Systems biology-based prediction for biomass formation in barley

Dissertation
zur Erlangung des akademischen Grades
doctor rerum naturalium (Dr. rer. nat.)

vorgelegt der
Naturwissenschaftlichen Fakultät I
Biowissenschaften, Institut für Biologie
der Martin-Luther-Universität Halle-Wittenberg

von

Herr **Mohammad Reza Ghaffari**
geb. am 23.08.1973 in Kashmar, Iran

verteidigt am
24.02.2014
begutachtet von:
Prof. Dr. Nicolaus von Wirén (IPK, Gatersleben)
Prof. Dr.Klaus Humbeck (Halle University)
Prof Dr. Philipp Franken (IGZ, Großbeeren)

Halle (Saale), Germany, 2014

Acknowledgments

First and foremost, I would like to thank my supervisor Dr. Mohammad Reza Hajirezaei for giving me the opportunity to pursue my PhD under his supervision, for all his kindness, help, and continuous support during four and half years at IPK. I would also like to thank my second supervisor Prof. Nicolaus von Wirén for many fruitful discussions, helpful suggestions, important advice, and constant encouragement during my PhD work. Furthermore, I would like to thank Prof. Philipp Franken and Prof. Klaus Humbeck for accepting to review the current thesis.

I thank Dr. Nese Sreenivasulu and his colleagues in abiotic stress genomics group in IPK, specially Mandy Pueffeld and Dr. Christiane Seiler for their helps in performing of microarray analysis. I also thank to former members of plant Bioinformatics group at IPK, Dr. Astrid Junker and Matthias Klapperstueck and Dr. Christian Klukas in image analysis group for their help in teaching and developing of “plug in” for visualization of metabolic and transcriptomic data in VANTED.

I gratefully acknowledge the technical, intellectual input and for their kindness of our laboratory colleagues especially Wally Wendt, Melanie Ruff, Andrea Knospe, Dagmar Boehmert, Heike Nierig, Christa Kallas and Susanne Reiner. I would like to express my grateful to my entire colleagues group for their support and very friendly and warm atmosphere and scientific condition especially in our lab meeting. I am especially thankful to Dr. Amir Hossein Ahkami, Seyed Abdollah Hosseini, Dr. Fahimeh Shahinnia, Young-Min Kim. I would like also to thank Dr. Britt Leps for all her patience, kindness, and support during our stay in Gatersleben.

I would like to express my sincere thanks to my parents, Gholamreza and Zahra for their endless love, their support, and their confidence in me. I am thankful to my brothers, Abbas and Hossein and my sisters, Masomeh, Ashraf and Azam. I thank my lovely wife, Maryam, for her unconditional love, patience, and especially for her constant support over the last four years.

Finally, I would like to thank the Leibniz-Institut für Pflanzengenetik und Kulturpflanzenforschung (IPK), Bundesministeriums für Bildung und Forschung (BMBF) and the Martin-Luther-Universität Halle-Wittenberg for their supports.

Mohammad Reza Ghaffari

Dedication

تقدیم به پدر و مادر عزیزم

و به همسر مهربانم مریم

Table of Contents

Data- CD

Excel File 1. Available passport data for spring and winter barley accessions

Excel File 2. Correlation matrices between selected metabolites and enzyme activities in each of the three developmental stages and biomass

Excel File 3. Correlation matrices between metabolic traits and biomass in winter barley

Excel File 4. The complete list of differentially expressed genes between highest and lowest biomass genotypes and their corresponding fold changes

Excel File 5. The complete list of barley differentially up- and down-regulated genes used for integration analysis

Excel File 6. Correlation matrices between metabolic traits and final biomass in spring barley

List of Figures

List of Tables

1. Summary

By increasing of world population and change in global climate, demand for food and clean energy has triggered worldwide efforts for improvement of energy crop. Understanding of fundamental importance for optimization of the storage of photosynthetically fixed energy in plant biomass are primarily promising and mitigate the competition for land between food and feed production on the one hand and bioenergy on the other hand. However starting point and strategies to achieve this objective is nevertheless limited based on the prior knowledge. Therefore, for the first time a detailed study was performed to make predictions for interaction between metabolites, enzymes, transcripts and biomass at the vegetative and generative stages in barley for future improvement of the shoot biomass and/or of the seed yield. To this end, an integration of network correlation analysis of key metabolites and enzymes of central metabolism measured by high throughput technologies in biochemical and transcriptional level along with biomass were applied at three different developmental stages of barley under agricultural conditions of which the key findings of this work are summarized in the following.

Spring barley

Three key stages of barley growth phase including tillering (vegetative, 30 days after planting), grain filling (generative, 15 days after anthesis) and repining phase (mature dry state) were the target for the investigation of metabolic and enzyme signatures determining shoot and seed biomass formation. To this end, a panel of 12 barley accessions, which were genetically and geographically diverse, were grown under agricultural conditions. Metabolite profiling, the measurement of maximum enzyme activity and transcriptome analysis provide information that may predict a close connection between enzymes, metabolites and biomass at different barley developmental stages. Deciphering of network correlation of metabolites and enzymes operating in tiller leaves of barley 30 days after planting revealed a large tendency of connections between metabolites and enzymes of TCA cycle, glycolytic pathway and sucrose metabolism, which was reflected in relations of a

metabolic signature consisting of 27 metabolites. These combinations were metabolites and enzymes of central metabolism, which suggest their role in biosynthesis of biomass compositions like lignin, cellulose, proteins, and fatty acids and the main growth regulator, starch. This finding also provided the idea that metabolites involved in source capacity have strong predictive power for the hypothesis that metabolic traits may be used as biomarker for screening of plants in breeding processes for biomass at vegetative stage in barley.

The network correlation analysis between metabolic traits and shoot biomass revealed coordinative changes of 35 metabolites and eight enzymes important for shoot biomass production 15 days after anthesis. These metabolites and enzymes were involved in glycolytic pathway and TCA. The correlation network also revealed a strong relationship between major carbohydrate Suc, and a minor carbohydrate Tre in starch biosynthesis as a major growth integrator in the cell. Transcriptome analysis of the flag leaves showed differentially expressed genes for enzymes involved in light and dark reactions, Suc and aromatic amino acid synthesis in plant with high biomass production. The combination of correlation network between metabolites signature and transcripts helped to define five candidate enzymes including PGM, PGI, AlaAT, cAldolase, and PK, which may play defining role in biomass formation at the generative stage.

The combinations of seed metabolite composition, enzyme activity with seed biomass revealed a well-organized large number of 60 metabolic traits. The vast majority of connected traits were the metabolites and enzymes involved in starch metabolism such as AGPase and sucrose metabolism like Suc and Susy. The combination of amino acids and seed biomass also supports the idea that remobilization of nitrogenous compounds is a major factor at the grain filling in seeds. Microarray analysis of seed transcripts exhibited an over-representation of genes occurring in starch, Suc, and trehalose metabolism in the plant with high biomass production. The integration of enzyme activities and transcripts identified five enzymes including AGPase, Susy, PGM, cAldolase and AlaAT demonstrating that carbon and nitrogen metabolism are closely related and those enzymes might be used for predication of seed biomass improvement at generative stage.

At the ripening stage, combination analysis of metabolite network led to the identification of three strong biomarkers including Glu, Gln, and starch. This finding emphasis on the role of two major amino acids in the synthesis of other nitrogenous compounds in seed and starch as major carbon reserves in seed at the final growth stage of barley.

A comparison between the signatures for tiller leaves, flag leaves and seeds in relation to shoot and seed biomass revealed that for tiller leaves Susy, 6PGDH, G6PDH, MDH, AspAT and Suc, ADPGlc, 2-oxo, Met, Lys, ATP, ADP and AMP were indicative signatures as enzymes and metabolites, respectively (vegetative stage). In contrast to the signatures of tiller leaves, additional enzymes including pAldolase, cAldolase, PK, PFK, AlaAT and PGI as enzyme signature and many other amino acids including Ser, Gln, Gly, Ala, Asp, Val, Ile, Thr, Tyr, His, Pro and metabolites including Tacnt, Glc6P, Mal, Glyct, Tre6P, Pglu, Qun, PPi, Tdca, Ddca, starch, Xyl, Fuc, Man and Raff as metabolite signature were identified in the flag leaves (generative stage). In addition to the signature in leaves, another enzyme and metabolite signatures were determined for seed biomass that included AGPase, CWInv, CytInv, GOGAT, ICDH as enzyme signature and Glc1P, UDP, Amdp, Put, OAA, dhdascb, GABA, Gal, PEP, ACC, Galcnt as metabolite signature.

Winter barley

The generative stage was the only major stage of winter barley that was investigated in detail to identify metabolic signature determining shoot and seed biomass formation. A combinatorial analysis of metabolic traits in flag leaves of winter barley mostly identified close relations of enzymes and intermediates of glycolysis, TCA cycle and nucleotides, which led to the integration of 24 biomarkers that in turn have predictive power for shoot biomass formation. In contrast to spring barley, winter barley genotypes showed a majority of negative connections of metabolites suggesting a strong need and depletion of metabolites in winter barley. Moreover, Man6P, 2-oxo, Mal, Benza, pFBPase and pAldolase were the strongest connections and can be used as metabolic signatures for shoot biomass at the generative stage in winter barley. The metabolites of the seed applied in a network correlation analysis with TKW revealed a metabolic signature

with 34 different metabolites. The combination of major nitrogenous compounds suggests that remobilization of nitrogen compounds is a major effect of TKW or seed weight in barley grain. Furthermore, the strongest connection of the enzymes, MDH, PEPC, AspAT, PGM, PGI and PK were profoundly connected to TKW and may show their role as putative biomarkers in seed or their major role in the regulation of seed metabolism in winter barley at the generative stage.

2. Introduction

2.1. Bioenergy: An opportunity for the future

Sustainable energy and food security are important presumptions for raising sustainable development at the globe. By 2050, the population of the world is expected to exceed 9 billion people. This means 34% increase in population and 70% increase in food and energy demand. Nevertheless, global climate change with increasing temperatures over the world and probably lower, more erratic rainfall will result in a decrease of food and energy production. Fossil fuels account for most of the world energy supply. The dependence on fossil fuels has increased and doubled in developed countries for 1971 to 2009. Greenhouse gas emissions effect is a result from increasing combustion of fossil energy carries. Amongst several pollutants causing climate change, CO_2 emission is the major factor in the earth's climate system. Furthermore, more than 2-3 billion people around the world have no access to modern energy carriers. On the other hand, the global community should aim to provide a minimum threshold of modern energy services for both consumption and productive uses. Solving this global problem requires a clean wider access to energy, which increases efficiency and concerns with both our present needs and the further needs of coming generations. Among clean energies, a shift should be made to the use of renewable energy sources such as wind, sun and biomass energy.

Biofuels derived from plant biomass are a potential source of sustainable energy that can contribute to future global demands. Agricultural products are expanding and with accelerating use of grains and plants, derived byproducts for bioenergy production in developed countries are placing new exigencies on global plant productions. To meet higher food and bioenergy demands, the improvement of plant productivity on existing farmland is probably the more realistic options compared to an increase of arable farmland with regard to the challenge to further boost plant biomass production. One important and particularly promising way is the optimization of the metabolic processes, which are responsible for the storage of photosynthetically fixed energy in plant biomass. Starch, cellulose beta-glucan,

and lipids are the important storage compounds in higher plants. Plant yield and quality ultimately depend on amount of and composition of the storage products. Production of the storage compounds with different composition and biomass is mainly governed by regulatory process within higher plants, targeting carbon partitioning. Improving cellulose, starch, and sucrose metabolic pathways is particularly important, because the photochemically fixed radiation energy is stored primarily in the form of reduced carbon compounds in the plants. Possible approaches and strategies for the modulation of such complex traits are the evaluation of biological interactions and the relation of biological networks and biological functions, which are reflected in recent developments such as systems biology.

2.2. Plant biomass

Biomass can be defined as biological material containing energy. Plants are photoautotrophic. They use derived chemical energy gained by photosynthesis to produce all the energy required for growth in forms of carbon like sugars, starch and other derived organic materials representing the biomass. Therefore, energy crop plants can be classified into three major group of biomass composition, which are end-products, comprising of sugars and starch, cellulose, lipids or oil.

2.2.1. The role of primary metabolism for biomass production

The growth and development of green plants depend on the efficiency of photosynthetic organs and the metabolic capacity. Plant growth is determined by the accumulation of biomass, which is the result of an ultimate expression of metabolic performance, cell proliferation with concomitant accumulation of biomass mainly through protein biosynthesis and cell expansion driven by water uptake (Anastasiou and Lenhard, 2007).

Plants function as integrated system, in which the availability of resource pools may have a direct influence on development and growth. The distribution of metabolites between growth, production of defense compounds and storage has to be tightly regulated. This regulation has been shown by several investigations including growth depression upon reduction of primary metabolism such as

sucrose synthesis (Hajirezaei et al., 1993) and fruit load reduction upon changing photoassimilate partitioning in tomato plants (Baldet et al., 2006). Sugars such as glucose (Glc) and sucrose (Suc) have been shown to act as metabolic signals in the control of plant growth and development (Rolland et al., 2002). High concentrations of sugars such as Glc and Suc, and low levels of mannose (Man) inhibit root growth of young *Arabidopsis* seedlings (Gibson, 2005). In addition to their role as signal molecules, the availability of sugars such as Suc has been shown to have an important effect on minerals toxicity (Abebe et al., 2010) and a positive effect on growth rate (Gibson, 2005). In addition to Suc, trehalose-6-phosphate (Tre6P) has recently been identified to act as a high carbon availability signal and is involved in signaling of the plant sugar status to control growth and development (Wingler et al., 2012).

Many transgenic approaches have targeted the production and/or distribution of primary metabolites within various parts of the plant such as the source or the growing areas and sink or the storage organs aiming at modification of plant growth and biomass production. Modified primary metabolism, such as sucrose biosynthesis or the tricarboxylic acid (TCA) cycle, has shown to have major and mostly negative effects on growth and development (Trethewey et al., 1998).

2.2.2. Regulation of primary metabolism: Role of enzymes in biomass formation

Carbon and nitrogen metabolism play an important role in plant growth and development. Certain enzymes have been shown to play a key role in reactions in carbon metabolism for storage of carbohydrates, and primary building blocks for phosphorylated intermediates and organic acids that provide the carbon skeletons for the assimilation of nitrate and ammonium and amino acid synthesis. Specific enzymes of nitrogen metabolism have facilitated the reduction and assimilation of nitrogen in cells (Zhang et al., 2010).

Cleavage of Suc, the major transport form of carbon in most species, initiates Suc utilization, which is catalyzed by invertase and/or sucrose synthase (Susy). Inhibition of invertase in maize resulted in an inhibited of transport of photoassimilates to seed organs and to a reduction in seed weight (Sturm and

Tang, 1999). Sucrose is also the precursor for starch synthesis which is formed by condensation of glucose-1-phosphate (Glc1P) with ATP. Starch synthase then transfers Glc residues from ADP-glucose (ADPGlc) produced by the action of ADP-glucose pyrophosphorylase (AGPase) to preexisting starch molecules. In addition to plastidic AGPase, a cytosolic AGPase was found and characterized in transgenic maize plants that showed enhanced seed weight (Wang et al., 2007). Starch and cellulose are also two important sources in biofuel production.

Enzymes catalyze metabolic reactions and their interconnections are strongly affected by genetic variation (Sulpice et al., 2010). The impact of genetic variation in metabolite levels and fluxes can be explained by changes in the activity of one or several enzymes. Thus, genomic variation can be reflected by enzyme activities. The genetic variation has provided a robust strategy to link the changes in metabolite levels to the genetic variation at the loci that encode enzymes (cis-acting) or regulate (trans-acting) the expression of enzymes. The study of 15 enzymes of central metabolism in an *Arabidopsis* population identified 15 co-located QTLs for structural genes, expression QTLs, and trans-acting loci (Keurentjes et al., 2008). In a recent study with 10 enzymes from central carbon and nitrogen metabolism, 73 significant QTLs were found that influence the activity of these 10 enzymes and eight QTLs that influence seeding biomass in the maize IBM mapping population (Zhang et al., 2010). In this study, all enzymes showed a negative correlation with seedling biomass. Furthermore, it has already been shown that the relative investment of protein in the enzymes has predictive power for biomass productivity and can be used for identification of genotypes with differences in photosynthesis and carbon use (Sulpice et al., 2010). Thus, the relative stability of enzyme activities and the possibility that they reflect long-term acclimation makes them ideal markers for the study of growth or complex traits in plants (Gibon et al., 2004)

2.3. Systems biology towards crop improvement

Modulation of complex traits is crucial for understanding their genetic bases, interactions between the key metabolites of the cells, organs, and systems and their interactions during life cycle of plant. Systems biology approaches offer the

chance to predict the behavior of multi component systems driving complex biological phenomena by describing causalities and interplays among the components, for example genes for new transcription factors or enzymes and regulatory metabolites and their interactions (positive and/or negative correlations and synergetic relationships) from large-scale omics datasets (Saito and Matsuda, 2010). Crop systems biology is defined as an approach to model complex traits to achieve a sustainable development for food and energy production. One major goal of crop systems biology is the formulation of complex traits at the crop level by the integration of information at the genome level and the whole metabolism. The genetic variation among the genotypes can be identified at the levels of the enzymome, proteome, transcriptome, or metabolome. In the plant functional genomics era, combined studies of physiological components with omics data facilitate the identification of the function of genes, biochemical pathways, and cellular processes that are affected in a coordinated manner (Stitt and Fernie, 2003). Summarizing, Crop systems biology can be designed for filling the gap between omics technologies and traits, improving biotechnological techniques and helping to reduce environmental problems and energy consumption. The roles of metabolomics, transcriptomics and the integration of Omics data in system biology approaches are detailed in the following sections.

2.3.1. Plant metabolomics: towards higher biomass and yield improvement

The generation of omics data originating from plant genome sequences is increasingly becoming popular in assessing plant phenotypes and genetic diversity. Metabolite engineering of sweet potato for enhanced protein quality and rice for enriched of iron and vitamin A content are known as examples for combating malnutrition (Al-Babili et al., 2001). These approaches have been used in an attempt to improve the metabolic composition of plants. With the advent of metabolite profiling technologies, investigation of metabolomes of plant populations has become feasible by targeted and non-targeted analysis of primary and secondary metabolites. Furthermore, plant metabolomics with analysis of hundreds and thousands of metabolites has provided the way for identifying broad changes in metabolite patterns in response to environmental or genetic perturbations (Stitt

and Fernie, 2003). Today, Metabolite profiling is being applied to an array of applications such as diagnostics for drug discovery and development; for instance, in the identification of drugs or biomarkers for organ-specific toxicities (Dagogo-Jack, 2012). In plant biology, most investigations with omics technologies have focused on the model species *Arabidopsis thaliana*, but today they are increasingly adapted to investigations in crop species (Fernie and Schauer, 2009).

2.3.2. Transcriptomics: towards gene identification and prediction of biomass formation

With the completion of the sequencing of plant genomes, the monitoring of global changes in gene expression is an increasingly attractive method for dissecting the molecular basis of interactions between plant and environmental conditions and for the improvement of breeding strategies based on identification of candidate genes involved in complex traits. Transcript profiling is a key element in functional genomics to understand gene interactions and active pathways involved in biological processes. (Sreenivasulu et al., 2008). Transcriptome analysis represents a valuable approach to discover molecular pathways and transcripts signatures that serve as useful biomarker in agricultural research (Heidecker and Hare, 2007). Taken together, combination of omics technologies like transcriptomics and metabolomics and other genetics approaches will improve our knowledge on the relationship between the genetic and the phenotypic architecture of agronomic traits and thus create a basis for knowledge-based molecular breeding (Hammer et al., 2006).

2.3.3. Integrative analyses of biological data

With advent of omics based and functional genomics approaches, the knowledge of regulatory systems has been expanded from the single gene level to gene networks. The disadvantage of each omics technique is that it addresses only one dimension within the biological organization of regulatory levels, e.g. transcripts, metabolites, or enzymes. With the increasing availability of omics data, more and more studies employ multidimensional integration of omics data to better annotate the function of gene products and their functional relationships in relevant

biological systems (Ge et al., 2003). This can increase the understanding of the functional activities in and between living cells by drawing biological networks and help to uncover genetic relationships in complex trait (Hirai et al., 2004). Moreover, a comparative analysis of two pairs of omics data can help to identify specific regulatory genes in two close genotypes to identify biological events causing effective development (Osorio et al., 2011).

Data generation, storage, integration, and visualization for systems biology

Model construction needs the most precise and comprehensive parameterization that is possible. To characterize the state of a metabolic system, information of metabolite concentrations, enzyme activities, transcripts, and metabolic fluxes are required. The large size of datasets in systems biology needs data storage capacities, statistic software, and visualization software tools. A large variety of databases on various aspects of genomics of barley has been reviewed recently (Sreenivasulu et al., 2008). Metacrop is a metabolic pathway database for crops and model plants, which provides information for managing, and exploring information about crop plant metabolism. This database contains currently information on 362 reactions of 54 pathways of barley (Schreiber et al., 2012). Furthermore, Metacrop allows the collection and export of metabolic models in the standardized data format, systems biology Markup languages (SBML) (Hucka et al., 2003). For the visualization of omics datasets, Mapman/PagMan, **V**isualization and **A**nalysis of **N**e**t**works containing **E**xperimental **D**ata (VANTED) and the **D**atabase for **A**nnotation, **V**isualization and **I**ntegrated **D**iscovery (DAVID) have been established as useful tools in the plant community (Thimm et al., 2004). MapMan/PagMan is a tool for the visualization and clustering of large transcriptome, metabolome and enzymome datasets onto diagrams of metabolic pathways or other processes. VANTED is a tool that enables the visualization of transcriptomics, proteomics, enzymomics, metabolomics, and other high-throughput data on the underlying networks (Junker et al., 2006). VANTED also contains various tools for statistical analysis such as correlation and clustering analysis and offers some useful plug-in to import metabolic pathways from the Metacrop database.

2.4. Barley as a crop species for biomass production

General information

As a member of the *Poaceae* family, Barley is one of the most important cereal crops due to its role in human food and animal feed production. Barley is estimated to have diverged from a common ancestor 50 to 70 million years ago (Kellogg, 2001) and is phylogenically related to wheat, *Brachypodium* and oat. The human use of the wild ancestor of barley, *Hordeum spontaneum*, started in the Near East about 17,000 BC (Zohary and Hopf, 2000). Barley plants can be classified based on their growing season, as spring or winter types, and on further characteristics such as two-rowed or six-rowed, hulled or hull-less or another end-use such as malting or feeding. Based on grain composition, barley is further divided into normal, waxy or high amylose starch types, high lysine, high beta-glucan, and pro-anthocyanidin- free type (Baik and Ullrich, 2008). The role of barley in human nutrition gradually changed from food grain to feed and malting grain but it is still one of the important dominant crops for some parts of Asia and Northern Africa (Newman and Newman, 2006). Recently, the interest in barley for food has increased mainly due to health claims associated to its soluble fiber and starch content. Barley contains beta-glucans and selenium as lowering blood cholesterol and antioxidant (Baik and Ullrich, 2008) and as a good sugar free source for people suffering from diabetes (Baik and Ullrich, 2008). In 2010, the barley world production was estimated to 123 million tones and, thus, ranked fourth in cereal crop world production after maize, wheat, and rice (faostat.fao.org). Germany with 10.4 million tones barley production per year was among the first top ten barley producers in the year 2010 (faostat.fao.org).

Barley as crop plant model for study of complex traits in C3 plants

Arabidopsis and rice are currently the best models for study of molecular biology and genetics. The small genome size of *Arabidopsis thaliana* (125 Mb) and *Oryza sativa* (272 Mb) and easy handling are the best criteria to use them as model plants. Barley with the almost genome size of 5000 Mb is one the best candidate as model among cereal. The genome sequencing of barley has started in 2006 and

a highly valuable draft genome sequence has been released at the end of the year 2012 (Mayer et al., 2012).

Several features including the simple diploid genome (2n=14) and self-pollination in addition to it valuable role in agronomy and economy have led to the establishment of barley as a model for genetics, physiology, and plant breeding long before *Arabidopsis* emerged as a model for plant scientists (Ullrich, 2011). Furthermore, barley is used as a genomic model in favor of its close relative, the hexaploid bread wheat (*Triticum aestivum*) (Devos and Gale, 2000). The advent of "omics"-approaches to study functions of genomes using genomics, proteomics, transcriptomics and metabolomics approaches, has further enhanced the value of barley as a model plant (Sreenivasulu et al., 2008). Recently, highly valuable online bioinformatics tools and databases including GrainGenes (http://wheat.pw.usda.gov), PlantGDB (www.plantgdb.org), HarvEST (http://harvest.ucr.edu), KEGG (http://www.genome.jp), and Plexdb (http://www.plexdb.org) have been installed to allow the exploration of thousands of genes and biochemical pathways in barley. The Availability of microarray platforms of Affymetrix Barley1 GeneChip and 44K Agilent microarray have also promoted transcriptome analyses in barley (Sreenivasulu et al., 2008). Taken together, with a combination of omics technologies and classical breeding, the expectation has increased to accelerate the improvement of complex traits in barley like, yield and tolerance against abiotic or biotic stress.

Possibilities for the improvement of a food and energy crop

Plant products in the form of wood, fibers, and food have long been utilized as raw materials by humans for numerous applications, such as burning for gaining energy, paper-making textiles and food. As already mentioned, plant biomass is currently also considered as renewable source of biofuels. Biofuels are mainly generated from starch, Suc, and cellulose. Plant crops such as maize, rice, sorghum and sugarcane have been considered as optional biomass sources for biofuel production. With the increasing population and world energy demands, interests in understanding the mechanisms involved in higher plant biomass production and yield enhancement have increased tremendously. Barley can serve

for all, food, feed and biomass production. Economically, plant biomass and seed yield are two important factors. Barley can be used as a bioenergy crop in two ways, either starch in seeds can be used to produce ethanol or the crop residues can potentially be used to generate lignocellulotic ethanol (Ullrich, 2011).

In the first case, however, the use of seeds for biofuel production stays in competition with its use as a food source. The relative strength of different source and sink organs will determine the flow of carbohydrates and nutrients within the plant and will largely delimit plant biomass and seed yield. So far, strategies based on single genes have not reached favorable conditions. Furthermore, it has been shown that the role of similar genes in monocots and dicots are completely different in plant growth and development. Over-expression of the DWARF4 gene in *Arabidopsis* has led to an increase of plant biomass and yield (Choe et al., 2001). Yet, the similar approach leads to a reduction of plant biomass and yield in rice (Van Camp, 2005). Thus, crop systems biology using a network of omics datasets might facilitate the identification of regulatory networks involved in plant biomass production. For instance, in sugarcane metabolite engineering of the sucrose pathway has led to an increase in Suc and starch that are two important sources for bioenergy production (Wu and Birch, 2007). Recently, it has been shown that modifying biosynthesis or signal transduction of genes involved in indole acetic acid (IAA), and brassinosteroid pathways have an impact on plant growth and crop biomass. Transgenic dwarf plants have higher cellulose contents and can be used for feeding and bioenergy production.

Seed yield relates to source strength, assimilate partitioning, and sink strength. In barley, altering sink strength and metabolism is a strategy for the improvement of seed production. Assimilation of nitrogen and metabolism of carbon is coordinated by a regulatory complex network of metabolites, enzymes, and transcripts. For example, inactivation of a single enzyme like AGPase has led to a decrease of ADPGlc, resulting in decreasing starch levels, storage protein accumulation and seed size in barley (Faix et al., 2011). This result showed that altering of a single gene could change the whole plant network and finally seed development. Taken together, interconnected relationships between whole plant carbon and nitrogen

metabolism with photosynthesis, sink-source-balance, and growth might be new targets for the improvement of plant productivity and seed yield.

2.5. Aims of the present work

The aims of this work are to assess regulatory network of metabolites, enzyme activities, and transcripts on primary carbon and nitrogen metabolism at vegetative, generative, and ripening stages in barley. Furthermore, the relationships between shoot and seed biomass are investigated in regard to metabolites, enzymes and transcripts that might serve as a signature for biomass formation.

In particular, different accession and introgression lines were classified in high and low biomass-accumulating plants based on their phenotypic differences. Correlation analysis is performed between the biomass and metabolites, enzymes and transcripts at different developmental stages of barley growth to elucidate possible signatures crucial for improving shoot and/or seed biomass. Finally, based on the results conclusions should be drawn whether specific organs and/or metabolic pathways would be feasible for a genetics and/ or breeding modification to increase the biomass or seed yield.

To realize these achievements, non-targeted and targeted metabolite profiling as well as transcriptome and enzymome approaches was applied. Bioinformatics tools were used to analyze and to combine the results of metabolites, enzyme activities, and transcripts for the prediction of biomass signatures. Different spring and winter accession lines and three spring introgression lines were chosen to obtain a broad genetic base and to increase the expected variability in biomass traits.

In general, total shoot biomass (TSB) was taken as plant biomass and total Seed weight (TSW) and thousand kernel weight (TKW) as seed yield biomass. At the final stage called ripening, biomass is defined as the final shoot biomass (FSB). The network correlation analysis was carried out among metabolites, enzymes activities, TSB, TSW, and TKW at the vegetative and generative stages and between only metabolites and FSB at the ripening stage. The correlation analysis between transcripts with TSB, and TSW were also performed at the generative stage in spring barley.

3. Materials and Methods

3.1. Origin of investigated plants

Spring barley

Ten accession lines of spring barley of diverse geographic origin were selected and obtained from the gene bank of the Institute of Plant Genetics and Crop Plant Research (IPK), Gatersleben, Germany. All the lines were evaluated for different traits including the days until heading, plant height, leaf width and length, tiller number, spike length, six-rowed or two-rowed, seed color and TKW in field trial conditions at IPK in Gatersleben in the year 2009. Out of 10 accession lines, eight lines were selected for a detailed metabolite profiling and morphological analysis. The accession IDs (identity), accession numbers, the origins and the colors of various lines are provided in the Appendix Table1.

In addition, ten introgression lines were donated by Prof. Pillen from the University of Halle-Wittenberg was and included in the analysis. Introgression lines were derived from a cross between the German barley cultivar Scarlett and the Israeli wild barley accession ISR42-8. They carry different QTLs for important agronomic traits including yield, grain per ear and TKW as documented in Schmalenbach et al. (2009). In order to select contrasting introgression lines, field test were conducted at the IPK in Gatersleben in the year 2009. Altogether, three introgression lines were differing in TKW, grains per ear. These lines were designated as S42IL107, S42IL122, S42IL119, and Scarlett and were selected for further investigation. Compared to the accession lines, all introgression lines showed yellow seed color (Figure 1).

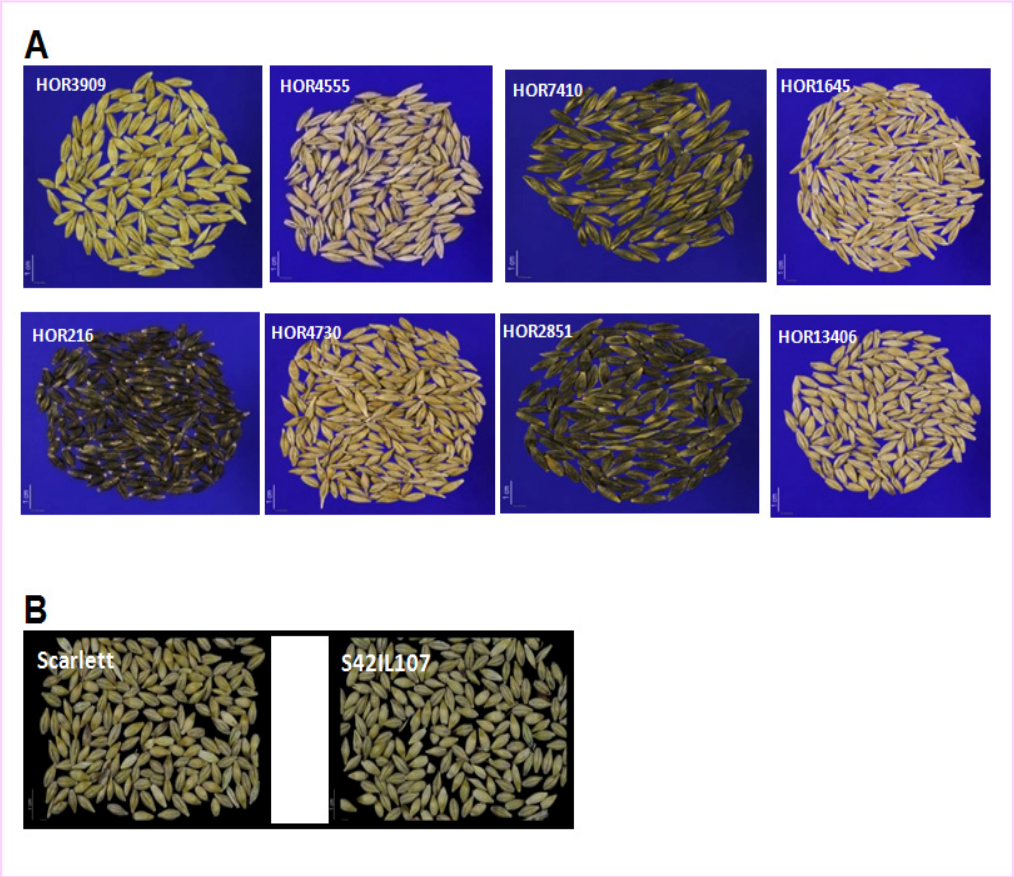


Figure 1. Seed color of accession lines and introgression lines. (A) Accession lines are shown in the upper panel and differ in their seed color being yellow (HOR3909), white (HOR 4555, 1645, 4730 and 3406); deep black (HOR216) and light black (HOR7410 and 2851). (B) All Introgression lines including S42IL107 and Scarlett showed yellow color.

Winter barley

In addition to spring barley lines, mature seeds of twenty accession lines of winter barley were obtained from the gene bank of IPK and were included in the biochemical investigation. In 2009, eight of twenty accession lines could not survive the hard winter whereas the other twelve lines were cold tolerant and used for further analysis. These genotypes were classified based on their spike density including Lax, intermediate or dense and origin based on Gene bank data at IPK (Appendix Table 2).

3.2. Growth conditions and definition of developmental stages of barley plants

Seeds of spring and winter barley were germinated separately in pots in a climate-controlled growth chamber for two weeks. The temperature was 12°C at night and 15°C during day with 12h: 12h light/dark cycle. Thereafter, seedlings were transferred in the field and were planted in a defined plot with 3.40m x 1.20 m and under standard agricultural practice (Figure 2A). The dates of barley planting were October 2009 and March 2010 for winter and spring barley, respectively. All plants were randomly distributed within the field. Each line had six plants. The spacing between plants and lines was 10 cm and 40 cm at planting, respectively. Each replication comprised of eighteen plants per accession. The experimental design was randomized. The average growth temperature was 16.8°C from May until August 2010 for spring and winter barley. Samples were harvested at three different developmental stages according to BBCH scale for cereals (Lancashire et al., 1991). These stages were vegetative, generative and repining and are described in details in the following sections (Figure 2B).

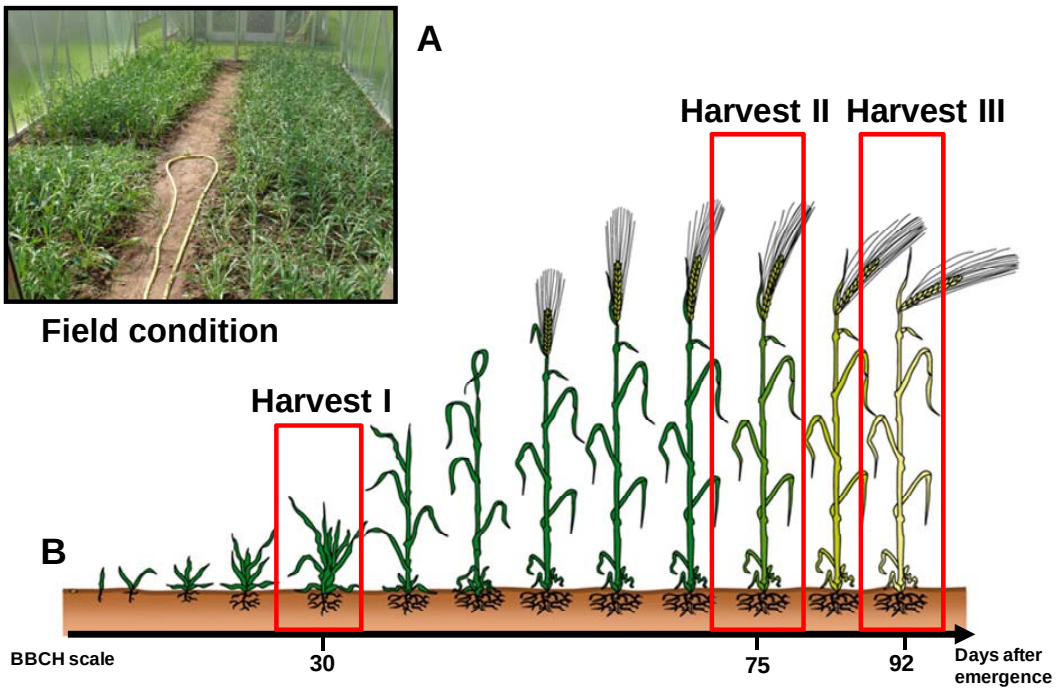


Figure 2. (A) Growth field and (B) three harvest time points based on BBCH scale for cereals. Samples were harvested 30 (tillering, harvest I), 75 (grain filling, harvest II) and 89 (ripening, harvest III) days after sowing for spring barley and 75 (harvest II) and 92 (harvest III) days for winter barley. Flag leaves, and seeds were harvested and subjected to metabolomics, transcriptomics and enzymomics analysis.

Vegetative stage

Tillers usually begin to emerge when barley leaves have three leaves corresponding to the principal stage of 2 of the BBCH scale and end when the plants have maximum detectable tillers 30 day after planting (Figure 2B). Samples of the leaves of whole tillers from HOR lines and introgression lines were harvested three hours after sunrise at the end of tillering phase (Figure 2B, Harvest I). Plant material was frozen immediately in liquid nitrogen for further analysis. Four phenotypic data including TSB, total leaf biomass (TLB), total dry weight (TDW) and leaf area (LA) were measured for spring barley.

Generative stage

At the BBCH 60, flowering ears were tagged for the determination of the days after anthesis. In this time point, the kernel rapidly increases in size and a milky fluid can be squeezed from the developing kernels. Fifteen days after anthesis, two different tissues including flag leaves, and seeds of spring and winter barley were harvested in the field (Figure 2B, harvest II). Harvested materials were frozen immediately in liquid nitrogen and stored at -80°C for further analysis. Total shoot biomass, TSW, flag leaf fresh weight (FLFW), flag leaf dry weight (FLDW), peduncle length (PDL), tiller number and flag leaf area (LA) of individual plants were scored simultaneously.

Ripening stage

Barley plants were harvested at BBCH 92 when kernels physically matured (Figure 2B, harvest III). Dried seeds were used for analysis of seed composition at the ripening stage of barley in HOR and introgression lines. TKW, grains yield (GY) and final shoot biomass (FSB) was measured at the end of growth period of spring and winter barley.

3.3. Preparation and measurement of soluble and insoluble carbohydrates

Soluble sugars and starch were determined in leaf and seed using a well-established method (Chen et al., 2005). 50 mg each of Frozen leaf, seed and peduncle of barley plants from different developmental stages were homogenized

in liquid nitrogen, dissolved in 0.75 ml of 80% (v/v) ethanol and incubated at 80°C for 60 minutes. Crude extracts were centrifuged at 14,000 rpm at 4ºC for 5 min and upper phase was concentrated in speed vacuum concentrator (Christ, Germany) at 45°C for 180 minutes. The pellet was re-suspended in 0.25 ml HPLC grade water and shaken for 15 minutes at 4°C. The remaining insoluble material was kept for starch measurement.

A buffer containing 100 mM imidazol-HCl (pH 6.9), 5 mM $MgCl_2$, 2.25 mM NAD, 1 mM ATP as final concentrations was used for the measurement of soluble sugars using EL808 ultramicroplate reader (BioTeK, Inc, Germany) at 340 nm. A sequential addition of auxiliary enzymes allowed detecting Glc, fructose (Fru) and Suc. Glucose-6-phosphate dehydrogenase (G6PDH) was first added to remove endogenous hexose-phosphates. Subsequently, hexokinase (HK), phophoglucoisomerase (PGI) and β-fructosidase were added successively to measure Glc, Fru and Suc as described in Hajirezaei et al. (2000).

The residue of sugar extraction was washed twice with one ml of 80 % (v/v) ethanol. Starch was decomposed with 0.4 ml of 0.2 N KOH for 16 h at 4°C and neutralized with 70 µl of 1 M acetic acid. Hydrolysis of starch was performed using a 1:1 ratio of sample and a buffer containing 50 mM sodium acetate, pH 5.2 and 7 units/mg of amyloglucosidase (Roche, Germany). The cocktail was incubated in 37°C for 16 hours. Determination of produced Glc was performed according to Hajirezaei et al. (2000).

3.4. Measurement of free amino acids

Free amino acids were extracted as described in section 3.3. To detect primary and secondary amino acids, a fluorescing reagent AQC (6-aminoquinolyl-N-hydroxysuccinimidylcarbamat) was used. ACQ was dissolved in 3 mg per ml of acetonitril and incubated in 55^{o}C for 10 min. Twenty microliters of sugar extract were derivatized in a cocktail containing 20 microliters of fluorescing reagent ACQ, 160 microliters of a 0.2 M boric acid buffer (pH 8.8) in a final volume of 200 microliters. The solution was incubated at 55^{o}C for 10 min.

The separation of derivatized samples was carried out with a reversed phase HPLC system (Waters, Germany) consisting of a gradient pump (Alliance 2795 HT,

Waters, Germany), a degassing module, an autosampler and a fluorescence detector (Waters 2475, Germany). A reversed phase column (XBridge; 150 mm, 5 µm) was used for separation and detection of amino acids at an excitation wavelength of 300 nm and an emission wavelength of 400 nm. The gradient was accomplished with a buffer A containing 140 mM sodium acetate, pH 5.8 (Suprapur, Merck) and 7 mM triethanolamine (Sigma, Germany). Acetonitril (Roti C Solv HPLC, Roth) and purest HPLC water (Geyer) were used as eluents B and C. Chromatograms were recorded using the software program Empower Pro.

3.5. Measurement of sugar alcohols

The concentrations of sugar alcohols were determined using an ion chromatography system (Dionex, Idstein, Germany) consisting of a gradient pump (GS50), an ED50 electrochemical detector, and an autosampler (AS50). Anionic compounds were separated on a CarboPac MA1 column (4×250 mm) connected to a guard column (4×10 mm) and an ATC-1 anion trap column that was placed between the eluent and separation column. The eluent (500 mM NaOH) was made from HPLC grade water (Millipore) and 50% NaOH (Merck, Germany). The column was equilibrated for one hour at a flow rate of 0.35 mL/min. The run time was 50 min and a linear gradient was accomplished with 52% water and 48% NaOH. The calibration and quantitative calculation of sugar alcohols was carried out using Chromeleon client software 6.6 SP, Build 1566 (Dionex, Germany).

3.6. Metabolite profiling by ICMS/GCMS

Extraction of primary metabolites

Primary metabolites were extracted from tiller leaves, flag leaves and seeds of twelve-spring barley plants at three developmental stages, vegetative, generative and ripening. For winter barley accessions, flag leaves and seeds were used for extraction at the generative stage. To this end, samples were harvested and frozen immediately in liquid nitrogen. Five independent biological replications were used in this study. Plant tissue was ground in liquid nitrogen and each 50 mg of fine powdered fresh materials were extracted using 1 ml (v/v)(1:1) ice-cold methanol

and chloroform. Subsequently, 0.4 ml of purchased LC-MS water was added to each tube. The mixture was mixed and kept on ice for twenty minutes. The samples were centrifuged for 10 minutes at 14000 rpm and 4°C. Thereafter, the upper phase containing methanol/water was transferred to new Eppendorf tubes and concentrated at 45°C for 2 hours in a speed vacuum concentrator (Christ, Germany). The remaining pellet was re-suspended in 0.4 ml of LC-MS water and was kept at -80°C for metabolomics analysis.

Non-targeted metabolite measurement by gas chromatography-mass spectrometry

In order to detect primary metabolites that can be either identified using publically available libraries or are of unknown origin, a sensitive gas chromatography system connected to a time of flight mass spectrometer (GC-MSTOF) was performed. The metabolite extracts prepared in section 3.6 were derivatized using an established method by (Lisec et al., 2006). Briefly, twenty microliters of extracts were concentrated in a vacuum concentrator to dryness for two hours at 45°C. The dried debris was solubilized and derivatized with 40 µL of methoxyamine hydrochloride (20 mg·mL-1 in pyridine) for 2 hours at 37°C. Seventy microliters of N-methyl-N-trimethylsilyltrifluoroacetamide (MSTFA) was added to each sample and incubated for 30 min at 37°C. Alkane's mixture (C7-C30) was added to all samples to facilitate chromatogram alignment and calculation of retention indices. Ribitol was used as an internal standard in all samples. Chromatography was conducted on a GC-TOF-MS system including a gas chromatograph A7890 (Agilent, Germany), an auto injector, Twister XXL (GERSTEL, Germany), and a TOF mass spectrometer GCT Premier (Waters Corporation, Germany). One microliter of derivatized sample was injected using splitless injection method at 80°C. The compounds were separated on a 30 m DB5ms capillary column (Phenomenex, Germany) with 0.25 mm inner diameter and 0.25 µm film thickness including a 5 m guardian column. Injection temperature was 240°C. Temperature program for GC separation was 3 min at 80°C isothermal followed by a ramp of 5°C min^{-1} to 310°C for 5 min. MS data in a range between 50 to 650 m/z was recorded with 1 scan/s over 50 min with Mass Lynx 4.1 (Waters Corporation).

GC-MS data mining

MeltDB web based software (Neuweger et al., 2008) was used for storage, analysis and annotation of GCMS chromatograms. Briefly, all chromatograms were converted to netCDF file using Databridge software (Waters Corporation) and imported to Web-based processing platform MeltDB. All chromatograms were located in a distinct chromatogram group. A MeltDB experiment can be created by selecting and organizing chromatogram groups. The spectral mass of each chromatograms in a MeltDB experiment were subjected to identification by searching and matching to **I**nstitute of **S**tandards and **T**echnology,(NIST), **G**olm **M**etabolite **D**atabase (GMD) and by comparison with retention indices which were manually identified for all chromatograms using a series of alkanes. Two different tools including MeltDB profiling pipeline for observation and annotation and MeltDB warped profiling pipeline for observation of peaks were used as MeltDB methods. The detection of peak in all samples was performed for high sensitivity (FWHM = 7, SN = 10) using the xcms tool of R (Smith et al., 2006). Extracted ion chromatograms of representative fragments of the identified compounds were computed. These peak intensities were normalized with respect to the ribitol and fresh weight extracted. The final data were exported as a CSV file for further analysis. Missing values were computed using an expectation-maximization algorithm (EM) module of SPSS software (Allison, 2002).

Targeted metabolite analysis by ion chromatography mass spectrometry (ICMSMS)

In order to analyze and to quantify primary metabolites, a targeted metabolite analysis was performed using external standards. 0.1 ml of samples extracted in section 3.6 was filtered at 2000 × g for 90 min using multiscreen filter plate (multiscreen ultracel-10 regenerated cellulose ultra-filtration membrane 10000 NMWL). The IC-MS instrumentation consisted of a Dionex ICS5000 (Dionex, Idstein, Germany) with a 6490 triple Quad LC/MSMS (Agilent, USA). Anionic compounds were separated on a 250×2 mm AS11-HC column (Dionex) connected to a 10×2 mm AG 11-HC guard column (Dionex) and an ATC-1 anion trap column. The gradient was produced with H_2O (buffer A; HPLC grade water) and KOH which

was generated by an EGCIII KOH eluent generator cartridge. The column was equilibrated with a mixture of buffer A (96%) and 4% KOH at a flow rate of 0.38 ml min^{-1} and heated to 37°C during the measurement. The gradient was produced by changes of KOH concentration as follows: 0–4 min: 4%; 4–15 min: 15%; 15–25 min: 25%; 25–28 min: 50%; 28–31 min: 80% and 31-40 min; 4%. Quantitative analysis of metabolites was performed using an Agilent 6490 triple quadruple (QQQ) mass spectrometer (Agilent, Germany). Electron spray ionization (ESI)-MS/MS condition was as follows: gas temperature 350°C, drying gas flow rate 12 L per min, nebulizer pressure 35 psi, capillary voltage ± 3.5 kV. The fragmentor voltage and collision energy were optimized for each compound individually by tuning standards with a defined concentration. Primary metabolites were detected in negative ion mode using multiple reactions monitoring (MRM) (Appendix Table 3). The data were extracted using MassHunter software version B.03.01 (Agilent Technologies, Germany). Quantification of metabolites was performed by creating a batch for each sample sets using Quantitative Analysis (QQQ) software (Agilent Germany). ^{13}C-pyruvate was used to normalize the data and was added to each sample as internal standard before analysis.

3.7. Protein extraction and enzyme activity measurement

Crude protein extracts were prepared from frozen leaves and seeds of barley with 0.5 ml of extraction buffer containing 50 mM Tris-HCl, pH 6.8, 5 mM $MgCl_2$, 5 mM mercaptoethanol, 15% glycerol, 1 mM EDTA, 1 mM EGTA and 0.1 mM pefablock phosphatase inhibitor. Pefablock phosphatase was always prepared freshly and added prior to extraction. The homogenate was centrifuged at 14,000 rpm for 5 min, and the supernatant was used for the assay of enzyme activities. The remaining debris were kept for a cell wall invertase assay. The maximum enzyme activity was determined spectrometrically using an ultramicroplate reader (BioTeK Inc, Germany) in coupled reactions by monitoring NADH or NADPH oxidation or NAD^+ or $NADP^+$ reduction at 340 nm.

Invertases (EC 3.2.1.26)

Invertases were distinguished based on their solubility, localization and pH optima, namely, cytoplasmic, acidic and cell wall invertases. The enzyme activity assays of invertases are described in detail in the following.

Cytoplasmic (CytInv) and acid invertase (AcidInv)

The activity of both CytInv and AcidInv was measured with 0.5 M purest sucrose and 20 µl of protein extract. The incubation buffer for CytInv contained 50 mM Hepes-KOH (pH 7.5), whereas the buffer for AcidInv was 50 mM sodium acetate (pH 5.2). Samples were incubated at 37ºC for 3 hours and the reaction was terminated at 95ºC for 5 min. Blank samples contained a reaction mixture of buffer and protein extract and were inactivated immediately for 5 minutes at 95°C (Hajirezaei et al., 2000; Zrenner et al., 1996). Produced Glc was measured enzymatically as previously described in section 3.3.

Cell wall invertase (CWInv)

Cell wall invertase was assayed using the remaining pellet of the samples described in section 3.7. Pellets were washed twice with a buffer containing 50 mM Tris-HCl (pH 6.8) and 5 mM $MgCl_2$ and centrifuged at 13000 rpm for 2 min. Incubation was performed in 90 microliters of a buffer containing 50 mM sodium acetate (pH 5.2), 0.5 M purest sucrose and complete pellets. The homogenate was incubated at 37°C for 180 min and neutralized immediately after incubation by adding 10 microliters of 1 M Tris-HCl (pH 8.0). The reaction was terminated at 95°C for 5 min (Zrenner et al., 1996). Samples were centrifuged at 13000 rpm for 5 min and produced Glc was measured enzymatically as previously described in section 3.3.

Phosphofructokinase (PFK, EC 2.7.1.11)

Phosphofructokinase was photometrically assayed using 75 mM Hepes (pH 7.9), 15 mM $MgCl_2$, 0.1 mM NADH, 5 mM 1 U aldolase, 1 U triose-phosphate isomerase (TPI) and 1 unit glycerol 3-phosphate dehydrogenase. A final concentration of 1 mM ATP was used as starter (Hajirezaei et al., 1993).

Glucose-6-phosphate dehydrogenase (G6PDH, EC 1.1.1.49)

Enzyme activity of G6PDH was assayed in a buffer containing 100 mM glycylglycine (pH 8.0), 150 mM $MgSO_4$ and 60 mM NADP. The assay was started by the addition of 15 mM Glc6P (Redinbaugh and H Campbell, 1998).

Glutamate synthase (GOGAT, EC 1.4.1.13)

Glutamate synthase (GOGAT) was measured in a reaction mixture containing 50 mM Tris-HCL (pH 7.9), 40 mM sorbitol (Sbt), 5 mM 2-oxoglutarate (2-oxo), 5 mM Glu, 10 mM Glc6P and 50 µl of fresh protein extract. Samples were incubated at 30ºC for 3 hours and stopped at 95ºC for 5 min. Blanks containing buffer and protein extract were inactivated at 95ºC for 10 min (Suzuki et al., 2001). The determination of produced glutamate (Glu) was performed by a reversed phase chromatography as essentially described in section 3.4.

Sucrose synthase (Susy, EC 2.4.1.13)

For the measurement a new MS-based method was established in which produced UDP-glucose (UDPGlc) was detected after incubation of protein extracts with appropriate substrates. The reaction assay medium contained 100 mM Hepes-KOH (pH 7.0), 0.5 M sucrose, 20 mM UDP and 20 µl of the protein extract in a final volume of 100 µl. Samples were incubated at 30ºC for 3 hours. The reactions was stopped at 95ºC for 5 min. Blanks were prepared with the above protocol but were heat-inactivated at 95ºC for 5 min (Zrenner et al., 1995). The determination of UDPGlc was performed by the same method as described in section 3.6 for metabolites using IC-MSMS with minor modification. The gradient of ion chromatography conditions was produced by changes of KOH concentration as follows: 0–4 min: 4%; 4–6 min: 25%; 6–12 min: 60%; and 12–16 min: 80%. Mass spectrometer parameters were set as described in section 3.6.

Glyceraldehyde-3-phosphate dehydrogenase (GAPDH, EC 1.2.1.9)

The activity of GAPDH was measured in 0.1 M Tris (pH 7.0) containing 5 mM $MgCl_2$, 1 mM NaF, 5 mM DTT, 5 mM 3-phosphoglycerate (3PGA), 0.4 mM NADPH, 2 units phosphoglycerate kinase (PGK) and protein extract. The reaction

was started by adding ATP in a final concentration of 2.5 mM (Sibley and Anderson, 1989).

Pyruvate kinase (PK, EC 2.7.1.40)

Pyruvate kinase activity was measured according to Ireland et al., (1980) with slight modification in 0.5 M TES (pH 7.9) containing 10 mM $MgCl_2$, 40 mM KCl, 0.1 mM NADH, 2 mM dithiothreitol, 0.5 mM PEP and 4 units lactate dehydrogenase (LDH). The reaction was started by the addition of 1 mM ADP in the assay medium.

Malate dehydrogenase (MDH, EC 1.1.1.37)

Malate dehydrogenase activity was assayed in a reaction medium containing 50 mM Heps-KOH (pH 8), 0.2 mM NADH and 0.5 mM EDTA. The reaction was started by the addition of 5 mM oxaloacetate (OAA) in solution (Jenner et al., 2001)

Phosphoenolpyruvate carboxylase (PEPC, EC 4.1.1.31)

PEPC activity was determined spectrophotometrically at 340 nm in a reaction buffer containing 50 mM Tris (pH 7), 10 mM $MgCl_2$, 10 mM $KHCO_3$, 0.2 mM NADH, 10 mM DTT and 2 units MDH. The reaction was started with the addition of 5 mM PEP (Rolletschek et al., 2004).

Isocitrate dehydrogenase (ICDH, EC 1.1.1.42)

Isocitrate dehydrogenase was assayed in a reaction mixture containing 20 mM Hepes pH (7.5), 1 mM $MnCl_2$, 0.2 mM NAD(P). 2 mM isocit was used as starter (Gibon et al., 2004).

6-Phosphogluconate dehydrogenase (6PGDH, EC 1.1.1.44)

Phosphogluconate dehydrogenase was analyzed in a reaction buffer containing 50 mM Hepes–NaOH (pH 7.2), 5 mM $MgCl_2$, 1 mM NADP. A final concentration of 1 mM 6-phoshogluconate was used as starter of the reaction (Redinbaugh and Campbell, 1998).

Phosphoglucoisomerase (PGI, EC 5.3.1.9)

The enzyme activity of PGI was assayed in a buffer containing 50 mM Hepes-NaOH, 5 mM $MgCl_2$, 1 mM NAD, 1 unit glucose-6-phosphate dehydrogenase. Reaction was started by adding of 5 mM fructose-6-Phosphate (Fru6P) as final concentration (Periappuram et al., 2000).

Phosphoglucomutase (PGM, EC 2.7.5.1)

The activity of PGM was determined in a reaction assay containing 50 mM Hepes-NaOH pH (7.2), 50mM $MgCl_2$, 5mM Glc6P, 1 mM NAD, and 2 units G6PDH. The reaction was started with 0.1 mM glucose-1,6-bisphosphate (Glc1,6BP) as substrate (Periappuram et al., 2000).

Alanine aminotransferase (AlaAT, EC 2.6.1.2)

Alanine aminotransferase was assayed in a reaction buffer containing 100 mM Tris-HCl pH (7.5), 50 mM alanine (Ala), 0.2 mM NADH, 5 units LDH and protein extract. The reaction was started by adding of 15 mM 2-oxo as substrate (Holtum and Winter, 1982).

Aspartate aminotransferase (AspAT, EC 2.6.1.1)

A buffer containing 50 mM Hepes-NaOH (pH 7.2), 5 mM Mg Cl_2, 5 mM Glc1P, 1 unit G6PDH and 1 mM NAD was used for the measurement of AspAT. 0.1 mM Glc1,6BP was used as starter in final concentration (Holtum and Winter, 1982).

ADP-Glucose Pyrophosphorylase (AGPase, EC 2.7.7.27)

The enzyme activity of AGPase was measured spectrometrically by monitoring NADH production at 340 nm. The reaction buffer contained 100 mM Heps-KOH (pH 7.8), 2 mM ADPGlc, 5 mM $MgCl_2$, 0.5 mM NAD, 2 units PGM, and 5 units G6PDH. The reaction was started by adding sodium pyrophosphate (PPi) at a final concentration of 2mM (Wang et al., 1998).

Cytosolic aldolase (cAldolase, EC 4.1.2.13)

Activity of cAldolase was measured in a buffer containing 100 mM Tris (pH 8.5), 5 mM $MgCl_2$, 0.2 mM NADH, 1 mM EDTA, 2 units glycerol 3-phosphate dehydrogenase and 2 U TPI. The reaction was started by the addition of 2 mM Fru1,6BP as starter of reaction. Produced NAD^+ was measured spectrometrically at 340 nm (Haake et al., 1998).

Cytosolic fructose-1,6-bisphosphatase (cFBPase, EC 3.1.3.11)

The activity of cFBPase was assayed in a mixture containing 100 mM imidazol (pH 7.1), 5 mM $MgCl_2$, 0.25 mM NADP, 1 U G6PDH (Yeast) and 1 U PGI. A final concentration of 0.12 mM of Fru1,6BP was used as starter. The NADPH production was measured at 340 nm using ultramicroplate reader (BioTeK, Inc) (Kelly et al., 1982).

Plastidic fructose-1,6-bisphosphatase (pFBPase, EC 3.1.3.11)

The activity of pFBPase was determined in a reaction assay containing 50 mM Tris-HCl (pH 8.2), 1 mM EDTA, 20 mM $MgCl_2$, 0.20 mM NADP, 5 mM DTT, 1 U G6PDH (Yeast) and 1 unit PGI. A final concentration of 0.12 mM of Fru1,6BP was used as starter (Kelly et al., 1982).

3.8. Graphical visualization and heat Maps

All graphs were designed using Microsoft Excel module of Microsoft office version 2007 (Microsoft Inc). Heat maps of correlation matrices in spring and winter barley were visualized in Microsoft Excel using a color-coding of significant negative and positive associations at the levels of pBH<0.01 or pBH<0.001.

3.9. Network correlation analysis

Correlation analyses were performed by calculating all pair-wise correlations (metabolites, enzyme activities, transcripts and phonotypical data) and their combinations in different tissues and at various developmental stages of barley using a web-based software platform, MetaboAnalyst (Xia et al., 2009). The data file was loaded into Metaboanalyst version 2 using comma-separated files. A

Pearson's product-moment correlation which applies a Benjamini-Hochberg (Benjamini and Hochberg, 1995) was used to assign the significant associations between different metabolites, enzymes and phonotypical data. The correlation matrices were imported to Excel file and visualized as described in section 3.8.

3.10. Visualization of network

VANTED is a state of the art tool for visualization and analysis of omics data in the context of networks (Junker et al., 2006). Data for analysis is designed in a predefined Excel-based form and imported into the software. The data can be mapped on drawn networks using software tools or the supported graphic files of pathways (xml and sbgn files). Node degree centrality was extracted using tool data tab after performing a pair-wise correlation analysis. Visualization of correlated data and ratio of node degree centrality on drawn biochemical network was carried as described in instruction manual (vanted.ipk-gatersleben.de*).*

3.11. Microarray analysis

A custom barley cDNA Microarray (Agilent Technologies, Germany) containing 56000 barley oligonucleotides was used for transcriptome analysis on flag leaves and seeds of spring barley at generative stage. A complete list of the genes on this cDNA microarray is provided in Excel File 4 (Data-CD).

cDNA microarray labeling, hybridization, and scanning

Total RNA was extracted from 200 mg leaves, and seeds using TRIzol reagent (Invitrogen, USA) following the manufacturer's protocols. RNA was purified using the RNasey MinElute clean up kit (Qiagen, USA). All total RNA samples were quality checked before starting processing steps by running out of each replication using a RNA chip technology that was assessed on an Agilent BioAnalyzer 2100 (Agilent Technology, Germany). All samples were determined to be of sufficient quality for next processes. Single-stranded, and double-stranded cDNA was synthesized from the poly (A^+) in the isolated RNA using the Affinity Script RT kit (Agilent Technology, Germany). cRNA was synthesized using T_7 RNA polymerase, Cy_3-CTP and nucleotides mix. Spike controls were included according to

manufacturer's instructions. Qiagen's RNeasy mini spin columns (Qiagen, Germany) were used for purification of the amplified cDNA samples. The labeled probes were hybridized with the barley cDNA microarray using a commercial hybridization buffer (Agilent Technologies, Germany) for 16 hours at 65°C in a hybridization chamber. After hybridization, the slides were washed using gene expression wash buffers, GE Wash Buffer 1 and GE Wash Buffer 2 (Agilent Technology, Germany) for 1 min at room temperature according to the manufacturer's protocols. The Cy3 fluorescent intensities for each spot were estimated by an Agilent Microarray Scanner. Scanner settings were as follows: Dye channel green, Scan Area (61 x 21.6 mm) and Scan resolution 3 µm. A 20-bit Tiff images was acquired for each image. Agilent Technologies FE 9.1 software was used for feature extraction of scanned images.

Analysis of cDNA microarray data

To identify potential differentially expressed genes, Gene spring software version 12.5 (Agilent Technology, Germany) was used. The data were imported to the software after creating a project and categorized based on tissues and genotypes. The normalization of data was carried out using quintile normalization. The values were filtered using coefficient of variation. The one way analysis of variation (ANOVA) which applied a Benjamini-Hochberg (P≤ 0.05) (Benjamini and Hochberg, 1995) for correction of *p*-value was used to assign significant differences in the expression levels.

Visualization of microarray data

Two different free software comprising of Mapman version 3.5.1R2 (Thimm et al., 2004) and VANTED version 2.1 were used for visualization of microarray data. Mapman is a free tool to display genomics and metabolic data sets onto diagrams of metabolic pathways and other biological processes. This tool uses a hierarchical system for classification and preparation of mapping data file. To map the data in corresponding metabolic pathway, the experimental data file which contained differentially expressed genes, was loaded into the Mapman using tab delimited file

(text). Metabolite overview and Ath-AGI-models-TAIR09.Jan09 were selected from image annotator module as metabolic pathway and mapping file, respectively. Furthermore, a file comprising of fold changes and modified gene ontology (GO) were loaded into the VANTED for functional classification of differentially expressed genes. A specific hierarchy tree was created after computing of all attributes and identification of two sets of up- and down-regulated genes. Tree layout (RT) algorithm was used to layout cluster graph. Moreover, Fisher's exact test (Fisher, 1987) was applied for evaluation of significant cluster at levels of less than 0.001.

4. Results

All spring barley accession lines from the IPK collection were delivered by Dr. Börner, and selected based on their striking genetic variability for plant height, tiller number and grain yield (GY) scored in 2009 (Excel File 1). Among them, only HOR9489 was excluded for further analysis due to failure in germination. In addition, the number of introgression lines provided by Dr. Pillen (University of Halle-Wittenberg, Institut für Agrar- und Ernährungswissenschaften, Germany) was also reduced from ten to four genotypes because most of the introgression lines except, S42IL107 and Scarlett did not show any remarkable differences in six measured morphological parameters (Excel File 1). Thus, the selection was carried out based on their genotypic information QTL(s) for thousand-kernel weight (S42IL109), yield, and grain per ear (S42IL122) (Schmalenbach et al., 2009). Other genotypes, which had no QTL for agronomic traits (S42IL131), or similar QTLs (S42IL118), or their QTLs were on different chromosomes (S42IL110, S42IL108, S42Il130 and S42IL118) were excluded for further investigations. The remaining genotypes were grown under field conditions, and analyzed in detail by molecular and biochemical means to investigate their correlations and relationships to shoot biomass at the vegetative stage, to seed filling and flag leaves during generative stage and to final mature seeds at ripening stage.

4.1. Investigation of relationships between metabolites and total shoot biomass at the vegetative stage

The effect of genetic variability on shoot biomass

Total shoot biomass which was expected serve as an indicator for plant biomass accumulation, was measured in the field in 2010. The TSB for twelve genotypes varied between 15.30 g for S42IL107 and 32.80 g and 32.14 g for HOR216 and Scarlett, respectively (Figure 3). The TSB was two-fold higher in the plant with highest shoot biomass (HOR216 and Scarlett) compared to the plant with lowest biomass (S42IL107). For further analysis, three intermediate genotypes, S42IL119, HOR4555, and HOR1645 were included. A list of all measured phenotypical data is provided in Excel File 1.

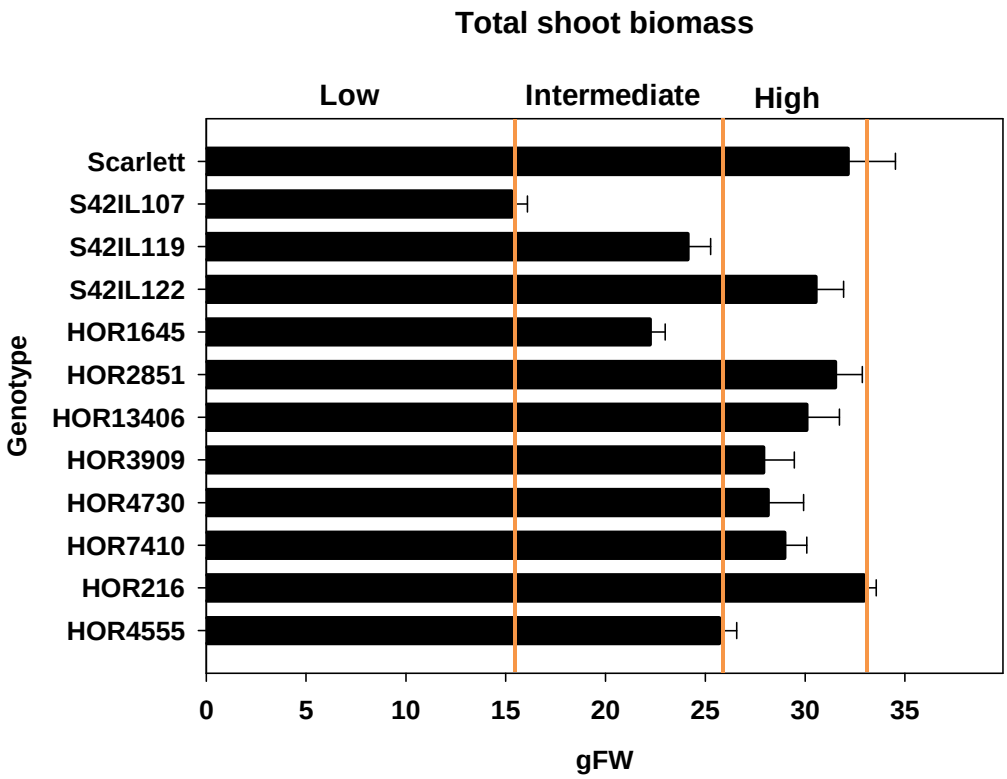


Figure 3. Genetic variability of total shoot biomass for twelve genotypes grown under field conditions. Genotypes are divided into three levels of low, intermediate, and high biomass production based on total shoot biomass. Bars show means of 12 independent plants and standard error (n=5).

4.1.1. Influence of metabolite composition and enzyme activities on shoot biomass

Non-targeted metabolite profiling of tiller leaf using GC-MS deconvoluted approximately 218 putative compounds per sample. In total, fifty-four out of 218 peaks were assigned into general compound classes based on representative masses, and seven peaks as unknown. The unknown metabolites and unassigned putative compounds were excluded from further analysis. Moreover, targeted metabolite profiling was used to identify crucial metabolites of primary metabolism. In total, 75 different metabolites were annotated that included amino acids, organic acids, mono- and disaccharides, sugar alcohols and starch. In addition to the metabolites, the maximum activities of 21 enzymes of central metabolism were measured to investigate their relationships with their substrates or products and morphological traits.

4.1.2. Network correlation analysis between intermediates of various pathways and their corresponding enzymes

Network correlation analysis was performed to identify the correlation between metabolites and enzyme activities and their relationships in correlation matrices. The r-values of correlations between pairs of metabolic traits are provided in Excel File 2. The results of Pearson's product-moment correlation are summarized in Appendix Table 4 and are described in details in the following sections. The ranges of significance levels of Benjamini-Hochberg, pBH was between 0.05 and 0.0001 were used to extract all significant pair-wise correlations of enzymes-to-enzymes, metabolites-to-metabolites and enzymes-to-metabolites (Appendix Table 4).

Evaluation of the relationships among metabolites from different metabolic pathways

Correlations between metabolites

The analysis of metabolite-to-metabolite relationships was performed to investigate the correlations between metabolites in the same metabolic pathway and the interconnections of metabolites between different sectors of central metabolism. The result of correlation analysis was 330 significant positive and 168 significant negative connections (Appendix Table 4 and Excel File 2). This was about 9.6 correlations per metabolite (pair of significant correlation / number of metabolic traits). These correlations were most obvious for TCA cycle intermediates with the strongest correlations among citrate (Cit), malate (Mal), 2-oxo, and isocitrate (Isocit) (pBH<0.0001), aliphatic amino acid, Valine (Val), and particulary, aromatic amino acids, phenylalanine (Phe) and tyrosine (Tyr) (Figure 4, Boxes A and C).

Glycolytic metabolites comprising of Fru6P, Glc6P, PGA and phosphoenolpyruvate (PEP) showed positive relationship with each other and with cis-and trans-aconitate (Tcant and Cacnt), Isocit and Mal (pBH< 0.001). Positive correlations were further observed among Suc, Glc, Fru and starch (pBH<0.0001). Glcucose and galactose (Gal) also showed negative correlations with TCA intermediates and glycolytic metabolites (pBH<0.001) (Figure 4 and Excel File 2). Interestingly, nucleotides, sugar nucleotides, sugar phosphates, and amino acids were

negatively correlated with triethanolamine (TEA) which is involved in glycerophospholipid biosynthesis, fatty acid derived compounds such as octadecanoic acid (Odca), hexadecanoic acid (Hxda), and dodecanoic acid (Ddca)(Figure 4, Box B). The results indicated that metabolites involved in glycolysis, TCA cycleand aromatic amino acid pathwayare highly connected to each other and a change in one pathway affects the production of the others.

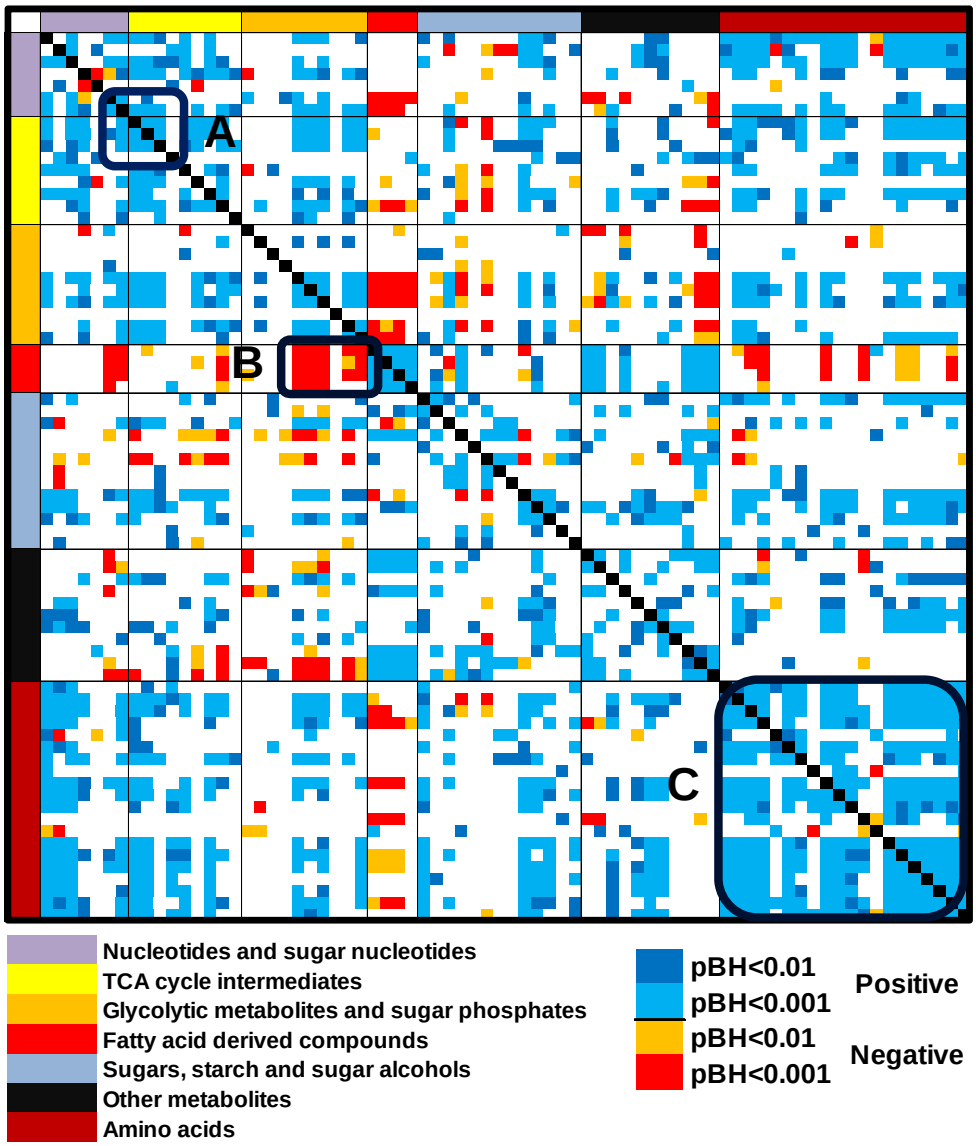


Figure 4. Heat map generated from the correlated metabolites of tiller leaves. Pearson correlation coefficients are shown in red and yellow colors as negative and dark and light blue as positive correlations. The black boxes indicate positive correlation between TCA cycle intermediates (A), and between amino acids (C).Negative correlation is indicated between fatty acid derived compounds and sugar phosphates (C).

Correlations among metabolites and amino acids

The correlation analysis between metabolites and amino acids revealed a high positive connectivity (Figure 5). Interrelated connections were found for Glu, proline (Pro) and arginine (Arg), aliphatic and aromatic amino acids, which positively correlated with glycolytic and TCA cycle intermediates and nucleotides (pBH<0.001). Among the TCA cycle intermediates, Isocit, Cit, Mal, OAA, 2-oxo, succinate (Succ), Cacnt indicated the largest number of connections to amino acids particularly to glutamine (Gln), Glu, Pro and Arg and aliphatic amino acids such as isoleucine (Ile), Val, leucine (Leu), and glycine (Gly), serine (Ser), Tyr and Phe (pBH<0.01 and pBH<0.001). Further strong correlations were found between nucleotides, ATP, AMP and ATP, two sugar nucleotides, ADPGlc and UDPGlc to most amino acids, (pBH<0.01 and pBH<0.001) (Figure 5). Sugar phosphates, Fru6P, Glc1P, Glc6P, and sucrose-6-phosphate (Suc6P) also positively connected to most amino acids (Figure 5).

Aliphatic amino acids such as Val, Ile and lysine (Lys) showed strong negative connections with Odca, Hxda (pBH<0.001) and with glycolytic compounds, PEP and PGA. Furthermore, Glc and Gal were negatively and strongly connected to Gln, Glu, Pro, Arg, Tyr, and Phe. Gln, Glu, Pro, Arg, along with other 10 amino acids except gamma aminobutyric acid (GABA), aspartate (Asp), methionine (Met), histidine (His) and 1-aminocyclopropane-1-carboxylic acid (ACC) correlated with 2-oxo, which is the carbon acceptor in the glutamine oxoglutarate aminotransferase (GOGAT) pathway (Figure 5).

Inconsistent with the observation of Sulpice et al. (2010) Suc and starch which are two main carbohydrate formed during photosynthesis were positively connected to amino acids (Figure 5, Suc, 15 and starch 2 connections). In summary, a large number of correlations, particularly, between ADP, ADPGlc, and ATP, four TCA intermediates, Isocit, Cit, 2-oxo, and Succ, two sugar alcohols, glycerol and inositol (Inost) with amino acids indicates that certain metabolites from central metabolism were strongly correlated indicating that their levels may play crucial roles in regulation of specific functions in vegetative stage in barley.

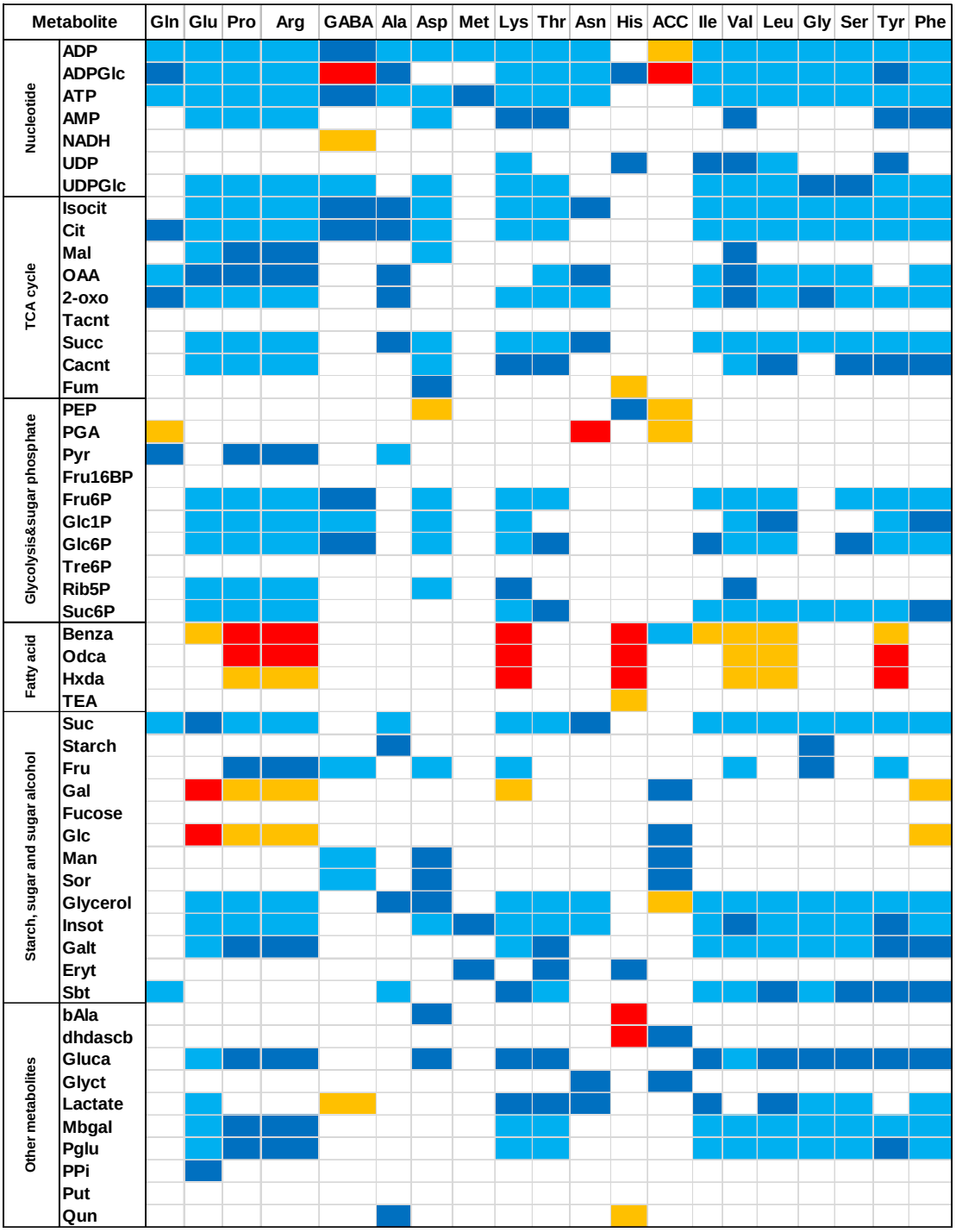


Figure 5. Heat map of bilateral Pearson correlations generated between metabolites and amino acids primary metabolism 30 day after planting (vegetative stage) under field conditions. The significance of associations at the levels of pBH<0.01 and pBH<0.001 are indicated with dark and blue colors for positive and yellow and red colors for negative correlations.

Elucidation of relationships between different enzymes and enzymes to metabolites in tiller leaves of barley

Relationships among enzymes

A correlation matrix was constructed to calculate the degree of connectivity between enzymes in tiller leaves. The results revealed 116 direct connections between 21 enzymes involved in different metabolic pathways (Appendix Table 4). Color-coding in Figure 6 was used to distinguish between positive and negative correlations, and to denote significances at a level of pBH<0.001. Consistent with the findings of Keurentjes et al. (2008) the correlation matrix revealed a large number of positive relationships among enzymes of different metabolic pathways indicating a strong co-regulation. As shown in Figure 6, relationships between enzymes from sucrose metabolism with TCA cycle, glycolysis, and the pentose phosphate pathway and enzymes of the Calvin cycle could be observed.

Enzymes of sucrose metabolism showed diverse correlations. PGI, PGM, Susy, CWInv and AcidInv showed positive correlations in the correlation matrix while CytInv showed only negative correlations with AspAT, cFBPase, PK, PGI, 6PGDH and pFBPase. Two enzymes, AspAT, involved in nitrogen metabolism and PGI occurring in sucrose metabolism showed the highest number of connections in the correlation network. Susy and AcidInv, two sucrose-degrading enzymes with two and six connections, were weakly connected enzymes in the correlation matrix (Figure 6). These result showed that sucrose-degrading enzymes interact with specific enzymes involved in glycolysis and nitrogen metabolism specifically AspAT which connect the the carbon and nitrogen metabolism.

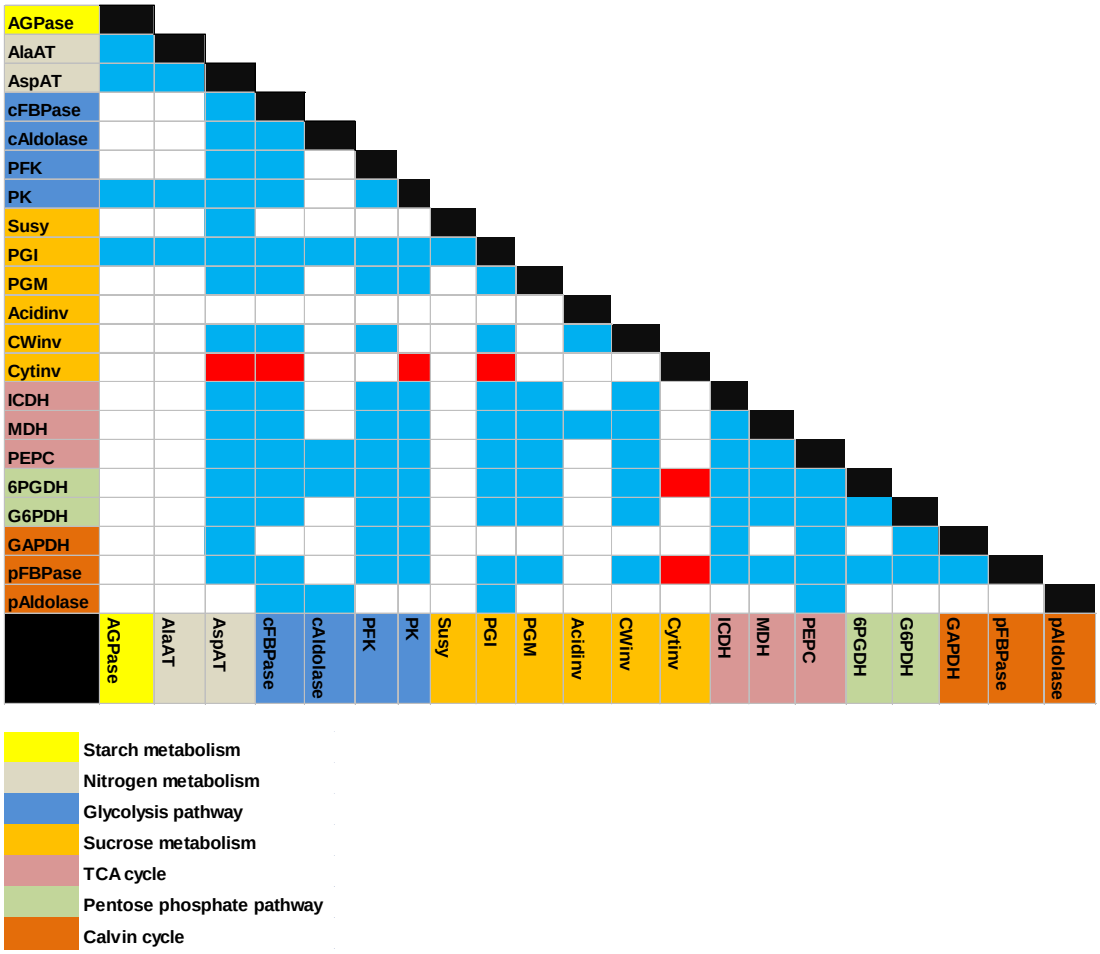


Figure 6. Heat map generated from the correlated enzymes of tiller leaves. The heat map of correlation matrix shows the significance of the Benjamin-Hochberg corrected p-values of the Pearson correlations at the level of pB<0.001. Color code indicates the association between the enzymes. Blue and red colors depict positive and negative associations, respectively.

The interrelations between enzyme activities and metabolites

The network correlation analysis between 75 detected metabolites and 21 enzymes revealed a close interaction between metabolites and enzymes of different pathways with 255 direct correlations at the level of 0.05 (Appendix Table 4). Figure 7 shows a summary of the correlation network of metabolites and enzymes. Dark and light blue colors show positive, and red and yellow colors negative associations. Inconsistent with the observation of Sulpice et al. (2010) and Keurentjes et al. (2008) metabolites and enzymes showed a large number of positive correlations than negative correlations indicating that metabolites are in the highest excess in tiller leaves at the vegetative stage. Two enzymes, cFBPase

and AspAT showed the largest number (36 connections) of connections with metabolites.

Among the metabolites of nucleotide metabolism, ATP (21 connections) ADP (21 connections) and ADPGlc (14 connections) were strongly connected to almost all enzymes (Figure 7) indicating their participation in many reactions throughout the network at the vegetative stage. Sucrose metabolizing enzymes, Susy, CWInv and CytInv showed strong correlation with the soluble sugars, Glc and Fru that are direct degradation products of the enzymes. ICDH also correlated with Isocit and Cit being the substrate or product of the appropriate enzyme (Figure 7). AspAT which is responsible for the conversion of aspartate and 2-oxo to OAA was positively correlated to Glu, 2-oxo, and Asp (Figure 7). However, PGA, ribose-5-phosphate (Rib5P), Cacnt and Glc6P were negatively correlated with enzymes of sucrose metabolism, glycolysis and TCA cycle indicating a notable tradeoff between metabolites and enzymes activities (Figure 7). The results demonstrated that enzymes from primary metabolism are regulated by their direct precursors and/or subsequent products.

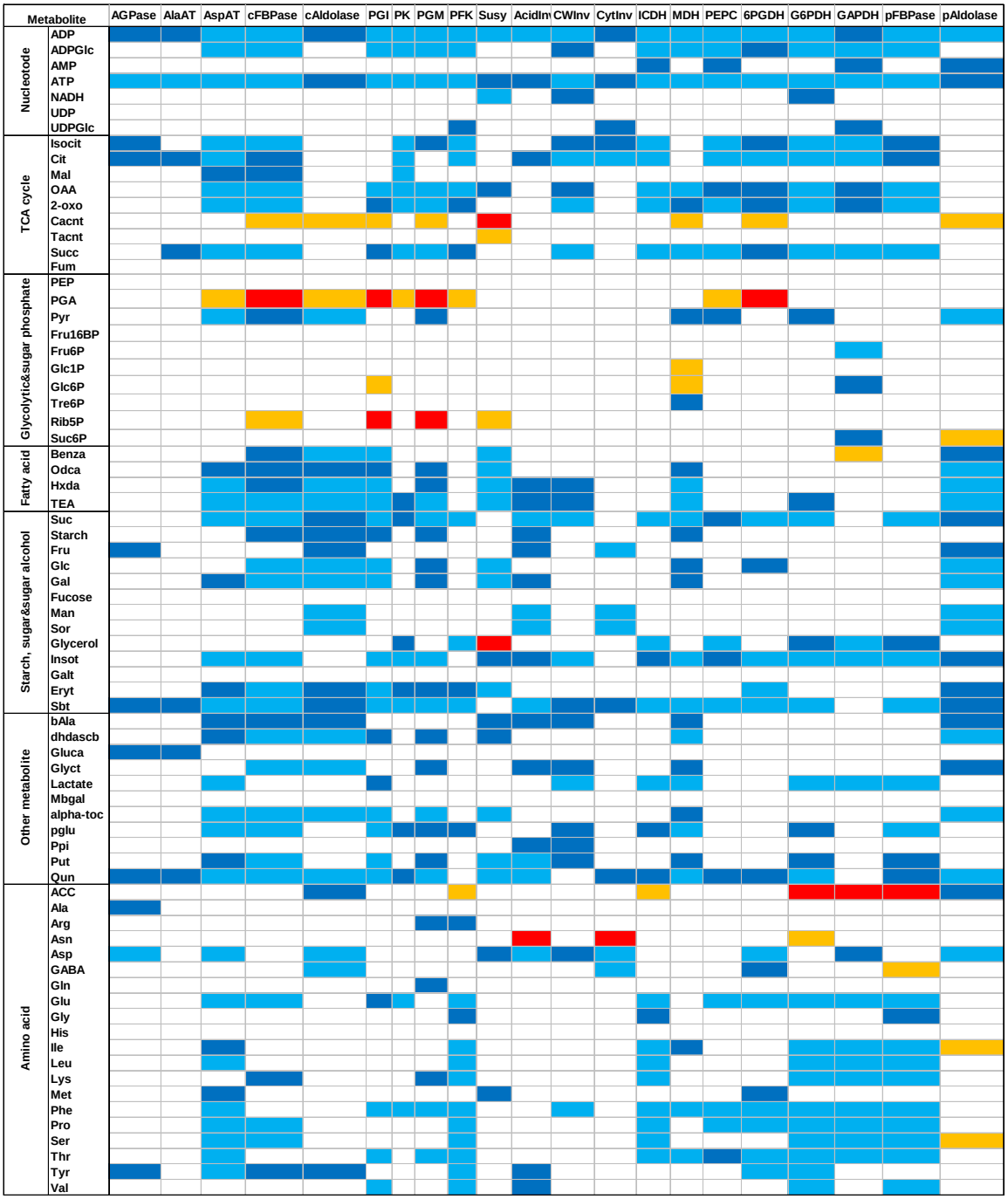


Figure 7. Heat map of a bilateral Pearson correlations generated between enzymes and metabolites of primary metabolism 30 day after planting (vegetative stage) under field conditions. The significance of correlations at the levels of pBH<0.01 and pBH<0.001 are indicated with dark and blue colors for positive and yellow and red colors for negative correlations.

4.1.3. Interaction between metabolites, enzyme activities and phonotypical traits

The correlation matrix between total shoot biomass, 75 metabolites, and 21 enzymes was generated by performing a pair wise correlation analysis. A detailed list of correlation is provided in Excel File 2 and as scatter plot for the strongest correlated metabolites and enzymes to TSB in Figure 8. The numbers of significant connections were confined to metabolites, enzymes, and amino acids, which passed multiple testing corrections with Benjamini and Hochberg (pBH) **F**alse **D**iscovery **R**ate (FDR) at the level of 0.0001, 0.001, and 0.05.

Relationship among metabolites, enzymes activities and total shoot biomass

A detailed correlation analysis was performed between metabolites and total shoot biomass at the vegetative stage to identify metabolites that might be crucial for the determination of shoot biomass. 19 positive and one negative relationship were found (Excel File 2). The strongest positive correlation to shoot biomass was observed with the metabolites 2-oxo, Pyr (Pyr), Succ, ADP ATP, ADPGlc, and AMP. Further strong correlations were identified for Odca, Hxda, TEA, alph-tocopherol (alpha-Toc), Suc, Sbt and Inost. In addition, three amino acids, Phe, asparagin (Asn), and Met showed positive and Lys negative associations with TSB. For all other metabolites, no significant correlation could be found (Figure 8 and Excel File 2).

Pair-wise correlation analysis also resulted in seven positive correlations between enzyme activities of primary metabolism and total shoot biomass (Figure 8 and Excel File 2). The highest connectivity was found for Susy and AspAT that are involved in sucrose and nitrogen metabolisms, respectively (Figure 8). Further relationships to total shoot biomass was identified for enzymes participating in carbohydrate metabolism including PGM, cFBPase, G6PDH, 6PGDH and MDH (Excel File 2).

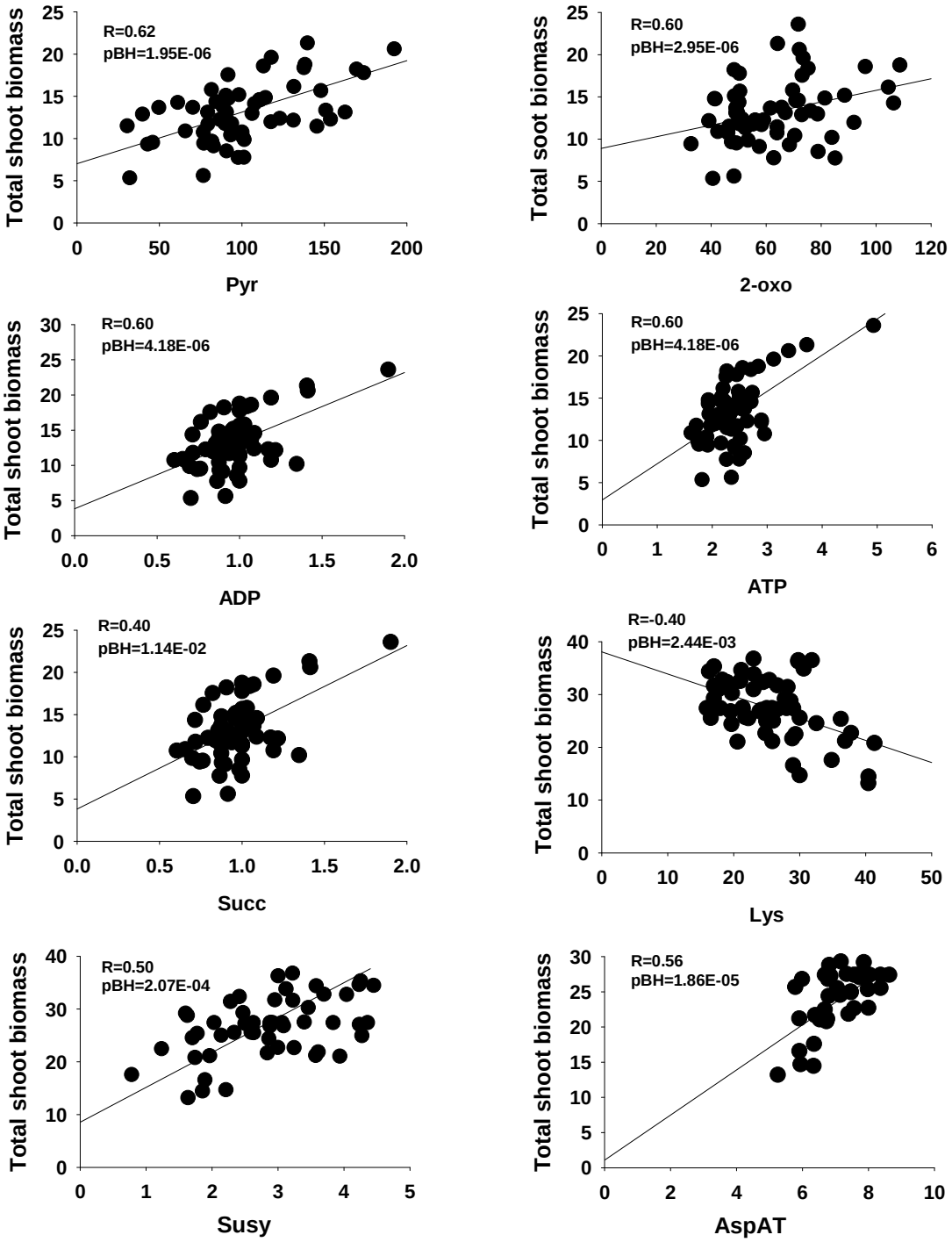


Figure 8. Scatter plots of relationships between selected metabolites and enzymes and total shot biomass at the vegetative stage and under field conditions. The significance of correlations at the levels of pBH<0.01 and pBH<0.001 and r-values greater than 0.40 are given in the sections The units are nmolgrFW^{-1} for all metabolites, μmol grFW^{1} min^{-1} for enzyme activities, and gr for total shoot biomass.

4.2. Investigation of phenotypical traits and their relation to biomass in barley plants at the generative stage

Total shoot and seed weight of twelve selected lines grown under in field conditions were measured as phenotypical parameters. A complete list scored phenotypical data is provided in Appendix Table 5. Among the genotypes, HOR216 and S42IL107 had the highest (58.2 gr) and lowest (15.9 gr) TSB, respectively (Appendix Table 4) which was 3.6-fold higher in HOR216 compared to S42IL107. Among six morphological parameters, TSB and TSW were selected as morphological parameters for the identification of metabolites, enzymes, and transcript signatures that were relevant for shoot and seed biomass at the generative stage.

Targeted and non-targeted metabolite profiling of flag leaves and seeds resulted in the identification of 87 and 83 metabolites, respectively. These metabolites covered the most important metabolic routes of primary carbon and nitrogen metabolism. Furthermore, maximum activities of twenty-two enzymes from carbon and nitrogen metabolisms were measured to investigate the relationship between enzymes, metabolites, and their correlations to total shoot and seed biomass.

4.2.1. Network correlation matrices of metabolites and enzyme activities

The correlation matrices were constructed based on 87 metabolites and 22 enzymes along with total shoot biomass. Pearson product-moment correlation was used to elucidate the interaction among metabolites to metabolites, among enzymes and between metabolites and enzymes. The relationship of the interactions is described in detail in the following sections.

Identification of metabolites crucial for biomass formation in flag leaves

The analysis of metabolite-to-metabolite correlations led to 1375 significant relationships at the levels of 0.05 and 0.001 with an average of 15.80 correlations per metabolite. The number of correlations (metabolite-to-metabolite) is indicated in the Figure 9 by large for high interactions and small node for weak interactions (number of positive or negative relationships to other metabolites). As shown in Figure 9, metabolites occurring in the TCA cycle or glycolytic pathway showed the

highest. A detailed analysis of metabolite-to- metabolites relationships indicated that Pyr and Tacnt had the highest number of correlations among glycolytic and TCA cycle compounds (Figure 9).

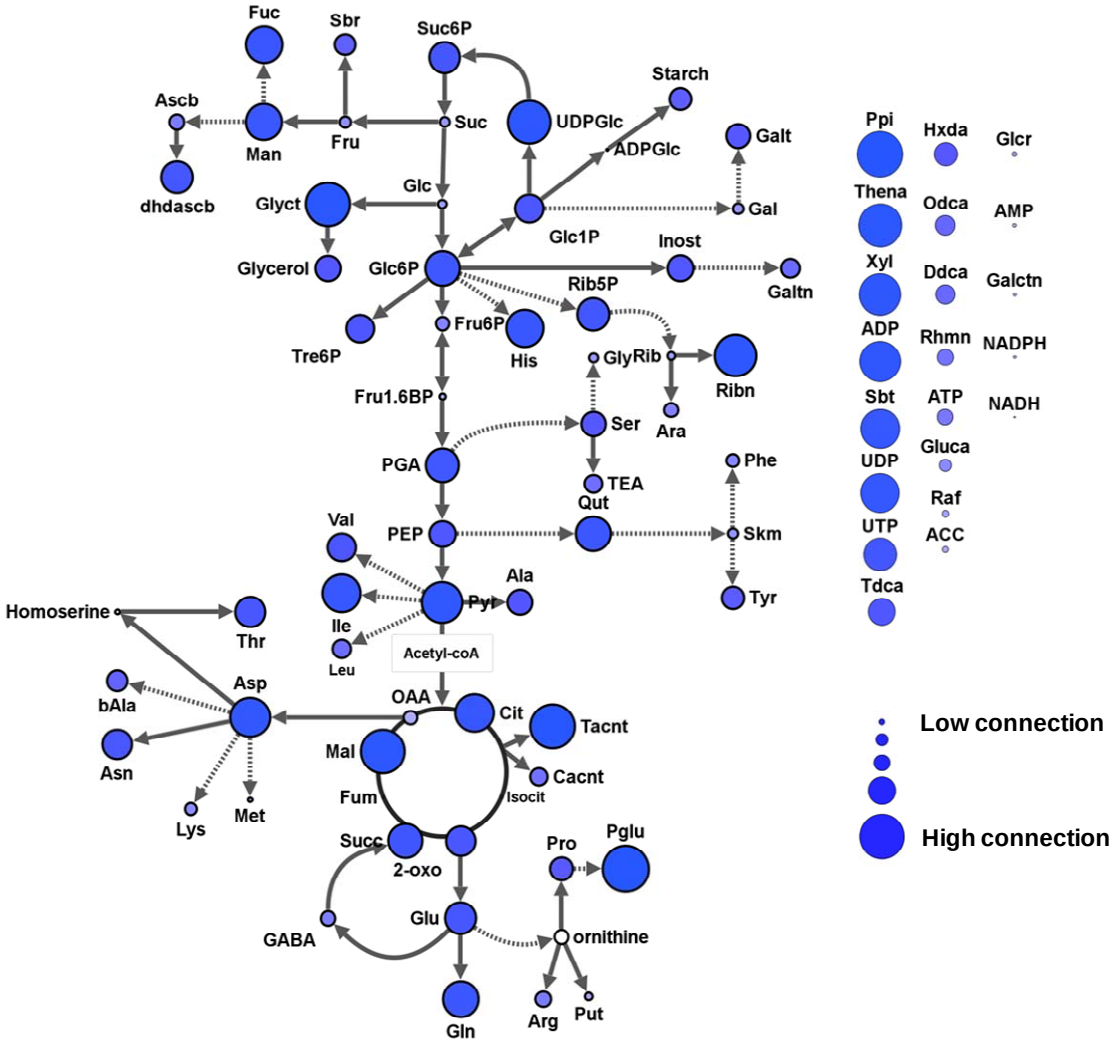


Figure 9. Relationships between metabolites at the generative stage in flag leaves under field conditions as shown in node degrees that were calculated based on the number of connections of a metabolite to other metabolites with a significance at the levels of pBH<0.01 and pBH<0.001. The size of the circles indicates the degree of connections. All abbreviations are defined in appendix.

The correlation network of metabolites to amino acids showed that TCA cycle intermediates, glycolytic metabolites and nucleotides highly correlated to amino acids (Figure 10) indicating a strong co-regulation of different metabolites may control amino acids levels in generative stage. Aliphatic amino acids like Val and Ile showed strong correlations with TCA cycle intermediates, glycolytic compounds and sugar phosphates. Moreover, 2-oxo, a TCA cycle intermediate was strongly correlated with Gln, Glu, Asp, Threonine (Thr), His, Ile and Val (Figure 10).

The negative connections were mainly found for sugars, Glc, Fru, Gal, Man, Sorbose, ribose (Rib), and amino acids. Among them, Man showed the highest number of negative correlations with seven amino acids, Glu, Gln, Pro, Thr, Asn, leu and Gly. Two major carbohydrates formed during photosynthesis, starch, and Suc poorly connected to amino acids whereas Glc and Fru showed strong relationships to Pro, Arg, Ala and Val, GABA. The results showed that TCA cycle intermediates and glycolytic metabolites strongly co-regulated with almost all metabolites and particularly with amino acids which provide nitrogen moiety for subsequent use in growth.

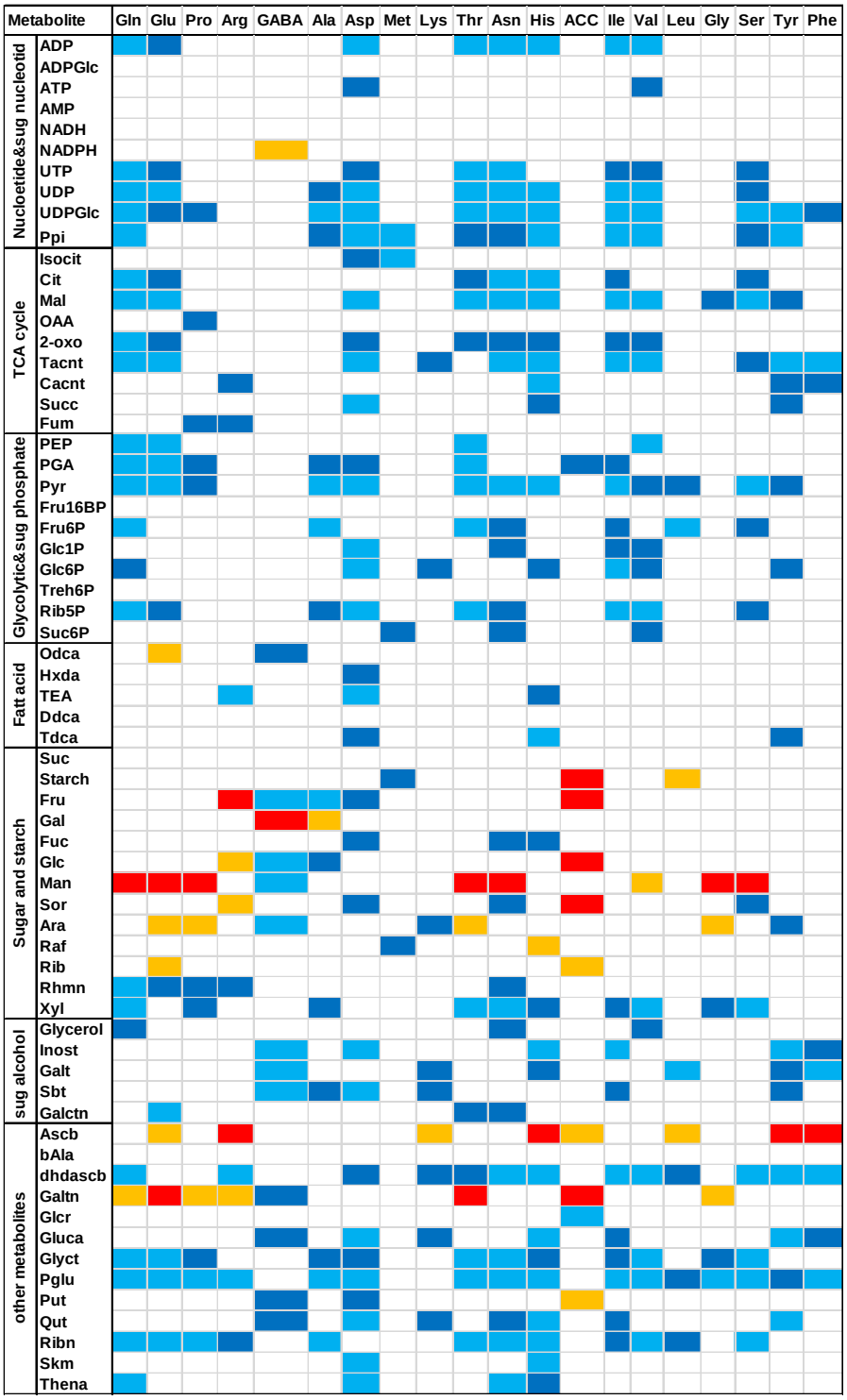


Figure 10. Heat map of a bilateral Pearson correlations generated between amino acids and metabolites of primary metabolism 30 day after planting (generative stage) under field conditions. The significance of correlations at the levels of pBH<0.01 and pBH<0.001 are indicated with dark and blue colors for positive and yellow and red colors for negative relationships.

Relationships between enzymes in primary metabolism

Network correlation analysis was generated to investigate the connectivity among enzymes and to check whether enzymes correlated with the enzymes in the same metabolic pathway than enzymes in the other pathways. The correlation analysis based on Pearson r coefficient revealed 119 significant relationships for enzymes at the levels of 0.05 and 0.001 (Figure 11). Consistent with the observations of Steinhauser et al. (2010) enzymes activities largely correlated in correlation matrix at the generative stage. The most highly connected enzymes were AlaAT, PK, PFK, and PGM (16 connections). Poorly connected enzymes also were cFBPase and GAPDH (two connections).

Enzymes of sucrose degradation pathway showed either positive or negative correlations together and with other enzymes. Among them, Susy was negatively correlated to AGPase, AlaAT, PFK, PK, PGI, PGM, MDH, G6PDH, pFBPase and pAldolase whereas AcidInv showed one negative connection to AGPase, in the correlation networks indicating markedly changes of correlation core set among sucrose-metabolizing enzymes during different stages (Figures 11 and 6). AspAT and AlaAT, two enzymes occurring in nitrogen metabolism showed strong correlation with G6PDH, PK, PGI, and PFK. Further strong relationship was found for GOGAT, ICDH, PEPC and PK ($pBH<0.001$; Figure 11). In summary, the obtained results indicated that PGM, two glycolytic enzymes, PK and PFK, followed by AlaAT were the key enzymes with high connections to the enzymes from different metabolic pathways.

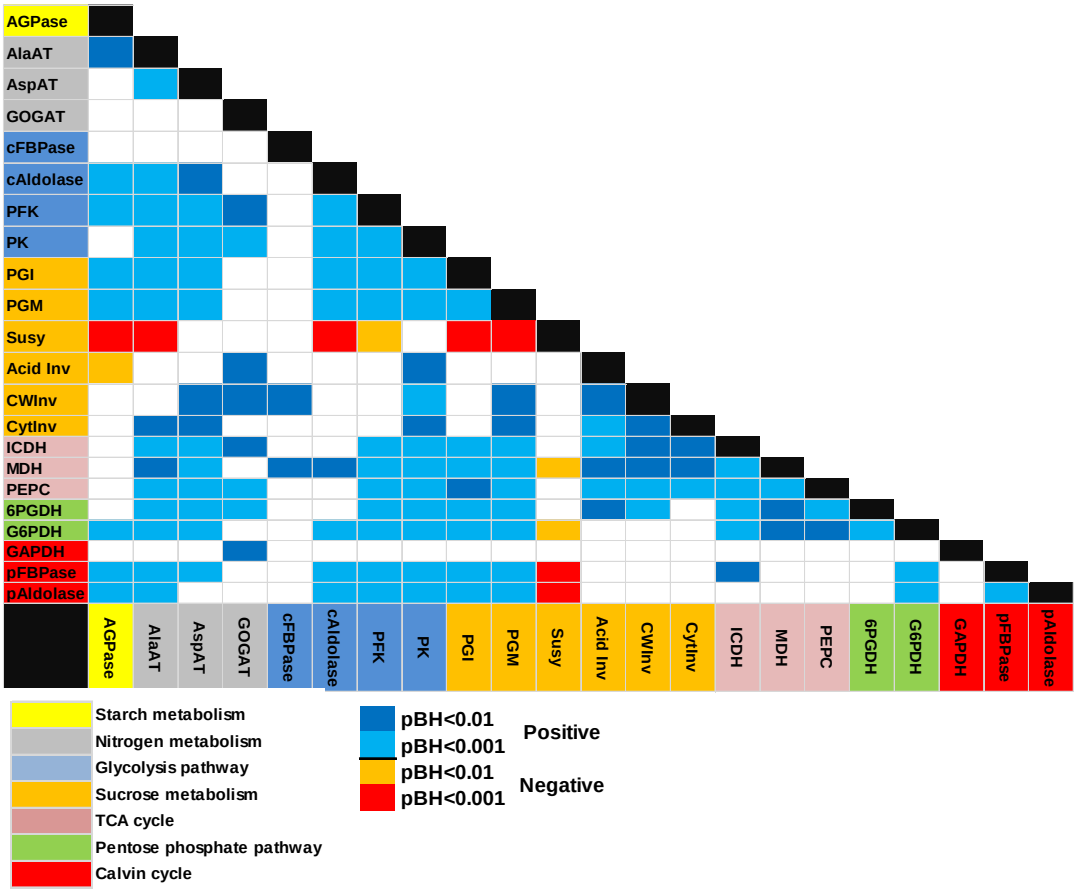


Figure 11. Heat map generated from the correlated enzymes of flag leaves. The heat map of correlation matrix shows the significance of the Benjamin-Hochberg corrected p-values of the Pearson correlations at the level of pB<0.001. Color code indicates the association between the enzymes. Blue and red colors depict positive and negative correlations.

Correlation analysis between enzyme activities and metabolites of primary metabolism

The correlation analysis between 22 enzymes and 85 metabolites are illustrated in Figure 12. Enzymes were selected from different metabolic routes of central metabolism and were correlated to soluble sugars, sugar phosphates, TCA intermediates, sugar alcohols, fatty acid precursors, nucleotides, and amino acids. Consistent with findings of Steinhauser et al. (2010) there were a low number of correlations of enzymes to metabolites. However, inconsistent with mentioned observations, the majority of these correlations were positive. Regarding the enzymes, AGPase was found to possess the highest number of negative correlations with metabolites indicating a distinguished tradeoff between a wide set of metabolite levels and AGPase activity (Figure 12).

Among the metabolites, amino acids, Thr, Asn, Gly, Ser, Val, Glu, Gln, Ala, and Pro showed the largest number of connections (up to 14 connections) to enzymes (Figure 12). Two enzymes of nitrogen metabolism, AlaAT and AspAT positively correlated with major amino acids Glu and Gln. Calvin cycle enzymes, pAldolase and pFBPase also had strong correlation with most amino acids, PEP, PGA, and Mal. A series of positive correlations was also found for enzymes and their substrates or products. There was a positive interaction between PK and Pyr ($pBH<0.05$), invertases and Fru or Glc, PFK and Glc6P, pFBPase and Fru6P, PGA and GAPDH and AlaAT and Ala. The strong negative correlation was observed for MDH and ICDH with several glycolytic compounds, including PEP, PGA, and Glc6P suggesting notable trade-off between TCA cycle and glycolysis (Figure 12).

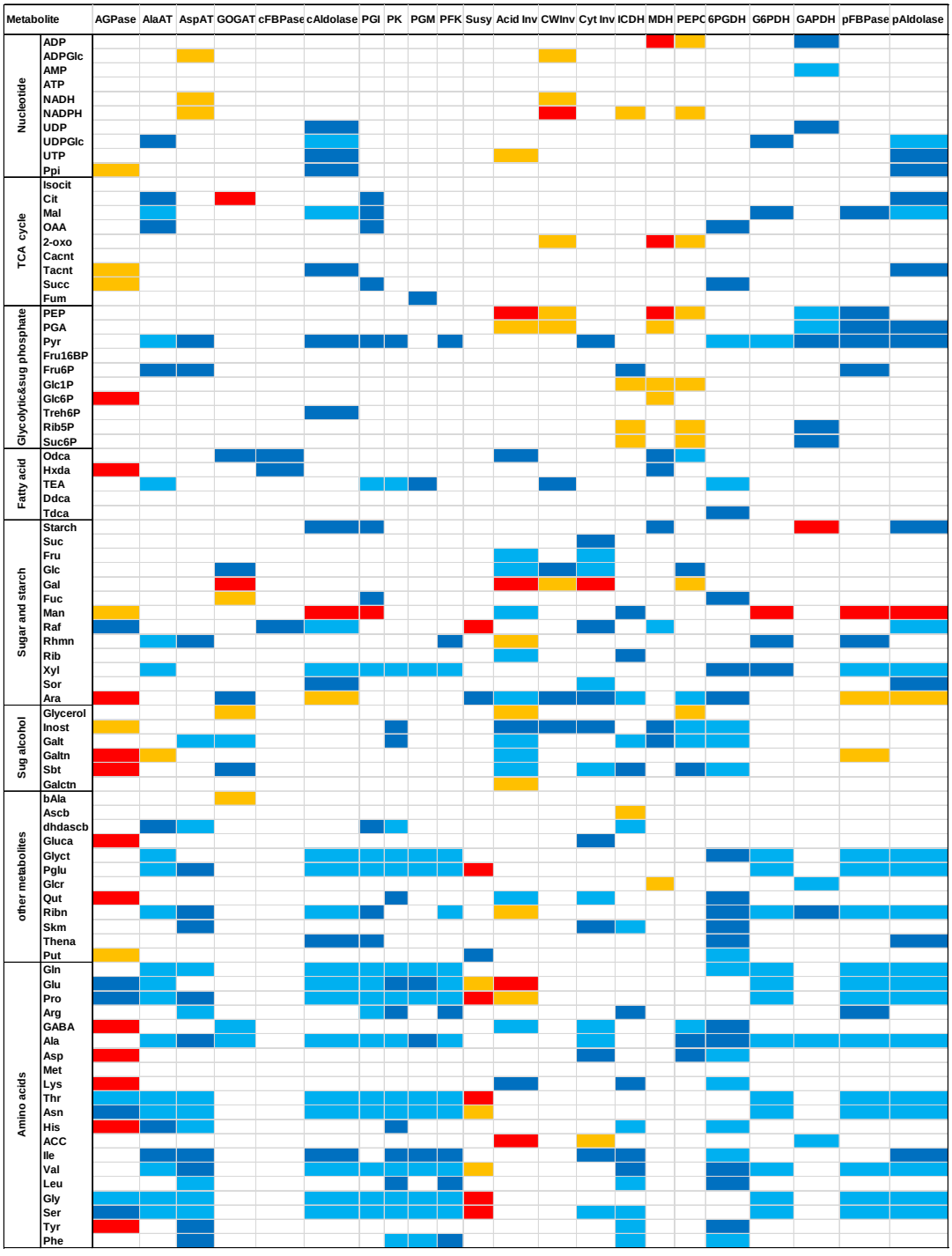


Figure 12. Heat map of a bilateral Pearson correlations generated between enzymes and metabolites of the flag leaves 15 days after anthesis at the generative stage. The significance of correlations at the levels of pBH<0.01 and pBH<0.001 are indicated with dark and blue colors for positive and yellow and red colors for negative correlations.

4.2.2. Interaction between metabolites, enzyme activities and shoot biomass

Influence of metabolites on biomass formation

In order to investigate whether metabolites and/or enzymes from various metabolic pathways show significant correlation to shoot biomass, a pair-wise correlation analysis was performed between detected metabolites and enzymes of central metabolism and plant shoot biomass. The basic matrix consisted of 87 metabolites, 22 enzymes and total shoot biomass. Significant correlations at a false discovery (FD) rate <1% and <5% are listed in Figures 13 and 14.

Metabolites with the strongest correlation to total shoot biomass were those involved in TCA cycle including Mal, Suc, Tacnt and Isocit. Moreover, TEA, a metabolite occurring in phospholipid biosynthesis, Ddca and Tdca, two fatty acids derived metabolites, PGA, Glc6P and Pyr, metabolites from glycolysis were positively correlated with total shoot biomass. In addition, starch and other intermediates such as Tre6P, Sbt and Inost as well as xylose (Xyl) and fucose (Fuc) exhibited strong connectivity to total shoot biomass. Among all metabolites, only Man that is involved in cell wall biosynthesis was negatively correlated to total shoot biomass (Figure 13). Further positive correlation to total shoot biomass was detected for Glyct, Pro-glutamate (Pglu) and quinate (Qut), an intermediate in the shikimate pathway., Furthermore, many amino acids including Asp, Gln, Val, Ile, Phe, Tyr, Ala, Thr, Gly, Pro, and His showed a strong positive correlation with total shoot biomass (Figure 13).

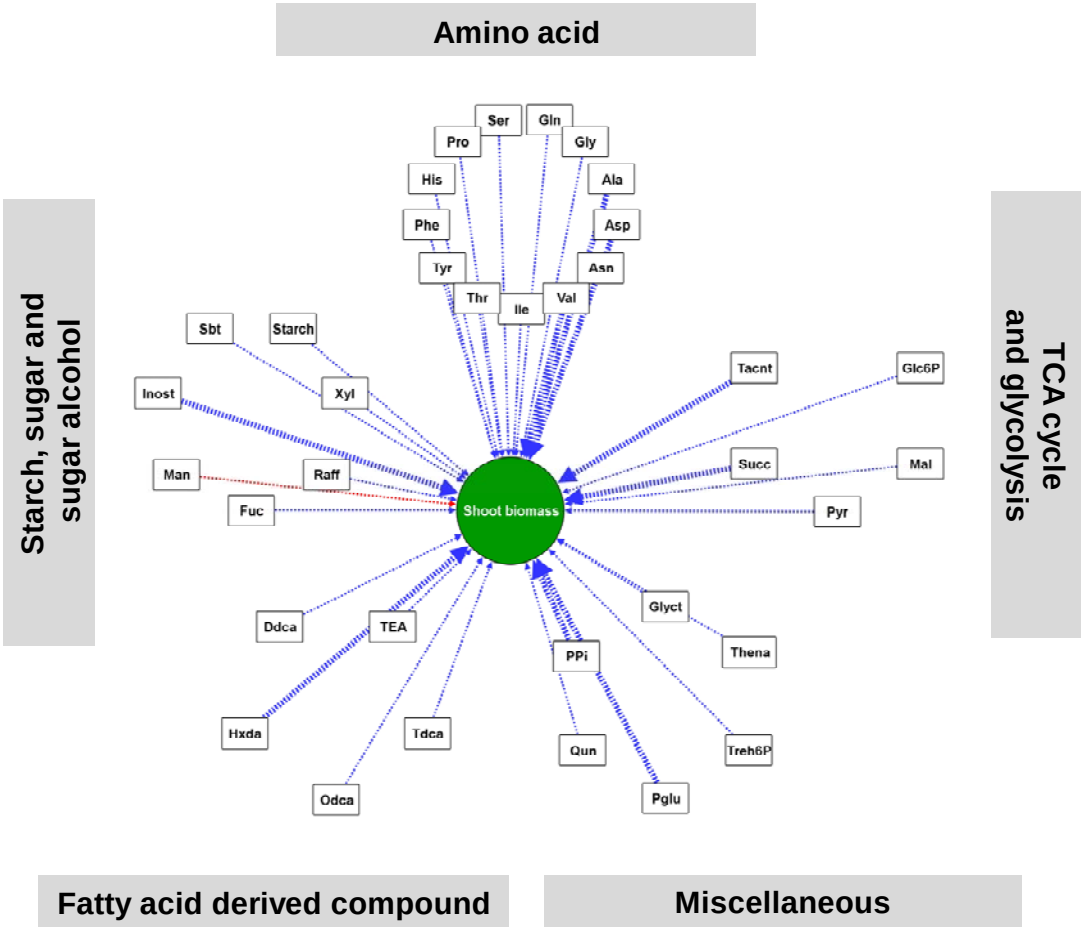


Figure 13. Metabolite signatures with high significant correlation to total shoot biomass at the generative stage, 15 days after anthesis in the leaves of spring barley. A Pearson correlation was applied to extract the correlations at the levels of 0.01 and 0.001. Metabolites with significant positive and negative relationships to total shoot biomass are shown with blue and red dotted lines respectively. Strong associations are indicated using thicker dotted lines.

Influence of enzymes on biomass formation

The correlation analysis between 22 enzymes and total shoot biomass revealed eight positive connections in the correlation matrix (Figure 14). Significant correlations were found for PFK, cAldolase, cFBPase, PK, PGM, PGI, 6PGDH, AlaAT and pAldolase, enzymes involved in various metabolic pathways (Figure 14).

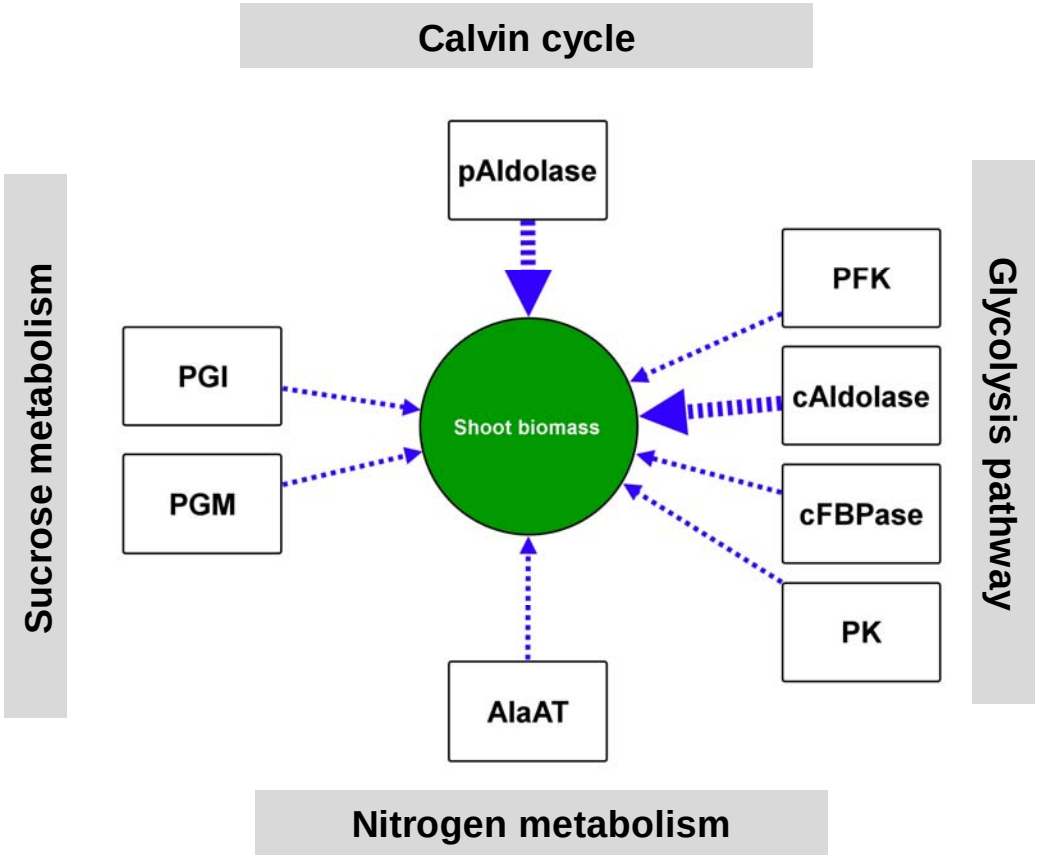


Figure 14. Enzyme signature for total shoots biomass at the generative stage, 15 days after anthesis in the leaves of spring barley. A pair-wise correlation analysis, which applies a correction for FDR, was used to extract significant metabolic traits at the levels of pBH<0.01 and pBH<0.001. Enzymes with significant positive association to shoot biomass are exhibited by dotted blue lines. Thicker dotted lines indicate strong correlations.

4.3. Identification of genes determining shoot and seed biomass in spring barley

In order to identify genes involved in different metabolic pathways which might determine the formation of shoot and/or seed biomass, a detailed trasnscriptome analysis was carried out in flag leaves and seeds of four selected genotypes of barley including HOR216 with high biomass, HOR4555 with intermediate biomass, S42IL107 with low biomass and Scarlett with intermediate biomass as internal control. To this end, a custom microarray Agilent platform was used that was

developed by Dr. Sreenivasulu at IPK Gatersleben and contained 56000 barley oligonucleotides. Results of differentially expressed genes and their expression correlations with emphasis on seed and shoot biomass production will be described in details in the following sections.

4.3.1. Annotation and functional classification of all identified genes

A comparison of sequence annotation with *Arabidopsis thaliana (*www.arabidopsis.org*)*, *Oryza sativa* (http://rice.plantbiology.msu.edu) and uniprot (www.uniprot.org) data bases with the search blast cut-of-value of 10^{-10} revealed that 6518 out of 12000 detectable entities putatively encoded particular proteins. MapMan ontology analysis, which groups probe-sets into hierarchical functional categories, termed bins, based on putative biological functions (Thimm et al., 2004) was used for functional categorization of the expressed gene by calculating the relative proportions of probe-sets in each bin. Hormone and lipid metabolisms with 130 (17%) and 133 (17%) unique genes and PS. Light reaction with 81(10%) genes constituted the largest portions of genes in their categories (Figure 15).

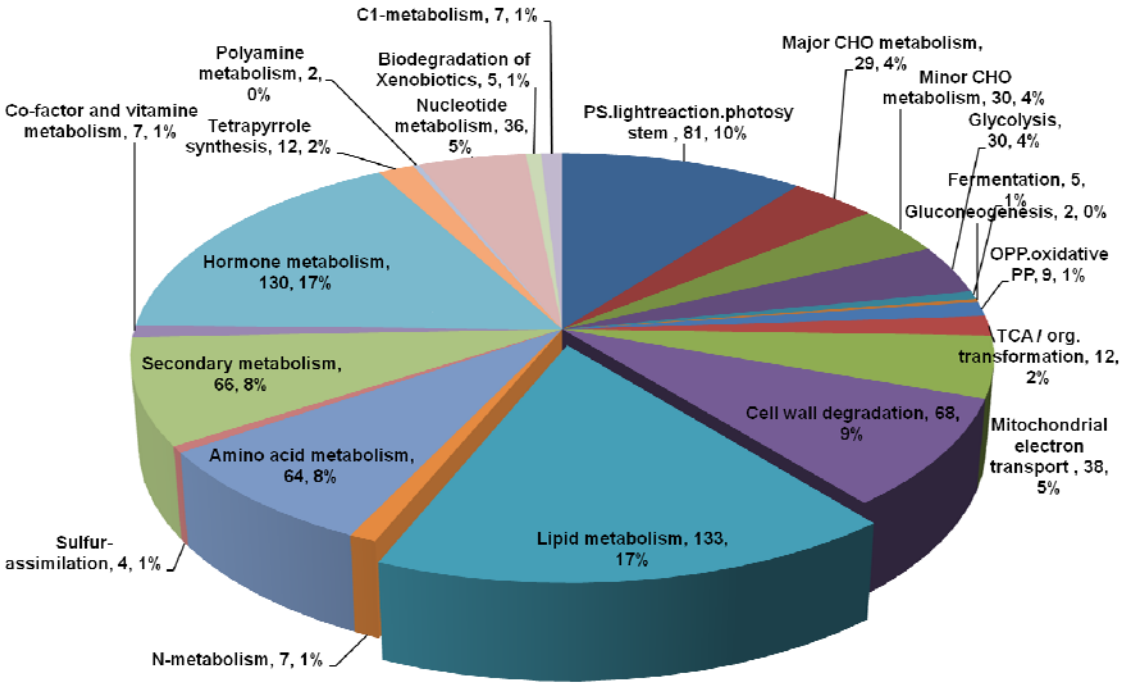


Figure 15. Classification of identified unigenes in barley assigned to bins in the overview using Mapman Gene Ontology (GO).

4.3.2. Identification of differentially expressed genes in the flag leaves

One-way ANOVA was applied to compare three contrasting genotypes, HOR4555, HOR216, and Scarlett with S42IL107 to identify genes differentially expressed. Using a FDR adjusted P value of 0.05 and filtering by at least 1.5-fold or greater changes, statistical analysis resulted in 3211, 3126, and 1176 differentially expressed genes in HOR216, HOR4555 and Scarlett compared to S42IL107. The complete table of differentially expressed genes is provided in Excel File 4. Based on MapMan ontology primary metabolism with an average of 21.6%, RNA and DNA metabolism with 16%, protein metabolism with 11% revealed the largest categories responsive with gene expression levels that differed among the genotypes (Appendix Figure 1 and Excel File 4).

4.3.3. Enrichment analysis of transcripts of co-expressed genes

Preferentially expressed genes of the plant with highest biomass (HOR216) compared to that with lowest biomass (S42IL107) were used to determine key genes that might be involved in the regulation of transcripts in primary metabolism in flag leaves. To this end, enrichment analysis described in the following sections was used to identify significant overrepresented clusters of genes in flag leaves. Fisher exact test was applied to extract significant overrepresented biological functions using GO terms and pathways at level of $p<0.0001$.

Functional enrichment analysis of genes involved in energy metabolism

Functional enrichment analysis was performed on 73 significant differentially expressed genes (section 4.3.2 and Appendix Figure 1) of plant with highest biomass compared to lowest biomass. These transcripts were putatively corresponding to genes participating in dark and light reactions, photorespiration and mitochondrial electron transport chain of energy metabolism. Visualization was carried out using VANTED with a modified Mapman function ontology (Figure 16). Fisher exact test analysis revealed that transcripts of two active clusters of the light reaction and the Calvin cycle were significantly overrepresented ($p<0.0001$) (Figure 16). Out of 37 up-regulated transcripts, 13 genes were detected in the Calvin cycle. Among them, genes coding for transketolase (Cross et al.), pFBPase, TPI,

seduheptulose bisphosphatase (SBP) and phosphoglycerate kinase (PGK) showed up-regulations. In contrast, Fructose-1,6-bisphosphate aldolase (FBP-Ald) and ribose-5-phosphate isomerase were down-regulated. Furthermore, the cytochrome-C-oxidase complex with six down-(up to 10-fold) and three up-regulated genes (up to two fold) showed the largest number of significant overrepresented genes in energy metabolism.

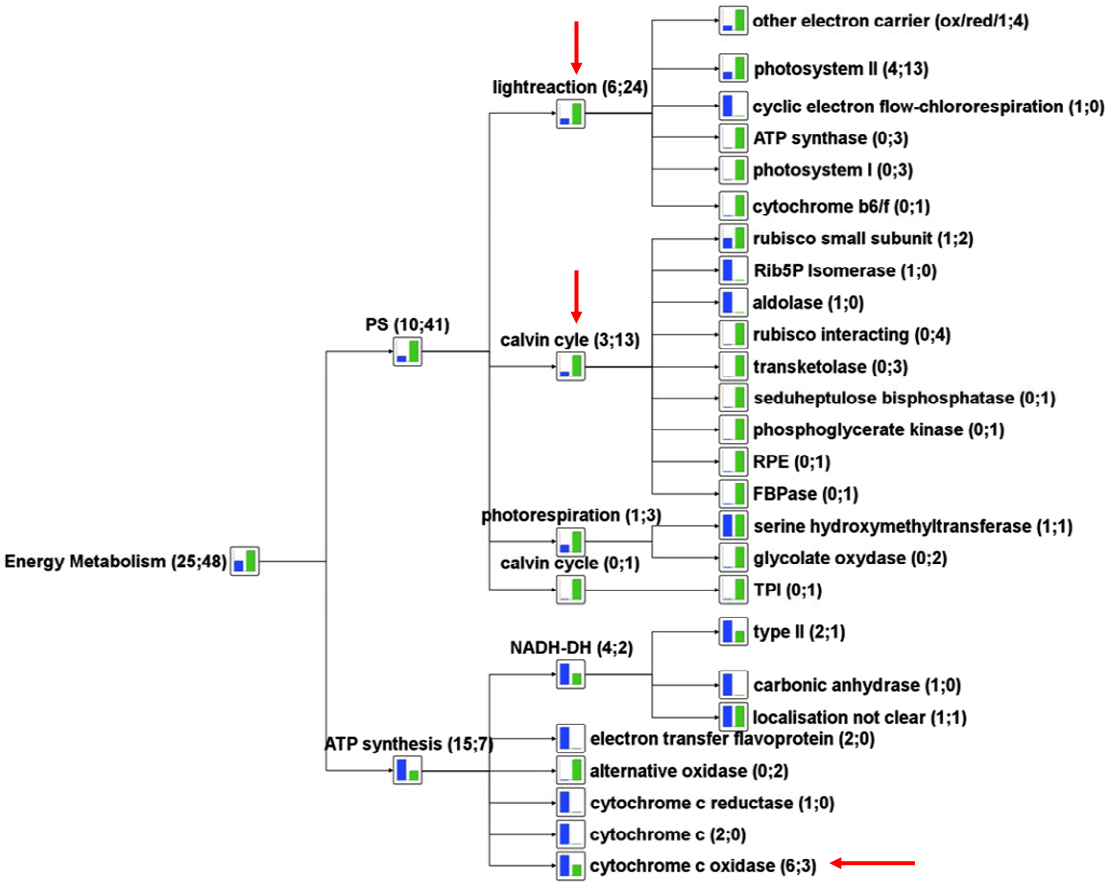


Figure 16. Categorization of genes with significantly different expression levels in line with highest biomass compared to lowest biomass using Gene ontology terms and pathways. In terms of gene ontology, energy metabolism is divided in two main bins comprising of ATP synthesis, and photosynthesis. Photosynthesis bin is containing three sub-bins such as light and dark reactions and photorespiration. Differentially expressed genes are imported in Vanted and clustered based on their functions. Three levels for energy metabolism were defined. The first level is containing two clusters while the second and third Levels comprise of 10 and 21 sub-clusters. In all clusters, green and blue colors show up- and down-regulation, respectively. Red arrows show significant overrepresented genes at the level of pBH<0.001.

Determination of differentially expressed genes involved in primary and secondary metabolism

In order to find out which genes and which specific pathways might contribute to the delivery of assimilates to the seeds after anthesis in barley, differentially expressed genes in the plant with the highest biomass formation compared to that with the lowest biomass formation, were used for the enrichment analysis using MapMan ontology (Figure 17). Detailed analysis resulted in a high-enriched number of 14 functional categories of metabolic pathways and cellular processes at the generative stage. Major CHO metabolism that includes sucrose synthesis and genes involved in the TCA cycle and organic acid metabolism was significantly overrepresented ($p<0.0001$). In accordance, transcripts of sucrose phosphate synthase (SPS) were overrepresented and up-regulated to 3-fold (Figure 17 and for more detailes Appendix Figure 2) indicating a relationship between strength of Suc and the accumulation rate of starch in seed for higher seed yield.

In addition, among transcripts of genes involved in TCA and organic acid metabolism only transcript pyruvate dehydrogenase complex involved in the conversion of Pyr to acetyl-coA, was overexpressed in flag leaves and up-regulated to 3-fold (Figure 17 and for more detailes Appendix Figure 2). Furthermore, transcripts of genes crucial for aromatic amino acid biosynthesis such as Phe, Tyr, and Trp were differentially overrepresented in the plant with high biomass. They included, indole -3-glycerolphosphate synthase (up to 172-fold), prephenate dehydratase (up to 2.8-fold) and shikimate kinase (up to 2.6-fold) (Figure 17 and for more detailes Appendix Figure 3). Moreover, several additional genes important for cell wall and secondary metabolism were also highly overrepresented. Among them, transcripts of genes involved in cellulose and cell wall modification (up to 26-fold), phenylpropanoids (20 genes, up to 53-fold) and 15 genes related to isoprenoid synthesis (up to 20-fold) were highly up-regulated (Figure 17 and Excel File 4). These results show that most genes involved in the synthesis of aromatic acids, phenylpropanoid and sucrose are major determinants for an increased biomass in flag leaves.

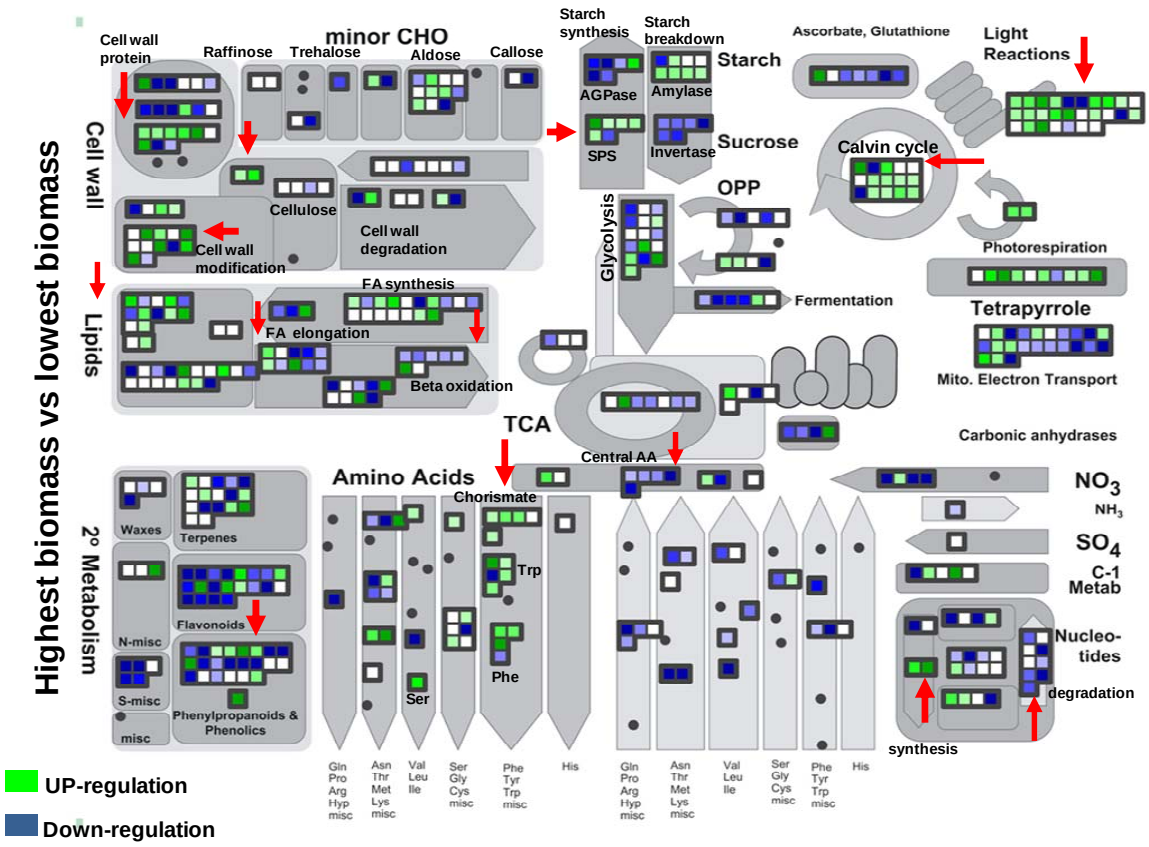


Figure 17. Schematic presentation of transcript levels in the plant with the highest biomass (HOR216) compared to that with the lowest biomass (S42II107) in flag leaves. Log 2 ratio was calculated to make average transcript abundance based on three independent replicates of Agilent Barley GeneChip normalized gene expression for the data of HOR216 versus S42IL107. The resulting file was loaded into the MapMan Image Annotator module to generate the metabolism overview map. On the logarithmic color, scale ranging from +2 to -2, dark blue and green represents at least 4-fold higher gene expression. The significant overrepresented differentially expressed genes are indicated with red arrows.

4.3.4. Relationship between transcripts and enzyme activities

Network correlation analysis was performed to investigate the relationship between transcripts and enzymes activities. To this end, a subset of 819 transcripts out of 6518 (section 4.3.1) that were mainly involved in energy, carbohydrate and amino acid metabolism and the activity of 22 enzymes from different metabolic pathways were used. The combination of network revealed a large number of positive linkages between transcripts and enzymes than negative linkages. All pair-wise correlations are provided in Excel File 5. Transcripts of genes encoding PGI, AlaAT, and AGPase showed the highest node degrees with 96, 93, and 77 connections with transcripts ($p<0.01$) (Excel File 5). However, out of 819 identified genes, only six genes showed direct positive correlations with the enzymes,

AGPase, AlaAT, PGI, 6PGDH, GAPDH and PGM (Table 1). Moreover, AGPase negatively correlated with transcript of AGPase large subunit 1 and positively with large subunit 2. The results demonstrate that enzyme activities can be affected by their corresponding transcripts (Table 1).

Table 1. Relationship between transcripts and corresponding enzyme of primary metabolism extracted after Pearson correlation analysis. Significant positive and negative associations are indicated at levels of 0.01 and 0.001.

Enzyme	Gene description	Correlation	p-value
AGPase	AGPase (Large subunite 1)	-0.91	<0.001
	AGPase (Large subunite 2)	0.80	<0.001
6PGDH	6PGDH	0.71	<0.01
AlaAT	AlaAT	0.74	<0.01
G6PDH	G6PDH	0.72	<0.01
GAPDH	GAPDH	0.74	<0.01
PGI	PGI	0.79	<0.001
PGM	PGM	0.83	<0.001

4.3.5. Identification of genes correlating with shoot biomass

Correlation analysis was conducted between transcripts of genes related to energy metabolism, carbohydrate and amino acid metabolisms to find out the genes associated with total shoot biomass at the generative stage. This matrix was analyzed by performing Pearson correlation for all pairs of measured transcripts across the selected genotypes and total shoot biomass. A stringent Pearson correlation ($r>0.70$) and accompanying *P*-values ($p<0.01$) was used for the identification of significant correlations. The results of correlation for each metabolic category are described in the following sections.

The correlation analysis between transcripts of genes involved in primary and secondary metabolism and shoot biomass showed strong associations (R-values greater or smaller than 0.70 and -0.70 at a level of at least 0.01). This was true for 19 positive and one negative connection for the transcripts of the genes involved in photosynthesis as shown in Table 2 and 18 positive and 16 negative associations summarized in Table 3. In term of GO, 17 genes belonged to light reactions, two were involved in the Calvin cycle pathway, and one was associated with photorespiration. Among genes involved in the light reaction, the gene encoding

cytochrome b (N-terminal)/b6/petB was negatively connected to total shoot biomass. A very strong positive relationship was found for Rubisco, the enzyme involved in the first step of carbon fixation and catalyzing the conversion of the ribulose-1,5-bis phosphate to 3PGA (Table 2).

Table 2. Pearson correlation analysis between total shoot biomass and transcripts of energy metabolism in the flag leaves. Correlations are indicates by r-value, which are significant at a level of 0.01 and 0.001. Significant negative correlations as given in bold.

Gene function	Pathway	Correlation
Calcium ion binding	Light reaction	0.89
Chlorophyll A-B binding protein	Light reaction	0.89
Cytochrome b6 gb	Light reaction	**-0.87**
Ferredoxin-related	Light reaction	0.88
Ferredoxin-related	Light reaction	0.87
Glycolate oxidase	Light reaction	0.87
Glycosyl hydrolase family 1 protein	photorespiration	0.87
PF00111 2Fe-2S iron-sulfur cluster binding domains	Light reaction	0.87
Photosystem I reaction center subunit V	Light reaction	0.86
Photosystem II 11 kDa protein-related	Light reaction	0.86
Photosystem II 11 kDa protein-related	Light reaction	0.85
Photosystem II 22kDa protein	Light reaction	0.86
photosystem II reaction center PsbP family protein	Light reaction	0.88
Plastid transcriptionally activate 14	Light reaction	0.87
PSAD-1	Light reaction	0.87
PsbP family protein	Light reaction	0.89
PSI-N PSI-N	Light reaction	0.85
Ribulose bisphosphate carboxylase/oxygenase activase	Calcin cycle	0.88
T01404 GA binding protein beta subunit	Light reaction	0.86
Trios phosphate isomerase	Calvin cycle	0.86

A tight co-regulation was found for transcripts of amino acid and cell wall biosynthesis genes (Table 3). Through those genes, aspartate aminotransferase, and transcripts of genes involved in cell wall metabolism like auxin response factor 7, cell wall precursor proteins and cellulose synthesis were more prominent. Concerning genes crucial for carbohydrate metabolism, three strong positive connections between total shoot biomass and transcripts of starch synthase (SS), SPS and trehalose-6-phosphate phosphatase (TPP) genes showed positive association to total shoot biomass (Table 3). Further associations with total shoot biomass were transcripts of cAldolase, PGM, citrate synthase and PK genes of which cAldolase and PGM positively connected to shoot biomass. As previously

shown in Figure 13, the activities of PGM, cAldoase and PK enzymes indicated strong positive correlations also to shoot biomass at the generative stage.

Moreover, transcripts of genes involved in secondary metabolism showed four strong positive and two negative correlations of which terpene synthase 4 and cinnamoyl-coA reductase 1 showed positive and negative correlations respectively. Further negative connections were found for transcripts of genes involved the TCA cycle, mainly for MDH and ICDH.

Table 3. Annotation of genes in relation to total shoot biomass at the generative stage. Genes located in their pathways were classified based on MapMan gene ontology.

Gene ID	Pathway	Function	Correlation	p-value
35_18216	amino acid metabolism	phosphoribosyl-AMP cyclohydrolase	0.75	<0.001
Contig6829_at		phosphoserine phosphatase	0.73	<0.001
35_9303		acetolactate synthase	0.83	<0.001
35_40404		aspartate aminotransferase	0.79	<0.001
Contig5255_at		alanine:glyoxylate aminotransferase 2	-0.71	<0.01
Contig4098_at		isovaleryl-CoA-dehydrogenase	-0.78	<0.001
Contig1748_s_at		aspartate aminotransferase	-0.80	<0.001
35_17044		isovaleryl-CoA-dehydrogenase	-0.76	<0.001
35_15359	cell wall	cell wall modification, expansin A10	0.71	<0.001
35_21030		cellulases and beta -1,4-glucanases	0.76	<0.001
35_9980		cell wall modification	0.75	<0.001
35_20887		cell wall precursor synthesis.UGE	0.77	<0.001
35_4931		xyloglucosyl transferase	0.80	<0.001
35_231		cell wall precursor synthesis.UGD	-0.81	<0.001
35_47336		cellulose synthesis	-0.73	<0.001
35_39662		cell wall modification	-0.77	<0.001
S0000800188A02F1_at	glycolysis/gluconeogenese	phosphoglucomutase	0.74	<0.001
35_1819		cytosolic Aldolase	0.74	<0.001
Contig4350_s_at		citrate synthase	-0.72	<0.001
35_16870		pyruvate kinase	-0.78	<0.001
HVSMEl0023l10r2_at	minor CHO metabolism	trehalose-6-phosphate phosphatase	0.73	<0.001
Contig1808_at		starch synthase	0.74	<0.001
35_41246		sucrose phosphate synthase	0.81	<0.001
35_526	nitrogen metabolism	glutamine synthase	-0.71	<0.001
35_528		glutamine synthase	-0.71	<0.01
35_8759	secondary metabolism	Terpene synthase 14	0.72	<0.001
35_29398		beta-carotene hydroxylase	0.75	<0.001
35_24887		Alkaloid-like strictosidine synthase	0.72	<0.001
35_22730		geranylgeranyl pyrophosphate synthase	0.80	<0.001
35_1498		cinnamoyl-coA reductase 1	-0.82	<0.001
35_19940		hXXXD-type acyl-transferase-like protein	-0.75	<0.001
35_962	TCA / organic acid	malate dehydrogenase	-0.73	<0.001
35_16234		isocitrate dehydrogenase	-0.73	<0.001
HVSMEi0006K11r2_at		malate dehydrogenase	-0.76	<0.001

4.4. Identification of co-regulated networks related to seed biomass in spring barley

4.4.1. Network correlation analysis of seed composition at the generative stage

The correlation matrices were generated based on 83 metabolites, 22 enzymes, and two phenotypical traits, TKW and total seed biomass (TSW). Network correlation analysis was computed using stringent Pearson product-moment correlation. Correlations for metabolites, enzymes and biomass with a significance at a false discovery (FD) rate <1% and <5% were selected as significant connections. The results of pair-wise relations are described in details in the following sections.

Elucidation of co-regulated metabolic pathways

Pearson r-coefficient was used for the correlation analysis in the seed. Out of 1402 direct relationships over the metabolites, 82 pairs resulted in significant negative and 1320 pairs in highly positive correlations (Appendix Table 6). Heat map was used to illustrate the correlations (Figure 18). In most cases, metabolites of the same metabolic pathway showed strong correlation with each other (Figure 18). Sugars, TCA cycle intermediates, and amino acids were highly correlated together in the network (Figure 18). There were also strong relationships between metabolites in different modules of carbon and nitrogen metabolism. For example, starch, the dominant accumulated carbohydrate in seed, showed the largest number of positive correlations with sugars, fatty acid derived compounds and TCA intermediates (see Excel File 2 for details). Amino acids were particularly correlated together and with sugars, sugar alcohols and fatty acid derived metabolites. The two major amino acids from nitrogen metabolism, Glu and Gln strongly and positively correlated with Pro, Ala, Asn and Asp (Figure 18, Excel File 2). These observations indicated that amino acids, sugars, sugar alcohols, starch, and fatty acid derived compound are major regulators in seed of spring barley at the generative stage.

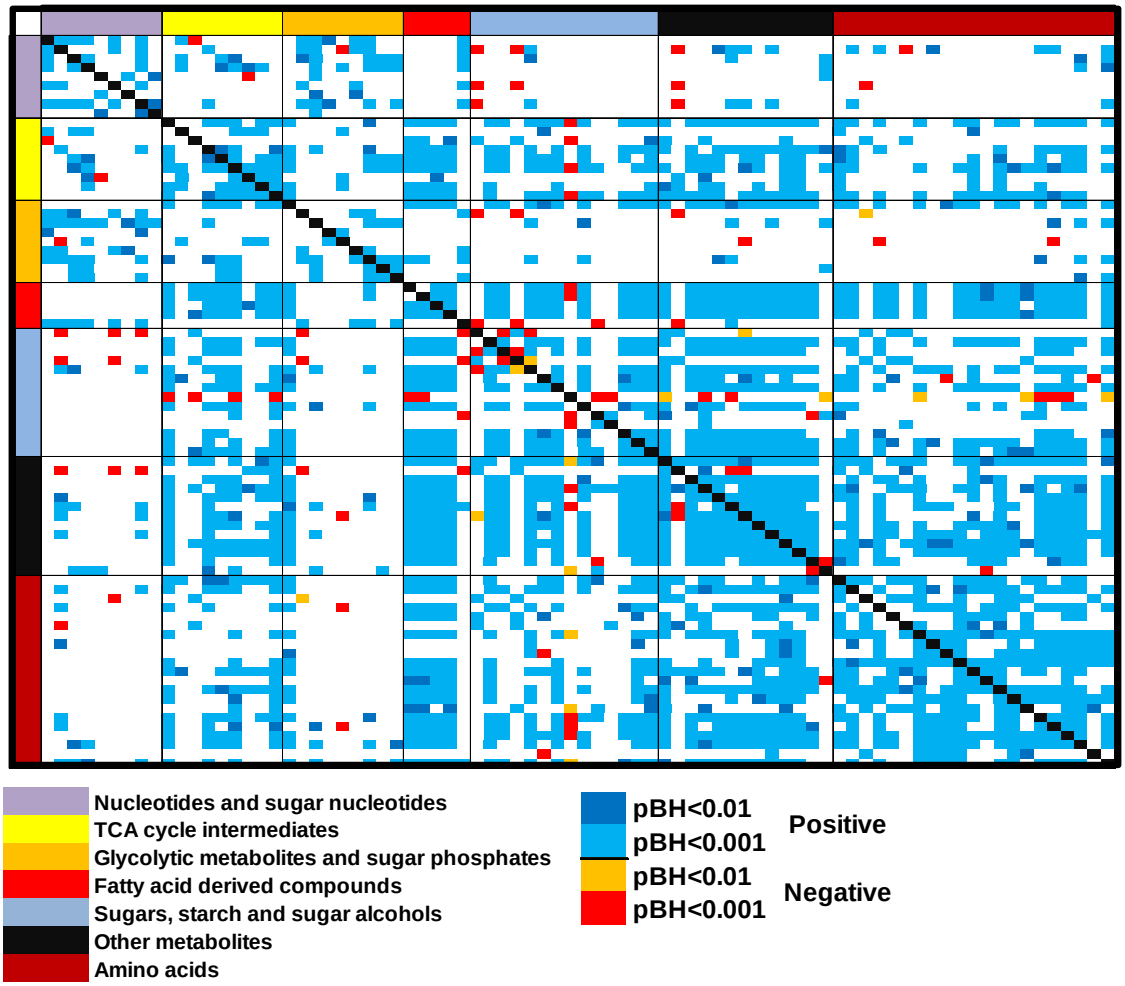


Figure 18. Heat map generated from the correlated metabolites of flag leaves. Pearson correlation coefficients are shown in red and yellow colors as negative and dark and light blue as positive correlations.

Correlations between enzyme activities

The network correlation assessment using a Pearson correlation analysis was conducted to investigate which enzymes show the highest co-regulations to other enzymes in seeds (Figure 19 and Excel File 2). The most highly connected enzymes were PGM, PGI, AlaAT and 6PGDH, followed by Susy and cFBPase (Figure 19). However, AcidInv, CWinv and CytInv showed weak connections. The investigation of interconnections among enzymes in the same pathway or different pathways revealed a positive correlation between enzymes of sucrose metabolism and starch indicating strong co-regulations between sucrose and starch biosynthesis (Figure 19).

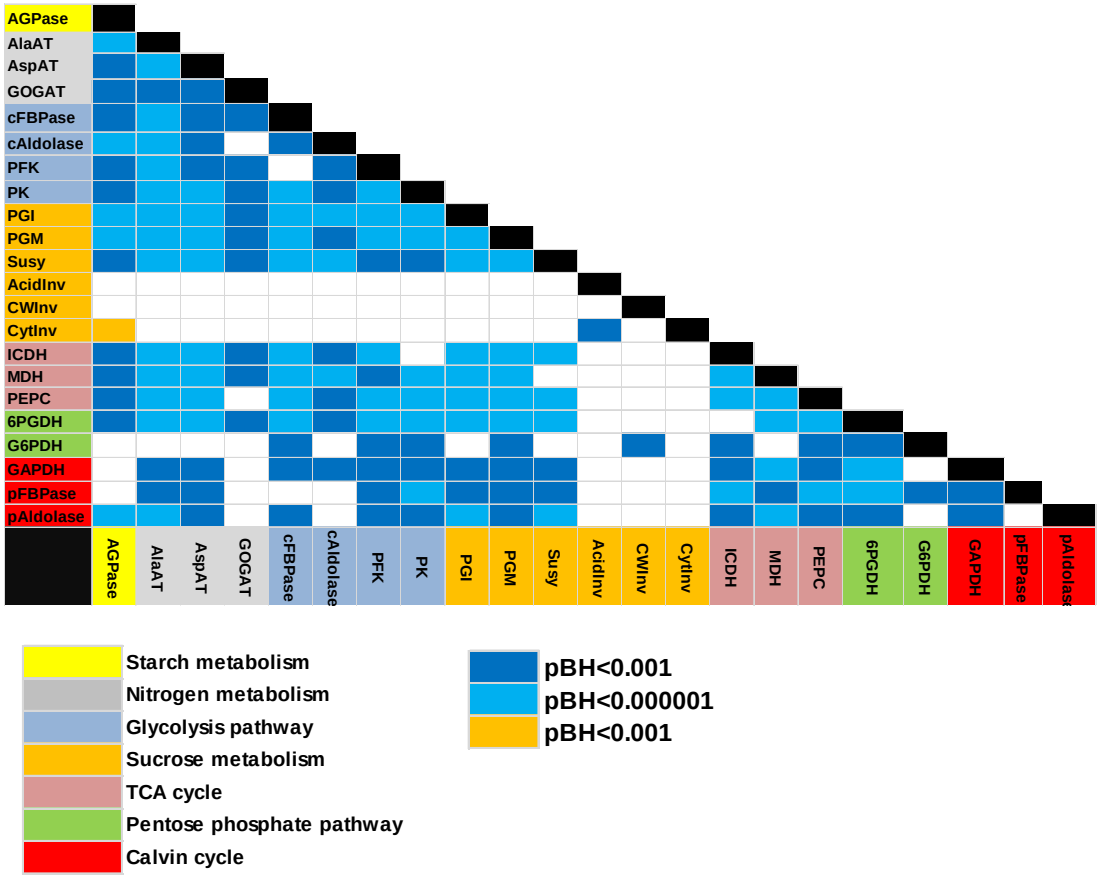


Figure 19. Heat map generated from the correlated enzymes of seeds. The heat map of correlation matrix shows the significance of the Benjamin-Hochberg corrected p-values of the Pearson correlations at the level of pB<0.001. Color code indicates the association between the enzymes. Dark and light blue and yellow colors depict positive and negative correlations, respectively.

Network correlation matrix between enzyme activities and metabolites

Twenty-two enzymes from carbon and nitrogen metabolisms were used for correlation matrix analysis with 86 annotated metabolites. A summary of all connections is provided in Appendix Table 6, and Figure 20. In total, 936 positive and 40 negative significant (pBH<0.01) connections were found. A high number of positive associations (pBH<0.0001) were found among amino acids and enzymes (Figure 20). GOGAT positively correlated with Gln and Glu. Further positive associations were detected for Asp and Ala to AspAT and AlaAT, respectively (Figure 20). AGPase was positively connected to starch whereas enzymes of sucrose metabolism were significantly associated with Suc and sugar derived

metabolites Glc1P, and Glc6P (Figure 20). The correlation between AGPase and Tre was negative. Further negative connections were found between two sucrose-metabolizing enzymes, AcidInv and CytInv and metabolites in the correlation matrix of which CytInv showed the strongest negative connection (pBH<0.001) to nucleotides ADP, ADPGlc, UDP and UDPGlc, to two glycolytic metabolites, PEP and PGA, and to two other metabolites aminoadipic acid (Amdp) and homoserine (Hser) (Figure 20).

Enzymes of TCA cycle, MDH, ICDH and PEPC were strongly associated to TCA cycle intermediates, Cit, Mal, Isocit, 2-oxo, OAA, Succ, and Fum, while Treh, Rib, ascorbate (Ascb) and arabinose (Ara) revealed negative connection to enzymes GOGAT, PFK, AcidInv, CytInv, and pFBPase. Further strong connections were found between PK and Pyr, Mal, Fum and MDH (pBH<0.001) and between Isocit and ICDH (pBH<0.001) (Figure 20). These results indicate a high co-regulation among metabolites and enzyme activities in all sectors of central metabolism particularly for amino acids, sugars and starch during seed set.

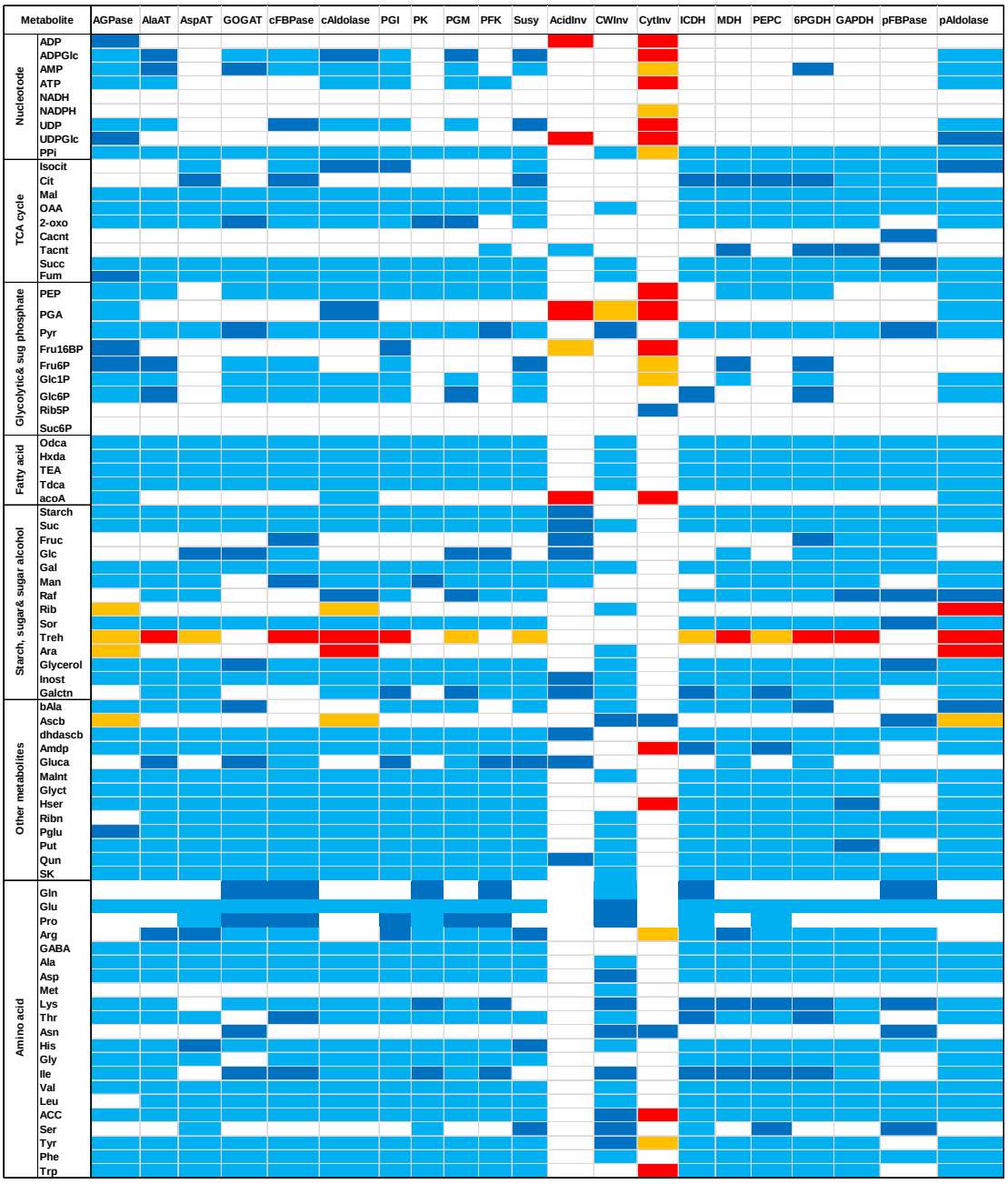


Figure 20. Heat map of a bilateral Pearson correlations generated between enzymes and metabolites of seeds at the generative stage. The significance of correlations at the levels of pBH<0.01 and pBH<0.001 are indicated with dark and blue colors for positive and yellow and red colors for negative correlations.

4.4.2. Identification of metabolite and enzyme signatures in seeds at the generative stage

Metabolite signatures

For the correlation analysis and the identification of important metabolites, which might determine the formation of seed biomass two morphological traits represented by TSW and TKW were used. The calculation was performed using a Pearson's product-moment r coefficient to extract metabolic traits significantly correlating with seed biomass. The significance test was based on stringent pBH-value at the levels of 0.05 and 0.001. A complete list of correlation analysis is provided in Excel File 2.

The first 53 metabolites extracted from the correlation analysis between metabolic traits and total seed biomass were summarized and integrated into biochemical pathways using MapMan (Figure 21). Metabolites, which showed high ranking according to extracted pBH-value (pBH<0.0001) are indicated with red color while white color represents significant correlation for metabolites with low ranking number (pBH<0.01). Sucrose showed the highest rank (pBH<0.0001) in the correlation matrix for seed biomass. Further significant associations were detected for starch, Glc1P, ADPGlc and Glc6P. Four TCA cycle intermediates including 2-oxo, Mal, Succ, Fum correlated with total seed biomass. Succ (pBH<0.001) indicated the highest rank among TCA intermediates (Figure 21). Out of glycolytic metabolites, PEP showed strong (pBH<0.001) and Pyr weak (pBH<0.01) correlation to total seed biomass. Among amino acids Val, His, Arg, Thr, Ile and Lys showed significant but with a low ranking correlation to total seed biomass (Figure 21). The metabolites involved in lipid metabolism including TEA, Ocda, Tdca, Hxda and glycerol indicated also significant relation (pBH<0.0001) to total seed biomass. Moreover, putrescine (Put), a polyamine showed a high ranking in the correlation matrix. In addition, two further metabolites, Ala and Dhasb were highly connected to total seed biomass (Figure 21).

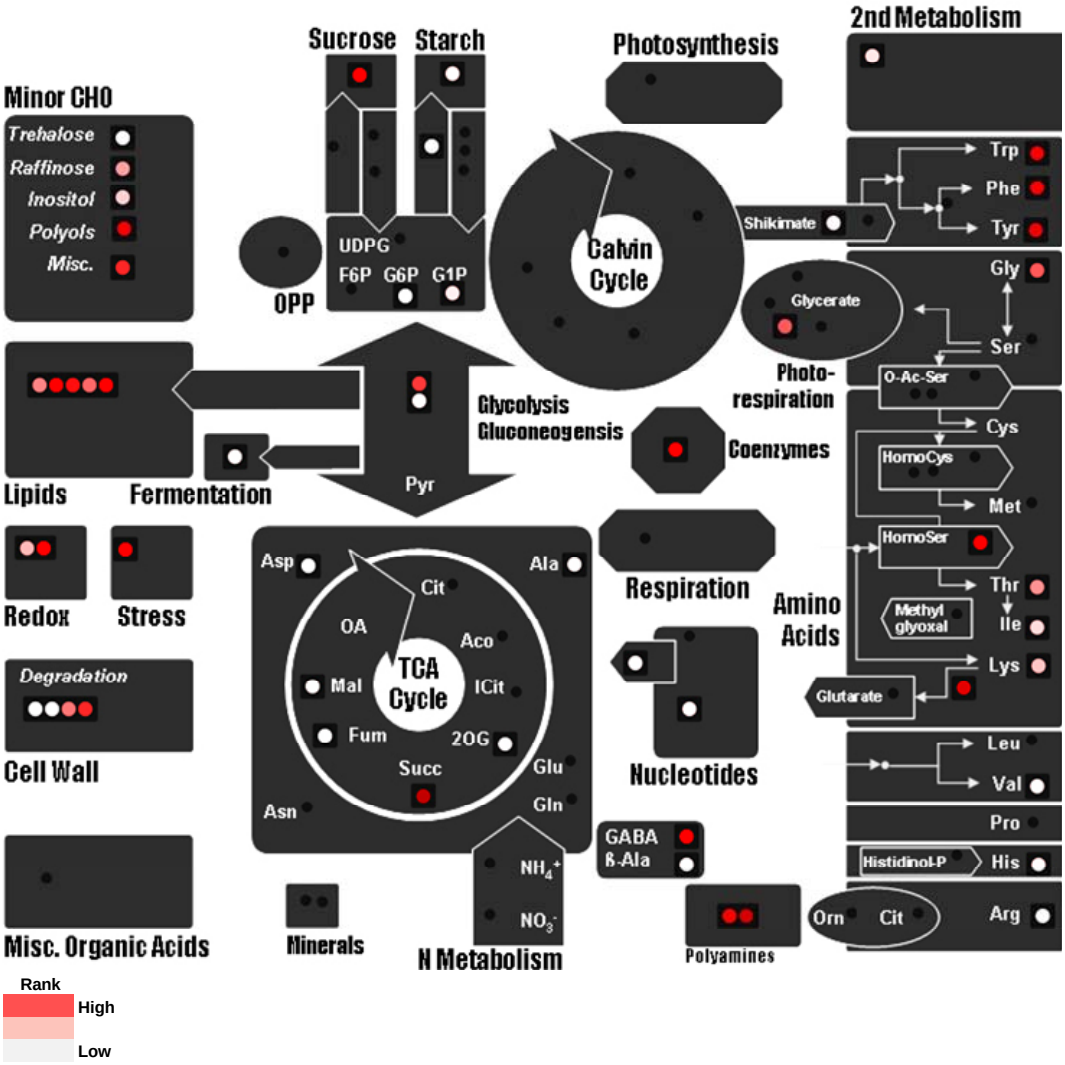


Figure 21. Schematic representation of the most important metabolites according to pBH-values on biochemical pathways. Correlation analysis was carried out between metabolites and seed biomass. 53 metabolites with high correlations to seed biomass were identified. The integration of metabolites in metabolic pathways was performed by MapMan. Red and white colors visualize metabolites, which correlated to total seed biomass. Red and white colors show metabolites with the highest and lowest impact on total seed biomass seed biomass formation respectively.

Enzyme signatures

In order to create an enzyme signature for a direct correlation between specific enzyme and total seed biomass, twenty-two enzyme activities of carbon and nitrogen metabolism were correlated against total seed biomass. Figure 22 illustrates a summary of enzymes from different metabolic pathways whose activities were associated with total seed biomass (see Excel File 2 for more details). Dark blue and large node size indicate significant relationships at the level of pBH<0.001, while light blue and small node size represent significant

relationships at the level of pBH<0.01. In total, six enzymes including AGPase, PGM, PGI, cAldolase and AlaAT were highly and significantly correlated to total seed biomass (pBH<0.00001). Further significant correlations were found for the sucrose-metabolizing enzyme Susy, two enzymes of pentose phosphate pathway comprising of 6PGDH and G6PDH, and an enzyme of the TCA cycle, MDH. GOGAT, which is involved in nitrogen and ammonia metabolism, was positively correlated to total seed biomass (Figure 22).

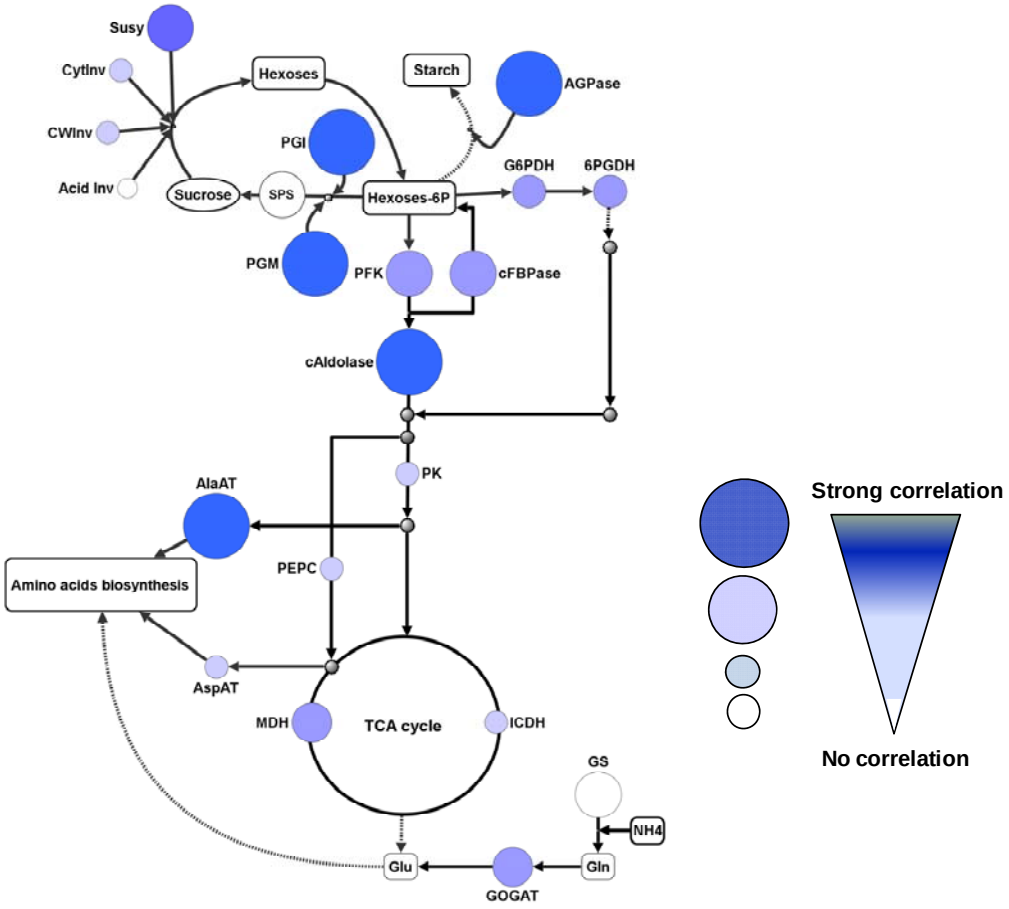


Figure 22. Schematic illustration of central metabolism and their corresponding enzymes that correlated to total seed biomass 15 days after anthesis under field conditions. Correlation analysis was computed between enzymes and total seed biomass using Pearson r coefficient. Nodes size shows correlation degrees. Blue color and large nodes followed by white color and small nodes visualize enzymes, which strongly or weakly correlated with seed biomass, respectively.

4.5. Evaluation of co-regulated genes at the generative stage in seeds

4.5.1. Statistical analysis for identification of differentially expressed genes

The one-way ANOVA was used to extract differentially expressed genes in four contrasting genotypes, HOR4555, HOR216, and Scarlett with S42IL107. Extraction of statistical results led to the identification of 2332, 2766, and 1391 differentially expressed genes in HOR216, HOR4555 and Scarlett compared to S42IL107 (Appendix Figure 4 and Excel File 4). RNA and protein metabolism and biotic stress responses with 523, 428, and 426 genes contained the largest number of differentially expressed, down-regulated transcripts (Appendix Figure 4 and Excel File 4). Furthermore, mineral nutrient acquisition, RNA and protein metabolism with 330, 319 and 468 genes were the most prominent categories of up-regulated genes (Excel File 4).

Transcriptional changes in primary and secondary metabolisms

Preferentially expressed genes of seeds (section 4.5.1 and Appendix Figure 4) in the plant with highest biomass formation compared to that with lowest biomass formation, were used for the enrichment analysis using MapMan ontology (Figure 23 and Excel File 4). Functional enrichment analysis indicated enrichments of 12 GO categories particularly for sucrose and starch synthesis, starch degradation, and minor carbohydrate metabolism. Detailed analysis showed an overrepresentation of 10 up-regulated genes in starch breakdown ($p<0.001$). They included starch phosphorylase (up to 2-fold), triose phosphate translocator (up to 2-fold), starch cleavage enzyme of beta-amylase (up to 15-fold) and D-enzyme or starch binding domain (up to 3-fold). Moreover, transcripts of SPS which is located in sub-cluster of sucrose synthesis were overrepresented indicating Suc being the main carbon supply for the endosperm (Figure 23 and for more details Appendix Figure 5).

Further, significantly overrepresented transcripts were found for genes involved in minor CHO metabolism. In total, 30 transcripts revealed a co-regulation of 19 up-regulated and 11 down-regulated genes in the plant with high biomass production. Transcript accumulation was detected for genes involved in trehalose metabolism

including TPP (up to 2-fold) and trehalose-6- phosphate synthase (TPS) as well as for alpha-galactosidase 1 gene (up to 2.8-fold) that occurs in galactose metabolism and finally for glucan synthase (up to 2-fold) which is involved in callose and polysaccharide biosynthesis. Moreover, three transcripts of genes for galactinol synthase (up to 4.8-fold), myo-inositol oxygenase (up to 5.2-fold) and three isoforms of xylose isomease (up to 7-fold) were up-regulated (Figure 23 and for more details Appendix Figure 5).

Further transcript accumulation was observed for genes involved in photosystem I (up to 6.2-fold) and II (up to 96-fold) and three down regulated genes involved in Calvin cycle (Figure 23). Furthermore, cellulose synthase transcript (up to 8-fold) and 21 additional transcripts from genes involved in cell wall modification (up to 15-fold) were up-regulated. Upon genes from nitrogen metabolism, glutamine synthetase (GS1 and GST 2) gene (up to 4-fold) was up-regulated. There were also up-regulated genes occurring in aromatic amino acid biosynthetic pathways such as shikimate kinase (up to 2.6-fold) (Figure 23 and Appendix Figure 6).

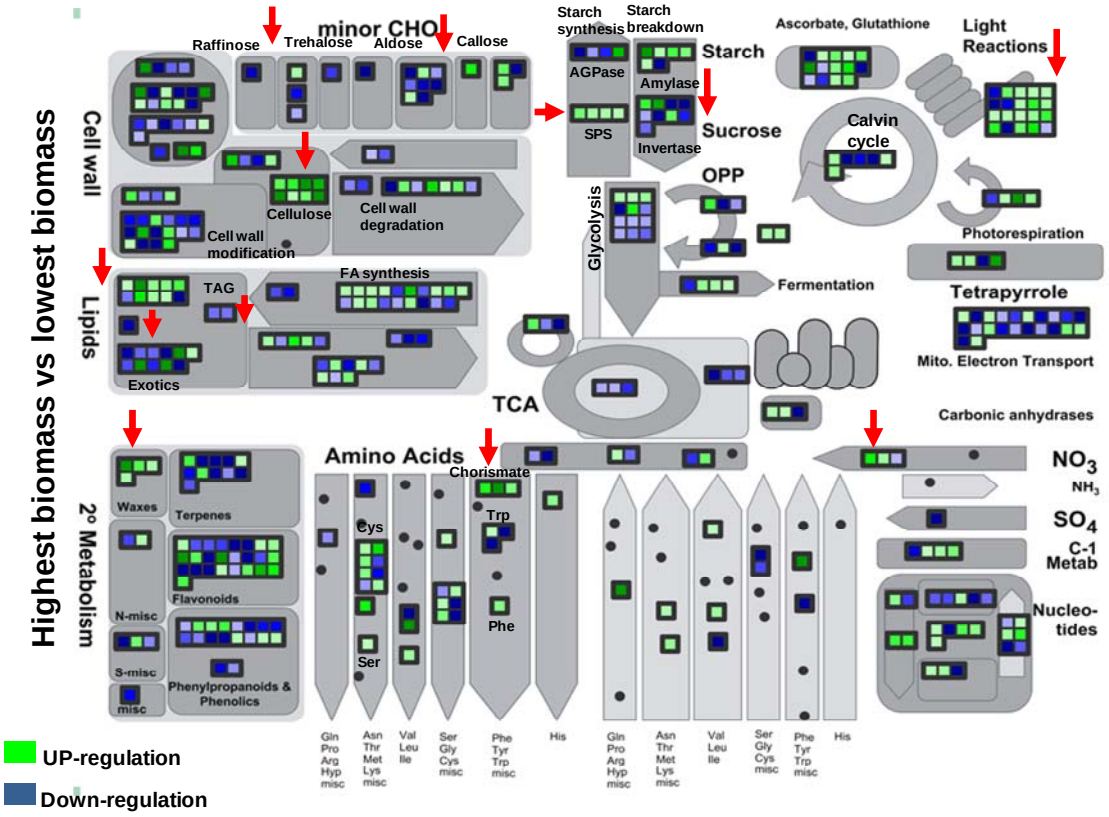


Figure 23. Schematic presentation of transcript levels in the plant with the highest biomass (HOR216) compared to that with the lowest biomass (S42II107) in seeds. Log 2 ratio was calculated to make average transcript abundance based on three independent replicates of Agilent Barley GeneChip normalized gene expression for the data of HOR216 versus S42IL107. The resulting file was loaded into the MapMan Image Annotator module to generate the metabolism overview map. On the logarithmic color, scale ranging from +2 to -2, dark blue and green represents at least 4-fold higher gene expression. The significant overrepresented differentially expressed genes are indicated with red arrows.

4.5.2. Network correlation between differentially expressed genes, enzyme activities and biomass

Relationship between enzyme activities and transcripts

Coordinated changes of 22 enzyme activities and 819 selected transcripts (section 4.5.1) were analyzed in seeds by running a Pearson's product moment correlation. The correlations for enzyme-transcript traits that were significant at the level of 0.01 are listed in Excel File 5. In summary, 391 positive and 432 negative correlations were found in the correlation matrix. In particular, the enzymes,

GOGAT (109 connections), cAldolase (21 connections) and two invertases, CytInv and AcidInv (110 connections) showed the largest number of significant positive correlations to transcripts (see Excel File 5 for details). However, out of 819 identified genes, only one gene showed direct positive correlations with the enzymes, G6PDH (Table 4). Moreover, AGPase and cAldolase showed high negative correlation to their coding transcripts (Table 4).

Table 4. Correlation analysis between enzyme activity and transcripts using pair-wise correlation analysis. Significant correlation for enzymes and their encoded transcripts are indicated at the level of 0.001.

Enzyme	Gene description	Correlation	p-value
AGPase	AGPase	-0.72	<0.001
cAldolase	cAldolase	-0.74	<0.001
G6PDH	G6PDH	0.75	<0.001

Relationships between seed biomass and transcripts

To identify genes, which might determine the formation of seed biomass at the generative stage, a correlation analysis between seed biomass and transcripts was performed. The correlation analysis resulted in 58 significant correlations of which 19 were positive and 39 negative (Table 5). In terms of GO, two clusters including transcripts of genes involved in cell wall and amino acid metabolism revealed the highest number of associated genes to total seed biomass. For amino acid metabolism, ten negative and one positive association were detected. The transcript of anthranilate synthase gene, which catalyzes the production of Glu, Pyr and anthranilate, cystein synthase, aspartate-semialdehyde dehydrogenase, lactoylglutathione lyase, Enoyl-CoA hydratase/isomerase and Indole-3-glycerol phosphate synthase, were significantly co-regulated with total seed biomass (Table 5). Furthermore, strong positive associations with total seed biomass were found for glutamate dehydrogenase gene that is involved in nitrogen metabolism and for genes participating in sugar alcohol biosynthesis like galactinol synthases, 1-epimerase-like protein, PfkB-like carbohydrate kinase family protein, Sorbitol dehydrogenase-like protein, and myo-inositol phosphatase. Transcripts of three

genes including myo-inositol synthase, TPP, and galactokinase were negatively connected to total seed biomass (Table 5).

Furthermore, four significant associations were found for transcripts of genes involved in TCA cycle and glycolysis like phosphoglycerate mutase, which showed positive and PGM, dihydrolipoyl dehydrogenase and carbonic anhydrase which showed negative correlation to total seed biomass. In addition, among the genes occurring in starch and sucrose metabolism, starch debranching enzyme and SPS showed the strongest positive and negative associations to total seed biomass, respectively. For two isoforms of sucrose metabolizing enzymes, Susy and CytInv, strong negative association with total seed biomass was detected. Further strong connections to total seed biomass were found for transcripts of genes involved in cell wall biosynthesis from which four positive and nine negative correlations were detected. The transcripts of hemicellulose synthase, xyluglucan hydrolyase and pectin esterase were significantly associated with total seed biomass. Transcripts of terpenoid synthase gene were positively and those of 4-coumarate-CoA ligase, cinnamoyl-CoA reductase, cinnamyl alcohol dehydrogenase and dihydroflavonol reductase were negatively connected to total seed biomass (Table 5).

Table 5. Pearson correlation analysis for total seed biomass and transcripts of the genes involved in primary and secondary metabolisms. Correlations indicate r-values, which are significant at a level of 0.01 and 0.001.

Probe set ID	Pathway	Putative function	Correlation	p-value
Contig11179_at	amino acid metabolism	anthranilate synthase	0.79	<0.001
35_907		lactoylglutathione lyase	-0.70	<0.01
35_29082		alanine aminotransferase	-0.71	<0.01
35_18158		indole-3-glycerol phosphate synthase	-0.71	<0.01
Contig6005_x_at		DPP6 N-terminal domain-like protein	-0.72	<0.01
35_4508		cysteine synthase	-0.75	<0.001
Contig7512_at		phosphoribosylanthranilate transferase like protein	-0.77	<0.001
35_22037		cysteine lyase	-0.77	<0.001
35_19930		aspartate-semialdehyde dehydrogenase	-0.82	<0.001
35_13256		enoyl-CoA hydratase/isomerase D	-0.84	<0.001
Contig17743_at		hypothetical protein	-0.84	<0.001
35_39662	cell wall	xyloglucan endotransglucosylase/hydrolase protein 29	0.76	<0.001
35_6707		hemicellulose synthesis	0.73	<0.01
35_25740		pectin esterases	0.73	<0.01
35_22838		pectin methylesterase inhibitor	0.71	<0.01
35_15987		nucleotide-rhamnose synthase/epimerase-reductase	-0.72	<0.01
35_17547		NDP sugar pyrophosphorylase	-0.74	<0.01
35_21030		cellulases and beta -1,4-glucanases	-0.75	<0.01
Contig2148_at		UDP-glucose dehydrogenase	-0.75	<0.001
35_2591		cell wall proteins, arabinogalactan	-0.77	<0.001
35_19659		pectate lyases and polygalacturonases	-0.77	<0.001
35_271		arabinogalactan protein	-0.78	<0.001
35_31565		xyloglucan endotransglucosylase/hydrolase	-0.79	<0.001
35_499		xyloglucan endotransglycosylase	-0.81	<0.001
Contig2924_s_at	fermentation	aldehyde dehydrogenase	0.76	<0.001
35_958		lactate dehydrogenase	0.72	<0.01
35_15578		aaldehyde dehydrogenase	-0.72	<0.01
Contig16253_s_at		oxalyl-CoA decarboxylase	-0.72	<0.01
35_2189	glycolysis/glyoxylate cycle	phosphoglycerate mutase	0.78	<0.001
Contig4785_at		Malate synthase	0.72	<0.01
Contig1650_at		PGM	-0.74	<0.01
35_21043	major CHO metabolism	starch debranching enzyme	0.85	<0.001
35_14485		putative sucrose synthase 3	-0.71	<0.01
35_2358		cytosolic invertase	-0.71	<0.01
Contig3114_at		triose phosphate translocator	-0.72	<0.01
35_5442		starch cleavage	-0.75	<0.01
35_28039		starch branching enzyme	-0.76	<0.001
35_14437		ADP/ATP carrier protein 1	-0.76	<0.001
Contig24678_at		putative Sucrose phosphate synthase 4F	-0.85	<0.001
35_1096	minor CHO metabolism	galactinol synthases, putative	0.80	<0.001
Contig4392_at		1-epimerase-like protein	0.78	<0.001
Contig9891_at		PfkB-like carbohydrate kinase family protein	0.75	<0.01
Contig4192_at		sorbitol dehydrogenase-like protein	0.73	<0.01
Contig12332_at		myo-inositol phosphatases	0.73	<0.01
35_35200		myo-inositol synthases	-0.73	<0.01
Contig24583_at		trehalose 6-phosphate phosphatase	-0.74	<0.01
35_17471		galactokinases	-0.75	<0.01
35_2280	nirogen metabolism	glutamate dehydrogenase	0.71	<0.01
35_6186	secondary metabolism	terpenoid synthase 6	0.79	<0.001
35_4011		4-coumarate-CoA ligase isoform	0.74	<0.01
35_43672		FAD/NAD(P)-binding oxidoreductase	0.74	<0.01
35_5011		putative cinnamoyl-CoA reductase	-0.72	<0.01
35_15396		dihydroflavonol 4-reductase	-0.74	<0.01
35_6378		3-hydroxy-3-methylglutaryl coenzyme A synthase	-0.74	<0.01
35_31433		cinnamyl alcohol dehydrogenase	-0.79	<0.001
35_17695		ubiquitin-protein ligase E3-alpha-like protein	-0.80	<0.001
35_2202	TCA cycle / organic acid	dihydrolipoyl dehydrogenase	-0.74	<0.01
35_8873		carbonic anhydrase	-0.80	<0.001

4.6. Relationship between phenotypical traits and metabolites in seeds at ripening stage

In order to analyze the relationship between final yield parameters at the ripening phase and metabolite composition, a pair-wise correlation analysis was conducted. The correlation was carried out using Pearson moment r coefficient between metabolites comprising of three sugars including Glc, Fru and Suc, and starch, 18 amino acids and three phenotypical traits e.g. FSB, grain yield (GY), and TKW (Figure 24). Nine significant correlations (pBH<0.05) were found between metabolites and FSB (Excel File 6) with starch having the strongest connection (pBH<0.01) (Figure 24 -box A). Sucrose was also positively correlated to starch, TKW, GY and FSB in the correlation matrix (pBH<0.001). Nitrogen-containing amino acids, Glu and Gln showed positive correlation with FSB indicating relationships between nitrogen reserves and growth. Two aliphatic amino acids of Ile and Val also were negatively correlated to all traits (Figure 24-box B).

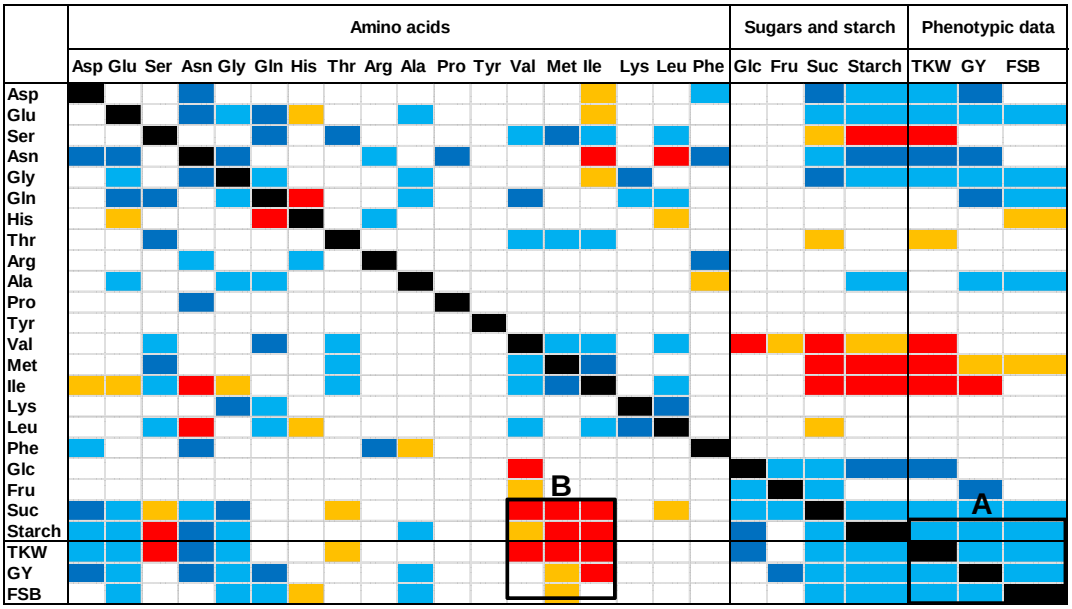


Figure 24. Correlation analysis between metabolites and final yield parameters in barley seeds at the ripening phase. Dark and light Blue and yellow and red colors indicate positive and negative associations, respectively. Black boxes indicate the positive connections between yield traits and their correlation with starch (A) and negative associations with amino acids (B).

4.7. Investigation of metabolic and biomass relationship at generative stage in winter barley

4.7.1. The effect of genetic variability on shoot biomass

For the following experiments, 12 diverse genotypes of winter barley were grown in field trial conditions in the year 2011. For the analysis, two morphological traits, TSB and TKW were used as indicators of final plant shoot and seed biomass. Five important phenotypical parameters are provided in Appendix Table 7. The TSB for twelve genotypes varied between 19 gr for BCC887 and 139 gr for BCC4555 (Appendix Table 7). This was seven-fold higher in plant with higher biomass (BCC4555) compared to the plant with low biomass (BCC887). Two genotypes, BCC1306 (64 gr) and BCC1315 (61 gr) were identified as intermediate performers. TKW was two-fold higher in BCC1596 (59 gr) compared to BCC662 (27 gr) (Appendix Table 7).

4.7.2. Influence of metabolites and enzyme activities on shoot biomass

Targeted and non-targeted metabolite profiling on flag leaves and seeds of 12 winter barley genotypes identified 80 (flag leaves) and 54 (seeds) metabolites of primary metabolism, which covered a range of sugars, sugar phosphates, nucleotides and sugar nucleotides, organic acids, amino acids, fatty acid derived metabolites and starch. In addition, 20 enzymes from sucrose, starch, and nitrogen metabolisms, TCA and Calvin cycle, glycolysis and pentose phosphate pathway were also added to the analysis. In total 100 (metabolites and enzymes) metabolic traits were measured in flag leaves and 74 (metabolites and enzymes) in seeds that were used to investigate their relations with their substrates or products to two morphological parameters, TSB and TKW.

Network correlation analysis between metabolites in flag leaves

The analysis of metabolite-to-metabolite relation indicated that TCA cycle intermediates, glycolytic metabolites, nucleotides and sugar nucleotides were positively connected. In contrast to metabolite-to-metabolite co-regulations in flag leaves of spring barley lines, amino acids were negatively associated to TCA cycle

intermediates, glycolytic metabolites nucleotides and sugar nucleotides (Figure 25 and Excel File 3). Fatty acid derived metabolites also negatively correlated with the TCA cycle intermediate and glycolytic metabolites. The correlations between sugars, sugar alcohols, and starch were negative to nucleotides, sugar nucleotides, TCA cycle intermediates, and glycolytic metabolites (Figure 25). The obtained results indicated an opposite metabolite interaction in flag leaves of winter barely compared to those of spring barley.

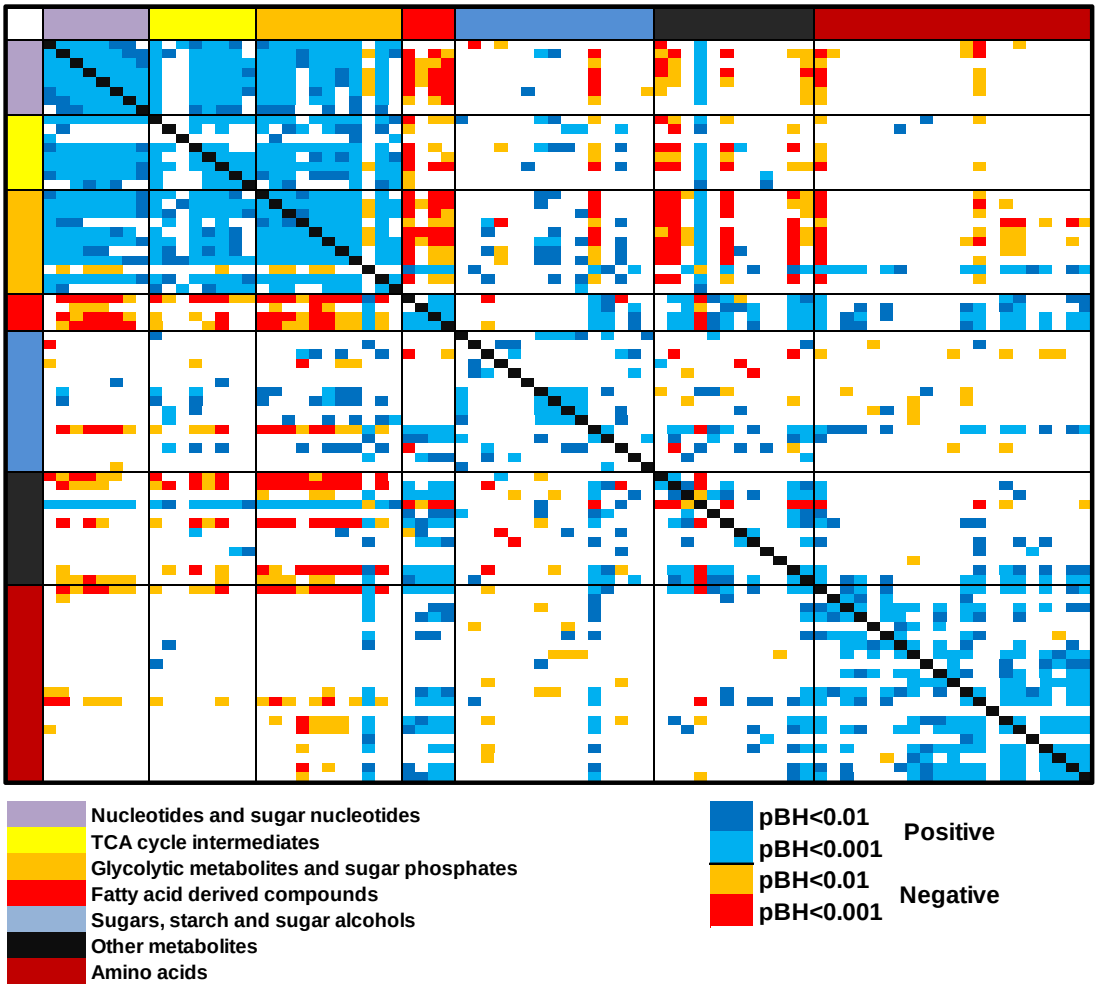


Figure 25. Metabolite to metabolite relationship resulted from a pair-wise analysis in flag leaves of winter barley at the generative stage 15 days after anthesis under field conditions. All metabolites are grouped in various classes that are indicated as color codes bellow the figure. Metabolites with the strong relations at the levels of 0.01 and 0.0001 are indicated with dark and light blue colors for positive, and yellow and red colors for negative interactions.

The interaction between enzyme activities in flag leaves

The network correlation analysis was performed to investigate which enzymes show the highest interaction to other enzymes and whether these interactions are within the same pathway or with other pathways. Consistent with the positive relationship between the enzymes in flag leaves of spring barley, winter barley also revealed a large number of positive than negative connections (Figure 26 and Excel File 3). This was obvious for the enzymes in the same metabolic route like TCA cycle, glycolysis, and pentose phosphate pathway (Figure 26). Inconsistent with the negative co-regulations of Susy and AcidInv in flag leaves of spring barley, CwInv and CytInv negatively co-regulated in other enzymes of flag leaves in winter barley at the generative stage. Poorly connected enzymes in correlation matrix included cAldolase and pAldolase (only one significant correlation), and PK, GAPDH, PGM and AcidInv (three significant correlations).

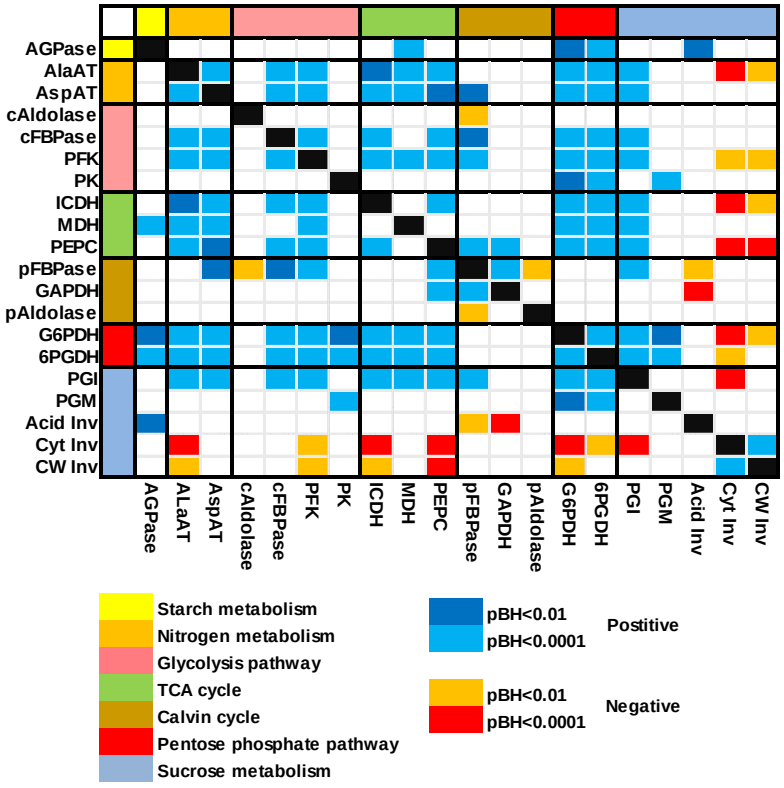


Figure 26. Correlation analysis between enzyme activities of flag leaves of winter barley at generative stage 15 days after anthesis under field conditions. The significant association at the level of 0.01 and 0.001 are indicated using dark and light blue colors for positive and yellow and red color for negative connections.

Relationships between enzyme activities and metabolites

The interrelation of enzymes and metabolites in flag leaves revealed a large number of negative relationships between enzymes, nucleotides, TCA cycle intermediates, glycolytic metabolites, and sugar phosphates (Figure 27). All the metabolites except Rib5P showed negative correlation with the majority of enzymes (Figure 27). A series of either positive or negative correlations was also found for enzymes and their substrates or products. Suc was positively associated to invertases ($pBH<0.01$) while negative correlations were identified between ICDH and Isocit, between PFK and Fru6P as well as G6PDH and Glc6P (Figure 27). Furthermore, three enzymes, AlaAT, PEPC and pFBPase involved in nitrogen metabolism, TCA, and Calvin cycle showed the largest number of associations in the correlation matrix (Figure 27). The results demonstrate that a metabolic depletion of primary pathways occurs in flag leaves of winter barley.

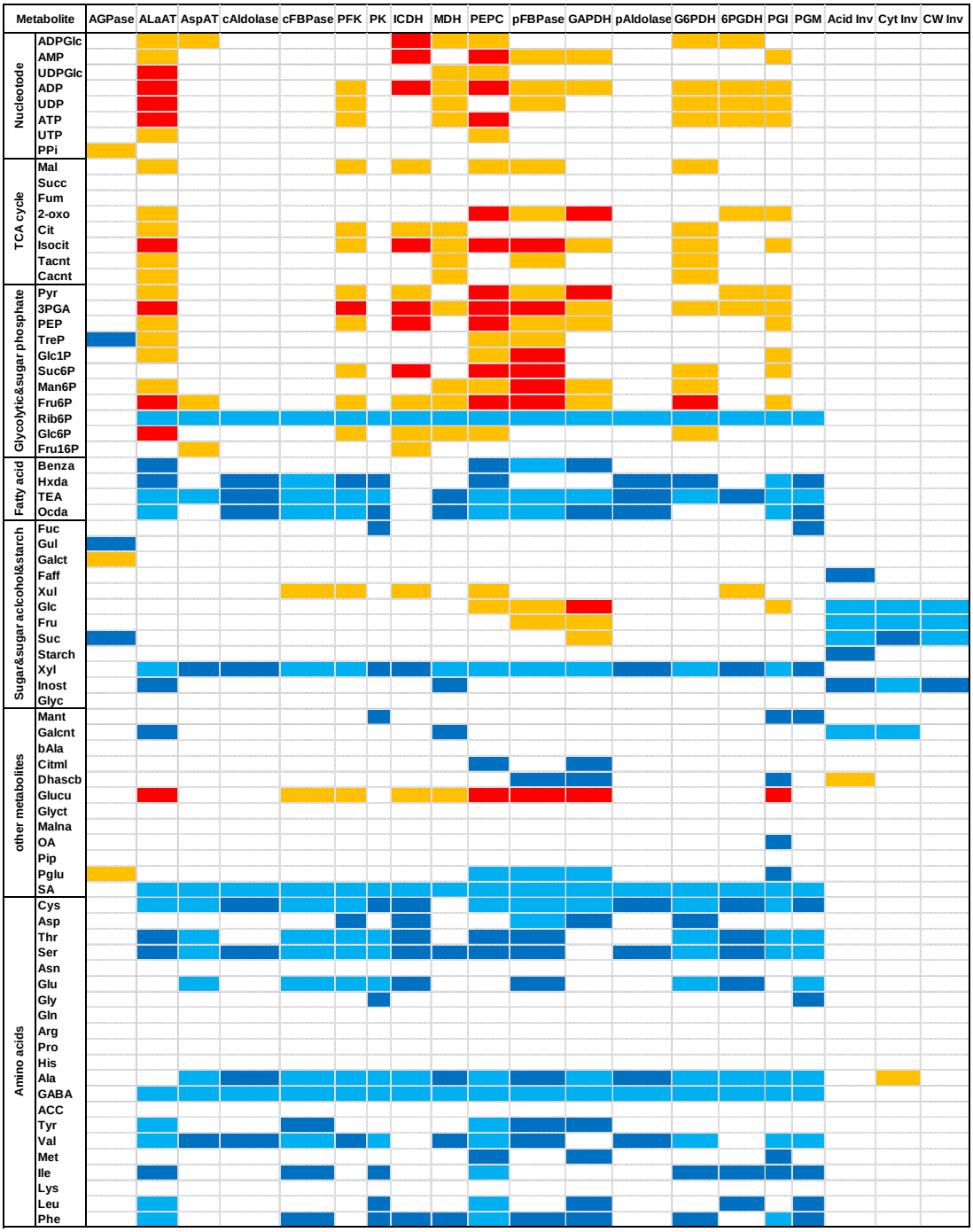


Figure 27. Heat map of a bilateral Pearson correlations generated between enzymes and metabolites of flag leaves of winter barley at generative stage. The significance of correlations at the levels of pBH<0.01 and pBH<0.001 are indicated with dark and blue colors for positive and yellow and red colors for negative correlations.

4.7.3. Relationship between total shoot biomass, enzyme activities and metabolites

Correlation analysis of 100 metabolic traits comprising of 20 enzymes and 80 metabolites with TSB revealed that 17 metabolites and 7 enzymes correlated strongly with total shoot biomass (Figure 28). TCA cycle intermediates, Mal, Cit, Isocit, and 2-oxo and sugar phosphates such as Glc6P, Fru6P, sucrose-6-phosphate (Suc6P), Glc1P, mannose-6-phosphate (Man6P) and PGA showed negative relations to total shoot biomass (Figure 28). Among metabolites mentioned above, Mal, Man6P and 2-oxo showed the strongest connections ($pBH<0.0001$) (Figure 28). Furthermore, AMP, ADP, and UDPGlc along with malonate (Malna) and glucuronic acid (Glucua), which is involved in nucleotide and cell wall biosynthesis as well as the amino acid Asn were negatively associated to total shoot biomass. Only one positive strong correlation could be found between benzoic acid and shoot biomass (Figure 28).

Out of twenty enzymes, seven enzymes exhibited significant relationships to shoot biomass (Figure 28). They included two negative connections, pAldolase and cAldose and five positive associations for PGI, cFBPase, GAPDH, G6PDH and pFBPase. Among the enzymes, pAldolase and pFBPase showed the strongest connections to shoot biomass ($pBH<0.001$) (Figure 28).

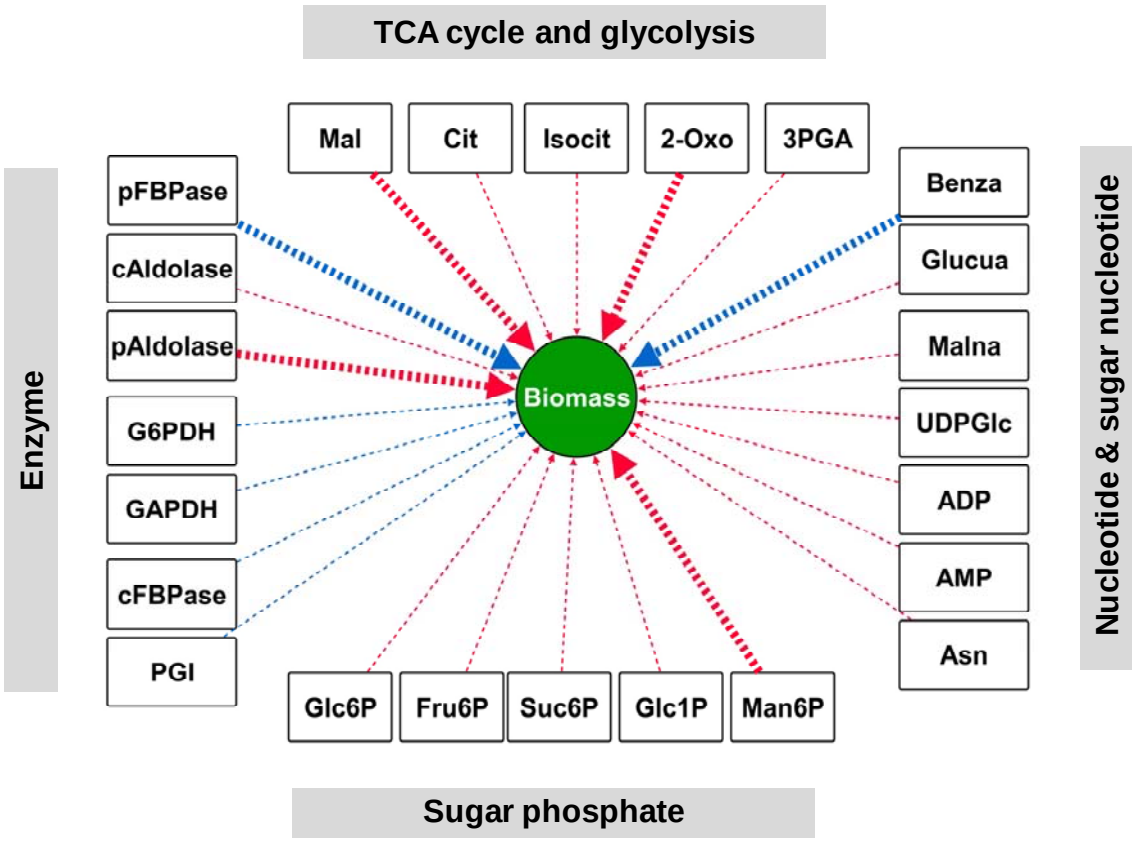


Figure 28. Correlations between metabolites, enzymes and total shoot biomass at the generative stage under field condition. Pair-wise correlation analysis resulted in strong associations at the level of 0.01 and 0.001. Blue and red dotted lines indicate positive and negative associations. Strong positive or negative connections are shown using thicker dotted lines.

4.8. Investigation of relationship between metabolites and final yield in seeds

The relationships between metabolites and enzymes and their connections to TKW were investigated at the generative stage in seeds of winter barley plants. For this analysis, 75 traits (54 metabolites, 20 enzymes, and TKW) were used to identify metabolite and enzyme signatures relevant for seed biomass. These correlations are described in details in the following sections.

4.8.1. Co-regulations of metabolites

As shown in Figure 29, there were two hot spots of positive correlations between nucleotides, sugar nucleotides, TCA cycle intermediates, and glycolytic metabolites and between sugars, sugar alcohols, and amino acids (Figure 29). These

observations were inconsistent with highly positive correlations between metabolites in seed of spring barley (Figure 18).

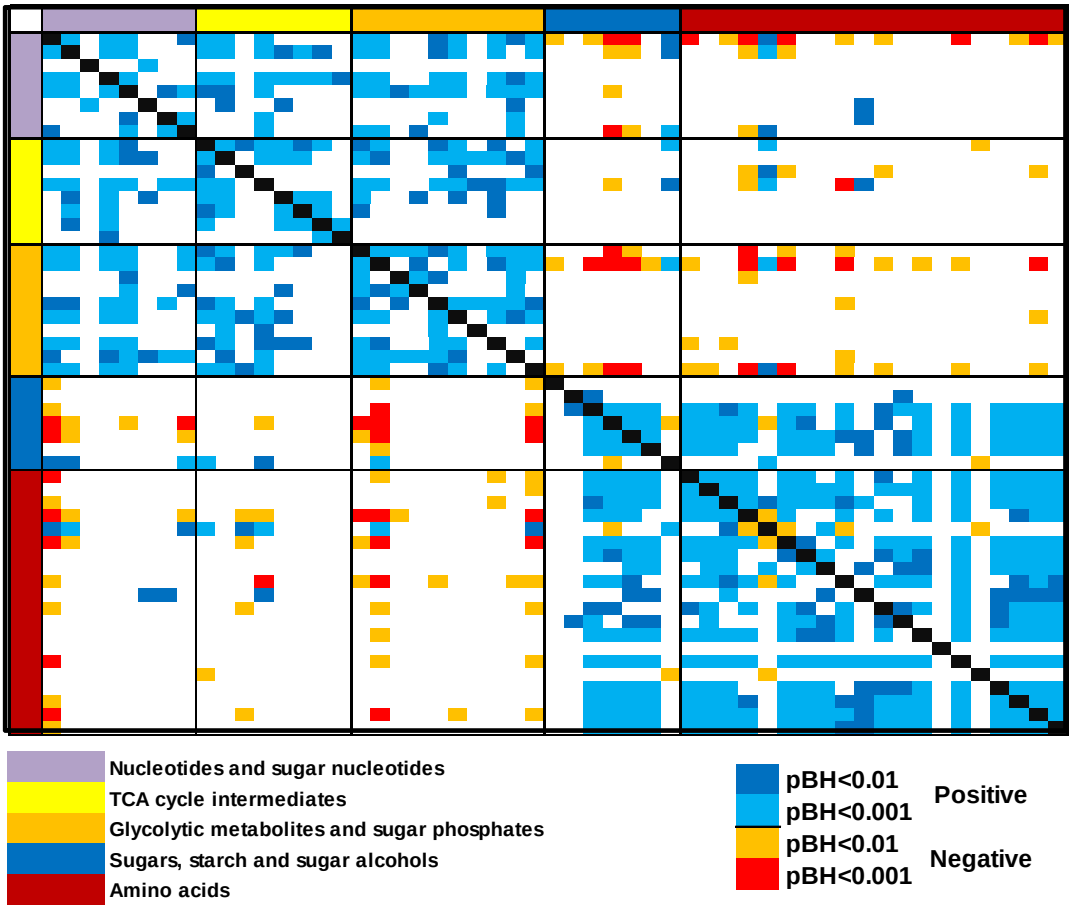


Figure 29. Pearson correlation analysis of metabolites in seeds of winter barley 15 days after anthesis at the generative stage under field conditions. Significant metabolite associations are indicated by color code and at the levels of 0.01 and 0.0001. Blue and dark colors show positive and yellow and red colors negative correlations.

4.8.2. Interrelations between enzyme activities

Figure 30 illustrates the interconnections between different enzymes in seeds at the generative stage. The majority of interconnections among enzymes were positive (Figure 30). Consistent with spring barley three sucrose-metabolizing enzymes, CWinv, CytInv and AcidInv, poorly and negatively connected in correlation matrix.

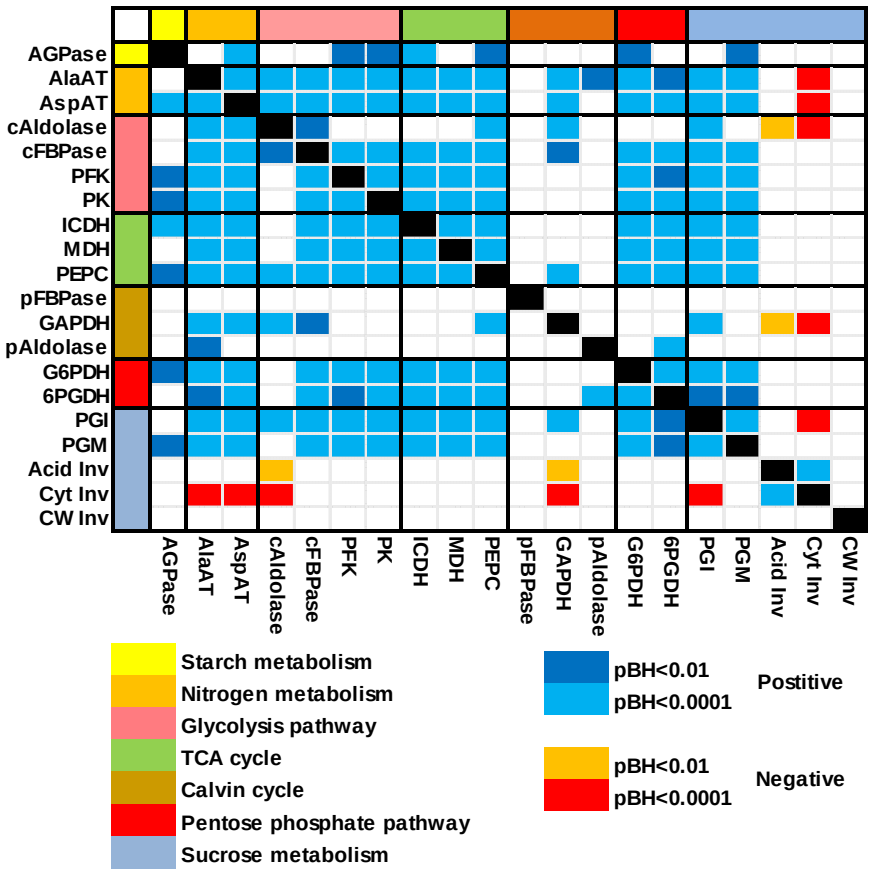


Figure 30. Heat map generated from the correlated enzymes of seed in winter barley. Color code indicates the association between the enzymes. Blue and red colors depict positive and negative associations, respectively.

4.8.3. Co-regulation between metabolites and enzyme activities

Figure 31 shows a summary of correlation network between all metabolites and enzymes. In contrast to the large number of correlations among enzyme activities and metabolites in spring barley, there were a low number of associations in winter barley (Figure 31). Nucleotides, TCA cycle intermediates, glycolytic compounds and sugar phosphates indicated a low number of connections whereas sugars, sugar alcohols, starch, and amino acids had strong connection to enzymes activities. Three sucrose-metabolizing enzymes, CytInv, CwInv and AcidInv, showed also a high number of connections to the metabolites in seed of winter barley (Figure 31).

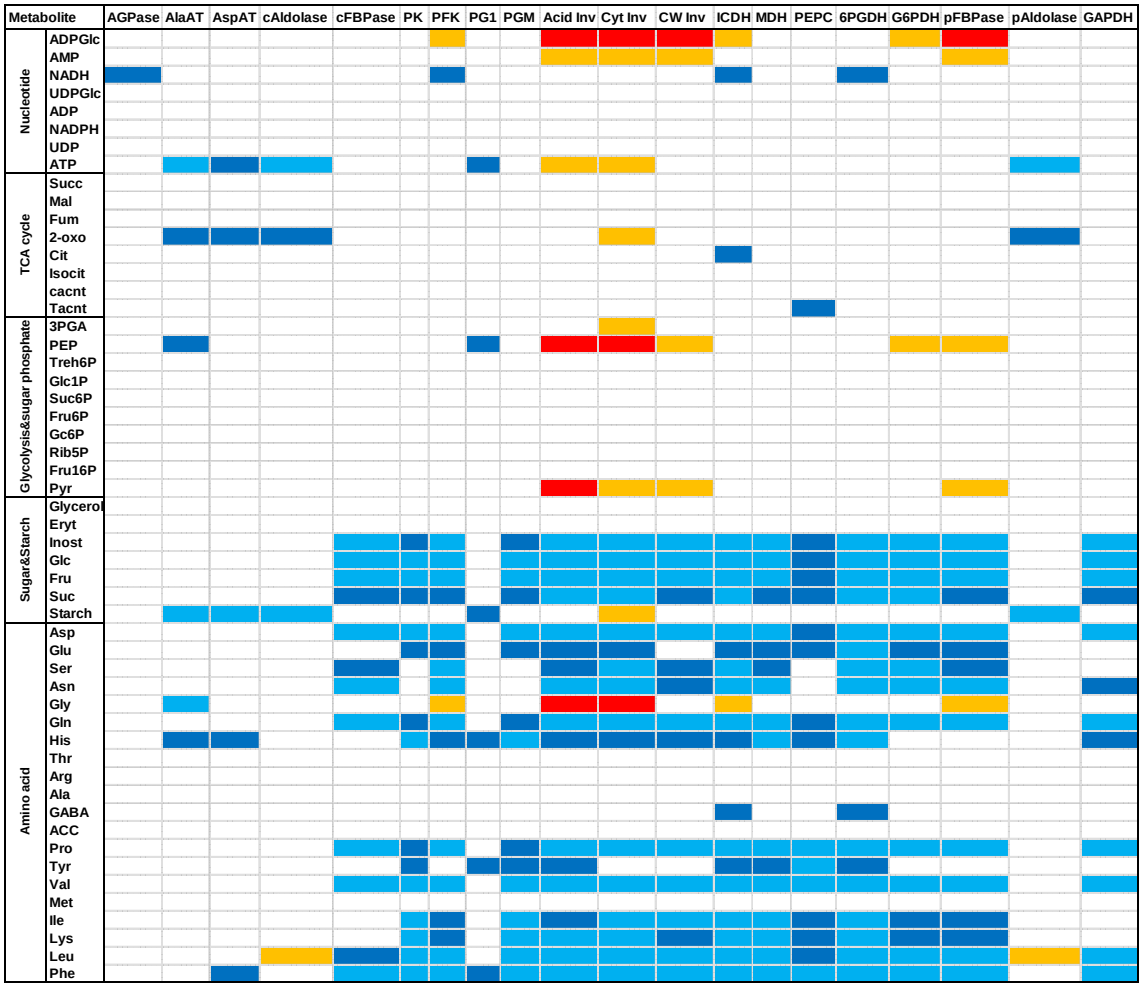


Figure 31. Heat map of a bilateral Pearson correlations generated between enzymes and metabolites of seeds of winter barley at the generative stage. The significance of correlations at the levels of pBH<0.01 and pBH<0.0001 are indicated with dark and blue colors for positive and yellow and red colors for negative correlations.

4.8.4. Enzyme and metabolite signatures for total seed biomass

A pairwise correlation analysis was applied to identify the relationship between enzymes, metabolites and TKW as indicator for total seed biomass in winter cultivars. The basic matrix comprised of 54 metabolites, 20 enzymes, and TKW. Only significant correlations at a false discovery (FD) rate <1% and <5% were selected as putative signatures (Figure 32).

Most amino acids were closely related to TKW. Threonine showed the strongest connection to TKW (pBH<0.001). Moreover, the end product, starch and soluble sugars Suc and Fru were positively related to TKW. Further positive connection was found for Succ, a TCA intermediate (Figure 32). Among enzymes, AGPase,

AlaAT, AspAT, cFBPase, PK, PFK, PGM, PGI, CWInv, PEPC, MDH, ICDH, G6PDH, 6PGDH and pGAPDH showed positive correlation to TKW (Figure 32).

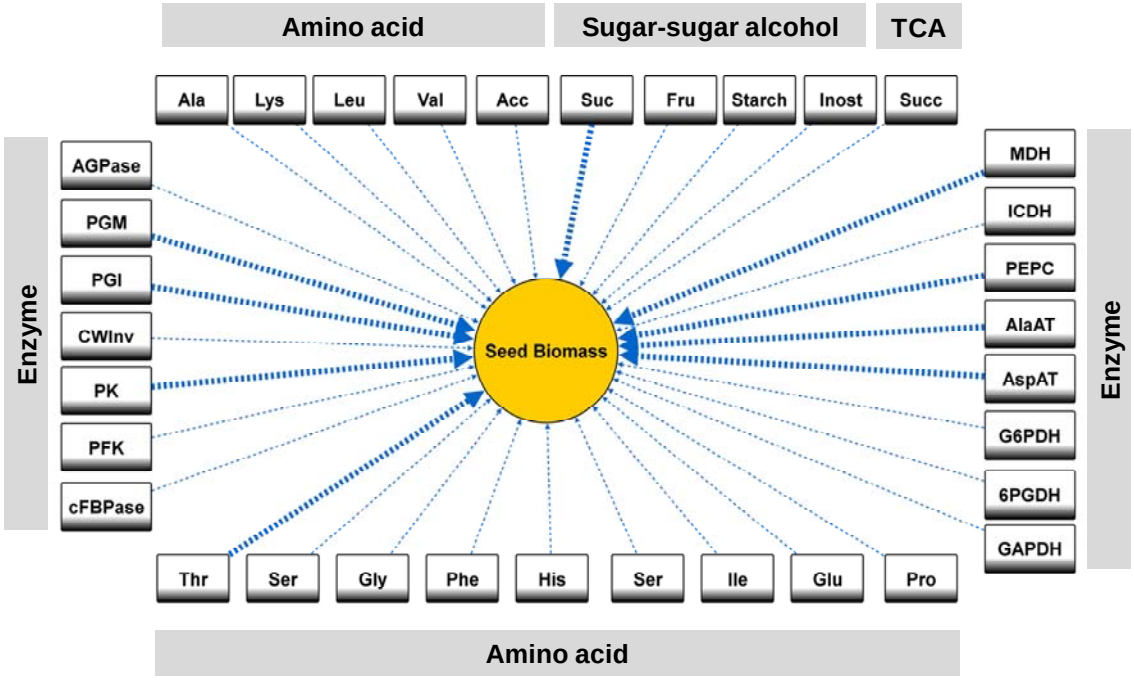


Figure 32. Correlations between metabolites, enzymes, and TKW at the generative stage in seeds of winter barley. Pearson product-moment correlation coefficient resulted in strong associations at the level of 0.01 and 0.001. Blue dotted lines indicate positive associations. Strong positive connections are shown using bold dotted lines.

5. Discussion

Biomass formation in plants is a highly complex process that has received much research attention in recent years related to the land delineation and distribution of crop grown for bioenergy production versus food and feed. However, in contrast to extensive studies on genetic markers for marker-assisted selection in breeding programs, metabolite based biomarkers are largely neglected in projects of plant breeding. Recently, few studies on breeding programs towards biomass formation have been performed using metabolomics and enzymomic approaches on shoot biomass in *Arabidopsis thaliana* (Bollina et al., 2011; Meyer et al., 2007). However, the question still remains to be addressed whether firstly different metabolite, enzyme and trancriptome signatures exist at various developmental stages of a crop plant such as barley, secondly whether these signatures differ from those found in *Arabidopsis* and finally whether these signatures can be used for a possible metabolic or genetic modification to improve plant biomass. To answer these questions, a comprehensive study of changes in metabolism occurring during three developmental stages of barley has been performed in the current study. To realize these achievements, different accession and introgression lines characterized by high or low biomass and morphological differences were used. All obtained results were correlated and combined with bioinformatics tools to find significant correlations in regard to biomass. In the following, the results for various stages of barley development will be discussed in detail.

5.1. Investigation of regulatory mechanisms at the vegetative stage of barley growth

5.1.1. Metabolites of tiller leaves correlate in structural and regulatory network levels

In the current study, the network correlation analysis identified a large number of connections between metabolites in tiller leaves (Figure 4). Although the metabolite network pattern in tiller leaves was reported for the first time in barley and under field trial conditions, correlation network regardless of positive or negative vectors

are similar to previous studies reported for tomato fruit (Carrari et al., 2006) and *Arabidopsis* rosettes recently published (Sulpice et al., 2010). These findings indicate that the high number of coordination between metabolites is probably due to global level (Sulpice et al., 2010), although experimental condition and the type of metabolites and their regulations strongly affect the network correlation pattern (Sulpice et al., 2013). Strong correlations within specific sectors of central metabolism were also detected in the current study, for instance amino acids of related functional class such as aromatic amino acids (Figure 4). This finding suggests that the significant correlations between mentioned amino acids are coordinated across biosynthetic families, and their usage and storage are strikingly co-regulated (Noctor et al., 2002).

Interestingly, strong correlations were found between Suc which is the main carbon source produced by photosynthesis, amino acids, and glycolytic metabolites (Figures 4 and 5). These results indicate a tight positive relationship of metabolic network between high carbohydrate contents, high glycolytic rate, and high amino acid levels in barley at the vegetative stage (Krapp and Stitt, 2006; Roessner-Tunali et al., 2003). The correlated network of metabolic traits also indicated negative relationships between metabolites involved in glycerophospholipid biosynthesis, fatty acids derived metabolites with nucleotides, sugar nucleotides, sugar phosphates (Figure 4) indicating that the biochemical regulation of biosynthetic routes drives the abundance of these intermediates in opposite directions as the site of carbon for the synthesis of building blocks for biomass production. Taken together a strong co-regulation of metabolites from different pathways in tiller leaves indicates that modification of certain metabolic route may have consequences for determination of the end-product of related pathways.

5.1.2. Enzyme activities and metabolites of central metabolism interact closely in tiller leaves

The network correlation pattern between enzyme activities of tiller leaves revealed a close positive connectivity for various pathways including glycolysis, TCA cycle, Calvin cycle, starch and sucrose metabolism (Figure 6). The strong coordination of enzymes in the metabolic network of tiller leaves in barley lines suggests that

modification of several enzymes within a pathway at the same time might be a more effective strategy for driving the flux instead of an individual enzyme. The minor effect of a single enzyme from Calvin cycle, sucrose and starch metabolism on pathway flux (Stitt et al., 2010) and altering of pathway flux by a large panel of enzymes of photosynthetic carbon metabolism have been previously shown in *Arabidopsis thaliana* and by the study of steady state of metabolic modelling (Zhu et al., 2007). In the current study, interestingly, the enzymes from the same metabolic pathway showed similar correlation pattern within and with other enzymes of different metabolic sectors. For example, enzymes of nitrogen and sucrose metabolisms, glycolysis and TCA cycle correlated with Calvin cycle enzymes demonstrating that Calvin cycle is a limiting step since the precursors for the synthesis of Suc are delivered by this pathway.

Metabolites and enzymes were also strongly related in the correlation matrix in tiller leaves with a large number of positive than negative relationships demonstrating that obviously enzymes of different pathways are directly connected to their substrates and/or end-products (Figure 7). These results show that the modification of a specific enzyme such as Susy may result in an increase of a desired end-product such as starch. Recently, this phenomenon has been supported in enzyme and metabolite profiling in tomato fruits (Steinhauser et al., 2010) and grape berry (Dai et al., 2013). Interestingly, most positive connections were between enzymes and related metabolites like invertases and Suc and their products, Glc, and Fru (Figure 7). This finding is in disagreement with earlier results presented in tomato under greenhouse conditions in which opposite relationships between substrates and products in the correlation matrices were observed (Steinhauser et al., 2010). Thus, it seems that metabolic networks of barley genotypes are strongly affected by growth and developmental stage. Therefore, it is suggested that the high frequency of positive correlations between products and substrates of enzymes are firstly confined to irreversible biochemical reactions, secondly they are dependent on levels of substrates and products that are not determined by the maximum enzymes activities (Sulpice et al., 2010) and finally they are due to flux alterations of many enzymes in the biochemical pathway (Kacser and Burns, 1973).

5.1.3. Enzyme activities and metabolites of central metabolism are tightly co-regulated with shoot biomass

The association of metabolic traits and biomass indicated strong relations of Suc and Susy, a sucrose-metabolizing enzyme, and PGM, which show interconnections between carbohydrate production and usage (Figure 8). The transport form of carbon in plants is Suc, which is formed normally in active green tissues and transported towards the sink organs. After cleavage into hexoses, Suc can be used directly as a carbon source or is converted to storage compounds such as starch. This double function in translocation and storage donates a crucial role for Suc in plant growth and development. This finding coincides with previous investigations indicating that a lack of Suc leads to growth depression upon starvation of Suc in tobacco (Chen et al., 2005) and potato (Fernie et al., 2002). However, other studies showed a negative connection between Suc and biomass formation in *Arabidopsis thaliana* under long day conditions which is disagreement with the findings in the current work (Meyer et al., 2007). The positive correlation between Suc and shoot biomass in barley at the vegetative stage is probably due to a higher activity of the sucrose metabolism or to its accumulation that may have a role in plant defence or adaptation to stress condition. Additionally, there are some evidences that Suc may act in a non-nutritive role in the regulation of gene expression level in cellular metabolism (Jang and Sheen, 1994; Koch, 2004). It has been shown in *Arabidopsis* that the application of Suc correlates with morphological traits like increase of proportion of leaves with hairs on their lower surface that are crucial for the transition to generative phase (Proveniers, 2013). In this study, tiller leaves were harvested at the end of vegetative stage. Thus, it is hypothesized that the increasing of metabolic activity and the accumulation of Suc presumably regulates the transition of barley from the vegetative phase to reproductive phase.

Another noteworthy relation was Susy, which has diverse roles in carbon partitioning in developing tissues. In poplar, it was shown that higher expression level of Susy correlates with higher cellulose and wood synthesis (Coleman et al., 2009; Hertzberg et al., 2001) and in accession lines of wheat, Susy was associated with increased cell wall polysaccharide levels (Xue et al., 2008). There is further

evidence that Susy correlates positively with plant biomass and flower development in *Nicotiana tabacum* (Coleman et al., 2010) and with an increase in height in *Populus alba* transformed with a modified mung bean Susy (Konishi et al., 2004). Similarly, down-regulation of Susy has been shown to produce smaller and fewer leaves in carrot plants (Tang and Sturm, 1999). Therefore, Susy plays a crucial role in the partitioning of carbon between source and sink organs leading to the suggestion that control of sink strength by increasing Susy activity may lead to biomass accumulation whereas suppression of Susy can lead to decreased growth.

Phosphoglucomutase (PGM) that catalyzes the readily reversible conversion of Glc1P and Glc6P correlated with biomass in this study. In potato plants with reduced PGM activity, the capacity of tuber for metabolizing of Suc was reduced resulting in a dramatically stunted phenotype, retardation of growth and marked reduction in tuber and leaf starch content (Fernie et al., 2002). Reduction in the PGM activity also leads to a marked reduction of photosynthetic rate, with a large effect on nucleotide biosynthesis (Fernie et al., 2002). Thus, it is predicted that the over-expression of PGM may results in a change of biomass allocation in both photosynthetic and non-photosynthetic organs and thus to enhanced biomass production.

Further highly correlated enzyme of central metabolism to shoot biomass was cFBPase, which along with SPS regulates sucrose synthesis in plants. cFBPase catalyzes the conversion of Fru1,6BP to Fru6P. A mutant line of *Flaveria linearis* lacking cFBPase partitioned two times less carbon into Suc and showed a decreased growth at the vegetative phase (Sharkey et al., 1992). A decrease in the activity of cFBPase also leads to a reduction of phosphorylated intermediates and in the ratio of cofactomes like ATP and ADP. Cho et al. (2012) indicated that transgenic *Arabidopsis* plants that simultaneously overexpressed cFBPase and triose phosphate/phosphate translocator showed increased photosynthetic activity and enhanced growth. Thus, increasing the efficiency of the conversion of assimilates into sugars and enhancing the photosynthetic activity of source leaves are key determinants to improve plant biomass and this may be applicable to develop crop plants with high yield.

The association of two enzymes of the pentose phosphate pathway, G6PDH and 6PGDH, which provide reducing equivalents and a variety of sugar molecules that are required for the biosynthesis of nucleic acids and amino acids revealed to be important for biomass formation. This is supported by previous reports that showed higher activities of G6PDH and 6PGDH enzymes triggering nucleotide biosynthesis and the activation of cellular components (Kanamori et al., 1979; Schröder et al., 2005). Therefore, G6PDH and 6PGDH might be good candidates for modulation of cofactomes and partitioning of biomass components in crop plants.

Malate dehydrogenase (MDH), the enzyme catalyzing the oxidation of Mal to OAA was found amongst positively correlated enzymes to shoot biomass. By contrast, Nunes-Nesi et al. (2005) reported a significant negative association between the activity of MDH and biomass formation in tomato plant. They indicated that transgenic tomato plants with reduced expression of MDH displayed an enhanced biomass, efficient photosynthesis and higher carbohydrate levels like Suc and starch but exhibited low respiration and produced low content of amino acids such as Asn, Glu and Met and organic components in long day condition (Nunes-Nesi et al., 2005). In coincidence with the relationship between MDH and biomass formation in the current research, a double knockout *of Arabidopsis* being deficient in mitochondrial mMDH1 and mMDH2 showed low biomass and low CO_2 assimilation (Tomaz et al., 2010) demonstrating that the function of MDH is species-dependent and must be evaluated in barley in future investigations. Interestingly, the significant relation of MDH to biomass coincided with the strong relationship of AspAT which is important for the conversion of Asp and 2-oxo to Glu and OAA. 2-oxo and Glu connect C to N metabolism. Glu in turn act as important amino group donor for the synthesis of other amino acids such as Gln, Pro and Arg, which serve as building block in protein synthesis (Roessner-Tunali et al., 2003).

Regarding metabolites, Phe, which is derived from chorismate and is used as precursor in the phenylpropanoid pathway for biosynthesis of monolignols, lignans and other intermediates, was found amongst correlated amino acids to biomass. It has been reported that the intermediates of phenylpropanoid pathway like, *p*-coumaric acid and caffeic acid were highly associated with lignin biosynthesis

(Riedelsheimer et al., 2012). Lignification is important for maintaining the integrity of plant cell walls and has a pivotal role in plant development and defence. The relation between metabolites involved in pathogen defence and biomass formation is plausible because resistant plants are likely to have more plant biomass to survive under field conditions especially during the vegetative phase.

Further highly positive relationships of metabolites to shoot biomass were found in the current study to be sugar alcohols Sbt and Inost. The latter one is a critical compound in the cell and has important roles in many processes, such as regulation of gene expression (Alcázar-Román and Wente, 2008), auxin receptor association (Tan et al., 2007) and stress tolerance. The oxidative form of Inost (glucuronic acid) is used for synthesis of pectin, hemicelluloses, and Xyl (Klinghammer and Tenhaken, 2007; Loewus et al., 1962).

5.1.4. Working model

The network correlation analysis between metabolic traits and shoot biomass resulted in the identification of 18 metabolites and seven enzymes that may play a crucial role in the biomass formation in barley. The possible roles of signatures in biomass formation are described in Figure 33. At the vegetative stage, the barley plants are photosynthetically active and produce large amounts of assimilates in their leaves. One of the most important assimilate produced is Suc which is a direct precursor of cellulose biosynthesis and thus it is associated with plant biomass (Kirst et al., 2004). The produced Suc is also metabolized by Susy to produce UDPGlc and Fru. UDPGlc is involved in cell wall polysaccharide biosynthesis such as hemicelluloses and pectins (Figure 33). The high association of PGM and ADPGlc which is produced from Glc1P is linked to the biosynthesis of leaf starch (Figure 33). Further candidates in primary metabolism with significant association to the shoot biomass were G6PDH and 6PGDH, which play a central role for the connection of glycolysis, lignin biosynthesis and fatty acid synthesis and for the delivery of redox equivalents to various pathways (Figure 33). Downstream of glycolysis, MDH and AspAT were identified as key enzymes for biomass production, which connect the TCA cycle and biosynthetic pathways of amino acids to ensure the biosynthesis of proteins (Figure 33). In addition, different metabolites

such as Pyr, Succ, 2-Oxo, Phe, Met, Asn and Lys (shown in red in Figure 33) were identified as key markers whose increase either by direct modification of the corresponding enzyme or by the expression of exogenous proteins may result in enhanced shoot biomass (Figure 33).

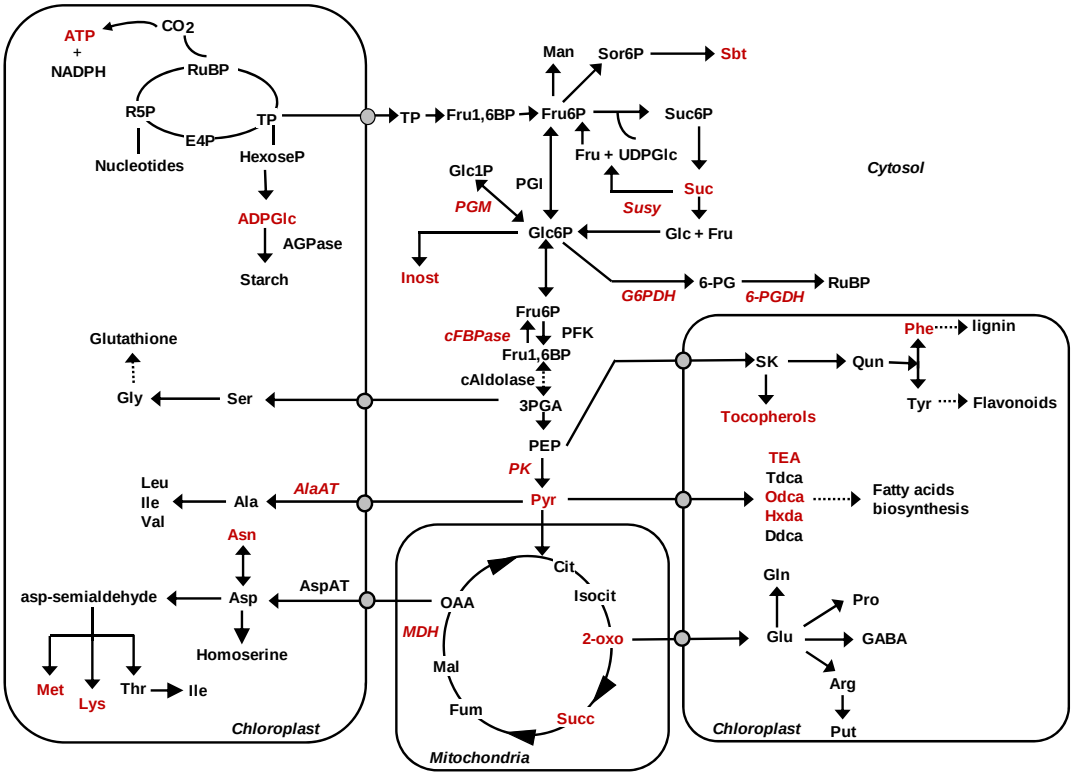


Figure 33. Schematic representation of the metabolic signatures involved in biomass formation at the vegetative stage of barley. The network correlation analysis identified the interconnection between biomass and central metabolism. Metabolites related to biomass are indicated with red colors. Enzymes signatures also are dipected italic and with red colors. Abbreviations are as follows: Suc: sucrose, Fru: fructose, Glc: glucose, Man: mannose, UDPGlc: UDP-glucose, Glc6P: glucose-6-phosphate, Fru1,6BP: Fructose-1,6-bisphosphate, Fru6P: fructose-6-phosphate, Suc6P: Sor6P: sorbose-6-phosphate, sucrose-6-phosphate, 3PGA: 3-phosphoglycerate, Glc1P: glucose-1-phosphate, ADPGlc: ADP-glucose, Pyr: Pyruvate, PEP: phosphoenolpyruvate, Succ: succinate, 2-oxo: 2-oxoglutarate,, Cit: citrate, Mal: malate, Isocit: isocitrate, Fum: fumarate, OAA: oxaloacetate, R5P: ribulose-5-phosphate, TP, triose-phosphate, RuBP:ribulose-1,5-bisphosphate, E4P: erythrose-4-phosphate, Ocda: octadecanoic acid, TEA: triethanolamine, Hxda, hexadecanoic acid, Tdca, tetradecanoic acid, Ddca:dodecanoic acid, Qun:quinate, SK:shikimate, Sbt: sorbitol, Inost: inositol, Gly:glycine, Ala:alanine, Ser: serine, Leu: leucine, Val:valine, Asp:aspartate, Thr: Threonine, Ile:isoleucine, Pro: proline, Gln:glutamine, Glu:glutamate, Phe: phenylalanine, Tyr:tyrosine, Arg:arginine, Put: putrescine, GABA: gamma aminobutyric acid, Lys: lysine, Met: Methionine, Asn: asparagines, 6-PG: 6-phosphogluconate, G6PDH: glucose-6-phosphate dehydrogenase,6-PGDH: 6-phosphogluconate dehydrogenase, AspAT: aspartate aminotransferase, AlaAT, alanine aminotransferase, cFBPase, cytosolic fructose-1,6-bisphosphatase, PGM, phosphoglucomutase, PGI:phosphoglucoisomerase, Susy: sucrose synthase, AGPase; ADP-glucose pyrophosphorylase, PK:pyruvate kinase, cAldolase:cytosolic aldolase, PFK:phosphofructokinase.

5.2. The role of metabolic networks in the regulation of shoot and seed biomass at generative stage in spring barley

5.2.1. Elucidation of metabolite and enzyme signatures relevant for shoot biomass

The strong relationship among metabolites are between metabolites of the same metabolic pathways or with similar chemical properties indicating a higher heritability and low variation between metabolic traits in flag leaves (Do et al., 2010). The close relationship among metabolites with the same chemical pattern have been previously observed in tomato plants on metabolites involved in fruit development (Carrari et al., 2006), and *Arabidopsis thaliana* in long day regime (Sulpice et al., 2010) or under high and low level nitrogen supply (Sulpice et al., 2013). In the current study, the highest number of positive associations of metabolites for specific metabolic routes was those for the glycolytic pathway, TCA cycle, and amino acid (Figures 9 and 10). The positive relationship between 2-oxo, Glu and Gln (Figure 10) suggests a genetic conservation of the regulatory network between carbon and nitrogen metabolism (Foyer et al., 2011).

The network relationship between enzymes indicated a large number of positive rather than negative associations (Figure 11). Suc is metabolized by a divergence of enzymes organized in vacuole, cytoplasm, and the cell wall. The different associations among sucrose metabolizing enzymes suggest that carbon-breakdown is catalyzed by different energy-conserving routes in a C3 plant like barley as previously shown in *Arabidopsis* (Keurentjes et al., 2008).

More importantly, enzymes in different metabolic sectors indicated coordinated changes of activity. For example, enzymes of the pentose phosphate pathway correlated strongly with enzymes of nitrogen metabolism. NO_3 is metabolized to NO_2 by the activity of nitrate reductase, which requires NADPH. The latter is generated from Glc6P by the action of G6PDH in pentose phosphate pathway (Wilson et al., 2007). The metabolic network shows that they are interconnected, although both are occurring in distinct pathways (Wilson et al., 2007).

Amongst enzymes of primary metabolism Calvin cycle and sucrose metabolizing enzymes exhibited significant relationship with the metabolites of TCA cycle,

glycolytic intermediates and amino acids (Figure 12). In plants, the main carbon source are starch, Suc, and Mal, which are mainly dependent on the activity of enzymes involved in Calvin cycle. Furthermore, the strong association between mentioned enzymes and metabolites reflects the fact that fluxes through these pathways are dependent on carbon fixation and breakdown via Calvin cycle enzymes and sucrose metabolizing enzymes. Furthermore, the network of enzymes for the upstream pathways including sucrose and glycolysis and metabolites of downstream pathways including TCA cycle were reflected by a positive correlation of cAldolase, PGI and PGM to TCA cycle intermediates demonstrating the importance of these enzymes as possible key reactions in driving the assimilates in flag leaves at the generative stage.

5.2.2. The relation of metabolites and enzyme signatures to biomass formation

It has been recently shown that metabolites and enzyme profiling have predictive power for biomass prediction under different growth conditions (Lisec et al., 2008; Riedelsheimer et al., 2012; Sulpice et al., 2010). Starch as a good indicator for growth rate is one of three important carbon sources, which is produced in the light and utilized in the night for growth and development. Mutants of *Arabidopsis*, deficient in starch biosynthesis or breakdown, exhibit a reduced plant growth (Stettler et al., 2009). It has further been shown that starch is an important growth regulator and strongly connected to biomass formation in *Arabidopsis thaliana* in a long day experiment and low carbon availability (Sulpice et al., 2009). In current study, the positive connection between starch and biomass, which is in contrast to its negative association and rosette fresh weight in *Arabidopsis* (Sulpice et al., 2009) reflects the fact that the correlation of metabolic traits and biomass is strongly dependent on species, conditions and developmental stages.

Trehalose-6-phosphate (Tre6P) which has an important role in the utilization of Suc and starch biosynthesis in the cell (Paul et al., 2008) was found in the current study to be highly and positively related to biomass. It has been shown that Tre6P regulates the biosynthesis of starch by increasing the activity of AGPase, the key enzyme controlling starch biosynthesis (Wingler et al., 2000). Furthermore, it has

been shown that Tre6P acts as signal in increasing of carbon availability and nitrogen remobilization to sink organ at the generative phase (O'Hara et al., 2013). Whether Tre6P is contributing to biomass determination as an intermediate or as a signalling molecule is still open and has to be elucidated in future experiments in barley plants. The high association of Tre6P and starch, or PGM, PGI and Glc6P to biomass (Figures 13 and 14) suggested a concomitance of active metabolic traits from photosynthesis to biosynthesis of starch and hexoses for cell component biosynthesis or for a vast majority of biochemical activities in the cell. This link is reinforced with the positive impact of Succ, Tacnt and Pyr on biomass (Figure 13).

Among amino acids, Ala and Asp were shown to accumulate to high level in source leaves. The synthesis of these amino acids is catalyzed by aminotransferases from which AlaAT was positively associated with biomass. it has been shown previously that the over-expression of barley AlaAT gene in *Brassica napus* (Good et al., 2007) and in rice (Shrawat et al., 2008) increases the absorption and efficient usage of nitrogen and leads to an increase of shoot biomass and grain yield in in-vitro and in the field under low nitrogen condition. The link between Pyr and shoot biomass and AlaAT also supports the idea that Pyr and translocated Ala allow a maintenance of carbon and nitrogen balance through the plant (Miyashita et al., 2007). Taken together, the relation between AlaAT and Asp, Gln which are the main amino acids of central metabolism may provide a link to connect other amino acids that are related to growth and development. Among them, Gly, which plays an important role in the biosynthesis of the antioxidant, glutathione, was identified. Glutathione is an important antioxidant and redox buffer in living organism that plays important roles in plant metabolism and defence. It has been shown that glutathione has a positive effect on growth and development of pollen in *Arabidopsis thaliana* (Zechmann et al., 2011). There is also evidence that the single point mutation of *Arabidopsis* for glutathione-S-transferase1 gene produces only 3% of the level of glutathione, and lose 50% of biomass compared to wild type (Xiang et al., 2001).

5.2.3. Working model

Figure 34 illustrates a simplified model for the interpretation of the predictive power of metabolic traits to shoot biomass at generative stage. In the current study, a metabolic signature including Suc starch, Glc6P, Tre6P, Pyr, Mal, Succ, 2-Oxo, Acnt and Qun could be identified that link different metabolic pathways starting with photosynthesis, Suc and starch biosynthesis, pentose phosphate pathway, glycolysis, shikimate pathway and the biosynthesis of cell wall components such as cellulose (Figure 34, red color). In addition, a metabolic signature was found between biomass and amino acids such as Gly necessary for glutathione synthesis and Ser, Ala, Ile, Val, Asp, Thr, Phe and Tyr. The latter two are used for the synthesis of lignin and flavonoids which are important for a better performance of plant growth towards stress and thus for the improvement of biomass production (Figure 34, red color). Furthermore, an enzyme signature including PGI, PGM, cAldolase, PFK, cFBPase and PK was found (Figure 34, red color and italic). All the enzyme activities are closely related to their corresponding metabolites indicating that both enzymes and metabolites can be used as possible indicators for biomass improvement.

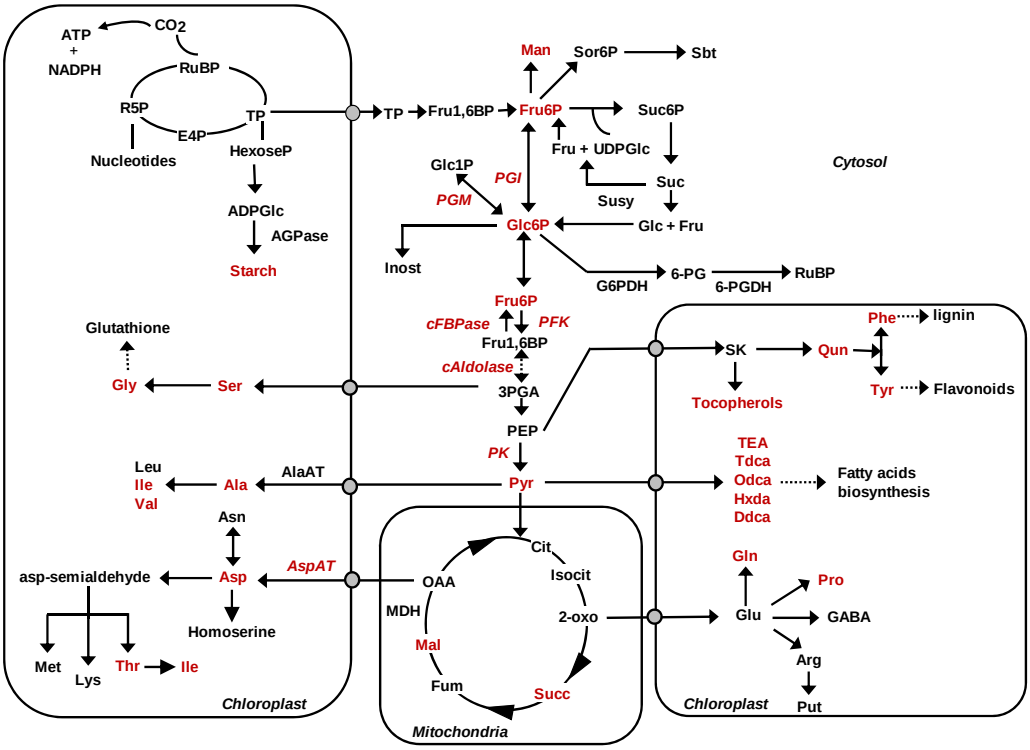


Figure 34. Schematic representation of the metabolic signatures involved in biomass formation at the generative stage of barley. The network correlation analysis identified the interconnection between biomass and central metabolism. Metabolites related to biomass are indicated with red colors. Enzymes signatures also are dipected italic and with red colors. Abbreviations are the same as given previously in Figure 33.

5.3. Comparative transcriptome profiling reveals the involvement of genes from different metabolic pathways for shoot and seed biomass formation

Most of studies on the role of leaves in growth and development in barley is limited to transcriptome analysis at the vegetative phase (Druka et al., 2006), or are applied to investigate the contributions of candidate genes in the molecular network underlying leaf senescence or nitrogen (Jukanti et al., 2008). In seeds, transcriptome analysis of seed have been confined to either a comparison of starch biosynthesis in seed compartment (Radchuk et al., 2009) or to a comparative study of seed transcripts at the early and late stages after pollination (Sreenivasulu et al., 2008). In a recent study, high throughput of RNA sequencing technology was applied to investigate a specific cluster of amino acid permeases in relation to sink-source organs in flag leaves and glumes in a barley plant cv.Barke, at the days before and after anthesis (Kohl et al., 2012).

In the current study, transcriptome analysis was profiled to firstly compare differentially expressed genes in barley plants contrasting in biomass production and secondly to identify carbon-regulated or carbon-responsive genes related to biomass formation at the generative phase in four different accessions of barley. To this end, a custom-made high-density microarray (Agilent technology) including a total number of 65,000 contigs sequences was used. With regard to the large number of the expressed genes, which are expressed in flag leaves and seeds, the focus was set on key genes involved in primary carbon and nitrogen metabolism. These differentially expressed genes may provide a predictive power for identification of candidate genes with integration of metabolome and enzymome data by means of transgenic approaches for future.

5.3.1. Flag leaves of genotype with highest biomass is enriched with the transcripts of energy metabolism

In flag leaves, a high number of up-regulated genes involved in photosynthesis were detected in plant with highest biomass compared to that with lowest biomass. This includes the pre-differentially expressed genes for the light reactions, Calvin cycle, and ATP synthesis (Figure 16). The light reactions take place in the thylakoid membrane of chloroplast and are divided in two photosystems. The energy of ATP and NADPH generated in light reaction is used to fix atmospheric CO_2 into the carbon skeletons via the Calvin cycle to fuel the rest of plant metabolism. Ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) was one of the highly represented transcripts followed by Sedoheptulose-1,6-phosphatase and pFBPase. The latter two enzymes are involved in the regeneration of Ribulose-1,6-bisPhosphate that is the immediate precursor for Rubisco (Harrison et al., 1997). Over-expression of Sedoheptulose 1,6-phosphatase and pFBPase resulted in high photosynthesis rate and dry biomass in tobacco supporting the idea that both enzymes are suitable candidate genes for modulation of shoot biomass in barley plants. Further overrepresented transcripts were highly up- and down-regulated genes of the complex IV in mitochondria. This contrasting regulation indicates that a complex transcriptional and

posttranscriptional process is occurring during mitochondria biogenesis as previously suggested at the germination phase in maize (Logan et al., 2001).

5.3.2. Genes involved in Suc and aromatic amino acid synthesis play a crucial role in biomass production in flag leaves

Out of the genes involved in carbohydrate metabolism, SPS was differentially expressed in the flag leaves (Figure 17 and Appendix Figure 2). SPS is a key regulatory enzyme in sucrose biosynthesis and catalyzes the reaction of UDPGlc and Fru6P to produce UDP and Suc6P. Suc6P is used as a substrate by the enzyme SPP to form Suc which occupies a central role as transport sugar, storage reserve, compatible solute, and signalling compound (Castleden et al., 2004). The enhancement of SPS activity may lead to an increase in photosynthetic activity and consequently to a higher partitioning of fixed carbon into Suc in source leaves which again increases the reproductive development (Micallef et al., 1995).

Other noteworthy up-regulated gene in the plant with highest biomass was mithocondrial pyruvate dehydrogenase complex, which links glycolytic carbon metabolism to the TCA cycle by catalyzing the oxidative decarboxylation of Pyr to acetyl-CoA. Pyruvate acts as the primary substrate for TCA, which, in turn, provides the reducing equivalents for ATP production by oxidative phosphorylation. The produced acetyl-coA is the starting point for the synthesis of fatty acids, isoprenoids and a number of secondary products (Marillia et al., 2003). The role of pyruvate dehydrogenase complex for fatty acid biosynthesis as biomass component in vegetative tissue has not been conclusively demonstrated. Thus, it might be suggested as new indicator for genetic modification for biomass formation.

Further up-regulated genes were transcripts involved in the biosynthesis of aromatic amino acids like Phe and Tyr (Figure 17 and Appendix Figure 3). Aromatic amino acids serve as important components in protein biosynthesis and are the major precursors for the biosynthesis of plant hormones, such as auxin and salicylate and of secondary metabolites. The role of Phe as precursor of phenylpropanoids pathway and lignin biosynthesis relies on its role in growth and defence response in plants and shows a close link with growth rate. The

overexpression of the genes involved in aromatic amino acid biosynthesis is in agreement with the higher expression of genes involved in biosynthesis of Gly-Ser-Cys, which are constitutes of glutathione. The level of glutathione associates with environmental stress, selective, and induction of genes involved in defence response and modulation of carbon partitioning and energy flux (Qi et al., 2010). Taken all together, these findings indicate close relationships between genes involved in storage reserve and defence response for higher biomass production.

5.3.3 Shoot biomass correlates strongly with genes involved in photosynthesis and sucrose metabolism

Energy metabolism

The correlation analysis between transcripts and shoot biomass revealed a complex of positive and negative correlations (Tables 2 and 3). This heterogeneity is presumable since the patterns of genes diurnally and developmentally change with switches in high and low expression level (Mangelsen et al., 2010). Of the highly relevant genes to biomass are those related to light reaction and Calvin cycle including cytochrome b6/f complexes, which passes the linear electron in the electron chain reaction to generate a proton gradient across the thylakoid membrane. This relation may indicate a tight feedback control between electron transfer and ATP synthase activity and biosynthesis and assembly of protein complex in nucleus and chloroplast which can latter on control carbon metabolism and light reactions (von Caemmerer and Evans, 2010).

Amino acid metabolism

Further highly connected transcript to shoot biomass is the gene encoding AspAT, which redistributes nitrogen and carbon pools between plant cell cytoplasm, mitochondria, and chloroplast compartments. Both negative and positive correlations for two isoforms of AspAT implies the fact that isforms of AspAT may respond differently to environmental conditions, and to carbon and nitrogen demand suggesting that they have specific and non-overlapping functions (Silvente et al., 2003). Transcripts of phosphoribosyl-AMP cyclohydrolase and 3-

phosphoserine phosphatase genes that play a role in the biosynthesis of the essential amino acid, His, were found amongst correlated transcripts to shoot biomass. It has been shown that His has a vital role in growth and development (Muralla et al., 2007) and resistance to metal toxicity in plant systems (Persans et al., 1999). Further transcript coding for acetoacetate synthase gene that correlated to shoot biomass was found. This enzyme catalyzes the synthesis of branched chain amino acids and leads to the production of Val, Leu, and Ile, which, in turn, increase protein synthesis and stimulate the cell division rate. The participation of acetoacetate synthase gene in herbicide resistance in barley (Lee et al., 2011) suggests that this gene is of great importance in stress responses (Ueda et al., 2008).

The transcript encoding glutamine synthetase, which catalyzes the first step in the conversion of inorganic nitrogen, ammonium, into its organic form Gln, was found amongst connected transcripts to shoot biomass. Glutamine synthetase plays a critical role in plant growth, plant productivity and crop yield (Mattsson et al., 1991). Isoforms of glutamine synthetase are reciprocally induced with Suc and amino acids in the cell. Thus, negative association between glutamine synthetase and shoot biomass probably indicates the reciprocal effects on glutamine synthetase expression and suggests that metabolite cues may activate cytosolic glutamine synthetase (Oliveira and Coruzzi, 1999).

Carbohydrate metabolism

A large number of the transcripts of genes involved in sucrose, starch metabolisms, and glycolysis strongly correlated with shoot biomass at generative stage. These include cAldolase, PK, and citrate synthase from glyoxylate and dicarboxylate metabolism, PGM and SPS from sucrose metabolism, starch synthase from starch metabolism, TPP from trehalose metabolism and ICDH and MDH (two informs) occurring in TCA cycle. Despite the posttranscriptional regulation of the enzymes of primary metabolism, transcripts and maximum enzyme activity of cAldolase, PK and PGM, correlated to shoot biomass.

Another highly associated transcript encoding TPP was found in this study. This gene is involved in trehalose metabolism and is supposed to be an important new

regulatory gene in plants (Vandesteene et al., 2012). Van Houtte et al. (2013) presented evidence that altering the gene expression of one of the plant endogenous TPP, the isoform *TPPB*, affects the shoot size of *Arabidopsis* plants during the vegetative and maturation phase. The resulting product of TPP action Tre operates as a major stress protectant of proteins and membranes during a wide range of stress conditions such as nutrient starvation, high salinity and dehydration (López-Gómez and Lluch, 2012) suggesting close relationship between plant growth and stress response.

Further interrelated transcripts were those encoding for the genes involved in TCA cycle such as citrate synthase, MDH and ICDH which serve to respire carbon and to generate NADH and ATP in the cell and to link carbon metabolism to nitrogen metabolism via 2-oxo. Citrate synthase catalyzes the first committed step of the TCA cycle with the condensation reaction occurring between the four-carbon oxaloacetate molecule and the two-carbon acetyl-CoA molecule. This activity is associated with citrate in root exudates in the interaction with soil (Delhaize et al., 2001; von Wirén et al., 1999) or fatty acid respiration and seed germination (Pracharoenwattana et al., 2005). Recently it was found that reduction of citrate synthase led to compromised nitrate assimilation without any significant effect on growth and photosynthesis performance (Sienkiewicz-Porzucek et al., 2008). Likewise, a reduced activity of MDH leads to a significant increase of growth rate and carbon partitioning in illuminated *Arabidopsis* leaf in contrast to citrate synthase and ICDH (Nunes-Nesi et al., 2005). These results indicate that MDH have greater consequence on growth and photosynthetic rate compared to citrate synthase and ICDH. In summary, in addition to already known enzymes of AGPase, 6PGDH, G6PDH, AlaAT, PGI, GAPDH and PGI, cAldolase and PK may also be used as new signatures to improve biomass production in barley.

5.4. Identification of metabolite and enzyme signatures for seed biomass in barley

The created network of metabolites in seeds resulted in a large number of highly correlated metabolites specifically for compounds occurring in the same metabolic routes (Figure 18). The correlation network analysis was a combination of

metabolites between sugars, amino acids, and organic compounds and led to a high coordination of positive correlations. On the other hand, these correlations apparently show a pattern of simultaneous changes of metabolites, which reflect a tight regulation of metabolic processes in the seed metabolism. There is evidence that the strong association of metabolic pattern during seed development is probably due to the maintenance of metabolic processes for storage allocation, balancing of carbon-nitrogen partitioning and reorganization of metabolism for maintaining of seed vigour (Toubiana et al., 2012). This is also in agreement with the significant coordinated change of metabolites found in *Arabidopsis* (Fait et al., 2006), and implies that any defect in seed metabolite regulation affects the basis of efficient regulation of metabolism in seeds (Borisjuk et al., 2002).

Enzymes in the seeds also showed a large number of positive associations (Figure 19). The connection of sucrose-metabolizing enzymes, AcidInv, CWInv, CytInv and Susy was different. While Susy strongly connected in the correlation matrix, AcidInv, CWInv and AcidInv showed weak connections (Figure 19). It has been previously reported that the activities of sucrose metabolising enzymes correlate in fava bean with developmental processes of sink tissues with the highest activity at the first days of developmental stage to low level after the storage processes (Weber et al., 1995). While the activity of three enzymes including CWInv, AcidInv and CytInv are relatively dependent on cellular differentiation of the seed, induction of Susy is an important determinant in storage associated metabolism (Borisjuk et al., 2004). There is also evidence that the activity of Susy correlates with sink strength of storage organs providing substrates for starch synthesis (Angeles-Núñez and Tiessen, 2010). The association between AGPase, a key enzyme in starch synthesis, PGM and PGI that provide substrate for AGPase and Susy also indicates a tight co-regulation of starch and sucrose metabolism.

Another significant associated enzyme to metabolites was pAldolase, which catalyzes the reversible conversion of dihydroxyacetone phosphate and 3PGA to Fru1,6BP (Figure 20). The role of pAldolase in carbon partitioning has been previously reported in potato (*Solanum tuberosum*) in which a small decrease in pAldolase activity resulted in the inhibition of photosynthesis and starch synthesis

(Haake et al., 1998). Thus, this enzyme might be also a suitable target for increasing the yield in monocotyledonous plants.

5.4.1. Interaction of metabolites and enzyme activities to seed biomass

Seed filling is an important phase of seed development process during which storage products accumulate and a high level of co-regulation in the maintenance and storage periods. The regulation of all these processes requires the concerted action of metabolites, which serve as metabolic signals, storage or structural compounds in cell process. In the current study, the investigation of highly correlated metabolites to seed biomass revealed that 53 metabolites and seven enzymes of the central metabolism were of great importance for the determination of final seed yield (Figure 21 and Figure 22). Similar to flag leaves, Suc was one of the highly related metabolites to seed biomass. The link between endosperm development and Suc synthesis is well documented (Hennen-Bierwagen and Myers, 2013) and supports the contribution of two important enzymes, SPS and sucrose phosphate phosphatase (SPP) to sink strength in the grain development (Bihmidine et al., 2013).

Further highly associated metabolite to seed biomass was starch, which is accumulated as energy reserve in cereals or serves as food, feed and renewable resource for bioenergy production (Sonnewald and Kossmann, 2013). ADPGlc, the direct precursor for AGPase was also found to be highly related to seed biomass. The maize mutant being deficient for ADPGlc transporter indicated reduced starch content in endosperm (Geigenberger, 2011). Thus, a combined approach to increase both the AGPase activity and the production of ADPGlc seems to be the most promising efficient tool for enhancing seed yield. The activity of AGPase during the seed development is consistent with the recently described enhancement of seed weight and starch content in transgenic maize plants (Li et al., 2011).

Further connected metabolites to seed biomass were fatty acid derived metabolites like glycerol, TEA, Ocda, Tdca whose final product constitutes 2% of seed weight in barley grain (Barthole et al., 2012). Fatty acids start to accumulate 13 days after flowering in embryo and endosperm of barley (Sreenivasulu et al., 2004). The

association of fatty acid intermediates and seed biomass is in accordance to the connections of glycolytic compounds such as PEP, which are the main carbon source for the biosynthesis of fatty acids. PK and enzymes of the pentose phosphate pathway, GAPDH and 6PGDH provide reducing equivalents NADPH required for fatty acid biosynthesis. This finding indicates that partitioning of carbon towards fatty acid biosynthesis and glycolytic metabolites are the driving force in fatty acid biosynthesis in barley grain.

Other noteworthy relations between metabolites and seed biomass are a high number of amino acids, which serve as important nitrogen sources and allocation forms of nitrogenous compounds between source and sink tissues (Lalonde et al., 2004). The availability of amino acids is an important prerequisite for plants to synthesize and to accumulate more protein and/or to use the amino acids as nitrogen source at the generative phase (Lawlor, 2002).

Regarding the enzymes Susy, along with Suc correlated strongly with the storage product starch. In the sink organ and the late storage phase in cereal, the hydrolysis of Suc is catalyzed by the activity of Susy which strongly connected to starch biosynthesis (Hennen-Bierwagen and Myers, 2013), polysaccharide biosynthesis of cell wall or fibre yield and reduction of seed abortion, a major problem in agriculture (Xu et al., 2012). The association of Susy to seed biomass is in contrast to the association of other sucrose hydrolyzing enzymes invertases, AcidInv and CWInv and supports the idea that the activity of different invertases correlate with the developmental processes in sink organ (Borisjuk et al., 2004). It has been shown that decreased activity of AcidInv correlates with the induction of cell differentiation at various developmental stages of carrot cells, whereas increased activity of the same enzyme can delay differentiation at the generative stage in barley (Weschke et al., 2003).

Three further enzymes involved in amino acid biosynthesis were identified to be highly associated with seed biomass. Two of them, AlaAT and AspAT play an important role in the synthesis of amino acids for the generation of storage proteins and accumulate during seed maturation (Kikuchi et al., 1999). It has been already shown that the heterologous expression of barley AlaAT produces a marked improvement in the nitrogen use efficiency of rice (Beatty et al., 2009). This finding

supports the growing argument that AlaAT has revolutionary importance for future crop production (Ridley, 2009). The role of AspAT as suitable indicator for seed biomass formation needs to be investigated in more detail. The other one GOGAT is the key enzyme of nitrogen assimilation during the barley grain development and is involved in the production of Glu, as a C and N source for the biosynthesis of most other amino acids (Goodall et al., 2013). It has been already shown that over-expression of GOGAT in rice resulted in 80% higher grain weight (Yamaya et al., 2002).

Another crucial enzyme was PEPC, which has an important role in carbon seed economy and fixes bicarbonate together with PEP to yield OAA that can be converted to Asp, Mal or other intermediates of the TCA cycle or can serve as a carbon acceptor for biosynthesis of Asp family, Thr, Ile, Met and Lys. The over-expression of PEPC in *Vicia narbonensis* increased protein content and might suggest that PEPC tightly regulated the carbon economy in seed (Rolletschek et al., 2004), probably also in barley seeds.

5.4.2. Working Model

Taken all together, observed results demonstrate that metabolites from various metabolic pathways of primary metabolism are associated strongly with seed biomass (Figure 35). As a metabolite signature Suc, Glc, Fru, Glc1P, Fruc6P, ADPGlc, Pyr, PEP, Mal, Succ, 2-Oxo and OAA, TEA, Tdca and Hxda was identified. For enzymes, a signature including PGM, PGI, AGPase, PK and MDH could be elucidated (Figure 35).

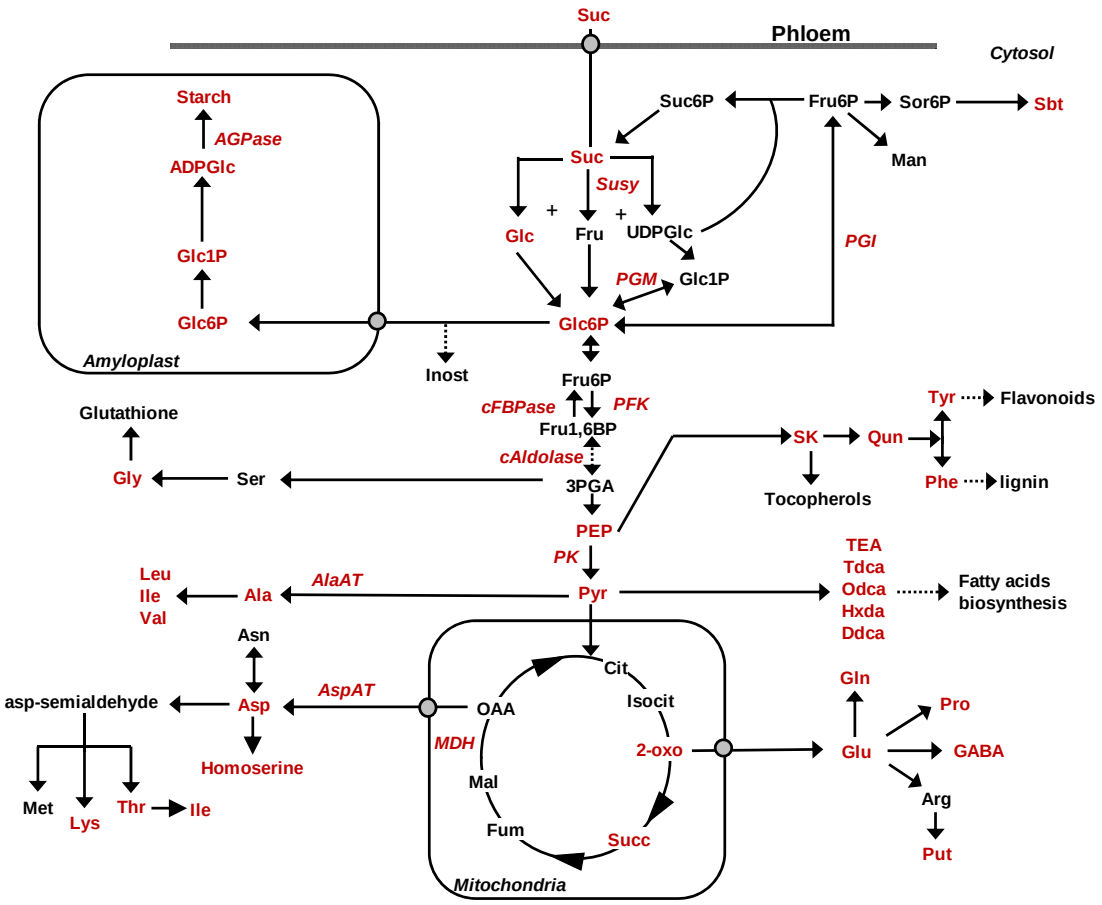


Figure 35. Schematic representation of the metabolic signatures involved in biomass formation at the generative stage of barley. The network correlation analysis identified the interconnection between biomass and central metabolism. Metabolites related to biomass are indicated with red colors. Enzymes signatures also are depicted italic and with red colors. Abbreviations are the same as given previously in Figure 33.

5.5. Seed development is enriched with the transcripts of starch, sucrose, amino acids and trehalose metabolisms in the plant with highest biomass

Transcript analysis of seeds revealed a very similar pattern as found for flag leaves by an over-representation of genes involved in light reactions, photosystem I and II, sucrose and starch metabolisms, trehalose metabolism (Figure 23) demonstrating that metabolic and regulatory mechanisms are similar in both flag leaves and seeds in barley. Further evaluation of the transcript analysis in seeds resulted in the identification of additional genes encoding for two isoforms of glutamine synthetases, GS1 and GS2, the key enzymes of nitrogen assimilation and remobilization which catalyze the fixation of ammonium to form the amino acid Gln

(Swarbreck et al., 2011). Hansen et al. (2009) showed that glutamine synthetase genes are highly expressed 15 days after flowering in wheat and had an interconnection with glutamate dehydrogenase, asparagine synthase and AspAT to produce intermediates needed for other amino acids and protein synthesis. The putative role of glutamine synthetase in crop productivity in maize was reported by (Hirel et al., 2001) who suggested a close relationship of GS to important agronomic traits such as TKW and seed weight. In addition, it was previously shown that the rice mutant lacking OsGST1;1, a cytosolic glutamine synthetase1;1 indicated reduced growth rate and grain filling (Tabuchi et al., 2005). Moreover, it has been already shown that GS1 in barley closely associated with a major QTL for grain protein content (Goodall et al., 2013). Thus, in addition to use as a marker for seed biomass, GS may also be used for the improvement of nutritional values of barley seeds.

Furthermore, other highly accumulated transcripts were found in seed that code for the genes of shikimate kinase, anthranilate phosphoribosyltransferase and Indole-3-glycerol phosphate synthase. Shikimate kinase catalyzes the conversion of shikimate to shikimate-3-phosphate whereas the latter two enzymes are involved in Trp and Phe synthesis. Aromatic amino acids Phe, Trp and Try are the major precursors for the synthesis of secondary metabolites and have in addition nutritional value in small grains (Petersen, 2013; Watanabe et al., 2013).

5.5.1 Correlation of seed biomass and transcripts in seeds are profoundly affected by sink-source interaction

Interaction between genes of amino acid metabolism and seed biomass

Transcripts of the genes implicated in amino acid metabolism extensively associated with seed biomass with large negative and low number of positive connections (Table 5). The large number of negative association for transcripts involved in the regulation of nitrogenous compounds is probably due to diurnal changes in metabolite composition between seed embryo and endosperm (Schauer et al., 2006). The transcript of the gene coding for anthranilate syntnase was highly associated with seed biomass. Anthranilate synthase catalyzes the

committed step of Trp biosynthesis converting chorismate and Gln to anthranilate, Glu and Pyr. The accumulation of Trp may affect the biosynthesis of other secondary metabolites and thus be expected to influence the synthesis of such metabolites and thereby elicit a pronounced change in metabolite profile. It has been already shown that the over-expression of *Oryza sativa* anthranilate syntnase (OSAD1) leads to the accumulation of Trp and stimulates the accumulation of other free amino acids in the seed of rice plants (Wakasa et al., 2006). Additional transcript accumulation was found for the gene encoding glutamate dehydrogenase, that is involved in the release of amino nitrogen from amino acids to give a keto-acid and NH_3 that can be separately recycled to be used in respiration and amide formation. This shows that the balance between C/N ratio has a major effect on seed biomass as remobilization of nitrogen is a dominate pathway during grain filling (Schlüter et al., 2012). It has been shown that higher activity of glutamate dehydrogenase in Triticale has a major effect on protein synthesis in seed coat (Kwinta et al., 1999). Two additional investigations indicated that the over-expression of *Aspergillus nidulans* glutamate dehydrogenase gene alters the amino acid content in tomato (Kisaka and Kida, 2003) and the photosynthesis rate, biomass and dry weight of tubers in potato plants (Egami et al., 2012) . In summary, the genes involved in the synthesis of aromatic amino acids and in the release of amino nitrogen are most probably determining the constitution of seed biomass in barley.

Interaction between genes of carbohydrate metabolism and seed biomass

Aldose-1-epimerase gene was highly associated with seed biomass. This enzyme catalyzes the epimerization of alpha-glucose to beta-glucose. Gomez et al. (2005) showed that the level of Tre6P correlates with the transcript of aldose-1-epimerase gene at the seed development phase in *Arabidopsis thaliana*. A large-scale analysis of 6500 *Arabidopsis* lines in eight level of Tre6P also uncovered a correlation between aldose-1-epimerase and Tre6P suggesting their associations in embryo development and seedling (Schluepmann et al., 2004).

Interestingly, high connection between transcripts and seed biomass was observed for a class of genes coding for PfkB-like carbohydrate kinase family protein whose

function are poorly characterized. It has been already reported that a member of PfkB-like carbohydrate kinase family, fructokinase-like proteins (FLN1 and FLN2) has a major role in chloroplast development in *Arabidopsis* and tobacco (Arsova et al., 2010). The study of natural variation for carbohydrate content in *Arabidopsis thaliana* also identified that PfkB-like carbohydrate kinase genes are co-localized with QTL for starch content (Calenge et al., 2006) suggesting their role as a putative candidate in starch metabolism.

Furthermore, the transcripts of genes coding for inositol-3-phosphate phosphatase and inositol synthase which are involved in inositol and phytic acid metabolisms, respectively, were found to be highly associated with seed biomass (Table 3). Phytic acid derived myo-inositol phosphate plays diverse roles in plants as signal transduction molecules, osmoprotectant, and cell wall constituents (Hegeman et al., 2001). Furthermore, phosphorus in seeds is stored primarily in the form of phytic acid. It is predicted that the expression of both enzymes together lead to a deposit of phytic acid which is used during germination to provide inorganic phosphate and *myo*-inositol to the growing seedling (Dionisio et al., 2007; Hegeman et al., 2001).

Transcript levels of genes coding for starch-related enzymes indicated a range of positive and negative associations showing their different potential during barley seed development or starch hydrolysis during seed germination (Radchuk et al., 2009). The gene encoding debranching enzyme was one of the prominent associated genes to seed biomass. The transcript of starch debranching enzyme was up-regulated 6 days after flowering in wheat, barley and rice and remained abundant until desiccation (Armonienė et al., 2013; Burton et al., 2002; Yun et al., 2011). Recently, it was demonstrated that a barley mutant lacking the isoamylase (ISA) activity was suppressed in phytoglycogen synthesis, and showed altered starch granule initiation and growth (Burton et al., 2002). Interestingly, transcripts of two genes encoding phosphoglycerate mutase (PGlyM) and citrate synthase that are involved in the glycolysis and TCA cycle, respectively, were highly associated with seed biomass. This result demonstrates that most probably a close co-regulation of carbon delivering and energy-producing pathways exist in seeds of barley. Taken all together, these results show clearly that not only genes involved

in biosynthetic pathways such as starch-metabolizing enzymes or phosphorous-producing enzymes but also genes participating in the regulation of seed development are of great importance for final seed yield.

5.6. Role of metabolic traits in shoot and seed biomass at the generative stage in winter barley

The metabolite profiling of winter barley was investigated to find out whether metabolite composition are influenced by the genetic background or different cultivation conditions and whether the network connectivity of metabolic traits is conserved in spring and winter barley. In accordance to spring barley, the network connectivity between metabolic traits showed a large number of positive associations with most metabolites (Figure 25) and enzymes in the same sectors of metabolic pathways (Figure 26). These connections were conserved between TCA intermediates, nucleotides, sugar nucleotides, and glycolytic metabolites in spring and winter barley showing a network connection in delivering of intermediates for vast majority of biochemical pathways. In contrast to spring barley, fatty acid derived metabolites showed a negative relationship to all metabolites mentioned above suggesting notable trade-offs of fatty acid intermediates in primary metabolism, particularly in winter barley. In addition, in winter barley, most of metabolic traits were negatively correlated with shoot biomass (Figure 28). Of the highly specific represented metabolites related to shoot biomass were Fru6P, Suc6P and Glc1P and the enzymes pFBPase, GAPDH and G6PDH which link carbon flow from photosynthesis, starch, and sucrose metabolism to cell wall biosynthesis. Pentose phosphate pathway also provides reducing power (NADPH) for nucleic acids, lignin and phenolic compounds, aromatic amino acids and fatty acid biosynthesis. It also delivers reductive agent for nitrogen assimilation in heterotrophic tissue (Bowsher et al., 1992). Members of TCA cycle such as, Cit, Isocit which provide energy and reducing power were found amongst unique metabolites. Further important metabolite was UDPGlc, which is a nucleotide sugar and acts as the donor for Glc in many glycolysation reactions, the formation of Suc, and biosynthesis of cellulose and callose. It is also used for the biosynthesis of glycoprotein and glycolipids or glycosylation of

compounds involved in stress responses or defence like betains, phenlypropanoids and terpenoids (Bennett and Wallsgrove, 1994; Kleczkowski et al., 2010). The biosynthesis of UDPGlc is linked to the activity of the sucrose hydrolysing enzymes, Susy and UGPase of which UGPase leads to an increase of cellulose content, sugar and starch and higher growth rate (Coleman et al., 2006). Other identified metabolites are Malna, Benza and glucua, which play a role in nucleotides and fatty acid biosynthesis. Amino acids indicated low contribution, with association of Asn which noteworthy play an important role in nitrogen assimilation and allocation within the plant and for Glu and Gln recovery in both source and sink tissues (Gaufichon et al., 2013).

In summary, the results for winter barley demonstrate that most metabolic traits are responding similarly in both spring and winter barley in regard to shoot biomass. However, additional important metabolic traits such as Fru6P, Suc6P, Glc1P, Man6P, UDPGlc, ADP, AMP, 3PGA and 2-oxo could be exclusively identified for winter barley as possible markers.

5.6.1. Thousand Kernel weight correlates with enzymes of central metabolism and amino acids in winter barley

Investigation of metabolite correlation to TKW showed that Suc and most amino acids were notably represented with Thr having the highest correlation (Figure 32). The integration of amino acids to TKW supports the idea that a link between nitrogen partitioning and crop yield exists (Toubiana et al., 2012). This is also exemplified with the noteworthy interaction of carbon-containing metabolites, amino acids and their network connections, which either supports the interaction of enzymes of central metabolism like sucrose, and starch metabolism, TCA cycle and pentose phosphate pathway to seed biomass or makes a cross-link between carbon and nitrogen metabolism during seed maturation. With other words, during the maturation phase, photoassimilates, such as Suc and amino acids are converted to the seed storage compounds starch, seed storage lipids and proteins (YongLing et al., 2006).

Among the highly associated enzymes to TKW were AGPase, PGM, PGI, PFK, PK, cFBPase, MDH, ICDH, PEPC which faciliate the biosynthesis of buliding block

for starch production and/or enegry nesessay for a vast majority of metaboic processes like protein or oil biosynthesis. In addition, AlaAT and AspAT were also identified with high association to TKW. These enzymes allow a balance between nitrogen and carbon. Taken all together, the high frequency of connections of amino acids and enzymes suggest that stored nitrogen and high N/C ratio is a dominate pathway at the generative stage and with high relationship to TKW in winter barley.

6. Zusammenfassung

Durch stätige Zunahme der Weltbevölkerung und die globale Veränderung des Klimas hat die Nachfrage nach Nahrungsmitteln und sauberer Energie eine weltweite Bemühung zur Verbesserung der Pflanzenenergie ausgelöst. Ein besseres Verständnis der fundamentalen Vorgänge zur Optimierung der photosynthetisch-fixierten Energie in Pflanzenbiomasse sind in erster Linie vielversprechend und vermindert den Wettstreit um Land zwischen Lebensmittel- und Futterproduktion auf der einen Seite und Bioenergie auf der anderen Seite. Basierend auf die Vorkenntnisse sind dennoch der Ausgangspunkt und die Strategien zur Erreichung dieses Ziels begrenzt. Deshalb wurde zum ersten Mal eine detaillierte Studie durchgeführt, um Vorhersagen für die Interaktionen zwischen Metaboliten, Enzymen und Transkripten in den vegetativen und generativen Phasen der Gerstenentwicklung zu treffen und diese eventuell in der Zukunft als Marker zum Optimieren der Pflanzenbiomasse und somit des Ertrags zu nutzen. Zu diesem Zweck wurden die wichtigsten Metabolite und Enzyme des zentralen Primärstoffwechsels und Transkriptveränderunegn durch Hochdurchsatz-Technologien an drei verschiedenen Entwicklungsstadien der Gerste unter natürlichen Bedingunegen im Feld bestimmt. Im folgenden sind die wichtigsten Ergebnisse dieser Arbeit zusammengefasst.

Sommergerste

Drei wichtige Entwicklungsphasen der Gerste einschließlich Bestockung (vegetativ, 30 Tage nach dem Einpflanzen), Kornfüllungsphase (generativ, 15 Tage nach der Blüte) und repining Phase (Reifungsphase) wurden für die Untersuchung der metabolischen und der enzymatischen Studien herangezogen. Zu diesem Zweck wurden 12 Gerste-Linien, die sich genetisch und geographisch unterschieden, unter agrarwirtschaflichen Anzuchtsbedingungen im Feld angebaut. Metabolitanalysen, die Messung der maximalen Enzymaktivität und die Transkriptom-Analyse lieferten Informationen, die eine enge Beziehung zwischen Enzymen, Metaboliten und Biomasse in verschiedenen Entwicklungsstadien als Vorhersage ermöglichten. Entschlüsselung der Netzwerk-Korrelation von Metaboliten und Enzymen in Tiller-Blättern der Gerste 30 Tage nach der Pflanzung

führte zur Identifizierung einer Reihe an Interaktionen zwischen den Metaboliten und Enzymen des TCA-Zyklus, der Glykolyse und dem Saccharose-Stoffwechsel. Dabei wurde eine metabolische Signatur aus 27 Metaboliten erstellt. Diese Interaktionen wurden vorwiegend bei Metaboliten und Enzymen gefunden, die eine wesentliche Rolle bei der Biosynthese von Biomasse-Bestandteilen wie Lignin, Zellulose, Proteine, Fettsäuren und Stärke spielen. Diese Erkenntnis führt zu der Schlussfolgerung, dass Metabolite in Source-Organen eine starke voraussagende Kraft als Biomarker besitzen und dementsprechend zur Selektion der Pflanzen bei der Züchtung eingesetzt werden könnten.

Die korrelative Netzwerkanalyse zwischen Metaboliten und Sprossbiomasse erbrachte 35 Metabolite und 8 Enzyme, die für die Produktion von Biomasse 15 Tage nach der Blüte wichtig sind. Diese Metabolite und Enzyme nehmen eine zentrale Stelle im Primärstoffwechsel ein. Die Netzwerkkorrelation zeigte auch eine starke Beziehung zwischen Saccharose als wichtige Kohlenhydratquelle und Trehalose als wichtiger Regulator der Stärke-Biosynthese. Transkriptom-Analyse des Fahnenblattes führte zur Identifizierung differentiell-exprimierter Gene, die an Licht- und Dunkelreaktionen und an der Synthese von Saccharose und aromatischen Aminosäuren beteiligt sind. Die Kombination der Netzwerkkorrelation zwischen Metaboliten und Genexpression führte zur Identifizierung von fünf Kandidaten, PGM, AlaAT, cAldolase und PK, die eine entscheidende Funktion bei der Biomassenbildung während der generativen Phase in Gerste spielen.

Die Analyse der Zusammensetzung der Metabolite und der Enzyme im Samen entschlüsselte 60 Marker. Die überwiegende Mehrheit ist am Stärke- (AGPase) und Saccharose-Stoffwechsel beteiligt (Suc und Susy). Die Beziehung zwischen den Aminosäuren und der Samenbiomasse unterstützt die Hypothese, dass Remobilisierung von Stickstoffhaltigen Verbindungen ein wichtiger Faktor in der Kornfüllungsphase ist. Transkriptom-Analyse der Samen führte zur Identifizierung der Gene, die vorwiegend an der Synthese von Stärke, Saccharose und Trehalose beteligt sind. Die Interaktion der metabolischen Merkmale und der Transkripte identifizierte fünf wichtige Enzyme inklusive AGPase, Susy, PGM, cAldolase und AlaAT. Diese Ergebnisse zeigen, dass Kohlenstoff- und Stickstoff-Stoffwechsel

eng miteinander verbunden sind und diese Enzyme als mögliche Biomarker zur Verbesserung der Biomasse in der generativen Phase fungieren könnten.
In der Reifungsphase wurden drei Biomarker inklusive Glutaminsäure, Glutamin und Stärke identifiziert. Diese Feststellung zeigt die bedeutende Funktion von zwei essentiellen Aminosäuren, die zur Synthese anderer stickstoffhaltigen Verbindungen im Samen und Stärke als wichtige Kohlenstoffquelle in der Endphase der Samenentwicklung.
Ein Vergleich zwischen den Signaturen für Tiller-Blätter, Fahnenblatt und Samen in Bezug auf Stengel- und Samen-Biomasse ergab, dass für Tiller-Blätter Susy, 6PGDH, G6PDH, MDH, AspAT und Suc, ADPGlc, 2-Oxo, Met, Lys, ATP, ADP und AMP als Enzym- und Metaboliten-Signaturen in der vegetativen Phase identifiziert werden konnten. Im Gegensatz zu den Signaturen der Tiller-Blätter konnten zusätzliche Enzyme, einschließlich pAldolase, cAldolase, PK, PFK, AlaAT und PGI als Enzym-Signatur und viele andere Aminosäuren, einschließlich Ser, Gln, Gly, Ala, Asp, Val, Ile, Thr, Tyr, His, Pro und Metabolite wie Tacnt, Glc6P, Mal, Glyct, Tre6P, Pglu, Qun, PPi, Tdca, Ddca, Stärke, Xyl, Fuc, Mann und Raff als Metabolit-Signatur in den Fahnenblättern gefunden werden (Generative Phase). Neben der Signatur in den Blättern wurden weitere Enzyme (AGPase, CWInv, CytInv, GOGAT, ICDH) und Metabolite (Glc1P, UDP, Amdp, Put, OAA, Dhdascb, GABA, Gal, PEP, ACC, Galcn) in den Samen identifiziert.

Wintergerstee

Die generative Phase war das einzige Entwicklungsstadium, das für die Analysen für Wintergerste ausgewählt wurde, um Metabolit- und Enzym-Signatuaren zu identifizieren, die die Bildung der Pflanzenspross- und Samen-Biomasse bestimmen. Eine kombinierte Analyse der metabolischen Merkmale im Fahnenblatt ergab eine enge Beziehung zwischen Enzymen und der Zwischenprodukte der Glykolyse, des TCA-Zyklus und der Nukleotide. Im Gegensatz zu Sommergerste zeigten die Wintersorten überwiegend eine negative Beziehung zwischen den Metaboliten, was auf einen starken Verbrauch der Metabolite in Wintergerste hindeutet.

Darüber hinaus konnten Man6P, 2-Oxo, Mal, Benza, pFBPase und pAldolase als metabolische Signatur in der Kornfüllungsphase in Wintergersten identifiziert werden. Die Analyse der Metabolite des Samens durch die Netzerkkorrelation-Methode ergab eine metabolische Signatur mit 34 verschiedenen Metaboliten zum Tausend-Korngewicht. Die zusammenhängende Funktion der Stickstoff-verbindungen deutet darauf hin, dass Remobilisierung von Stickstoffverbindungen einen wichtigen Beitrag zur Bestimmung des Tausend-Korngewichtes in Wintergerste leistet. Weiterhin wurde eine Enzym-Signatur für MDH, PEPC, AspAT, PGM, PGI und PK identifiziert, die in enger Beziehung zu Tausend-Korngewicht steht und als putative Biomarker im Samen benutzt werden könnte.

7. Literatures

Abebe, T., Melmaiee, K., Berg, V. and Wise, R.P. (2010) Drought response in the spikes of barley: gene expression in the lemma, palea, awn, and seed. *Functional & integrative genomics* **10**, 191-205.

Al-Babili, S., Ye, X., Lucca, P., Potrykus, I. and Beyer, P. (2001) Biosynthesis of β-Carotene (Provitamin a) in Rice Endosperm Achieved by Genetic Engineering. pp. 219-232. Wiley Online Library.

Alcázar-Román, A.R. and Wente, S.R. (2008) Inositol polyphosphates: a new frontier for regulating gene expression. *Chromosoma* **117**, 1-13.

Allison, P.D. (2002) Missing data: Quantitative applications in the social sciences. *British Journal of Mathematical and Statistical Psychology* **55**, 193-196.

Anastasiou, E. and Lenhard, M. (2007) Growing up to one's standard. *Current opinion in plant biology* **10**, 63-69.

Angeles-Núñez, J.G. and Tiessen, A. (2010) Arabidopsis sucrose synthase 2 and 3 modulate metabolic homeostasis and direct carbon towards starch synthesis in developing seeds. *Planta* **232**, 701-718.

Armonienė, R., Jonavičienė, K., Ruzgas, V. and Brazauskas, G. (2013) Isolation of differentially expressed genes in wheat caryopses with contrasting starch granule size. *Central European Journal of Biology* **8**, 297-305.

Arsova, B., Hoja, U., Wimmelbacher, M., Greiner, E., Üstün, Ş., Melzer, M., Petersen, K., Lein, W. and Börnke, F. (2010) Plastidial thioredoxin z interacts with two fructokinase-like proteins in a thiol-dependent manner: evidence for an essential role in chloroplast development in Arabidopsis and Nicotiana benthamiana. *The Plant Cell Online* **22**, 1498-1515.

Baik, B.K. and Ullrich, S.E. (2008) Barley for food: Characteristics, improvement, and renewed interest. *Journal of Cereal Science* **48**, 233-242.

Baldet, P., Hernould, M., Laporte, F., Mounet, F., Just, D., Mouras, A., Chevalier, C. and Rothan, C. (2006) The expression of cell proliferation-related genes in early developing flowers is affected by a fruit load reduction in tomato plants. *Journal of experimental botany* **57**, 961-970.

Barthole, G., Lepiniec, L., Rogowsky, P.M. and Baud, S. (2012) Controlling lipid accumulation in cereal grains. *Plant Science* **185**, 33-39.

Beatty, P.H., Shrawat, A.K., Carroll, R.T., Zhu, T. and Good, A.G. (2009) Transcriptome analysis of nitrogen-efficient rice over-expressing alanine aminotransferase. *Plant Biotechnology Journal* **7**, 562-576.

Benjamini, Y. and Hochberg, Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B (Methodological)*, 289-300.

Bennett, R.N. and Wallsgrove, R.M. (1994) Tansley review no. 72. Secondary metabolites in plant defence mechanisms. *New Phytologist*, 617-633.

Bihmidine, S., Hunter III, C.T., Johns, C.E., Koch, K.E. and Braun, D.M. (2013) Regulation of assimilate import into sink organs: Update on molecular drivers of sink strength. *Frontiers in plant science* **4**, 177.

Bollina, V., Kushalappa, A.C., Choo, T.M., Dion, Y. and Rioux, S. (2011) Identification of metabolites related to mechanisms of resistance in barley against Fusarium graminearum, based on mass spectrometry. *Plant molecular biology* **77**, 355-370.

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

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

Bowsher, C., Boulton, E., Rose, J., Nayagam, S. and Emes, M. (1992) Reductant for glutamate synthase in generated by the oxidative pentose phosphate pathway in non-photosynthetic root plastids. *The Plant Journal* **2**, 893-898.

Burton, R.A., Jenner, H., Carrangis, L., Fahy, B., Fincher, G.B., Hylton, C., Laurie, D.A., Parker, M., Waite, D. and Van Wegen, S. (2002) Starch granule initiation and growth are altered in barley mutants that lack isoamylase activity. *The Plant Journal* **31**, 97-112.

Calenge, F., Saliba-Colombani, V., Mahieu, S., Loudet, O., Daniel-Vedele, F. and Krapp, A. (2006) Natural variation for carbohydrate content in Arabidopsis. Interaction with complex traits dissected by quantitative genetics. *Plant physiology* **141**, 1630-1643.

Carrari, F., Baxter, C., Usadel, B., Urbanczyk-Wochniak, E., Zanor, M.I., Nunes-Nesi, A., Nikiforova, V., Centero, D., Ratzka, A. and Pauly, M. (2006) Integrated analysis of metabolite and transcript levels reveals the metabolic shifts that underlie tomato fruit development and highlight regulatory aspects of metabolic network behavior. *Plant physiology* **142**, 1380-1396.

Castleden, C.K., Aoki, N., Gillespie, V.J., MacRae, E.A., Quick, W.P., Buchner, P., Foyer, C.H., Furbank, R.T. and Lunn, J.E. (2004) Evolution and function of the sucrose-phosphate synthase gene families in wheat and other grasses. *Plant physiology* **135**, 1753-1764.

Chen, S., Hajirezaei, M., Peisker, M., Tschiersch, H., Sonnewald, U. and Börnke, F. (2005) Decreased sucrose-6-phosphate phosphatase level in transgenic tobacco inhibits photosynthesis, alters carbohydrate partitioning, and reduces growth. *Planta* **221**, 479-492.

Cho, M.-H., Jang, A., Bhoo, S.H., Jeon, J.-S. and Hahn, T.-R. (2012) Manipulation of triose phosphate/phosphate translocator and cytosolic fructose-1, 6-bisphosphatase, the key components in photosynthetic sucrose synthesis, enhances the source capacity of transgenic Arabidopsis plants. *Photosynthesis research* **111**, 261-268.

Choe, S., Fujioka, S., Noguchi, T., Takatsuto, S., Yoshida, S. and Feldmann, K.A. (2001) Overexpression of DWARF4 in the brassinosteroid biosynthetic pathway results in increased vegetative growth and seed yield in Arabidopsis. *The Plant Journal* **26**, 573-582.

Coleman, H.D., Beamish, L., Reid, A., Park, J.-Y. and Mansfield, S.D. (2010) Altered sucrose metabolism impacts plant biomass production and flower development. *Transgenic research* **19**, 269-283.

Coleman, H.D., Ellis, D.D., Gilbert, M. and Mansfield, S.D. (2006) Up-regulation of sucrose synthase and UDP-glucose pyrophosphorylase impacts plant growth and metabolism. *Plant Biotechnology Journal* **4**, 87-101.

Coleman, H.D., Yan, J. and Mansfield, S.D. (2009) Sucrose synthase affects carbon partitioning to increase cellulose production and altered cell wall ultrastructure. *Proceedings of the National Academy of Sciences* **106**, 13118-13123.

Cross, J.M., Von Korff, M., Altmann, T., Bartzetko, L., Sulpice, R., Gibon, Y., Palacios, N. and Stitt, M. (2006) Variation of enzyme activities and metabolite levels in 24 Arabidopsis accessions growing in carbon-limited conditions. *Plant physiology* **142**, 1574-1588.

Dagogo-Jack, S. (2012) Metabolomic Prediction of Diabetes and Cardiovascular Risk. *Medical Principles and Practice* **21**, 401-403.

Dai, Z.W., Léon, C., Feil, R., Lunn, J.E., Delrot, S. and Gomès, E. (2013) Metabolic profiling reveals coordinated switches in primary carbohydrate metabolism in grape berry (*Vitis vinifera* L.), a non-climacteric fleshy fruit. *Journal of experimental botany* **64**, 1345-1355.

Delhaize, E., Hebb, D.M. and Ryan, P.R. (2001) Expression of a Pseudomonas aeruginosa citrate synthase gene in tobacco is not associated with either enhanced citrate accumulation or efflux. *Plant physiology* **125**, 2059-2067.

Devos, K.M. and Gale, M.D. (2000) Genome relationships: the grass model in current research. *The Plant Cell* **12**, 637.

Dionisio, G., Holm, P.B. and Brinch-Pedersen, H. (2007) Wheat (*Triticum aestivum* L.) and barley (Hordeum vulgare L.) multiple inositol polyphosphate phosphatases (MINPPs) are phytases expressed during grain filling and germination†. *Plant Biotechnology Journal* **5**, 325-338.

Do, P.T., Prudent, M., Sulpice, R., Causse, M. and Fernie, A.R. (2010) The influence of fruit load on the tomato pericarp metabolome in a Solanum chmielewskii introgression line population. *Plant physiology* **154**, 1128-1142.

Druka, A., Muehlbauer, G., Druka, I., Caldo, R., Baumann, U., Rostoks, N., Schreiber, A., Wise, R., Close, T. and Kleinhofs, A. (2006) An atlas of gene expression from seed to seed through barley development. *Functional & integrative genomics* **6**, 202-211.

Egami, T., Wakayama, M., Aoki, N., Sasaki, H., Kisaka, H., Miwa, T. and Ohsugi, R. (2012) The effects of introduction of a fungal glutamate dehydrogenase gene (gdhA) on the photosynthetic rates, biomass, carbon and nitrogen contents in transgenic potato. *Plant Biotechnology* **29**, 57-64.

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

Faix, B., Radchuk, V., Nerlich, A., Hümmer, C., Radchuk, R., Emery, R., Keller, H., Götz, K.P., Weschke, W. and Geigenberger, P. (2011) Barley grains, deficient in cytosolic small subunit of ADP-glucose pyrophosphorylase, reveal coordinate adjustment of C: N metabolism mediated by an overlapping metabolic-hormonal control. *The Plant Journal*.

Fernie, A.R. and Schauer, N. (2009) Metabolomics-assisted breeding: a viable option for crop improvement? *Trends in Genetics* **25**, 39-48.

Fernie, A.R., Tauberger, E., Lytovchenko, A., Roessner, U., Willmitzer, L. and Trethewey, R.N. (2002) Antisense repression of cytosolic phosphoglucomutase in potato (*Solanum tuberosum*) results in severe growth retardation, reduction in tuber number and altered carbon metabolism. *Planta* **214**, 510-520.

Fisher, D.H. (1987) Knowledge acquisition via incremental conceptual clustering. *Machine learning* **2**, 139-172.

Foyer, C.H., Noctor, G. and Hodges, M. (2011) Respiration and nitrogen assimilation: targeting mitochondria-associated metabolism as a means to enhance nitrogen use efficiency. *Journal of experimental botany* **62**, 1467-1482.

Gaufichon, L., Masclaux, D., Celine, T., Guillaume, R., Cren, M., Sakakibara, Y.H., Toshiharu, C., Gilles , A., Jean, C.G. and Olivier Marmagne, A. (2013) Arabidopsis thaliana ASN2 encoding asparagine synthetase is involved in the control of nitrogen assimilation and export during vegetative growth. *Plant, Cell & Environment* **36**, 328-342.

Ge, H., Walhout, A.J.M. and Vidal, M. (2003) Integrating ' omic'information: a bridge between genomics and systems biology. *Trends in Genetics* **19**, 551-560.

Geigenberger, P. (2011) Regulation of starch biosynthesis in response to a fluctuating environment. *Plant physiology* **155**, 1566-1577.

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

Gibson, S.I. (2005) Control of plant development and gene expression by sugar signaling. *Current opinion in plant biology* **8**, 93-102.

Gómez, L.D., Baud, S., Gilday, A., Li, Y. and Graham, I.A. (2006) Delayed embryo development in the Arabidopsis trehalose 6-phosphate synthase 1 mutant is associated with altered cell wall structure, decreased cell division and starch accumulation. *The Plant Journal* **46**, 69-84.

Good, A.G., Johnson, S.J., De Pauw, M., Carroll, R.T., Savidov, N., Vidmar, J., Lu, Z., Taylor, G. and Stroeher, V. (2007) Engineering nitrogen use efficiency with alanine aminotransferase. *Botany* **85**, 252-262.

Goodall, A.J., Kumar, P. and Tobin, A.K. (2013) Identification and expression analyses of cytosolic glutamine synthetase genes in barley (*Hordeum vulgare* L.). *Plant and Cell Physiology* **54**, 492-505.

Haake, V., Zrenner, R., Sonnewald, U. and Stitt, M. (1998) A moderate decrease of plastid aldolase activity inhibits photosynthesis, alters the levels of sugars and starch, and inhibits growth of potato plants. *The Plant Journal* **14**, 147-157.

Hajirezaei, M., Sonnewald, U., Viola, R., Carlisle, S., Dennis, D. and Stitt, M. (1993) Transgenic potato plants with strongly decreased expression of pyrophosphate: fructose-6-phosphate phosphotransferase show no visible phenotype and only minor changes in metabolic fluxes in their tubers. *Planta* **192**, 16-30.

Hajirezaei, M.R., Takahata, Y., Trethewey, R.N., Willmitzer, L. and Sonnewald, U. (2000) Impact of elevated cytosolic and apoplastic invertase activity on carbon metabolism during potato tuber development. *Journal of experimental botany* **51**, 439-445.

Hammer, G., Cooper, M., Tardieu, F., Welch, S., Walsh, B., van Eeuwijk, F., Chapman, S. and Podlich, D. (2006) Models for navigating biological complexity in breeding improved crop plants. *Trends in plant science* **11**, 587-593.

Hansen, M., Friis, C., Bowra, S., Holm, P.B. and Vincze, E. (2009) A pathway-specific microarray analysis highlights the complex and co-ordinated transcriptional networks of the developing grain of field-grown barley. *Journal of experimental botany* **60**, 153-167.

Harrison, E.P., Willingham, N.M., Lloyd, J.C. and Raines, C.A. (1997) Reduced sedoheptulose-1, 7-bisphosphatase levels in transgenic tobacco lead to decreased photosynthetic capacity and altered carbohydrate accumulation. *Planta* **204**, 27-36.

Hegeman, C.E., Good, L.L. and Grabau, E.A. (2001) Expression of D-myo-inositol-3-phosphate synthase in soybean. Implications for phytic acid biosynthesis. *Plant physiology* **125**, 1941-1948.

Heidecker, B. and Hare, J.M. (2007) The use of transcriptomic biomarkers for personalized medicine. *Heart failure reviews* **12**, 1-11.

Hennen-Bierwagen, T.A. and Myers, A.M. (2013) Genomic Specification of Starch Biosynthesis in Maize Endosperm. *Seed Genomics*, 123-137.

Hertzberg, M., Aspeborg, H., Schrader, J., Andersson, A., Erlandsson, R., Blomqvist, K., Bhalerao, R., Uhlén, M., Teeri, T.T. and Lundeberg, J. (2001) A transcriptional roadmap to wood formation. *Proceedings of the National Academy of Sciences* **98**, 14732-14737.

Hirai, M.Y., Yano, M., Goodenowe, D.B., Kanaya, S., Kimura, T., Awazuhara, M., Arita, M., Fujiwara, T. and Saito, K. (2004) Integration of transcriptomics and metabolomics for understanding of global responses to nutritional stresses in Arabidopsis thaliana. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 10205.

Hirel, B., Bertin, P., Quilleré, I., Bourdoncle, W., Attagnant, C., Dellay, C., Gouy, A., Cadiou, S., Retailliau, C. and Falque, M. (2001) Towards a better understanding of the genetic and physiological basis for nitrogen use efficiency in maize. *Plant physiology* **125**, 1258-1270.

Holtum, J.A. and Winter, K. (1982) Activity of enzymes of carbon metabolism during the induction of Crassulacean acid metabolism in Mesembryanthemum crystallinum L. *Planta* **155**, 8-16.

Hucka, M., Finney, A., Sauro, H.M., Bolouri, H., Doyle, J.C., Kitano, H., Arkin, A.P., Bornstein, B.J., Bray, D. and Cornish-Bowden, A. (2003) The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics* **19**, 524-531.

Jang, J.-C. and Sheen, J. (1994) Sugar sensing in higher plants. *The Plant Cell Online* **6**, 1665-1679.

Jenner, H.L., Winning, B.M., Millar, A.H., Tomlinson, K.L., Leaver, C.J. and Hill, S.A. (2001) NAD malic enzyme and the control of carbohydrate metabolism in potato tubers. *Plant physiology* **126**, 1139-1149.

Jukanti, A.K., Heidlebaugh, N.M., Parrott, D.L., Fischer, I.A., McInnerney, K. and Fischer, A.M. (2008) Comparative transcriptome profiling of near-isogenic barley (Hordeum vulgare) lines differing in the allelic state of a major grain protein content locus identifies genes with possible roles in leaf senescence and nitrogen reallocation. *New Phytologist* **177**, 333-349.

Junker, B., Klukas, C. and Schreiber, F. (2006) VANTED: A system for advanced data analysis and visualization in the context of biological networks. *BMC bioinformatics* **7**, 109.

Kacser, H. and Burns, J. (1973) The control of flux. In: *Symp Soc Exp Biol* pp. 65-104.

Kanamori, I., Ashihara, H. and Komamine, A. (1979) Changes in the Activities of the Pentose Phosphate Pathway and Pyrimidine Nucleotide Biosynthesis during the Growth of *Vinca rosea* Cells in Suspension Culture. *Zeitschrift für Pflanzenphysiologie* **93**, 437-448.

Kellogg, E.A. (2001) Evolutionary history of the grasses. *Plant physiology* **125**, 1198-1205.

Kelly, G.J., Zimmermann, G. and Latzko, E. (1982) [62] Fructose-bisphosphatase from spinach leaf chloroplast and cytoplasm. *Methods in enzymology* **90**, 371-378.

Keurentjes, J.J.B., Sulpice, R., Gibon, Y., Steinhauser, M.C., Fu, J., Koornneef, M., Stitt, M. and Vreugdenhil, D. (2008) Integrative analyses of genetic variation in enzyme activities of primary carbohydrate metabolism reveal distinct modes of regulation in Arabidopsis thaliana. *Genome Biol* **9**, R129.

Kikuchi, H., Hirose, S., Toki, S., Akama, K. and Takaiwa, F. (1999) Molecular characterization of a gene for alanine aminotransferase from rice (*Oryza sativa*). *Plant molecular biology* **39**, 149-159.

Kirst, M., Myburg, A.A., De León, J.P., Kirst, M.E., Scott, J. and Sederoff, R. (2004) Coordinated genetic regulation of growth and lignin revealed by quantitative trait locus analysis of cDNA microarray data in an interspecific backcross of eucalyptus. *Plant physiology* **135**, 2368-2378.

Kisaka, H. and Kida, T. (2003) Transgenic tomato plant carrying a gene for NADP-dependent glutamate dehydrogenase (*gdhA*) from *Aspergillus nidulans*. *Plant Science* **164**, 35-42.

Kleczkowski, L.A., Kunz, S. and Wilczynska, M. (2010) Mechanisms of UDP-glucose synthesis in plants. *Critical reviews in plant sciences* **29**, 191-203.

Klinghammer, M. and Tenhaken, R. (2007) Genome-wide analysis of the UDP-glucose dehydrogenase gene family in Arabidopsis, a key enzyme for matrix polysaccharides in cell walls. *Journal of experimental botany* **58**, 3609-3621.

Koch, K. (2004) Sucrose metabolism: regulatory mechanisms and pivotal roles in sugar sensing and plant development. *Current opinion in plant biology* **7**, 235-246.

Kohl, S., Hollmann, J., Blattner, F., Radchuk, V., Andersch, F., Steuernagel, B., Schmutzer, T., Scholz, U., Krupinska, K. and Weber, H. (2012) A putative role for amino acid permeases in sink-source communication of barley tissues uncovered by RNA-seq. *BMC plant biology* **12**, 154.

Konishi, T., Ohmiya, Y. and Hayashi, T. (2004) Evidence that sucrose loaded into the phloem of a poplar leaf is used directly by sucrose synthase associated with various β-glucan synthases in the stem. *Plant physiology* **134**, 1146-1152.

Krapp, A. and Stitt, M. (2006) Influence of high carbohydrate content on the activity of plastidic and cytosolic isoenzyme pairs in photosynthetic tissues. *Plant, Cell & Environment* **17**, 861-866.

Kwinta, J., Bartoszewicz, K. and Bielawski, W. (1999) Glutamate dehydrogenase and glutamine synthetase activities during the development of triticale grains. *Acta physiologiae plantarum* **21**, 271-275.

Lalonde, S., Wipf, D. and Frommer, W.B. (2004) Transport mechanisms for organic forms of carbon and nitrogen between source and sink. *Annu Rev Plant Biol* **55**, 341-372.

Lancashire, P., Bleiholder, H., Boom, T., Langelüddeke, P., Stauss, R., WEBER, E. and Witzenberger, A. (1991) A uniform decimal code for growth stages of crops and weeds. *Annals of Applied Biology* **119**, 561-601.

Lawlor, D.W. (2002) Carbon and nitrogen assimilation in relation to yield: mechanisms are the key to understanding production systems. *Journal of experimental botany* **53**, 773-787.

Lee, H., Rustgi, S., Kumar, N., Burke, I., Yenish, J.P., Gill, K.S., von Wettstein, D. and Ullrich, S.E. (2011) Single nucleotide mutation in the barley acetohydroxy acid synthase (AHAS) gene confers resistance to imidazolinone herbicides. *Proceedings of the National Academy of Sciences* **108**, 8909-8913.

Li, N., Zhang, S., Zhao, Y., Li, B. and Zhang, J. (2011) Over-expression of AGPase genes enhances seed weight and starch content in transgenic maize. *Planta* **233**, 241-250.

Lisec, J., Meyer, R.C., Steinfath, M., Redestig, H., Becher, M., Witucka-Wall, H., Fiehn, O., Törjék, O., Selbig, J. and Altmann, T. (2008) Identification of metabolic and biomass QTL in Arabidopsis thaliana in a parallel analysis of RIL and IL populations. *The Plant Journal* **53**, 960-972.

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

Loewus, F.A., Kelly, S. and Neufeld, E.F. (1962) Metabolism of myo-inositol in plants: conversion to pectin, hemicellulose, D-xylose, and sugar acids. *Proceedings of the National Academy of Sciences of the United States of America* **48**, 421.

Logan, D.C., Millar, A.H., Sweetlove, L.J., Hill, S.A. and Leaver, C.J. (2001) Mitochondrial biogenesis during germination in maize embryos. *Plant physiology* **125**, 662-672.

López-Gómez, M. and Liuch, C. (2012) Trehalose and abiotic stress tolerance. In: *Abiotic Stress Responses in Plants* pp. 253-265. Springer.

Mangelsen, E., Wanke, D., Kilian, J., Sundberg, E., Harter, K. and Jansson, C. (2010) Significance of light, sugar, and amino acid supply for diurnal gene regulation in developing barley caryopses. *Plant physiology* **153**, 14-33.

Marillia, E.F., Micallef, B.J., Micallef, M., Weninger, A., Pedersen, K.K., Zou, J. and Taylor, D.C. (2003) Biochemical and physiological studies of Arabidopsis thaliana transgenic lines with repressed expression of the mitochondrial pyruvate dehydrogenase kinase1. *Journal of experimental botany* **54**, 259-270.

Mattsson, M., Johansson, E., Lundborg, T., Larsson, M. and Larsson, C.-M. (1991) Nitrogen Utilization in N-limited Barley during Vegetative and Generative Growth I. Growth and nitrate uptake kinetics in vegetative cultures growth at different relative addition rates of nitrate-N. *Journal of experimental botany* **42**, 197-205.

Mayer, K., Waugh, R., Brown, J., Schulman, A., Langridge, P., Platzer, M., Fincher, G.B., Muehlbauer, G.J., Sato, K. and Close, T.J. (2012) A physical, genetic and functional sequence assembly of the barley genome. *Nature* **491**, 711-716.

Meyer, R.C., Steinfath, M., Lisec, J., Becher, M., Witucka-Wall, H., Törjék, O., Fiehn, O., Eckardt, Ä., Willmitzer, L. and Selbig, J. (2007) The metabolic signature related to high plant growth rate in Arabidopsis thaliana. *Proceedings of the National Academy of Sciences* **104**, 4759.

Micallef, B.J., Haskins, K.A., Vanderveer, P.J., Roh, K.-S., Shewmaker, C.K. and Sharkey, T.D. (1995) Altered photosynthesis, flowering, and fruiting in transgenic tomato plants that have an increased capacity for sucrose synthesis. *Planta* **196**, 327-334.

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

Muralla, R., Sweeney, C., Stepansky, A., Leustek, T. and Meinke, D. (2007) Genetic dissection of histidine biosynthesis in Arabidopsis. *Plant physiology* **144**, 890-903.

Neuweger, H., Albaum, S.P., Dondrup, M., Persicke, M., Watt, T., Niehaus, K., Stoye, J. and Goesmann, A. (2008) MeltDB: a software platform for the analysis and integration of metabolomics experiment data. *Bioinformatics* **24**, 2726-2732.

Newman, C. and Newman, R. (2006) A brief history of barley foods. *Cereal Foods World* **51**, 4-7.

Noctor, G., Novitskaya, L., Lea, P.J. and Foyer, C.H. (2002) Coordination of leaf minor amino acid contents in crop species: significance and interpretation. *Journal of experimental botany* **53**, 939-945.

Nunes-Nesi, A., Carrari, F., Lytovchenko, A., Smith, A.M., Loureiro, M.E., Ratcliffe, R.G., Sweetlove, L.J. and Fernie, A.R. (2005) Enhanced photosynthetic performance and growth as a consequence of decreasing mitochondrial malate dehydrogenase activity in transgenic tomato plants. *Plant physiology* **137**, 611-622.

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

Oliveira, I.C. and Coruzzi, G.M. (1999) Carbon and amino acids reciprocally modulate the expression of glutamine synthetase in Arabidopsis. *Plant physiology* **121**, 301-310.

Osorio, S., Alba, R., Damasceno, C.M.B., Lopez-Casado, G., Lohse, M., Zanor, M.I., Tohge, T., Usadel, B., Rose, J.K.C. and Fei, Z. (2011) Systems biology of tomato fruit development: Combined transcript, protein, and metabolite analysis of tomato transcription factor (nor, rin) and ethylene receptor (Nr) mutants reveals novel regulatory interactions. *Plant physiology* **157**, 405-425.

Paul, M.J., Primavesi, L.F., Jhurreea, D. and Zhang, Y. (2008) Trehalose metabolism and signaling. *Annu Rev Plant Biol* **59**, 417-441.

Periappuram, C., Steinhauer, L., Barton, D.L., Taylor, D.C., Chatson, B. and Zou, J. (2000) The plastidic phosphoglucomutase from Arabidopsis. A reversible enzyme reaction with an important role in metabolic control. *Plant physiology* **122**, 1193-1200.

Persans, M.W., Yan, X., Patnoe, J.-M.M., Krämer, U. and Salt, D.E. (1999) Molecular dissection of the role of histidine in nickel hyperaccumulation in *Thlaspi goesingense* (*Halacsy*). *Plant physiology* **121**, 1117-1126.

Petersen, M. (2013) Rosmarinic acid: new aspects. *Phytochemistry Reviews*, 1-21.

Pracharoenwattana, I., Cornah, J.E. and Smith, S.M. (2005) Arabidopsis peroxisomal citrate synthase is required for fatty acid respiration and seed germination. *The Plant Cell Online* **17**, 2037-2048.

Proveniers, M. (2013) Sugars speed up the circle of life. *eLife* **2**.

Qi, Y., Liu, W., Qiu, L., Zhang, S., Ma, L. and Zhang, H. (2010) Overexpression of glutathione S-transferase gene increases salt tolerance of Arabidopsis. *Russian Journal of Plant Physiology* **57**, 233-240.

Radchuk, V.V., Borisjuk, L., Sreenivasulu, N., Merx, K., Mock, H.-P., Rolletschek, H., Wobus, U. and Weschke, W. (2009) Spatiotemporal profiling of starch biosynthesis and degradation in the developing barley grain. *Plant physiology* **150**, 190-204.

Redinbaugh, M.G. and H Campbell, W. (1998) Nitrate regulation of the oxidative pentose phosphate pathway in maize (*Zea mays* L.) root plastids: induction of 6-phosphogluconate dehydrogenase activity, protein and transcript levels. *Plant Science* **134**, 129-140.

Ridley (2009) Nitrogen-use efficiency, the next green revolution. . *The Economist*.

Riedelsheimer, C., Lisec, J., Czedik-Eysenberg, A., Sulpice, R., Flis, A., Grieder, C., Altmann, T., Stitt, M., Willmitzer, L. and Melchinger, A.E. (2012) Genome-wide association mapping of leaf metabolic profiles for dissecting complex traits in maize. *Proceedings of the National Academy of Sciences* **109**, 8872-8877.

Roessner-Tunali, U., Urbanczyk-Wochniak, E., Czechowski, T., Kolbe, A., Willmitzer, L. and Fernie, A.R. (2003) De novo amino acid biosynthesis in potato tubers is regulated by sucrose levels. *Plant physiology* **133**, 683-692.

Rolland, F., Moore, B. and Sheen, J. (2002) Sugar sensing and signaling in plants. *The Plant Cell Online* **14**, S185-S205.

Rolletschek, H., Borisjuk, L., Radchuk, R., Miranda, M., Heim, U., Wobus, U. and Weber, H. (2004) Seed specific expression of a bacterial phosphoenolpyruvate carboxylase in Vicia narbonensis increases protein content and improves carbon economy. *Plant Biotechnology Journal* **2**, 211-219.

Sahrawy, M., Ávila, C., Chueca, A., Cánovas, F.M. and López-Gorgé, J. (2004) Increased sucrose level and altered nitrogen metabolism in Arabidopsis thaliana transgenic plants expressing antisense chloroplastic fructose-1, 6-bisphosphatase. *Journal of experimental botany* **55**, 2495-2503.

Saito, K. and Matsuda, F. (2010) Metabolomics for functional genomics, systems biology, and biotechnology. *Annual review of plant biology* **61**, 463-489.

Schauer, N., Semel, Y., Roessner, U., Gur, A., Balbo, I., Carrari, F., Pleban, T., Perez-Melis, A., Bruedigam, C. and Kopka, J. (2006) Comprehensive metabolic profiling and phenotyping of interspecific introgression lines for tomato improvement. *Nature biotechnology* **24**, 447-454.

Schluepmann, H., van Dijken, A., Aghdasi, M., Wobbes, B., Paul, M. and Smeekens, S. (2004) Trehalose mediated growth inhibition of Arabidopsis seedlings is due to trehalose-6-phosphate accumulation. *Plant physiology* **135**, 879-890.

Schlüter, U., Mascher, M., Colmsee, C., Scholz, U., Bräutigam, A., Fahnenstich, H. and Sonnewald, U. (2012) Maize source leaf adaptation to nitrogen deficiency affects not only nitrogen and carbon metabolism but also control of phosphate homeostasis. *Plant physiology* **160**, 1384-1406.

Schmalenbach, I., Léon, J. and Pillen, K. (2009) Identification and verification of QTLs for agronomic traits using wild barley introgression lines. *TAG Theoretical and Applied Genetics* **118**, 483-497.

Schreiber, F., Colmsee, C., Czauderna, T., Grafahrend-Belau, E., Hartmann, A., Junker, A., Junker, B.H., Klapperstück, M., Scholz, U. and Weise, S. (2012) MetaCrop 2.0: managing and exploring information about crop plant metabolism. *Nucleic Acids Research* **40**, D1173-D1177.

Schröder, M., Giermann, N. and Zrenner, R. (2005) Functional analysis of the pyrimidine de novo synthesis pathway in solanaceous species. *Plant physiology* **138**, 1926-1938.

Sharkey, T.D., Savitch, L.V., Vanderveer, P.J. and Micallef, B.J. (1992) Carbon partitioning in a Flaveria linearis mutant with reduced cytosolic fructose bisphosphatase. *Plant physiology* **100**, 210-215.

Shrawat, A.K., Carroll, R.T., DePauw, M., Taylor, G.J. and Good, A.G. (2008) Genetic engineering of improved nitrogen use efficiency in rice by the tissue-specific expression of alanine aminotransferase. *Plant Biotechnology Journal* **6**, 722-732.

Sibley, M.H. and Anderson, L.E. (1989) Light/dark modulation of enzyme activity in developing barley leaves. *Plant physiology* **91**, 1620-1624.

Sienkiewicz-Porzucek, A., Nunes-Nesi, A., Sulpice, R., Lisec, J., Centeno, D.C., Carillo, P., Leisse, A., Urbanczyk-Wochniak, E. and Fernie, A.R. (2008) Mild reductions in mitochondrial citrate synthase activity result in a compromised nitrate assimilation and reduced leaf pigmentation but have no effect on photosynthetic performance or growth. *Plant physiology* **147**, 115-127.

Silvente, S., Camas, A. and Lara, M. (2003) Molecular cloning of the cDNA encoding aspartate aminotransferase from bean root nodules and determination of its role in nodule nitrogen metabolism. *Journal of experimental botany* **54**, 1545-1551.

Smith, C.A., Elizabeth, J., O'Maille, G., Abagyan, R. and Siuzdak, G. (2006) XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. *Analytical Chemistry* **78**, 779-787.

Sonnewald, U. and Kossmann, J. (2013) Starches—from current models to genetic engineering. *Plant Biotechnology Journal* **11**, 223-232.

Sreenivasulu, N., Altschmied, L., Radchuk, V., Gubatz, S., Wobus, U. and Weschke, W. (2004) Transcript profiles and deduced changes of metabolic pathways in maternal and filial tissues of developing barley grains. *The Plant Journal* **37**, 539-553.

Sreenivasulu, N., Usadel, B., Winter, A., Radchuk, V., Scholz, U., Stein, N., Weschke, W., Strickert, M., Close, T.J. and Stitt, M. (2008) Barley grain maturation and germination: metabolic pathway and regulatory network commonalities and differences highlighted by new MapMan/PageMan profiling tools. *Plant physiology* **146**, 1738-1758.

Steinhauser, M.-C., Steinhauser, D., Koehl, K., Carrari, F., Gibon, Y., Fernie, A.R. and Stitt, M. (2010) Enzyme activity profiles during fruit development in tomato cultivars and Solanum pennellii. *Plant physiology* **153**, 80-98.

Stettler, M., Eicke, S., Mettler, T., Messerli, G., Hörtensteiner, S. and Zeeman, S.C. (2009) Blocking the metabolism of starch breakdown products in Arabidopsis leaves triggers chloroplast degradation. *Molecular plant* **2**, 1233-1246.

Stitt, M. and Fernie, A.R. (2003) From measurements of metabolites to metabolomics: an 'on the fly'perspective illustrated by recent studies of carbon–nitrogen interactions. *Current opinion in biotechnology* **14**, 136-144.

Stitt, M., Lunn, J. and Usadel, B. (2010) Arabidopsis and primary photosynthetic metabolism–more than the icing on the cake. *The Plant Journal* **61**, 1067-1091.

Sturm, A. and Tang, G.Q. (1999) The sucrose-cleaving enzymes of plants are crucial for development, growth and carbon partitioning. *Trends in plant science* **4**, 401-407.

Sulpice, R., Nikoloski, Z., Tschoep, H., Antonio, C., Kleessen, S., Larhlimi, A., Selbig, J., Ishihara, H., Gibon, Y. and Fernie, A.R. (2013) Impact of the carbon and nitrogen supply on relationships and connectivity between metabolism and biomass in a broad panel of Arabidopsis accessions. *Plant physiology* **162**, 347-363.

Sulpice, R., Pyl, E.-T., Ishihara, H., Trenkamp, S., Steinfath, M., Witucka-Wall, H., Gibon, Y., Usadel, B., Poree, F. and Piques, M.C. (2009) Starch as a major integrator in the regulation of plant growth. *Proceedings of the National Academy of Sciences* **106**, 10348-10353.

Sulpice, R., Trenkamp, S., Steinfath, M., Usadel, B., Gibon, Y., Witucka-Wall, H., Pyl, E.T., Tschoep, H., Steinhauser, M.C. and Guenther, M. (2010) Network analysis of enzyme activities and metabolite levels and their relationship to biomass in a large panel of Arabidopsis accessions. *The Plant Cell Online* **22**, 2872-2893.

Suzuki, A., Rioual, S., Lemarchand, S., Godfroy, N., Roux, Y., Boutin, J.P. and Rothstein, S. (2001) Regulation by light and metabolites of ferredoxin-dependent glutamate synthase in maize. *Physiologia Plantarum* **112**, 524-530.

Swarbreck, S.M., Defoin-Platel, M., Hindle, M., Saqi, M. and Habash, D.Z. (2011) New perspectives on glutamine synthetase in grasses. *Journal of experimental botany* **62**, 1511-1522.

Tabuchi, M., Sugiyama, K., Ishiyama, K., Inoue, E., Sato, T., Takahashi, H. and Yamaya, T. (2005) Severe reduction in growth rate and grain filling of rice mutants lacking OsGS1; 1, a cytosolic glutamine synthetase1; 1. *The Plant Journal* **42**, 641-651.

Tan, X., Calderon-Villalobos, L.I.A., Sharon, M., Zheng, C., Robinson, C.V., Estelle, M. and Zheng, N. (2007) Mechanism of auxin perception by the TIR1 ubiquitin ligase. *Nature* **446**, 640-645.

Tang, G.-Q. and Sturm, A. (1999) Antisense repression of sucrose synthase in carrot (Daucus carota L.) affects growth rather than sucrose partitioning. *Plant molecular biology* **41**, 465-479.

Thévenot, C., Simond-Côte, E., Reyss, A., Manicacci, D., Trouverie, J., Le Guilloux, M., Ginhoux, V., Sidicina, F. and Prioul, J.-L. (2005) QTLs for enzyme activities and soluble carbohydrates involved in starch accumulation during grain filling in maize. *Journal of experimental botany* **56**, 945-958.

Thimm, O., Bläsing, O., Gibon, Y., Nagel, A., Meyer, S., Krüger, P., Selbig, J., Müller, L.A., Rhee, S.Y. and Stitt, M. (2004) mapman: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. *The Plant Journal* **37**, 914-939.

Tomaz, T., Bagard, M., Pracharoenwattana, I., Lindén, P., Lee, C.P., Carroll, A.J., Ströher, E., Smith, S.M., Gardeström, P. and Millar, A.H. (2010) Mitochondrial malate dehydrogenase lowers leaf respiration and alters photorespiration and plant growth in Arabidopsis. *Plant physiology* **154**, 1143-1157.

Toubiana, D., Semel, Y., Tohge, T., Beleggia, R., Cattivelli, L., Rosental, L., Nikoloski, Z., Zamir, D., Fernie, A.R. and Fait, A. (2012) Metabolic profiling of a mapping population exposes new insights in the regulation of seed metabolism and seed, fruit, and plant relations. *PLoS genetics* **8**, e1002612.

Trethewey, R.N., Geigenberger, P., Riedel, K., Hajirezaei, M.R., Sonnewald, U., Stitt, M., Riesmeier, J.W. and Willmitzer, L. (1998) Combined expression of glucokinase and invertase in potato tubers leads to a dramatic reduction in starch accumulation and a stimulation of glycolysis. *The Plant Journal* **15**, 109-118.

Ueda, A., Li, P., Feng, Y., Vikram, M., Kim, S., Kang, C.H., Kang, J.S., Bahk, J.D., Lee, S.Y. and Fukuhara, T. (2008) The Arabidopsis thaliana carboxyl-terminal domain phosphatase-like 2 regulates plant growth, stress and auxin responses. *Plant molecular biology* **67**, 683-697.

Ullrich, S.E. (2011) Significance, Adaptation, Production, and Trade of Barley. *Barley*, 3-13.

Van Camp, W. (2005) Yield enhancement genes: seeds for growth. *Current opinion in biotechnology* **16**, 147-153.

Van Houtte, H., Lopez-Galvis, L., Vandesteene, L., Beeckman, T. and Van Dijck, P. (2013) Redundant and non-redundant roles of the trehalose-6-phosphate phosphatases in leave growth, root hair specification and energy-responses in Arabidopsis. *Plant signaling & behavior* **8**, e23209.

Vandesteene, L., López-Galvis, L., Vanneste, K., Feil, R., Maere, S., Lammens, W., Rolland, F., Lunn, J.E., Avonce, N. and Beeckman, T. (2012) Expansive evolution of the trehalose-6-phosphate phosphatase gene family in Arabidopsis. *Plant physiology* **160**, 884-896.

von Caemmerer, S. and Evans, J.R. (2010) Enhancing C3 photosynthesis. *Plant physiology* **154**, 589-592.

von Wirén, N., Klair, S., Bansal, S., Briat, J.-F., Khodr, H., Shioiri, T., Leigh, R.A. and Hider, R.C. (1999) Nicotianamine chelates both FeIII and FeII. Implications for metal transport in plants. *Plant physiology* **119**, 1107-1114.

Wakasa, K., Hasegawa, H., Nemoto, H., Matsuda, F., Miyazawa, H., Tozawa, Y., Morino, K., Komatsu, A., Yamada, T. and Terakawa, T. (2006) High-level tryptophan accumulation in seeds of transgenic rice and its limited effects on agronomic traits and seed metabolite profile. *Journal of experimental botany* **57**, 3069-3078.

Wang, S.-m., Lue, W.-l., Yu, T.-s., Long, J.-h., Wang, C.-n., Eimert, K. and Chen, J. (1998) Characterization of ADG1, an Arabidopsis locus encoding for ADPG pyrophosphorylase small subunit, demonstrates that the presence of the small subunit is required for large subunit stability. *The Plant Journal* **13**, 63-70.

Wang, Z., Chen, X., Wang, J., Liu, T., Liu, Y., Zhao, L. and Wang, G. (2007) Increasing maize seed weight by enhancing the cytoplasmic ADP-glucose pyrophosphorylase activity in transgenic maize plants. *Plant cell, tissue and organ culture* **88**, 83-92.

Watanabe, M., Balazadeh, S., Tohge, T., Erban, A., Giavalisco, P., Kopka, J., Mueller-Roeber, B., Fernie, A.R. and Hoefgen, R. (2013) Comprehensive dissection of spatio-temporal metabolic shifts in primary, secondary and lipid metabolism during developmental senescence in Arabidopsis thaliana. *Plant physiology*.

Weber, H., Borisjuk, L., Heim, U., Buchner, P. and Wobus, U. (1995) Seed coat-associated invertases of fava bean control both unloading and storage functions: cloning of cDNAs and cell type-specific expression. *The Plant Cell Online* **7**, 1835-1846.

Weschke, W., Panitz, R., Gubatz, S., Wang, Q., Radchuk, R., Weber, H. and Wobus, U. (2003) The role of invertases and hexose transporters in controlling sugar ratios in maternal and filial tissues of barley caryopses during early development. *The Plant Journal* **33**, 395-411.

Wilson, R.A., Jenkinson, J.M., Gibson, R.P., Littlechild, J.A., Wang, Z.-Y. and Talbot, N.J. (2007) Tps1 regulates the pentose phosphate pathway, nitrogen metabolism and fungal virulence. *The EMBO journal* **26**, 3673-3685.

Wingler, A., Delatte, T.L., O'Hara, L.E., Primavesi, L.F., Jhurreea, D., Paul, M.J. and Schluepmann, H. (2012) Trehalose 6-phosphate is required for the onset of leaf senescence associated with high carbon availability. *Plant physiology*.

Wingler, A., Fritzius, T., Wiemken, A., Boller, T. and Aeschbacher, R.A. (2000) Trehalose induces the ADP-glucose pyrophosphorylase gene, ApL3, and starch synthesis in Arabidopsis. *Plant physiology* **124**, 105-114.

Wu, L. and Birch, R.G. (2007) Doubled sugar content in sugarcane plants modified to produce a sucrose isomer. *Plant Biotechnology Journal* **5**, 109-117.

Xia, J., Psychogios, N., Young, N. and Wishart, D.S. (2009) MetaboAnalyst: a web server for metabolomic data analysis and interpretation. *Nucleic Acids Research* **37**, W652-W660.

Xiang, C., Werner, B.L., E'Lise, M.C. and Oliver, D.J. (2001) The biological functions of glutathione revisited in Arabidopsis transgenic plants with altered glutathione levels. *Plant physiology* **126**, 564-574.

Xu, S.-M., Brill, E., Llewellyn, D.J., Furbank, R.T. and 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.

Xue, G.-P., McIntyre, C.L., Jenkins, C.L., Glassop, D., van Herwaarden, A.F. and Shorter, R. (2008) Molecular dissection of variation in carbohydrate metabolism related to water-soluble carbohydrate accumulation in stems of wheat. *Plant physiology* **146**, 441-454.

Yamaya, T., Obara, M., Nakajima, H., Sasaki, S., Hayakawa, T. and Sato, T. (2002) Genetic manipulation and quantitative-trait loci mapping for nitrogen recycling in rice. *Journal of experimental botany* **53**, 917-925.

YongLing, R., Chourey, P.S. and Basra, A. (2006) Carbon partitioning in developing seed. *Handbook of seed science and technology*, 125-152.

Yun, M.-S., Umemoto, T. and Kawagoe, Y. (2011) Rice debranching enzyme isoamylase3 facilitates starch metabolism and affects plastid morphogenesis. *Plant and Cell Physiology* **52**, 1068-1082.

Zamboni, A., Di Carli, M., Guzzo, F., Stocchero, M., Zenoni, S., Ferrarini, A., Tononi, P., Toffali, K., Desiderio, A. and Lilley, K.S. (2010) Identification of putative stage-specific grapevine berry biomarkers and omics data integration into networks. *Plant physiology* **154**, 1439-1459.

Zechmann, B., Koffler, B. and Russell, S. (2011) Glutathione synthesis is essential for pollen germination in vitro. *BMC plant biology* **11**, 54.

Zhang, N., Gibon, Y., Gur, A., Chen, C., Lepak, N., Höhne, M., Zhang, Z., Kroon, D., Tschoep, H. and Stitt, M. (2010) Fine quantitative trait loci mapping of carbon and nitrogen metabolism enzyme activities and seedling biomass in the maize IBM mapping population. *Plant physiology* **154**, 1753-1765.

Zhu, X.-G., de Sturler, E. and Long, S.P. (2007) Optimizing the distribution of resources between enzymes of carbon metabolism can dramatically increase photosynthetic rate: a numerical simulation using an evolutionary algorithm. *Plant physiology* **145**, 513-526.

Zohary, D. and Hopf, M. (2000) *Domestication of plants in the old world: the origin and spread of cultivated plants in West Asia, Europe, and the Nile Valley*:Oxford University Press, USA.

Zrenner, R., Salanoubat, M., Willmitzer, L. and Sonnewald, U. (1995) Evidence of the crucial role of sucrose synthase for sink strength using transgenic potato plants (Solanum tuberosum L.). *The Plant Journal* **7**, 97-107.

Zrenner, R., Schüler, K. and Sonnewald, U. (1996) Soluble acid invertase determines the hexose-to-sucrose ratio in cold-stored potato tubers. *Planta* **198**, 246-252.

8. Abbreviations

Full Name	Abbreviation
1-Aminocyclopropane-1-carboxylic acid	ACC
2-oxoglutarate	2-oxo
3-phosphoglycerate	3PGA
6-phosphogluconate	6-PG
6-phosphogluconate dehydrogenase	6PGDH
ADP-glucose	ADPGlc
ADP-glucose pyrophosphorylase	AGPase
Alanine	Ala
Alanine aminotransferase	AlaAT
alph-tocopherol	alpha-Toc
Aminoadipic	Amdp
Arabinose	Ara
Arginine	Arg
Ascorbate	Ascb
Asparagin	Asn
Aspartate	Asp
Aspartate aminotransferase	AspAT
Cis-aconitate	Cacnt
Citrate	Cit
Cytoplasmic invertase	CytInv
Cytosolic aldolase	cAldolase
Cytosolic fructose-1,6-bisphosphatase	cFBPase
Dehydroascorbate	Dhdascb
Dodecanoic acid	Ddca
Erythrose-4-phosphate	E4P
Final shoot biomass	FSB
Flag leaf area	FLA
Flag leaf fresh weight	FLFW
Fructose	Fru
Fructose-1,6-bis phosphate	Fru1,6BP
Fructose-6-phosphate	Fru6P
Fucose	Fuc
Fumarate	Fum
Galactose	Gal
Gamma aminobutyric acid	GABA
Gene Ontology	GO
Glucose	Glc
Glucose 1,6-bisphosphate	Glc1,6BP
Glucose-1-phosphate	Glc1P
Glucose-6-phosphate	Glc6P

Glucose-6-phosphate dehydrogenase	G6PDH
Glucuronic acid	Glucua
Glutamate	Glu
Glutamine	Gln
Glyceraldehyde-3-phosphate dehydrogenase	GAPDH
Glycerate	Glyct
Glycine	Gly
Grain yield	GY
Hexadecanoic acid	Hxda
Hexokinase	HK
Histidine	His
Homoserine	Hser
Indole acetic acid	IAA
Inositiol	Inost
Isocitrate	Isocit
Isocitrate dehydrogenase	ICDH
Isoleucine	Ile
Lactate dehydrogenase	LDH
Leaf area	LA
Leucine	Leu
Lysine	Lys
Malate	Mal
Malate dehydrogenase	MDH
Mannose	Man
Mannose-6-phosphate	Man6P
Methionine	Met
Multiple reaction monitoring	MRM
Octadecanoic acid	Odca
Oxaloacetate	OAA
Peduncle length	PL
Phenylalanine	Phe
Phophoglucoisomerase	PGI
Phosphoenolpyruvate	PEP
Phosphofructokinase	PFK
Phosphoglucoisomerase	PGI
Phosphoglucomutase	PGM
Phosphoglycerate kinase	PGK
Plastidic fructose-1,6-bisphosphatase	pFBPase
Pro-glutamate	Pglu
Proline	Pro
Putrescine	Put
Pyrophosphate	PPi
Pyruvate	Pyr
Pyruvate dehydrogenase	PDH

Pyruvate kinase	PK
Quinate	Qun
Ribonate	Ribn
Ribose	Rib
Ribose-5-phosphate	Rib5P
Ribulose-1,5-bisphosphate	RuBP
Ribulose-5-phosphate	R5P
Seduheptulose bisphosphatase	SBP
Serine	Ser
Sorbitol	Sbt
Sorbose-6-phosphate	Sor6P
Starch synthase	SS
Succinate	Succ
Sucrose	Suc
Sucrose -6-phosphate phosphatase	SPP
Sucrose synthase	Susy
Sucrose-6-phosphate	Suc6P
Sucrose-6-phosphate synthase	SPS
Systems biology Markup languages	SBML
Tetradecanoic acid	Tdca
The Database for Annotation, Visualization and Integrated Discovery	DAVID
Thousands kernel weight	TKW
Threonine	Thr
Total dry weight	TDW
Total leaf biomass	TLB
Total seed weight	TSW
Total shoot biomass	TSB
Trans-aconitate	Tacnt
Transketolase	TK
Trehalose-6-phosphate	Tre6P
Trehalose-6-phosphate phosphatase	TPP
Trehalose-6-phosphate synthase	TPS
Tricarboxylic acid	TCA
Triethanolamine	TEA
Triose-phosphate isomerase	TPI
Tyrosine	Tyr
UDP-glucose	UDPGlc
Valine	Val
Visualization and Analysis of Networks containing Experimental Data	VANTED
Xylose	Xyl

9. Appendix

Appendix Table 1 List of eight accession lines of spring barley selected for the biochemical study. Accession lines originated from different regions are characterized on their origin, spike morphology and seed color. HOR: Hordeum

Accession ID	Accession number	Genus/ species	Origin	Six - two rowed	Seed Color
20517	**HOR4555**	*Hordeum vulgare*	Russia	Two-rowed	Yellow
20663	**HOR4730**	*Hordeum vulgare*	India	Two-rowed	Yellow
23234	**HOR7410**	*Hordeum vulgare*	Ethiopia	Two-rowed	Black
19888	**HOR3909**	*Hordeum vulgare*	Unknown	Two-rowed	Yellow
18878	**HOR2851**	*Hordeum vulgare*	Iran	Two-rowed	Black
16472	**HOR216**	*Hordeum vulgare*	Sweden	Six-rowed	Black
98426	**HOR13406**	*Hordeum vulgare*	Unknown	Six-rowed	Yellow
17788	**HOR1645**	*Hordeum agriocrithon*	china	Six-rowed	Yellow

Appeddix Table 2: Classification of the accession lines of winter barley. Twelve out of twenty accession lines were resistant to freezing temperatures and were used for the analysis at the grain filling stage. BCC: Barley Core Collection.

Accession ID	Accession Name	Genus/ Species	Origin	Two - Six rowed
638586	**BCC1568**	*Hordeum vulgare*	Germany	Two-rowed
638324	**BCC1306**	*Hordeum vulgare*	Great Britain	Two-rowed
638333	**BCC1315**	*Hordeum vulgare*	Great Britain	Two-rowed
638356	**BCC1338**	*Hordeum vulgare*	Germany	Six-rowed
638377	**BCC1359**	*Hordeum vulgare*	Korea	Six-rowed
638576	**BCC1558**	*Hordeum vulgare*	Germany	Six-rowed
637841	**BCC455**	*Hordeum vulgare*	China	Six-rowed
637972	**BCC586**	*Hordeum vulgare*	Japan	Six-rowed
638048	**BCC662**	*Hordeum vulgare*	Korea	Six-rowed
638246	**BCC880**	*Hordeum vulgare*	USA	Six-rowed
638586	**BCC887**	*Hordeum vulgare*	USA	Six-rowed
638614	**BCC1596**	*Hordeum vulgare*	Italy	Six-rowed

Appendix Table 3. Multiple reaction monitoring (MRM) of potentials of MS/MS transitions used in negative mode. Nitrogen was used as carrier gas, nebulizer gas, heater gas and collision gas. Ion spray voltage and the capillary temperature were -3500 V and 350°C, respectively. Pre ion=precursor ion, Pro ion=product ion, Dwell=dwell time, Frag=fragmentor, CE= collision energy, CAV= Cell accelerator voltage.

Compound Name	Prec Ion	Pro Ion	Dwell	Frag	CE	CAV	Polarity
Acetyl-CoA	808.1	460.9	20	380	33	5	Negative
Acetyl-CoA	808.1	403.9	20	380	20	5	Negative
NADPH	744	159	20	380	53	5	Negative
NADPH	744	79	20	380	80	5	Negative
NADH	664	408	20	380	33	5	Negative
NADH	664	346	20	380	29	5	Negative
ADPGlc	588	346	20	380	29	5	Negative
ADPGlc	588	241	20	380	29	5	Negative
ADPGlc	565	322.9	20	380	25	5	Negative
ADPGlc	565	158.9	20	380	53	5	Negative
ATP	505.9	408	20	380	21	5	Negative
ATP	505.9	159	20	380	33	5	Negative
UTP	483	402.9	20	380	21	5	Negative
UTP	483	384.2	20	380	17	5	Negative
ADP	426	328	20	380	17	5	Negative
ADP	426	159	20	380	21	5	Negative
Tre6P	421	240.9	20	380	25	5	Negative
Tre6P	421	139	20	380	29	5	Negative
Suc6P	421	96.8	20	380	33	5	Negative
Suc6P	421	79.1	20	380	53	5	Negative
UDP	403	272.8	20	380	21	5	Negative
UDP	403	110.9	20	380	17	5	Negative
AMP	346	134.2	20	380	37	5	Negative
AMP	346	106.9	20	380	61	5	Negative
Fru 1.6BP	339	97	20	380	21	5	Negative
Fru1.6BP	339	79	20	380	57	5	Negative
Glc1P	259	96.8	20	380	9	5	Negative
Glc1P	259	78.9.8	20	380	44	5	Negative
Fru6P	259	96.8	20	380	9	5	Negative
Fru6P	259	78.9	20	380	44	5	Negative
Glc6P	259	78.9	20	380	44	5	Negative
Glc6P	259	96.8	20	380	44	5	Negative
Rib5P	229	97	20	380	5	5	Negative
Rib5P	229	79	20	380	44	5	Negative
Gluca	193	113	20	380	5	5	Negative
Gluca	193	85	20	380	13	5	Negative
Isocit	191	110.8	20	380	9	5	Negative
Isocit	191	87	20	380	13	5	Negative

Continue

Compound Name	Prec Ion	Pro Ion	Dwell	Frag	CE	CAV	Polarity
3PGA	185.1	97	20	380	13	5	Negative
3PGA	185.1	79	20	380	41	5	Negative
Cacnt	173.1	128.9	20	380	1	5	Negative
Cacnt	173.1	85	20	380	9	5	Negative
Tacnt	173.1	111	20	380	5	5	Negative
Tacnt	173.1	128.9	20	380	1	5	Negative
PEP	167	78.9	20	380	13	5	Negative
PEP	167	63.2	20	380	77	5	Negative
2-oxo	145.1	101	20	380	5	5	Negative
Mal	133	115	20	380	5	5	Negative
Mal	133	71	20	380	9	5	Negative
Succ	116.9	73	20	380	5	5	Negative
Fum	115	70.9	20	380	1	5	Negative
13C-Pyr	89.9	89.9	20	380	1	5	Negative
13C-Pyr	89.9	45.1	20	380	5	5	Negative
Pyr	87.1	87	20	380	5	5	Negative

Appendix Table 4. Number of relationships between metabolites and enzymes at vegetative stage in leaves of spring barley 30 days after planting. Numbers of significant associations are indicated at the levels of 0.01 and 0.001.

Parameter pair	Number of correlations		Total number of traits	Total number of correlations
	Positive	Negative		
Enzyme-enzyme	110	6	21	116
Metabolite-metabolite	330	168	52	498
Amino acid-amino acid	90	2	20	92
Enzyme-metabolite	137	118	80	255

Appendix Table 5. Phenotypical parameters determined in barley plants at the generative stage under field conditions. Values are the average and standard error of five individual plants. PH: plant height (cm), FLA: flag leaf area (cm^2), FLDW: flag leaf dry weight (gr), FLFW: flag leaf fresh weight (gr), TSB: total shoot biomass (gr), TSW: total seed biomass (gr)

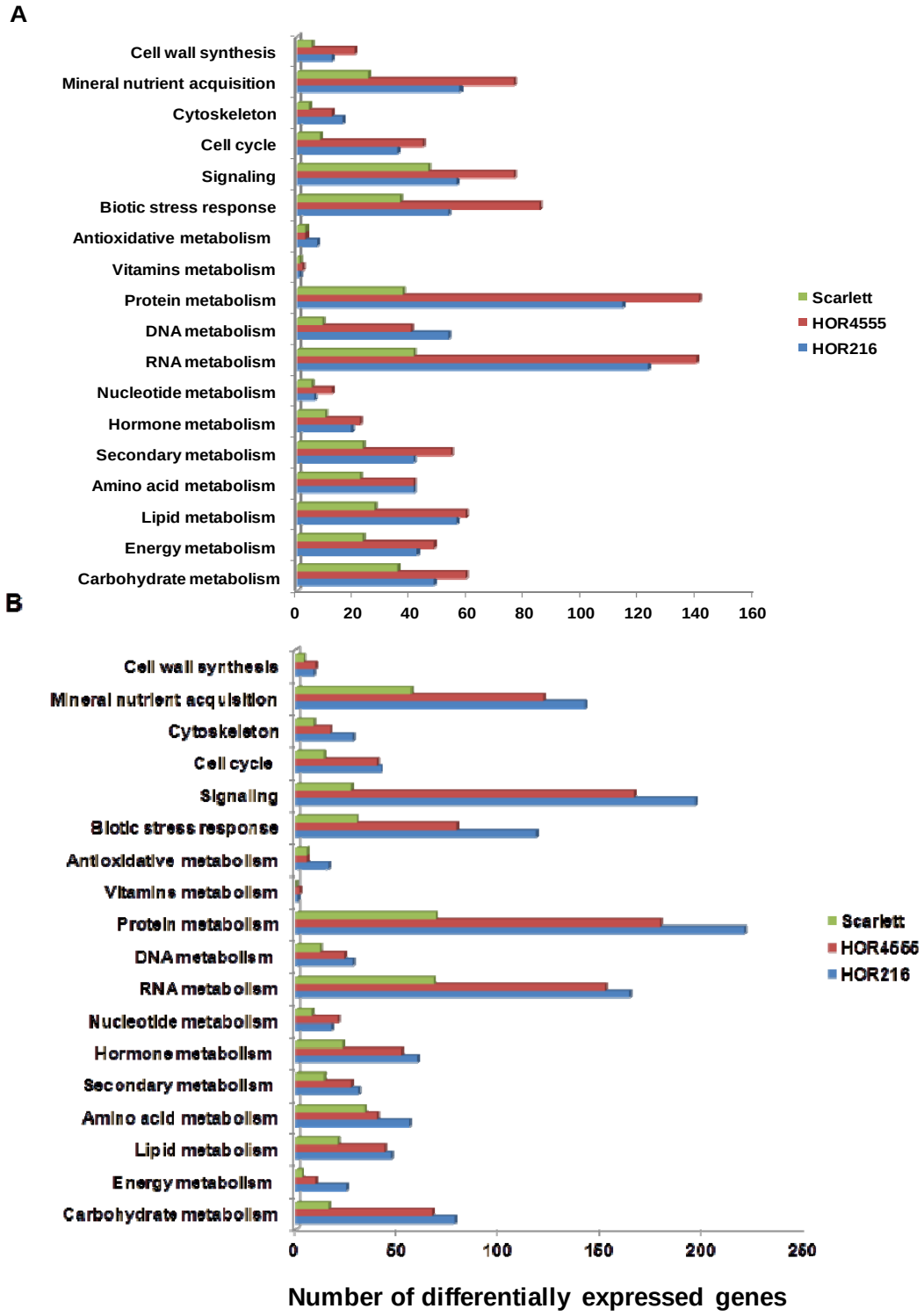
Genotype	PH	FLDW	FLFW	TSB	FLA
HOR4555	144±1.9	0.66±0.06	2.43±0.21	30.87±2.41	141±14
HOR216	157±3.1	1.4±0.13	3.56±0.47	58.24±5.43	205±14
HOR7410	147±2.6	0.64±08	2.02±0.3	46.48±3.18	185±5
HOR4730	141±3	1.02±0.05	3.81±0.21	45.31±3.18	239±20
HOR3909	129±4	1.07±0.09	3.87±0.34	45.78±3.74	258±30
HOR13406	136±2	0.63±0.09	2.55±0.32	34.36±3.17	159±7
HOR2851	138±4	0.45±0.01	1.88±0.34	43.33±1.70	142±15
HOR1645	137±3	0.72±0.05	2.59±0.39	44.33±1.49	193±4
S42IL122	105±4	0.52±0.07	1.76±0.25	36.72±1.60	154±20
S42IL119	116±2	0.46±0.05	1.79±0.15	38.88±0.74	139±4
S42IL107	104±2	0.28±0.04	1.14±0.16	15.86±1.36	75±16
Scarlett	113±2	0.55±0.07	2.10±0.23	32.13±3.37	139±5

Appendix Table 6. Number of relationships between metabolites and enzymes at the generative stage in flag leaves of spring barley. Numbers of significant correlations are indicated at the levels of 0.01 and 0.001.

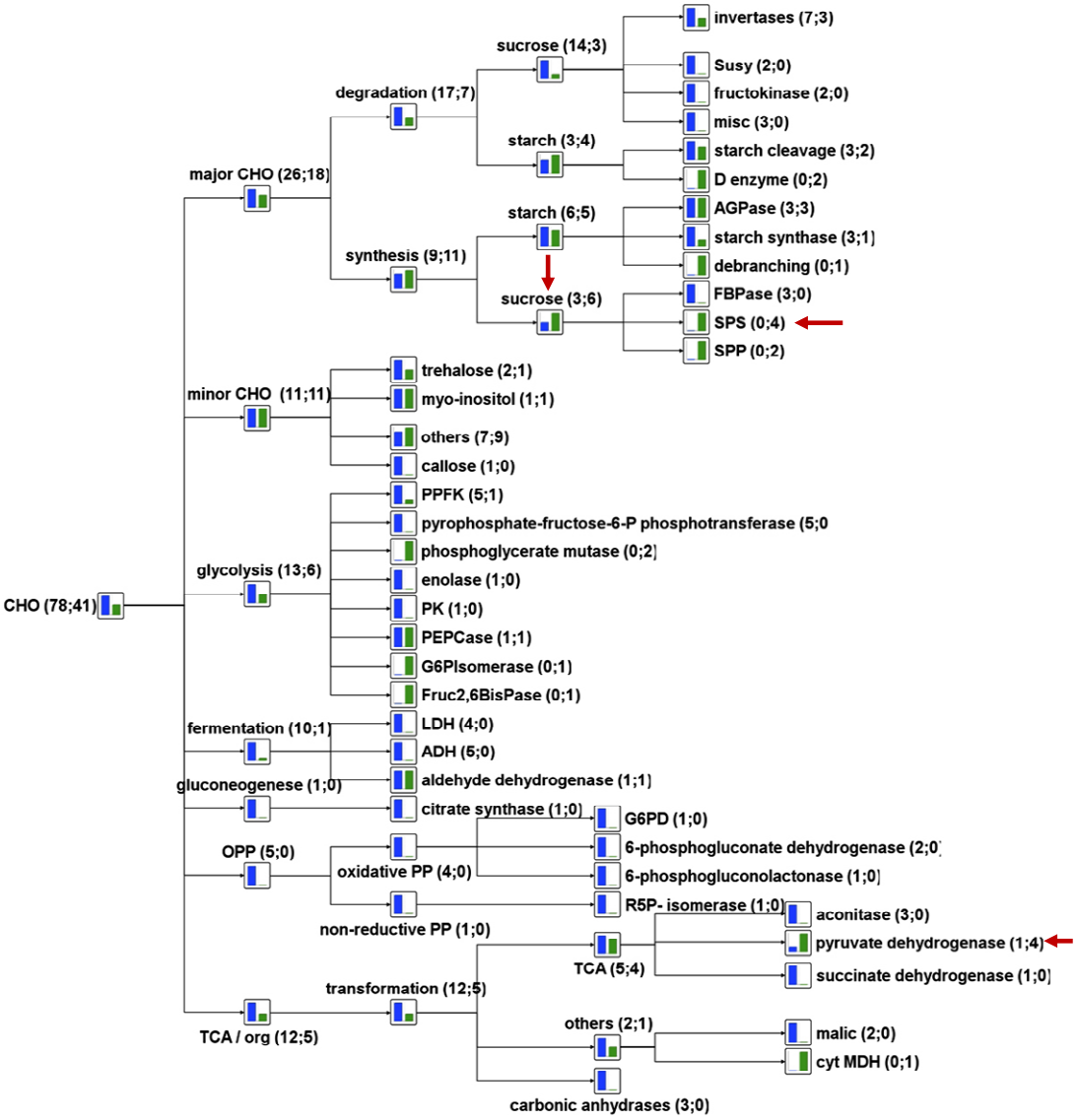
Parameter	Number of correlations	Number of positive correlations	Number of negative correlations
Metabolites-to-metabolite	1402	1320	82
Enzyme-to-metabolite	976	937	40

Appendix Table 7. Average values (±SE) of morphological parameters measured under field conditions for twelve-winter barley accessions in the year 2011. PH: plant height (cm), FLDW (gr): flag leaf dry weight (gr), TSB: total shoot biomass (gr), FLA: flag leaf area (cm^2), TKW: thousand kernel weight (gr), SE: standard error (n=5)

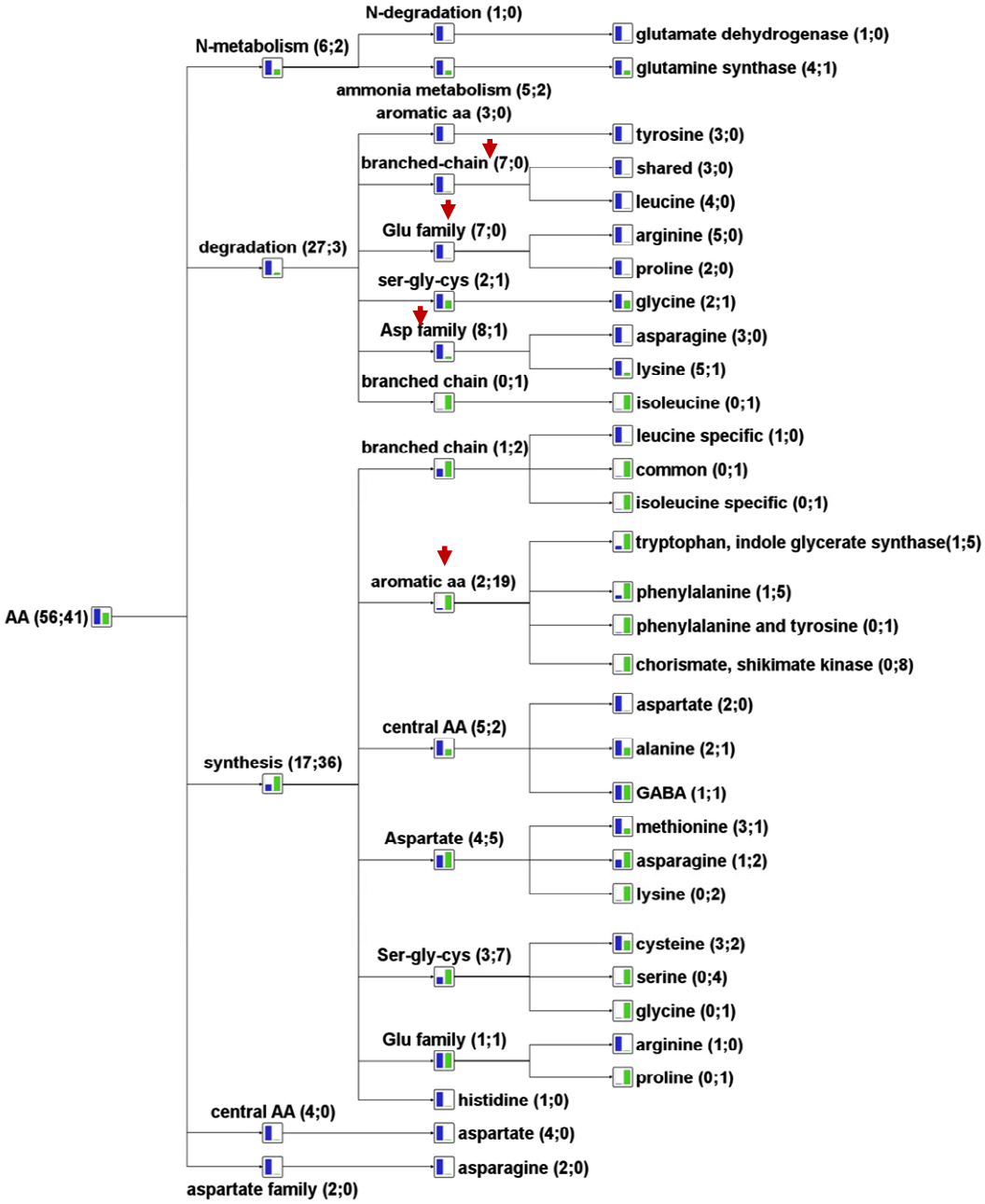
Genotype	PH	FLDW	TSB	FLA	TKW
BCC887	152 ± 3.26	0.13 ± 0.06	19 ± 1.97	51 ± 13.88	38 ± 1.29
BCC1306	143 ± 5.81	0.96 ± 0.23	64 ± 12.57	260 ± 48.53	48 ± 2.34
BCC1315	127 ± 1.86	1.16 ± 0.17	61 ± 9.02	241 ± 34.08	45 ± 2.88
BCC1338	146 ± 8.31	1.17 ± 0.18	53 ± 8.35	325 ± 32.67	43 ± 1.24
BCC1359	169 ± 7.05	0.39 ± 0.07	33 ± 5.26	98 ± 16.09	48 ± 3.33
BCC1558	167 ± 2.04	0.89 ± 0.10	81 ± 10.49	229 ± 27.33	39 ± 2.93
BCC1568	175 ± 2.05	1.05 ± 0.22	88 ± 13.10	255 ± 55.38	40 ± 1.81
BCC1596	133 ± 4.79	1.02 ± 0.16	34 ± 2.14	248 ± 41.19	59 ± 2.39
BCC455	150 ± 2.80	3.53 ± 0.39	139 ± 20.26	975 ± 112.48	28 ± 1.41
BCC586	92 ± 2.54	1.37 ± 0.21	46 ± 5.33	277 ± 37.98	32 ± 0.72
BCC662	106 ± 2.62	1.06 ± 0.09	38 ± 7.34	242 ± 27.78	27 ± 1.62
BCC880	151 ± 4.61	2.28 ± 0.12	93 ± 8.53	542 ± 29.62	38 ± 0.30



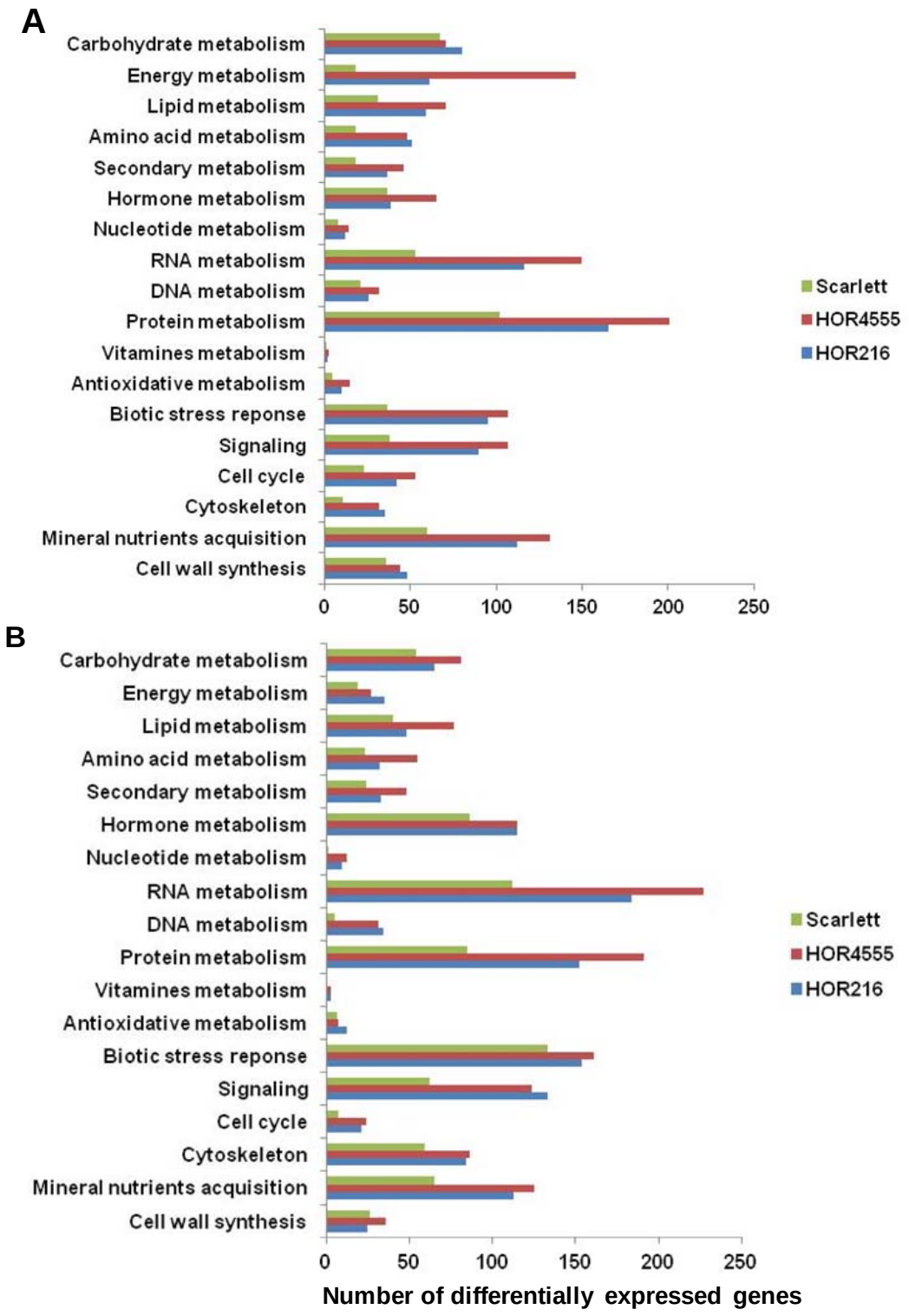
Appendix Figure 1. Categorization of genes with significantly different expression levels in HOR216, HOR4555 and Scarlett lines to lowest biomass line in, S52IL107 in flag leaves using Gene ontology terms and pathways. A and B show significant up and down regulations



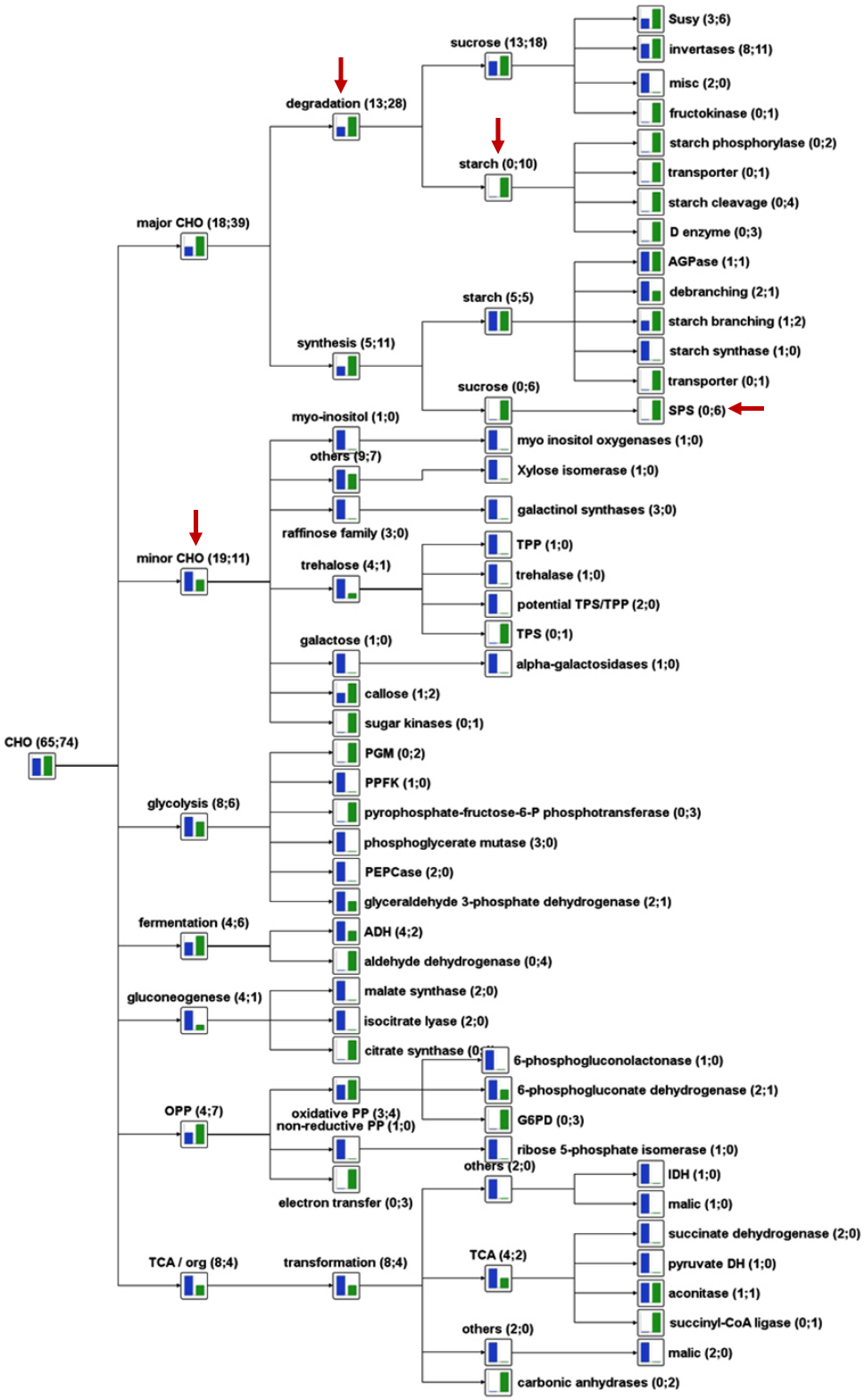
Appendix Figure 2. Differentially up- and down regulated genes of carbohydrate metabolism (CHO) in plant with highest biomass in compared to lowest biomass in flag leaves. Blue and green colors show up- and down-regulations. Arrows show significant overrepresented clusters and genes at a significant threshold of $p<0.0001$.



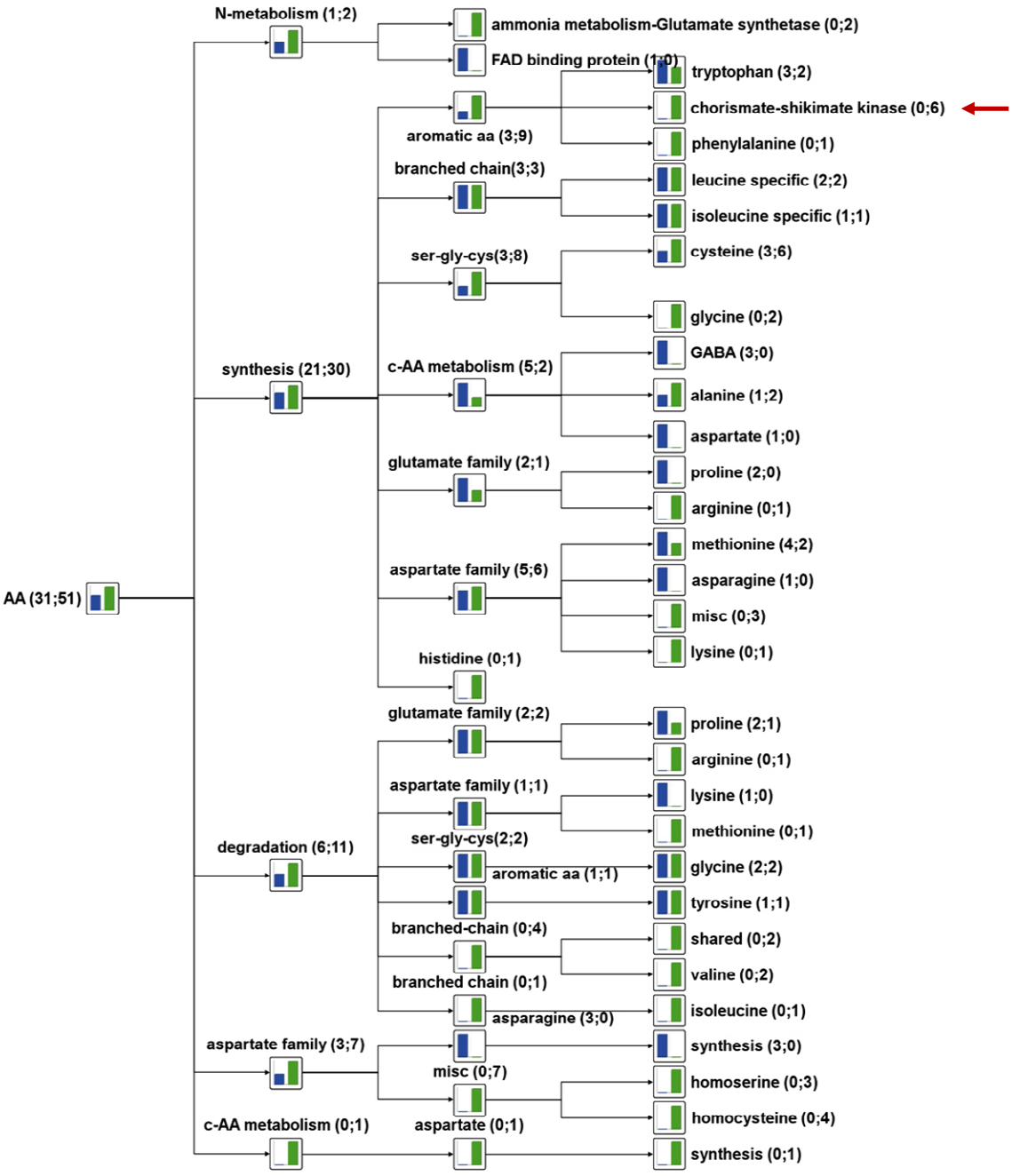
Appendix Figure 3. Differentially up- and down regulated genes of amino acid metabolism (AA) in plant with highest biomass in compared to lowest biomass in flag leaves. Blue and green colors show up- and down- regulations. Arrows show significant overrepresented clusters and genes at a significant threshold of $p<0.0001$.



Appendix Figure 4. Categorization of genes with significantly different expression levels in HOR216, HOR4555 and Scarlett lines to lowest biomass line in, S52IL107, in seeds using Gene ontology terms and pathways.



Appendix Figure 5. Differentially up- and down regulated genes of carbohydrate metabolism (CHO) in plant with highest biomass in compared to lowest biomass in seeds. Blue and green colors show up- and down- regulations. Arrows show significant overrepresented clusters and genes at a significant threshold of p<0.0001.



Appendix Figure 6. Differentially up- and down regulated genes of amino acid metabolism (AA) in plant with highest biomass in compared to lowest biomass in seeds. Blue and green colors show up- and down-regulations. Arrows show significant overrepresented clusters and genes at a significant threshold of $p<0.0001$.

10. Publications and Proceedings related to the submitted thesis

1. Astrid Junker, Mathias Franke, **Mohammad-Reza Ghaffari**, Eva Grafahrend-Belau, Jan Hüge, Christian Krach, André Eschenröder , Thomas Herter, Heike Riegler, Katrin Knorr, Katrin Lotz, Mohammad-Reza Hajirezaei, Björn Junker, Johannes Müller, Björn Usadel, Michael Leps, Falk Schreiber. BioEnergy 2021 MMM – Multiscale Metabolic Modeling of cereals: An integrative systems biology approach for biomass research. 2012. Feb, Potsdam, Germany

2. **Ghaffari**, M.R., Usadel, B., Junker, B., Junker, A., Schreiber, F., von Wiren, N and Hajirezaei, M. Metabolite-based prediction for biomass formation during grain filling in barley. Botanikertagung 2011, September 18 - 23, Berlin, Germany

3. **Ghaffari**, M.R., Usadel, B., Junker, B., Schreiber, F., von Wiren, N and Hajirezaei, M. Targeted and non-targeted metabolite analysis to predict biomass formation during the grain filling phase in barley. Trends in Metabolomics - Analytics and Applications 2011, May 19-20, Frankfurt am Main, Germany

4. Riegler, H, Herter T, Hajirezaei M.R, **Ghaffari** M.R and, Usadel, B .Multiscale Metabolic Modelling (MMM) of cereals: An integrative systems biology approach for biomass research (V) Cell Wall Analysis. GABI Meeting 2010, February, Potsdam, Germany

5. M. R. Hajirezaei, M.R. **Ghaffari**, B.H. Junker, J. Muller, B. Usadel, M. Leps, R. Lemke, F. Schreiber . Multiscale Metabolic Modeling of cereals – an integrated systems biology approach for research biomass. Renewable Resources and Biotechnology for Material Applications. Book Chapter 34. pp. 325-331 Nova Science Publishers, New York

Herr **Mohammad Reza Ghaffari**
Institute of Plant Genetics and Crop Plant Research
Molecular Plant Nutrition
Correnssstraße 3
06466 Gatersleben

11. Eidesstattliche Erklärung

Hiermit erkläre ich, dass diese Arbeit von mir bisher weder der Naturwissenschaftliche Fakultät I – Biowissenschaften, der Martin-Luther-Universität Halle-Wittenberg noch einer anderen wissenschaftlichen Einrichtung zum Zweck der Promotion eingereicht wurde.
Ich erkläre ferner, dass ich diese Arbeit selbständig und nur unter Zuhilfenahme der angegebenen Hilfsmittel und Literatur angefertigt habe.

12. Curriculum Vitae

Name: Mohammad Reza

Given name: Ghaffari

Address: Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Physiology and Cell Biology Dept. Molecular Plant Nutrition Group, Corrensstr. 3, 06466, Gatersleben, Germany

Phone number: +49-39482-5452

E-mail: ghaffari@ipk-gatersleben.de

E-mail: ghaffari@abrii.ac.ir

Nationality: Iranian

Date of Birth: 23/08/1973

Place of Birth: Kashmar / Iran

Gender: Male

Marital status: Married

Academic Education:

Ph.D.; Systems biology, Martin Luther University, Halle, Germany (Jun. 2009 - November 2013), Thesis prepared at Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Gatersleben, Germany

M.Sc., Biochemistry, Faculty of basic science, Guilan University, Rasht, Iran (Sep.1999 - Jan. 2002.), Thesis prepared at Agricultural Biotechnology Research Institute of Iran, Karaj (ABRII).

B.Sc.; Biology, Faculty of science, Shiraz University, Shiraz, Iran (Sep.1995 - June 1999).

Diploma; Life Sciences, Kashmar, Iran (1993-1997).

Research:

- Multiscale Metabolic Modelling of cereals: an integrative systems biology approach for biomass research" (Bioenergy 2021), Funded by BMBF- Jun 2009- Dec 2012, Germany

- Development of molecular marker based on NPR1 candidate resistance gene in wheat.

- Proteomic study of wheat spike infected by *Fusarium graminearum*. Iran

- Analysis of differentially expressed genes responsible to DON (Deoxynivalenol) in resistant and susceptible varieties of wheat. Iran

- QTL mapping and marker assisted breeding for improved Fusarium head blight resistance in wheat. Iran

- Development of molecular markers based on EST linked to Fusarium head blight resistance gene in wheat. Iran

- Transcriptome analysis of wheat spike infected by *Fusarium graminearum* using cDNA-AFLP of gene expression profiling. Iran

- Identification and development of PCR based molecular markers linked to Septoria leaf blotch resistance genes in wheat. Iran

- Functional genomics of wheat: Development of PCR based molecular markers for candidate Fusarium resistance genes, University of agricultural sciences, Vienna and Agricultural biotechnology research institute of Iran. 2002-2004

Job Experiences:

- Working on Ph.D. thesis as an employee at the Institute of Plant Genetics and Crop Plant Research (IPK), Gatersleben, Germany (Jun.2006-May 2013).
- Lecturer, Agricultural biotechnology Research Institute of Iran, 2002-2009

International Refereed Journals

1. Ahkami AH, Melzer M, **Ghaffari** MR, Pollmann S, Ghorbani Javid M, Shahinnia F, Hajirezaei MR, Druege U. 2013 Distribution of indole-3-acetic acid in Petunia hybrid shoot tip cuttings and relationship between auxin transport, carbohydrate metabolism and adventitious root formation , Planta 2013. Jun 14. [Epub ahead of prin]

2. M.R. **Ghaffari**, M. Mardi, F. Ehya, L. Karimi Farsad, S. Hosseini, B. Ghareyazie, 2010. Mapping and expression analysis of a Fusarium head blight resistance gene candidate pleiotropic drug resistance 5 (*PDR5*) in wheat. Iranian Journal of Biotechnology, 8: 112-116.

3. P. Mahmoodi, S.M. Khayam Nekoui, M. Mardi, S.M. Pirseyedi, M.R. **Ghaffari**, M. Ramroodi, B.A. Siahsar. 2010. Isolation and characterization of new microsatellite marker in *Taxus baccata* L. Conservation Genetics Resources, 2: 195-199. DOI: 10.1007/s12686-009-9133-5.

4. S.M. Pirseyedi, S. Valizadehghan, M. Mardi, M.R. **Ghaffari**, P. Mahmoodi, M. Zahravi, M. Zeinalabedini, S.M. Khayam Nekoui. 2010. Isolation and Characterization of Novel Microsatellite Markers in Pomegranate (*Punica granatum* L.). Int. J. Mol. Sci. 11: 2010-2016. doi:10.3390/ijms11052010.

5. L. Pazouki, M. Mardi, P. Salehi Shanjani, M. Hagidimitriou, S.M. Pirseyedi, M.R. Naghavi, D. Avanzato, E. Vendramin, S. Kafkas, B. Ghareyazie, M. R. **Ghaffari**, S.M. Khayam Nekoui, 2010. Genetic diversity and relationships among Pistacia species and cultivars. Conservation Genetics, 11:311–318. DOI 10.1007/s10592-009-9812-5.

6. P. Salehi Shanjani, M. Mardi, L. Pazouki, M. Hagidimitriou, D. Avanzato, S.M. Pirseyedi, M.R. **Ghaffari**, S.M. Khayam Nekoui, 2009. Analysis of the molecular variation between and within cultivated and wild Pistacia species using AFLPs. Tree Genetics and Genomes, 5:447–458.

7. M. Mazaheri, M.R. Naghavi, M.R. **Ghaffari**, S.M. Pirseyedi, B. Ghareyazie, S. Abdemishani, M. Mardi, 2008. Resistance Gene Analog Polymorphism (RGAP) Markers Co-Localize with the Major QTL of FHB Resistance, *Qfhs.ndsu-3BS*, in Wheat. Iranian Journal of Biotechnology, 6: 119-123.

8. Fathi, B. Ghareyazi, A. Haghnazari, M.R. **Ghaffari**, S.M. Pirseyedi, S. Kadkhodaei, M.R. Naghavi, M. Mardi, 2008. Assessment of the genetic diversity of almond (*Prunus dulcis*) using microsatellite markers and morphological traits. Iranian Journal of Biotechnology, 6: 98-106.

9. A.M. Naji, M. Moghaddam, M.R. **Ghaffari**, H. Pour Irandoost, L. Karimi Farsad, S.M. Pirseyedi, S.A. Mohammadi, B. Ghareyazie, M. Mardi, 2008. Validation of EST-derived STS markers localized on *Qfhs.ndsu-3BS* for Fusarium head blight resistance in wheat using a Wangshuibai‟ derived population. J. Genetics and Genomics 35: 625-629.

10. Eivazi A.R., Naghavi M. R., Hajheidari M., Pirseyedi S.M., **Ghaffari** M.R., Mohammadi S. A., Majidi E., Salekdeh G.H., Mardi M., 2008. Comparative genetic diversity based on quality traits, AFLP, SSR and proteome analyses in wheat (*Triticum aestivum* L.). Annals of Applied Biology 152: 81–91.

11. Naghavi, M.R., Mardi, M., Pirseyedi, S.M., Kazemi, M.B., Potki, P., and **Ghaffari**, M.R. 2007. Comparison of genetic variation among accessions of Aegilops tauschii using AFLP and SSR markers. Genetic Resources and Crop Evolution. 54:237-240

12. Mardi, M., Ghareyazie, B., **Ghaffari**, M.R., Mohammadi, S.A., Buerstmayr, H., and Lemmens, M. 2007. Estimation of the genetic components of *Fusarium* head blight resistance in wheat (*Triticum aestivum* L.). *J. Agric. Sci.* (In press)

13. M. Mardi, M.R. Ghaffari, et al, 2006. Molecular breeding to improve Fusarium head blight resistance in Iran. Canadian Journal of Plant Pathology. 28 (2):38

14. Mardi, M., **Ghaffari**, M.R., Adam, G., Buerstmayer, H., and Ghareyazie, B. 2006. Functional genomics of wheat: development of PCR-based molecular markers for candidate Fusarium resistance genes. Annual wheat newsletter 5.2:69

15. Ehya, F., **Ghaffari**, M.R., Mardi, M., Zaheri, S. and Ghareyazie, B. 2005. Development of PCR based marker for candidate *Fusarium* head blight resistance gene (*PDR5*) in wheat. Iranian J. Crop Science, 6 (4): 395 – 401

Book chapter

1- M. R. Hajirezaei, M.R. **Ghaffari**, B.H. Junker, J. Muller, B. Usadel, M. Leps, R. Lemke, F. Schreiber . Multiscale Metabolic Modeling of cereals – an integrated systems biology approach for research biomas M. R. Hajirezaei, M.R. Ghaffari, B.H. Junker, J. Muller, B.

Usadel, M. Leps, R. Lemke, F. Schreiber. Renewable Resources and Biotechnology for Material Applications. Book Chapter 34. pp. 325-331 Nova Science Publishers, New York

Supervisor and Advisory of Thesis

Mona Mazaheri, MSc.Tehran University, Iran, 2004

Farveh Ehya, MSc., Tabriz University, Iran, 2004

Ali Fathi, MSc., Zanjan University, Iran, 2005

Yousef Nami, MSc., Mazandaran University, Iran, 2006

Awards

Honor Scientific Researcher, Agricultural Biotechnology Research Institute of Iran 2007

Honor Nanotechnology Scientific Technician, Iranian Nanotechnology Initiative, 2008

References:

Prof. Dr. Mohsen Mardi, Agricultural Biotechnology Research Institute of Iran, mardi@abrii.ac.ir

Prof. Dr. Nicolaus von Wiren, Leibniz-Institut für Pflanzengenetik und Kulturpflanzenforschung (IPK), vonwiren@ipk-gatersleben.de

Dr. Mohammad Reza Hajirezaei, Leibniz-Institut für Pflanzengenetik und Kulturpflanzenforschung (IPK), mohammad@ipk-gatersleben.de

_____

Mohammad Reza Ghaffari