: new views of plasmodesmal structure and function. Current Opinion Plant Biologie (13): 684-692.

Yan D., Duermeyer L., Leoveanu C., Nambara E. (2015) The functions of the endosperm during seed germination. Plant Cell Physiol. 55(9): 1521-1533.

Young, T. E., Gallie, D. R. (2000) Programmed cell death during endosperm development. Plant Molecular Biology 44 (3): 283-301.

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 W., Jeffrey W. Smith, Chun-Ming Huang (2010) Mass Spectrometry-Based Label-Free Quantitative Proteomics. Journal of Biomedicine and Biotechnology, Article ID 840518, 6 pages.

Chapter 2: Publications and Manuscripts

Publication 1

2.1 *B. napus* seed endosperm – metabolism and signaling in a dead end tissue

Christin Lorenz[1], Hardy Rolletschek[2], Stephanie Sunderhaus[1], Hans-Peter Braun[1]

[1] Department of Plant Proteomics, Institute for Plant Genetics, Faculty of Natural Sciences, Leibniz Universität Hannover

[2] Department of Molecular Genetics, Leibniz Institute of Plant Genetics and Crop Plant Research (IPK) Gatersleben

Approximately 50 % of this work was obtained during my M. Sc thesis

Type of authorship:	First author
Type of article:	Research article
Share of the work:	85 %
Contribution to the publication:	Planned and performed all experiments, analysed data, prepared all figures and wrote the paper
Journal:	Journal of Proteomics
Impact factor:	3.888
Date of publication:	Published in August 2014
Number of citations: (Google scholar, Nov. 07th, 2015)	1
DOI:	10.1016/j.jprot.2014.05.024
PubMed-ID:	24906024

JOURNAL OF PROTEOMICS 108 (2014) 382–426

Available online at www.sciencedirect.com

ScienceDirect

www.elsevier.com/locate/jprot

Brassica napus seed endosperm — Metabolism and signaling in a dead end tissue

Christin Lorenz[a], *Hardy Rolletschek*[b], *Stephanie Sunderhaus*[a], *Hans-Peter Braun*[a,*]

[a]*Institute of Plant Genetics, Faculty of Natural Sciences, Leibniz Universität Hannover, 30419 Hannover, Germany*
[b]*Department of Molecular Genetics, Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Corrensstr. 3, D-06466 Gatersleben, Germany*

ARTICLE INFO

Article history:
Received 1 March 2014
Accepted 27 May 2014
Available online 4 June 2014

Keywords:
Brassica napus
Endosperm
Proteomics
Mass spectrometry
Seed development
Embryo

ABSTRACT

Oilseeds are an important element of human nutrition and of increasing significance for the production of industrial materials. The development of the seeds is based on a coordinated interplay of the embryo and its surrounding tissue, the endosperm. This study aims to give insights into the physiological role of endosperm for seed development in the oilseed crop *Brassica napus*. Using protein separation by two-dimensional (2D) isoelectric focusing (IEF)/SDS polyacrylamide gel electrophoresis (PAGE) and protein identification by mass spectrometry three proteome projects were carried out: (i) establishment of an endosperm proteome reference map, (ii) proteomic characterization of endosperm development and (iii) comparison of endosperm and embryo proteomes. The endosperm proteome reference map comprises 930 distinct proteins, including enzymes involved in genetic information processing, carbohydrate metabolism, environmental information processing, energy metabolism, cellular processes and amino acid metabolism. To investigate dynamic changes in protein abundance during seed development, total soluble proteins were extracted from embryo and endosperm fractions at defined time points. Proteins involved in sugar converting and recycling processes, ascorbate metabolism, amino acid biosynthesis and redox balancing were found to be of special importance for seed development in *B. napus*. Implications for the *seed filling* process and the function of the endosperm for seed development are discussed.

Biological significance
The endosperm is of key importance for embryo development during seed formation in plants. We present a broad study for characterizing endosperm proteins in the oilseed plant *B. napus*. Furthermore, a project on the biochemical interplay between the embryo and the endosperm during seed development is presented. We provide evidence that the endosperm includes a complete set of enzymes necessary for plant primary metabolism. Combination of our results with metabolome data will further improve systems-level understanding of the *seed filling* process and provide rational strategies for plant bioengineering.

* *Corresponding author*. Tel.: +49 511 7622674; fax: +49 511 7623608.
E-mail address: braun@genetik.uni-hannover.de (H.-P. Braun).

http://dx.doi.org/10.1016/j.jprot.2014.05.024
1874-3919/

1. Introduction

The endosperm is a central compartment of the developing plant seed. It embeds the embryo and is surrounded by the seed coat. Seed development is initiated by a double fertilization: (i) fertilization of an egg cell by a sperm cell, which leads to the embryo and (ii) fertilization of a diploid central cell by another sperm cell, which gives rise to the triploid endosperm.

Seed formation of dicotyledonous species such as *Arabidopsis thaliana* and *Brassica napus* is characterized first by a stepwise differentiation of the endosperm including transition from a syncytial to a cellular phase [1–4] followed by embryo development. The formation of a syncytium is initiated by several rounds of mitosis of the fertilized central cell without cytokinesis leading to a multinuclear cell [1]. In the course of seed development the endosperm is divided into three mitotic domains: (1) the micropylar, (2) the peripheral and (3) the chalazal region in which nuclei divide simultaneously [5]. The syncytium becomes cellularized by the formation of radial microtubule systems starting from the micropylar region [1,4,6]. The process of cellularization has a great impact on the accurate embryo development [7] and leads to the initiation of cotyledon growth [4]. After cellularization the outermost cells of the endosperm differentiate into the aleurone layer. The characteristics of this cell type are thick cell walls and the deposition of storage compounds and enzymes [8]. In dicotyledons the endosperm is of transient nature and only the aleurone layer remains until seed maturity whereas the bulk of the endosperm is replaced by the growing embryo. Programmed cell death leads to the degradation of endosperm cells and is probably controlled by ethylene and interactions between embryo and endosperm [9–11]. The interplay of different seed compartments on various molecular and genetic levels is crucial for coordinated seed development [12–16].

Investigations of mutants carrying endosperm specific defects have given insights into development and function of this compartment [7,17–21]. Most of these studies were focused on plant species exhibiting a constitutive endosperm like cereals. In these monocotyledonous plants, the endosperm has a prime role in storage compound accumulation [22–28], while in dicotyledonous species the endosperm rather seems to play a transient role. It nourishes the embryo and supports development and growth at early seed developmental stages [4,9,29,30]. Here, storage compounds are rather accumulated in the cotyledons [31]. The endosperm is the main source of carbohydrates and amino acids for the growing embryo [32]. At very early stages of seed development sucrose imported from the mother plant is cleaved into hexoses in the endosperm and provided to the embryo. Accumulation of hexoses subsequently drives water and nutrient uptake of the seed [33]. During later stages of development sucrose accumulates in the seeds. The switch from high hexose to high sucrose level is related to differentiation, cell expansion and storage activity in embryo [34,35]. The endosperm represents the direct environment of the dicot embryo and ensures constant and optimal conditions for its differentiation and growth [32]. However, the metabolic functions ensuring its evolutionary mission are just beginning to emerge [36–38]. It is particularly unknown to which extend endosperm metabolism can modulate the incoming stream of assimilates, and thereby affect differentiation, growth and storage functions of the embryo.

In this study a comparative proteomic strategy was applied to developing seeds of *B. napus*. This plant represents one of the most important oilseed crops. Based on (i) the establishment of an endosperm proteome reference map, (ii) the proteomic characterization of endosperm development, and (iii) the comparison of endosperm and embryo proteomes, we aimed to get insights into both physiology and metabolic functions of endosperm in dicot seeds.

2. Material and methods

2.1. Cultivation of *B. napus* seedlings and seed tissue preparation

B. napus plants were grown under the following conditions: 16-h light (16 klx) at 22 °C, 8-h dark at 18 °C with a relative humidity of 55%. Our experimental set up was the following: For each experiment, 10 rapeseed plants were used. Approximately 25 siliques were harvested per plant and approximately 30 seeds per silique (overall approximately 7500 seeds per experiment). Embryo and endosperm fractions were prepared manually. The procedure was repeated at three times points per fraction (endosperm: 10, 15 and 20 days after pollination [DAP], embryos 15, 20 and 25 DAP). All 10-plant experiments for all time points were carried out twice. The endosperm fractions were isolated with a microsyringe from the seeds. Embryos were isolated by removing the seed coat and subsequent washing to remove the potentially remaining endosperm. The endosperm was directly used for phenolic extraction of proteins. In contrast, the embryos were first pulverized using a bead mill.

2.2. Phenolic extraction of proteins and two dimensional (2D) IEF/SDS PAGE

Total soluble proteins were extracted from the endosperm and embryo fraction according to Colditz et al. [39]. Briefly, endosperm and pulverized embryo fractions were homogenized in extraction buffer (700 mM sucrose, 500 mM Tris, 50 mM EDTA, 100 mM KCl, 2% (v/v) β-mercaptoethanol and 2 mM PMSF, pH adjusted to 8.0). Then, saturated phenol (pH 6.6/7.9; Amresco, Solon, USA) was added. After several rounds of centrifugation proteins were precipitated with 100 mM ammonium acetate in methanol at −20 °C overnight. The resulting protein pellets were resuspended in resuspension buffer (8 M urea, 2 M thiourea, 2% (w/v) CHAPS, 100 mM DTT, 12 μl/ml DeStreak-reagent, 0.5% (v/v) IPG-buffer pH 3-11 NL, GE Healthcare, Freiburg, Germany) and directly loaded onto an IPG strip (24 cm, pH 3–11 NL, GE Healthcare, Freiburg, Germany). Isoelectric focusing was performed as described in Mihr and Braun [40]. Second gel dimensions (precast 12.5% polyacrylamide gels, Serva Electrophoreses, Heidelberg, Germany) were performed using the HPE FlatTop Tower system (Serva Electrophoresis, Heidelberg, Germany) according to the manufacturer's guidelines. 2D analyses for endosperm

and embryo fractions (three time points each) were repeated at least three times.

2.3. Gel staining procedure

All 2D gels were fixed with 10% (v/v) acetate in 40% (v/v) methanol for 45 min and stained with Coomassie blue CBB G-250 (Merck, Darmstadt, Germany) as described by Neuhoff et al. [41,42].

2.4. Quantitative gel analyses of endosperm and embryo proteomes

Coomassie colloidal stained gels were first scanned and then analyzed by using the Delta2D software 4.3 (Decodon, Greifswald, Germany) with three replicates for each seed tissue (endosperm and embryo) and developmental stage (10, 15, 20 DAP for endosperm and 15, 20, 25 DAP for embryo fractions) according to Berth et al. [43]. Spots were detected automatically (minor corrections of obvious gel disturbances were performed manually). To determine significant alterations in spot abundance of endosperm and embryo fractions, a Student's t-test (p-value ≤ 0.05) was applied on the basis of normalized relative spot volume. Alterations in protein volumes between the compared fractions $\geq$ factor 1.5 were considered to represent true differences on protein levels. Mean of normalized relative spot volume and coefficient of variation for each analyzed sample group are given in Supplementary Table 1. In order to analyze the similarity of the sample replicates a principal component analysis (PCA) was performed using Delta2D analysis following the manufacturer's instructions (for 3D plot of the samples see Supplementary Fig. 11). Additionally a heatmap using hierarchical cluster analysis (HCL) was generated using Delta2D software (Supplementary Fig. 12).

2.5. Mass spectrometry analysis of *B. napus* seed tissue proteins

The establishment of an endosperm proteome reference map and the analyses of changes in protein abundances during endosperm and embryo development were based on protein identifications by mass spectrometry (MS). Tryptic digestion and MS analysis was performed according to Klodmann et al. [44] using the EASY-nLC System (Proxeon) coupled to a MicroTOF-Q II mass spectrometer (Bruker Daltonics, Bremen, Germany). Identification of proteins was carried out using the MASCOT search algorithm (www.matrixscience.com) against the (i) SwissProt (www.uniprot.org), (ii) NCBInr (www.ncbi.nig.gov) and (iii) TAIR (www.arabidopsis.org, TAIR release 10) databases. In some cases, *B. napus* sequences are available. However, most proteins were identified via *A. thaliana*. *B. napus* and *A. thaliana* both belong to the Brassicaceae family and sequence identity within exons on average is above 90%. Therefore, the protein sequences of *Arabidopsis* are very useful for protein identifications in rapeseed. Identified proteins in *B. napus* are considered to represent homologs of the corresponding proteins in *Arabidopsis* (which does not exclude that they have differing functional roles). For details on peptides used for protein identification see Supplementary Table 2. Identified proteins were functionally classified according to the KEGG PATHWAY Database (http://www.genome.jp/kegg/pathway.html). The pathway categories were adjusted for seed metabolism by adding four functional groups: (i) storage, (ii) defense, (iii) desiccation and (iv) detoxification.

3. Results and discussion

3.1. Characterization of the endosperm proteome

Soluble proteins were extracted from the endosperm harvested 20 days after pollination (DAP), separated by high resolution 2D IEF/SDS PAGE using the HPE FlatTop Tower gel electrophoresis system and resolved proteins were finally identified by mass spectrometry. Delta2D analysis revealed separation of 964 protein spots on our gels. 385 spots were selected on a random basis and represent the most abundant protein spots of all areas of the gel and resulting in the identification of 930 distinct endosperm proteins (Supplementary Table 1). It should be noted that, based on our experimental approach, the proteome of the endosperm is only partially covered. Results of our study were used to generate a 'reference map' of the *B. napus* endosperm proteome (Fig. 1). Identified proteins were systematically analyzed by usage of various database resources.

3.1.1. Insights into overall proteome architecture

The *B. napus* endosperm proteome (Fig. 1) is composed of proteins covering a wide range of molecular masses (Fig. 2A) and pIs (Fig. 2B). Proteins within the 10 to 70 kDa range are predominant in the endosperm. Analysis of isoelectric points (pIs) of the identified proteins revealed two peaks, a major one at acidic to the neutral pH (pH 4.5 to 7.5) and a minor one at basic pH (pH 8.5–11). Apparent molecular masses of the 930 proteins very much correlate with calculated molecular masses (Fig. 2C). However, as expected, several proteins have a slightly reduced apparent molecular mass if compared to the calculated mass, which is due to the cleavage of targeting peptides of proteins transported to the endosplasmic reticulum (ER), chloroplasts or mitochondria. Only a few proteins on our 2D gels are significantly smaller than predicted, most likely reflecting proteolytic processing. A few proteins are larger than expected which could be the consequence of protein modification. Likewise, measured and calculated isoelectric points (pIs) of the 930 proteins correlate (Fig. 2D). However, several proteins clearly exhibit more acidic pIs than calculated, which most likely is due to protein modifications like phosphorylation.

3.1.2. Functional classification of proteins abundant in the endosperm

The endosperm of *B. napus* includes a broad spectrum of proteins (Supplementary Table 1). To gain insights into endosperm metabolism, identified proteins were grouped into functional categories according to KEGG PATHWAY Database (Fig. 3). In the following sections, details on the predominant functional categories (i) genetic information processing (31% of the total set of identified proteins), (ii) carbohydrate metabolism (17%), (iii) environmental information processing

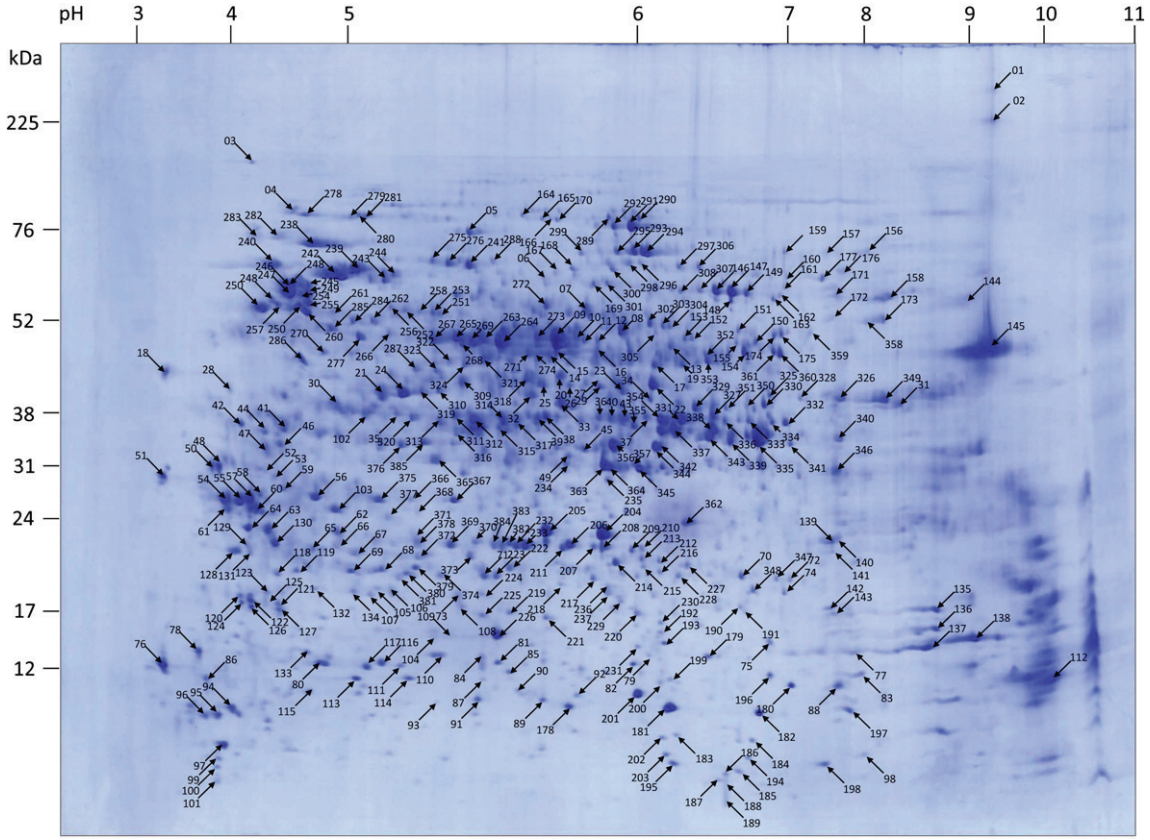


Fig. 1 – Two-dimensional endosperm proteome reference map. Total soluble endosperm proteins were separated on a 3–11 NL pH gradient IPG strip and subsequently on a 12.5% SDS-PAGE. The gel was stained with colloidal Coomassie brilliant blue G250. The arrows indicate spots that were picked for MS analysis (for results see Supplementary Table 1). Molecular masses of standard proteins are given to the left and pIs above the gel.

(12%), (iv) energy metabolism (9%), (v) cellular processes (7%) and (vi) amino acid metabolism (6%) are given:

(i) *Genetic information processing*

The endosperm proteome includes a high number of proteins taking part in transcription and transcriptional regulation, like DEAD box RNA helicases, G-rich RNA-binding proteins, DNA-binding protein BIN4, regulator of ribonuclease-like protein 1, FACT complex subunit and rho termination factor. BIN4 is a component of the plant DNA topoisomerase. It has been discussed that expression of bin4 drives cell expansion [45]. Several transcription factors (AP2-like ethylene-responsive transcription factor BBM, basic transcription factor 3, GRAS transcription factor, WRKY transcription factor, squamosa promoter binding protein-like, SSXT family protein) were found in the endosperm. AP2-like ethylene-responsive transcription factor BBM promotes cell proliferation, differentiation and morphogenesis and leads to the induction of embryo development [46]. Also WRKY and GRAS transcription factors are proposed to control developmental processes [47,48]. Transcription factors expressed in the endosperm are known to control final seed size and to coordinate embryo development [49]. The factors identified here might be involved in this regulatory cascade and/or related processes in *B. napus* seeds. Many proteins involved in translation, such as ribosomal proteins, initiation factors, elongation factors, and poly(A) binding protein, were identified in the course of our study. Proteins related to protein folding and sorting were found in several protein spots, for instance diverse chaperones, heat shock proteins and protein disulfide isomerases. The presence of numerous proteins involved in protein degradation, e.g. subunits of the proteasome and several monomeric proteases, indicates high protein turnover in the endosperm. The cell division control protein 48, an ATPase associated with versatile cellular functions [50], and the aspartic protease phytepsin are known to play important roles in targeted degradation of proteins.

(ii) *Carbohydrate metabolism*

Proteins involved in carbohydrate metabolism, which for instance play a role in glycolysis/gluconeogenesis, tricarboxylic acid (TCA) cycle and other sugar converting processes, are highly abundant in the endosperm. All glycolytic enzymes as well as the pyruvate dehydrogenase complex were identified. Glycolytic activity is one of the key features

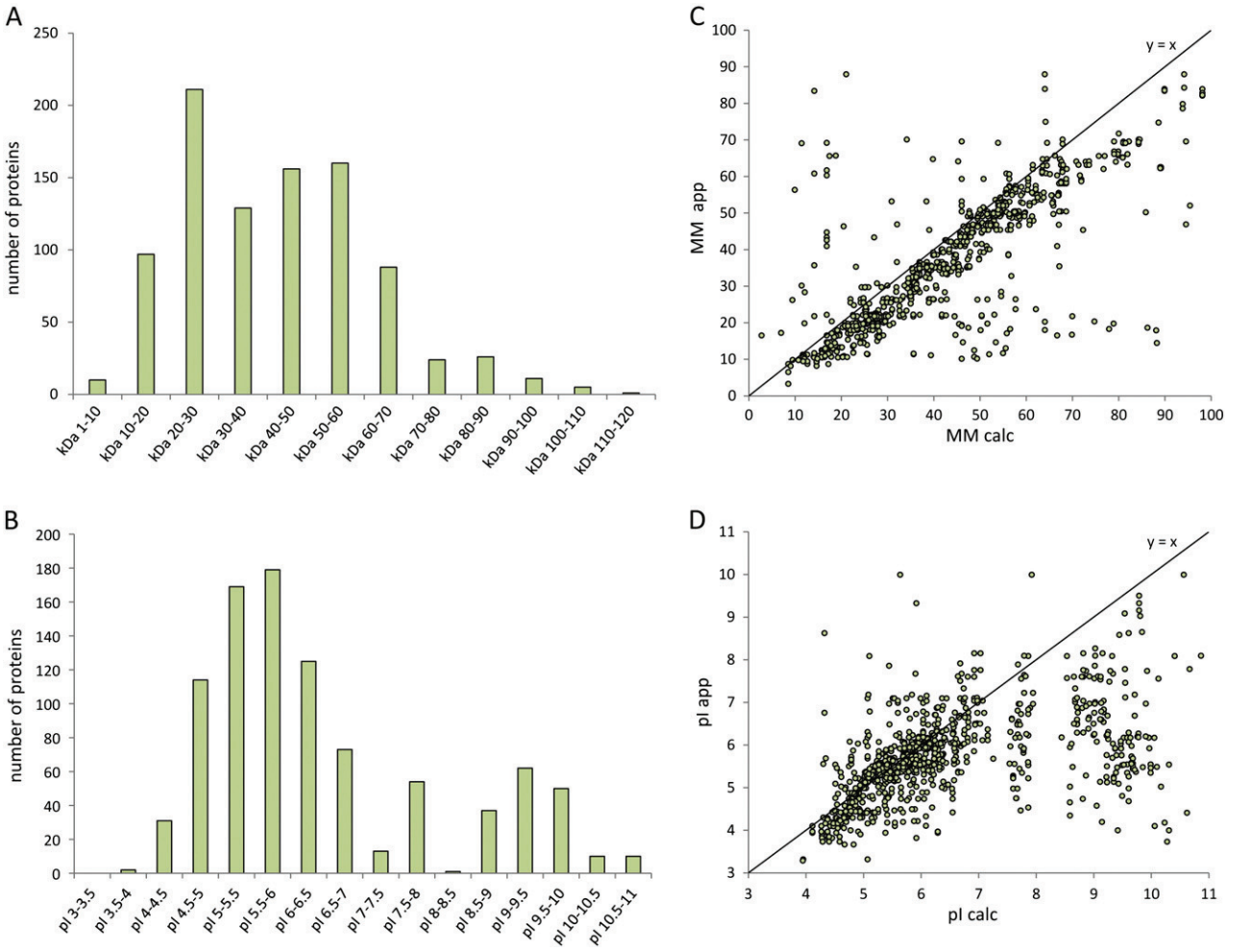


Fig. 2 – Proteome structure of *B. napus* endosperm. A: Distribution of calculated molecular masses of identified endosperm proteins. B: Distribution of calculated isoelectric points of identified endosperm proteins. C: Correlation of apparent and molecular masses. D: Correlation of apparent and calculated isoelectric points (pIs). Evaluations are based on a set of 930 identified proteins. Regression lines (y = x) are given in C and D.

of endosperm metabolism in *B. napus* seeds because it generates metabolic intermediates, reductants and energy [51]. *B. napus* embryos were reported to have a reduced cyclic flux around the TCA cycle [52]. In our study we only identified five out of the eight enzymes of the TCA cycle in the endosperm (aconitase, citrate synthase, isocitrate dehydrogenase, malate dehydrogenase and succinyl-CoA-ligase). Possibly the *B. napus* endosperm also has a reduced TCA cycle or the missing enzymes are beyond the detection limit of our study. Identification of enzymes contributing to the acetyl-CoA pool (2-oxoacid dehydrogenase acyltransferase, citrate lyase) might point to the high demand of acetyl-CoA for the synthesis of elongated fatty acids (and other compounds) which are highly abundant in rapeseed endosperm [53]. Furthermore the endosperm exhibits a high potential to produce reductant (NADH) via the pentose phosphate pathway as indicated by the detection of ribose-5-phosphate isomerase, phosphogluconate dehydrogenase, glucose-6-phosphate 1-dehydrogenase, 6-phosphogluconolactonase, transaldolase and transketolase. Two enzymes related to ascorbate biosynthesis were identified (ascorbate peroxidase, dehydroascorbate reductase). The endosperm also exhibits an active starch metabolism as indicated by the presence of enzymes required for starch synthesis (AGPase) and starch degradation (beta amylase). These data altogether indicate an active biosynthetic machinery, and correspond to the histological detection of starch granules, storage protein vacuoles and oil bodies in rapeseed endosperm [53].

(iii) *Environmental information processing*

The endosperm is the direct environment for the growing embryo and therefore is proposed to have a protective role. Accordingly, it includes a large number of proteins involved in signal transduction, plant hormone biosynthesis, stress response and defense related enzymes. The endosperm signaling pathways include annexin, calmodulin and 14-3-3

JOURNAL OF PROTEOMICS 108 (2014) 382–426

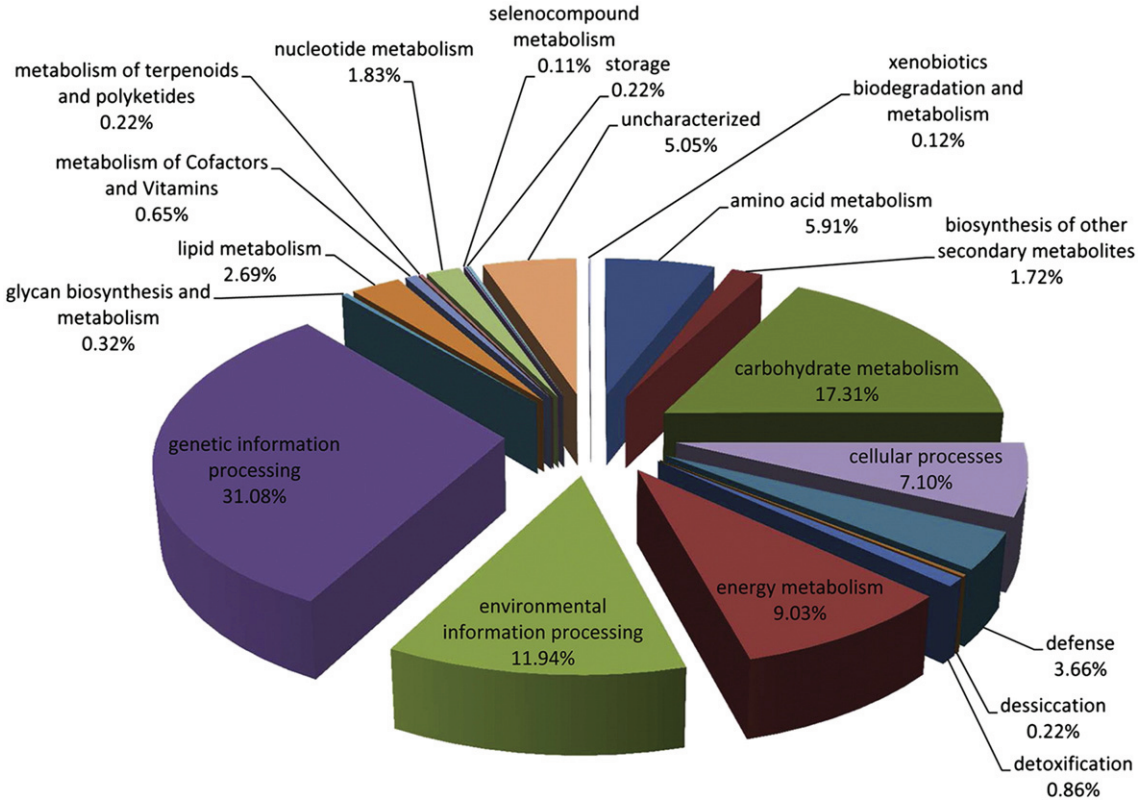


Fig. 3 – Functional classification of identified endosperm proteins of the proteome reference map. Proteins were assigned to the functional categories of the KEGG PATHWAY Database or one of the four following extra categories: (i) storage, (ii) defense, (iii) desiccation and (iv) detoxification.

Fig. 4 – *B. napus* seed development. Images of pods, whole seeds, seed cross-sections and embryos were taken from 7 to 56 DAP. Proteome analyses were carried out for 10 DAP (endosperm), 15 and 20 DAP (endosperm and embryo) and 25 DAP (embryo).

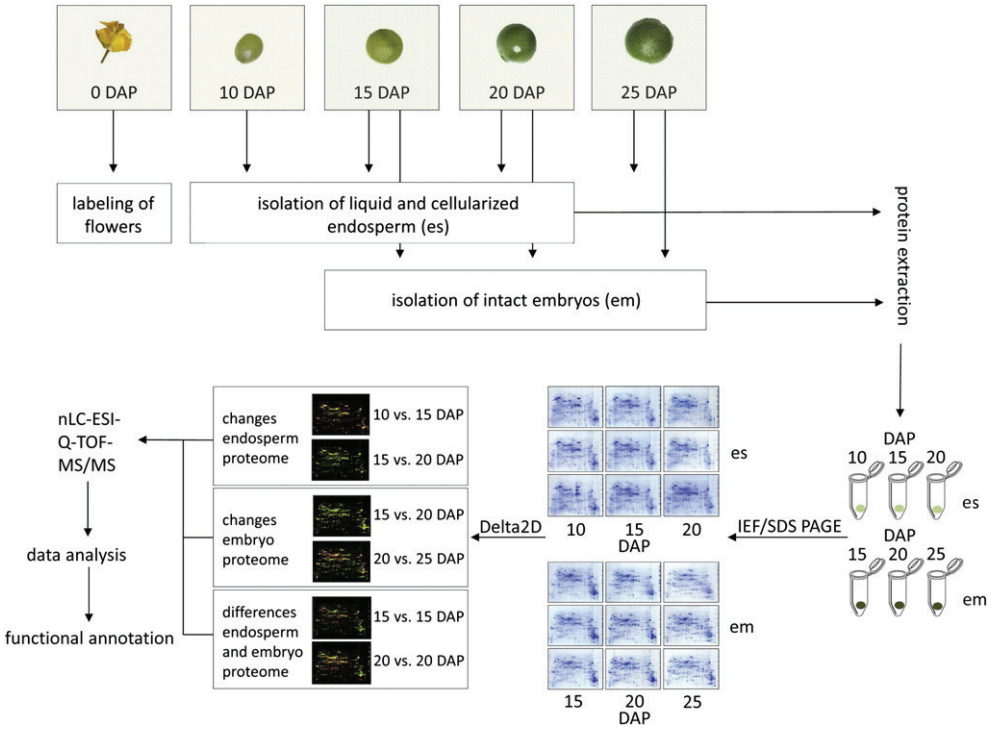


Fig. 5 – Experimental strategy for proteomic characterization of early seed development. Flowers were tagged directly after pollination (0 DAP) to define developmental stages. Endosperm (es) and embryo (em) fractions obtained at 10–25 DAP were used for phenolic protein extractions followed by 2D IEF/SDS gel electrophoreses (three replicates for each fraction) Delta2D analysis was used for protein quantification. Proteins differing in abundance between the compared fractions were identified by MS analyses and assigned to functional categories.

proteins. Calmodulin is a calcium binding protein and plays a role in varying signaling pathways [54]. Annexin, another calcium-binding protein, is able to build ion channels by forming a hexameric structure [55]. It possibly represents a regulator of ion exchange between subregions of the seed. Calcium-mediated signal transduction pathways might represent an important feature of the endosperm to coordinate developmental processes. Also, conserved regulatory 14-3-3 proteins were identified in this study which possibly are involved in seed development and germination [56]. The endosperm is also a source for plant hormones at maturation [57] but their role for endosperm development remained unknown so far. The following proteins involved in plant hormone biosynthesis were identified: auxin-responsive GH3 family protein, cytokinin riboside 5′-monophosphate phosphoribohydrolase LOG8, cytokinin-O-glucosyltransferase 1 and jasmonic acid-amido synthetase JAR1. Endosperm function in protecting the growing embryo might rely, besides other factors, on the occurrence of glutathione S-transferase, glutathione synthetase, glutaredoxin, catalase, peroxiredoxin, superoxide dismutase, translationally controlled tumor protein and adenine nucleotide alpha hydrolase, which all were identified in the course of our study. Also, plant defense related proteins, such as beta glucosidase, actin depolymerizing factors, protein LIR18A and thiocyanate methyltransferase were abundant in the endosperm. It is known that the termination of endosperm development is initiated by a gradual degradation of endosperm cells due to programmed cell death [11]. Accordingly, metacaspase 7, a protein involved in programmed cell death, was detected in the course of our study.

(iv) *Energy metabolism*

Embryos of green seeds perform photosynthesis to support biosynthetic activity and to diminish oxygen shortage [58]. Several proteins involved in photosynthetic processes were found in the endosperm, including various isoforms of the RuBisCO subunits, PSII and PSI subunits, light harvesting proteins, subunits of the cytochrome b_6f complex, chlorophyll a–b binding protein, ferredoxin NADP reductase, subunits of chloroplast ATP synthase and early light induced protein. Additionally, the oxygen evolving enhancer protein, which might promote oxygen production, was found at high abundance in the endosperm.

(v) *Cellular processes*

Several proteins involved in cell wall biogenesis were detected in the course of our study, including alpha-1,4-glucan-protein synthase, xyloglucan galactosyltransferase, UDP arabinopyranose mutase, UDP-glucose-6-dehydrogenase and fibrillin. Furthermore, an endoglucanase involved in cell wall degradation was identified. The endosperm includes cytoskeleton proteins such as tubulin, actin, and profilin to

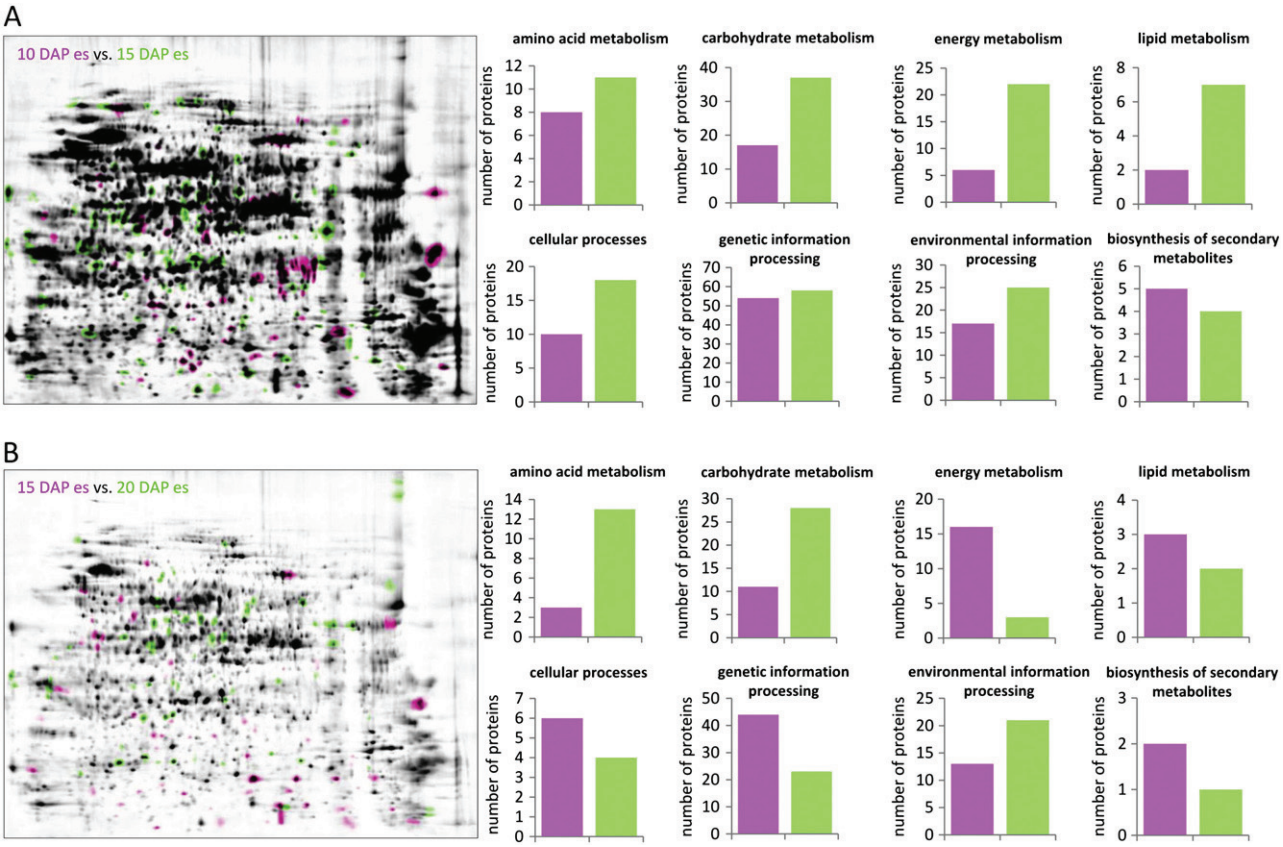


Fig. 6 – Changes in protein abundance during endosperm development. Three developmental stages of the *B. napus* endosperm were investigated (10, 15, 20 DAP) by 2D IEF/SDS PAGE and Delta2D analysis. To determine significant changes in spot volume, a Student's t-test (p-value ≤ 0.05) was applied on the basis of normalized relative spot volume. Changes in spot volume ≥1.5 were considered to represent true alterations. Protein spots of increased volume at 15 DAP (with respect to 10 DAP, A) and at 20 DAP (with respect to 15 DAP, B) are illustrated at the fused gel image of both groups in green, protein spots of reduced volume at 15 DAP (with respect to 10 DAP) and at 20 DAP (with respect to 15 DAP) in pink. The number of changed spots with respect to selected functional protein categories is given to the right of the gels (same color code).

ensure vesicle and membrane transport activities. The identification of proteins related to membrane trafficking and protein transport processes such as glycolipid transfer protein, plastid lipid associated protein, patellin, clathrin adaptor complexes, dynamin, importin, GOLD family protein, Mov34/MPN/PAD-1 family protein, NSF attachment protein, TOC75-3, RAB GTPase, Ran binding protein and RAB GDP dissociation inhibitor indicates occurrence of regulated transport processes and membrane biogenesis in the endosperm tissue.

(vi) *Amino acid metabolism*

The endosperm is known to receive amino acids from the mother plant, to transiently store and deliver them toward the growing embryo [32]. We here identified a large set of proteins involved in (i) *alanine*, *aspartate* and *glutamate metabolism* (alanine aminotransferase, aspartate aminotransferase, glutamate dehydrogenase), (ii) *arginine metabolism* (acetylornithine deacetylase), (iii) *cysteine* and *methionine metabolism* (adenosylhomocysteinase, methionine synthase, S-adenosylmethionine synthase, 1,2-dihydroxy-3-keto-5-methylthiopentene dioxygenase, 5-methyltetrahydropteroyltriglutamate-homocysteine methyltransferase), (iv) *glycine*, *serine* and *threonine metabolism* (serine hydroxymethyltransferase, PLP-dependent transferase), (v) *lysine biosynthesis* (lysine decarboxylase, LL-diaminopimelate aminotransferase) and (vi) *valine*, *leucine* and *isoleucine metabolism* (3-isopropylmalate dehydrogenase, methylmalonate-semialdehyde dehydrogenase). These findings support the view that the endosperm is not only a (passive) port of transit for amino acids but possesses a pronounced amino acid metabolism, supporting its own demands and able to respond to embryo's requirements.

3.2. Developmental changes in the endosperm proteome in comparison to embryo

Comparative proteomics was employed to obtain insights into the dynamic changes of proteins during *B. napus* seed development. For this approach, two experimental lines were followed: (i) comparison of different

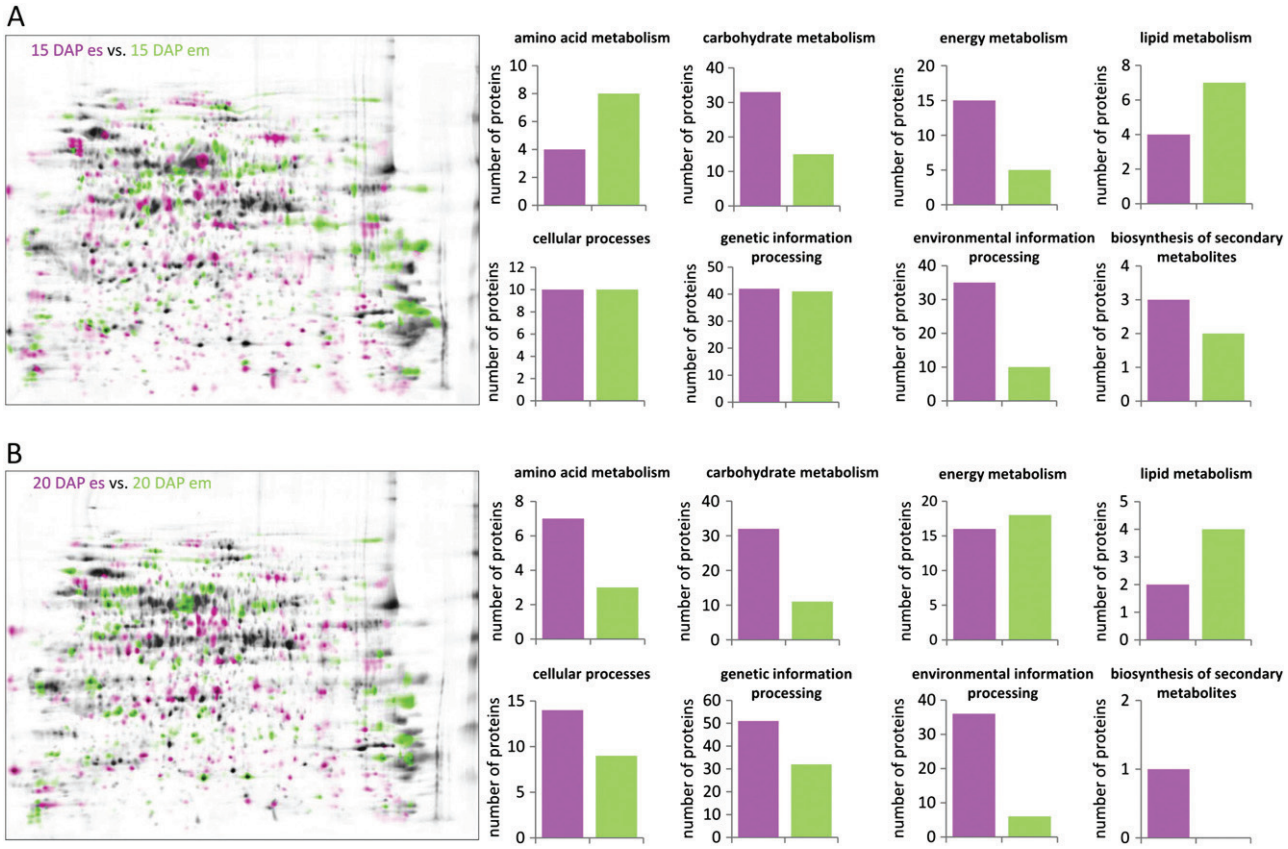


Fig. 7 – Differences of endosperm and embryo proteomes. Embryo and endosperm proteomes were investigated at 15 DAP (A) and at 20 DAP (B) by 2D IEF/SDS PAGE and Delta2D analysis. To determine significant differences in spot volume, a Student's t-test (p-value ≤0.05) was applied on the basis of normalized relative spot volume. Changes in spot volume ≥1.5 were considered to represent true alterations. Protein spots of increased volume in the endosperm (with respect to the embryo) are illustrated at the fused gel image of both groups in pink, protein spots of reduced volume in the endosperm (with respect to the embryo) in green. The number of changed spots with respect to selected functional protein categories is given to the right of the gels (same color code).

endosperm developmental stages and (ii) comparison of endosperm versus embryo. In a pre-experiment, seed development was systematically characterized by visual inspection to determine endosperm and embryo stages suitable for proteome analyses (Fig. 4). The embryo and also the endosperm of *B. napus* seeds are green during development. From 7 to 10 DAP the seed consists of endospermal liquid, very thin endospermal cell layers, and a seed coat while the embryo is hardly visible by eye. At the beginning of the cotyledon stage (14 DAP) the embryo becomes visible. During further development the largely fluid endosperm is gradually replaced by the growing embryo. At 25 DAP the endosperm is no more visible and its development is terminated while the

Table 1 – Number of spots with altered volumes and identified proteins within the comparisons (i)–(vi).

Comparison	Days after pollination	Tissue	Number of spots with increased (↑) or decreased (↓) volume	Number of identified proteins
(i)	10 → 15	Endosperm	↑105	205
			↓71	147
(ii)	15 → 20	Endosperm	↑65	106
			↓61	112
(iii)	15 → 20	Embryo	↑41	80
			↓45	48
(iv)	20 → 25	Embryo	↑78	108
			↓68	113
(v)	15 → 15	Endosperm/embryo	↑351	182
			↓276	111
(vi)	20 → 20	Endosperm/embryo	↑271	194
			↓212	93

embryo is still expanding. Accompanied by the color transition of the seed coat from green to brown also the embryo changes its color. Reduction of green pigments starts at the radicle and then spreads to the inner and outer cotyledons. At late maturation, embryo growth stops and desiccation of the seed is initiated. Based on this developmental sequence, six different fractions were taken for further proteomic analyses: endosperm tissue at 10, 15 and 20 DAP and embryo tissue at 15, 20 and 25 DAP.

Total soluble proteins were extracted from the six selected stages and analyzed by comparative proteomics. The workflow was based on three gel replicates per fraction and Delta2D analyses (summarized in Fig. 5). Proteins of differential abundance within the compared fractions were identified using the reference map or MS analyses. While overall proteome architecture with respect to the molecular mass and pI profiles of proteins of the six compared fractions are largely unchanged (Supplementary Fig. 10), the protein composition within the fractions specifically differs. We carried out six comparisons:

(i) Changes of the endosperm proteome at 15 DAP with respect to 10 DAP (Fig. 6, Supplementary Figs. 1, 2, Table 2)
(ii) Changes of the endosperm proteome at 20 DAP with respect to 15 DAP (Fig. 6, Supplementary Figs. 1, 3, Table 3)
(iii) Changes of the embryo proteome at 20 DAP with respect to 15 DAP (Supplementary Figs. 4, 5, 7, Supplementary Table 1).
(iv) Changes of the embryo proteome at 25 DAP with respect to 20 DAP (Supplementary Figs. 4, 6, 7, Supplementary Table 1)
(v) Differences between the endosperm and embryo proteomes at 15 DAP (Fig. 7A, Supplementary Fig. 8, Table 4)
(vi) Differences between the endosperm and embryo proteomes at 20 DAP (Fig. 7B, Supplementary Fig. 9, Table 5)

Table 1 summarizes the number of spots which have reduced or increased volumes and the corresponding proteins identified by MS in the six comparisons (changes in abundance of ≥1.5 fold).

In the following sections, based on the six proteomic comparisons, developmental changes of processes taking place in both endosperm and embryo are discussed. Since many spots include more than one protein and since some proteins are present in more than one spot, quantitative regulation patterns for individual enzymes cannot be concluded from our experimental approach. However, tendencies for dynamic changes of functional processes are clearly visible (Figs. 6 and 7):

(i) Changes of the endosperm proteome at 15 DAP with respect to 10 DAP: The number of proteins included in spots with altered volumes rather increases than decreases for most functional categories (Fig. 6A, Supplementary Figs. 1, 2, Table 2). Notably, proteins involved in energy metabolism, carbohydrate metabolism, amino acid and lipid metabolism increase. This underlines the rising metabolic activity (storage metabolism) of endosperm and the fundamental role of endosperm in nutrient supply for the embryo at 15 DAP. Also, proteins involved in environmental information processing increase. In contrast, no clear tendency in up- or down-regulation of proteins involved in genetic information processing is visible at 15 DAP with respect to 10 DAP. The number of proteins involved in the biosynthesis of secondary metabolites decreases from 10 to 15 DAP.
(ii) Changes of the endosperm proteome at 20 DAP with respect to 15 DAP: During development from 15 to 20 DAP, many processes in the endosperm significantly change. In contrast to the situation from 10 DAP to 15 DAP, further development of the endosperm until 20 DAP is accomplished with a relative reduction of proteins involved in energy metabolism (Fig. 6B, Supplementary Figs. 1, 3, Table 3). Interestingly, also lipid metabolism of the endosperm rather goes down from 15 to 20 DAP. In contrast, amino acid and carbohydrate metabolism as well as environmental information processing further increase from 15 to 20 DAP in the endosperm.
(iii)/(iv) Changes of the embryo proteome at 20 DAP with respect to 15 DAP and at 25 DAP with respect to 20 DAP: Changes in processes during seed development clearly differ between the endosperm and the embryo (Supplementary Figs. 4–7, Supplementary Table 1). In the embryo, energy metabolism significantly increases up from 15 to 20 DAP. Also, lipid metabolism and biosynthesis of secondary metabolites are enhanced and even further increase from 20 to 25 DAP. Carbohydrate and amino acid metabolism first goes up from 15 to 20 DAP in the embryo but decreases from 20 to 25 DAP. As a consequence, metabolism is shifted at 25 DAP from carbohydrate/amino acid metabolism to lipid/secondary metabolism in the embryo. The latter might reflect the shift from transient starch synthesis to both starch degradation and oil storage [53].
(v)/(vi) Differences between the endosperm and embryo proteomes at 15 DAP and at 20 DAP (Fig. 7A, B, Supplementary Figs. 8 and 9, Tables 4 and 5): If embryo and endosperm proteomes are directly compared at 15 and 20 DAP, respectively, fundamental differences are visible with respect to various cellular and metabolic processes: energy metabolism is clearly more pronounced in the endosperm than in the embryo at 15 DAP. This changes until 20 DAP when the embryo is fully competent in photosynthesis. At 15 and 20 DAP, environmental information processing rather takes place in the endosperm than in the embryo. This might underline the protective role of the endosperm for the embryo at both time points. Also carbohydrate metabolism is more pronounced in the endosperm at 15 as well as 20 DAP. In contrast, amino acid metabolism in the embryo at 15 DAP is prominent but decreases with respect to the endosperm until 15 DAP. Finally, lipid metabolism is especially pronounced in the embryo at 15 and 20 DAP.

As stated above, our experimental approach does not allow to quantify the changes in abundance for individual enzymes because most spots on our 2D gels include more than one protein. However, in some cases, quantitative changes are visible for individual proteins and in several other tendencies can be seen (see Supplementary discussion).

3.3. Concluding remarks

This study revealed for the first time the characteristics of the *B. napus* endosperm proteome, allowing to deduce dynamic changes in its metabolic and cellular processes. It became obvious that the metabolic architecture of endosperm comprises the entire set of central metabolic pathways. Its enzymatic machinery turns the endosperm into a self-sustaining, metabolically competent tissue, which of course receives major assimilates from the mother plant, but which is able to produce all or most of the needed metabolic intermediates by itself. It remains to be seen how this ability affects the assimilate supply/composition toward the developing embryo. The endosperm undergoes a unique developmental sequence from proliferation, growth and transient storage product accumulation toward programmed cell death. This developmental shift is clearly reflected in the proteome by the sequential appearance of proteins involved in energy metabolism, carbohydrate, amino acid and lipid metabolism. Notably, our analysis of the endosperm proteome also revealed a high number of proteins involved in redox balancing. It is tempting to speculate that a tight redox control in the endosperm is of highest significance, providing the embryo with a non-oxidizing environment. Such conditions were recently proposed to be key for germ cell fate and sexual reproduction [59]. Finally, proteome analysis indicated a high number of transcription factors and regulatory components in endosperm. It is very likely that their involvement in the various signaling cascades makes a fundamental contribution to coordinated seed development. A more targeted approach could help to identify their specific, individual contribution to growth control and eventually crop yield. As a next step, we plan absolute quantification of proteins involved in seed metabolism, e.g. by Selected Reaction Monitoring (SRM) mass spectrometry. This will allow to even deeper investigate the metabolic infrastructure of endosperm as well as its interplay with the developing embryo. In combination with metabolite profiling and isotope-based flux analysis, this will further improve our systems-level understanding of the *seed filling* process and provide rational strategies for bioengineering.

Transparency document

The transparency document associated with this article can be found, in the online version.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.jprot.2014.05.024.

REFERENCES

[1] Olsen O-A. Nuclear endosperm development in cereals and *Arabidopsis thaliana*. Plant Cell 2004;16:214–22.
[2] Berger F. Endosperm: the crossroad of seed metabolism. Curr Opin Plant Biol 2003;6(1):42–50.
[3] Brown RC, Lemmon BE, Nguyen H. Events during the first four rounds of mitosis establish three developmental domains in the syncytial endosperm of *Arabidopsis thaliana*. Protoplasma 2003;222(3–4):167–74.
[4] Brown RC, Lemmon BE, Nguyen H, Olsen OA. Development of endosperm in *Arabidopsis thaliana*. Plant Reprod 1999;12(1):32–42.
[5] Boisnard-Lorig C, Colon-Carmona A, Bauch M, Hodge S, Doerner P, Bancharel E, et al. Dynamic analyses of the expression of the HISTONE::YFP fusion protein in *Arabidopsis* show that syncytial endosperm is divided in mitotic domains. Plant Cell 2001;13(3):495–509.
[6] Sørensen MB. Cellularization in the endosperm of *Arabidopsis thaliana* is coupled to mitosis and shares multiple components with cytokinesis. Development 2002;129(24):5567–76.
[7] Hehenberger E, Kradolfer D, Köhler C. Endosperm cellularization defines an important developmental transition for embryo development. Plant Cell Physiol 2012;53(1):16–27.
[8] Bethke PC, Libourel IGL, Aoyama N, Chung Y-Y, Still DW, Jones R. The *Arabidopsis* aleurone layer responds to nitric oxide, gibberellin, and abscisic acid and is sufficient and necessary for seed dormancy. Plant Physiol 2007;143(3):1173–88.
[9] Berger F, Grini PE, Schnittger A. Endosperm: an integrator of seed growth and development. Curr Opin Plant Biol 2006;9:664–70.
[10] Berger F. Endosperm development. Curr Opin Plant Biol 1999;2:28–32.
[11] Young TE, Gallie DR. Programmed cell death during endosperm development. Plant Mol Biol 2000;44(3):283–301.
[12] Bauer MJ, Fischer RL. Genome demethylation and imprinting in the endosperm. Curr Opin Plant Biol 2011;14(2):162–7.
[13] Wolff P, Weinhofer I, Seguin J, Roszak P, Beisel C, Donoghue MT, et al. High-resolution analysis of parent-of-origin allelic expression in the *Arabidopsis* endosperm. PLoS Genet 2011;7(6):e1002126.
[14] Huh JH, Bauer MJ, Hsieh T-F, Fischer R. Endosperm gene imprinting and seed development. Curr Opin Genet Dev 2007;17(6):480–5.
[15] Berger F. Imprinting — a green variation. Science 2004;303:483–5.
[16] Gehring M, Choi Y, Fischer RL. Imprinting and seed development. Plant Cell 2004;16:203–13.
[17] Lu J, Zhang C, Baulcombe DC, Chen ZJ. Maternal siRNAs as regulators of parental genome imbalance and gene expression in endosperm of *Arabidopsis* seeds. Proc Natl Acad Sci 2012;109(14):5529–34.
[18] Cavel E, Pillot M, Pontier D, Lahmy S, Bies-Etheve N, Vega D, et al. A plant-specific transcription factor IIB-related protein, pBRP2, is involved in endosperm growth control. PLoS One 2011;6(2):e17216.
[19] Le BH, Cheng C, Bui AQ, Wagmaister JA, Henry KF, Pelletier J, et al. Global analysis of gene activity during *Arabidopsis* seed development and identification of seed-specific transcription factors. Proc Natl Acad Sci 2010;107(18):8063–70.

[20] Ohto MA, Fischer RL, Goldberg RB, Nakamura K, Harada JJ. Control of seed mass by APETALA2. Proc Natl Acad Sci 2005;102:3123–8.
[21] Garcia D. *Arabidopsis* haiku mutants reveal new controls of seed size by endosperm. Plant Physiol 2003;131(4):1661–70.
[22] Lu X, Chen D, Shu D, Zhang Z, Wang W, Klukas C, et al. The differential transcription network between embryo and endosperm in the early developing maize seed. Plant Physiol 2013;162(1):440–55.
[23] Hands P, Kourmpetli S, Sharples D, Harris RG, Drea S. Analysis of grain characters in temperate grasses reveals distinctive patterns of endosperm organization associated with grain shape. J Exp Bot 2012;63(17):6253–66.
[24] Wilson SM, Burton RA, Collins HM, Doblin MS, Pettolino FA, Shirley N, et al. Pattern of deposition of cell wall polysaccharides and transcript abundance of related cell wall synthesis genes during differentiation in barley endosperm. Plant Physiol 2012;159(2):655–70.
[25] Alonso AP, Val DL, Shachar-Hill Y. Central metabolic fluxes in the endosperm of developing maize seeds and their implications for metabolic engineering. Metab Eng 2011;13(1):96–107.
[26] Rolletschek H, Melkus G, Grafahrend-Belau E, Fuchs J, Heinzel N, Schreiber F, et al. Combined noninvasive imaging and modeling approaches reveal metabolic compartmentation in the barley endosperm. Plant Cell 2011;23(8):3041–54.
[27] Mechin V, Thevenot C, Le Guilloux M, Prioul J-L, Damerval C. Developmental analysis of maize endosperm proteome suggests a pivotal role for pyruvate orthophosphate dikinase. Plant Physiol 2007;143(3):1203–19.
[28] Lai J. Characterization of the maize endosperm transcriptome and its comparison to the rice genome. Genome Res 2004;14(10a):1932–7.
[29] Novack MK, Unguru A, Bjerkan KN, Grini PE, Schnittger A. Reproductive cross-talk: seed development in flowering plants. Biochem Soc Trans 2010;38:604–12.
[30] Chaudhury AM, Koltunow A, Payne T, Luo M, Tucker MR, Dennis ES, et al. Control of early seed development. Annu Rev Cell Dev Biol 2001;17(1):677–99.
[31] Tiwari S, Spielman S, Day RC, Scott RJ. Proliferative phase endosperm promoters from *Arabidopsis thaliana*. Plant Biotechnol J 2006;4:393–407.
[32] Melkus G, Rolletschek H, Radchuk R, Fuchs J, Rutten T, Wobus U, et al. The metabolic role of the legume endosperm: a noninvasive imaging study. Plant Physiol 2009;151:1139–54.
[33] Moreley-Smith ER, Pike MJ, Findlay K, Köckenberger W, Hill LM, Smith AM, et al. The transport of sugars to developing embryos is not via the bulk endosperm in oilseed rape seeds. Plant Physiol 2008;147:2121–30.
[34] Rolland F, Baena-Gonzalez E, Jen S. Sugar sensing and signaling in plants: conserved and novel mechanisms. Annu Rev Plant Biol 2006;57(1):675–709.
[35] Weber H, Borisjuk L, Wobus U. Molecular physiology of legume seed development. Annu Rev Plant Biol 2005;56:253–79.
[36] Lafon-Placette C, Köhler C. Embryo and endosperm, partners in seed development. Curr Opin Plant Biol 2014;17:64–9.
[37] Belmonte MF, Kirkbride RC, Stone SL, Pelletier JM, Bui AQ, Yeung EC, et al. Comprehensive developmental profiles of gene activity in regions and subregions of the *Arabidopsis* seed. Proc Natl Acad Sci 2013;110:E435–44.
[38] Huang Y, Chen L, Wang L, Vijayan K, Phan S, Liu Z, et al. Probing the endosperm gene expression landscape in *Brassica napus*. BMC Genomics 2009;10:256.
[39] Colditz F, Nyamsuren O, Niehaus K, Eubel H, Braun HP, Krajinski F. Proteomic approach: identification of *Medicago truncatula* proteins differentially expressed after infection with the pathogenic oomycete *Aphanomyces euteiches*. Plant Mol Biol 2004;55:109–20.
[40] Mihr C, Braun HP. Proteomics in plant biology. In: Michael P, editor. Handbook of proteomics methods. Totowa: Humana Press Inc; 2003. p. 409–41.
[41] Neuhoff V, Stamm R, Pardowitz I, Arold N, Ehrhardt W, Taube D. Essential problems in quantification of proteins following colloidal staining with Coomassie brilliant blue dyes in polyacrylamide gels, and their solution. Electrophoresis 1990;11(2):101–17.
[42] Neuhoff V, Stamm R, Eibl H. Clear background and highly sensitive protein staining with Coomassie blue dyes in polyacrylamide gels: a systematic analysis. Electrophoresis 1985;6(9):427–48.
[43] Berth M, Moser FM, Kolbe M, Bernhardt J. The state of the art in the analysis of two-dimensional gel electrophoresis images. Appl Microbiol Biotechnol 2007;76(6):223–1243.
[44] Klodmann J, Sunderhaus S, Nimtz M, Jansch L, Braun HP. Internal architecture of mitochondrial complex I from *Arabidopsis thaliana*. Plant Cell 2010;22(3):797–810.
[45] Breuer C, Stacey NJ, West CE, Zhao Y, Chory J, Tsukaya H, et al. BIN4, a novel component of the plant DNA topoisomerase VI complex, is required for endoreduplication in *Arabidopsis*. Plant Cell 2007;19(11):3655–68.
[46] Boutilier K, Offringa R, Sharma VK, Kieft H, Ouellet T, Zhang L, et al. Ectopic expression of BABY BOOM triggers a conversion from vegetative to embryonic growth. Plant Cell 2002;14(8):1737–49.
[47] Rushton PJ, Somssich IE, Ringler P, Shen QJ. WRKY transcription factors. Trends Plant Sci 2010;15(5):247–58.
[48] Hirsch S, Oldroyd GED. GRAS-domain transcription factors that regulate plant development. Plant Signal Behav 2009;4(8):698–700.
[49] Ingouff M, Jullien PE, Berger F. The female gametophyte and the endosperm control cell proliferation and differentiation of the seed coat in *Arabidopsis*. Plant Cell 2006;18:3491–501.
[50] Stolz A, Hilt W, Buchberger A, Wolf DH. Cdc48: a power machine in protein degradation. Trends Biochem Sci 2011;36(10):515–23.
[51] Hill LM, Morley-Smith ER, Rawsthorne S. Metabolism of sugars in the endosperm of developing seeds in oilseed rape. Plant Physiol 2003;131(1):228–36.
[52] Schwender J, Shachar-Hill Y, Ohlrogge JB. Mitochondrial metabolism in developing embryos of *Brassica napus*. J Biol Chem 2006;281(45):34040–7.
[53] Borisjuk L, Neuberger T, Schwender J, Heinzel N, Sunderhaus S, Fuchs J, et al. Seed architecture shapes embryo metabolism in oilseed rape. Plant Cell 2013;25:1625–40.
[54] Perochon A, Aldon D, Galaud J-P, Ranty B. Calmodulin and calmodulin-like proteins in plant calcium signaling. Biochimie 2011;93(12):2048–53.
[55] Laohavisit A, Davies JM. Annexins. New Phytol 2011;189(1):40–53.
[56] Fulgosi H, Soll J, de Faria Maraschin S, Korthout HAAJ, Wang M, Testerink C. 14-3-3 proteins and plant development. Plant Mol Biol 2002;50(6):1019–29.
[57] Lopes MA, Larkins BA. Endosperm origin, development, and function. Plant Cell 1993;5:1383–99.
[58] Borisjuk L, Rolletschek H. The oxygen status of the developing seed. New Phytol 2009;182(1):17–30.
[59] Kelliher T, Walbot V. Hypoxia triggers meiotic fate acquisition in maize. Science 2012;337(6092):345–8.

Table 2 – Identification of proteins from spots with changed abundance in the endosperm between 10 DAP and 15 DAP. To determine significant changes in spot volume, a Student's t-test (p-value ≤ 0.05) was applied on the basis of normalized relative spot volume. Changes in spot volume ≥ 1.5 were considered to represent true alterations. Identification of proteins was carried out using the MASCOT search algorithm (www.matrixscience.com) against the (i) SwissProt (www.uniprot.org), (ii) NCBInr (www.ncbi.nig.gov) and (iii) TAIR (www.arabidopsis.org, TAIR release 10) databases. Identified proteins were functionally classified according to the KEGG PATHWAY Database (http://www.genome.jp/kegg/pathway.html). The pathway categories were adjusted for seed metabolism by adding four functional groups: (i) storage, (ii) defense, (iii) desiccation and (iv) detoxification.

ID[a]	Stage[b]	Reg[c]	Accession[d]	Name	Organism	MM calc[e]	pI calc[f]	Score[g]	Pep[h]	SC%[i]
Amino acid metabolism										
1170	10 DAP es	1.63	P46248	Aspartate aminotransferase, chloroplastic	*A. thaliana*	49.80	8.91	1907	22	43.27
1179	10 DAP es	2.35	P46248	Aspartate aminotransferase, chloroplastic	*A. thaliana*	49.80	8.91	1233	17	41.28
1172	10 DAP es	2.53	Q43314	Glutamate dehydrogenase 1	*A. thaliana*	44.50	6.42	1027	16	33.58
1172	10 DAP es	2.53	P46248	Aspartate aminotransferase, chloroplastic	*A. thaliana*	49.80	8.91	87	2	4.86
892	10 DAP es	1.68	P93732	Proline iminopeptidase	*A. thaliana*	43.03	5.87	246	3	7.11
900	10 DAP es	1.86	AT1G79230.1	Mercaptopyruvate sulfurtransferase 1	*A. thaliana*	41.87	5.94	79	1	2.37
214	10 DAP es	2.29	AT1G11930.1	Predicted pyridoxal phosphate-dependent enzyme, YBL036C type	*A. thaliana*	27.99	4.98	86	2	6.61
893	10 DAP es	2.93	AAK57436	Nitrilase-like protein	*B. napus*	38.39	5.51	65	1	4.00
Biosynthesis of secondary metabolites										
893	10 DAP es	2.93	XP_002888957	Strictosidine synthase family protein	*A. lyrata*	34.25	9.29	98	2	10.12
616	10 DAP es	2.00	Q9FT25	Pyridoxal biosynthesis protein PDX1	*P. vulgaris*	33.38	5.72	88	2	7.05
616	10 DAP es	2.00	AAZ67141	Pyridoxine biosynthesis protein	*L. japonicus*	33.04	6.12	60	1	7.42
200	10 DAP es	3.11	AAV80204	Caffeoyl-CoA 3-O-methyltransferase	*B. napus*	17.83	4.80	256	3	16.25
892	10 DAP es	1.68	ACP20256	Cinnamyl-alcohol dehydrogenase	*B. rapa*	31.93	4.83	63	2	8.33
Carbohydrate metabolism										
898	10 DAP es	1.69	Q84TF0	Aldo-keto reductase family 4 member C10	*A. thaliana*	34.89	6.21	84	1	2.55
1189	10 DAP es	1.76	Q42605	UDP-glucose 4-epimerase 1	*A. thaliana*	39.13	6.14	1177	13	37.04
740	10 DAP es	4.47	AT1G76550.1	Phosphofructokinase family protein	*A. thaliana*	67.52	6.92	64	4	7.29
103	10 DAP es	2.26	O04499	2,3-Bisphosphoglycerate-independent phosphoglycerate mutase 1	*A. thaliana*	60.54	5.20	1520	18	31.42
1179	10 DAP es	2.35	AT3G52930.1	Aldolase superfamily protein	*A. thaliana*	38.52	6.04	2345	22	49.16
1179	10 DAP es	2.35	P04796	Glyceraldehyde-3-phosphate dehydrogenase, cytosolic	*S. alba*	36.90	8.71	213	5	19.82
1186	10 DAP es	1.51	AT3G52930.1	Aldolase superfamily protein	*A. thaliana*	38.52	6.04	1282	17	43.02
1189	10 DAP es	1.76	P04796	Glyceraldehyde-3-phosphate dehydrogenase, cytosolic	*S. alba*	36.90	8.71	358	8	30.77
1189	10 DAP es	1.76	O65735	Fructose-bisphosphate aldolase, cytoplasmic isozyme	*C. arietinum*	38.43	6.22	100	3	10.31
214	10 DAP es	2.29	P48491	Triosephosphate isomerase, cytosolic	*A. thaliana*	27.15	5.27	74	3	16.14
900	10 DAP es	1.86	XP_002893304	Oxidoreductase (SP: alcohol dehydrogenase activity)	*A. lyrata*	40.81	8.64	100	3	10.88
1172	10 DAP es	2.53	Q96533	Alcohol dehydrogenase class-3	*A. thaliana*	40.67	6.59	482	7	16.36
1172	10 DAP es	2.53	P04796	Glyceraldehyde-3-phosphate dehydrogenase, cytosolic	*S. alba*	36.90	8.71	59	4	11.24
1170	10 DAP es	1.63	Q07511	Formate dehydrogenase, mitochondrial	*S. tuberosum*	42.01	6.73	120	4	12.86
1321	10 DAP es	2.51	AAK50346	Putative 6-phosphogluconolactonase	*B. carinata*	29.19	6.75	51	1	3.49
135	10 DAP es	3.70	Q9M462	Glucose-1-phosphate adenylyltransferase small subunit, chloroplastic	*B. napus*	57.01	5.82	394	11	20.58
893	10 DAP es	2.93	P86074	Malate dehydrogenase, mitochondrial (Fragments)	*C. annuum*	3.73	3.73	165	1	34.29

Cellular processes										
1239	10 DAP es	1.75	XP_002501166	Multidrug/oligosaccharidyl-lipid/ polysaccharide flippase	*Micromonas* sp.	62.39	10.40	50	1	1.18
740	10 DAP es	4.47	P42697	Dynamin-related protein 1A	*A. thaliana*	68.13	9.15	649	15	24.43
296	10 DAP es	4.32	XP_002862306	Hydroxyproline-rich glycoprotein family protein	*A. lyrata*	36.76	6.32	127	4	8.56
341	10 DAP es	2.66	AT5G55820.1	Inner centromere protein, ARK-binding region	*A. thaliana*	202.90	5.17	55	1	0.55
383	10 DAP es	1.75	AAB18643	Actin	*P. sativum*	31.06	5.27	77	1	6.76
1278	10 DAP es	1.66	AT1G04820.1	Tubulin alpha-4 chain	*A. thaliana*	49.51	4.79	81	4	10.67
296	10 DAP es	4.32	ABB97039	Unknown (EMBL-EBI: Ran Binding Protein 1)	*B. rapa*	25.39	4.75	55	2	9.33
366	10 DAP es	3.29	AT1G27310.1	Nuclear transport factor 2A	*A. thaliana*	13.52	6.05	74	1	8.20
391	10 DAP es	1.92	Q9C7F5	Nuclear transport factor 2	*A. thaliana*	13.99	5.89	94	1	7.94
898	10 DAP es	1.69	AT4G35220.1	Cyclase family protein	*A. thaliana*	29.97	5.59	291	4	11.40
Defense										
135	10 DAP es	3.70	O04309	Myrosinase-binding protein-like	*A. thaliana*	48.47	4.99	1092	9	15.96
135	10 DAP es	3.70	AT3G16470.1	Mannose-binding lectin superfamily protein	*A. thaliana*	48.47	4.99	1064	9	15.96
677	10 DAP es	2.29	AT1G52400.1	Beta glucosidase 18	*A. thaliana*	60.42	6.82	264	5	7.01
686	10 DAP es	1.59	Q9SE50	Beta-glucosidase 18	*A. thaliana*	60.42	6.82	329	7	7.01
1170	10 DAP es	1.63	AT1G52400.1	Beta glucosidase 18	*A. thaliana*	60.42	6.82	328	3	4.73
1245	10 DAP es	7.23	Q84WV2	Beta-glucosidase 20	*A. thaliana*	61.64	5.65	225	2	2.43
1246	10 DAP es	8.25	Q84WV2	Beta-glucosidase 20	*A. thaliana*	61.64	5.65	206	4	2.43
1257	10 DAP es	5.83	Q84WV2	Beta-glucosidase 20	*A. thaliana*	61.64	5.65	155	4	2.43
1238b	10 DAP es	1.73	Q9SE50	Beta-glucosidase 18	*A. thaliana*	60.42	6.82	160	2	3.41
147	10 DAP es	4.15	O04309	Myrosinase-binding protein-like At3g16470	*A. thaliana*	48.47	4.99	400	4	8.43
695	10 DAP es	1.64	Q9SE50	Beta-glucosidase 18	*A. thaliana*	60.42	6.82	101	2	3.41
1172	10 DAP es	2.53	Q9SE50	Beta-glucosidase 18	*A. thaliana*	60.42	6.82	180	4	4.73
1262	10 DAP es	6.83	Q84WV2	Beta-glucosidase 20	*A. thaliana*	61.64	5.65	100	1	2.43
1278	10 DAP es	1.66	Q84WV2	Beta-glucosidase 20	*A. thaliana*	61.64	5.65	110	1	2.43
1321	10 DAP es	2.51	BAJ33862	Unnamed protein product (homology to Beta-glucosidase 1)	*T. halophila*	60.69	7.25	67	3	5.66
1321	10 DAP es	2.51	AT1G75940.1	Glycosyl hydrolase superfamily protein	*A. thaliana*	61.64	5.65	55	1	2.43
Desiccation										
393	10 DAP es	2.83	AT5G07190.1	Seed gene 3	*A. thaliana*	23.07	6.36	58	2	3.76
Detoxification										
1172	10 DAP es	2.53	ACR40091	S-Nitrosoglutathione reductase	*B. juncea*	40.07	8.72	298	1	17.65
Energy metabolism										
1102	10 DAP es	1.85	P05346	Ribulose bisphosphate carboxylase small chain, chloroplastic	*B. napus*	20.21	9.22	200	5	24.31
341	10 DAP es	2.66	P08135	Ribulose bisphosphate carboxylase small chain, chloroplastic	*R. sativus*	20.31	9.21	156	2	8.29
946	10 DAP es	3.81	AT1G12250.1	Pentapeptide repeat-containing protein	*A. thaliana*	30.05	9.64	57	2	9.64
341	10 DAP es	2.66	Q945M1	NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 9	*A. thaliana*	13.61	9.35	140	1	9.40
686	10 DAP es	1.59	AT5G04590.1	Sulfite reductase	*A. thaliana*	71.90	9.17	139	5	7.63
900	10 DAP es	1.86	ACB59214	Cytoplasmic thiosulfate:cyanide sulfur transferase	*B. oleracea*	41.51	6.93	87	2	5.01

(continued on next page)

Table 2 (*continued*)

ID [a]	Stage [b]	Reg [c]	Accession [d]	Name	Organism	MM calc [e]	pI calc [f]	Score [g]	Pep [h]	SC% [i]
Environmental information processing										
892	10 DAP es	1.68	Q9SYT0	Annexin D1	*A. thaliana*	36.18	5.09	695	12	38.80
391	10 DAP es	1.92	AT4G30380.1	Barwin-related endoglucanase	*A. thaliana*	13.25	9.71	121	1	16.26
383	10 DAP es	1.75	CAA46591	BnD22 drought induced protein	*B. napus*	23.52	5.84	70	1	5.50
1278	10 DAP es	1.66	CAA46591	BnD22 drought induced protein	*B. napus*	23.52	5.84	140	1	5.50
214	10 DAP es	2.29	AT2G37970.1	SOUL heme-binding family protein	*A. thaliana*	24.91	9.25	162	2	8.00
270	10 DAP es	1.68	AT1G30580.1	GTP binding protein	*A. thaliana*	44.44	6.38	912	9	28.43
1070	10 DAP es	1.63	Q6EUP4	14-3-3-like protein GF14-E	*Oryza sativa* subsp. *Japonica*	29.67	4.56	1268	12	30.15
1071	10 DAP es	1.57	P48347	14-3-3-like protein GF14 epsilon	*A. thaliana*	28.90	4.57	1285	14	34.65
1261	10 DAP es	1.62	O04266	GTP-binding protein SAR1A	*B. campestris*	21.95	7.74	1063	13	66.84
1439	10 DAP es	1.56	Q9FZ48	Ubiquitin-conjugating enzyme E2 36	*A. thaliana*	17.21	7.64	950	11	65.36
262	10 DAP es	3.56	CAA07494	Heat stress-induced protein	*B. oleracea*	23.47	9.23	489	7	24.31
317	10 DAP es	1.77	ABV89642	Universal stress protein family protein	*B. rapa*	17.58	9.02	260	4	42.14
317	10 DAP es	1.77	AT3G03270.1	Adenine nucleotide alpha hydrolases-like superfamily protein	*A. thaliana*	22.59	5.44	143	2	15.42
1080	10 DAP es	2.91	AT5G20500.1	Glutaredoxin family protein	*A. thaliana*	14.82	5.62	163	2	11.85
214	10 DAP es	2.29	BAC54102	Water-soluble chlorophyll protein	*B. oleracea*	24.73	6.12	130	2	7.56
946	10 DAP es	3.81	CAA07494	Heat stress-induced protein	*B. oleracea*	23.47	9.23	904	8	24.31
946	10 DAP es	3.81	BAB72020	Water-soluble chlorophyll protein	*R. sativus*	23.83	9.53	341	1	14.41
Genetic information processing										
651	10 DAP es	2.90	AT5G26710.1	Glutamyl/glutaminyl-tRNA synthetase	*A. thaliana*	81.01	6.68	410	14	15.86
740	10 DAP es	4.47	AT5G26710.1	Glutamyl/glutaminyl-tRNA synthetase	*A. thaliana*	81.01	6.68	360	10	10.71
483	10 DAP es	1.65	XP_003569976	Vacuolar-processing enzyme beta-isozyme-like	*B. distachyon*	54.56	6.15	99	1	2.65
214	10 DAP es	2.29	AT2G19480.1	Nucleosome assembly protein 1;2	*A. thaliana*	43.52	4.17	237	3	6.86
1070	10 DAP es	1.63	Q43124	Proliferating cell nuclear antigen	*B. napus*	29.15	4.46	410	8	33.46
36	10 DAP es	2.05	Q9SCN8	Cell division control protein 48 homolog A	*A. thaliana*	89.34	4.99	2501	36	40.79
756	10 DAP es	1.64	Q9LIP2	Proteasome subunit beta type-5-B	*A. thaliana*	29.47	5.79	1305	10	26.74
947	10 DAP es	3.08	Q8LD27	Proteasome subunit beta type-6	*A. thaliana*	25.14	5.21	104	2	10.73
1045	10 DAP es	4.30	AT5G49910.1	Chloroplast heat shock protein 70-2	*A. thaliana*	76.95	5.03	1614	17	24.37
1065	10 DAP es	2.33	AT1G11910.1	Aspartic proteinase	*A. thaliana*	54.58	5.25	75	2	6.13
1070	10 DAP es	1.63	AAL60579	Senescence-associated cysteine protease	*B. oleracea*	50.58	5.61	142	3	8.91
1186	10 DAP es	1.51	O22263	Protein disulfide-isomerase like 2-1	*A. thaliana*	39.47	5.73	1051	12	26.04
1324	10 DAP es	1.62	Q9LIP2	Proteasome subunit beta type-5-B	*A. thaliana*	29.47	5.79	223	3	10.26
1238b	10 DAP es	1.73	AT5G26360.1	TCP-1/cpn60 chaperonin family protein	*A. thaliana*	60.30	5.45	360	5	10.99
147	10 DAP es	4.15	O23894	26S protease regulatory subunit 6A homolog	*B. campestris*	47.46	4.77	139	4	8.96
214	10 DAP es	2.29	O23708	Proteasome subunit alpha type-2-A	*A. thaliana*	25.69	5.40	163	4	18.72
330	10 DAP es	1.90	AT4G25370.1	Double Clp-N motif protein	*A. thaliana*	26.03	9.84	82	2	4.20
341	10 DAP es	2.66	AT5G48580.1	Peptidyl-Prolyl cis-trans isomerase FKBP15-2	*A. thaliana*	17.65	5.14	53	1	7.36
639	10 DAP es	3.73	AT1G32940.1	Subtilase family protein	*A. thaliana*	82.86	5.85	106	1	1.16
639	10 DAP es	3.73	AAD03430	similar to the subtilase family of serine proteases	*A. thaliana*	74.15	6.57	78	1	1.31
946	10 DAP es	3.81	XP_002301464	Peptidyl-prolyl cis-trans isomerase	*P. trichocarpa*	29.07	10.04	126	2	4.51
1097	10 DAP es	1.63	AT3G53990.1	Adenine nucleotide alpha hydrolases-like superfamily protein	*A. thaliana*	17.78	5.56	126	3	15.00

1239	10 DAP es	1.75	NP_199848	Adenylate kinase 2	*A. thaliana*	27.32	7.74	65	1	4.44
616	10 DAP es	2.00	AT3G01340.1	Transducin/WD40 repeat-like superfamily protein	*A. thaliana*	32.61	5.61	515	5	14.90
1070	10 DAP es	1.63	AT5G05010.1	Clathrin adaptor complexes medium subunit family protein	*A. thaliana*	57.68	5.54	133	2	5.12
330	10 DAP es	1.90	AT1G24450.1	Ribonuclease III family protein	*A. thaliana*	20.73	10.29	194	3	7.33
952	10 DAP es	1.63	O48646	Probable phospholipid hydroperoxide glutathione peroxidase 6, mitochondrial	*A. thaliana*	25.57	9.95	495	6	20.26
616	10 DAP es	2.00	AT5G11170.1	DEAD/DEAH box RNA helicase family protein	*A. thaliana*	48.31	5.33	60	1	2.34
42	10 DAP es	2.24	ABL97965	Putative nuclear transport factor 2	*B. rapa*	13.63	5.67	434	5	49.59
269	10 DAP es	3.85	Q40468	Eukaryotic initiation factor 4A-15	*N. tabacum*	46.69	5.26	717	11	17.43
612	10 DAP es	2.77	AT1G56070.1	Ribosomal protein S5/Elongation factor G/III/V family protein	*A. thaliana*	93.83	5.85	1642	17	14.00
740	10 DAP es	4.47	P42731	Polyadenylate-binding protein 2	*A. thaliana*	68.63	8.82	1011	17	19.87
964	10 DAP es	4.13	Q9XHS0	40S ribosomal protein S12	*H. vulgare*	15.28	5.25	378	3	14.69
979	10 DAP es	1.68	Q93VG5	40S ribosomal protein S8-1	*A. thaliana*	24.98	10.86	500	6	28.83
984	10 DAP es	1.88	P49204	40S ribosomal protein S4-2	*A. thaliana*	29.86	10.70	1085	19	59.16
1065	10 DAP es	2.33	O04202	Eukaryotic translation initiation factor 3 subunit F	*A. thaliana*	31.84	4.78	516	9	34.47
1248	10 DAP es	13.09	Q93VH9	40S ribosomal protein S4-1	*A. thaliana*	29.78	10.70	342	5	18.77
1324	10 DAP es	1.62	Q43467	Elongation factor Tu, chloroplastic	*G. max*	52.06	6.22	558	4	8.77
341	10 DAP es	2.66	Q9CAV0	40S ribosomal protein S3a-1	*A. thaliana*	29.83	10.38	101	1	3.82
383	10 DAP es	1.75	AT5G59950.2	RNA-binding (RRM/RBD/RNP motifs) family protein	*A. thaliana*	18.88	9.71	91	4	17.98
391	10 DAP es	1.92	AT3G53890.1	Ribosomal protein S21e	*A. thaliana*	9.07	9.23	50	1	18.29
392	10 DAP es	3.98	AT3G60245.1	Zinc-binding ribosomal protein family protein	*A. thaliana*	10.23	11.10	314	5	46.74
483	10 DAP es	1.65	AT1G75350.1	Ribosomal protein L31	*A. thaliana*	16.02	10.49	53	1	8.33
528	10 DAP es	1.8	Q9SF40	60S ribosomal protein L4-1	*A. thaliana*	44.67	10.86	1074	22	38.42
898	10 DAP es	1.69	Q42112	60S acidic ribosomal protein P0-2	*A. thaliana*	34.11	4.82	65	1	4.69
900	10 DAP es	1.86	AT1G56070.1	Ribosomal protein S5/Elongation factor G/III/V family protein	*A. thaliana*	93.83	5.85	1507	20	13.64
900	10 DAP es	1.86	BAJ33766	Unnamed protein product (Elongation factor EF-2)	*T. halophila*	93.88	5.85	940	19	14.23
946	10 DAP es	3.81	AT1G54270.1	eif4a-2	*A. thaliana*	46.73	5.35	52	1	4.37
952	10 DAP es	1.63	Q9XI91	Eukaryotic translation initiation factor 5A-1	*A. thaliana*	17.35	5.34	339	3	27.22
1249	10 DAP es	4.02	P0DH99	Elongation factor 1-alpha 1	*A. thaliana*	49.47	9.79	337	6	12.03
1278	10 DAP es	1.66	AT1G03810.1	Nucleic acid-binding, OB-fold-like protein	*A. thaliana*	15.66	10.23	55	2	5.59
1321	10 DAP es	2.51	AT1G72370.1	40s ribosomal protein SA	*A. thaliana*	32.27	4.88	60	1	6.04
1321	10 DAP es	2.51	CAA48794	Laminin receptor homolog	*A. thaliana*	32.28	4.97	60	1	6.04
1435	10 DAP es	7.60	ABL97959	Ribosomal protein L7Ae-like	*B. rapa*	13.88	7.52	57	1	6.25
Lipid metabolism										
892	10 DAP es	1.68	AT1G09480.1	Enoyl-[acyl-carrier protein] reductase I	*A. thaliana*	41.25	5.02	62	2	6.23
898	10 DAP es	1.69	P80030	Enoyl-[acyl-carrier-protein] reductase [NADH], chloroplastic	*B. napus*	40.45	9.39	484	9	17.14
Uncharacterized										
269	10 DAP es	3.85	ACJ84369	Unknown	*M. truncatula*	28.22	6.15	360	2	28.10
612	10 DAP es	2.77	BAJ33766	Unnamed protein product	*T. halophila*	93.88	5.85	1067	8	14.71
677	10 DAP es	2.29	BAJ33862	Unnamed protein product	*T. halophila*	60.69	7.25	251	8	19.06
1189	10 DAP es	1.76	ABQ50551	Hypothetical protein	*B. rapa*	34.03	5.84	1521	21	69.58
214	10 DAP es	2.29	XP_002322949	Predicted protein	*P. trichocarpa*	22.34	7.21	136	3	9.85
898	10 DAP es	1.69	XP_002308181	Predicted protein	*P. trichocarpa*	35.30	6.23	67	3	8.52

(continued on next page)

Table 2 (*continued*)

ID [a]	Stage [b]	Reg [c]	Accession [d]	Name	Organism	MM calc [e]	pI calc [f]	Score [g]	Pep [h]	SC% [i]
946	10 DAP es	3.81	NP_001047560	Os02g0643500	*O. sativa Japonica group*	28.94	10.10	128	2	5.42
1097	10 DAP es	1.63	ABL97944	Hypothetical protein	*B. rapa*	17.65	5.92	760	12	52.83
1239	10 DAP es	1.75	AT4G12900.1	Gamma interferon responsive lysosomal thiol (GILT) reductase family protein	*A. thaliana*	26.01	6.15	93	1	4.33
1239	10 DAP es	1.75	BAK00512	Predicted protein	*H. vulgare*	56.47	9.43	54	1	1.98
Amino acid metabolism										
1191	15 DAP es	1.69	P46643	Aspartate aminotransferase, mitochondrial	*A. thaliana*	47.73	9.13	957	14	31.63
1176	15 DAP es	1.90	P32289	Glutamine synthetase nodule isozyme	*V. aconitifolia*	39.08	5.7	78	2	7.87
1176	15 DAP es	1.90	S18603	Glutamate-ammonia ligase, cytosolic	*A. thaliana*	40.70	5.3	60	2	8.02
882	15 DAP es	2.38	ABD65618	Acetylornithine deacetylase, putative	*B. oleracea*	44.79	5.1	79	4	17.20
1336	15 DAP es	3.03	AT1G75330.1	Ornithine carbamoyltransferase, chloroplast	*A. thaliana*	40.98	7.9	66	2	4.27
913	15 DAP es	2.27	Q9LDQ7	S-Adenosylmethionine synthase	*C. sinensis*	42.77	5.2	112	2	9.41
1157	15 DAP es	1.67	AT4G01850.1	S-Adenosylmethionine synthetase 2	*A. thaliana*	43.23	5.6	246	5	27.48
1291	15 DAP es	1.61	Q9LEV3	CBS domain-containing protein CBSX3, mitochondrial	*A. thaliana*	22.71	9.8	1073	15	36.89
122	15 DAP es	2.28	P30184	Leucine aminopeptidase 1	*A. thaliana*	54.48	5.6	271	4	8.85
912	15 DAP es	2.29	AAR13689	Peptide methionine sulfoxide reductase	*B. oleracea*	22.72	5.3	248	2	42.16
882	15 DAP es	2.38	Q93ZN9	LL-diaminopimelate aminotransferase, chloroplastic	*A. thaliana*	50.36	7.7	891	10	22.56
Biosynthesis of secondary metabolites										
796	15 DAP es	1.93	CAA57285	ACC oxidase	*B. oleracea*	36.63	4.9	66	2	11.84
1336	15 DAP es	3.03	AAY53488	Epithiospecifier protein	*B. oleracea*	37.72	5.7	51	2	7.87
856	15 DAP es	2.24	AT1G53280.1	Glutamine amidotransferase-like superfamily protein	*A. thaliana*	46.96	9.0	287	2	7.99
925	15 DAP es	2.05	XP_002864577	Cinnamoyl-CoA reductase family	*A. lyrata*	35.39	5.8	57	2	12.69
Carbohydrate metabolism										
1157	15 DAP es	1.67	AT1G08200.1	UDP-D-apiose/UDP-D-xylose synthase 2	*A. thaliana*	43.76	5.5	81	2	5.40
767	15 DAP es	1.63	XP_002867164	L-Galactose dehydrogenase	*A. lyrata*	34.62	5.7	56	3	9.40
913	15 DAP es	2.27	Q9LFA3	Probable monodehydroascorbate reductase, cytoplasmic isoform 3	*A. thaliana*	46.46	6.5	62	3	9.68
1157	15 DAP es	1.67	AAK72107	Monodehydroascorbate reductase	*B. rapa*	46.43	5.7	108	7	29.95
1254	15 DAP es	2.85	AAY47048	Dehydroascorbate reductase	*S. lycopersicum*	23.52	6.4	309	5	23.81
1297	15 DAP es	3.12	AAY47048	Dehydroascorbate reductase	*S. lycopersicum*	23.52	6.4	243	4	21.90
925	15 DAP es	2.05	Q42592	L-Ascorbate peroxidase S, chloroplastic/mitochondrial	*A. thaliana*	40.38	9.0	358	5	14.52
122	15 DAP es	2.28	AT5G57655.2	Xylose isomerase family protein	*A. thaliana*	53.69	5.5	284	6	13.00
1336	15 DAP es	3.03	AT5G51830.1	pfkB-like carbohydrate kinase family protein	*A. thaliana*	37.01	4.8	294	4	10.79
122	15 DAP es	2.28	Q9SEE5	Galactokinase	*A. thaliana*	54.31	5.6	150	2	4.64
726	15 DAP es	1.99	AT1G76550.1	Phosphofructokinase family protein	*A. thaliana*	67.52	6.9	64	4	7.29
729	15 DAP es	7.43	AT1G76550.1	Phosphofructokinase family protein	*A. thaliana*	67.52	6.9	246	2	21.56
122	15 DAP es	2.28	P25696	Bifunctional enolase 2/transcriptional activator	*A. thaliana*	47.69	5.5	714	11	20.72
620	15 DAP es	3.18	AT3G52930.1	Aldolase superfamily protein	*A. thaliana*	38.52	6.0	1790	21	49.72
620	15 DAP es	3.18	AEC13713	Alcohol dehydrogenase 1	*B. rapa*	41.14	5.9	350	8	44.33

620	15 DAP es	3.18	AT1G13440.1	Glyceraldehyde-3-phosphate dehydrogenase C2	*A. thaliana*	36.89	6.8	87	3	11.24
780	15 DAP es	2.64	Q38799	Pyruvate dehydrogenase E1 component subunit beta, mitochondrial	*A. thaliana*	39.15	5.6	1319	11	31.13
913	15 DAP es	2.27	P25696	Bifunctional enolase 2/transcriptional activator	*A. thaliana*	47.69	5.5	1458	11	30.63
1191	15 DAP es	1.69	AT3G52930.1	Aldolase superfamily protein	*A. thaliana*	38.52	6.0	56	2	7.26
1416	15 DAP es	3.22	Q9M5K3	Dihydrolipoyl dehydrogenase 1, mitochondria	*A. thaliana*	53.95	7.2	111	5	13.81
779	15 DAP es	2.45	Q38799	Pyruvate dehydrogenase E1 component subunit beta, mitochondrial	*A. thaliana*	39.15	5.6	76	1	3.03
1176	15 DAP es	1.90	AT3G52930.1	Aldolase superfamily protein	*A. thaliana*	38.52	6.0	228	8	22.07
1176	15 DAP es	1.90	AT1G01090.1	Pyruvate dehydrogenase E1 alpha	*A. thaliana*	47.14	7.9	222	3	4.91
908	15 DAP es	1.61	AT3G60750.1	Transketolase	*A. thaliana*	79.92	5.9	628	7	11.20
1416	15 DAP es	3.22	Q43839	Glucose-6-phosphate 1-dehydrogenase, chloroplastic	*S. tuberosum*	65.65	7.1	123	2	3.12
910	15 DAP es	1.71	Q39366	Putative lactoylglutathione lyase	*B. oleracea*	31.63	5.0	906	11	36.88
732	15 DAP es	2.63	AAP80614	Beta amylase	*T. aestivum*	30.88	9.3	214	4	18.59
1089	15 DAP es	3.63	AAP37972	Seed specific protein Bn15D33A (SP: galactosyltransferase activity)	*B. napus*	12.66	9.2	338	5	31.90
925	15 DAP es	2.05	BAB01287	Sucrose cleavage protein-like	*A. thaliana*	34.23	5.5	70	1	2.87
152	15 DAP es	2.82	Q9SJH7	Citrate synthase 3, peroxisomal	*A. thaliana*	56.14	7.8	447	3	7.07
180	15 DAP es	1.98	Q9SN86	Malate dehydrogenase, chloroplastic	*A. thaliana*	42.38	9.4	1340	10	27.79
182	15 DAP es	1.55	Q43744	Malate dehydrogenase, mitochondrial	*B. napus*	35.69	9.5	91	2	10.26
732	15 DAP es	6.45	AT5G49460.1	ATP citrate lyase subunit B 2	*A. thaliana*	65.79	8.5	2074	32	44.24
1165	15 DAP es	2.80	AT1G65930.1	Cytosolic NADP+-dependent isocitrate dehydrogenase	*A. thaliana*	45.72	6.1	432	8	17.07
1352	15 DAP es	9.11	Q9SJH7	Citrate synthase 3, peroxisomal	*A. thaliana*	56.14	7.8	304	3	7.07
1467	15 DAP es	2.34	Q42560	Aconitate hydratase 1	*A. thaliana*	98.09	6.0	978	26	29.51
1336	15 DAP es	3.03	P93819	Malate dehydrogenase, cytoplasmic 1	*A. thaliana*	35.55	6.1	808	12	35.24
Cellular processes										
726	15 DAP es	1.99	P42697	Dynamin-related protein 1A	*A. thaliana*	68.13	9.2	649	15	24.43
1151	15 DAP es	3.00	Q9LIA8	UDP-glucose 6-dehydrogenase 1	*A. thaliana*	53.17	5.6	1482	13	27.10
314	15 DAP es	15.27	AAP37968	Seed specific protein Bn15D12A (SP: pectin esterase activity)	*B. napus*	17.03	9.3	1409	13	77.71
1306	15 DAP es	47.89	AAP37968	Seed specific protein Bn15D12A (SP: pectin esterase activity)	*B. napus*	17.03	9.3	535	10	64.97
1176	15 DAP es	1.90	AT2G39050.1	Hydroxyproline-rich glycoprotein family protein	*A. thaliana*	35.63	6.2	212	2	9.78
148	15 DAP es	6.91	BAA99394	Vacuolar calcium binding protein	*R. sativus*	27.09	4.0	245	4	12.90
764	15 DAP es	2.60	AT3G08900.1	Reversibly glycosylated polypeptide 3 (SP: UDP-arabinopyranose mutase 3)	*A. thaliana*	41.25	5.3	140	4	7.46
1176	15 DAP es	1.90	AT5G15650.1	Reversibly glycosylated polypeptide 2 (SP: UDP-arabinopyranose mutase 2)	*A. thaliana*	40.86	5.7	448	14	32.78
928	15 DAP es	2.58	O81644	Villin-2	*A. thaliana*	107.78	5.1	263	4	5.12
1151	15 DAP es	3.00	P20363	Tubulin alpha-3/alpha-5 chain	*A. thaliana*	49.62	4.8	404	9	25.11
1176	15 DAP es	1.90	XP_002867495	Band 7 family protein	*A. lyrata*	45.21	6.4	71	2	5.11
339	15 DAP es	3.41	ABA95924	Expressed protein (SP function:hydrogen ion transmembrane transporter activity)	*O. sativa japonica group*	12.75	11.9	75	1	5.98
488	15 DAP es	3.32	AAS68185	Lipid transfer-like protein	*B. napus*	11.18	9.2	203	3	33.33
813	15 DAP es	2.24	Q56WK6	Patellin-1	*A. thaliana*	64.01	4.7	830	13	18.32
181	15 DAP es	2.56	AT4G35220.1	Cyclase family protein	*A. thaliana*	29.97	5.6	86	3	11.76

(continued on next page)

Table 2 (*continued*)

ID [a]	Stage [b]	Reg [c]	Accession [d]	Name	Organism	MM calc [e]	pI calc [f]	Score [g]	Pep [h]	SC% [i]
Cellular processes										
897	15 DAP es	1.69	AT4G35220.1	Cyclase family protein	*A. thaliana*	29.97	5.6	445	7	19.12
908	15 DAP es	1.61	P23344	Actin-11	*D. carota*	41.77	6.6	106	2	11.02
339	15 DAP es	3.41	Q9C9L5	Wall-associated receptor kinase-like 9 (serine/threonine-protein kinase)	*A. thaliana*	88.54	9.0	58	1	1.14
Defense										
315	15 DAP es	2.14	Q9ZVF2	MLP-like protein 329	*A. thaliana*	17.59	5.2	383	4	26.49
1060	15 DAP es	2.96	Q944W6	Translationally-controlled tumor protein homolog	*B. oleracea*	19.03	4.5	784	9	42.26
1061	15 DAP es	1.65	Q944W6	Translationally-controlled tumor protein homolog	*B. oleracea*	19.03	4.5	556	8	39.29
1304	15 DAP es	4.42	AT2G21130.1	Cyclophilin-like peptidyl-prolyl cis-trans isomerase family protein	*A. thaliana*	18.45	9.4	196	4	19.54
Detoxification										
766	15 DAP es	1.52	AT4G33670.1	NAD(P)-linked oxidoreductase superfamily protein	*A. thaliana*	34.51	5.4	73	2	5.64
767	15 DAP es	1.63	AT4G33670.1	NAD(P)-linked oxidoreductase superfamily protein	*A. thaliana*	34.51	5.4	89	2	5.64
Energy metabolism										
1151	15 DAP es	3.00	ATCG00490.1	Ribulose-bisphosphate carboxylases	*A. thaliana*	52.92	5.9	205	5	11.27
1173	15 DAP es	1.73	P25857	Glyceraldehyde-3-phosphate dehydrogenase B, chloroplast	*A. thaliana*	47.63	6.4	372	6	14.32
170	15 DAP es	2.24	P46283	Sedoheptulose-1,7-bisphosphatase, chloroplastic	*A. thaliana*	42.39	6.2	1191	22	32.06
1339	15 DAP es	1.55	P29196	Phosphoenolpyruvate carboxylase	*S. tuberosum*	110.24	5.4	234	5	4.87
778	15 DAP es	1.83	Q56WN1	Glutamine synthetase cytosolic isozyme 1-1	*A. thaliana*	39.09	5.2	396	5	20.51
909	15 DAP es	1.78	AT3G01500.1	Carbonic anhydrase 1, chloroplastic	*A. thaliana*	29.49	5.4	314	5	13.70
601	15 DAP es	3.11	Q9SJ12	Probable ATP synthase 24 kDa subunit, mitochondrial	*A. thaliana*	27.58	6.3	359	7	22.92
601	15 DAP es	3.11	AT2G18230.1	Pyrophosphorylase 2	*A. thaliana*	24.66	5.7	71	2	13.30
794	15 DAP es	1.70	P11574	V-type proton ATPase subunit B1	*A. thaliana*	54.07	4.8	1537	13	28.60
180	15 DAP es	1.98	P41344	Ferredoxin-NADP reductase, leaf isozyme, chloroplastic	*O. sativa*	39.98	9.4	74	2	4.70
239	15 DAP es	6.55	P21276	Oxygen-evolving enhancer protein 2, chloroplastic	*S. alba*	27.91	7.6	289	4	23.08
284	15 DAP es	1.54	XP_002502857	Early light induced protein-like 5, chloroplast	*M. sp. RCC299*	27.77	4.3	70	2	3.72
312	15 DAP es	2.41	A4QK04	ATP synthase subunit b, chloroplastic	*A. hirsuta*	21.13	9.1	689	8	38.59
584	15 DAP es	1.86	P11594	Oxygen-evolving enhancer protein 2, chloroplastic	*S. alba*	27.91	7.6	234	5	25.00
785	15 DAP es	1.52	P23321	Oxygen-evolving enhancer protein 1-1, chloroplastic	*A. thaliana*	35.12	5.4	1562	17	52.41
1276	15 DAP es	1.52	XP_002878030	Photosystem II reaction center PsbP family protein	*A. lyrata*	25.55	9.9	215	6	22.27
170	15 DAP es	2.24	P12333	Chlorophyll a–b binding protein, chloroplastic	*S. oleracea*	28.41	5.2	62	2	6.74
385	15 DAP es	4.81	AT1G03600.1	photosystem II family protein	*A. thaliana*	18.82	10.3	310	4	17.24
405	15 DAP es	2.27	Q41932	Oxygen-evolving enhancer protein 3-2, chloroplastic	*A. thaliana*	24.63	10.1	89	3	15.22
1336	15 DAP es	3.03	AT4G04640.1	ATPase, F1 complex, gamma subunit protein	*A. thaliana*	40.89	9.1	80	1	2.68
1336	15 DAP es	3.03	O23324	ATP-sulfurylase 3, chloroplastic	*A. thaliana*	52.00	7.1	89	1	1.94
1336	15 DAP es	3.03	AT4G14680.1	Pseudouridine synthase/archaeosine transglycosylase-like family protein	*A. thaliana*	52.00	7.1	88	1	1.94
Environmental information processing										
766	15 DAP es	1.52	Q9SYT0	Annexin D1	*A. thaliana*	36.18	5.1	1504	21	51.74
767	15 DAP es	1.63	Q9SYT0	Annexin D1	*A. thaliana*	36.18	5.1	1588	17	47.00

780	15 DAP es	2.64	Q9SYT0	Annexin D1	*A. thaliana*	36.18	5.1	200	3	8.52
939	15 DAP es	3.30	XP_002512810	Diphosphoinositol polyphosphate phosphohydrolase, putative (SP: nudix hydrolase family)	*R. communis*	26.65	5.4	288	2	10.43
1205	15 DAP es	1.59	Q9LX08	Annexin D6	*A. thaliana*	36.55	8.7	548	11	26.42
1425	15 DAP es	2.37	Q9XEE2	Annexin D2	*A. thaliana*	36.24	5.7	1276	11	29.65
1355	15 DAP es	3.25	ADD74397	annexin 4	*B. juncea*	35.50	7.9	416	9	36.83
779	15 DAP es	2.45	AT1G62380.1	ACC oxidase 2	*A. thaliana*	36.16	4.8	68	2	4.38
412	15 DAP es	10.44	P49310	Glycine-rich RNA-binding protein GRP1A	*S. alba*	16.01	5.1	165	2	16.27
433	15 DAP es	2.68	Q05966	Glycine-rich RNA-binding protein 10	*B. napus*	16.29	5.4	78	1	5.92
722	15 DAP es	3.18	Q9SVD7	Ubiquitin-conjugating enzyme E2 variant 1D	*A. thaliana*	16.52	6.2	919	12	65.07
1175	15 DAP es	1.80	AT1G30580.1	GTP binding protein	*A. thaliana*	44.44	6.4	800	10	32.23
1205	15 DAP es	1.59	Q39336	Guanine nucleotide-binding protein subunit beta-like protein	*B. napus*	35.70	9.1	750	14	40.98
315	15 DAP es	2.14	ABV89642	Universal stress protein family protein	*B. rapa*	17.58	9.0	260	4	42.14
315	15 DAP es	2.14	AT3G03270.1	Adenine nucleotide alpha hydrolases-like superfamily protein	*A. thaliana*	22.59	5.4	143	2	15.42
417	15 DAP es	2.50	AT3G17210.1	Heat stable protein 1	*A. thaliana*	12.18	5.4	354	2	20.18
601	15 DAP es	3.11	AT1G10370.1	Glutathione S-transferase family protein	*A. thaliana*	25.29	6.2	116	4	12.33
618	15 DAP es	4.89	AT1G10370.1	Glutathione S-transferase family protein	*A. thaliana*	25.29	6.2	259	7	19.38
969	15 DAP es	5.82	ABD36807	Glutathione S-transferase	*B. napus*	24.72	5.8	104	2	9.68
1089	15 DAP es	3.63	Q9FNE2	Glutaredoxin-C2	*A. thaliana*	11.75	7.7	209	2	9.91
1157	15 DAP es	1.67	AT4G27585.1	PHB domain-containing membrane-associated protein family	*A. thaliana*	44.99	6.4	89	3	8.03
1276	15 DAP es	1.52	AT2G47710.1	Adenine nucleotide alpha hydrolases-like superfamily protein	*A. thaliana*	17.29	8.9	95	3	16.05
909	15 DAP es	1.78	AT1G53580.1	Glyoxalase II 3	*A. thaliana*	32.31	6.6	159	2	3.74
909	15 DAP es	1.78	Q9C8L4	Hydroxyacylglutathione hydrolase 3, mitochondrial	*A. thaliana*	32.31	6.6	159	2	3.74
779	15 DAP es	2.45	CAC85247	Salt tolerance protein 5	*B. vulgaris*	33.28	4.8	81	1	3.05
Genetic information processing										
939	15 DAP es	3.30	AT1G29880.1	Nudix hydrolase homolog 12 (SP: Glycyl-tRNA synthetase 1)	*A. thaliana*	23.85	4.6	189	5	14.78
1411	15 DAP es	3.06	O23627	Glycyl-tRNA synthetase 1, mitochondrial	*A. thaliana*	81.89	6.6	200	6	8.09
552	15 DAP es	2.18	Q05966	G.-rich RNA-binding protein 10	*B. napus*	16.29	5.4	523	4	19.53
856	15 DAP es	2.24	AAF31402	Putative G.-rich RNA binding protein 1	*C. roseus*	14.15	9.6	50	2	10.95
412	15 DAP es	10.44	CAA16739	Pollen-specific protein-like	*A. thaliana*	90.12	5.8	99	1	1.19
43	15 DAP es	2.14	AT5G50920.1	CLPC homolog 1 (SP: Chaperone)	*A. thaliana*	103.39	6.4	3615	52	49.73
234	15 DAP es	1.61	AT3G22630.1	20S proteasome beta subunit D1	*A. thaliana*	22.53	5.9	402	6	26.47
396	15 DAP es	2.64	Q9SCN8	Cell division control protein 48 homolog A	*A. thaliana*	89.34	5.0	2709	33	40.30
601	15 DAP es	3.11	O23715	Proteasome subunit alpha type-3	*A. thaliana*	27.36	5.9	307	6	24.10
722	15 DAP es	3.18	AT5G67360.1	Subtilase family protein	*A. thaliana*	79.37	5.9	77	3	4.23
794	15 DAP es	1.70	P34066	Proteasome subunit alpha type-1-A	*A. thaliana*	30.46	4.8	429	5	20.14
794	15 DAP es	1.70	AAL60581	Senescence-associated cysteine protease	*B. oleracea*	40.19	7.8	120	2	9.24
856	15 DAP es	2.24	Q9LZF6	Cell division control protein 48 homolog E	*A. thaliana*	89.90	4.9	974	17	23.09
1028	15 DAP es	2.66	AAL60582	Senescence-associated cysteine protease	*B. oleracea*	39.31	5.4	93	2	13.93
1028	15 DAP es	2.66	AT5G60360.1	Aleurain-like protease	*A. thaliana*	38.93	6.3	53	2	8.10
1104	15 DAP es	1.53	XP_002522624	Immunophilin, putative (SP: peptidy-prolyl isomerase)	*R. communis*	12.00	9.2	499	4	14.29
1134	15 DAP es	1.62	AT3G11830.1	TCP-1/cpn60 chaperonin family protein	*A. thaliana*	59.74	6.0	1427	15	28.73
1151	15 DAP es	3.00	AT3G44110.1	DNAJ homolog 3	*A. thaliana*	46.41	5.7	51	2	6.43

(continued on next page)

Table 2 (*continued*)

ID [a]	Stage [b]	Reg [c]	Accession [d]	Name	Organism	MM calc [e]	pI calc [f]	Score [g]	Pep [h]	SC% [i]
Genetic information processing										
1416	15 DAP es	3.22	AT4G04460.1	Saposin-like aspartyl protease family protein	*A. thaliana*	55.54	6.9	51	2	3.54
1425	15 DAP es	2.37	Q9LT08	26S proteasome non-ATPase regulatory subunit 14	*A. thaliana*	34.33	6.3	245	4	12.99
1477	15 DAP es	4.08	CAZ40339	Putative aminopeptidase	*R. sativus*	99.22	5.8	130	9	11.29
405	15 DAP es	2.27	NP_566406	Adenine nucleotide alpha hydrolases-like protein	*A. thaliana*	21.44	5.5	111	2	18.09
1176	15 DAP es	1.90	O22263	Protein disulfide-isomerase like 2-1	*A. thaliana*	39.47	5.7	190	3	9.70
1237	15 DAP es	2.30	AT3G51260.1	20S proteasome alpha subunit PAD1	*A. thaliana*	27.32	7.7	684	11	37.60
1237	15 DAP es	2.30	Q9FK35	Adenylate kinase 2	*A. thaliana*	27.32	7.7	149	4	14.92
1287	15 DAP es	5.13	Q9SAA2	ATP-dependent Clp protease proteolytic subunit 6, chloroplastic	*A. thaliana*	29.36	10.0	128	1	3.32
170	15 DAP es	2.24	AT3G63460.1	Transducin family protein/WD-40 repeat family protein	*A. thaliana*	119.80	4.9	57	1	1.00
1028	15 DAP es	2.66	Q9LHG9	Nascent polypeptide-associated complex subunit alpha-like protein 1	*A. thaliana*	21.97	4.1	184	4	19.21
234	15 DAP es	1.61	Q9SFD8	Nuclear transcription factor Y subunit B-9	*A. thaliana*	26.05	5.6	171	2	11.34
722	15 DAP es	3.18	Q84W89	DEAD-box ATP-dependent RNA helicase 37	*A. thaliana*	67.58	6.7	1356	22	28.59
897	15 DAP es	1.69	AT1G07660.1	Histone superfamily protein	*A. thaliana*	11.40	12.0	98	2	21.36
939	15 DAP es	3.30	AAF78493	Contains similarity to AP2/EREBP-like transcription factor	*A. thaliana*	47.48	4.8	122	3	5.84
1352	15 DAP es	9.11	AT2G18510.1	RNA-binding (RRM/RBD/RNP motifs) family protein (SP: putative spliceosome associated protein)	*A. thaliana*	39.86	7.9	274	3	9.92
41	15 DAP es	2.44	AT1G56070.1	Ribosomal protein S5/Elongation factor G/III/V family protein	*A. thaliana*	93.83	5.9	2348	38	37.60
466	15 DAP es	2.99	AT1G23410.1	Ribosomal protein S27a	*A. thaliana*	17.66	10.4	796	9	41.67
474	15 DAP es	4.16	AT1G33120.1	Ribosomal protein L6 family	*A. thaliana*	22.00	10.1	323	3	14.90
764	15 DAP es	2.60	Q9ZSR8	40S ribosomal protein SA	*B. napus*	32.01	5.0	888	12	39.38
897	15 DAP es	1.69	AT2G40010.1	Ribosomal protein L10 family protein	*A. thaliana*	33.65	5.0	193	2	7.57
1157	15 DAP es	1.67	O04487	Probable elongation factor 1-gamma 1	*A. thaliana*	46.63	5.2	1476	11	22.46
1383	15 DAP es	697.77	Q8LBI1	60S ribosomal protein L5-1	*A. thaliana*	34.34	9.7	510	8	24.25
417	15 DAP es	2.50	Q9LUV2	Probable protein Pop3	*A. thaliana*	12.18	5.4	361	2	20.18
601	15 DAP es	3.11	O04663	Eukaryotic translation initiation factor 4E-2	*A. thaliana*	22.50	5.4	118	4	13.64
722	15 DAP es	3.18	O64380	Polyadenylate-binding protein 3	*A. thaliana*	72.83	9.3	105	3	3.64
726	15 DAP es	1.99	P42731	Polyadenylate-binding protein 2	*A. thaliana*	68.63	8.8	1011	18	19.87
729	15 DAP es	7.43	P42731	Polyadenylate-binding protein 2	*A. thaliana*	68.63	8.8	867	13	15.42
882	15 DAP es	2.38	P41376	Eukaryotic initiation factor 4A-1	*A. thaliana*	46.67	5.4	1884	29	47.09
913	15 DAP es	2.27	P41376	Eukaryotic initiation factor 4A-1	*A. thaliana*	46.67	5.4	221	4	11.65
954	15 DAP es	2.89	AAR91929	Eukaryotic translation initiation factor-5A	*B. napus*	17.09	5.7	474	9	57.86
1024	15 DAP es	6.67	ABK78691	Putative elongation factor 1-beta	*B. rapa*	25.04	4.3	1465	14	63.48
1028	15 DAP es	2.66	Q84WM9	Elongation factor 1-beta 1	*A. thaliana*	24.77	4.4	453	12	15.35
1096	15 DAP es	4.62	Q9C505	Eukaryotic translation initiation factor 5A-3	*A. thaliana*	17.20	5.5	478	5	24.05
1416	15 DAP es	3.22	AT3G62120.1	Class II aaRS and biotin synthetases superfamily protein (SP: multifunctional aminoacyl-tRNA ligase-like protein)	*A. thaliana*	60.72	6.1	1833	30	50.38
170	15 DAP es	2.24	Q9C5Z2	Eukaryotic translation initiation factor 3 subunit H	*A. thaliana*	38.35	4.7	215	4	9.79
381	15 DAP es	6.62	NP_197529	40S ribosomal protein S8-1	*A. thaliana*	24.98	10.9	214	4	16.67
433	15 DAP es	2.68	Q9SUM2	Probable small nuclear ribonucleoprotein F	*A. thaliana*	9.86	4.1	93	1	9.09
1219	15 DAP es	42.2	Q8LBI1	60S ribosomal protein L5-1	*A. thaliana*	34.34	9.7	89	2	4.98

1287	15 DAP es	5.13	Q40467	Eukaryotic initiation factor 4A-14	*N. tabacum*	46.85	5.3	270	7	11.62
1355	15 DAP es	3.25	Q8LBI1	60S ribosomal protein L5-1	*A. thaliana*	34.34	9.7	68	2	7.31
Lipid metabolism										
122	15 DAP es	2.28	AAK60339	Biotin carboxylase	*B. napus*	58.38	6.6	51	2	5.23
464	15 DAP es	5.39	Q39315	Acyl-CoA-binding protein	*B. napus*	10.17	5.3	424	6	68.48
897	15 DAP es	1.69	P80030	Enoyl-acyl-carrier-protein reductase NADH, chloroplastic	*B. napus*	40.45	9.4	618	11	25.45
433	15 DAP es	2.68	P08971	Acyl carrier protein, chloroplastic	*B. napus*	14.69	4.9	211	1	11.19
1176	15 DAP es	1.90	Q9MA55	Acyl-CoA-binding domain-containing protein 4	*A. thaliana*	73.03	5.0	70	2	3.59
1355	15 DAP es	3.25	AT1G76150.1	Enoyl-CoA hydratase 2	*A. thaliana*	34.06	7.7	212	5	19.42
1477	15 DAP es	4.08	AAO03559	Lipoxygenase 2	*B. napus*	101.38	5.1	195	4	5.16
Metabolism of terpenoids and polyketides										
796	15 DAP es	1.93	Q42553	Isopentenyl-diphosphate Delta-isomerase II	*A. thaliana*	32.59	6.1	507	14	44.37
Nucleotide metabolism										
1241	15 DAP es	2.21	NP_199848	Adenylate kinase 2	*A. thaliana*	27.30	7.7	83	2	6.50
1416	15 DAP es	3.22	AT1G63660.1	GMP synthase (glutamine-hydrolyzing), putative/ glutamine amidotransferase, putative	*A. thaliana*	59.28	6.1	78	4	7.30
170	15 DAP es	2.24	AT4G14930.1	Survival protein SurE-like phosphatase/ nucleotidase	*A. thaliana*	34.10	4.9	97	1	3.49
Storage										
796	15 DAP es	1.93	AAM65577	Globulin-like protein	*A. thaliana*	38.29	5.8	376	6	13.48
1276	15 DAP es	1.52	P27740	Napin-B	*B. napus*	20.10	9.0	345	5	34.27
779	15 DAP es	2.45	AT1G07750.1	RmlC-like cupins superfamily protein	*A. thaliana*	38.29	5.8	401	8	14.04
Uncharacterized										
216	15 DAP es	2.25	ACU18911	Unknown	*G. max*	28.49	6.2	60	2	11.46
234	15 DAP es	1.61	ACU13447	Unknown	*G. max*	22.58	5.9	104	2	20.59
552	15 DAP es	2.18	BAJ34207	Unnamed protein product	*T. halophila*	16.24	5.4	451	7	28.31
813	15 DAP es	2.24	XP_002887418	Hypothetical protein ARALYDRAFT_895064	*A. lyrata*	62.02	4.7	474	15	19.86
813	15 DAP es	2.24	A5BX00	Hypothetical protein V.V_028074	*V. vinifera*	123.63	5.6	62	2	2.21
1276	15 DAP es	1.52	BAJ33672	Unnamed protein product	*T. halophila*	17.49	6.9	167	5	42.59
170	15 DAP es	2.24	CAB10272	Hypothetical protein	*A. thaliana*	30.20	5.3	84	1	4.00
170	15 DAP es	2.24	CAB87799	Putative protein	*A. thaliana*	119.00	4.8	57	1	1.00
405	15 DAP es	2.27	XP_002500296	Predicted protein	*Micromonas* sp.	271.99	6.0	75	1	0.55
1176	15 DAP es	1.90	BAB08430	GAMM1 protein-like	*A. thaliana*	39.31	5.5	61	2	5.20

[a] Spot number in accordance with Supplementary Fig. 2.
[b] Developmental stage of endosperm (es) 10 and 15 DAP.
[c] Regulation of spots, spot volume at least 1.5 fold increased in the corresponding developmental stage.
[d] Accession numbers as given by SwissProt (http://www.uniprot.org), NCBInr (http://www.ncbi.nlm.nih.gov/protein) and TAIR (http://www.Arabidopsis.org/).
[e] Calculated molecular masses of the identified proteins as deduced from the corresponding genes.
[f] Calculated isoelectric point of the identified proteins as deduced from the corresponding genes.
[g] Probability score for the protein identifications based on MS/MS analysis and MASCOT search.
[h] Number of unique matching peptides.
[i] Sequence coverage of a protein by identified peptides.

Table 3 – Identification of proteins from spots with changed abundance in the endosperm between 15 DAP and 20 DAP. To determine significant changes in spot volume, a Student's t-test (p-value ≤ 0.05) was applied on the basis of normalized relative spot volume. Changes in spot volume ≥ 1.5 were considered to represent true alterations. Identification of proteins was carried out using the MASCOT search algorithm (www.matrixscience.com) against the (i) SwissProt (www.uniprot.org), (ii) NCBInr (www.ncbi.nig.gov) and (iii) TAIR (www.arabidopsis.org, TAIR release 10) databases. Identified proteins were functionally classified according to the KEGG PATHWAY Database (http://www.genome.jp/kegg/pathway.html). The pathway categories were adjusted for seed metabolism by adding four functional groups: (i) storage, (ii) defense, (iii) desiccation and (iv) detoxification.

ID[a]	Stage[b]	Reg[c]	Accession[d]	Name	Organism	MM calc[e]	pI calc[f]	Score[g]	Pep[h]	SC%[i]
Amino acid metabolism										
900	15 DAP es	2.36	AT1G79230.1	Mercaptopyruvate sulfurtransferase 1	*A. thaliana*	41.87	5.94	79	1	2.37
122	15 DAP es	1.79	P30184	Leucine aminopeptidase 1	*A. thaliana*	54.48	5.58	271	4	8.85
143	15 DAP es	1.56	AT1G12050.1	Fumarylacetoacetase, putative	*A. thaliana*	46.07	5.21	125	3	10.93
Biosynthesis of secondary metabolites										
527	15 DAP es	1.74	AT1G74470.1	Pyridine nucleotide-disulfide oxidoreductase family protein	*A. thaliana*	51.80	9.63	793	6	15.20
881	15 DAP es	1.50	P16127	Magnesium-chelatase subunit chlI, chloroplastic	*A. thaliana*	46.24	6.06	55	2	7.55
Carbohydrate metabolism										
849	15 DAP es	2.02	Q93WJ8	Probable monodehydroascorbate reductase, cytoplasmic isoform 4	*A. thaliana*	47.45	5.11	210	2	5.98
122	15 DAP es	1.79	AT5G57655.2	Xylose isomerase family protein	*A. thaliana*	53.69	5.52	284	6	13.00
985	15 DAP es	2.01	Q9SID0	Probable fructokinase-1	*A. thaliana*	35.25	5.19	1238	11	32.00
122	15 DAP es	1.79	Q9SEE5	Galactokinase	*A. thaliana*	54.31	5.61	150	2	4.64
122	15 DAP es	1.79	P25696	Bifunctional enolase 2/transcriptional activator	*A. thaliana*	47.69	5.45	714	11	20.72
779	15 DAP es	1.72	Q38799	Pyruvate dehydrogenase E1 component subunit beta, mitochondrial	*A. thaliana*	39.15	5.57	76	1	3.03
900	15 DAP es	2.36	XP_002893304	Oxidoreductase (SP: alcohol dehydrogenase activity)	*A. lyrata*	40.81	8.64	100	3	10.88
1146	15 DAP es	2.24	AAY86035	Pyruvate kinase	*C. sinensis*	55.53	9.33	61	2	4.31
422	15 DAP es	2.45	AAP37972	Seed specific protein Bn15D33A	*B. napus*	12.66	9.24	55	1	7.76
143	15 DAP es	1.56	AT1G10670.1	ATP-citrate lyase A-1	*A. thaliana*	46.65	5.26	358	7	17.02
232	15 DAP es	1.56	ADK98071	Oxoglutarate dehydrogenase	*S. adamantis*	16.30	5.60	55	2	7.00
Cellular processes										
370	15 DAP es	3.81	XP_002501166	Multidrug/oligosaccharidyl-lipid/polysaccharide flippase	*Micromonas* sp.	62.39	10.40	51	1	1.18
849	15 DAP es	2.02	Q8L5U0	COP9 signalosome complex subunit 4	*A. thaliana*	44.93	4.72	751	7	17.63
383	15 DAP es	2.45	AAB18643	Actin	*P. sativum*	31.06	5.27	77	1	6.76
881	15 DAP es	1.50	O81644	Villin-2	*A. thaliana*	107.78	5.05	72	2	3.07
1278	15 DAP es	2.18	AT1G04820.1	Tubulin alpha-4 chain	*A. thaliana*	49.51	4.79	81	4	10.67
488	15 DAP es	2.20	AAS68185	Lipid transfer-like protein	*B. napus*	11.18	9.19	203	3	33.33
Defense										
685	15 DAP es	2.00	Q9SE50	Beta-glucosidase 18	*A. thaliana*	60.42	6.82	310	7	8.71
1278	15 DAP es	2.18	Q84WV2	Beta-glucosidase 20	*A. thaliana*	61.64	5.65	110	1	2.43
1278	15 DAP es	2.18	AT1G75940.1	Glycosyl hydrolase superfamily protein	*A. thaliana*	61.64	5.65	109	1	2.43
Detoxification										
985	15 DAP es	2.01	AT2G37790.1	NAD(P)-linked oxidoreductase superfamily protein (SP: Aldo-keto reductase family 4 member C10)	*A. thaliana*	34.89	6.21	109	2	5.41
Energy metabolism										
878	15 DAP es	1.82	P25697	Phosphoribulokinase, chloroplastic	*A. thaliana*	44.44	5.65	1048	13	25.32
881	15 DAP es	1.50	P25697	Phosphoribulokinase, chloroplastic	*A. thaliana*	44.44	5.65	347	6	12.41

1076	15 DAP es	1.81	P05346	Ribulose bisphosphate carboxylase small chain, chloroplastic	*B. napus*	20.21	9.22	1200	12	58.01
1286	15 DAP es	2.06	AT5G38430.1	Ribulose bisphosphate carboxylase (small chain) family protein181	*A. thaliana*	20.27	8.86	723	7	31.49
1313	15 DAP es	2.11	P10796	Ribulose bisphosphate carboxylase small chain 1B, chloroplastic	*A. thaliana*	20.27	8.86	447	5	17.13
170	15 DAP es	2.27	P46283	Sedoheptulose-1,7-bisphosphatase, chloroplastic	*A. thaliana*	42.39	6.18	1191	22	32.06
946	15 DAP es	5.08	AT1G12250.1	Pentapeptide repeat-containing protein	*A. thaliana*	30.05	9.64	57	2	9.64
239	15 DAP es	1.55	P21276	Oxygen-evolving enhancer protein 2, chloroplastic	*S. alba*	27.91	7.60	289	4	23.08
312	15 DAP es	2.30	A4QK04	ATP synthase subunit b, chloroplastic	*A. hirsuta*	21.13	9.07	689	8	38.59
442	15 DAP es	2.11	P49107	Photosystem I reaction center subunit N, chloroplastic	*A. thaliana*	18.42	10.00	575	6	33.33
968	15 DAP es	1.63	P11594	Oxygen-evolving enhancer protein 2, chloroplastic	*S. alba*	27.91	7.60	601	8	33.85
170	15 DAP es	2.27	P12333	Chlorophyll a–b binding protein, chloroplastic	*S. oleracea*	28.41	5.18	62	2	6.74
170	15 DAP es	2.27	AT2G34430.1	Light-harvesting chlorophyll–protein complex II subunit B1	*A. thaliana*	28.15	5.01	62	2	9.40
435	15 DAP es	1.58	AT5G64040.1	Photosystem I reaction center subunit PSI-N, chloroplast, putative/PSI-N, putative (PSAN)	*A. thaliana*	18.42	10.00	60	1	5.85
685	15 DAP es	2.00	AT5G04590.1	Sulfite reductase	*A. thaliana*	71.90	9.17	365	13	18.85
900	15 DAP es	2.36	ACB59214	Cytoplasmic thiosulfate:cyanide sulfur transferase	*B. oleracea*	41.51	6.93	87	2	5.01
Environmental information processing										
939	15 DAP es	2.57	XP_002512810	Diphosphoinositol polyphosphate phosphohydrolase, putative (SP: Nudix Hydrolase family)	*R. communis*	26.65	5.37	288	2	10.43
383	15 DAP es	2.45	CAA46591	BnD22 drought induced protein	*B. napus*	23.52	5.84	70	1	5.50
1278	15 DAP es	2.18	CAA46591	BnD22 drought induced protein	*B. napus*	23.52	5.84	140	1	5.50
779	15 DAP es	1.72	AT1G62380.1	ACC oxidase 2	*A. thaliana*	36.16	4.84	68	2	4.38
494	15 DAP es	2.19	Q9SHE7	Ubiquitin-NEDD8-like protein RUB1	*A. thaliana*	17.39	5.71	189	4	21.15
1071	15 DAP es	3.06	P48347	14-3-3-like protein GF14 epsilon	*A. thaliana*	28.90	4.57	1285	14	34.65
1086	15 DAP es	2.66	P69310	Ubiquitin	*A. sativa*	8.52	7.58	510	8	77.63
1094	15 DAP es	1.60	P69310	Ubiquitin	*A. sativa*	8.52	7.58	428	8	77.63
262	15 DAP es	2.11	CAA07494	Heat stress-induced protein	*B. oleracea*	23.47	9.23	489	7	24.31
417	15 DAP es	2.11	AT3G17210.1	Heat stable protein 1	*A. thaliana*	12.18	5.35	354	2	20.18
1082	15 DAP es	1.51	AT5G20500.1	Glutaredoxin family protein	*A. thaliana*	14.82	5.62	130	2	11.85
946	15 DAP es	5.08	CAA07494	Heat stress-induced protein	*B. oleracea*	23.47	9.23	904	8	24.31
946	15 DAP es	5.08	BAB72020	Water-soluble chlorophyll protein	*R. sativus*	23.83	9.53	341	1	14.41
Genetic information processing										
939	15 DAP es	2.57	AT1G29880.1	Nudix hydrolase homolog 12 (SP: Glycyl-tRNA synthetase 1)	*A. thaliana*	23.85	4.59	189	5	14.78
483	15 DAP es	2.19	XP_003569976	Vacuolar-processing enzyme beta-isozyme-like	*B. distachyon*	54.56	6.15	99	1	2.65
80	15 DAP es	1.67	AT5G10540.1	Zincin-like metalloproteases family protein	*A. thaliana*	78.99	5.34	122	4	5.28
143	15 DAP es	1.56	AT5G58290.1	Regulatory particle triple-A ATPase 3	*A. thaliana*	45.72	5.30	1298	22	50.98
354	15 DAP es	1.53	AT1G29990.1	Prefoldin 6	*A. thaliana*	14.85	9.57	357	3	22.48
800	15 DAP es	5.16	Q9SGW3	26S proteasome non-ATPase regulatory subunit RPN12A	*A. thaliana*	30.69	4.67	962	8	39.70
836	15 DAP es	1.54	Q9SRG3	Protein disulfide isomerase-like 1-2	*A. thaliana*	56.33	4.76	993	8	15.94
837	15 DAP es	5.67	Q9XI01	Protein disulfide isomerase-like 1-2	*A. thaliana*	56.33	4.76	1104	9	19.09
906	15 DAP es	4.42	AT5G66190.1	TCP-1/cpn60 chaperonin family protein	*A. thaliana*	58.89	5.84	200	4	8.41
947	15 DAP es	2.18	Q8LD27	Proteasome subunit beta type-6	*A. thaliana*	25.14	5.21	104	2	10.73
1104	15 DAP es	1.53	XP_002522624	Immunophilin, putative (SP: peptidy-prolyl isomerase)	*R. communis*	12.00	9.17	499	4	14.29
422	15 DAP es	2.45	NP_567632	Subtilase family protein	*A. thaliana*	82.61	7.88	88	2	1.17
435	15 DAP es	1.58	XP_002863802	Ubiquitin family protein	*A. lyrata*	8.58	9.01	351	10	68.49
946	15 DAP es	5.08	XP_002301464	Peptidyl-prolyl cis-trans isomerase	*P. trichocarpa*	29.07	10.04	126	2	4.51
1146	15 DAP es	2.24	AT5G20890.1	TCP-1/cpn60 chaperonin family protein	*A. thaliana*	57.25	5.52	277	10	14.99

(continued on next page)

Table 3 (*continued*)

ID [a]	Stage [b]	Reg [c]	Accession [d]	Name	Organism	MM calc [e]	pI calc [f]	Score [g]	Pep [h]	SC% [i]
Genetic information processing										
170	15 DAP es	2.27	AT3G63460.1	Transducin family protein/WD-40 repeat family protein	*A. thaliana*	119.80	4.86	57	1	1.00
779	15 DAP es	1.72	CAC85247	Salt tolerance protein 5	*B. vulgaris*	33.28	4.84	81	1	3.05
952	15 DAP es	1.55	O48646	Probable phospholipid hydroperoxide glutathione peroxidase 6, mitochondrial	*A. thaliana*	25.57	9.95	495	6	20.26
939	15 DAP es	2.57	AAF78493	Contains similarity to AP2/EREBP-like transcription factor	*A. thaliana*	47.48	4.80	122	3	5.84
143	15 DAP es	1.56	P41379	Eukaryotic initiation factor 4A-2	*N. plumbaginifolia*	46.80	5.26	321	7	19.61
269	15 DAP es	2.45	Q40468	Eukaryotic initiation factor 4A-15	*N. tabacum*	46.69	5.26	717	11	17.43
417	15 DAP es	2.11	Q9LUV2	Probable protein Pop3	*A. thaliana*	12.18	5.35	361	2	20.18
493	15 DAP es	2.11	AT2G43460.1	60S ribosomal protein L38	*A. thaliana*	8.12	10.64	414	5	60.87
511	15 DAP es	1.58	Q9LVC9	60S acidic ribosomal protein P3-2	*A. thaliana*	11.86	4.29	311	2	15.83
527	15 DAP es	1.74	P29521	Elongation factor 1-alpha	*D. carota*	49.27	9.83	217	3	6.24
530	15 DAP es	1.73	P13905	Elongation factor 1-alpha	*A. thaliana*	49.47	9.79	411	2	9.80
533	15 DAP es	2.06	AT2G03870.1	Small nuclear ribonucleoprotein family protein	*A. thaliana*	10.77	4.89	119	4	25.25
1041	15 DAP es	7.54	P48006	Elongation factor 1-delta 1	*A. thaliana*	25.12	4.28	770	7	27.71
1096	15 DAP es	3.74	Q9C505	Eukaryotic translation initiation factor 5A-3	*A. thaliana*	17.20	5.51	478	5	24.05
1383	15 DAP es	3.10	Q8LBI1	60S ribosomal protein L5-1	*A. thaliana*	34.34	9.71	510	8	24.25
170	15 DAP es	2.27	Q9C5Z2	Eukaryotic translation initiation factor 3 subunit H	*A. thaliana*	38.35	4.68	215	4	9.79
370	15 DAP es	3.81	BAB02821	Similarity to protein translation inhibitor	*A. thaliana*	15.15	5.33	60	3	23.08
381	15 DAP es	2.76	NP_197529	40S ribosomal protein S8-1	*A. thaliana*	24.98	10.86	214	4	16.67
383	15 DAP es	2.45	AT5G59950.2	RNA-binding (RRM/RBD/RNP motifs) family protein	*A. thaliana*	18.88	9.71	91	4	17.98
422	15 DAP es	2.45	NP_190957	40S ribosomal protein S21-1	*A. thaliana*	9.07	9.23	59	1	13.41
435	15 DAP es	1.58	AT1G54270.1	eif4a-2	*A. thaliana*	46.73	5.35	153	1	2.43
483	15 DAP es	2.19	AT1G75350.1	Ribosomal protein L31	*A. thaliana*	16.02	10.49	53	1	8.33
900	15 DAP es	2.36	AT1G56070.1	Ribosomal protein S5/Elongation factor G/III/V family protein	*A. thaliana*	93.83	5.85	1507	20	13.64
900	15 DAP es	2.36	BAJ33766	Unnamed protein product (Elongation factor EF-2)	*T. halophila*	93.88	5.85	940	19	14.23
946	15 DAP es	5.08	AT1G54270.1	eif4a-2	*A. thaliana*	46.73	5.35	52	1	4.37
952	15 DAP es	1.55	Q9XI91	Eukaryotic translation initiation factor 5A-1	*A. thaliana*	17.35	5.34	339	3	27.22
1219	15 DAP es	2.67	Q8LBI1	60S ribosomal protein L5-1	*A. thaliana*	34.34	9.71	89	2	4.98
1274	15 DAP es	2.66	AT2G40660.1	Nucleic acid-binding, OB-fold-like protein	*A. thaliana*	42.06	7.70	53	1	2.06
1278	15 DAP es	2.18	AT1G03810.1	Nucleic acid-binding, OB-fold-like protein	*A. thaliana*	15.66	10.23	55	2	5.59
Lipid metabolism										
122	15 DAP es	1.79	AAK60339	Biotin carboxylase	*B. napus*	58.38	6.57	51	2	5.23
464	15 DAP es	4.51	Q39315	Acyl-CoA-binding protein	*B. napus*	10.17	5.28	424	6	68.48
985	15 DAP es	2.01	P80030	Enoyl-[acyl-carrier-protein] reductase [NADH], chloroplastic	*B. napus*	40.45	9.39	174	3	10.65
Nucleotide metabolism										
881	15 DAP es	1.50	Q9LZG0	Adenosine kinase 2	*A. thaliana*	37.82	5.01	1729	16	59.13
170	15 DAP es	2.27	AT4G14930.1	Survival protein SurE-like phosphatase/nucleotidase	*A. thaliana*	34.10	4.91	97	1	3.49
Storage										
779	15 DAP es	1.72	AT1G07750.1	RmlC-like cupins superfamily protein	*A. thaliana*	38.29	5.79	401	8	14.04
Uncharacterized										
80	15 DAP es	1.67	XP_002876580	Hypothetical protein ARALYDRAFT_486548 (SP: Transketolase)	*A. lyrata*	79.80	5.82	1544	22	26.86
269	15 DAP es	2.45	ACJ84369	Unknown	*M. truncatula*	28.22	6.15	360	2	28.10

170	15 DAP es	2.27	CAB10272	Hypothetical protein	*A. thaliana*	30.20	5.25	84	1	4.00
170	15 DAP es	2.27	CAB87799	Putative protein	*A. thaliana*	119.00	4.84	57	1	1.00
946	15 DAP es	5.08	NP_001047560	Os02g0643500	*O. sativa japonica group*	28.94	10.10	128	2	5.42
1274	15 DAP es	2.66	XP_002965025	Hypothetical protein SELMODRAFT_82474	*S. moellendorffii*	35.06	8.77	85	1	2.51
1274	15 DAP es	2.66	CAN60565	Hypothetical protein VITISV_013997	*V. vinifera*	57.79	9.96	76	2	4.97
Amino acid metabolism										
882	20 DAP es	1.62	ABD65618	Acetylornithine deacetylase, putative	*B. oleracea*	44.79	5.10	79	17.2	4.00
1349	20 DAP es	1.83	AT4G38220.1	Peptidase M20/M25/M40 family protein	*A. thaliana*	47.71	5.91	124	8.8	3.00
885	20 DAP es	1.66	Q5DNB1	S-Adenosylmethionine synthase	*B. rapa*	43.16	5.63	617	26.0	11.00
913	20 DAP es	1.81	Q9LDQ7	S-Adenosylmethionine synthase	*C. sinensis*	42.77	5.25	112	9.4	2.00
1156	20 DAP es	2.05	AT2G36880.1	Methionine adenosyltransferase 3	*A. thaliana*	42.47	5.73	1916	61.5	22.00
1157	20 DAP es	2.08	AT4G01850.1	S-Adenosylmethionine synthetase 2	*A. thaliana*	43.23	5.63	246	27.5	5.00
1162	20 DAP es	1.51	Q5DNB1	S-Adenosylmethionine synthase	*B. rapa*	43.16	5.63	393	20.9	5.00
1329	20 DAP es	3.25	Q9FUZ1	S-Adenosylmethionine synthase 2	*B. juncea*	42.85	5.27	873	36.1	10.00
956	20 DAP es	1.78	Q9SRZ4	Peroxiredoxin-2C	*A. thaliana*	17.40	5.22	197	6.2	2.00
129	20 DAP es	1.50	AT4G13930.1	Serine hydroxymethyltransferase 4	*A. thaliana*	51.69	6.98	1322	34.2	16.00
129	20 DAP es	1.50	AT4G13890.1	Pyridoxal phosphate (PLP)-dependent transferases superfamily protein	*A. thaliana*	52.23	5.60	314	8.7	5.00
912	20 DAP es	2.26	AAR13689	Peptide methionine sulfoxide reductase	*B. oleracea*	22.72	5.29	248	42.2	2.00
882	20 DAP es	1.62	Q93ZN9	LL-diaminopimelate aminotransferase, chloroplastic	*A. thaliana*	50.36	7.71	891	22.6	10.00
Biosynthesis of secondary metabolites										
925	20 DAP es	1.58	XP_002864577	cinnamoyl-CoA reductase family	*A. lyrata*	35.39	5.81	57	12.7	2.00
Carbohydrate metabolism										
1157	20 DAP es	2.08	AT1G08200.1	UDP-D-apiose/UDP-D-xylose synthase 2	*A. thaliana*	43.76	5.51	81	5.4	2.00
885	20 DAP es	1.66	AT1G08200.1	UDP-D-apiose/UDP-D-xylose synthase 2	*A. thaliana*	43.76	5.51	51	2.8	1.00
913	20 DAP es	1.81	Q9LFA3	Probable monodehydroascorbate reductase, cytoplasmic isoform 3	*A. thaliana*	46.46	6.47	62	9.7	3.00
1156	20 DAP es	2.05	AT5G28840.1	GDP-D-mannose 3′,5′-epimerase	*A. thaliana*	42.73	5.81	50	6.4	2.00
1157	20 DAP es	2.08	AAK72107	Monodehydroascorbate reductase	*B. rapa*	46.43	5.74	108	30.0	7.00
1162	20 DAP es	1.51	AT5G28840.1	GDP-D-mannose 3′,5′-epimerase	*A. thaliana*	42.73	5.81	227	12.7	5.00
1315	20 DAP es	1.52	Q42564	L-ascorbate peroxidase 3, peroxisomal	*A. thaliana*	31.55	6.53	235	17.8	5.00
885	20 DAP es	1.66	AAK72107	monodehydroascorbate reductase	*B. rapa*	46.43	5.74	129	15.9	5.00
925	20 DAP es	1.58	Q42592	L-Ascorbate peroxidase S, chloroplastic/mitochondrial	*A. thaliana*	40.38	9.01	358	14.5	5.00
99	20 DAP es	2.33	Q9M9K1	Probable 2,3-bisphosphoglycerate-independent phosphoglycerate mutase 2	*A. thaliana*	60.73	5.46	446	15.7	6.00
103	20 DAP es	1.50	O04499	2,3-Bisphosphoglycerate-independent phosphoglycerate mutase 1	*A. thaliana*	60.54	5.20	1520	31.4	18.00
891	20 DAP es	2.07	AT2G21330.1	Fructose-bisphosphate aldolase 1	*A. thaliana*	42.90	6.18	162	10.0	4.00
913	20 DAP es	1.81	P25696	Bifunctional enolase 2/transcriptional activator	*A. thaliana*	47.69	5.45	1458	30.6	11.00
1186	20 DAP es	1.58	AT3G52930.1	Aldolase superfamily protein	*A. thaliana*	38.52	6.04	1282	43.0	17.00
1343	20 DAP es	2.18	O49299	Probable phosphoglucomutase, cytoplasmic 1	*A. thaliana*	63.13	5.88	1014	22.5	12.00
1401	20 DAP es	1.77	Q1WIQ6	NADP-dependent glyceraldehyde-3-phosphate dehydrogenase	*A. thaliana*	53.03	6.24	1024	25.6	16.00
1401	20 DAP es	1.77	AT3G16950.1	Lipoamide dehydrogenase 1	*A. thaliana*	60.72	9.00	87	4.6	2.00
123	20 DAP es	1.58	Q96348	Inositol-3-phosphate synthase	*A. thaliana*	56.34	5.27	868	25.7	13.00
910	20 DAP es	1.77	Q39366	Putative lactoylglutathione lyase	*B. oleracea*	31.63	5.02	906	36.9	11.00
1116	20 DAP es	1.67	AAP37972	Seed specific protein Bn15D33A (SP: galactosyltransferase activity)	*B. napus*	12.66	9.24	835	68.1	8.00
1195	20 DAP es	1.64	AT5G59290.1	UDP-glucuronic acid decarboxylase 3	*A. thaliana*	38.54	9.08	963	45.3	18.00
925	20 DAP es	1.58	BAB01287	Sucrose cleavage protein-like	*A. thaliana*	34.23	5.51	70	2.9	1.00
94	20 DAP es	1.68	O82663	Succinate dehydrogenase [ubiquinone] flavoprotein subunit 1, mitochondrial	*A. thaliana*	69.61	5.84	958	26.2	13.00

(continued on next page)

Table 3 (*continued*)

ID [a]	Stage [b]	Reg [c]	Accession [d]	Name	Organism	MM calc [e]	pI calc [f]	Score [g]	Pep [h]	SC% [i]
Carbohydrate metabolism										
152	20 DAP es	1.89	Q9SJH7	Citrate synthase 3, peroxisomal	*A. thaliana*	56.14	7.79	447	7.1	3.00
154	20 DAP es	1.94	Q9SJH7	Citrate synthase 3, peroxisomal	*A. thaliana*	56.14	7.79	375	7.1	3.00
668	20 DAP es	2.12	Q9SJH7	Citrate synthase 3, peroxisomal	*A. thaliana*	56.14	7.79	561	7.1	3.00
1147	20 DAP es	1.90	AT1G65930.1	Cytosolic NADP+-dependent isocitrate dehydrogenase	*A. thaliana*	45.72	6.13	196	14.1	5.00
1352	20 DAP es	1.71	Q9SJH7	Citrate synthase 3, peroxisomal	*A. thaliana*	56.14	7.79	304	7.1	3.00
Cellular processes										
174	20 DAP es	2.49	Q9FFD2	Probable UDP-arabinopyranose mutase 5	*A. thaliana*	38.56	4.92	1592	31.3	14.00
813	20 DAP es	2.67	Q56WK6	Patellin-1	*A. thaliana*	64.01	4.67	830	18.3	13.00
123	20 DAP es	1.58	AT1G75780.1	Tubulin beta-1 chain	*A. thaliana*	50.18	4.54	67	6.5	3.00
218	20 DAP es	1.53	Q9LW57	Probable plastid-lipid-associated protein 6, chloroplastic1	*A. thaliana*	30.44	5.76	91	11.3	3.00
Defense										
311	20 DAP es	2.33	Q9ZSK3	Actin-depolymerizing factor 4	*A. thaliana*	16.02	6.18	219	20.1	3.00
315	20 DAP es	2.56	Q9ZVF2	MLP-like protein 329	*A. thaliana*	17.59	5.21	383	26.5	4.00
524	20 DAP es	2.00	Q9LV33	Beta-glucosidase 44	*A. thaliana*	58.95	9.81	311	13.1	7.00
959	20 DAP es	1.54	Q9ZVF2	MLP-like protein 329	*A. thaliana*	17.59	5.21	346	23.2	3.00
1100	20 DAP es	1.78	Q9ZVF3	MLP-like protein 328	*A. thaliana*	17.50	5.35	498	37.1	6.00
1401	20 DAP es	1.77	XP_002876291	Glycosyl hydrolase family 20 protein	*A. lyrata*	61.54	5.85	90	9.4	6.00
Energy metabolism										
778	20 DAP es	1.54	Q56WN1	Glutamine synthetase cytosolic isozyme 1-1	*A. thaliana*	39.09	5.17	396	20.5	5.00
607	20 DAP es	1.58	Q9SJ12	Probable ATP synthase 24 kDa subunit, mitochondrial	*A. thaliana*	27.58	6.30	502	19.2	7.00
944	20 DAP es	2.05	Q9C5C8	Peptide methionine sulfoxide reductase B2, chloroplastic	*A. thaliana*	21.95	9.96	97	4.5	6.00
Environmental information processing										
1224	20 DAP es	4.23	Q9LX08	Annexin D6	*A. thaliana*	36.55	8.66	328	13.5	6.00
236	20 DAP es	2.20	AT1G30580.1	GTP binding	*A. thaliana*	44.44	6.38	382	15.7	4.00
944	20 DAP es	2.05	Q9FLP6	Small ubiquitin-related modifier 2	*A. thaliana*	11.65	5.23	408	27.2	3.00
1027	20 DAP es	1.59	P48347	14-3-3-like protein GF14 epsilon	*A. thaliana*	28.90	4.57	1679	52.0	24.00
1116	20 DAP es	1.67	P69310	Ubiquitin	*A. sativa*	8.52	7.58	510	77.6	8.00
1281	20 DAP es	1.72	Q9SVD7	Ubiquitin-conjugating enzyme E2 variant 1D	*A. thaliana*	16.52	6.24	919	65.1	12.00
1315	20 DAP es	1.52	P38548	GTP-binding nuclear protein Ran/TC4	*V. faba*	25.27	6.44	114	11.3	2.00
1224	20 DAP es	4.23	Q39336	Guanine nucleotide-binding protein subunit beta-like protein	*B. napus*	35.70	9.12	711	31.5	13.00
1224	20 DAP es	4.23	AT3G18130.1	receptor for activated C kinase 1C	*A. thaliana*	35.81	6.81	209	14.4	2.00
177	20 DAP es	2.37	AT1G52820.1	2-Oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein	*A. thaliana*	36.40	5.68	278	8.8	2.00
218	20 DAP es	1.53	ABD36807	Glutathione S-transferase	*B. napus*	24.72	5.78	376	45.6	5.00
315	20 DAP es	2.56	ABV89642	Universal stress protein family protein	*B. rapa*	17.58	9.02	260	42.1	4.00
315	20 DAP es	2.56	AT3G03270.1	Adenine nucleotide alpha hydrolases-like superfamily protein	*A. thaliana*	22.59	5.44	143	15.4	2.00
969	20 DAP es	2.40	ABD36807	Glutathione S-transferase	*B. napus*	24.72	5.78	104	9.7	2.00
969	20 DAP es	2.40	AT3G11930.1	Adenine nucleotide alpha hydrolases-like superfamily protein	*A. thaliana*	21.44	5.46	97	21.1	3.00
970	20 DAP es	2.72	ABD36807	Glutathione S-transferase	*B. napus*	24.72	5.78	192	18.4	4.00
970	20 DAP es	2.72	AT3G11930.1	Adenine nucleotide alpha hydrolases-like superfamily protein	*A. thaliana*	21.44	5.46	169	18.1	2.00
1157	20 DAP es	2.08	AT4G27585.1	PHB domain-containing membrane-associated protein family	*A. thaliana*	44.99	6.37	89	8.0	3.00
1162	20 DAP es	1.51	AT4G27585.1	PHB domain-containing membrane-associated protein family	*A. thaliana*	44.99	6.37	414	15.8	5.00

1283	20 DAP es	1.72	AT3G03270.1	Adenine nucleotide alpha hydrolases-like superfamily protein	*A. thaliana*	22.59	5.44	985	38.3	10.00
1283	20 DAP es	1.72	ABV89642	Universal stress protein family protein	*B. rapa*	17.58	9.02	804	75.5	14.00
Genetic information processing										
941	20 DAP es	1.67	P49311	G.-rich RNA-binding protein GRP2A	*S. alba*	16.35	5.41	682	37.3	6.00
154	20 DAP es	1.94	AT1G45000.1	AAA-type ATPase family protein	*A. thaliana*	44.73	9.02	67	6.0	2.00
199	20 DAP es	1.89	AT4G04460.1	Saposin-like aspartyl protease family protein	*A. thaliana*	55.54	6.93	183	7.5	3.00
602	20 DAP es	1.62	AT4G04460.1	Saposin-like aspartyl protease family protein	*A. thaliana*	55.54	6.93	231	7.1	3.00
1027	20 DAP es	1.59	Q9SGW3	26S proteasome non-ATPase regulatory subunit RPN12A	*A. thaliana*	30.69	4.67	521	32.2	7.00
1033	20 DAP es	3.06	Q9M4T8	Proteasome subunit alpha type-5	*G. max*	25.96	4.55	788	40.1	14.00
1186	20 DAP es	1.58	O22263	Protein disulfide-isomerase like 2-1	*A. thaliana*	39.47	5.73	1051	26.0	12.00
1422	20 DAP es	1.54	AT1G79930.1	Heat shock protein 91	*A. thaliana*	91.69	5.01	485	7.7	2.00
1224	20 DAP es	4.23	AT1G18080.1	Transducin/WD40 repeat-like superfamily protein	*A. thaliana*	35.73	8.79	565	23.2	11.00
44	20 DAP es	2.00	AT1G56070.1	Ribosomal protein S5/Elongation factor G/III/V family protein	*A. thaliana*	93.83	5.85	2221	33.2	34.00
463	20 DAP es	2.86	O22860	60S ribosomal protein L38	*A. thaliana*	8.12	10.64	344	58.0	4.00
668	20 DAP es	2.12	Q9SF40	60S ribosomal protein L4-1	*A. thaliana*	44.67	10.86	317	17.0	10.00
1157	20 DAP es	2.08	O04487	Probable elongation factor 1-gamma 1	*A. thaliana*	46.63	5.23	1476	22.5	11.00
1352	20 DAP es	1.71	AT2G18510.1	RNA-binding (RRM/RBD/RNP motifs) family protein (SP:putative spliceosome associated protein)	*A. thaliana*	39.86	7.94	274	9.9	3.00
595	20 DAP es	2.61	AT3G01340.1	Transducin/WD40 repeat-like superfamily protein	*A. thaliana*	32.61	5.61	579	16.6	6.00
882	20 DAP es	1.62	P41376	Eukaryotic initiation factor 4A-1	*A. thaliana*	46.67	5.36	1884	47.1	29.00
913	20 DAP es	1.81	P41376	Eukaryotic initiation factor 4A-1	*A. thaliana*	46.67	5.36	221	11.7	4.00
331	20 DAP es	3.08	AT2G36160.1	Ribosomal protein S11 family protein	*A. thaliana*	16.25	11.26	118	7.3	1.00
434	20 DAP es	3.20	AT2G43460.1	Ribosomal L38e protein family	*A. thaliana*	8.12	10.64	64	36.2	2.00
562	20 DAP es	1.66	P55852	Small ubiquitin-related modifier 1	*A. thaliana*	10.97	4.77	471	20.0	4.00
885	20 DAP es	1.66	AT1G09640.1	Translation elongation factor EF1B, gamma chain	*A. thaliana*	46.63	5.23	782	13.0	8.00
885	20 DAP es	1.66	Q9SAB3	Polyadenylate-binding protein RBP45B	*A. thaliana*	44.09	5.49	53	2.2	1.00
1224	20 DAP es	4.23	AAZ67604	80A08_19	*B. rapa*	36.41	6.88	534	40.3	17.00
Glycan biosynthesis and metabolism										
1401	20 DAP es	1.77	AT3G55260.1	Beta-hexosaminidase 1	*A. thaliana*	61.19	5.85	130	7.2	5.00
Lipid metabolism										
891	20 DAP es	2.07	P29108	Acyl-[acyl-carrier-protein] desaturase, chloroplastic	*B. napus*	45.32	5.69	3001	52.3	27.00
1147	20 DAP es	1.90	AT5G46290.1	3-Ketoacyl-acyl carrier protein synthase I	*A. thaliana*	50.38	9.25	896	29.8	16.00
Uncharacterized										
813	20 DAP es	2.67	XP_002887418	Hypothetical protein ARALYDRAFT_895064	*A. lyrata*	62.02	4.75	474	19.9	15.00
813	20 DAP es	2.67	A5BX00	Hypothetical protein V.V_028074	*V. vinifera*	123.63	5.57	62	2.2	2.00
1147	20 DAP es	1.90	XP_002875287	Hypothetical protein ARALYDRAFT_904762	*A. lyrata*	50.29	9.20	373	12.9	7.00
1401	20 DAP es	1.77	E4MX61	Unnamed protein product	*T. halophila*	53.21	6.48	925	39.3	22.00

[a] Spot number in accordance with Supplementary Fig. 3.
[b] Developmental stage of endosperm (es) 15 and 20 DAP.
[c] Regulation of spots, spot volume at least 1.5 fold increased in the corresponding developmental stage.
[d] Accession numbers as given by SwissProt (http://www.uniprot.org), NCBInr (http://www.ncbi.nlm.nih.gov/protein) and TAIR (http://www.Arabidopsis.org/).
[e] Calculated molecular masses of the identified proteins as deduced from the corresponding genes.
[f] Calculated isoelectric point of the identified proteins as deduced from the corresponding genes.
[g] Probability score for the protein identifications based on MS/MS analysis and MASCOT search.
[h] Number of unique matching peptides.
[i] Sequence coverage of a protein by identified peptides.

Table 4 – Identification of proteins from spots with changed abundance in the endosperm compared to the embryo at 15 DAP. To determine significant changes in spot volume, a Student's t-test (p-value ≤ 0.05) was applied on the basis of normalized relative spot volume. Changes in spot volume ≥ 1.5 were considered to represent true alterations. Identification of proteins was carried out using the MASCOT search algorithm (www.matrixscience.com) against the (i) SwissProt (www.uniprot.org), (ii) NCBInr (www.ncbi.nig.gov) and (iii) TAIR (www.arabidopsis.org, TAIR release 10) databases. Identified proteins were functionally classified according to the KEGG PATHWAY Database (http://www.genome.jp/kegg/pathway.html). The pathway categories were adjusted for seed metabolism by adding four functional groups: (i) storage, (ii) defense, (iii) desiccation and (iv) detoxification.

ID [a]	Stage [b]	Reg [c]	Accession [d]	Name	Organism	MM calc [e]	pI calc [f]	Score [g]	Pep [h]	SC% [i]
Amino acid metabolism										
34	15 DAP es	1.89	Q43314	Glutamate dehydrogenase 1	*A. thaliana*	44.5	6.42	116	4	9.00
23	15 DAP es	1.60	O04937	Glutamate dehydrogenase A	*N. plumbaginifolia*	44.8	6.69	60	1	2.43
266	15 DAP es	2.20	P30184	Leucine aminopeptidase 1	*A. thaliana*	54.5	5.58	271	4	8.85
384	15 DAP es	1.64	AT1G07780.1	phosphoribosylanthranilate isomerase 1	*A. thaliana*	29.6	9.52	68	1	3.27
Biosynthesis of secondary metabolites										
313	15 DAP es	1.81	AT5G54160.1	Quercetin 3-O-methyltransferase 1	*A. thaliana*	39.6	5.56	138	4	12.40
313	15 DAP es	1.81	Q9S818	Naringenin,2-oxoglutarate 3-dioxygenase	*A. thaliana*	40.3	5.17	104	3	8.10
62	15 DAP es	2.11	Q42553	Isopentenyl-diphosphate delta-isomerase II	*A. thaliana*	32.6	6.10	507	14	44.37
Carbohydrate metabolism										
301	15 DAP es	4.24	Q9C5I1	UDP-sugar pyrophosphorylase	*A. thaliana*	67.8	6.06	77	1	1.63
6	15 DAP es	5.38	AT5G52560.1	UDP-sugar pyrophosphorylase	*A. thaliana*	67.8	6.06	357	6	10.91
348	15 DAP es	2.02	AAY47048	Dehydroascorbate reductase	*S. lycopersicum*	23.5	6.36	309	5	23.81
74	15 DAP es	6.88	AAY47048	Dehydroascorbate reductase	*S. lycopersicum*	23.5	6.36	243	4	21.90
384	15 DAP es	1.64	CAA55209	L-Ascorbate peroxidase	*R. sativus*	27.7	5.41	788	18	62.80
266	15 DAP es	2.20	AT5G57655.2	Xylose isomerase family protein	*A. thaliana*	53.7	5.52	284	6	13.00
266	15 DAP es	2.20	Q9SEE5	Galactokinase	*A. thaliana*	54.3	5.61	150	2	4.64
350	15 DAP es	2.29	Q96533	Alcohol dehydrogenase class-3	*A. thaliana*	40.7	6.59	454	8	22.69
11	15 DAP es	1.57	Q9SU63	Aldehyde dehydrogenase family 2 member B4, mitochondrial	*A. thaliana*	58.6	7.78	1210	14	29.74
266	15 DAP es	2.20	P25696	Bifunctional enolase 2/transcriptional activator	*A. thaliana*	47.7	5.45	714	11	20.72
315	15 DAP es	1.75	AT4G38970.1	fructose-bisphosphate aldolase 2	*A. thaliana*	43.0	7.54	1696	15	41.71
314	15 DAP es	1.75	P16096	Fructose-bisphosphate aldolase, chloroplastic	*S. oleracea*	42.4	7.62	127	4	9.90
315	15 DAP es	1.75	AT5G57330.1	Galactose mutarotase-like superfamily protein	*A. thaliana*	35.4	5.66	308	7	27.56
315	15 DAP es	1.75	NP_200543	Glucose-6-phosphate 1-epimerase	*A. thaliana*	35.4	5.66	208	7	27.56
33	15 DAP es	2.60	P25857	Glyceraldehyde-3-phosphate dehydrogenase B, chloroplast	*A. thaliana*	47.6	6.36	372	6	14.32
260	15 DAP es	2.20	AT4G24620.1	Phosphoglucose isomerase 1	*A. thaliana*	67.0	5.37	86	3	5.22
314	15 DAP es	1.75	AT1G79550.1	Phosphoglycerate kinase	*A. thaliana*	42.1	5.37	2543	17	53.87
313	15 DAP es	1.81	AT1G79550.1	Phosphoglycerate kinase	*A. thaliana*	42.1	5.37	1619	8	50.12
169	15 DAP es	43.07	Q9SGC1	Probable phosphoglucomutase, cytoplasmic 2	*A. thaliana*	63.4	5.48	1202	17	26.50
301	15 DAP es	4.24	Q9SGC1	Probable phosphoglucomutase, cytoplasmic 2	*A. thaliana*	63.4	5.48	1372	15	24.96
6	15 DAP es	5.38	Q9SGC1	Probable phosphoglucomutase, cytoplasmic 2	*A. thaliana*	63.4	5.48	208	6	10.43
174	15 DAP es	2.94	AT3G52990.1	Pyruvate kinase family protein	*A. thaliana*	57.5	6.74	148	5	11.39
151	15 DAP es	2.80	AT5G63680.1	Pyruvate kinase family protein	*A. thaliana*	55.0	6.25	197	10	14.90
384	15 DAP es	1.64	P48491	Triosephosphate isomerase, cytosolic	*A. thaliana*	27.2	5.27	88	3	8.27
210	15 DAP es	1.59	P48491	Triosephosphate isomerase, cytosolic	*A. thaliana*	27.2	5.27	58	1	4.33
195	15 DAP es	5.43	AAP37972	Seed specific protein Bn15D33A	*B. napus*	12.7	9.24	338	10	31.90
290	15 DAP es	2.08	AT4G35830.1	Aconitase 1	*A. thaliana*	98.1	5.96	1323	31	32.18

166	15 DAP es	2.41	Q42560	Aconitate hydratase 1	*A. thaliana*	98.1	5.96	1110	29	30.29
164	15 DAP es	1.51	Q42560	Aconitate hydratase 1	*A. thaliana*	98.1	5.96	978	26	29.51
165	15 DAP es	3.04	XP_002870438	Aconitate hydratase 2, mitochondrial	*A. lyrata*	106.7	7.25	319	1	17.50
170	15 DAP es	1.90	Q9SIB9	Aconitate hydratase 2, mitochondrial	*A. thaliana*	108.1	6.79	1532	31	27.37
23	15 DAP es	1.60	AT1G65930.1	Cytosolic NADP+-dependent isocitrate dehydrogenase	*A. thaliana*	45.7	6.13	112	2	6.34
234	15 DAP es	2.17	Q43744	Malate dehydrogenase, mitochondrial	*B. napus*	35.7	9.48	91	2	10.26
313	15 DAP es	1.81	Q8RU27	Alpha-1,4-glucan-protein synthase UDP-forming 2	*S. tuberosum*	41.6	5.66	131	5	11.20
Cellular processes										
113	15 DAP es	2.47	Q9SVJ4	Endoglucanase 22	*A. thaliana*	55.0	6.36	58	1	1.62
23	15 DAP es	1.60	Q8H038	Xyloglucan galactosyltransferase KATAMARI1 homolog	*O. sativa*	66.7	5.54	51	1	1.02
220	15 DAP es	1.52	AT3G07680.1	GOLD family protein	*A. thaliana*	24.3	5.93	517	10	40.87
216	15 DAP es	1.63	AT4G20360.1	RAB GTPase homolog E1B	*A. thaliana*	51.6	5.79	57	1	2.31
151	15 DAP es	2.80	AT4G34490.1	cyclase associated protein 1	*A. thaliana*	50.9	6.24	619	8	15.13
94	15 DAP es	3.30	AT4G29350.1	Profilin 2	*A. thaliana*	14.0	4.77	301	4	16.79
95	15 DAP es	1.84	Q64LH1	Profilin-1	*A. artemisiifolia*	14.1	4.58	520	5	17.56
96	15 DAP es	1.84	Q64LH1	Profilin-1	*A. artemisiifolia*	14.1	4.58	439	3	16.03
345	15 DAP es	5.81	AT5G23540.1	Mov34/MPN/PAD-1 family protein	*A. thaliana*	34.3	6.34	236	4	12.99
Defense										
80	15 DAP es	1.79	AT3G46000.1	Actin depolymerizing factor 2	*A. thaliana*	15.7	5.10	151	1	8.76
186	15 DAP es	5.41	CAA47357	Allergen Car b I	*C. betulus*	17.2	5.80	53	1	14.47
114	15 DAP es	2.59	ABV89615	Bet v, allergen family protein	*B. rapa*	17.1	5.15	161	4	29.03
147	15 DAP es	6.03	AT1G52400.1	Beta glucosidase 18	*A. thaliana*	60.4	6.82	264	5	7.01
148	15 DAP es	6.03	AT1G52400.1	Beta glucosidase 18	*A. thaliana*	60.4	6.82	363	5	5.68
10	15 DAP es	1.57	AT3G03640.1	Beta glucosidase 25	*A. thaliana*	59.8	6.08	132	6	7.91
149	15 DAP es	3.67	Q9SE50	Beta-glucosidase 18	*A. thaliana*	60.4	6.82	310	7	8.71
146	15 DAP es	6.83	Q9SE50	Beta-glucosidase 18	*A. thaliana*	60.4	6.82	329	7	7.01
307	15 DAP es	9.90	Q9SE50	Beta-glucosidase 18	*A. thaliana*	60.4	6.82	201	5	5.68
144	15 DAP es	1.64	Q9LV33	Beta-glucosidase 44	*A. thaliana*	58.9	9.81	311	7	13.09
270	15 DAP es	3.98	AT3G16470.3	Mannose-binding lectin superfamily protein	*A. thaliana*	32.1	5.15	438	3	13.80
81	15 DAP es	5.30	Q9ZVF2	MLP-like protein 329	*A. thaliana*	17.6	5.21	383	4	26.49
270	15 DAP es	3.98	O04309	Myrosinase-binding protein	*A. thaliana*	48.5	4.99	457	3	9.09
23	15 DAP es	1.60	P52778	Protein LlR18A	*L. luteus*	16.8	5.07	54	1	7.69
6	15 DAP es	5.38	P52778	Protein LlR18A	*L. luteus*	16.8	5.07	55	1	7.69
18	15 DAP es	16.86	P52778	Protein LlR18A	*L. luteus*	16.8	5.07	55	1	7.69
Detoxification										
91	15 DAP es	3.10	AAD05576	Cu/Zn superoxide dismutase	*R. sativus*	15.1	5.41	508	6	34.21
Energy metabolism										
182	15 DAP es	6.09	AT5G38430.1	Ribulose bisphosphate carboxylase (small chain) family protein181	*A. thaliana*	20.3	8.86	723	7	31.49
10	15 DAP es	1.57	P48686	Ribulose bisphosphate carboxylase large chain	*B. oleracea*	52.9	5.85	890	23	42.80
11	15 DAP es	1.57	P48686	Ribulose bisphosphate carboxylase large chain	*B. oleracea*	52.9	5.85	611	13	30.27
247	15 DAP es	1.92	P28380	Ribulose bisphosphate carboxylase large chain (Fragment)	*A. graveolens*	50.8	6.19	57	2	3.49
197	15 DAP es	3.52	P10796	Ribulose bisphosphate carboxylase small chain 1B, chloroplastic	*A. thaliana*	20.3	8.86	447	5	17.13
197	15 DAP es	3.52	Q9ZST4	Nitrogen regulatory protein P-II homolog	*A. thaliana*	21.3	9.93	77	1	5.61
10	15 DAP es	1.57	AT2G07698.1	ATP synthase subunit alpha	*A. thaliana*	85.9	5.31	59	2	2.83

(continued on next page)

Table 4 (*continued*)

ID [a]	Stage [b]	Reg [c]	Accession [d]	Name	Organism	MM calc [e]	pI calc [f]	Score [g]	Pep [h]	SC% [i]
Energy metabolism										
345	15 DAP es	5.81	AT1G50940.1	Electron transfer flavoprotein alpha	*A. thaliana*	38.4	6.56	59	1	5.23
208	15 DAP es	3.16	Q9SJ12	Probable ATP synthase 24 kDa subunit, mitochondrial	*A. thaliana*	27.6	6.30	359	7	22.92
208	15 DAP es	3.16	AT2G18230.1	Pyrophosphorylase 2	*A. thaliana*	24.7	5.68	71	2	13.30
384	15 DAP es	1.64	Q9XF89	Chlorophyll a–b binding protein CP26, chloroplastic	*A. thaliana*	30.1	5.98	55	2	7.14
136	15 DAP es	3.42	XP_002502857	Early light induced protein-like 5, chloroplast	*M. sp. RCC299*	27.8	4.32	70	2	3.72
23	15 DAP es	1.60	Q9LIK9	ATP sulfurylase 1, chloroplastic	*A. thaliana*	51.4	6.36	721	13	21.60
149	15 DAP es	3.67	AT5G04590.1	Sulfite reductase	*A. thaliana*	71.9	9.17	365	13	18.85
146	15 DAP es	6.83	AT5G04590.1	Sulfite reductase	*A. thaliana*	71.9	9.17	139	5	7.63
Environmental information processing										
363	15 DAP es	5.11	Q9XEE2	Annexin D2	*A. thaliana*	36.2	5.71	2585	18	40.06
345	15 DAP es	5.81	Q9XEE2	Annexin D2	*A. thaliana*	36.2	5.71	1276	11	29.65
96	15 DAP es	1.84	XP_002880102	Calcium-binding EF hand family protein	*A. lyrata*	15.8	4.29	135	2	15.49
169	15 DAP es	43.07	AT2G46370.1	Auxin-responsive GH3 family protein	*A. thaliana*	64.3	5.38	346	8	16.35
169	15 DAP es	43.07	Q9SKE2	Jasmonic acid-amido synthetase JAR1	*A. thaliana*	64.3	5.38	323	8	16.35
216	15 DAP es	1.63	AT3G46810.1	Cysteine/Histidine-rich C1 domain family protein	*A. thaliana*	78.9	8.44	173	2	1.02
186	15 DAP es	5.41	P69310	Ubiquitin	*A. sativa*	8.5	7.58	510	8	77.63
188	15 DAP es	2.05	P69310	Ubiquitin	*A. sativa*	8.5	7.58	428	8	77.63
189	15 DAP es	2.05	P69310	Ubiquitin	*A. sativa*	8.5	7.58	333	6	69.74
79	15 DAP es	2.46	Q9SVD7	Ubiquitin-conjugating enzyme E2	*A. thaliana*	16.5	6.24	825	11	54.11
83	15 DAP es	4.16	AT3G03270.1	Adenine nucleotide alpha hydrolases-like superfamily protein	*A. thaliana*	22.6	5.44	143	2	15.42
196	15 DAP es	5.64	AT3G03270.1	Adenine nucleotide alpha hydrolases-like superfamily protein	*A. thaliana*	22.6	5.44	985	10	38.31
231	15 DAP es	1.89	AT3G17020.1	Adenine nucleotide alpha hydrolases-like superfamily protein	*A. thaliana*	17.8	6.52	65	1	13.50
71	15 DAP es	2.60	AT3G11930.1	Adenine nucleotide alpha hydrolases-like superfamily protein	*A. thaliana*	21.4	5.46	97	3	21.11
80	15 DAP es	1.79	AT4G39260.1	Cold, circadian rhythm, and RNA binding 1	*A. thaliana*	16.6	5.44	443	5	14.79
85	15 DAP es	1.77	ADR01108	Copper/zinc superoxide dismutase	*B. rapa*	19.0	6.32	69	1	10.81
78	15 DAP es	4.42	AAK01359	Dehydration stress-induced protein	*B. napus*	19.7	4.67	315	6	30.90
202	15 DAP es	1.86	AT5G20500.1	Glutaredoxin family protein	*A. thaliana*	14.8	5.62	130	2	11.85
195	15 DAP es	5.43	Q9FNE2	Glutaredoxin-C2	*A. thaliana*	11.7	7.68	209	2	9.91
71	15 DAP es	2.60	ABD36807	Glutathione S-transferase	*B. napus*	24.7	5.78	104	2	9.68
348	15 DAP es	2.02	Q9FRL8	Glutathione S-transferase DHAR2	*A. thaliana*	23.4	5.75	127	2	7.04
74	15 DAP es	6.88	Q9FRL8	Glutathione S-transferase DHAR2	*A. thaliana*	23.4	5.75	177	3	10.33
384	15 DAP es	1.64	AT1G10370.1	Glutathione S-transferase family protein	*A. thaliana*	25.3	6.23	80	2	6.61
204	15 DAP es	3.16	AT1G10370.1	Glutathione S-transferase family protein	*A. thaliana*	25.3	6.23	293	6	16.74
208	15 DAP es	3.16	AT1G10370.1	Glutathione S-transferase family protein	*A. thaliana*	25.3	6.23	116	4	12.33
205	15 DAP es	2.60	AT1G10370.1	Glutathione S-transferase family protein	*A. thaliana*	25.3	6.23	259	7	19.38
228	15 DAP es	1.77	AT2G30860.1	Glutathione S-transferase PHI 9	*A. thaliana*	24.1	6.19	556	7	37.67
10	15 DAP es	1.57	AT3G24170.1	Glutathione-disulfide reductase	*A. thaliana*	53.8	6.39	163	7	15.63
180	15 DAP es	1.87	Q9SK52	Peroxidase 18	*A. thaliana*	35.6	5.08	63	1	2.74
350	15 DAP es	2.29	ACR40091	S-Nitrosoglutathione reductase	*B. juncea*	40.1	8.72	185	1	24.33
124	15 DAP es	1.51	CAD31719	Translationally controlled tumor-like protein Cicer arietinum	*C. arietinum*	7.0	6.08	117	1	26.23
120	15 DAP es	1.51	Q944W6	Translationally-controlled tumor protein homolog	*B. oleracea*	19.0	4.48	556	8	39.29
83	15 DAP es	4.16	ABV89642	Universal stress protein family protein	*B. rapa*	17.6	9.02	260	4	42.14
71	15 DAP es	2.60	XP_002882756	Universal stress protein family protein	*A. lyrata*	21.7	5.28	53	3	25.50
196	15 DAP es	5.64	ABV89642	Universal stress protein family protein	*B. rapa*	17.6	9.02	804	14	75.47

Genetic information processing										
270	15 DAP es	3.98	Q940A6	Pentatricopeptide repeat-containing protein	*A. thaliana*	94.6	9.49	51	1	0.72
345	15 DAP es	5.81	Q9LT08	26S proteasome non-ATPase regulatory subunit 14	*A. thaliana*	34.3	6.34	245	4	12.99
210	15 DAP es	1.59	Q8H0X6	Cysteine proteinase inhibitor 6	*A. thaliana*	26.3	5.84	130	3	11.11
246	15 DAP es	7.23	AT1G60420.1	DC1 domain-containing protein	*A. thaliana*	65.1	4.75	292	5	9.69
197	15 DAP es	3.52	Q38935	FK506-binding protein 2-1	*A. thaliana*	16.3	9.29	103	3	23.53
197	15 DAP es	3.52	AAC49390	Immunophilin	*A. thaliana*	15.7	9.29	85	3	24.66
184	15 DAP es	3.01	XP_002522624	Immunophilin, putative	*R. communis*	12.0	9.17	499	4	14.29
23	15 DAP es	1.60	AT1G29150.1	Non-ATPase subunit 9	*A. thaliana*	46.7	6.25	1215	21	39.14
108	15 DAP es	3.45	AT4G31300.1	N-terminal nucleophile aminohydrolases superfamily protein	*A. thaliana*	25.1	5.21	106	2	10.73
216	15 DAP es	1.63	AT3G60820.1	N-terminal nucleophile aminohydrolases superfamily protein	*A. thaliana*	24.6	7.71	518	7	24.22
246	15 DAP es	7.23	AT1G77510.1	PDI-like 1-2	*A. thaliana*	56.3	4.76	693	8	12.20
247	15 DAP es	1.92	AT1G77510.1	PDI-like 1-2	*A. thaliana*	56.3	4.76	662	7	15.16
255	15 DAP es	2.20	AT1G77510.1	PDI-like 1-2	*A. thaliana*	56.3	4.76	771	5	11.61
143	15 DAP es	3.14	P34790	Peptidyl-prolyl cis-trans isomerase	*A. thaliana*	18.4	9.01	587	7	46.51
246	15 DAP es	7.23	O80763	Probable nucleoredoxin 1	*A. thaliana*	65.1	4.75	310	5	9.69
372	15 DAP es	1.57	O23708	Proteasome subunit alpha type-2-A	*A. thaliana*	25.7	5.40	431	5	31.91
208	15 DAP es	3.16	O23715	Proteasome subunit alpha type-3	*A. thaliana*	27.4	5.90	307	6	24.10
210	15 DAP es	1.59	O23715	Proteasome subunit alpha type-3	*A. thaliana*	27.4	5.90	658	10	44.98
72	15 DAP es	2.67	Q7DLR9	Proteasome subunit beta type-4	*A. thaliana*	27.6	6.10	1636	12	53.66
245	15 DAP es	4.84	ABB17025	Protein disulfide isomerase	*B. carinata*	55.7	4.86	383	12	18.66
260	15 DAP es	2.20	ABB17025	Protein disulfide isomerase	*B. carinata*	55.7	4.86	2150	28	55.80
314	15 DAP es	1.75	O22263	Protein disulfide-isomerase like 2-1	*A. thaliana*	39.5	5.73	857	11	26.32
314	15 DAP es	1.75	AAC36164	Putative serpin	*A. thaliana*	23.2	4.30	52	1	4.23
135	15 DAP es	5.22	AT4G34870.1	Rotamase cyclophilin 5 RD	*A. thaliana*	18.4	9.84	106	1	8.72
174	15 DAP es	2.94	AT3G18190.1	TCP-1/cpn60 chaperonin family protein	*A. thaliana*	57.7	8.68	86	3	6.34
314	15 DAP es	1.75	AT2G47470.1	Thioredoxin family protein	*A. thaliana*	39.5	5.73	747	11	26.32
96	15 DAP es	1.84	Q9XGS0	Thioredoxin M-type, chloroplastic	*B. napus*	19.3	10.27	142	1	9.04
247	15 DAP es	1.92	Q9LT79	U-box domain-containing protein 25	*A. thaliana*	46.0	7.86	59	1	1.66
184	15 DAP es	3.01	Q9LT79	U-box domain-containing protein 25	*A. thaliana*	46.0	7.86	56	1	1.66
132	15 DAP es	9.56	AAF78493	Contains similarity to AP2/EREBP-like transcription factor	*A. thaliana*	47.5	4.80	66	2	6.33
80	15 DAP es	1.79	Q05966	G.-rich RNA-binding protein 10	*B. napus*	16.3	5.42	523	4	19.53
210	15 DAP es	1.59	AAF31402	Putative G.-rich RNA binding protein 1	*C. roseus*	14.2	9.61	52	1	10.95
245	15 DAP es	4.84	AAF31402	Putative G.-rich RNA binding protein 1	*C. roseus*	14.2	9.61	57	1	10.95
114	15 DAP es	2.59	AAF31402	Putative G.-rich RNA binding protein 1	*C. roseus*	14.2	9.61	72	1	7.30
132	15 DAP es	9.56	AT2G02090.1	SNF2 domain-containing protein	*A. thaliana*	86.2	5.42	55	1	0.92
71	15 DAP es	2.60	AT5G28640.1	SSXT family protein	*A. thaliana*	22.4	5.78	56	1	6.67
90	15 DAP es	1.81	A6YGA0	30S ribosomal protein S18, chloroplastic	*Leptosira terrestris*	11.6	11.68	73	1	8.00
95	15 DAP es	1.84	Q9LH85	60S acidic ribosomal protein	*A. thaliana*	11.7	4.40	57	1	7.83
96	15 DAP es	1.84	Q9LH85	60S acidic ribosomal protein P2-3	*A. thaliana*	11.7	4.40	64	1	7.83
216	15 DAP es	1.63	P17745	Elongation factor Tu, chloroplastic	*A. thaliana*	51.6	5.79	58	1	2.31
208	15 DAP es	3.16	O04663	Eukaryotic translation initiation factor 4E-2	*A. thaliana*	22.5	5.39	118	4	13.64
174	15 DAP es	2.94	Q9T034	Probable phenylalanyl-tRNA synthetase alpha chain	*A. thaliana*	55.8	9.02	496	11	15.67
Glycan biosynthesis and metabolism										
202	15 DAP es	1.86	CAJ15148	Sialyltransferase-like protein	*L. japonicus*	49.4	9.77	66	1	1.58

(continued on next page)

Table 4 (*continued*)

ID [a]	Stage [b]	Reg [c]	Accession [d]	Name	Organism	MM calc [e]	pI calc [f]	Score [g]	Pep [h]	SC% [i]
Lipid metabolism										
315	15 DAP es	1.75	Q01771	Acyl-acyl-carrier-protein desaturase, seed specific, chloroplastic	*B. napus*	45.3	6.04	292	8	22.81
266	15 DAP es	2.20	AAK60339	biotin carboxylase	*B. napus*	58.4	6.57	51	3	5.23
314	15 DAP es	1.75	CAA52786	Stearoyl-acyl carrier protein desaturase	*B. napus*	45.5	6.04	56	1	3.24
78	15 DAP es	4.42	AT4G39730.1	Lipase/lipooxygenase family protein	*A. thaliana*	20.1	4.83	134	2	8.84
Metabolism of cofactors and vitamins										
315	15 DAP es	1.75	Q9LR75	Coproporphyrinogen-III oxidase, chloroplastic	*A. thaliana*	43.8	6.25	478	12	28.76
Nucleotide metabolism										
132	15 DAP es	9.56	Q52K88	Nudix hydrolase 13, mitochondrial	*A. thaliana*	23.2	4.51	190	4	14.36
90	15 DAP es	1.81	AT3G56490.1	HIS triad family protein 3 (SP: protein kinase C inhibitor-like protein)	*A. thaliana*	16.0	6.79	294	5	50.34
180	15 DAP es	1.87	AT4G23895.3	Nucleoside diphosphate kinase	*A. thaliana*	52.0	5.84	507	2	12.85
180	15 DAP es	1.87	Q8RXA8	Nucleoside diphosphate kinase 4, chloroplastic	*S.oleracea*	25.7	9.71	643	8	28.09
Uncharacterized										
90	15 DAP es	1.81	P42855	14 kDa zinc-binding protein (Fragment)	*B. juncea*	12.6	6.62	119	4	54.87
94	15 DAP es	3.30	AT1G42960.1	Expressed protein localized to the inner membrane of the chloroplast	*A. thaliana*	17.8	9.42	64	2	13.69
74	15 DAP es	6.88	AAF79440	F18O14.33	*A. thaliana*	50.1	6.92	149	1	2.50
197	15 DAP es	3.52	AT4G01900.1	GLNB1 homolog	*A. thaliana*	21.3	9.93	73	1	5.61
199	15 DAP es	9.92	AT2G39050.1	Hydroxyproline-rich glycoprotein family protein	*A. thaliana*	35.6	6.21	137	3	12.30
135	15 DAP es	5.22	EEC81610	Hypothetical protein OsI_25113	*O. sativa*	69.9	11.07	77	1	2.37
81	15 DAP es	5.30	AT1G61600.1	Protein of unknown function	*A. thaliana*	48.7	10.31	51	1	1.90
94	15 DAP es	3.30	AT1G61600.1	Protein of unknown function (DUF1262)	*A. thaliana*	48.7	10.31	50	1	1.90
136	15 DAP es	3.42	AT2G24620.1	S-Locus glycoprotein family protein	*A. thaliana*	18.2	9.61	53	1	5.10
62	15 DAP es	2.11	AT4G14930.1	Survival protein SurE-like phosphatase/nucleotidase	*A. thaliana*	34.1	4.91	146	1	3.49
91	15 DAP es	3.10	Q9ZUX4	Uncharacterized protein At2g27730, mitochondrial	*A. thaliana*	11.9	10.10	91	1	7.08
18	15 DAP es	16.86	BAA99394	Vacuolar calcium binding protein	*R. sativus*	27.1	3.95	245	4	12.90
Xenobiotics biodegradation and metabolism										
378	15 DAP es	1.57	AT2G32520.1	Carboxymethylenebutenolidase	*A. thaliana*	25.9	5.14	114	3	11.30
Amino acid metabolism										
309	15 DAP em	2.24	ABD65618	Acetylornithine deacetylase, putative	*B. oleracea*	44.79	5.10	79	6.0	17.20
318	15 DAP em	1.95	AT4G01850.1	S-Adenosylmethionine synthetase 2	*A. thaliana*	43.23	5.63	246	7.0	27.48
319	15 DAP em	2.04	AT3G14990.1	Glutamine amidotransferase-like superfamily protein	*A. thaliana*	41.83	5.18	90	1.0	2.81
309	15 DAP em	2.24	AT4G33680.1	Pyridoxal phosphate (PLP)-dependent transferases superfamily protein (SP: Serine hydroxy methyltransferase)	*A. thaliana*	50.36	7.71	880	10.0	22.56
150	15 DAP em	1.51	Q9SZJ5	Serine hydroxymethyltransferase, mitochondrial	*A. thaliana*	57.36	8.80	776	17	30.75
319	15 DAP em	2.04	P29102	3-Isopropylmalate dehydrogenase, chloroplastic	*B. napus*	43.32	6.05	916	15	38.92
309	15 DAP em	2.24	Q93ZN9	LL-Diaminopimelate aminotransferase, chloroplastic	*A. thaliana*	50.36	7.71	891	10	22.56
8	15 DAP em	1.55	Q0WM29	Methylmalonate-semialdehyde dehydrogenase acylating, mitochondrial	*A. thaliana*	65.88	9.66	65	1	1.81
Biosynthesis of secondary metabolites										
318	15 DAP em	1.95	AAK68820	Similar to dihydroflavonol reductase	*A. thaliana*	43.67	5.51	60	2	5.40
319	15 DAP em	2.04	AAP96742	ThiJ-like protein	*B. rapa*	41.66	5.24	224	3	9.18

Carbohydrate metabolism										
318	15 DAP em	1.95	AT1G08200.1	UDP-D-apiose/UDP-D-xylose synthase 2	*A. thaliana*	43.76	5.51	81	2	5.40
318	15 DAP em	1.95	AAK72107	Monodehydroascorbate reductase	*B. rapa*	46.43	5.74	108	9	29.95
322	15 DAP em	2.11	P52417	Glucose-1-phosphate adenylyltransferase small subunit 2, chloroplastic	*V. faba*	56.02	6.20	226	7	15.04
13	15 DAP em	2.13	AT2G24270.1	Aldehyde dehydrogenase 11A3	*A. thaliana*	53.03	6.24	976	17	29.23
352	15 DAP em	2.37	AT4G16155.1	Dihydrolipoyl dehydrogenases	*A. thaliana*	67.05	8.66	404	9	16.83
316	15 DAP em	2.26	AT2G21330.1	Fructose-bisphosphate aldolase 1	*A. thaliana*	42.90	6.18	162	4	10.03
13	15 DAP em	2.13	AT3G16950.1	Lipoamide dehydrogenase 1	*A. thaliana*	60.72	9.00	87	2	4.56
352	15 DAP em	2.37	AT3G16950.1	Lipoamide dehydrogenase 1	*A. thaliana*	60.72	9.00	393	3	19.12
13	15 DAP em	2.13	Q1WIQ6	NADP-dependent glyceraldehyde-3-phosphate dehydrogenase	*A. thaliana*	53.03	6.24	1024	16	25.60
162	15 DAP em	1.58	AT1G76550.1	Phosphofructokinase family protein	*A. thaliana*	67.52	6.92	72	4	7.29
259	15 DAP em	1.75	AT4G24620.1	Phosphoglucose isomerase 1	*A. thaliana*	67.01	5.37	247	7	12.56
352	15 DAP em	2.37	AT3G02360.1	6-Phosphogluconate dehydrogenase family protein	*A. thaliana*	53.54	7.67	751	17	32.30
241	15 DAP em	2.34	AT3G60750.1	Transketolase	*A. thaliana*	79.92	5.92	2075	20	24.83
322	15 DAP em	2.11	AT5G48300.1	ADP glucose pyrophosphorylase 1	*A. thaliana*	56.62	6.13	177	6	12.31
17	15 DAP em	2.95	AT1G65930.1	Cytosolic NADP+-dependent isocitrate dehydrogenase	*A. thaliana*	45.72	6.13	196	5	14.15
Cellular processes										
41	15 DAP em	2.62	AT5G16510.1	Alpha-1,4-glucan-protein synthase family protein	*A. thaliana*	38.56	4.92	179	2	5.75
274	15 DAP em	1.56	Q9FM01	Probable UDP-glucose 6-dehydrogenase 2	*A. thaliana*	53.06	5.51	354	8	19.38
161	15 DAP em	4.92	Q8H038	Xyloglucan galactosyltransferase KATAMARI1 homolog	*O. sativa*	66.67	5.54	53	1	1.02
274	15 DAP em	1.56	P29511	Tubulin alpha-6 chain	*A. thaliana*	49.51	4.79	2390	27	57.56
271	15 DAP em	1.56	P29511	Tubulin alpha-6 chain	*A. thaliana*	49.51	4.79	1230	19	50.44
50	15 DAP em	1.87	AT3G48890.1	Membrane-associated progesterone binding protein 3	*A. thaliana*	25.37	4.38	52	1	4.29
177	15 DAP em	1.59	Q9STE8	Protein TOC75-3, chloroplastic	*A. thaliana*	89.13	9.55	1985	20	22.13
310	15 DAP em	3.61	AT5G09810.1	Actin 7	*A. thaliana*	41.71	5.20	147	6	15.92
319	15 DAP em	2.04	Q96292	Actin-2	*A. thaliana*	41.85	5.27	1940	26	54.91
19	15 DAP em	1.58	NP_180328	Vacuolar protein-sorting-associated protein 4	*A. thaliana*	48.56	6.55	55	1	5.06
Defense										
13	15 DAP em	2.13	XP_002876291	Glycosyl hydrolase family 20 protein	*A. lyrata*	61.54	5.85	90	6	9.38
161	15 DAP em	4.92	P52778	Protein LlR18A	*L. luteus*	16.85	5.07	51	1	7.69
19	15 DAP em	1.58	P52778	Protein LlR18A	*L. luteus*	16.85	5.07	57	1	7.69
380	15 DAP em	2.57	XP_002887781	Metacaspase 7	*A. lyrata*	45.48	4.61	80	1	2.87
Detoxification										
218	15 DAP em	1.81	AAC15842	Superoxide dismutase	*R. sativus*	23.79	5.97	187	6	20.28
Energy metabolism										
8	15 DAP em	1.55	P92549	ATP synthase subunit alpha, mitochondrial	*A. thaliana*	55.01	6.22	1196	24	28.40
380	15 DAP em	2.57	P11594	Oxygen-evolving enhancer protein 2, chloroplastic	*S. alba*	27.91	7.60	610	7	25.77
198	15 DAP em	3.11	P49108	Photosystem II 10 kDa polypeptide, chloroplastic	*B. campestris*	14.64	9.27	355	4	31.91
75	15 DAP em	1.52	XP_002878030	photosystem II reaction center PsbP family protein	*A. lyrata*	25.55	9.90	215	6	22.27
162	15 DAP em	1.58	AT5G04590.1	Sulfite reductase	*A. thaliana*	71.90	9.17	454	12	17.60
Environmental information processing										
346	15 DAP em	1.60	Q9LX08	Annexin D6	*A. thaliana*	36.55	8.66	548	11	26.42
330	15 DAP em	1.73	AT1G30580.1	GTP binding protein	*A. thaliana*	44.44	6.38	800	10	32.23
346	15 DAP em	1.60	Q39336	Guanine nucleotide-binding protein subunit beta-like protein	*B. napus*	35.70	9.12	750	14	40.98

(continued on next page)

Table 4 (*continued*)

ID [a]	Stage [b]	Reg [c]	Accession [d]	Name	Organism	MM calc [e]	pI calc [f]	Score [g]	Pep [h]	SC% [i]
Environmental information processing										
104	15 DAP em	3.40	AT1G23260.1	MMS ZWEI homolog 1	*A. thaliana*	17.84	5.04	89	4	20.25
104	15 DAP em	3.40	Q93YP0	Ubiquitin-conjugating enzyme E2 variant 1A	*A. thaliana*	17.84	5.04	111	4	20.25
75	15 DAP em	1.52	AT2G47710.1	Adenine nucleotide alpha hydrolases-like superfamily protein	*A. thaliana*	17.29	8.93	77	3	32.10
150	15 DAP em	1.51	P25890	Catalase	*P. sativum*	57.31	6.78	52	3	7.49
225	15 DAP em	1.57	CAA07494	Heat stress-induced protein	*B. oleracea*	23.47	9.23	52	2	15.14
318	15 DAP em	1.95	AT4G27585.1	PHB domain-containing membrane-associated protein family	*A. thaliana*	44.99	6.37	89	3	8.03
300	15 DAP em	2.43	AT1G62740.1	Stress-inducible protein, putative571	*A. thaliana*	64.48	5.80	445	14	16.46
Genetic information processing										
150	15 DAP em	1.51	Q39230	Seryl-tRNA synthetase	*A. thaliana*	51.60	6.27	982	14	24.39
380	15 DAP em	2.57	O65282	20 kDa chaperonin, chloroplastic	*A. thaliana*	26.79	9.35	199	5	25.30
19	15 DAP em	1.58	Q9SSB5	26S protease regulatory subunit 7 homolog A	*A. thaliana*	47.77	6.36	2483	34	57.51
282	15 DAP em	1.81	AT2G04030.1	Chaperone protein htpG family protein	*A. thaliana*	88.61	4.78	974	18	22.95
5	15 DAP em	3.76	AT5G50920.1	CLPC homolog 1	*A. thaliana*	103.39	6.35	3615	52	49.73
5	15 DAP em	3.76	P42762	ERD1 protein, chloroplastic	*A. thaliana*	103.17	5.84	164	1	2.12
300	15 DAP em	2.43	Q43468	Heat shock protein STI	*G. max*	63.55	5.76	373	5	4.04
221	15 DAP em	3.60	P34791	Peptidyl-prolyl cis-trans isomerase CYP20-3, chloroplastic	*A. thaliana*	28.19	9.68	196	5	22.69
225	15 DAP em	1.57	P34791	Peptidyl-prolyl cis-trans isomerase CYP20-3, chloroplastic	*A. thaliana*	28.19	9.68	417	6	16.92
322	15 DAP em	2.11	Q9ZU25	Probable mitochondrial-processing peptidase subunit alpha-1	*A. thaliana*	54.37	5.91	905	13	16.90
259	15 DAP em	1.75	ABB17025	Protein disulfide isomerase	*B. carinata*	55.71	4.86	2355	34	63.65
177	15 DAP em	1.59	XP_002520530	Sorting and assembly machinery (sam50) protein, putative	*R. communis*	88.97	9.55	874	4	14.60
309	15 DAP em	2.24	AAN74635	DEAD box RNA helicase	*P. sativum*	46.85	5.27	917	2	38.01
324	15 DAP em	1.82	AT5G11170.1	DEAD/DEAH box RNA helicase family protein	*A. thaliana*	48.31	5.33	748	18	31.62
322	15 DAP em	2.11	Q56XG6	DEAD-box ATP-dependent RNA helicase 15	*A. thaliana*	48.31	5.33	681	15	31.15
366	15 DAP em	1.93	Q9FLU1	DNA-binding protein BIN4	*A. thaliana*	49.45	5.02	66	1	1.98
218	15 DAP em	1.81	ACG26589	G.-rich protein 2b	*Z. mays*	20.29	5.91	95	1	7.69
117	15 DAP em	1.69	Q03250	G.-rich RNA-binding protein 7	*A. thaliana*	16.88	5.81	231	2	16.48
157	15 DAP em	2.29	P59259	Histone H4	*A. thaliana*	11.40	11.96	111	1	9.71
50	15 DAP em	1.87	Q9FL92	Probable WRKY transcription factor 16	*A. thaliana*	155.60	5.92	54	1	0.80
300	15 DAP em	2.43	Q9FL92	Probable WRKY transcription factor 16	*A. thaliana*	155.60	5.92	56	1	0.80
274	15 DAP em	1.56	Q9FL92	Probable WRKY transcription factor 16	*A. thaliana*	155.60	5.92	67	1	0.80
203	15 DAP em	1.71	Q3E902	40S ribosomal protein S21-2	*A. thaliana*	9.46	9.90	549	2	17.65
318	15 DAP em	1.95	AT5G36230.1	ARM repeat superfamily protein	*A. thaliana*	47.19	5.38	63	2	5.60
50	15 DAP em	1.87	P48006	Elongation factor 1-delta 1	*A. thaliana*	25.12	4.28	966	11	35.50
319	15 DAP em	2.04	Q43467	Elongation factor Tu, chloroplastic	*G. max*	52.06	6.22	308	5	13.36
310	15 DAP em	3.61	P17745	Elongation factor Tu, chloroplastic	*A. thaliana*	51.60	5.79	1807	18	34.03
104	15 DAP em	3.40	AT1G69410.1	Eukaryotic elongation factor 5A-3	*A. thaliana*	17.20	5.51	131	1	6.33
309	15 DAP em	2.24	P41376	Eukaryotic initiation factor 4A-1	*A. thaliana*	46.67	5.36	1884	29	47.09

JOURNAL OF PROTEOMICS 108 (2014) 382–426

241	15 DAP em	2.34	AT5G44320.1	Eukaryotic translation initiation factor 3 subunit 7	*A. thaliana*	66.18	5.32	145	4	8.16
157	15 DAP em	2.29	AT5G26710.1	Glutamyl/glutaminyl-tRNA synthetase	*A. thaliana*	81.01	6.68	320	11	11.27
156	15 DAP em	3.71	AT5G26710.1	Glutamyl/glutaminyl-tRNA synthetase	*A. thaliana*	81.01	6.68	360	10	10.71
50	15 DAP em	1.87	AT1G74230.1	Glycine-rich RNA-binding protein 5	*A. thaliana*	28.71	4.45	86	1	2.77
298	15 DAP em	68.72	O23627	Glycyl-tRNA synthetase 1, mitochondrial	*A. thaliana*	81.89	6.63	972	24	24.69
161	15 DAP em	4.92	AT2G23350.1	Poly(A) binding protein 4	*A. thaliana*	71.61	6.44	459	9	14.80
162	15 DAP em	1.58	P42731	Polyadenylate-binding protein 2	*A. thaliana*	68.63	8.82	55	5	7.95
318	15 DAP em	1.95	O04487	Probable elongation factor 1-gamma 1	*A. thaliana*	46.63	5.23	1476	11	22.46
291	15 DAP em	2.11	AT1G56070.1	Ribosomal protein S5/Elongation factor G/III/V family protein	*A. thaliana*	93.83	5.85	2348	38	37.60
292	15 DAP em	2.37	AT1G56070.1	Ribosomal protein S5/Elongation factor G/III/V family protein	*A. thaliana*	93.83	5.85	2221	34	33.21
346	15 DAP em	1.60	AT1G18080.1	Transducin/WD40 repeat-like superfamily protein	*A. thaliana*	35.73	8.79	631	12	38.53
104	15 DAP em	3.40	AT3G21000.1	Gag-Pol-related retrotransposon family protein	*A. thaliana*	46.29	5.89	55	1	2.47
Glycan biosynthesis										
13	15 DAP em	2.13	AT3G55260.1	Beta-hexosaminidase 1	*A. thaliana*	61.19	5.85	130	5	7.21
Lipid metabolism										
17	15 DAP em	2.95	AT5G46290.1	3-Ketoacyl-acyl carrier protein synthase I	*A. thaliana*	50.38	9.25	896	16	29.81
104	15 DAP em	3.40	AAZ66933	117M18_14	*B. rapa*	24.79	9.58	238	7	29.02
19	15 DAP em	1.58	P52410	3-Oxoacyl-acyl-carrier-protein synthase I, chloroplastic	*A. thaliana*	50.38	9.25	263	4	8.88
316	15 DAP em	2.26	P29108	Acyl-acyl-carrier-protein desaturase, chloroplastic	*B. napus*	45.32	5.69	2497	25	56.78
271	15 DAP em	1.56	AAK60339	Biotin carboxylase	*B. napus*	58.38	6.57	1013	21	42.62
274	15 DAP em	1.56	AAK60339	Biotin carboxylase	*B. napus*	58.38	6.57	72	7	15.89
104	15 DAP em	3.40	AT5G10160.1	3R-Hydroxymyristoyl ACP dehydrase	*A. thaliana*	24.11	9.99	123	3	12.33
Nucleotide metabolism										
352	15 DAP em	2.37	ABE87611	Phosphoribosyltransferase; Orotidine 5-phosphate decarboxylase	*M. truncatula*	51.24	7.02	183	2	5.33
352	15 DAP em	2.37	AT3G54470.1	Uridine 5′-monophosphate synthase	*A. thaliana*	51.82	6.94	288	4	7.56
Uncharacterized										
225	15 DAP em	1.57	AAY86774	C2 domain-containing protein	*N. caerulescens*	18.48	5.15	50	1	6.06
150	15 DAP em	1.51	AT1G17370.1	Oligouridylate binding protein 1B	*A. thaliana*	46.02	7.89	68	2	6.21
104	15 DAP em	3.40	CAE05556	OSJNBb0116K07.9	*O. sativa*	53.41	7.75	61	1	1.79
241	15 DAP em	2.34	XP_001753354	Predicted protein	*P. patens*	18.79	9.99	54	1	5.10
309	15 DAP em	2.24	BAD36713	PRLI-interacting factor L	*O. sativa*	49.14	6.07	114	1	7.61

[a] Spots in accordance with Supplementary Fig. 8.
[b] Developmental stage of endosperm (es) and embryo (em) at 15 DAP.
[c] Regulation of spots, spot volume at least 1.5 fold increased in the corresponding developmental stage.
[d] Accession numbers as given by SwissProt (http://www.uniprot.org), NCBInr (http://www.ncbi.nlm.nih.gov/protein) and TAIR (http://www.Arabidopsis.org/).
[e] Calculated molecular masses of the identified proteins as deduced from the corresponding genes.
[f] Calculated isoelectric point of the identified proteins as deduced from the corresponding genes.
[g] Probability score for the protein identifications based on MS/MS analysis and MASCOT search.
[h] Number of unique matching peptides.
[i] Sequence coverage of a protein by identified peptides.

Table 5 – Identification of proteins from spots with changed abundance in the endosperm compared to the embryo at 20 DAP. To determine significant changes in spot volume, a Student's t-test (p-value ≥ 0.05) was applied on the basis of normalized relative spot volume. Changes in spot volume ≥ 1.5 were considered to represent true alterations. Identification of proteins was carried out using the MASCOT search algorithm (www.matrixscience.com) against the (i) SwissProt (www.uniprot.org), (ii) NCBInr (www.ncbi.nig.gov) and (iii) TAIR (www.arabidopsis.org, TAIR release 10) databases. Identified proteins were functionally classified according to the KEGG PATHWAY Database (http://www.genome.jp/kegg/pathway.html). The pathway categories were adjusted for seed metabolism by adding four functional groups: (i) storage, (ii) defense, (iii) desiccation and (iv) detoxification.

ID[a]	Stage[b]	Reg[c]	Accession[d]	Name	Organism	MM calc[e]	pI calc[f]	Score[g]	Pep[h]	SC%[i]
Amino acid metabolism										
332	20 DAP es	1.63	P46645	Aspartate aminotransferase, cytoplasmic isozyme 1	*A. thaliana*	44.24	6.97	743	13	30.37
334	20 DAP es	2.26	P46643	Aspartate aminotransferase, mitochondrial	*A. thaliana*	47.73	9.13	957	14	31.63
34	20 DAP es	1.96	Q43314	Glutamate dehydrogenase 1	*A. thaliana*	44.50	6.42	116	4	9.00
23	20 DAP es	1.76	O04937	Glutamate dehydrogenase A	*N. plumbaginifolia*	44.77	6.69	60	1	2.43
381	20 DAP es	2.26	Q8GXE2	1,2-Dihydroxy-3-keto-5-methylthiopentene dioxygenase 2	*A. thaliana*	22.57	4.91	168	3	17.19
20	20 DAP es	1.87	AT2G36880.1	Methionine adenosyltransferase 3	*A. thaliana*	42.47	5.73	1916	22	61.54
384	20 DAP es	2.12	AT1G07780.1	Phosphoribosylanthranilate isomerase 1	*A. thaliana*	29.62	9.52	68	1	3.27
Biosynthesis of secondary metabolites										
341	20 DAP es	1.61	AT1G72680.1	Cinnamyl-alcohol dehydrogenase	*A. thaliana*	38.65	6.76	97	2	6.76
Carbohydrate metabolism										
301	20 DAP es	3.34	Q9C5I1	UDP-sugar pyrophosphorylase	*A. thaliana*	67.81	6.06	77	1	1.63
383	20 DAP es	1.57	BAB84009	Ascorbate peroxidase	*B. oleracea*	27.54	5.42	922	17	63.60
348	20 DAP es	2.07	AAY47048	Dehydroascorbate reductase	*S. lycopersicum*	23.52	6.36	309	5	23.81
74	20 DAP es	2.84	AAY47048	Dehydroascorbate reductase	*S. lycopersicum*	23.52	6.36	243	4	21.90
376	20 DAP es	1.84	NP_195093	D-Threo-aldose 1-dehydrogenase	*A. thaliana*	34.51	5.40	55	2	5.64
20	20 DAP es	1.87	AT5G28840.1	GDP-D-mannose 3′,5′-epimerase	*A. thaliana*	42.73	5.81	50	2	6.37
384	20 DAP es	2.12	CAA55209	L-Ascorbate peroxidase	*R. sativus*	27.67	5.41	788	18	62.80
350	20 DAP es	4.25	Q96533	Alcohol dehydrogenase class-3	*A. thaliana*	40.67	6.59	454	8	22.69
334	20 DAP es	2.26	AT3G52930.1	Aldolase superfamily protein	*A. thaliana*	38.52	6.04	56	2	7.26
33	20 DAP es	1.63	P25857	Glyceraldehyde-3-phosphate dehydrogenase B, chloroplast	*A. thaliana*	47.63	6.36	372	6	14.32
341	20 DAP es	1.61	P04796	Glyceraldehyde-3-phosphate dehydrogenase, cytosolic	*S. alba*	36.90	8.71	632	13	37.28
358	20 DAP es	2.83	AT1G76550.1	Phosphofructokinase family protein	*A. thaliana*	67.52	6.92	450	13	25.45
311	20 DAP es	2.56	P12783	Phosphoglycerate kinase, cytosolic	*T. aestivum*	42.10	5.55	77	2	7.48
169	20 DAP es	7.52	Q9SGC1	Probable phosphoglucomutase, cytoplasmic 2	*A. thaliana*	63.44	5.48	1202	17	26.50
301	20 DAP es	3.34	Q9SGC1	Probable phosphoglucomutase, cytoplasmic 2	*A. thaliana*	63.44	5.48	1372	15	24.96
358	20 DAP es	2.83	P21342	Pyrophosphate-fructose 6-phosphate 1-phosphotransferase subunit alpha	*S. tuberosum*	67.11	7.03	260	6	14.94
376	20 DAP es	1.84	Q38799	Pyruvate dehydrogenase E1 component subunit beta, mitochondrial	*A. thaliana*	39.15	5.57	65	1	3.03
383	20 DAP es	1.57	AT3G55440.1	Triosephosphate isomerase	*A. thaliana*	27.15	5.27	1049	9	28.74
369	20 DAP es	1.57	AT3G55440.1	Triosephosphate isomerase	*A. thaliana*	27.15	5.27	1238	11	45.67
384	20 DAP es	2.12	P48491	Triosephosphate isomerase, cytosolic	*A. thaliana*	27.15	5.27	88	3	8.27
362	20 DAP es	1.87	AAK50346	Putative 6-phosphogluconolactonase	*B. carinata*	29.19	6.75	235	10	36.43

195	20 DAP es	4.08	AAP37972	Seed specific protein Bn15D33A	*B. napus*	12.66	9.24	338	10	31.90
185	20 DAP es	3.62	AAP37972	Seed specific protein Bn15D33A	*B. napus*	12.66	9.24	835	5	68.10
16	20 DAP es	2.13	AT5G15490.1	UDP-glucose 6-dehydrogenase family protein	*A. thaliana*	53.08	5.68	2273	4	45.63
340	20 DAP es	2.67	AT5G59290.1	UDP-glucuronic acid decarboxylase 3	*A. thaliana*	38.54	9.08	963	18	45.32
170	20 DAP es	1.67	Q9SIB9	Aconitate hydratase 2, mitochondrial	*A. thaliana*	108.13	6.79	1532	31	27.37
326	20 DAP es	2.13	Q9SJH7	Citrate synthase 3, peroxisomal	*A. thaliana*	56.14	7.79	447	3	7.07
349	20 DAP es	1.72	Q9SJH7	Citrate synthase 3, peroxisomal	*A. thaliana*	56.14	7.79	561	3	7.07
328	20 DAP es	2.00	Q9SJH7	Citrate synthase 3, peroxisomal	*A. thaliana*	56.14	7.79	304	3	7.07
23	20 DAP es	1.76	AT1G65930.1	Cytosolic NADP+-dependent isocitrate dehydrogenase	*A. thaliana*	45.72	6.13	112	2	6.34
234	20 DAP es	1.90	Q43744	Malate dehydrogenase, mitochondrial	*B. napus*	35.69	9.48	91	2	10.26
311	20 DAP es	2.56	O82662	Succinyl-CoA ligase GDP-forming subunit beta, mitochondrial	*A. thaliana*	45.32	6.32	1249	20	44.42
Cellular processes										
311	20 DAP es	2.56	P80607	Alpha-1,4-glucan-protein synthase UDP-forming	*Z. mays*	41.18	5.70	452	10	23.35
113	20 DAP es	2.11	Q9SVJ4	Endoglucanase 22	*A. thaliana*	54.98	6.36	58	1	1.62
52	20 DAP es	2.83	AAD03693	Fibrillin	*B. napus*	25.86	4.98	266	1	44.73
311	20 DAP es	2.56	AT3G08900.1	UDP-arabinopyranose mutase 3	*A. thaliana*	41.25	5.28	563	17	40.61
23	20 DAP es	1.76	Q8H038	Xyloglucan galactosyltransferase KATAMARI1 homolog	*O. sativa*	66.67	5.54	51	1	1.02
16	20 DAP es	2.13	P20363	Tubulin alpha-3/alpha-5 chain	*A. thaliana*	49.62	4.82	404	9	25.11
4	20 DAP es	2.67	Q56WK6	Patellin-1	*A. thaliana*	64.01	4.67	830	13	18.32
52	20 DAP es	2.83	Q94FZ9	Plastid lipid-associated protein 1, chloroplastic	*B. campestris*	35.62	5.19	449	7	32.11
220	20 DAP es	1.64	AT3G07680.1	GOLD family protein	*A. thaliana*	24.32	5.93	517	10	40.87
52	20 DAP es	2.83	Q8RWG8	Ran-binding protein 1 homolog b	*A. thaliana*	24.37	4.68	220	3	14.75
94	20 DAP es	7.34	AT4G29350.1	Profilin 2	*A. thaliana*	13.99	4.77	301	4	16.79
95	20 DAP es	2.75	Q64LH1	Profilin-1	*A. artemisiifolia*	14.10	4.58	520	5	17.56
96	20 DAP es	2.75	Q64LH1	Profilin-1	*A. artemisiifolia*	14.10	4.58	439	3	16.03
345	20 DAP es	4.44	AT5G23540.1	Mov34/MPN/PAD-1 family protein	*A. thaliana*	34.33	6.34	236	4	12.99
Defense										
114	20 DAP es	4.11	ABV89615	Bet v, allergen family protein	*B. rapa*	17.15	5.15	161	4	29.03
147	20 DAP es	3.92	AT1G52400.1	Beta glucosidase 18	*A. thaliana*	60.42	6.82	264	5	7.01
148	20 DAP es	3.92	AT1G52400.1	Beta glucosidase 18	*A. thaliana*	60.42	6.82	363	5	5.68
149	20 DAP es	2.47	Q9SE50	Beta-glucosidase 18	*A. thaliana*	60.42	6.82	310	7	8.71
146	20 DAP es	5.58	Q9SE50	Beta-glucosidase 18	*A. thaliana*	60.42	6.82	329	7	7.01
307	20 DAP es	4.34	Q9SE50	Beta-glucosidase 18	*A. thaliana*	60.42	6.82	201	5	5.68
144	20 DAP es	3.41	Q9LV33	Beta-glucosidase 44	*A. thaliana*	58.95	9.81	311	7	13.09
270	20 DAP es	5.11	AT3G16470.3	Mannose-binding lectin superfamily protein	*A. thaliana*	32.07	5.15	438	3	13.80
81	20 DAP es	8.75	Q9ZVF2	MLP-like protein 329	*A. thaliana*	17.59	5.21	383	4	26.49
270	20 DAP es	5.11	O04309	Myrosinase-binding protein	*A. thaliana*	48.47	4.99	457	3	9.09
129	20 DAP es	2.41	P52778	Protein LlR18A	*L. luteus*	16.85	5.07	67	4	7.69
23	20 DAP es	1.76	P52778	Protein LlR18A	*L. luteus*	16.85	5.07	54	1	7.69
18	20 DAP es	352.11	P52778	Protein LlR18A	*L. luteus*	16.85	5.07	55	1	7.69
20	20 DAP es	1.87	P52778	Protein LlR18A	*L. luteus*	16.85	5.07	52	1	7.69

(continued on next page)

Table 5 (*continued*)

ID [a]	Stage [b]	Reg [c]	Accession [d]	Name	Organism	MM calc [e]	pI calc [f]	Score [g]	Pep [h]	SC% [i]
Dessiccation										
52	20 DAP es	2.83	AT2G44060.1	Late embryogenesis abundant protein, group 2	*A. thaliana*	36.01	4.52	53	4	16.31
Detoxification										
92	20 DAP es	1.63	P09678	Superoxide dismutase Cu–Zn	*B. oleracea*	15.16	5.76	496	4	25.66
91	20 DAP es	3.80	AAD05576	Cu/Zn superoxide dismutase	*R. sativus*	15.09	5.41	508	6	34.21
Energy metabolism										
182	20 DAP es	7.38	AT5G38430.1	Ribulose bisphosphate carboxylase (small chain) family protein181	*A. thaliana*	20.27	8.86	723	7	31.49
197	20 DAP es	2.88	P10796	Ribulose bisphosphate carboxylase small chain 1B, chloroplastic	*A. thaliana*	20.27	8.86	447	5	17.13
16	20 DAP es	2.13	ATCG00490.1	Ribulose-bisphosphate carboxylases	*A. thaliana*	52.92	5.85	205	5	11.27
197	20 DAP es	2.88	Q9ZST4	Nitrogen regulatory protein P-II homolog	*A. thaliana*	21.26	9.93	77	1	5.61
345	20 DAP es	4.44	AT1G50940.1	Electron transfer flavoprotein alpha	*A. thaliana*	38.38	6.56	59	1	5.23
208	20 DAP es	2.84	Q9SJ12	Probable ATP synthase 24 kDa subunit, mitochondrial	*A. thaliana*	27.58	6.30	359	7	22.92
208	20 DAP es	2.84	AT2G18230.1	Pyrophosphorylase 2	*A. thaliana*	24.66	5.68	71	2	13.30
311	20 DAP es	2.56	AT1G12840.1	Vacuolar ATP synthase subunit C	*A. thaliana*	42.59	5.28	647	14	32.80
4	20 DAP es	2.67	A4QKH8	ATP synthase subunit b, chloroplastic	*C. bursa-pastoris*	21.04	9.05	98	1	4.35
131	20 DAP es	5.03	P0CJ48	Chlorophyll a-b binding protein 2, chloroplastic	*A. thaliana*	28.21	5.16	52	1	3.00
384	20 DAP es	2.12	Q9XF89	Chlorophyll a-b binding protein CP26, chloroplastic	*A. thaliana*	30.14	5.98	55	2	7.14
340	20 DAP es	2.67	Q9LTL8	Cytochrome P450 71B24	*A. thaliana*	56.77	6.64	50	1	1.41
381	20 DAP es	2.26	P11594	Oxygen-evolving enhancer protein 2, chloroplastic	*S. alba*	27.91	7.60	51	2	11.92
23	20 DAP es	1.76	Q9LIK9	ATP sulfurylase 1, chloroplastic	*A. thaliana*	51.43	6.36	721	13	21.60
149	20 DAP es	2.47	AT5G04590.1	Sulfite reductase	*A. thaliana*	71.90	9.17	365	13	18.85
146	20 DAP es	5.58	AT5G04590.1	Sulfite reductase	*A. thaliana*	71.90	9.17	139	5	7.63
Environmental information processing										
376	20 DAP es	1.84	Q9SYT0	Annexin D1	*A. thaliana*	36.18	5.09	1504	21	51.74
385	20 DAP es	2.70	Q9SYT0	Annexin D1	*A. thaliana*	36.18	5.09	2250	22	51.74
363	20 DAP es	5.54	Q9XEE2	Annexin D2	*A. thaliana*	36.24	5.71	2585	18	40.06
345	20 DAP es	4.44	Q9XEE2	Annexin D2	*A. thaliana*	36.24	5.71	1276	11	29.65
96	20 DAP es	2.75	XP_002880102	Calcium-binding EF hand family protein	*A. lyrata*	15.82	4.29	135	2	15.49
76	20 DAP es	1.79	P93087	Calmodulin	*C. annuum*	16.82	3.94	1283	10	69.13
169	20 DAP es	7.52	AT2G46370.1	Auxin-responsive GH3 family protein	*A. thaliana*	64.31	5.38	346	8	16.35
169	20 DAP es	7.52	Q9SKE2	Jasmonic acid-amido synthetase JAR1	*A. thaliana*	64.31	5.38	323	8	16.35
362	20 DAP es	1.87	P38548	GTP-binding nuclear protein Ran/TC4	*V. faba*	25.27	6.44	1246	12	52.49
88	20 DAP es	2.50	P35134	Ubiquitin-conjugating enzyme E2 11	*A. thaliana*	16.54	8.92	93	2	30.41
179	20 DAP es	1.93	Q9SVD7	Ubiquitin-conjugating enzyme E2 variant 1D	*A. thaliana*	16.52	6.24	919	12	65.07
83	20 DAP es	4.48	AT3G03270.1	Adenine nucleotide alpha hydrolases-like superfamily protein	*A. thaliana*	22.59	5.44	143	2	15.42
196	20 DAP es	4.01	AT3G03270.1	Adenine nucleotide alpha hydrolases-like superfamily protein	*A. thaliana*	22.59	5.44	985	10	38.31

231	20 DAP es	1.77	AT3G17020.1	Adenine nucleotide alpha hydrolases-like superfamily protein	*A. thaliana*	17.76	6.52	65	1	13.50
381	20 DAP es	2.26	AT3G11930.1	Adenine nucleotide alpha hydrolases-like superfamily protein	*A. thaliana*	21.44	5.46	53	1	5.53
71	20 DAP es	6.53	AT3G11930.1	Adenine nucleotide alpha hydrolases-like superfamily protein	*A. thaliana*	21.44	5.46	97	3	21.11
374	20 DAP es	3.60	AT3G11930.1	Adenine nucleotide alpha hydrolases-like superfamily protein	*A. thaliana*	21.44	5.46	169	2	18.09
376	20 DAP es	1.84	AT4G33670.1	Aldo-keto reductase family 4 member C8	*A. thaliana*	34.51	5.40	73	2	5.64
129	20 DAP es	2.41	Q9SAR5	Ankyrin repeat domain-containing protein 2	*A. thaliana*	36.96	4.34	191	4	14.33
78	20 DAP es	3.84	AAK01359	Dehydration stress-induced protein	*B. napus*	19.72	4.67	315	6	30.90
195	20 DAP es	4.08	Q9FNE2	Glutaredoxin-C2	*A. thaliana*	11.75	7.68	209	2	9.91
71	20 DAP es	6.53	ABD36807	Glutathione S-transferase	*B. napus*	24.72	5.78	104	2	9.68
374	20 DAP es	3.60	ABD36807	Glutathione S-transferase	*B. napus*	24.72	5.78	192	4	18.43
348	20 DAP es	2.07	Q9FRL8	Glutathione S-transferase DHAR2	*A. thaliana*	23.39	5.75	127	2	7.04
74	20 DAP es	2.84	Q9FRL8	Glutathione S-transferase DHAR2	*A. thaliana*	23.39	5.75	177	3	10.33
384	20 DAP es	2.12	AT1G10370.1	Glutathione S-transferase family protein	*A. thaliana*	25.29	6.23	80	2	6.61
204	20 DAP es	2.84	AT1G10370.1	Glutathione S-transferase family protein	*A. thaliana*	25.29	6.23	293	6	16.74
208	20 DAP es	2.84	AT1G10370.1	Glutathione S-transferase family protein	*A. thaliana*	25.29	6.23	116	4	12.33
205	20 DAP es	3.82	AT1G10370.1	Glutathione S-transferase family protein	*A. thaliana*	25.29	6.23	259	7	19.38
70	20 DAP es	2.10	P46422	Glutathione S-transferase PM24	*A. thaliana*	24.11	5.91	105	3	11.79
381	20 DAP es	2.26	O04885	Lactoylglutathione lyase	*B. juncea*	20.77	5.49	707	16	55.14
110	20 DAP es	1.63	Q9SRZ4	Peroxiredoxin-2C	*A. thaliana*	17.40	5.22	197	2	6.17
350	20 DAP es	4.25	ACR40091	S-Nitrosoglutathione reductase	*B. juncea*	40.07	8.72	185	1	24.33
83	20 DAP es	4.48	ABV89642	Universal stress protein family protein	*B. rapa*	17.58	9.02	260	4	42.14
71	20 DAP es	6.53	XP_002882756	Universal stress protein family protein	*A. lyrata*	21.69	5.28	53	3	25.50
196	20 DAP es	4.01	ABV89642	Universal stress protein family protein	*B. rapa*	17.58	9.02	804	14	75.47
Genetic information processing										
239	20 DAP es	2.33	Q9LKR3	Luminal-binding protein 1	*A. thaliana*	73.58	4.94	3095	30	40.96
270	20 DAP es	5.11	Q940A6	Pentatricopeptide repeat-containing protein	*A. thaliana*	94.55	9.49	51	1	0.72
52	20 DAP es	2.83	Q94JW0	Deoxyhypusine hydroxylase	*A. thaliana*	34.07	4.72	59	1	3.50
98	20 DAP es	1.89	P34893	10 kDa chaperonin	*A. thaliana*	10.81	7.68	470	1	61.22
381	20 DAP es	2.26	O65282	20 kDa chaperonin, chloroplastic	*A. thaliana*	26.79	9.35	169	3	13.04
129	20 DAP es	2.41	AT1G53850.1	20S proteasome alpha subunit E1	*A. thaliana*	25.93	4.55	52	1	5.06
131	20 DAP es	5.03	CAD10778	20S proteasome subunit alpha V	*P. patens*	26.24	4.70	68	1	3.80
106	20 DAP es	2.16	Q38806	22.0 kDa heat shock protein	*A. thaliana*	21.98	5.48	167	5	22.05
345	20 DAP es	4.44	Q9LT08	26S proteasome non-ATPase regulatory subunit 14	*A. thaliana*	34.33	6.34	245	4	12.99
52	20 DAP es	2.83	XP_002862370	26S proteasome regulatory subunit	*A. lyrata*	31.90	4.66	529	7	32.42
246	20 DAP es	3.23	AT1G60420.1	DC1 domain-containing protein	*A. thaliana*	65.13	4.75	292	5	9.69
16	20 DAP es	2.13	AT3G44110.1	DNAJ homolog 3	*A. thaliana*	46.41	5.72	51	2	6.43

(continued on next page)

Table 5 (*continued*)

ID [a]	Stage [b]	Reg [c]	Accession [d]	Name	Organism	MM calc [e]	pI calc [f]	Score [g]	Pep [h]	SC% [i]
197	20 DAP es	2.88	Q38935	FK506-binding protein 2-1	*A. thaliana*	16.34	9.29	103	3	23.53
4	20 DAP es	2.67	AT4G24190.1	Hsp90-7	*A. thaliana*	94.15	4.79	74	1	1.70
197	20 DAP es	2.88	AAC49390	Immunophilin	*A. thaliana*	15.68	9.29	85	3	24.66
184	20 DAP es	3.16	XP_002522624	Immunophilin, putative	*R. communis*	12.00	9.17	499	4	14.29
23	20 DAP es	1.76	AT1G29150.1	Non-ATPase subunit 9	*A. thaliana*	46.72	6.25	1215	21	39.14
246	20 DAP es	3.23	AT1G77510.1	PDI-like 1-2	*A. thaliana*	56.33	4.76	693	8	12.20
143	20 DAP es	3.21	P34790	Peptidyl-prolyl cis-trans isomerase	*A. thaliana*	18.36	9.01	587	7	46.51
246	20 DAP es	3.23	O80763	Probable nucleoredoxin 1	*A. thaliana*	65.13	4.75	310	5	9.69
372	20 DAP es	1.67	O23708	Proteasome subunit alpha type-2-A	*A. thaliana*	25.69	5.40	431	5	31.91
208	20 DAP es	2.84	O23715	Proteasome subunit alpha type-3	*A. thaliana*	27.36	5.90	307	6	24.10
383	20 DAP es	1.57	Q9XG77	Proteasome subunit alpha type-6	*N. tabacum*	27.29	5.89	72	3	11.79
106	20 DAP es	2.16	Q9XI05	Proteasome subunit beta type-3-A	*A. thaliana*	22.78	5.17	454	5	29.41
70	20 DAP es	2.10	Q7DLR9	Proteasome subunit beta type-4	*A. thaliana*	27.63	6.10	1371	14	47.97
227	20 DAP es	1.60	Q9LIP2	Proteasome subunit beta type-5-B	*A. thaliana*	29.47	5.79	1305	10	26.74
245	20 DAP es	3.09	ABB17025	Protein disulfide isomerase	*B. carinata*	55.71	4.86	383	12	18.66
110	20 DAP es	1.63	ACD56608	Putative protein disulfide isomerase family protein	*G. kirkii*	55.56	4.84	81	1	1.82
349	20 DAP es	1.72	NP_199115	Regulatory particle triple-A ATPase 4A	*A. thaliana*	44.79	9.02	175	5	12.78
135	20 DAP es	2.52	AT4G34870.1	Rotamase cyclophilin 5 RD	*A. thaliana*	18.37	9.84	106	1	8.72
176	20 DAP es	1.68	AT5G67360.1	Subtilase family protein	*A. thaliana*	79.37	5.90	77	3	4.23
96	20 DAP es	2.75	Q9XGS0	Thioredoxin M-type, chloroplastic	*B. napus*	19.26	10.27	142	1	9.04
70	20 DAP es	2.10	Q9LT79	U-box domain-containing protein 25	*A. thaliana*	46.03	7.86	50	1	1.66
381	20 DAP es	2.26	Q9LT79	U-box domain-containing protein 25	*A. thaliana*	46.03	7.86	55	1	1.66
184	20 DAP es	3.16	Q9LT79	U-box domain-containing protein 25	*A. thaliana*	46.03	7.86	56	1	1.66
16	20 DAP es	2.13	CAD41363	OSJNBa0088A01.2	*O. sativa*	20.50	12.28	116	1	6.84
328	20 DAP es	2.00	AT2G18510.1	RNA-binding (RRM/RBD/RNP motifs) family protein	*A. thaliana*	39.86	7.94	274	3	9.92
227	20 DAP es	1.60	AT1G17880.1	Basic transcription factor 3	*A. thaliana*	17.94	7.55	91	1	11.52
176	20 DAP es	1.68	Q84W89	DEAD-box ATP-dependent RNA helicase 37	*A. thaliana*	67.58	6.70	1356	22	28.59
245	20 DAP es	3.09	AAF31402	Putative G.-rich RNA binding protein 1	*C. roseus*	14.15	9.61	57	1	10.95
114	20 DAP es	4.11	AAF31402	Putative G.-rich RNA binding protein 1	*C. roseus*	14.15	9.61	72	1	7.30
110	20 DAP es	1.63	Q9M8R9	Regulator of ribonuclease-like protein 1	*A. thaliana*	17.83	5.62	56	2	12.05
71	20 DAP es	6.53	AT5G28640.1	SSXT family protein	*A. thaliana*	22.45	5.78	56	1	6.67
95	20 DAP es	2.75	Q9LH85	60S acidic ribosomal protein	*A. thaliana*	11.73	4.40	57		7.83
96	20 DAP es	2.75	Q9LH85	60S acidic ribosomal protein P2-3	*A. thaliana*	11.73	4.40	64	1	7.83
349	20 DAP es	1.72	Q9SF40	60S ribosomal protein L4-1	*A. thaliana*	44.67	10.86	317	10	17.00

52	20 DAP es	2.83	P48006	Elongation factor 1-delta 1	*A. thaliana*	25.10	4.30	133	4	17.70
208	20 DAP es	2.84	O04663	Eukaryotic translation initiation factor 4E-2	*A. thaliana*	22.50	5.39	118	4	13.64
358	20 DAP es	2.83	P42731	Polyadenylate-binding protein 2	*A. thaliana*	68.63	8.82	1012	19	21.94
176	20 DAP es	1.68	O64380	Polyadenylate-binding protein 3	*A. thaliana*	72.83	9.33	105	3	3.64
129	20 DAP es	2.41	AT3G55620.1	Translation initiation factor IF6	*A. thaliana*	26.47	4.48	315	5	22.86
Lipid metabolism										
141	20 DAP es	1.51	P33207	3-Oxoacyl-acyl-carrier-protein reductase, chloroplastic	*A. thaliana*	33.53	10.12	65	2	5.02
78	20 DAP es	3.84	AT4G39730.1	Lipase/lipooxygenase family protein	*A. thaliana*	20.12	4.83	134	2	8.84
Metabolism of cofactors and vitamins										
119	20 DAP es	7.05	XP_002890203	Soul heme-binding family protein	*A. lyrata*	25.50	4.54	153	2	6.90
131	20 DAP es	5.03	AT4G29270.1	HAD superfamily acid phosphatase	*A. thaliana*	28.73	9.14	62	1	4.30
Metabolism of terpenoids and polyketides										
340	20 DAP es	2.67	AT3G46440.1	UDP-XYL synthase 5	*A. thaliana*	38.36	7.82	812	1	37.83
Nucleotide metabolism										
139	20 DAP es	1.51	Q9FK35	Adenylate kinase 2	*A. thaliana*	27.32	7.74	64	2	6.45
Storage										
381	20 DAP es	2.26	AT4G14710.1	RmlC-like cupins superfamily protein	*A. thaliana*	23.34	4.86	154	3	16.58
Uncharacterized										
239	20 DAP es	2.33	NP_001149638	cePP protein	*Z. mays*	45.27	4.67	55	1	2.01
94	20 DAP es	7.34	AT1G42960.1	Expressed protein localized to the inner membrane of the chloroplast	*A. thaliana*	17.82	9.42	64	2	13.69
74	20 DAP es	2.84	AAF79440	F18O14.33	*A. thaliana*	50.13	6.92	149	1	2.50
197	20 DAP es	2.88	AT4G01900.1	GLNB1 homolog	*A. thaliana*	21.26	9.93	73	1	5.61
199	20 DAP es	6.49	AT2G39050.1	Hydroxyproline-rich glycoprotein family protein	*A. thaliana*	35.63	6.21	137	3	12.30
135	20 DAP es	2.52	EEC81610	Hypothetical protein OsI_25113	*O. sativa*	69.89	11.07	77	1	2.37
25	20 DAP es	2.19	NP_001045523	Os01g0969100	*O. sativa*	44.30	5.82	92	3	6.03
81	20 DAP es	8.75	AT1G61600.1	Protein of unknown function	*A. thaliana*	48.74	10.31	51	1	1.90
94	20 DAP es	7.34	AT1G61600.1	Protein of unknown function (DUF1262)	*A. thaliana*	48.74	10.31	50	1	1.90
383	20 DAP es	1.57	Q9SJ32	Putative F-box/FBD/LRR-repeat protein	*A. thaliana*	50.35	9.65	50	1	1.37
91	20 DAP es	3.80	Q9ZUX4	Uncharacterized protein At2g27730, mitochondrial	*A. thaliana*	11.94	10.10	91	1	7.08
18	20 DAP es	352.11	BAA99394	Vacuolar calcium binding protein	*R. sativus*	27.09	3.95	245	4	12.90
Xenobiotics biodegradation and metabolism										
378	20 DAP es	1.67	AT2G32520.1	Carboxymethylenebutenolidase	*A. thaliana*	25.91	5.14	114	3	11.30
Amino acid metabolism										
69	20 DAP em	2.22	XP_002530662	3-isopropylmalate dehydratase, putative	*R. communis*	26.91	6.54	89	3	5.51
21	20 DAP em	1.91	P29102	3-isopropylmalate dehydrogenase, chloroplastic	*B. napus*	43.32	6.05	51	3	11.08
69	20 DAP em	2.22	AT2G43090.1	Aconitase/3-isopropylmalate dehydratase protein	*A. thaliana*	26.77	6.41	295	4	21.51

(continued on next page)

Table 5 (*continued*)

ID [a]	Stage [b]	Reg [c]	Accession [d]	Name	Organism	MM calc [e]	pI calc [f]	Score [g]	Pep [h]	SC% [i]
Carbohydrate metabolism										
273	20 DAP em	2.04	P25696	Bifunctional enolase 2/transcriptional activator	*A. thaliana*	47.69	5.45	60	1	2.70
352	20 DAP em	1.58	AT4G16155.1	Dihydrolipoyl dehydrogenases	*A. thaliana*	67.05	8.66	404	9	16.83
352	20 DAP em	1.58	AT3G16950.1	Lipoamide dehydrogenase 1	*A. thaliana*	60.72	9.00	393	3	19.12
7	20 DAP em	1.51	AT1G12000.1	Phosphofructokinase family protein	*A. thaliana*	61.42	5.74	577	12	14.31
259	20 DAP em	1.67	AT4G24620.1	Phosphoglucose isomerase 1	*A. thaliana*	67.01	5.37	247	7	12.56
21	20 DAP em	1.91	Q9LD57	Phosphoglycerate kinase 1, chloroplastic	*A. thaliana*	50.08	5.86	2382	21	44.49
352	20 DAP em	1.58	AT3G02360.1	6-Phosphogluconate dehydrogenase family protein	*A. thaliana*	53.54	7.67	751	17	32.30
21	20 DAP em	1.91	ACK56136	Transaldolase	*D. longan*	47.88	6.38	192	4	10.23
276	20 DAP em	3.85	XP_002876580	Transketolase	*A. lyrata*	79.80	5.82	1544	21	26.86
241	20 DAP em	3.95	AT3G60750.1	Transketolase	*A. thaliana*	79.92	5.92	2075	20	24.83
17	20 DAP em	1.94	AT1G65930.1	Cytosolic NADP+-dependent isocitrate dehydrogenase	*A. thaliana*	45.72	6.13	196	5	14.15
Cellular processes										
41	20 DAP em	2.48	AT5G16510.1	Alpha-1,4-glucan-protein synthase family protein	*A. thaliana*	38.56	4.92	179	2	5.75
274	20 DAP em	1.56	Q9FM01	Probable UDP-glucose 6-dehydrogenase 2	*A. thaliana*	53.06	5.51	354	8	19.38
161	20 DAP em	3.88	Q8H038	Xyloglucan galactosyltransferase KATAMARI1 homolog	*O. sativa*	66.67	5.54	53	1	1.02
274	20 DAP em	1.56	P29511	Tubulin alpha-6 chain	*A. thaliana*	49.51	4.79	2390	27	57.56
271	20 DAP em	1.56	P29511	Tubulin alpha-6 chain	*A. thaliana*	49.51	4.79	1230	19	50.44
240	20 DAP em	2.38	AT5G49910.1	Chloroplast heat shock protein 70-2 |	*A. thaliana*	76.95	5.03	1960	27	34.26
21	20 DAP em	1.91	P53492	Actin-7	*A. thaliana*	41.71	5.20	1305	20	45.62
365	20 DAP em	1.73	AT4G35220.1	Cyclase family protein	*A. thaliana*	29.97	5.59	445	7	19.12
284	20 DAP em	2.06	AT4G16143.1	Importin alpha isoform 2	*A. thaliana*	58.87	4.86	59	2	4.86
Defense										
161	20 DAP em	3.88	P52778	Protein LlR18A	*L. luteus*	16.85	5.07	51	1	7.69
380	20 DAP em	2.14	XP_002887781	Metacaspase 7	*A. lyrata*	45.48	4.61	80	1	2.87
Detoxification										
258	20 DAP em	2.06	Q9S795	Betaine aldehyde dehydrogenase 1, chloroplastic	*A. thaliana*	54.40	5.04	144	2	4.99
218	20 DAP em	1.94	AAC15842	Superoxide dismutase	*R. sativus*	23.79	5.97	187	6	20.28
Energy metabolism										
273	20 DAP em	2.04	P48686	Ribulose bisphosphate carboxylase large chain	*B. oleracea*	52.92	5.85	1836	27	42.80
178	20 DAP em	1.88	P05346	Ribulose bisphosphate carboxylase small chain, chloroplastic	*B. napus*	20.21	9.22	943	12	58.01
181	20 DAP em	1.86	P05346	Ribulose bisphosphate carboxylase small chain, chloroplastic	*B. napus*	20.21	9.22	1200	12	58.01
256	20 DAP em	1.57	P21240	RuBisCO large subunit-binding protein subunit beta, chloroplastic	*A. thaliana*	63.77	6.18	488	7	15.00
252	20 DAP em	1.58	P08927	RuBisCO large subunit-binding protein subunit beta, chloroplastic	*P. sativum*	62.95	5.76	1230	17	29.08
121	20 DAP em	1.89	Q9FT52	ATP synthase subunit d, mitochondrial	*A. thaliana*	19.57	4.94	557	6	30.95
299	20 DAP em	1.68	Q9FGI6	NADH-ubiquinone oxidoreductase 75 kDa subunit, mitochondrial	*A. thaliana*	81.47	6.24	1235	18	26.74
103	20 DAP em	1.92	Q39258	V-type proton ATPase subunit E1	*A. thaliana*	26.04	6.02	672	10	29.57

284	20 DAP em	2.06	P56757	ATP synthase subunit alpha, chloroplastic	*A. thaliana*	55.29	5.04	1639	18	37.28
56	20 DAP em	1.71	P23321	Oxygen-evolving enhancer protein 1-1, chloroplastic	*A. thaliana*	35.12	5.39	1562	17	52.41
103	20 DAP em	1.92	Q9S841	Oxygen-evolving enhancer protein 1-2, chloroplastic	*A. thaliana*	35.00	5.84	1450	17	48.34
379	20 DAP em	3.31	P11594	Oxygen-evolving enhancer protein 2, chloroplastic	*S. alba*	27.91	7.60	601	8	33.85
69	20 DAP em	2.22	P11594	Oxygen-evolving enhancer protein 2, chloroplastic	*S. alba*	27.91	7.60	780	9	35.77
380	20 DAP em	2.14	P11594	Oxygen-evolving enhancer protein 2, chloroplastic	*S. alba*	27.91	7.60	610	7	25.77
68	20 DAP em	1.70	Q96334	Oxygen-evolving enhancer protein 2, chloroplastic (fragment)	*B. juncea*	23.33	4.76	80	2	14.29
68	20 DAP em	1.70	AT1G61520.1	Photosystem I light harvesting complex	*A. thaliana*	29.16	9.11	665	6	17.58
198	20 DAP em	2.51	P49108	Photosystem II 10 kDa polypeptide, chloroplastic	*B. campestris*	14.64	9.27	355	4	31.91
69	20 DAP em	2.22	AT1G06680.1	Photosystem II subunit P-1	*A. thaliana*	28.08	7.67	513	6	25.48
Environmental information processing										
123	20 DAP em	4.16	BAE72093	9-cis-epoxycarotenoid dioxygenase 4	*L. sativa*	64.02	8.59	56	1	1.21
54	20 DAP em	1.62	AAK26634	GF14 omega	*B. napus*	29.11	4.53	475	13	47.69
61	20 DAP em	1.62	AT5G12110.1	Glutathione S-transferase, C-terminal-like	*A. thaliana*	24.77	4.37	413	1	15.35
225	20 DAP em	2.11	CAA07494	heat stress-induced protein	*B. oleracea*	23.47	9.23	52	2	15.14
69	20 DAP em	2.22	O04885	Lactoylglutathione lyase	*B. juncea*	20.77	5.49	106	4	26.49
56	20 DAP em	1.71	Q8W593	Probable lactoylglutathione lyase, chloroplast	*A. thaliana*	39.14	7.66	92	1	3.14
Genetic information processing										
240	20 DAP em	2.38	ACL01101	Zinc finger protein	*P. edulis*	17.43	10.62	54	1	6.10
380	20 DAP em	2.14	O65282	20 kDa chaperonin, chloroplastic	*A. thaliana*	26.79	9.35	199	5	25.30
68	20 DAP em	1.70	O65282	20 kDa chaperonin, chloroplastic	*A. thaliana*	26.79	9.35	1000	13	50.20
54	20 DAP em	1.62	AT5G60360.1	Aleurain-like protease	*A. thaliana*	38.93	6.29	53	2	8.10
61	20 DAP em	1.62	AT5G60360.1	Aleurain-like protease	*A. thaliana*	38.93	6.29	453	7	17.04
240	20 DAP em	2.38	ABE79560	Chaperone DnaK	*M. truncatula*	75.71	5.05	886	1	30.13
252	20 DAP em	1.58	P35480	Chaperonin CPN60, mitochondrial	*B. napus*	62.32	9.22	895	17	30.49
256	20 DAP em	1.57	P29197	Chaperonin CPN60, mitochondrial	*A. thaliana*	61.24	5.53	2238	31	46.62
5	20 DAP em	3.36	AT5G50920.1	CLPC homolog 1	*A. thaliana*	103.39	6.35	3615	52	49.73
5	20 DAP em	3.36	P42762	ERD1 protein, chloroplastic	*A. thaliana*	103.17	5.84	164	1	2.12
221	20 DAP em	4.31	P34791	Peptidyl-prolyl cis-trans isomerase CYP20-3, chloroplastic	*A. thaliana*	28.19	9.68	196	5	22.69
225	20 DAP em	2.11	P34791	Peptidyl-prolyl cis-trans isomerase CYP20-3, chloroplastic	*A. thaliana*	28.19	9.68	417	6	16.92
219	20 DAP em	1.71	P34791	Peptidyl-prolyl cis-trans isomerase CYP20-3, chloroplastic	*A. thaliana*	28.19	9.68	377	5	13.08
259	20 DAP em	1.67	ABB17025	Protein disulfide isomerase	*B. carinata*	55.71	4.86	2355	34	63.65
252	20 DAP em	1.58	AAF05855	Putative T-complex protein 1, theta subunit	*A. thaliana*	56.92	5.15	54	4	9.09
54	20 DAP em	1.62	AAL60582	Senescence-associated cysteine protease	*B. oleracea*	39.31	5.38	93	3	13.93
7	20 DAP em	1.51	P28769	T-complex protein 1 subunit alpha	*A. thaliana*	59.19	5.89	749	16	29.17

(continued on next page)

Table 5 (*continued*)

ID[a]	Stage[b]	Reg[c]	Accession[d]	Name	Organism	MM calc[e]	pI calc[f]	Score[g]	Pep[h]	SC%[i]
284	20 DAP em	2.06	XP_002864384	Trigger factor type chaperone family protein	*A. lyrata*	61.69	5.01	51	1	2.36
276	20 DAP em	3.85	AT5G10540.1	Zincin-like metalloproteases family protein	*A. thaliana*	78.99	5.34	122	4	5.28
61	20 DAP em	1.62	Q9ZUU4	Ribonucleoprotein, chloroplastic	*A. thaliana*	30.70	4.91	129	2	6.23
218	20 DAP em	1.94	ACG26589	G.-rich protein 2b	*Z. mays*	20.29	5.91	95	1	7.69
365	20 DAP em	1.73	AT1G07660.1	Histone superfamily protein	*A. thaliana*	11.40	11.96	98	2	21.36
274	20 DAP em	1.56	Q9FL92	Probable WRKY transcription factor 16	*A. thaliana*	155.60	5.92	67	1	0.80
54	20 DAP em	1.62	Q84WM9	Elongation factor 1-beta 1	*A. thaliana*	24.77	4.37	453	4	15.35
61	20 DAP em	1.62	Q9SCX3	Elongation factor 1-beta 2	*A. thaliana*	24.19	4.27	478	6	25.45
56	20 DAP em	1.71	P17745	Elongation factor Tu, chloroplastic	*A. thaliana*	51.60	5.79	76	1	2.31
241	20 DAP em	3.95	AT5G44320.1	Eukaryotic translation initiation factor 3 subunit 7	*A. thaliana*	66.18	5.32	145	4	8.16
167	20 DAP em	2.59	Q9ZPI1	Lysyl-tRNA synthetase	*A. thaliana*	70.84	5.85	365	9	11.82
61	20 DAP em	1.62	AT3G12390.1	Nascent polypeptide-associated complex (NAC), alpha subunit family protein	*A. thaliana*	21.97	4.11	72	3	10.84
54	20 DAP em	1.62	Q9LHG9	Nascent polypeptide-associated complex subunit alpha-like protein 1	*A. thaliana*	21.97	4.11	184	4	19.21
161	20 DAP em	3.88	AT2G23350.1	Poly(A) binding protein 4	*A. thaliana*	71.61	6.44	459	9	14.80
365	20 DAP em	1.73	AT2G40010.1	Ribosomal protein L10 family protein	*A. thaliana*	33.65	5.05	193	2	7.57
Lipid metabolism										
17	20 DAP em	1.94	AT5G46290.1	3-Ketoacyl-acyl carrier protein synthase I	*A. thaliana*	50.38	9.25	896	16	29.81
271	20 DAP em	1.56	AAK60339	Biotin carboxylase	*B. napus*	58.38	6.57	1013	21	42.62
274	20 DAP em	1.56	AAK60339	Biotin carboxylase	*B. napus*	58.38	6.57	72	7	15.89
365	20 DAP em	1.73	P80030	Enoyl-acyl-carrier-protein reductase NADH, chloroplastic	*B. napus*	40.45	9.39	618	11	25.45
Nucleotide metabolism										
273	20 DAP em	2.04	AT2G41680.1	NADPH-dependent thioredoxin reductase C	*A. thaliana*	57.91	6.32	178	6	11.72
352	20 DAP em	1.58	ABE87611	Phosphoribosyltransferase; Orotidine 5-phosphate decarboxylase	*M. truncatula*	51.24	7.02	183	2	5.33
352	20 DAP em	1.58	AT3G54470.1	Uridine 5′-monophosphate synthase	*A. thaliana*	51.82	6.94	288	4	7.56
Uncharacterized										
225	20 DAP em	2.11	AAY86774	C2 domain-containing protein	*N. caerulescens*	18.48	5.15	50	1	6.06
56	20 DAP em	1.71	AAG00253	F1N21.10	*A. thaliana*	40.00	7.65	82	1	3.08
241	20 DAP em	3.95	XP_001753354	Predicted protein	*P. patens*	18.79	9.99	54	1	5.10

[a] Spots in accordance with Supplementary Fig. 9.
[b] Developmental stage of endosperm (es) and embryo (em) at 20 DAP.
[c] Regulation of spots, spot volume at least 1.5 fold increased in the corresponding developmental stage.
[d] Accession numbers as given by SwissProt (http://www.uniprot.org), NCBInr (http://www.ncbi.nlm.nih.gov/protein) and TAIR (http://www.Arabidopsis.org/).
[e] Calculated molecular masses of the identified proteins as deduced from the corresponding genes.
[f] Calculated isoelectric point of the identified proteins as deduced from the corresponding genes.
[g] Probability score for the protein identifications based on MS/MS analysis and MASCOT search.
[h] Number of unique matching peptides.
[i] Sequence coverage of a protein by identified peptides.

Publication 2

2.2 The native structure and composition of the cruciferin complex in *Brassica napus*

Thomas Nietzel[1], Natalya V. Dudkina[2], Christin Haase[1], Peter Denolf[3], Dmitry A. Semchonok[2], Egbert J. Boekema[2], Stephanie Sunderhaus[1], Hans-Peter Braun[1]

[1]Department of Plant Proteomics, Institute for Plant Genetics, Faculty of Natural Sciences, Leibniz Universität Hannover
[2]Electron Microscopy Group, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen
[3]Bayer CropScience NV, BioScience - Oilseeds Research, Zwijnaarde

Type of authorship:	Co-author
Type of article:	Research article
Share of the work:	20 %
Contribution to the publication:	Performed western blotting experiments and mass spectrometry, analysed data, prepared figures and tables
Journal:	Journal of Biological Chemistry
Impact factor:	4.573
Date of publication:	Published in January 2013
Number of citations: (Google scholar, Nov. 07th, 2015)	2
DOI:	10.1074/jbc.M112.356089
PubMed-ID:	23192340

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 288, NO. 4, pp. 2238–2245, January 25, 2013

The Native Structure and Composition of the Cruciferin Complex in *Brassica napus*

Received for publication, February 24, 2012, and in revised form, November 19, 2012 Published, JBC Papers in Press, November 28, 2012, DOI 10.1074/jbc.M112.356089

Thomas Nietzel[‡], Natalya V. Dudkina[§], Christin Haase[‡], Peter Denolf[¶], Dmitry A. Semchonok[§1], Egbert J. Boekema[§], Hans-Peter Braun[‡2], and Stephanie Sunderhaus[‡3]

From the [‡]Department of Plant Proteomics, Institute for Plant Genetics, Faculty of Natural Sciences, Leibniz University of Hannover, Herrenhäuser Strasse 2, 30419 Hannover, Germany, the [§]Electron Microscopy Group, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Nijenborgh 47, 9747 AG Groningen, The Netherlands, and [¶]BioScience, Oilseeds Research, Bayer CropScience NV, Technologiepark 38, 9052 Zwijnaarde, Belgium

Background: Cruciferin represents the most abundant protein in *Brassica napus* seeds where its efficient storage is essential under minimized space conditions.
Results: The cruciferin complex has an octameric barrel-like structure of ~420 kDa.
Conclusion: The barrel-like structure represents a compact building block optimized for maximal storage of amino acids.
Significance: Novel insights into structure and packing of seed storage proteins.

Globulins are an important group of seed storage proteins in dicotyledonous plants. They are synthesized during seed development, assembled into very compact protein complexes, and finally stored in protein storage vacuoles (PSVs). Here, we report a proteomic investigation on the native composition and structure of cruciferin, the 12 S globulin of *Brassica napus*. PSVs were directly purified from mature seeds by differential centrifugations. Upon analyses by blue native (BN) PAGE, two major types of cruciferin complexes of ~ 300–390 kDa and of ~470 kDa are resolved. Analyses by two-dimensional BN/SDS-PAGE revealed that both types of complexes are composed of several copies of the cruciferin α and β polypeptide chains, which are present in various isoforms. Protein analyses by two-dimensional isoelectric focusing (IEF)/SDS-PAGE not only revealed different α and β isoforms but also several further versions of the two polypeptide chains that most likely differ with respect to posttranslational modifications. Overall, more than 30 distinct forms of cruciferin were identified by mass spectrometry. To obtain insights into the structure of the cruciferin holocomplex, a native PSV fraction was analyzed by single particle electron microscopy. More than 20,000 images were collected, classified, and used for the calculation of detailed projection maps of the complex. In contrast to previous reports on globulin structure in other plant species, the cruciferin complex of *Brassica napus* has an octameric barrel-like structure, which represents a very compact building block optimized for maximal storage of amino acids within minimal space.

Seed storage proteins accumulate during the seed filling process and serve as a source of nitrogen and amino acids for the germinating embryo. In dicotyledonous plants, *e.g.* sunflower, soybean, *Arabidopsis*, and rapeseed, the most abundant storage proteins are the 2 S albumins and 7 S and 12 S globulins, the latter ones representing the phylogenetically most widely distributed group of storage proteins. Mature 12S globulins consist of two polypeptide chains, which are covalently linked by a disulfide bond. Both chains stem from the same precursor molecule, which is co-translationally synthesized into the rough endoplasmic reticulum. During import, the endoplasmic reticulum signal peptide is cleaved off and a disulfide bond is formed between the N- and C-terminal parts of the polypeptide chain (1, 2). These proglobulins are then assembled into trimers (3), transported through the Golgi apparatus, and finally stored as mature hexamers in protein storage vacuoles (PSVs).[4] In this process, at least a second cleavage of the proglobulin occurs, which forms the mature protein with α and β polypeptide chains. The latter step is a prerequisite for hexamer assembly (4, 5). The site of this second cleavage is between an asparagine and glycine residue and both amino acids are highly conserved among the group of the 12 S globulins (5, 6). Several proteases, so-called vacuolar processing enzymes, which belong to a novel family of cysteine proteinases, are responsible for this second cleavage of the monomers (7). Recently, it was shown that vacuolar processing enzymes as well as the corresponding storage proteins are sorted in the Golgi apparatus simultaneously in two different types of vesicles. Subsequently, these vesicles fuse to form a prevacuolar compartment, the multivesicular body in which the proteolytic process takes place (8, 9). Later, multivesicular bodies fuse into a PSV. Most members of the 12 S globulins are believed to undergo these two events of posttranslational cleavage (4, 6, 7), and the resulting polypeptide chains are referred to as the α polypeptide chain (the original N-terminal

[S] This article contains supplemental Table 1 and Figs. 1 and 2.
[1] Supported by HARVEST Marie Curie Research Training Network Grant PITN-GA-2009-238017.
[2] To whom correspondence may be addressed: Dept. of Plant Proteomics, Institute for Plant Genetics, Faculty of Natural Sciences, Leibniz University of Hannover, Herrenhäuser Str. 2, 30419 Hannover, Germany. Tel.: 49-511-762-2674; Fax: 49-511-762-3608; E-mail: braun@genetik.uni-hannover.de.
[3] To whom correspondence may be addressed: Dept. of Plant Proteomics, Institute for Plant Genetics, Faculty of Natural Sciences, Leibniz University of Hannover, Herrenhäuser Str. 2, 30419 Hannover, Germany. Tel.: 49-511-762-5290; Fax: 49-511-762-3608; E-mail: sunderhaus@genetik.uni-hannover.de.

[4] The abbreviations used are: PSV, protein storage vacuole; BN, blue native; AA, amino acid(s); DAP, days after pollination; BN PAGE, blue native PAGE; IEF, isoelectric focusing.

part of the precursor protein) and the β polypeptide chain (the original C-terminal part).

Also in *Brassica napus*, a 12 S globulin, cruciferin, is the major storage protein and accounts for ~60% of the total seed protein (10). Five different *B. napus* cruciferins belong to three different families, which are listed in the reviewed UniProtKB database. The P1 family contains only one protein, CRU1, with a sequence length of 509 amino acids (AA) (α_1 = 296 AA and β_1 = 190 AA). In contrast, the P2 family contains three proteins: BnC1 with a sequence length of 490 AA (α_2 = 277 AA and β_2 = 190 AA); BnC2 with a sequence length of 496 AA (α_3 = 283 AA and β_3 = 190 AA) and CRU2/3 with a sequence length of 488 AA ($\alpha_{2/3}$ = 275 and $\beta_{2/3}$ = 190 AA). The P3 family contains one protein, CRU4 with a sequence length of 465 AA (α_4 = 254 AA and β_4 = 189 AA). Similar to other 12S globulins, all cruciferin family members are secreted into the endoplasmic reticulum, and all exhibit the typical conserved 12S globulin cleavage site between asparagine and glycine. They are thought to be transported as trimers via the Golgi apparatus and stored as hexamers in PSVs (11–15).

Currently, little is known about the structure and organization of *B. napus* cruciferin within the PSV, although this most likely is important for ensuring efficient use of storage space in developing seeds. The first reports on the quaternary structure of a 12S globulin were given by Badley *et al.* (16) in 1975. Using electron microscopy, it was shown that the soybean 12S globulin glycinin extracted from soy flour is composed of two hexagonal rings with alternating α and β polypeptide chains. This dodecameric model was also supported by Marcone *et al.* (17) for amaranth globulin isolated from defatted flour. However, for rapeseed and sunflower globulin a triagonal antiprism of six subunits was suggested (18). In subsequent x-ray crystallographic studies, soybean proglycinin and rapeseed procruciferin overexpressed in *Escherichia coli* were reported to exhibit triangular patterns (19–21). The crystal structure of a mature soybean glycinin extracted from a mutant soybean cultivar indicates a hexamic particle formed by two face-to-face stacking trimers (22). Two conclusions can be drawn from these results: (i) quaternary structures of 12S globulins may vary between species, and (ii) the quaternary structure of 12S globulins also can differ for the same species, depending on the expression system used for globulin biosynthesis (*e.g.* soybean *versus E. coli*). Consequently, detailed structural information of native cruciferin is best obtained directly from the plant species of interest without the use of heterologous expression systems.

In this study, *B. napus* cruciferin complexes were directly isolated from mature seeds under native conditions. The resulting fractions were highly enriched with cruciferin. Purified proteins were used for investigations by various gel electrophoresis systems, mass spectrometry and single particle EM. A large variation with respect to isoelectric points and molecular mass was observed for the cruciferin α and β polypeptide chains. In contrast to 12S globulin complexes from other plant species, the cruciferin complex exhibits a unique octameric structure that allows packaging of the amino acids under most space-saving conditions.

EXPERIMENTAL PROCEDURES

Cultivation of B. napus—B. napus plants were cultivated in growth chambers under the following conditions: 16-h light (16 klux) at 22 °C, 8-h dark at 18 °C with a relative humidity of 55%. Seeds were harvested 53 to 60 days after pollination (DAP) and directly used for the isolation of PSVs.

Isolation of Protein Storage Vacuoles—Native proteins from PSVs were isolated as described previously (23). All steps were carried out at 4 °C or on ice. 14 g of mature seeds were harvested and directly homogenized in 100 ml of glycerol (100% (v/v)) using a mortar and a pestle. The homogenate was filtered through one layer of Miracloth (20–25 μm, Calbiochem) and centrifuged for 10 min at 1100 × *g*. To collect PSVs, the supernatant was centrifuged for 20 min at 41,000 × *g*. The pellet was again resuspended in glycerol (30 ml, 100% (v/v)) and centrifuged for 20 min at 41,000 × *g*. This pellet contained the PSV fraction and was further resuspended in 10 ml of TE buffer (1 mM EDTA, 5 mM Tris-HCl, pH 8.5) to disrupt the PSV and loaded onto a discontinuous sucrose step gradient (30, 45, and 68% (w/v) sucrose in TE buffer) for 2 h at 78,000 × *g*. The fraction above the 30% sucrose layer represents the vacuolar matrix containing the storage proteins and was directly used or frozen in liquid nitrogen and stored at −80 °C.

Gel Electrophoresis Procedures—Proteins were analyzed by one-dimensional blue native polyacrylamide gel electrophoresis (BN PAGE) and by two-dimensional BN/SDS-PAGE using 100 μl of the frozen PSV matrix fraction supplemented with 5 μl of "blue loading buffer" (5% (w/v) Coomassie Blue in 750 mM aminocaproic acid) according to the protocol outlined in Wittig *et al.* (24). Mitochondria from *Arabidopsis thaliana* were prepared as outlined in Sunderhaus *et al.* (25) and solubilized by 5 g/g digitonin. One-dimensional SDS-PAGE was carried out as described by Schägger and von Jagow (26) using 30 μl of the frozen PSV matrix fraction. For two-dimensional isoelectric focusing/SDS-PAGE (IEF/SDS-PAGE), 50 μl of the PSV fraction were mixed with 300 μl of resuspension buffer (8 M urea, 2 M thiourea, 50 mM DTT, 2% CHAPS (w/v), 5% IPG buffer 3–11 nl (v/v), 12 μl/ml DeStreak reagent, bromphenol blue) and directly applied into a strip holder. Isoelectric focusing was carried out with the Ettan IPGphor 3 apparatus (GE Healthcare) using 3–11 nl of Immobiline DryStrip gels (18 cm). Rehydration took place at 30 V for 12 h and focusing during 4 steps at 500 V (1 h), 500–1000 V (1 h), 1000–8000 V (3 h), and 8000 V (6 h). Afterward, strips were equilibrated for 15 min in equilibration solution (6 M urea, 30% glycerol (87%, v/v), 2% SDS, 50 mM Tris-HCl, pH 8.8, bromphenol blue) with (i) 1% DTT (w/v) and (ii) 2.5% iodacetamide (w/v). IPG strips were finally transferred horizontally onto a 16.5% tricine gel, and electrophoresis was carried out for 20 h at 35 mA/mm gel layer.

Gel Staining Procedures—All polyacrylamide gels were stained with Coomassie colloidal (27, 28).

Generation of Antibodies against Cruciferin—Two different polyclonal antibodies directed against peptides of the α or β polypeptide chains of *B. napus* cruciferin were generated (Eurogentec S.A. Antisera Production, Seraing, Belgium). For details, see supplemental Fig. 1.

Native Structure and Composition of Cruciferin in B. napus

Western Blotting—Proteins separated on polyacrylamide gels were blotted onto nitrocellulose membranes for antibody staining using the Trans Blot Cell from Bio-Rad. Blotting was carried out as described in Kruft *et al.* (29). Immunostainings were performed using the VectaStain ABC kit (Vector Laboratories, Burlingame, CA).

MS Analyses—Tryptic digestion of proteins and peptide extraction were carried out as published in Sunderhaus *et al.* (25). MS analyses were performed using an EASY-nLC-system (Proxeon) coupled to a MicrOTOF-Q-II mass spectrometer (Bruker Daltonics). Identification of proteins was carried out using the MASCOT search algorithm against (i) SwissProt (ii) NCBI non-redundant protein database and (iii) The *Arabidopsis* Information Resource (Tair release 9).

Electron Microscopy and Image Analyses—Fractions containing cruciferin were applied to carbon-coated copper grids and negatively stained with 2% uranyl acetate. Electron microscopy was performed on a Philips CM120 electron microscope. Data were collected with a 4K slow-scan CCD camera (Gatan) at a magnification of 130,000 with a pixel size (after binning the images) of 0.23 nm at the specimen level. Single particle analysis was performed with the Groningen Image Processing (GRIP) software package on a CPU cluster as outlined by Dudkina *et al.* (30). The three-dimensional model of cruciferin was created with a Blender three-dimensional creation suite. X-ray structure of procruciferin from *B. napus* (Protein Data Bank code 3KGL (20)) was used for the superimposing on projection maps of cruciferin. The octamer of cruciferin was generated from the hexameric x-ray structure of procruciferin maintaining monomer-monomer interfaces by using the Chimera program from the University of California, San Francisco (31).

RESULTS

PSVs were isolated from freshly harvested seeds at 53 to 60 DAP under most gentle conditions using the glycerol protocol published by Jiang *et al.* (23). The total protein fraction was directly extracted from PSVs.

Characterization of Cruciferin Complexes from B. napus Storage Vacuoles—BN PAGE was originally developed for the characterization of native mitochondrial protein complexes but also proved to be a very suitable procedure for the investigation of protein complexes from other cellular compartments (24). To characterize the protein composition of the PSVs, their soluble fraction was directly supplemented with Coomassie blue and subsequently loaded onto a one-dimensional BN gel. As a control and molecular mass standard, respiratory protein complexes of *A. thaliana* mitochondria were solubilized by 5 g/g digitonin and separated on the same gel (Fig. 1). Three bands are visible in the lane of the PSV fraction: a central broad band sharply confined at its lower border at 300 kDa with a more diffuse upper border at ~390 kDa is flanked by two bands of lower abundances at ~150 and 470 kDa. Judging from the molecular masses of the bands, it can be expected that they represent multisubunit protein complexes.

To investigate the subunit compositions of the PSV protein complexes, two-dimensional BN/SDS-PAGE was carried out for separating the subunits of the resolved complexes (Fig. 2*A*). Under the denaturing conditions of the second gel dimension,

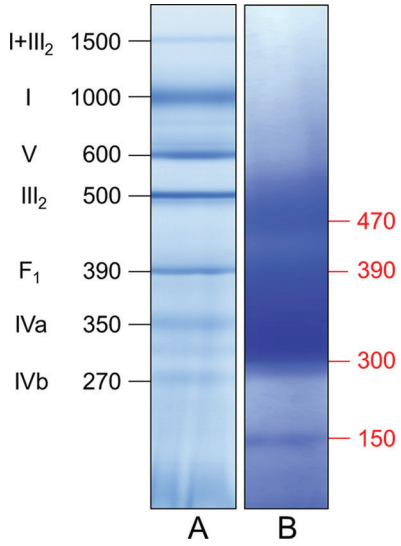


FIGURE 1. **Separation of soluble proteins from PSVs by BN PAGE indicates presence of three major classes of protein complexes of 150, 300–390, and 470 kDa.** The proteins of storage vacuoles of *B. napus* (*lane B*) and of a mitochondrial fraction of *A. thaliana* (molecular mass standard; *lane A*) were separated on a native gradient gel according to molecular mass. Molecular masses are given in kDa. Designations of *A. thaliana* mitochondrial protein complexes are given as follows: $I+III_2$, supercomplex containing NADH dehydrogenase and dimeric cytochrome *c* reductase; *I*, NADH dehydrogenase; *V*, F_0F_1 ATP synthase; III_2, dimeric cytochrome *c* reductase; F_1, F_1 part of the ATP synthase; *IVa* and *IVb*, larger and smaller form of cytochrome *c* oxidase.

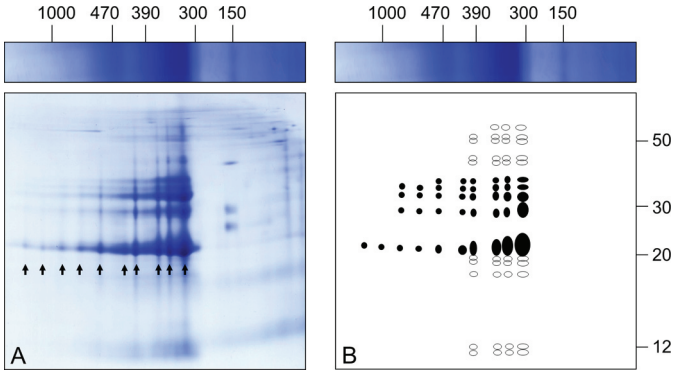


FIGURE 2. Separation of proteins isolated from PSVs by two-dimensional BN/SDS-PAGE (*A*) suggests specified associations of cruciferin. *Arrows* indicate the occurrence of 10 defined cruciferin complexes with a maximum of 15 different protein species (given on the scheme in *B*). Highly abundant subforms are indicated in *filled circles*, less abundant subforms are in *open circles*. Molecular masses are given in kDa *above* and to the *right* of the gel.

the smaller 150-kDa complex dissociated into three subunits of ~40, 25, and 20 kDa. The two larger complexes dissociated into ~15 different subunits of 60 to 12 kDa. The protein patterns of both bands indicate that the two complexes have the same composition but differ in subunit stoichiometry. In contrast to the one-dimensional BN PAGE, even larger forms of this complex are visible on the two-dimensional gel, which is schematically summarized in Fig. 2*B*.

To verify that separated fractions contain cruciferin, peptide-specific antibodies were generated against the α and β polypeptide chain of the protein (see supplemental Fig. 1). Two-dimensional BN/SDS gels were then blotted onto nitrocellulose membranes and incubated with antibodies directed against both chains (Fig. 3).

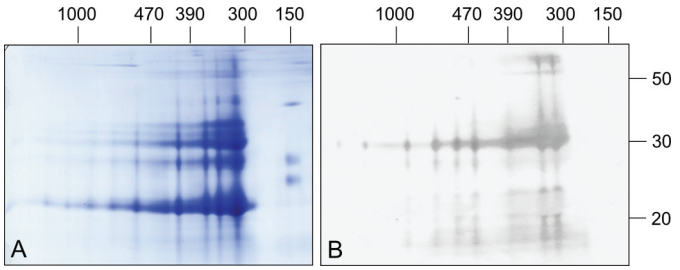


FIGURE 3. **Immunoblotting analyses of a *B. napus* PSV fraction separated by two-dimensional BN/SDS-PAGE reveals high enrichment of cruciferin.** Gels were either stained by Coomassie colloidal (*A*) or blotted onto nitrocellulose (*B*). Western blots were incubated with antibodies directed against the α chain of cruciferin. Molecular masses of standard proteins are shown *above* and to the *right* of the gel (in kDa).

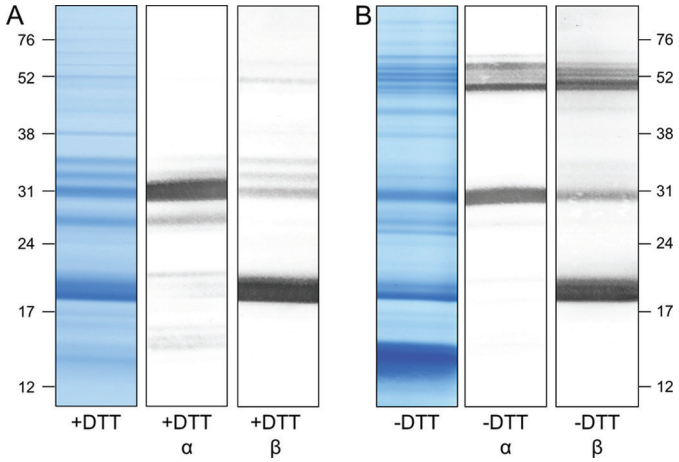


FIGURE 4. **Immunoblotting analyses of a *B. napus* PSV fraction separated by one-dimensional SDS-PAGE reveals cruciferin complexes of ~48 to 58 kDa under non-reducing conditions.** PSV fractions were separated under reducing conditions (*A*; +DTT) and non-reducing conditions (*B*; −DTT). Gels were either stained Coomassie colloidal or blotted onto nitrocellulose membranes and incubated with antibodies directed against the α or the β chains of cruciferin as indicated. Molecular masses of standard proteins are shown to the *right* and to the *left* in kDa.

The obtained results illustrate the purity of the PSV fraction. A strong signal is present in the 30-kDa range, representing α polypeptide chains from the different cruciferin families. Additionally, a strong signal is visible in the 55-kDa range indicating that few cruciferins still have an intact disulfide bond between both polypeptide chains after the β-mercaptoethanol treatment preceding the second dimension. Some weaker signals in the molecular mass range below 20 kDa most likely represent α polypeptide chain breakdown products. Blots developed with antibodies against the β polypeptide chain of cruciferin show a strong signal in the 20 to 25 kDa range as well as in the 60 kDa range (data not shown). The two-dimensional gel shows that the broad band observed in the one-dimensional gel is actually composed of several individual bands and that cruciferin is present in all those bands, also in the 470-kDa band. Additionally, cruciferin at least partially occurs in larger forms up to 1500 kDa on one-dimensional BN gels but does not form part of the 150-kDa complex.

Analyses of Cruciferin α and β Polypeptide Chains at Reducing/Non-reducing Conditions—To investigate the size of the monomeric cruciferin complex as well as its α and β polypeptide chains, cruciferin isolated from PSVs was dissolved in either denaturating buffer containing DTT to reduce the disulfide bond between the N and the C terminus or in buffer without DTT to keep the protein in its native monomeric form. Fractions were separated using one-dimensional SDS gels and blotted onto nitrocellulose membranes, which were incubated with antibodies against either the α or the β chain (Fig. 4).

Under denaturing conditions, antibodies directed against the α polypeptide chain produced a strong signal visible in the 30-kDa range. Some additional but much weaker signals are observable in the range of 27 and 33 kDa and in the low molecular mass range between 14 and 20 kDa. No signals could be detected above 35 kDa, indicating that all cruciferin is cleaved into α and β polypeptide chains by this treatment. In addition, incubation with antibodies directed against the β polypeptide chain produced a strong signal at 18 to 20 kDa as well as some minor bands in the range of 30–35 and at ~50 kDa (Fig. 4*A*). Under non-reducing conditions both antibodies gave the same signals in the 27–33 kDa and 18–20 kDa range, respectively, but with notably reduced intensities. Additionally, a number of new signals are visible in the range of ~52 kDa (Fig. 4*B*). Because these proteins are detected by the antibodies directed against the α and the β polypeptide chains, these bands most likely represent cruciferins with intact disulfide bonds.

Because isoforms of cruciferin precursor proteins as well as their respective α and β polypeptide chains cannot be separated on one-dimensional SDS gels due to similar molecular masses, the PSV fraction was analyzed by two-dimensional IEF/SDS-PAGE (Fig. 5). IEF is often capable of separating proteins with the same or similar molecular masses due to different occurrences of ionizable amino acids within isoforms.

As shown in Fig. 5 and by Western blotting (data not shown), α polypeptide chains are separated in a range between 27 and 36 kDa and isoelectric points between 6.7 and 8.8 (*red box*). The corresponding β polypeptide chains are separated in a range between 19 and 21 kDa and isoelectric points between 5.9 and 9.5 (*green box*).

For assigning cruciferin polypeptide chains from the three individual cruciferin families to each detected spot/band on our one-dimensional BN (Fig. 1), two-dimensional BN/SDS (Fig. 3), one-dimensional SDS (Fig. 4), and two-dimensional IEF/SDS (Fig. 5) gels, protein identifications by mass spectrometry were carried out. A total number of 95 spots, representing the most abundant proteins, were picked from the four gel systems and analyzed by LC-ESI-Q-TOF-MS/MS (supplemental Fig. 2 and supplemental Table 1). Members of all five cruciferin families were detected as well as the 2 S storage protein Napin. The latter one is also known to accumulate during the seed filling process in *B. napus* PSVs. Furthermore, five proteins were identified, which do not belong to the family of storage proteins.

Altogether, mass spectrometry enabled the assignment of spots on two-dimensional IEF/SDS gels to distinct cruciferin families and polypeptide chains (Fig. 6). With respect to α polypeptide chains, four spots in the range of 34-kDa spots were identified as αCRU1 (P1); another four spots in the range of 32 kDa were assigned to αBnC1, αBnC2, and αCRU2/3 (P2); and two spots in the range of 27 kDa represent αCRU4 (P3). All six

Native Structure and Composition of Cruciferin in B. napus

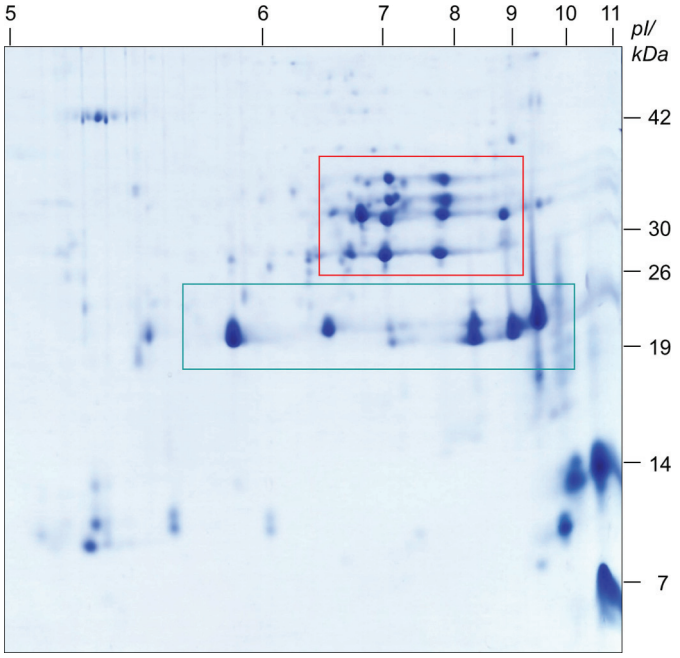


FIGURE 5. **Resolving a PSV fraction according to isoelectric point and molecular mass reveals existence of cruciferin α chains in a pH range from 6.7 to 8.8 (*red box*) and β-chains in a pH range from 5.9 to 9.5 (*green box*).** Molecular masses of standard proteins as well as the pI range are given on the *right* and *above* the gel, respectively.

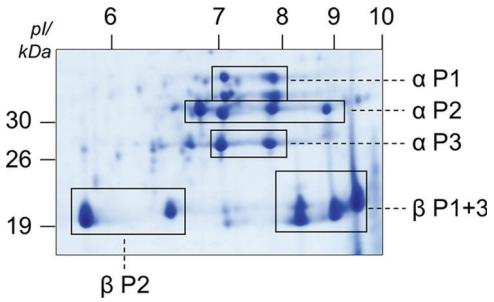


FIGURE 6. **Analyses of cruciferin subunits by LC-MS/MS allows assigning clusters of spots on two-dimensional IEF/SDS gels to the three distinct cruciferin families.** Names of families and polypeptide chain attribution are given to the *right* and *below* the gel. Molecular masses and pIs are given to the *left* and *above* the gel, respectively. *P1*, family 1 containing cruciferin Cru 1; *P2*, family 2 containing cruciferin BnC1, BnC2, and Cru 2/3; *P3*, family 3 containing cruciferin Cru 4.

spots in the range of 19 to 21 kDa were distinguishable exclusively by their isoelectric points, and they all represent β polypeptide chains. Four spots exhibiting isoelectric points of ~8 to 9.5 were identified as βCRU1 and βCRU4 (P1 and P3), and two spots exhibiting isoelectric points from ~5.8 to 6.8 were assigned to βBnC1, βBnC2, and βCRU2/3 (P2).

The Cruciferins of Rapeseed Are Phosphorylated—Phosphorylation of cruciferins was investigated using antibodies directed against phosphoserine, phosphothreonine, and phosphotyrosine. For this approach, storage vacuole protein fractions were separated by two-dimensional IEF/SDS-PAGE and blotted onto filter membranes (Fig. 7). Several cruciferin forms belonging to both the α and the β chains are recognized by all three sera. Overall, the α chain of cruciferin reacts stronger with the sera than the β chain. The strongest immune signals were obtained with the serum directed against phosphoserine (Fig. 7*D*).

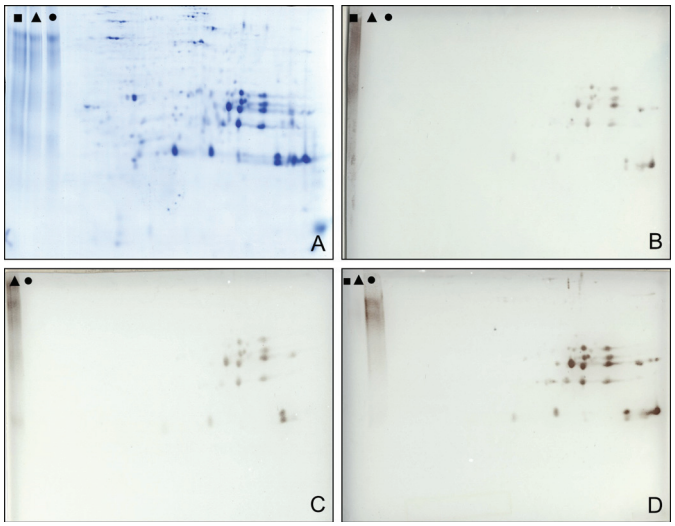


FIGURE 7. **Immunological analysis of cruciferin phosphorylation.** Protein storage vacuole fractions were resolved by two-dimensional IEF/SDS-PAGE and blotted onto membranes, and blots were developed using antibodies directed against three different phosphorylation sites. As a control, phosphorylated BSA (phosphorylated at tyrosine (■); phosphorylated at threonine (▲); phosphorylated at serine (●)) was separated within the same gel. *A*, Coomassie-stained reference gel. *B*, immunoblot developed with an IgG directed against phosphorylated tyrosine; *C*, immunoblot developed with an IgG directed against phosphorylated threonine; *D*, immunoblot developed with an IgG directed against phosphorylated serine. Note that the phosphotyrosine-BSA control on *C* is not indicated because it partially was beyond the limit of the blot. However, the specificity of the IgG was verified by independent experiments (not shown).

Structure of the Native Cruciferin Storage Form in Mature PSVs—For a structural analysis of the native cruciferin oligomers, PSV fractions purified from developing (53 DAP) or mature seeds (60 DAP) were examined and analyzed by single particle EM (Fig. 8). Altogether, 23,000 protein projections were analyzed to obtain final averages at ~20 Å resolution after several iterations of alignment and classification. As can be seen from class averages, the majority of molecules appeared to be oriented in a top or bottom view positions (Fig. 8, *a*, *b*, *d*, *e*, *f*, *h*, *i*, *j*, and *l*), whereas only a smaller group of projections represented side views (Fig. 8, *c*, *g*, and *k*). The top and bottom views clearly demonstrate a 4-fold symmetry. By combining the information obtained on the top and side views, it can be concluded that the structure is composed of two layers, each consisting of four identical subunits in a tetrameric configuration enclosing a pore in the center of the complex. A groove divides the side-view projection into two halves, each of which appears to be mirror the other one in the projection. This indicates that the two sets of tetramers are facing each other in antiparallel positions.

From the single particle EM data shown in Fig. 8, we generated a model of the cruciferin storage complex (Fig. 8*m*). It is composed of eight subunits arranged in an octameric configuration rather than the hexamer, which was proposed by other studies for globulins from other organisms (11–15, 20, 22). From the dimensions observed in the projection maps, we estimated the mass of the octamer to be ~420 kDa. Estimation was done by using a formula for the cylinder-shaped complex as $m = \Pi R^2 h/\rho$, where m is the mass of the protein, R and h are the radius and the height of the cylinder, and ρ is the average pro-

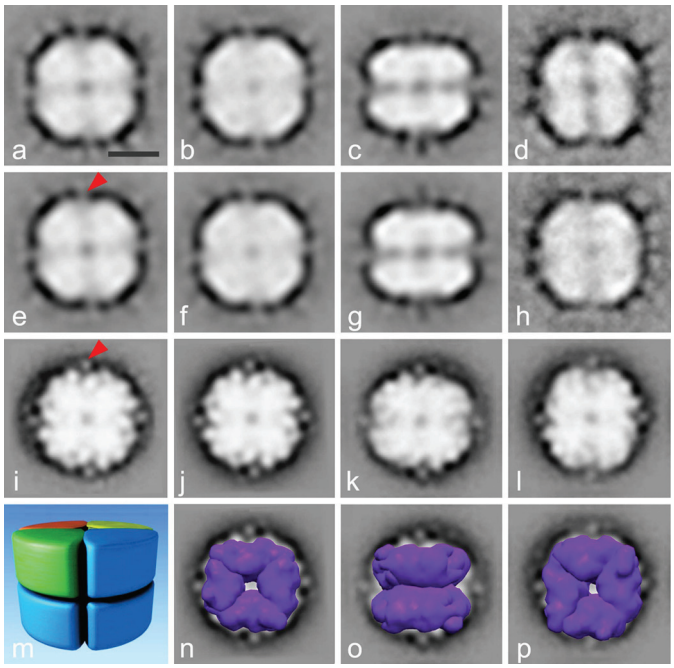


FIGURE 8. **Structural analyses of the cruciferin complex by single particle electron microscopy.** *Top* view (*a*, *b*, *d–f*, and *h*) and *side* view (*c* and *g*) projection maps of negatively stained cruciferin complexes isolated 53 DAP and *top* view (*i*, *j*, *l*) and *side* view (*k*) projections of cruciferin isolated 60 DAP. *Panels e*, *f*, *g*, and *j* demonstrate the views of *a*, *b*, *c*, and *i* correspondingly after symmetrization. *a* represents a sum of 705 projections; *b*, 582 projections; *c*, 372 projections, *d*, 575 projections; *h* projections, 335 projections; *i*, 5000 projections; *k*, 5000 projections; *l*, 5000 projections. *m* shows the model of the octameric cruciferin complex based on the negatively stained projection maps. *n–p* demonstrate the superimposing of procruciferin density maps (*purple*) on projection maps from *j–l* correspondingly. The *bar* is 5 nm. The *red arrows* point to additional mass in every monomer in our projections, which is not visible in the monomers of the hexameric cruciferin complex solved x-ray diffraction.

tein density, which is 1.5 g/cm^3 (32). Thus, each monomer, composed of an α and β polypeptide chain, should have a molecular mass of ~50 kDa. The obtained model is in agreement with the observations from electrophoretic analyses showing average molecular masses of α and β chains of ~32 kDa and 20 kDa, respectively. To interpret the projection maps, we generated the octamer of cruciferin from the x-ray structure of procruciferin and superimposed the density map of the octamer on our projection maps (Fig. 8, *n–p*). Importantly, by fitting the crystal structure, we gain a resolution by a factor of 5 (33). Based on the close fit, we conclude that the analyzed projections indeed represent the top and side views of the cruciferin complex in an octameric configuration. There is an additional mass in every monomer in our projections, seen as spikes (Fig. 8, *e* and *i*, *red arrows*), which do not belong to the monomers of the hexameric cruciferin complex solved by x-ray diffraction. This could be a result of incompleteness of the hexamer structure, which has missing residues as seen from the deposited atomic coordinates. Alternatively, crystallization conditions might cause changes in the protein configuration or subunit arrangements.

DISCUSSION

In this study, the structural properties of cruciferin proteins as well as the supramolecular structure of the complex formed by cruciferin subunits is investigated. In previous studies, storage protein complexes were analyzed *e.g.* by the expression of the proteins in a heterologous expression system, which allowed purification of the complexes to a level of purity rendering them amenable for crystallographic structural analysis. Here, we chose a different approach. Protein complexes were purified directly from freshly harvested seeds and analyzed using gel electrophoretic approaches, mass spectrometry, and single particle electron microscopy. Although the latter method is compromised in the resolution obtained at the structural level, it allows investigations of the protein from native environment. Differences in the results obtained by both approaches and their implications are discussed below.

Suitability of the Vacuolar Fraction for Biochemical and Structural Analyses—The purity of the isolated cruciferin fraction was demonstrated by the different gel electrophoresis procedures in combination with antibodies directed against cruciferin α and β polypeptide chains (Figs. 3 and 4) as well as by mass spectrometry. Results from the 95 analyzed spots/bands revealed that only five proteins do not belong to the family of storage proteins. These are as follows: glyceraldehyde-3-phosphate dehydrogenase, elongation factor 1-α, myrosinase binding protein, protein disulfide isomerase, and oleosin (supplemental Fig. 2 and supplemental Table 1). The first three proteins most likely represent cytosolic constituents, whereas the protein disulfide isomerase is reported to be an integral part of protein storage vacuoles (34). Oleosin is a component of oil bodies in *B. napus* seeds (35). The purity of the vacuolar fraction is additionally confirmed by electron microscopy. Negatively stained specimens were found to exclusively contain oligomeric Cruciferin, either in side or in top views. Therefore, the PSV fractions obtained are well suited for a detailed analysis of native *B. napus* cruciferin.

Heterogeneity of B. napus Cruciferin—The analyses of cruciferin and its polypeptide chains were carried out using different electrophoretic approaches. The results obtained by one-dimensional BN, one-dimensional SDS, two-dimensional BN/SDS, and two-dimensional IEF/SDS/PAGE, respective immunoblots, and mass spectrometry analyses suggest a strong diversity in physicochemical properties of the cruciferin polypeptide chains, proteins, and protein complexes (Table 1). In all gel systems, α polypeptide chains show molecular masses between 27 and 36 and β polypeptide chains between 19 and 21 kDa. Partially, molecular masses differ from the calculated ones in a range of 6 kDa, especially for α chains. Isoelectric points of some cruciferin polypeptides differ significantly compared with their calculated pIs in databases. In general, the β polypeptide chains show a very diverse pI range on IEF/SDS gels between 5.9 and 9.5. This is also reflected by the calculated pIs (6.1 and 8.6). Hence, β polypeptide chains can be considered to be rather basic. Surprisingly, α chains, which previously were observed to have an acidic pI, are found at higher pH values (36, 37). Their apparent isolelectric points on IEF/SDS gels are all between pH 6.8 and 8.6. This clearly indicates a neutral rather than an acidic pI (Table 1 and Fig. 6).

A certain level of heterogeneity was also observed on the level of the cruciferin holo-complex. The main band for cruciferin complexes on BN gels smeared over a range of nearly 100 kDa.

Native Structure and Composition of Cruciferin in B. napus

TABLE 1
Characteristics of *B. napus* cruciferins and their α and β polypeptide chains

	No AA α^a	No AA β^a	No AA $\alpha+\beta^a$	MM cal $\alpha+\beta^b$	MM cal α^b	MM app α^c	MM cal β^b	MM app β^c	pI cal α^b	pI app α^c	pI cal β^b	pI app β^c
Cru 1	296	190	486	54.1	32.9	31.0–36.0	21.2	19.0–21.0	7.17	6.8–7.8	6.94	8.3–9.5
BnC1	254	190	477	51.4	30.6	27.0–36.0	20.8	19.0–21.0	6.78	6.8–7.8	6.16	5.9–7.0
BnC2	283	190	473	51.8	30.8	31.0–32.0	21.0	19.0–21.0	7.95	7.1–7.8	8.6	8.3–8.9
Cru 2/3	275	190	465	51.3	30.5	27.0–32.0	20.8	19.0–21.0	6.78	7.1	6.16	5.9–7.0
Cru 4	254	189	443	48.9	28.1	27.0–32.0	20.9	19.0–21.0	6.66	6.7–7.0	8.6	8.3–9.5

[a] Number of amino acids, excluding the signal peptide, derived from the corresponding UniProt database entry.
[b] Calculated isoelectric points and calculated molecular masses, excluding the signal peptide, using the ExPASy MM/pI tool.
[c] Apparent isoelectric points and molecular masses upon analyses by two-dimensional IEF/SDS PAGE (Fig. 5).

This band includes protein complexes ranging from 300 to 390 kDa (Fig. 1), whereas the estimated mass of the cruciferin octamer based on EM analyses is ~420 kDa. The difference between the estimated mass of the complex and the one observed on BN gels is most likely caused by the barrel-like structure of the octamer. Indeed, the three-dimensional structure of cruciferin is considerably compact. In contrast, the structures of the respiratory complexes of *A. thaliana* mitochondria are much more irregular (complex I, for instance, has an L-like shape). Because the BN gel was calibrated by the masses of the respiratory complexes, the apparent molecular mass of the cruciferin complex is very likely to be underestimated. Apart from that, calculation of the mass from the EM projections does not take into account the presence of the central pore and space between the monomers, thereby providing slight overestimation for the mass of cruciferin.

Phosphorylation of Cruciferin—To explain pI variations on two-dimensional IEF/SDS gels, an immune blotting approach was carried out (Fig. 7). Western blots developed with antibodies directed against phosphoserine, phosphothreonine, and phosphotyrosine indicate stable phosphorylations on several but not all cruciferin forms separated by IEF/SDS-PAGE (Fig. 7). Similar results were previously reported for *B. napus* and *A. thaliana* (39–41). In fact, it was shown that cruciferin is one of the most phosphorylated proteins in *A. thaliana* seeds. In contrast, our MS analyses did not reveal direct evidence for cruciferin phosphorylation in *B. napus*. This negative result probably is due to the fact that phosphopeptides were not enriched during sample preparation for MS (*e.g.* by Immobilized metal ion affinity chromatography (IMAC) or TiO_2 affinity chromatography). Most likely, phosphorylation sites in cruciferins are substoichiometrically phosphorylated. As a consequence, non-phosphorylated peptides dominated the mass spectra.

The Supramolecular Structure of Directly Isolated Cruciferin—By combining a mild purification procedure with analyses by BN PAGE and single particle EM, we were able to structurally characterize cruciferin oligomers from mature seeds. Because proteins for crystallization experiments often are produced by overexpression in heterologous systems, mainly *E. coli*, several factors ensuring a correct assembly may be missing, *e.g.* signal peptide cleavage, the proteolytic cleavage of the α and β polypeptide chains and their assembly in the endoplasmic reticulum as well as the following formation into oligomers. Furthermore, many results on seed storage proteins are obtained using defatted flours for industrial purposes or NaCl extraction, which may strongly disturb the native conformation of the proteins (17, 18, 42–44). Therefore, further analyses should be carried out to verify hexameric globulin structure in seeds as reported before for other plants. Finally, globulin complexes from additional species should be characterized to obtain more general insights into their oligomeric organization.

Acknowledgment—We thank Dr. Holger Eubel for critically reading the manuscript.

REFERENCES

1. Ereken-Tumer, N., Richter, J. D., and Nielsen, N. C. (1982) Structural characterization of the glycinin precursors. *J. Biol. Chem.* **257,** 4016–4018
2. Sengupta, C., Deluca, V., Bailey, D. S., and Verma, D. P. (1981) Post-translational processing of 7S and 11S components of soybean storage proteins. *Plant Mol. Biol.* **1,** 19–34
3. Chrispeels, M. J., Higgins, T. J., and Spencer, D. (1982) Assembly of storage protein oligomers in the endoplasmic reticulum and processing of the polypeptides in the protein bodies of developing pea cotyledons. *J. Cell Biol.* **93,** 306–313
4. Dickinson, C. D., Hussein, E. H., and Nielsen, N. C. (1989) Role of post-translational cleavage in glycinin assembly. *Plant Cell* **1,** 459–469
5. Jung, R., Scott, M. P., Nam, Y. W., Beaman, T. W., Bassüner, R., Saalbach, I., Müntz, K., and Nielsen, N. C. (1998) The role of proteolysis in the processing and assembly of 11S seed globulins. *Plant Cell* **10,** 343–357
6. Müntz, K. (1996) Proteases and proteolytic cleavage of storage proteins in developing and germinating dicotyledonous seeds. *J. Exp. Bot.* **47,** 605–622
7. Shimada, T., Yamada, K., Kataoka, M., Nakaune, S., Koumoto, Y., Kuroyanagi, M., Tabata, S., Kato, T., Shinozaki, K., Seki, M., Kobayashi, M., Kondo, M., Nishimura, M., and Hara-Nishimura, I. (2003) Vacuolar processing enzymes are essential for proper processing of seed storage proteins in *Arabidopsis thaliana*. *J. Biol. Chem.* **278,** 32292–32299
8. Otegui, M. S., Herder, R., Schulze, J., Jung, R., and Staehelin, L. A. (2006) The proteolytic processing of seed storage proteins in *Arabidopsis* embryo cells starts in the multivesicular bodies. *Plant Cell* **18,** 2567–2581
9. Vitale, A., and Hinz, G. (2005) Sorting of proteins to storage vacuoles: how many mechanisms? *Trends Plant Sci.* **10,** 316–323
10. Crouch, M. L., and Sussex, I. M. (1981) Development and storage-protein synthesis in *Brassica napus* L. embryos *in vivo* and *in vitro*. *Planta* (*Berl.*) **153,** 64–74
11. Breen, J. P., and Crouch, M. L. (1992) Molecular analysis of a cruciferin storage protein gene family of *Brassica napus*. *Plant Mol. Biol.* **19,** 1049–1055
12. Inquello, V., Raymond, J., and Azanza, J. L. (1993) Disulfide interchange reactions in 11S globulin subunits of *Cruciferae* seeds. Relationships to gene families. *Eur. J. Biochem.* **217,** 891–895
13. Rödin, J., Ericson, M. L., Josefsson, L. G., and Rask, L. (1990) Characterization of a cDNA clone encoding a *Brassica napus* 12 S protein (cruciferin) subunit. Relationship between precursors and mature chains. *J. Biol. Chem.* **265,** 2720–2723
14. Rödin, J., Sjödahl, S., Josefsson, L. G., and Rask, L. (1992) Characterization of a *Brassica napus* gene encoding a cruciferin subunit: estimation of sizes

of cruciferin gene families. *Plant Mol. Biol.* **20,** 559–563
15. Sjödahl, S., Rödin, J., and Rask, L. (1991) Characterization of the 12S globulin complex of *Brassica napus*. Evolutionary relationship to other 11–12S storage globulins. *Eur. J. Biochem.* **196,** 617–621
16. Badley, R. A., Atkinson, D., Hauser, H., Oldani, D., Green, J. P., and Stubb, J. M. (1975) The structure, physical and chemical properties of the soy bean protein glycinin. *Biochim. Biophys. Acta* **412,** 214–228
17. Marcone, M. F., Beniac, D. R., Harauz, G., and Yada, R. Y. (1994) Quaternary structure and model for the oligomeric seed globulin from *Amaranthus hypochondriacus* K343. *J. Agric. Food Chem.* **42,** 2675–2678
18. Plietz, P., Damaschun, G., Müller, J. J., and Schwenke, K. D. (1983) The structure of 11-S globulins from sunflower and rape seed. A small-angle X-ray scattering study. *Eur. J. Biochem.* **130,** 315–320
19. Adachi, M., Takenaka, Y., Gidamis, A. B., Mikami, B., and Utsumi, S. (2001) Crystal structure of soybean proglycinin A1aB1b homotrimer. *J. Mol. Biol.* **305,** 291–305
20. Tandang-Silvas, M. R., Fukuda, T., Fukuda, C., Prak, K., Cabanos, C., Kimura, A., Itoh, T., Mikami, B., Utsumi, S., and Maruyama, N. (2010) Conservation and divergence on plant seed 11S globulins based on crystal structures. *Biochim. Biophys. Acta* **1804,** 1432–1442
21. Tandang, M. R., Adachi, M., and Utsumi, S. (2004) Cloning and expression of rapeseed procruciferin in *Escherichia coli* and crystallization of the purified recombinant protein. *Biotechnol. Lett.* **26,** 385–391
22. Adachi, M., Kanamori, J., Masuda, T., Yagasaki, K., Kitamura, K., Mikami, B., and Utsumi, S. (2003) Crystal structure of soybean 11S globulin: glycinin A3B4 homohexamer. *Proc. Natl. Acad. Sci. U.S.A.* **100,** 7395–7400
23. Jiang, L., Phillips, T. E., Rogers, S. W., and Rogers, J. C. (2000) Biogenesis of the protein storage vacuole crystalloid. *J. Cell Biol.* **150,** 755–770
24. Wittig, I., Braun, H. P., and Schägger, H. (2006) Blue native PAGE. *Nat. Protoc.* **1,** 418–428
25. Sunderhaus, S., Klodmann, J., Lenz, C., and Braun, H. P. (2010) Supramolecular structure of the OXPHOS system in highly thermogenic tissue of *Arum maculatum*. *Plant Physiol. Biochem.* **48,** 265–272
26. Schägger, H., and von Jagow, G.(1987) Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal. Biochem.* **166,** 368–379
27. Neuhoff, V., Stamm, R., Pardowitz, I., Arold, N., Ehrhardt, W., and Taube, D. (1990) Essential problems in quantification of proteins following colloidal staining with coomassie brilliant blue dyes in polyacrylamide gels, and their solution. *Electrophoresis* **11,** 101–117
28. Neuhoff, V., Stamm, R., and Eibl, H. (1985) Clear background and highly sensitive protein staining with Coomassie Blue dyes in polyacrylamide gels: A systematic analysis. *Electrophoresis* **6,** 427–448
29. Kruft, V., Eubel, H., Jänsch, L., Werhahn, W., and Braun, H. P. (2001) Proteomic approach to identify novel mitochondrial proteins in *Arabidopsis*. *Plant Physiol.* **127,** 1694–1710
30. Dudkina, N. V., Eubel, H., Keegstra, W., Boekema, E. J., and Braun, H. P. (2005) Structure of a mitochondrial supercomplex formed by respiratory-chain complexes I and III. *Proc. Natl. Acad. Sci. U.S.A.* **102,** 3225–3229
31. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., Ferrin, T. E. (2004) UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25,** 1605–1612
32. Quillin, M. L., and Matthews, B. W. (2000) Accurate calculation of the density of proteins. *Acta Crystallogr. D Biol. Crystallogr.* **56,** 791–794
33. Rossmann, M. G. (2000) Fitting atomic models into electron-microscopy maps. *Acta Crystallogr. D Biol. Crystallogr.* **56,** 1341–1349
34. Andème Ondzighi, C., Christopher, D. A., Cho, E. J., Chang, S. C., and Staehelin, L. A. (2008) *Arabidopsis* protein disulfide isomerase-5 inhibits cysteine proteases during trafficking to vacuoles before programmed cell death of the endothelium in developing seeds. *Plant Cell* **20,** 2205–2220
35. Huang, A. H. (1996) Oleosins and oil bodies in seeds and other organs. *Plant Physiol.* **110,** 1055–1061
36. Moreira, M. A., Hermodson, M. A., Larkins, B. A., and Nielsen, N. C. (1979) Partial characterization of the acidic and basic polypeptides of glycinin. *J. Biol. Chem.* **254,** 9921–9926
37. Shewry, P. R., Napier, J. A., and Tatham, A. S. (1995) Seed storage proteins: structures and biosynthesis. *Plant Cell* **7,** 945–956
38. Jagow, G., and Schägger, H. (2003) *Membrane Protein Purification and Crystallization: A Practical Guide*, 2nd Ed., pp. 105–130 Academic Press
39. Agrawal, G. K., and Thelen, J. J. (2006) Large scale identification and quantitative profiling of phosphoproteins expressed during seed filling in oilseed rape. *Mol. Cell Proteomics* **5,** 2044–2059
40. Wan, L., Ross, A. R., Yang, J., Hegedus, D. D., and Kermode, A. R. (2007) Phosphorylation of the 12 S globulin cruciferin in wild-type and abi1–1 mutant *Arabidopsis thaliana* (thale cress) seeds. *Biochem. J.* **404,** 247–256
41. Meyer, L. J., Gao, J., Xu, D., and Thelen, J. J. (2012) Phosphoproteomic analysis of seed maturation in *Arabidopsis*, rapeseed, and soybean. *Plant Physiol.* **159,** 517–528
42. Bhatty, R. S., McKenzie, S. L., and Finlayson, A. J. (1968) The proteins of rapeseed (*Brassica napus* L.) soluble in salt solutions. *Can. J. Biochem.* **46,** 1191–1197
43. Lawrence, M. C., Suzuki, E., Varghese, J. N., Davis, P. C., Van Donkelaar, A., Tulloch, P. A., and Colman, P. M. (1990) The three-dimensional structure of the seed storage protein phaseolin at 3 A resolution. *EMBO J.* **9,** 9–15
44. Schwenke, K. D., Raab, B., Plietz, P., and Damaschun, G. (1983) The structure of the 12 S globulin from rapeseed (*Brassica napus* L.). *NAHRUNG* **27,** 165–175

Publication 3

2.3 Proteomic and histological analyses of endosperm development in *Cyclamen persicum* as a basis for optimization of somatic embryogenesis

Jenniffer Wamaitha Mwangi[1], Christina Rode[1], Frank Colditz[1], Christin Haase[1], Hans-Peter Braun[1], Traud Winkelmann[2]

[1]Department of Plant Proteomics, Institute for Plant Genetics, Faculty of Natural Sciences, Leibniz Universität Hannover
[2]Institute of Floriculture and Woody Plant Science, Leibniz Universität Hannover

Type of authorship:	Co-author
Type of article:	Research article
Share of the work:	15 %
Contribution to the publication:	Analysed data, prepared figures
Journal:	Plant Science
Impact factor:	3.607
Date of publication:	Published in March 2013
Number of citations: (Google scholar, Nov. 07th, 2015)	1
DOI:	10.1016/j.plantsci.2012.11.004
PubMed-ID:	23352402

Plant Science 201–202 (2013) 52–65

Contents lists available at SciVerse ScienceDirect

Plant Science

journal homepage: www.elsevier.com/locate/plantsci

Proteomic and histological analyses of endosperm development in *Cyclamen persicum* as a basis for optimization of somatic embryogenesis

Jenniffer Wamaitha Mwangi[a], Christina Rode[a], Frank Colditz[a], Christin Haase[a], Hans-Peter Braun[a], Traud Winkelmann[b,*]

[a] Institute of Plant Genetics, Leibniz Universitaet Hannover, Herrenhaeuser Str. 2, D-30419 Hannover, Germany
[b] Institute of Floriculture and Woody Plant Science, Leibniz Universitaet Hannover, Herrenhaeuser Str. 2, D-30419 Hannover, Germany

ARTICLE INFO

Article history:
Received 18 September 2012
Received in revised form
13 November 2012
Accepted 15 November 2012
Available online 28 November 2012

Key words:
Cyclamen persicum
Endosperm
Mass spectrometry
Proteomics
Seed development

ABSTRACT

The endosperm plays an important role for the development of zygotic embryos, while somatic embryos lack a seed coat and endosperm and often show physiological disorders. This study aims at elucidating the cellular and physiological processes within the endosperm of the ornamental species *Cyclamen persicum* Mill. Histological analyses were performed from 0 to 11 weeks after pollination (WAP). At 3 WAP, a syncytium was clearly visible with a globular zygotic embryo. From 4 WAP, cellularization of the endosperm, at 5 WAP a small torpedo shaped embryo, and from 7 WAP cell expansion was observed. By 11 WAP the endosperm appeared fully differentiated. Total soluble proteins were extracted from the endosperm at 4, 5, 7, 9 and 11 WAP and resolved using two dimensional isoelectric focussing/sodium dodecyl sulphate–polyacrylamide gel electrophoresis (2D IEF/SDS–PAGE). A shift from high-molecular-mass proteins to low-molecular-mass proteins during endosperm development was observed. A total of 1137 protein spots/gel were detected in the three protein fractions extracted at 7, 9 and 11 WAP. Mass spectrometry analysis of the 48 predominant protein spots in endosperm at 7, 9 and 11 WAP resulted in the identification of 62 proteins, ten of which were described for the first time in *Cyclamen*. Additionally, 186 proteins were identified using the *C. persicum* embryo proteome reference map. Proteins involved in abscisic acid signalling and oxidative stress responsive proteins were found to be important for seed development in *Cyclamen*. The new insights into endosperm physiology including storage compounds are discussed.

1. Introduction

Within many plant seeds, the endosperm functions mainly to provide nutrients to the embryo, thereby supporting its development and also its later germination [1–4]. In addition, the endosperm insulates the embryo from mechanical pressure imposed by the seed coat [5], and plays a role in signalling towards the embryo [4]. Recently, the function of the endosperm especially as an integrator of seed growth and development based on signalling between endosperm and embryo as well as mechanical barrier has been emphasized [6].

The endosperm, depending on the species, may be a transient tissue which is largely reabsorbed during late seed development e.g. in *Arabidopsis*, or it may be enlarged and persistent even upon seed maturity e.g. in *Cyclamen* and cereals. In the case of *Arabidopsis*, the cotyledon offers a comparable storage function as represented by the endosperm [3,4]. Endosperm development generally progresses through several characteristic stages, i.e. syncytium formation by several nuclear divisions, cellularization, growth and differentiation and finally maturation including accumulation of storage compounds, development of desiccation tolerance and dormancy [7]. Plants are able to accumulate carbohydrates [8], proteins [9] and fatty acids [10] as storage compounds in their endosperm. Therefore the endosperm functions as the storage organ and developmental control unit for the embryo and for the germinating seed [2,11].

Cyclamen persicum is a popular ornamental crop with high economic relevance. For commercial propagation, the F_1 hybrid cultivars are of predominant meaning. However, relatively high costs for seed production due to inbreeding depression of parent lines and intensive manual work are challenging problems. Therefore, there is an interest in an alternative vegetative propagation system. Somatic embryogenesis has been reported to represent an efficient in vitro propagation technique in *Cyclamen* [12–15]. However, somatic embryos often show physiological disorders,

Abbreviations: CBB, Coomassie Brilliant Blue; IEF, isoelectric focussing; SDS–PAGE, sodium dodecyl sulphate–polyacrylamide gel electrophoresis; 2D, two-dimensional.

* Corresponding author. Fax: +49 511 762 3608.
E-mail address: traud.winkelmann@zier.uni-hannover.de (T. Winkelmann).

0168-9452/$ – see front matter © 2012 Elsevier Ireland Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.plantsci.2012.11.004

asynchronous development and misshaping during development. In addition, unlike zygotic embryos, somatic embryos lack a seed coat and the endosperm. One marked difference being the availability of storage compounds i.e. carbohydrates, proteins and lipids in the endosperm.

For *Cyclamen* seeds, the storage polysaccharide xyloglucan [16,17], the storage proteins 11S and 7S globulin [17] and truncated forms of enolase proposed to function as storage proteins [18] have been reported. But so far, a profound knowledge about the components and physiological processes within the endosperm development of *Cyclamen* seeds is lacking.

Proteomic studies have been shown to be a powerful tool for monitoring the present physiological status of cells and tissues under specific developmental conditions [19] and during development [20,21] and have been performed successfully for seeds and seed compartments in many crops recently, e.g. in maize [22], coffee [23] or rice [24].

In this study, we aimed (i) to identify the key events during endosperm development at the cellular level by histological analyses, and (ii) to elucidate on the basis of alterations in the protein profiles of *C. persicum* endosperm during seed development the predominant role of the endosperm for the embryo development.

2. Materials and methods

2.1. Plant material

The diploid *C. persicum* commercial F_1 cultivar 'Maxora Light Purple' bred by the company Varinova (Berkel en Rodenrijs, Netherlands) was grown in the greenhouse at a heating set point of 17 °C and a ventilation set point of 20 °C. Closed flower buds were emasculated and self-pollinated after 48 h. Fruits were harvested weekly from 0 (before pollination) to 11 weeks after pollination (WAP). Three biological replications were collected for each stage.

2.2. Histological analyses

Histological analyses were performed from day 0 (before pollination) to 11 WAP. For each stage, histological analyses were repeated twice to ensure two biological and technical replications. The samples were fixed in FAA (formaldehyde acetic acid) solution containing 67% ethanol, 20% H_2O, 1.8% formaldehyde and 5% glacial acetic acid for at least 24 h and then stored at 4 °C. The tissues were dehydrated in a vacuum using graded alcohol series (70–96% ethanol and 100% isopropanol) and embedded in paraffin wax [25].

Sections of 3–8 μm were cut using a microtome and stained using FCA (Fuchsine Chrysoidine Astra blue, Morphisto, Germany) solution according to Hoenemann et al. [25] and Morphisto histology manual (www.morphisto.de). The slides were visualized using a light microscope (Carl Zeiss, Germany) and the photographs were taken using AxioCam MR3 and edited with the AxioVision software (Carl Zeiss, Germany).

2.3. Proteomic analyses

2.3.1. Phenolic protein extraction

Proteomic analyses were performed for tissue harvested at 4, 5, 7, 9 and 11 WAP. Three biological replicates were collected for each stage. The samples were prepared under a stereo microscope (Carl Zeiss, Germany) directly frozen in liquid nitrogen and stored at −80 °C until the time of analysis.

Total proteins were extracted and precipitated according to Hurkman and Tanaka [26] protocol modified by Colditz et al. [27]. Whole seeds were used for proteome analyses at 4 and 5 WAP as the endosperm was still liquid (Fig. 1a and b). For 7, 9 and 11 WAP, endosperm and seed coat were collected (Fig. 1c). 60 mg (9 and 11 WAP) and 80 mg (4, 5, and 7 WAP) of endosperm tissue were pulverized in a bead mill and homogenized in extraction buffer (700 mM sucrose, 500 mM Tris, 50 mM EDTA, 100 mM KCl, 2 mM PMSF and 2% (v/v) (-mercaptoethanol, pH 8.0). Saturated phenol (pH 6.6/7.9) was added to the samples and proteins were precipitated in the phenolic phase with 100 mM ammonium acetate in methanol at −20 °C overnight.

2.3.2. Two dimensional (2D) IEF/SDS–PAGE

For first dimension isoelectric focussing (IEF), immobilized dry strips (18 cm) with pH gradients 3–11 were rehydrated with protein samples in rehydration buffer (8 M urea, 2% (w/v) CHAPS, 100 mM DTT, 0.5% (v/v) IPG buffer). Isoelectric focussing was done using IPGphor system (GE Healthcare). IPG strips were equilibrated in equilibration solution I (30% (v/v) glycerol, 50 mM Tris–HCl pH 8.8, 6 M urea, 2% (w/v) SDS, a trace of bromophenol blue, 0.01 g DTT ml^{-1} (w/v)) and equilibration solution II (same compounds like equilibration solution I, but DTT substituted by 0.025 g iodoacetamide ml^{-1}).

For second dimension of sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS–PAGE) IPG strips were fixed horizontally onto SDS–tricine–polyacrylamide gels of 12% acrylamide. Electrophoresis was carried out for 20 h at 30 mA mm^{-1} using Biorad Protean IIXL gel system (Biorad, München, Germany). Gels were subsequently stained overnight using colloidal Coomassie Brilliant Blue (CBB-250 G, Merck, Darmstadt, Germany) after treatment with the fixing solution (40% (v/v) methanol, 10% (v/v) acetate) for at least 2 h [28,29].

2.3.3. Quantitative gel evaluation

The gels were scanned on an ImageScanner III (GE Healthcare, Freiburg, Germany) and evaluated using Delta 2D software, version 4.0 (Decodon, Greifswald, Germany) with three replicates for each group (4, 5, 7, 9 and 11 WAP). Spots detection was done automatically and occasionally corrected manually. In gel normalization was performed using the Delta 2D software for the overlays of three replicate gels each. Spots with a relative spot volume of less than 0.005% were deleted. Significant changes in spot patterns of the different endosperm groups were determined using Student's *t*-test (confidence interval ≥95%) based on the relative spot volume.

2.4. Mass spectrometry

Protein spots were cut out from the 2D IEF/SDS gel using a manual spot picker (Genetix; spot diameter, 1.4 mm) and were assigned the unique spot number identifier according to Delta 2D software. Mass spectrometry analyses were carried out according to [30]. The gel pieces were de-stained by washing in water, dried under vacuum and dehydrated using acetonitrile and then incubated with 0.1 M NH_4HCO_3. Trypsin digestion (2 mg/ml resuspension buffer [Promega] in 0.1 M NH_4HCO_3) was performed at 37 °C overnight. The resulting tryptic peptides were extracted by incubation with 5% formic acid in 50% acetonitrile for 15 min at 37 °C. The supernatant was set aside and a second extraction with 100% acetonitrile was performed. Finally, the supernatant was combined with the first one. Extracted peptides were dried by vacuum centrifugation and stored at −20 °C. The peptides were analysed via LC–MS/MS using a Proxeon Easy – nLC (pre-column: 100 μm diameter/2 cm length, main column: 75 μm diameter/10 cm length; Acclaim® PepMap, Germany) and a micrOTOF – QII ESI-MS/MS (Bruker Daltonics, Bremen, Germany). Spectra were generated using the "HyStar compass post processing" software (Bruker Daltonics, Bremen, Germany). Proteins were identified using the MASCOT search algorithm against the NCBI non-redundant protein database (http://www.ncbi.nlm.nih.gov, RefSeq collection (Release 47, May 2011), plant sub-database),

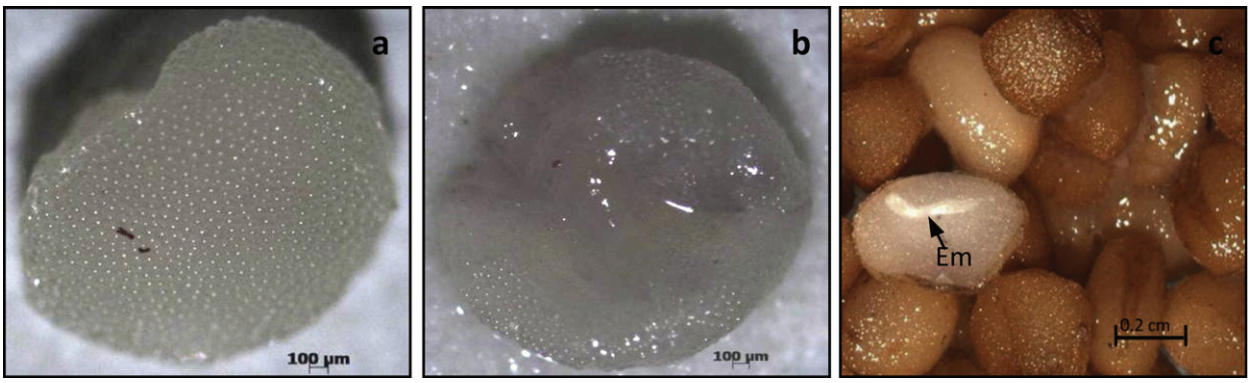


Fig. 1. *C. persicum* seed: (a) at 4 WAP; (b) pierced seed at 4 WAP with liquid endosperm and (c) at 11 WAP seed showing the zygotic embryo (Em).

the SWISS-PROT (http://www.expasy.org) and TAIR 10 databases (www.arabidopsis.org). Only proteins with a MASCOT score of at least 40 were selected. Hypothetical/unknown protein sequences were further blasted through SWISS-PROT. The identified proteins were grouped according to their physiological pathway based on KEGG PATHWAY modified by [31].

3. Results

3.1. Histological analysis of endosperm development

Histological analyses applying FCA staining served as a basis for the proteomic analyses to identify the key events during endosperm development. At the beginning, ovules were observed before pollination had been carried out (0 WAP) in which the embryo sac was clearly visible (Fig. 2, 0 WAP). One week after pollination, the chalazal and micropylar ends of the embryo sac could be identified and the outer cell layers forming the testa were stained brown to orange indicating secondary cell wall formation and lignin/suberin accumulation. By 2 WAP, endosperm development had already begun characterized by a mass of multinucleate cytoplasm in the centre of the seed (Fig. 2, 2 WAP). At 3 WAP, the syncytium was clearly visible with a tiny globular embryo in the middle (not shown). At this stage the endosperm was at a transition stage to become cellular. Cellularization of the endosperm begun at 4 WAP from the periphery of the embryo sac. At 5 WAP, a small torpedo shaped embryo was observed within the cellularized endosperm. The development of the endosperm from 6 to 11 WAP was characterized by changes in cell shape and cell wall thickness. At 7 WAP, the endosperm was observed to consist of loosely packed cells which by 11 WAP formed a compact mass of cells. The endosperm surrounding region was observed to be a distinct development domain (Fig. 2).

3.2. Comparison of zygotic embryo and endosperm proteomes

Proteomic analyses were performed in order to characterize the *C. persicum* endosperm proteome. For this purpose, 2D IEF/SDS gel electrophoresis was carried out with protein samples derived from five seed developmental stages at 4, 5, 7, 9 and 11 WAP (Fig. 3). Since the endosperm at 4 and 5 WAP was still liquid and separation from other tissue was not possible, whole seeds were analysed. After solidification of the endosperm, it was easy to isolate from the embryo as both tissues are separated by the epidermis of the embryo. Hence, proteomic analyses at 7, 9 and 11 WAP were carried out with isolated endosperm tissue. Proteomic analyses led to the resolution of approximately 1000 separate protein spots per single gel after CBB staining. From the endosperm protein gels, protein identification was achieved via two different methods:

(i) Identification based on similar protein spot patterns when compared with the *C. persicum* digital proteome reference map for zygotic and somatic embryos [18,31] (http://www.gelmap.de/cyclamen). Because of a high similarity level (186 of 490 spots) in protein pattern observed between the zygotic embryo and endosperm gels (Suppl. Fig. S1), this reference map was used for protein identification in 7, 9 and 11 WAP endosperm gels. The transfer of protein spot labels from the zygotic embryo to endosperm gel was performed automatically using Delta 2D software and each spot was further verified manually. A total of 186 protein spots out of 490 detected spots were identified. Among them, 19 spots had multiple proteins identified (see Appendix A in Supporting Information).
(ii) Additional protein identification via LC–ESI-MS/MS of selected protein spots that revealed altered abundance patterns during the time-course analysed but were not annotated in the *C. persicum* digital proteome reference map (Table 1, see also Appendix B in Supporting Information).

Mass spectrometric analyses were performed for 48 protein spots predominantly appearing in 7, 9 and 11 WAP endosperm gels. In 39 out of these 48 spots, 62 proteins were identified while 9 spots had no identity matches. This represents 81% identification rate. Ten proteins were described for the first time in *Cyclamen*, including zeaxanthin epoxidase, F-box protein, PR 10 family protein, glucose and ribitol dehydrogenase, seed maturation protein, vacuolar processing enzyme 1a, ethylene receptor, sugar carrier protein C, phytoene desaturase and wax synthase.

3.3. Overall shift in proteins' molecular mass during endosperm development

The whole seed proteomes at 4 and 5 WAP had a high proportion of high-molecular-mass proteins that were partly observed also in 7 WAP endosperm gel but were characteristically lacking in the 11 WAP endosperm gel (Fig. 3a). Additionally, at early seed development stages (4 and 5 WAP), the low-molecular-mass proteins were absent (Fig. 3a). Therefore, a shift from high-molecular-mass proteins in the early seed development stages to low-molecular-mass proteins during seed maturation stage was observed (Suppl. Fig. S2).

A total of 174 protein spots were of different abundance between 7 and 11 WAP (112 spots at least 1.5-fold higher in abundance at 7 WAP, Fig. 4), an indication of the metabolic difference between these stages. In contrast, only 45 spots were of different abundance between 9 and 11 WAP (36 spots at least 1.5-fold higher in abundance at 11 WAP, Fig. 5).

Table 1
C. persicum endosperm proteins identified by LC–MS/MS after IEF/SDS–PAGE (a full list of the proteins identified can be found in the appendix).

								7 WAP			9 WAP			11 WAP		
Spot No[a]	Score[b]	SC[c] (%)	M_r[d] kDa	pI[e]	MW[f]	Protein name[g]	Physiological pathway[h]	v[i]	cov[j]	ratio[k]	v[i]	cov[j]	ratio[k]	v[i]	cov[j]	
346	55.5	6	61.4	8.6	l	Phytoene desaturase	abscisic acid signaling pathway	0.08	53	6.8	0.64	27	0.9	0.56	13	
15	53	1.4	72.4	7.8	m	Zeaxanthin epoxidase	abscisic acid signaling pathway	0.09	5	3.3	0.30	31	1.0	0.29	28	
349	79	1.4	72.4	7.8	m	Zeaxanthin epoxidase	abscisic acid signaling pathway	2.03	11	0.5	1.27	25	0.8	1.07	11	
416	46	1.4	72.4	7.8	s	Zeaxanthin epoxidase	abscisic acid signaling pathway	2.47	16	0.9	2.21	26	1.1	2.33	1	
313	53	1.6	54.9	5.4	l	BR-signaling kinase 2/Protein kinase-like protein	brassinosteroid biosynthesis	4.06	18	0.5	2.56	15	0.8	1.94	10	
331	56	5.7	35.8	6.9	m	Annexin	cell division and growth	0.64	38	0.2	0.13	39	0.8	0.11	29	
179	117	9.8	17	4.8	m	PR 10 like protein	defense response	0.92	7	1.2	1.02	7	1.0	1.05	3	
232	44	2.7	31.7	6.7	l	Seed maturation protein	embryogenesis/seed germination	0.08	46	4.7	0.19	9	1.9	0.35	16	
312	40	3.4	82.2	8.9	l	Ethylene receptor 1	ethylene signaling pathway	0.61	20	0.8	0.54	1	0.9	0.50	13	
461	1658	23.6	47.9	5.9	m	Enolase	glycolysis	0.87	9	0.6	0.57	15	1.0	0.56	12	
314	62	4.4	31.6	6.7	m	Glucose and ribitol dehydrogenase	glycolysis	1.12	20	0.8	0.97	33	0.9	0.88	21	
351	91	2.1	60.1	8.0	m	Dihydrolipoyl dehydrogenase	lipid pathway	1.61	12	0.4	0.70	8	0.8	0.57	22	
427	44	1.7	59.2	6.5	m	Wax synthase	lipid pathway	0.18	41	0.1	0.02	23	0.8	0.02	30	
349	120	8	47.7	6.1	s	26S proteasome ATPase subunit	protein degradation	2.03	11	0.5	1.27	25	0.8	1.07	11	
287	80	4.6	31.2	4.6	m	Vacuolar processing enzyme	protein degradation	0.53	18	0.4	0.28	27	0.8	0.22	3	
59	44	2.1	41.8	5.2	m	F-box protein	protein degradation/ signalling	0.32	17	0.3	0.08	34	1.3	0.11	36	
196	49	2.1	41.8	5.2	l	F-box protein	protein degradation/ signalling	0.46	18	2.4	0.76	3	1.4	1.09	6	
313	43	2.1	41.8	5.2	l	F-box protein	protein degradation/ signalling	4.06	18	0.5	2.56	15	0.8	1.94	10	
326	40	2.1	41.8	5.2	l	F-box protein	protein degradation/ signalling	4.07	19	1.0	4.35	16	0.9	3.92	12	
349	40	2.1	41.8	5.2	s	F-box protein	protein degradation/ signalling	2.03	11	0.5	1.27	25	0.8	1.07	11	
427	41	2.1	41.8	5.2	l	F-box protein	protein degradation/ signalling	0.18	41	0.1	0.02	23	0.8	0.02	30	
59	458	8.9	57.3	6.8	m	Catalase	stress response	0.32	17	0.3	0.08	34	1.3	0.11	36	
176	75	10.5	17.4	5.5	m	Peroxiredoxin, putative	stress response	0.74	6	1.8	1.18	3	1.1	1.35	10	
313	46	1.3	57.7	9.5	l	Sugar carrier protein C	transport	4.06	18	0.5	2.56	15	0.8	1.94	10	
314	61	1.3	57.7	9.5	l	Sugar carrier protein C	transport	1.12	20	0.8	0.97	33	0.9	0.88	21	
326	41	1.3	57.7	9.5	l	Sugar carrier protein C	transport	4.07	19	1.0	4.35	16	0.9	3.92	12	
346	44.4	1.3	57.7	9.5	l	Sugar carrier protein C	transport	0.08	53	6.8	0.64	27	0.9	0.56	13	
291	68	1.7	33.1	5.0	m	Xyloglucan endotransglucosylase	xyloglucan biosynthesis	3.01	62	1.8	7.33	17	0.7	5.45	11	

[a] Protein spot number as assigned by Delta 2D spot detection.
[b] Mascot score.
[c] Sequence coverage.
[d] Calculated molecular mass (kDa).
[e] Calculated isoelectric point.
[f] Molecular weight (MW) in gel as compared to the theoretically expected MW: (m) MW in gel corresponding to the theoretically expected MW ± 15 kDa; (s) MW in gel lower than theoretically expected and (l) MW in gel larger than theoretically expected.
[g] Protein name according to the best hit of MASCOT search against NCBInr, SwissProt and Tair 10 databases. For "hypothetical protein", "unknown" or "unnamed protein product" a BLAST search of the "unnamed" protein sequence was performed against SwissProt database.
[h] Physiological pathway according to KEGG PATHWAY modified by Rode et al. [31].
[i] Mean relative spot volume obtained in the three gel replicates of 7, 9 and 11 WAP. These values are illustrated by graphs at the right side of each row for all shown spots. The first bar (purple) represents 7 WAP, the second (red) 9 WAP and the third (green) 11 WAP.
[j] Coefficient of variation as calculated using Delta 2D software.
[k] Ratio of protein spot abundance (ratio = v (11 WAP)/v (7 WAP or 9 WAP)). Statistically significant values of at least 1.5 fold more or 1.5 fold less (all values ≤0.6) abundant protein spots are given. Green figures are indicating values of significant higher abundances in 11 WAP endosperm, purple indicates spots that are significant higher abundant in 7 WAP endosperm and red in 9 WAP endosperm.

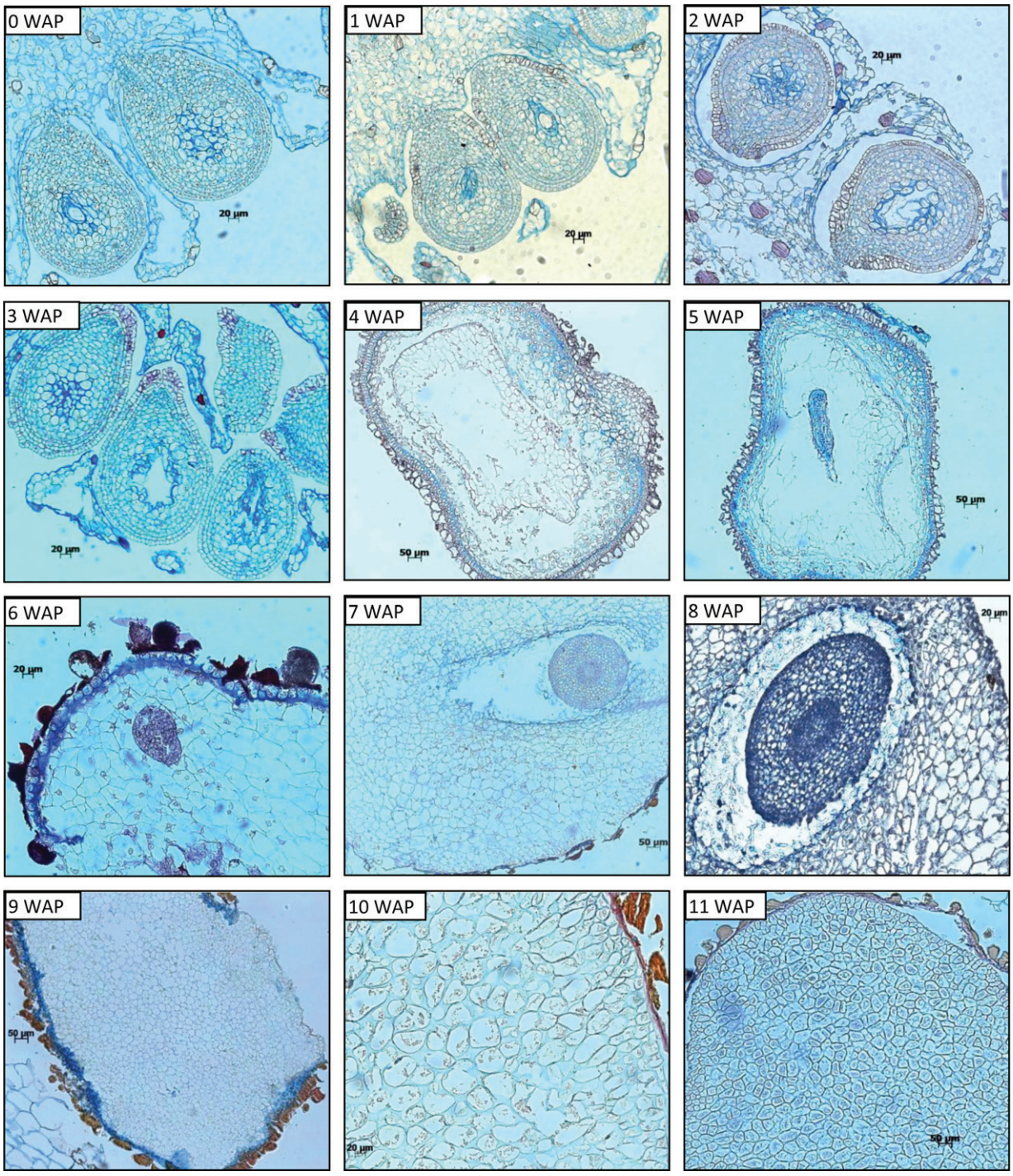


Fig. 2. Histological analysis of development of *C. persicum* seed, 0–11 WAP, using Fuchsine Chrysoidine Astra blue which stains cell walls blue, lignified or cutinized cell walls in orange to brownish red and nuclei in light purple. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of the article.)

3.4. Proteins changing in abundance during endosperm development

3.4.1. Cell division and growth proteins

Several proteins involved in cell division and growth were identified, such as annexin (2 spots), cell division cycle protein (2 spots), alpha tubulin and actin 1 (4 spots) (Table 2). Annexin (spots 110 and 331) and actin 1 (spots 100 and 351) were 1.5 fold higher abundant in 7 WAP as compared to 11 WAP indicating that at there was a decrease in cell division activity from 7 WAP to 9 and 11 WAP.

3.4.2. Stress related proteins

A number of proteins were identified that play a role in the control of the oxidative state of cells and in the elimination of reactive oxygen species. They are catalase, aluminium induced protein, glutathione reductase and osmotin-like protein

Table 2
C. persicum endosperm proteins identified using the *C. persicum* proteome reference map by Rode et al. [18,31].

				7 WAP			9 WAP			11 WAP	
Spot No.[a]	Name[b]	ID[c]	Physiological pathway[d]	v[e]	cov[e]	ratio[f]	v[d]	cov[e]	ratio[f]	v[d]	cov[e]
413	Aspartate aminotransferase	ID1937	Amino acid metabolism	1.00	14	1.2	0.75	17	1.5	1.15	9
454	D-isomer specific 2-hydroxyacid dehydrogenase	ID368	Amino acid metabolism	0.02	14	1.4	0.02	8	1.4	0.03	9
34	Peptidase	ID290	Amino acid metabolism	0.24	18	0.3	0.05	16	1.6	0.08	39
315	Peptidyl-prolyl cis-trans isomerase	ID942	Amino acid metabolism	0.03	17	9.0	0.07	32	4.3	0.29	31
410	Serine hydroxymethyltransferase	ID462	Amino acid metabolism	0.16	25	0.3	0.06	30	0.8	0.05	22
346	Auxin-amidohydrolase	ID853	auxin action regulation	0.08	53	6.8	0.64	27	0.9	0.56	13
110	Annexin	ID653	cell division and growth	0.31	14	0.5	0.14	30	1.0	0.14	5
100	Actin 1	ID586	cytoskeleton	0.26	12	0.3	0.10	12	0.9	0.09	31
351	Actin 1	ID472	cytoskeleton	1.61	12	0.4	0.70	8	0.8	0.57	22
247	Em-like protein	ID1059	embryogenesis/seed germination	0.03	39	4.0	0.10	12	1.4	0.13	22
289	LEA family protein	ID621	embryogenesis/seed germination	0.16	22	0.5	0.08	41	1.0	0.08	19
290	Alcohol dehydrogenase	ID1288	glycolysis	0.14	10	0.6	0.03	66	3.4	0.09	19
72	Enolase	ID11657	glycolysis	0.27	8	0.6	0.16	9	1.0	0.16	13
195	Enolase	ID976	glycolysis	0.15	8	1.6	0.17	8	1.4	0.24	10
196	Enolase	ID981, ID8980	glycolysis	0.46	18	2.4	0.76	3	1.4	1.09	6
201	Enolase	ID997	glycolysis	1.65	6	0.8	1.54	9	0.8	1.29	7
222	Enolase	ID1043	glycolysis	0.06	32	6.2	0.12	7	3.0	0.35	10
224	Enolase	ID17377	glycolysis	0.13	26	2.3	0.30	8	1.0	0.30	6
226	Enolase	ID1038	glycolysis	0.12	22	1.6	0.07	21	2.5	0.18	10
227	Enolase	ID1032	glycolysis	0.15	49	4.9	0.40	31	1.9	0.75	2
239	Enolase	ID1051	glycolysis	0.13	31	1.8	0.21	11	1.1	0.22	9
257	Enolase	ID1083	glycolysis	0.12	34	1.7	0.13	9	1.5	0.20	12
301	Enolase	ID799	glycolysis	0.08	23	0.2	0.03	16	0.4	0.01	74
365	Enolase	ID353	glycolysis	0.23	13	0.7	0.14	32	1.0	0.15	8
394	Enolase	ID18597	glycolysis	0.04	10	4.1	0.09	29	1.7	0.15	14
409	Enolase	ID763	glycolysis	0.80	8	0.6	0.54	8	1.0	0.51	19
449	Enolase	ID475	glycolysis	0.39	7	0.5	0.19	19	1.0	0.20	16
461	Enolase	ID480	glycolysis	0.87	9	0.6	0.57	15	1.0	0.56	12
199	Enolase	ID1022	glycolysis	1.19	4	0.8	1.29	2	0.8	1.00	5
393	Enolase	ID18597	glycolysis	0.13	11	4.3	0.27	55	2.0	0.54	31

Table 2 (*Continued*)

				7 WAP			9 WAP			11 WAP		
Spot No.[a]	Name[b]	ID[c]	Physiological pathway[d]	v[e]	cov[e]	ratio[f]	v[d]	cov[e]	ratio[f]	v[d]	cov[e]	
32	Fructose-bisphosphate aldolase	ID314,	glycolysis	0.17	18	0.3	0.03	6	2.0	0.05	23	
97	Fructose-bisphosphate aldolase	ID572	glycolysis	0.06	3	0.5	0.03	15	1.0	0.03	16	
377	Fructose-bisphosphate aldolase	ID568	glycolysis	0.39	8	0.6	0.22	6	1.0	0.21	24	
437	Fructose-bisphosphate aldolase	ID566	glycolysis	0.74	11	0.6	0.46	18	1.0	0.46	19	
28	Glyceraldehyde-3-phosphate dehydrogenase	ID199	glycolysis	0.21	22	0.0	0.03	39	0.3	0.01	88	
215	Glyceraldehyde-3-phosphate dehydrogenase	ID1025	glycolysis	0.09	22	4.7	0.08	18	5.2	0.40	26	
304	Glyceraldehyde-3-phosphate dehydrogenase	ID793	glycolysis	0.07	26	2.2	0.12	8	1.2	0.14	12	
359	Glyceraldehyde-3-phosphate dehydrogenase	ID580	glycolysis	0.19	4	0.7	0.14	23	0.9	0.12	24	
362	Glyceraldehyde-3-phosphate dehydrogenase	ID600	glycolysis	0.17	11	0.2	0.05	30	0.7	0.04	26	
393	Glyceraldehyde-3-phosphate dehydrogenase	ID18847	glycolysis	0.13	11	4.3	0.27	55	2.0	0.54	31	
414	Glyceraldehyde-3-phosphate dehydrogenase	ID607	glycolysis	0.23	12	0.3	0.08	42	0.8	0.06	24	
415	Glyceraldehyde-3-phosphate dehydrogenase	ID602	glycolysis	0.21	8	0.5	0.09	21	1.1	0.09	40	
419	Glyceraldehyde-3-phosphate dehydrogenase	ID17504	glycolysis	0.26	13	1.5	0.22	25	1.8	0.40	11	
439	Glyceraldehyde-3-phosphate dehydrogenase	ID606	glycolysis	0.33	17	0.2	0.10	47	0.8	0.07	19	
419	Malate dehydrogenase	ID648	glycolysis	0.26	13	1.5	0.22	25	1.8	0.40	11	
420	Malate dehydrogenase	ID665	glycolysis	0.46	8	0.5	0.20	15	1.1	0.21	11	
93	Phosphoglycerate kinase	ID550	glycolysis	0.05	15	0.5	0.02	46	1.1	0.03	24	
208	Phosphoglycerate kinase	ID998	glycolysis	0.07	27	2.3	0.07	16	2.3	0.16	4	
219	Phosphoglycerate kinase	ID9067	glycolysis	0.15	46	3.7	0.43	10	1.3	0.56	15	
359	Pyruvate dehydrogenase	ID578	glycolysis	0.19	4	0.7	0.14	23	0.9	0.12	24	
386	Triosephosphate isomerase	ID771	glycolysis	0.93	25	1.3	0.95	6	1.3	1.24	10	
459	UDP-glucose pyrophosphorylase	ID441	glycolysis	0.22	12	0.5	0.09	22	1.4	0.12	21	
354	3-Oxoacyl-[acyl-carrier-protein] synthase I	ID516	lipid pathways	0.96	11	0.3	0.40	32	0.7	0.28	21	
394	Enoyl-ACP reductase	ID677	lipid pathways	0.04	10	4.1	0.09	29	1.7	0.15	14	
210	GDSL esterase/lipase,	ID9002	lipid pathways	0.17	31	1.9	0.28	7	1.2	0.33	13	
388	Glutamine synthetase	ID597	nitrogen fixation	0.07	14	0.5	0.04	29	0.9	0.03	16	
29	Beta-xylosidase/alpha-L-arabinosidase	ID240	other processes	0.02	15	0.2	0.01	49	0.6	0.01	11	
124	Epimerase/Dehydratase, NAD-dependent	ID782	other processes	0.13	21	1.2	0.08	7	1.8	0.15	4	

Table 2 (*Continued*)

				7 WAP			9 WAP			11 WAP	
Spot No.[a]	Name[b]	ID[c]	Physiological pathway[d]	v[e]	cov[e]	ratio[f]	v[d]	cov[e]	ratio[f]	v[d]	cov[e]
411	Formate dehydrogenase, mitochondrial	ID574	other processes	2.28	6	0.7	1.68	3	0.9	1.57	8
431	Monodehydroascorbate reductase	ID473	other processes	0.07	13	0.5	0.05	16	0.8	0.04	27
287	Polygalacturonase	ID3278	other processes	0.53	18	0.4	0.28	27	0.8	0.22	3
288	Polygalacturonase	ID3278	other processes	0.19	29	0.3	0.09	33	0.7	0.06	15
349	6-Phosphogluconate dehydrogenase	ID476	oxidative pentose phosphate pathway	2.03	11	0.5	1.27	25	0.8	1.07	11
73	ATPase, mitochondrial	ID9601	oxydative phosphorylation	0.75	22	0.1	0.09	18	1.2	0.11	22
167	ATPase, mitochondrial	ID920	oxydative phosphorylation	2.58	9	0.8	2.89	9	0.7	2.04	5
15	NADH-ubiquinone oxidoreductase	ID2114	oxydative phosphorylation	0.09	5	3.3	0.30	31	1.0	0.29	28
124	Proteasome	ID801	protein degradation	0.13	21	1.2	0.08	7	1.8	0.15	4
349	Proteasome	ID412	protein degradation	2.03	11	0.5	1.27	25	0.8	1.07	11
351	Proteasome	ID3077	protein degradation	1.61	12	0.4	0.70	8	0.8	0.57	22
482	Proteasome	ID852	protein degradation	0.05	72	8.9	0.31	16	1.3	0.40	13
164	Heat shock protein 20	ID947	protein folding	0.49	9	0.7	0.44	7	0.7	0.33	9
199	Heat shock protein 20	ID985	protein folding	1.19	4	0.8	1.29	2	0.8	1.00	5
313	Heat shock protein 20	ID879	protein folding	4.06	18	0.5	2.56	15	0.8	1.94	10
404	Heat shock protein 20	ID874	protein folding	0.03	10	3.2	0.09	7	1.1	0.10	28
405	Heat shock protein 20	ID874	protein folding	0.12	32	1.0	0.18	13	0.7	0.12	12
406	Heat shock protein 20	ID898	protein folding	0.12	30	2.5	0.27	16	1.1	0.30	28
43	Heat shock protein 60	ID355	protein folding	0.68	25	0.1	0.06	17	0.7	0.04	7
46	Heat shock protein 60	ID340	protein folding	0.06	19	0.3	0.02	22	0.8	0.02	8
167	Heat shock protein 70	ID7469	protein folding	2.58	9	0.8	2.89	9	0.7	2.04	5
288	Heat shock protein 70	ID596	protein folding	0.19	29	0.3	0.09	33	0.7	0.06	15
297	Heat shock protein 70	ID766	protein folding	0.04	31	1.6	0.03	41	2.4	0.06	20
305	Heat shock protein 70	ID766	protein folding	0.01	132	6.1	0.03	10	1.3	0.04	25
384	Heat shock protein 70	ID761	protein folding	2.53	29	0.8	2.44	8	0.8	1.98	6
442	Heat shock protein 70	ID750	protein folding	0.07	42	5.2	0.23	13	1.6	0.37	21
484	Heat shock protein 70	ID850	protein folding	0.49	54	1.0	0.32	23	1.6	0.51	7
127	Aluminium induced protein	ID825	stress response	0.05	20	0.4	0.01	16	1.3	0.02	37

Table 2 (*Continued*)

Spot No.[a]	Name[b]	ID[c]	Physiological pathway[d]	7 WAP v[e]	7 WAP cov[e]	7 WAP ratio[f]	9 WAP v[d]	9 WAP cov[e]	9 WAP ratio[f]	11 WAP v[d]	11 WAP cov[e]
59	Catalase	ID386	stress response	0.32	17	0.3	0.08	34	1.3	0.11	36
61	Catalase	ID17829	stress response	0.06	2	0.4	0.02	26	1.0	0.02	28
426	Catalase	ID389	stress response	0.14	16	0.4	0.05	19	1.3	0.06	23
298	Copper/zinc superoxide dismutase	ID767	stress response	0.10	39	1.5	0.09	19	1.8	0.16	18
350	Glutathione reductase	ID414	stress response	0.11	28	0.4	0.05	23	1.0	0.05	14
306	Osmotin-like protein	ID3770	stress response	0.04	6	0.2	0.06	104	0.1	0.01	43
115	S-formylglutathione hydrolase	ID1376	stress response	0.07	19	3.0	0.10	19	2.0	0.21	24
176	Thioredoxin peroxidase 1	ID961	stress response	0.74	6	1.8	1.18	3	1.1	1.35	10
70	Universal stress protein	ID9992	stress response	0.02	54	5.6	0.06	18	1.5	0.09	13
140	Universal stress protein	ID881	stress response	0.05	7	2.8	0.11	15	1.3	0.14	10
232	Endoribonuclease	ID1057	transcription	0.08	46	4.7	0.19	9	1.9	0.35	16
6	Elongation factor 2	ID148	translation	0.18	10	0.2	0.02	27	1.4	0.03	35
60	Ubiquitin	ID9986	translation	0.04	34	0.0	0.00	34	0.7	0.00	96
67	Ubiquitin	ID9986	translation	0.06	20	2.5	0.12	19	1.2	0.14	15
11	Aconitase	ID160	tricarboxylic acid cycle	0.04	26	0.3	0.01	30	0.9	0.01	28
351	Dihydrolipoamide S-acetyltransferase	ID8206	tricarboxylic acid cycle	1.61	12	0.4	0.70	8	0.8	0.57	22
37	Malic enzyme	ID308	tricarboxylic acid cycle	0.07	26	0.3	0.01	14	1.2	0.02	20
238	Uncharacterized	ID1060	uncharacterized	0.06	17	1.9	0.05	21	2.4	0.12	11

[a]Protein spot number as assigned during Delta 2D protein spot detection. Corresponding spots of all gels are labelled with the same spot number.
[b]Proteins identified using *C. persicum* embryo proteome reference map [31].
[c]Spot ID represents the protein spot number as presented in the *C. persicum* embryo proteome reference map.
[d]Mean relative spot volume obtained in the three gel replicates of 7, 9 and 11 WAP. These values are illustrated by graphs at the right side of each row for all shown spots. The first bar (purple) represents 7 WAP, the second (red) 9 WAP and the third (green) 11 WAP.
[e]Coefficient of variation as calculated using Delta 2D software.
[f]Ratio of protein spot abundance (ratio = v (11 WAP)/v (7 WAP or 9 WAP)). Statistically significant values of at least 1.5 fold more or 1.5 fold less (all values ≤0.6) abundant protein spots are given. Green figures are indicating values of significant higher abundances in 11 WAP endosperm, purple indicates spots that are significant higher abundant in 7 WAP endosperm and red in 9 WAP endosperm.

and were higher abundant in 7 WAP as compared to 11 WAP, while S-formylglutathione hydrolase, universal stress protein and thioredoxin peroxidase 1 were 1,5 fold higher abundant in 11 WAP as compared to 7 WAP.

Heat shock proteins (HSP101, HSP20, HSP60 and HSP70) had predominant distribution with 28 spots indicating high protein turnover. Of these, only 15 spots were statistically significant. HSP 60 (spots 43 and 46) was 1.5 fold higher abundant in 7 WAP as compared to 11 WAP. HSP 20 and HSP 70 did not show any clear pattern of expression.

3.4.3. ABA pathway

Several proteins involved in ABA biosynthesis signalling pathway were found to be differentially abundant: zeaxanthin epoxidase, ZEP (spots 15, 349, and 416), phytoene desaturase (spot 346) and SAL1 phosphatase. ZEP was identified in three spots with different molecular weight and isoelectric point, of which, 2 spots were found increased from 7 to 11 WAP and one was reduced 7 to 11 WAP. Phytoene desaturase was 1.5 fold higher abundant in 11 WAP as compared to 7 WAP.

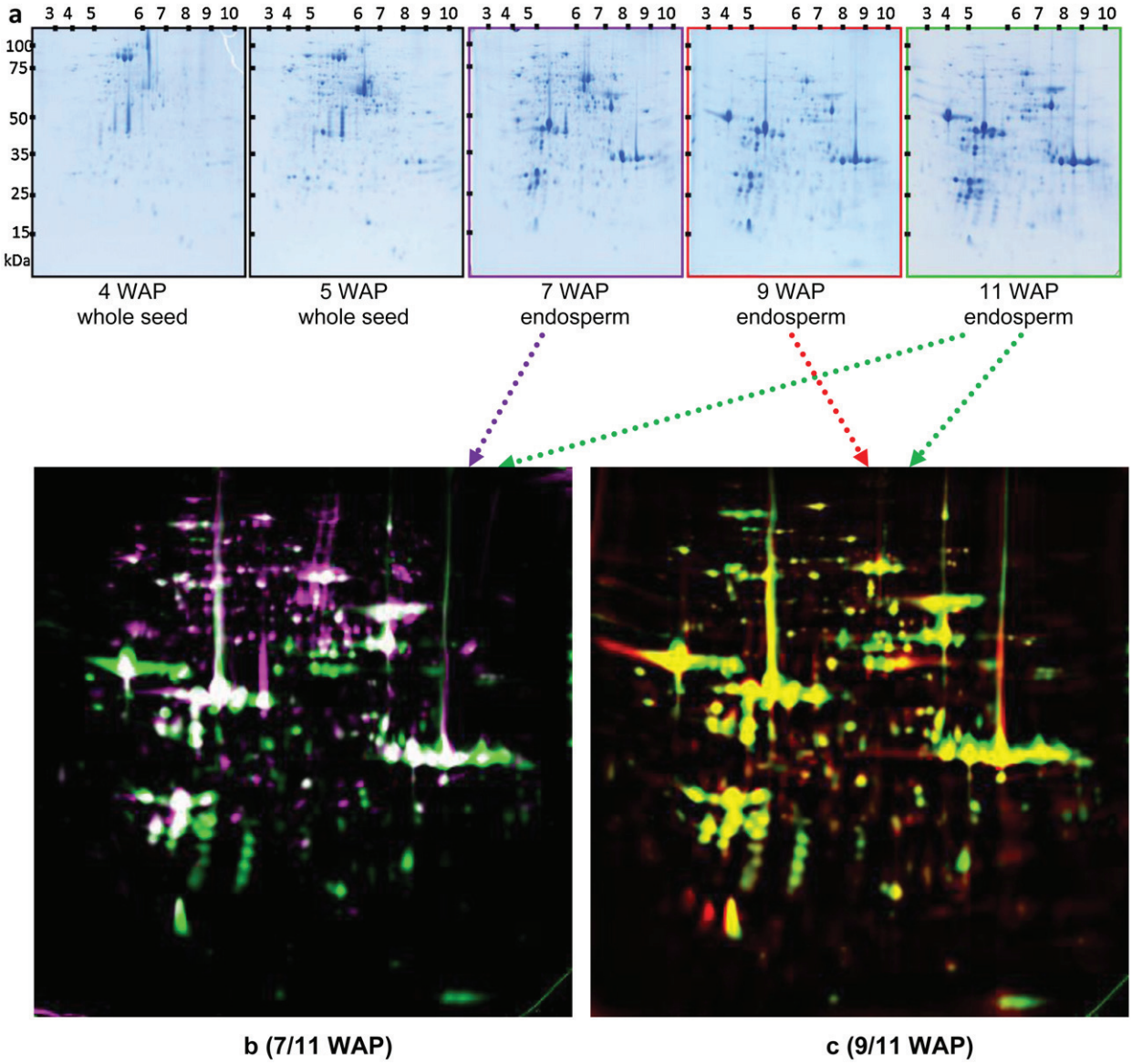


Fig. 3. (a) IEF/SDS gels of the five seed development stages selected for proteomic analyses. (b) IEF/SDS overlay image of 7 and 11 WAP endosperm proteomes (each three replicates) using Delta 2D software. Green spots were at least 1.5-fold more abundant at 11 WAP and purple spots were at least 1.5-fold more abundant at 7 WAP. Spots with similar abundance in both endosperm stages appear in white. (c) IEF/SDS overlay image of 9 and 11 WAP endosperm proteomes (each three replicates) using Delta 2D software. Green spots (36) were at least 1.5-fold more abundant at 11 WAP and red spots (9) were at least 1.5-fold more abundant at 9 WAP. Spots with similar abundance in both endosperm stages appear in yellow. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of the article.)

Further evidence for plant hormone signalling being involved in endosperm and probably embryo development came from the identification of BR signalling kinase (spot 313) (brassinolide signalling) and ethylene receptor 1 (spot 312) (both higher abundant in endosperm 11WAP, Fig. 5).

3.4.4. Storage reserves in Cyclamen seed

One protein spot was identified as 7S globulin (spot 416) which had been previously described in *Cyclamen* endosperm [17], but no further classical storage proteins were identified in the mass spectrometric analyses. The enzyme xyloglucan endotransglycosylase (XTH) which is involved in the production of xyloglucans, the special storage carbohydrates of *Cyclamen*, was also identified at the same spot as previously described by Winkelmann et al. [17]. This spot increased in volume from 7 to 11 WAP.

Several proteins were identified that play a role in the fatty acid/lipid pathway. They are 3-oxoacyl-[acyl-carrier-protein] synthase I, Enoyl-ACP reductase, GDSL esterase/lipase, and wax synthase. These results suggest that lipids are storage reserve in *Cyclamen*. Wax synthase (spot 427) and 3-oxoacyl-[acyl-carrier-protein] synthase I (spot 354) were 1.5 fold higher abundant at 7 WAP while Enoyl-ACP reductase (spot 394) was 1.5 fold higher abundant at 11 WAP.

A large proportion of the proteins identified belong to the enolase group (27 spots). These include both enolase 1 and 2 as described in Rode et al. [18]. In the comparison of 11 and 7 WAP groups, 6 enolase spots were significantly higher abundant (at least 1.5-fold) in 7 WAP while 9 enolase spots were significantly higher abundant in 11 WAP. The enolase spots that were higher abundant in 11 WAP correspond to the 'small enolases' which were first described for *Cyclamen* by Rode et al. [18]. Truncated forms of enolase lacking the catalytic groups, found in high abundance in *C. persicum* zygotic embryos, have been proposed to be candidates for a new group of storage proteins so far not described in plants [18].

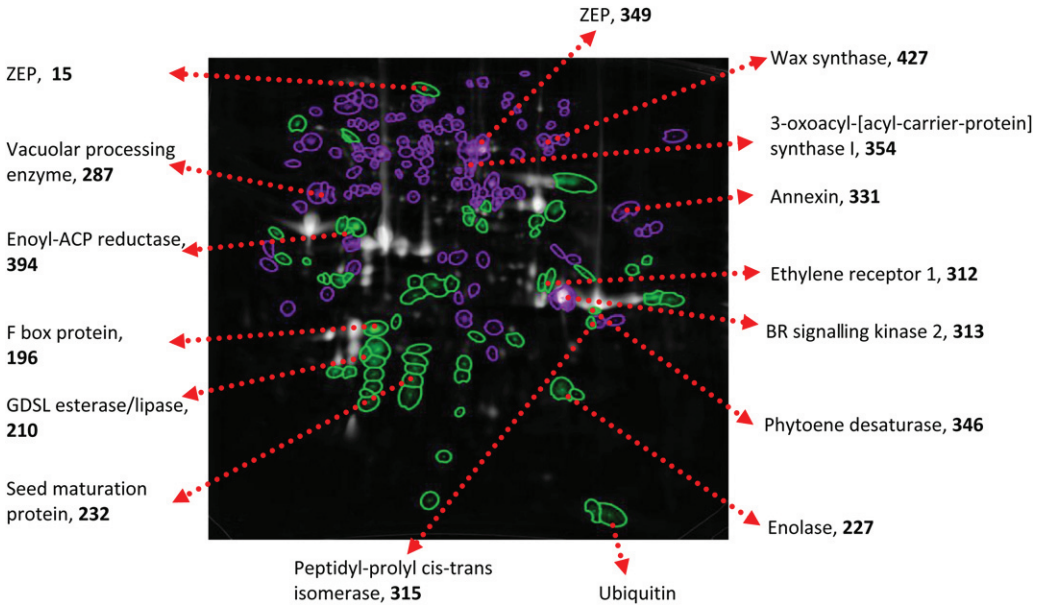


Fig. 4. Proteins of significantly changed abundance between the 7 and 11 WAP endosperm proteomes. Green spots were at least 1.5-fold more abundant at 11 WAP and purple spots were at least 1.5-fold more abundant at 7 WAP. Spots with similar abundance in both endosperm stages appear in white. Proteins discussed in the text are indicated. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of the article.)

4. Discussion

4.1. Key events in endosperm development in C. persicum

After several nuclear divisions, the syncytium stage of endosperm development was clearly visible at 3 WAP, while cellularization of the endosperm was observed from 4 WAP (Fig. 2). This correlates with Hoenemann et al. [25] who also reported cellularization of *Cyclamen* seed at 30 days after pollination. Endosperm cellularization was observed to begin from the periphery of the seed, by formation of cell walls in the multinucleate cytoplasm, characterized by addition of new layers of cells until closure in the centre. This indicates that *Cyclamen* also undergoes nuclear type of endosperm development as observed in many cereals and legumes [4,32,33]. In *Arabidopsis*, cellularization was reported to begin from the micropylar domain when the embryo reaches the globular stage of development [34]. This could also be true for *Cyclamen* as the globular embryo was first observed at 3 WAP while cellularization of the endosperm was seen to begin from 4 WAP. Endosperm cellularization appeared to be complete between 7 and 8 WAP. After cellularization, the endosperm was observed to increase in size through cell division and cell elongation as was previously reported

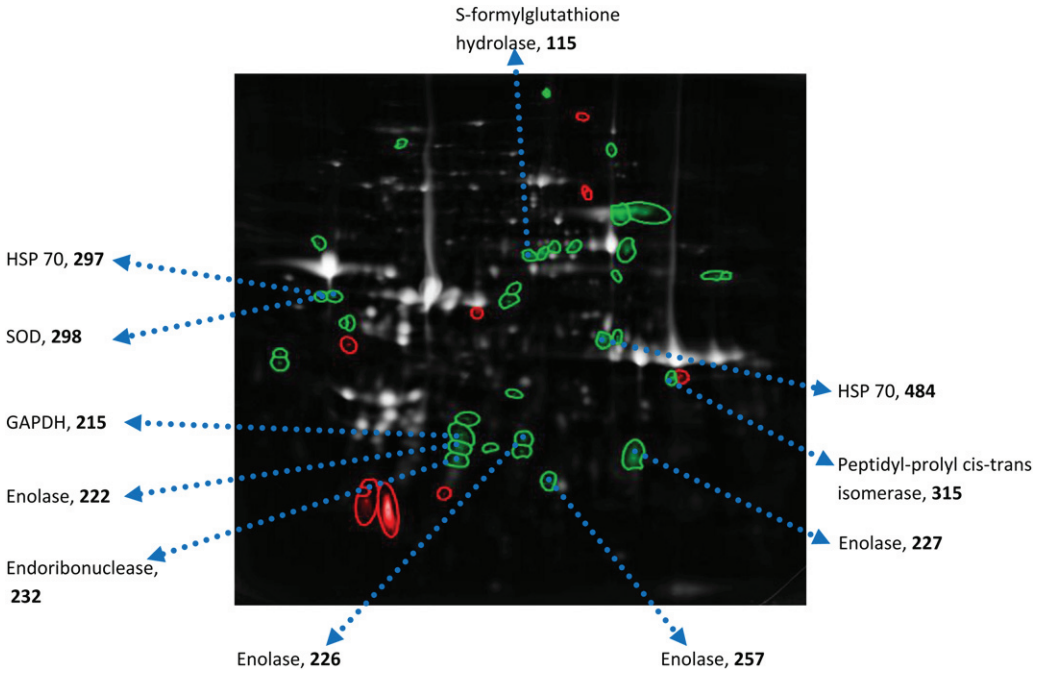


Fig. 5. Proteins of significantly changed abundance between the 9 and 11 WAP endosperm proteomes. Green spots (36) were at least 1.5-fold more abundant at 11 WAP and red spots (9) were at least 1.5-fold more abundant at 9 WAP. Spots with similar abundance in both endosperm stages appear in white. Proteins discussed in the text are indicated. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of the article.)

in a flow cytometric study by Schmidt et al. [35] and by 11 WAP, the endosperm appeared to be fully differentiated consisting of a compact mass of cells.

Small globular zygotic embryos were observed at 3 WAP within the multinucleate endosperm. In a histological analysis of *C. persicum* zygotic embryos, globular embryos were first observed at 30 days after pollination [25]. At 5 WAP, a tiny torpedo-shaped embryo was detected and from 9 to 11 WAP, mature torpedo-shaped embryos were observed.

The endosperm cell wall was observed to be thick, characterized by the deep blue colour of the FCA stain (Fig. 2). In *C. purpurascens*, using electron microscopy, the endosperm cell wall was reported to be 'pitted' with unevenly thickened cell walls and dark hollows on the inside cell wall surface [36]. Endosperm cell walls contain little cellulose and large amounts of hemicelluloses, e.g. xyloglucan, which has also been reported to be a storage polysaccharide in *Cyclamen* [16,17]. These specialized cell walls are an important source of carbohydrates during germination, and they function in water retention, signalling, mechanical protection of the embryo and dormancy [34].

4.2. High similarity of zygotic embryo and endosperm proteomes

There was high similarity level (186 of 490 spots) in protein pattern observed between the zygotic embryo and endosperm proteome (Suppl. Fig. S1). These included the 7S and 11S globulin, the small enolases and XTH enzyme among others. The same similarity between the endosperm and zygotic embryo has also been reported for tomato [37] and *Jatropha curcas* [38]. These indicate the great similarity in regard to function and metabolism between the zygotic embryo and endosperm. Additionally, the existing differences still indicate the novel role played by each of these tissues. However, it has to be stressed, that the applied gel-based proteomic approach considered only proteins of relatively high abundance.

4.3. Decreasing cell division and primary metabolism is in line with previous reports of seed development in Cyclamen

A decrease in abundance of cell division related proteins was observed between 7 and 9/11 WAP pointing to a decrease in cell division activity. Gallardo et al. [39] reported a decreased abundance of annexin before accumulation of major storage proteins in *M. truncatula*. Schmidt et al. [35] using flow cytometry likewise showed the mitotic activity of *C. persicum* endosperm to be high at 28 days after pollination, decreasing sharply up to 49 days and remaining relatively low up to 119 days after pollination. The decline in mitotic activity is an indication that cell growth occurring at these stages is a result of cell elongation and expansion, characteristic of seed filling stage.

Six enzymes were identified that are involved in TCA cycle, including 2-oxoacid dehydrogenase family protein, aconitase, isocitrate dehydrogenase, dihydrolipoamide s-acetyltransferase, malic enzyme and succinate dehydrogenase. These proteins were decreasing in abundance from 7 to 11 WAP. Mechin et al. [22] also reported a similar trend for TCA cycle enzymes in maize endosperm, between 4 and 21 days after pollination. Furthermore, pyruvate dehydrogenase which catalyses conversion of pyruvate to acetyl-CoA, linking glycolytic pathway to TCA cycle [40], was also decreasing in abundance from 7 to 11 WAP. This is an indication that in later stages of seed development, a proportion of the pyruvate synthesized by glycolysis does not join the TCA cycle. It is important to note that pyruvate can be converted back to carbohydrates through gluconeogenesis, to fatty acids through acetyl-CoA or used in amino acid metabolism. This decrease in TCA cycle enzymes in later stages of seed development could suggest that pyruvate synthesized at later stages of seed development is used for accumulation of seed storage reserves.

4.4. Storage proteins in Cyclamen

The endosperm is a storage tissue, characterized by the accumulation of storage compounds in high amounts. Indeed, in this study we could identify 7S globulin and the putative storage protein group of small enolases to be especially high abundant in the latest stages of endosperm development (11 weeks > 9 weeks > 7 weeks). This is in accordance to Winkelmann et al. [17] and Rode et al. [18,41] who had shown these proteins in endosperm and matured embryos of *Cyclamen*, respectively. Reinhardt [42] illustrated that *Cyclamen* seeds accumulate all three categories of energy rich molecules, such as starch, fatty acids and proteins in comparable ratios. Concerning this fact, there should be more storage proteins in the endosperm regarding quality and quantity than those two protein species identified in the recent study. The lack of storage protein identification could be due to technical reasons. The proteins could not be purified from the tissue or resolved by the gel e.g. due to their hydrophobic character (e.g. the major seed storage proteins in grasses, prolamines are soluble in 60–70% ethanol [9]) or size (proteins with a higher molecular weight than 150 kDa and lower than 5 kDa cannot be resulted by the gel system used here). In addition, the lack of sequence information for the non model organism *Cyclamen* is a bottleneck for the identification of species specific storage proteins. An enrichment of low molecular proteins in the later stages of endosperm development could be an indicator of "small" units of so far unknown storage protein complexes.

4.5. Other storage compounds in Cyclamen

Besides the above discussed storage proteins that accumulate in protein bodies [42], *Cyclamen* seeds contain carbohydrates and fatty acids/lipids as further storage reserves as demonstrated in thin sections by staining [42]. Regarding carbohydrates, xyloglucans, the special storage carbohydrates found in secondary cell walls of *Cyclamen* [16] were most likely assembled in endosperm cells in the later phase of seed development by the enzyme xyloglucan endotransglycosylase (XTH). Since XTHs had been identified before in endosperm tissue of another genotype [17], this finding proves the reproducibility and stability of the proteomic data. Xyloglucans are hemicelluloses being present in the primary cell wall of somatic cells, but also serving as densely packed and easy to mobilize carbohydrate stores in endosperm cell walls, e.g. in *Myrsinaceae* to which also the genus *Cyclamen* belongs [34].

Other evidence for the important role of carbohydrates as storage reserves can be deduced from the high percentage (34%) of identified endosperm proteins that are involved in carbohydrate metabolism. Sugars not only serve as carbon and energy sources, but also as signalling molecules, osmotic regulators [43] and membrane protectants during desiccation stress [44]. Sugars are considered to be supplied by the endosperm to the growing embryo. This requires sugar transporters one type of which – sugar carrier C – was identified in the course of the present study. Sugar transporters are membrane-bound and highly hydrophobic and generally not separable by 2D IEF/SDS PAGE. However, the phenolic protein extraction used for *Cyclamen* endosperm tissue led to their detection in four different spots (313, 314, 326, and 346). The exact localization and direction of transport needs further investigations.

Cyclamen endosperm serves as storage organ containing proteins or amino acids and carbohydrates providing energy, carbon, nitrogen and further components to the embryo during later germination. Moreover, fatty acids/lipids are storage reserves in *Cyclamen* seed as demonstrated by Reinhardt [42] by the detection of lipid

bodies using sudan black staining and supported in this study by identification of enzymes involved in fatty acid/lipid metabolism.

4.6. Stress and ABA signalling are important for seed development

Several enzymes involved in stress regulation processes were found to be induced during the analysed time course of endosperm development. This list of proteins includes proteins which generally function in cellular stress response (e.g. universal stress proteins; spots 68, 70, 140, and 181) as well as the predominant abundant heat shock proteins of different classes, which are involved in protein folding and processing mechanisms (molecular chaperones) as well as stress-related protein structural alterations. In addition, they contribute to processes that have been associated with seed longevity, such as thermotolerance, tolerance to embryo desiccation or oxidative stress resistance [45]. Next to several HSP members, enzymes involved in detoxification of reactive oxygen species (ROS) showed also predominant values of abundance, such as catalases (spots 59, 61, 426, and 428), peroxiredoxin type II (spot 179), thioredoxin peroxidase 1 (spot 176), manganese/copper/zinc and iron superoxide dismutases (spots 312; 194, 213, 298, and 422), glutathione reductase (spot 350), S-formylglutathione hydrolase (spot 115), apoptosis-related protein (spot203), aluminium induced protein (spot 127), anti-oxidant 1 (spot 262) and osmotin-like proteins (spots 306 and 478). Thus, these oxidative stress responsive proteins play a major role during endosperm development in *C. persicum*. For the induced catalases and peroxiredoxins, antioxidant-responsive elements (ARE) core sequences were identified in the corresponding promoter regions of the *Arabidopsis* and maize genes and putative protein-DNA interactions may allow fast responses to altered ROS levels [46,47]. Future studies will aim at gaining insights into the role of ROS during embryogenesis and endosperm development in *C. persicum*.

In addition, the Arabidopsis *AtPER1* and its homologue from barley, *Per1*, have been described to be dormancy-related and thus highly involved in seed physiology during desiccation and also resumption of respiration upon rehydration of germinating seeds, providing protection against ROS [48,49]. On the other hand, more recent results indicate that these peroxiredoxins function as molecular chaperones under oxidative stress conditions via inhibition of germination rather than regulators of seed dormancy [50,51]. However, these proteins are specific for seed physiology since they are localized to embryo, endosperm and aleurone cells [51].

Although the induction patterns of the protein members of the anti-oxidative system are quite diverse during the analysed stages of *C. persicum* endosperm development, by trend most enzymes are higher abundant at 11 WAP than at the previous 9 and 7 WAP stages. This is true for the superoxide dismutases, glutathione reductase, peroxiredoxin, thioredoxins and anti-oxidant 1. In contrast, catalases and apoptosis-related protein are higher abundant at 7 WAP. Interestingly, the enzymes related to ABA biosynthesis show a similar diversity of abundance patterns. In general, endogenous ABA acts as a positive regulator for the late maturation program in embryogenesis while at the same time is a negative regulator of the germination programs [52]. The expression of peroxiredoxins in endosperm and embryo from the cotyledon stage of development are found to be ABA-dependent via ABA-responsive elements (ABRE), following their induction via ARE promoter elements at earlier stages [53].

In conclusion, this study has identified the time course of the major cellular events during endosperm development in *C. persicum*. Proteomic analyses supported the histological as well as previous flow cytometric studies and revealed several insights into major metabolic changes and storage reserve synthesis. Accomplished by metabolomic analyses the in vitro system of somatic embryogenesis may be optimized by adjusting culture medium and culture conditions in a way that mimics the situation in the endosperm. Therefore, the application of abscisic acid for maturation of somatic embryos should be tested and optimized in terms of concentration and time course. Secondly the role of reactive oxygen species during development of somatic embryos will be investigated in future studies.

Acknowledgements

The authors thank the DAAD (German Academic Exchange Service) for financial support of Jenniffer Mwangi, Kathrin Lindhorst for excellent technical assistance, Dr. Annette Hohe and Katja Krüger of Leibniz-Institute of Vegetable and Ornamental Crops in Erfurt for support with respect to histological analyses.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.plantsci.2012.11.004.

References

[1] F. Corbineau, N. Neveur, D. Come, Seed germination and seedling development in *Cyclamen persicum*, Ann. Bot. 63 (1989) 87–96.
[2] M.A. Lopes, B.A. Larkins, Endosperm origin, development, and function, Plant Cell 5 (1993) 1383–1399.
[3] P.A. Sabelli, B.A. Larkins, The development of endosperm in grasses, Plant Physiol. 149 (2009) 14–26.
[4] A. Linkies, K. Graeber, C. Knight, G. Leubner-Metzger, The evolution of seeds, New Phytol. 186 (2010) 817–831.
[5] G. Melkus, H. Rolletschek, R. Radchuk, J. Fuchs, T. Rutten, U. Wobus, T. Altmann, P. Jakob, L. Borisjuk, The metabolic role of the legume endosperm: a non-invasive imaging study, Plant Physiol. 151 (2009) 1139–1154.
[6] F. Berger, P.E. Grini, A. Schnittger, Endosperm: an integrator of seed growth and development, Curr. Opin. Plant Biol. 9 (2006) 664–670.
[7] R.C. Brown, B.E. Lemmon, The developmental biology of cereal endosperm, in: O.A. Olsen (Ed.), Endosperm, Plant Cell Monographs, vol. 8, Springer-Verlag, Berlin, 2007, pp. 1–20.
[8] L.C. Hannah, M. James, The complexities of starch biosynthesis in cereal endosperms, Curr. Opin. Biotechnol. 19 (2008) 160–165.
[9] P.R. Shewry, J.A. Napier, A.S. Tatham, Seed storage proteins: structure and biosynthesis, Plant Cell 7 (1995) 945–956.
[10] I.A. Graham, Seed storage oil mobilization, Annu. Rev. Plant Biol. 59 (2008) 115–142.
[11] F.D. Boesewinkel, F. Bouman, The seed: structure and function, in: J. Kigel, G. Galili (Eds.), Seed Development and Germination, Marcel Dekker, New York, 1995, pp. 1–25.
[12] G. Wicart, A. Mouras, A. Lutz, Histological study of organogenesis and embryogenesis in *Cyclamen persicum* Mill. tissue cultures: evidence for a single organogenic pattern, Protoplasma 119 (1984) 159–167.
[13] E. Kiviharju, U. Tuominen, T. Törmälä, The effect of explants material on somatic embryogenesis of *Cyclamen persicum* Mill., Plant Cell Tissue Organ Cult. 28 (1992) 178–194.
[14] H.G. Schwenkel, T. Winkelmann, Plant regeneration via somatic embryogenesis from ovules of *Cyclamen persicum* Mill., Plant Tissue Cult. Biotechnol. 4 (1998) 28–34.
[15] T. Winkelmann, Clonal propagation of *Cyclamen persicum* via somatic embryogenesis, in: S.M. Jain, S.J. Ochatt (Eds.), Protocols for In Vitro Propagation of Ornamental Plants, Methods in Molecular Biology, vol. 589, Springer-Verlag, Berlin, 2010, pp. 281–290.
[16] I. Braccini, C. Hervé du Penhoat, V. Michon, R. Goldberg, M. Clochard, M.C. Jarvis, Z.H. Huang, D.A. Gage, Structural analysis of *Cyclamen* seed xyloglucan oligosaccharides using cellulose digestion and spectroscopic methods, Carbohydr. Res. 276 (1995) 167–181.
[17] T. Winkelmann, D. Heintz, A. Van Dorsselaer, M. Serek, H.P. Braun, Proteomic analyses of somatic and zygotic embryos of *Cyclamen persicum* Mill. reveal new insights into seed and germination physiology, Planta 224 (2006) 508–519.
[18] C. Rode, S. Gallien, D. Heintz, A. Van Dorsselaer, H.P. Braun, T. Winkelmann, Enolases: storage compounds in seeds? Evidence from a proteomic comparison of zygotic and somatic embryos of *Cyclamen persicum* Mill., Plant Mol. Biol. 75 (2011) 305–319.
[19] J.K.C. Rose, S. Bashir, J.J. Giovannoni, M.M. Jahn, R.S. Saravanan, Tackling the plant proteome: practical approaches, hurdles and experimental tools, Plant J. 39 (2004) 715–733.
[20] F. Hochholdinger, M. Sauer, D. Dembinsky, N. Hoecker, N. Muthreich, M. Saleem, Y. Liu, Proteomic dissection of plant development, Proteomics 6 (2006) 4076–4083.

[21] T. Takáč, T. Pechan, J. Šamaj, Differential proteomics of plant development, J. Proteomics 74 (2011) 577–588.
[22] V. Mechin, C. Thevenot, M. Le Guilloux, J.L. Prioul, C. Damerval, Developmental analysis of maize endosperm proteome suggests a pivotal role for pyruvate orthophosphate dikinase, Plant Physiol. 143 (2007) 1203–1219.
[23] O.L. Franco, P.B. Pelegrini, C.P.C. Gomes, A. Souza, F.T. Costa, G. Domont, B.F. Quirino, M.T. Eira, A. Mehta, Proteomic evaluation of coffee zygotic embryos in two different stages of seed development, Plant Physiol. Biochem. 47 (2009) 1046–1050.
[24] S.B. Xu, H.T. Yu, L.F. Yan, T. Wang, Integrated proteomic and cytological study of rice endosperms at the storage phase, J. Proteome Res. 9 (2010) 4906–4918.
[25] C. Hoenemann, S. Richardt, K. Krüger, A.D. Zimmer, A. Hohe, S.A. Rensing, Large impact of the apoplast on somatic embryogenesis in *Cyclamen* persicum offers possibilities for improved developmental control in vitro, BMC Plant Biol. 10 (2010) 77.
[26] W.J. Hurkman, C.K. Tanaka, Solubilization of plant membrane proteins for analysis by two-dimensional gel electrophoresis, Plant Physiol. 8 (1986) 802–806.
[27] F. Colditz, O. Nyamsuren, K. Niehaus, H. Eubel, H.P. Braun, F. Krajinski, Proteomic approach: identification of *Medicago truncatula* proteins induced in roots after infection with the pathogenic oomycete *Aphanomyces euteiches*, Plant Mol. Biol. 55 (2004) 109–120.
[28] V. Neuhoff, R. Stamm, H. Eibl, Clear background and highly sensitive protein staining with Coomassie Blue dyes in polyacrylamide gels: a systematic analysis, Electrophoresis 6 (1985) 427–448.
[29] V. Neuhoff, R. Stamm, I. Pardowitz, N. Arold, W. Ehrhardt, D. Taube, Essential problems in quantification of proteins following colloidal staining with Coomassie Brilliant Blue dyes in polyacrylamide gels, and their solution, Electrophoresis 11 (1990) 101–117.
[30] J. Klodmann, S. Sunderhaus, M. Nimtz, L. Jänsch, H.P. Braun, Internal architecture of mitochondrial complex I from *Arabidopsis thaliana*, Plant Cell 22 (2010) 797–810.
[31] C. Rode, M. Senkler, J. Klodmann, T. Winkelmann, H.P. Braun, GelMap – a novel software tool for building and presenting proteome reference maps, J. Proteomics 74 (2011) 2214–2219.
[32] S.K. Floyd, W.E. Friedman, Evolution of endosperm developmental patterns among basal flowering plants, Int. J. Plant Sci. 161 (2000) 57–81.
[33] V. Raghavan, Double Fertilization. Embryo and Endosperm Development in Flowering Plants, Springer-Verlag, Berlin, 2006, pp. 152–168.
[34] M.S. Otegui, Endosperm cell walls: formation, composition, and functions, in: O.A. Olsen (Ed.), Endosperm, Plant Cell Monographs, vol. 8, Springer-Verlag, Berlin, 2007, pp. 159–177.
[35] T.H. Schmidt, A. Ewald, M. Seyring, A. Hohe, Comparative analysis of cell cycle events in zygotic and somatic embryos of *Cyclamen persicum* indicates strong resemblance of somatic embryos to recalcitrant seeds, Plant Cell Rep. 25 (2006) 643–650.
[36] M. Morozowska, A. Czarna, M. Kujawa, A.M. Jagodzinski, Seed morphology and endosperm structure of selected species of *Primulaceae*, *Myrsinaceae*, and *Theophrastaceae* and their systematic importance, Plant Syst. Evol. 291 (2011) 159–172.
[37] T. Berry, J.D. Bewley, Seeds of tomato (*Lycopersicon esculentum* Mill.) which develop in a fully hydrated environment in the fruit switch from a developmental to a germinative mode without a requirement for desiccation, Planta 186 (1991) 27–34.
[38] H. Lui, Y.Z. Lui, M.F. Yang, S.H. Shen, A comparative analysis of embryo and endosperm proteome from seeds of *Jatropha curca*, J. Integr. Plant Biol. 51 (2009) 850–857.
[39] K. Gallardo, C. Le Signor, J. Vandekerckhove, R.D. Thompson, J. Burstin, Proteomics of *Medicago truncatula* seed development establishes the time frame of diverse metabolic processes related to reserve accumulation, Plant Physiol. 133 (2003) 664–682.
[40] W.C. Plaxton, The organization and regulation of plant glycolysis, Annu. Rev. Plant Physiol. Plant Mol. Biol. 47 (1996) 185–214.
[41] C. Rode, K. Lindhorst, H.P. Braun, T. Winkelmann, From callus to embryo – a proteomic view on the development and maturation of somatic embryos in *Cyclamen persicum* Mill., Planta 235 (2012) 995–1011.
[42] S. Reinhardt, Vergleichende morphologische-anatomische und physiologische Untersuchung zur Samenanlagenqualitaet von *Cyclamen persicum* Mill. unter besonderer Beruecksichtigung der Saatgutqualitaet, Dissertation, Friedrich-Schiller-University, Jena, Germany, 2006.
[43] L.M. Hill, E.R. Morley-Smith, S. Rawsthorne, Metabolism of sugars in the endosperm of developing seeds of oilseed rape, Plant Physiol. 131 (2003) 228–236.
[44] S.A. Blackman, S.H. Wettlaufer, R.L. Obendorf, A.C. Leopold, Maturation proteins associated with desiccation tolerance in soybean, Plant Physiol. 96 (1991) 868–874.
[45] N. Wehmeyer, E. Vierling, The expression of small heat shock proteins in seeds responds to discrete developmental signals and suggests a general protective role in desiccation tolerance, Plant Physiol. 122 (2000) 1099–1108.
[46] R.B. Aalen, Peroxiredoxin antioxidants in seed physiology, Seed Sci. Res. 9 (1999) 285–295.
[47] A.N. Polidoros, J.G. Scandalios, Role of hydrogen peroxide and different classes of antioxidants in the regulation of catalase and glutathione S-transferase gene expression in maize (*Zea mays* L.), Physiol. Plant. 106 (1999) 112–120.
[48] R.A. Stacy, E. Munthe, T. Steinum, B. Sharma, R.B. Aalen, A peroxiredoxin antioxidant is encoded by a dormancy-related gene, *Per1*, expressed during late development in the aleurone and embryo of barley grains, Plant Mol. Biol. 31 (1996) 1205–1216.
[49] C. Haslekås, R.A. Stacy, V. Nygaard, F.A. Culiáñez-Macià, R.B. Aalen, The expression of a peroxiredoxin antioxidant gene, AtPer1, in *Arabidopsis thaliana* is seed-specific and related to dormancy, Plant Mol. Biol. 36 (1998) 833–845.
[50] R.A. Stacy, T.W. Nordeng, F.A. Culiáñez-Macià, R.B. Aalen, The dormancy-related peroxiredoxin anti-oxidant, PER1, is localized to the nucleus of barley embryo and aleurone cells, Plant J. 19 (1999) 1–8.
[51] S.Y. Kim, S.K. Paeng, G.M. Nawkar, P. Maibam, E.S. Lee, K.S. Kim, D.H. Lee, D.J. Park, S.B. Kang, M.R. Kim, et al., The 1-Cys peroxiredoxin, a regulator of seed dormancy, functions as a molecular chaperone under oxidative stress conditions, Plant Sci. 181 (2011) 119–124.
[52] F. Parcy, C. Valon, M. Raynal, P. Gaubier-Comella, M. Delseny, J. Giraudat, Regulation of gene expression programs during *Arabidopsis* seed development: roles of the *ABI3* locus and of endogenous abscisic acid, Plant Cell 6 (1994) 1567–1582.
[53] C. Haslekås, P.E. Grini, S.H. Nordgard, T. Thorstensen, M.K. Viken, V. Nygaard, R.B. Aalen, ABI3 mediates expression of the peroxiredoxin antioxidant AtPER1 gene and induction by oxidative stress, Plant Mol. Biol. 53 (2003) 313–326.

Publication 4

2.4 Comparative proteomic analysis of early somatic and zygotic embryogenesis in *Theobroma cacao* L.

Alexandre Mboene Noah[1], Nicolas Niemenak[1], Stephanie Sunderhaus[2], Christin Haase[2], Denis Ndoumou Omokolo[1], Traud Winkelmann[3], Hans-Peter Braun[2]

[1]Laboratory of Plant Physiology and Biochemistry, Department of Biological Science, University of Yaounde
[2]Institute of Plant Genetics, Leibniz Universität Hannover
[3]Institute of Floriculture and Woody Plant Science, Leibniz Universität Hannover

Type of authorship:	Co-author
Type of article:	Research article
Share of the work:	25 %
Contribution to the publication:	Performed experiments and MS analysis, analysed data, prepared figures and tables, wrote parts of the manuscript
Journal:	Journal of Proteomics
Impact factor:	3.888
Date of publication:	Published in January 2013
Number of citations: (Google scholar, Nov. 07th, 2015)	7
DOI:	10.1016/j.jprot.2012.11.007
PubMed-ID:	23178419

JOURNAL OF PROTEOMICS 78 (2013) 123–133

Available online at www.sciencedirect.com

SciVerse ScienceDirect

www.elsevier.com/locate/jprot

Comparative proteomic analysis of early somatic and zygotic embryogenesis in *Theobroma cacao* L.

Alexandre Mboene Noah[a,*], *Nicolas Niemenak*[a], *Stephanie Sunderhaus*[b], *Christin Haase*[b], *Denis Ndoumou Omokolo*[a], *Traud Winkelmann*[c], *Hans-Peter Braun*[b]

[a]*Laboratory of Plant Physiology and Biochemistry, Department of Biological Science, Higher Teachers' Training College, University of Yaounde I, P.O. Box 47 Yaounde, Cameroon*

[b]*Institute of Plant Genetics, Leibniz Universität Hannover, Herrenhäuser Str. 2, 30419 Hannover, Germany*

[c]*Institute of Floriculture and Woody Plant Science, Leibniz Universität Hannover, Herrenhäuser Str. 2, 30419 Hannover, Germany*

ARTICLE INFO

Article history:
Received 24 August 2012
Accepted 5 November 2012
Available online 22 November 2012

Keywords:
Theobroma cacao
Embryogenesis
Proteomics
Metabolic pathways

ABSTRACT

Somatic embryogenesis can efficiently foster the propagation of *Theobroma cacao*, but the poor quality of resulted plantlet hinders the use of this technique in the commercial scale. The current study has been initiated to systematically compare the physiological mechanisms underlying somatic and zygotic embryogenesis in *T. cacao* on the proteome level. About 1000 protein spots per fraction could be separated by two-dimensional isoelectric focusing/SDS PAGE. More than 50 of the protein spots clearly differed in abundance between zygotic and somatic embryos: 33 proteins spots were at least 3-fold higher in abundance in zygotic embryos and 20 in somatic embryos. Analyses of these protein spots differing in volume by mass spectrometry resulted in the identification of 68 distinct proteins. Many of the identified proteins are involved in genetic information processing (21 proteins), carbohydrate metabolism (11 proteins) and stress response (7 proteins). Somatic embryos especially displayed many stress related proteins, few enzymes involved in storage compound synthesis and an exceptional high abundance of endopeptidase inhibitors. Phosphoenolpyruvate carboxylase, which was accumulated more than 3-fold higher in zygotic embryos, represents a prominent enzyme in the storage compound metabolism in cacao seeds. Implications on the improvement of somatic embryogenesis in cacao are discussed.

1. Introduction

Cacao (*Theobroma cacao* L) is a tropical tree of the Malvaceae family [1] that has a high economic value because it is the main source of chocolate. Furthermore, there is a growing interest in cacao due to its high content of polyphenols which have been suggested to prevent brain and cardiovascular diseases [2–5]. Despite its utility, the propagation of cacao is still limited because of its allogamous mating system [6] that results in inbreeding depression of parental lines. For this reason, somatic embryogenesis has been developed as an alternative strategy for the large scale propagation of elite

Abbreviations: 2,4-D, 2,4-dichlorophenoxyacetic acid; ICGD, International Cocoa Germplasm Database; PA, Parinari; SE, somatic embryos; TDZ, Thidiazuron; ZE, zygotic embryos.

* *Corresponding author at:* Laboratory of Plant Physiology and Biochemistry, Department of Biological Science, Higher Teachers' Training College, University of Yaounde I, P.O. Box 47 Yaounde, Cameroon. Tel.: +237 75 35 60 76.
E-mail address: noahmbalr@yahoo.fr (A.M. Noah).

1874-3919/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jprot.2012.11.007

genotypes of cacao. Studies carried out to date on somatic embryogenesis in *T. cacao* were focused at the morphological, biochemical and gene expression level. It was the aim of most of these studies to define a medium for optimal induction of embryos [7–12]. Despite all these efforts, plantlets resulting from somatic embryos still exhibit a disturbed development and poor conversion ability. Insufficient accumulation of storage compounds and enzymatic imbalances in somatic embryos were earlier suspected to cause developmental irregularities in cacao [13] and in carrot [14].

The early seedling growth of angiosperms depends both on the storage compounds accumulated during maturation of the embryo and also on the enzymatic system that mobilizes nutrient elements to provide optimal germination. Protein mobilization releases reduced nitrogen needed for growth and development [15]. Sugar compounds such as starch and sucrose act as energy suppliers with lipid reserves [16]. In *Arabidopsis*, the use of metabolic inhibitors revealed that transcription is not required for the completion of germination; the potential of germination consequently is considered to be programmed during embryo and seed maturation [17]. Therefore, maturation is a critical point to evaluate the potential quality and viability of somatic embryogenesis-derived seedlings. In order to establish somatic embryogenesis in cacao as a standard technique in commercial scale, it is of prime interest to investigate the physiological processes that govern maturation with the hope to enhance conversion ability of somatic embryos.

Proteomics is a set of biochemical tools which are used to describe the expression, function, and regulation of the entire set of proteins encoded by an organism. Proteins are molecules that directly influence cellular biochemistry; thus, the proteome is thought to reflect more accurately the cellular state than the transcriptome which only represents an information intermediate during the process of protein biosynthesis [18–20]. The continuing development of MS-based proteomics led to gradually improved capacities of protein separation, identification and quantification. From the pioneer 2-D gel electrophoresis procedures to the most accurate linear quadrupole ion trap–orbitrap system, impressive strides have been achieved in terms of resolution, mass accuracy and throughput [21]. Hence, proteomics, which first was applied to plant biology about two decades ago, has yielded valuable information for the comprehensive understanding of a broad range of physiological processes [20,22–24].

The efficiency of proteomics to analyze physiological processes makes it a promising approach to investigate somatic embryogenesis. This is reflected by the increasing number of proteome studies in the field of somatic embryogenesis. Several plant systems such as *Picea glauca* [25], *Cyclamen persicum* [26–29] and *Phoenix dactylifera* [30] have been investigated. In cacao, proteome studies are still at an early stage [31]. The few studies in cacao based on this technology mostly focused on improving cacao derived products through post-harvest treatments [32–34]. In the present study, two dimensional gel electrophoresis coupled to protein identification by mass spectrometry is used to compare the proteomes of *T. cacao* somatic embryos with their zygotic counterparts at the torpedo stage. New insights into embryogenesis and germination physiology in *T. cacao* are presented which are relevant for the further improvement of this prominent regeneration strategy (somatic embryogenesis).

2. Materials and methods

2.1. Plant material

In the frame of our investigation, zygotic and somatic embryos were investigated by a comparative proteomic analysis. Both types of embryos were investigated at the torpedo stage, a stage comparable to stage IIIz defined by Alemanno et al. 1997 [13]. Studies were carried out on the cacao genotype "PA 150" (PA stands for Parinari, a local name in Peru), which is included in the gene-bank of the Institute of Agricultural Research for Development at Nkolbisson (Yaounde, Cameroon). This genotype is known for its high productivity and its resistance to *Phytophthora* [35].

Zygotic embryos were harvested at the torpedo stage from 12-week-old fruits (Fig. 1A). At this stage the endosperm surrounding the zygotic embryos was liquid (Fig. 1B). The protocol used to produce somatic embryos is taken from Minyaka et al. 2008 [11]. Driver and Kuniyuki [36] mineral solution (known as DKW solution) was used to prepare all culture media. The somatic embryogenesis process is divided into three steps. The first step lasted 14 days: petals and staminodes (starting material) of immature floral buds (Fig. 1C) were cultured in PCG (primary callus growth) medium. This medium was supplemented with 250 mg l^{-1} glutamine, 100 mg l^{-1} myo-inositol, 1 ml l^{-1} DKW vitamin stock (100 mg ml^{-1} myo-inositol, 2 mg ml^{-1} thiamine-HCl, 1 mg ml^{-1} nicotinic acid and 2 mg ml^{-1} glycine), 20 g l^{-1} glucose, 18 μM 2,4 dichlorophenoxyacetic acid (2,4-D) and 45.4 nM TDZ (Thidiazuron). As a second step, developing callus cultures were transferred to SCG (secondary callus growth) medium for 14 days. This SCG medium was supplemented with 0.5 ml l^{-1} DKW vitamin, 20 g l^{-1} glucose, 9 μM 2,4-D and 1.2 μM kinetin. At the third step, calluses were maintained for 21 days in ED (embryo development) medium which is a growth regulator free medium supplemented with 1 ml l^{-1} DKW vitamins, 30 g l^{-1} sucrose and 1 g l^{-1} glucose. Some torpedo embryos were visible after this first cultivation on ED medium, but two additional rounds of subcultures of three weeks each were required to increase the number of embryos.

2.2. Protein extraction

Three independent extractions (biological replicates) for each embryo type were carried out. The entire embryo was used as starting material for total protein extractions using a phenol based procedure [37]: 100 mg of harvested plant material was shock frozen in liquid nitrogen, then pulverized in a bead mill and subsequently dissolved in extraction buffer (700 mM sucrose, 500 mM Tris, 50 mM EDTA, 100 mM KCl, 2% (v/v) β-mercaptoethanol and 2 mM PMSF, pH adjusted to 8.0). The dissolved plant material was incubated on ice for 30 min and saturated phenol (pH 6.6/7.9; Amresco, Solon, USA) was added. After incubation, several centrifugation steps were carried out according to Colditz et al. [37]. Finally, proteins

JOURNAL OF PROTEOMICS 78 (2013) 123–133

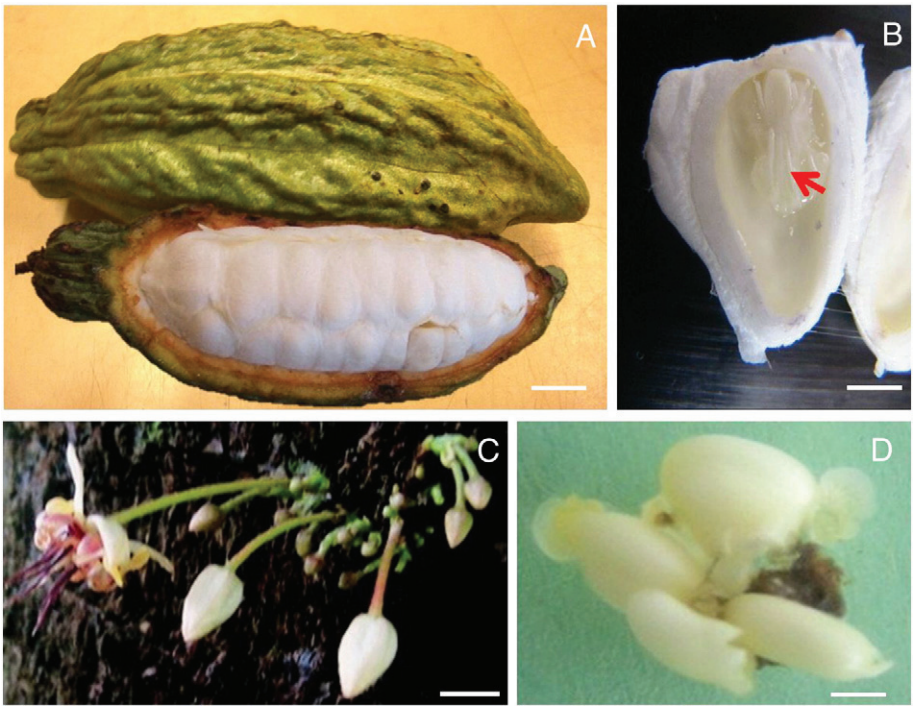


Fig. 1 – Zygotic and somatic embryos of *T. cacao*, genotype PA 150. (A) 12-week-old fruit, a closed fruit beside an opened (bar=1.5 cm). (B) Longitudinal section of one seed displaying a zygotic embryo (labeled with a red arrow) embedded in liquid endosperm (bar=0.25 cm). (C) Immature floral buds used as starting material for somatic embryogenesis (bar=0.5 cm). (D) Somatic embryos after 6 weeks of differentiation on solid medium free of growth regulator (bar=0.5 cm).

were precipitated with 100 mM ammonium acetate in methanol at −20 °C overnight. The resulting protein pellet was washed thrice at 4 °C with 100 mM ammonium acetate in methanol and finally once with 80% acetone.

2.3. *Two-dimensional IEF/SDS–PAGE and protein staining*

Three independent electrophoresis runs were performed from each type of material. Proteins were separated in the first gel dimension according to their isoelectric point. The soluble protein fraction extracted from 100 mg fresh embryo tissue was resuspended in 350 μl rehydration buffer encompassing 8 M urea, 2 M thiourea, 2% (w/v) CHAPS, 100 mM DTT, 0.5% (v/v) IPG-buffer (pH 3–11, GE Healthcare, Freiburg, Germany) and a trace of bromophenol blue. The dissolved proteins were then loaded on Immobiline dry strip gels (18 cm, pH 3–11 non-linear, GE healthcare, Freiburg, Germany), overlaid with mineral oil and electrofocused for 24 h with voltages set from 30 to 8000 V according to Mihr and Braun [38] using an IPGphor apparatus (GE Healthcare, Munich, Germany). Following IEF, the IPG strips were immediately treated with equilibration buffer (50 mM Tris–HCl, pH 8.8, 6 M urea, 30% [v/v] glycerin and 2% [w/v] SDS) containing 1% DTT for 15 min, followed by another treatment for 15 min with equilibration buffer containing 2.5% iodoacetamide. The second electrophoretic dimension was performed on 12% polyacrylamide Tricine–SDS gels [39,40]. Gels were run at 30 mA for 20 h using the Biorad Protean II-XL gel apparatus (Biorad, Munich, Germany). After gel electrophoresis runs, the gels were fixed with 10% (v/v) acetate in 40% (v/v) methanol for at least 2 h and stained with Coomassie Blue CBB G-250 (Merck, Darmstadt, Germany) [41,42].

2.4. *Quantitative analysis of zygotic and somatic embryo proteomes*

Three gels from each type of the studied material were considered for the quantitative analysis. Evaluation of 2D gels was carried out according to Berth et al. [43]. Briefly, scans of Coomassie stained gels were analyzed with DELTA 2D software, version 4.0 (Decodon, Greifswald, Germany). Spots were detected automatically and some minor corrections were performed manually. A student's *t*-test (p-value≥0.05) based on the normalized relative spot volume was applied to determine significant alteration in spot pattern between zygotic and somatic embryo proteomes. Spots of at least 3 fold alteration in abundance with a p-value≥0.05 were considered to represent significant differences in protein levels and analyzed in more detail.

2.5. *Mass spectrometry analysis and physiological clustering of identified proteins*

Spots of interest were picked manually from the Coomassie stained gels using a GelPal Protein Excision System (Genetix, Queensway, U.K.) and in-gel proteolytic digestion with trypsin was carried out as previously described by Klodmann et al. [44]. In brief, acetonitrile was used for dehydration of gel pieces which were then processed for cysteine alkylation through two

successive incubations of 30 min, first at 56 °C in 20 mM DTT and secondly at room temperature in 55 mM iodoacetamide. The proteolytic digestion was based on incubation of the dehydrated gel pieces overnight at 37 °C in trypsin solution (2 μg trypsin [Promega, Madison, WI, USA] per ml resuspension buffer [0.1 M NH_4HCO_3]). Afterwards, tryptic peptides were extracted with acetonitrile at 37 °C; basic peptides were extracted with 5% formic acid at the same temperature.

Mass spectrometric analysis of resulting tryptic peptides was performed with the EASY-nLC System (Proxeon) coupled to a MicroTOF-Q II mass spectrometer (Bruker, Daltonics). Proteins were identified using the MASCOT search algorithm (www.matrixscience.com) against UniProtKB (www.expasy.org) and NCBInr (www.ncbi.nig.gov) databases. Only proteins with a ppm below 20, a MASCOT score above 50 and coverage by at least 2 unique peptides were considered and functionally classified according to the KEGG PATHWAY Database (http://www.genome.jp/kegg/pathway.html). Categories were adjusted for seed metabolism by adding functional groups e.g. “stress response” and “defense”.

3. Results

Somatic embryo production in *T. cacao* is nowadays well mastered as demonstrated by the studies of Li et al. 1998 [7], Minyaka et al. 2008 [11] and Niemenak et al. 2008 [12]. The most challenging issue at the moment is the conversion of these embryos into healthy plantlets. Any improvement in this respect would represent an important step towards usage of somatic embryogenesis in *T. cacao* on a commercial scale. The present study aims at elucidating the physiological mechanisms that govern embryogenesis of this economically important plant species. This issue is addressed by a comparative proteomic-based investigation of zygotic and somatic embryos. Both types of embryos were cultivated up to an equivalent developmental stage: the torpedo stage (Fig. 1B, D). Somatic embryos showed an asynchronous development and were frequently misshaped, while their zygotic counterparts appeared morphologically well-shaped and homogeneous. Only well-shaped somatic embryos were processed for proteomic analyses.

3.1. *Two-dimensional analysis of proteomes from zygotic and somatic embryos*

Total protein fractions extracted from somatic and zygotic embryos were separated by 2D IEF/SDS–PAGE. Visual inspection of the gels revealed that both fractions were composed of similar components (spots on the 2D gels). However, individual protein abundances largely differed between the two fractions. Overall, about 1000 spots were detectable per gel in somatic and zygotic embryos (Fig. 2A). To accurately evaluate quantitative changes in protein abundance between the two types of embryos, protein gels were analyzed by the DELTA-2D software package. Firstly, three biological repetitions for each type of embryo were used to calculate average gels. Secondly, the average gels for somatic and zygotic embryos were compared. This comparison revealed that 33 protein spots were at least 3-fold more abundant in the zygotic embryo proteome, whereas 20 protein spots were at least 3-fold higher in abundance in the somatic embryo proteome (Fig. 2B).

3.2. *Protein identification*

Protein spots of at least 3-fold higher abundance in one or the other type of embryo were used for protein identifications by mass spectrometry (53 distinct spots). For 32 spots, protein identifications were successful (58%). Cacao is a non-model plant; its genome has just been described [45]. Hence, only a limited number of cacao proteins have been correctly annotated up to date. Table 1 shows proteins of increased abundance in the zygotic embryo proteomes. Those of increased abundance in somatic embryo proteomes are presented in Table 2. Detailed information on MS parameters of each identified protein is given in Table 3 (supplementary material). Within the 32 protein spots successfully analyzed, 68 distinct proteins were identified, 30 of which belong to the SE proteomes and 38 to the ZE proteomes. Within 19 spots, more than one protein was identified. This indicates that proteins overlapped during electrophoretic separation (Table 3). Functional classification of the identified proteins, which was carried out in accordance to KEGG PATHWAY Database, allowed the distinction of seven functional categories: (1) carbohydrate metabolism, (2) energy metabolism, (3) amino acid metabolism, (4) genetic information processing, (5) cellular processes, (6) stress response and (7) defense. Within these functional categories, carbohydrate metabolism, genetic information processing and amino acid metabolism likewise were of significant dimension in both proteomes but represented by different proteins. Proteins of stress response and energy metabolism were of higher abundance in SE. Two protein species of an ABC transporter were of increased abundance in ZE proteomes. In both proteomes, three protein categories were especially important, which overall include 56% of all identified proteins. These categories are: genetic information processing (21 proteins [14 proteins in ZE and 7 in SE]), carbohydrate metabolism (11 proteins [6 proteins in ZE and 5 in SE]) and stress response (7 proteins) (Table 1 and 2).

A few spots were found to include identical proteins. These proteins are considered to represent different protein species resulting from posttranslational modifications which introduce changes of molecular weight (MW) and/or isoelectric point (pI). Examples are a class III peroxidase (spots 29 and 31), ABC transporter E family member 2 (spots 2 and 25), trypsin inhibitor (spots 44 and 45), 21 kDa seed protein (spots 44 and 45) and phosphoglycerate dehydrogenase (spots 37 and 38).

Proteins involved in genetic information processing and carbohydrate metabolism represent 30 and 16% of the proteins which are of increased abundance in one of the two types of embryos. In the ZE proteome, except for sucrose synthase (spot 51), all other proteins are glycolytic enzymes: phosphoenolpyruvate carboxylase (spots 50), phosphoglucomutase (spot 52), pyrophosphate-fructose 6-phosphate 1-phosphotransferase (spot 53), pyruvate decarboxylase isozyme 2-like (spot 52) and pyruvate kinase (spot 38) (Table 1). In the SE proteome, proteins of carbohydrate metabolism belong to different pathways such as (i) tricarboxylic acid cycle (1 protein), (ii) glyoxylate cycle (1 protein), (iii) pyruvate metabolism (2 proteins) and

JOURNAL OF PROTEOMICS 78 (2013) 123–133

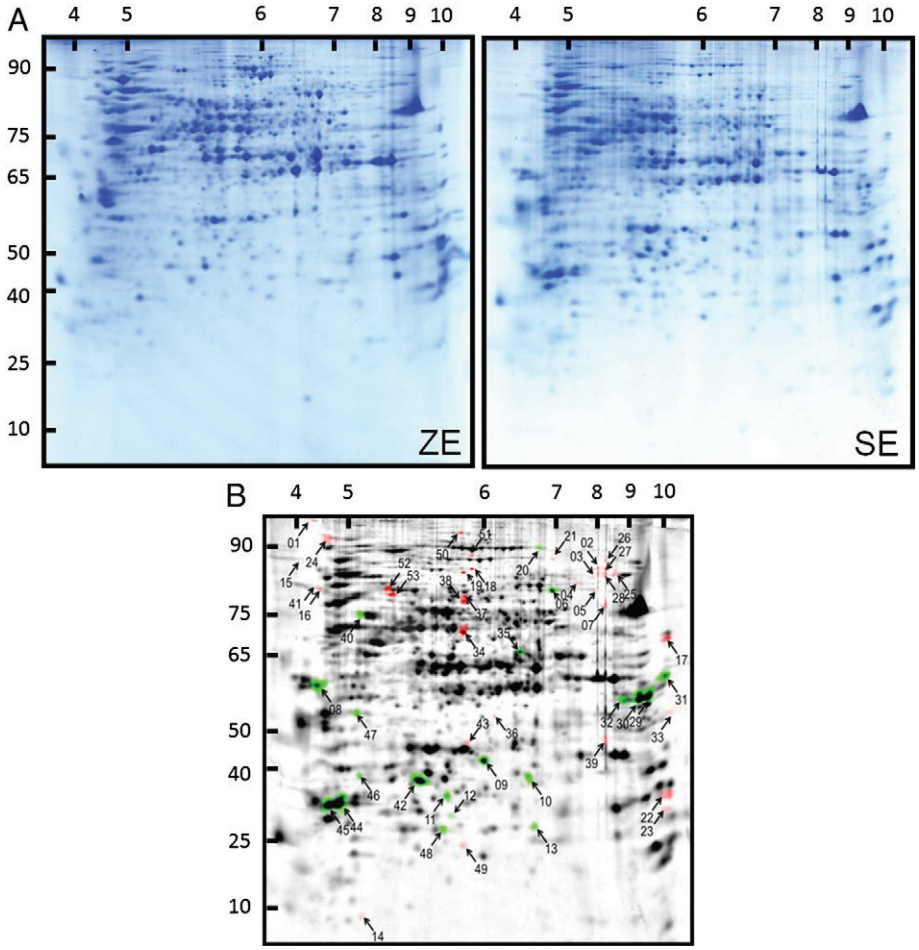


Fig. 2 – Proteome of *T. cacao* embryos. (A) IEF–SDS PAGE of total soluble protein fractions extracted from 100 mg fresh embryo tissue. Molecular masses range from 5 to 100 kDa are given on the left side. Isoelectric points, pI 3–11 are given above the gel image. Gels were stained with CBB-G 250. ZE=Zygotic embryos. SE=Somatic embryos. (B) Overlay image of IEF–SDS PAGE of zygotic and somatic embryo proteomes using the DELTA 2D software. Red-labeled spots were at least 3-fold more abundant in zygotic embryos and green-labeled spots were at least 3-fold more abundant in somatic embryos.

(iv) ascorbate metabolism (1 protein) (Table 2). Five proteins in the zygotic embryo proteome were considered of special interest regarding their physiological function. These proteins are phosphoenolpyruvate carboxylase (spot 50), sucrose synthase (spot 51), aspartic proteinase (spot 17), ubiquitin (spot 50) and histone H2A (spot 22). They were respectively 4.82, 4.02, 3.04, 4.82 and 8.04 fold more abundant than in SE proteomes. In the SE proteome, besides proteins involved in carbohydrate metabolism and genetic information processing, stress related proteins were of special abundance (7 proteins). Another interesting finding within the SE proteome was the relatively high abundance of a trypsin inhibitor found in spots 44 and 45. These two proteins species were 9.67 and 3.84 fold more abundant than in the ZE proteome.

4. Discussion

T. cacao, due to its allogamous status [6], requires an efficient vegetative propagation system. Somatic embryogenesis was reported the first time in the study of Esan 1975 [46]. It then has been improved successively over the two last decades [7,11,12]. Currently, this clonal technique is considered to represent a prominent in vitro regeneration system for cacao. Several advantages are associated with this technique such as multiplication of elite genotypes, mass propagation [12] and the development of genetic transformation systems [47]. Unfortunately, the complexity of this developmental system and the numerous unknown factors that control it discouraged its systematical exploration. The current limitation encountered with this propagation system especially is the poor plantlet growth and low survival rate after transplanting [48]. Proteomics, which nowadays became a powerful biochemical technique to unravel in depth physiological processes, has been applied to embryogenesis in several plant systems [25,26,28–30,49]. These studies led to valuable information. In this study, we applied a comparative proteomic approach to describe physiological differences between zygotic and somatic embryos that may explain the poor development of plantlets derived from somatic embryogenesis in *T. cacao*. Three independent electrophoresis runs from

Table 1 – Proteins of increased abundance in the zygotic embryo proteomes.

Protein [a]	Category [b]	Physiological function [b]	Abundance [c]	ID [d]
Sucrose synthase 1	**Carbohydrate metabolism**	**Starch and sucrose metabolism**	**4.02**	**51**
Phosphoenolpyruvate carboxylase	**Carbohydrate metabolism**	**Glycolysis/gluconeogenesis**	**4.82**	**50**
Phosphoglucomutase, cytoplasmic-like	Carbohydrate metabolism	Glycolysis/gluconeogenesis	3.29	52
Pyrophosphate-fructose 6-phosphate 1-phosphotransferase subunit beta-like	carbohydrate metabolism	Glycolysis/gluconeogenesis	3.71	53
Pyruvate decarboxylase isozyme 2-like	**Carbohydrate metabolism**	**Glycolysis/gluconeogenesis**	**3.29**	**52**
Pyruvate kinase, putative	Carbohydrate metabolism	Glycolysis/gluconeogenesis	5.20	38
Phosphoglycerate dehydrogenase	**Amino acid metabolism**	**Glycine, serine and threonine metabolism**	**4.13**	**37**
Phosphoglycerate dehydrogenase	**Amino acid metabolism**	**Glycine, serine and threonine metabolism**	**5.20**	**38**
DNA replication licensing factor MCM3-like protein	Genetic information processing	Replication and repair	4.02	51
Histone H2B like protein	**Genetic information processing**	**Transcription**	**8.04**	**22**
Histone H2A	**Genetic information processing**	**Transcription**	**8.04**	**22**
Splicing factor 3A subunit, putative	Genetic information processing	Transcription	3.04	17
60S ribosomal protein L4-1	Genetic information processing	Translation	3.88	7
40S ribosomal protein S17 (RPS17A)	Genetic information processing	Translation	9.04	23
Elongation factor 1 gamma, putative	Genetic information processing	Translation	3.23	34
Elongation factor	Genetic information processing	Translation	4.02	51
Leucyl-tRNA synthetase, cytoplasmic-like	Genetic information processing	Translation	4.82	50
Lysyl-tRNA synthetase-like	Genetic information processing	Translation	5.85	19
Ubiquitin	**Genetic information processing**	**Folding, sorting and degradation**	**4.82**	**50**
T complex protein	**Genetic information processing**	**Folding, sorting and degradation**	**3.71**	**53**
Aspartic proteinase nepenthesin-1 precursor, putative	**Genetic information Processing**	**Folding, sorting and degradation**	**3.04**	**17**
Plasminogen activator inhibitor 1 RNA-binding protein	Genetic information processing	Folding, sorting and degradation	3.88	7
ABC transporter E family member 2	Cellular processes	Transport and catabolism	4.66	2
ABC transporter E family member 2	Cellular processes	Transport and catabolism	4.80	25

[a] Proteins identified in the current study. Protein name according to the best hit of MASCOT search against SwissProt and NCBI databases. Bold printed proteins are of major interest and discussed in detail in the text. Detailed information on mass spectrometry parameters for each spot analyzed is given in the supplementary material (Table 3).
[b] Functional protein classification according to the KEGG PATHWAY Database. Additional categories are included when necessary.
[c] Protein spot abundance: statistically significant values of at least 3 fold more abundant protein spots are given for the zygotic embryos proteomes (relative spot volume ZE/relative spot volume SE). These protein spots are given in red on Fig. 2B.
[d] Spot ID as given on Fig. 2B.

either SE or ZE material enabled the resolution of about 1000 protein spots. The analysis of scanned gels confirmed the reproducibility of the experiments. Several proteins were found to be of differential abundance between ZE and SE including enzymes involved in protein folding processes. Enzymes of special interest are discussed in the following sections.

4.1. *Carbohydrate metabolism, the backbone physiological pathway in embryogenesis*

Zygotic and somatic embryos both highly accumulate enzymes of carbohydrate metabolism. Only one enzyme involved in sucrose metabolism, sucrose synthase (spot 51), was found in the frame of the current study. The other

Table 2 – Proteins of increased abundance in the somatic embryo proteomes.

Protein [a]	Category [b]	Physiological function [b]	Abundance [c]	ID [d]
Lactoylglutathione lyase	Carbohydrate metabolism	Pyruvate metabolism	4.11	46
Lactoylglutathione lyase	Carbohydrate metabolism	Pyruvate metabolism	5.00	47
Malate synthase, glyoxysomal	**Carbohydrate metabolism**	**Glyoxylate and dicarboxylate metabolism**	**4.50**	**6**
Malate dehydrogenase, glyoxysomal	**Carbohydrate metabolism**	**TCA cycle**	**46.06**	**32**
Putative L-galactose-1-phosphate phosphatase	Carbohydrate metabolism	Ascorbate metabolism	5.00	47
Glutamate dehydrogenase 2	**Amino acid metabolism**	**Alanine, Aspartate and Glutamate Metabolism**	**11.26**	**35**
Eukaryotic translation initiation factor 3 subunit G-B	Genetic information processing	Translation	46.06	32
60S ribosomal protein L18-2	Genetic information processing	Translation	9.67	44
30S ribosomal protein 2, chloroplastic	Genetic information processing	Translation	15.00	42
Trypsin inhibitor	**Genetic information processing**	**Folding, sorting and degradation**	**9.67**	**44**
Trypsin inhibitor	**Genetic information processing**	**Folding, sorting and degradation**	**3.84**	**45**
21 kDa seed protein	**Genetic information processing**	**Folding, sorting and degradation**	**9.67**	**44**
21 kDa seed protein	**Genetic information processing**	**Folding, sorting and degradation**	**3.84**	**45**
NADH-ubiquinone oxidoreductase-related	Energy metabolism	Oxidative phosphorylation	3.84	45
ATP synthase beta subunit 1	Energy metabolism	Oxidative phosphorylation	3.79	40
Flavoprotein wrbA isoform 1	Energy metabolism	Oxidative phosphorylation	4.00	11
Probable protein phosphatase 2C 59	Defense	Defense	5.00	47
Bacterial-induced peroxidase	Stress response	Stress response	18.53	29
Class III peroxidase	Stress response	Stress response	18.53	29
Class III peroxidase	Stress response	Stress response	6.38	31
Peroxidase N1-like	Stress response	Stress response	46.06	32
Peroxidase	Stress response	Stress response	11.26	35
Osmotin-like pathogenesis-related protein	Stress response	Stress response	15.00	42
Probable glutathione S-transferase-like	**Stress response**	**Stress response**	**11.01**	**9**

[a] Proteins identified in the current study. Protein name according to the best hit of MASCOT search against SwissProt and NCBI databases. Bold printed proteins are of major interest and discussed in detail in the text. Detailed information on mass spectrometry parameters for each spot analyzed is given in the supplementary material (Table 3).
[b] Functional protein classification according to the KEGG PATHWAY Database. Additional categories are included when necessary.
[c] Protein spot abundance: statistically significant values of at least 3 fold more abundant protein spots are given for the somatic embryos proteomes (relative spot volume SE/relative spot volume ZE). These protein spots are given in green on Fig. 2B.
[d] Spot ID as given on Fig. 2B.

proteins of this functional category are involved in energy metabolism (glycolysis, TCA cycle, glyoxylate cycle). Indeed, embryogenesis is a complex developmental process which is extensively based on carbohydrate metabolism (including sugar conversion), as previously reported in *Picea marina*, *P. glauca* [50], *C. persicum* [26,28,29] and in *P. dactylifera* [30]. The explanation for extensive carbohydrate metabolism is the heavy energy demand required for metabolic processes that occur during cell division and elongation.

In this study, although carbohydrate and energy metabolism was expressed in both types of embryos investigated, a clear change in specific protein abundance was observed. In the course of our investigation, glycolytic proteins were found of higher abundance in ZE while SE proteomes included higher abundance of proteins involved in other metabolic pathways of carbohydrate metabolism such as lactoylglutathione lyase (spot 46, 47), malate synthase (spot 6), a putative L-galactose-1-phosphate phosphatase (spot 47) and malate dehydrogenase (spot 32). The latter protein, which belongs to the tricarboxylic acid (TCA) cycle, is one of the most abundant proteins found in SE proteomes. It was found to be of 46.02 fold higher abundance than in ZE proteomes. Furthermore, oxidative phosphorylation, represented by NADH–ubiquinone oxidoreductase (spots 11), ATP synthase (spot 40) and flavoprotein wrbA (spot 45) were 3.84, 3.79 and 4 fold more abundant in SE, indicating that these embryos undergo extensive oxidation processes. The higher level of TCA cycle and oxidative phosphorylation enzymes in SE suggests a more active aerobic/respiration pathway. This intensive energy metabolism might be the reason for the absence of dormancy (common feature of SE) as previously suggested by Sghaier-Hammami et al. [30]. Moreover, increased accumulation of pyruvate decarboxylase in ZE (PDC; EC 4.1.1.1; spot 52), which catalyzes the decarboxylation of pyruvate to acetaldehyde in the frame of an ethanol fermentation pathway [51], reinforces ideas of anaerobic fermentation as the major energy pathway in ZE. This might be the consequence of limited oxygen availability for the ZE, which is surrounded by seed coat and endosperm. Furthermore, the high abundance of malate synthase (spot 32) in SE, which represents a marker of beta-oxydation and which normally takes place in seed germination [52,53], indicates

the breakdown of lipid compounds. This finding reinforces the idea of a high energy demand in SE. Furthermore, it reveals involvement of lipid compounds as energy supplier in cacao embryos. Interestingly, these observations demonstrate that, compared to their zygotic counterpart, SE of *T. cacao* may undergo precocious mobilization of storage compounds which might be the reason for the less vigor of derived plantlets compared to those originating from ZE. An earlier study carried out on globulin mobilization in SE of *Elaeis guineensis* resulted in a similar conclusion [54].

Phosphoenolpyruvate carboxylase (PEPC; EC 4.1.1.31) is found in increased abundance in ZE. This glycolytic protein is a CO_2-fixing enzyme that catalyzes the irreversible β-carboxylation of phosphoenolpyruvate (PEP) in the presence of HCO_3^- and Mg^{2+} to yield oxaloacetate (OAA) and inorganic phosphorus (Pi) [55]. In earlier studies, this enzyme was only related to photosynthesis but, over the last decade, a growing number of studies revealed that, under non-photosynthesis conditions, especially in seeds, PEPC re-fixes HCO_3^- liberated by respiration and, together with PEP, yield oxaloacetate that can be converted to aspartate, malate or other intermediates of the TCA cycle [56–58]. Thus, this metabolic pathway maintains the pool of carbon residues necessary for the biosynthesis of oil and storage proteins that take place in later stage of embryo development [59–61]. Intriguingly, the consistent accumulation of PEPC in cacao embryos raises the question of the role of this enzyme in cacao seed maturation and germination which, according to its high lipid content, might display an extensive PEPC activity as demonstrated with castor oilseed [62,63].

Sucrose synthase (SuSy, EC 2.4.1.13; spot 51 in Fig. 2B) is the only enzyme of sugar metabolism found in the frame of our study. This enzyme was 4.02 fold higher in abundance in ZE proteomes. SuSy is a cytoplasmic enzyme that catalyzes the reversible break down of sucrose. This enzyme has been studied by several research groups and has been associated with a broad range of physiological functions mostly related to carbohydrate metabolism [64,65]. Iraqi and Tremblay [50] suggested that this enzyme catalyzes the synthesis of sucrose from glucose and fructose in order to maintain the optimal sucrose level in plant tissues during embryo maturation. Other studies reported its involvement in starch and cell wall synthesis [66]. The low abundance of SuSy in SE suggests a possible disturbance in carbohydrate metabolism that might result in irregularities in storage compound and cell wall metabolism.

4.2. *High abundance of stress related proteins in somatic embryos of T. cacao*

In the frame of our study, stress induced proteins were found in higher abundance in SE proteomes where they represented the dominant group of proteins (more than 23% of the total proteins classified as more abundant in SE [Table 2]). The most representative proteins among this group were peroxidases (spots 29, 31, 32 and 35), pathogenesis related protein (spot 42) and glutathione S-transferase (GSTs) (spot 9). This observation is in accordance with the one of Rode et al. 2011 [28] who noticed that SE of *C. persicum* were more stressed than their ZE counterpart. Indeed, somatic cells of plants require a number of stress factors to switch their developmental program to a specific physiological state that allows the reprogramming of gene expression towards acquisition of embryonic competence [67–70]. In *T. cacao*, a set of stress factors encompassing wounding, salt component ($MgSO_4$) and several growth regulators (TDZ, 2,4-D and Kinetin) are required to turn somatic cells of floral explant into competent embryonic callus cultures [7,8,11]. Among plant growth regulators of auxin family, 2,4-D is especially recognized as oxidative stress inducer [67,68,71]. For instance, GSTs which appear in increased abundance in SE proteomes in this study have been earlier described as auxin-induced proteins involved in the acquisition of cellular totipotency during somatic embryogenesis [68,71]. Further explanations for this extensive activity of oxidative stress enzymes come from oxidative phosphorylation described in the previous section. Indeed, oxidative phosphorylation is the source of reactive hydrogen, a poisonous compound for plant tissues. Peroxidases and GSTs, which are endowed with xenobiotic functions [72,73], are of great importance for eliminating H_2O_2 resulting from oxidative phosphorylation.

4.3. *Aspartic proteinase, dominant enzyme in protein mobilization in T. cacao embryo*

Aspartic proteinase precursor (spot 17) was found 3.04 fold more abundant in ZE proteomes. This enzyme represents the only protein involved in a catabolic pathway, which was of increased abundance in this study. Aspartic proteinase (EC 3.4.23) is the dominant endopeptidase component involved in storage protein processing found in cacao seeds [74]. In addition, this enzyme was found to be exceptionally more active in cacao seeds than in seeds of other plant species [75]. Regarding the crucial role of this enzyme, which is responsible for the initial breakdown of proteins during germination [15,76], its relatively early accumulation in ZE in the torpedo stage, might be an indication of the onset of maturation which did not take place in SE.

4.4. *Exceptional high accumulation of endopeptidase inhibitor in somatic embryos*

In contrast to the active protein biosynthesis and the significant accumulation of aspartic proteinase precursor noticed in ZE, SE highly accumulate peptidase inhibitors. Indeed, two types of endopeptidase inhibitors were found in high abundance in SE proteomes. From a total of four protein species identified as members of this family, two belonged to the group of trypsin inhibitors while the two others were identified as 21 kDa seed proteins. The latter was also described as homolog of soybean trypsin inhibitor [77]. Thus, all these endopeptidase inhibitors can be considered as trypsin inhibitors. Trypsin inhibitor has been earlier reported in *Daucus carota* to be associated with somatic embryogenesis [78]. Unfortunately, no biological function of this enzyme related to somatic embryogenesis has been established yet. Recently, Guilloteau et al. [76] found trypsin inhibitor to be part of an active aspartic proteinase complex in cacao seeds. They suggested this trypsin inhibitor sub-unit to protect storage proteins from precocious hydrolysis that might result from the high aspartic proteinase content in cacao seeds.

Nevertheless, the role of this enzyme in cacao seeds remains far from being elucidated.

4.5. *Differences in protein folding processes in zygotic and somatic embryos*

Ubiquitin (spot 50) and T complex protein (spot 53) are the two proteins of increased abundance found in our experimental system to be involved with protein folding processes. They appeared in ZE proteomes and they were respectively 4.82 and 3.71 fold more abundant than in SE proteomes. Ubiquitin has been previously described as a biochemical tag endowed with the capacity to link covalently with other proteins for degradation and refolding of proteins [79]. The low accumulation of this protein in SE proteomes suggests a possible disturbance on protein folding in these embryos. On the other hand, T complex protein is a protein binding component whose physiological role is not yet clearly elucidated. However, it has been described as a binding factor that transfers DNA fragments from *Agrobacterium tumefaciens* into plant nucleus [80]. Recently, Hurov et al. [81] reported the intervention of proteins belonging to T complex family in protecting cells from spontaneous DNA damage. The high abundance of T complex protein in ZE proteomes suggests the involvement of this protein in a specific pathway important in embryo development in *T. cacao*. Furthermore, due to its presumable involvement in protecting DNA from damage, this protein, which might not be sufficiently accumulated in SE, might cause susceptibility of DNA in this type of embryo and therefore explain the cause of somaclonal variations that frequently occur during somatic embryogenesis. The increased accumulation of histones (histone H2A and histone H2B; spot 22) that both were 8.04 fold higher in abundance in ZE proteomes, corroborate this view because of their participation in DNA integrity.

5. Concluding remarks

The data presented here clearly demonstrate the efficiency of MS-based proteomic investigations in describing in depth physiological processes. The comparison between SE and ZE proteomes come-up with new insights into embryogenesis of *T. cacao*. The absence of dormancy and the occurrence of somaclonal variations, current features of somatic embryogenesis, have been connected to changes in specific proteins abundances. Furthermore, the consequences of the exposure of plant material to stress factors during this specific developmental process have been described. Trypsin inhibitor, whose function in seed biology is not yet clearly understood, appeared in this study to be intrinsically associated with somatic embryogenesis in *T. cacao*. ZE proteomes revealed the importance of aspartic proteinase, phosphoenolpyruvate carboxylase and sucrose synthase in *T. cacao* embryogenesis. The most pronounced difference among the two types of embryos concerns carbohydrate metabolism: ZE proteomes display a high glycolytic activity while SE are characterized by intensive aerobic/respiration pathway activity as documented by the exceptional increased of TCA cycle proteins as well as proteins of oxidative phosphorylation. Thus, regarding these specific physiological features, somatic embryogenesis, which is an induced developmental process, is driven by an adapted metabolism. These new findings shed light to the procedure of inducing SE.

Both types of embryos were harvested in similar morphological shapes; however their developmental stages might have been slightly different. Thus, it cannot be excluded that some of the observed differences in protein abundances were related to this developmental variation. Further proteomic approaches should therefore also include series of zygotic and somatic embryo stages.

Finally, a large number of uncharacterized proteins have also been found in this study (Table 3; supplementary material). The reason for this fact is the poor availability of cacao proteins in databases. Although the cacao genome has recently been sequenced, there is still a great need for protein characterization in this plant species.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.jprot.2012.11.007.

Acknowledgments

This research has been carried out at the Institute of Plant Genetics (Department of Plant Proteomics) of Hannover University, Germany. Work of Alexandre Mboene Noah was financed by a short-term Wood–Whelan Fellowship from the International Union of Biochemistry and Molecular Biology (IUBMB) and by the Institute of Plant Genetics. Support of Alexandre Mboene Noah during his stay in Germany by Kerstin Hartung and Dr. Wolfgang Geller is gratefully acknowledged.

REFERENCES

[1] Alverson WS, Whitlock BA, Nyffeler R, Bayer C, Baum DA. Phylogeny of the core Malvales: evidence from ndhF sequence data. Am J Bot 1999;86:1474–86.
[2] Kris-Etherton PM, Keen CL. Evidence that the antioxidant flavonoids in tea and cocoa are beneficial for cardiovascular health. Curr Opin Lipidol 2002;13:41–9.
[3] Lee KW, Kim YJ, Lee HJ, Lee CY. Cocoa has more phenolic phytochemicals and a higher antioxidant capacity than teas and red wine. J Agric Food Chem 2003;51:7292–5.
[4] Rimbach G, Melchin M, Moehring J, Wagner AE. Polyphenols from cocoa and vascular health — a critical review. Int J Mol Sci 2009;10:4290–309.
[5] Spencer JP. Flavonoids and brain health: multiple effects underpinned by common mechanisms. Genes Nutr 2009;4: 243–50.
[6] Pandey KK. Incompatibility system in *Theobroma cacao*. Am Nat 1960;94:379–81.
[7] Li Z, Traore A, Maximova S, Guiltinan M. Somatic embryogenesis and plant regeneration from floral explant of cocao (*Theobroma cacao* L.) using thidiazuron. In Vitro Cell Dev Biol Plant 1998;34:293–9.
[8] Traore A, Maximova S, Guiltinan M. Micropropagation of *Theobroma cacao* using somatic embryo-derived plants. In Vitro Cell Dev Biol Plant 2003;39:332–7.
[9] Santos MO, Romanoa E, Yotoko KSC, Tinocoa MLP, Diasa BBA, Aragão FJL. Characterisation of the cacao somatic

embryogenesis receptor-like kinase (SERK) gene expressed during somatic embryogenesis. Plant Sci 2004;168:723–9.
[10] Alemanno L, Devic M, Niemenak N, Sanier C, Guilleminot J, Rio M, et al. Characterization of *leafy cotyledon1-like* during embryogenesis in *Theobroma cacao* L. Planta 2008;227: 853–66.
[11] Minyaka E, Niemenak N, Fotso, Sangare A, Omokolo ND. Effect of $MgSO_4$ and K_2SO_4 on somatic embryo differentiation in *Theobroma cacao* L. Plant Cell Tissue Organ Cult 2008;94:149–60.
[12] Niemenak N, Saare-Surminski K, Rohsius C, Omokolo ND, Lieberei R. Regeneration of somatic embryos in *Theobroma cacao* L. in temporary immersion bioreactor and analysis of free amino acids in different tissues. Plant Cell Rep 2008;27: 667–76.
[13] Alemanno L, Berthouly M, Micheaux-Ferrière N. A comparison between *Theobroma cacao* L., zygotic embryogenesis and somatic embryogenesis from floral explants. In vitro Cell Dev Biol Plant 1997;33:163–72.
[14] Dodeman VL, Ducreux G, Kreis M. Zygotic embryogenesis versus somatic embryogenesis. J Exp Bot 1997;48:1493–509.
[15] Callis J. Regulation of protein degradation. Plant Cell 1995;7: 845–57.
[16] Elamrani AJ, Raymond P, Saglio P. Nature and utilization of seed reserves during germination and heterotrophic growth of young sugar beet seedlings. Seed Sci Res 1992;2: 1–8.
[17] Rajjou L, Gallardo K, Debeaujon I, Vandekerckhove J, Job C, Job D. The effect of α-amanitin on the *Arabidopsis* seed proteome highlights the distinct roles of stored and neosynthesized mRNAs during germination. Plant Physiol 2004;134:1598–613.
[18] Gygi SP, Rochon Y, Franza BR, Aebersold R. Correlation between protein and mRNA abundance in yeast. Mol Cell Biol 1999;19:1720–30.
[19] Greenbaum D, Colangelo C, Williams K, Gerstein M. Comparing protein abundance and mRNA expression levels on a genomic scale. Genome Biol 2003;4:117–24.
[20] Rose JKC, Bashir S, Giovannoni JJ, Jahn MM, Saravanan RS. Tackling the plant proteome: practical approaches, hurdles and experimental tools. Plant J 2004;39:715–33.
[21] Manna M, Kelleher NL. Precision proteomics: the case for high resolution and high mass accuracy. Proc Natl Acad Sci U S A 2008;105:18132–8.
[22] Thiellement H, Bahrman N, Damerval C, Plomion C, Rossignol M, Santoni V, et al. Proteomics for genetic and physiological studies in plants. Electrophoresis 1999;20:2013–26.
[23] Kersten B, Burkle L, Kuhn EJ, Giavalisco P, Konthur Z, Lueking A, et al. Large-scale plant proteomics. Plant Mol Biol 2002;48: 133–41.
[24] Roberts JK. Proteomics and a future generation of plant molecular biologists. Plant Mol Biol 2002;48:143–54.
[25] Lippert D, Zhuang J, Ralph S, Ellis DE, Gilbert M, Olafson R, et al. Proteome analysis of early somatic embryogenesis in *Picea glauca*. Proteomics 2005;5:461–73.
[26] Winkelmann T, Heintz D, Van Dorsselaer A, Serek M, Braun HP. Proteomic analyses of somatic and zygotic embryos of *Cyclamen persicum* Mill. reveal new insights into seed and germination physiology. Planta 2006;224:508–19.
[27] Bian F, Zheng C, Qu F, Gong X, You C. Proteomic analysis of somatic embryogenesis in *Cyclamen persicum* Mill. Plant Mol Biol Rep 2010;28:22–31.
[28] Rode C, Gallien S, Heintz D, Van-Dorsselaer A, Braun HP, Winkelmann T. Enolases: storage compounds in seeds? Evidence from a proteomic comparison of zygotic and somatic embryos of *Cyclamen persicum* Mill. Plant Mol Biol 2011;75:305–19.
[29] Rode C, Lindhorst K, Braun HP, Winkelmann T. From callus to embryo — a proteomic view on the development and maturation of somatic embryos in *Cyclamen persicum*. Planta 2012;35:995–1101.
[30] Sghaier-Hammami B, Drira N, Jorrín-Novo VJ. Comparative 2-DE proteomic analysis of date palm (*Phoenix dactylifera* L.) somatic and zygotic embryos. J Proteomics 2009;1:161–77.
[31] Awang A, Karim R, Mitsui T. Proteomic analysis of *Theobroma cacao* pod husk. J Appl Glycosci 2010;57:245–67.
[32] Voigt J, Biehl B, Kamaruddin S. The major seed proteins of *Theobroma cacao* L. Food Chem 1993;47:145–51.
[33] Abecia-Soria L, Pezoa-Garcia NH, Amaya-Farfan J. Soluble albumin and biological value of protein in cocoa (*Theobroma cacao* L.) beans as a function of roasting time. J Food Sci 2005;70:S294–8.
[34] Bertazzo A, Agnolin F, Comai S, Zancato M, Costa CVL, Seraglia R, et al. The protein profile of *Theobroma cacao* L. seeds as obtained by matrix-assisted laser desorption/ionization mass spectrometry. Rapid Commun Mass Spectrom 2011;25:2035–42.
[35] ICGD (International Cocoa Germplasm Database). NYSE Liffe/CRA Ltd./University of Reading, UK. Available:http://www.icgd.reading.ac.uk. (Accessed January 2nd, 2012).
[36] Driver JA, Kuniyuki AH. In vitro propagation of Paradox walnut root stock. HortScience 1984;19:507–9.
[37] Colditz F, Nyamsuren O, Niehaus H, Eubel H, Braun HP, Krajinski F. Proteomic approach: Identification of *Medicago truncatula* proteins induced in roots after infection with the pathogenic oomycete Aphanomyces euteiches. Plant Mol Biol 2004;55:109–20.
[38] Mihr C, Braun HP. Proteomics in plant biology. In: Conn PM, editor. Handbook of proteomic methods. Totowa: Humana Press; 2003. p. 409–16.
[39] Schägger H, von Jagow G. Tricine–sodium dodecyl sulfate polyacrylamide gel electrophoresis for the separation of proteins in the range from 1–100 kDalton. Anal Biochem 1987;166:368–79.
[40] Schägger H. Tricine–SDS–PAGE. Nat Protoc 2006;1:16–22.
[41] Neuhoff V, Stamm R, Eibl H. Clear background and highly sensitive protein staining with Coomassie Blue dyes in polyacrylamide gels: a systematic analysis. Electrophoresis 1985;6:427–48.
[42] Neuhoff V, Stamm R, Pardowitz I, Arold N, Ehrhardt W, Taube D. Essential problems in quantification of proteins following colloidal staining with Coomassie Brilliant Blue dyes in polyacrylamide gels, and their solution. Electrophoresis 1990;11:101–17.
[43] Berth M, Moser FM, Kolbe M, Bernhardt J. The state of the art in the analysis of two-dimensional gel electrophoresis images. Appl Microbiol Biotechnol 2007;76:1223–43.
[44] Klodmann J, Sunderhaus S, Nimtz M, Jansch L, Braun HP. Internal architecture of mitochondrial complex I from *Arabidopsis thaliana*. Plant Cell 2010;22:797–810.
[45] Argout X, Salse J, Aury J-M, Guiltinan MJ, Droc G, Gouzy J, et al. The genome of *Theobroma cacao*. Nat Genet 2011;43:101–8.
[46] Esan E. Tissue culture studies on cacao (*Theobroma cacao* L.). Int Cacao Research Conference. Ibadan(Nigeria); 1975. p. 119–25.
[47] Maximova SN, Alemanno L, Young A, Ferriere N, Traore A, Guiltinan M. Efficiency, genotypic variability, and cellular origin of primary and secondary somatic embryogenesis of *Theobroma cacao* L. In Vitro Cell Dev Biol Plant 2002;38:252–9.
[48] Vieitez A. Somatic embryogenesis in *Camellia* spp. Somatic embryogenesis in woody plants, vol. 2. , In: Jain SM, Gupta PK, Newton R, editors. Angiosperm; 1995. p. 63–70.
[49] Sghaier-Hammami B, Valledor L, Drira N, Jorrin-Novo VJ. Proteomics, analysis of the development and germination of date palm (*Phoenix dactylifera* L.) zygotic embryos. Proteomics 2009;9:2543–54.
[50] Iraqi D, Tremblay FM. Analysis of carbohydrate metabolism enzymes and cellular contents of sugar and proteins during spruce somatic embryogenesis suggests a regulatory role of exogenous sucrose in embryo development. J Exp Bot 2001;52: 2301–11.

[51] Rudy KPI, Mary De Pauw D, Dennis ES, Good AG. Enhanced low oxygen survival in *Arabidopsis* through increased metabolic flux in the fermentation pathway. Plant Physiol 2003;132:1292–302.
[52] Baker A, Graham IA, Holdsworth M, Smith SM, Theodoulou FL. Chewing the fat: beta-oxidation in signalling and development. Trends Plant Sci 2006;11:124–32.
[53] Pracharoenwattana I, Zhou W, Smith SM. Fatty acid beta-oxidation in germinating *Arabidopsis* seeds is supported by peroxisomal hydroxypyruvate reductase when malate dehydrogenase is absent. Plant Mol Biol 2010;72:101–9.
[54] Aberlenc-bertossi F, Chabrillange N, Duval Y, Tregear J. Contrasting globulin and cysteine proteinase gene expression patterns reveal fundamental developmental differences between zygotic and somatic embryos of oil palm. Tree Physiol 2008;28:1157–67.
[55] Chollet R, Vidal J, O'Leary MH. Phosphoenolpyruvate carboxylase: a ubiquitous, highly regulated enzyme in plants. Annu Rev Plant Physiol Plant Mol Biol 1996;47:273–98.
[56] Golombek S, Heim U, Horstmann C, Wobus U, Weber H. Phosphoenolpyruvate carboxylase in developing seeds on *Vicia faba*. Gene expression and metabolic regulation. Planta 1999;208:66–72.
[57] O'Leary B, Park J, Plaxton WC. The remarkable diversity of plant phosphoenolpyruvate carboxylase (PEPC): recent insights into the physiological functions and post-translational controls of nonphotosynthetic PEPCs. Biochem J 2011;436:15–34.
[58] Leblová S, Strakošová A, Vojtcehová M. Regulation of the activity of phosphoenolpyruvate carboxylase isolated from germinating maize (*Zea mays* L.) Seeds by Some Metabolites. Biol Plant 1991;33:66–74.
[59] Smith AJ, Rinne RW, Seif RD. PEP carboxylase and pyruvate kinase involvement in protein and oil biosynthesis during soybean seed development. Crop Sci 1989;29:349–53.
[60] Gonzalez MC, Osuna L, Echevarria C, Vidal J, Cejudo FJ. Expression and localization of PEP carboxylase in developing and germinating wheat grains. Plant Physiol 1998;116:1249–58.
[61] Rolletschek H, Borisjuk L, Radchuk R, Miranda M, Heim U, Wobus U, et al. Seed-specific expression of a bacterial phosphoenolpyruvate carboxylase in *Vicia narbonensis* increases protein content and improves carbon economy. Plant Biotechnol J 2004;2:211–9.
[62] Sangwan RS, Singh N, Plaxton WC. Phosphoenolpyruvate carboxylase activity and concentration in the endosperm of developing and germinating castor oil seeds. Plant Physiol 1992;99:445–9.
[63] Blonde JD, Plaxton WC. Structural and kinetic properties of high and low molecular mass phosphoenolpyruvate carboxylase isoforms from the endosperm of developing castor oilseeds. J Biol Chem 2003;278:11867–73.
[64] Sturm A, Tang GQ. The sucrose-cleaving enzymes of plants are crucial for development, growth and carbon partitioning. Trends Plant Sci 1999;4:401–7.
[65] Winter H, Huber SC. Regulation of sucrose metabolism in higher plants: localization and regulation of activity of key enzymes. Crit Rev Plant Sci 2000;19:31–67.
[66] King SP, Lunn JE, Furbank RT. Carbohydrate content and enzyme metabolism in developing canola (*Brassica napus* L.) siliques. Plant Physiol 1997;114:153–60.
[67] Pasternak TP, Prinsen E, Ayaydin F, Miskolczi P, Potters G, Asard H, et al. The role of auxin, pH, and stress in the activation of embryogenic cell division in leaf protoplast-derived cells of *Alfalfa*. Plant Physiol 2002;129:1807–19.
[68] Fehér A, Pasternak TP, Dudits D. Transition of somatic plant cells to an embryogenic state. Plant Cell Tissue Organ Cult 2003;74:201–28.
[69] Verdeil JL, Alemanno L, Niemenak N, Tranbager TJ. Pluripotent versus totipotent plant cells: dependence versus autonomy. Trends Plant Sci 2007;12:245–52.
[70] Thibaud-Nissen F, Shealy RT, Khanna A, Vodkin LO. Clustring of microarray data reveals transcript patterns associated with somatic embryogenesis in soybean. Plant Physiol 2003;132:118–36.
[71] Marrs KA. The functions and regulation of glutathione S-transferases in plants. Annu Rev Plant Physiol Plant Mol Biol 1996;47:127–58.
[72] Lamoureux GL, Rusness DG. Glutathione in the metabolism and detoxification of xenobiotics in plants. In: De Kok LJ, et al, editor. Sulfur Nutrition and Assimilation in Higher Plants. SPB: Academic Publishing; 1993. p. 221–37.
[73] Edwards R, Dixon DP, Walbot V. Plant glutathione S-transferases: enzymes with multiple functions in sickness and in health. Trends Plant Sci 2000;5:193–8.
[74] Voigt J, Kamaruddin S, Heinrichs H, Wrann D, Senyuk V, Biehl B. Developmental stage-dependent variation of the levels of globular storage protein and aspartic endoprotease during ripening and germination of *Theobroma cacao* L. seeds. J Plant Physiol 1995;145:299–307.
[75] Voigt G, Biehl B, Heinrichs H, Voigt J. Aspartic proteinase levels in seeds of different angiosperms. Phytochemistry 1997;44:389–92.
[76] Guilloteau M, Laloi M, Michaux S, Bucheli P, McCarthy J. Identification and characterisation of the major aspartic proteinase activity in *Theobroma cacao* seeds. J Sci Food Agric 2005;85:549–62.
[77] Tai H, McHenry L, Fritz PJ, Furtek DB. Nucleic acid sequence of a 21 kDa cocoa seed protein with homology to the soybean trypsin inhibitor (Kunitz) family of protease inhibitors. Plant Mol Biol 1991;16:913–5.
[78] Carlberg I, Jonsson L, Bergenstråhle A, Söderhäll K. Purification of a trypsin inhibitor secreted by embryogenic carrot cells. Plant Physiol 1987;84:197–290.
[79] Ingvardsen C, Veierskov B. Ubiquitin and proteasome dependent proteolysis in plants. Physiol Plant 2001;112:451–9.
[80] Loyter A, Rosenbluh J, Zakai N, Li J, Kozlovsky SV, Tzfira T, et al. The plant VirE2 interacting protein 1. A molecular link between the agrobacterium T-Complex and the host cell chromatin? Plant Physiol 2005;138:1318–21.
[81] Hurov KE, Ramusino CC, Elledge SJ. A genetic screen identifies the Triple T complex required for DNA damage signaling and ATM and ATR stability. Genes Dev 2010;24:1939–50.

Manuscript 1

2.5 Dealing with the sulfur part of cysteine: four enzymatic steps degrade L-cysteine to pyruvate and thiosulfate in *Arabidopsis* mitochondria preventing toxic effects of persulfides

Saskia Höfler[1], Christin Lorenz[1], Tjorven Busch[1], Mascha Brinkkötter[1], Takayuki Tohge[2], Alisdair R. Fernie[2], Hans-Peter Braun[1], Tatjana M. Hildebrandt[1]

[1]Institut für Pflanzengenetik, Leibniz Universität Hannover, Herrenhäuser Str. 2, 30419 Hannover, Germany
[2]Max-Planck-Institute of Molecular Plant Physiology, Am Mühlenberg 1, 14476 Potsdam-Golm, Germany

Type of authorship:	Co-author
Type of article:	Research article
Share of the work:	25 %
Contribution to the publication:	Performed experiments, analysed data, prepared figures, wrote parts of the manuscript
Status of publication:	In preparation

*****Note that the supplementary material for chapter 2.5 is not printed but included in the compact disc attached to this work.

Dealing with the sulfur part of cysteine: four enzymatic steps degrade L-cysteine to pyruvate and thiosulfate in *Arabidopsis* mitochondria preventing toxic effects of persulfides

Saskia Höfler[1], Christin Lorenz[1], Tjorven Busch[1], Mascha Brinkötter[1], Takayuki Tohge[2], Alisdair R. Fernie[2], Hans-Peter Braun[1], and Tatjana M. Hildebrandt [1]

[1]Institut für Pflanzengenetik, Leibniz Universität Hannover, Herrenhäuser Str. 2, 30419 Hannover, Germany
[2]Max-Planck-Institute of Molecular Plant Physiology, Am Mühlenberg 1, 14476 Potsdam-Golm, Germany

Corresponding author: Tatjana Hildebrandt: hildebrandt@genetik.uni-hannover.de

Abstract

Amino acid catabolism is essential for adjusting pool sizes of free amino acids and takes part in energy production as well as nutrient remobilization. The carbon skeletons are generally converted to precursors or intermediates of the tricarboxylic acid cycle. In the case of cysteine, the reduced sulfur derived from the thiol group also has to be oxidized in order to prevent accumulation of toxic concentrations. Here we present a mitochondrial sulfur catabolic pathway catalyzing complete oxidation of L-cysteine to pyruvate and thiosulfate. After transamination to 3-mercaptopyruvate the sulfhydryl group from L-cysteine is transferred to glutathione by sulfurtransferase 1 and oxidized to sulfite by the sulfur dioxygenase ETHE1. Sulfite is then converted to thiosulfate by addition of a second persulfide group by sulfurtransferase 1. This pathway is most relevant during early embryo development and for vegetative growth under light limiting conditions. Characterization of a double mutant produced from *Arabidopsis thaliana* T-DNA insertion lines for ETHE1 and sulfurtransferase 1 revealed that a persulfide intermediate interferes with amino acid catabolism and induces early senescence.

Introduction

Amino acid catabolism in plants is involved in the regulation of steady state levels of free amino acids and is particularly important in conditions of increased protein turnover such as germination and senescence. It is critical for nutrient redistribution from senescing leaves to newly formed sink organs such as young leaves and developing seeds. The degradation of amino acids also contributes to energy production in situations of carbohydrate starvation, which may occur during drought or unfavorable light conditions. Pool sizes of free amino

acids are highly diverse and dynamically change in response to environmental or developmental factors (Hildebrandt et al. 2015). However, the intracellular concentration of cysteine is amongst the lowest known for protein amino acids. The thiol group of cysteine is highly reactive, it can deplete cells of pyridoxal phosphate by forming thiazolidine derivatives and in addition autooxidation in the presence of transition metals generates reactive oxygen species (Osman et al. 1997). Therefore, intercellular concentrations have to be tightly controlled in order to avoid toxic effects. However, cysteine levels also must be sufficiently high to support protein synthesis and the production of other essential molecules such as glutathione, coenzyme A and reduced sulfur for the biosynthesis of iron sulfur clusters, biotin, thiamin, and molybdenum cofactor (Balk and Pilon 2011, Van Hoewyk et al. 2008). Thus, synthesis as well as the degradation of cysteine has to be tightly regulated.

During degradation, the carbon skeleton of cysteine is converted to pyruvate via removal of the amino and sulfhydryl group. These reactions can be catalyzed by different pathways. Cysteine desulfhydrases deaminate cysteine to pyruvate, ammonia, and hydrogen sulfide (H_2S). Specific isoforms using L-Cys (DES1, EC 4.4.1.1) and D-Cys (D-CDES, EC 4.4.1.15) as a substrate are present in the cytosol and the mitochondrial matrix, respectively (Alvarez et al. 2010, Riemenschneider et al. 2005). Both enzymes are induced during senescence and therefore have been implicated in cysteine catabolism for nutrient remobilization. DES1 deficient *Arabidopsis* mutants accumulate cysteine and in addition show alterations in autophagosome formation and stomatal opening, which are reversible by supplementation with H_2S (Alvarez et al. 2010, 2012, Jin et al. 2013). Taken together, these results demonstrate that the desulfhydration reaction is relevant for cysteine homeostasis and also regulates production of the signaling molecule H_2S. However, the further fate of H_2S, which is highly toxic already at low micromolar concentrations, has not been analyzed yet. Cysteine desulfurases (NifS-like proteins, EC 2.8.1.7) provide sulfur for the synthesis of iron-sulfur clusters, biotin, and thiamin (Couturier et al. 2013).

We recently demonstrated a role of the sulfur dioxygenase ETHE1 (EC 1.13.11.18) in plant cysteine catabolism (Krüßel et al. 2014). ETHE1 is localized in the mitochondrial matrix and oxidizes glutathione persulfide (GSSH) to sulfite. Knockout of the ETHE1 gene (AT1G53580) in *Arabidopsis* leads to an arrest of embryo development at early heart stage (Holdorf et al. 2012). Knockdown mutants are viable but show a delay in embryo development (Krüßel et al. 2014). Interestingly, ETHE1 is highly induced by carbohydrate

starvation and mutants develop premature leaf senescence under light limiting growth conditions indicating a role in the use of amino acids as alternative energy source.
Here we present the additional steps of the ETHE1 dependent mitochondrial sulfur catabolic pathway catalyzing complete oxidation of L-cysteine to pyruvate and thiosulfate. After transamination to 3-mercaptopyruvate the sulfhydryl group from L-cysteine is transferred to glutathione by sulfurtransferase 1. Interestingly, this enzyme also converts sulfite to the final product thiosulfate by addition of a second persulfide group.

Sulfurtransferases catalyze the transfer of a sulfur atom from a suitable sulfur donor to nucleophilic sulfur acceptors. 20 putative sulfurtransferase isoforms are annotated in the *Arabidopsis* genome and some of them have already been characterized on a protein basis (Bartels et al. 2007). They might play a role in plant development and stress response, however, the exact function has not been identified yet. *In vitro*, the activity is measured using either thiosulfate or 3-mercaptopyruvate as a sulfur donor and cyanide as an acceptor. Mao et al. (2011) recently demonstrated that the mitochondrial sulfurtransferase Str1 (AT1G79230) contributes the main mercaptopyruvate sulfurtransferase activity (about 80 %) in *Arabidopsis*. Knockout plants had no visible phenotype under long-day growth conditions, but seed development was severely compromised. The majority of seeds were shrunken and not able to germinate. Embryo development arrested at the heart stage producing abnormal morphological shapes and eventually aborted.

In order to identify the role of mitochondrial cysteine catabolism and the individual reaction steps in seed development as well as plant energy metabolism, we produced and characterized a double mutant from the ETHE1 knockdown line *ethe1-1* (Krüßel et al. 2014) and the Str1 knockout line *str1-1* (Mao et al. 2011).

Materials and methods

Plant material and growth conditions

All *Arabidopsis thaliana* plants used for this study were of the Columbia ecotype (Col-O). Plants were grown in climate chambers under long-day conditions (16h light/8h dark) or short-day conditions (8h light/16h dark) at 22 °C, 85 µmol s^{-1} m^{-2} light and 65 % humidity. The T-DNA insertion line SALK_021573 (*ethe1-1*) for the gene AT1G53580 has been

characterized in our lab before (Krüßel et al. 2014). Seeds of the line SALK_015593 (*str1-1*) for the gene AT1G79230 were obtained from the Nottingham *Arabidopsis* Stock Centre (University of Nottingham, UK). This line has been described by Mao et al. (2011). The *ethe1-1* x *str1-1* double mutant was produced by crossing of the two T-DNA insertion lines described above.Homozygous mutant lines were identified by genomic PCR using gene-specific primers (5'-TGGAATTGGGTTATATGGTGG-3' and 5'-CGGATCAATCAACTGCTCATC-3' for *ethe1-1* and 5'-AAAGGGGATCTTTAGTGCAGC-3' and 5'- GTGGGAAGGAAGCAAATTCTC-3' for *str1-1*) and the T-DNA left border primer LBb1.3 (5'-ATTTTGCCGATTTCGGAAC-3'). Complete rosettes of wild type and mutant plants were harvested at dawn and used for metabolite analysis.

Expression and purification of ETHE1 and Str1

Expression and purification of recombinant ETHE1 protein was performed like described before (Krüßel et al. 2014). The plasmid containing *Arabidopsis* Str1 was generously provided by Jutta Papenbrock (Institute for Botany, Leibniz University Hannover) and expressed in the same way as ETHE1.

Cell suspension cultures and isolation of mitochondria

Arabidopsis cell suspension cultures were established and maintained as described by May and Leaver (1993) and Sunderhaus et al. (2006). Mitochondria were prepared following to the procedure outlined by Werhahn et al. (2001).

Sulfur dioxygenase activity test

SDO activity was measured at 25°C in a Clarke-type oxygen electrode (Oroboros Oxygraph and Hansatech DW1 Oxygrapy) following the procedure described in Hildebrandt and Grieshaber (2008). The reaction contained 1 to 2 µg/ml purified enzyme or 150 to 300 µg/ml mitochondrial protein in 0.1 M potassium phosphate buffer pH 7.4. For the standard activity test, 1 mM GSH (final concentration) was added, followed by 15 µl/ml of a saturated elemental sulfur solution in acetone. Acetone did not inhibit enzyme activity. Rates were

measured during the linear phase of oxygen depletion, which occurred in the first 2 to 3 minutes.

Phenotype analysis

For general phenotype analysis a modified version of the procedure described in Boyes et al. (2001) was used. Plants were grown under long-day or short-day conditions and growth parameters were measured once per week.
Embryo morphology was analyzed with a microscope equipped with Normarski optics. Seeds were dissected from the siliques and cleared in Hoyer's solution (15 ml distilled water, 3.75 g gum arabic, 2.5 ml glycerine, 50 g chloral hydrate) overnight before analysis.
Germination rates were determined for seeds of wild type, *ethe1-1* and three different morphological types of *str1-1* and *ethe1-1xstr1-1*. The seeds were surface sterilized with 6% sodium hypochloride and 100% ethanol followed by five washing steps with sterilized water. For vernalization seeds were incubated for 2 days at 4°C in the dark. Approximately 20 seeds were sown per plate (3 replicates per sample) on MS-medium without a carbon source and incubated for another 2 days at 4°C in the dark. Afterwards the plates were placed into a growth chamber (24°C, 16 h light/8 h dark). After 24, 48 and 72h germinated seeds were counted. A seed is considered to be germinated when the radicle raptures the endosperm and the testa.

Metabolite analysis

Products of the sulfur dioxygenase reaction were analyzed by HPLC (Hildebrandt and Grieshaber, 2008).

For metabolite profiling complete rosettes of wild type and mutant plants grown under short day conditions were harvested at the end of the dark period at the age of 42 days and 100 days. Metabolite analysis by GC-MS was performed essentially as described by Lisec et al. (2006) and by extracting of metabolites for injection of extracts from 1mg fresh weight of plant material into the GC/TOF-MS. Chromatograms and mass spectra were evaluated by using TagFinder 4.0 software (Luedemann et al. 2008) and Xcalibur 2.1 software (Thermo Fisher Scientific, Waltham, USA). Metabolites were identified in comparison to database entries of authentic standards (Kopka et al. 2005, Schauer et al. 2005). Peak areas of the mass

(m/z) fragments were normalized to the internal standard (ribitol) and fresh weight of the samples. Identification and annotation of detected peaks are shown in Supplemental Table S1 following recent recommendations for reporting metabolite data (Fernie et al. 2011).

Results

Enzymatic steps of the mitochondrial cysteine catabolic pathway

In a previous study we demonstrated that the mitochondrial sulfur dioxygenase ETHE1 oxidizes persulfide groups derived from 3-mercaptopyruvate to sulfite and postulated a role for this reaction sequence in L-cysteine catabolism (Krüßel et al. 2014). In order to identify the complete pathway and provide evidence for the individual reaction steps (Fig. 1A), we measured sulfur dioxygenase activity in mitochondria after the addition of L-cysteine as a substrate (Fig. 1B) and in addition reconstituted part of the pathway with isolated recombinant enzymes (Fig. 1 C, D). Our results indicate that L-cysteine is oxidized by the ETHE1 pathway after transamination to 3-mercaptopyruvate. Mitochondrial oxygen consumption with L-cysteine as a substrate was detectable only if 2-oxoglutarate was included in the reaction mixture as an amino group acceptor for transamination. The activity was very low in the *ethe1-1* knockdown mutant showing that the oxygen consumption rate in wild type mitochondria was indeed specific for the ETHE1 dependent pathway (Fig. 1B).

Next, we tested whether ETHE1 in combination with the main mitochondrial sulfurtransferase Str1 is able to catalyze the remaining part of the postulated pathway, i.e. transfer of the sulfhydryl group from 3-mercaptopyruvate to glutathione, oxidation to sulfite, and addition of a second persulfide group to produce thiosulfate (Fig. 1A). Isolated recombinant ETHE1 enzyme oxidizes glutathione persulfide (GSSH), which is non-enzymatically produced from reduced glutathione (GSH) plus elemental sulfur (S8), to sulfite but cannot use the sulfhydryl group of 3-mercaptopyruvate as a substrate (left side of Fig. 1C and D). However, in the presence of isolated recombinant sulfurtransferase 1 protein, the sulfur dioxygenase activity is comparable with elemental sulfur or 3-mercaptopyruvate added as a substrate (right side of Fig. 1C). Thus, the sulfurtransferase is able to produce GSSH as a substrate for ETHE1 by transferring the sulfhydryl group of 3-mercaptopyruvate to GSH. Thiosulfate did not serve as a substrate for this reaction sequence (data not shown). Our results also provide evidence for a second function of Str1 in the cysteine catabolic pathway. The main product of persulfide

oxidation by ETHE1 in the presence of Str1 is thiosulfate, indicating that the sulfurtransferase adds an additional persulfide group to the sulfite produced by the sulfur dioxygenase reaction (left side of Fig. 1D).

Physiological roles of mitochondrial cysteine catabolism

Having identified the individual steps of mitochondrial cysteine catabolism we next addressed the physiological role of this reaction sequence during the *Arabidopsis* life cycle. T-DNA insertion lines for both, ETHE1 and Str1 have already been characterized (Holdorf et al. 2012, Krüßel et al. 2014, Mao et al. 2011) and show different phenotypes. Therefore, in order to unravel the metabolic reason for the individual effects we produced a double mutant from the strong knockdown line *ethe1-1* (Krüßel et al. 2014) and the knockout line *str1-1* (Mao et al. 2011).

Seed and embryo development is severely impaired in the ethe1-1 x str1-1 double mutant

Since embryo development is effected in ETHE1 as well as Str1 deficient mutants, we first analyzed the seeds of the *ethe1-1 x str1-1* double mutant to see whether down-regulation of both enzymes had additive effects. In summary, embryo development of the double mutant was comparable to *str1-1* (Fig. 2). Wild type embryos developed very uniformly reaching the mature stage at seven days after pollination (DAP), whereas several growth stages were simultaneously present in *ethe1-1* siliques, and not until 10 DAP all of them were mature (Fig. 2B). In contrast, embryogenesis was severely delayed in *str1-1* and the *ethe1-1 x str1-1* double mutant. Most of the mutant embryos were still in the pre-globular or globular stage at 5 DAP when the wild type had already reached cotyledon stage (Fig. 2B). At 7 DAP wild type seeds were already dark green indicating that the embryos had developed their full photosynthetic capacity (Fig. 2A). In contrast, seeds of *str1-1* plants and the double mutant were smaller, of a white, light green or even light brown color, and contained embryos in the globular to cotyledon stage. Starting from about 10 DAP different degrees of a brown, shriveled seed phenotype became apparent (Fig. 2A, Supp Figs. S1 and S2). Embryos developed morphological abnormalities such as giant heart stage or asymmetric cotyledons (Fig. 2A, Supp Figs. S3 and S4). Less than 50 % of the *str1-1* and *ethe1-1 x str1-1* embryos finally reached the mature stage whereas the rest aborted mainly at heart stage.

While mature *ethe1-1* seeds were morphologically indistinguishable from the wild type, only about one fifth of the seeds from *str1-1* plants and the double mutant appeared normal. The rest of the seeds were to about equal parts of a rectangular shape or severely shrunken (Fig. 3). Germination was slightly delayed in *ethe1-1* with 82 ± 3 % germinated seeds after 24 h compared to 93 ± 8 % in the wild type (Fig. 3). However, already after 48 h there was no significant difference between germination rates of *ethe1-1* and wild type seeds. In contrast, the germination capacity of *str1-1* seeds was severely compromised. Even normal looking seeds were not completely germinated after 72 h, and in rectangular and shriveled seeds rates dropped to 43 ± 8 % and 20 ± 5 % respectively. Interestingly, seeds of the double mutant germinated significantly better than those of the *str1-1* mutant (marked by crosses in Fig. 3). The normal looking seeds behaved like *ethe1-1* seeds and also higher percentages of the misshaped seeds compared to the *str1-1* mutant were able to germinate (63 ± 10 % of rectangular and 35 ± 5 % of shriveled seeds).

Patterns of early leaf senescence under light limitation

When grown under long-day conditions (16h light/8h dark) the phenotype of the three analyzed mutant lines was comparable to the wild type (Supp Fig S5). Shortening of the light period to 8 hours per day led to reduced growth and early leaf senescence in *ethe1-1* plants as described before (Krüßel et al. 2014), whereas the *str1-1* plants were indistinguishable from the wild type (Fig. 4). While development proceeded very uniformly in the wild type as well as in both individual mutant lines, the phenotype of the *ethe1-1 x str1-1* double mutant was diverse. Growth rates of the nine double mutant plants analyzed differed by a factor of 5 between 0.6 and 3 mm · day^{-1} (Fig. 4A inserted graph). Plants number 3, 4, 6, 7, and 8 grew fast and were indistinguishable from the wild type until 10 weeks after sowing. Then, leafs progressively developed chlorotic patches in two of the large plants (number 4 and 6). In contrast, growth rates of 4 double mutant plants (number 1, 2, 5, and 9) were reduced in a similar way as in *ethe1-1*. However, only two of these small plants (number 2 and 5) started to become chlorotic after seven weeks of growth while the other two (number 1 and 9) remained dark green until leaves were harvested at the age of 100 days.

In order to identify the metabolic reason for these phenotype variations we analyzed metabolite profiles of wild type and mutant plants grown under short-day conditions (Supp. Tab. S2). Samples were harvested immediately after the dark period at two different

developmental stages: After 42 days neither wild type nor mutant plants showed any visible signs of senescence ("young plants" Supp. Fig. S6). However, *ethe1-1* rosettes were significantly smaller than the wild type (6.0 ± 0.8 compared to 9.3 ± 0.9 cm) and contained less leaves (18.7 ± 1.2 vs. 23.9 ± 2.5). The rosette diameter of *ethe1-1 x str1-1* double mutants varied between 1.6 and 8.5 cm, and small as well as big plants were used for the metabolite analysis. The second set of leaf samples was harvested after 100 days ("old plants"; Fig. 4B), when all plants had already started flowering. *Ethe1-1* leaves were about 50 % chlorotic and the double mutant displayed the four different phenotypes described above (see Fig. 4B), with an *ethe1-1*-like phenotype (plants 2 and 5), a wild-type-like phenotype (plants 3, 7, and 8), and two intermediate forms, i.e. small green plants (plants 1 and 9), and large plants with yellowing leaves (plants 4 and 6).

Already in the young plants harvested at 42 days after sowing a number of specific changes in the metabolite profiles of the mutants compared to the wild type became apparent (Figs. 5 and 6, Supp Tab S2). Nitrogen-rich amino acids accumulated in all mutants, and the concentrations of secondary metabolites involved in nitrogen turnover such as putrescine and spermidine were also significantly increased (Fig. 5A; upper panel). In addition, we detected significantly higher concentrations of the amino acids taking part in photorespiration, glycine and serine, in *str1-1* and the double mutant (Fig. 5C). Interestingly, serine was also elevated in *ethe1-1* plants without the concomitant increase in glycine leading to a significant decrease of the glycine/serine ratio, which is often used as an indicator of photorespiratory activity. Specific effects on the TCA cycle also became apparent. In all three mutants succinate and fumarate concentrations were significantly lower than in the wild type (1.9 to 6.5fold), and malate was additionally decreased in *str1-1* and the double mutant (Fig. 6).

The metabolite profile of young *ethe1-1* plants revealed two characteristic changes that were not present in *str1-1* or the double mutant. There was a significant increase in the amino acids associated with energy metabolism, i.e. branched-chain amino acids, lysine, and aromatic amino acids (Fig. 5B, upper panel). In addition, carbohydrate metabolism was effected with a strong accumulation of galactinol (5.3fold) and raffinose (8fold) and lower but still significant increases in sucrose and fructose concentrations while glucose was slightly decreased (Fig. 6). Metabolite profiles of the old plants harvested 100 days after sowing reflected patterns typical for leaf senescence. In particular, a general increase relative to the young plants in concentrations of free amino acids most pronounced in those involved in nitrogen

remobilization, a strong decrease in the glycine/serine ratio, and an accumulation of sugars were detected in wild type as well as mutant plants (Figs. 5 and 6).

Next, we analyzed the metabolite profiles of the individual plants using the hierarchical clustering tool available in MeV (Fig. 7). The metabolite profile of *str1-1* was very similar to the wild type in the old plants so that the clustering algorithm was not able to distinguish between these two groups. However, all five *ethe1-1* plants were clearly grouped together on the basis of their metabolite profiles and, most interestingly, 8 of the 9 *ethe1-1 x str1-1* plants clustered correctly according to their phenotype. The small and chlorotic double mutant plants (number 2 and 5), which looked like *ethe1-1* plants of the same age, also had the most similar metabolite profiles followed by the big plants with chlorotic leaves (number 6 and 7). Four of the five double mutant plants that were still entirely green at the time of harvest clustered together, and again profiles of the small rosettes (number 1 and 9) were clearly separated from those of the big rosettes (number 7 and 8). Only one of the double mutant plants with no visible phenotype (number 3) was mixed up with the wild type and *str1-1* plants. Since not only the plants but also the metabolites were clustered it was possible to derive patterns typical for the individual phenotypes (marked by the black boxes in Fig. 7). A strong accumulation of galactinol and raffinose was present only in plants with the *ethe1-1*-phenotype, while sugar levels in general were decreased compared to the wild type in *ethe1-1* as well as all double mutant plants (see also Fig. 6). Concentrations of urea as well as the amino acids and metabolites associated with nitrogen remobilization were low in all plants with small rosettes (see also Fig. 5A, lower panel). Features characteristic for plants with chlorotic leaves were an increase of serine and glycine concentrations, a decrease in ascorbate and dehydroascorbate content, and specific changes in TCA cycle intermediates with a relative increase in citrate and 2-oxoglutarate and a decrease in succinate and fumarate (see also Fig. 6). Proline and alanine accumulated in all *ethe1-1* and double mutant plants compared to the wild type and *str1-1*.

Discussion

We identified a mitochondrial cysteine catabolic pathway and revealed the physiological relevance of individual reaction steps for seed development and plant growth under limited light conditions.

Three mitochondrial enzymes oxidize L-cysteine to pyruvate and thiosulfate

Our results show that L-cysteine can be transaminated to 3-mercaptopyruvate in *Arabidopsis* mitochondria. Subsequently, the thiol group is transferred to GSH by sulfurtransferase 1, oxidized to sulfite by the sulfur dioxygenase ETHE1 and transferred to an additional persulfide to produce thiosulfate by sulfurtransferase 1. A similar reaction sequence has been described in animal mitochondria (Hildebrandt and Grieshaber 2008). However, persulfide groups are derived from the oxidation of hydrogen sulfide catalyzed by sulfide:quinone oxidoreductase, which is not present in plants.

A cysteine aminotransferase is present in Arabidopsis mitochondria

The protein catalyzing the initial transamination step of the cysteine degradation pathway has not been identified so far, and no cysteine aminotransferase is annotated in the *Arabidopsis* genome. The same activity has been found in animal mitochondria and postulated to be an additional function of an aspartate aminotransferase (Ubuka et al. 1978). 31 aminotransferases with different substrate specificities have been described in *Arabidopsis* so far, and 10 of them are probably localized in the mitochondrial matrix (Hildebrandt et al. 2015). These mitochondrial aminotransferases are annotated to be specific for ornithine, aspartate, branched-chain amino acids, alanine, and glyoxylate, and it will be the subject of future studies to find out whether one of them or an additional enzyme not identified as an aminotransferase so far catalyzes the transamination of L-cysteine.

Two physiological roles for plant mitochondrial sulfurtransferase

Sulfurtransferases are ubiquitous enzymes present in all the different compartments of plant, animal and also prokaryotic cells. However, their metabolic functions have not been completely understood yet. The detoxification of cyanide has initially been considered to be

the main role of sulfurtransferases since cyanide is a very good acceptor for persulfide groups *in vitro* and sulfur addition converts it to less toxic thiocyanate (Westley 1973). In plants, cyanide is produced in equimolar amounts during the biosynthesis of the hormone ethylene (Peiser et al. 1984). However, most of it is metabolized to asparagine via β-cyanoalanine, and the contribution of sulfurtransferases to cyanide detoxification seems to be rather low (Machingura et al. 2014, Meyer et al. 2003). Additional physiologically relevant functions of sulfurtransferases might be control of the sulfane sulfur pool, maintenance of iron–sulfur clusters, and molybdenum cofactor synthesis (Matthies et al. 2004, Ogata and Volini 1990, Westley et al. 1983). Here we demonstrate that the mitochondrial sulfurtransferase Str1 is part of a cysteine catabolic pathway in *Arabidopsis* and even catalyzes two of the four reaction steps. In the *in vitro* assay using cyanide as a sulfur acceptor Str1 has a higher activity with 3-mercaptopyruvate than with thiosulfate as a substrate and is therefore called mercaptopyruvate sulfurtransferase (Papenbrock and Schmidt 2000). In accordance with this result, we found that Str1 readily transfers the thiol group from 3-mercaptopyruvate but not from thiosulfate to GSH to produce GSSH as a substrate for the sulfur dioxygenase reaction. As already described in animals, sulfurtransferases can also produce thiosulfate by transferring a persulfide group from a suitable donor such as GSSH to sulfite (Hildebrandt and Grieshaber 2008). We detected the same activity in *Arabidopsis* Str1 representing the final step in mitochondrial L-cysteine degradation.

Cysteine degradation and additional persulfide transfer reactions are essential during embryo development

Complete knockout of either sulfurtransferase 1 or ETHE1 in *Arabidopsis* severely affects embryo development indicating that cysteine catabolism is highly important during this developmental stage. The reason for the seed defects could either be accumulation of a toxic intermediate or nutrient deficiency. Since both mutant lines are affected, the potential toxic intermediate is most likely produced prior to the first transsulfuration step, i.e. it might be 3-mercaptopyruvate, L-cysteine or the dimer L-cystine, which forms non-enzymatically from oxidized cysteine but, at least in animals, is supposed to be less toxic (Baker 2006). An alternative explanation would be the presence of two independent toxic intermediates in the individual mutant lines, since the sulfurtransferase probably has additional functions to its role in cysteine degradation such as cyanide detoxification. The seed phenotype of *ethe1-1 x str1-1* plants resembled the *str1-1* single mutant and was not worsened but in contrast even slightly

improved by the additional down-regulation of sulfur dioxygenase activity. This result would be in agreement with the postulated accumulation of toxic cyanide due to Str1 knockout and toxic persulfide due to ETHE1 knockdown. Persulfide accumulation is most likely reduced in the double mutant since the sulfurtransferase producing it is missing. In addition, residual persulfides potentially resulting from other reactions would readily react with cyanide and thus lessen the toxic effect.

As heterotrophic organs seeds depend on nutrient supply by the mother plant. Carbon is imported as sucrose, which is also the main substrate for energy production in the developing seed. Nitrogen is derived from remobilization of amino acids by protein degradation in the senescing leaves and mainly transported as glutamine and asparagine (Masclaux-Daubresse et al. 2010). It is currently not clear, whether the degradation of amino acids other than glutamine and asparagine substantially contributes to the nutrition of the developing embryo, and knowledge on the general role of amino acid catabolism in seeds is very limited. However, there is some evidence indicating that amino acids, analogous to the situation in starving leaves, can serve as alternative respiratory substrates to generate energy (Galili et al. 2014). Knockdown of ETHE1 leads to an accumulation of most amino acids in the seeds (Krüßel et al. 2014). Thus, defects in cysteine degradation probably interfere with protein catabolism in general and might create a shortage of nitrogen supply. However, this option has to be further investigated.

Defects in cysteine catabolism induce metabolic shifts characteristic for developmental senescence

Plant senescence is a highly coordinated physiological process essential for nutrient remobilization to the growing seeds (Lim et al. 2007). This developmental stage is associated with a shift from anabolic metabolism to catabolic pathways catalyzing the degradation of proteins, carbohydrates, lipids, and nucleic acids in the leaves. Recently, comprehensive metabolite profiles of the senescence process in *Arabidopsis* have been published, providing a catalog of metabolite patterns characteristic for the onset of developmental senescence that can be used as markers (Watanabe et al. 2013). Interestingly, specific shifts in the metabolite profile already occur in presenescent tissues before chlorosis becomes apparent.

We detected several senescence specific metabolite changes in young rosettes of all mutant lines analyzed. The concentrations of amino acids as well as secondary metabolites associated with nitrogen turnover were elevated in the mutants compared to the wild type. High asparagine/aspartate and glutamine/glutamate ratios are typical metabolite markers for the onset of senescence in *Arabidopsis*, and nitrogen-rich amino acids are closely associated with nitrogen remobilization and transport (Watanabe et al. 2013). Polyamines are synthesized from arginine or ornithine and thus interconnected with the major N pools. Due to their high intracellular concentrations they are considered to be major sinks and storage forms of assimilated N and also might play a role in the detoxification of excess ammonium (Moschou et al. 2012). Thus, the increased concentrations of putrescine and spermidine in the mutant plants could be an indication for higher rates of protein turnover. However, in contrast to the nitrogen-rich amino acids high levels of polyamines seem to delay the onset of senescence (Moschou et al. 2012).

A change in the spatiotemporal distribution of TCA cycle intermediates within individual leaves has been shown to occur in *Arabidopsis* plants gown under short day conditions before the onset of senescence (Watanabe et al. 2013). Concentrations of succinate and fumarate were consistently reduced in all three mutants related to cysteine catabolism. Since for the present study samples were harvested at the end of the dark period, the TCA cycle can be assumed to be in the cyclic mode in wild type leaves to meet the high energy demand during the night. Thus, the strong decrease in the central part of the cycle would indicate a change in the flux mode towards production of precursors for other pathways and may be connected to a higher nitrogen turnover (Sweetlove et al. 2010).

Serine and glycine levels have been shown to increase during senescence at least in plants grown under short day conditions (Watanabe et al. 2013). We observed the same trend in young rosettes of the three mutant lines. Glycine and to a lesser extent also serine accumulate during the day due to their role in photorespiration leading to an increased glycine/serine ratio (Timm et al. 2013). High glycine and serine levels in the morning might indicate changes in the photorespiratory pathway that prevent efficient removal at night. In addition, one carbon-metabolism could be affected in the mutants since catabolic reactions of serine as well as glycine transfer methyl groups to tetrahydrofolate (Hildebrandt et al. 2015).

According to these results, defects in the mitochondrial cysteine degradation pathway lead to several metabolic shifts that normally occur during senescence. However, since *str1-1* plants developed normally additional effects were obviously necessary to actually induce premature leaf senescence.

A persulfide intermediate inhibits amino acid catabolism, affects the C/N status, and induces early senescence under light limitation

Our results show that the sulfur dioxygenase ETHE1, which is involved in cysteine degradation, is relevant for vegetative growth and the onset of senescence particularly under light limiting conditions. This finding is in line with recent studies demonstrating the important function of autophagy and protein turnover for energy production during the night as well as for nitrogen remobilization (Guiboileau et al. 2012, 2013, Izumi et al. 2013). Autophagy-defective *Arabidopsis* mutants display reduced growth under short-day conditions and earlier and more rapid leaf senescence than the wild type (Guiboileau et al. 2012, Izumi et al. 2013). Thus, regular protein turnover for nutrient remobilization during the night is essential mainly under short light periods. Knockdown ETHE1 could lead to a general disturbance in amino acid catabolism and thus cause similar defects as a block of earlier steps in the autophagic process. Indeed, branched-chain amino acids, aromatic amino acids, and lysine accumulated in young *ethe1-1* plants at the end of the dark period. Energy yield from the oxidation of these amino acids is particularly high, and they typically increase during carbohydrate starvation probably to serve as an alternative substrate for ATP production (Araujo et al. 2011, Hildebrandt et al. 2015). However, we did not detect a depletion of carbohydrates in the *ethe1-1* plants, but in contrast sugar levels except glucose were even increased compared to the wild type. Thus, the accumulation of amino acids was probably not caused by an increase in protein turnover as a reaction to carbohydrate starvation. An alternative explanation could be inhibition or down-regulation of individual steps in the catabolic pathways. Interestingly, the accumulation of amino acids connected to energy metabolism was specific for *ethe1-1* and not present in the *ethe1-1 x str1-1* double mutant indicating that most likely this effect was caused by a persulfide intermediate that is not produced in the absence of the sulfurtransferase. Persulfides are highly reactive (nucleophilic and reducing) and have been identified as critical components of redox cell signaling in mammals (Ida et al. 2014). GSSH as well as protein bound cysteine persulfides are endogenously produced and maintained in animal cells, and our results indicate that these

reactive sulfur species might have a similar function in plant tissues. The catabolic pathways for branched-chain and aromatic amino acids are rather complicated and only partially known in plants (Hildebrandt et al. 2015). Therefore, it is difficult to speculate about which particular step might be inhibited or regulated by a persulfide intermediate, and further, more detailed metabolite analyzes will be necessary to identify the position of the block.

In contrast to the situation in seeds, the function of sulfurtransferase 1 does not seem to be essential for vegetative growth in *Arabidopsis*. The *str1-1* mutant showed no phenotype differences compared to the wild type under the conditions tested indicating that all necessary physiological functions can probably be compensated by one of the additional 19 sulfurtransferase isoforms (Bartels et al. 2007). Most notably, the early senescence phenotype characteristic for *ethe1-1* plants grown under short-day conditions was not present in the Str1 deficient mutant. Thus, like the inhibitory effect on amino acid catabolism premature senescence is most likely caused not by a general decrease in cysteine degradation but rather by accumulation of a persulfide intermediate. In this case it could be expected that the phenotype can be rescued in the double mutant, which we observed only in some of the plants. The high variability in *str1-1* seed development already indicates that the compensation mechanism by sulfurtransferase isoforms does not always work equally well. Therefore, the different phenotypes of the *ethe1-1 x str1-1* double mutant might be due to distinct expression profiles of sulfurtransferases which are able to replace Str1 to a different extent. Considering their potential role in cell signaling the balance between persulfide production and removal can be expected to be critical for plant development, and persulfide accumulation will have different effects depending on the extent, timing and intracellular localization. This interesting aspect has to be further investigated.

Another question still remaining open is which factors actually induce premature senescence under short day conditions in ETHE1 deficient plants. A shift in the C/N-status reflected by an accumulation of hexoses in combination with increased concentrations of amino acids and other metabolites related to nitrogen remobilization is discussed to be an internal signal for the induction of leaf senescence (Aguera et al. 2010, Wingler et al 2009). While metabolites associated with nitrogen remobilization were high in all mutant plants analyzed, hexoses as well as the stress-related sugars galactinol and raffinose particularly accumulated in *ethe1-1* plants. Thus, the increase in free sugar levels in combination with a low nitrogen status might act as a metabolic signal for the onset of senescence in young ETHE1 deficient plants.

However, the mechanistic context between persulfide and sugar accumulation still has to be established.

After 100 days of growth under short day conditions metabolite profiles of wild type and *str1-1* plants clearly indicated that the senescence process had already started although only minor chlorotic patches were visible in the leaves. Some of the characteristic differences detected in *ethe1-1* and *ethe1-1 x str1-1* plants can be attributed to the advanced senescence state of these mutants compared to the wild type (Watanabe et al. 2014). A high degree of chlorosis correlated with increased concentrations of the stress related sugars raffinose and galactinol, as well as the amino acids connected to photorespiration glycine and serine. Levels of ascorbate and dehydroascorbate were low, which might indicate a more oxidized state of the dying cells. In addition, changes in the TCA cycle mode probably reflect the lower energy demand and increased nutrient remobilization during late senescence.

Interestingly, some metabolite patterns detected in *ethe1-1* and double mutant plants were contrary to the characteristic senescence profile of *Arabidopsis* leaves and could therefore reflect specific effects of persulfides. The accumulation of nitrogen-rich amino acids as well as hexoses was clearly reduced in *ethe1-1* plants. Hexose concentrations were also low in the *ethe1-1 x str1-1* double mutant indicating that carbohydrate starvation was induced by the defect in ETHE1 protein and independent of the presence of the sulfurtransferase. Interestingly, nitrogen remobilization correlated with the size of the double mutant plants. Concentrations of arginine, asparagine, and ornithine were low in *ethe1-1* and the small plants of the double mutant but comparable to wild type levels in the big double mutant plants, indicating that growth repression might be connected to consistent changes in the nitrogen status.

Conclusions

None of the previously published pathways of cysteine catabolism consider the ultimate fate of the thiol group. However, our results clearly demonstrate that a tight regulation of the concentrations of reactive persulfide intermediates is highly important. An exciting question for further research will be whether persulfides also have a defined signaling function in plants as described in animals. By now it is already obvious that either toxic or regulatory effects of persulfides interfere with central aspects of plant metabolism during seed

development as well as vegetative growth under light limiting conditions. Thus, a better understanding of the underlying mechanism will provide a basis for improving plant yield and fitness.

Author contributions

S.H. and M.B. performed and evaluated experiments addressing the individual steps of the cysteine catabolic pathway, C.L. and T.B. characterized the *ethe1-1 x str 1-1* double mutant, T.T. and A.R.F. performed metabolite profiling, H.P.B. contributed to the scientific concept, T.M.H. designed the experiments, evaluated data and wrote the manuscript.

Acknowledgements

We thank Jutta Papenbrock and Pamela von Trzebiatowski (Institute for Botany, Leibniz University Hannover) for providing the plasmid for Str1 expression and valuable assistance during heterologous expression and purification of Str1 and ETHE1. We are also grateful to Antje Bolze (Max-Planck-Institute of Molecular Plant Physiology) for skillful technical assistance during metabolite analysis and Herbert Geyer as well as Jens-Peter Barth (Leibniz University Hannover) for growing the plants.

References

Agüera E, Cabello P, La Haba P de (2010) Induction of leaf senescence by low nitrogen nutrition in sunflower (Helianthus annuus) plants. Physiologia plantarum 138(3):256–267. doi: 10.1111/j.1399-3054.2009.01336.x

Alvarez C, Calo L, Romero LC, García I, Gotor C (2010) An O-acetylserine(thiol)lyase homolog with L-cysteine desulfhydrase activity regulates cysteine homeostasis in *Arabidopsis*. Plant Physiol. 152(2):656–669. doi: 10.1104/pp.109.147975

Álvarez C, García I, Moreno I, Pérez-Pérez ME, Crespo JL, Romero LC, Gotor C (2012) Cysteine-generated sulfide in the cytosol negatively regulates autophagy and modulates the transcriptional profile in *Arabidopsis*. Plant Cell 24(11):4621–4634. doi: 10.1105/tpc.112.105403

Araújo WL, Tohge T, Ishizaki K, Leaver CJ, Fernie AR (2011) Protein degradation - an alternative respiratory substrate for stressed plants. Trends Plant Sci. 16(9):489–498. doi: 10.1016/j.tplants.2011.05.008

Baker DH (2006) Comparative species utilization and toxicity of sulfur amino acids. The Journal of nutrition 136(6 Suppl):1670S-1675S

Balk J, Pilon M (2011) Ancient and essential: the assembly of iron-sulfur clusters in plants. Trends in plant science 16(4):218–226. doi: 10.1016/j.tplants.2010.12.006

Bartels A, Mock H, Papenbrock J (2007) Differential expression of *Arabidopsis* sulfurtransferases under various growth conditions. Plant physiology and biochemistry : PPB / Société française de physiologie végétale 45(3-4):178–187. doi: 10.1016/j.plaphy.2007.02.005

Boyes DC, Zayed AM, Ascenzi R, McCaskill AJ, Hoffman NE, Davis KR, Görlach J (2001) Growth stage-based phenotypic analysis of *Arabidopsis*: a model for high throughput functional genomics in plants. The Plant cell 13(7):1499–1510

Couturier J, Touraine B, Briat J, Gaymard F, Rouhier N (2013) The iron-sulfur cluster assembly machineries in plants: current knowledge and open questions. Front Plant Sci 4:259. doi: 10.3389/fpls.2013.00259

Fait A, Angelovici R, Less H, Ohad I, Urbanczyk-Wochniak E, Fernie AR, Galili G (2006) *Arabidopsis* seed development and germination is associated with temporally distinct metabolic switches. Plant physiology 142(3):839–854. doi: 10.1104/pp.106.086694

Fernie AR, Aharoni A, Willmitzer L, Stitt M, Tohge T, Kopka J, Carroll AJ, Saito K, Fraser PD, DeLuca V (2011) Recommendations for reporting metabolite data. The Plant cell 23(7):2477–2482. doi: 10.1105/tpc.111.086272

Galili G, Avin-Wittenberg T, Angelovici R, Fernie AR (2014) The role of photosynthesis and amino acid metabolism in the energy status during seed development. Frontiers in plant science 5:447. doi: 10.3389/fpls.2014.00447

Guiboileau A, Avila-Ospina L, Yoshimoto K, Soulay F, Azzopardi M, Marmagne A, Lothier J, Masclaux-Daubresse C (2013) Physiological and metabolic consequences of autophagy deficiency for the management of nitrogen and protein resources in *Arabidopsis* leaves depending on nitrate availability. The New phytologist 199(3):683–694. doi: 10.1111/nph.12307

Guiboileau A, Yoshimoto K, Soulay F, Bataillé M, Avice J, Masclaux-Daubresse C (2012) Autophagy machinery controls nitrogen remobilization at the whole-plant level under both limiting and ample nitrate conditions in *Arabidopsis*. The New phytologist 194(3):732–740. doi: 10.1111/j.1469-8137.2012.04084.x

Hildebrandt TM, Grieshaber MK (2008) Three enzymatic activities catalyze the oxidation of sulfide to thiosulfate in mammalian and invertebrate mitochondria. The FEBS journal 275(13):3352–3361. doi: 10.1111/j.1742-4658.2008.06482.x

Hildebrandt TM, Nunes Nesi A, Araújo WL, Braun H (2015) Amino Acid Catabolism in Plants. Molecular plant 8(11):1563–1579. doi: 10.1016/j.molp.2015.09.005

Holdorf MM, Owen HA, Lieber SR, Yuan L, Adams N, Dabney-Smith C, Makaroff CA (2012) *Arabidopsis* ETHE1 encodes a sulfur dioxygenase that is essential for embryo and endosperm development. Plant physiology 160(1):226–236. doi: 10.1104/pp.112.201855

Ida T, Sawa T, Ihara H, Tsuchiya Y, Watanabe Y, Kumagai Y, Suematsu M, Motohashi H, Fujii S, Matsunaga T, Yamamoto M, Ono K, Devarie-Baez NO, Xian M, Fukuto JM, Akaike T (2014) Reactive cysteine persulfides and S-polythiolation regulate oxidative stress and redox signaling. Proceedings of the National Academy of Sciences of the United States of America 111(21):7606–7611. doi: 10.1073/pnas.1321232111

Izumi M, Hidema J, Makino A, Ishida H (2013) Autophagy contributes to nighttime energy availability for growth in *Arabidopsis*. Plant physiology 161(4):1682–1693. doi: 10.1104/pp.113.215632

Jin Z, Xue S, Luo Y, Tian B, Fang H, Li H, Pei Y (2013) Hydrogen sulfide interacting with abscisic acid in stomatal regulation responses to drought stress in *Arabidopsis*. Plant Physiol. Biochem. 62:41–46. doi: 10.1016/j.plaphy.2012.10.017

Kopka J, Schauer N, Krueger S, Birkemeyer C, Usadel B, Bergmüller E, Dörmann P, Weckwerth W, Gibon Y, Stitt M, Willmitzer L, Fernie AR, Steinhauser D (2005) GMD@CSB.DB: the Golm Metabolome Database. Bioinformatics (Oxford, England) 21(8):1635–1638. doi: 10.1093/bioinformatics/bti236

Krüßel L, Junemann J, Wirtz M, Birke H, Thornton JD, Browning LW, Poschet G, Hell R, Balk J, Braun H, Hildebrandt TM (2014) The mitochondrial sulfur dioxygenase ETHYLMALONIC ENCEPHALOPATHY PROTEIN1 is required for amino acid catabolism during carbohydrate starvation and embryo development in *Arabidopsis*. Plant Physiol. 165(1):92–104. doi: 10.1104/pp.114.239764

Lim PO, Kim HJ, Nam HG (2007) Leaf senescence. Annual review of plant biology 58:115–136. doi: 10.1146/annurev.arplant.57.032905.105316

Lisec J, Schauer N, Kopka J, Willmitzer L, Fernie AR (2006) Gas chromatography mass spectrometry-based metabolite profiling in plants. Nature protocols 1(1):387–396. doi: 10.1038/nprot.2006.59

Luedemann A, Strassburg K, Erban A, Kopka J (2008) TagFinder for the quantitative analysis of gas chromatography--mass spectrometry (GC-MS)-based metabolite profiling experiments. Bioinformatics (Oxford, England) 24(5):732–737. doi: 10.1093/bioinformatics/btn023

Machingura M, Ebbs SD (2014) Functional Redundancies in Cyanide Tolerance Provided by β-Cyanoalanine Pathway Genes in *Arabidopsis thaliana*. International Journal of Plant Sciences 175(3):346–358. doi: 10.1086/674450

Mao G, Wang R, Guan Y, Liu Y, Zhang S (2011) Sulfurtransferases 1 and 2 play essential roles in embryo and seed development in *Arabidopsis thaliana*. The Journal of biological chemistry 286(9):7548–7557. doi: 10.1074/jbc.M110.182865

Masclaux-Daubresse C, Daniel-Vedele F, Dechorgnat J, Chardon F, Gaufichon L, Suzuki A (2010) Nitrogen uptake, assimilation and remobilization in plants: challenges for sustainable and productive agriculture. Annals of botany 105(7):1141–1157. doi: 10.1093/aob/mcq028

Matthies A, Rajagopalan KV, Mendel RR, Leimkühler S (2004) Evidence for the physiological role of a rhodanese-like protein for the biosynthesis of the molybdenum cofactor in humans. Proceedings of the National Academy of Sciences of the United States of America 101(16):5946–5951. doi: 10.1073/pnas.0308191101

Mattoo AK, Minocha SC, Minocha R, Handa AK (2010) Polyamines and cellular metabolism in plants: transgenic approaches reveal different responses to diamine putrescine versus higher polyamines spermidine and spermine. Amino acids 38(2):405–413. doi: 10.1007/s00726-009-0399-4

May MJ, Leaver CJ (1993) Oxidative Stimulation of Glutathione Synthesis in *Arabidopsis thaliana* Suspension Cultures. Plant physiology 103(2):621–627

Meyer T, Burow M, Bauer M, Papenbrock J (2003) *Arabidopsis* sulfurtransferases: investigation of their function during senescence and in cyanide detoxification. Planta 217(1):1–10. doi: 10.1007/s00425-002-0964-5

Moschou PN, Wu J, Cona A, Tavladoraki P, Angelini R, Roubelakis-Angelakis KA (2012) The polyamines and their catabolic products are significant players in the turnover of nitrogenous molecules in plants. Journal of experimental botany 63(14):5003–5015. doi: 10.1093/jxb/ers202

Ogata K, Volini M (1990) Mitochondrial rhodanese: membrane-bound and complexed activity. The Journal of biological chemistry 265(14):8087–8093

Osman LP, Mitchell SC, Waring RH (1997) Cysteine, its Metabolism and Toxicity. Sulfur reports 20(2):155–172. doi: 10.1080/01961779708047918

Papenbrock J, Schmidt A (2000) Characterization of a sulfurtransferase from *Arabidopsis thaliana*. European journal of biochemistry / FEBS 267(1):145–154

Peiser GD, Wang TT, Hoffman NE, Yang SF, Liu HW, Walsh CT (1984) Formation of cyanide from carbon 1 of 1-aminocyclopropane-1-carboxylic acid during its conversion to ethylene. Proceedings of the National Academy of Sciences of the United States of America 81(10):3059–3063

Riemenschneider A, Wegele R, Schmidt A, Papenbrock J (2005) Isolation and characterization of a D-cysteine desulfhydrase protein from *Arabidopsis thaliana*. FEBS J. 272(5):1291–1304. doi: 10.1111/j.1742-4658.2005.04567.x

Schauer N, Steinhauser D, Strelkov S, Schomburg D, Allison G, Moritz T, Lundgren K, Roessner-Tunali U, Forbes MG, Willmitzer L, Fernie AR, Kopka J (2005) GC-MS libraries for the rapid identification of metabolites in complex biological samples. FEBS letters 579(6):1332–1337. doi: 10.1016/j.febslet.2005.01.029

Sunderhaus S, Dudkina NV, Jänsch L, Klodmann J, Heinemeyer J, Perales M, Zabaleta E, Boekema EJ, Braun H (2006) Carbonic anhydrase subunits form a matrix-exposed domain attached to the membrane arm of mitochondrial complex I in plants. The Journal of biological chemistry 281(10):6482–6488. doi: 10.1074/jbc.M511542200

Sweetlove LJ, Beard KFM, Nunes-Nesi A, Fernie AR, Ratcliffe RG (2010) Not just a circle: flux modes in the plant TCA cycle. Trends in plant science 15(8):462–470. doi: 10.1016/j.tplants.2010.05.006

Timm S, Florian A, Wittmiß M, Jahnke K, Hagemann M, Fernie AR, Bauwe H (2013) Serine acts as a metabolic signal for the transcriptional control of photorespiration-related genes in *Arabidopsis*. Plant physiology 162(1):379–389. doi: 10.1104/pp.113.215970

Ubuka T, Umemura S, Yuasa S, Kinuta M, Watanabe K (1978) Purification and characterization of mitochondrial cysteine aminotransferase from rat liver. Physiological chemistry and physics 10(6):483–500

van Hoewyk D, Pilon M, Pilon-Smits EA (2008) The functions of NifS-like proteins in plant sulfur and selenium metabolism. Plant Science 174(2):117–123. doi: 10.1016/j.plantsci.2007.10.004

Watanabe M, Balazadeh S, Tohge T, Erban A, Giavalisco P, Kopka J, Mueller-Roeber B, Fernie AR, Hoefgen R (2013) Comprehensive dissection of spatiotemporal metabolic shifts in primary, secondary, and lipid metabolism during developmental senescence in *Arabidopsis*. Plant physiology 162(3):1290–1310. doi: 10.1104/pp.113.217380

Werhahn W, Niemeyer A, Jänsch L, Kruft V, Schmitz UK, Braun H (2001) Purification and characterization of the preprotein translocase of the outer mitochondrial membrane from *Arabidopsis*. Identification of multiple forms of TOM20. Plant physiology 125(2):943–954

Westley J (1973) Rhodanese. Advances in enzymology and related areas of molecular biology 39:327–368

Westley J, Adler H, Westley L, Nishida C (1983) The sulfurtransferases. Fundamental and applied toxicology : official journal of the Society of Toxicology 3(5):377–382

Wingler A, Masclaux-Daubresse C, Fischer AM (2009) Sugars, senescence, and ageing in plants and heterotrophic organisms. Journal of experimental botany 60(4):1063–1066. doi: 10.1093/jxb/erp067

Figures

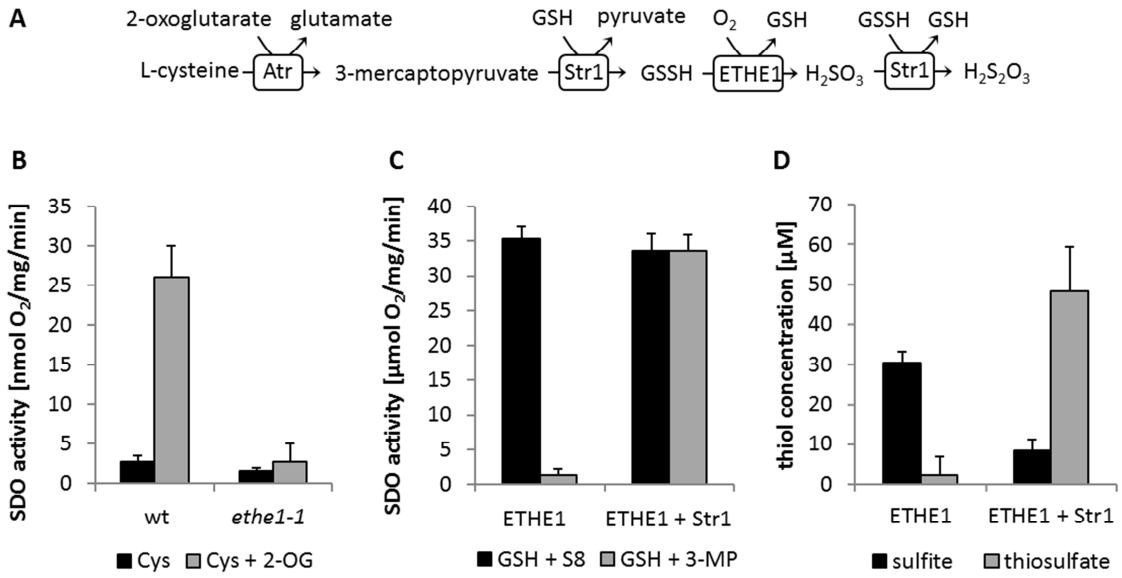


Fig. 1: Reaction steps of the mitochondrial L-cysteine catabolic pathway. A, Reaction scheme of the oxidation of L-cysteine to pyruvate and thiosulfate catalyzed by a cysteine transaminase (Atr), sulfurtransferase 1 (Str1), and the sulfur dioxygenase ETHE1. B, Sulfur dioxygenase (SDO) activity [nmol/mg protein/min] of mitochondria isolated from wild type and *ethe1-1* cell culture after addition of either L-cysteine (5 mM, black bars) or L-cysteine (5 mM) plus 2-oxoglutarate (5 mM, grey bars) as a substrate. C, Sulfur dioxygenase activity [µmol O_2/mg protein/min] of isolated recombinant ETHE1 enzyme in the absence and presence of isolated recombinant sulfurtransferase 1 (Str1) after addition of reduced glutathione (1 mM) plus elemental sulfur (15 µl/ml) (GSH + S8), which non-enzymatically produce glutathione persulfide (black bars), or reduced glutathione (1 mM) plus 3-mercaptopyruvate (5 mM) (GSH + 3-MP, grey bars). D, Concentrations of the persulfide oxidation products sulfite (black bars) and thiosulfate (grey bars) in the reaction mixture containing 2 µg/ml purified recombinant ETHE1 enzyme in the absence or presence of 20 µg/ml purified recombinant sulfurtransferase 1 ten minutes after the addition of GSH plus sulfur as a substrate.

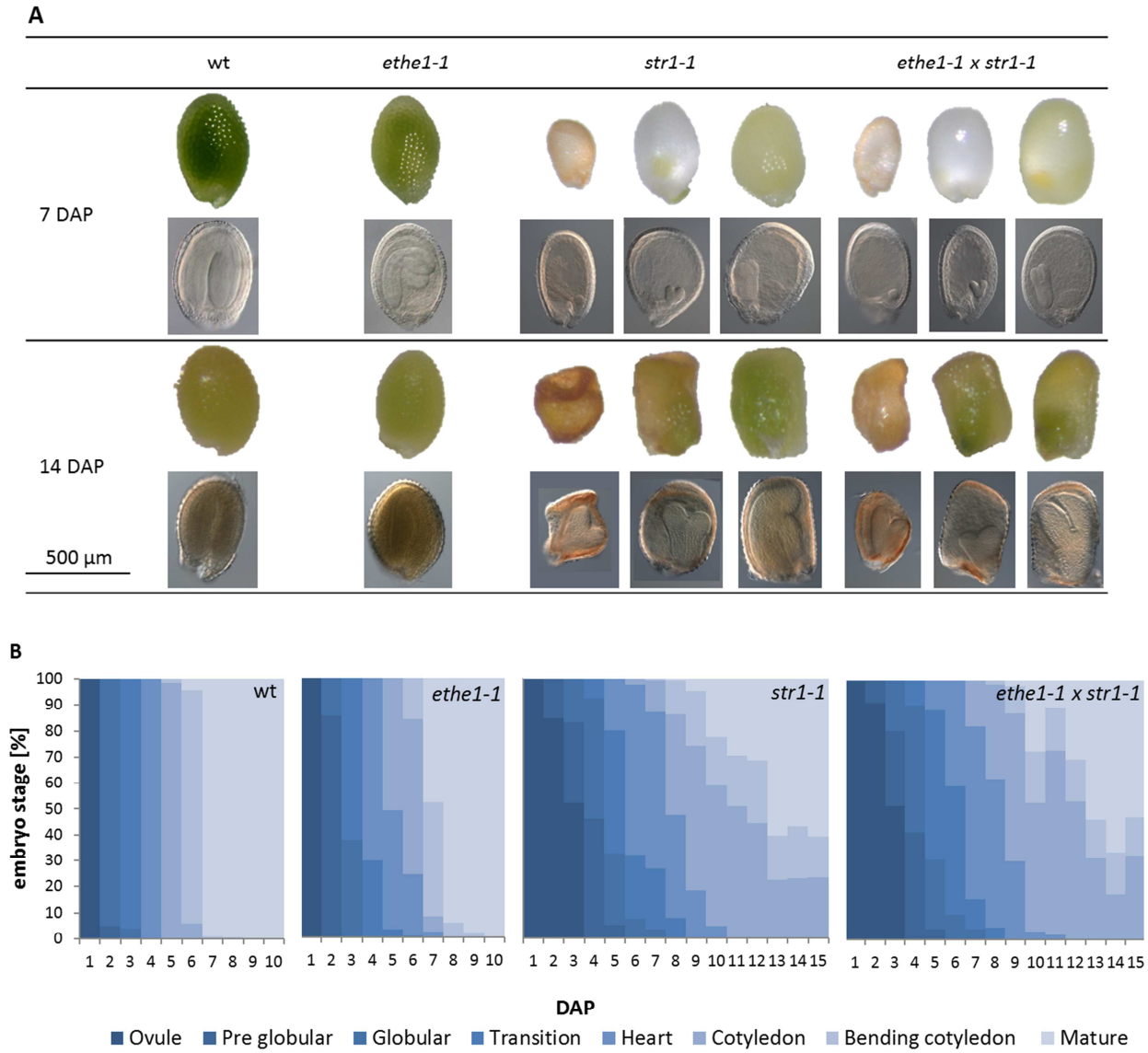


Fig. 2: Embryo development in the *ethe1-1 x str1-1* double mutant compared to *ethe1-1* and *str1-1* single mutants and the wild type. A, Representative images of seeds and embryos at 7 and 14 days after pollination (DAP). B, Progression of embryo development from ovule to mature stage. Approximately 220 seeds were analyzed for each day and genotype, and bars respresent the percentage of embryos in the developmental stage indicated.

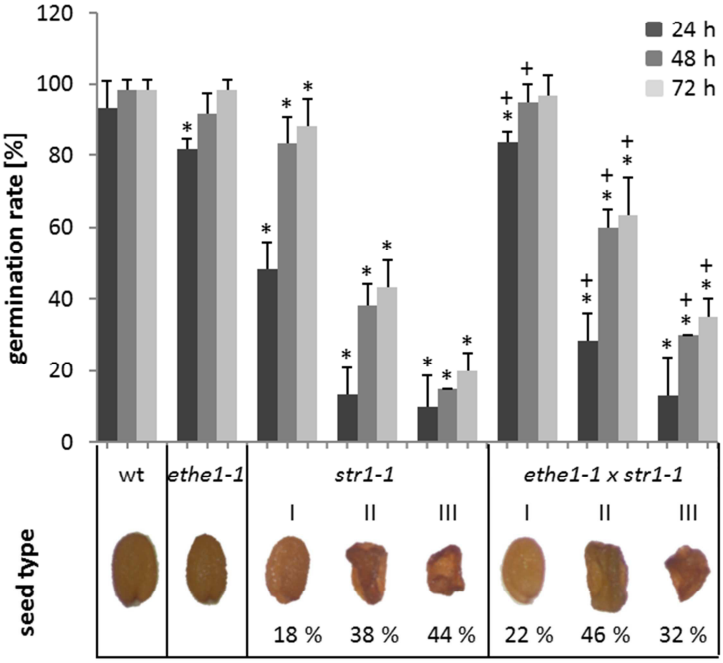


Fig. 3: Phenotype and germination rates of mature seeds in the *ethe1-1 x str1-1* double mutant compared to *ethe1-1* and *str1-1* single mutants and the wild type. Whereas wild type and *ethe1-1* plants consistently produced intact seeds, three different phenotypes (I, intact; II, rectangular, III, shriveled) were present in *str1-1* plants and the double mutant in a ratio indicated below the respective seed type. Bars represent the germination rates [%] of the different seed types after 24 h (dark grey), 48 h (medium grey), and 71 h (light grey). Asterisks indicate values significantly different from the wild type, and values that are significantly different in the double mutant from the corresponding data points in the str1 mutant are marked with a plus ($p < 0.05$, Student's t-test).

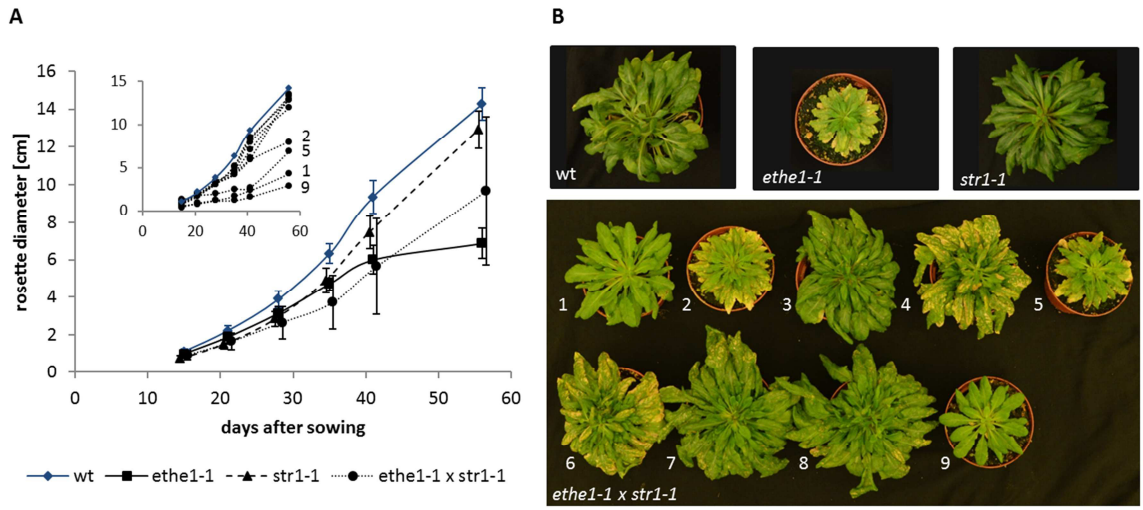


Fig. 4: Phenotype of the *ethe1-1 x str1-1* double mutant compared to *ethe1-1* and *str1-1* single mutants and the wild type under short-day growth conditions (8 h light/16 h dark). A, Means ± standard deviations of maximal rosette diameters from 9 to 10 plants [cm]. Since double mutant plants had a very heterogeneous phenotype, growth curves of the individual plants are shown in the inserted graph. B, Rosettes of the 9 double mutant plants analyzed and representative plants of the *ethe1-1* and *str1-1* single mutants and the wild type after 100 days of growth under short-day conditions.

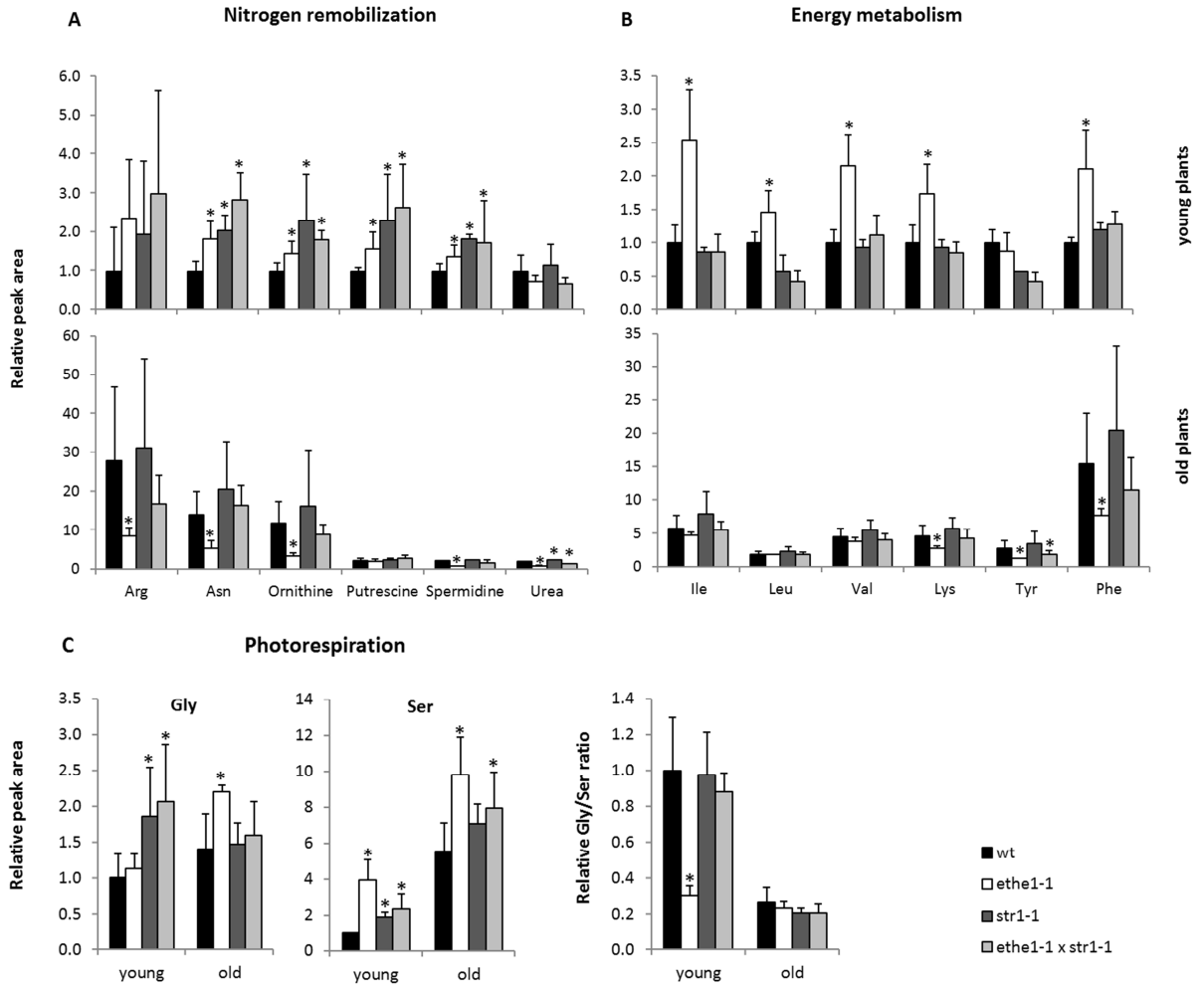


Fig. 5: Amino acid profile of the *ethe1-1 x str1-1* double mutant compared to *ethe1-1* and *str1-1* single mutants and the wild type under short-day growth conditions (8 h light/16 h dark). Complete rosettes of 3-5 plants were harvested at dawn at the age of 42 days (young plants) and 100 days (old plants) and analyzed by GC/MS. Amino acid contents are given as relative peak areas normalized to young wild type plants. *Significantly different from corresponding wild type samples ($p<0.05$, Student's t-test).

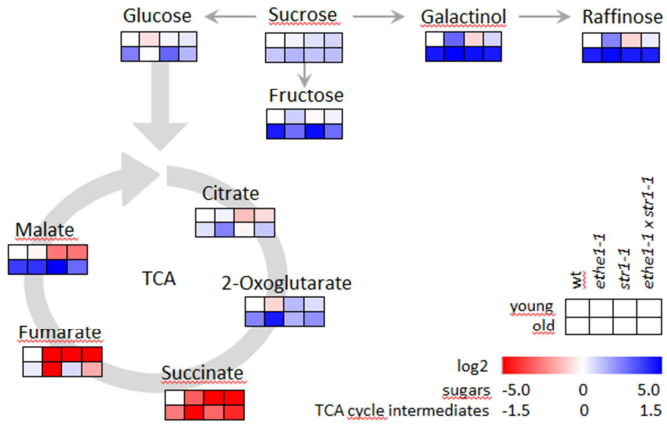


Fig. 6: Heatmap of metabolite changes related to sugar metabolism and the TCA cycle in young and old wild type and mutant plants. Log_2 ratios of fold changes in relation to young wild type samples are indicated by shades of blue or red color according to the scale bar. Data represent mean values of three to nine biological replicates for each data point. Rosettes were harvested at dawn after 42 days (young) and 100 days (old) of growth under short-day conditions (8h light/16h dark). The complete dataset (means ± standard deviations, statistical analysis) is presented in Supplemental Table S2.

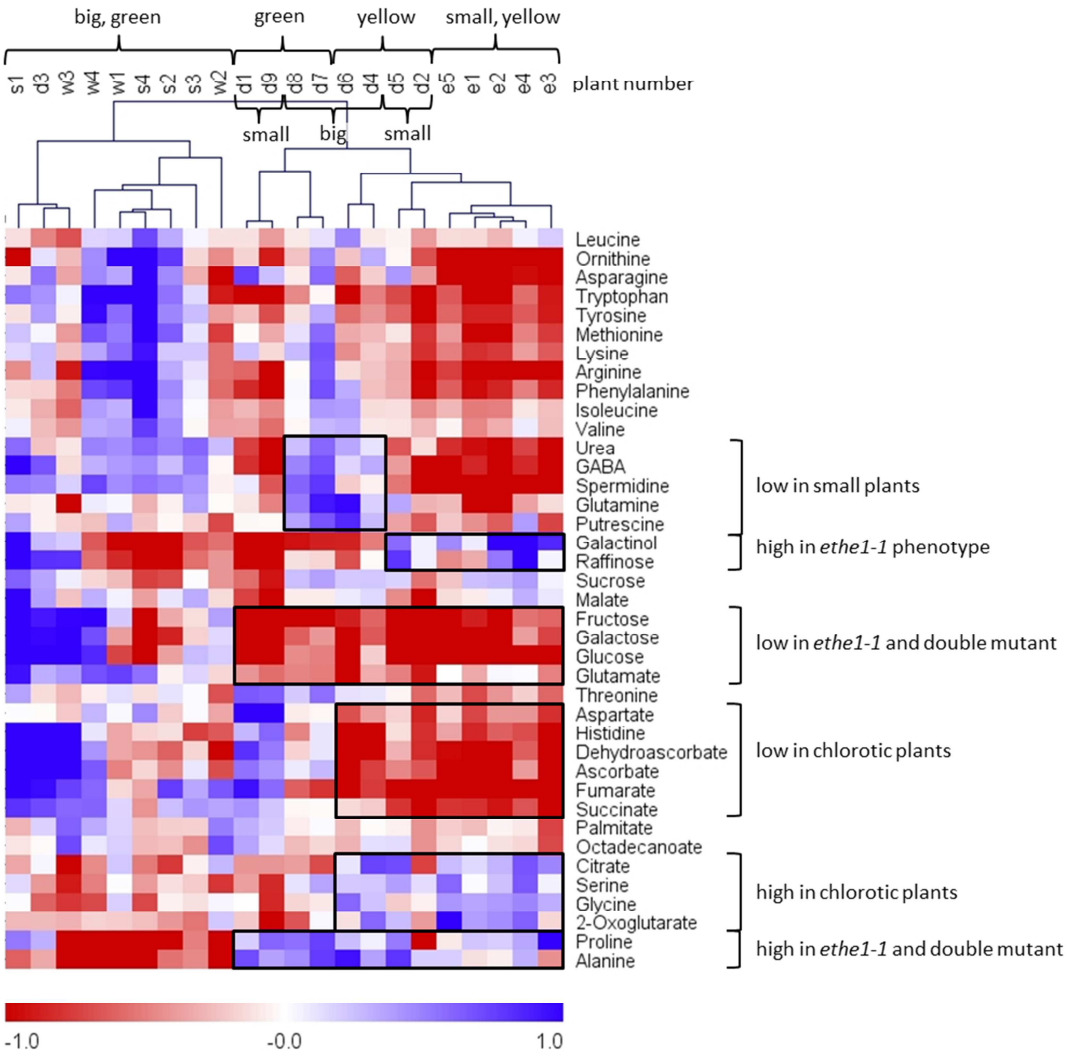


Fig. 7: Clustering of individual mutant and wild type plants harvested after 100 days of growth under short-day conditions according to their metabolite profiles. Log_2 ratios of fold changes in relation to the mean peak area of each metabolite are indicated by shades of blue or red color according to the scale bar. Rosettes of 4 wild type plants (w1-4), 4 *str1-1* plants (s1-4), 5 *ethe1-1* plants (e1-5), and the nine *ethe1-1 x str1-1* plants shown in Fig. 4B (d1-9) were harvested at dawn after 100 days of growth under short-day conditions (8h light/16h dark). Hierarchical clustering was performed using MeV software. Plant numbers are given in the upper part of the figure together with some phenotype information. Black boxes mark metabolite patterns characteristic for individual phenotypes. The complete dataset (means ± standard deviations, statistical analysis) is presented in Supplemental Table S2.

Manuscript 2

2.6 Amino acid metabolism and the role of ETHE1 in *Arabidopsis thaliana* seeds

Christin Lorenz[1], Ljudmilla Borisjuk[2], Hardy Rolletschek[2], Nicolas Heinzel[2], Takayuki Tohge[3], Alisdair R. Fernie[3], Hans-Peter Braun[1], Tatjana M. Hildebrandt[1]

[1] Department of Plant Proteomics, Institute for Plant Genetics, Faculty of Natural Sciences, Leibniz Universität Hannover
[2] Department of Molecular Genetics, Leibniz Institute of Plant Genetics and Crop Plant Research (IPK) Gatersleben
[3]Max-Planck-Institute of Molecular Plant Physiology, Potsdam-Golm

Type of authorship:	First author
Type of article:	Research article
Share of the work:	85 %
Contribution to the publication:	Planned and performed experiments, analysed data, prepared all figures and wrote the paper
Status of publication:	In preparation

*****Note that the supplementary material for chapter 2.6 is not printed but included in the compact disc attached to this work.

Amino acid metabolism and the role of ETHE1 in *Arabidopsis thaliana* seeds

Lorenz C.[1], Borisjuk L. [2], Rolletschek H.[2], Heinzel N.[2], Tohge T.[3], Fernie A.R.[3], Braun H.-P.[1], Hildebrandt T.M.[1]

[1]Department of Plant Proteomics, Institute for Plant Genetics, Faculty of Natural Sciences, Leibniz Universität Hannover

[2] Department of Molecular Genetics, Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Gatersleben

[3]Max-Planck-Institute of Molecular Plant Physiology, Potsdam-Golm

Abstract

The sulfur dioxygenase ETHE1 oxidizes persulfides in the mitochondrial matrix and is involved in the degradation of L-cysteine. ETHE1 has an essential but yet undefined function in early embryo development of *Arabidopsis thaliana*. In leaves, ETHE1 is strongly induced by extended darkness and takes part in the use of amino acids as alternative respiratory substrates during carbohydrate starvation. Thus, we tested the effect of darkness on seed development in an ETHE1 deficient mutant compared to the wild type. Since ETHE1 knock out is embryo lethal, the knock down line *ethe1-1* with about 1 % residual sulfur dioxygenase activity was used for this study. We performed phenotype analysis, metabolite profiling and comparative proteomics to investigate the general effect of extended darkness on seed metabolism and further define the particular function of the mitochondrial sulfur dioxygenase ETHE1 in seeds. Shading of the siliques had no morphological effect on embryogenesis in wild type plants. However, the developmental delay that was already visible in *ethe1-1* seeds under control conditions was further enhanced in the darkness. Interestingly, dark conditions strongly affected seed quality parameters of wild type as well as mutant plants. The lipid content of the seeds was severely decreased to about 20 % of light conditions, and there was also a clear reduction in free sugar concentrations. In contrast, amino acids accumulated with the strongest effect on the nitrogen-rich amino acids asparagine and glutamine. Proteome as well as metabolite data revealed a major regulatory effect of light on nitrogen turnover in the developing seeds. In addition, first evidence for a function of branched-chain amino acid catabolism in seed energy production was provided. Knockdown of ETHE1 lead to alterations in endosperm cellularization as well as mitochondrial

structure. Interestingly, the effect on amino acid profiles was clearly different from that found in leaves indicating that in seeds the ETHE1 dependent pathway interacts with alanine and glycine rather than branched-chain amino acid metabolism.

Introduction

Seed development of plants starts with the unique event of double fertilization. The resulting seed compartments endosperm and embryo undergo specific developmental sequences with dramatic changes on cellular and molecular level from tissue differentiation and growth to storage deposition during seed filling. The endosperm is essential for embryo development. It embeds the embryo and is surrounded by the seed coat. In *A. thaliana* the endosperm is of transient nature and only the aleurone layer remains until seed maturity (Novack et al. 2010). Its development is characterized by a stepwise differentiation including transition from a syncytial to a cellular phase (Olsen 2004, Berger 2003, Brown et al. 2003, Brown et al. 1999). The process of cellularization has a great impact on the accurate embryo development (Hehenberger et al. 2012) and conducts the initiation of embryo growth (Brown et al. 1999). During seed maturation the embryo massively accumulates storage compounds within the cotyledons and becomes a highly specialized storage tissue. At early stages starch and hexoses accumulate transiently, but their amount gradually decreases accompanied by a rapid increase of protein and lipid storage (Borisjuk et al. 2005, Baud et al. 2002). The synthesis of storage compounds requires provision of sufficient amounts of precursor molecules, reductants and energy (Gallardo et al. 2008, Fait et al. 2006). Nutrients such as sugars and amino acids delivered from the mother plant are precursors for storage product biosynthesis in seeds (Melkus et al. 2009). An adaption of energy metabolism including photosynthesis and respiration is essential to enable efficient storage product accumulation. Seed plastids hold special structures and have adapted their metabolism to cope with reduced light levels available for photosynthetic reactions (Borisjuk et al. 2013, Borisjuk et al. 2004). The permeability of gases into seeds is low, and plastidial activity during the maturation phase contributes to oxygen allocation and reassimilation of CO_2 (Borisjuk and Rolletschek 2009, Rolletschek et al. 2005, Borisjuk et al. 2005, Ruuska et al. 2004). Highest photosynthetic activity is observed during storage product synthesis, suggesting a correlation of metabolic processes and photosynthesis (Fait et al. 2006, Ruuska et al. 2002). Recently it was postulated that amino acids are not only precursors for storage proteins within seeds, but might also serve as alternative substrates for mitochondrial metabolism during situations of high energy

demand (Galili et al. 2014). In vegetative tissues amino acid catabolism is induced by carbon starvation situations e.g. periods of extended darkness (Hildebrandt et al. 2015, Araujo et al. 2011 and 2010) First studies denote a potential effect of amino acid degradation also on the energy status of seeds (Credali et al. 2013, Angelovici et al. 2011, Gu et al. 2010, Angelovici et al. 2009, Weigelt et al. 2008, Zhu and Galili 2003). However, biosynthesis as well as catabolism of amino acids in seeds is still largely unknown. The mitochondrial sulfur dioxygenase ETHE1 (AT1G53580) is part of a sulfur catabolic pathway that catalyzes the oxidation of sulfide or persulfides derived from amino acids to thiosulfate and sulfate. In *Arabidopsis* leaves ETHE1 has a key function in amino acid catabolism in situations of carbohydrate starvation such as extended darkness (Krüßel et al. 2014). In seeds, ETHE1 knock out leads to alterations in endosperm formation and finally causes seed abortion (Holdorf et al. 2012). It has been shown that a sulfur dioxygenase activity of 1% present in an ETHE1 knock down mutant (*ethe1-1*) is sufficient for embryo survival, but development is severely delayed (Krüßel et al. 2014), which underlines the importance of this enzyme for seed metabolism.

This study aimed to investigate the physiological role of the mitochondrial sulfur dioxygenase ETHE1 in seeds. One major aspect was to establish whether the functional context of ETHE1 in leaves, the use of amino acids as alternative respiratory substrates, is relevant during embryo development. Therefore, our experimental approach included shading of the siliques combined with metabolite and proteome analysis. Our results show a strong effect of extended darkness on amino acid metabolism in seeds and indicate an additional role of ETHE1 in cell structure formation.

Results

A combination of phenotype analysis, metabolite profiling and comparative proteomics was used to investigate amino acid metabolism in general and the particular function of the mitochondrial sulfur dioxygenase ETHE1 in seeds. Since ETHE1 knock out is embryo lethal (Holdorf et al. 2012), the knock down line *ethe1-1* described in Krüßel et al. (2014) was used for all experiments.

ETHE1 knock down causes a delay in embryogenesis which is further enhanced in the darkness

In leaves ETHE1 is strongly induced by extended darkness and has a key function in the use of amino acids as alternative respiratory substrates during carbohydrate starvation (Krüßel et al. 2014). Thus, we tested the effect of darkness on seed development in the *ethe1-1* mutant by covering the young siliques with aluminium foil one day after pollination (DAP, Supp. Fig. 1). The role of photosynthesis in seeds is still largely unknown. Since carbohydrates are mainly provided by the mother plant, additional functions compared to leaves can be expected. Therefore, we first tested the effect of darkness on seed development in wild type plants. Overall development of the siliques and the number of seeds produced was not influenced by the shading procedure (data not shown). However, the seeds did not turn green at 4-5 DAP like under light conditions but appeared yellowish (Figure 1A). Seeds grown under light and dark conditions were harvested over a period of 1 to 9 DAP. For each time point (1-9 DAP) and light condition (light, dark) embryo development in seeds from 5 siliques was analyzed using a seed clearing method. Interestingly, there were no morphological differences between wild type embryos growing under light or dark conditions. Development from globular to mature stage proceeded very uniformly, and all seeds in a silique were in the same developmental stage (Figure 1B). In contrast, we observed a delay of embryogenesis in *ethe1-1* seeds with several embryo stages present within one silique. Dark conditions induced a stronger phenotype in *ethe1-1* seed, which became apparent starting from 3 DAP were most of the embryos were still in the globular stage, while light grown *ethe1-1* embryos were in the transition stage. At 9 DAP most of the embryos were fully matured, but some were still in the cotyledon and bending cotyledon stage. The

developmental delay of *ethe1-1* seeds compared to the wild type observed under control conditions is further increased by about two days in darkness.

Darkness leads to reduced seed weight and a changed biomass composition

Mature seeds of wild type as well as *ethe1-1* plants grown under light or dark conditions were of the same size and morphologically indistinguishable. However, the final seed weight was drastically reduced in darkness for both genotypes (Figure 2A). To investigate whether the lower weight is caused by a shift in storage product accumulation, we analyzed the final biomass composition of the dormant seeds (Figure 2B). For both, wildtype and *ethe1-1* seeds, lipid contents were decreased to about 20% of light grown seeds. In addition shading of the siliques induced an accumulation of free amino acids. Most notably contents of the nitrogen-rich amino acids Asn and Gln increased more than 20-fold in the darkness compared to control conditions (Figure 2C) In contrast, the protein content was slightly reduced in dark grown seeds, whereas final starch levels remained unaffected. However, concentrations of hexoses and sucrose were strongly decreased after the dark treatment in wild type as well as *ethe1-1* seeds. In agreement with the postulated impact of photosynthesis on germination timing (Allorent et al. 2015), we observed a reduced germination capacity of dark grown seeds (Supp. Fig. 2). Only minor differences were observed regarding biomass composition and germination rate between wildtype and *ethe1-1* seeds.

Proteomics and metabolite profiles

To further investigate the role of ETHE1 in seed metabolism as well as the influence of light on storage product accumulation we performed comparative proteomics and analyzed metabolite profiles of wild type and *ethe1-1* seeds grown under light and dark conditions. Samples were taken at two different time points (Figure 1A, indicated by blue arrows). We selected 5 DAP, since the developmental delay in the mutant seeds was most pronounced at this stage. The endosperm still has an important function and covers a large part of the seed. A second set of samples was collected at 9 DAP, when the seed was completely filled by the embryo and there was no morphological difference between the genotypes or light conditions. Seeds from 8 batches including 10 plants each were pooled and used for shotgun proteomics as well as metabolite

analysis via GC/MS. In total 3958 unique proteins were identified in the shotgun proteomics approach. Technical replicates of the samples clustered very well in a principal component analysis and the different groups were clearly separated so that it was possible to identify sets of differentially abundant proteins (Figure 3). Since the effect of silique shading on the storage compound composition of mature seeds was similar in wild type and *ethe1-1* we first concentrated on the wild type data set to identify the role of light in the seed filling process (Figure 3, comparison 1). We considered 9 DAP to be the most relevant stage for addressing this question because the metabolic activity of the embryo is directed to lipid and storage protein synthesis (Le et al. 2010). Next we analyzed the effect of ETHE1 deficiency on seed metabolism by comparing wild type and mutant seed proteomes at an early (5 DAP) as well as a later (9 DAP) developmental stage (Figure 3, comparison 2a and 2b).

Effects of light on seed energy metabolism and storage product accumulation

The metabolite profiles of wild type as well as *ethe1-1* seeds revealed only a mild shortage of carbohydrates in the absence of photosynthesis. The level of sucrose, which is the main sugar present at 9 DAP, was nearly identical, and also glucose and fructose which are dominant earlier in development were only moderately reduced to 65-95 % of the level present in seeds grown under light conditions. The TCA cycle intermediates citrate, 2-oxoglutarate, succinate, fumarate, and malate were even increased up to 4-fold (Figure 4). In contrast, we detected strong changes in the amino acid composition of seeds grown in shaded siliques (Figure 4). Like already detected in the mature seeds, the most pronounced effect was on amino acids used for nitrogen storage and transport with a 67 fold increase in asparagine concentrations in dark-grown seeds compared to control conditions at 9 DAP. At 5 DAP there was also a clear accumulation of branched-chain and aromatic amino acids detectable in the darkness, which became less pronounced later in development. The proteome comparison indicated a large impact of the availability of light on various metabolic pathways. In wild type seeds harvested at 9 DAP, 338 proteins were of significantly higher abundance in the dark treated samples compared to control conditions and 324 proteins were reduced. To get a global overview of pathways potentially affecting seed composition, proteins were annotated to functional categories, and all spectra obtained from proteins of significantly changed abundance were summed up (Figure 5A). As expected, there was a clear reduction of photosynthesis related proteins including light reaction and Calvin cycle

in dark grown seeds. Interestingly, plastidial proteins with other functions were mainly upregulated. We also detected a strong influence of light on mitochondrial protein composition with several up- and downregulated proteins. A high turnover of lipids was indicated by an increase of proteins related to lipid biosynthesis as well as degradation. Protein synthesis was also increased under dark conditions, while protein degradation as well as amino acid metabolism was rather down regulated. Since metabolite profiling had indicated a strong effect of darkness on asparagine concentrations, we specifically analyzed proteins involved in asparagine metabolism. Aspartate amino transferase 2, which produces aspartate as the precursor for asparagine synthesis, was significantly increased in darkness at 5 and 9 DAP. We also detected 2 peptides of asparagine synthase, but the number of spectra was not sufficient for solid quantification. To identify the most explicit changes induced by darkness, log 2 ratios of differentially abundant proteins were plotted to the number of spectral counts. By considering a fold change of at least 1.5 and a minimum of 5 spectral counts to represent major changes, the number of regulated proteins decreased to 76 proteins with higher abundance and 99 proteins with lower abundance in dark grown seeds of wildtype compared to light conditions (Figure 5B). As expected, the majority of proteins with a decreased abundance in darkness are related to light reactions of photosynthesis such as photosystem I and II and cytochrome b6f subunits. In addition, dark grown seeds show a reduced abundance of proteins balancing redox levels such as stromal ascorbate peroxidase, glutathione-disulfide reductase, copper/zinc superoxide dismutase 1, peroxiredoxin-2E and thioredoxin 3, whereas proteins related to fatty acid degradation (transducin family protein, cycloartenol synthase 1) and proteins involved in nucleotide metabolism (pyrimidine 1, purine biosynthesis 4, Inosine-5'-monophosphate dehydrogenase 2, carbamoyl phosphate synthetase B) showed an increased abundance. Several enzymes related to protein turnover and fatty acid synthesis were strongly affected by changing light conditions. However, proteins annotated to these pathways showed a regulation in both directions suggesting that instead of regulating a whole pathway, individual sub-steps are modulated.

Effects of ETHE1 deficiency on seed metabolism and development

While the composition of mature seeds was very similar in the ETHE1 knockdown mutant and in the wild type, specific differences in the metabolite profile became apparent at earlier developmental stages (Figure 4). At 5 DAP alanine and glycine concentrations were 4 to 5 fold

increased compared to the wild type under light as well as dark growth conditions. Comparison to the metabolite profile of wild type seeds harvested at 4 DAP confirmed, that this effect was not due to the earlier developmental stage present in the mutant. There was a trend towards a general increase in free amino acids in *ethe1-1* seeds compared to the wild type (on average 1.7 fold in the light and 1.9 fold in the darkness) (Figure 4C). Hexose levels in the mutant were also elevated at 5 DAP, which could however be linked to the earlier developmental stage of the *ethe1-1* embryo, since in the wild type hexose concentrations decreased between 4 and 5 DAP (Figure 4A). Sucrose as well as TCA cycle intermediates were unaffected (Figure 4C). At 9 DAP the metabolite profiles of the ETHE1 deficient seeds were more similar to the wild type than at the earlier stage. Amino acid concentrations were only slightly increased (on average 1.3 fold), hexose levels were 1.5-2.5 fold higher, and sucrose was decreased to 70 % of wild type level in the light and to 60 % in shaded siliques.

The proteome analyses of *ethe1-1* seeds compared to wildtype revealed 175 (5 DAP) and 291 (9 DAP) proteins significantly higher in abundance whereas 176 (5 DAP) and 199 (9 DAP) proteins showed a reduced abundance (a full list of proteins with significantly changed abundance is given in the supplementary material). To get a global overview of pathways potentially influenced by ETHE1, proteins were annotated to functional categories and all spectra obtained from proteins of significantly changed abundance at 5 DAP and 9 DAP were summed up (Figure 6A, Figure 7A). To characterize the proteome changes induced by an ETHE1 knock down in more detail, log 2 ratios of differentially abundant proteins observed at 5 DAP and 9 DAP were plotted to the number of spectral counts. A fold change of at least 1.5 and a minimum of 5 spectral counts were considered representing major changes (Figure 6B, Figure 7B).

Overall, the proteome profiles of *ethe1-1* seeds show similar patterns of regulated pathways at the two stages investigated, mainly affecting mitochondrial metabolism, protein turnover, amino acid metabolism and cell organisation. However, the *ethe1-1* seed proteome at 9 DAP differs from that at 5 DAP in respect to light reaction of photosynthesis, Calvin cycle and other plastidial proteins, which are decreased at 5 DAP and increased at 9 DAP compared to wildtype proteomes. Proteins related to sucrose metabolism such as sucrose synthase 1 and sucrose phosphate synthase 1F are exclusively increased in *ethe1-1* seeds at 5 DAP. Protein turnover might be enhanced in *ethe1-1* seeds during development indicated by increased abundance of proteins

related to protein biosynthesis as well as degradation at both stages. In contrast proteins related to amino acid metabolism are reduced at 5 DAP (tryptophan synthase beta-subunit 1, acetylornithine aminotransferase) and 9 DAP (glutamate synthase 2, glutamine synthase, S-adenosyl-l-homocysteine hydrolase 2, dehydroquinate dehydratase, alanine:glyoxylate aminotransferase 3, aspartate aminotransferase 3 and cysteine synthase C1). However, although alanine concentrations were strongly increased in the mutant seeds at 5 DAP, there was no significant effect on protein abundance of alanine aminotransferase 1 catalyzing synthesis as well as degradation of alanine (1.16 fold at 5 DAP, 1.19 fold at 9 DAP). Several enzymes involved in cell wall formation were significantly changed in abundance at 5 DAP (UDP-arabinopyranose mutase 1, root hair defective 3 GTP-binding protein, beta-1,3-glucanase 3, beta glucosidase 18 and PHE ammonia lyase 1) and 9 DAP (FASCICLIN-like arabinogalactan protein 8, beta-D-xylosidase 4, pectinacetylesterase family protein, UDP-D-apiose/UDP-D-xylose synthase 2). In addition *ethe1-1* seeds show an accumulation of tubulin at 5 DAP which is further increased at 9 DAP.

ETHE1 knock down leads to impaired endosperm cellularization

Endosperm cellularization has been shown to be affected by ETHE1 knockout (Holdorf et al. 2012), and our proteomic results indicate that changes in cell wall establishment and cellular organization are also present in the knockdown line *ethe1-1* at 5 DAP as well as 9 DAP. Therefore, we measured the size of central endosperm cells obtained at 7 DAP, in microscopic sections of wild type as well as mutant seeds stained with Toluidine Blue O (Figure 8A and B). The area of about 50 cells per sample type was calculated. The average cell size in wild type endosperm was 156±40 µm². In contrast *ethe1-1* seeds showed characteristically enlarged endosperm cells with a size of 305±96 μm^2. Darkness further increased the endosperm cell size in the mutant, while endosperm of wildtype seeds was not affected.

Aberrant mitochondria in ethe1-1seeds

Previous results obtained with the ETHE1 knockout mutant (Holdorf et al. 2012) as well as our proteome study indicate potential effects of ETHE1 deficiency on seed mitochondria. Therefore, we next investigated the ultracellular structure of embryo and endosperm cells in *ethe1-1* at

5 DAP by using transmission electron microscopy (TEM). The overall organellar structure was relatively normal compared to wildtype, with the exception of the appearance of mitochondria (Figure 8C). The main population of mitochondria in both wildtype and *ethe1-1* were similar in structure (e.g. well distinguished double membrane, lamellar cristae, and relative small area of electron-transparent matrix). The maximal differences in size of mitochondria in embryo of wildtype cells (0.08-0.28 µm²) and of *ethe1-1* (0.04-0.38 µm²) displayed high heterogeneity of population in both genotypes. Similar results were achieved by comparison of mitochondria in endosperm, namely mitochondria size in wildtype (0.05-0.42 µm²) and *ethe1-1* (0.04-0.34 µm²). Extremely high variability of individual organelles did not allow defining any statistically significant differences. However, embryo cells of *ethe1-1* showed aberrant mitochondria, which differed from wildtype at the ultrastructural level in means of strongly enlarged size (up to 5-fold in length), less pronounced lamellar cristae, larger electron transparent areas in the matrix and some electron-dense insertions (Figure 8C).

Discussion

The sulfurdioxygenase ETHE1, which is involved in cysteine catabolism, has an essential but yet undefined function in *Arabidopsis* seed development. In this study, metabolomics and shotgun mass spectrometry based proteomics were implemented to (i) investigate the general role of amino acid metabolism in seeds and the specific impact of changing light conditions on seed filling and (ii) the characterization of ETHE1 function in developing seeds.

Impact of changing light conditions on seed metabolism

Darkness reduces the seed filling capacity

In serveral dicotyledonous species such as *Arabidopsis* embryos become green during development. Seed plastids differ in their structure from those present in leaves, and since carbohydrates are mainly supplied by the mother plant, one can assume that plastidial metabolism is changed as well (Ruuska et al. 2002). Nevertheless, it has been proposed that seed photosynthesis has a great impact on seed metabolism and supports efficient storage product accumulation (Andriotis et al. 2010, Fait et al. 2006, Goffman et al. 2005). Our results confirm

that shading of siliques, which obviously prevents photosynthesis, has a great impact on storage metabolism characterized by reduced lipid and sugar levels, slightly reduced protein content and an accumulation of free amino acids in the seeds. It is very likely that the reduced storage capacity in darkness leads to the lower seed weight we observed, since under light conditions storage product accumulation coincides with increasing seed weight (Baud and Lepiniec 2009). Similar results were obtained in a previous study on *B. napus* seeds grown in darkness (Borisjuk et al. 2013). Since gas diffusion into seeds is low, a major function of seed photosynthesis in fatty acid synthesis might be to supply oxygen required for mitochondrial ATP production (Borisjuk and Rolletschek 2009, Borisjuk et al. 2005). Interestingly, we found an accumulation of TCA cycle intermediates in dark grown seeds, which could be an indication for decreased respiration rates. The role of photosynthesis in seed development can also be studied by using inhibitors (Allorent et al. 2015). Brushing of *Arabidopsis* siliques with DCMU to block photosystem II activity had no effect on storage protein or lipid content. Also, only minor differences in primary metabolites were detected, with the exception of an increase in GABA, proline and galactinol. Therefore, it can be assumed that darkness applied to the growing seeds might cause additional effects on development and metabolism.

Darkness induces an accumulation of nitrogen-rich amino acids

Nitrogen loading of the seeds is achieved via the import of amino acids (asparagine, glutamine, and alanine), which during seed filling are mainly used for storage protein synthesis (Higashi et al. 2006). We observed that the nitrogen-rich amino acids asparagine (Asn) and glutamine (Gln) were most dramatically increased under dark conditions. One reason for this effect could be an induction of Asn synthesis in the absence of light. In leaves asparagine synthetase shows a diurnal pattern of activity and is inhibited by light as well as high levels of reduced carbon (Nozawa et al. 1999). So far no information is available on the relevance of Asn synthesis vs. import in seeds, but it is postulated that Asn has a role in storage product accumulation and might act as signal of the internal nitrogen status (Hernandez-Sebastia et al. 2005). We detected asparagine synthase in our seed proteome data set, but the number of detected spectra was too low for quantification. Another reason for the accumulation of Asn during darkness might be reduced nitrogen uptake by the embryo. It is known that nitrogen uptake of the embryo is realized by amino acid permease 1 (AAP1), and a defect of this transporter results in accumulation of

amino acids in the endosperm at early developmental stages due to limited amino acid shuffling to the embryo (Sanders et al. 2009). Asn is imported from the mother plant and subsequent degradation and metabolism might be important to produce other essential amino acids. It has been shown that defective Asn catabolism causes abnormal seeds and reduced seed weight in *L. japonicus* (Credali et al. 2013). However, we were not able to identify asparaginase in our seed samples.

Darkness induces accumulation of branched chain and aromatic amino acids at early stage of seed development

In vegetative tissues, amino acids serve as alternative substrate for mitochondrial respiration in situations of carbon starvation induced by extended darkness (Araujo et al. 2011). The branched-chain amino acids valine, leucine, and isoleucine as well as lysine are particularly relevant in this functional context since their catabolic pathways produce high amounts of ATP and directly transfer electrons into the respiratory chain. Today only limited information on amino acid metabolism in seeds is available. Recently, results related to amino acid catabolism obtained from leaves were hypothetically transferred to seeds. It is proposed that in response to specific developmental processes and stress situations, amino acid catabolism together with photosynthesis contributes to energy supply in seeds (Galili et al. 2014). The metabolite profiles of dark grown seeds revealed only a mild shortage of carbohydrates in the absence of photosynthesis, which was expected since sugars are delivered from the mother plant and further metabolized (Melkus et al. 2009). Thus, shading of the silique does not necessarily induce carbon starvation conditions, but we could show that it nevertheless induces amino acid accumulation. Interestingly, at least in young seeds (at 5 DAP) levels of branched-chain and aromatic amino acids were clearly elevated. This first experimental evidence highlights that amino acids are presumably not only used for protein biosynthesis but might also contribute to the energy status in developing seeds by alternative respiration.

Physiological function of ETHE1 in seeds

ETHE1 is required for cell organisation and cell wall establishment

Phenotype analysis revealed that *ethe1-1* mutants display a higher sensitivity to changing light conditions, indicated by an enhanced delay of embryogenesis during darkness. However, the final biomass composition of the dormant seeds was not affected by a knock down of ETHE1. In *ethe1-1* seeds, the endosperm is composed of enlarged irregularly shaped cells during cellularization. Our proteomic results confirm that ETHE1 is involved in the establishment of cell walls during development. Proteins related to cell wall degradation such as glucanase and glucosidase were increased in abundance in the mutant compared to the wild type and cell wall proteins were reduced. Overall, we found high amounts of tubulin in *ethe1-1* seeds. In plants, tubulins accumulate during high activity of cell division, cell expansion processes, and also during development (Wasteneys 2002). A high turnover of cell wall elements and defective cell structure establishment might not only affect endosperm cellularization but also nutrient transport to the embryo during development of *ethe1-1* seeds. It has been shown that a mutant blocking transfer cell formation (pea E2748) is defective in normal establishment of embryonic epidermis, which induces enlarged cell wall layers and impaired embryo growth (Borisjuk et al. 2002). Therefore, proper tissue formation is essential for nutrient transport. Our data might indicate that *ethe1-1* mutant embryos suffer from nutrient starvation due to impaired tissue formation. It is proposed that a switch from high hexose to high sucrose level is essential to ensure proper differentiation and cell expansion that elevates storage activity of the embryo (Rolland et al. 2006, Weber et al. 2005). At 5 DAP high concentrations of glucose and fructose and comparatively low levels of sucrose were detected. Enhanced abundance of sucrose synthase activity might induce an increase of relative levels of hexoses observed in *ethe1-1* seeds. However, analyzing wildtype seeds with an equal embryo stage (4 DAP) indicates that the difference observed are rather linked to the general delay of seed development than to a specific ETHE1 function. The switch to high sucrose levels required for proper embryo development is clearly visible at 9 DAP, though sucrose levels are decreased in *ethe1-1* compared to wildtype seeds.

ETHE1 knock down leads to accumulation of amino acids and impaired mitochondrial structure

Amino acid concentrations in *ethe1-1* seeds show a general trend towards accumulation compared to the wildtype. In particular, *ethe1-1* seeds accumulated high amounts of alanine and glycine. The endosperm is the first sink within the seed, and amino acids concentration is high during early stages. Melkus et al. (2009) demonstrated that a decrease of amino acid levels in the endosperm coincides with storage activity in the embryo. At 5 DAP the abundance of photosynthesis related proteins was reduced in *ethe1-1* seeds. Therefore oxygen supply at that time can be expected to be rather low and might facilitate anoxic conditions. Evidence is given that alanine accumulates under hypoxic conditions (van Dongen et al. 2013) and it is proposed that the catalytic activity of alanine aminotransferase is involved in low-oxygen tolerance due to degradation of alanine after hypoxia (Miyashita et al. 2007). Indeed, protein abundance of alanine aminotransferase was slightly, albeit not significant, increased in the mutant (1.16 fold at 5 DAP and 1.19 fold at 9 DAP). Another line of evidence is given by the presence of unusually shaped, enlarged mitochondria in *ethe1-1* seeds. Recently it has been shown that in leaves the establishment of giant mitochondria is induced by darkness, at reduced carbon levels and under hypoxia (Jaipargas et al. 2015). However, for seeds only little information on mitochondrial structure is available. Giant mitochondria were observed in the egg cells during double fertilization, but so far not in other stages of embryogenesis (Kuroiwa and Kuroiwa 1992). ETHE1 knock down leads to a severe defect in mitochondrial structure, underlining its importance in seed energy metabolism.

Material and Methods

Plant growth and seed harvesting

A. thaliana wildtype (ecotype Columbia) and *ethe1-1 (SALK_021573,* Nottingham Arabidopsis Stock Centre, University of Nottingham) plants were grown in a climate chamber under longday conditions (16h light/8h dark, 22°C, 85 µmol s^{-1} m^{-2} light intensity and 65% humidity). Flowers were labeled at the day of pollination. Subsequently siliques were harvested from 1 to 9 DAP. For dark treatment, siliques were shaded with aluminium foil 24 hours after flower tagging while the rest of the plant and control siliques were grown under normal light conditions.

Phenotypic analysis of seed samples

Whole-mount preparations were used for microscopic analysis of *A. thaliana* wildtype and *ethe1-1* embryo development grown under light and dark conditions. Seeds from at least 5 siliques each were cleared by incubation with Hoyer´s solution [15 mL distilled water, 3.75 g gum Arabic, 2.5 mL glycerine, 50 g chloral hydrate] overnight. Embryo development was analysed by using a Normarski optics and light microscopy. Observed embryo stages were counted. For analyzing seed tissue differentiation histological sections were prepared form 7 DAP *ethe1-1* and wildtype seeds grown under light and dark conditions. Seed samples were fixed with 1.25% formaldehyde and 0.1 M phosphate buffer overnight followed by dehydration of the samples by a series of ethanol aqueous solutions (30-99%). After dehydration samples were infiltrated with Historesin (Leica). Sections of 3.5 µm were prepared using a microtome (HYRAX M55, Zeiss) and Sec55 low profile blades (MICROM). Sections were stained with 0.5% Toluidine Blue O in 200mM phosphate buffer. Photographs of cleared seeds and sections were taken with a Zeiss microscope (Axioskop2) and AxioCam MRc5 camera. Seed size and weight were measured for dormant seeds of *ethe1-1* and wildtype grown under light and dark conditions for 100 seeds each. Significant differences were calculated using a Student´s t-test (p-value $\leq$0.05). For electron microscopy of mitochondria seeds of wildtype and *ethe1-1* (5 DAP) were fixed and embedded in accordance to Tschiersch et al. (2011). Digital records of seed mitochondria were made on a Zeiss 902 electron microscope at 80 kV.

Germination rates

Arabidopsis seeds of wildtype (ecotype Columbia, Col0) and *ethe1-1* grown under light and dark conditions were surface sterilized with 6% sodium hypochloride and 100% ethanol followed by five washing steps with sterilized water . For vernalization seeds were incubated for 2 days at 4°C in the dark. Approximately 20 seeds were sown per plate (3 replicates per sample) on MS-medium [60mM sucrose, 1% Agar, 0.5% MS-medium (Duchefa), pH 5.7-5.8 with KOH] and MS medium without a carbon source and incubated for another 2 days at 4°C in the dark. Afterwards the plates were placed to a growth chamber (24°C, 16 h light/8 h dark). After 72h germinated seeds were counted. A seed is considered to be germinated when the radicle raptures the endosperm and the testa.

Silique culture

Siliques were harvested from wildtype and *ethe1-1* plants at 2 DAP and cultured in a medium [1.5% MS and 2mM MES, pH 5.6] (control). For amino acid treatments 0.1mM, 0.5mM and 1 mM cysteine and leucine were added to the silique culture. A concentration of 10 mM resulted in seed abortion for both genotypes (data not shown). Siliques were cultured for 3 days in a climate chamber under longday conditions (seed stage equivalent to 5 DAP). Subsequently, seeds were harvested and embryo stages investigated by seed clearing and light microscopy as described above.

Metabolite profiling

Metabolites of *ethe1-1* and wildtype seeds (4, 5 and 9 DAP) grown under light and dark conditions were extracted and subsequently analysed by GC-TOF MS as described by Lisec et al. (2006). Chromatograms and mass spectra were evaluated by using TagFinder 4.0 software (Luedemann et al., 2008) and Xcalibur 2.1 software (Thermo Fisher Scientific, Waltham, USA). Metabolites were identified in comparison to database entries of authentic standards (Kopka et al., 2005; Schauer et al., 2005). Peak areas of the mass (m/z) fragments were normalized to the internal standard (ribitol) and fresh weight of the seed samples. Identification and annotation of detected peaks followed recent recommendations for reporting metabolite data (Fernie et al.,

2011). Metabolites (starch, total proteins, total lipids and amino acids) of dormant wildtype and *ethe1-1* seeds grown under light and dark conditions were extracted and analyzed as described in Schwender et al. (2015).

Sample preparation for mass spectrometry

Proteins of wildtype and *ethe1-1* (5DAP, 9 DAP, light, dark) were extracted from 30 mg of pulverized seeds (pooled sample) by adding 150 µL extraction buffer [4% SDS (w/v), 125mM TRIS, 20% glycerol (v/v)] and incubating at 60°C for 5 min. Another 150 µL of ddH2O were added to the sample followed by centrifugation at 18000xg for 10 min. The protein concentration of the supernatant was measured by using the Pierce™ BCA Protein Assay Kit (Thermo Fisher Scientific, Dreieich, Germany). To each sample (total protein 50 µg) 2-mercapto ethanol with a trace of bromophenol blue was added to a final concentration of 5%. SDS gel electrophoresis using a 1D glycine gel (stacking gel 4% acrylamide, separating gel 14% acrylamide) was performed according to Leammli (1970). The gel run was stopped before proteins entered the separation gel. The gel was fixed with 10% (v/v) acetate in 40% (v/v) methanol for 45 min and stained with Coomassie blue CBB G-250 (Merck, Darmstadt, Germany) for 30 min as described by Neuhoff et al. (1985 and 1990). Gel bands were cut using a scalpell and diced into 1.0-1.5 mm cubes. Carbamidomethylated followed by tryptic digestion and extraction of proteins was performed according to Klodmann et al. (2010). Resulting peptides were resolved in 20 µL of 2% [v/v] ACN, 0.1% [v/v] formic acid (FA) prior to MS analysis.

Shot gun mass spectrometry and relative protein quantification

Shot gun mass spectrometry was performed by using a Q-Exactive (Thermo Fisher Scientific, Dreieich, Germany) mass spectrometer coupled to an Ultimate 3000 (Thermo Fisher Scientific, Dreieich, Germany) UPLC. Three times (technical replicates) four microliter of peptide solution per seed sample were injected into a 2 cm, C18, 5 µm, 100 Å reverse phase trapping column (Acclaim PepMap100, Thermo Fisher Scientific, Dreieich, Germany). Peptide separation was done on a 50 cm, C18, 3 µm, 100 Å reverse phase analytical column (Acclaim PepMap100, Thermo Fisher Scientific, Dreieich, Germany). Peptides were eluted by using a non-linear 2% [v/v] to 34% [v/v] acetonitrile gradient in 0.1% [v/v] formic acid of 60 minutes. For MS analysis

a spray voltage of 2.2 kV, capillary temperature to 275°C and S-lens RF level to 50% was tuned. For full MS scans, the number of microscans was adjusted to 1, resolution to 70,000, AGC target to 1e6, maximum injection time to 400 ms, number of scan ranges to 1 and scan range to 400 to 1600m/z. For dd-MS2, the number of microscans was adjusted to 1, resolution to 17,500, AGC target to 1e5, maximum injection time to 120 ms, Loop count to 10, MSX count to 1, isolation window to 3.0 m/z, fixed first mass to 100.0 m/z and NCE to 27.0. Data dependent (dd) settings were adjusted to underfill ratio of 0.5%, intensity threshold to 2.0e3; apex trigger to 10 to 60 s, charge exclusion to unassigned 1, 5, 5 – 8, >8, peptide match to preferred, exclude isotopes to on and dynamic exclusion to 45.0 s. Protein identification and was done by using Proteome Discoverer (Thermo Fisher Scientific). MS/MS data were queried against the Tair10plus database using the MASCOT search engine (peptide confidence high, minimum peptide count 1). For quality control of MS shotgun run a principle component analysis (PCA) using Perseus software was performed. Protein abundance was quantified by spectral counting and significant differences between samples were calculated using a Student´s t-test (p-value ≤ 0.05).

Figures

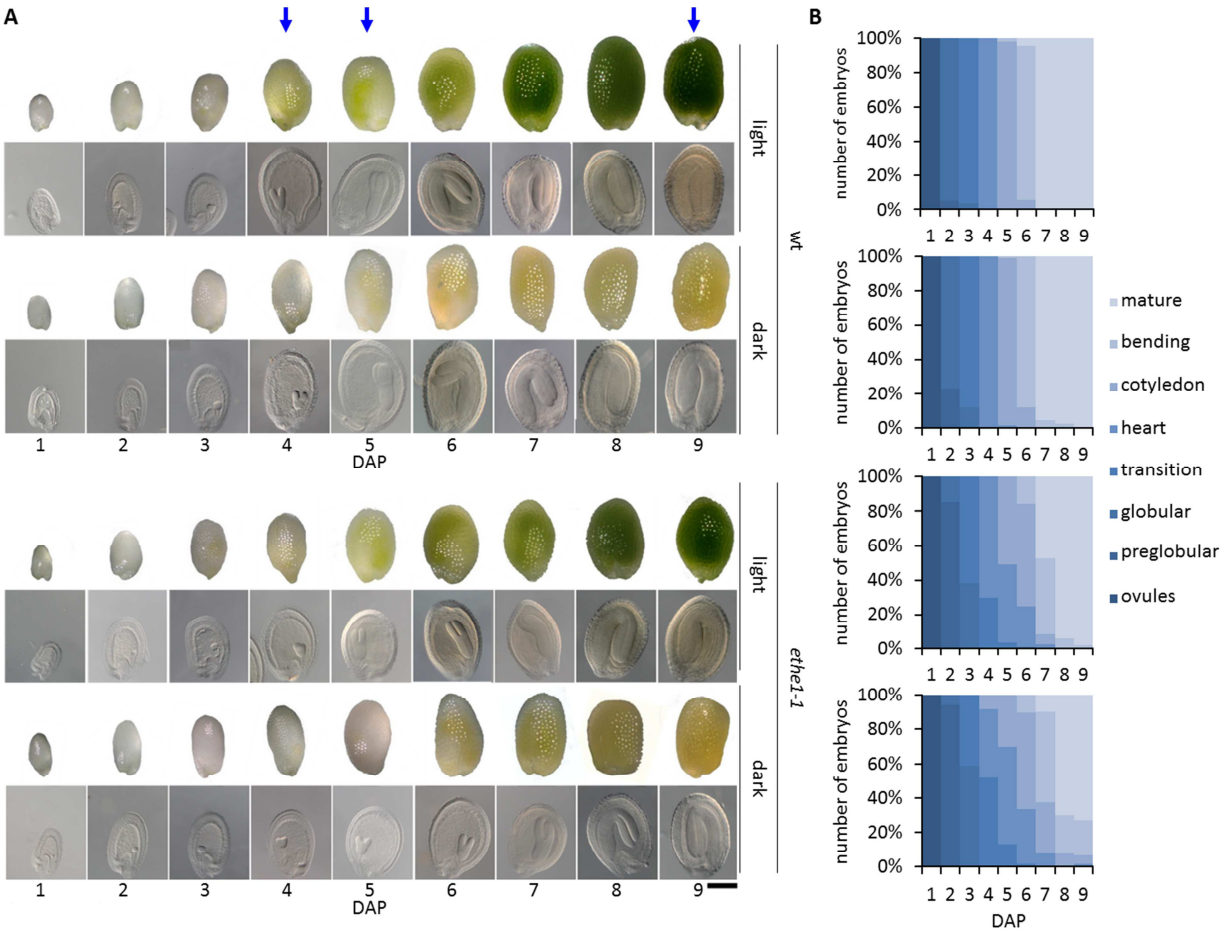


Figure 1: Seed development of *A. thaliana* wildtype and *ethe1-1* seeds under light and dark conditions. A: Representative seeds and embryos of wildtype (wt) and *ethe1-1* from 1 to 9 days after pollination (DAP) grown under light and dark conditions, bars = 100 µm. **B:** Progression of embryo development in wt and *ethe1-1* seeds. A total of 180 siliques and 7781 seeds between 1 and 9 DAP were analyzed of wildtype and *ethe1-1* grown under light and dark conditions. Blue arrows indicates stages investigated by metabolomics and proteomics.

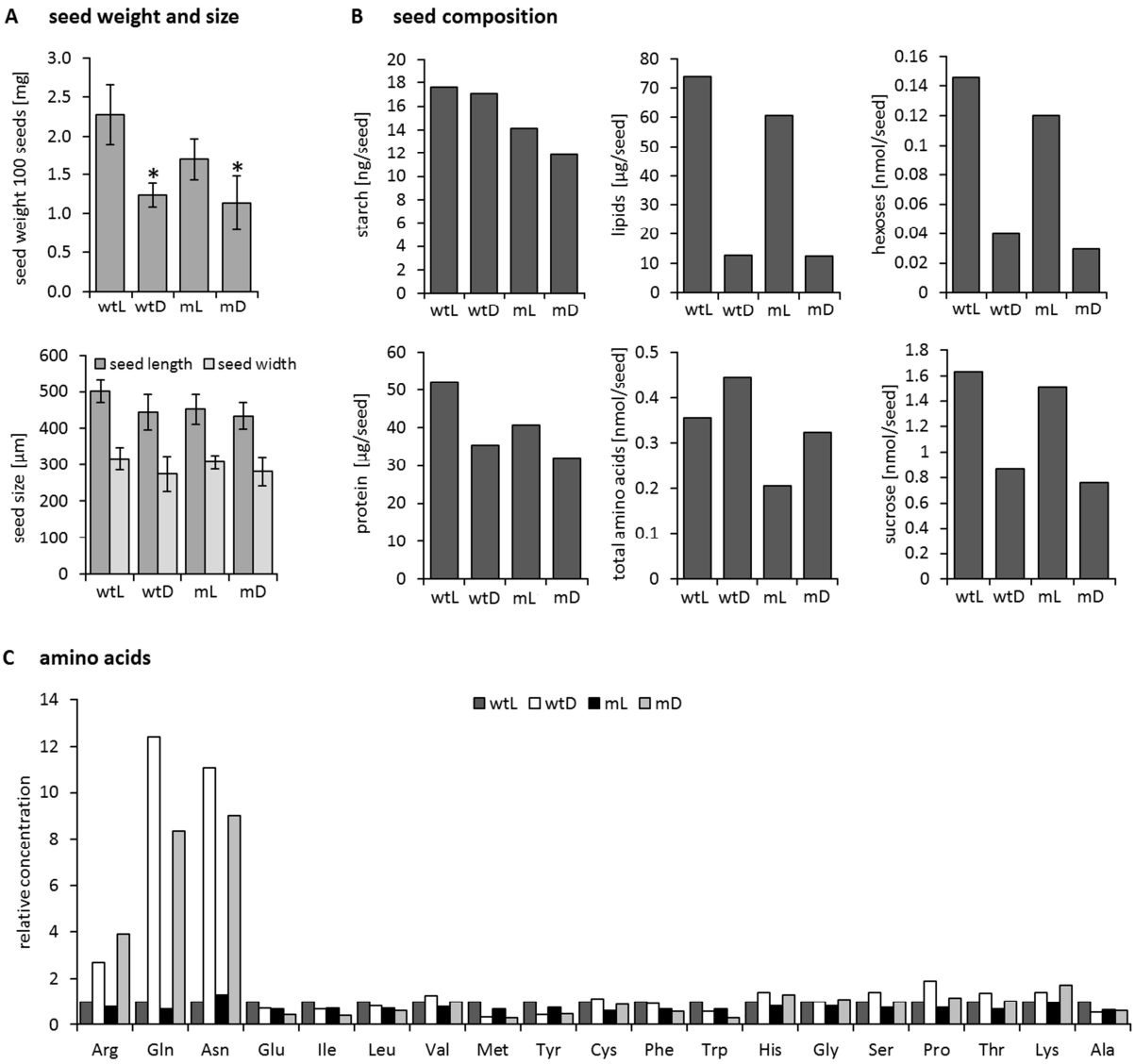


Figure 2: Seed weight, seed size and biomass composition of dormant seeds. A: Seed size and weight of 100 dormant seeds from wildtype and *ethe1-1* plants grown under light and dark conditions. Asterisks indicate significant changes based on a Student´s t-test (p-value ≤0.05) **B:** Storage compounds and amino acids were extracted from dormant seeds from wildtype and *ethe1-1* grown under light and dark conditions and quantified. **C:** amino acid profiles of dormant seeds from wildtype and *ethe1-1*. **wtL** wildtype seeds grown under light conditions, **wtD** wildtype seeds grown under conditions dark conditions, **mL** *ethe1-1* seeds grown under light conditions, **mD** *ethe1-1* seeds grown under dark conditions.

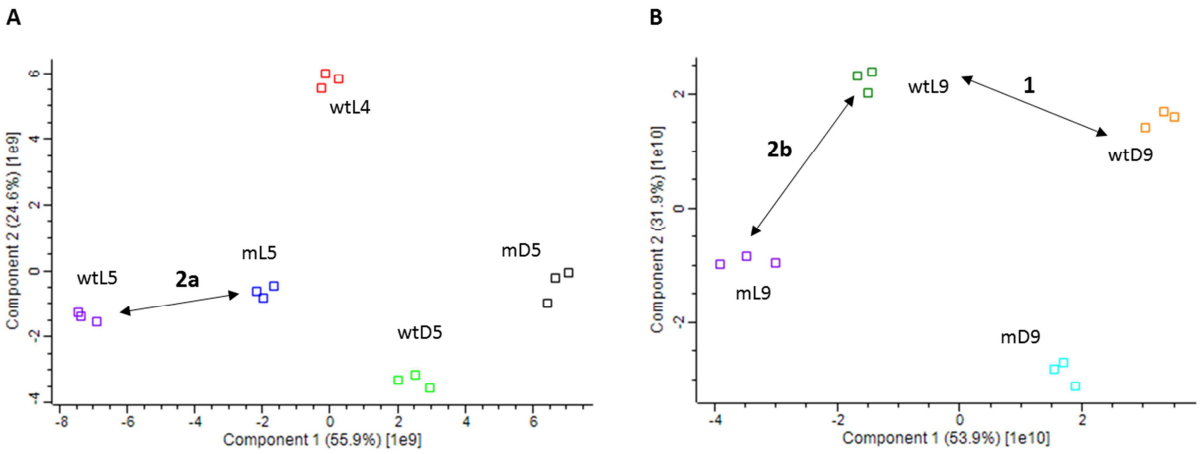


Figure 3: Priciple component analysis and overview of proteomic comparisons. For quality of shotgun mass spectrometry data principle component analysis (PCA) was performed by using Perseus software. Proteomic comparisons are indicated by numbers (**1, 2a, 2b**) in the PCA plots of 4 and 5 DAP seeds (**A**) and 9 DAP seeds (**B**).

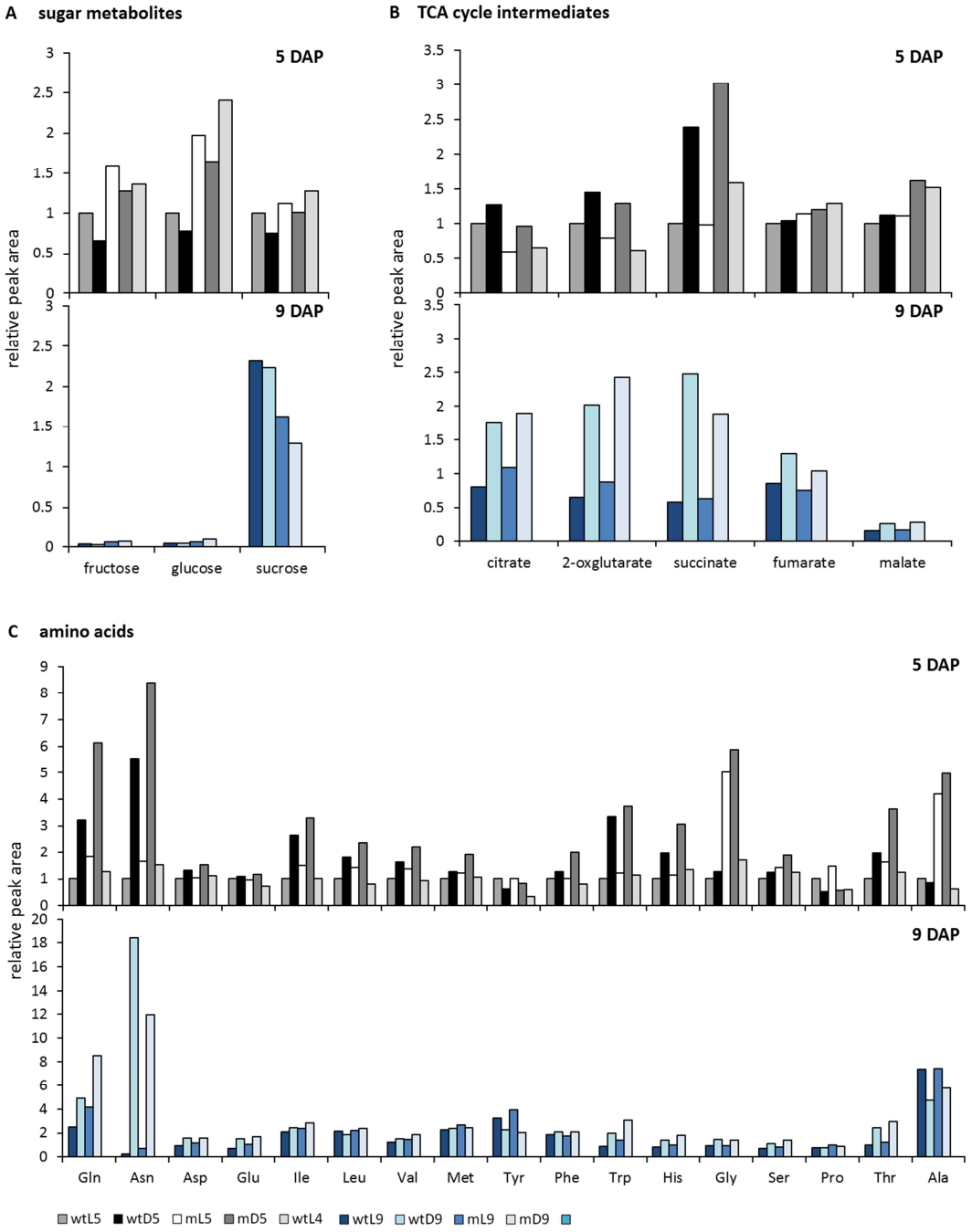


Figure 4: Metabolite profiles. Metabolites (sugar metabolites, TCA cycle intermediates, amino acids) of *ethe1-1* and wildtype seeds (4, 5 and 9 DAP) grown under light and dark conditions were extracted and subsequently analysed by GC-TOF MS. Peak areas of the mass (m/z) fragments were normalized to the internal standard (ribitol) and fresh weight of the seed samples. **wtL** wildtype seeds grown under light conditions, **wtD** wildtype seeds grown under conditions dark conditions, **mL** *ethe1-1* seeds grown under light conditions, **mD** *ethe1-1* seeds grown under dark conditions.

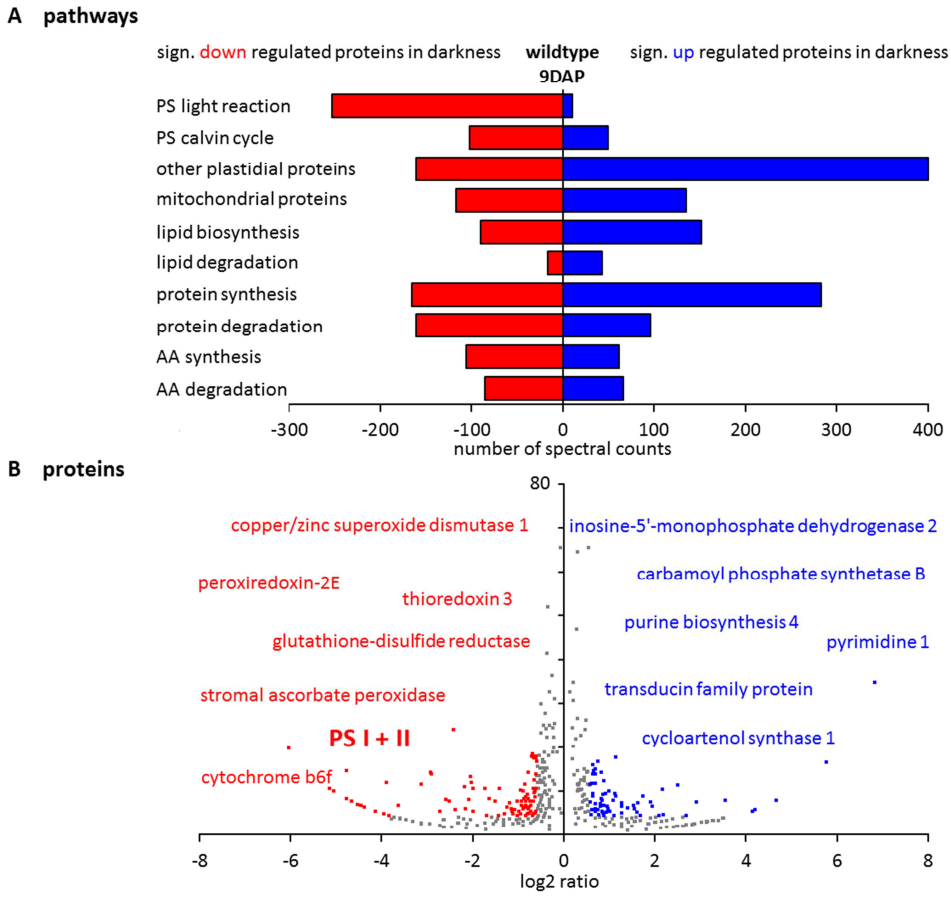


Figure 5: Regulated pathways and proteins in dark grown seeds. A: Distribution of regulated proteins showing a potentially abundance in dark grown seeds across functional categories, **B:** Signigicantly regulated proteins in dark grown seeds. **red**= decreased abundance with a ratio >1.5 (log2 = 0.58) and >5 spectra; **blue** = increased abundance with a ratio >1.5 and >5spectra**, grey** = signigicantly abundant proteins with a ratio <1.5 and <5 spectra. Sigifincance levels were calcultated with a Student´s t-test (p-value $\leq$ 0.05). Differences between the samples are given in log2 ratios.

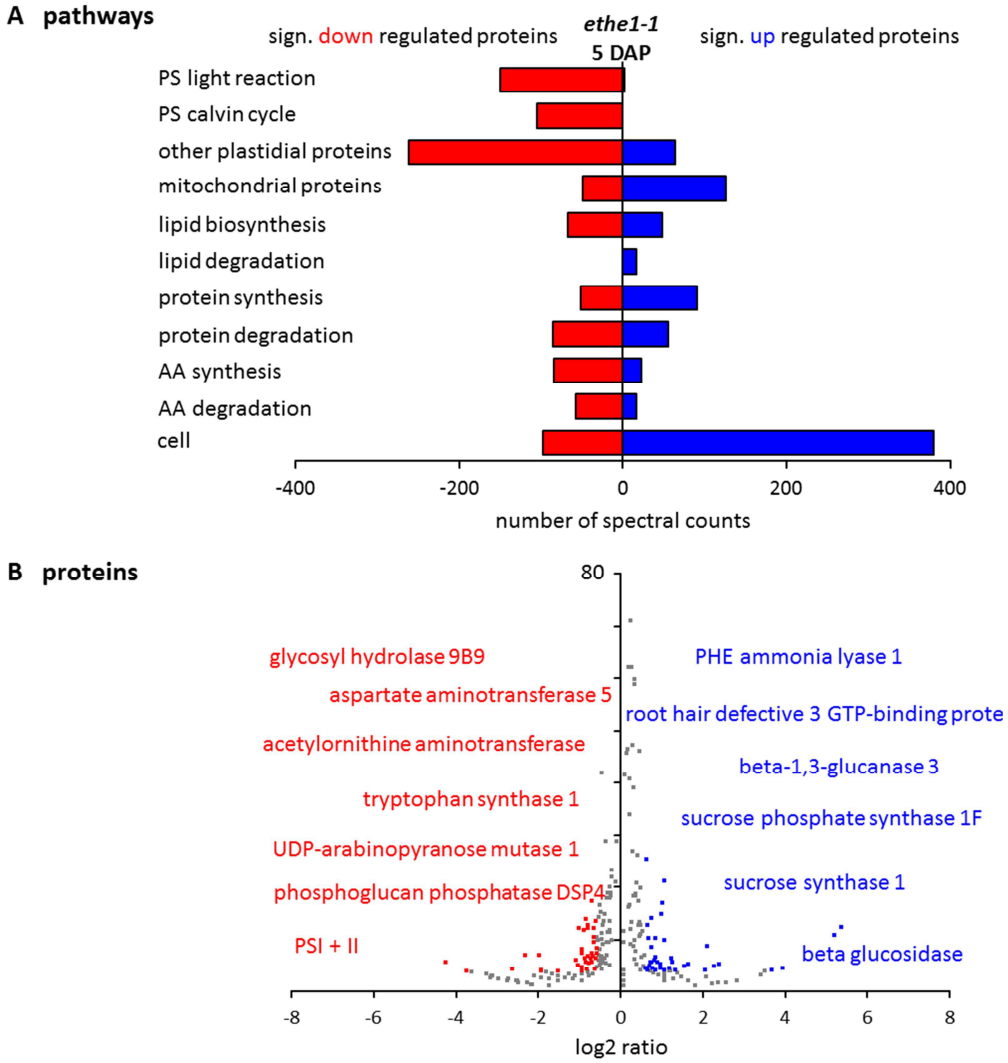


Figure 6: Regulated pathways and proteins in *ethe1-1* seeds at 5 DAP compared to wildtype. A: Distribution of regulated proteins showing a potentially abundance in *ethe1-1* seeds at 5 DAP across functional categories, **B:** Signigicantly regulated proteins in *ethe1-1* seeds. **red**= decreased abundance with a ratio >1.5 (log2 = 0.58) and >5 spectra; **blue** = increased abundance with a ratio >1.5 and >5spectra**, grey** = signigicantly abundant proteins with a ratio <1.5 and <5 spectra. Sigifincance levels were calcultated with a Student´s t-test (p-value ≤ 0.05). Differences between the samples are given in log2 ratios.

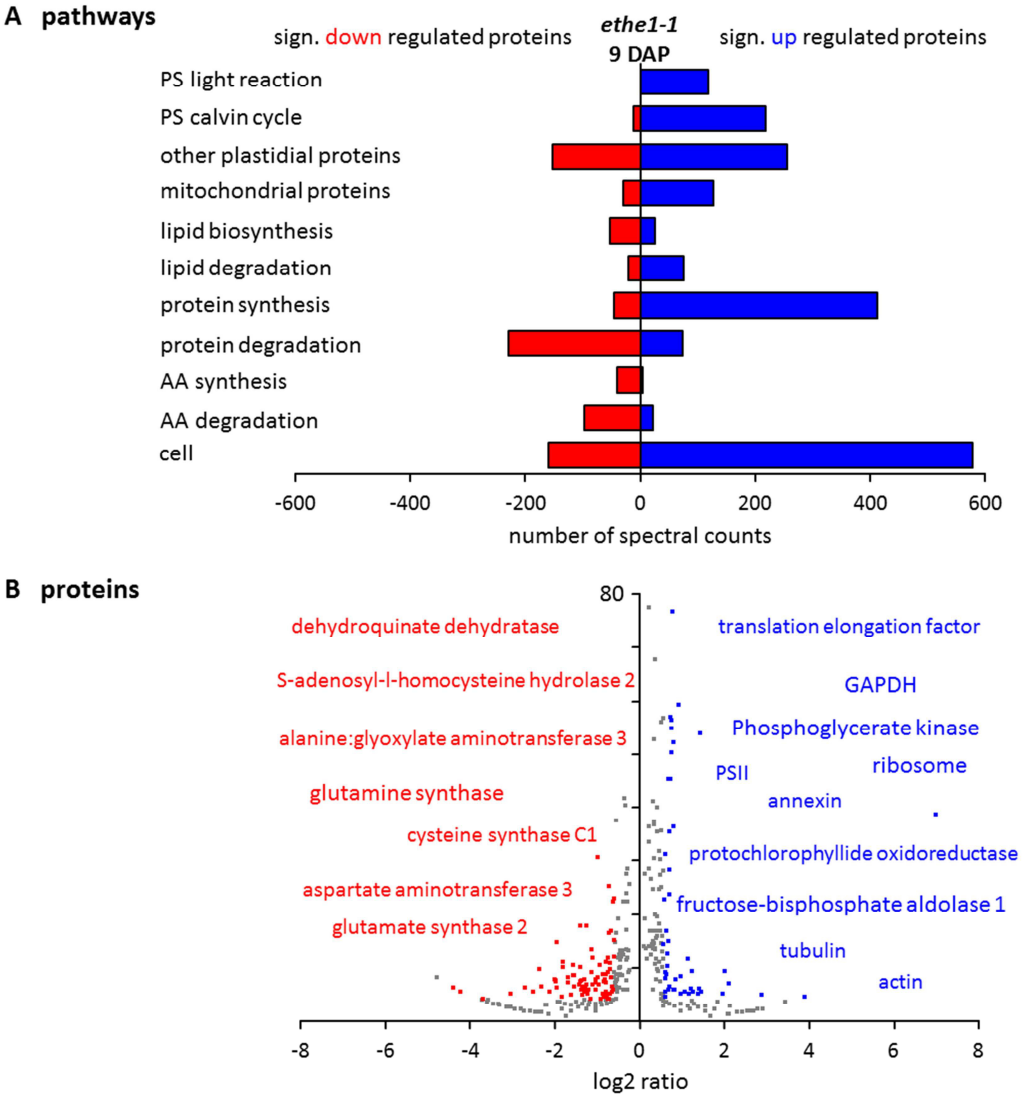


Figure 7: Regulated pathways and proteins in *ethe1-1* seeds at 9 DAP compared to wildtype. A: Distribution of regulated proteins showing a potentially abundance in *ethe1-1* seeds at 9 DAP across functional categories, **B:** Signigicantly regulated proteins in *ethe1-1* seeds. **red**= decreased abundance with a ratio >1.5 (log2 = 0.58) and >5 spectra; **blue** = increased abundance with a ratio >1.5 and >5spectra**, grey** = signigicantly abundant proteins with a ratio <1.5 and <5 spectra. Sigifincance levels were calcultated with a Student´s t-test (p-value ≤ 0.05). Differences between the samples are given in log2 ratios.

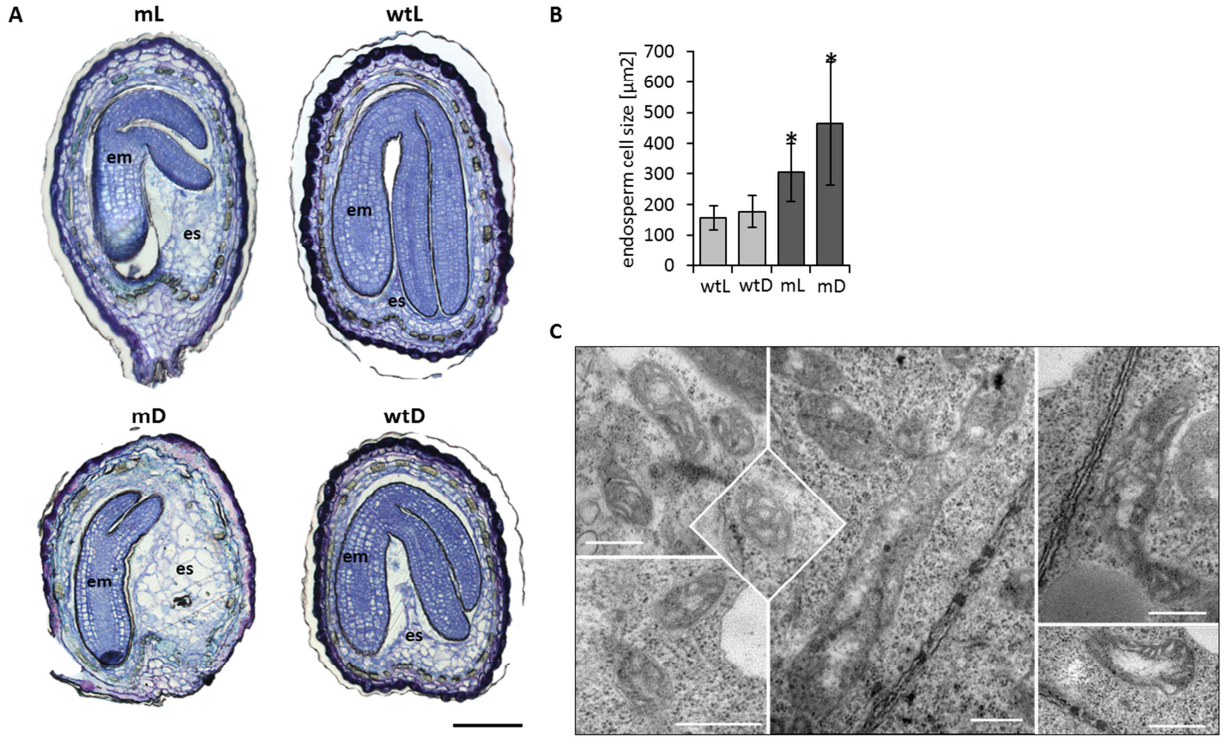
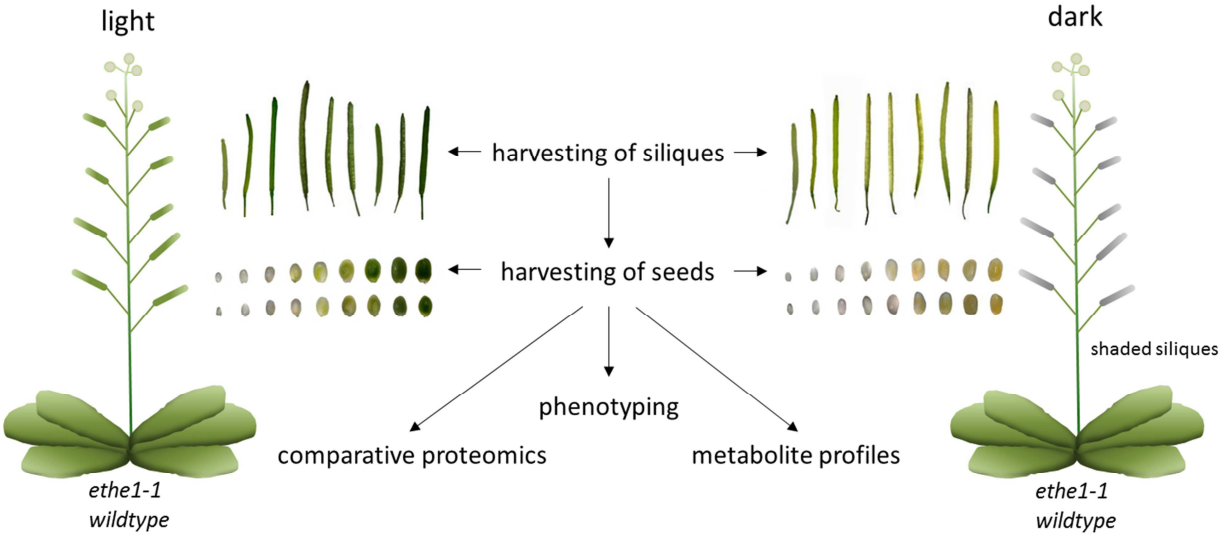
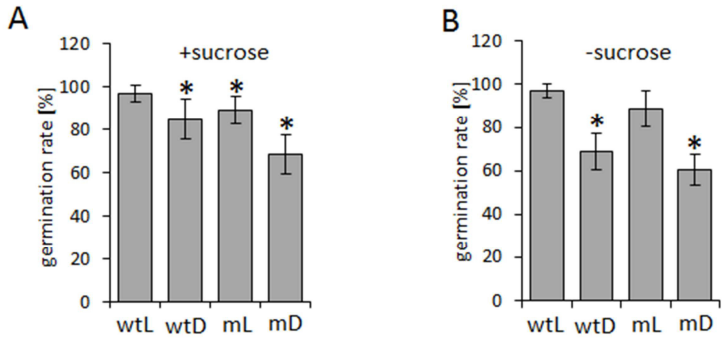


Figure 8: Endosperm cellularization and mitochondria structure. A: Seed sections (7DAP) of wildtype and *ethe1-1* seeds grown under light and dark conditions. Sections were stained with Toluidine Blue O to visualize endospserm cellularization, bars = 100 µm **B:** Endosperm cell size of ca. 50 cells is given as area [μm^2] obtained from microscopy images using AxioVision software (Version 4.8.1). Asterisks indicate significant changes to wildtype light (wtL) based on a Student´s t-test (p-value ≤0.05) **C:** Electron microsopic analysis of mitochondrial structure from *ethe1-1* seeds compared to wildtype bars = 500 nm. **wtL** wildtype seeds grown under light conditions, **wtD** wildtype seeds grown under conditions dark conditions, **mL** *ethe1-1* seeds grown under light conditions, **mD** *ethe1-1* seeds grown under dark conditions.



Supplementary figure 1: Experimental design. *A. thaliana* wildtype (ecotype Columbia) and *ethe1-1* plants were grown in a climate chamber under longday conditions (16h light/8h dark, 22°C, 85 µmol s^{-1} m^{-2} light intensity and 65% humidity). Flowers were labeled at the day of pollination. Subsequently siliques were harvested from 1 to 9 DAP. For dark treatment, siliques were shaded with aluminium foil 24 hours after flower tagging while the rest of the plant and control siliques were grown under normal light conditions. Seeds were investigated by phenotyping, comparative proteomics and measuring metabolite profiles.



Supplementary figure 2: Germination rates. Sterilized seeds of wildtype and *ethe1-1* grown under light and dark conditions were were sown per plate (3 replicates per sample) on MS-medium [60mM sucrose, 1% Agar, 0.5% MS-medium (Duchefa), pH 5.7-5.8 with KOH] and MS medium without sucrose and incubated for another 2 days at 4°C in the dark. Afterwards the plates were placed to a growth chamber (24°C, 16 h light/8 h dark). After 72h germinated seeds were counted. A seed is considered to be germinated when the radicle raptures the endosperm and the testa.

References

Allorent G., Osorio S, Vu J.L., Falconet D., Jouhet J., Kuntz M., Fernie A.R., Lerbs-Mache S., Macherel D., Courtois F., Finazzi G. (2015) Adjustments of embryonic photosynthetic activity modulate seed fitness in Arabidopsis thaliana. New Phytologist 205(2): 707-719.

Andriotis V.M., Kruger N.J., Pike M.J., Smith A.M (2010) Plastidial glycolysis in developing Arabidopsis embryos. New Phytologist 185(3): 649-662.

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

Angelovici R., Fait A., Zhu X., Szymanski J., Feldmesser E., Fernie A.R., Galili G. (2009) Deciphering transcriptional and metabolic networks associated with lysine metabolism during Arabidopsis seed development. Plant Physiology 151(4): 2058-2072.

Araujo W.L., Ishizaki K., Nunes-Nesi A., Larson T.R., Tohge T., Krahnert I., Bauwe H., Hagemann M., Fernie A.R. (2010) Photorespiration: players, partners and origin. Trends in Plant Science 15: 330-336.

Araujo W.L., Tohge T., Ishizaki K., Leaver C.J., Fernie A.R. (2011) Protein degradation-an alternative respiratory substrate for stressed plants. Trends in Plant Science 16:489-498.

Baud S., Boutin J-P., Miquel M., Lepiniec L., Rochat C. (2002) An integrated overview of seed development in Arabidopsis thaliana ecotype Ws. Plant Physiology Biochemistry 40: 151-160.

Baud S., Lepiniec L. (2009) Regulation of de novo fatty acid synthesis in maturing oilseeds of Arabidopsis. Plant Physiology and Biochemistry 47(6): 448-455.

Berger F. (2003) Endosperm: The crossroad of seed metabolism. Current Opinion in Plant Biology 6 (1): 42-50.

Borisjuk L., Neuberger T., Schwender J., Heinzel N., Sunderhaus S., Fuchs J., Hay J.O., Tschiersch H., Braun H.P., Denolf P., Lambert B., Jakob P.M., Rolletschek H. (2013) Seed architecture shapes embryo metabolism in oilseed rape. Plant Cell 25: 1625-1640.

Borisjuk L., Nguyen T.H., Neuberger T., Rutten T., Tschiersch H., Claus B., Feussner I., Webb A.G., Jakob P., Weber H., Wobus U., Rolletschek H. (2005) Gradients of lipid storage, photosynthesis and plastid differentiation in developing soybean seeds. New Phytologist 167(3): 761-776.

Borisjuk L., Rolletschek H. (2009) The oxygen status of the developing seed. New Phytologist 182(1): 17-30.

Borisjuk L., Rolletschek H., Radchuk R., Weschke W., Wobus U., Weber H. (2004) Seed development and differentiation: a role for metabolic regulation. Plant Biology 6(4): 375-386.

Borisjuk L., Wang T.L., Rolletschek H., Wobus U., Weber H.A (2002) Pea seed mutant affected in the differentiation of the embryonic epidermis is impaired in embryo growth and seed maturation. Development 129(7):1595-1607.

Brown R. C., Lemmon B. E., Nguyen H. (2003) Events during the first four rounds of mitosis establish three developmental domains in the syncytial endosperm of Arabidopsis thaliana. Protoplasma 222(3-4): 167-174.

Brown R. C., Lemmon B. E., Nguyen H., Olsen O. A. (1999) Development of endosperm in Arabidopsis thaliana. Plant Reproduction 12(1): 32-42.

Credali A., García-Calderón M., Dam S., Perry J., Díaz-Quintana A., Parniske M., Wang T.L., Stougaard J., Vega J.M., Márquez A.J. (2013) The K+-dependent asparaginase, NSE1, is crucial for plant growth and seed production in Lotus japonicus. Plant Cell Physiology 54(1): 107-118.

Fait A., Angelovici R., Less H., Ohad I., Urbanczyk-Wochniak E., Fernie. A.R., Galili G. (2006) Arabidopsis seed development and germination is associated with temporally distinct metabolic switches. Plant Physiology (142): 839-854.

Fernie A.R., Aharoni A., Willmitzer L., Stitt M., Tohge T., Kopka J., Carroll A.J., Saito K., Fraser P.D., DeLuca V. (2011) Recommendations for reporting metabolite data. The Plant Cell 23(7): 2477-2482.

Foyer CH, Noctor G. (2003) Redox sensing and signalling associated with reactive oxygen in chloroplasts, peroxisomes and mitochondria. Physiol. Plant. 119:355-64

Galili G., Avin-Wittenberg T., Angelovici R., Fernie A.R. (2014) The role of photosynthesis and amino acid metabolism in the energy status during seed development. Frontiers in Plant Science 5: 447

Gallardo K., Thompson R.D., Burstin J. (2008) Reserve accumulation in legume seeds. Comptes Rendus Biologies 331: 755-762.

Goffman F.D., Alonso A.P., Schwender J., Shachar-Hill Y., Ohlrogge J.B. (2005) Light enables a very high efficiency of carbon storage in developing embryos of rapeseed. Plant Physiology 138(4): 2269-2279.

Golombek S, Rolletschek H, Wobus U, Weber H. (2001) Control of storage protein accumulation during legume seed development. J. Plant Physiol. 158:457-464

Gu L., Jones A.D., Last R.L. (2010) Broad connections in the Arabidopsis seed metabolic network revealed by metabolite profiling of an amino acid catabolism mutant. Plant Journal 61: 579-590.

Hehenberger E., Kradolfer D., Köhler C. (2012) Endosperm cellularization defines an important developmental transition for embryo development. Plant Cell Physiology 53(1): 16-27.

Hernández-Sebastià C., Marsolais F., Saravitz C., Israel D., Dewey R.E., Huber S.C.(2005) Free amino acid profiles suggest a possible role for asparagine in the control of storage-product accumulation in developing seeds of low- and high-protein soybean lines. Journal of Experimental Botany 56(417):1951-1963.

Higashi Y., Hirai M.Y., Fujiwara T, Naito S, Noji M, Saito K. (2006) Proteomic and transcriptomic analysis of Arabidopsis seeds: molecular evidence for successive processing of seed proteins and its implication in the stress response to sulfur nutrition. Plant Journal 48(4): 557-71.

Hildebrandt T.M., Nesi A.N., Araújo W.L., Braun H.P. (2015) Amino acid catabolism in plants. Molecular Plant pii: S1674-2052(15)00366-4.

Holdorf M.M., Owen H.A., Lieber S.R., Yuan L., Adams N., Dabney-Smith C., Makaroff C.A. (2012) Arabidopsis ETHE1 encodes a sulfur dioxygenase that is essential for embryo and endosperm development. Plant Physiology 160(1): 226-236.

Jaipargas E.A., Barton K.A., Mathur N., Mathur J. (2015) Mitochondrial pleomorphy in plant cells is driven by contiguous ER dynamics. Front Plant Science 6:783.

Klodmann J., Sunderhaus S., Nimtz M., Jansch L., Braun H.P. (2010) Internal architecture of mitochondrial complex I from Arabidopsis thaliana. Plant Cell, 22 (3): 797-810

Kopka J., Schauer N., Krueger S., Birkemeyer C., Usadel B., Bergmüller E., Dörmann P., Weckwerth W., Gibon Y., Stitt M., Willmitzer L., Fernie A.R., Steinhauser D. (2005) GMD@CSB.DB: the Golm Metabolome Database. Bioinformatics (Oxford, England) 21(8): 1635-1638.

Krüßel L., Junemann J., Wirtz M., Birke H., Thornton J.D., Browning L.W., Poschet G., Hell R., Balk J., Braun H.P., Hildebrandt T.M. (2014) The mitochondrial sulfur dioxygenase ETHYLMALONIC ENCEPHALOPATHY PROTEIN1 is required for amino acid catabolism during carbohydrate starvation and embryo development in Arabidopsis. Plant Physiology 165(1): 92-104.

Kuroiwa H, Kuroiwa T.(1992) Giant mitochondria in the mature egg cell of Pelargonium zonale. Protoplasma 168: 184-188.

Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227(5259): 680-685.

Le B. H., Cheng C., Bui A. Q., Wagmaister J. A., Henry K. F., Pelletier J., Kwong L., Belmonte M., Kirkbride R., Horvath S., Drews G. N., FischerR. L., Okamuro K. K., Harada J. J., Goldberg R. B. (2010) Global analysis of gene activity during Arabidopsis seed development and identification of seed-specific transcription factors. Proceedings of the National Academy of Sciences of the USA 107(18): 8063-8070.

Le B.H., Cheng C., Bui A.Q., Wagmaister J.A., Henry K.F., Pelletier J., Kwong L., Belmonte M., Kirkbride R., Horvath S., Drews G.N., Fischer R.L., Okamuro J.K., Harada J.J., Goldberg R.B. (2010) Global analysis of gene activity during Arabidopsis seed development and identification of seed-specific transcription factors. Proc Natl Acad Sci USA 107(18):8063-8070.

Lisec J., Schauer N., Kopka J., Willmitzer L., Fernie A.R. (2006) Gas chromatography mass spectrometry-based metabolite profiling in plants. Nature protocols 1(1): 387-396.

Luedemann A., Strassburg K., Erban A., Kopka J. (2008) TagFinder for the quantitative analysis of gas chromatography--mass spectrometry (GC-MS)-based metabolite profiling experiments. Bioinformatics (Oxford, England) 24(5): 732-737.

Melkus G., Rolletschek H., Radchuk R., Fuchs J., Rutten T., Wobus U., Altmann T., Jakob P., Borisjuk L. (2009) The metabolic role of the legume endosperm: A noninvasive imaging study. Plant Physiology 151: 1139-1154.

Miyashita Y., Dolferus R., Ismond K.P., Good A.G. (2007) Alanine aminotransferase catalyses the breakdown of alanine after hypoxia in Arabidopsis thaliana. Plant Journal 49(6):1108-1121.

Neuhoff V., Stamm R., Eibl H. (1985) Clear background and highly sensitive protein staining with Coomassie blue dyes in polyacrylamide gels: a systematic analysis. Electrophoresis, 6 (9): 427-448

Neuhoff V., Stamm R., Pardowitz I., Arold N., Ehrhardt W., Taube D. (1990) Essential problems in quantification of proteins following colloidal staining with Coomassie brilliant blue dyes in polyacrylamide gels, and their solution. Electrophoresis, 11 (2): 101-117.

Novack M.K., Unguru A., Bjerkan K.N., Grini P.E., Schnittger A. (2010) Reproductive cross-talk: seed development in flowering plants. Biochemical Society Transactions, 38: 604-612.

Nozawa A., Ito M., Hayashi H., Watanabe A. (1999) Dark-induced expression of genes for asparagine synthetase and cytosolic glutamine synthetase in radish cotyledons is dependent on the growth stage. Plant Cell Physiol 40(9): 942-948.

Olsen O.-A. (2004) Nuclear Endosperm Development in Cereals and Arabidopsis thaliana. The Plant Cell 16: 214-222.

Rolland F., Baena-Gonzalez E., Jen S. (2006) SUGAR SENSING AND SIGNALING IN PLANTS: Conserved and Novel Mechanisms. Annual Review of Plant Biology 57(1): 675-709.

Rolletschek H., Radchuk R., Klukas C., Schreiber F., Wobus U., Borisjuk L.(2005) Evidence of a key role for photosynthetic oxygen release in oil storage in developing soybean seeds. New Phytologist 167(3): 777-786.

Ruuska S.A., Girke T., Benning C., Ohlrogge J.B. (2002) Contrapuntal networks of gene expression during Arabidopsis seed filling. Plant Cell 14(6): 1191-1206.

Sanders A., Collier R., Trethewy A., Gould G., Sieker R., Tegeder M. (2009) AAP1 regulates import of amino acids into developing Arabidopsis embryos. Plant Journal 59(4): 540-52.

Tschiersch H., Borisjuk L., Rutten T., Rolletschek H. (2011) Gradients of seed photosynthesis and its role for oxygen balancing. BioSystems 103:302-308.

Schwender J., Hebbelmann I., Heinzel N., Hildebrandt T., Rogers A., Naik D., Klapperstück M., Braun H.P., Schreiber F., Denolf P., Borisjuk L., Rolletschek H. (2015) Quantitative Multilevel Analysis of Central Metabolism in Developing Oilseeds of Oilseed Rape during in Vitro Culture. Plant Physiology 168(3): 828-848.

Van Dongen J.T., Schurr U., Pfister M., Geigenberger P. (2003) Phloem metabolism and function have to cope with low internal oxygen. Plant Physiology 131: 1529-1543.

Wasteneys G.O. (2002) Microtubule organization in the green kingdom: chaos or self-order? Journal of Cell Science 115: 1345-1354.

Weber H., Borisjuk L., Wobus U. (2005) Molecular physiology of legume seed development. Annual Review of Plant Biology 56: 253-279.

Weigelt K., Küster H., Radchuk R., Müller M., Weichert H., Fait A., Fernie A.R., Saalbach I.,

Weber H. (2008) Increasing amino acid supply in pea embryos reveals specific interactions of N and C metabolism, and highlights the importance of mitochondrial metabolism. Plant Journal 55(6): 909-1026.

Curriculum Vitae

Name:	Christin Lorenz (née Haase)
Date of Birth:	12.01.1987
Place of Birth:	Schwerin, Germany
Citizenship:	Citizen of Germany

SCHOOL EDUCATION

1993 - 1997	Primary school in Plate
1997 - 2006	Secondary school 'Gymnasium am Sonnenberg' in Crivitz Qualification: 'Abitur'

ACADEMIC EDUCATION

Oct. 2006 - Sept. 2009	Student at the Leibniz Universität Hannover, B. Sc. Plant Biotechnology (grade: "good") Thesis: „Proteomische Charakterisierung der Samenentwicklung von *Brassica napus*"
Oct. 2009 - Dec. 2012	Student at the Leibniz Universität Hannover, M. Sc. Plant Biotechnology (Grade: "with honors") Thesis: „Proteomische Charakterisierung des Endosperms von *Brassica napus* Samen"
Since Feb. 2012	PhD student at the Leibniz Universität Hannover, Institute of Plant Genetics, Department of Plant Proteomics, Prof. Dr. Hans-Peter Braun

PUBLICATIONS

Lorenz C., Borisjuk L., Rolletschek H., Heinzel N., Tohge T., Fernie A.R., Braun H.P., Hildebrandt T. M. (*in preparation*) Amino acid metabolism and the role of ETHE1 in *Arabidopsis thaliana* seeds.

Höfler S., **Lorenz C.**, Busch T., Brinkkötter M., Tohge T., Fernie A.R., Braun H.P., Hildebrandt T.M. (*in preparation*) Dealing with the sulfur part of cysteine: four enzymatic steps degrade L-cysteine to pyruvate and thiosulfate in *Arabidopsis* mitochondria preventing toxic effects of persulfides.

Lorenz C., Rolletschek H., Sunderhaus, S., Braun H.P. (2014) *B. napus* seed endosperm – metabolism and signaling in a dead end tissue. Journal of Proteomics 108: 382-426.

Lorenz C. (2014) Weltkongress der Pflanzenproteomik. Transkript 20, Heft 11, p 46.

Nietzel T., Dudkina N.V., **Haase C.**, Denolf P., Semchonok D., Boekema E.J., Braun H.P., Sunderhaus S. (2013) The native structure and composition of the cruciferin complex in *Brassica napus*. Journal of Biological Chemistry (288): 2238-2245.

Mwangi J.W., Rode C., Colditz F., **Haase C.**, Braun H.P., Winkelmann T. (2013) Proteomic and histological analyses of endosperm development in *Cyclamen persicum* as a basis for optimization of somatic embryogenesis. Plant Science 201(20): 52-65.

Noah A.M., Niemenak N., Sunderhaus S., **Haase C.**, Omokolo D.N., Winkelmann T., Braun H.P. (2013) Comparative proteomic analysis of early somatic and zygotic embryogenesis in *Theobroma cacao* L. Journal of Proteomics 78: 123-133.

Fromm S., Göing J., **Lorenz C.**, Peterhänsel C., Braun H.P. (2016) Depletion of the "gamma-type carbonic anhydrase-like" subunits of complex I specifically affects central mitochondrial metabolism in *Arabidopsis thaliana*. Biochimica et Biophysica Acta 1857: 60-71.

Petereit J., Katayama K., **Lorenz C.**, Schertl P., Kitsche A., Wada H., Frentzen M., Braun H.P., and Eubel H. (*in preparation*) Cardiolipin deficiency has pleiotropic effects on the plant mitochondrial electron chain.

CONFERENCE CONTRIBUTION

26. Tagung Molekularbiologie der Pflanzen in Dabringhausen, Germany (February 2013)
Poster: "The native structure and composition of the Cruciferin complex in *Brassica napus*"

1st INNPO World Congress on Plant Proteomics in Hamburg, Germany (June 2014)
Poster "Metabolism and signaling in a dead end tissue *Brassica napus* seed endosperm"

8th European Summer School Advanced Proteomics in Brixen, Italy (September 2014)
Poster: "Metabolism and signaling in a dead end tissue *Brassica napus* seed endosperm"

Proteomic Forum in Berlin, Germany (March 2015)
Poster: "Physiological role of ETHE1 for *A. thaliana* seed metabolism"

Danksagung

Lieber **Hans-Peter**, an dieser Stelle möchte ich mich als erstes bei Dir dafür bedanken, dass Du mir die Möglichkeit gegeben hast in Deiner Abteilung zu promovieren. Vielen Dank für die großartige Betreuung und Unterstützung während der gesamten Zeit in Deiner Arbeitsgruppe. Sieben einzigartige Jahre, in denen ich eine Menge über die Pflanzenforschung und Proteomik gelernt habe. Vielen Dank, dass Du mich auf diesem Weg begleitet hast. Ich möchte Dir weiterhin danken, dass Du es mir ermöglicht hast nationale und internationale Konferenzen zu besuchen. Es hat mir viel Freude gemacht meine Forschung mit anderen Wissenschaftlern zu diskutieren.

Vielen Dank **Prof. Schmitz**, das Sie sich als Korreferent dieser Dissertation zur Verfügung gestellt haben.

Der gesamten **AG Pflanzenproteomik** möchte ich ganz herzlich für die schöne Zeit danken. Danke für den freundlichen Umgang miteinander und eure Hilfsbereitschaft!

Liebe **Dagmar**, ich danke Dir dafür, dass Du mir schon im Bachelor-Studium das 1x1 der Proteomik beigebracht hast und mir bei Problemen oder Fragen jederzeit geholfen hast.

Ich möchte mich ganz herzlich bei **Stephanie Sunderhaus** für die Betreuung und Unterstützung am Anfang meiner Doktorarbeit danken.

Ein ganz besonderer Dank gilt **Tatjana**. Danke, dass ich gemeinsam mit Dir arbeiten durfte. Es hat mir sehr viel Freude gemacht. Ich habe wahnsinnig viel gelernt und Du hast mich mit der Begeisterung für Dein Thema infiziert. Danke ganz besonders für die große Unterstützung beim Fertigstellen der Manuskripte in den letzten Wochen!

Lieber **Holger**, ich danke Dir für Deinen unermüdlichen Einsatz im Labor, besonders wenn das Massenspektrometer mal wieder Schwierigkeiten gemacht hat. Ich möchte Dir danken, dass Deine grüne Bürotür jederzeit für mich offen stand und Du auch im größten Chaos immer Zeit gefunden hast, mich zu unterstützen.

Lieber **Hardy**, liebe **Ljudmilla**, vielen Dank für eure Unterstützung bei meinen Projekten als Experten der Samenbiologie. Es hat mir immer viel Freude gemacht, mit euch über meine neusten Ergebnisse zu diskutieren.

Bei meinen lieben Bürokolleginnen **Beate** und **Christine** möchte ich mich für die schöne Zeit in unserem „Wohnzimmer“ bedanken. Es war eben nicht nur ein Büro, sondern viel mehr. Ich möchte euch danken, dass trotz der vielen Arbeit der Spaß nie zu kurz kam. Ich werde euch sehr vermissen…**007** die Lizenz zum Forschen.

Liebe **Jenni**, vielen Dank, dass Du zu jeder Zeit für mich da warst und immer ein offenes Ohr für mich hattest.

Lieber **Michael**, ohne Deine kompetente Hilfe in allen Softwarefragen, wäre ich wohl heute noch nicht mit meiner Doktorarbeit fertig, sondern würde immer noch versuchen meine Massenspektrometrie-Daten per Hand auszuwerten. Vielen Dank für Deine Unterstützung.

Ich möchte **Tjorven** danken, die mit viel Einsatz und Ausdauer Samen präpariert hat und mich im Labor unterstützt hat.

Liebe **Edda**, danke für Deinen Einsatz und Deine Unterstützung während meiner Zeit in dieser Abteilung. Ich danke Dir weiterhin für die vielen wunderbaren Gespräche.

Vielen Dank and **Herrn Geyer** und **Herrn Barth**, die immer dafür gesorgt haben, dass es meinen Versuchspflanzen gut geht!

Ich möchte meiner **Familie** danken, die immer hinter mir steht und mich bei Allem unterstützt. Auch wenn es manchmal schwierig war, wart ihr jederzeit für mich da. Ich danke Euch für Eure Liebe.

Lieber **Bastian**, danke dass es Dich gibt!