Genetic Control of the Response to Sulfur, Nitrogen, and Phosphorus Supply in *Arabidopsis thaliana*

Patrycja Baraniecka

A thesis submitted to the University of East Anglia for the degree of Doctor of Philosophy

John Innes Centre
Norwich
November 2014

Acknowledgements

The last three years were very productive and exciting. This PhD position was a dream come true and I am grateful for all the opportunities I was provided with at the John Innes Centre and within Bionut ITN.

I would like to thank my first supervisor, Stan Kopriva, for offering me the position and bringing me to the environment where I could develop my scientific skills. I would also like to thank my second supervisor, Alison Smith, for taking me to her group and for the continuous support. I sincerely thank the members of my supervisory committee, Tony Miller and Malcolm Hawkesford, for their invaluable advice and constructive discussions. I cannot forget Hannah Daugaard-Hansen. Thank you for taking such good care of me.

I am thankful to Marie Skłodowska-Curie Actions and the Bionut ITN for the funding and for the numerous training and networking opportunities I was offered. Many thanks to all the PIs and students from Bionut network for all the courses, interesting discussions, advice, and a wonderful time we had together. I also thank all my collaborators for the opportunity to work in their labs and the warm welcome.

I would like to acknowledge all members of Kopriva and Smith labs. Special thanks go to Marilyn Pike for her help, support, and a good word she always had for me, and to Cristina Pignocchi for amazing atmosphere in the lab and all the time she spent answering my never-ending questions. Thank you all very much for the time we spent together and all the interesting discussions we had in the office.

My years in Norwich would not be so enjoyable without my wonderful friends. An especially big thank you to my current and previous housemates at La Casa City Road for amazing atmosphere, for all the dinners we had together, and for being great! Special thank goes to Richard Parsons for being there for me over last months and to Michał Bochenek for helping me go through probably the hardest part of that PhD project. I had a lot of fun with all of you and I am really glad that I met you!

I dedicate this work to my mother:

Mamo, byłam tutaj dzięki temu, że nigdy we mnie nie zwątpiłaś. Dziękuję Ci za te trzy lata, za wsparcie i miłość jakich mi nigdy nie szczędzisz. Dziękuję Ci za Twoje zaangażowanie we wszystko co robię. Bez Ciebie nie byłabym tym, kim jestem. Kocham Cię!

Declaration

I certify that the work contained in this thesis is my original work except where due reference is made to other authors:

- Professor David E. Salt and Dai-Yin Chao, University of Aberdeen, Institute of Biological and Environmental Sciences, Cruickshank Building St Machar Drive, Aberdeen, Scotland AB24 3UU, United Kingdom

David E. Salt and Dai-Yin Chao performed the GWAS using EMMA script implemented in R package (Chapter 5). Additionally, they have provided seeds of some accessions used for GWAS as well as mutants of *apr2* used in the analysis described in Chapter 3. The experiment described in Chapter 3 was designed based on the results provided by David E. Salt's laboratory concerning total sulfur and selenium concentration in arabidopsis (Chao et al. 2014).

- Dr. Marco Giovannetti, University of Turin, Department of Life Science and Stystem Biology, Viale Mattioli 25, 10125 Turin, Italy

The analysis of ATPS1 activity, *ATPS1* gene expression, and anion concentration in selected accessions was performed together with Dr. Marco Giovannetti.

- Dr. Fabien Chardon and Giorgiana Chietera, French National Institute of Agricultural Research (INRA) Versailles, 10 Route de Saint-Cyr, 78026 Versailles, France

Dr. Fabien Chardon and Giorgiana Chietera grew 25 arabidopsis accessions in hydroponic cultures with limited source of sulfate. The analysis of these accessions is described in Chapter 4. Additionally, Dr. Fabien Chardon performed the statistical analysis of the data described in Chapter 4 where referred in the description of the figure.

- Dr. Rainer Höfgen, Dr. Sarah Whitcomb, and Dr. Joachim Kopka, Max Planck Institute of Molecular Plant Physiology (MPIMP), Wissenschaftspark Golm, Am Mühlenberg 1, 14476 Potsdam, Germany

The collaborators from MPIMP were involved in the primary metabolite profiling described in Chapter 4. Dr. Sarah Whitcomb (from the group of Dr. Rainer Höfgen) together with me performed the samples extraction. The extracts were analysed using GC-MS by the group of Dr. Joachim Kopka. Subsequently, Dr. Sarah Whitcomb performed the statistical analysis of the metabolite data described in Chapter 4 where referred in the description of the figure.

- Dr. Vasilios Andriotis, John Innes Centre, Department of Metabolic Biology, Colney Lane, NR47UH Norwich, Norfolk, United Kingdom

Dr. Vasilios Andriotis advised me about the analysis of embryo lethality of the knock-out mutant of AT5G03430 described in Chapter 6.

- Professor Ralf R. Mendel, Dr. Tobias Kruse, and Corinna Probst, Braunschweig University of Technology, Department of Plant Biology, 38106 Braunschweig, Germany

The group of Professor Ralf Mendel was involved in the analysis of AT5G03430 and its involvement in molybdenum cofactor (MoCo) metabolism described in Chapter 6. They will continue the work on that gene after my PhD is finished to investigate its exact function.

Apart from the information provided in this section the exact contribution of different collaborators to the results described in particular chapters together with their affiliations is explained in great detail in the last paragraph of the introduction section of each chapter. Additionally, the references are made under particular figures in the results section of each chapter.

Abstract

Sulfur deficiency is a relatively new problem in Europe and the studies on sulfur use efficiency are still lagging behind those on the other major nutrients such as nitrogen or phosphorus. Therefore, the main aim of this work was to improve the understanding of the sulfate assimilation pathway, its regulation and interaction with other elements. In the course of this project natural variation was used to characterise further the regulation of the pathway and to identify new regulatory components. This analysis revealed that the first two enzymes involved in sulfate reduction – ATP sulfurylase and APS reductase – are nearly equally involved in its control but through different mechanisms. Moreover, a Genome-Wide Association Study was conducted on the accumulation of nitrate, phosphate, and sulfate in more than 200 arabidopsis accessions. This analysis resulted in identification of new functions of already known genes which were not previously related to plant nutrition. Additionally, previously undescribed genes were identified disruption of which results in changes in the anion accumulation phenotype.

To characterise arabidopsis response to sulfate and/or nitrate deficiency a collection of genetically divergent accessions grown under different nutrition regimes was examined for a number of morphological and metabolic traits. This analysis resulted in dissection of four different patterns of plant response to sulfate availability. Individual accessions were characterised as best adapted to nutrient deficiency. Traits such as biomass allocation or root architecture were suggested as potential targets in the process of developing new crop varieties. This analysis is unique since, to my knowledge, it is the first one which provides the characterisation of arabidopsis response to nutrient availability based on the analysis of such a large number (25) of natural accessions. The results described here provided new insight into sulfate metabolism and can be used to develop new breeding strategies and improve crop yield and quality.

Table of contents

List of Figures

<u>Supplementary Figures</u>

List of Tables

Supplementary Tables

Abbreviations

AO	Aldehyde oxidase
APK	APS kinase
APR	APS reductase
APS	Adenosine 5'-phosphosulfate
ASF	Absolute sulfate flux
ATPS	ATP sulfurylase
CER	Controlled Environment Room
CGS	Cystathionine synthetase
CSC	Cysteine synthase complex
EMMA	Efficient Mixed Model Association
FDR	False Discovery Rate
FWER	Family Wise Error Rate
GAPIT	Genome Association and Prediction Integrated Tool
GGT	γ-Glutamyltransferase
GSH	Glutathione
GSHS	Glutathione synthase
GST	Glutathione S-transferase
GWAS	Genome-Wide Association Mapping
HPLC	High-Performance Liquid Chromatography
IP	(sulfate) Incorporation into proteins [nmol h^{-1} g^{-1}]
ITh	(sulfate) Incorporation into thiols [nmol h^{-1} g^{-1}]
LD	Linkage Disequilibrium
MoBP	MoCo binding proteins
MoCo	Molybdenum cofactor
MOT	Molybdenum transporter

MPT	Molybdopterin
MS	Methionine synthase
NR	Nitrate reductase
OAS	O-acetylserine
OAS-TL	O-acetylserine (thiol) lyase
PAPR	PAPS reductase
PAPS	3'-phosphoadenosine 5'-phosphosulfate
PRL	Primary root length [cm]
PSK	Phytosulfokines
QTL	Quantitative Trait Locus
RIL	Recombinant Inbred Lines
ROS	Reactive oxygen species
RSF	Relative sulfate flux [%]
RT	Root thickness
S/RCys	Shoot or root cysteine concentration [nmol mg^{-1} FW or DW]
S/RDM	Shoot or root dry matter [mg]
S/RFM	Shoot or root fresh matter [mg]
S/RGSH	Shoot or root GSH concentration [nmol mg^{-1} FW or DW]
S/RH2O	Shoot or root water concentration [%]
S/RNO3	Shoot or root nitrate concentration tent [nmol mg^{-1} FW or DW]
S/RPO4	Shoot or root phosphate concentration [nmol mg^{-1} FW or DW]
S/RSO4	Shoot or root sulfate concentration [nmol mg^{-1} FW or DW]
SAM	S-adenosylmethionine
SAT	Serine acetytransferase
SD	Standard deviation
SE	Standard error
SiR	Sulfite reductase

SLIM1	SULFUR LIMITATION 1
SNP	Single Nucleotide Polymorphism
SO	Sulfite oxidase
SRCys	Shoot to root ratio of cysteine concentration
SULTR	Sulfate transporter
TS	Threonine Synthase
TTS	(sulfate) Translocation to the shoots [nmol h^{-1} g^{-1}]
TU	Total sulfate uptake [nmol h^{-1} g^{-1}]
v/v	Volume/volume
w/v	Weight/volume
WT	Wild-type
XDH	Xantine dehydrogenase
γ-EC	γ-glutamylcysteine
γ-ECS	γ-glutamylcysteine synthetase

Chapter 1:

General Introduction

1.1 Sulfur nutrition

Plants are autotrophic organisms and as such they are able to assimilate mineral nutrients from the environment (White & Brown 2010). They can also convert the non-mineral nutrients – oxygen, carbon, and hydrogen – into sugars during photosynthesis. Additionally, for sufficient nutrition plants require at least 14 mineral elements that can be acquired from the soil. Six of them – nitrogen, phosphorus, potassium, calcium, magnesium and sulfur – are required in large amounts and these are called macronutrients. The requirement for the rest of them – chlorine, boron, iron, manganese, copper, zinc, nickel, and molybdenum – is much lower and these are called micronutrients. In case of deficiency of essential nutrients in the soil, they can be supplied to plants as fertilizers to increase plant yield and quality. This practice applies mainly to crop production (White & Brown 2010). However, it needs to be used carefully in agriculture as inappropriate application of fertilisers may disturb the mineral balance of the soil solution which may cause inhibition of plant growth and reduction of crop yield (White & Brown 2010).

Adequate sulfur supply is required for proper growth and health of all living organisms (Takahashi et al. 2011). Sulfur is cycled in the global ecosystem and can be converted to its organic compounds by photosynthetic organisms and microorganisms. Plants, most bacteria and fungi use inorganic sulfate as a source of sulfur. Plants can also use atmospheric sulfur in form of sulfur dioxide and hydrogen sulfide (Durenkamp & De Kok 2004). In contrast, animals and humans need to obtain organic sulfur compounds e.g. methionine with their food (Nimni et al. 2007). The most common form of sulfur in nature is sulfate (SO_4^{2-}) – the most oxidized form in which sulfur is in the +VI redox state. Sulfate is taken up from the soil and reduced to sulfide in an energy-dependent reduction pathway. Sulfide can be incorporated into O-acetylserine to form cysteine which is the first stable form of bound reduced sulfur in plants.

Reduced sulfur is present in a wide variety of metabolites with specific biological functions (Takahashi et al. 2011). Apart from cysteine, sulfur is present in methionine, peptides and proteins (in sulfur containing amino acids or in iron-sulfur clusters), and other metabolites. Moreover, a number of molecules such as sulfolipids, sulfated hormones i.e. phytosulfokines or polysaccharides contain sulfur in its oxidised form. Sulfate also plays a crucial role in stabilising protein structure and functions (disulfude bonds between the sulfur containing amino acids). Since it can be found in a number of coenzymes and other catalytic biomolecules it also has an important catalytic function. Glutathione (GSH), which is the second main product of the sulfate reduction pathway, has a crucial role in the control of redox state of

plant cells and the elimination of reactive oxygen species (Hawkesford & De Kok 2006, Leustek et al. 2000). Sulfur is also present in vitamins and cofactors such as coenzyme A, thiamine and biotin (Hell et al. 2002). Sulfur containing secondary metabolites have important functions in defence against biotic stress (Halkier & Gershenzon 2006). The breakdown products of alliins and glucosinolates deter plant pests and are responsible for characteristic smell and taste of many vegetables (Halkier & Gershenzon 2006). Both of these classes of natural products are also very beneficial for human health. Alliins, found in large amounts in garlic, have antimicrobial properties (Chung 2006), and glucosinolate degradation products induce enzymes that prevent tumour formation in humans (Wu et al. 2009).

It has been reported that sulfur availability is continuously decreasing in many areas of Europe mainly because of the changes in fertiliser practice and strong decrease of sulfur inputs from atmospheric deposition due to reduction in the emission of sulfur dioxide (Blake-Kalff et al. 2001, McGrath et al. 1996, Zhao et al. 1996). Moreover, intensive agriculture programmes and optimisation of plant breeding strategies led to increase in plant demands (Abdallah et al. 2010). Sulfur deficiency in crop plants not only reduces the yield but has also a very negative effect on the crop quality which results in a decrease of the nutritional value of food and affects human health. This is particularly important in case of wheat breeding and the maintenance of baking quality (Shahsavani & Gholami 2008). The protein fraction is known to play an essential role in bread-making quality of wheat. The gluten proteins, gliadins and glutenins, represent about 80-85% of total flour protein. These are responsible for elasticity and extensibility that are essential for functionality of wheat flours (Hussain et al. 2012, Kuktaite et al. 2004). However, in sulfur limited conditions the synthesis of sulfur-poor storage proteins such as ω-gliadin and the high molecular weight subunits of glutenin is favoured at the expense of sulfur-rich proteins, which may cause unpredictable and unwanted variations in wheat quality (Flæte et al. 2005, Moss et al. 1981). Another problem, discovered relatively recently, is the formation of acrylamide during high-temperature processing of potato and wheat (Tareke et al. 2002). Acrylamide is potentially carcinogenic to humans. It also has negative neurological and reproductive effects (Friedman 2003). The major determinant of acrylamide forming potential is the concentration of free asparagine (Curtis et al. 2010). The accumulation of free asparagine in wheat grain during severe sulfur deprivation may be 30-fold greater level than in sulfur sufficient conditions, which makes asparagine up to 50% of the total free amino acid pool. For that reason, even very small amounts of such grain entering the food chain could have a significant effect on acrylamide formation, which makes the application of sulfur fertilizers very important (Halford et al. 2012).

Studies on model species provide an important tool for exploring metabolic processes mainly because of a large amount of publically available information (genomic, metabolic, transcriptomic) and the rapid life cycle of the plants. The knowledge obtained from research on model plants can subsequently be transferred to crops and used for improving crop breeding strategies. For these reasons the work described in this thesis is focused on sulfur metabolism in model plant *Arabidopsis thaliana (*L.) Heynh.

1.2 Sulfate transport

Sulfate uptake from the soil is the first stage of plant sulfate metabolism (Takahashi et al. 2011). After entry into plants, sulfate needs to be delivered to the plastids for assimilation or to the vacuoles for storage. The cell-to-cell transport as well as long-distance transport between organs required to fulfil the source/sink demands during plant growth involve specific sulfate transporter proteins (Buchner et al. 2004b). Genes encoding these proteins belong to the sulfate transporter gene family and are divided into five groups (Figure 1.1; Hawkesford 2003). Members of different groups vary in their kinetics of transport and in expression pattern indicating different functions in the process of sulfate uptake and distribution.

The influx of sulfate through the plasma membrane is well characterised at the physiological and functional level. Plants essentially use a proton/sulfate co-transport system to mediate sulfate flux (Lass & Ullrich-Eberius 1984). A proton gradient is generated by a plasma membrane proton ATPase (Saito 2004) and the transport process is pH dependent with $3H^+$/ sulfate stoichiometry (Hawkesford et al. 1993, Smith et al. 1995a). Products encoded by sulfate transporter genes usually possess 12 putative membrane-spanning domains and belong to a large family of cation/solute co-transporters (Saito 2000). Additionally, the analysis of the C-terminal region of these transporters revealed the presence of STAS (sulfate transporters and antisigma factor antagonists) domains (Aravind & Koonin 2000). It was shown that deletion of the STAS domain affects the localization of transporters to the membrane (Shibagaki & Grossman 2004) and results in a loss of sulfate transport activity (Rouached et al. 2005).

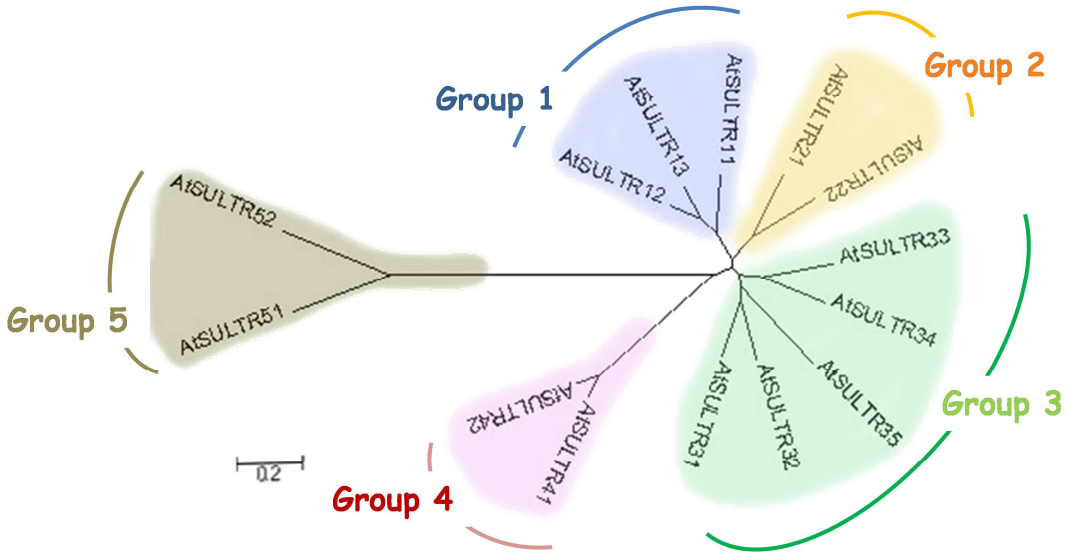


Figure 1.1 *Sulfate transporter family*

Unrooted phylogenetic tree of the *Arabidopsis thaliana* members of the sulfate transporter family was drawn using MEGA5.1 software. Different colours represent different subfamilies (groups).

1.2.1 Functions of sulfate transporters in Arabidopsis thaliana

The first sulfate transporter was identified by functional complementation of a sulfate transporter-deficient yeast mutant (Smith et al. 1995a, Smith et al. 1995b). Subsequently, yeast deletion mutants became the main tool for functional characterisation of sulfate transporters. Up to date a number of sulfate transporters from a variety of plant species have been described (Buchner et al. 2004a, Howarth et al. 2003, Vidmar et al. 2000, Yoshimoto et al. 2003, Yoshimoto et al. 2002). The availability of fully sequenced genomes of *Arabidopsis thaliana* and rice has enabled analyses which led to identification of 14 putative sulfate transporter genes in each genome.

The analysis of the first isolated sulfate transporters identified high and low-affinity transport patterns (Smith et al. 1995a). More detailed analysis revealed that the high-affinity components facilitate uptake of sulfate into the plant root (Barberon et al. 2008, Yoshimoto et al. 2007). In *Arabidopsis thaliana* the high-affinity sulfate transporters comprise 3 genes (SULTR1;1-3) and belong to the clade which forms group 1 of sulfate transporters (Figure 1.1). The primary sulfate acquisition in roots is mediated by two of them: SULTR1;1 and SULTR1;2 (Figure 1.2). Various studies confirmed the expression of these two genes in root hairs, root epidermal and root cortical cells (Rae & Smith 2002, Takahashi et al. 2000, Yoshimoto et al. 2002) suggesting the capacity of these tissues for high-affinity sulfate influx into the symplast.

SULTR1;3 appears to be an exception in this group because it is located only in the phloem in both roots and cotyledons (Yoshimoto et al. 2003). SULTR1;2 mediates sulfate uptake in normal conditions and in sulfur deficiency. Its expression is relatively independent of sulfate supply. In contrast, SULTR1;1 is strongly induced by sulfate limitation, but almost absent under sufficient sulfur supply (Clarkson et al. 1983, Howarth et al. 2003, Yoshimoto et al. 2002). Another study showed that in mutants deficient in SULTR1;2 the expression of SULTR1;1 is slightly up-regulated, however reduced growth of *sultr1;2* mutants (relative to WT plants) suggests that the SULTR1;1 is not able to compensate for the missing SULTR1;2. This might indicate that SULTR1;2 is the major component for sulfate acquisition (Maruyama-Nakashita et al. 2003).

Sulfate taken up by the epidermis cells needs to be transferred across the root cells to the xylem (Figure 1.2; Takahashi et al. 2011). This step is necessary to deliver it to the target cells in shoot organs for reduction or storage in the vacuole. The horizontal sulfate transfer from the epidermis to the central cylinder cells may occur via plasmodesmata. This is the likely strategy to cross the barrier of the Casparian strip at the endodermal cell layers. During this process sulfate may leak from the symplast to apoplast. This mechanism seems to be passive, but has not yet been characterised (Takahashi et al. 2011). The transporter for efflux of sulfate into the xylem vessels is still unknown. However, the expression patterns of arabidopsis low-affinity sulfate transporters from group 2 (Figure 1.1) in the central cylinder cells suggest that they may contribute to long distance sulfate transport (Figure 1.2). In roots SULTR2;1 is expressed in the xylem parenchyma and pericycle cells whereas SULTR2;2 is restricted to the root phloem. In contrast, in leaves SULTR2;1 is expressed in xylem parenchyma and phloem cells and SULTR2;2 in the cells surrounding the xylem vessels (Takahashi et al. 2000). Additionally, SULTR2;1 was assumed to be involved in sulfate transport into developing seeds (Awazuhara et al. 2005). Localization of SULTR2;2 suggests its role in sulfate transport via the phloem. The efflux of sulfate to the apoplast of the root vascular tissue leads to a high sulfate concentration. SULTR2;1 expressed in the xylem parenchyma cells can reabsorb it, regulating the amount of sulfate which is transported to the shoots (Buchner et al. 2004a). In the leaf, the expression of SULTR2;2 in the closest cells to the xylem vessels suggests its role in sulfate uptake from the vessels. Subsequently, sulfate is probably transferred to cells where it will be assimilated (Figure 1.2). The expression of SULTR2;1 in the phloem suggests its role in sulfate transfer to other organs, and in xylem parenchyma – reabsorption for further xylem transport (Buchner et al. 2004b). Taken together, it seems that sulfate transporters from group 2 are

involved in the balancing of the sulfate flux through the plant under changing sulfate availability (Takahashi et al. 2000).

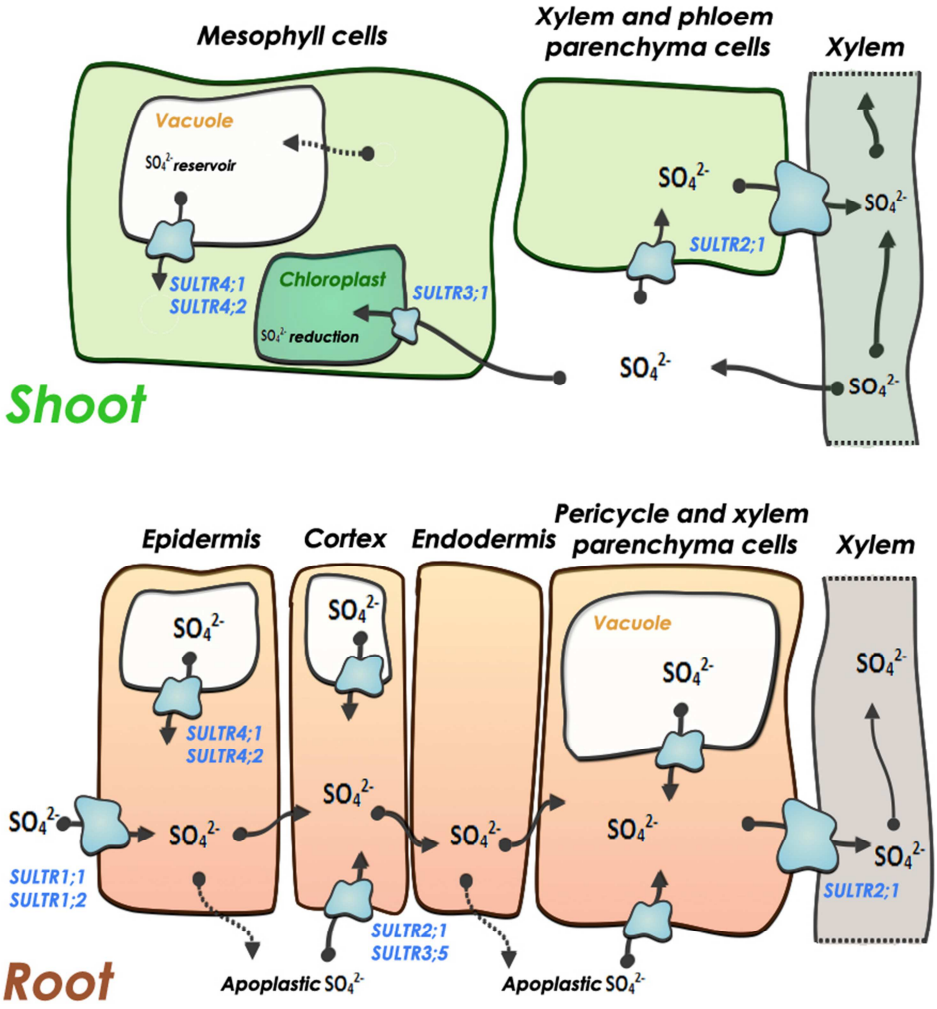


Figure 1.2 ***Sulfate transport system in Arabidopsis thaliana***

Blue shapes and labels indicate sulfate transporters (SULTRs) mediating transport across plasma membranes. Dashed arrows indicate yet unknown transport pathways. Figure is based on Takahashi et al. 2011.

Plastids are the final destination for sulfate where it is metabolised (Figure 1.2; Takahashi et al. 2011). A chloroplast sulfate transporter SULTR3;1 has been characterised only recently (Cao et al. 2013). Its chloroplast subcellular localization was confirmed with the analysis of plants expressing *SULTR3;1*-GFP constructs and the sulfate transport functionality with a validated *in organello* assay. However, chloroplast sulfate uptake did not change despite the disruption of *SULTR3;1* suggesting the existence of other chloroplast sulfate transporters.

The analysis of other members of SULTR3 subfamily revealed that *SULTR3;2, SULTR3;3, SULTR3;4,* but not *SULTR3;5* might be also chloroplast sulfate transporters (Cao et al. 2013). These results are consistent with those of Kataoka et al. (2004a) who have shown an essential role for SULTR3;5 in the vascular root-to-shoot transport.

Alternatively, sulfate can also be transported to the vacuoles, which play the role as sulfate reservoirs in cells (Figure 1.2; Takahashi et al. 2011). It was shown that vacuoles isolated from the arabidopsis *sultr4;1/sultr4;2* double mutant contain more sulfate than the wild type suggesting that the efflux of sulfate from the vacuole is mediated by sulfate transporters from group 4. Moreover, enhanced expression of these two genes under sulfur stress conditions indicates that sulfate is released from vacuole in response to sulfur demands in the cells (Kataoka et al. 2004b). The vacuole influx transporters have not yet been identified.

Members of group 5 differ significantly from the other groups (Figure 1.1; Takahashi et al. 2011). This group contains two isoforms which are also dissimilar to each other. SULTR5;2 was described to be involved in molybdenum transport (Baxter et al. 2008, Tomatsu et al. 2007) and renamed as a MOT1 (Molybdenum transporter 1) transporter. Tomatsu et al. (2007) have localised this transporter to the endomembrane system and plasma membrane whereas Baxter et al. (2008) suggest its localization in the mitochondrial membrane. This discrepancy might be explained by different sites of GFP fusion, either to the N-terminal end of *MOT1* (Tomatsu et al. 2007) or to its C-terminus (Baxter et al. 2008). Nevertheless, the exact location of MOT1 requires further investigation. *SULTR5;1* was shown to be expressed in most plant tissues, but it was not affected by sulfate supply (Shinmachi et al. 2010). Its function in the export of molybdate from the vacuole was shown only recently and it was renamed as MOT2 (Gasber et al. 2011). As there are no reports indicating sulfate transport function of these two transporters and because of the absence of the STAS domain which is present in all other sulfate transporters it could be considered to exclude these two genes from the sulfate transporter family.

1.3 Sulfate assimilation and metabolism in *Arabidopsis thaliana*

In nature sulfur occurs in many oxidation states in inorganic, organic and bioorganic compounds (Takahashi et al. 2011). Various organisms such as algae, bacteria, plants and fungi are able to reduce sulfate and incorporate it into amino acids in the assimilatory sulfate reduction pathway. This process is very well described on both biochemical and molecular levels. In photosynthetic organisms it occurs in plastids (Brunold & Suter 1989). The only

exception is *Euglena gracilis* where sulfate reduction takes place in mitochondria (Brunold & Schiff 1976).

Because of a very low oxidation/reduction potential of sulfate compared to other cellular reductants it needs to be activated before it can enter the reduction pathway (Figure 1.3; Takahashi et al. 2011). In the activation reaction a high energy anhydride bond is produced between inorganic sulfate and phosphate delivered from ATP. It yields adenosine 5'-phosphosulfate (APS) and releases pyrophosphate. This reaction is catalysed by ATP sulfurylase (ATPS; EC: 2.7.7.4; Schmidt & Jager 1992) and is the first step of sulfur assimilation. Moreover, it is the only step of the pathway which is common to primary and secondary sulfate metabolism. APS forms a branching point in the sulfate reduction pathway (Figure 1.3). It can be directly reduced to sulfite by APS reductase (APR; EC: 1.8.99.2) or phosphorylated by APS kinase (APK; EC: 2.7.1.25) to form 3'-phosphoadenosine 5'-phosphosulfate (PAPS) which serves as a donor of activated sulfate for various cellular reactions modifying proteins, saccharides or synthesis of secondary metabolites including glucosinolates. This process is called sulfation and it is important in regulation of plant growth and development. However, sulfate reduction is a dominant route for assimilation (Leustek et al. 2000) and it is fulfilled in two steps. In the first step APR transfers two electrons to APS to produce sulfite. The electrons are derived from glutathione (GSH; Bick et al. 1998). Subsequently, sulfite is reduced to sulfide by ferredoxin dependent sulfite reductase (SiR; EC: 1.8.7.1). This reaction requires a transfer of six electrons from ferredoxin to sulfite. Sulfide is then incorporated into the amino acid skeleton of O-acetylserine (OAS) to form cysteine. This reaction is catalysed by OAS thiol-lyase (OAS-TL; EC: 2.5.1.47; Kopriva 2006, Leustek et al. 2000, Takahashi et al. 2011). A detailed description of isoform 1 of ATPS and isoform 2 of APR, and the characterisation of their regulatory functions over the sulfate assimilation pathway are the subject of Chapter 3 of this thesis.

APK catalyses the transfer of phosphate from ATP to APS leading to the formation of PAPS (Mugford et al. 2009). In arabidopsis APK is encoded by four genes, all located on different chromosomes and with a high level of similarity. Three of the isoforms contain chloroplast transit peptides at the N-termini and these have been confirmed to be localized in plastids. The APK3 isoform does not contain the N-terminal extension and it is likely to be responsible for cytosolic activity. Little is known about the biochemistry and the functions of the individual plant APKs. However, they have a significant effect on sulfur metabolism. It was shown that *apk1apk2* double mutant has a dramatically low glucosinolate concentration and also substantially higher concentration of cysteine and glutathione than wild type plants (Mugford et al. 2009). This suggests a compensation of low glucosinolate concentration by an

increase in cysteine and glutathione. This might indicate that primary sulfate metabolism is up-regulated in this mutant implying an important role of APK in controlling sulfur distribution in plants (Kopriva et al. 2012, Mugford et al. 2009).

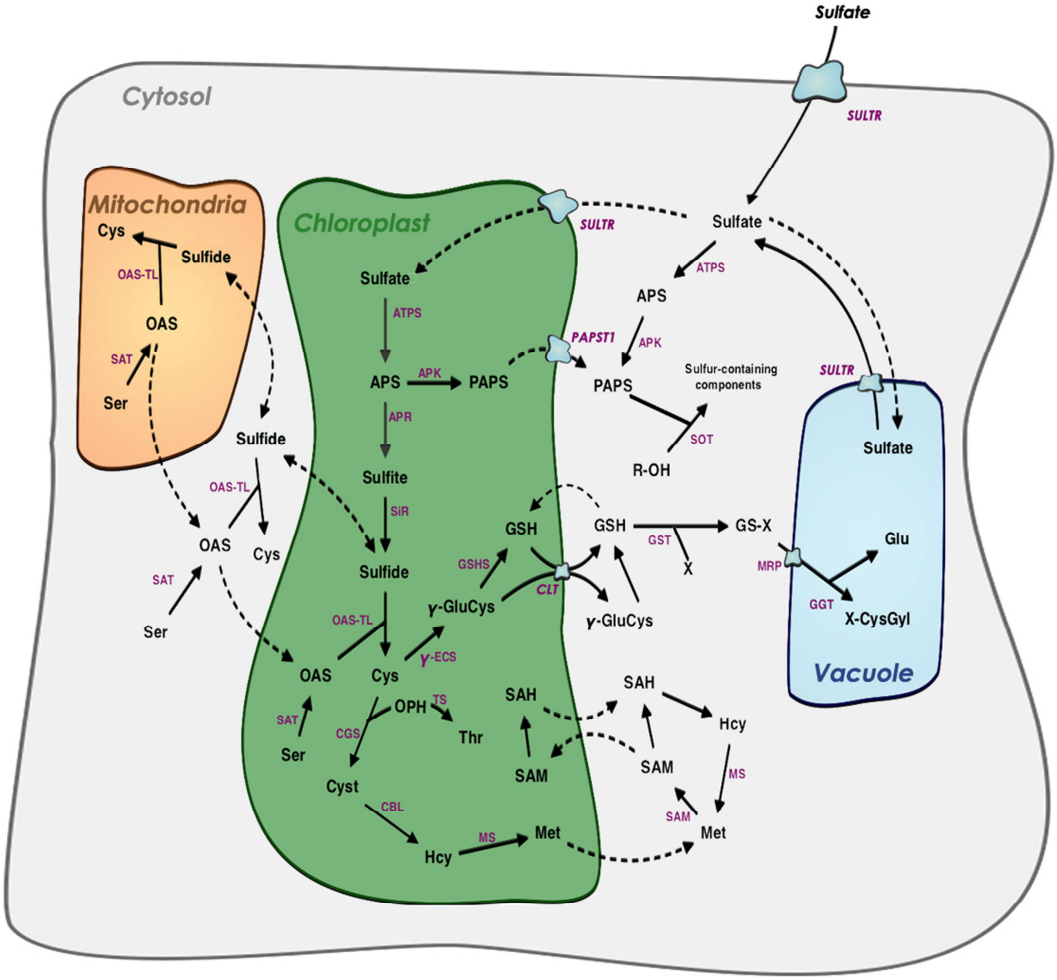


Figure 1.3 ***Cellular organization of sulfate metabolism in Arabidopsis thaliana***

Enzymes and transporters are indicated in purple characters. Abbreviations of enzymes and transporters: ATPS, ATP sulfurylase; APK, APS kinase; APR, APS reductase; SiR, sulfite reductase; OAS-TL, OAS(thiol)lyase; SAT, serine acetyltransferase; CGS, cystathionine γ-synthase; CBL, cystathionine β-lyase; TS, threonine synthase; MS, methionine synthase; SAM, S-adenosylmethionine synthetase; γ-ECS, γ-glutamylcysteine synthetase; GSHS, glutathione synthetase; CLT, thiol transporter; GST, glutathione-S-transferase; MRP, multidrug resistance-associated protein; GGT, γ-glutamyltransferase; SOT, sulfotransferase; SULTR, sulfate transporter; PAPST1, plastidic PAPS transporter. Abbreviations of metabolites: APS, adenosine 5'-phosphosulfate; Cys, cysteine; Cyst, cystathionine; Hcy, homocysteine; OPH, O-phosphohomoserine; Thr, threonine; Met, methionine; SAM, S-adenosylmethionine; SAH, S-adenosylhomocysteine; γ-GluCys, γ-glutamylcysteine; GSH, glutathione; GS-X, glutathione conjugate; Glu, glutamate; X-CysGyl, cysteinylglycine conjugate; Ser, serine; OAS, O-acetylserine; PAPS, 3'-phosphoadenosine 5'-phosphosulfate; R-OH, hydroxylated precursor. Figure is based on Takahashi et al. 2011.

The six electron reduction of sulfite to sulfide is catalysed by sulfite reductase (SiR) in plastids (Nakayama et al. 2000). Plant SiR is a 65 kDa monomer. It requires the presence of siroheme and iron-sulfur cluster as cofactors, and ferredoxin as an electron donor. In contrast to other enzymes of the sulfur metabolism pathway, SiR is encoded by a single gene in arabidopsis. The amino acid sequence and protein structure are very similar to nitrite reductase (NiR) which catalyses a six electron reduction of nitrite to ammonia in the nitrate assimilation pathway.

1.3.1 Cysteine biosynthesis

Cysteine is the key sulfur containing compound in plants (Takahashi et al. 2011). It is synthesised by incorporation of sulfide into the β-position of the serine carbon skeleton in the terminal step of sulfur assimilation (Saito 2004). Before the incorporation of sulfide, serine needs to be activated to O-acetylserine (OAS). This process occurs by acetyl transfer from acetyl coenzyme A which is catalysed by serine acetyltransferase (SAT; Serat; EC 2.1.3.30). Subsequently, OAS and sulfide are the substrates for O-acetylserine (thiol) lyase (OAS-TL; EC: 2.5.1.47) which catalyses the β-replacement reaction (Hell & Wirtz 2011, Takahashi et al. 2011). The enzymes involved in the cysteine biosynthesis process, SAT and OAS-TL, are localized in plastids, mitochondria and cytosol (Figure 1.3; Saito 2000).

The analysis of whole plant protein extracts showed that SAT activity is always associated with OAS-TL, and that an excess of free active OAS-TL is present (Hell & Wirtz 2011). This and other results indicate that these two enzymes form a hetero-oligomeric cysteine synthase complex (Hell & Wirtz 2008, Hell & Wirtz 2011, Saito 2004, Takahashi et al. 2011, Wirtz et al. 2001, Wirtz & Hell 2006). The binding of OAS-TL to SAT stabilizes SAT. In the complex, SAT is the only active enzyme. The product OAS causes the release of OAS-TL from the complex which enables the conversion of OAS to cysteine by the free enzyme and reduces the rate of OAS formation (Hesse et al. 2004b). Therefore, formation of the cysteine synthase complex appears to be the main regulatory step in cysteine synthesis in arabidopsis (Hell & Wirtz 2008).

The crystallisation of the SAT protein revealed that it is a hexamer composed from 29 kDa subunits (Olsen et al. 2004). It is folded in a left-handed parallel β-helix which is characteristic for this protein family. The SAT gene family includes five members in arabidopsis two of which were recognized only recently (Hell & Wirtz 2008, Kawashima et al. 2005). Among the five SAT proteins three of them, SAT2, 4, and 5 (Serat 3.1, 3.2, and 1.1 respectively)

are located in the cytosol whereas SAT1 (Serat 2.1) was found in plastids and SAT3 (Serat 2.2) in mitochondria.

OAS-TL belongs to the β-replacement enzyme family which requires pyridoxal-5'-phosphate as a cofactor (Hell & Wirtz 2008). The protein is a homodimer composed of two 35 kDa subunits. The arabidopsis gene family contains nine members which encode eight functionally transcribed proteins. Three of them, called OAS-TL A, B, and C, are thought to be the main OAS-TL proteins in plant cells. Similarly to SAT they are localised in the cytosol, plastids and mitochondria, respectively (Wirtz et al. 2004). OAS-TL proteins seem to have a wide range of functions. It is likely that apart from cysteine synthesis they are also responsible for other processes such as sulfide and cyanide detoxification in mitochondria (Alvarez et al. 2012) or determination of antioxidative capacity in the cytosol (López-Martín et al. 2008). They also seem to be involved in the synthesis of secondary metabolites in various plant species.

A recent study of SAT and OAS-TL mutants suggests the cytosol as the main cell compartment for cysteine production and mitochondria as the main place for OAS synthesis (Figure 1.3; Haas et al. 2008, Krueger et al. 2009, Watanabe et al. 2008). Plants with decreased mitochondrial SAT activity show strongly reduced OAS levels and reduced flux into cysteine and glutathione (Haas et al. 2008). The analyses of OAS concentration and SAT activity in non-aqueous gradients showed the largest amount of OAS in mitochondria and the smallest in plastids, whereas the OAS-TL activity was localised in the cytosol and plastids (Krueger et al. 2009). OAS can, however, be transferred between all three compartments. Additionally, the analysis of compartment-specific OAS-TL mutants revealed reduced cysteine concentration only in a mutant lacking the cytosolic OAS-TL isoform (Haas et al. 2008, Watanabe et al. 2008). Consequently, OAS has to be transported from mitochondria to the cytosol for efficient cysteine biosynthesis. Taking into account that plastids are the main compartment for sulfide production, the presence of sulfide in cytosol also requires transport across chloroplast envelope membrane. Taken together, the very low SAT activity in plastids, the presence of sulfide in the cytosol and the cytosolic localisation of cysteine strongly suggest the cytosol as the main cellular compartment for cysteine biosynthesis in arabidopsis (Krueger et al. 2009).

1.3.2 Glutathione biosynthesis and functions

Formation of cysteine is the terminal step of the sulfate assimilation pathway and the starting point for production of methionine, glutathione and many other sulfur-containing compounds (Figure 1.3; Takahashi et al. 2011,). Glutathione (GSH) is the main thiol-containing molecule in plant cells and it is present in much higher concentrations than cysteine (Noctor et

al. 2012). It has a broad range of functions which include removal of reactive oxygen species (ROS), detoxification of heavy metals and xenobiotics, sulfur donation, transport and storage (in catalytic reactions), redox signalling and many others. It is synthesised from glutamate, cysteine, and glycine by two enzymes: γ-glutamylcysteine synthetase (γ-ECS; EC: 6.3.2.2) and glutathione synthetase (GSHS; EC: 6.3.2.3). The reaction consumes two ATP molecules. The enzyme that catalyses the first step of GSH biosynthesis – γ-ECS – is redox sensitive; its oxidised form has high activity, whereas activity of the reduced form is much lower. The increase in γ-ECS transcript abundance in response to various environmental changes suggests its role as regulatory factor (Xiang & Oliver 1998). It has been shown that γ-ECS is also inhibited by higher concentrations of GSH.

Synthesis of GSH is regulated by cysteine availability (Noctor et al. 2002). In arabidopsis GSH is synthesised in cytosol and plastids where γ-ECS is localised to plastids only and GSHS activity is distributed between plastids and cytosol (Figure 1.3). GSH from leaves is transported to roots, seeds and fruits via the phloem (Leustek et al. 2000) consistent with its important role as a sulfur donor. Therefore, GSH degradation is a very important process in plants; however, the mechanism is not well understood. The main enzymes responsible for GSH degradation include glutathione reductase, γ-glutamyltransferase (GGT EC: 2.3.2.2), glutathione S-transferase (GST), and glutaredoxin. It seems that GSH turnover in cells is maintained mainly by GGT activities (Takahashi et al. 2011). However, it also has been shown that the intracellular degradation of GSH is initiated by γ-glutamylcyclotransferase (Ohkama-Ohtsu et al. 2008).

1.3.3 Methionine biosynthesis

Methionine is a sulfur-containing amino acid that belongs to the aspartate family of amino acids together with lysine, threonine, leucine and isoleucine (Ravanel et al. 2004). It is an essential amino acid for mammals and must be delivered entirely from the diet. Plants are able to synthesise it *de novo* from cysteine or homocysteine. Methionine plays an important role as a protein component and it is involved in initiation of translation. S-adenosylmethionine (SAM) which is produced from methionine is a methyl-group donor and precursor of important secondary metabolites. It was shown that 80% of methionine is used for SAM synthesis whereas 20% is incorporated into proteins (Giovanelli et al. 1985). Synthesis of methionine requires products of three metabolic pathways: the carbon skeleton originates from aspartate, the sulfur atom from cysteine, and the methyl group from serine (Ravanel et al. 1998).

In higher plants methionine synthesis starts from a γ-replacement reaction catalysed by cystathionine γ-synthase (CGS) leading to formation of cystathionine from cysteine and O-phosphohomoserine (OPH; Figure 1.3). Cystathionine is converted to homocysteine by α,β-elimination catalysed by cystathionine β-lyase. The final step includes the transfer of methyl group from N^5-methyl-tetrahydrofolate to homocysteine which in plants is catalysed by methionine synthase (MS; Hesse et al. 2004a, Ravanel et al. 1998). In arabidopsis three MS isoforms are known. Two of them are in the cytosol and the third is located in plastids (Ravanel et al. 2004). Subsequently, methionine is converted to SAM by SAM synthetase which requires ATP. It was shown that accumulation of SAM inhibits the enzyme activity (Ravanel et al. 1998). When SAM is used for synthesis of ethylene or polyamines, methylthioadenosine (MTA) is produced as intermediate. MTA can be used for synthesis of another methionine molecule increasing SAM availability as a methyl-group donor (Burstenbinder et al. 2007).

1.4 Secondary sulfate metabolism

Sulfur is also present in plant metabolites as sulfo-groups modifying carbohydrates, proteins, and many natural products. Many sulfated metabolites play distinct roles in plant defence against biotic and abiotic stresses.

1.4.1 Glucosinolates

Glucosinolates are the best known group of secondary sulfate containing compounds (Halkier & Gershenzon 2006). They play an important role in protection against herbivores. They are also responsible for taste and flavour of many *Brassica* vegetables (e.g. cabbage, broccoli). Products of glucosinolate degradation, isothiocyanates, possess an anticarcinogenic activity in mammalian cells (Mithen et al. 2003).

In general, glucosinolates are synthesised in a three-phase pathway (Figure 1.4; Halkier & Gershenzon 2006). The elongation of certain amino acids which are the precursors of aliphatic or aromatic glucosinolates by sequential insertions of few (up to nine) methylene groups into the side chain is the first phase of the process. Subsequently, the amino acid moiety elongated or not, is converted to form the core structure. The biosynthesis of glucosinolate structure involves intermediates common to all glucosinolates. In this process sulfur from cysteine is incorporated into the structure via a yet unknown enzyme. The final block concerns various modifications, such as hydroxylation, O-methylation, desaturation, acylation and others, which leads to formation of a broad range of structures (Halkier & Gershenzon 2006). To date, more than 140 different glucosinolates have been described

(Fahey et al. 2001). In arabidopsis nearly 30 different glucosinolates have been found in most organs, at various developmental stages (Brown et al. 2003). The chemical structure of most of them consists of a β-D-thioglucose group linked to a (Z)-N-hydroximinosulfate ester via a single sulfur atom (Halkier & Gershenzon 2006).

Although knowledge about glucosinolate biosynthesis is important, it is their degradation products that are responsible for most of their biological functions (Figure 1.4; Halkier & Gershenzon 2006). The process of hydrolysis begins with breakdown of the thioglucoside bond, which leads to the formation of glucose and unstable aglycone. The latter can then isomerise to different products depending on the structure of side chain and availability of various cofactors. The initiation process is catalysed by myrosinase (EC: 3.2.3.1) and has been investigated in a number of biochemical and molecular studies. Myrosinase is separated from glucosinolates in idioblast cells to avoid unnecessary glucosinolate hydrolysis. These two components however, mix very quickly after the loss of cellular integrity as a result of wounding or insect or pathogen attack, to activate the binary glucosinolate – myrosinase system leading to the generation of glucosinolate hydrolysis products which serve as plant defence molecules (Renwick 2001). The system however seems to work in two ways: it deters usual pathogens but it may also attract some specialized herbivores. Many of them use glucosinolates as a food or use glucosinolate-containing plants for oviposition (Halkier & Gershenzon 2006).

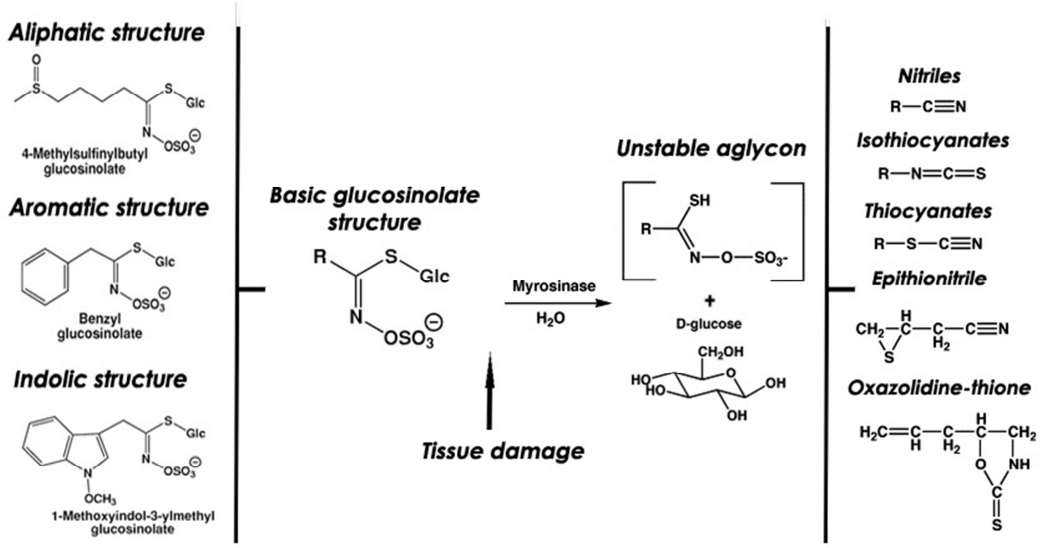


Figure 1.4 ***Glucosinolate metabolism***

Schematic representation of glucosinolate hydrolysis (middle) including examples of aliphatic, aromatic and indolic glucosinolates (left) and most common degradation products (right). Figure is based on Dinkova-Kostova & Kostov 2012.

1.4.2 Phytosulfokines

Phytosulfokines (PSK) are other sulfur-containing metabolites of great importance to plant growth and health (Chen et al. 2000). They promote various stages of plant growth such as somatic embryogenesis, adventitious bud, root formation, and pollen germination (Chen et al. 2000, Kobayashi et al. 1999, Matsubayashi & Sakagami 1996, Matsubayashi et al. 1997, Yamakawa et al. 1998). PSK-α was the first sulfated peptide found in plants (Matsubayashi & Sakagami 1996). It has been shown to strongly promote cell proliferation at low concentrations (Matsubayashi et al. 1997). PSK-α is universally distributed in the plant kingdom (Yang et al. 2000). PSK is synthesised from an ~80 amino acid precursor peptide (PP-PSK) which has a 22-amino acid hydrophobic secretion signal at the N-terminus (Yang et al. 1999). Predicted mature form of PSK-α is a peptide of 67 amino acids (Yang et al. 1999). The sulfation of PP-PSK is crucial for its biological activity (Matsubayashi & Sakagami 1996). Tyrosine is the only amino acid which has been shown to undergo sulfation and sulfated tyrosines are usually surrounded by acidic residues (Dorner & Kaufman 1990, Yang et al. 2000). There are two such tyrosines (surrounded by acidic residues) in PP-PSK (Yang et al. 2000). It has been suggested that both of them undergo sulfation (Yang et al. 1999). The analysis of a synthetic PP-PSK oligopeptide revealed that the acidic amino acid residues are crucial for the sulfation reaction and that the reaction is catalysed by tyrosylprotein sulfotransferase (Hanai et al. 2000) – an enzyme homologous to that found in human (Ouyang et al. 1998).

In arabidopsis four genes encoding the precursor peptide have been identified, and two of them characterised in detail (Yang et al. 2001). The expression of PSK genes in arabidopsis is not limited to tissues characterised by active cell division and differentiation. It has been detected in most plant organs including mature leaves, stems, roots, and callises which indicate that it is not a simple mitogen or differentiation initiator. PSK precursor overexpression causes no apparent changes in plant growth or development under normal growth conditions (Matsubayashi & Sakagami 2006).

1.5 Regulation of sulfate assimilation

Understanding the tight regulation of the sulfate assimilation pathway is extremely important for two reasons: the essential role of sulfur for plant growth and crops quality, and the potential cytotoxicity of sulfite and sulfide, which are intermediates in sulfate assimilation. Various stages of the assimilatory pathway are regulated in both positive and negative feedback mechanisms in a demand-driven manner (Figure 1.5).

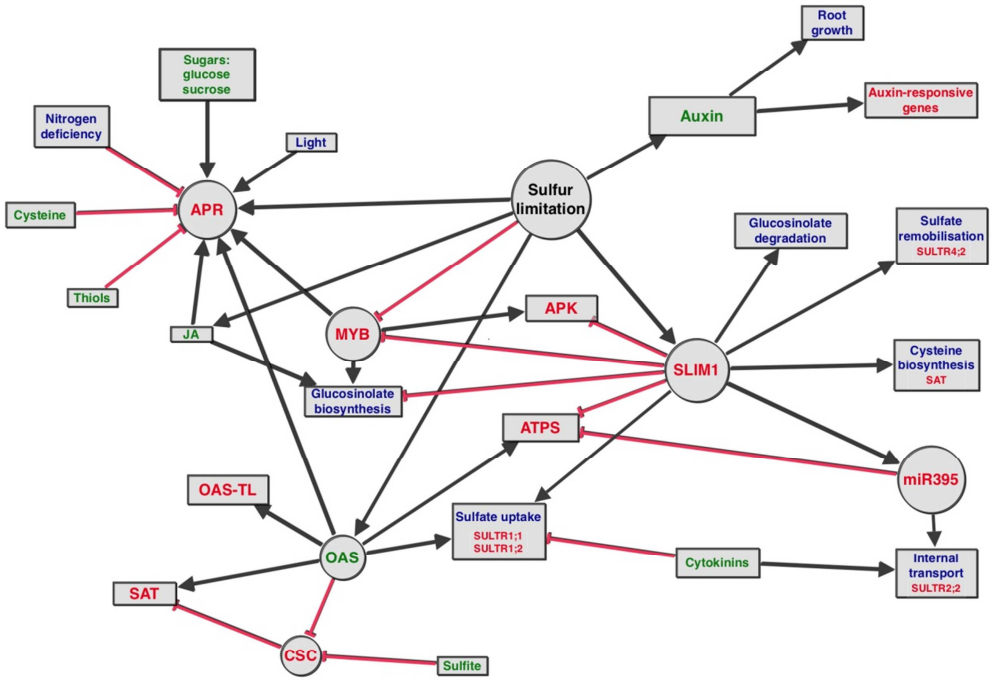


Figure 1.5 ***Complete network of regulation of sulfate metabolism in Arabidopsis thaliana***

Black arrows indicate positive regulation (induction or accumulation) and red arrows indicate negative regulation (repression). The grey circles correspond to the key regulatory factors of the pathway; grey boxes correspond to all the other components involved in regulation. Blue colour indicates processes, red – genes and proteins, green – metabolites. Abbreviations: APR, APS reductase, MYB, MYB transcription factors; APK, APS kinase; SLIM1; Sulfur Limiting factor 1; SULTRs, sulfate transporters; SAT; serine acetyltransferase; ATPS, ATP sulfurylase; OAS-TL, OAS(thiol)lyase; OAS, O-acetylserine; CSC, cysteine synthase complex. The figure is based on the information described in section 1.5.

1.5.1 Regulation on the level of sulfate uptake and transport

The efficient acquisition of sulfate from the soil and its distribution in the plant is of great importance especially under sulfur limiting conditions (Takahashi et al. 2011). In a number of molecular biological studies it was shown that the rate of sulfate transport during

low sulfate supply is driven mainly by regulation of the two high-affinity sulfate transporter genes, *SULTR1;1* and *SULTR1;2* (Shibagaki et al. 2002, Takahashi et al. 2000, Vidmar et al. 2000, Yoshimoto et al. 2002). The study of promoter-reporter constructs indicated that both of them are regulated in response to sulfate nutrition (Maruyama-Nakashita et al. 2004a; see below). Further studies led to the identification of SLIM1 (sulfur limitation 1) transcription factor which is responsible for regulation of sulfate uptake and metabolism during insufficient sulfur supply (Maruyama-Nakashita et al. 2006; Figure 1.5). The SLIM1 transcription factor belongs to the family of ethylene insensitive-like (EIL) transcription factors. EIL3 is the only member of the family which has a specific function in regulation of sulfate uptake and metabolism (Maruyama-Nakashita et al. 2006). Further analysis revealed that EIL3 is able to restore the wild type phenotype of *slim1* mutants suggesting that EIL3 is in fact SLIM1 (Maruyama-Nakashita et al. 2006). S*lim1* mutants showed about 30% reduction in root lengths and 60% decrease of sulfate uptake rates in sulfur-limiting conditions. SLIM1 regulates the expression of the majority of sulfate-limitation responsive genes in the pathway which suggests a hub-like function in the regulation system. Interestingly, APR which is the key enzyme of sulfur metabolism seems not to be subjected to the control of SLIM1 (Maruyama-Nakashita et al. 2006).

Sulfur-responsive cis-acting element (SURE) is another factor involved in both negative and positive regulation of the sulfate reduction pathway (Maruyama-Nakashita et al. 2005). It was first identified in the sequence of the promoter of the *SULTR1;1* sulfate transporter gene. SURE is a 7 nucleotide long specific sequence localised in the 5'-region of *SULTR1;1*. This sequence contains the core sequence of the auxin response factor (ARF; Hagen & Guilfoyle 2002). However, SURE is sulfur specific and is not involved in auxin signalling. The analysis of sulfur limitation inducible expression of *SULTR1;1* revealed that the SURE element was an essential target for sulfur-limitation response in arabidopsis roots (Maruyama-Nakashita et al. 2005). Interestingly, SURE was identified in *SULTR1;1* but not in *SULTR1;2* suggesting different mechanisms of regulation of these two transporters (Maruyama-Nakashita et al. 2005). Indeed, *SULTR1;1* is known to be controlled more specifically by sulfur limitation whereas the control of *SULTR1;2* is more dependent on metabolic demand (Rouached et al. 2008). The *in silico* promoter analysis with GeneChip microarrays of 15 genes which are known to be up-regulated by sulfur limitation revealed the presence of a SURE core sequence (GAGAC or GTCTC) in the promoters of all of these genes. Similar sequences are present in the NIT3 nitrilase (Kutz et al. 2002) and β-conglycinin β-subunit (Awazuhara et al. 2002). Therefore it was concluded that the SURE core sequences are conserved in promoters activated by sulfate

limitation and may play a key role in the induction of the sulfur starvation response. However, a number of genes regulated by sulfur starvation do not have a SURE element (Maruyama-Nakashita et al. 2005). Despite the important role of SURE elements in the sulfur-limitation inducible response there are still many gaps in our knowledge. Many important questions, such as specific SURE-binding candidates, still require further investigation.

Sulfur metabolism is also controlled post-transcriptionally (Takahashi et al. 2011). The main player in post-transcriptional regulation is microRNA395 (miR395; Figure 1.5). MiRNAs are a group of small RNAs which are formed from noncoding double-stranded RNA precursors. They are able to negatively regulate their target genes by cleavage or by binding to complementary messenger RNA sequence of their targets repressing their translation (Bartel 2004). Analysis of the arabidopsis genome revealed the low-affinity sulfate transporter SULTR2;1 and ATPS1 and ATPS4 isoforms of ATPS as target genes for miR395 (Jones-Rhoades & Bartel 2004). The accumulation of miR395 increases during sulfate starvation and this process is dependent on SLIM1 (Kawashima et al. 2011). Further studies revealed that miR395 regulates sulfate accumulation in shoots by cleavage of ATPS1 and ATPS4 mRNA during sulfate starvation (Kawashima et al. 2009). It has been shown that in arabidopsis lines overexpressing miR395, the mRNA abundance of *ATPS4* strongly decreased under sulfate deficiency suggesting a canonical regulation of *ATPS4* by miR395 (Kawashima et al. 2011).

MiR395 was also shown to play a role in sulfate translocation between roots and shoots by targeting SULTR2;1 (Liang et al. 2010). It has been shown that the mRNA abundance of *SULTR2;1* in roots increase during sulfate starvation and its expression is limited to the xylem parenchyma cells (Kawashima et al. 2011, Kawashima et al. 2009). The increase in *SULTR2;1* expression in the xylem parenchyma and reduction of flux through sulfate assimilation in the roots caused by miR395 indicate the importance of the SLIM1 dependent induction of miR395 for the increased translocation of sulfate to the shoots when sulfate is limited. More efficient translocation of sulfate between roots and shoots which is a result of increased expression of SULTR2;1 might improve the efficiency of sulfate assimilation in leaves (Kawashima et al. 2011). Very recently Matthewman et al. (2012) have shown that the complex regulation of miR395 is linked not only to SLIM1-dependent regulation during sulfate starvation, but also to glutathione and more generally thiol levels and/or cell redox state. Increased expression of miR395 in mutants affected in sulfate accumulation, *fou8* and *sultr1;2* suggests that miRNA395 is regulated by internal sulfate level irrespective of external sulfate availability (Lee et al. 2012). All these results confirm that the miR395 is an integral component of the sulfate assimilation regulatory network and complex regulation mechanism.

1.5.2 Regulation of sulfate assimilation pathway

Sulfate assimilation is tightly regulated in response to sulfate demands and environmental changes (Davidian & Kopriva 2010). The control mechanisms are on different steps of the pathway and involve regulation of specific enzymes. Additionally, the whole pathway may be controlled as a process. Experiments focused on the first step of the pathway revealed that the expression of ATPS is regulated by sulfate availability (Logan et al. 1996). It was also shown that ATPS activity is inhibited by glutathione. However, a number of studies revealed that the key regulatory step is sulfate reduction by APR (Figure 1.5) as it is strongly affected by various treatments and environmental changes (Hesse et al. 2003, Koprivova et al. 2008, Vauclare et al. 2002). The application of GSH to the root culture medium decreases APR expression and activity indicating that GSH rather than cysteine (which is the final product of the pathway) is involved in the regulation of sulfate assimilation pathway (Vauclare et al. 2002). They also estimated the flux control coefficient for APR as 0.57 considering that the coefficient of all enzymes in the sulfate uptake and reduction pathway adds up to 1 which confirms the importance of APR in control of flux through sulfate assimilation pathway. The quantitative trait loci (QTL) analysis of a recombinant inbred line (RIL) population derived from the cross between two wild type arabidopsis accessions – Bay-0 and Sha – led to the identification of a single nucleotide polymorphism in the gene encoding the isoform 2 of APR. The substitution of alanine with glutamate in a conserved domain of the APR2 protein resulted in significant differences in enzyme activity leading to sulfate accumulation (Loudet et al. 2007). Additionally, reduction of APR activity and mRNA accumulation under low nitrogen availability shown in this study confirmed an interconnection of the two assimilatory pathways (Loudet et al. 2007).

Overexpression of *APR2* leads to accumulation of sulfite and thiosulfate which are toxic for plants and strongly affect plant health (Martin et al. 2005). Such an effect is not observed when overexpressing other enzymes of sulfate reduction pathway suggesting that APR regulation affects the entire pathway. APR activity increases during the day and decreases during the night which shows that it has a diurnal rhythm (Kopriva et al. 1999). A recently discovered transcription factor, Long Hypocotyl 5 (HY5) seems to be responsible for APR regulation by light (Lee et al. 2011). In the dark adapted arabidopsis seedlings a rapid increase in the transcript abundance of all three APR isoforms was observed to different extents. Transcript accumulation of the *APR2* isoform was 12-fold higher after 90 minutes of illumination than in control plants kept in the dark. However, in *hy5* mutant seedlings no light induction was observed for *APR1*, and *APR3* induction was lower. Further analysis also

revealed that HY5 is involved in APR regulation by OAS and nitrogen deficiency which alter the demand for reduced sulfur (Lee et al. 2011). Interestingly, the analysis of the sulfite reductase mutant *sir1-1* revealed the downregulation of genes encoding ATPS4, APR2, and SULTR2;1 relative to the wild type plants (Khan et al. 2010). The most likely reason for the downregulation of these genes, especially *ATPS4* and *APR2*, is to avoid the accumulation of toxic sulfite which cannot be incorporated into cysteine as a result of reduced SiR activity in the mutant. These results suggest that SiR can contribute to the control of sulfate reduction pathway (Kutz et al. 2002).

1.5.3 Regulation of cysteine synthesis – protein-protein interactions

Cysteine synthesis plays an important role in the regulation of sulfate metabolism (Hell & Writz 2011). The regulation of SAT and OAS-TL, the main enzymes in cysteine biosynthesis, is mainly due to a protein-protein interaction in the cysteine synthase complex (CSC; Figure 1.5). As mentioned before, SAT is strongly activated by OAS-TL which, however, is inactive in the complex and has only a regulatory role (Droux et al. 1998). Formation of the complex is strongly dependent on the availability of OAS and sulfide. When SAT is bound to OAS-TL, the access of OAS to the complex is strongly inhibited (Francois et al. 2006). OAS which cannot be bound to the complex is released and metabolised by free OAS-TL dimers. During sulfur deficiency there is not enough sulfide for cysteine synthesis and therefore OAS accumulates in cells. Accumulated OAS dissociates from the CSC complex, which rapidly decreases SAT activity (Hell & Wirtz 2008). High OAS concentration increases the expression of genes encoding sulfate transporters, APR, SAT, and OAS-TL. This leads to increased sulfate uptake and reduction and equilibrates the system (Hopkins et al. 2005, Koprivova et al. 2000, Smith et al. 1997). Hubberten et al. (2012b) have shown recently that OAS may serve as a signalling molecule and change the transcription levels of specific genes irrespective of the sulfur status in the plant.

The feedback inhibition of the cytosolic isoform of SAT by cysteine concentration serves as another form of regulation in arabidopsis. It is important to notice that in arabidopsis plastidial and mitochondrial isoforms of SAT remain mostly insensitive to changes in cysteine concentration (Noji et al. 1998). Cytosolic SAT activity may be considered as important for control of OAS concentration. The cysteine insensitive SAT isoforms in organelles may allow independent formation of cysteine (Noji et al. 1998). Although the regulation of cysteine biosynthesis is very important for sulfur homeostasis in plants and therefore it has been

intensively studied over the past few years, many important aspects such as the role of CSC in the regulation of OAS inducible genes still require further investigation.

1.5.4 Control of methionine biosynthesis – posttranscriptional regulation

Methionine biosynthesis undergoes a complex regulation (Ravanel et al. 2004). The whole pathway is controlled by feedback inhibition of aspartate kinase by lysine, threonine or lysine together with SAM. It seems that the competition between CGS and threonine synthase (TS) for their common substrate, O-phosphohomoserine, has an important regulatory role in the flux of carbon into methionine (Hesse et al. 2004a). It has been shown that the K_m values of activated TS are 250-500 fold lower than those of CGS (Ravanel et al. 2004). Therefore, when methionine levels are high the carbon flux is directed towards threonine synthesis. A recent study of arabidopsis *mto1* mutants (with disabled CGS at the MTO1 region which results in the overaccumulation of methionine – see below) provided evidence that the regulation of methionine biosynthesis also occurs at the posttranscriptional level (Chiba et al. 1999, Chiba et al. 2003). The analysis of this process is focused on the MTO1 region in exon 1 of CGS (Suzuki et al. 2001). The MTO1 mRNA region may act in *cis* and destabilize CGS mRNA in response to high concentrations of methionine or SAM (Chiba et al. 1999, Chiba et al. 2003, Lambein et al. 2003). Computational analysis revealed the possible formation of a stable stem-loop structure in the MTO1 region which could support the posttranscriptional mechanism of regulation (Amir et al. 2002). More recently the presence of a truncated form of CGS transcript in arabidopsis was shown (Hacham et al. 2006). This transcript lacks about 90 nucleotides from the first exon. Overexpression of this CGS transcript causes an even higher level of methionine than the overexpression of full-length CGS. This may suggest that this truncated transcript is not subject to feedback regulation by methionine (Hacham et al. 2006).

1.5.5 Control of glucosinolate biosynthesis – MYB transcription factors

Glucosinolate biosynthesis is highly regulated by a network of transcription factors in response to biotic and abiotic stress (Figure 1.5; Gigolashvili et al. 2007, Gigolashvili et al. 2008, Sønderby et al. 2007). They belong to two groups of the R2R3-MYB family of transcription factors. The first group consists of three members: MYB28, MYB76 and MYB29, and is involved in the control of aliphatic glucosinolate biosynthesis (Gigolashvili et al. 2008). The second group consists of MYB34, MYB51, and MYB122 and is involved in the control of the biosynthesis of indolic glucosinolates (Celenza et al. 2005, Malitsky et al. 2008). T-DNA insertions, RNAi or overexpression of these transcription factors affects the expression of

genes encoding glucosinolate biosynthesis enzymes and, consequently the level of glucosinolates (Gigolashvili et al. 2007, Gigolashvili et al. 2008).

Studies on MYB28, MYB29 and MYB76 showed that they are able to transactivate each other in control of biosynthesis of aliphatic glucosinolates (Hirai et al. 2007, Gigolashvili et al. 2008). Analysis of MYB28 and MYB29 revealed that MYB28 is essential for the synthesis of aliphatic glucosinolates whereas MYB29 induces biosynthetic genes in response to plant hormone methyl jasmonate (Hirai et al. 2007). They were also shown to downregulate the expression of genes encoding enzymes involved in the synthesis of indolic glucosinolates (Gigolashvili et al. 2008). MYB34, MYB51 and MYB122 have different functions in regulation of the pathway (Gigolashvili et al. 2007). Although all of them can up-regulate the genes from the biosynthesis pathway, MYB34 and MYB122 also function as stimulators of auxin biosynthesis whereas MYB51 additionally activates the glucosinolate biosynthesis pathway (Gigolashvili et al. 2007). Maruyama-Nakashita et al. (2006) showed that mutation of SLIM1 additionally affects the expression of MYB34, suggesting that this factor may be negatively controlled by SLIM1 in response to sulfur deficiency. More recently it was indicated that MYB factors are involved not only in the regulation of the glucosinolate biosynthesis genes, but also other genes in the sulfur reduction pathway. Yatusevitch et al. (2010) concluded that *APR1* and *APR3* are under control of all MYB factors whereas *APR2* is regulated only by some of them. They also suggested that *APK1* and *APK2* are a part of the glucosinolate synthesis network controlled by MYB factors, whereas *APK3* and *APK4* have much lower input to the network.

1.5.6 Regulation by hormone signals

Metabolic regulation is not the only way to control metabolic pathways. Plant hormones play very important roles in many developmental processes. Recent studies show involvement of plant hormones in regulation of different nutrient metabolic pathways (Figure 1.5).

Cytokinins are adenine derived plant hormones which are responsible for regulation of cell division and differentiation in plants together with auxin (Sakakibara 2006). The best known example of regulation of the sulfate reduction pathway by cytokinin is the downregulation of genes encoding high-affinity transporters SULTR1;1 and SULTR1;2 (Maruyama-Nakashita et al. 2004b). Addition of cytokinins to wild type plants decreases sulfate uptake and mRNA abundance for the two transporters. It is interesting that SULTR1;2 is much more responsive to the cytokinins than SULTR1;1. It was suggested that in this regulation the cytokinin response 1 (CRE1)/wooden leg (WOL)/arabidopsis histidine kinase 4

(AHK4) cytokinin receptor is involved. The arabidopsis *cre1-1* mutant was unable to regulate the high-affinity sulfate transporters in response to cytokinins (Maruyama-Nakashita et al. 2004b). These results suggest the independent regulation of high-affinity sulfate transporters by cytokinin (repression) and by sulfate (induction). In contrast, Ohkama et al. (2002) have shown that sulfur responsive genes *APR1* and *SULTR2;2* were upregulated by cytokinin. The authors have concluded that this regulation depends on an increase in sucrose concentration in plant tissues.

Auxin-dependent signalling in the regulation of sulfate assimilation may be connected to the response to sulfate deficiency by indole glucosinolate hydrolysis (Kutz et al. 2002). During sulfate deficiency the aglycone is released from indole glucosinolates. Indole acetic acid (IAA) is generated from the remaining indole acetonitrile, catalysed by nitrilase *NIT3*. IAA stimulates root growth (Hell & Hillebrand 2001). Indeed, increase in length and numbers of lateral roots is a common phenotype of sulfur deficient plants. Additionally, various studies indicate positive regulation of auxin-responsive genes during sulfate starvation (Maruyama-Nakashita et al. 2003, Nikiforova et al. 2003). Decrease in cysteine production leads to accumulation of OAS and its precursor serine and sulfate deficiency induces tryptophan synthase. Both of these events result in increased tryptophan biosynthesis as in plants it is synthesized from indole and serine through the activity of the tryptophan synthase β-subunit (Nikiforova et al. 2003). In consequence increased biosynthesis of tryptophan increases production of auxin.

Jasmonic acid (JA) is another plant hormone which participates in regulation of sulfate metabolism (Xiang & Oliver 1998). JA is involved in response to oxidative stress and synthesis of defence molecules. The cellular GSH concentration rapidly decreases during sulfur deficiency and this may lead to oxidative stress in cells. It was shown however, that JA increases the expression of glutathione synthesis pathway enzymes (Xiang & Oliver 1998). Additionally microarray studies showed induced expression of genes encoding enzymes involved in jasmonate biosynthesis during sulfur limitation and in *sultr1;2* mutants (Maruyama-Nakashita et al. 2003, Nikiforova et al. 2003). JA was also shown to induce the expression of the genes involved in glucosinolate biosynthesis (Brader et al. 2001, Doughty et al. 1995). JA-deficient mutants showed normal responses to sulfur limitation, e.g. sulfate transporters and APR were induced and glucosinolate biosynthesis genes were repressed under sulfur starvation (Takahashi & Saito 2008). JA significantly induced gene expression and protein activity of APR, mainly isoforms 1 and 3 (Koprivova et al. 2008). In overview, JA promotes the synthesis of both antioxidants and glucosinolates and has important general role in plant

pathogen defence and detoxification. It also co-ordinately induces multiple genes of sulfate assimilation suggesting its positive effect on sulfur homeostasis in plants (Jost et al. 2005).

1.6 The aims of the project

Since sulfur deficiency in Europe appeared relatively recently, the understanding of sulfur use efficiency still lags behind that of the other major nutrients. Many research projects have been focused on investigation of sulfur assimilation and the regulation of sulfur metabolism in recent years. However, there are still many gaps in our knowledge especially in terms of regulation of sulfate assimilation and plant response to sulfate limitation. Moreover, there is increasing interest in investigation of the interconnection between different elements since it has been shown that the disruption of homeostasis of one nutrient may result in a decrease in use efficiency of other nutrient(s). Therefore, in the course of my PhD project I investigated different aspects of sulfate metabolism in *Arabidopsis thaliana* (L.)Heynh.

- First, I was interested in the investigation of natural variation in plant response to low sulfur supply. Through collaboration with two other research groups an experiment was conducted in which 25 arabidopsis accessions were analysed in their response to sulfate deficiency, nitrate deficiency, and sulfate/nitrate double deficiency. Results of this work are described in Chapter 4.

- I used natural variation among worldwide population of *Arabidopsis thaliana* to look for new genes and regulatory factors that contribute to homeostasis and regulation of nitrate, phosphate, and sulfate metabolism. Results of Genome-Wide Association Studies (GWAS) which delivered a number of candidate genes potentially involved in the control of these three metabolic pathways are described in Chapter 5.

- Together with members of the Kopriva group and different collaborators I used natural variation among *Arabidopsis thaliana* and various biochemical and molecular techniques to get a better understanding of regulatory functions of already known enzymes of the sulfur reduction pathway which is described in Chapter 3.

- My contribution to the functional characterisation of a previously undescribed gene delivered by GWAS on nitrate and sulfate accumulation data is described in Chapter 6.

Chapter 2:

Materials and Methods

2.1 Materials

This chapter describes general materials and methods that were used for the laboratory experiments. It includes the general methods of molecular biology (e.g. DNA extraction, PCR reaction, restriction digestion) as well as the principal methods for investigation of the traits related to sulfur metabolism. Other chapters include a Materials and Methods section with plant growth conditions, experimental design, and techniques specific for the analyses described in that chapter.

2.1.1 Bacterial strains

Table 2.1 ***Bacterial strains***

Strain	Antibiotic resistance	Properties	Source and Reference	Comment
		Escherichia coli		
DH5α		*F- φ80lacZΔM15 Δ (lacZYA-argF)U169 recA1 endA1 hsdR17(rk-, mk+) phoA supE44 thi-1 gyrA96 relA1*	Hanahan, 1983	Generaly used for plasmid propagation and construction
TOP10	Streptomycin	*F- mcrA (mrr-hsdRMS-mcrBC) 80lacZM15 lacX74 recA1 ara139 (ara-leu)7697 galU galK rpsL (StrR) endA1 nupG*	Invitrogen	Generaly used for plasmid propagation and construction
		Agrobacterium tumefaciens		
AGL-1	Rifampicin Carbenicillin	*pTiBo542DT-DNA, genes for succinamopine biosynthesis*	-	Infiltration of *N. benthamiana* and transformation of *A. thaliana*

2.1.2 Antibiotics

Table 2.2 ***Concentration and usage of antibiotics***

Antibiotic	Final concentration	Solvent	Targets
Ampicillin	100 µg ml^{-1}	H_2O	gram-negative bacteria
Carbenicillin	100 µg ml^{-1}	H_2O	bacteria, fungi, plants
Kanamycin	50 µg ml^{-1}	H_2O	bacteria, fungi, plants
Rifampicin	50 µg ml^{-1}	methanol	bacteria
Hygromycin B	25 µg ml^{-1}	H_2O	bacteria, fungi, plants
Spectinomycin	100 µg ml^{-1}	H_2O	bacteria

2.1.3 Plasmids

Table 2.3 ***Cloning vectors***

Vactor	Resistance	Description
pCR8/GW/TOPO TA	Spectinomycin	Vector for cloning of DNA fragments with terminal 3' A-overhangs (Invitrogen)
pGWB2	Kanamycin, Hygromycin	GATEWAY Binary vector for the expression of proteins in plants driven by the 35S promoter (Research Institute of Molecular Genetics, Shimane University, Japan)
pK7FWG2,0	Chloramphenicol Kanamycin	GATEWAY C-terminal GFP vector under the control of 35S promoter (Invitrogen)

2.1.4 Growth media

Table 2.4 ***Composition of growth media for plants and bacteria***

Medium	Composition
	Plant growth medium
Murashige and Skoog (MS)	½ MS plant salt mixture (Duchefa Biochemie, Ipswich, UK); 166 mg l^{-1} $CaCl_2$, 85 mg l^{-1} KH_2PO_4, 950 mg l^{-1} KNO_3, 90 mg l^{-1} $MgSO_4$, 825 mg l^{-1} NH_4NO_3, 0.0125 mg l^{-1} $CoCl_2.6H_2O$, 0.0125 mg l^{-1}, $CuSO_4.5H_2O$, 18 mg l^{-1} FeNaEDTA, 3 mg l^{-1} H_3BO_3, 0.4 mg l^{-1} KI, 8 mg l^{-1} $MnSO_4.H_2O$, 0.125 mg l^{-1} $Na_2MoO_4.2H_2O$, 4 mg l^{-1} $ZnSO_4$.7H2O, 50 mg l^{-1} myo-inositol, 0.5 mg l^{-1} thiamine, 0.25 mg l^{-1} pyroxidine, 0.25 mg l^{-1} nicotinic acid, 0.25 g l^{-1} 2-[N-morpholino]-ethanesulphonic acid (MES), 0.8% (w/v) agar (ForMedium AGA03), where stated 3% sucrose (w/v), pH 5.7 (with 1 M KOH)
	Bacteria growth medium
Luria-Bertani Broth (LB)	10 g l^{-1} tryptone, 5 g l^{-1} yeast extract, 10 g l^{-1} NaCl, pH 7.0, agar was added for solid medium (final concentration 1.5% (w/v); Formedium AGA03)
Super Optimal broth with Catabolite repression (SOC)	20 g l^{-1} tryptone, 5 g l^{-1} yeast extract, 0.5 g l^{-1} NaCl, 0.186 g l^{-1} KCl, pH 7.0. The medium was autoclaved before addition of glucose at a final concentration of 20 mM and $MgCl_2$ at a final concentration of 2 mM

2.1.5 Oligonucleotides

All oligonucleotides were obtained from Sigma-Aldrich Ltd (Haverhill, UK). The sequences were obtained by using T-DNA Express, Quant Prime software (for qPCR primers), or they were designed by me. All primers were re-suspended in dH_2O in the amount appropriate to obtain 100 µM stock solution according to manufacturer's instruction. The sequences of particular primers are listed in the Materials and Methods sections specific for each chapter.

2.2 Methods

2.2.1 Plant growth on plates

Dry seeds were surface sterilised for up to four hours in a vacuum desiccator using chlorine gas which was generated by mixing 125 ml of sodium hypochlorite with 2.5 ml of 12 M HCl. Seeds were then mixed with 0.1% sterile agarose (Sigma Aldrich www.sigmaaldrich.com) and seeded along the edge of square plastic plates (10 cm x 10 cm; R & L Slaughter, Essex, UK) with appropriate medium (depending on the experiment – see Materials and Method section of particular chapters) using pipette. Subsequently, plates were stored in 4°C in the dark for three days. After that time they were transferred to a CER at 22°C under 16-h-light/8-h-dark cycles where they were grown vertically for up to three weeks with the photon flux density 140 µmol m^{-2} $s^{-}1$. The plants were grown on the surface of the agar. Exact growth conditions for particular experiments including the media composition, growth period, and experimental design are described in the Material and Methods section of following chapters.

2.2.2 Plant growth in the soil

Dry seeds were surface sterilised for up to four hours in a vacuum desiccator using chlorine gas which was generated by mixing 125 ml of sodium hypochlorite with 2.5 ml of 12 M HCl. Seeds were then mixed with 0.1% sterile agarose (Sigma Aldrich) and seeded on square plates (10 cm x 10 cm) with MS medium divided for nine little squares. Subsequently, plates were stored in 4°C in the dark for three days. After that time they were transferred to a CER at 22°C under 16-h-light/8-h-dark cycles where they were left to germinate – up to one week (horizontally). One week old seedlings were transferred from plates to the 40-cell tray (21x35cm) with Levington Horticulture soil mix and grown for another four weeks in CER 10-h-light/14-h-dark cycles at constant temperature 22°C, 60% relative humidity, and light intensity of 160 µE s^{-1} m^{-2}. Exact growth conditions for particular experiments including the media composition, growth period, and experimental design are described in the Material and Methods section of following chapters.

2.2.3 Homogenization of plant material

Unless stated differently plants were usually frozen immediately after harvesting in liquid nitrogen and stored at -80°C for further analysis as fresh matter. Before analysis plant material was either homogenised in the extraction buffer by using disposable pestles (for example for RNA isolation) or ground into a fine powder in 2 ml microcentrifuge tubes

containing 4 mm diameter stainless steel ball (Bearing Supplies; www.bearing-supplies.co.uk) by using a Genogrinder (SPEX Sample Prep, Metuchen, USA). For the dry matter analyses plants were freeze dried and homogenized to fine powder using stainless steel balls.

2.2.4 DNA isolation from plant samples

Unless stated differently, DNA was isolated from one leaf using Edwards buffer (200 mM Tris HCl, pH 7.5, 250 mM NaCl, 25 mM ethylenediaminetetraacetic acid EDTA, 0.5% (w/w) SDS). For large numbers of extractions, DNA isolation was carried out in 96-well racks containing strips of 8-well collection tubes (Qiagen, http://www.qiagen.com/). Plant tissue was harvested into the tubes which contained a 4 mm diameter stainless steel ball. The racks were frozen at -80°C. The tissue was then homogenised by shaking the racks in a Genogrinder at 200 strokes per minute for 30 s. Homogenised tissue was resuspended in 300 µl of Edwards buffer. Cell debris was pelleted by centrifugation at 4000 g for 10 min at room temperature. The supernatant (250 µl) was transferred into new tubes containing 250 µl of isopropanol (Sigma Aldrich; http://www.sigmaaldrich.com/united-kingdom.html). The tubes were inverted several times, incubated at room temperature for 2 min and centrifuged at 4000 g for 10 min at room temperature. The DNA pellet was washed in 300 µl 70% (v/v) ethanol and re-pelleted using the same centrifugation conditions. The tubed with pellet were left under the laminar flow hood for 20 min to let the pellet dry. Subsequently, the pellet was resuspended in 50 µl dH_2O. For the extraction of a smaller number of samples, DNA was extracted in Eppendorf microcentrifuge tubes (http://www.eppendorf.com/UK-en/). The same volume of Edwards buffer was used (300 µl), but centrifugation was carried out at 14,000 g in a microcentrifuge 10 min as stated above.

2.2.5 Isolation of plasmid DNA

The *Escherichia coli* plasmid DNA was isolated from up to 10 ml of liquid overnight bacterial culture grown in LB medium (see section 2.1.4) at 37°C in a shaking incubator. The plasmid DNA was isolated using the QIAprep Spin Miniprep Kit (Qiagen) according to the manufacturer's instruction. The DNA was eluted from the spin column using 30 µl of dH_2O and its quality and quantity were assessed using the NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, http://www.nanodrop.com/).

2.2.6 Polymerase chain reaction

For colony PCR and genotyping the GoTaq Flexi Kit (Promega; Southampton, UK) was used for PCR. For cloning, Platinum™ High-Fidelity DNA polymerase (Invitrogen; Paisley, UK)

was used. PCR reactions contained 50-500 ng of gDNA or 1 pg - 50 ng plasmid DNA as template, 1x polymerase reaction buffer, 0.1 U of GoTaq polymerase (or 0.05 U of Platinum™ High-Fidelity DNA polymerase), 0.4 µM of each primer and 500 µM dNTPs. PCR amplifications were carried out in a DYAD Thermal Cycler (Biorad, Hertfordshire, UK) or Mastercycler Pro (Eppendorf). The standard PCR amplification protocol with GoTaq Polymerase included initial denaturation step of 95°C for 2 min, followed by 35 cycles of: 94°C for 30 s, 50-60°C (Primer Tm-3°C) for 30 s, 72°C for 1 min for each 1kb of target DNA product and a final elongation step at 72°C for 10 min. The reaction was carried out in 10 µl total volume. For reactions containing Platinum™ High-Fidelity DNA polymerase, the initial denaturation step was carried out for 1 min at 94°C, followed by 25-35 cycles of 94°C for 30 s, 50-60°C (Primer Tm-3°C) for 20 s, 68°C for 1 min for each 1kb of target DNA product and 68°C for 10 min. The reaction was carried out in 50 µl total volume. For colony PCR, GoTaq polymerase was used and instead of DNA template, parts of a bacterial colony were added with a pipette tip to the PCR mix. The PCR products were separated by agarose gel electrophoresis using 1% gel.

2.2.7 Restriction digestion

The restriction digestion for confirmation of successful candidates was carried out overnight according to the enzyme manufacturer's instructions (Roche Diagnostics, West Sussex, UK; New England Biolabs, Herts, UK). Usually, the reactions were carried out in 10 µl final volume using 500 ng of DNA and 1 U of enzyme, the restriction buffer and bovine serum albumin (BSA) if required according to manufacturer's instruction. The restriction enzymes used most often are listed in Table 2.5.

Table 2.5 ***Restriction enzymes***

Enzyme	Concentration	Source
EcoRI	5 U/µl	Roche Diagnostics
BspHI	10 U/µl	New England Biolabs
ClaI	10 U/µl	Roche Diagnostics
BglI	10 U/µl	New England Biolabs

2.2.8 Agarose gel electrophoresis

PCR DNA fragments or the DNA after restriction digestion were mixed with 5 x loading dye (50% (v/v) glycerol, 0.05% (w/v) Orange G) and separated on 1% (w/v) agarose gels, prepared using TAE buffer (40 mM Tris-HCl, 20 mM acetic acid, 1 mM EDTA, pH 8.0) containing 0.01% (v/v) ethidium bromide. The DNA separated on the gel was visualised through

fluorescence of the ethidium bromide-DNA complex when exposed to ultraviolet (UV) light from a transilluminator and photographed using a Gel Doc 1000 system (Bio-Rad).

2.2.9 LR reaction

Recombination between sequences within attL and attR sites was catalysed by LR clonase enyzme (Invitrogen). The reaction was performed according to the manufacturer's instruction. For the reaction, 4 µl of the pCR8 plasmid containing the fragment of interest were mixed with the same amount of destination vector and 2 µl LR clonase. Samples were incubated at room temperature overnight. Subsequently, the reaction was terminated by addition of 0.5 µl proteinase K solution (Invitrogen) and incubation at 37°C for 10 min. Subsequently, *E.coli* TOP10 or DH5α cells were transformed with 3 µl of the reaction mix (see section 2.2.8).

2.2.10 Transformation of Escherichia coli

The *Escherichia coli* cells were transformed by heat shock. Chemically competent cells (30 µl), stored at -80°C were thawed on ice and 3 µl of the LR reaction mix (see section 2.2.7) were added and the mix was incubated on ice for 30 min. For the heat shock, the cells were heated to 42°C for 30 s and immediately transferred to ice. After 2 min, 1 ml of SOC medium (see section 2.1.4) was added and bacteria grown for one hour at 37°C with gentle shaking. Subsequently, they were plated on LB agar containing the appropriate antibiotics. Plates were incubated upside down at 37°C overnight. The single colonies were screened by colony PCR to identify successful transformants or they were used to set up the liquid culture for isolation of plasmid DNA.

2.2.11 Transformation of Agrobacterium tumefaciens

The *Agrobacterium tumefaciens* was transformed via electroporation or heat shock. The MicroPulser Electroporator (Biorad) was set at 125V. The liquid suspension of electro-competent Agrobacterium cells (50 µl) was transferred to sterile electroporation cuvette with 1 mm electrode gap width, dried and placed in the machine. The voltage was passed through for 5 s. Immediately after that 1 ml of LB medium was added to the cuvette and entire mix was transferred to a screw cap tube. Bacteria were grown in 28°C for 3 h. Subsequently, they were spread on plates with LB medium containing appropriate antibiotics. Successful transformants were used for transformation of arabidopsis or infiltration of *Nicotiana benthamiana.*

Alternatively, 5 µl of plasmid DNA isolated from *Escherichia coli* was added to the frozen chemically competent agrobacterium cells. The bacteria were incubated in a water bath

in 37°C for 5 min. After that time 1 ml of LB medium was added and the bacteria were grown at 28°C for 3 h. Subsequently, they were spread on plates with LB medium containing appropriate antibiotics. Successful transformants were used for transformation of arabidopsis or infiltration of *Nicotiana benthamiana*.

2.2.12 Transformation of Arabidopsis thaliana

Plants were grown in a controlled environment room (CER) until flowering and production of secondary inflorescences. *Agrobacterium tumefaciens* cells transformed with the desired plasmid (see section 2.2.9) were grown overnight at 28°C in 500 ml LB medium containing the appropriate antibiotics. Cells were pelleted by centrifugation at room temperature at 5,000 g for 15 min and then re-suspended in 500 ml infiltration medium (5% (w/v) sucrose, 0.05% Silwet® L-77 (http://www.siliconeforbuilding.com/; GE silicones), 3 mM MES, 0.1 mM acetosyringone, pH 5.5). Plant inflorescences were dipped into an agrobacterium suspension in a beaker for 2 min, then placed into clear plastic bags and shaded for 24 h. Subsequently, they were removed from the bags and grown in the glasshouse under ambient light conditions with supplemental lightning as required to obtain a 16 h photoperiod. Transformants were selected by growing seeds on plates supplemented with hygromycin (see section 2.1.8).

2.2.13 Transient transformation of Nicotiana benthamiana epidermal cells

Nicotiana benthamiana plants were grown in the glasshouse under ambient light conditions with supplemental lightning as required to obtain a 16 h photoperiod. *Agrobacterium tumefaciens* cells transformed with the desired plasmid (see section 2.2.9) were grown overnight at 28°C in 50 ml LB medium containing the appropriate antibiotics. Subsequently, the cells were pelleted by centrifugation at room temperature at 5,000 g for 15 min and then re-suspended in 5 ml of 10 mM $MgCl_2$ to give an OD_{600} of 0.8. The final dilutions were kept in the dark for 30 min at room temperature. After that, a *Nicotiana benthamiana* leaf was pricked with a needle and the bacterial suspension was infiltrated into the leaf using a syringe. The infiltrated area of the leaf was marked and the plants were grown for three days in the glass house. Subsequently, discs from the infiltrated area of the leaf were looked at under the confocal microscope Leica SP5 (II).

2.2.14 HPLC analysis of concentration of nitrate, phosphate and sulfate

Around 25 mg of washed and ground polyvinylpolypirrolidone (PVPP; Sigma Aldrich) was soaked with 1 ml of sterile water overnight at 4°C to remove the phenols from the plant

samples. Subsequently, 30-50 mg of frozen plant tissue or up to 10 mg of freeze dried tissue (for anion measurements described in Chapter 4) was homogenised fine powder by using the Genogrinder as described in section 2.2.3. Water incubated with PVPP (200 µl) was added to the plant sample and shaken vigorously to suspend the tissue powder. The whole extract was then returned to the tube with PVPP and shaken for approximately 1 h at 4°C. The extracts were then incubated at 95°C for 15 min, then centrifuged for 15 min at 4°C at 14,000 g. Anions were separated by High-Performance Liquid Chromatography (HPLC; Waters 2695 separation module and conductivity detector Waters 432) using the IC-Pak Anion HO 4.6 x 75 mm column with the suitable guard column and isocratic method with 0.8ml/min flow rate for 16 – 20 min. Solvent was composed of 120 ml of acetonitrile, 20 ml 50 x lithium gluconate/borate and 860 ml of dH_2O. Each set of samples (maximum 117) included 0.5 mM, 1 mM and 2 mM standards for nitrate, phosphate and sulfate.

Alternatively, for some measurements described in Chapter 4 the 0.1 M HCl extracts used for the low molecular weight thiols analysis (section 2.2.13) were diluted 1:50 (v/v) using dH_2O to avoid overloading the column with chloride ions from HCl, centrifuged 15 min at 4°C at 14,000 g and transferred to HPLC vials. The anions were subsequently separated and detected as described above.

2.2.15 HPLC analysis of low molecular weight thiols

Low molecular weight thiols, cysteine and glutathione, were analysed as previously described (Koprivova et al. 2008). Frozen or freeze dried tissue (for thiol measurements described in Chapter 4) was homogenized as described above and extracted with up to 50-fold volume of 0.1 M HCl (depending on the sample weight). To remove cell debris, the extract was centrifuged at 14,000 g for 10 min at 4°C. Subsequently, 25 µl of the supernatant was neutralised by 25 µl of 0.1 M NaOH. To reduce disulfides, the neutralised extract was incubated with 1 µl of 100 mM dithiothreitol (DTT) at 37°C for 15 min in the dark. Subsequently, 35 µl water, 10 µl 1 M Tris/HCl pH 8.0, and 5 µl of 100 mM monobromobimane (MBB; Thiolyte® MB, Calbiochem) were added and the samples were incubated in the dark at 37°C for 15 min to allow the derivatization of thiols. The reaction was stopped and the conjugates stabilized by the addition of 100 µl of 9% acetic acid. Bimane conjugates were separated by HPLC using the fluorescence detector (Waters 474) with excitation at a wavelength of 390 nm and emission at 480 nm, reverse-phase column (Spherisorb™ ODS2, 250 x 4.6 mm, 5 µm) and following solvents:

- 10% (v/v) methanol, 0.25% (v/v) acetic acid (pH 3.9) as solvent A and 90% (v/v) methanol, 0.25% (v/v) acetic acid (pH 9.3) as solvent B or,
- 100% (v/v) acetonitrile as solvent B and 0.25% (v/v) acetic acid (pH 9.3), 0.05% (v/v) 3M KOH as solvent C.

The elution protocol employed a linear gradient of the two solvents from 96% to 82% within 20 minutes with a constant flow rate of 1ml/min. Each set of samples (maximum 117) involved 0.025 mM, 0.0625 mM,, and 0.125 mM standards for GSH and cysteine.

2.2.16 The activity assay of adenosine 5'-phosphosulfate reductase (APR)

Adenosine 5'-phosphosulfate (APS) reductase (APR) activity was measured as the formation of [^{35}S]sulfite from [^{35}S]APS and dithiothreitol (DTT; Brunold & Suter 1991) with a modification of procedure described by Kopriva et al. (1999). The plant material was homogenized as described above (section 2.2.1) using disposable pestles and extraction buffer containing 50 mM Na/KPO_4 pH 8.0, 30 mM Na_2SO_3, 500 µM 5'-monophosphate (AMP), and 10 mM DTT. To remove the cell debris, the extracts were centrifuged 1 min at 8,000 g. Subsequently, 10 µl of extract was added to a 1.5 ml tube without a lid containing 240 µl of reaction assay mix (Table 2.6), mixed, and incubated in a water bath at 37°C for 30 min. After that time, 100 µl of 1 M Na_2SO_3 was added and the tubes were mixed and transferred to 20 ml scintillation vials filled with 1 ml of 1 M triethanolamine (TEA) solution. Subsequently, 200 µl of 1 M H_2SO_4 was added to the tubes. The scintillation vials were closed immediately and incubated over night at room temperature. The acidification of the reaction mix causes the formation of [^{35}S]SO_2 which is trapped in the TEA solution. Next day the bottoms of the tubes were washed from outside with 200 µl of H_2O and they were removed from the scintillation vials. The radioactivity was measured in each vial in the scintillation counter (Perkin Elmer Tri-Carb 2910 TR) after addition of 2.5 ml of Optiphase HiSafe3 scintillation cocktail (Perkin Elmer, Bucks, UK).

Table 2.6 ***APR activity reaction assay***

Concentration	Component	Volume/240 µl
1 M	Tris/HCl pH 9.0	25 µl
2 M	$MgSO_4$	100 µl
0.2 M	DTT	10 µl
3.75 mM	[^{35}S]APS with specific activity 1 kBq/10µl	5 µl
-	H_2O	100 µl for shoots 85 µl for roots
-	plant extract	10 µl for shoots 25 µl for roots

The concentration of protein in the extract was determined by using the Bio-Rad protein assay kit (Bio-Rad Laboratories, München, Germany) based on the method described by Bradford (1976). Each sample was composed of 5 µl of plant extract, 200 µl of Bio-Rad protein assay, and 795 µl of dH_2O (1 ml total volume). The mixture was incubated for 15 min at room temperature. The protein concentration was measured as absorbance at 595 nm using a UV-VIS spectrophotometer (Lambda Bio, Bucks, UK). Bovine serum albumin was used as a protein standard.

The APR activity was calculated as nmol min^{-1} mg^{-1} of protein according to the formula:

$$\mathrm{APR_{activity}} = \frac{37.5 \times \mathrm{cpm}}{\mathrm{cpm_{APS}} \times \mathrm{C_{prot}} \times \mathrm{V_E} \times \mathrm{t}}$$

where:

cpm_{APS} – specific activity (counts per minute) of APS

C_{prot} – protein concentration in the extract (mg ml^{-1})

V_E – volume of extract in the assay (ml)

t – incubation time (min)

2.2.17 The activity assay of ATP sulfurylase (ATPS)

The activity of ATPS was measured in reverse reaction as the formation of ATP dependent on APS and pyrophosphate as described previously (Cumming et al. 2007). One leaf was homogenized as described above (section 2.2.1) using the disposable pestles in 1:20 (w/v) of the extraction buffer containing 50 mM Na/K phosphate buffer pH 8.0, 30 mM Na_2SO_3, 0.5 mM 5'-AMP, and 10 mM DTT. The homogenate was subsequently centrifuged for 30 s at 2,000 g to remove cell debris and the protein concentration was determined using the Bio-Rad

protein assay with bovine serum albumin as a standard as described above (section 2.2.14). To measure the activity, 40 µl of protein extract was combined with 230 µl of reaction assay mix (Table 2.7) in a flat bottom 96-well polystyrene microplate (Greiner Bio-One). Production of NADH was measured at 340 nm using a SpectraMax 340PC^{384} microplate spectrophotometer. The background absorbance was measured for 3 min before the reaction was initiated by the addition of 30 µl of 10 mM sodium pyrophosphate. Progress of the reaction was measured for a further 3 min at 340 nm. Activity was calculated as nmol min^{-1} mg^{-1} of protein.

Table 2.7 ***ATPS activity reaction assay***

Concentration	Component	Volume/230 µl
1 M	Tris/HCl pH 8.0	15 µl
100 mM	$MgCl_2$	15 µl
100 mM	D-glucose	15 µl
6 mM	β-nicotinamide adenine dinucleotide hydrate (NAD^+)	15 µl
3.75 mM	adenosine 5'-phosphosulfate sodium salt (APS)	7.5 µl
1u $µl^{-1}$	Hexokinase from *S. cerevisiae,* Type III (HK)	5 µl
1u $µl^{-1}$	Glucose-6-phosphate dehydrogenase (g-6-PDH)	5 µl
-	dH_2O	152.5 µl

2.2.18 Determination of [^{35}S]sulfate incorporation in sulfate, thiols, and proteins

The uptake of [^{35}S]sulfate and its incorporation into sulfate, thiols, and proteins was measured as described previously (Kopriva et al. 1999, Vauclare et al. 2002). Briefly, arabidopsis plants grown on plates for 2.5-3 weeks were transferred into 24-well plates containing 1 ml of nutrient solution adjusted to sulfate concentration of 0.2 mM (Table 2.8) and supplemented with [^{35}S]sulfuric acid solution to obtain ca. 70,000 dpm/10 µl. Only roots were submerged in the solution. Plants were incubated for four hours in light at room temperature. After that time seedlings were washed in water and carefully blotted with paper towel. Roots and shoots were weighed separately in 1.5 ml tubes and frozen immediately in liquid nitrogen. Subsequently, plant tissue was homogenized as described above (section 2.2.1) with disposable pestles and extracted 1:10 (w/v) in 0.1 M HCl.

To determine sulfate uptake, 10 µl of plant extract was mixed with 1 ml of Optiphase HiSafe3 scintillation cocktail (Perkin Elmer) and the radioactivity was measured in a scintillation counter (Perkin Elmer Tri-Carb 2910 TR).

Table 2.8 ***Macroelement nutrient solution***

Component	Final concentration
$Ca(NO_3)_2$ $4H_2O$	1.5 mM
KNO_3	1 mM
KH_2PO_4	0.75 mM
$MgSO_4$ $7H_2O$	0.2 mM
Fe-EDTA	0.1 mM
pH adjusted to 6 with KOH	

To measure incorporation of [^{35}S]sulfate in proteins, total protein in 50 µl of plant extract was precipitated with 12.5 µl of 100% trichloroacetic acid (TCA) on ice as described previously (Kopriva et al. 1999). After that time the precipitate was collected by centrifugation for 10 min at 14,000 g and washed once in 100 µl of 1% TCA and once in 200 µl of 100% EtOH. The precipitate was subsequently dissolved in 100 µl of 0.1 M NaOH and the radioactivity was determined after addition of 1 ml of scintillation cocktail (Perkin Elmer) in a scintillation counter (Perkin Elmer Tri-Carb 2910 TR).

To determine the radioactivity in thiols, 50 µl of plant extract was mixed with 50 µl of 0.1 M NaOH and 1 µl of DTT, and incubated in the dark at 37°C for 15 min. Subsequently, 11.5 µl of 1M Tris/HCl pH 8.0 and 5 µl of 100 mM monobromobimane (MBB; Invitrogen) was added, and the samples were mixed and incubated in the dark at 37°C for 15 min. After this incubation the reaction was stopped by adding 11.5 µl of 50% acetic acid. Subsequently, samples were centrifuged for 15 min at 14,000 g. Standard thiols analysis was performed as described previously (section 2.2.13) with an injection volume of 100 µl. The HPLC was connected to a fraction collector (Frac-920, Amersham Biosciences, Buckinghamshire, UK) and fractions of 0.8 ml were collected in 6 ml scintillation vials. The radioactivity in each fraction was detected after addition of 2.5 ml of Optiphase HiSafe3 scintillation cocktail (Perkin Elmer) in a scintillation counter (Perkin Elmer Tri-Carb 2910 TR).

Chapter 3:

ATPS1 and APR2 – Two Consecutive Enzymes Contributing to Control of the Sulfate Reduction Pathway

3.1 Introduction

Sulfate is taken up by the plant from the soil solution via sulfate transporters SULTR1;1 and SULTR1;2 (Buchner et al. 2004). Most of the sulfate taken up by the root is transported to the shoot where it enters the reduction pathway as described in Chapter 1, being incorporated into cysteine in the final step of the process (Chapter 1; Takahashi et al. 2011). The remaining sulfate, not entering the assimilation pathway, is stored in the vacuole – the main reservoir of sulfate in plant cells (Martinoia et al. 2000). The pool of sulfate stored in the vacuole can be remobilized across the tonoplast membrane and enter the assimilation pathway (Kataoka et al. 2004b). The efficiency of sulfate remobilisation is of great importance especially during sulfur deficiency. The large variation in this process depends on the environmental conditions (Hawkesford 2000) and differs between species as well as within species (Durenkamp et al. 2007, Koralewska et al. 2007). Kataoka et al. (2004b) in their study on vacuolar sulfate transporters characterised differences in their mRNA accumulation depending on sulfate availability in the soil. They revealed that the expression of SULTR4;1 is independent of the sulfur supply whereas the expression of SULTR4;2 increases during sulfate deficiency. On the other hand, the mobilisation of the vacuolar pool in some species such as *Brassica napus* is known to be slower than in other species (Blake-Kalff et al. 1998). Moreover, sulfate stored in the vacuoles can buffer or optimize the flux of sulfate through the plant, balancing sulfate homeostasis at the whole plant level (Kataoka et al. 2004b). Despite the importance of sulfate remobilisation, very little is known about its control (see Chapter 1). Crop plants that accumulate high concentrations of sulfate such as *Brassica napus* often require sulfate fertilisation. In fact, the production of one tonne of wheat grain requires 2-3 kg of sulfur (Zhao et al. 1999), whereas production of one tonne of oilseed rape seeds requires 16 kg of sulfur (McGrath et al. 1996). Therefore, the investigation of control of plant sulfate homeostasis at the genetic level is of great importance.

Sulfate is chemically very stable and therefore it requires activation before it can be reduced. ATP sulfurylase (ATPS; EC: 2.7.7.4) activates sulfate by adenylation to adenosine 5'-phosphosulfate (APS). Plant ATPS is a homotetramer composed from 52-54 kDa polypeptides (Murillo & Leustek 1995). Most ATPS is plastid localised, where it is responsible for sulfate activation for further reduction. However, its activity was also detected in the cytosol (Lunn et al. 1990, Renosto et al. 1993, Rotte & Leustek 2000). During arabidopsis development ATPS activity in chloroplasts declines whereas it increases in the cytosol (Rotte & Leustek 2000). The reduction in total ATPS activity in leaves is not dependent on the leaf age since it was shown to decrease proportionally in all leaves in the plant (Rotte & Leustek 2000). ATPS is encoded by a

small multigene family. Most plant species possess two ATPS isoforms. However, a four-member gene family was identified in arabidopsis, which may indicate some level of genetic redundancy (Kopriva et al. 2009). Surprisingly, all of them contain a predicted chloroplast transit peptide (Hatzfeld et al. 2000, Murillo & Leustek 1995). The most likely explanation for the existence of a cytosolic ATPS isoform in arabidopsis is possible use of an alternate translational start codon for one of the genes, so that the transit peptide is not translated (Hatzfeld et al. 2000). However, this hypothesis remains to be investigated.

In the next step of sulfate assimilation, APS is either reduced to sulfite or phosphorylated to form 3'-phosphoadenosine-5'-phosphosulfate (PAPS). However, in plants reduction is the dominant route for assimilation (Leustek et al. 2000). APS reductase (APR) catalyses the reduction by transfer of two electrons delivered from glutathione (GSH) to APS to form sulfite. APR (EC: 1.8.99.2) is a key enzyme of the sulfur assimilation pathway. The expression of the *APR2* gene as well as protein activity are regulated by various environmental factors and signalling molecules (Kopriva 2006, Koprivova et al. 2008). Similarly to ATPS, APR is encoded by a small multigene family and three isoforms are known in arabidopsis. It was shown in various studies that *APR1* and *APR3* are co-regulated on the level of transcription and share the highest sequence similarity. However, the expression of *APR2* responds differently to hormone treatments (Koprivova et al. 2008) which indicate specific functions of particular APR isoforms. The amino acid sequence of APR suggests a multidomain structure. Its precursor is synthesized with an N-terminal plastid transit peptide. In the mature protein the N-terminal domain is similar in predicted amino acid sequence to PAPS reductase (PAPR) from bacteria and the C-terminal domain to thioredoxin (Trx). Because of the lack of a C-terminal domain, bacterial PAPR requires thioredoxin or glutaredoxin as a cofactor (Kopriva et al. 2007). This may suggest the role of the C-terminal domain of APR as redox cofactor (Leustek et al. 2000). arabidopsis APR is a dimer of 45 kDa subunits (Kopriva & Koprivova 2004) binding a $[Fe_4S_4]$ iron-sulfur cluster (Kopriva et al. 2001). Bacterial PAPR consists of two 28 kDa subunits without any prosthetic groups. A single conserved cysteine residue is responsible for its activity and dimerization.

Various studies have revealed that APR is the crucial regulatory step in the sulfate reduction pathway (Hesse et al. 2003, Koprivova et al. 2008, Vauclare et al. 2002). The expression of the *APR2* gene and the activity of APR protein vary strongly in response to environmental fluctuations. Evidence for its importance in controlling sulfate accumulation comes from studies of the Recombinant Inbred Lines (RIL) population derived from the cross between two wild arabidopsis accessions Bay-0 and Shahdara (Sha) which differ in sulfate

accumulation (Loudet et al. 2007). The Quantitative Trait Loci (QTL) analysis of RILs led to the identification of two major and several minor QTLs responsible for the variation in the sulfate accumulation within the population. The cloning of one of the major QTLs revealed that the underlying gene encoded the APR2 isoform of APS reductase (Loudet et al. 2007). The analysis of APR2 in both Bay-0 and Sha revealed a single nucleotide polymorphism (SNP) which leads to a substitution of alanine (Ala^{399}) with glutamate (Glu^{399}) in a thioredoxin active site of the protein. Due to this amino acid change Sha lost over 99% of APR2 enzymatic activity compared to Bay-0. This loss of sulfate reduction capacity led to sulfate accumulation, particularly under nitrogen limitation (Loudet et al. 2007).

Further evidence of the importance of APR2 comes from David E. Salt (University of Aberdeen) who investigated the leaf ionome of 349 arabidopsis accessions collected from across the species range. This analysis revealed a large variation in the elemental composition among the population. All the data obtained in these experiments have been deposited on the iHUB (Baxter et al. 2007; www.ionomicshub.org) and are publically available. In this analysis the Hod accession collected in the Czech Republic near to Hodonín showed the highest total sulfur concentration among entire collection. The causal locus was mapped to a region on the long arm of chromosome 1, containing 30 genes (Chao et al. 2014). This region includes the gene encoding isoform 2 of APS reductase (APR2) which was thus a strong candidate for the causal gene of the high total sulfur concentration in Hod accession. The analysis of three independent T-DNA insertion lines and transgenic complementation confirmed the hypothesis that *APR2* is the causal locus for the high total sulfur concentration in Hod (Chao et al. 2014). However, Hod did not share the amino acid change which was responsible for the loss of function in Sha.

In this chapter I describe further analyses of the molecular basis of sulfur homeostasis in arabidopsis. Using biochemical analysis of the enzyme as well as protein haplotype analysis it was established that the natural variation in leaf sulfate concentration is controlled by several rare APR2 alleles across the species range. Additionally, in a paper on which I am a co-author, it has been established recently that the gene underlying the second major QTL revealed in the analysis of Bay-0 x Sha RIL population encodes the ATPS1 isoform of ATP sulfurylase (Koprivova et al. 2013). In contribution to these studies I investigated the natural variation in *ATPS1* and have established that it was due to differences in expression of the gene. Results described here indicate that the differences in expression are linked to the two deletions in the *ATPS1* common for several arabidopsis accessions. The experiments on *ATPS1* were performed together with Marco Giovannetti from University of Turin.

3.2 Materials and Methods

3.2.1 Plant material and growth conditions

Three independent T-DNA lines: GABI_108G02 (*apr2-1*), SALK_119683 (*apr2-2*), and SALK_035546 (*apr2-3*) as well as transgenic lines expressing *APR2* genomic fragment from Col-0 introduced into Hod were kindly provided by Dai-Yin Chao (University of Aberdeen). The *apr1* mutant and the *apr1apr2* double mutant were obtained in the laboratory of Stanislav Kopriva at the John Innes Centre. These lines as well as the wild type Sha and Hod accessions were grown on vertical plates with Murashige Skoog media without sucrose (MS) supplemented with 0.8% agarose (section 2.2.1). The plates were placed in a controlled environment room at 20°C under 16-h-light/8-h-dark cycles for two weeks. Plants used for the measurement of APR activity and sulfate concentration as well as sulfate flux through the pathway in the analysis of Hod related genotypes were provided to me as frozen samples.

The seeds of natural arabidopsis accessions used for the investigation of the effect of natural variation on the function and activity of APR2 and ATPS1 were either provided to the laboratory of Stanislav Kopriva by David E. Salt (University of Aberdeen) or obtained from the Nottingham Arabidopsis Stock Centre (NASC). Dry seeds were surface sterilised for up to four hours in a vacuum desiccator using chlorine gas which was generated by mixing 125 ml of sodium hypochlorite with 2.5 ml of 12 M HCl. Seeds were then mixed with 0.1% sterile agarose (Sigma) and sown on plates with MS medium (section 2.2.2). Subsequently, plates were stored in 4°C in the dark for three days. After that time they were transferred to a CER at 22°C under 16-h-light/8-h-dark cycles where they were left to germinate – up to one week (horizontally). One week old seedlings were transferred from plates to the 40-cell tray (21x35cm) with Levington Horticulture soil mix (Figure 5.2) and grown for another four weeks in CER 10-h-light/14-h-dark cycles at constant temperature 22°C, 60% relative humidity, and light intensity of 160 µE s^{-1} m^{-2}.

3.2.2 RNA extraction and expression analysis

Total RNA was isolated from leaves of five week old plants by the extraction with phenol : chloroform : isoamylalcohol mix (in proportion 25:24:1 units respectively) and LiCl precipitation (Sambrook et al. 1989). Each sample was homogenized as described in Chapter 2 with 500 µl of extraction buffer (80 mM Tris pH 9.0; 5% SDS; 150 mM LiCl; 50 mM EDTA). Subsequently an equal amount of phenol mix was added and the sample was mixed for 15 sec. All samples were then centrifuged 25 min at 14,000 g. The supernatant was transferred to new tubes and 500 µl µl of the phenol mix was added. The sample was mixed for 15 sec and

centrifuged again for 20 min at 14,000 g. The extraction step was repeated one more time. Subsequently 130 µl of 8 M LiCl was added to the supernatant and the sample was placed at -20°C overnight. The next day all samples were centrifuged at 4°C at 14,000 g. The supernatant was discarded and 300 µl of water was added to the sample. The pellet was re-dissolved by incubation at 65°C for 10 min. Subsequently 100 µl of 8 M LiCl was added and the samples were placed at -20°C overnight. The next day all the samples were centrifuged at 4°C at 14,000 g and the pellet was washed with 400 µl of 70% ETOH. The samples were then centrifuged for 5 min at 14,000 g, the supernatant was discarded and the pellet was dried for 5 min under the laminar flow hub. Subsequently 30 µl of water was added to the samples and the pellet was re-dissolved by incubation at 65°C for 10 min. Afterwards, 5 µl of RNA was mixed with 500 µl of RNAse free water and the concentration was measured by the Lambda Bio spectrophotometer (Perkin Elmer) with the quartz cuvette and originally installed settings for measurement of RNA concentration. The first-strand complementary DNA (cDNA) was synthesised from 1 µg of RNA using the QuantiTect Reverse Transcription Kit (Qiagen) according to the manufacturer's instruction. This kit includes the DNase treatment necessary to remove DNA contamination. Transcript abundance was analysed by real-time quantitative RT-PCR (qPCR), using the fluorescent intercalating dye SYBR Green (Applied Biosystems) in a DNA engine OPTICON2 continuous fluorescence detector (Bio-Rad). The qPCR was performed using gene specific primers and the results were normalised to the *TIP41* gene (Table 3.1M). The qPCR was performed in duplicate for each of three independent samples.

Table 3.1M ***Sequences of primers used for the analysis of transcript abundance by qPCR***

gene	5' primer	3' primer
TIP41	GTGAAAACTGTTGGAGAGAAGCA	TCAACTGGATACCCTTTCGC
ATPS1	CACTCGGAGGTTTCATGAGAG	AGACGTAGCGAGTTAAAATGAAGAG
APR2	AAAAGAGCTCCACGGGCTAT	CGACATGAGTGAATCAACATCTC

3.2.3 Statistical analysis

The statistical analysis of the data presented in this chapter was performed using the DSAADTAT (http://accounts.unipg.it/~onofri/DSAASTAT/DSAASTAT.htm), a Microsoft Excel (2010) VBA MACRO for basic statistical analyses released by Andrea Onofri.

3.3 Results

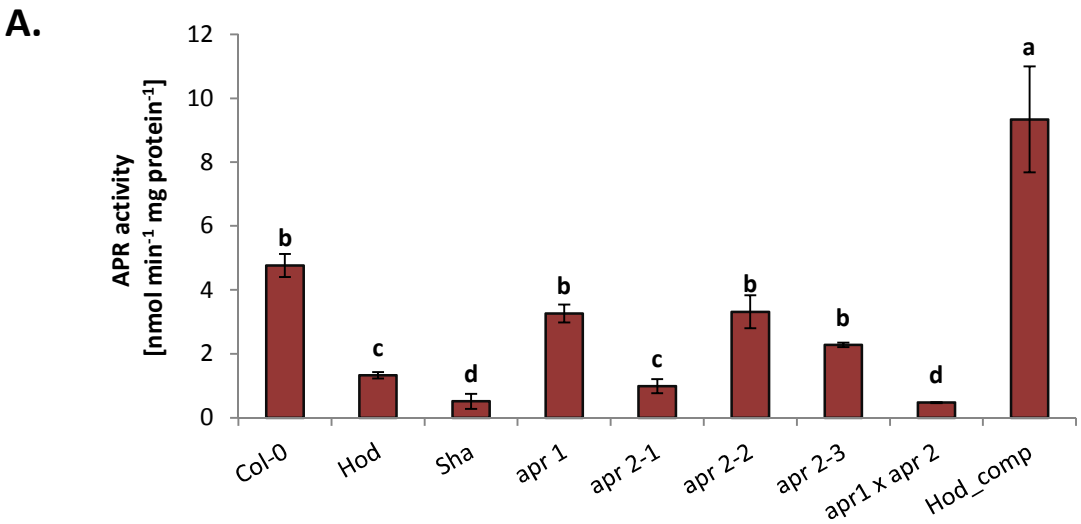
3.3.1 The link between sulfate concentration and APR activity

My contribution to the collaborative analysis of the high sulfur phenotype in Hod was biochemical analysis of sulfate assimilation. Since the results delivered by Dai-Yin Chao (University of Aberdeen) and the data deposited on the iHUB concerned total sulfur concentration, I first investigated the sulfate concentration in the accessions and mutants as well as the link between sulfate concentration and APR activity. I analysed the three independent T-DNA lines with the insertion in the 4th exon of *APR2* (*apr2-1,* GABI_108G02), and 169 bp (*apr2-2*, SALK_119683) and 116 bp (*apr2-3*, SALK_035546) upstream of the *APR2* translation start site. Additionally, I analysed the transgenic lines expressing the *APR2* genomic fragment from Col-0 (including a 1.5 kb promoter region, the gene body and 896 bp downstream sequences) introduced into the Hod accession. The *apr2* T-DNA lines and the transgenic lines listed above were provided by Dai-Yin Chao. Additionally, I analysed the wild type Sha and Hod accessions using Col-0 as a control. To explore further the contribution of different APR isoforms to the total activity and possible natural variation in isoforms I also analysed the *apr1* mutant, obtained previously in the Kopriva's laboratory by TILLING (Kopriva et al. 2009), as well as a double mutant *apr1apr2* which was generated by crossing the *apr1* TILLING line and *apr2-1* T-DNA line. All of these plant materials were provided to me as frozen samples on which I performed the analysis of sulfate concentration and the biochemical analysis of APR activity.

The analysis of two week old seedlings revealed a robust negative link between sulfate concentration and APR activity. In the *apr2-1* mutant I observed 80% reduction in APR activity and an increase in sulfate concentration of about 60% compared to Col-0 (Figure 3.1). The *apr2-2* and *apr2-3* mutants with the T-DNA insertion in the *APR2* promoter also showed a decrease in APR activity, but to a much lesser extent than in *apr2-1* which was previously determined as a loss-of-function allele (Loudet et al. 2007). The qPCR analysis revealed a complete loss of transcript in *apr2-1* as expected and partial loss in *apr2-2* and *apr2-3* (Chao et al. 2014). Sha showed 90% lower APR activity than Col-0 and no difference in sulfate concentration compared to Col-0 which is consistent with the observation of Loudet et al. (2007). Similarly, the Hod accession showed 70% reduction in APR activity compared to Col-0 suggesting that it carries a loss-of-function *APR2* allele (Figure 3.1A). Interestingly, the expression of *APR2* in Hod was indistinguishable from this ebserved in Col-0 (Chao et al. 2014). These data confirm that the low APR activity in Hod is not caused by the low expression of the

gene, but is probably driven by other changes in the *APR2* gene sequence. However, in contrast to the *apr2-1* mutant I did not observe higher sulfate concentration in Hod accession compared to Col-0 (Figure 3.1B). It was observed in Kopriva's laboratory previously that the differences in sulfate concentration are more pronounced in older plants. Therefore, the lack of difference in sulfate accumulation in Hod accession might be due to the developmental stage of the plants. Moreover, complementation of the Hod accession with the *APR2* allele from Col-0 resulted in a significant increase in APR activity and sulfate concentration comparable to Col-0 (Figure 3.1) confirming further that variation in *APR2* gene is responsible for the variation in enzyme activity and subsequent sulfate accumulation. All these results indicate that high sulfate phenotype in Hod is due to variation in *ARP2* gene. This is in agreement with the findings of Chao et al. (2014) who showed that the high total sulfur concentration in Hod is due to variation in the *APR2* gene.

The T-DNA insertion in *APR1* did not result in a significant decrease in APR activity (Figure 3.1A). The sulfate concentration in that line was also comparable with Col-0 (Figure 3.1B). Additionally, the APR activity in an *apr1xapr2* double mutant was strongly reduced and sulfate concentration was elevated to similar extent as in the *apr2* mutants. These results indicate that the APR2 isoform of APR contributes much more to total APR activity than the APR1 isoform as established previously (Loudet et al. 2007).



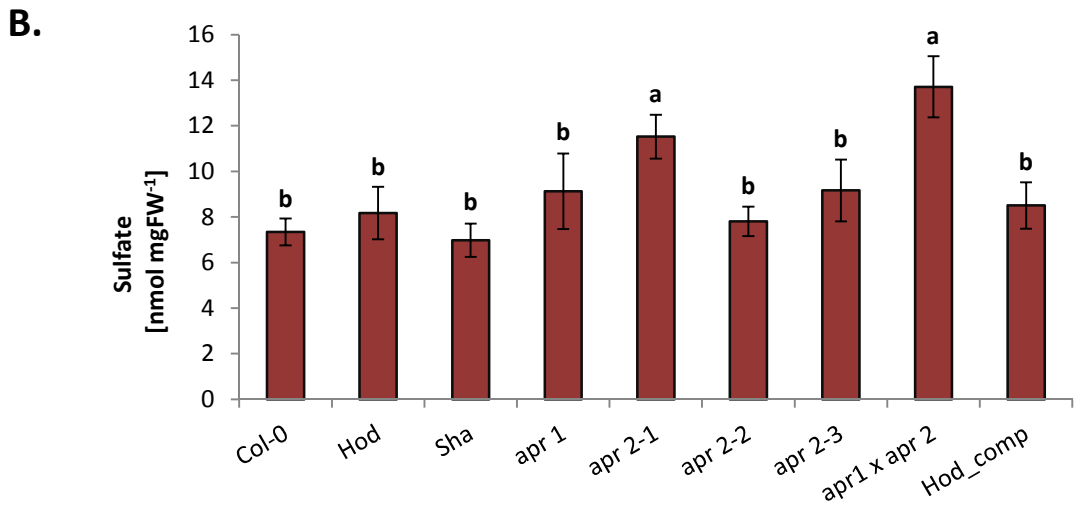


Figure 3.1 ***Analysis of the link between APR activity and sulfate concentration***

APR activity (A) and sulfate concentration (B) in the arabidopsis accessions, mutants in *APR1* and *APR2* and Hod lines complemented with APR2^{Col-0} allele. All data except Hod lines complemented with APR2^{Col-0} (Hod_comp) represent the mean values of 3 biological replicates ± standard error. The data for Hod line complemented with APR^{Col-0} represent mean value of three independent lines ± standard error. Letters above each bar indicate statistically significant groups using one-way ANOVA with Newman-Keuls multiple range test using 95% confidence interval.

3.3.2 Sulfate uptake and flux through the reduction pathway

To understand the link between a low APR activity and leaf sulfate concentration I performed an experiment on two week old seedlings of accessions and mutants with changes in APR2 (relative to Col-0) including Hod, Sha, *apr1*, *apr2-1*, *apr2-2*, *apr2-3*, and Hod complemented with *APR2^{Col-0}* using Col-0 as a control. These seedlings were incubated with [^{35}S] sulfate for four hours before harvesting (see Chapter2, section 2.2.18). The concentration of radioactive sulfate in plant extracts is referred to as sulfate uptake and was determined essentially as described by Mugford et al. (2011). Sulfur flux through sulfate assimilation pathway was measured as incorporation of ^{35}S from [^{35}S]sulfate to thiols and proteins. The incorporation of radioactive sulfate into thiols and proteins (expressed as concentration) were measured essentially as described by Kopriva et al. (1999) and Vauclare et al. (2002). The percentage of the radioactivity detected in thiols and proteins together (relative to the total radioactivity detected in plant extracts) is referred to as relative sulfate flux through the assimilation pathway. It should be stressed that these traits are expressed as concentrations (Chao et al. 2014) and therefore should be interpreted with caution. The analysis revealed a significantly lower relative sulfate flux (lower percentage of [^{35}S] incorporated into thiols and

proteins) in Hod and Sha compared to Col-0. Similarly, relative sulfate flux was also lower in the *apr2* mutants, but higher than in the two accessions (Hod and Sha). The *apr1apr2* double mutant showed the lowest relative sulfate flux (Figure 3.2A) suggesting a link between APR activity and relative sulfate flux. Indeed, the APR activity showed a strong positive correlation (r^2=0.7) with relative sulfate flux (Figure 3.2B).

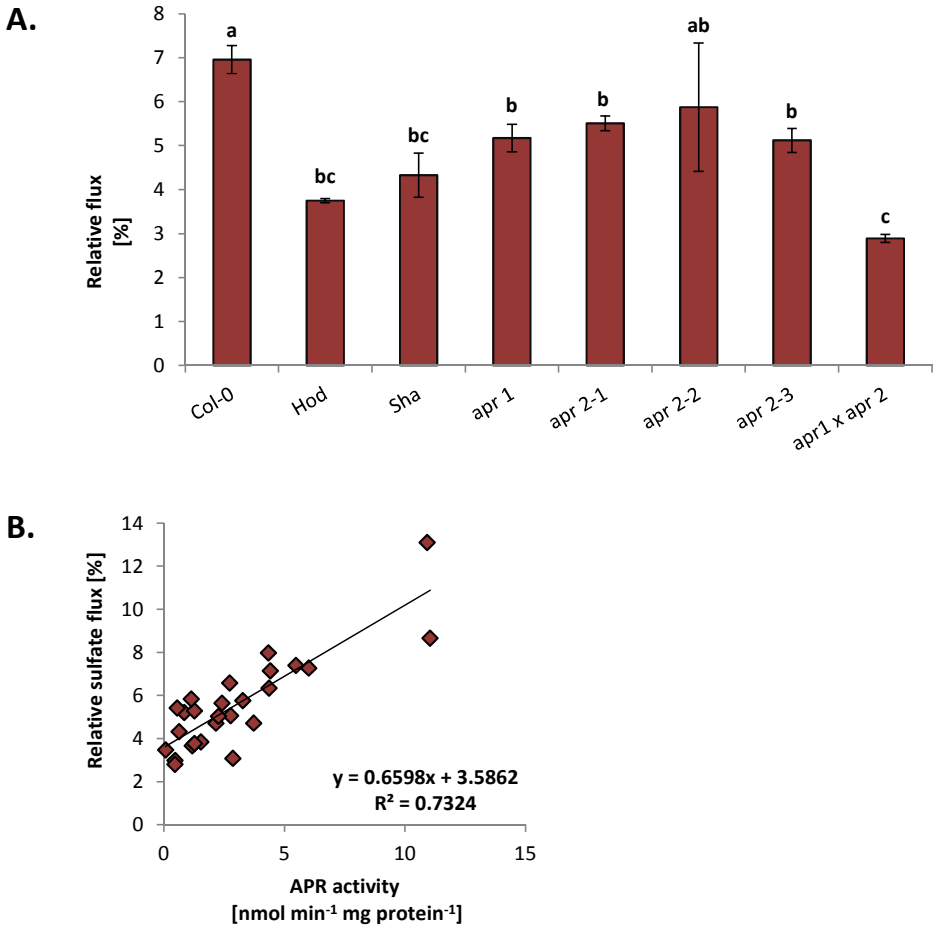


Figure 3.2 ***Analysis of sulfate flux through the assimilation pathway***

Relative sulfate flux through assimilation pathway (A) in the arabidopsis accessions and mutants with a weak *APR2* allele shown as a mean value of 3 biological replicates ± standard error. The letters indicate significantly different groups using one-way ANOVA with Newman-Keuls multiple range test using 95% confidence interval. (B) Correlation between APR activity and relative sulfate flux. Data points correspond to values from individual plants of all genotypes analysed.

Sulfate uptake is known to be strongly up-regulated in response to sulfate starvation (Davidian & Kopriva 2010) and this response is mediated by the concentration of cysteine and glutathione (Datko & Mudd 1984, Vauclare et al. 2002). Therefore, I measured sulfate uptake and GSH concentration in all accessions and mutants with *APR2* allele different than in Col-0 to examine whether increased sulfate accumulation may lead to the induction of a sulfate

starvation response in these plants (Figure 3.3). The *apr2-1* mutant and *apr1xapr2* double mutant showed the greatest concentration of [^{35}S]sulfate (sulfate uptake) among all the lines tested (Figure 3.3A), and Hod and Sha showed the smallest sulfate uptake among all lines tested. No other lines were different from Col-0 in sulfate uptake. Both Sha and Hod had a low GSH concentration (Figure 3.3B) which most likely is caused by the low sulfate uptake (Figure 3.3A), low APR activity (Figure 3.1A), and low incorporation of sulfate into thiols and proteins (Figure 3.2A) observed in these two accessions. Interestingly, the GSH concentration in *apr2-1* mutant was comparable with Col-0 whereas it was lower in the *apr1xapr2* double mutant than in Col-0.

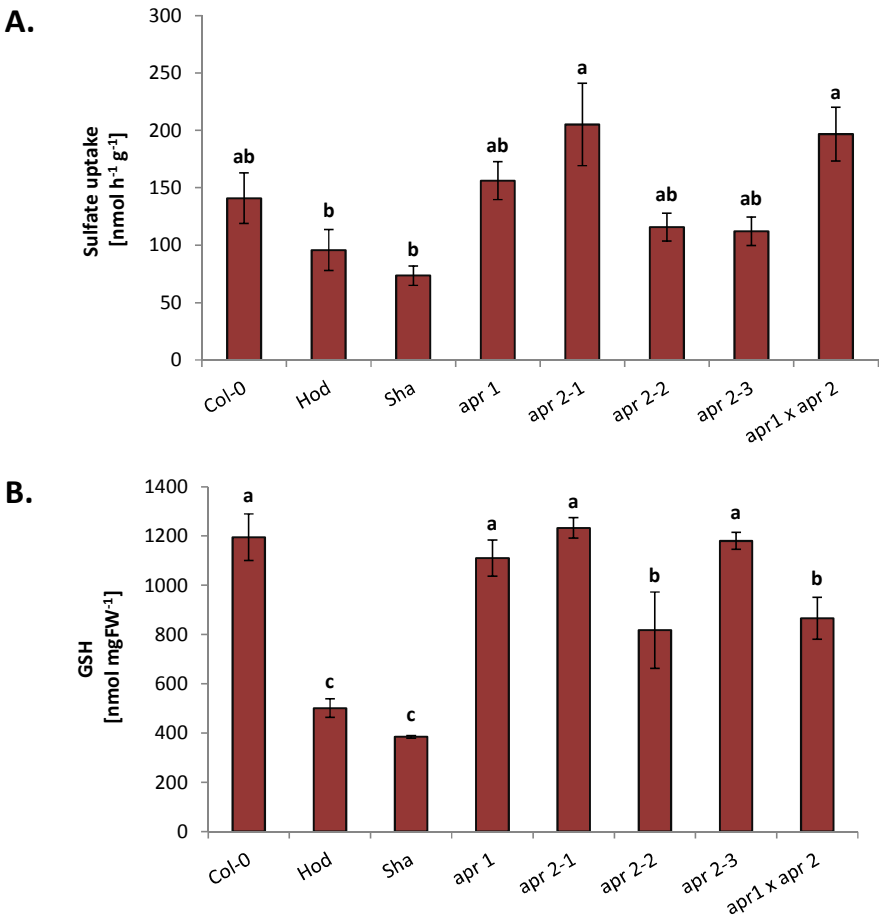


Figure 3.3 ***Analysis of sulfate uptake and concentration of glutathione (GSH)***

Sulfate uptake (A) and leaf GSH concentration (B) in arabidopsis accessions and mutants with weak *APR2* allele. Data represent the mean values of three biological replicates ± standard error. Letters above each bar indicate statistically significant groups using one-way ANOVA with Newman-Keuls multiple range test using 95% confidence interval.

3.3.3 *Rare APR2 alleles with a strong effect on the enzymatic activity of the protein*

The severe loss of APR activity in Sha is caused by a single amino acid substitution of Ala^{399} with Glu^{399} (Loudet et al. 2007). In Hod, Gly^{216} localised in the conserved Arg-loop in the central region of the APR2 protein is changed to an Arg. This polymorphism was suggested to cause the low APR activity in this accession (Chao et al. 2014). Moreover, the analysis of 855 APR2 protein haplotypes for which sequences were obtained as a part of the 1001 genomes project (www.1001genomes.org) did not reveal any other accessions with the Sha or Hod *APR2* alleles confirming that these two alleles are very rare in a global A*rabidopsis thaliana* population. In order to examine whether there is a more general link between *APR2* and sulfate accumulation two groups of accessions were selected as candidates. The first group comprised the four accessions with the highest shoot total sulfur concentration from the collection of 349 accessions screened by David E. Salt and Dai-Yin Chao from Aberdeen excluding Hod: LAC-5, Tamm-2, Oy-0, and CS28332 (publically available at www.ionomicshub.org). The second group comprised eight accessions with non-synonymous substitutions in the *APR2* coding region (Table 3.1). They were selected based on the sequence analysis of the 855 accessions re-sequenced within 1001 genomes project. Since the differences in sulfate accumulation are more pronounced in older plants, I grew these 12 accessions in a controlled environment room (CER) for five weeks to maximise the chances of observing differences. I analysed the shoots for sulfate concentration and APR activity.

Table 3.1 ***Non-synonymous amino acid substitutions in APR2 coding region***

									PAPS reductase domain									Trx-like domain			APR
Position	5	16	21	24	40	56	58	65	107	111	155	182	216	242	265	343	349	385	399	427	activity
Col-0	Val	Ser	Gly	Ser	Thr	-	Ser	Thr	Arg	Gln	Ala	Gly	Gly	Asp	Phe	Asn	**Lys**	Ile	Ala	Pro	8.91
ICE61	**Ala**	Ser	Gly	Ser	Thr	-	Ser	Thr	Arg	Gln	**Val**	Gly	Gly	Asp	Phe	Asn	Arg	Ile	Ala	Pro	6.27
Lc-0	Val	Ser	**Arg**	**Ala**	**Asn**	-	**Thr**	Thr	**Lys**	**Glu**	Ala	**Asp**	Gly	Asp	Phe	Asn	Arg	Ile	Ala	Pro	9.06
Qar-8a	Val	Ser	Gly	Ser	Thr	-	Ser	Thr	Arg	Gln	Ala	Gly	Gly	**Gly**	Phe	**Ser**	Arg	Ile	Ala	Pro	12.24
Lov-1	Val	Ser	Gly	Ser	Thr	-	Ser	Thr	Arg	Gln	Ala	Gly	Gly	Asp	**Ser**	Asn	Arg	Ile	Ala	Pro	0.95
Lov-5	Val	Ser	Gly	Ser	Thr	-	Ser	Thr	Arg	Gln	Ala	Gly	Gly	Asp	**Ser**	Asn	Arg	Ile	Ala	Pro	0.71
Fäl-1	Val	Ser	Gly	Ser	Thr	-	Ser	Thr	Arg	Gln	Ala	Gly	Gly	Asp	**Ser**	Asn	Arg	Ile	Ala	Pro	0.81
Tfa-08	Val	Ser	Gly	Ser	Thr	-	Ser	Thr	Arg	Gln	Ala	Gly	Gly	Asp	**Ser**	Asn	Arg	Ile	Ala	Pro	0.98
Sei-0	Val	Ser	Gly	Ser	Thr	-	Ser	Thr	Arg	Gln	Ala	Gly	Gly	Asp	Phe	Asn	Arg	Ile	Ala	**Ser**	8.22
Sha	Val	Ser	Gly	Ser	Thr	-	Ser	Thr	Arg	Gln	Ala	Gly	Gly	Asp	Phe	Asn	Arg	**Val**	**Glu**	Pro	1.67
Hod	Val	**Thr**	Gly	Ser	**Asn**	**Ser**	**Thr**	**Leu**	**Lys**	**Glu**	Ala	Gly	**Arg**	Asp	Phe	Asn	Arg	**Val**	Ala	Pro	0.96

Amino acid changes in various accessions are highlighted in dark grey. The residues in box with the positions between 107 and 265 are localised in the PAPS reductase domain and those between positions 385 and 427 – in thioredoxin-like domain. APR activity unit is nmol min^{-1} mg $protein^{-1}$. Data obtained from 1001 Genomes database.

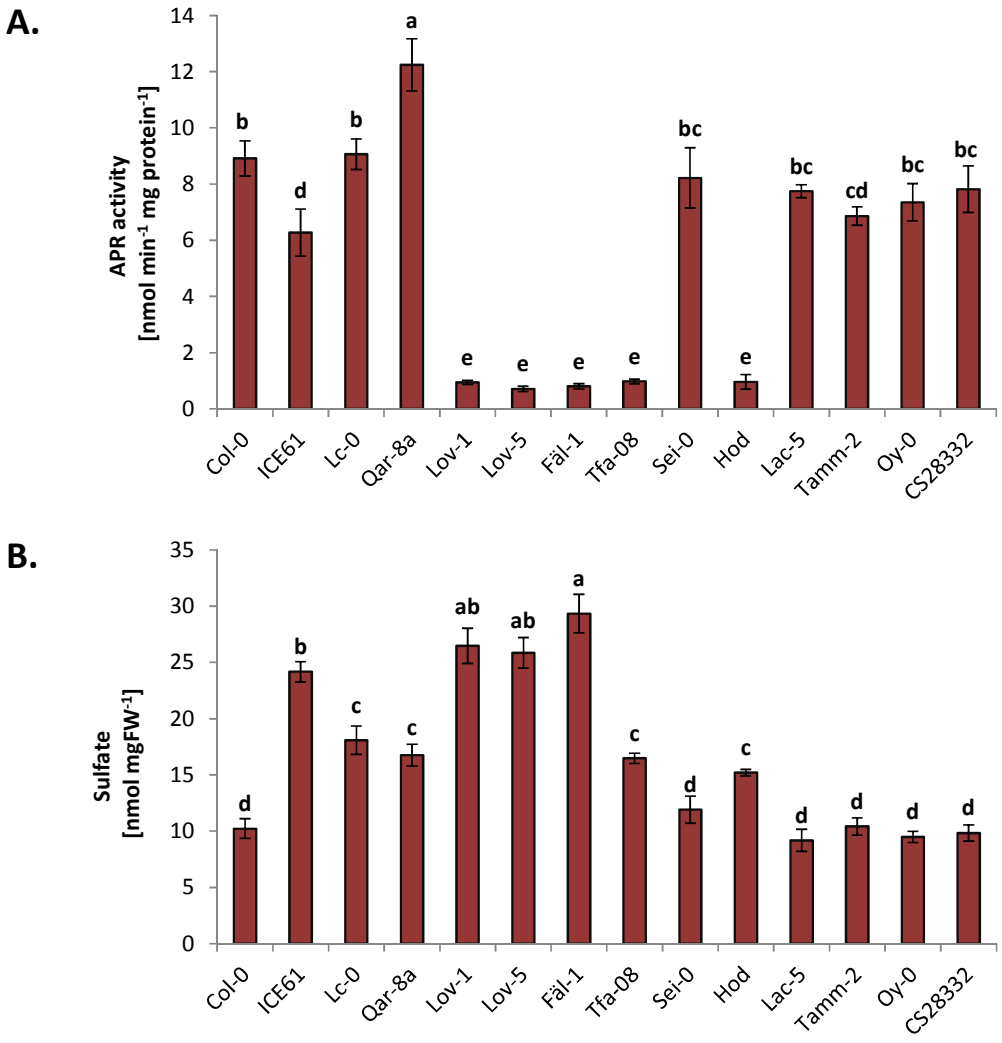


Figure 3.4 ***Analysis of the natural variation in APR activity and sulfate concentration***

APR activity (A) and sulfate concentration (B) in arabidopsis accessions. Data represent the mean values of four biological replicates ± standard error. Letters above each bar indicate statistically significant groups using one-way ANOVA with Newman-Keuls multiple range test using 95% confidence interval.

The analysis of five week old plants again revealed a very low APR activity in Hod accession. I also observed 55% higher shoot sulfate concentration in Hod than in Col-0 (Figure 3.4). These results confirm that the differences in sulfate accumulation between accessions might not be noticeable in seedlings, but they are more pronounced in mature plants. Apart from Hod five other accessions showed lower APR activity compared to Col-0. Four of them: Lov-1, Lov-5, Fäl-1, and Tfa-08; were collected in western Sweden and share the same amino acid substitution of Phe^{265} to Ser^{265} in the conserved PAPS reductase domain of the APR2 enzyme (Table 3.1). Similarly to Hod, I observed about 86% lower APR activity for all the Swedish accessions compared to Col-0 (Figure 3.4A). Furthermore, three of them: Lov-1, Lov-5

and Fäl-1 showed two-fold higher sulfate concentration compared to Col-0 (Figure 3.4B). Surprisingly, the sulfate concentration in Tfa-08 accession was much less elevated than in the other Swedish accessions even though they share the same amino acid substitution. The ICE61 accession, collected in south western Russia, was the fifth accession for which I observed lower APR activity. However it was not so strongly pronounced as in Hod and the Swedish accessions (30%; Figure 3.4A). Sulfate concentration in this accession was about 60% higher compared to Col-0 (Figure 3.4B). ICE61 represents a singleton haplotype which has an amino acid change of Ala^{155} to Val^{155}, also localised in the PAPS reductase domain (Table 3.1). Interestingly, the Qar-8a accession, collected from Lebanon, with the substitution of Asp^{242} to Gly^{242} in the PAPS reductase domain showed higher APR activity than Col-0. Moreover, this accession also had over 60% more sulfate than Col-0. There were no significant differences in APR activity in the following accessions with high sulfate concentration: LAC-5, Tamm-2, Oy-0, and CS28332, compared to Col-0, suggesting that the mechanism underlying the high sulfur phenotype in these accessions is independent from APR activity (Figure 3.4).

3.3.4 Natural variation in the ATPS1 gene

As mentioned before, *ATPS1* was identified as a gene underlying the second major QTL responsible for the variation in sulfate accumulation revealed in the analysis of the Bay-0 x Sha RIL population (Koprivova et al. 2013, Loudet et al. 2007). To confirm that ATPS1 contributes to the variation in sulfate concentration between Bay-0 and Sha, Koprivova et al. (2013) analysed Col-0, Sha and Bay-0 as well as the *atps1* mutant for sulfate concentration, ATPS activity and expression of *ATPS1* (Figure 3.5). The *atps1* mutant in the Col-0 background had significantly higher sulfate concentration compared to all the wild accessions tested (Figure 3.5A), as shown previously (Kawashima et al. 2011, Liang et al. 2010). Additionally, Bay-0 had greater sulfate concentration compared to Col-0 and Sha. ATPS activity was about 60% lower in the mutant and about 40% lower in Bay-0 compared to both Col-0 and Sha (Figure 3.5B). The analysis of *ATPS1* mRNA transcript (Figure 3.5C) confirmed a loss-of-function phenotype in the *apts1* mutant and revealed approximately 30% lower transcript abundance in Bay-0 compared to Col-0 (Koprivova et al. 2013). This suggests that the differences in ATPS activity between Bay-0 and Sha are more likely to be due to differences in expression rather than kinetics of the enzyme. Indeed, the sequencing of the *ATPS1* gene from Bay-0 and Sha did not reveal any amino acid differences. However, two deletions and various nucleotide substitutions were observed in the Bay-0 ATPS1 gene, compared to the Sha and Col-0 backgrounds (Figure S3.1).

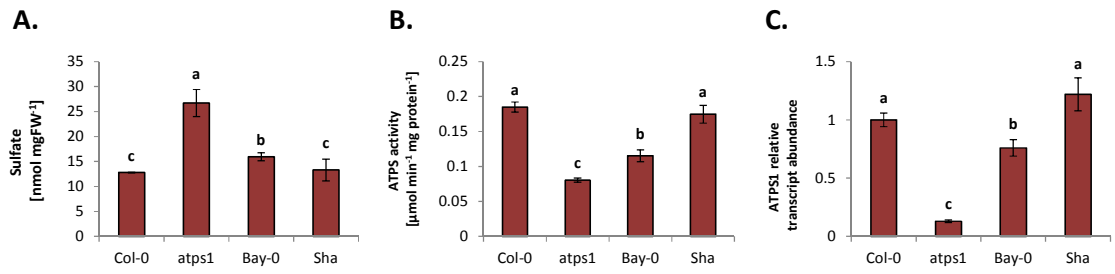


Figure 3.5 ***Analysis of the contribution of ATPS1 to sulfate homeostasis***

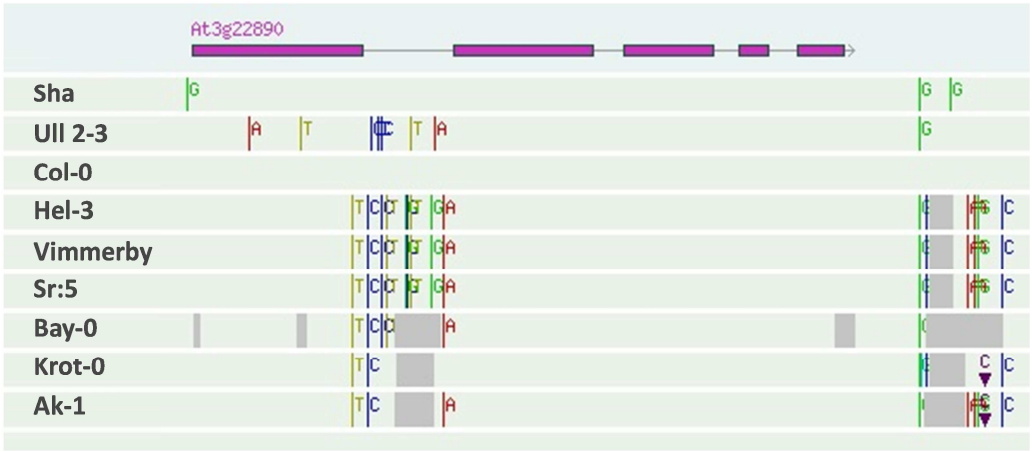
(A) Sulfate concentration, (B) ATPS activity, and (C) *ATPS1* transcript level shown as relative to Col-0 for which the value was set to 1. All measurements were carried out on five week old plants. Results are presented as means from three individual plants ± standard error. The transcript accumulation were analysed in duplicate. Letters above each bar indicate significantly different values at P<0.05. These data are from Koprivova et al. (2013).

Following on from the work done by Koprivova et al. (2013), the differences in *ATPS1* sequence in the global population of A*rabidopsis thaliana* were examined. Since the presence of the two deletions in Bay-0, in intron 1 and downstream of the gene, is probably the main difference between $ATPS1^{Bay\text{-}0}$ and $ATPS1^{Sha}$ (Koprivova et al. 2013) the 1001 Genomes database was screened to investigate whether there are more accessions sharing this allele. In Bay-0 the deletion localised in a middle part of the first intron is 13-bp long and it is accompanied by highly polymorphic region around it. The second deletion, located 249 bp downstream of the *ATPS1* stop codon is 71-bp long and it is also accompanied by few additional substitutions (Figure 3.6A). Among all the 500 accessions available in the database at the time of the analysis, 50 shared the deletion downstream of the *ATPS1* gene. Among these, 28 also had the deletion in the first intron. To confirm the presence of these deletions and verify their exact position, corresponding DNA fragments from six representative accessions were amplified and sequenced by Koprivova et al. (2013; Figure S3.2). The deletions indicated in the database were confirmed. However, the exact borders of the deletions differed slightly from those specified in the database (Koprivova et al. 2013). Nevertheless, a subset of 12 accessions with no deletion (including Sha and Col-0), 7 accessions with the deletion downstream of *ATPS1* and 10 accessions with two deletions (including Bay-0) was selected for further analyses.

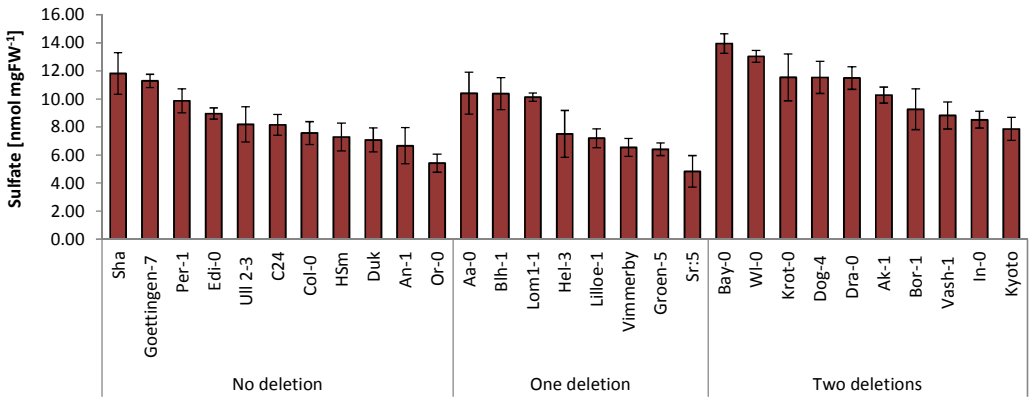
To examine how the deletions affect the mRNA transcript accumulation in these accessions a qPCR analysis was conducted. To account for the differences in the expression of the sulfate reduction pathway, the *ATPS1* transcript abundance is shown as relative to

constitutively expressed *APR2* (Figure 3.6C). The mean value for all 10 accessions with two deletions was significantly lower than the mean value from the accessions with one deletion and with no deletions (Figure 3.6E). Additionally, sulfate concentration was analysed in these accessions to investigate how the variation in transcript abundance affects accumulation of sulfate (Figure 3.6B). The 10 accessions sharing the two deletion haplotype showed higher sulfate concentration compared to the accessions with one deletion and to the accessions with no deletion. The mean value from all the 10 accessions was about 30% higher than the mean values from the two other haplotype groups (Figure 3.6D). These results support the hypothesis that the deletion in the intron (but not the deletion downstream of *ATPS1* stop codon) results in a decrease of expression of *ATPS1* which leads to a decrease in overall ATPS activity and increase in sulfate accumulation.

A.



B.



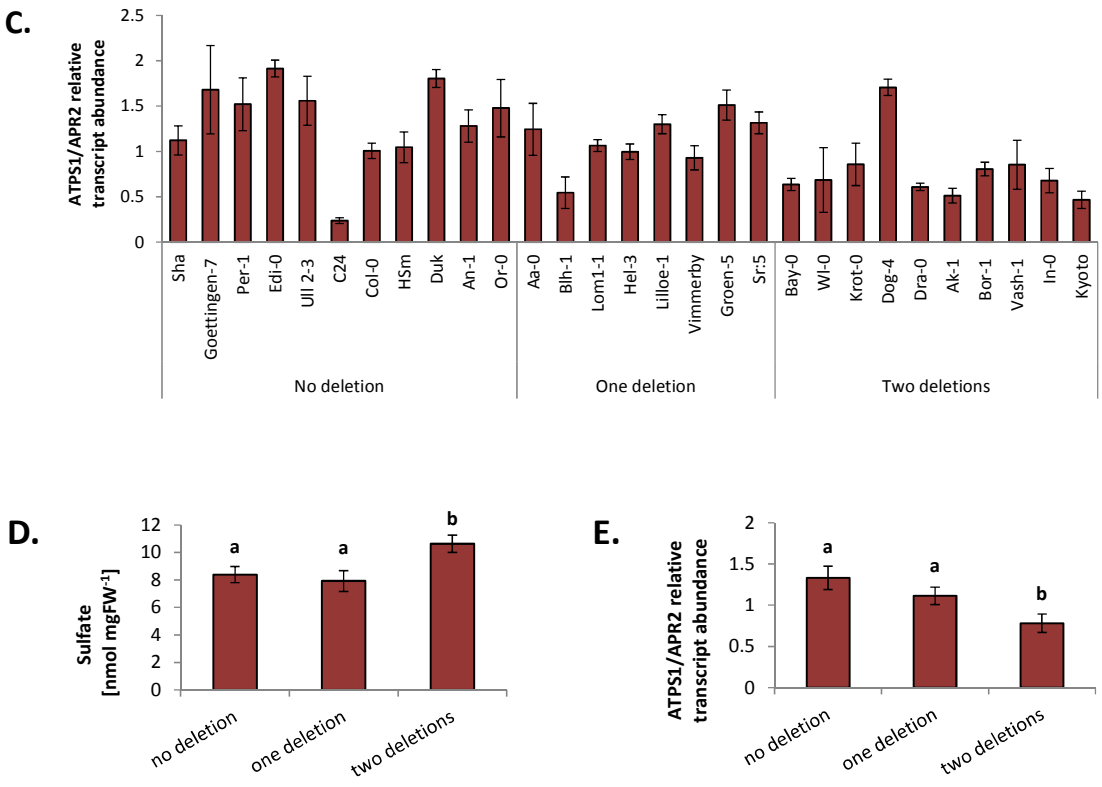


Figure 3.6 ***Natural variation in ATPS1 among Arabidopsis thaliana accessions***

Schematic presentation of the one and two deletion haplotypes in representative accessions (A); the data were obtained from 1001 Genomes database. The grey boxes represent the deletions. Sulfate concentration (B) and relative transcript abundance of ATPS1 in leaves of individual accessions representing three different haplotypes grouped according to the number of deletions in ATPS1. Results are presented as means of three biological replicates ± standard error. The transcript accumulation was measured in duplicate. Sulfate concentration (D) and relative transcript abundance of ATPS1 (E) in leaves of three haplotypes of arabidopsis accessions shown as means of all accessions analysed per haplotype (as in B and C respectively) ± standard error. Different letters mark significantly different values at $p<0.05$.

The analysis of 1001 Genomes database also revealed a number of accessions with various non-synonymous SNPs leading to amino acid substitutions which were not found in Bay-0 and Sha. To investigate how this variation in ATPS1 affects the enzyme activity and sulfate accumulation sulfate concentration was first analysed in representative accessions for each haplotype (Figure 3.7A). Additionally, sulfate accumulation in Col-0 and the *atps1* mutant was analysed to use it as a reference. All of the accessions analysed showed the same or higher sulfate accumulation than Col-0. Fjae1-2 was the only accession with sulfate concentration lower then Col-0. Three accessions including Dra2-1, Vie-0, and Naes-2 showed the highest sulfate accumulation comparable to that observed in *atps1* mutant (Figure 3.7A). Since the variation in sulfate accumulation is dependent on various factors, the total ATPS activity in these accessions was determined (Figure 3.7B) to investigate if there is any correlation

between sulfate accumulation and ATPS activity as was observed for accessions with different APR2 haplotypes. This analysis revealed that the ATPS activity in most of the accessions was comparable to or even higher than that of Col-0 suggesting that the amino acid changes found in these accessions do not affect negatively the enzyme activity and therefore sulfate accumulates due to other factors. Moreover, T800 had the highest ATPS activity among all accessions tested, but the sulfate concentration in this accession did not differ significantly from Col-0. The only accession with significantly lower ATPS activity and significantly higher sulfate concentration compared to Col-0 was Naes-2 indicating that the substitution of G^{342} into D^{342} found in this accession may affect the catalytic activity of the enzyme (Figure 3.7B).

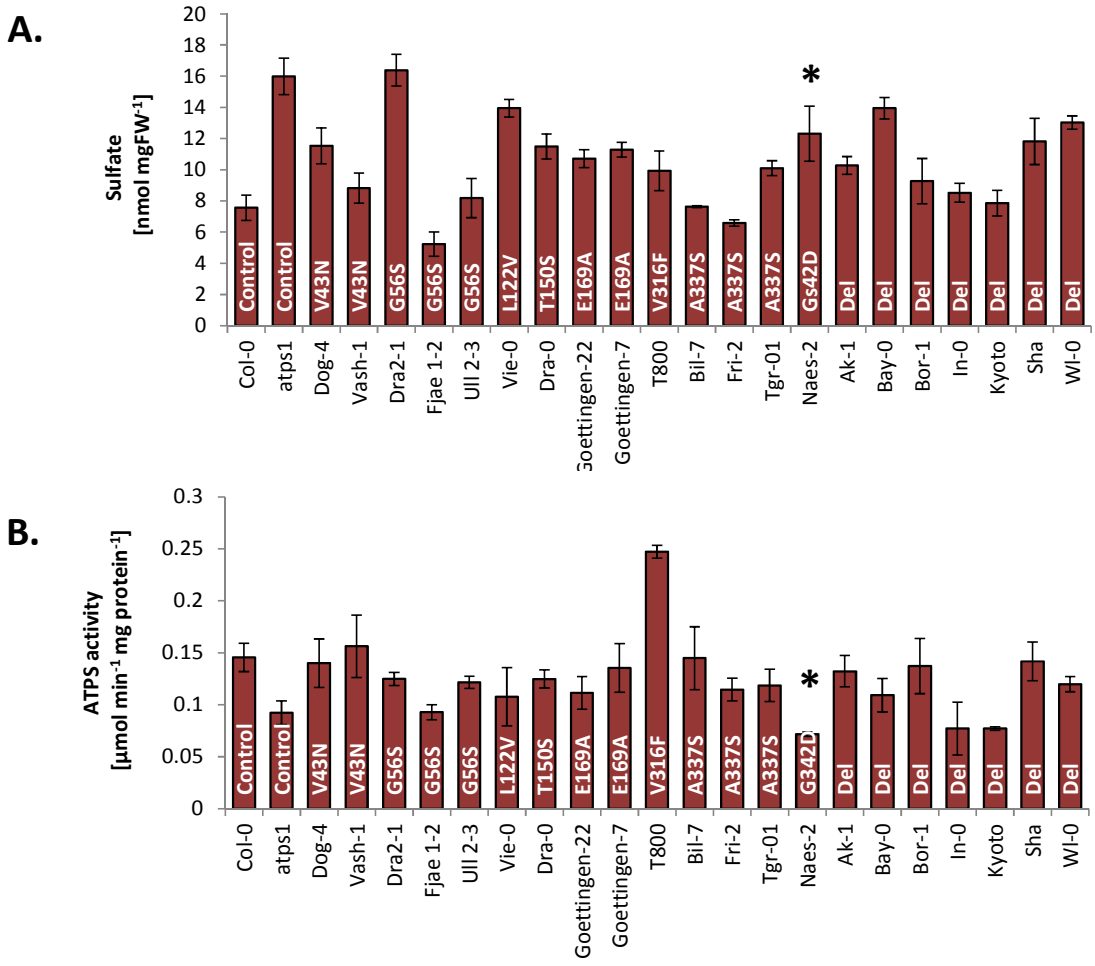


Figure 3.7 ***Analysis of the natural variation in ATPS activity and sulfate concentration***

Sulfate concentration (A) and ATPS activity (B) were measured in the accessions representing various ATPS1 haplotypes. Additionally, Col-0 and the *atps1* mutant were measured as a control. The bars represent mean values of three biological replicates ± standard error. White symbols on the bottom of each bar correspond to amino acid substitution specific for this accession. Asterisks indicate an accession with sulfate concentration higher than Col-0 and ATPS activity comparable with *atps1* mutant.

Recently, Herrmann et al. (2014) determined the crystal structure of the ATP sulfurylase isoform 1 from soybean. Since there is 79% amino acid sequence identity between the ATPS1 in arabidopsis (AtATPS1) and soybean (GmATPS), it was possible to map the amino acid substitutions found in arabidopsis accessions onto the GmATPS structure (see Herrmann et al. (2014), on which I am a co-author). The information about particular amino acid substitutions, their localisation in the protein and possible effect on the enzymatic activity are summarised in Table 3.2. Two of these substitutions: V43N and G56S were localised in the region corresponding to the disordered region of GmATPS and therefore they were not mapped to the structural model (Herrmann et al. 2014).

Table 3.2 ***Summary of variation in arabidopsis ATPS1 haplotypes***

Mutation	Accessions	ATPS1 activity [µmol/min/mg protein]	Localisation	Comment
V43N	Dog-4	0.140	transit peptide	mutation not mapped to *A. thaliana* ATPS1
	Vash-1	0.156		
G56S	Dra2-1	0.125	n.a	
	Fjae 1-2	0.093		mutation not mapped to *A. thaliana* ATPS1
	Ull 2-3	0.121		
L122V	Vie-0	0.108	dimer interface	L122 very conserved among species; may cause alteration of oligomerization and/or disruption of protein folding
T150S	Dra-0	0.125	loop region	mutations in loop region usually don't alter enzyme function
E169A	Goettingen-22	0.111	dimer interface	may cause alteriation of oligomerization and/or disruption of protein folding
	Goettingen-7	0.135		
V316F	T800	0.247	internal packing	may alter protein structure
A337S	Bil-7	0.145	loop region	
	Fri-2	0.115		mutations in loop region usually don't alter enzyme function
	Tgr-01	0.119		
G342D	Naes-2	0.072	internal packing	may alter protein structure

The information in comments is taken from Herrmann et al. (2014). The ATPS1 activity for accessions listed is the same as shown on Figure 3.7.

3.4 Discussion

Most agronomically important traits including control of plant elemental composition and nutrient use efficiency are tightly regulated by various genes and regulatory factors as well as the products and intermediates of the same or interacting metabolic pathways (White et al. 2014). The control of sulfate homeostasis has a crucial role for plant fitness and development (Koprivova et al. 2014, White & Brown 2010). It is also of great importance from human perspective since we utilize plants for food. Investigation of the genetic basis of plant mineral composition is a first critical step to understand the regulatory networks and improve crop yield and quality.

Sulfate homeostasis depends mainly on the balance between sulfate uptake and reduction and is controlled in a demand-driven manner (Lappartient & Touraine 1996). When plants experience sulfate starvation, a simultaneous increase in mRNA accumulation is observed for sulfate transporters, ATPS and APR (Hopkins et al. 2004). On the other hand, increase in reduced sulfur compounds such as cysteine or glutathione causes a decreased accumulation of ATPS mRNA and protein as well as enzymatic activity (Bolchi et al. 1999, Vauclare et al. 2002). The demand-driven control mechanisms appear on different steps of the pathway and involve regulation of specific enzymes (see Chapter 1). Moreover, the whole pathway may be controlled as a process e.g. in response to stress (Bick et al. 2001, Koprivova et al. 2008), via phytohormones (Harada et al. 2000, Jost et al. 2005), or within the interconnection with nitrogen and carbon metabolism (Kopriva et al. 2002, Koprivova et al. 2000).

The main research on genetic control of sulfate assimilation pathway so far has been conducted on the Bay-0 x Sha RIL population and led to the identification of two major QTLs responsible for sulfate concentration in arabidopsis (Loudet et al. 2007). One of these QTLs was underlined by the APR2 isoform of the key enzyme of sulfate assimilation pathway, APS reductase. Biochemical analysis revealed that a single nucleotide polymorphism led to an amino acid substitution in $APR2^{Sha}$ resulting in over 99% loss of enzyme activity (Loudet et al. 2007). Similarly, an amino acid substitution in *APR2* in the Hod accession, described in this chapter and identified independently from QTL analysis, led to a severe reduction in APR activity (Figure 3.1A). The transgenic complementation of Hod lines with the $APR2^{Col-0}$ allele restored APR activity confirming further that it is disruption in *APR2* which affects sulfate accumulation in this accession. The analysis of leaf sulfate concentration in the *apr* mutants revealed a link between very low APR activity and sulfate concentration (Figure 3.1). The

reduction in APR activity of about 80% compared to Col-0 resulted in 50% increase in leaf sulfate concentration in the loss-of-function *apr2-1* mutant. In the *apr1* mutant the APR activity was only slightly lower than in Col-0. Additionally, sulfate concentration in this mutant did not differ from that observed in Col-0 which is in agreement with previous results of Loudet et al. (2007) indicating that the APR1 isoform contributes much less to the total APR activity than APR2.

As described previously (Mugford et al. 2011, Scheerer et al. 2010, Vauclare et al. 2002) the disruption of APR activity reduces the sulfate flux through the pathway into reduced organic forms (such as cysteine and glutathione) leading to accumulation of the pathway substrate – sulfate. The analysis described in this chapter revealed strong positive correlation between APR activity and sulfate flux through the reduction pathway (measured as the percentage of radioactive sulfate incorporated into thiols and proteins) providing evidence that supports this model (Figure 3.2). It was true for the Hod accession, analysis of which is described here, but not for Sha which despite possessing the weak *APR2* and reduced flux, accumulated sulfate only under nitrate limitation (Loudet et al. 2007). Under normal nitrate supply, sulfate concentration in this accession was comparable with that observed in Col-0. These results suggest the existence of additional mechanism(s) of regulation of sulfate assimilation beyond APR. In support of this hypothesis, APR activity in Hod lines complemented with $APR2^{Col-0}$ was significantly higher than in Col-0, but there was no difference in sulfate concentration compared to Col-0.

The accumulation of sulfate due to disruption of the first steps of the pathway may lead to a decrease in concentration of the final products of the pathway: cysteine and glutathione (GSH). Moreover, evidence was provided that GSH, rather than cysteine, is the thiol compound used as a signal in this process (Lappartient & Touraine 1996, Lappartient et al. 1999, Vauclare et al. 2002). Decreased concentration of the products of the pathway can result in an induction of sulfate starvation response in plants leading to an increase in sulfate uptake or elevated *APR2* transcript abundance. Indeed, it has been shown that the adjustment of sulfate uptake and assimilation to the plant's demand for sulfur is mediated by reduced sulfur containing compounds such as cysteine and glutathione (Datko & Mudd 1984, Vauclare et al. 2002). However, there was no difference in the *APR2* expression between Col-0 and Hod (Chao et al. 2014). Moreover, there was no difference in sulfate uptake in the *apr* mutants compared to Col-0 (Figure 3.3A). In contrast, Hod and Sha showed the lowest sulfate uptake among all lines tested, confirming further the existence of some additional regulatory mechanisms. These results indicate that sulfate accumulation in the accessions with low APR activity is

uncoupled from the control of sulfate uptake. This is in contrast to low sulfate accumulation phenotype which is known to induce the sulfate starvation response leading to the induction of sulfate uptake (Lee et al. 2012, Matthewman et al. 2012). The most likely explanation is that in the accessions with high sulfate concentration it accumulates in the vacuole and is not available as a signal (Koprivova et al. 2013).

Based on the results described in this chapter it might be hypothesised that the APR1 isoform of APS reductase, in addition to APR2, might exercise significant control over sulfate flux through the reduction pathway. Indeed, it was shown previously that the activity of APR1 is modulated by the accumulation of the oxidised form of GSH (Bick et al. 2001). Moreover, Loudet et al. (2007) showed that reduced APR activity in Sha results in lower accumulation of glutathione in this accession compared to Bay-0 which possess a stronger *APR2* allele. The same was observed for Sha and Hod in the analysis described in this chapter when compared to Col-0. Since GSH synthesis depends on cysteine availability (Galant et al. 2011, Strohm et al. 1995), the lower GSH concentration in Sha and Hod compared to Col-0 indicates an indirect effect of reduction in APR activity. Despite the reduced APR activity and sulfate flux in the *apr2-1* mutant the GSH concentration was comparable with this in Col-0, even though the unused substrate of the pathway – sulfate – was accumulated (Figure 3.3). However, the *apr1apr2* double mutant showed significantly lower GSH concentration compared to Col-0. Furthermore, the $APR2^{Hod}$ allele caused sulfate accumulation, especially in mature plants, whereas in Sha sulfate was accumulated only under nitrate deficiency, but both of these accessions showed significantly low GSH accumulation. According to the 1001 genomes database both Bay-0 and Sha have two deletions in the third exon of *APR1* (data not shown), one of which is localised in slightly different place in the two accessions. However, neither of these deletions was observed in Hod. This observation further supports the hypothesis of APR1 being involved in the regulation of sulfate flux through the metabolic pathway. Moreover it indicates that the natural variation in *APR1* might affect the GSH accumulation. However, this hypothesis requires further investigation.

Taken together, the results described here indicate that the control of thiol accumulation is partly uncoupled from the control of sulfate reduction. It also seems that GSH accumulation is well maintained by balancing synthesis with utilisation since its concentration in the *apr2* mutant did not differ from this in Col-0 despite dramatically low APR activity and sulfate flux in this mutant. The decreased GSH concentration in the double mutant may suggest that either APR1 is involved in the control of GSH accumulation or the disruption of total APR activity in the double mutant is too severe to maintain the appropriate concentration

of this compound. These findings might be useful for developing new strategies to increase its concentration in plants and thus improve stress tolerance and detoxification capacity and provide a better understanding of the complex regulation of the sulfate assimilation pathway.

The analysis of the *APR2* sequence among the worldwide population of *Arabidopsis thaliana* revealed additional rare alleles with a strong effect on APR activity (Table 3.1, Figure 3.4). Chao et al. (2014) have found four accessions sharing the same amino acid substitution in a conserved PAPS reductase domain in the APR2 enzyme. All of these accessions were collected in western Sweden around the Härnösand. It might be suggested that this amino acid substitution is a result of a recent evolutionary adaptation to the local environment. Since sulfate assimilation is an energy consuming process, there may be a fitness advantage for plants with decreased APR2 activity and lower assimilation, growing in soils with elevated sulfate concentration. Such local adaptations that maximized plant fitness were previously described in arabidopsis, usually in relation to flowering time (Brachi et al. 2012, Li et al. 2010). However, in this case more work testing more accessions would be required to establish if the rare alleles of APR2 are a result of local adaptation.

It has been demonstrated that the variation in the sequence of enzymes involved in sulfate reduction is reflected in the natural variation in sulfate accumulation in different arabidopsis accessions (Koprivova et al. 2013, Loudet et al. 2007). However, this correlation is not always straightforward and the same amino acid substitution may cause different phenotypes. The four Swedish accessions are an excellent example of such a difference. All of them were collected in the same location and share the same amino acid substitution leading to dramatically low APR activity. However, Tfa-08, accumulates significantly less sulfate than the other three Swedish accessions, but still more than Col-0. Therefore it seems that Tfa-08 contains additional loci that limit sulfate accumulation. Making crosses between these accessions might be a good way to discover additional genes involved in the control of sulfur homeostasis in plants. A similar phenomenon was observed for ICE61. It had only 30% lower APR activity than Col-0, however sulfate concentration was comparable with this observed in the three Swedish accessions which lost more than 80% of APR activity (Figure 3.4). A very interesting phenotype was identified in Qar-8a which had elevated APR activity and sulfate accumulation. Further investigation of this phenotype could bring new insights into regulatory mechanisms of sulfate reduction. This could be useful for developing new lines used in agriculture as well as new breeding strategies.

Taken together, there is a strong link between decreased APR activity and increased sulfate concentration (relative to Col-0) in the leaves. However, results described here showed that the two traits are not absolutely correlated with each other. The reduction in APR activity did not always affect sulfur accumulation (e.g. ICE61) and the accumulation of sulfate in leaves was not always driven by low APR activity (e.g. LAC-5). On the other hand all the accessions with severely disrupted APR activity accumulated sulfate.

The analysis of the purified recombinant APR2 protein heterologously expressed in *E. coli* revealed that the extremely low APR activity in Sha was caused by a decrease in the efficiency of interaction of the protein with its electron donor (Loudet et al. 2007). The K_M (Michaelis constant) for glutathione (GSH) as a substrate was much higher for $APR2^{Sha}$ than $APR2^{Bay\text{-}0}$. The K_M was the same in the two accessions for APS as a substrate. The newly identified, extremely rare alleles of *APR2* provided an opportunity to increase our understanding of the correlation between structure of the protein and its function. The polymorphism in $APR2^{Hod}$ allele severely reduces the enzymatic activity of the protein. Based on current knowledge the Gly/Arg polymorphism found in the $APR2^{Hod}$ protein is a part of conserved Arg-loop that interacts with the APS substrate. However, binding affinity for the APS as a substrate is not affected in this protein (Chao et al. 2014). Therefore it might be proposed that Gly^{216} might rather be involved in the catalytic mechanism. On the other hand the Phe^{265}/Ser^{265} found in $APR^{Lov\text{-}5}$ and the other three Swedish accessions is localised in helix alpha7 of the PAPS reductase domain. According to the literature, this region does not have any predicted catalytic function (Chartron et al. 2006). Nevertheless, protein sequence analysis among 855 accessions examined revealed that this residue is conserved as Phe or Tyr across the global *A. thaliana* population (Chao et al. 2014). Its substitution to Ser results in over 1000-fold reduction in the catalytic capacity of the enzyme. Therefore, it has been suggested that this amino acid substitution might have a more global effect on the enzyme structure such as maintaining the tertiary or quaternary structure of the protein (Chao et al. 2014).

As mentioned before the analysis of the Bay-0xSha RIL population led to the identification of two major QTLs for sulfate accumulation in this population (Loudet et al. 2007). The *APR2* gene encoding an isoform of APR was identified as underling the first QTL. Further analysis of the natural variation in this gene and its contribution to the natural variation in sulfate accumulation among arabidopsis accessions was described in the first part of this chapter. Recently, *ATPS1* was identified as the gene underlying the second major QTL (Koprivova et al. 2013) and my contribution to the investigation of the variation in this gene

and its role in the control of sulfate accumulation is described in the second part of this chapter.

It was previously thought that APR possesses much stronger control over sulfate flux through the pathway than ATPS (Scheerer et al. 2010, Vauclare et al. 2002). Therefore, it was expected that the Bay-0 allele of *ATPS1* will have a greater effect on the genotypes with the Sha allele of *APR2*. However, Koprivova et al. (2013) revealed that the two enzymes act independently and they have additive effects on sulfate accumulation, indicating the important contribution of ATPS to the control of sulfate flux through the pathway (Koprivova et al. 2013). While the allelic variation in *APR2* alters the enzymatic capacity of the protein the variation in *ATPS1* leads to decrease in transcript abundance. Moreover, the results described in this chapter indicate that the main cause of low transcript abundance among these accessions is the presence of the deletion in the first intron of *ATPS1* gene sequence. These two deletions were first identified as the main cause of the sulfate accumulation in Bay-0 (Koprivova et al. 2013). Subsequently, a number of accessions that shared similar variants of *ATPS1* as the one found in Bay-0 was identified (Figure 3.6). In general, accessions with the two deletions showed lower *ATPS1* transcript abundance and increased accumulation of sulfate. It seems that there is a direct correlation between *ATPS1* transcript abundance and ATPS activity (Koprivova et al. 2013). Additionally, these results confirm that the variation in sulfate concentration in different accessions is at least partially due to variation in ATPS1.

Disruption in ATPS1 causes a decrease in sulfate flux through the metabolic pathway and accumulation of sulfate (Kawashima et al. 2011, Koprivova et al. 2013). However, the analysis of ATPS activity and sulfate concentration in accessions with amino acid substitutions in the ATPS1 protein sequence did not reveal a specific link between ATPS enzymatic activity and sulfate concentration (Figure 3.7). The only accession which showed elevated sulfate concentration and decreased ATPS activity was Naes-1. The single nucleotide polymorphism in this accession resulted in an amino acid substitution of Gly with Asp in position 342. Asparagine has a different size and charge of side chain than glycine and it can change local amino acid interactions. This might be detrimental for the function of the protein (Herrmann et al. 2014). Localisation of amino acid substitutions that affect ATPS1 and sulfur metabolism may provide new understanding of the nature of the relationship between enzyme structure and its function, as was already shown in the analysis of APR2. However, in the case of ATPS1 the effect of variation in the sequence on enzymatic activity was not that pronounced. Therefore, it might be necessary to investigate the *ATPS1* sequence and sulfate concentration in a greater number of accessions to get better understanding of this phenomenon.

3.5 Conclusions

The two enzymes described in this chapter are two consecutive steps of the same pathway and were identified in the QTL analysis of the same RIL population (Koprivova et al. 2013, Loudet et al. 2007). The work described in this chapter has revealed that limitation in the catalytic ability of two sequential sulphate assimilatory enzymes is an important source for the variation in sulfate accumulation among *Arabidopsis thaliana* accessions. However, the mechanism by which the two enzymes alter sulfate concentration is different. Reduced catalytic activity of APR2 results in a reduction of the flux through the sulfate assimilation pathway. Consequently, unused sulfate accumulates in the cells. The reduction in ATPS1 activity is caused by low transcript accumulation of the main isoform of the enzyme. Low transcript accumulation is associated with the deletion in the first intron of the *ATPS1* gene. This deletion also leads to sulfate accumulation. Results described here and by Koprivova et al. (2013) indicate that both of these enzymes make an important contribution to the control of sulfate flux through the metabolic pathway. This analysis complements our knowledge about the regulation of sulfate assimilation pathway which, until now, was thought to be controlled almost exclusively by APR (Vauclare et al. 2002).

Both of the accessions used to create the RIL population (Sha and Bay-0) possess changes in one of the enzymes that contribute to sulfate accumulation: Sha possesses the disrupted *APR2* allele and Bay-0 – *APTS1*. Therefore, it could be expected that the accessions with a high APR activity and high sulfate concentration will carry a reduced function *ATPS1* allele, but this is not the case. An excellent example of such an accession used in these experiments is Qar-8a. It accumulates sulfate even though it has the highest APR activity from all the accessions tested in the experiment shown on Figure 3.4. However, it has no specific changes (insertion/deletions or non-synonymous SNPs) in the *ATPS1* sequence. This indicates that the control of sulfate flux is much more complex than expected. It seems that the regulation of sulfate homeostasis involves a much broader range of factors than the two enzymes characterised in this chapter. Taking into account the importance of sulfate homeostasis for plant fitness, more research is needed to further investigate what these components may be. A contribution to the discovery of new important genes involved in sulfate homeostasis as well as sulfate and nitrate use efficiency is described in the next chapters of this dissertation.

Chapter 4:

Natural Variation in Arabidopsis Response to Sulfate and/or Nitrate Availability

4.1 Introduction

Because sulfur deficiency in Europe appeared relatively recently, research on sulfur use efficiency still lags behind that on the other major nutrients (Abdallach et al. 2010). Exploring sulfur assimilation and the regulation of sulfur metabolism is of great interest for agriculture and plant science. Therefore, a series of experiments on the collection of natural *Arabidopsis thaliana* accessions was conducted to characterise plant response to sulfate and nitrate deficiency. General arabidopsis response profile to sulfate limitation was described. Four different patterns of plants response to sulfate limitation were characterised and the accessions which were best adapted to the analysed conditions were identified. In this section I briefly introduce the main problems addressed in the course of these studies.

The major goal of this project was to get better understanding of the processes involved in plant response to sulfate deficiency. Unbalanced plant mineral nutrients in soil are a major limiting factor for crop growth and development (Blake-Kalff et al. 2001). The need to feed a rapidly growing population puts extra pressure on limited land resources and is the main justification for such studies. Breeding high yielding varieties resulted in a shortage of important plant mineral nutrients in the soil around the globe (Mba et al. 2012). Sulfur deficiency in agricultural soils has been reported frequently in recent years (Ahmad et al. 2005, Blake-Kalff et al. 2001, Scherer 2001). It was also suggested that the demand for sulfur in many crops has increased due to intensive agriculture and optimization during plant breeding programmes (Abdallah et al. 2010). Crop requirements for sulfur vary in different species. Generally, wheat requires approximately 2-3 kg of sulfur for each tonne of grain produced (Zhao et al. 1999) whereas the production of one tonne of oilseed rape seeds requires about 16 kg of sulfur (McGrath & Zhao 1996). It was suggested that this high sulfur demand in oilseed rape is due to an accumulation of large amounts of sulfate in the vacuole under sulfate sufficient conditions. However, this vacuolar pool of sulfate is not easily available for redistribution during sulfate limitation which results in an early induction of sulfate starvation response (Blake-Kalff et al. 1998). Therefore oilseed rape is particularly sensitive to sulfur deficiency or limitation which reduces both seed quality and yield (Malhi et al. 2007). Sulfur deficiency in crop plants has been recognized as a limiting factor not only for crop growth and seed yield, but also for grain quality which is particularly important for wheat breeding and the maintenance of baking quality (Shahsavani & Gholami 2008).

The symptoms of sulfur deficiency appear shortly after it occurs in the soil, especially in oilseed rape (McGrath & Zhao 1996). In general, sulfur deficiency results in uniform pale

green chlorosis through the plant (Nikiforova et al. 2003). A considerable reduction in growth may be suffered without the appearance of any other visible symptoms. Clear symptoms include severe stunting, reduced leaf size and activity of axillary buds which results in less branching. Physiologically, in wheat plants sulfur deficiency affects CO_2 assimilation rates and Rubisco enzyme activity as well as protein abundance which results in general inhibition of de novo synthesis of the photosynthetic apparatus (Gilbert et al. 1997). Additionally, depression of root hydraulic conductivity was observed in sulfur deficient barley plants. It was suggested that this response can have a role in signalling nutrient starvation from shoots to roots (Karmoker et al. 1991).

In the course of this study the interaction between sulfate, phosphate, and nitrate metabolism was characterised. Nitrogen deficiency often occurs in the field (Rossato et al. 2001). It is the most limiting resource in many areas of the world. Nitrogen availability can fluctuate greatly due to precipitation or changes in temperature, wind, soil type and pH (Masclaux-Daubresse et al. 2010). Plants are able to sense the concentration of external and internal nitrogen and adapt to changing nitrogen conditions by modifying different aspects of nitrogen metabolism e.g. gene expression or enzyme activities (Sakakibara et al. 2006). In the first step of the response to nitrogen depletion plants reduce the rate of leaf elongation leading to a decrease in shoot growth rate without affecting photosynthesis (Anandacoomaraswamy et al. 2002). Root elongation is maintained by the direction of additional carbon towards the subterranean part of the plant (Smolders & Merckx 1992). All these reactions together with increased nitrogen remobilisation allow the plant to cope with short periods of nitrogen starvation e.g. in the intervals between supply of fertilizer. However, extended nitrogen deficiency is known to hasten senescence (Gombert et al. 2006) and decrease the photosynthetic capacity of the plant due to Rubisco breakdown, ultimately inhibiting whole plant growth (Walker et al. 2001). Therefore, the plant's ability to lengthen the nitrogen recycling step described above, without affecting photosynthetic capacity is one of the key factors in maintaining growth during temporary nitrogen shortage (Richard-Molard et al. 2008). Chemical nitrogen fertilization is a common agronomic practice to improve crop yield and quality (Kant et al. 2011). However, nitrogen fertilisation is the highest input cost for many crops (Rothstein 2007). Therefore, current research on nitrogen use efficiency (NUE) is directed towards developing genotypes that use nitrogen more efficiently. This is expected to improve plant growth and development while sustaining the quality of the environment (Ikram et al. 2012).

It is long known that nitrogen deficiency results in changes in expression pattern of genes related to sulfate metabolism (Yamaguchi et al. 1999). Koprivova et al. (2000) have shown that in nitrate deficient plants the APR activity in leaves decreases about 30% and in roots about 50%. This is caused by a decrease in APR mRNA and protein accumulation which further confirms that nitrogen availability regulates sulfate assimilation on the transcriptional level. Similarly, sulfate deficiency results in a decrease in nitrate uptake and the activity of nitrate reductase which leads to a decrease in amino acid accumulation (Prosser et al. 2001). In oilseed rape, which is known to have a high demand for sulfate, the growth during short periods of sulfate deficiency is maintained by an optimization of nitrogen uptake and remobilising sulfate from internal reservoirs (Abdallah et al. 2010). The unavailability of sulfate for crops is known to decrease the efficiency with which plants use nitrogen fertilizers (Ceccotti 1995). Therefore, sufficient sulfate availability is necessary to ensure efficient nitrogen use and maintain adequate oil concentration and fatty acid quality of seeds in oilseed brassicas (Fismes et al. 2000). Additionally, positive interaction between sulfur and nitrogen was reported to be beneficial for tolerance to various stress factors (Anjum et al. 2012). Glutathione (GSH) is considered to be the main factor in the intracellular defence against oxidative stress (Gill & Tuteja 2010). Its synthesis depends on the availability of cysteine (provided through assimilation of sulfate) and glutamate (provided through assimilation of nitrate) and therefore might be influenced by the supply of the two macronutrients (Anjum et al. 2012). These observations clearly indicate a strong interconnection between the two metabolic pathways and emphasise the need for further investigation of this coordination in order to understand the regulation of nutrient metabolism in plants and thus improve crop yield and quality.

For the experiments described in this chapter a large number (25) of natural arabidopsis accessions was used to uncover the genetic variation underlying adaptation to environmental gradients. This type of analysis leads to better understanding of plant responses to fluctuating environments and development of agricultural systems that are more resilient to such changes. The genetic diversity of *Arabidopsis thaliana* makes it a suitable model for studying genetic variability of plant adaptation to nutrient deficiency. Indeed, arabidopsis has a broad geographical distribution and as such is subjected to diverse nutritional environments (Montesinos et al. 2009, Pico et al. 2008, Weigel 2012). Therefore, it is a useful model for studying possible contrasting adaptation to limited nutrient availability. A number of studies have been conducted that use natural variation within *Arabidopsis* species to investigate the genetic diversity within the genus. The analysis of a Bay-0 and Shahdara RIL population

revealed a number of loci associated with nitrate homeostasis (Loudet et al. 2003) and key regulatory factors of sulfate metabolism (Loudet et al. 2007). The same population was used to investigate natural variation in adaptation to growth under low nitrogen supply (North et al. 2009, Richard-Molard et al. 2008). However, these studies used a limited number of accessions that were characterised for only a few selected traits. New genomic information has made it possible to generate collections of accessions with a maximum possible genetic diversity and minimum of repetitiveness (Cao et al. 2011, McKhann et al. 2004). These collections are extensively used to uncover the natural variation in plant response to environmental changes (Chevalier et al. 2003, De Pessemier et al. 2013, Ikram et al. 2012, Kellermeier et al. 2013) and investigate the best adapted individuals.

No studies on the extent of diversity of arabidopsis growth responses to sulfate limitation and starvation have been published. Therefore, to find accessions showing contrasting responses to different sulfur conditions a detailed analysis of the response of 25 arabidopsis accessions originating from various areas of the world was conducted (McKhann et al. 2004). These accessions were characterised for seven morphological and ten metabolic traits in normal, limited, and sulfate-starved conditions. It was possible to distinguish four different classes of accessions depending on the pattern of plant response to different sulfur supply. In general, this part of the experiment was inspired by and based on the analysis of the response of the same collection of accessions to nitrate availability published recently (Ikram et al. 2012). Therefore, there is a high level of similarity in plant growth conditions and measured traits as well as statistical analysis and abbreviations used in the two experiments.

Subsequently, one representative accession from each group was selected and analysed in more detail to identify different adaptive strategies within the *Arabidopsis* species. At this stage additional nutrition regimes were added including sulfate/nitrate double limitation to further investigate the interconnection between the two elements. In this part I analysed the plants for additional traits such as sulfate uptake and incorporation into thiols and cysteine. Additionally, primary metabolite profiling was conducted to investigate the global changes in plant metabolism under each nutritional condition.

This experiment was conducted in collaboration with two other BioNut partners (http://bionutitneu2.fatcow.com/; Chapter 7): Fabien Chardon and Giorgiana Chietera from INRA Versailles in France and Rainer Höfgen and Sarah Whitcomb from the Max Planck Institute of Molecular Plant Physiology (MPIMP) in Golm, Germany. Since there is a well-established plant hydroponic culture facility in the institute in Versailles, plants for the first

part of this experiment were grown there. I spent two weeks in the laboratory of Fabien Chardon and I participated in all the stages of setting up the hydroponic cultures and subsequent harvesting together with collection of morphological data. I analysed the samples for metabolic data in Norwich. The statistical analysis of the data was conducted by Fabien Chardon. For the detailed analysis of the four representative accessions of each group I grew the plants on plates and performed all the experiments with radioactive labelling in Norwich. The samples for metabolite profiling were analysed in Golm. I visited the laboratory of Rainer Höfgen for a week and together with Sarah Whitcomb processed the samples. The actual analysis of the samples was performed by the Applied Metabolome Analysis group of Dr Joachim Kopka who provides GC-MS based metabolite profiling services to scientists from MPIMP and to cooperating partner groups. They also provided primary data normalisation and initial statistical analysis. Subsequently, Sarah Whitcomb conducted more sophisticated analysis of significance of delivered metabolites followed by assigning these metabolites onto general plant metabolism.

4.2 Materials and Methods

4.2.1 Plant material

The 24 *Arabidopsis thaliana* accessions used in this study were obtained from the Versailles stock centre (http://dbsgap.versailles.inra.fr/publiclines). The 24 accessions (Akita, Alc-0, Bay-0, Bl-1, Blh-1, Bur-0, Ct-1, Cvi-0, Edi-0, Ge-0, Gre-0, Jea, Kn-0, Mh-1, Mt-0, N13, Oy-0, Pyl-1, Sakata, Sha, St-0, Stw-0, Tsu-0, Ws) belong to the core collection selected by McKhann et al. (2004) and were used in the previous study on nitrate limitation (Ikram et al. 2012). This collection was selected on the basis of genetic variability and it was used in this and the previous study because it maximizes allelic richness. Moreover, Col-0 which is a parental line for most of RIL populations available at the resource centre in Versailles was added to the collection. All these accessions are characterized by similar flowering time in short day conditions (McKhann et al. 2004). Col-0 was added to the collection since it is used widely as a control accession, making it 25 accessions in the initial collection. Additionally, C24 was added in part of experiment concerning the detailed analysis of representative accessions to create a range in the primary concentration of sulfate and nitrate.

4.2.2 Growth conditions in hydroponic cultures

The growth conditions used in this experiment were as described by Ikram et al. (2012). Briefly, seeds were surfaced sterilized using ethanol-"bayrochlor" (BAYROL, Germany; 95/5%, v/v) prior to stratification at 4°C on the top of one Eppendorf tube with the cut base filled with 0.65% agar. Tubes were inserted into 96-well 1000µl tip boxes filled with deionized water. After 3 days at 4°C the boxes containing tubes were transferred to a growth chamber with 8-h-light/16-h-dark cycle and 21°C day and 17°C night temperatures. The photon flux density was 140 µmol m^{-2} s^{-1}. Boxes remained closed by a transparent plastic lid and they were opened after two days. On the seventh day of growth, the tubes with seedlings were transferred to six plastic tanks (two tanks per nutrition condition) having the capacity to grow 104 plants each. Eight plants per nutrition condition (four in each tank), were used for the evaluation of natural variation among the 25 accessions core collection. Samples were composed of two plants resulting in four biological replicates per treatment. Plastic tanks were filled with 33 l of nutrient solution. The plants were cultivated in hydroponics for 35 days (the entire vegetative growth period). Cycles of 8-h-light/16-h-dark were chosen to prevent early flowering. Relative humidity in the growth chamber was 65%. One set of plants was fed on complete nutrient solution (S+N+; referred to as normal condition) containing 0.15 mM $MgSO_4$ as a sulfate source and the second set of plants was fed on limited nutrient solution (S-N+;

referred to as sulfate limitation) containing 0.05mM $MgSO_4$ as a sulfate source for 28 days (Table 4.1M). The third set of plants was supplied with complete nutrient solution for 21 days and subsequently moved for 7 days to nutrient solution with no sulfate source (S0N+; referred to as induced sulfate starvation). The exact media composition is shown in table 4.1M (grey box). Solutions were renewed once during each week of culture up to harvest. Shoots and roots of each plant were separated at the time of harvest (10am–1pm with a daylight period of 9am– 5pm) and weighed. Roots were patted dry with a paper towel before weighing. All samples were frozen in liquid nitrogen and stored at –80°C to preserve biological material and were ground to a fine powder using the steel balls (section 2.2.3) which was then lyophilized before analysis.

Table 4.1M *Media composition*

	Control condition	Sulfate limitation	Sulfate starvation	Nitrate limitation	Sulfate/nitrate double limitation
Elements	**S+N+**	**S-N+**	**S0N+**	**S+N-**	**S-N-**
			[mM]		
KNO_3	3.00	3.00	3.00	1.00	1.00
$CaCl_2$	1.70	1.70	1.70	1.70	1.70
$MgCl_2$	1.85	1.95	2.00	3.45	3.55
$MgSO_4$	0.15	0.05	0.00	0.15	0.05
KH_2PO_4	2.00	2.00	2.00	2.00	2.00
			[µM]		
H_3BO_3	24.00	24.00	24.00	24.00	24.00
MnCl	10.00	10.00	10.00	10.00	10.00
ZnCl	3.00	3.00	3.00	3.00	3.00
KI	0.50	0.50	0.50	0.50	0.50
Na_2MoO_4	0.24	0.24	0.24	0.24	0.24
CuCl	0.90	0.90	0.90	0.90	0.90
$CoCl_2$	0.01	0.01	0.01	0.01	0.01
FeEDTH	22.40	22.40	22.40	22.40	22.40
Na_2EDTA	22.30	22.30	22.30	22.30	22.30
KCl	0.00	0.00	0.00	0.50	0.50
$Ca(NO_3)_2$	0.50	0.50	0.50	0.00	0.00

The composition of media used in all the experiments described in this chapter. Grey boxes correspond to media used in hydroponic cultures; the purple box corresponds to media used for plant growth on plates. All the values for macronutrients are shown in mM and the values for micronutrients are shown in µM. This media composition is based on the media used by Ikram et al. (2012).

4.2.3 Growth conditions on plates

Dry seeds were surface sterilised for up to four hours in a vacuum desiccator using chlorine gas which was generated by mixing 125 ml of sodium hypochlorite with 2.5 ml of 12 M HCl. Seeds were then mixed with 0.1% sterile Low EEO agarose (Sigma Aldrich) and seeded

on plates with specific media. One set of plants was fed on complete nutrient solution containing 3 mM KNO_3 as nitrate source and 0.15 mM $MgSO_4$ as sulfate source (S+N+; referred to as normal condition). A second set of plants was fed on sulfate limited nutrient solution containing 3 mM KNO_3 and 0.05 mM $MgSO_4$ (S-N+; referred to as sulfate limitation). A third set of plants was supplied with complete nutrient solution for 17 days and subsequently moved to the medium with no sulfate source for 4 days: 3 mM KNO_3 and 0 mM $MgSO_4$ (S0N+; referred to as induced sulfate starvation). A fourth set of plants was fed on nitrate limited medium containing 1 mM KNO_3 and 0.15 mM $MgSO_4$ (S+N-; referred to as nitrate limitation). The last set of plants was fed on sulfate/nitrate double limited medium containing 1 mM KNO_3 and 0.05 mM $MgSO_4$ (S-N-; referred to as sulfate/nitrate double limitation). The exact media composition is shown in table 4.1M (purple box). All media were supplied with 0.8% of EEO agarose. Plates were first stored at 4°C in the dark for three days. Subsequently they were transferred to a controlled environmental chamber at 22°C under 16-h-light/8-h-dark cycles where they were grown vertically for 21 days.

4.2.4 Primary metabolite profiling

Metabolite profiling was performed by gas chromatography coupled to electron impact ionization/time-of-flight mass spectrometry (GC-EI/TOF-MS) using an Agilent 6890N24 gas chromatograph (http://www.agilent.com Agilent Technologies, Böblingen, Germany) with split and splitless injection onto a FactorFour VF-5ms capillary column, 30 m length, 0.25 mm inner diameter, and 0.25 µm film thickness (Varian-Agilent Technologies), which was connected to a Pegasus III time-of-flight mass spectrometer (LECO Instrumente GmbH, Mönchengladbach, Germany; http://www.leco.de) as described previously (Dethloff et al. 2014, Erban et al. 2007, Wagner et al. 2003). Briefly, plant material was frozen in the liquid nitrogen immediately after harvesting to stop the metabolism. Frozen tissue was homogenised as described in Chapter 2 using the Genogrinder and aliquoted into 2 ml Eppendorf tubes. The soluble metabolites were extracted by a two-step methanol and chloroform extraction. Briefly, the extraction buffer was pre-mixed: for each 100 µl of ice-cold MeOH (to stop most enzymatic activity), 10 µl nonadecanoic acid methylester "C19" (fresh prepared, dissolved in chloroform: 2 mg ml^{-1}; internal standard for organic phase), and 10 µl sorbitol (from stock, dissolved in MeOH: 2 mg ml^{-1}; internal standard for polar phase) was added. The appropriate amount of pre-mix was added to the sample depending on its fresh weight (Table 4.2M). The sample was then mixed and left at room temperature until entire batch (including the blank samples) was prepared. Subsequently, samples were shaken at 70°C for 15 min (950 rpm in the thermomixer). After that they were cooled down to room temperature and the appropriate

amount of chloroform was added (Table 4.2M). Subsequently, the samples were shaken at 37°C for 5 min (950 rpm in the thermomixer). Afterwards, appropriate amount of water was added (Table 4.2M) and the samples were well mixed and centrifuged at 14,000 g for 5 min to separate the bottom organic phase from the top polar phase. Subsequently two 100 µl aliquots of polar phase and one (if possible) 100 µl aliquot of organic phase were dried in a vacuum concentrator without heat. The pellet was stored at -80°C until further processing.

Table 4.2M ***The composition of extraction mix depending on the sample fresh weight***

Sample weight [mg]	Pre Mix [µl]	Chloroform [µl]	Water [µl]
30 - 45	240	133	267
45 - 55	300	167	333
55 - 65	360	200	400
65 - 75	420	233	467
75 -90	480	267	533
>90	600	333	666

Dried extracts were given to the Applied Metabolome Analysis group of Dr Joachim Kopka for GC-MS based metabolite profiling. Metabolites were methoxyaminated and trimethylsilylated manually prior to GC-EI/TOF-MS analysis (Erban et al. 2007, Fiehn et al. 2000, Lisec et al. 2006, Roessner et al. 2000, Wagner et al. 2003). Retention indices were calibrated by addition of a C_{10}, C_{12}, C_{15}, C_{18}, C_{19}, C_{22}, C_{28}, C_{32}, and C_{36} n-alkane mixture to each sample (Strehmel et al. 2008).

GC-EI/TOF-MS chromatograms were acquired, visually controlled, baseline corrected and exported in NetCDF file format using ChromaTOF software (Version 4.22; LECO, St. Joseph, USA). GC-MS data processing into a standardized numerical data matrix and compound identification were performed using the TagFinder software (Allwood et al. 2009, Luedemann et al. 2008). Compounds were identified by mass spectral and retention time index matching to the reference collection of the Golm metabolome database (GMD, http://gmd.mpimpgolm.mpg.de/; Hummel et al. 2010, Kopka et al. 2005, Schauer et al. 2005) and to (http://www.nist.gov/srd/mslist.htm) – the mass spectra of the NIST08 database. Guidelines for manually supervised metabolite identification were the presence of at least 3 specific mass fragments per compound and a retention index deviation < 1.0% (Strehmel et al. 2008).

All metabolites identified in an experiment were normalized by sample fresh weight and internal standard. For quantification purposes all metabolites were evaluated for best specific, selective and quantitative representation of observed analytes. Laboratory and reagent contaminations were evaluated by non-sample control experiments. Metabolites were routinely assessed by relative changes expressed as response ratios, i.e. x-fold factors in comparison to a control condition or in comparison to the overall median of each metabolite measurement.

4.2.5 Statistical analysis

Statistical analysis described in the first two parts of the chapter (excluding metabolite profiling) was performed by Fabien Chardon as previously described (Ikram et al. 2012). Briefly, ANOVA of phenotypic data was performed using the general linear models (GLM) procedure of SAS software (http://www.sas.com). Phenotypic correlations were calculated for all combinations of traits in each sulfate condition as described previously (Ikram et al. 2012) by using XLSTAT software (http://www.xlstat.com). Significant correlations were visualised by using Cytoscape software (http://www.cytoscape.org). Hierarchical ascendant clustering was performed by using the XLSTAT software according to the Ward method. Differences between accessions and conditions were determined using XLSTAT ANOVA comparisons and Student-Newman-Keuls multiple comparison method.

Statistical analysis of the primary metabolite profiling was performed by Sarah Whitcomb. All the analysis was performed in Microsoft Excel 2010 using the $\log_{10}$-transformed response ratios. The assignment of significantly altered metabolites was based on information from Kyoto Encyclopaedia of Genes and Genomes (KEGG) – a database of manually drawn pathway maps (http://www.genome.jp/kegg/).

4.3 Results

4.3.1 The general arabidopsis response to sulfur availability

In order to characterise general plant responses to sulfate limitation and starvation, 25 arabidopsis accessions were grown in hydroponic culture for 35 days which is the vegetative growth period (section 4.2.2). One set of plants was grown on complete nutrient solution containing 0.15 mM $MgSO_4$ as a source of sulfate and is referred to as the normal or control condition (S+N+). The second set of plants was supplied with 0.05 mM $MgSO_4$ as a source of sulfate and is referred to as sulfate limitation (S-N+). The starvation condition (no source of sulfate; S0N+) in this experiment was induced for the last seven days of plant growth. Initially these plants were grown in normal sulfur conditions. The nutrient solutions were changed once every week in order to maintain constant concentration of elements in the growing medium over the entire growth period. The composition of the growth media was based on the experiment published previously by Ikram et al. (2012).

To characterise developmental variation in response to various sulfur regimes among 25 arabidopsis accessions ten morphological traits that contribute to plant growth were investigated. To quantify the differences in vegetative plant growth between normal conditions, sulfate limitation, and sulfate starvation shoot and root fresh matter (SFM and RFM respectively) and primary root length (PRL) were measured at the time of harvesting. Subsequently, the ratio of RFM and PRL was calculated, which is termed root thickness (RT) and corresponds to the changes in the amount of lateral roots (Ikram et al. 2012). After the lyophilisation of the samples the shoot and root dry matter (SDM and RDM respectively) was measured. The shoot and root water concentration (SH2O and RH2O respectively) was calculated by subtracting dry weight of samples from their fresh weight and expressing it as a percentage of the total fresh weight. The shoot to root fresh and dry matter ratio (SRFM and SRDM respectively) was calculated to obtain the information about biomass allocation.

In order to characterise sulfate metabolism ten metabolic traits were investigated. All the metabolites described in this part of experiment are shown as nmol mg^{-1} DW. I measured shoot and root sulfate concentration (SSO4 and RSO4 respectively) and the concentration of sulfur containing metabolites in both shoots and roots: cysteine (SCys and RCys respectively) and glutathione (SGSH and RGSH respectively) to examine the effect of different sulfur nutrition regimes on those traits. Additionally, I measured shoot and root nitrate concentration (SNO3 and RNO3 respectively) and phosphate concentration (SPO4 and RPO4 respectively) to verify the natural variation and the effect of sulfate availability on those

metabolites. All the metabolic traits were measured from the dry tissue samples using high-performance liquid chromatography (HPLC).

To investigate the general arabidopsis response to normal, limited, and sulfur-starved conditions and identify the most responsive traits in these three conditions the averages of all accessions for each trait in each of the three nutrition regimes tested were computed. The phenotypic profile of arabidopsis response to sulfate limitation and starvation is shown in Figure 4.1 as a percentage of the values obtained in control condition. The significance of arabidopsis response to different nutrition regimes is shown in Figure S4.1.

The analysis of morphological traits revealed that in *Arabidopsis thaliana* sulfate limitation results in significantly higher root biomass (21%), but lower shoot biomass (15%, Figure 4.1). Since there was no change in root length, but root thickness was significantly higher (32%) compared to control condition, it can be concluded that plants produce more lateral roots in sulfate limitation compared to control condition. This is also reflected in low SRDM which indicates an increase in allocation of new biomass to the roots compared to control conditions.

The differences in morphological traits during induced sulfate starvation were not that clearly pronounced. The root biomass did not change compared to control condition and the root thickness slightly increased (15%), but the difference was not as clearly pronounced as in continuous sulfate limitation (Figure 4.1). The decrease in shoot biomass in induced starvation resulted in a decrease in SRDM which again suggests higher allocation of new biomass into roots during sulfate withdrawal. The less pronounced differences may be a result of redistribution of sulfate accumulated in cells when it was sufficiently supplied. However, in this experiment total sulfate concentration was measured. Therefore, at this stage it is not possible to quantify the amount of sulfate from different subcellular pools.

Sulfate limitation resulted in a lower sulfate (16%) and phosphate (32%), but slightly higher nitrate concentration (8%) in roots compared to control condition (Figure 4.1). In contrast, sulfate limitation resulted in higher concentrations of all measured anions in shoots compared to control condition. The concentration of cysteine and GSH was significantly lower in both shoots and roots in sulfate limitation compared to control condition. These results indicate that the response of plants to continuous sulfate limitation differs from the response to induced sulfate starvation.

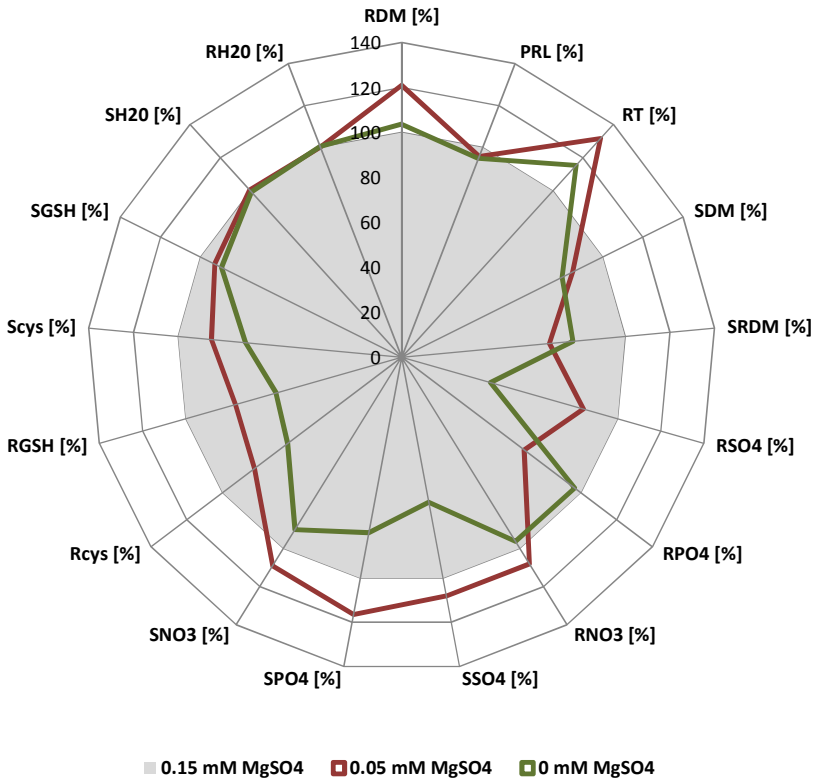


Figure 4.1 ***General arabidopsis response profile to sulfate limitation and starvation***

Phenotypic response profiling of arabidopsis to sulfate limitation (0.05 mM $MgSO_4$ – red line) and starvation (0 mM $MgSO_4$ – green line) shown as a percentage of the value obtained in control condition (0.15 mM $MgSO_4$ – grey circle). The percentage was calculated for both sulfate limitation and starvation from the average of all accessions for each trait. The average value of all accessions in normal sulfate condition was set as 100%. Abbreviations: root dry matter (RDM), shoot dry matter (SDM), shoot to root ratio of dry matter (SRDM), primary root length (PRL), root thickness (RT), shoot water concentration (SH2O), root water concentration (RH2O), shoot nitrate concentration (SNO3), root nitrate concentration (RNO3), shoot phosphate concentration (SPO4) root phosphate concentration (RPO4), shoot sulfate concentration (SSO4), root sulfate concentration (RSO4), shoot cysteine concentration (SCys), root cysteine concentration (RCys), shoot glutathione concentration (SGSH), root glutathione concentration (RGSH); This analysis was performed by Fabien Chardon.

Induced sulfate starvation resulted in a dramatic decrease in sulfate concentration in roots (59%; Figure 4.1). However, root nitrate and phosphate concentration did not differ compared to control condition. The concentration of the three anions decreased significantly in the shoots. Nevertheless, sulfate concentration in the shoots decreased to much lesser extent than in roots (34%). This is not surprising since sulfate reduction occurs mainly in the leaves. Induced sulfate starvation resulted also in a significant decrease in the concentration of both thiols compared to control condition (Figure 4.1). However, this phenotype was much

more pronounced in roots, especially for GSH. The low concentration of thiols in shoots during sulfate starvation is most likely a result of unavailability of pathway substrate - sulfate.

4.3.2 The analysis of natural variation among 25 arabidopsis accessions

In order to determine the proportion of variance explained by genetic and environmental effects, ANalysis Of VAriance (ANOVA) was performed by Fabien Chardon for all the traits studied under normal, limited, and sulfur-starved conditions (Figure 4.2). In this analysis genotype and nutrition were considered as the two main effects with potential genotype x nutrition interaction. The ANOVA analysis revealed that nutrition had a significant effect on all investigated traits except RH2O and the genetic effect was highly significant for all studied traits (Table S4.1). Two traits– SRFM and SH2O – did not show a significant effect of genotype x nutrition interaction (Figure 4.2). The global nutrition effect is shown in Figure S4.1 and the global genetic effect is shown in Figure S4.2. The plant response in different nutrition regimes was significant for majority of traits analysed and the percentage variation explained by each factor which was computed from the sum of squares differed among the studied traits (Table S4.1).

Genotype explains most of the variation in morphological traits (Figure 4.2A). For six out of ten traits measured the percentage variation explained by genotype was over 50%. The root length is the morphological trait with the highest variation due to genotype (73%) and the lowest effect of nutrition (1%) among morphological traits. Additionally, the genotype explained also most of the variation in shoot and root biomass and root thickness (*c.a.* 55%). The variation in biomass ratios was mainly due to nutrition (42% and 26% respectively for SRFM and SRDM). These were also the two morphological traits with the highest nutrition effect. Because both fresh matter and dry matter showed similar patterns of response the fresh matter data were excluded from further analyses in order to avoid unnecessary duplication of the data. Interestingly, the variation in SH2O was mainly explained by genotype (48%), whereas the variation in RH2O was mainly explained by the genotype x nutrition interaction (31%). A significant effect of genotype x nutrition interaction was observed for most of the investigated traits indicating that the effect of sulfate nutrition on plant growth is genotype dependent (Figure 4.2).

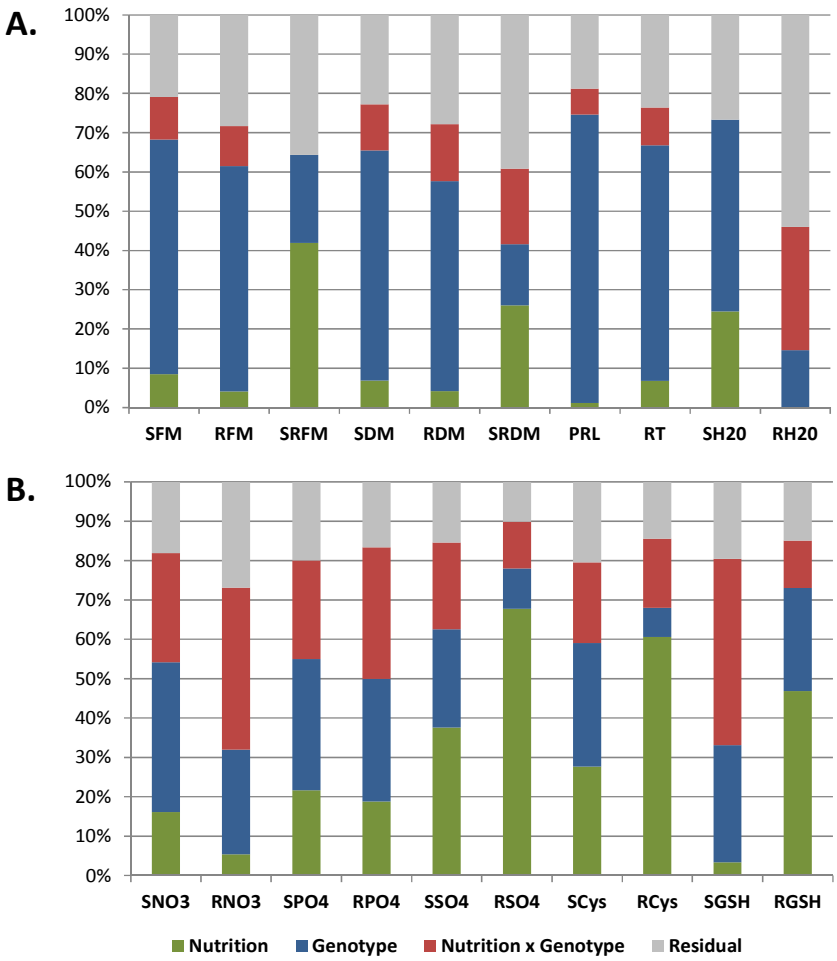


Figure 4.2 ***Analysis of variance (ANOVA) among 25 arabidopsis accessions***

ANOVA of ten morphological traits and 15 metabolic traits among 25 arabidopsis accessions grown under normal sulfate condition, sulfate limitation and sulfate starvation. (A) The morphological traits studied are shoot fresh matter (SFM), root fresh matter (RFM), shoot to root ratio of fresh matter (SRFM), shoot dry matter (SDM), root dry matter (RDM), shoot to root ratio of dry matter (SRDM), primary root length (PRL), root thickness (RT), shoot water concentration (SH2O), and root water concentration (RH2O); (B) The metabolic traits studied are shoot nitrate concentration (SNO3), root nitrate concentration (RNO3), shoot phosphate concentration (SPO4) root phosphate concentration (RPO4), shoot sulfate concentration (SSO4), root sulfate concentration (RSO4), shoot cysteine concentration (Scys), root cysteine concentration (Rcys), shoot glutathione concentration (SGSH), root glutathione concentration (RGSH); Histograms show the effects of nutrition, genotype, and nutrition x genotype interaction as percentages of explained variation. This analysis was performed by Fabien Chardon.

Both genotype and nutrition have a significant effect on all investigated metabolic traits (Figure 4.2B). The percentage of explained variation differs depending on the trait. Nutrition has the highest effect on variation among sulfate related metabolites in roots. Moreover, it is the main source of variation in shoot sulfate concentration (38%), but the

variation in shoot cysteine concentration is mainly due to genotype (31%). Genotype x nutrition effect explains most of the variation in SGSH (47%).

Genotype is the main source of variation for nitrate and phosphate concentration in shoots (38% and 33% respectively). Apart from SGSH genotype x nutrition effect was the main source of variation for nitrate concentration in roots (41%) and phosphate concentration in shoots (33%). A significant effect of genotype x nutrition interaction indicates that the variation in plant response to sulfur nutrition is genotype dependent. The significant variation due to different factors (nutrition, genotype or both) among all measured traits can be used to dissect the genetic variation of plant response to sulfur availability and eventually different adaptive strategies to environmental changes among arabidopsis accessions.

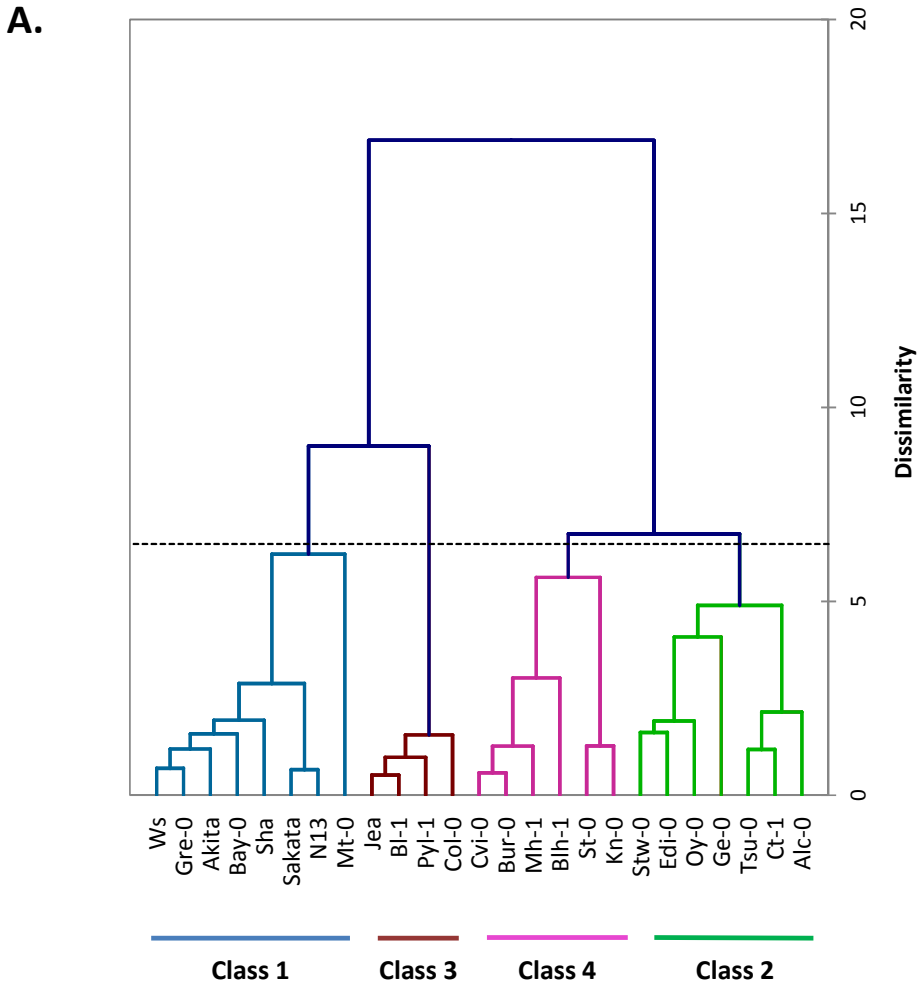
4.3.3 Classification of accessions according to their response to sulfate limitation

The analysis of morphological and metabolic traits among the collection of 25 accessions described above provided a general response profile of *Arabidopsis thaliana* to changing sulfur availability. The analysis of the response of particular accessions to sulfate availability (described in previous sections of this chapter and shown in detail in Figure S4.2) revealed that some of them were following the general patterns of response shown in Figure 4.1. However, some of the accessions showed different response patterns to sulfate deficiency. Therefore, in order to characterise these specific responses a hierarchical ascendant clustering (HAC) was performed by Fabien Chardon (Figure 4.3A). For this analysis he used the values obtained in the two stress conditions shown as a percent of values obtained in control condition as described before (Ikram et al. 2012). Based on this analysis the behaviour of all accessions from the collection was classified into four distinct classes.

Class 1 included Ws-0, Gre-0, Akita, Bay-0, Sha, Sakata, N13, and Mt-0; class 2 included: Stw-0, Edi-0, Oy-0, Ge-0, Tsu-0, Ct-1 and Alc-0; class 3 included: JEA, Bl-1, Pyl-1, Col-0; and class 4 included: Cvi-0, Bur-0, Mh-1, Blh-1, St-0, and Kn-0. The response profile of each class was first characterised based on the plant phenotype under sufficient sulfate supply (Figure 4.3B). The investigation of morphological traits allowed the separation of class 1 and 4 from class 2 and 3. Classes 1 and 4 group relatively small plants (with class 1 including the smallest plants in the collection), whereas classes 2 and 3 comprise significantly bigger plants (with class 2 including the biggest plants from entire collection) in terms of shoot and root biomass as well as root length and thickness. Class 3 consisting of big plants showed the lowest SRDM suggesting a high biomass allocation to the roots in accessions from this class. This

result is in contrast to class 4 containing slightly smaller plants, which has the highest SRDM ratio among all four classes indicating higher biomass allocation to the shoots.

The comparison of class 2 and 3 including morphologically bigger accessions revealed interesting differences in metabolic traits (Figure 4.3B). In general, class 3 which includes slightly smaller plants than those from class 2 showed the highest anion concentration among all classes. On average accessions from class 3 have 26% higher SNO3, 10% higher SSO4, and 23% higher SPO4 compared to accessions from class 2. The differences in anion concentration in roots are slightly less pronounced. Interestingly, the thiol concentration is low in both classes, with the lowest thiol concentration in class 2 among all four classes. Both shoot and root sulfate concentration is lower in class 2 than in class 3. There are substantial differences in phosphate concentration between shoots and roots in the two classes. The accessions from class 2 have more phosphate in roots, whereas accessions from class 3 have more phosphate in the shoots.



B.

MORPHOLOGICAL TRAITS

	Class 1	Class 2	Class 3	Class 4	GR
RDM	0.944	1.121	1.079	0.857	0.994
SDM	0.917	1.112	0.972	0.999	1.000
SRDM	0.999	0.970	0.874	1.157	1.009
RL	0.970	1.003	1.017	1.009	0.996
RT	1.006	1.141	1.036	0.820	1.003
SH20	1.001	0.999	1.001	0.997	1.000
RH20	1.002	1.006	1.000	0.994	1.000

METABOLIC TRAITS

	Class 1	Class 2	Class 3	Class 4	
RNO3	1.052	0.970	1.011	0.967	1.002
SNO3	0.905	0.948	1.192	0.955	0.975
RSO4	0.945	0.937	1.062	1.056	0.988
SSO4	0.840	1.030	1.130	1.000	0.978
RPO4	0.793	1.252	0.805	1.150	1.009
SPO4	0.876	0.918	1.127	1.080	0.977
RCys	1.060	0.994	0.983	0.964	1.006
SCys	1.013	0.953	0.967	1.067	1.002
RGSH	0.912	0.966	1.270	0.853	0.970
SGSH	0.962	0.915	0.966	1.158	0.996

Figure 4.3 ***Classification of the collection into four distinct classes***

Hierarchical ascendant classification of accessions performed based on the different patterns of plant response to sulfate limitation and starvation. (A) Dendrogram of 25 accessions. (B) Heat map of the differences in morphological and metabolic traits between the classes in normal sulfate condition. The values indicate class means shown as relative to average value of all classes for each trait. Additionally, general arabidopsis response (GR) is shown in the panel on right as relative to average of all classes. Colours indicate the differences between classes with the lowest value in blue and the highest in red. GR is color-coded independently from the classes. Abbreviations: root dry matter (RDM), shoot dry matter (SDM), shoot to root ratio of dry matter (SRDM), primary root length (PRL), root thickness (RT), shoot water concentration (SH2O), root water concentration (RH2O), shoot nitrate concentration (SNO3), root nitrate concentration (RNO3), shoot phosphate concentration (SPO4) root phosphate concentration (RPO4), shoot sulfate concentration (SSO4), root sulfate concentration (RSO4), shoot cysteine concentration (SCys), root cysteine concentration (RCys), shoot glutathione concentration (SGSH), root glutathione concentration (RGSH);

Similarly, the investigation of metabolic traits among class 1 and 4 including small plants revealed that class 1 has lower metabolite concentration than class 4 which has morphological phenotype closer to average (Figure 4.3B). Class 1 is characterised by the lowest anion concentration among the classes. The only exception is nitrate concentration in roots which in this class is the highest among all four classes. Thiol concentration in this class is also lower than in class 4. In fact, class 4 is characterised by the lowest root thiol concentration.

Although the diversification of the four classes according to morphological traits imply the grouping of class 2 and 3 including bigger plants and class 1 and 4 including smaller plants, according to metabolic traits class 1 and 2 include accessions with metabolite concentration below the average whereas class 3 and 4 include accessions with metabolite concentration above the average. Additionally, class 3 is characterised by higher anion concentration and class 4 by higher thiol concentration which is the main difference between the two classes. There was no apparent difference in water concentration between the classes (Figure 4.3B).

4.3.4 Analysis of different patterns of response to low sulfur supply

In order to characterise plant response to different sulfur nutrition regimes the main differences between the classes in response to sulfate limitation and induced sulfate starvation were characterised (Figure 4.4). This analysis emphasised higher root biomass production characteristic for class 4 and significantly lower concentration of all metabolites in class 3 in response to sulfate limitation. Apart from class 3 all the other classes responded to sulfate limitation with an increase in anion concentration, especially in shoots. Accessions from all classes responded to sulfate starvation with a significant decrease in anion concentration. These results indicate that arabidopsis responds differently to continuous sulfate limitation and to induced sulfate starvation (Figure 4.4).

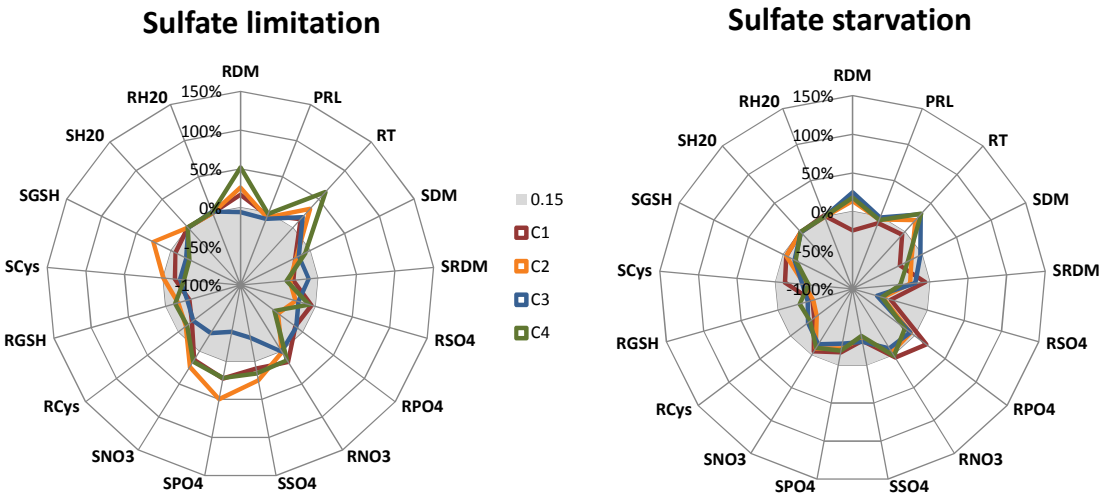


Figure 4.4 ***Characteristic of the four classes in response to sulfate nutrition***

Phenotypic response profiling of four classes is shown as a mean of all accessions belonging to the class. The response of individual traits is shown as a percentage of the value obtained in control condition (grey circle). Red line corresponds to class 1, orange line corresponds to class 2, blue line corresponds to class 3, and green line corresponds to class 4; annotation is in Figure 4.3.

In order to characterise the behaviour of accessions from each class in response to sulfate limitation and induced sulfate starvation detailed response profiles for each class were performed, using the same strategy as in previous work on plant response to nitrate availability (Ikram et al. 2012). For each trait the average of all accessions within the class was calculated. The percentage data are shown as the difference in response to sulfate limitation and/or starvation and control condition (Figure 4.6).

Class 4 is the most similar to the general arabidopsis response profile to sulfate availability shown in Figure 4.1. The heat map shown in the Figure 4.3B indicates clearly pronounced lower concentration of anions and slightly higher concentration of sulfur-containing compounds as a common characteristic of class 4 and the general arabidopsis response profile. They are also characterised by low root thiol concentration and high root phosphate concentration. The similarities between class 4 and general arabidopsis response profile occur not only in normal sulfate condition but also in response to sulfate deficiency. A higher root biomass and amount of lateral roots (expressed as root thickness) is a common response to sulfate limitation for the two groups.

Accessions from class 2 responded to sulfate limitation with a higher root biomass and thickness which was common for all classes except class 3 which did not show a significant difference in root biomass compared to control condition (Figure 4.6). Moreover, class 2 showed higher concentration of all anions tested and GSH in the shoots, whereas class 3 was characterised by a low concentration of all metabolites. The response of both class 2 and 3 to induced sulfate starvation was similar, especially in morphological traits

Accessions from class 4 responded to sulfate limitation with the highest root biomass and root thickness among all the classes (Figure 4.6). This phenotype was not that clearly pronounced in class 1. High shoot concentration of all anions and root nitrate concentration was a common response to sulfate limitation of both class 1 and class 4.

The response of accessions from class 1 and 4 to induced sulfate starvation differed compared to their response to sulfate limitation. Class 1 is the only one among all classes which responded to induced sulfate starvation with a decrease in root biomass and thickness (Figure 4.5). Moreover, it is the only class which showed an increase in root phosphate concentration in response to induced sulfate starvation. The traits not affected in any of the classes by any of the conditions were SH2O, RH2O and PRL suggesting that neither water concentration nor root length play a significant role in response to sulfur deficiency (Figure

4.5). In fact, the variation due to nutrition was not significant in RH2O and very low in PRL (Figure 4.2).

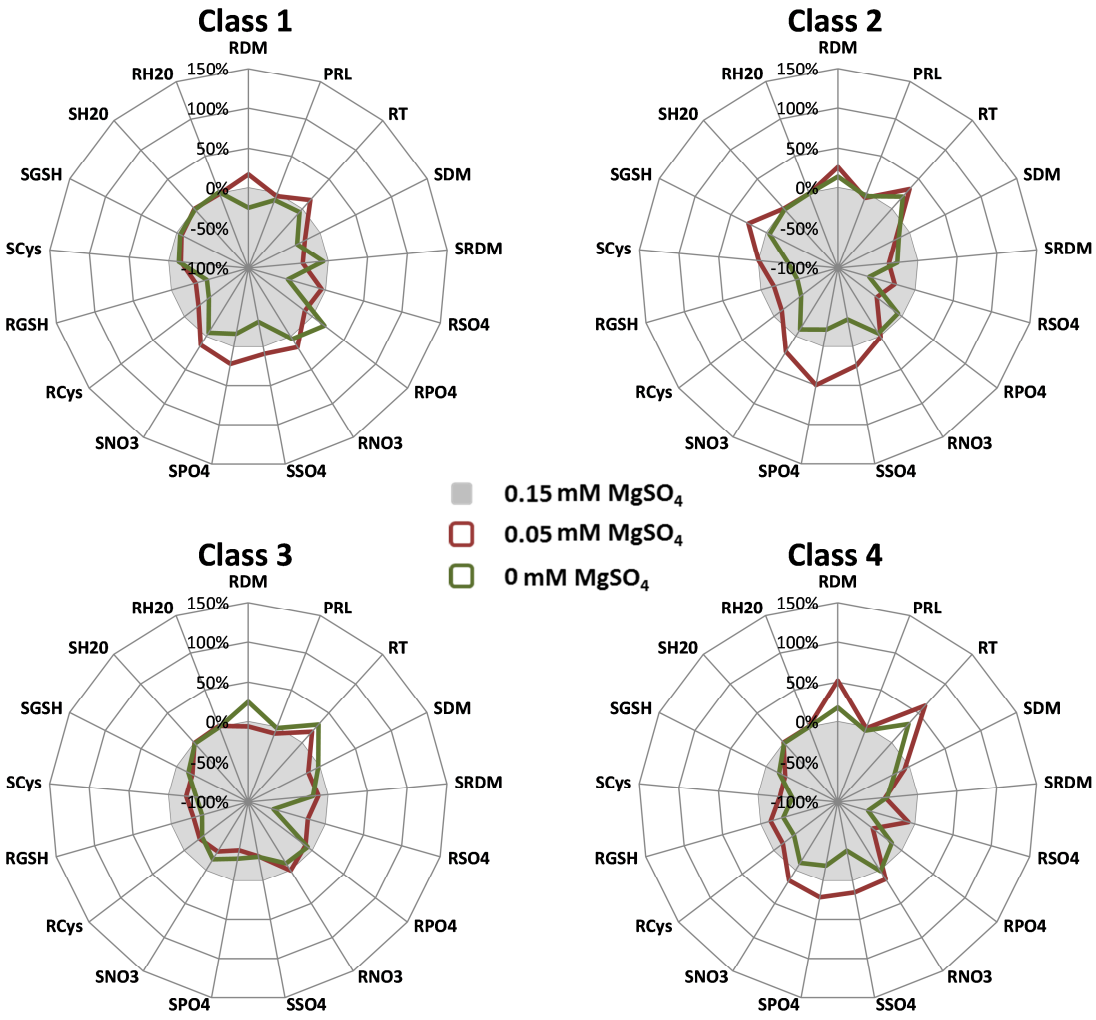


Figure 4.5 ***Characteristic of the four classes***

The response profiles of each class are shown as an average of all accessions from the class. The response of individual traits is shown as a percentage of the value obtained in control condition (grey circle). Red line indicates response to sulfate limitation (0.05 mM $MgSO_4$) and green line to sulfate starvation (0 mM $MgSO_4$). Abbreviations are in Figure 4.3.

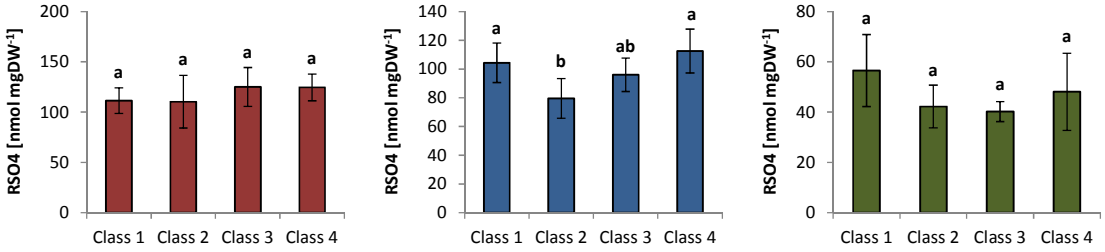
4.3.5 Perturbations in sulfate utilization in sulfate deficient medium

Because sulfate was the only source of sulfur in the growing media, it was possible to analyse its utilisation by the plants in different conditions. To better illustrate this phenomenon I show in Figure 4.6 part of the data from figure 4.5 (the morphological traits as well as accumulation of nitrate and phosphate are excluded) in a way which demonstrates the

statistical significance of the differences between the classes and not plant response to different nutrition regimes as in Figure 4.5.

As shown before, class 3 is characterised by high anion concentration under normal sulfate supply (Figure 4.3B). Indeed, class 3 has the highest shoot sulfate concentration among all classes (Figure 4.6). There were no changes in shoot thiol concentration in this class indicating that increased sulfate concentration does not occur due to disruptions in sulfate assimilation. However, the GSH concentration in roots is significantly elevated in this class. In sulfate limitation class 3 showed the lowest sulfate concentration, however GSH concentration in roots was still high compared to other classes. In induced sulfate starvation this class showed high shoot sulfate concentration and root GSH concentration similarly to what was observed in normal sulfate condition. GSH is known to be involved in long distance transport of small molecules between shoots and roots (Li et al. 2006). Moreover, it has an important role in regulation of the sulfate assimilation pathway and response to sulfate deficiency. GSH is involved in the modulation of root architecture in response to sulfate deficiency. As mentioned in the introduction to this chapter reduction in GSH concentration results in an inhibition in root growth. Therefore, it might be hypothesised that increased GSH concentration in roots in class 3 can be a way to maintain the production of lateral roots – an organ involved in acquiring limited nutrients. In fact, this class showed no changes in root biomass, but an increase in the amount of lateral roots was observed (Figure 4.5).

Under normal sulfate supply accessions from class 4 did not show specific difference in shoots and roots sulfate concentration compared to other classes (Figure 4.6). However, the highest shoot GSH and cysteine concentration was observed in these accessions under sufficient sulfur supply among all the classes. Therefore, it might be suggested that this class has the most effective mechanism of sulfate reduction among the four classes. However, this phenomenon was observed under normal sulfur supply only. There were no significant changes in GSH and cysteine concentration in this class under sulfate limitation or induced starvation.



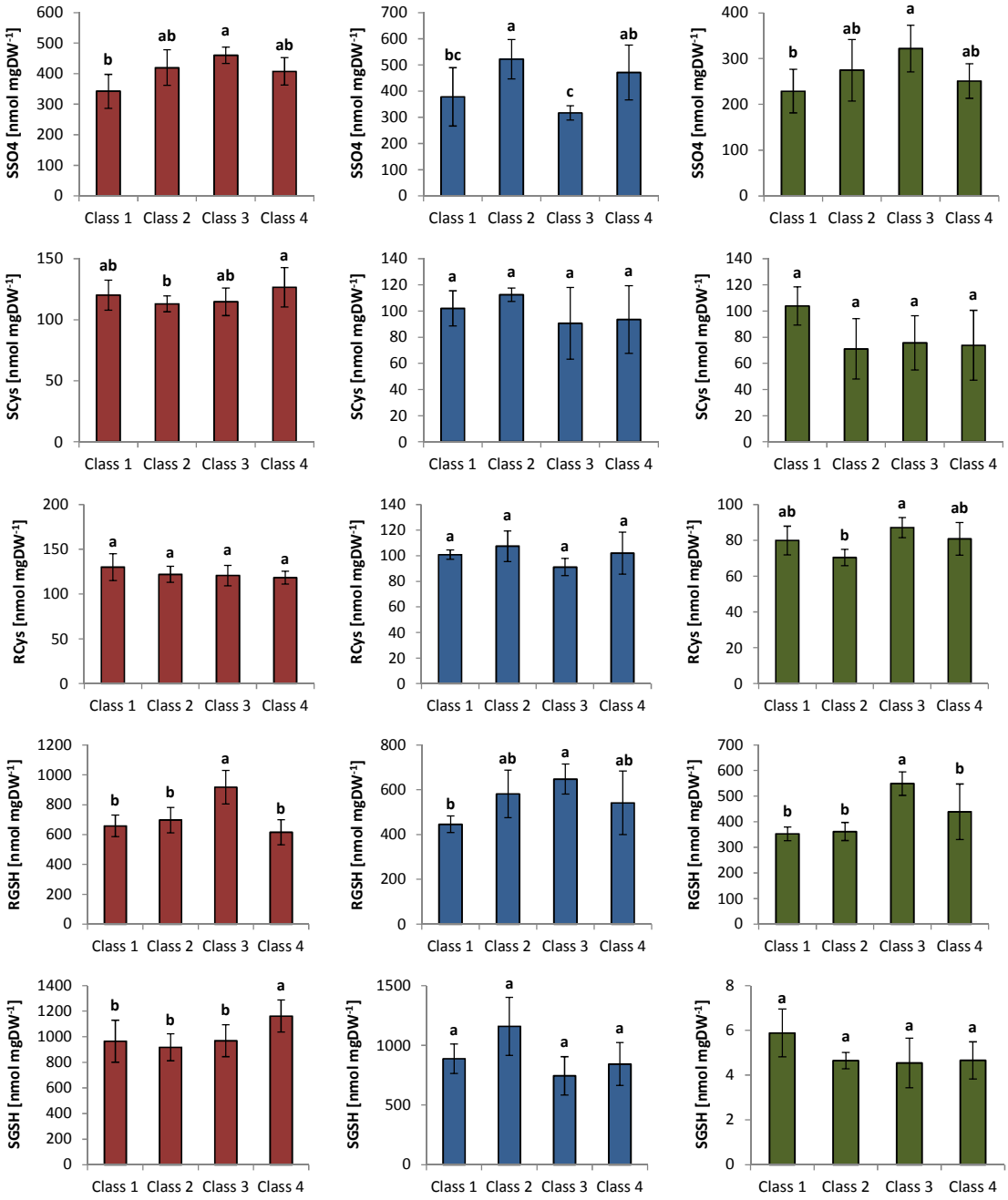


Figure 4.6 ***Changes in sulfate metabolism in response to differences in sulfate supply***

The differences between sulfate related metabolites in the four classes are shown as average values of all accession in a class. The red bars indicate response in normal sulfate condition, the blue bars indicate response to sulfate limitation, and the green bars indicate response to sulfate starvation ± standard deviation. Different letters above the bars indicate values significantly different at P-value<0.05 obtained from ANOVA with Newman-Keuls (SNK) multiple comparison grouping test.

4.4 The more in depth analysis of different patterns of response

Clustering of the 25 arabidopsis accessions revealed four different patterns of behaviour in sulfur limitation and starvation. In order to characterise particular patterns of response to different nutrition regimes one representative accession from each class was selected and analysed in more detail. The representative accessions were selected based on the comparison of individual behaviours of accessions with the average behaviours for the class and the clustering of accessions based on their behaviour in response to nitrate deficiency (Ikram et al. 2012). Sakata was selected as a representative of class 1, Edi-0 as a representative of class 2, Col-0 as a representative of class 3, and Cvi-0 as a representative of class 4. Class 3 which was characterised by the highest anion concentration among all classes in normal sulfate condition was most severely affected by sulfate deficiency (Figure 4.6). This may suggest that the variation in anion concentration in different accessions is dependent on plant demand and plant response to sulfate availability differs depending on the initial concentration of anions. Therefore, to further investigate this hypothesis I added C24 to this reduced set of accessions to create a range in initial concentration of nitrate and sulfate which would allow me to investigate this problem further. C24 was not included in the global analysis of 25 accessions (see section 4.2.1).

As mentioned before, the analysis of the response of 25 arabidopsis accessions to sulfate limitation and starvation was based on previous analysis of the same collection of accessions which provided information about plant response to nitrate limitation (Ikram et al. 2012). Additionally, an importance of interaction between sulfate and nitrate metabolism in response to limitation of one of the two elements was highlighted in the introduction to this chapter. Therefore, it was of interest to characterise the effect of nitrate limitation and the interaction between nitrate and sulfate in more detail. Hence, I added two additional nutrition regimes to this analysis: nitrate limitation with sufficient sulfate and nitrate/sulfate double limitation. This experimental design allowed independent testing of the effect of sulfate status, nitrate status, and sulfate x nitrate interactions on plant growth.

I grew the plants on agarose plates, vertically for 21 days as described in section 4.2.3. Briefly, plants were grown in 16-h-light/8-h-dark cycles. One set of plants was grown on complete nutrient solution containing 0.15 mM $MgSO_4$ as a source of sulfate and 3 mM KNO_3 as a source of nitrate and is referred to as normal or control condition (S+N+). The second set of plants was supplied with 0.05 mM $MgSO_4$ as a source of sulfate and is referred to as sulfate limitation (S-N+). The starvation condition (no source of sulfate) in this experiment was

induced for the last four days of plant growth by transferring plants from the normal condition to new plates with no sulfate (S0N+). Up to this point these plants were grown in normal sulfur condition. Both sulfate limitation and starvation condition media were supplied with 3 mM KNO_3 as a source of nitrate. The nitrate limited plants were supplied with 0.15 mM $MgSO_4$ and 1 mM KNO_3 (S+N-). In double limitation plants were supplied with 0.05 mM $MgSO_4$ and 1 mM KNO_3 (S-N-). In this experiment the media were not changed over the growth period except for plants exposed to sulfate starvation which were moved to new plates for the last four days before harvesting. The composition of growth medium used in this experiment was based on the medium used in hydroponic cultures. In order to adjust the amount of particular nutrients to the requirements of experimental set up I ran a pre-experiment. I grew Col-0 on media with different concentrations of nitrate and sulfate. I measured anion concentration in these plants to make sure that the conditions I used would be sufficient to see the differences in plant response to different nutrition regimes (data not shown).

After 21 days of growth plants were incubated with medium containing 100 µCi [^{35}S] sulphuric acid (specific activity 70,000 dpm / 10 µl) in the light for four hours; only roots were flooded (pH 6; see Chapter 2 section 2.2.18). The experiments with radioactive labelling were conducted essentially as previously described (Kopriva et al. 1999, Koprivova et al. 2000, Vauclare et al. 2002). To investigate the concentration of radioactive sulfate taken up by the plant I determined the radioactivity in plant extract (TU). Additionally, I measured the incorporation of [^{35}S]sulfate into thiols (ITh) and proteins (IP). The concentration of [^{35}S]sulfate detected into thiols and proteins together is referred to as absolute sulfate flux (ASF) through the sulfate assimilation pathway. Additionally, I calculated the percentage of the radioactivity detected in thiols and proteins together (relative to the total radioactivity detected in plant extracts) which is referred to as relative sulfate flux through the assimilation pathway. The ratio of radioactivity detected in the shoots and the radioactivity detected in the entire plant indicates sulfate translocation from the roots to the shoots (TTS). It should be stressed that these traits are expressed as concentrations (Chao et al. 2014, Mugford et al. 2011) and therefore should be interpreted with caution. Because the plant growth conditions in this experiment were different from those in the global analysis of 25 accessions described above I also measured the shoot anion (SNO3, SPO4, SSO4) and thiol (SCys and SGSH) concentration in these plants.

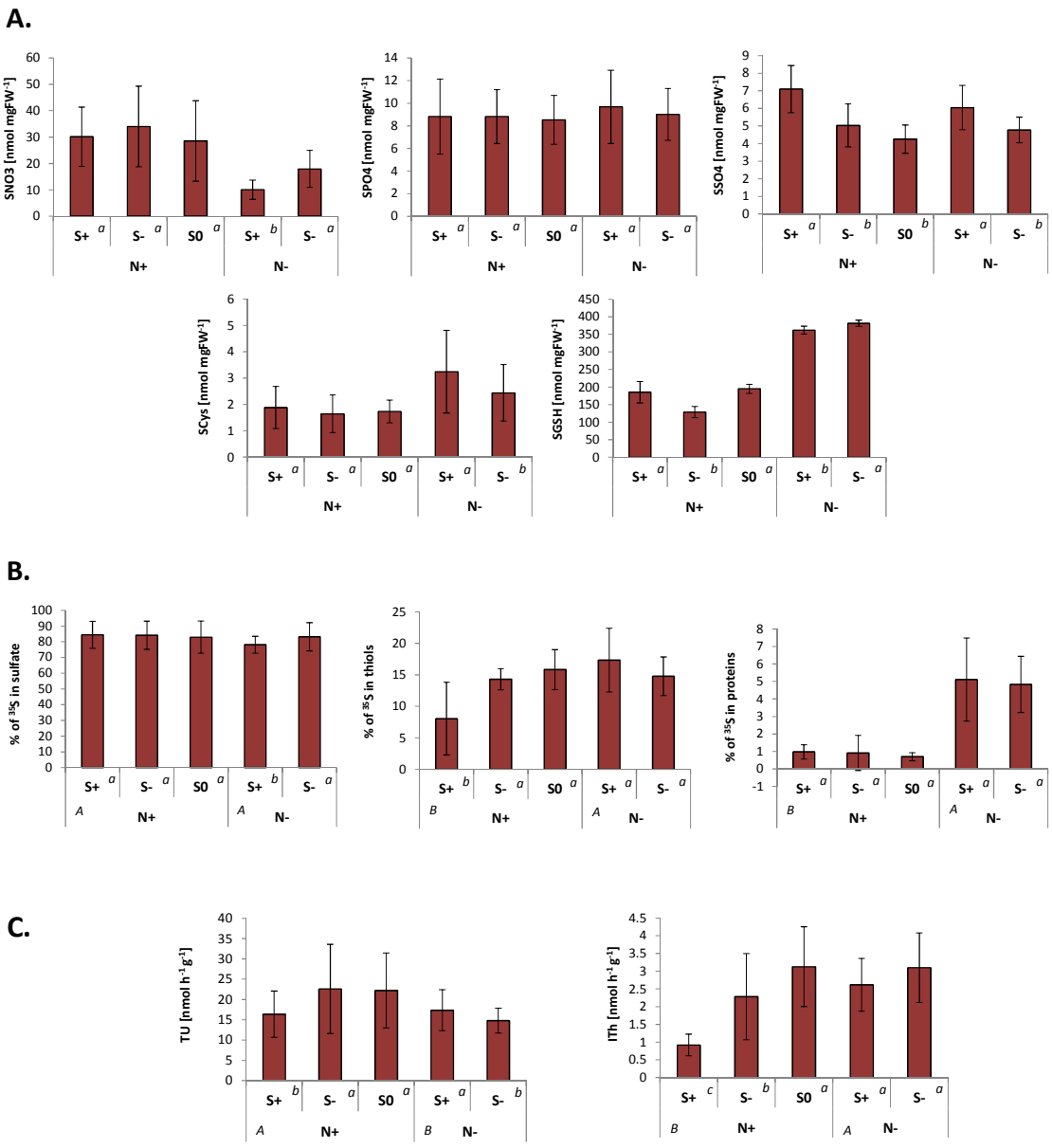
4.4.1 General effect of sulfate and nitrate availability on sulfate metabolism

To characterise the effect of sulfate and nitrate availability on the anion concentration and sulfate metabolism, I calculated the averages of all accessions for each trait in each condition tested (Figure 4.8). The experimental design allowed independent analysis of sulfate limitation, nitrate limitation, and sulfate x nitrate interactions.

Under sufficient nitrate the distribution of radioactive sulfate between sulfate and proteins shown as a percentage of [^{35}S]sulfate taken up by the plant did not change depending on sulfate availability (Figure 4.8B). About 80% of [^{35}S]sulfate remained in the form of sulfate and about 1% was incorporated into proteins. In contrast, the percentage of [^{35}S]sulfate which ended up in thiols was higher in sulfate limitation (14%) and induced sulfate starvation (16%) compared to control condition (8%). The remaining radioactive sulfate was probably incorporated into methionine and its derivatives, and the products of secondary sulfate metabolism which were not measured in this experiment. It is known that sulfate uptake is strongly up regulated in sulfate limitation (Davidian & Kopriva 2010). Indeed, in this experiment sulfate uptake was higher in plants exposed to sulfate limitation and induced starvation (Figure 4.8C). The total sulfate concentration of plants exposed to sulfate limitation and starvation not incubated with [^{35}S]sulfate was lower compared to plants grown under normal sulfate supply. There were no changes in cysteine and only slightly lower GSH concentration in these plants in relation to plants grown under normal sulfate supply (Figure 4.8A). These results suggest that sulfate flux through the reduction pathway increased as a result of sulfate limitation. In fact, there was a higher incorporation of [^{35}S]sulfate into thiols as well as absolute and relative sulfate flux in plants exposed to sulfate limitation (Figure 4.8C). It is also long known that the activity of APR2 – the key enzyme of the sulfate assimilation pathway – is induced by sulfate limitation (Davidian & Kopriva 2010, Lappartient et al. 1999, Lee et al. 2011).

In this experiment sulfate uptake was significantly lower in plants exposed to nitrate limitation compared to plants grown under sufficient nitrate (Figure 4.7C). Moreover, the percentage of [^{35}S]sulfate taken up by the plant which remained as sulfate was also lower under nitrate limitation and sufficient sulfate compared to plants grown under sufficient nitrate (Figure 4.7B). However, the total sulfate concentration in plants not exposed to radioactive sulfate (Figure 4.8A) did not change due to nitrate limitation, but cysteine and glutathione were higher compared to phenotypes observed under sufficient nitrate supply. The percentage of [^{35}S]sulfate taken up by the plant that was converted into proteins reached

about 4% in nitrate limitation compared to 1% under nitrate sufficient condition (Figure 4.8B). This phenotype was common for roots and shoots (Figure 4.7C) and indicates higher assimilation of sulfate under nitrate deficiency. This is not in agreement with the literature as it was shown that nitrate deficiency inhibits sulfate reduction (Koprivova et al. 2000, Lee et al. 2011). However, in these experiments five week old plants grown in hydroponic cultures (Koprivova et al. 2000) or two week old plants grown on plates but with higher sulfate and lower phosphate (compared to media used in the experiment described here) were analysed. Although it is possible that these differences in growth conditions might have caused these discrepancies, more analysis would be required to investigate this issue further.



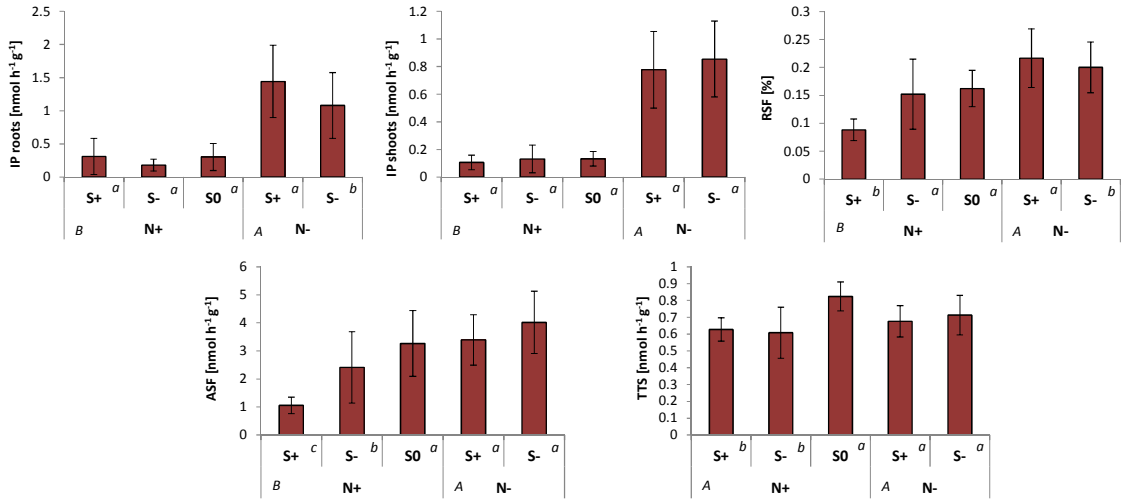


Figure 4.7 ***General effect of sulfate and nitrate limitation on sulfate metabolism***

(A) Concentration of anions and thiols in response to nutrient limitation. (B) The percentage of [^{35}S] taken up by the plant and remained in form of sulfate and incorporated into thiols and proteins. (C) The analysis of sulfate flux and uptake expressed as a concentration of [^{35}S]sulfate detected in plant extracts (TU) and different fractions of plant extracts (IP, ITh etc.). Arabidopsis response to different nutrition regimes is shown as average of all accessions for each trait in normal sulfate condition (S+), sulfate limitation (S-), sulfate starvation (S0) as well as normal nitrate condition (N+) and nitrate limitation (N-).Different letters indicate values significantly different at P-value ≤ 0.05 obtained from ANOVA with Newman-Keuls (SNK) multiple comparison grouping method. Capital letters correspond to differences between nitrate nutrition regimes and small letters correspond to differences between sulfate nutrition regimes – treated separately under sufficient and limited nitrate. The error bars correspond to standard deviation.

4.4.2 The analysis of natural variation among five arabidopsis accessions

To take the analysis described above further the variation in response to different nitrate and sulfate availability among five accessions representing different patterns of behaviour was investigated. ANOVA was run by Fabien Chardon to determine the proportion of variance explained by genetic and environmental effects (Figure 4.8). In this analysis there were three main effects: genotype, nitrate nutrition, and sulfate nutrition; with potential interactions: nitrate nutrition x sulfate nutrition, genotype x nitrate nutrition, and genotype x sulfate nutrition. The ANOVA analysis revealed significant effects of genotype for all the traits except SSO4 (Table S4.3). There was also significant variation due to nitrate nutrition for all the traits analysed and significant effect of sulfate nutrition for most of the traits except SPO4 and IP (Figure 4.8). Only five traits showed significant effect of nitrate x sulfate nutrition interaction: SPO4, SSO4, SGSH, ITh and RSF; six traits showed significant effect of genotype x nitrate nutrition interaction: SNO3, SPO4, SSO4, SCys, SGSH, and TU; and five traits showed significant effect of genotype x sulfate nutrition interaction: SNO3, SPO4, SSO4, SCys, SGSH.

The percentage variation explained by each factor was computed from the sum of squares and differed among studied traits (Table S4.3).

Genotype was the main source of variation for SPO4 (79%), SCys (35%), and TU (17%). Interestingly, for five out of 11 traits measured, nitrate nutrition was the main source of variation (Figure 4.9). It explained most of the variation observed for SGSH concentration (92%) and sulfate incorporation into proteins (63%). Moreover, it accounted for 42% of the variation observed for nitrate concentration, 41% of the variation in relative sulfate flux and 35% of the variation in absolute sulfate flux. Sulfate nutrition was the main source of variation among three traits closely related to sulfate metabolism: 45% for SSO4, 38% for TTS and 33% for ITh. The two traits with the highest effect of nitrate x sulfate nutrition interaction are RSF (8%) and SGSH (2.9%). However, SGSH and RSF behave in an opposite manner. Sulfate limitation causes lower SGSH concentration in sufficient nitrate availability. But when nitrate is limited, SGSH concentration is lower due to sulfate limitation. In contrast, sulfate limitation causes higher RSF when nitrate is sufficient. When nitrate is limited, sulfate limitation causes lower RSF (see below).

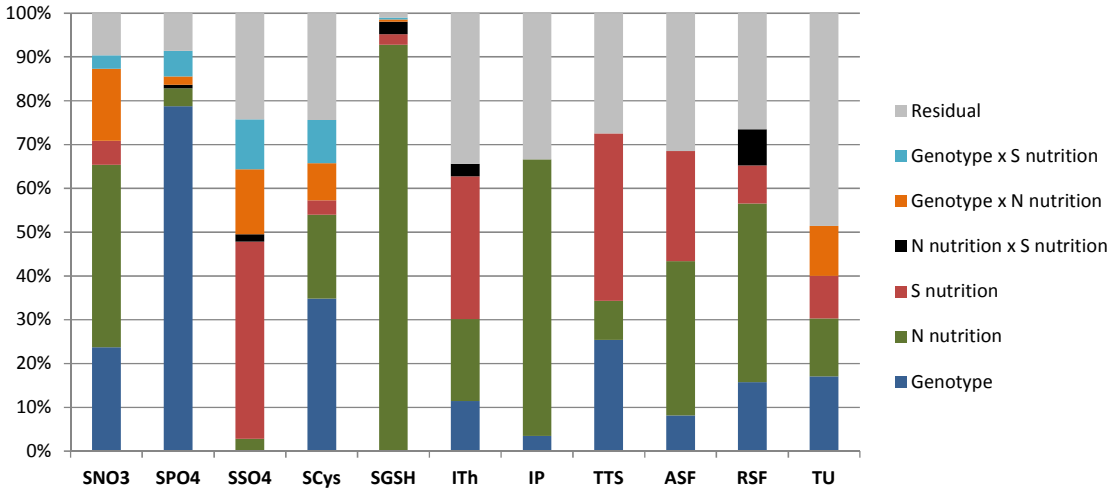


Figure 4.8 ***Analysis of variance (ANOVA) among five Arabidopsis accessions***

The analysis of variance (ANOVA) of the morphological traits which correspond to anion concentration and sulfate metabolism among five accessions representing different patterns of response to sulfate availability; Abbreviations: shoot nitrate concentration (SNO3), shoot phosphate concentration (SPO4), shoot sulfate concentration (SSO4), shoot cysteine concentration (SCys), shoot glutathione concentration (SGSH), sulfate incorporation into thiols (ITh), sulfate incorporation into proteins (IP), sulfate translocation to the shoots (TTS), absolute sulfate flux (ASF), relative sulfate flux (RSF), total sulfate uptake (TU).

4.4.3 Characteristics of accessions

In order to characterise the differences in response to the various nutrition regimes used in this experiment among all five accessions the genetic variation among the traits was investigated (Figure 4.9). Edi-0 was selected as a representative of class 2 which includes the largest plants in the collection. In this experiment it had the lowest total sulfate uptake among all accessions tested. Consequently, the absolute flux which is a function of both uptake and assimilation is very low in this accession. However, relative flux which refers to assimilation only is the highest. Therefore, it is possible to conclude that this accession has the highest rate of sulfate reduction. However, taking into account low incorporation of [^{35}S]sulfate into thiols in this accession as well as low accumulation of cysteine and glutathione most of the sulfate probably enters secondary metabolism. A quantification of methionine and glucosinolates would be necessary to verify this statement.

Col-0 as a representative of class 3 is characterised by the highest absolute sulfate flux and high sulfate incorporation into thiols among all accessions (Figure 4.9). There was no significant difference between accessions in terms of sulfate incorporation into proteins which might suggest that sulfate incorporation in proteins is not genotype dependent. However, Col-0 showed the lowest concentration of SGSH, but the highest concentration of cysteine in shoot among all accessions tested. The low SGSH concentration is in agreement with the general characteristic of class 3 which showed rather low concentration of thiols. However, high cysteine concentration in this accession is common with the general arabidopsis response profile, rather than class 3 specifically. Interestingly, C24 showed very similar pattern of response in this experiment as Col-0 suggesting that it might belong to the same class (Figure 4.9).

Sakata was selected as a representative of class 1 that includes the smallest plants in the entire collection. Moreover, class 1 was also characterised by low concentration of anions and thiols under sufficient sulfate supply (Figure 4.3). In the experiment with [^{35}S]sulfate labelling it showed the highest sulfate uptake and sulfate translocation to the shoots (Figure 4.10). However it was characterised by the lowest relative flux which is a function of sulfate assimilation. It also showed low absolute flux and sulfate incorporation in thiols. This, together with low cysteine and glutathione concentration in this accession could suggest low rate of sulfate assimilation in this accession. The apparent differences between the general characteristic of class 1 shown on Figure 4.3 and the results shown on figure 4.9 e.g. concerning nitrate concentration and sulfate allocation between shoots and roots (both traits

are the lowest among all the classes in first experiment and the highest here) are most likely to occur due to big differences in growth conditions. In the first experiment the plants were grown in hydroponic cultures for five weeks and here they were grown on plates for three weeks. Moreover, the growth conditions differed in photoperiod and light conditions. It has been shown previously that the phenotypes of mutants lacking different isoforms of O-acetylserine(thiol)lyase (OAS) differ significantly depending on growth conditions and photoperiod (Álvarez et al. 2012, Bermúdez et al. 2012).

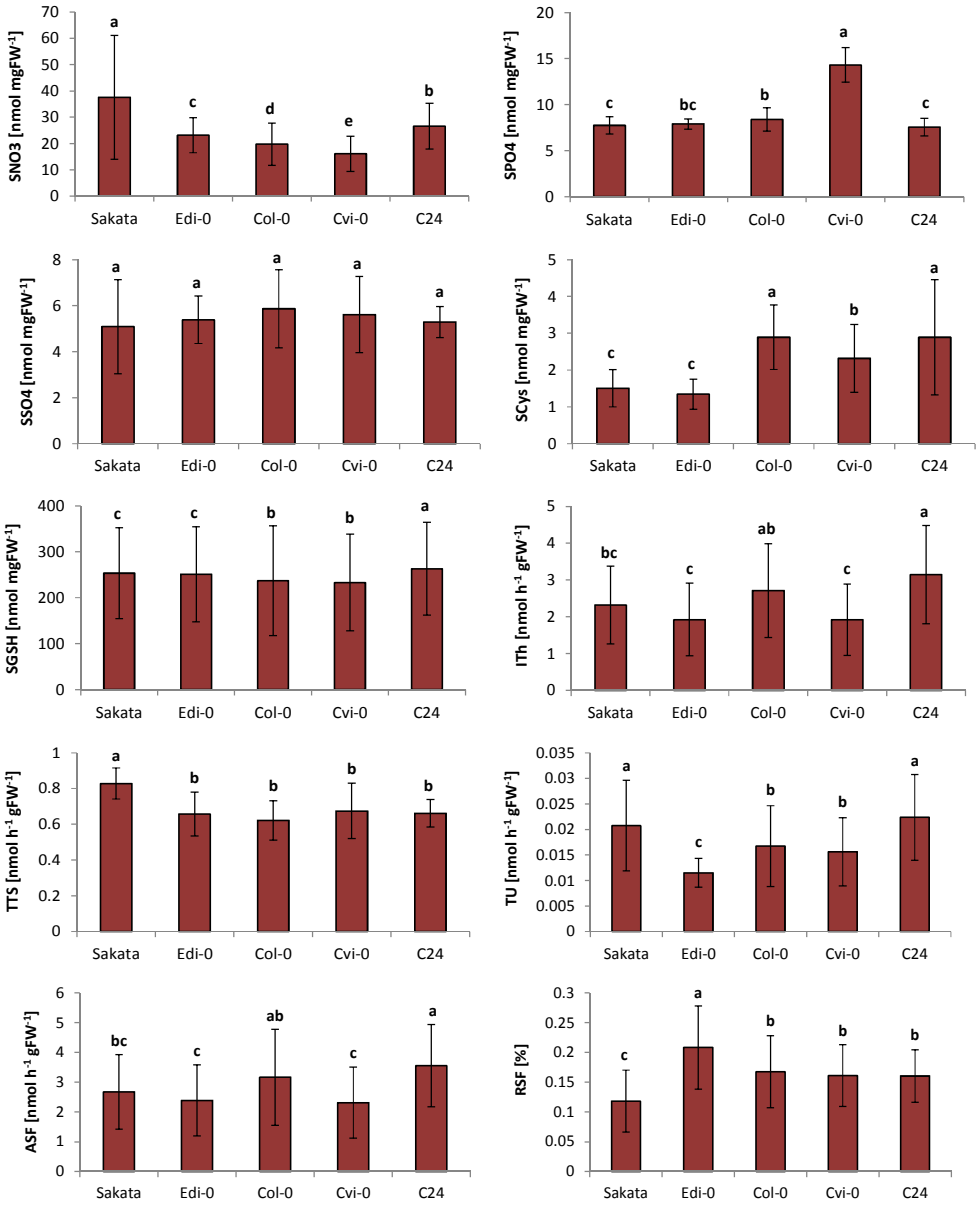


Figure 4.9 T*he response of the five accessions to sulfate and nitrate supply*

The difference between arabidopsis accessions is shown as an average of the values for all five nutrition regimes for each trait ± standard deviation. Different letters above each bar indicate values significantly different at P-value < 0.05 obtained from ANOVA with Newman-Keuls (SNK) multiple comparison grouping test. Abbreviations are in Figure 4.8.

Cvi-0, selected as a representative of class 4 in the experiment with [^{35}S]sulfate labelling, showed the lowest absolute sulfate flux and sulfate incorporation into thiols. Moreover, it was also characterised by low concentration of SGSH and shoot cysteine compared to other accessions. This is opposite to the general characteristics of class 4 shown in Figure 4.3 which indicates a high concentration of thiols in this class. The most likely reason is the difference in growth conditions mentioned above. However, the pattern of anion concentration in these two experiments (general characterisation of the class and the characterisation of one accession as a representative) is similar. Cvi-0 showed the highest phosphate concentration among all the accessions tested (Figure 4.9). The variation in phosphate concentration in response to different nutrition regimes is strongly genetically driven (Figure 4.8). Both of these results are in agreement with the general characterisation of class 4 (Figure 4.2 and 4.3). Moreover, both class 4 and Cvi-0 are characterised by the low nitrate concentration. These results might suggest some correlation in accumulation of these two anions.

It should be stressed that these data are expressed as concentration of [^{35}S]sulfate and therefore they should be treated as an indication only and they should be confirmed by the analysis of concentration of radioactive sulfate ([^{35}S]sulfate per shoot or root).

4.4.4 Response to nutrient availability

In order to provide more in depth analysis of the response of different accessions to the availability of sulfate and nitrate Fabien Chardon performed hierarchical ascendant clustering (HAC) analysis. To perform the HAC he used values obtained in all stress conditions as a percent of values obtained in control condition (Figure 4.10). The response of different accessions to changes in sulfate and nitrate availability was gathered into four distinct clusters of response dependent on both genotype and nutrition.

This analysis revealed that nitrate availability is the most important factor that contributes to plant response in this experiment. Nitrate availability separates clusters 1 and 2 that include plant response under sufficient nitrate supply from clusters 3 and 4 that include plant response under nitrate limitation. Furthermore, sulfate availability is the second factor that contributes to plant response under sufficient nitrate. Cluster 1 includes plant response to sulfate limitation and cluster 2 includes plant response to sulfate starvation under sufficient nitrate supply. Sakata and C24 are clustered irrespective of sulfate availability making these the two exceptions in these two clusters (Figure 4.10). Cluster 1 includes the response of C24

to sulfate starvation and cluster 2 includes the response of Sakata to sulfate limitation. In this case the genetic effect is stronger than the environmental effect.

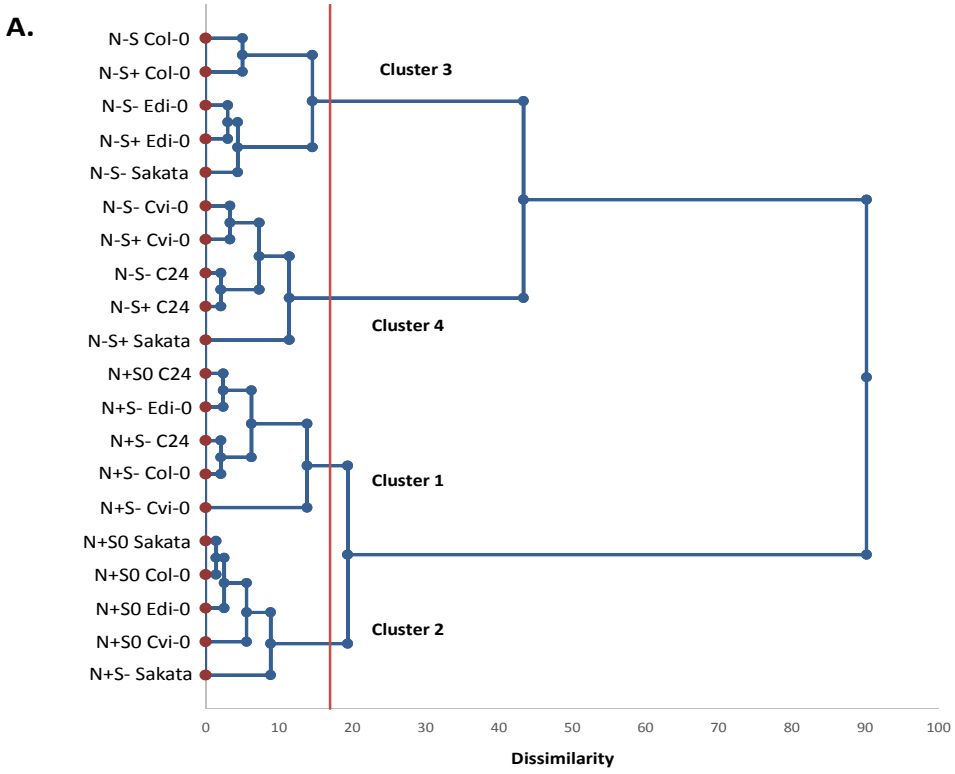


Figure 4.10 ***Key factors in plant response to sulfate and nitrate availability***

Hierarchical classification of the response to different nutrition regimes among five accessions; the dendrogram was created based on the percentage of value obtained in control condition. This analysis was performed by Fabien Chardon.

In contrast, the results of HAC revealed that under limited nitrate, plant response to different sulfate supply is mainly driven by genotype and not by sulfate nutrition. In contrast to clusters 1 and 2, the difference between clusters 3 and 4 is due to genetic effect. Cluster 3 comprises the response of Col-0 and Edi-0 and cluster 4 comprises the response of Cvi-0 and C24. Because both clusters (3 and 4) correspond to accessions response under limited nitrate these results suggest that the response of Cvi-0 and C24 as well as Col-0 and Edi-0 to nitrate limitation was similar. In this case Sakata is again an exception. Cluster 3 comprises its response to sulfate limitation and cluster 4 comprises its response to nitrate limitation under sufficient sulfate suggesting that in this accession sulfate nutrition has a stronger effect than genotype.

4.5 Primary metabolite profiling

Up to this point a general profile of the arabidopsis response to sulfate limitation and starvation has been described. Additionally, four different patterns of response to sulfate limitation and starvation were distinguished in the collection of 25 accessions. This and previous studies showed that sulfate deficiency affects not only sulfate concentration, but also other metabolic pathways. Therefore it was of interest to unravel the primary metabolites and biochemical pathways that are involved in the response to sulfate limitation at a more global scale and further characterise the patterns of response to sulfate limitation revealed earlier. Primary metabolite profiling was performed on two independent sets of samples. First of all, I grew the four arabidopsis accessions selected as representatives of different patterns of response to sulfate limitation and starvation on plates for three weeks in five different sulfate/nitrate regimes as described above (section 4.2.3). Because I did not obtain enough root tissue from three week old seedlings, in this set only shoot samples were analysed. However, the plants for the global analysis of 25 arabidopsis accessions were grown in hydroponic cultures. This way of plant cultivation allows longer periods of plant growth delivering bigger plants and more tissue for analysis. Therefore, a second set of plants for metabolite profiling was grown in hydroponic cultures in Versailles by Giorgiana Chietera. In both cases samples were analysed at MPIMP in Golm. Plant extracts were prepared by two step extraction with methanol and chloroform and analysed by a gas chromatography-mass spectrometry (GC-MS). Both of these plant sets were analysed only recently and the statistical analysis of the data is not complete yet. However, some interesting results have already been revealed.

All results described below concern the shoots of three week old seedlings grown on plates. The statistical analysis of plants grown on hydroponic cultures is still at an early stage. However, it has already shown a significant overlap with metabolites detected in the first experiment suggesting that the two analyses may well complement each other. Moreover, the analysis of the plants grown in hydroponic cultures included the analysis of roots. Therefore, it will provide a more complete picture of the different patterns of response to changes in nutrient availability among the four accessions analysed and eventually, better understanding of plant adaptation strategies to different environmental conditions. It will also allow investigation of differences between the metabolic changes in shoots and roots and identification of specific metabolic pathways which were affected by nutrition stress.

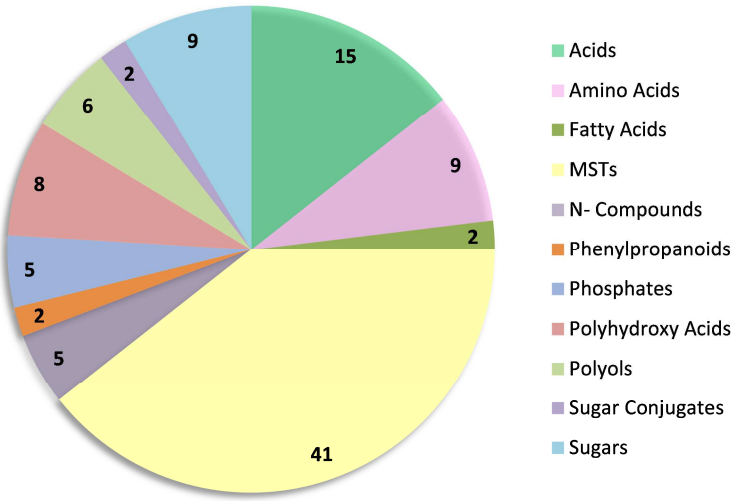


Figure 4.11 ***Metabolites identified in the shoots of three week old seedlings***

A pie chart presenting the amount of identified metabolites classified based on the chemical structure. MST – non-identified but repeatedly observed mass spectral metabolites tags.

4.5.1 General characterisation of changes in primary metabolites

The GC-MS analysis of the shoots of three week old seedlings grown on plates resulted in the identification of 107 metabolites. Among these 63 were known metabolites and 41 were non-identified, but repeatedly observed mass spectral metabolite tags (MSTs) which correspond to 40% of all detected metabolites in this experiment. The profiled metabolites are classified into ten categories based on their chemical structure: acids, amino acids, fatty acids, nitrogenous compounds, phenylpropanoids, phosphates, polyhydroxyacids, polyols, sugars, and sugar conjugates (Figure 4.11). Because of the high cost of standards for particular metabolites the data obtained from the GC-MS analysis are usually expressed as a pool size and not concentration or content (Dethloff et al. 2014). A pool size refers to the data obtained from GC-MS (peak area) normalised to an internal standard (in this experiment C19 for organic phase and sorbitol for polar phase; section 4.2.4) and fresh matter of the sample. However, if the pool size is 10 for metabolite A and 30 for metabolite B it is not possible to say that in a given sample there is more metabolite B than A. Therefore, to investigate the changes among known metabolites identified in this experiment in response to different nutrition regimes a ratio of response for each of them was calculated by Sarah Whitcomb (MPIMP Golm). To obtain the ratio of response a pool size of a given metabolite identified in a sample was divided by average pool size of that metabolite in the samples from control condition from the same

accession. This allows to clearly compare the changes in each accession due to nutrition regime (Dethloff et al. 2014). Additionally, the response ratios were $\log_2$ transformed to make the data distribution more symmetric and to convert multiplicative relations into additive relations (to reduce heteroscedasticity, Kvalheim et al. 1994, van den Berg et al. 2006, Veyel et al. 2014).

Log2(response ratio)

Metabolite	Sakata				Edi-0				Col-0				Cvi-0			
	Class 1				Class 2				Class 3				Class 4			
	S-	S0	N-	N-S-	S-	S0	N-	N-S-	S-	S0	N-	N-S-	S-	S0	N-	N-S-
Citric acid	-0.221	-1.041	-1.000	-1.880	-0.910	-0.478	-0.965	-0.832	0.911	1.025	0.186	0.264	-1.016	0.268	-1.796	-1.472
Fumaric acid	0.856	0.350	0.130	0.417	-0.265	0.394	-0.251	-0.447	0.571	0.803	0.020	-0.475	0.061	0.275	-0.565	-0.358
Malic acid	0.691	-0.457	-0.026	0.000	-0.231	0.008	-0.003	-0.566	0.474	-0.152	-0.398	-0.587	-0.097	0.513	-1.305	-0.901
Succinic acid	0.206	-0.445	0.219	0.204	-0.421	0.326	0.356	0.162	0.179	0.347	-0.313	-0.353	-0.326	0.987	-0.619	-0.628
Benzoic acid,	0.741	-0.368	-0.183	0.042	0.132	-0.092	0.063	0.309	0.258	0.237	0.339	0.844	0.170	0.150	0.243	0.655
Lactic acid	0.820	0.453	0.461	0.324	0.590	-0.244	0.027	0.201	0.145	0.012	0.304	0.337	0.210	0.511	1.716	0.687
Boric acid	0.441	-0.136	-0.019	0.236	0.580	-0.688	0.247	0.322	0.341	0.184	0.209	0.567	0.369	0.305	0.169	0.623
4-hydroxy-benzoic acid	0.380	0.064	0.000	-0.010	-0.128	-0.555	-0.319	0.395	0.394	0.044	0.205	0.475	0.401	0.186	0.718	0.454
2,4-dihydroxy-butanoic acid	1.751	-0.645	-0.112	0.116	0.061	-0.278	-0.036	-0.127	0.420	0.376	0.308	1.196	0.644	-0.078	0.229	0.854
4-amino-butanoic acid	-1.232	-1.314	-0.332	0.159	-0.303	-0.387	-0.082	-0.144	0.482	-1.152	0.749	0.659	0.522	0.691	1.002	0.289
Valine	0.061	-0.283	-0.679	-0.560	-0.013	-1.142	-0.266	0.142	1.181	0.979	1.520	1.108	-0.266	-0.955	-0.442	-0.201
Glycine	0.565	-0.026	-0.081	-0.159	0.327	-0.038	-0.140	0.045	0.493	-0.036	1.033	0.583	0.506	1.079	0.806	0.594
β-Alanine	0.323	-0.947	-0.386	-0.255	0.037	-1.317	-0.536	0.066	1.109	0.464	1.222	0.982	-0.464	-0.534	0.222	0.289
Serine	0.108	-1.040	0.227	-0.576	-0.400	-0.630	-0.298	0.006	0.645	0.226	0.754	1.098	-0.252	0.596	1.111	0.679
Leucine	-0.195	-0.492	-0.596	-0.730	-0.223	-0.729	-0.228	-0.035	0.873	0.860	0.980	0.957	-1.025	-0.514	0.082	-0.013
Octadecanoic acid	0.605	-0.073	0.152	0.179	-0.118	-0.270	0.096	0.154	0.378	0.052	0.081	0.224	-0.332	0.272	0.199	0.414
Hexadecanoic acid	0.618	-0.196	0.094	0.114	-0.004	-0.274	0.115	0.216	0.322	-0.027	0.036	0.303	-0.362	0.197	0.029	0.337
Ethanolamine	0.513	-0.250	-0.009	0.154	0.226	-0.270	0.124	0.496	0.210	0.167	0.195	0.450	0.220	0.229	0.237	0.487
2-hydroxy-pyridine	0.396	-0.282	-0.053	0.128	0.292	-0.262	0.026	0.502	0.050	0.299	0.218	0.455	0.366	0.212	0.086	0.372
Indole-3-acetonitrile	0.756	0.087	0.089	-0.010	-0.330	-0.260	-0.024	-0.455	0.189	-0.158	0.709	0.349	-0.894	-0.131	-0.030	-0.253
cis - sinapic acid	0.995	0.054	0.320	0.235	-0.032	-0.128	0.091	-0.732	1.204	0.380	0.477	0.746	-0.298	0.737	-0.372	-0.111
Phosphoric acid	0.407	-0.456	-0.142	0.044	-0.034	-0.387	-0.309	-0.154	0.442	0.524	0.799	0.863	0.214	0.342	0.370	0.662
Phosphoric acid monomethyl ester	0.457	-0.565	-0.229	-0.523	-0.609	-0.522	-0.604	-0.829	1.388	0.859	0.587	0.486	0.544	0.806	0.865	1.130
Glyceric acid	-1.499	-0.927	-1.977	-0.931	-0.349	0.478	0.111	-0.174	1.040	1.858	0.281	0.670	0.219	1.608	0.569	0.153
Dehydroascorbic acid dimer	0.743	-0.198	-0.277	0.003	-0.001	-0.141	0.277	0.205	0.510	0.640	0.713	0.818	0.559	0.870	0.346	0.199
Threonic acid	0.743	0.766	0.148	0.193	0.249	0.449	0.034	-0.302	0.405	0.386	0.149	-0.030	-0.657	0.785	-0.836	-0.689
Threonic acid-1,4-lactone	0.615	0.258	-0.034	0.093	0.143	0.300	-0.065	0.087	0.394	0.820	0.359	0.351	-0.073	0.891	-0.445	-0.348
Galactonic acid	0.190	-1.096	-0.049	-0.439	-0.653	-0.876	-0.253	-0.536	0.222	-0.627	0.537	0.147	-0.427	-0.074	0.335	0.235
myo -Inositol	0.571	0.365	0.169	0.764	-0.161	0.635	0.181	0.354	0.419	1.208	0.539	0.654	0.609	1.508	-0.052	-0.302
Glycerol	0.197	0.312	-0.089	0.553	1.213	0.159	-0.466	0.260	0.621	0.836	0.497	1.417	-0.533	-0.290	0.555	0.276
Erythritol	0.432	-0.593	-0.160	0.046	-0.060	-0.427	-0.183	0.196	0.329	0.111	0.302	0.449	0.417	0.187	0.278	0.163
Sorbitol	0.337	-0.079	-0.543	-0.308	-0.126	0.037	-0.093	-0.510	0.698	0.797	0.468	0.189	0.593	1.208	0.455	0.098
Ribitol	0.386	-1.193	-0.150	0.113	0.015	-0.864	0.063	0.287	0.362	-0.088	0.555	0.532	0.514	0.088	0.594	0.721
Mannitol	0.885	-0.122	-0.031	-0.086	1.314	0.410	0.274	0.034	0.299	0.352	0.065	0.206	-0.044	0.485	0.110	0.037
Fructose	0.703	-0.546	0.338	0.817	-1.394	0.907	-0.321	-0.040	-0.051	0.097	-0.245	-0.318	1.292	1.549	0.875	-0.733
Sucrose	0.839	0.477	0.435	0.555	-0.436	0.488	-0.076	0.006	0.644	1.043	1.018	1.149	-0.018	0.940	0.115	-0.220
Glucose	0.868	-0.112	0.070	0.494	-0.740	1.149	-0.009	-0.060	0.392	0.389	0.540	0.645	0.778	1.002	0.829	0.389
Galactose	0.330	-0.411	-0.266	0.130	-0.508	0.454	-0.131	0.188	0.693	0.611	0.422	0.415	0.698	0.868	0.292	-0.051
Ribose	0.286	-0.705	0.058	0.135	-0.093	-0.310	0.204	0.220	0.678	0.505	0.455	0.590	0.665	0.595	0.832	0.581
β - 1,6-anhydro glucose	0.464	-0.106	0.156	0.340	-0.030	-0.010	0.240	0.328	0.650	0.215	0.555	0.671	0.042	0.089	0.383	0.652
Xylose	0.817	0.152	-0.077	0.379	-0.061	0.018	0.227	0.177	0.457	0.883	0.258	0.779	0.330	1.264	0.400	0.508

Figure 4.12 ***Changes in concentration of identified metabolites in different accessions***

The heat map shows the changes in metabolite response ratio of four accessions in response to different nutrition regimes tested. The values are shown as a relative to response ratio and in $\log_2$ space. Colours indicate the differences in metabolite accumulation between different accessions and in different nutrition regimes with the lowest value in blue (decrease) and the highest in red (increase). This analysis was made by Sarah Whitcomb.

In order to investigate how the primary metabolites varied in response to different nutrition regimes among the four accessions analysed, Sarah Whitcomb (MPIMP Golm) created a heat map for metabolites which were detected in 70% of all samples tested (Figure 4.12). In general, Sakata and Edi-0 selected as representatives of class 1 and 2 respectively showed decrease in response ratio of most metabolites detected in response to most of the conditions analysed when compared to control condition. In contrast, most of the metabolites in Col-0 and Cvi-0 selected as representatives of class 3 and 4 respectively showed an increase in response ratio of most metabolites detected in response to the nutrition regimes tested when compared to control condition. Interestingly, Sakata showed increase in response ratio of most metabolites in response to sulfate limitation. However, in induced sulfate starvation it showed a decrease in response ratio of most detected metabolites. Edi-0, selected as a representative of class 2, showed decrease in response ratio of most metabolites in response to changes in sulfate and nitrate supply. In contrast, Col-0 which is a representative of class 3 showed a general increase in response ratio of metabolites detected in response to different nutrition regimes. In Cvi-0 the response to different nutrition regimes varied depending on the metabolite (Figure 4.13).

In order to reveal the metabolites which response ratios changed most significantly a *t*-test significance analysis was carried out by Sarah Whitcomb on these response ratios. Metabolite response ratios showing *t*-test significance P-value ≤ 0.05 were considered as different. Altogether, the response ratio of 69 metabolites and MSTs changed significantly in response to at least one nutrition regime tested. In general, Col-0 had the highest amount of metabolites which response ratio changed significantly in response to sulfate limitation (21), nitrate limitation (22), and sulfate/nitrate double limitation (15; Figure 4.13). Edi-0 was the accession which had the lowest number of metabolites which response ratio changed significantly in response to these three nutrition regimes (2, 4, and 2 metabolites, respectively). The nutrition regime which resulted in the biggest number of metabolites which response ratio changed significantly for all accessions tested was induced sulfate starvation (Figure 4.14B). In this condition Cvi-0 had 26 metabolites which response ratio changed significantly, Edi-0 had 25 metabolites which response ratio changed significantly, Col-0 had 20 metabolites which response ratio changed significantly, and Sakata had 15 metabolites which response ratio changed significantly. The only chemical compound category which was not affected by any of the nutrition regimes tested in any of the four accessions was phenylpropanoids. This is surprising since it has been shown previously that these compounds accumulate often in nitrogen limitation (Fritz et al. 2006, Yaeno & Iba 2008). MSTs were the

group with the largest amount of metabolites which response ratio changed significantly. However, these are not useful in identification of particular response process in arabidopsis accessions to different nutrition regimes. Therefore, at this stage MSTs were excluded from further analysis. After elimination of MSTs, amino acids and sugars were the two chemical compound categories with the largest number of metabolites which response ratio changed significantly.

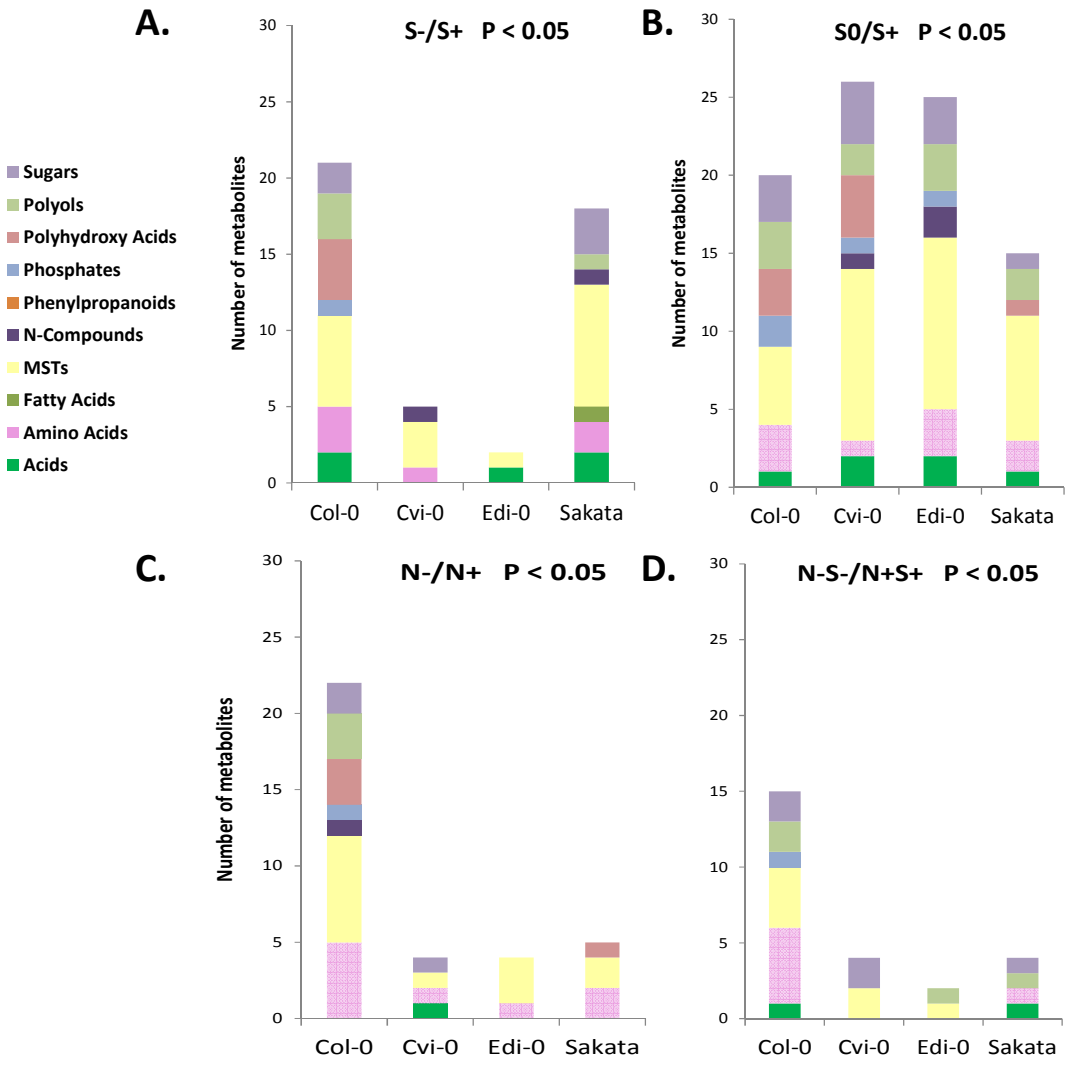


Figure 4.13 ***Number of metabolites which response ratios changed significantly***

Metabolites which response ratio changed significantly in response to sulfate limitation (A), sulfate starvation (B), nitrate limitation (C), and sulfate/nitrate limitation (D) at P-value<0.05. Data are shown as a relative to normal condition with sufficient sulfate and nitrate supply. Different colours correspond to different chemical categories.

4.5.2 Changes in metabolite response ratios among accessions

The profiling of primary metabolites revealed substantial differences between accessions. The repertoire of metabolites which response ratio changed significantly differed between accessions as well as between the nutrition regimes. Col-0, as a representative of class 3, was the accession with the highest number of metabolites which response ratio changed significantly in all regimes tested. Moreover, the response ratio of only one metabolite – 4-amino-butanoic acid, a non-proteinogenic amino acid involved in the metabolism of alanine, aspartate, and glutamate – was reduced in response to induced sulfate starvation. The response ratios of all the other metabolites which response ratio changed significantly in this accession increased in response to different nutrition regimes. The response ratio of the two proteinogenic amino acids – leucine and valine – significantly increased in this accession in response to all nutrition regimes tested compared to control condition. Valine and β-alanine (not changed in response to sulfate starvation) were the two metabolites which response ratios increased most strongly in this accession. Additionally, *myo*-inositol – a molecule that builds a number of lipid signalling molecules in the cell – and glyceric acid were the two metabolites which response ratios increased most strongly in response to induced sulfate starvation. Col-0 was also characterised by strong increase of response ratio of sucrose in response to all nutrition regimes except sulfate limitation.

Similarly to Col-0, Cvi-0 selected as a representative of class 4 was also characterised by general increase in response ratio of significantly changed metabolites in response to different nutrition regimes. In this case only the response ratio of leucine and indole-3- decreased in response to sulfate limitation compared to control condition. These were also the only two metabolites which response ratio changed significantly in response to sulfate limitation. The response ratio of most metabolites in this accession was higher in response to sulfate starvation compared to control condition. In this accession galactinol was the metabolite with the most strongly elevated response ratio. The response ratio of only three metabolites was higher in response to nitrate limitation, and only two in response to sulfate/nitrate double limitation.

Induced sulfate starvation was also the condition with the highest number of metabolites which response ratio changed significantly in Edi-0, selected as a representative of class 2. It is also the accession with the lowest number of metabolites which response ratio changed significantly. In Edi-0 65% of all metabolites which response ratio changed significantly were reduced in response to different nutrition regimes compared to control

condition. Amino acids – alanine, glutamate and valine – were the metabolites which response ratio was the lowest in response to induced sulfate starvation. In contrast, glucose was the metabolite which response ratio was the highest in response to induced sulfate starvation.

Table 4.1 ***Metabolites which response ratio changed significantly***

Metabolite	Col-0				Class
	S-	S0	N-	N-S-	
4-hydroxy-benzoic acid				0.475	Acids
2,4-dihydroxy-butanoic acid	0.420				Acids
Fumaric acid	0.571	0.803			Acids
β-Alanine	1.109		1.222	0.982	Amino Acids
4-amino-butanoic acid		-1.152	0.749	0.659	Amino Acids
Leucine	0.873	0.860	0.980	0.957	Amino Acids
Serine			0.754	1.098	Amino Acids
Valine	1.181	0.979	1.520	1.108	Amino Acids
Indole-3-acetonitrile			0.709		N- Compounds
Phosphoric acid		0.524	0.799	0.863	Phosphates
Phosphoric acid monomethyl ester	1.388	0.859			Phosphates
Dehydroascorbic acid dimer			0.713		Polyhydroxy Acids
Erythronic acid	0.737	0.616			Polyhydroxy Acids
Galactonic acid			0.537		Polyhydroxy Acids
Glyceric acid	1.040	1.858			Polyhydroxy Acids
Threonic acid	0.405				Polyhydroxy Acids
Threonic acid-1,4-lactone	0.394	0.820	0.359		Polyhydroxy Acids
Erythritol				0.449	Polyols
Glycerol	0.621	0.836	0.497		Polyols
myo -Inositol	0.419	1.208	0.539		Polyols
Ribitol				0.532	Polyols
Sorbitol	0.698	0.797	0.468		Polyols
Galactose	0.693	0.611			Sugars
β - 1,6-anhydro glucose			0.555	0.671	Sugars
Ribose	0.678				Sugars
Sucrose		1.043	1.018	1.149	Sugars
Xylose		0.883			Sugars

Metabolite	Edi-0				Class
	S-	S0	N-	N-S-	
Boric acid		-0.688			Acids
Fumaric acid		0.394			Acids
Lactic acid	0.590				Acids
β-Alanine		-1.317	-0.536		Amino Acids
Glutamic acid		-0.959			Amino Acids
Valine		-1.142			Amino Acids
Ethanolamine		-0.270			N- Compounds
Putrescine		0.984			N- Compounds
Phosphoric acid		-0.387			Phosphates
Erythritol		-0.427			Polyols
myo -Inositol		0.635			Polyols
Ribitol		-0.864			Polyols
Sorbitol				-0.510	Polyols
Fructose		0.907			Sugars
Glucose		1.149			Sugars
Ribose		-0.310			Sugars

Metabolite	Cvi-0				Class
	S-	S0	N-	N-S-	
Boric-acid		0.305			Acids
Lactic acid			1.716		Acids
Succinic acid		0.987			Acids
4-amino-butanoic acid		0.691	1.002		Amino Acids
Leucine	-1.025				Amino Acids
Indole-3-acetonitrile	-0.894				N- Compounds
2-hydroxy-pyridine		0.212			N- Compounds
Fructose-6-phosphate		1.706			Phosphates
Dehydroascorbic acid dimer		0.870			Polyhydroxy Acids
Glyceric acid		1.608			Polyhydroxy Acids
Threonic acid		0.785			Polyhydroxy Acids
Threonic acid-1,4-lactone		0.891			Polyhydroxy Acids
myo -Inositol		1.508			Polyols
Sorbitol		1.208			Polyols
Galactinol		2.209			Sugar Conjugates
Galactose		0.868			Sugars
β - 1,6-anhydro glucose				0.652	Sugars
Raffinose			1.575		Sugars
Ribose		0.595		0.581	Sugars
Sucrose		0.940			Sugars

Metabolite	Sakata				Class
	S-	S0	N-	N-S-	
Benzoic acid,	0.741				Acids
Citric acid		-1.041		-1.880	Acids
Fumaric acid	0.856				Acids
β-Alanine		-0.947			Amino Acids
4-amino-butanoic acid	-1.232	-1.314			Amino Acids
Glycine	0.565				Amino Acids
Isoleucine			-0.415		Amino Acids
Valine			-0.679	-0.560	Amino Acids
Hexadecanoic acid	0.618				Fatty Acids
Indole-3-acetonitrile	0.756				N- Compounds
Galactonic acid		-1.096			Polyhydroxy Acids
Glyceric acid			-1.977		Polyhydroxy Acids
Erythritol		-0.593			Polyols
myo -Inositol				0.764	Polyols
Mannitol	0.885				Polyols
Ribitol		-1.193			Polyols
Fructose				0.817	Sugars
β - 1,6-anhydro glucose	0.464				Sugars
Ribose		-0.705			Sugars
Sucrose	0.839				Sugars
Xylose	0.817				Sugars

The data correspond to averages of response ratios from up to eight biological replicates shown as $\log_2$ transformed values. Only metabolites which response ratios changed significantly (*t*-test significance $P<0.05$) are shown. Values higher than those obtained in control condition and marked in green, lower – in red.

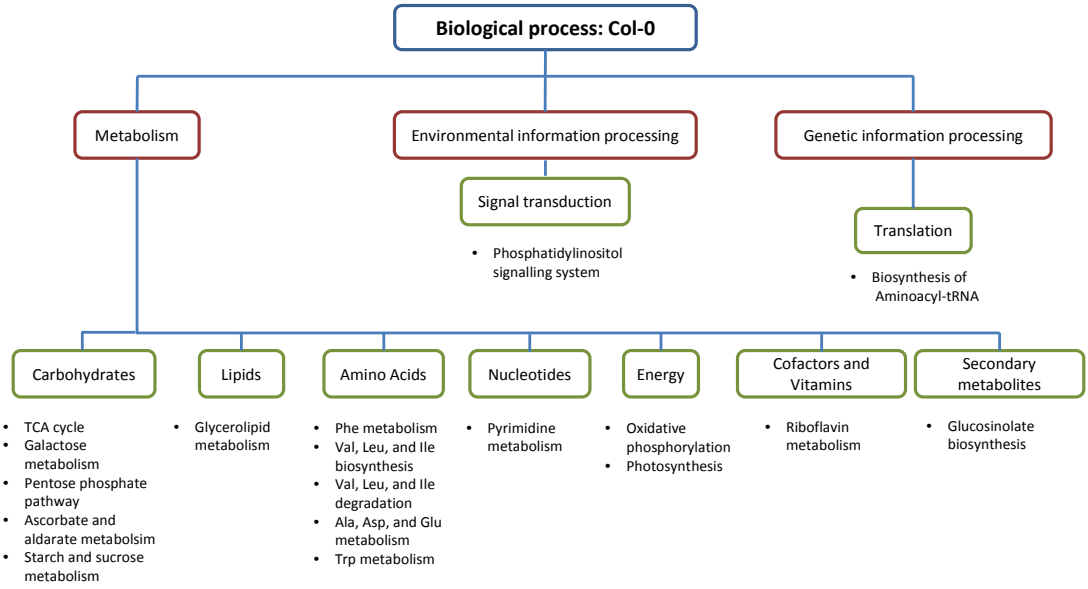
In Sakata, selected as a representative of class 1, 54% of all metabolites which response ratio changed significantly were reduced in response to different nutrition regimes compared to control condition. In this accession response ratio of metabolites that changed significantly in response to sulfate limitation was higher than in control condition whereas response ratio of metabolites that changed significantly in response to induced sulfate starvation was lower compared to control condition. Mannitol – the sugar alcohol that is involved in plant response to biotic and abiotic stress – was the metabolite which response ratio increased the most in this accession (in response to sulfate limitation). Glyceric acid was the metabolite which response ratio was most strongly reduced in this accession (in response to nitrate limitation).

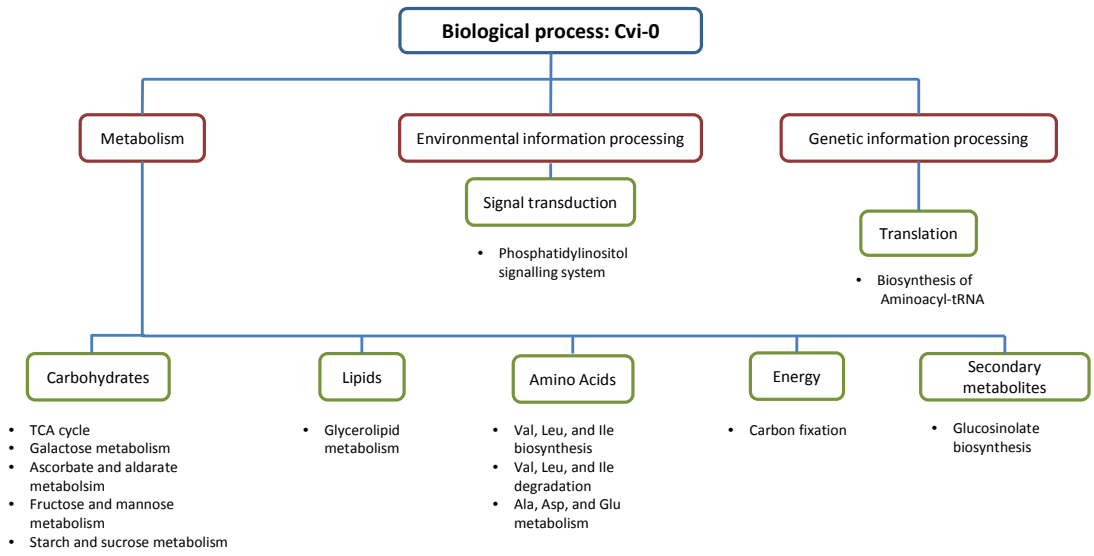
4.5.3 Alterations of metabolic pathways among four accessions

After identifying the metabolites which response ratio changed significantly it was possible to assign them to their respective metabolic pathways and processes they are involved in. This assignment was performed by Sarah Whitcomb (MPIMP) and it is based on information from Kyoto Encyclopaedia of Genes and Genomes (KEGG) – a database of manually drawn pathway maps including the current understanding of the molecular interactions and processes in plant cells. This analysis revealed that Col-0, selected as a representative of class 3, showed the most complex response to the analysed nutrition regimes in terms of the amount of altered processes in the cell (Figure 4.14). The metabolism of amino acids and carbohydrates were the two processes mainly altered in this accession. The metabolism of amino acids was mainly altered by nitrate and sulfate/nitrate double limitation. However, response ratio of some of the amino acids was significantly higher in response to sulfate limitation and induced sulfate starvation. In contrast, carbohydrate metabolism was altered in response to all nutrition regimes tested. The distortion of carbohydrate metabolism is a common change in plants under nutrient stress and it is most likely to occur due to reduction in photosynthetic capacity (Honsel et al. 2012, Lunde et al. 2008).

Similarly to Col-0, carbohydrate metabolism was also the main cellular process altered in Cvi-0 – a representative of class 4 (Figure 4.14). Moreover, not only the metabolic process, but also particular pathways related to carbohydrate metabolism and altered by nutrition regimes analysed in this experiment were similar in these two accessions. Cvi-0 is the accession with the smallest number of metabolic processes altered in response to different nutrition regimes and most of these changed in response to induced sulfate starvation.

Edi-0 was selected as a representative of class 2 and was characterised by the lowest number of significantly altered metabolites among all four accessions (Figure 4.14). Similarly to Col-0, amino acid and carbohydrate metabolism are the two metabolic processes mainly altered in this accession. Like in Cvi-0, most of metabolic processes are altered by sulfate starvation. In contrast, apart from amino acid and carbohydrate metabolism, in Sakata metabolism of the lipids, particularly fatty acid metabolism, was significantly altered. In this accession both sulfate limitation and induced sulfate starvation are the main drivers of observed metabolic changes.


Biological process: Col-0
Metabolism
Environmental information processing
Genetic information processing
Signal transduction
Phosphatidylinositol signalling system
Translation
Biosynthesis of Aminoacyl-tRNA
Carbohydrates
Lipids
Amino Acids
Nucleotides
Energy
Cofactors and Vitamins
Secondary metabolites
TCA cycle
Galactose metabolism
Pentose phosphate pathway
Ascorbate and aldarate metabolsim
Starch and sucrose metabolism
Glycerolipid metabolism
Phe metabolism
Val, Leu, and Ile biosynthesis
Val, Leu, and Ile degradation
Ala, Asp, and Glu metabolism
Trp metabolism
Pyrimidine metabolism
Oxidative phosphorylation
Photosynthesis
Riboflavin metabolism
Glucosinolate biosynthesis


Biological process: Cvi-0
Metabolism
Environmental information processing
Genetic information processing
Signal transduction
Phosphatidylinositol signalling system
Translation
Biosynthesis of Aminoacyl-tRNA
Carbohydrates
Lipids
Amino Acids
Energy
Secondary metabolites
TCA cycle
Galactose metabolism
Ascorbate and aldarate metabolsim
Fructose and mannose metabolism
Starch and sucrose metabolism
Glycerolipid metabolism
Val, Leu, and Ile biosynthesis
Val, Leu, and Ile degradation
Ala, Asp, and Glu metabolism
Carbon fixation
Glucosinolate biosynthesis

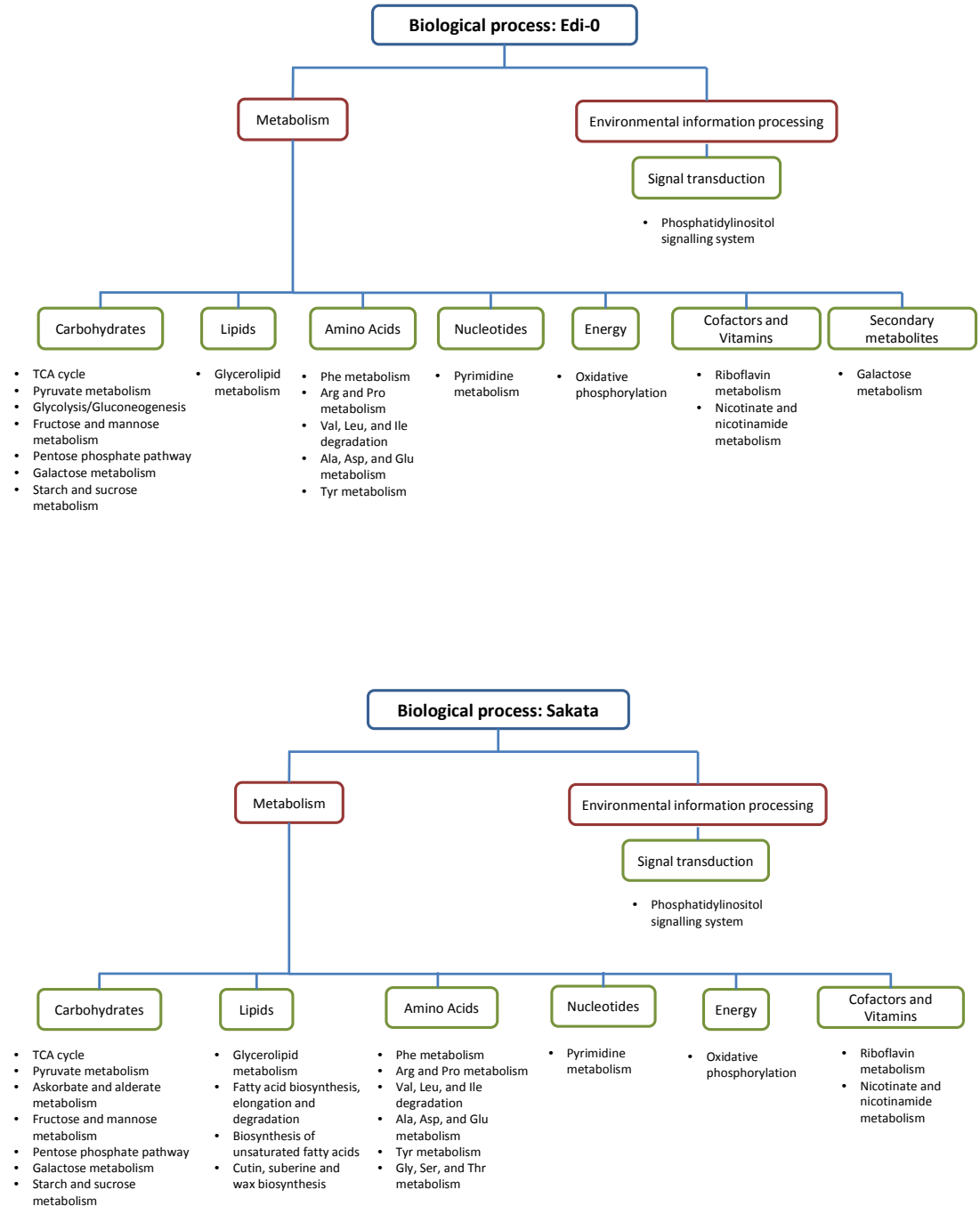


Figure 4.14 ***Metabolic pathways altered in response to sulfate and/or nitrate deficiency***

The figures were created based on the assignment of significantly altered metabolites in particular accessions to their respective metabolic pathways performed by Sarah Whitcomb. The assignment was performed using the Kyoto Encyclopaedia of Genes and Genomes (KEGG).

4.6 Discussion

In this study the extent of variation in growth response to different sulfate supply in *Arabidopsis thaliana* was investigated with the goal to identify accessions showing contrasted responses and ultimately different growth adaptive strategies. The analysis described in this chapter is unique since it provides the general arabidopsis response profile to changes in sulfate availability based on the analysis of 25 genetically different accessions grown in highly controlled environments. Such collections of accessions were successfully used to dissect the natural variation in response to nitrate (De Pessemier et al. 2013, Ikram et al. 2012), phosphate (Chevalier et al. 2003), and potassium (Kellermeier et al. 2013) availability. Several studies were conducted to provide a general profile of arabidopsis response to sulfate availability. However, in these studies, only Col-0 (Hirai et al. 2004, Nikiforova et al. 2003, Nikiforova et al. 2005) or *sel1* mutants (Zhang et al. 2014) were used, thus limiting the overall outcome of a study. The analysis of natural variation allows dissection of traits important for plant growth in response to combined genetic and environmental variation.

The global response profile computed as an average of values obtained from 25 arabidopsis accessions revealed that plants respond to sulfate limitation with a higher root biomass and a bigger amount of lateral roots (Figure 4.1). This is in agreement with previous studies which reported the changes in root morphology leading to an increase of total absorptive surface of the system (Kutz et al. 2002, Lopez-Bucio et al. 2003). Moreover, it has been shown that during limited sulfate supply the available resources are not transported to the shoots, but remain in the roots to increase the biomass production of organs involved in acquiring of depleted nutrients (Hawkesford & De Kok 2006, Stuiver et al. 1997). In the experiment described here this phenomenon was shown as significantly lower SRDM ratio in both sulfur stress conditions (Figure 4.1). Kutz et al. (2002) have shown that in arabidopsis the lateral roots develop closer to root tip (earlier) and at increased frequency under sulfate limitation. This adaptation response is dependent on the induction of *NITRILASE3* (*NIT3*) which initiates the production of additional auxin leading to increased root growth and branching at the expense of shoot growth. Subsequently, Koprivova et al. (2010) have shown that depletion of GSH by inhibition of its synthesis with buthionine sulphoximine (BSO) results in loss of auxin accumulation in the root apex leading to an inhibition in root growth. In the experiment described here the classes characterised by smaller decrease of GSH concentration in response to sulfate limitation showed also a smaller reduction in root growth in response to sulfate limitation compared to other classes. These results highlight the importance of GSH in plant response to sulfate deficiency.

Accessions from classes 1 and 2 were characterised by no changes or even increase (respectively) in GSH concentration in shoots in response to sulfate limitation (Figure 4.4). GSH serves as the main storage of reduced sulfur in the plant. It is also the main form (together with S-methylmethionine) of reduced sulfur transported via phloem (Rennenberg et al. 1979). Moreover, it has been postulated to act as a signal of the sulfur status from shoots to roots (Herschbach et al. 2000, Lappartient et al. 1999). Apart from its role in sulfur metabolism, GSH also plays an important role in regulation of plant growth and development including redox homeostasis (Lewandowska & Sirko 2008). Since nutrient deficiency is known to induce oxidative stress (Anjum et al. 2012, Waraich et al. 2012) the maintenance of GSH concentration and partitioning between different plant organs has an important role in plant response to nutrient depletion. The shoot is assumed to be the predominant site of sulfate reduction in the plant (although all the enzymes of sulfate reduction pathway are present in roots; Hubberten et al. 2012a, Saito 2000). It has been shown that APS reductase, the key regulatory component of the sulfate reduction pathway, is post-translationally induced by oxidative stress (Bick et al. 2001). The induction of APS reductase can lead to an increase in GSH concentration (Bick et al.). Therefore, higher concentration of GSH in shoots is not surprising and could be postulated as useful plant strategy to maintain the growth during limited sulfate availability. Additionally, GSH may also serve as a donor of cysteine for protein synthesis since it can be degraded completely to cysteine, glutamate, and glycine (Noctor et al. 1998) which further explains its higher concentration in the aerial parts of the plant than in roots. However, to verify this hypothesis the analysis of translocation of GSH between shoots and roots and the expression analysis of genes encoding the enzymes involved in GSH biosynthesis and degradation would be necessary.

The increase in the concentration of all anions tested in the shoots (SNO3, SPO4, SSO4) in response to sulfate limitation was observed in the general arabidopsis response profile as well as in the characterisation of all classes except class 3 (Figure 4.1 and 4.4 respectively). These results suggest an interconnection between these three elements in response to sulfate limitation. It has been reported that the changes in concentration of phosphate and/or sulfate cause coordinated downstream metabolic responses (Essigmann et al. 1998, Sugimoto et al. 2007). The most common is the rapid replacement of sulfolipids by phospholipids under sulfate deficiency and vice versa (Essigmann et al. 1998, Sugimoto et al. 2007). Additionally, deficiency of one of these elements often results in modification of expression level of genes encoding proteins involved in the regulation of the homeostasis of the other elements

(proteins catalysing assimilation and controlling assimilatory flux of these elements and regulatory factors acting in the signalling pathways; Misson et al. 2005, Rouached 2011).

Nitrate availability is the most important factor that contributes to the plant response in this experiment (Figure 4.10). Similarly to the interconnection between phosphate and sulfate accumulation the assimilatory pathways of sulfur and nitrogen are also considered functionally convergent and well-coordinated (Koprivova et al. 2000, Anjum et al. 2012). However, the mechanism of this coordination is different than in case of phosphate. In the case of sulfur and nitrogen the availability of one element regulates the metabolism of the other (Anjum et al. 2012, Habtegebrial & Singh 2006). The interconnection between sulfate and nitrate assimilation is long established. Such coordination is not surprising since the molar ratio of sulfur to nitrogen in the proteins is relatively stable (1:25; Koprivova et al. 2000, Rennenberg 1984, Smith 1980). Sulfur deficiency is known to reduce nitrogen use efficiency and vice versa (Fismes et al. 2000). It has been shown that nitrate uptake and activity of nitrate reductase are strongly reduced in maize, spinach, and oilseed rape under sulfate deficiency (Abdallah et al. 2010, Friedrich & Schrader 1978, Prosser et al. 2001). Moreover, sulfate deficiency has been shown to cause accumulation of nitrate in maize, wheat, and oilseed rape (Dietz 1989, Gilbert et al. 1997, McGrath et al. 1996). The results described in this chapter indicate that this phenomenon is also common among arabidopsis accessions in response to sulfate limitation. Koprivova et al. (2000) have shown that, in arabidopsis, the flux through the sulfate metabolic pathway measured as incorporation of [^{35}S]sulfate into proteins (after plant incubation with [^{35}S]sulfate) increased after addition of nitrogen containing compounds to the growing medium. Addition of O-acetylserine (OAS) – a precursor necessary for biosynthesis of cysteine – to the growing medium was shown to cause the accumulation of APR mRNA and protein (Koprivova et al. 2000). Therefore, OAS was proposed to play a signalling role in the co-ordination of nitrate and sulfate metabolic pathways.

The primary metabolite profiling in three week old seedlings grown in different nutrition regimes revealed that in this experiment sulfate limitation altered mainly the metabolism of carbohydrates and amino acids (Table 4.1). This is in agreement with previous results showing that continued sulfate limitation results in general reduction of metabolic activities which are shown as decrease in chlorophyll, RNA, and total protein followed by reduction in photosynthetic activity and decreased biomass production (Hirai et al. 2004, Nikiforova et al. 2005). The decrease in the total protein during sulfate deficiency is not surprising since about half of all internal sulfur is allocated into proteins (Blake-Kalff et al. 1998).

Sugars were altered in all accessions analysed, but by different environmental conditions and in various ways (Table 4.1). Apart from the crucial role of sugars in the global metabolism of the cell (they are the starting point for almost all biosynthetic and energy-producing pathways in the cell) they play also an important role in the regulation of various processes such as seed and embryo development, timing of flowering, photosynthesis, β-oxidation and more (Cakmak et al. 1994, Lloyd & Zakhleniuk 2004, Smeekens & Hellmann 2014). Fixed carbon is required to convert inorganic nitrogen and sulfur into amino acids, nucleotides and cofactors. In plant metabolism, OAS is formed from acetate and L-serine, thus linking the assimilatory sulfate reduction with carbohydrate and nitrogen metabolism (Kopriva et al. 2002). Additionally, it has been long known that the sulfate reduction pathway is regulated by light (Kopriva et al. 1999). The induction of APR mRNA accumulation and activity in plants kept in the dark could be mimicked by addition of sucrose to the growing medium (Kopriva et al. 1999). Similarly, glucose was also able to induce the expression of APR (Hesse et al. 2003).

Sucrose was accumulated in response to either sulfate limitation or induced sulfate starvation in all accessions except Edi-0 (Table 4.1). In nitrogen metabolism, when nitrate is limited, carbohydrates, starch, and soluble sugars accumulate in photosythetically active organs because they cannot be used for further synthesis of nitrogenous compounds (Sun et al. 2002). The accumulation of carbon results in an inhibition of photosynthesis (Sun et al. 2002). Since carbon skeletons are necessary to complete sulfate assimilation by production of cysteine and sulfate metabolism is closely related to nitrate metabolism the mechanism described above could be also involved, at least to some extent, in sulfate metabolism. This is further supported by the fact that sulfate deficiency is known to reduce nitrogen use efficiency (Fismes et al. 2000; see above). Interestingly, in Edi-0 selected as a representative of class 2, sucrose was not significantly altered by any of the nutrition regimes analysed. Instead, fructose and glucose were accumulated in response to sulfate starvation. Glucose and fructose are not direct end products of photosynthesis, but they are produced by the action of invertases on sucrose (Koch 2004). Therefore, it seems that the role of sugars in the regulation of sulfate assimilation differs in Edi-0 (and perhaps other accessions from class 2) compared to the accessions from the other classes. However, analysis of starch and chlorophyll concentration in these accessions would be necessary to better understand this phenomenon.

Myo-inositol accumulates in all accessions (Table 4.1) suggesting an induction of complex regulatory processes in response to sulfate deficiency. It accumulates either in response to sulfate starvation only or in both sulfate starvation and limitation. In plant cells it

forms a structural basis of a number of lipid signalling molecules that function in diverse pathways. It is involved in stress response and the regulation of cell death (Valluru & Van den Ende 2011). *Myo*-inositol is involved in the biosynthesis of L-ascorbic acid (Davey et al. 2000, Gallie 2013). It is also involved in cell wall biosynthesis and auxin perception (Gillaspy 2011, Valluru & Van den Ende 2011). The attempts to genetically engineer *myo*-inositol and its derivatives have been undertaken to improve crop tolerance to abiotic stress (Sengupta et al. 2012). Therefore, if *myo*-inositol will be also significantly altered in the metabolite profiling of older plants grown in hydroponic cultures, further investigation of its accumulation could uncover results useful for improving crop productivity. However, this will not be straightforward because of the multiple functions of this molecule.

Accessions from class 4 as well as Cvi-0, which was selected as a representative of this class, are well adapted to sulfate deficiency. Under sufficient sulfur supply accessions from class 4 include rather small plants and they accumulate sulfate containing compounds rather than anions, especially in shoots (Figure 4.3B). Under sulfate deficiency they showed the highest production of root biomass and lowest decrease in shoot biomass especially in response to sulfate limitation (Figure 4.5). They also showed the highest biomass allocation to the roots. It has been known for a long time that sulfate metabolism is controlled not only by the external availability of sulfate but also by metabolic demand for reduced sulfate (Lappartient & Touraine 1996, Lewandowska & Sirko 2008, Rouached et al. 2008, Westerman et al. 2001). Plants supplied with additional forms of reduced sulfur, such as H_2S, cysteine or GSH show a strong decrease in sulfate uptake and accumulation (Lappartient et al. 1999, Westerman et al. 2001). Therefore, it might be hypothesised that the high concentration of GSH and cysteine in accessions from class 4 may mask the demand for external sulfate which results in a delay of the induction of the sulfate starvation response in accessions from this class. However, more analyses are needed to examine this hypothesis. In contrast, accessions from class 3, with Col-0 as a representative, included rather big plants. However, they seemed to be the least adapted to sulfate deficiency. Despite the high accumulation of anions (rather than sulfur containing compounds) under normal sulfate supply, this class is the most severely affected by sulfate deficiency in almost all traits analysed.

The comparison of the experiments described in this chapter with the analysis of natural variation in response to nitrogen deficiency described by Ikram et al. (2012), on which the experiments described here were based, revealed that plant response to sulfur and nitrogen deficiency are independent. The hierarchical ascendant clustering of similar set of accessions analysed in both experiments revealed different composition of particular classes in

the two experiments (Ikram et al. 2012; Figure 4.3A). Accessions classified in class 2 (Edi-0, Alc-0, Ge-0) in the analysis of natural variation in response to nitrogen limitation described by Ikram et al. (2012) are the only exception. In the analysis of natural variation in response to sulfate availability described in this chapter all of these accession belong to the class 4. However, there are some significant differences in the characterisation of the two classes. In the analysis of natural variation in response to nitrogen limitation these accessions are characterised as big plants (Ikram et al. 2012) whereas in the experiments described here they belong to a class including small plants. On the other hand, class 2 from the analysis of Ikram et al. (2012) is considered as well adapted to nitrogen starvation because of high rate of biomass allocation to the roots. In the experiment described in this chapter class 4 was also suggested to group best adapted accessions from the entire collection. Moreover, in the analysis of Ikram et al. (2012) Col-0 belongs to class 4 which, similarly to the analysis described in this chapter (Col-0 belongs to class 3), was characterised as the least well adapted. However, in the analysis of Ikram et al. (2012) Col-0 was considered as a small plant, whereas in the analysis described in this chapter it was a big plant.

The main difference in morphological traits between the two experiments was the biomass allocation in response to sulfate and nitrate limitation (Ikram et al. 2012; Figure 4.1). In the experiment described in this chapter both sulfate limitation and starvation resulted in higher root biomass and thickness at the expense of shoot biomass (Figure 4.1). However, in the experiment of Ikram et al. (2012) nitrate limitation resulted in low root biomass and thickness and low shoot biomass whereas nitrate starvation resulted in an increase in root biomass and thickness and a severe decrease in shoot biomass. The authors concluded that this contrasting response to limitation *versus* starvation is due to plant response not only to the external availability of nitrate, but also its allocation between shoots and roots (Ikram et al. 2012, Zhang & Forde 2000). Moreover, since nitrate is known to act as a signal of nitrogen status of the plant (Vidal et al. 2010, Zhang & Forde 2000) which is partially mediated by auxin (Krouk et al. 2010) they assumed that the difference in response to nitrate limitation and starvation is due to alteration in the sensing of the nitrate signal (Ikram et al. 2012). Additionally, the decrease in root biomass and thickness in nitrate limitation was directly, negatively correlated with SRNO3 (Ikram et al. 2012). This phenomenon was not observed in sulfate limitation where the morphological traits did not show significant correlations with metabolic traits. Neither nitrate deficiency (Ikram et al. 2012) nor sulfate deficiency resulted in changes in primary root length.

4.7 Conclusion

The analysis described in this chapter provided a general profile of adaptation to sulfate deficiency in *Arabidopsis thaliana*. It is one of the first case studies in which global metabolic changes were analysed in such a big collection of accessions. The analysis of morphological traits revealed that increased root growth, often at the expense of shoot biomass, is a common plant response to sulfate limitation. The hierarchical ascendant clustering of the collection of 25 arabidopsis accessions revealed four different classes of accessions characterised by different patterns of response to different sulfate availability. Subsequently, accessions selected as characteristic for each of these patterns were analysed in more detail to get better understanding of each of these strategies. Traits specific for sulfate metabolism were analysed and the repertoire of primary metabolites that change in response to sulfate and nitrate deficiency was explored. It was possible to identify accessions that are better adapted to nutrition stress than others. Moreover, the metabolic changes which may lead to changes in the expression of the genes encoding enzymes and regulatory factors were identified. The inhibition or induction of particular pathways due to changes in expression profiles and the activity of regulatory molecules newly synthesised in response to environmental fluctuations may lead to further changes in biochemical reactions making the general plant response more complex on each level. This study provided unprecedented insight into metabolic changes as well as reprograming and regulation of metabolism during sulfate limitation. Additionally, important targets for improving crop tolerance to sulfate deficiency were identified. However, these targets would require further analysis before commercial application.

Chapter 5:

Genome-Wide Association Study on the Accumulation of Nitrate, Phosphate, and Sulfate in Arabidopsis thaliana

5.1 Introduction

The analysis of natural variation in the response of arabidopsis to changes in the availability of nitrate and sulfate described in Chapter 4 is important in the investigation of plant adaptation strategies to the different environmental conditions. It allowed identification of crucial processes and metabolites involved in the control of sulfate and nitrate homeostasis. This knowledge may be used to develop new breeding strategies leading to even more effective food production. However, at this stage these analyses will not result in the discovery of new genes and regulatory factors responsible for these adaptations unless additional approaches such as gene-expression profiling are undertaken. Therefore, to examine the genetic architecture underlying the adaptation mechanisms genome-wide association study (GWAS) was conducted to identify regions of the genome at which genetic variation is associated with the accumulation of nitrate, phosphate or sulfate in arabidopsis leaves.

The relationship between genetic and phenotypic variation observed between individuals within the species is interesting from two general points of view. First, it can be useful from an ecological and evolutionary perspective since it leads to the identification of allelic variants crucial for the adaptation to specific environmental conditions. Secondly, analysis of natural variation allows the discovery and characterisation of individual genes (Koornneef et al. 2004). Genetic variation among arabidopsis accessions was identified for a number of traits including plant response to biotic and abiotic stress (Hannah et al. 2006, McKay et al. 2003, Nemri et al. 2010, Todesco et al. 2010), traits related to plant development (Alonso-Blanco et al. 1999, Coupland 1995, Juenger et al. 2000), plant physiology (Borevitz et al. 2002, Botto et al. 2003, Loudet et al. 2003) and metabolism (El-Soda et al. 2014, Huang et al. 2012a). The accessions that differ in genotype are screened for phenotypic differences. Subsequently, any identified phenotypic differences are linked back with their causative loci via various mapping approaches (Korte & Farlow 2013).

Quantitative trait loci (QTL) mapping is one of the classic approaches to identify regions of the genome that co-segregate with a trait of interest in Recombinant Inbred Line (RIL) families (Koornneef et al. 2004). In principle, a trait of interest is quantified in the progenies derived from the cross between two different parental accessions. The values are then compared with the genetic markers of the progeny to search for the regions in the genome that show statistically significant associations between polymorphism and the variation in the trait (Koornneef et al. 2004). This approach was successfully used to identify regions in the genome responsible for the variation in sulfate accumulation in the RIL

population derived from the cross between Bay-0 and Sha (Loudet et al. 2007). Some aspects of the characterisation of the QTLs identified in this analysis are described in Chapter 3 of this thesis. However, despite the fact that QTL mapping remains a very powerful tool to discover genes responsible for natural variation, it has its limitations. One of the major obstacles is the fact that in the RIL population the genetic variation is restricted to only the two parental accessions (Koornneef et al. 2004). Additionally, the genotyping density of the RIL population and the relatively low number of recombination events in the mapping population are important factors limiting the resolution of QTL. Various approaches have been undertaken to overcome these problems. RIL populations have been generated by using an advanced inter-cross design (AI-RIL) to increase number of recombination events which improves the resolution of QTL (Balasubramanian et al. 2009). Additionally, to increase allelic diversity within the population multiple genetically diverse accessions are inter-crossed before establishing the RILs as in MAGIC (Multi-parent Advanced Generation Inter-Cross) or AMPRIL (arabidopsis multi-parent RIL) lines (Huang et al. 2011, Kover et al. 2009). Nevertheless, the allele frequencies and combinations present in such populations will always be different from those in the natural population. Therefore, more advanced approaches are needed to overcome the major limitations of QTL mapping. An alternative approach suggested as a solution for the QTL limitations was GWAS (Korte & Farlow 2013, Weigel 2012).

Taking advantage of a large number of historic recombination events and mutations that have occurred within the population of arabidopsis, the GWAS approach is based on looking for associations between the DNA sequence variants present in an individual's genome and the phenotype of interest (Nordborg & Weigel 2008, Weigel 2012). In other words, GWAS associates phenotype with single-nucleotide polymorphism (SNP) data from a diverse population of accessions (100 accessions or more) to identify loci that correlate allele frequency to phenotypic variation (Atwell et al. 2010) or adaptation to particular environmental conditions (Lasky et al. 2012). This technique was pioneered in human genetics (Hirschhorn & Daly 2005) nearly ten years ago. Since then it was successfully applied in model organisms including arabidopsis (Aranzana et al. 2005, Verslues et al. 2014) and mouse (Flint & Eskin 2012) and non-model organisms including rice (Huang et al. 2012b), maize (Tian et al. 2011), and cattle (Olsen et al. 2011).

GWAS, also called Linkage Disequilibrium (LD) mapping, overcomes the two major limitations of QTL mapping. It can take account of a large portion of natural variation in a species and localize the associations to much smaller genomic regions. The array-based re-sequencing study in arabidopsis provided large collections of hundreds of thousands of SNPs

that could be used for GWAS (Horton et al. 2012, Kim et al. 2007). The whole-genome sequencing revealed c.a. 7 million SNPs segregating within a worldwide population (Cao et al. 2011). However, Kim et al. (2007) by using the tag SNP selection algorithms and "hide-the-SNP" simulations have shown that GWAS will require only 40%-50% of the observed SNPs. Therefore they designed an Affymetrix genotyping array that includes 250, 000 SNPs segregating within the population (Kim et al. 2007). This genotyping array is currently available as a collection of genotyped SNPs in arabidopsis. Although it is only a small percent of all observed SNPs, the 250, 000 SNPs lead on average to one SNP every 600bp and tag almost all of the non-repetitive genome (Kim et al. 2007, Korte & Farlow 2013). The high density of SNPs has made GWA study in arabidopsis much more powerful than in humans.

GWAS generates lists of SNPs associated with variation in the studied phenotype (Korte & Farlow 2013). Each association has a probability (P)-value defining the strength of the association. However, it is possible that some of the SNPs associated with a given phenotype might be false positive. Taking into account that GWAS is based on LD – a non-random association of alleles at different loci – which is a crucial aspect of genetic variation in natural populations (Kim et al. 2007) it has been established that in arabidopsis the SNP might be linked to a causal gene within 20kb around it (Atwell et al. 2010). Therefore, for some phenotypes, it is possible to predict which genes might be responsible for natural variation based on existing functional knowledge and select the candidates for follow-up studies.

While the general principles of GWAS are straightforward, there are a number of confounding factors that might affect its outcome (Korte & Farlow 2013). False-positive associations might arise from population structure as the individuals in a population are not equally distantly related to each other. This is a limitation especially for traits important in the adaptation to the local environment (Nordborg & Weigel 2008, Weigel 2012). Mixed models are a powerful method that corrects for population structure by accounting for the amount of phenotypic co-variance which is due to genetic relatedness (Yu et al. 2006). Mixed models were first developed in the field of animal breeding. They were successfully applied to GWAS and have led to a significant reduction of false-positive associations (Atwell et al. 2010, Zhao et al. 2007). However, it should be stressed that these methods do not always work and the additional control of GWAS outcome is necessary (Weigel 2012). False-positive associations may also arise due to multiple testing of thousands of SNPs. Conversely, traits with complex architecture, where many different genes affect the same trait, may cause false-negative associations (Weigel 2012). The significance of a particular allele in the phenotypic variation

within the population depends on its frequency and effect. Consequently, rare alleles present difficulties to GWAS even if they have a large phenotypic effect (Nordborg & Weigel 2008). Non-additive effects among loci as well as alleles arising repeatedly at a single locus with a similar effect on gene function (allelic heterogeneity) may also potentially lead to false-negative associations (Bergelson & Roux 2010).

Several studies have been published which show that GWAS is often successful in arabidopsis and may result in discovery of previously undescribed genes (Chao et al. 2012, Huang et al. 2012b, Todesco et al. 2010). Additional information such as functional data from mutant studies or involvement of a gene of interest in complex regulatory networks might help to prioritize GWAS candidates and increase its utility (Aranzana et al. 2005, Atwell et al. 2010, Verslues et al. 2014). Additionally, QTL mapping in an experimental population can indicate the region of the genome which should be considered for the location of the locus of interest (Brachi et al. 2010, Nemri et al. 2010). Therefore, QTL mapping and GWAS can be used as complementary approaches that mitigate each other's limitations when applied together. Such a combined approach – referred to as nested-association-mapping – takes advantage of the increased resolution of GWAS and allows GWAS limitations to be overcome because QTL mapping is more robust to confounding factors (Nordborg & Weigel 2008). Nested association mapping was successfully developed in maize (Yu et al. 2006) and in theory could be also applied in arabidopsis (Nordborg & Weigel 2008). Although QTL mapping and GWAS proved themselves to be successful, especially in arabidopsis, both of them are based on correlations in form of linkage disequilibrium described above. Therefore, they cannot prove the causality by themselves and they need to be followed by the functional analysis of the candidate genes (Weigel 2012).

Arabidopsis is an ideal organism for such a study. Its life cycle can be as short as six weeks and most importantly it is self-compatible (Weigel 2012). Therefore, various accessions can be maintained as inbred lines. Thus, it is possible to genotype them once and phenotype repeatedly for the same or different phenotypes. Both of these properties greatly facilitate genetic studies (Korte & Farlow 2013). Arabidopsis is broadly distributed throughout the northern hemisphere (Figure 5.1). It originated from Eurasia and North Africa (Hoffmann 2002). Subsequently, it has been introduced throughout much of the rest of the world. It grows in a range of climatic habitats and is continuously exposed to various selective pressures (Hoffmann 2002). Therefore, its genome contains extensive diversity within the global population. Part of this genetic diversity is associated with large phenotypic variation (Atwell et al. 2010) and adaptation to the local conditions (Hancock et al. 2011). Currently, the analysis

of natural variation in arabidopsis is used continuously to reveal new insights into various biological processes and the entire genus is used to explore fundamental questions of evolution (Bergelson & Roux 2010).

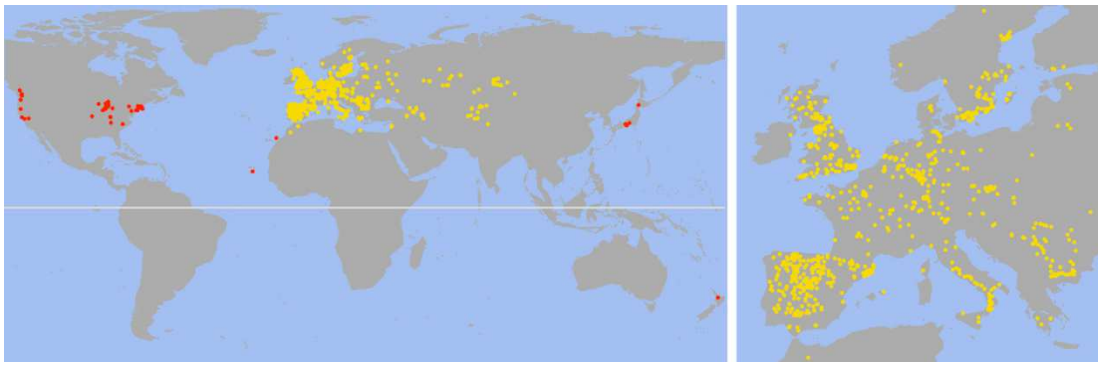


Figure 5.1 ***Worldwide distribution of natural accessions of Arabidopsis thaliana***

Distribution of 7,000 *Arabidopsis thaliana* accessions collected from worldwide locations available for the analysis of natural variation. Locations marked in yellow are considered to be native and the ones marked in red are presumably introduced (Weigel 2012).

In this chapter I describe GWAS on 317 *Arabidopsis thaliana* accessions which are a subset of 360 accessions from the Justin Borevitz core collection (Baxter et al. 2010). This analysis was conducted to explore the genetic architecture of traits related to nutrient use efficiency in the model organism. I screened all these accessions for the leaf concentration of nitrate, phosphate, and sulfate. To identify the regions of the genome at which genetic variation is associated with changes in concentration of the three macronutrients in leaves GWAS was conducted using two different tools: an online pipeline called Matapax (Childs et al. 2012) and Efficient Mixed-Model Association (EMMA) method (Kang et al. 2008). The Matapax GWAS was conducted by me. The EMMA analysis was performed by David E. Salt and Dai-Yin Chao from University of Aberdeen. To verify the output of the GWAS, I analysed T-DNA insertion lines to investigate whether disruption of selected candidate genes affects accumulation of the three macronutrients tested.

5.2 Materials and Methods

5.2.1 Growth conditions in the soil

Plants used for generation of the dataset for GWAS and T-DNA insertion lines grown for anion analysis shown in Figures 5.7 and 5.8 were grown for five weeks in a Controlled Environment Room (CER). First, the dry seeds were surfaced sterilised for up to four hours in a vacuum desiccator using chlorine gas which was generated by mixing 125ml of sodium hypochlorite with 2.5 ml of 12 M HCl. Seeds were then mixed with 0.1% sterile agarose (Sigma Aldrich) and seeded on plates with MS medium (see Chapter 2). Subsequently, plates were stored at 4°C in the dark for three days. After that time they were transferred to a CER at 22°C under 16-h-light/8-h-dark cycles where they were left horizontally up to one week during which the seeds germinated. One week old seedlings were transferred from plates to the 40-cell tray (21 x 35cm) filled with Levington Horticulture soil mix (Figure 5.2) and grown for another four weeks in CER 10-h-light/14-h-dark cycles at constant temperature 22°C, 60% relative humidity, and light intensity of 160 µE s^{-1} m^{-2}. Twice a week each tray was supplied with 1 l of ¼ Hoagland solution (Table 5.1M) in order to provide sufficient nutrients to the plants over the growth period.

Table 5.1M ***The composition of ¼ Hoagland solution***

¼ Hoagland solution			
Macroelements [mM]		**Microelements [µM]**	
Element	**Concentration**	**Element**	**Concentration**
		H_3BO_3	12.5
KNO_3	1.25	$MnCl_2$	1
$Ca(NO_3)_2$	1.25	$ZnSO_4$	1
KH_2PO_4	0.25	$CuSO_4$	0.25
$MgSO_4$	0.5	$NaMoO_4$	0.03
		KCl	12.5
		FeEDTA	10

The concentration is shown per 1 l of solution.

5.2.2 Experimental design

Each accession was grown in five biological replicates. In the soil, plants were grown in seven independent sets in blocks containing up to seven accessions each. Overall, each accession was grown once. In each block Col-0 was included as internal reference and always grown in the first column. Additionally, up to seven accessions were grown in remaining columns within the block. The number of blocks within a set varied from four to ten depending on the space availability (Figure 5.2). After four weeks of vegetative growth a young leaf was harvested individually for each plant. Only four plants were used for sample composition. If five plants were available one was left out randomly. If some of the plants looked different than the others from the same accession they were not harvested. Plant material was weighed, frozen in liquid nitrogen and stored at -80°C for further analysis.

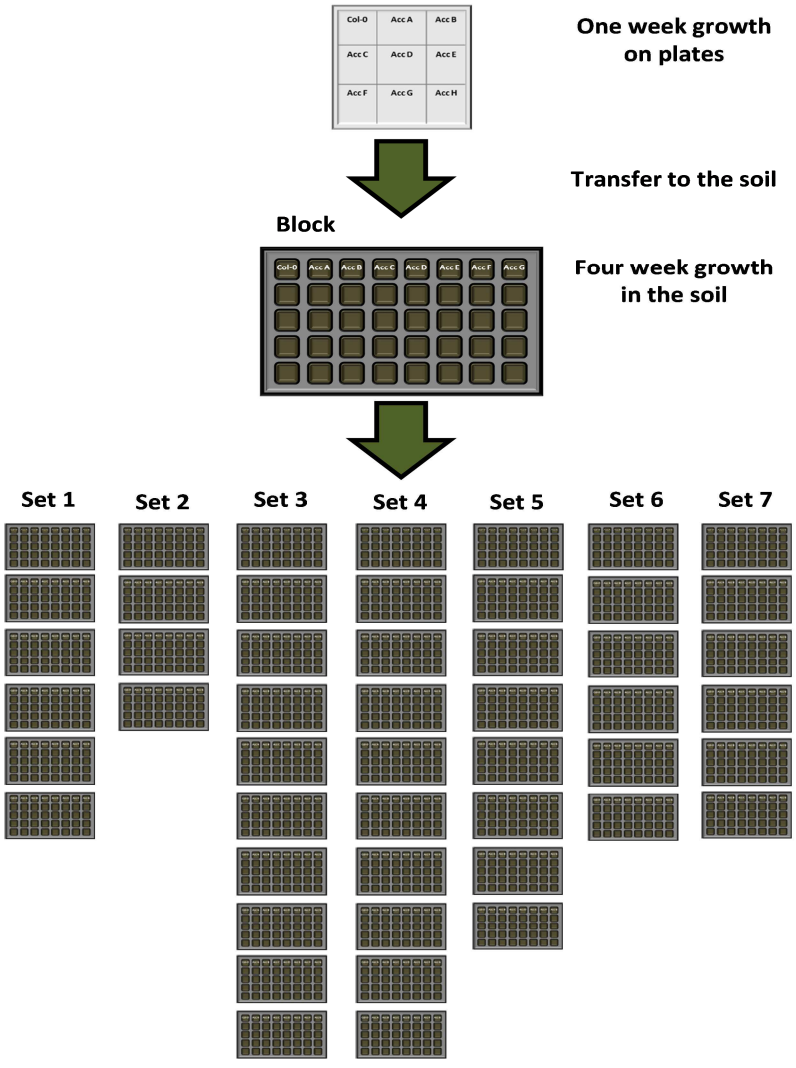


Figure 5.2 ***Experimental design of growth of 317 Arabidopsis thaliana accessions***

5.2.3 Growth conditions on plates

The seeds of all the T-DNA insertion lines were obtained from The European arabidopsis Stock Centre (NASC). The T-DNA insertion lines grown for anion analysis shown in Figures 5.6 and 5.10 were grown on plates for two weeks. First, dry seeds were surfaced sterilised as described above. They were mixed with 0.1% sterile agarose (Sigma) and seeded on plates with MS medium (see Chapter 2). Medium for plants grown for the selection of homozygous lines (shown in Figure 5.6) was supplemented with 50 µg ml^{-1} kanamycin. Medium for plants grown on sucrose (shown in Figure 5.10) was supplemented with 0.8% sucrose. Subsequently, plates were stored at 4°C in the dark for three days. After that time they were transferred to the CER at 22° under 16-h-light/8-h-dark cycles where they were grown vertically for two weeks. The photon flux density was *ca.* 125 µmol s^{-1} m^{-2}.

5.2.4 Genome-Wide Association mapping methods

Two different tools were used for GWAS: web-based pipeline called Matapax (Childs et al. 2012) with GAPIT method for correction for population structure was used by me; David E. Salt and Dai-Yin Chao from University of Aberdeen ran GWAS using R package with the EMMA method for correction for population structure. EMMA takes an advantage of linear mixed-models which were shown to effectively correct for the population structure in model organisms (Yu et al. 2006). Taking into account that two genetically related individuals are more likely to be correlated than genetically different individuals, linear mixed-models incorporate pairwise genetic relatedness between every pair of individuals directly in the statistical model. This strategy allows reduction of the rate of false positive associations and increases the power of the analysis (Yu et al. 2006, Zhao et al. 2007). EMMA is the most frequently used tool for GWAS in arabidopsis. However, it was not possible to use it within the Matapax pipeline with datasets greater than 200 accessions (Childs et al. 2012) and using it within R package requires specialist statistical knowledge.

Using a web-based tool such as Matapax is much more user friendly and does not require a specialist statistical and programing knowledge. Genome Association and Prediction Integrated Tool (GAPIT) was used to correct for population structure (Lipka et al. 2012). GAPIT is a mixed-model association tool that implements EMMA and enhances it with additional algorithms that increase the statistical power and decrease the computational time (Childs et al. 2012). In the GWAS described here the kinship matrix (K matrix) was used to control for the population structure. It contains the pairwise relationship between all represented genotypes and calculates the relatedness of each genotype to others (Kang et al. 2008). Additionally, the

marker dataset from arabidopsis POLYmorphism DataBase (AtPolyDB; Atwell et al. 2010, Horton et al. 2012) was used, which is one of the three marker sets available within Matapax. It contains the SNPs determined using a 250k Affymetrix genotyping chip providing a high resolution map of genetic variation across a great number of accessions. The output from Matapax is a graphical genome browser with the gene annotations from The Arabidopsis Information Resource (TAIR) database and detailed results table. The results table includes chromosome, position, p-value, gene annotation and a specific polymorphism. Matapax output does not include the Manhattan plots which are a common way to visualize the data and inspect the inflation of P-value (Korte & Farlow 2013). However, the result table was downloaded and the Manhattan plots were created in Microsoft Excel.

Identifying meaningful associations from such a large amount of pair-wise tests as these obtained from GWAS requires the application of additional approaches (Noble 2009). First of all, since GWAS involves a large number of accession-phenotype tests, the produced P-values require an application of a multiple testing correction method to control the number of false-positive associations (Noble 2009). To address that issue a False Discovery Rate (FDR) estimation (Banjamini-Hochberg method) was employed in the Matapax pipeline (Childs et al. 2012). In principle, the FDR procedures are intend to control the proportion of incorrectly rejected null hypotheses within a particular analysis (Benjamini & Hochberg 1995). The simple FDR estimation method is sufficient for many studies. However, there are cases in which a more strict method is required (Noble 2009). The most stringent method of multiple correction testing commonly used in GWAS is a Bonferroni adjustment. It is based on minimizing the Family-Wise Error Rate (FWER) to one percent. In other words, it gives 99% assurance that none of the observed scores is drawn according to the null hypothesis (Noble 2009). Usually, 5% significance threshold is used for Bonferroni adjusted P-values to extract the true associations (Atwell et al. 2010) and was applied also in the analysis described in this chapter. To calculate it one needs to divide the threshold by the number of separate tests used in the analysis (Noble 2009).

5.2.5 Genotyping of T-DNA lines

The DNA from plant tissue was isolated as described in Chapter 2. The genotyping of T-DNA lines was carried out on genomic DNA by PCR amplification using the gene specific primers listed in Table 5.2M and the left-border SALK primer for Salk lines or left-border SAIL primer for Sail lines. Genotyping of T-DNA insertion mutants was carried out in two steps. In the first reaction, gene specific primers listed in the table were used to amplify genomic DNA

of the wild-type sequence. This PCR gave a product in heterozygous mutants or wild-type plants, but not in homozygous mutants. In the second step, one gene specific primer was used in combination with SALK_LB1 or SAIL_LB1 primer respectively. The gene specific primer depended on the orientation of the T-DNA insert in the gene and is highlighted in the Table 5.2M.

5.2.6 Genotyping polymerase chain reaction

The GoTaq Flexi Kit (Promega) was used for genotyping PCR. PCR reactions contained 50-500 ng of gDNA or 1 pg - 50 ng plasmid DNA as template, 1x polymerase reaction buffer, 0.1 U of GoTaq polymerase, 0.4 µM of each primer and 500 µM dNTPs. PCR amplifications were carried out in a DYAD Thermal Cycler (Biorad) or Mastercycler Pro (Eppendorf). The standard PCR amplification protocol included an initial denaturation step of 95°C for 2 min, followed by 35 cycles of: 94°C for 30 s, 50-60°C (Primer Tm-3°C) for 30 s, 72°C for 1 min for each 1kb of target DNA product and a final elongation step at 72°C for 10 min. After PCR reaction the PCR products were separated by agarose gel electrophoresis using 1% gel.

Table 5.2M ***Gene specific primers used for genotyping***

Name	T-DNA line	Gene	Sequence 5' -> 3'
1Fw	SALK_098851	AT3G45570	ATGATCAGGCATATCTATGC
1Rw			gcgaggaaaccacagagc
2Fw	SALK_009859	AT5G23000	AACAACAACACCTCGAACG
2Rw			aagcatttgtattctagtagg
4Fw	SALK_021437	AT1G10670	AAATGAGCTAGTTGAAAAGG
4Rw			ggtgtcagggaagcacc
5Fw	SALK_030899	AT2G17260	CAATCATCTTTACTCATTTGG
5Rw			accaagaacaatgatgaagc
6Fw	SALK_036267	AT4G17500	CGTTCCTAACCAAACCCTAGC
6Rw			tcctactcttctccctgctcc
7Fw	SALK_071872	AT3G16600	GCAACAATGTTCCTTTTACG
7Rw			taaagacaatagtatcagtgg
8Fw	SALK_092114	AT2G01760	AGTTGGTCTTCTTGAAATGG
8Rw			ctcagagtcattactctgc
10Fw	SALK_112450	AT4G14370	GCATCTATTAAGATGTGACG
10Rw			atggtatcaacaaaacttacc
11Fw	SALK_116156	AT4G15530	CAGTCCCTGAAGTGTCC
11Rw			gcttgaacgaccaagtcg
12Fw	SALK_033906C	AT1G06450	TGTATTGGAAGTTTTGCCATTATC
12Rw			ccgtgagtctctggtagcttg
13Fw	SALK_070088C	AT4G31990	CGCGGTAACAGAGAAGTGG
13Rw			atcacaatccactagtatcc
15Fw	SALK_124733C	AT1G06470	TCCATTACAGGGGTACTCGTG
15Rw			aagaagaaatattggggctgc
16Fw	SALK_092343C	AT4G31360	GAGGTTGAGAAGGAAGAGC
16Rw			tcttgtcttgaaagcattgc
17Fw	SALK_080084C	AT5G08590	TTGTTGGCTGAAAGTTGAGG
17Rw			tttgaacatacacactgagc
18Fw	SALK_044043C	AT5G22250	CCTCAACGAAATACAAATCC
18Rw			tctccgccgtggaaacc
19Fw	SALK_061581C	AT3G51770	CAAACTGGTTTTCTTCCACC
19Rw			gcttgtgattcgcatttggc
20Fw	SALK_110194C	AT5G43370	TTCTTTACCGATGCGTACG
20Rw			gcgtattcagacatgatgg
21Fw	SALK_043036C	AT1G32450	TCTACTAATAAATCATCAAACC
21Rw			caattatagtacttatgtaaagc
22Fw	SALK_065920C	AT2G43900	TTACGGAATTGACAGGACC
22Rw			caccggaggagatacagg
23Fw	SALK_081127C	AT5G44720	TCTGTTCCTCAGGCTACC
23Rw			tgaaactgtagtcttgatgc

Name	T-DNA line	Gene	Sequence 5' -> 3'
24Fw	SALK_099679C	AT5G54310	TTTTCTCTGCCGTTTGACC
24Rw			gcaacgagaataaatcttcc
25Fw	SALK_138643C	AT2G38940	AAGTACTTCCAGTGTCTCC
25Rw			cttggccacgtcaagtgc
27Fw	SALK_088586C	AT5G43350	GCCCATGGCGTCTAAGG
27Rw			gtcatgaccaccctttgc
28Fw	SALK_117672C	AT5G10800	CTCCCCTTCTTTGAAATGG
28Rw			gagttctttctcgatcacc
29Fw	SALK_129778C	AT2G38920	GGTATCTCATTTGATCCAAG
29Rw			atatctacatggagaagagg
30Fw	SALK_138009C	AT2G32830	ATGAATTTGAATTTGATTGTGC
30Rw			gctcggtttcggcatagg
31Fw	SALK_103821C	AT5G39080	CAATGAACCATTCCCCTCC
31Rw			attatgcatactaatacatacg
33Fw	SALK_122868C	AT5G03730	CGCTGTTATGGTTGCTCC
33Rw			atatgaatagatatcatagtcg
34Fw	SALK_141801C	AT5G03430	CAGATTCTTCTTCTTCTCG
34Rw			cgatgatcaattggcttcg
35Fw	SALK_111336C	AT3G14400	AAATGGCTTTATCAATCTTCC
35Rw			ggatctaactaccgatgc
36Fw	SALK_151903C	AT1G11390	CAAAGCTGGAATTAAAGTCG
36Rw			gtcctatctttctgtcttcg
37Fw	SALK_097431	AT1G12110	ATCAGATACTTATCTTTAATCC
37Rw			atagacccgtgctcatcc
38Fw	SALK_120831	AT1G12140	CGAGTCAACTCACTCAACG
38Rw			gaacatgagcaactctagg
39Fw	SAIL_41_C07	AT5G10790	CGACGTGGAACAAAGTGC
39Rw			agttattgaaagatttgtaacc
40Fw	SAIL_371_D04	AT4G23750	GAGAATCAAATTCACAGAGC
40Rw			tccgacccgaatagatcc
43Fw	SALK_135953	AT1G29230	CGTATATTGCACCCGAGG
43Rw			caaactcccattctttcttcc
44Fw	SALK_007262	AT3G55780	TGTTGAGAAAGATGCTAAGC
44Rw			ttctccatcttcaatgtacc
46Fw	SALK_107550	AT3G48850	ACCGCAGCTCAGCATCC
46Rw			tgttgacaagtaaagagtcg
48Fw	SALK_133599C	AT1G65840	TCGTTTCCAGATAATCTTCC
48Rw			ccatatcaacaggacaacc
49Fw	SALK_009687C	AT1G52990	GTGAGGATGATCAAGTCC
49Rw			tgatatgggcaatacaagc

Name	T-DNA line	Gene	Sequence 5' -> 3'
50Fw	SALK_021421C	AT1G03495	CCCAACCTCAAACATTCC
50Rw			tctgaccatgttctcagg
51Fw	SALK_057196C	AT1G52430	ACTCTCACTTGCAAGAGC
51Rw			tcttgcactcgaggagg
52Fw	SALK_018778C	AT5G09930	CAAGTGACATTCTCATTTGC
52Rw			ttgagaacattagtaagagc
53Fw	SALK_009893C	AT4G14580	AGATCCAGCTCGGAAACC
53Rw			aggaagcggaagcttcg
54Fw	SALK_097424C	AT4G40010	CGTGACAATAGCTAACTCC
54Rw			cagagagagatcataaacc
55Fw	SALK_142820C	AT4G31805	GAGGATAGACAAACAAAGC
55Rw			gcggtgtatcgaacatgc
57Fw	SALK_024894C	AT1G44350	CGCCATCAAATGAAGTAACC
57Rw			aggagttatggtctttagg
58Fw	SALK_003255C	AT4G26890	TGGGATGTACGATGATCG
58Rw			gtcgtcagaggaatattcc
59Fw	SALK_115780C	AT5G59670	GTCTGGATCTAGGCTACG
59Rw			gcttatctgggacttcagg
60Fw	SALK_021591C	AT4G21230	CCCTTGCTGAAACCTTCC
60Rw			aatgtgttgtaagttgtaacc
62Fw	SALK_096310C	AT1G52890	ACATAGAACCCAATCATCC
62Rw			ttacaaattcgatccatgg
63Fw	SALK_121775C	AT5G55240	GTCCTCAAGTAGGCAAGG
63Rw			tgtccaaatcgaagaaagc
64Fw	SALK_004669C	AT3G26340	ATTGTACCAAGCATATAAGG
64Rw			tttgggtacctctcaagc
65Fw	SALK_099224C	AT1G52950	GACATTATGGATAACACAGG
65Rw			gtctgagagttgtcctcc
66Fw	SALK_039947C	AT4G30510	GGATCAAGGTTCCTTGC
66Rw			cttctatactcgctgttcg
67Fw	SALK_052482C	AT5G65500	AAGTAATGTATTTGGACAGC
67Rw			tggaaattcacctgtttcc
68Fw	SALK_047774C	AT1G67600	CACCAGCTTTGGATTTCG
68Rw			gccacactgttcttgagg
69Fw	SALK_019928C	AT5G57180	GGGATCGGAAGCTGTCG
69Rw			cgaacaaatcaatctgagc
70Fw	SAIL_512_B08	AT4G28140	CATAGGACAAGAGAGAGG
70Rw			tcctagtcgtcctaaacc
71Fw	SAIL_572_G08	AT5G57240	GGACGTTGAGGTGAAGG
71Rw			ggaaataatcagagttattcg

Name	T-DNA line	Gene	Sequence 5' -> 3'
72Fw	SAIL_418_H08	AT1G48630	TAACTCCATCAGTGTATCC
72Rw			agtcccaataggtactgg
73Fw	SALK_095121	AT2G45650	CAAGCTCTACGAGTTTGG
73Rw			ttataccgttcgattgtgc
74Fw	SALK_128101	AT5G03430	ATTGAAACTAAAGGCGACC
74Rw			ccaatctactaaatagttcg
75Fw	SALK_087356C	AT4G17490	CAGAAACTCCGTCAAATCC
75Rw			gtaccggttgtagcagc
76Fw	SAIL_549_F06	AT2G03590	CTCTTGCCTTGCCTTTGC
76Rw			tatatctgttagggtaaagg
77Fw	SAIL_675_A03	AT1G11930	AGAAGTGAAGTGTGTAAGG
77Rw			ttccaaatccatagtattgg
78Fw	SAIL_869_F08	AT2G28810	GCTGAACCAGAAAGCACG
78Rw			tgtaaatggtctatatatacc
79Fw	SALK_117130	AT2G37440	GGTATTGACTTATACGTGC
79Rw			ttctcaacaggataagttcc
80Fw	SALK_072213C	AT4G14880	CGCTAGAGCACGGTTCC
80Rw			gccatgtgactgacacc
PAPS1_1Fw	SALK_031937	AT5G03430	AGTTGCAGATTGTAACTGCCG
PAPS1_1Rw			gcccattattagtgtccgagac
PAPS1_2Fw	SALK_043060	AT5G03430	GCTCAAAATTTTCTGCACGTC
PAPS1_2Rw			cgccatcttcgtcattaaaag
PAPS1_3Fw	SALK_128101	AT5G03430	GAATTGCTCTTGACCAACCTG
PAPS1_3Rw			tgcagatactccgaaatttgg
PAPS1_4Fw	SALK_140720	AT5G03430	AGGCACCTGATGAAGAGTTTG
PAPS1_4Rw			tgtgaccggtacatagaagcc
PAPS1_5Fw	SALK_144629C	AT5G03430	GAGGTCGTGTGGGATATGTTG
PAPS1_5Rw			acagaaacgagagcagagcag
CAF1_1Fw	SALK_101782	AT1G03450	TTCATGTCCAAGACCAAAACAC
CAF1_1Rw			gatcagcttggtgttctcgac
CAF1_2Fw	SAIL_533_G10	AT1G03450	GATTAAGGTGGAGCTTCTGGG
CAF1_2Rw			ccaattaattttaacgcggtg
CAF1_3Fw	SAIL_1306_A08	AT1G03450	TGTATTGGAAGTTTTGCCATTATC
CAF1_3Rw			ccgtgagtctctggtagcttg
CAF1_4Fw	SALK_009021C	AT1G03450	TGCCAGTCAGTCATTAGCTTG
CAF1_4Rw			ttcggtcaagtttacgacacc
ST1_1Fw	SALK_088510	AT1G03470	TGGATTACATCTGATTTCG
ST1_1Rw			ctccctcattggatatgg
ST1_2Fw	SAIL_156_B12	AT1G03470	GCTGGTGTATGACGCAAG
ST1_2Rw			ccacgtaaattcgtcatgg
ST1_3Fw	SALK_126893C	AT1G03470	GTTGTCAAAGAGGCTGTCACC
ST1_3Rw			GACCAAATTTGCAGCCACTAC

Name	T-DNA line	Gene	Sequence 5' -> 3'
SALK_LB1	Salk lines	insertion	GCGTGGACCGCTTGCTGCAACT
SAIL_LB1	Sail lines	insertion	TAGCATCTGAATTTCATAA ACCACTCGATACAC

One of the primers from each pair which is underlined and bold was used in combination with either of the left-border primers to identify homozygous lines.

5.2.7 Bioinformatics tools

The arabidopsis Information Resource (TAIR) database was used for the functional annotation of the candidate genes (http://www.arabidopsis.org/); the Gene Expression Omnibus (GEO) database was used to obtain the microarray data for the *slim1* mutant (www.ncbi.nlm.nih.gov/geo; the 1001 Genomes database was used for the investigation of the natural variation in the architecture of candidate genes (http://signal.salk.edu/atg1001/3.0/gebrowser.php); the "Stress Response" tool within the gene expression search engine Genevestigator (Zimmermann et al. 2004; https://www.genevestigator.com/gv/plant.jsp) was used to investigate the candidate genes which in the analysis of T-DNA insertion lines showed differences in anion accumulation; the T-DNA Express search tool was used to obtain the T-DNA line with insertions in genes of interest (Alonso et al. 2003); the basic statistical analysis of the data presented in this chapter was performed using Microsoft Excel 2010.

5.3 Results

5.3.1 Analysis of the anion concentration in Arabidopsis thaliana accessions

To generate the dataset that could be used for GWAS I screened 317 arabidopsis accessions for concentration of nitrate, phosphate and sulfate. The population of 317 accessions used for analysis of anion concentration is a part of the collection of 360 accessions selected by Justin Borevitz and collaborators within the HapMap project (Baxter et al. 2010; Table S5.1). The set of 360 accessions was selected from 5,810 worldwide accessions based on the genotypes at 149 SNPs developed by Platt et al. (2010) and spread across the genome. This population was selected to minimize redundancy and close family relatedness. As mentioned in the introduction, these factors may affect population structure which leads to an increased rate of false-positive associations in GWAS results. More detailed description of the entire collection of 360 accessions can be found online (http://naturalvariation.org/hapmap) and in Baxter et al. (2010)..

The subset of 317 accessions used in this analysis was obtained based on seed availability and germination rate (Table S5.1). Initially, seeds of 71 Swedish accessions were provided by Matthew Box and Susan Duncan from the Caroline Dean lab at the John Innes Centre and seeds of 273 accessions, originating from a wide range of locations, were provided by David E. Salt from the University of Aberdeen. From this initial set of 344 accessions 28 did not germinate or did not develop mature plants (Table S5.1). Finally, 317 accessions were screened for anion concentration to generate the dataset which was later used for GWAS (Table S5.1). All these accessions were previously genotyped using custom Affymetrix SNP-tilling array Atsnptile 1 which contains probe sets for 248,584 SNPs (Atwell et al. 2010, Baxter et al. 2010). The SNP data are publically available and widely used for various GWA analyses (Chao et al. 2012, Li et al. 2010).

All the accessions used for this experiment were grown for five weeks in CER as described in the Material and Methods section of this chapter. Leaves of five week old plants were analysed for concentration of nitrate, phosphate and sulfate in four biological replicates using HPLC (Chapter 2). This analysis provided a large quantitative dataset which was subsequently used for GWAS to investigate the genetic architecture of nutrient use efficiency in arabidopsis. However, before that a basic verification and normalisation was applied to the dataset to examine its quality.

Table 5.1 Col-0 performance within the datasets

Anion	Average anion concentration in Col-0 [nmol mg^{-1} FW]	Standard Deviation
Nitrate	137	± 22
Phosphate	14	± 4
Sulfate	14	± 4

The average values correspond to the concentration of particular anions in Col-0 from 50 blocks grown in seven independent sets containing four biological replicates each (200 samples).

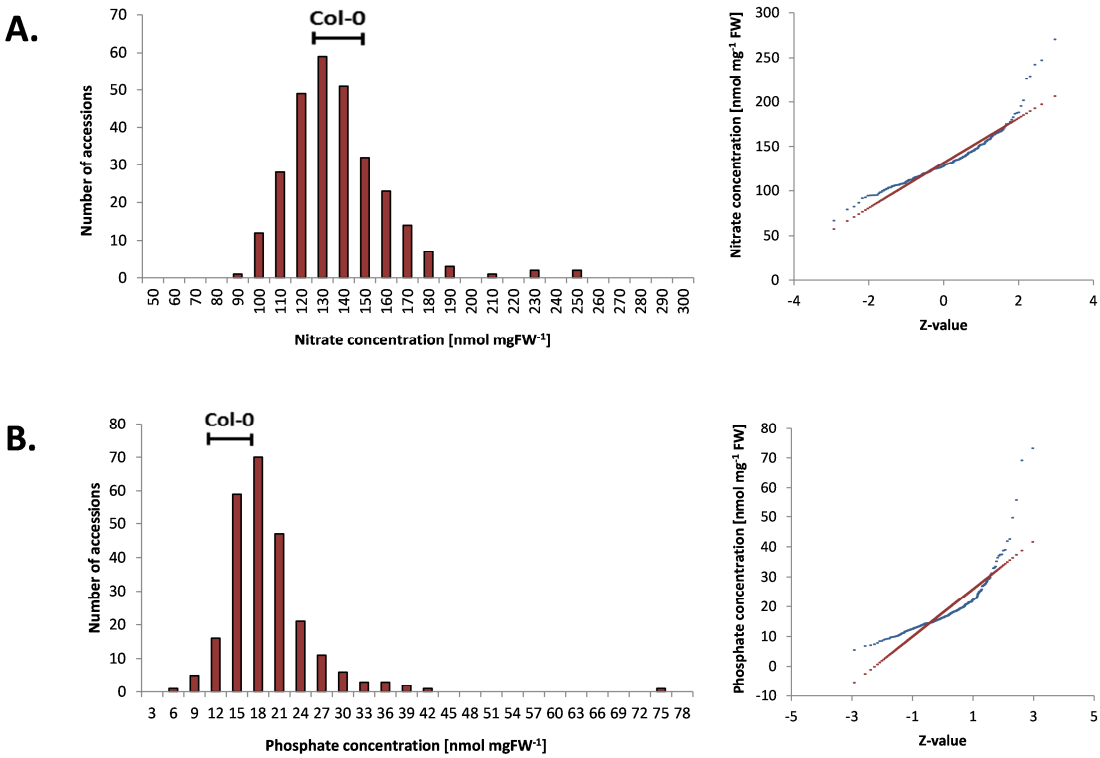
To investigate the variation between the sets of accession and between the blocks within the sets I first examined the variation between the Col-0 which was used as a reference accession (Table 5.1). There was variation between Col-0 values from block to block measured in independently grown sets. Therefore, to correct for this variation, I normalised values for all accessions within a block on the basis of Col-0 value for that block. Subsequently, I removed from each of the three datasets the accessions for which standard deviation was higher than 20% of average value. Finally, I used 284 accessions to compose the nitrate accumulation dataset, 246 accessions to compose the phosphate accumulation dataset, and 261 accessions to compose the sulfate accumulation dataset (Table 5.2).

Table 5.2 ***Summary of the macronutrient accumulation datasets***

Dataset	Accessions with lowest concentration [nmol mg^{-1} FW]		Accession with highest concentration [nmol mg^{-1} FW]		Mean of all accessions [nmol mg-1 FW] ± SD	Variation	Number of accessions in the dataset
Nitrate	Se-0	86.4	TOU.A1.12	270.3	132 ± 25	4 fold	284
Phosphate	DNAII-1	5.4	UKID-22	73.1	18 ± 8	13 fold	246
Sulfate	CS28053	6.7	Can-0	32.8	15 ± 5	5 fold	261

After the normalisation of the data I examined the frequency distribution of leaf anion concentration in all accessions analysed. The mean value of nitrate concentration fell between 130-140 nmol mg^{-1} fresh weight. Col-0 (137.17 nmol mg^{-1} fresh weight) fell in the centre of distribution (Figure 5.3A). The leaf phosphate concentration across 246 accessions varied from 5.43 nmol mg^{-1} fresh weight in DraII-1 to 73.15 nmol mg^{-1} fresh weight in UKID-22 giving 13

fold variation between accessions (Table 5.2). However, phosphate concentration in UKID-22 is exceptionally high. The accession with the second highest phosphate concentration is Kr-0 which has 41.80 nmol mg^{-1} fresh weight which is nearly two fold lower than in UKID-22. Therefore, the mean of the frequency distribution of phosphate data which is between 15-18 nmol mg^{-1} fresh weight is shifted to the left. The average Col-0 phosphate concentration (14.35 nmol mg^{-1} fresh weight) is also slightly shifted to the left side of the graph (Figure 5.3B). The leaf sulfate concentration across 261 accessions showed fivefold variation (Table 5.2). The mean of the frequency distribution of sulfate data fell between 14-16 nmol mg^{-1} fresh weight. The average Col-0 sulfate concentration (13.74 nmol mg^{-1} fresh weight) is slightly shifted to the left (Figure 5.3). In general, all these data showed a prominent variation, sufficient for GWA analysis. The analysis of the data revealed that the population is suitable for GWAS of the anion traits. However, the normal probability plot, which is the simple normality test for Microsoft Excel, showed that the datasets are not normally distributed, but slightly skewed to the right (Figure 5.3). Therefore, methods implying the correction for population structure need to be used in order to limit the rate of false-positive associations (Atwell et al. 2010, Zhao et al. 2007).



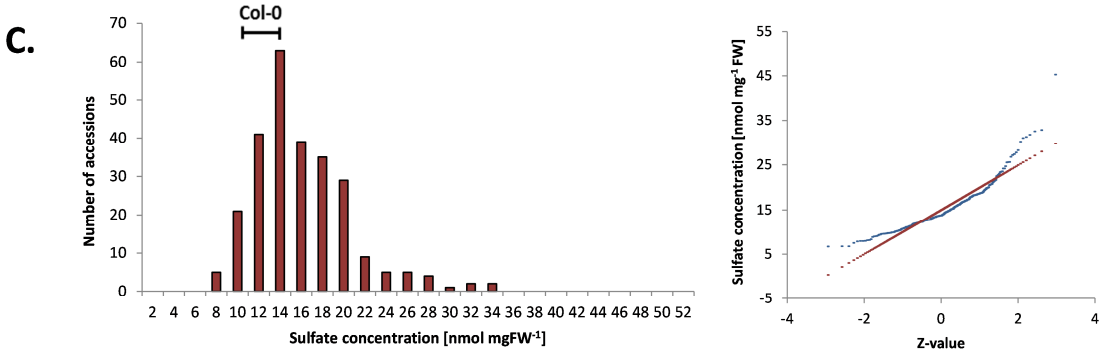


Figure 5.3 ***Anion concentration in a worldwide collection of arabidopsis accessions***

The frequency distribution (left panel) and normality testing (right panel) of the leaf nitrate concentration in 284 arabidopsis accessions (A); leaf phosphate concentration in 246 arabidopsis accessions (B); leaf sulfate concentration in 261 arabidopsis accessions (C) grown in a Controlled Environment Room (CER). Horizontal bars represent the Col-0 interval. The normal probability plot was performed in Microsoft Excel 2010. The red line represents the expected value, the blue line represents the actual data. The expected value was calculated by using cumulative distribution function (CDF) and the Excel formula NORMINV (CDF at sample point, sample mean, sample standard deviation; Harmon 2011).

5.3.2 Genome-Wide Association Study

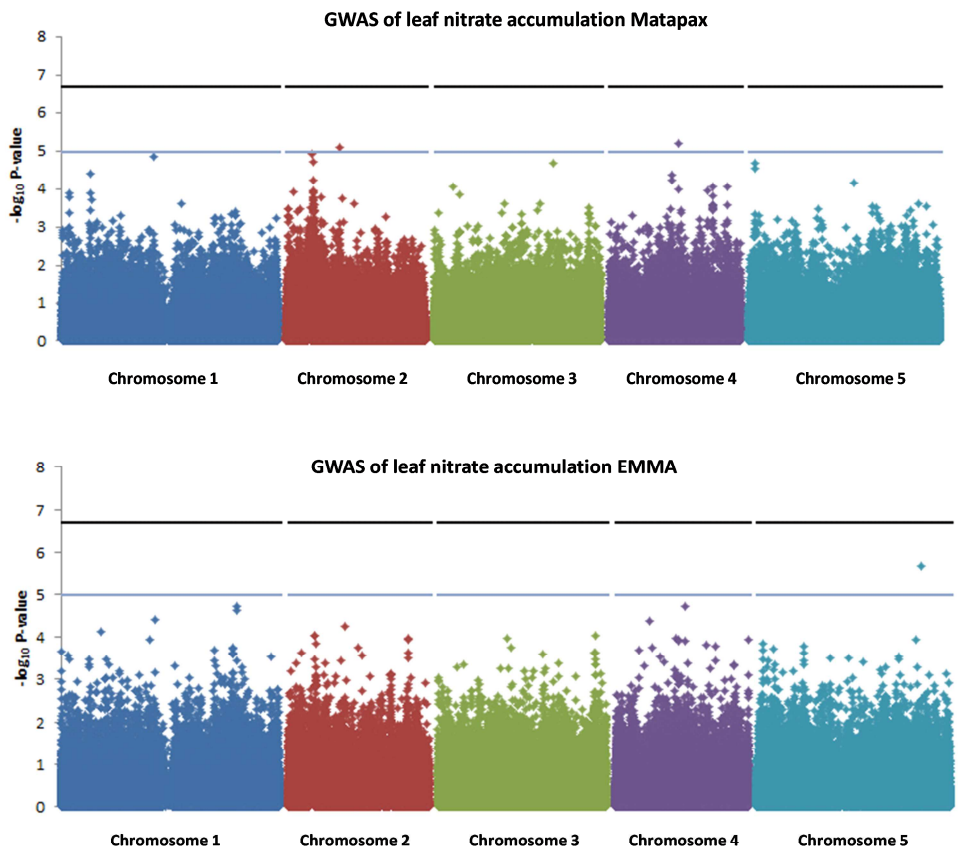
GWAS on the nitrate, phosphate, and sulfate accumulation datasets was run by using the MArker-Trait Association Platform And eXplorer (Matapax). Matapax is a Web-based platform to support GWAS in arabidopsis (Childs et al. 2012; section 5.2.4). Since it was not possible to use EMMA (within Matapax) with the datasets greater than 200 accessions, I used GAPT method that includes correction for kinship and population structure (Childs et al. 2012). It is the main association tool in the Matapax pipeline. In general, it implements EMMA and enhances it with algorithms designed to increase statistical power and further decrease computational time (Childs et al. 2012). It is able to handle large datasets by subdividing the genotypic data into multiple smaller files (Lipka et al. 2012).

Despite the convenience of using the Web-based tool Matapax for GWAS, it would be desirable to use independent and more flexible tools for this purpose. In addition, GAPIT has not been extensively used for GWAS whereas EMMA was proved to produce meaningful associations (Atwell et al. 2010) and since then has been successfully used for GWAS of various, agronomically important traits (Baxter et al. 2010, Chao et al. 2012). Therefore, to verify the results obtained from Matapax David E. Salt and Dai-Yin Chao performed the GWAS on the datasets using EMMA algorithm implemented in the R package.

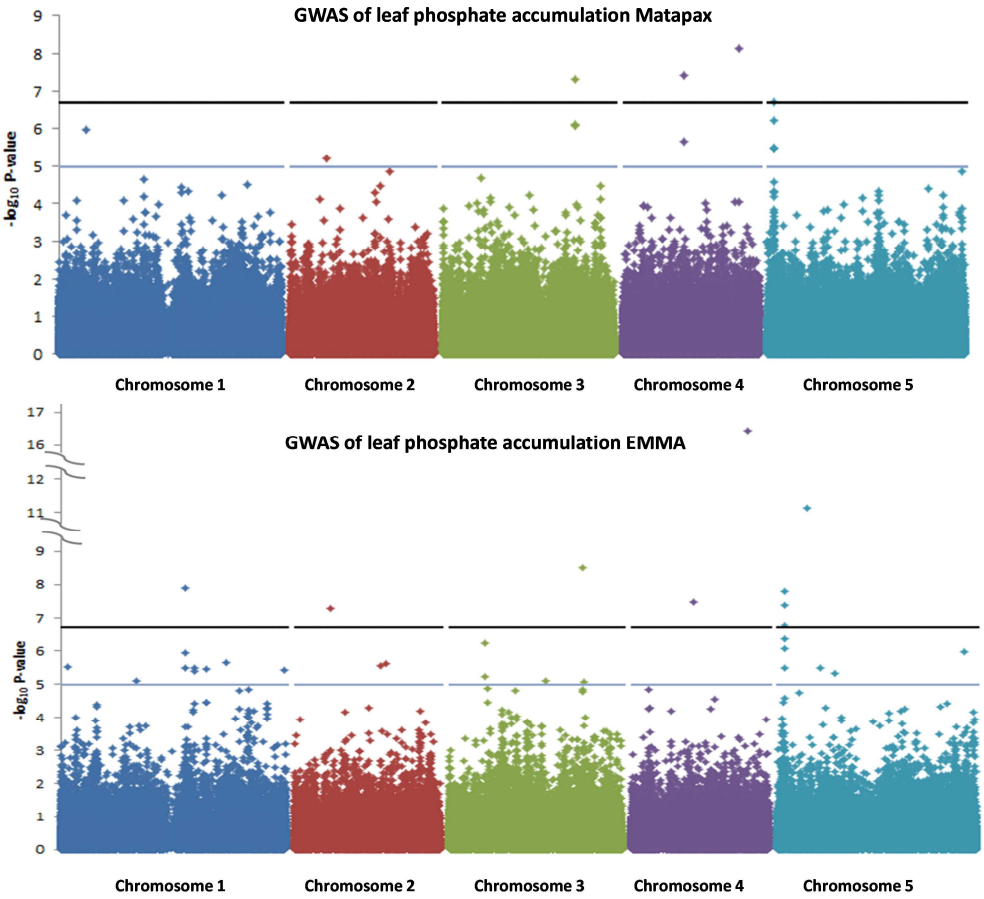
To identify significant associations I created Manhattan plots for the P-values obtained by using both the Matapax pipeline and the EMMA algorithm (Figure 5.4). The calculated 5% significant threshold with Bonferroni correction for 250k tests was set at $-\log_{10}$ P-value = 6.7. Because the SNP data used for both EMMA and Matapax GWAS were the same, the threshold is also the same for all the results. There were no associations that exceeded this threshold for either nitrate or sulfate accumulation data with either of the two methods used (Figure 5.4A, C). Twelve significantly associated SNPs were identified for phosphate accumulation data. Three of them were derived by Matapax and nine by EMMA (Figure 5.4B). A major peak of SNPs associated with leaf phosphate accumulation was localised on chromosome 5. It was identified using both Matapax and EMMA methods. However, it was not significant in the Matapax results with the Bonferroni threshold (Figure 5.4B). Within 20 kb either side of the most significantly associated SNP from that region there were nine genes including AT5G03730 which is annotated in TAIR database as Constitutive Triple-Response 1 (CTR1) kinase and is involved in the negative regulation of the ethylene signal transduction pathway. Ethylene was shown to be involved in the regulation of plant response to phosphate deficiency (Nagarajan & Smith 2012). Therefore, this gene was selected as a candidate for further analysis which is described in next sections of this chapter.

As mentioned before, the Bonferroni threshold is the strictest one from the available and commonly used methods. Chao et al. (2012) in a GWAS of leaf cadmium accumulation used 337 arabidopsis accessions which also were a subset of the Borevitz collection and the same set of SNPs as in my analysis. In this study they set the threshold at $-\log_{10}$ P-value = 5, (even though the major peak revealed by this analysis would be significant with the higher threshold). Therefore, I decided to lower the threshold in this analysis to the same level. The lower threshold revealed three SNPs significantly associated with nitrate accumulation data (two derived by Matapax and one by EMMA; Figure 5.4A), three SNPs significantly associated with sulfate accumulation (two derived by Matapax and one by EMMA; Figure 5.4C), and 31 additional SNPs significantly associated with phosphate accumulation (10 derived by Matapax and 21 by EMMA; Figure 5.4B). However, the meaning of the threshold depends on the aim of the study. If the aim of the analysis is to identify candidate genes for follow up studies, a stringent FDR and a strict threshold are recommended (Korte & Farlow 2013). This strategy prevents the substantial cost consequences of extensive wasted effort and money in further biological and functional characterisation of spurious associations. Conversely, if the aim of the study is to investigate the genetic architecture of a trait, one may consider a less stringent FDR and subsequent lower threshold value (Korte & Farlow 2013).

A.


GWAS of leaf nitrate accumulation Matapax
-log10 P-value
Chromosome 1
Chromosome 2
Chromosome 3
Chromosome 4
Chromosome 5
GWAS of leaf nitrate accumulation EMMA
-log10 P-value
Chromosome 1
Chromosome 2
Chromosome 3
Chromosome 4
Chromosome 5

B.


GWAS of leaf phosphate accumulation Matapax
-log10 P-value
Chromosome 1
Chromosome 2
Chromosome 3
Chromosome 4
Chromosome 5
GWAS of leaf phosphate accumulation EMMA
-log10 P-value
Chromosome 1
Chromosome 2
Chromosome 3
Chromosome 4
Chromosome 5

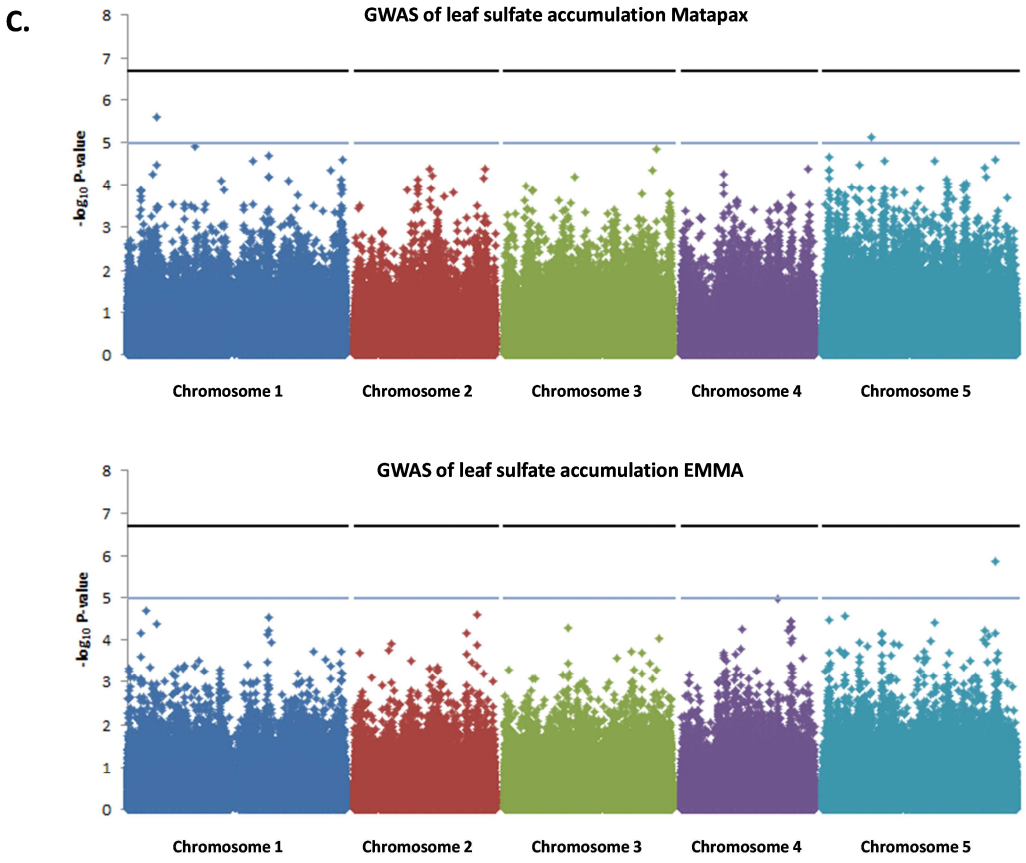


Figure 5.4 ***Manhattan plots of GWAS results on anion accumulation data***

GWAS of leaf nitrate concentration in 284 arabidopsis accessions (A); leaf phosphate concentration in 246 arabidopsis accessions (B); leaf sulfate concentration in 261 arabidopsis accessions (C) grown in CER. In each case results derived by the Matapax pipeline are presented in the upper panel and results derived by the EMMA algorithm are presented in the lower panel. Black horizontal lines indicate a nominal 5% significance threshold with Bonferroni correction for 250,000 tests $-\log_{10}$ P-value = 6.7. Blue horizontal lines indicate genome-wide significance threshold of $-\log_{10}$ P-value = 5. SNPs are accurately plotted according to their position along the appropriate chromosome. Different colours facilitate the visualisation of individual chromosomes.

5.3.3 Selection of candidate genes

Keeping in mind the low significance of the GWAS outcome the most strongly associated markers obtained as a result of the two methods were analysed. For the GWAS on phosphate and sulfate concentration data all the markers with P-value ≤ 0.0001 were analysed whereas for nitrate all the markers with the P-value ≤ 0.005 were analysed. As mentioned before, the significantly associated markers are often not a direct effect of a variation in the tested trait (Myles et al. 2009). Hence, the genes with a causal polymorphism can be found within the LD of the markers. It has been shown that in arabidopsis this LD is between 10-20

kb from a marker (Atwell et al. 2010, Kim et al. 2007). Therefore, an LD decay distance of 20 kb either side from each analysed marker was assumed.

The possible candidate genes that seemed likely to be involved in the control of natural variation in the accumulation of the three macronutrients tested were selected based on two main information sources. First of all, based on the gene annotations from the TAIR database and the existing functional knowledge candidates involved in the uptake and assimilation of the three nutrients tested were selected. Additionally, signalling and regulatory factors that could be involved in the control of the homeostasis of the three macronutrients were of interest. Secondly, publically available microarray data was used as additional sources of information. It was assumed that the profiles of expression of genes involved in the control of sulfate homeostasis will be altered under sulfate deficiency. For that reason, the microarray data for Col-0 WT grown in sulfate deficient media were evaluated. These data were published previously (Maruyama-Nakashita et al. 2006) and they can be obtained from the Gene Expression Omnibus (GEO; www.ncbi.nlm.nih.gov/geo) database under accession number GSE445.

The P-values produced by both Matapax and EMMA were comparable for nitrate and sulfate accumulation data suggesting that both of them produced similar P-value distributions. For phosphate accumulation data some P-values produced by EMMA were two fold lower than those produced by Matapax. A good example of such elevated P-value is SNP *Chr4:15754220*. It is the most significantly associated SNP on the chromosome 4 in the phosphate accumulation dataset produced by both methods. However, the P-value for that SNP produced by EMMA is only half of the P-value produced by Matapax (Figure 5.4B). It was shown previously that P-values produced by EMMA are not always well estimated and should be interpreted with caution (Atwell et al. 2010). Therefore, using the results from both Matapax and EMMA allowed a mutual complementation of the data which gave confidence in the results.

These analyses resulted in a selection of 59 genes for further analysis of their involvement in the natural variation of accumulation of the analysed nutrients. Seventeen of these genes were derived from the nitrate accumulation data (Table 5.3), 20 were derived from the phosphate accumulation data (Table 5.4), and 22 were derived from the sulfate accumulation data (Table 5.5). A large number of analysed markers were related to unknown genes, pseudogenes or transposable elements. Therefore, they seemed unlikely to influence the nutrient accumulation. Due to time limitation of my PhD project they were excluded from

further analyses. On the other hand, a number of genes was likely to be involved in nutrient homeostasis. For example, genes encoding mitochondrial phosphate transporter 2 (MPT2) derived from the phosphate accumulation dataset and the ATPS1 isoform of ATPS (an enzyme involved in the first step of sulfate reduction in the plant cells described in chapter 3 of this thesis) and sulfate transporter SULTR3;3 derived from the sulfate accumulation dataset suggest that there is real information in these data. Among the genes selected as candidates only 12 directly included the GWAS markers. All the other genes were selected using the LD. Therefore, in some cases two candidates were selected from the same marker, e.g. *Chr2:1081648* especially when they belonged to the same class of genes as in case of *Chr4:9749937* or when the marker was significant in both Matapax and EMMA.

Table 5.3 ***Candidate genes derived from GWAS on leaf nitrate concentration data***

Rank	Chr	Position	Score	Method	Candidate	TAIR annotation
1	4	9749937	5.1894	both	AT4G17500	Ethylene response factor 1 AtERF-1
1	4	9749937	5.1894	both	AT4G17490	Ethylene response factor 6 AtERF-6
2	2	7507230	5.0797	Matapax	AT2G17260	Glutamate receptor 2 GLR-2
4	1	24480901	4.6135	EMMA	AT1G65840	Peroxisomal polyamide oxidase AtPAO4
7	3	16728880	4.6849	both	AT3G45570	RING/U-box protein with C6HC-type zinc finger domain
10	4	8877666	4.3534	both	AT4G15530	Pyruvate orthophosphate dikinase PPDK
12	3	9643862	3.9868	both	AT3G26340	N-terminal nucleophile aminohydrolase
13	4	8513274	3.9838	EMMA	AT4G14880	Cytosolic isoform of cytosolic O-acetylserine(thiol)lyase OAS-TL
18	4	13967619	3.9570	Matapax	AT4G28140	Member of ERF/AP2 transcription factor family
19	4	18550616	3.9339	EMMA	AT4G40010	SNF1 related protein kinase 2.7 SNRK2.7
53	5	18067534	3.4955	Matapax	AT5G44720	Molybdenum cofactor sulfurase family protein
56	2	319733	3.4782	Matapax	AT2G01760	Response regulator 14 ARR14
60	1	4023029	3.4630	both	AT1G11930	Predicted pyridoxal phosphate-dependent enzyme
152	5	845487	3.0073	both	AT5G03430	putative PAPS reductase
192	2	1081648	2.9727	Matapax	AT2G03600	Hypothetical protein similar to AtUPS1
192	2	1081648	2.9727	Matapax	AT2G03590	Member of a class of allantoin transporters AtUPS1
242	2	16740121	2.7982	EMMA	AT2G40090	Member of ATH transporters subfamily AtATH9

Rank column corresponds to the significance of the association in the total dataset (and not on individual chromosome); in cases where the marker was significant in both methods ("both" in the method column) the higher number is given; grey boxes indicate genes which directly included the GWAS marker; the Chr = Chromosome; Score corresponds to $-\log_{10}$ P-value.

Table 5.4 ***Candidate genes derived from GWAS on leaf phosphate concentration data***

Rank	Chr	Position	Score	Method	Candidate	TAIR annotation
6	1	8342207	7.4881	EMMA	AT4G14580	SNF1 related protein kinase 3.3 SNRK3.3
3	3	18150590	7.3103	both	AT3G48850	Mitochondrial phosphate transporter 2 MPT2
8	1	3828144	5.9743	Matapax	AT1G11390	Protein kinase superfamily protein
16	2	12364981	5.6052	both	AT2G28810	Dof-type zinc finger DNA-binding family protein
18	1	873731	5.5202	EMMA	AT1G03495	HXXXD-type acyl-transferase family protein
10	5	955628	5.4877	both	AT5G03730	Constitutive Triple-Response 1 kinase CTR1
20	1	16834866	5.4851	both	AT1G44350	IAA-leucine resistant (ILR)-like gene 6 ILL6
23	1	19751436	5.4497	EMMA	AT1G52990	Thioredoxin family protein
24	1	30284235	5.4174	EMMA	AT1G80580	Ethylene response transcription factor AtERF B-1
27	3	4795786	5.2239	both	AT3G14400	Ubiquitin-specific protease 25 UBP25
28	1	10195762	5.1123	EMMA	AT1G29230	Member of SNF1 related kinase gene family AtCIPK18
40	4	11310888	4.5607	both	AT4G21230	Cysteine-rich receptor-like protein kinase CRK27
43	1	19687330	4.4621	EMMA	AT1G52890	NAC transcription factor ANAC019
49	5	22060060	4.3181	EMMA	AT5G54310	A member of ARF GAP domain AGD5
35	1	11693580	4.1888	Matapax	AT1G32450	Nitrate transporter 1.5 NRT1.5
67	1	17982197	4.1822	both	AT1G48630	RACK1B
75	1	25320168	4.0275	both	AT1G67600	Acid phosphatase/vanadium-dependent haroperoxidase related protein
66	3	5637573	3.8274	both	AT3G16600	SNF2 domain-containing protein
68	5	7694730	3.8114	Matapax	AT5G23000	ATMYB37
75	5	26173474	3.7688	Matapax	AT5G65500	U-box domain-containing protein kinase family protein

Rank column corresponds to the significance of the association in the total dataset (and not on an individual chromosome); in cases where the marker was significant in both methods ("both" in the method column) the higher number is given; grey boxes indicate genes which directly included the GWAS marker; the Chr = Chromosome; Score corresponds to $-\log_{10}$ P-value.

Table 5.5 ***Candidate genes derived from GWAS on leaf sulfate concentration data***

Rank	Chr	Position	Score	Method	Candidate	TAIR annotation
1	5	24051345	5.8611	both	AT5G59670	Leucine-rich repeat protein kinase family protein up-regulated in *slim1* microarray data
2	4	13513943	4.9556	both	AT4G26890	Member of MEKK subfamily MAPKKK16
3	1	2629873	4.6802	both	AT1G08340	Rho GTP-ase activating protein
5	5	3100575	4.5520	both	AT5G09930	Member of GCN subfamily ABCF2 up-regulated in *slim1* microarray data
5	1	19737375	4.6938	both	AT1G52950	Nucleic acid-binding, OB-fold-like protein up-regulated in *slim1* microarray data
7	5	830621	4.4844	both	AT5G03430	putative PAPS reductase
10	5	15642814	4.5619	both	AT5G39040	Member of TAP subfamily of ABC transporters
10	5	15642814	4.5619	both	AT5G39080	HXXXD-type acyl-transferase family protein
12	4	15366769	4.3249	both	AT4G31805	WRKY family transcription factor up-regulated in *slim1* microarray data
13	4	15231390	4.2762	EMMA	AT4G31360	Selenium binding protein
15	2	18177061	4.3817	Matapax	AT2G43900	5 inositol polyphosphate 5-phosphatase 12 5PTASE12
18	4	14920764	4.2093	both	AT4G30510	*Arabidopsis thaliana* homolog of yeast autophagy 18 (ATG18) B
19	1	1966415	4.1612	both	AT1G06450	CCR4-associated factor 1 CAF1
19	1	1966415	4.1612	both	AT1G06470	Nucleotide/sugar transporter family portein
20	2	15728631	4.1599	EMMA	AT2G37440	DNAse I-like superfamily protein
22	1	3520775	4.2670	Matapax	AT1G10670	Subunit A of the trimeric protein ATP Citrate Lyase ACLA1
24	1	19515429	4.1282	EMMA	AT1G52430	Ubiquitin carboxyl-terminal hydrolase-related protein
25	5	23177993	4.0906	EMMA	AT5G57180	Chloroplast import apparatus 2 CIA2 down-regulated in *slim1* microarray data
25	5	23177993	4.0906	EMMA	AT5G57240	OSBP(oxysterol binding protein)-related protein 4C ORP4C up-regulated in *slim1* microarray data
28	5	22407229	4.0031	EMMA	AT5G55240	*Arabidopsis thaliana* peroxygenase 2 ATPXG2
38	4	5931214	4.0191	both	AT4G09350	Chlororespiratory reduction J CRRJ up-regulated in *slim1* microarray data
85	3	8139451	3.9860	EMMA	AT3G22890	ATPS1

Rank column corresponds to the significance of the association in the total dataset (and not on an individual chromosome); in cases where the marker was significant in both methods ("both" in the method column) the higher number is given; grey boxes indicate genes which directly included the GWAS marker; the Chr = Chromosome; Score corresponds to $-\log_{10}$ P-value.

Since there is evidence of interconnection between different nutrient assimilation pathways (Chapter 4 of this thesis; Kopriva & Rennenberg 2004, Koprivova et al. 2000) I calculated the ratios between nitrate, phosphate, and sulfate (N/P, N/S, S/P) and performed GWAS on these data. These data were analysed by using only the Matapax pipeline (Figure 5.5). In this case all the markers with P-value ≤ 0.001 for each of the three datasets were analysed. Subsequently, SNPs which were common between accumulation and ratio data were excluded from the further analysis of ratio data. Similarly to the analysis of accumulation data, the LD decay distance of 20kb either side of a given SNP was assumed to look for the candidate genes that may explain the natural variation in the accumulation of analysed anions. The rationale for candidate genes selection was as described above. Finally two genes derived from N/P ratio data, two genes derived from N/S ratio data and eight genes from S/P ratio data were selected for further analysis of their involvement in the regulation of the homeostasis of particular anions (Table 5.6).

Table 5.6 ***Candidate genes derived from GWAS analysis of the ratio data***

Rank	Chr	Position	Score	Dataset	Candidate	TAIR annotation
3	2	18798180	5.28433086	np	AT2G45650	Agamous-like 6 transcription factor AGL6
7	4	15459090	4.97510404	np	AT4G31990	Aspartate aminotransferase 5 ASP5
37	3	19182836	3.6476245	ns	AT3G51770	Ethylene overproducer 1 ETO1
54	4	8303998	3.43616308	ns	AT4G14370	Phosphoinositide binding protein
3	5	23672391	5.14935376	sp	AT5G58550	ETO1-like 2 EOL2
9	5	3417901	4.46142627	sp	AT5G10800	RNA recognition motif
9	5	3417901	4.46142627	sp	AT5G10790	Ubiquitin-specific protease 22 UBP22
12	5	17420543	4.33198703	sp	AT5G43370	Phosphate transporter 1;2 PHT1;2
12	5	17420543	4.33198703	sp	AT5G43350	Phosphate transporter 1;1 PHT1;1
16	2	13915316	4.25212229	sp	AT2G32830	Phosphate transporter 1;5 PHT1;5
23	2	16249798	4.07810553	sp	AT2G38940	Phosphate transporter 1;4 PHT1;4
60	5	2776844	3.56161589	sp	AT5G08590	SNF1 protein related kinase 2.1 SNRK2.1

Rank column corresponds to the significance of the association in the total dataset (and not on an individual chromosome); grey box indicates gene which directly included the GWAS marker; Chr = Chromosome; Score corresponds to $-\log_{10}$ P-value; np indicates nitrate/phosphate ratio, ns – nitrate/sulfate ratio, and sp – sulfate/phosphate ratio

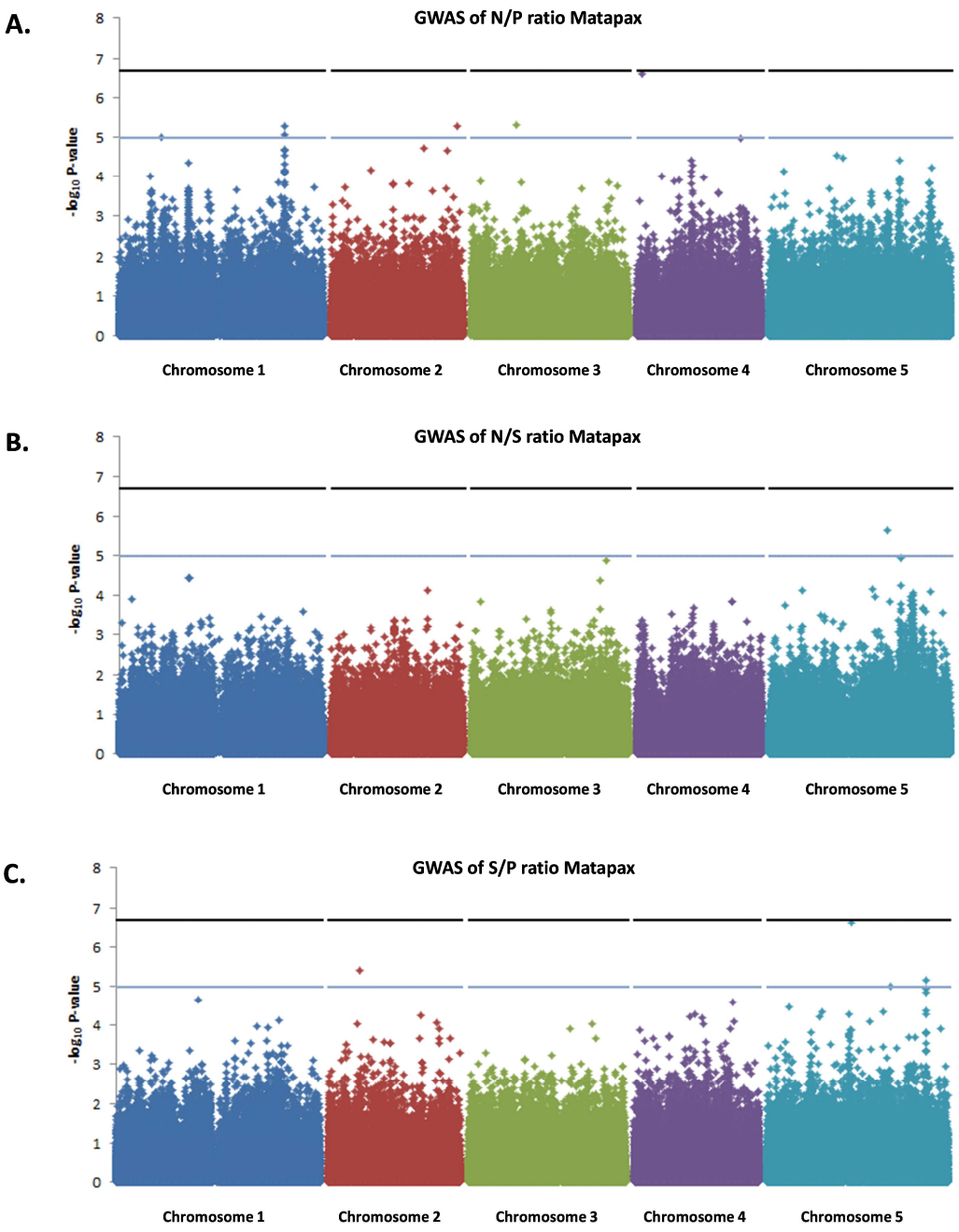


Figure 5.5 ***Manhattan plots of GWAS results on ratio data***

GWAS of the ratio between nitrate/phosphate (N/P) leaf concentration in 246 arabidopsis accessions (A); ratio between nitrate/sulfate (N/S) leaf concentration in 261 arabidopsis accessions (B); ratio between sulfate/phosphate (S/P) leaf concentration in 246 arabidopsis accessions (C) grown in CER. Black horizontal lines indicate a nominal 5% significance threshold with Bonferroni correction for 250,000 tests $-\log_{10}$ P-value = 6.7. Blue horizontal lines indicate genome-wide significance threshold of $-\log_{10}$ P-value = 5. SNPs are accurately plotted according to their position along the appropriate chromosome. Different colours facilitate the visualisation of individual chromosomes.

5.3.4 The analysis of T-DNA insertion lines

Because of the low statistical significance of the GWAS results the selected candidates required careful verification. Assuming that disruptions in genes involved in the regulation of nutrient homeostasis will affect the leaf anion concentration, one Salk or Sail T-DNA insertion line per gene selected as a candidate (Tables 5.3 to 5.6) was obtained from The European Arabidopsis Stock Centre (NASC; Table 5.7). Since the homozygous T-DNA insertion lines are resistant to kanamycin on the initial stage of the analysis I grew all the lines on plates with MS medium supplemented with kanamycin in order to be able to distinguish the homozygous lines from heterozygous line and lines with no insertion for the primary analyses.

Table 5.7 ***DNA insertion lines analysed to verify GWAS candidate genes***

Line number	NASC line number	Gene	TAIR annotation	Comment
		NITRATE DATASET		
Line 1	SALK_098851	AT3G45570	RING/U-box protein with C6HC-type zinc finger domain	heterozygous
Line 5	SALK_030899	AT2G17260	Glutamate receptor 2 GLR-2	heterozygous
Line 6	SALK_036267	AT4G17500	Ethylene response factor 1 AtERF-1	Km^R homozygous
Line 8	SALK_092114	AT2G01760	Response regulator 14 ARR14	homozygous
Line 11	SALK_116156	AT4G15530	Pyruvate orthophosphate dikinase PPDK	heterozygous
Line 23	SALK_081127C	AT5G44720	Molybdenum cofactor sulfurase family protein	homozygous
Line 26	SALK_064134C	AT2G03600	Hypothetical protein similar to AtUPS1	Km^R not genotyped
Line 32	SALK_056596C	AT2G40090	Member of ATH transporters subfamily AtATH9	Km^R not genotyped
Line 34	SALK_141801C	AT5G03430	PAPS reductase	Km^R homozygous
Line 48	SALK_133599C	AT1G65840	Peroxisomal polyamide oxidase AtPAO4	homozygous
Line 54	SALK_097424C	AT4G40010	SNF1 related protein kinase 2.7 SNRK2.7	Km^R homozygous
Line 64	SALK_004669C	AT3G26340	N-terminal nucleophile aminohydrolases	homozygous
Line 70	SAIL_512_B08	AT4G28140	Member of ERF/AP2 transcription factor family	homozygous
Line 75	SALK_087356C	AT4G17490	Ethylene response factor 6 AtERF-6	Km^R homozygous
Line 76	SAIL_549_F06	AT2G03590	Member of a class of allantoin transporters ATUPS1	homozygous
Line 77	SAIL_675_A03	AT1G11930	Predicted pyridoxal phosphate-dependent enzyme	homozygous
Line 80	SALK_082213C	AT4G14880	Cytosolic isoform of O-acetylserine(thiol)lyase OAS-TL	homozygous

Line number	NASC line number	Gene	TAIR annotation	Comment
			PHOSPHATE DATASET	
Line 2	SALK_009859	AT5G23000	ATMYB37	heterozygous
Line 7	SALK_071872	AT3G16600	SNF2 domain-containing protein	homozygous
Line 21	SALK_043036C	AT1G32450	Nitrate transporter 1.5 NRT1.5	homozygous
Line 24	SALK_099679C	AT5G54310	A member of ARF GAP domain AGD5	homozygous
Line 33	SALK_122868C	AT5G03730	Constitutive Triple-Response 1 kinase CTR1	heterozygous
Line 35	SALK_111336C	AT3G14400	Ubiquitin-specific protease 25 UBP25	homozygous
Line 36	SALK_151903C	AT1G11390	Protein kinase superfamily protein	heterozygous
Line 43	SALK_135953	AT1G29230	Member of SNF1 related kinase gene family AtCIPK18	Km^R homozygous
Line 46	SALK_107550	AT3G48850	Mitochondrial phosphate transporter 2 MPT2	homozygous
Line 47	SALK_142536	AT1G80580	Ethylene response transcription factor AtERF B-1	didn't germinate
Line 49	SALK_009687C	AT1G52990	Thioredoxin family protein	Km^R homozygous
Line 50	SALK_021421C	AT1G03495	HXXXD-type acyl-transferase family protein	Km^R homozygous
Line 53	SALK_009893C	AT4G14580	SNF1 related protein kinase 3.3 SNRK3.3	Km^R homozygous
Line 57	SALK_024894C	AT1G44350	IAA-leucine resistant (ILR)-like gene 6 ILL6	homozygous
Line 60	SALK_021591C	AT4G21230	Cysteine-rich receptor-like protein kinase CRK27	no insertion
Line 62	SALK_096310C	AT1G52890	NAC transcription factor ANAC019	heterozygous
Line 67	SALK_052482C	AT5G65500	U-box domain-containing protein kinase family protein	homozygous
Line 68	SALK_047774C	AT1G67600	Acid phosphatase/vanadium-dependent haroperoxidase related protein	heterozygous
Line 72	SAIL_418_H08	AT1G48630	RACK1B	homozygous
Line 78	SAIL_869_F08	AT2G28810	Dof-type zinc finger DNA-binding family protein	heterozygous

Line number	NASC line number	Gene	TAIR annotation	Comment
			SULFATE DATASET	
Line 3	SALK_011884	AT5G39040	Member of TAP subfamily of ABC transporters	Km^R not genotyped
Line 4	SALK_021437	AT1G10670	Subunit A of the trimeric protein ATP Citrate Lyase ACLA1	homozygous
Line 12	SALK_033906C	AT1G06450	CCR4-associated factor 1 CAF1	Km^R homozygous
Line 15	SALK_124733C	AT1G06470	Nucleotide/sugar transporter family portein	Km^R homozygous
Line 16	SALK_092343C	AT4G31360	Selenium binding protein	homozygous
Line 22	SALK_065920C	AT2G43900	5 inositol polyphosphate 5-phosphatase 12 5PTASE12	homozygous
Line 31	SALK_103821C	AT5G39080	HXXXD-type acyl-transferase family protein	heterozygous
Line 44	SALK_007262	AT3G55780	Glycosyl hydrolase	homozygous
Line 51	SALK_057196C	AT1G52430	Ubiquitin carboxyl-terminal hydrolase-related protein	heterozygous
Line 52	SALK_018778C	AT5G09930	Member of GCN subfamily ABCF2 up-regulated in *slim1* microarray data	Km^R homozygous
Line 55	SALK_142820C	AT4G31805	WRKY family transcription factor up-regulated in *slim1* microarray data	Km^R homozygous
Line 56	SALK_111394C	AT4G09350	Chlororespiratory reduction JCRRJ up-regulated in *slim1* microarray data	didn't germinate
Line 58	SALK_003255C	AT4G26890	Member of MEKK subfamily MAPKKK16	Km^R homozygous
Line 59	SALK_115780C	AT5G59670	Leucine-rich repeat protein kinase family protein up-regulated in *slim1* microarray data	homozygous
Line 61	SALK_080212C	AT1G08340	Rho GTP-ase activating protein	didn't germinate
Line 63	SALK_121775C	AT5G55240	*Arabidopsis thaliana* peroxygenase 2 ATPXG2	heterozygous
Line 65	SALK_099224C	AT1G52950	Nucleic acid-binding, OB-fold-like protein up-regulated in *slim1* microarray data	homozygous
Line 66	SALK_039947C	AT4G30510	*Arabidopsis thaliana* homolog of yeast autophagy 18 (ATG18) B	no insertion
Line 69	SALK_019928C	AT5G57180	Chloroplast import apparatus 2 CIA2 down-regulated in *slim1* microarray data	homozygous
Line 71	SAIL_572_G08	AT5G57240	OSBP(oxysterol binding protein)-related protein 4C ORP4C up-regulated in *slim1* microarray data	homozygous
Line 74	SALK_128101	AT5G03430	PAPS reductase	heterozygous
Line 79	SALK_117130	AT2G37440	DNAse I-like superfamily protein	homozygous

Line number	NASC line number	Gene	TAIR annotation	Comment
			N/P RATIO DATASET	
Line 13	SALK_070088C	AT4G31990	Aspartate aminotransferase 5 ASP5	homozygous
Line 73	SALK_095121	AT2G45650	Agamous-like 6 transcription factor AGL6	no insertion
			N/S RATIO DATASET	
Line 10	SALK_112450	AT4G14370	Phosphoinositide binding protein	homozygous
Line 19	SALK_061581C	AT3G51770	Ethylene overproducer 1 ETO1	heterozygous
			S/P RATIO DATASET	
Line 14	SALK_114207C	AT5G58550	ETO1-like 2 EOL2	Km^R not genotyped
Line 17	SALK_080084C	AT5G08590	SNF1 protein related kinase 2.1 SNRK2.1	homozygous
Line 20	SALK_110194C	AT5G43370	Phosphate transporter 1;2 PHT1;2	heterozygous
Line 25	SALK_138643C	AT2G38940	Phosphate transporter 1;4 PHT1;4	no insertion
Line 27	SALK_088586C	AT5G43350	Phosphate transporter 1;1 PHT1;1	heterozygous
Line 28	SALK_117672C	AT5G10800	RNA recognition motif	heterozygous
Line 30	SALK_138009C	AT2G32830	Phosphate transporter 1;5 PHT1;5	homozygous
Line 39	SAIL_41_C07	AT5G10790	Ubiquitin-specific protease 22 UBP22	homozygous

Line numbers correspond to the working numbers of lines; Km^R corresponds to lines which showed resistance to kanamycin; N/P indicates nitrate/phosphate ratio; N/S indicates nitrate/sulfate ratio; S/P indicates sulfate/phosphate ratio.

After 11 days of growth on plates 17 lines (out of 79) remained green and healthy indicating resistance to kanamycin (Table 5.7). I assumed that these plants are homozygous and transferred them to the soil for four weeks. Subsequently, I harvested leaves of these plants and measured anion concentration (Figure 5.6). This preliminary analysis revealed three T-DNA lines with significantly lower nitrate concentration (Figure 5.6A), one line with significantly higher phosphate concentration (Figure 5.6B), and two lines with contrasting sulfate concentration; one line with significantly lower and one line with significantly higher sulfate concentration (Figure 5.6C). However, these results should be treated with caution. Due to poor column performance, hence poor separation of nitrate on the HPLC some batches of plants gave much lower values for nitrate concentration than those reported in other chapters of this thesis (right panel of Figure 5.6A). The separation of phosphate and sulfate

was as usual. Because all the T-DNA insertion lines are in the Col-0 background in this preliminary experiment, I compared the lines with each other.

The lines with low nitrate have insertions in genes annotated in TAIR as putative PAPS reductase family protein (line 34), arabidopsis ethylene response factor 6 AtERF6 (line 75), and arabidopsis ethylene response factor 1 AtERF1 (line 6; Figure 5.6A). The line with higher phosphate has an insertion in gene encoding HXXXD-type acyl-transferase family protein (line 50) according to TAIR annotation (Figure 5.6B). The line with high sulfate concentration has an insertion in gene encoding a phosphate/sugar transporter (line 15) and the line with low sulfate has an insertion in gene encoding CCR4-NOT associated factor 1 CAF1 (line 12; Figure 5.6C). Due to the limited length of my PhD project I decided to investigate the differences between the lines which have insertions in genes selected based on GWAS for a particular anion e.g. in lines with an insertion in genes selected based on GWAS for nitrate accumulation I looked only at nitrate concentration. I did not look at the variation in phosphate or sulfate concentration in these lines which could also be an interesting subject of future research.

The results shown in Figure 5.6 indicate that some of the genes selected as candidates might have an effect on anion accumulation. However, since these were preliminary results, it was necessary to verify them. Moreover, some of them showed a large standard error on values which might be due to bad column performance. It was also seen before in Kopriva's lab that kanamycin resistance is not always the best way of selection for homozygous plants (A. Koprivova, personal communication). Therefore, I decided to genotype the lines which did not show resistance to kanamycin as well as lines which showed differences in anion concentration in the preliminary experiment (Figure 5.6) and analyse them again for anion concentration. Four of the lines from the batch on the right side of Figure 5.6 which showed resistance to kanamycin, but no changes in anion concentration were not genotyped and excluded from further analyses (Table 5.7).

All the remaining lines (except the four removed after the preliminary experiment) were genotyped using the standard PCR procedure. The genotyping revealed 43 homozygous lines and 17 heterozygous lines; four lines did not have an insertion and three lines did not germinate (Table 5.7). Subsequently, I measured the anion concentration in five week old plants of 39 homozygous lines (the four lines on the right panel of figure 5.6 which showed differences in anion concentration were not included in this analysis).

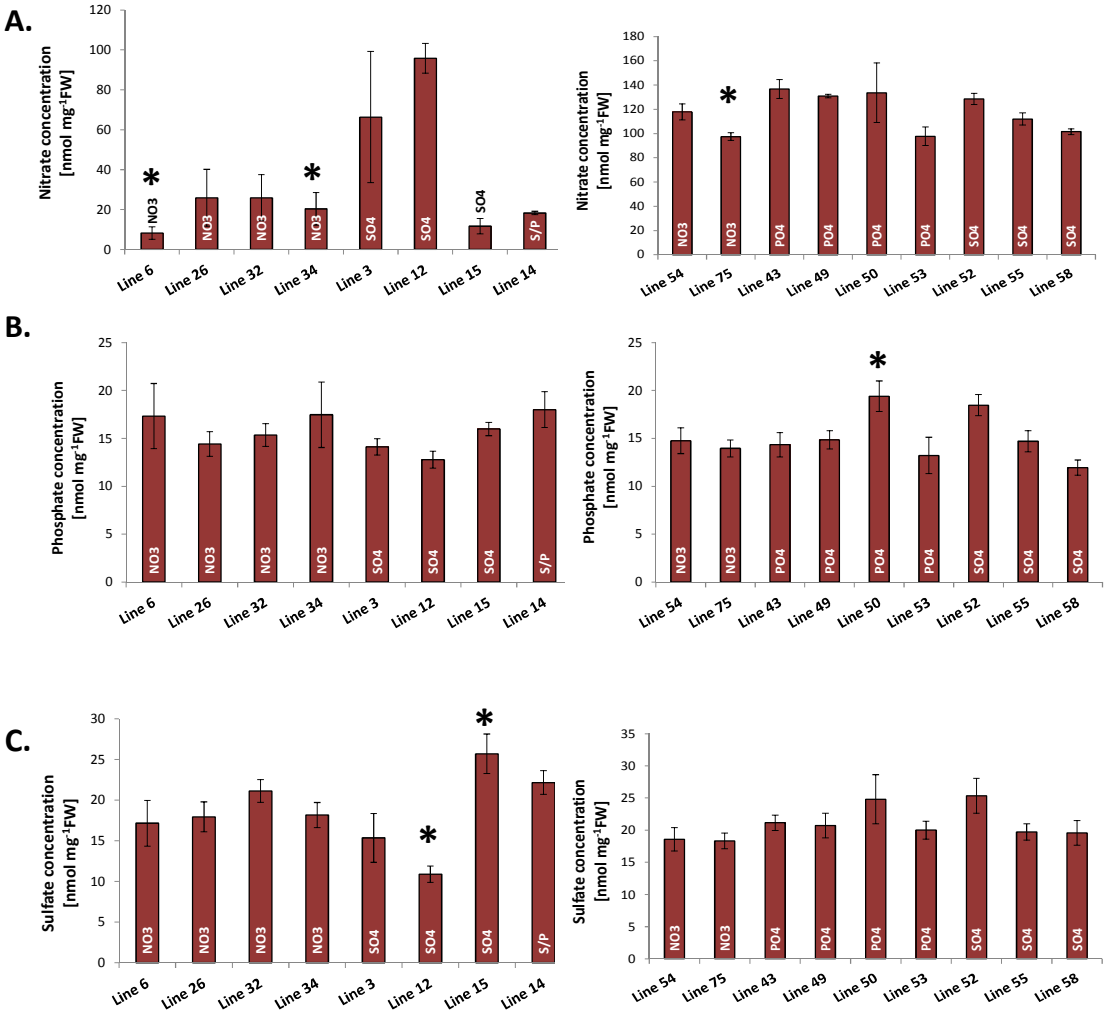


Figure 5.6 ***Anion concentration in T-DNA insertion lines resistant to kanamycin***

Primary analysis of nitrate (A), phosphate (B), and sulfate (C) concentration in the T-DNA insertion lines which showed resistance to kanamycin after 11 days of growth; Left and right panels correspond to measurements in different batches of plants; The bars represent mean values of three biological replicates ± standard error; Asterisks indicate lines significantly different from the other lines within a measurement (*t*-test P-value ≤ 0.05); White labels on the bottom of each bar correspond to the dataset from which the candidate gene was derived; NO3 indicates nitrate dataset; PO4 indicates phosphate dataset; SO4 indicates sulfate dataset; S/P indicates sulfate/phosphate ratio dataset.

There was large variation in nitrate concentration most probably due to the poor HPLC separation (see above). Thus it is likely that these results may be unreliable. Therefore I decided not to use them in gene selection (Figure 5.7A). In contrast, the analysis of phosphate accumulation revealed two lines with significantly higher phosphate concentration and one line with significantly lower phosphate concentration (Figure 5.7B). The lines with high

phosphate concentration have an insertion in the genes encoding Ubiquitin-specific protease 25 UBP25 (Line 35) and SNF1 related kinase gene family AtCIPK18 (Line 43). The line with low phosphate concentration has an insertion in the gene encoding IAA-leucine resistant (ILR)-like gene 6 ILL6 (Line 57). Line 50 which in the preliminary experiment described above (Figure 5.6) showed high phosphate concentration, in this experiment did not differ from other lines. Therefore, I excluded this line from further analysis. Additionally, I selected one line with significantly lower sulfate concentration (Figure 5.7C). This line has an insertion in the gene encoding MAPKKK16 kinase (Line 58). In this experiment I compared the selected lines to all the lines measured.

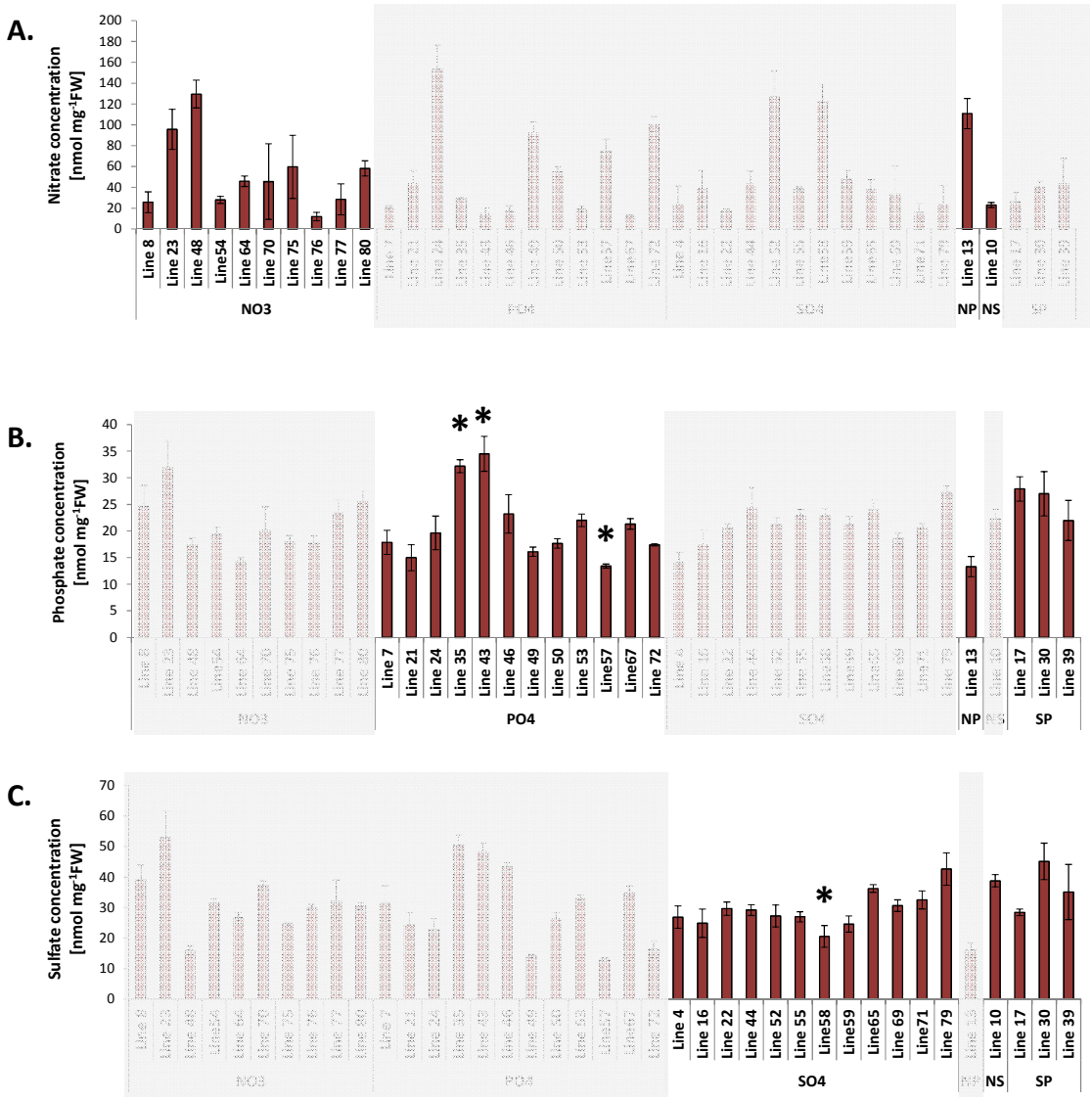


Figure 5.7 Anion concentration in the homozygous T-DNA insertion lines
Analysis of nitrate (A), phosphate (B), and sulfate (C) concentration in the homozygous T-DNA lines with insertions in the genes selected based on the GWAS results; The bars represent mean values of three biological replicates ± standard error; Asterisks indicate lines with insertions in a gene selected based on GWAS for a given anion significantly different within a

dataset to all the other lines (*t*-test P-value ≤ 0.05) – light grey boxes on each histogram shade the data which are not particularly important for the analysis described in this chapter, but they were used for statistical analysis; NO3 indicates nitrate dataset; PO4 indicates phosphate dataset; SO4 indicates sulfate dataset; NP indicates nitrate/phosphate ratio dataset; NS indicates nitrate/sulfate dataset; SP indicates sulfate/phosphate dataset.

In order to finally examine the anion concentration phenotypes of lines selected so far based on the preliminary experiment (Figure 5.6; lines 6, 34, 75, 12, and 15) and the analysis of 39 homozygous lines (Figure 5.7; lines 35, 43, 57, and 58) and select the best candidates for further analyses I analysed these lines once again. I increased the number of biological replicates to five and included WT Col-0 as a control (Figure 5.8). The analysis of five week old plants confirmed the low nitrate accumulation phenotype obtained in the preliminary analysis for lines 6 and 34 (Figure 5.6) with the insertions in genes selected based on GWAS for nitrate accumulation. Line 75 which in the preliminary analysis showed low nitrate concentration, in this experiment was not significantly different from Col-0 (Figure 5.8A). Therefore this line was excluded from the further analyses. The verification experiment did not confirm any of the phosphate accumulation phenotypes revealed in previous experiments. Line 43 was the only line with phosphate concentration significantly different than Col-0 (Figure 5.8B), but it showed a different phenotype in every experiment in which it was included. Because none of the phosphate related lines showed consistent phosphate accumulation phenotype, all of them were excluded from further analyses. Similarly to the preliminary experiment, line 12 showed lower sulfate accumulation (Figure 5.8C). This result was not significant when compared to Col-0, but it was significant when compared to other lines tested (as was done in the preliminary experiment). Line 15 showed opposite response in sulfate accumulation compared to the preliminary experiment where it had a high sulfate concentration compared to other lines tested (Figure 5.6C). In this experiment it had lower sulfate concentration compared to other lines (Figure 5.8C). However, this line has an insertion in a gene encoding a tonoplast localised transporter gene. Since the genes encoding vacuolar sulfate influx transporters have not been identified yet I decided to include this gene in further analyses. Even though the sulfate accumulation results obtained so far were contrary to each other, they were always significantly different from other lines tested.

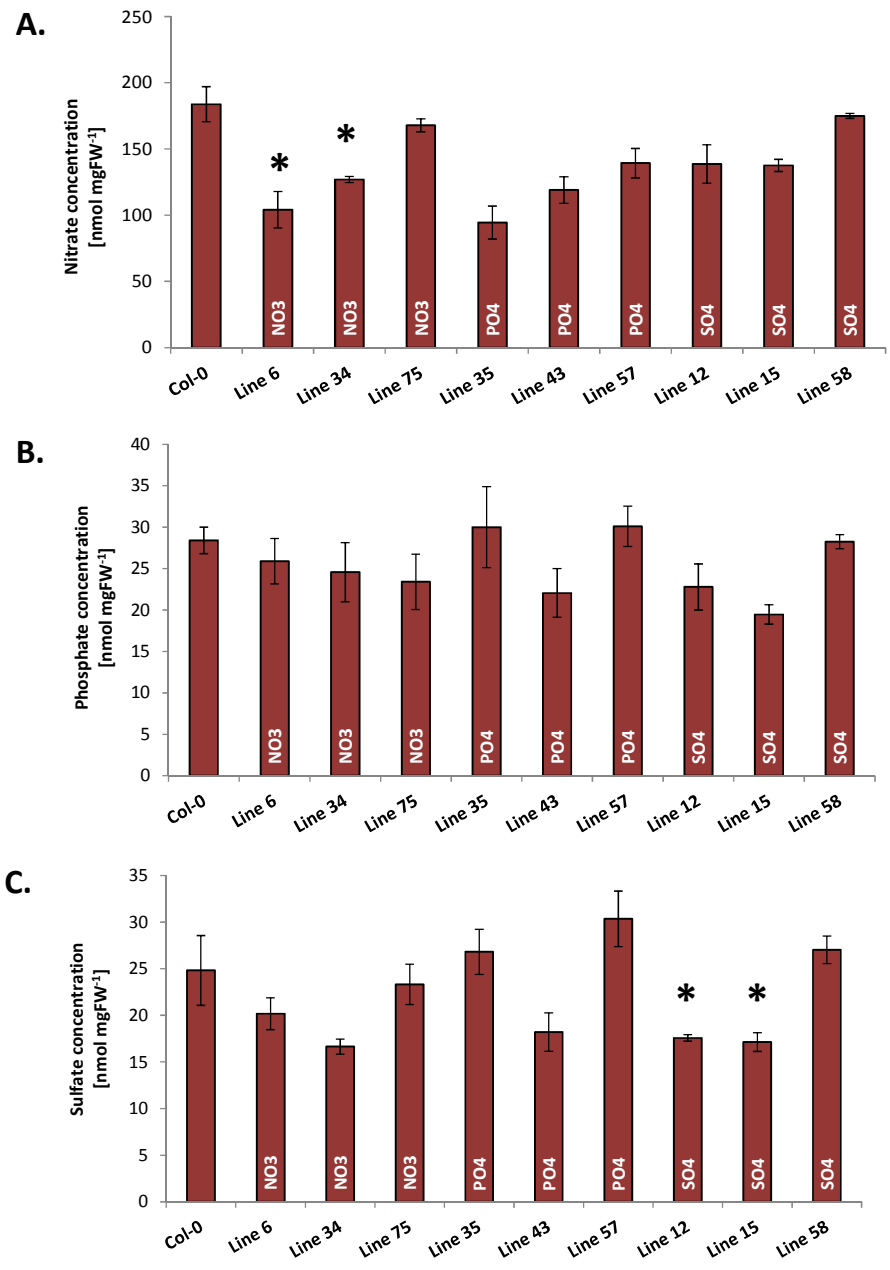


Figure 5.8 ***Anion concentration in T-DNA insertion lines – verification***

The analysis of nitrate (A), phosphate (B), and sulfate (C) concentration in the T-DNA insertion lines selected based on previous experiments. The bars represent mean values of five biological replicates ± standard error. Asterisks indicate lines significantly different from Col-0 or the other lines (indicated in text; *t*-test P-value ≤ 0.05); White labels on the bottom of each bar correspond to the dataset from which the candidate gene was derived; NO3 indicates nitrate dataset; PO4 indicates phosphate dataset; SO4 indicates sulfate dataset.

5.3.5 Outcome of Genome-Wide Association Study

Finally, three candidate genes were selected for further analysis based on the analysis of anion accumulation in the homozygous T-DNA insertion lines. Line 34 has an insertion in the AT5G03430 gene annotated in TAIR as putative PAPS reductase family protein. PAPS reductase

domain is similar to APS reductase and enzymes of primary sulfate assimilation from bacteria and yeast (see Chapter 6). In bacteria, enzymes with PAPS reductase activity reduce PAPS (the main donor of activated sulfate for sulfation reactions) to free sulfite in a two-step reaction (Savage et al. 1997). This gene was revealed in both nitrate and sulfate accumulation GWAS. The T-DNA line showed low nitrate concentration and sulfate concentration similar to Col-0. The low nitrate phenotype in this line was observed in a number of independent experiments performed by me and Anna Koprivova from the Kopriva lab on two week old plants as well as on five week old plants. Line 12 has the insertion in AT1G06450 annotated in TAIR as CCR4-NOT associated factor 1 (CAF1). CAF1 is a key subunit of the CCR4-NOT complex and it is involved in the regulation of plant development and biotic stress resistance (Sarowar et al. 2007). The gene was revealed in sulfate accumulation GWAS and the T-DNA line with insertion in this gene showed low sulfate concentration in a number of independent experiments. The last gene selected as a candidate was also revealed in sulfate accumulation GWAS – line 15. It is annotated in TAIR database as a phosphate/sugar translocator (AT1G06470). However, the phenotype of the T-DNA line with insertion in this gene is not consistent.

To investigate the architecture of these genes among arabidopsis natural accessions 1001 Genomes Database was used. Subsequently, the data generated for GWAS were reconsidered to investigate whether the variation in the gene architecture is likely to affect anion accumulation phenotypes. The analysis of 1001 Genomes Database led to an identification of various amino acid changes and insertions/deletions (indels) across ca. 150 accessions which were common between accessions which sequences were deposited in the database and the accessions from Borevitz collection (Figure 5.9). A subset of the accessions with various amino acid substitutions and deletions was identified in the gene encoding putative PAPS reductase family protein which were not found in Col-0. Some of these deletions included the second exon, which may significantly affect the protein function (accessions Ag-0 and Sp-0 in Figure 5.9A). The analysis of GWAS nitrate concentration data revealed that the accessions with these changes have lower nitrate concentration compared to accessions with the gene architecture similar to Col-0 (Figure 5.9B). There was no difference in sulfate concentration between the two groups of accessions (Figure 5.9C).

Similarly, a subset of accessions with E170D amino acid substitution was identified in the gene encoding CAF1 (Figure 5.9D; accessions C24, Ha-0, Per-1, and Ws-0). The analysis of sulfate accumulation data from 150 accessions used for GWAS (a subset which overlaps with the 1001 Genomes database) revealed that the accessions with the D allele in position 170 have significantly lower sulfate concentration compared to the accessions with the E allele in

this position as in Col-0 (Figure 5.9E). Additionally, a subset of accessions with S63V amino acid substitution was identified in the first exon of the gene encoding phosphate/sugar transporter (Figure 5.9F; accessions Bla-1, Boot-1, Pla-0, Rou-0). The analysis of sulfate accumulation data revealed that the accessions with V allele in position 63 have significantly higher sulfate concentration compared to the accessions with S allele in this position as in Col-0 (Figure 5.9G). These data suggest that the natural variation in selected candidate genes may have a real effect on nutrient homeostasis in A*rabidopsis thaliana*.

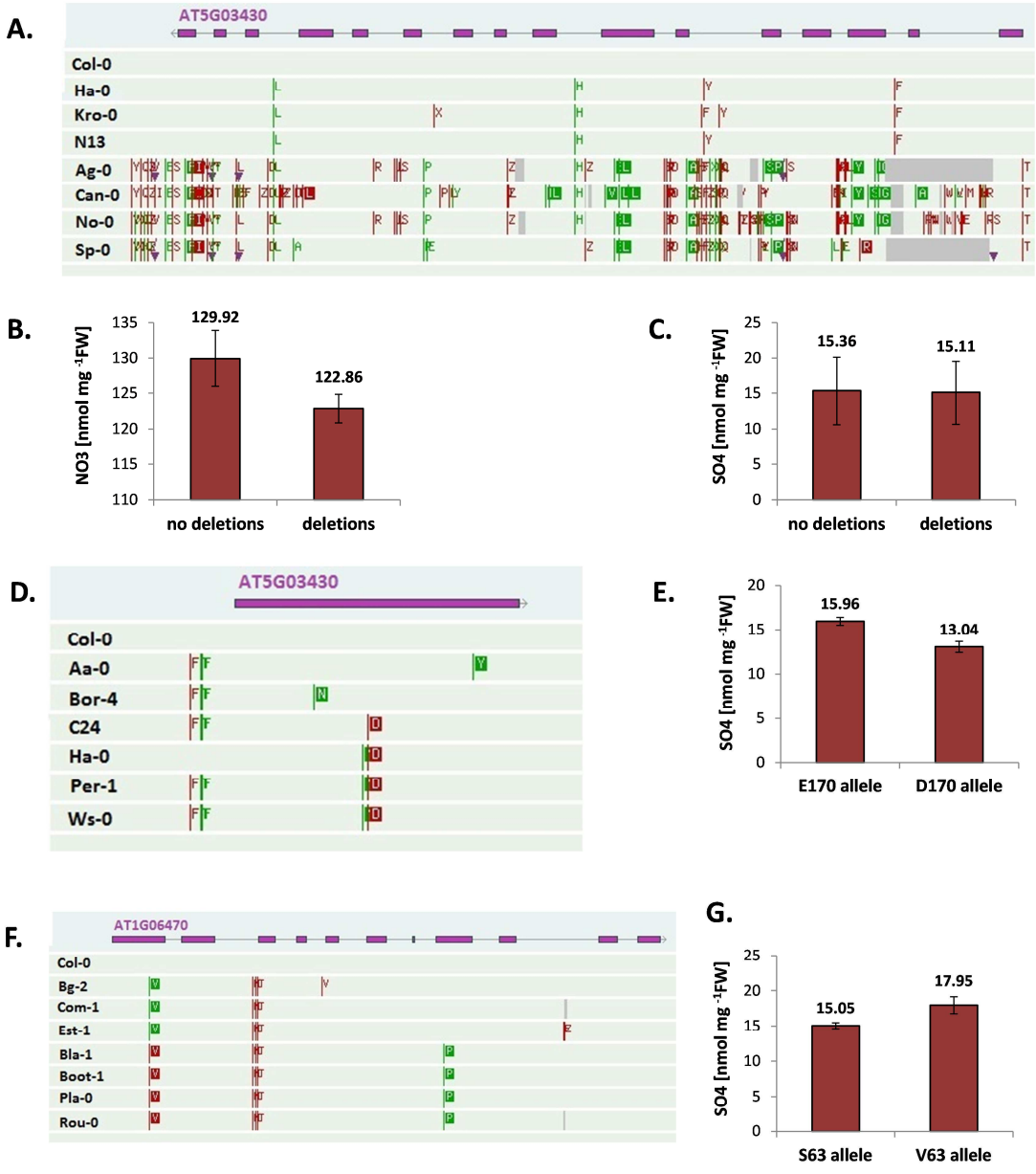


Figure 5.9 ***Natural variation in selected candidate genes***

(A) Natural variation in the architecture of AT5G03430; amino acids marked in green indicate non-synonymous substitutions; amino acids marked in red indicate synonymous substitutions; grey boxes indicate deletions; (B) Average nitrate concentration (NO3) of accessions sharing similar gene architecture of At5G03430; (C) Average sulfate concentration (SO4) of accessions

sharing similar gene architecture of AT5G03430; (D) Natural variation in the architecture of AT1G06450, the changes in the gene are shown in the same way as in A; (E) Average sulfate concentration (SO4) in accessions sharing the same allele in AT1G06450; (F) Natural variation in the architecture of AT1G06470, the changes in the gene are shown in the same way as in A; (G) Average sulfate concentration (SO4) in accessions sharing the same allele in AT1G06470; the natural variants are subtracted from the 1001 Genomes database; the error bars on the histograms correspond to standard deviation; the differences between different groups of accessions are statistically significant with *t*-test P-value ≤ 0.05 except C where is no significant difference.

To further examine the involvement of these genes in anion homeostasis I analysed additional T-DNA insertion lines in each of these genes for anion accumulation. I obtained five additional insertion lines for the putative PAPS reductase, four lines for CAF1, and three lines for phosphate/sugar transporter from NASC (Table 5.8). PCR genotyping revealed two homozygous lines with an insertion in CAF1, one homozygous line with an insertion in the phosphate/sugar transporter and no additional homozygous lines for PAPS reductase (Table 5.8). I grew all the homozygous lines with insertions in the selected candidate genes on plates with MS medium for two weeks. Additionally, one set of plants was grown on MS medium supplemented with sucrose to investigate sugar effect on the anion accumulation in these lines. The analysis of anion concentration revealed no significant difference in nitrate concentration in the T-DNA line with the insertion in PAPS reductase gene in either of the two conditions tested (Figure 5.10A). This might be due to the fact that anions accumulate in plant cells during development and the differences are more pronounced in older plants. Nevertheless, addition of sucrose revealed a significant difference in sulfate concentration in the homozygous lines with insertion in both CAF1 and phosphate/sugar transporter, even in two week old seedlings (Figure 5.10B). There were no differences in anion accumulation between the lines grown on medium without sucrose which, as mentioned before, might be due to developmental stage of the plants (anion accumulation phenotype more pronounced in older plants). The differences in plant response depending on the presence of sucrose in the growing medium are not surprising. It is long known that plant growth and development as well as the sensing of external and internal environmental signals are influenced by sucrose availability (Koch 2004, Ohto et al. 2001, Rolland et al. 2002). These results further confirm that disruption in genes selected as candidates may affect the anion accumulation. However, more detailed investigation of the involvement of these genes in nutrient metabolism is required in order to characterise their specific function.

Table 5.8 ***Additional T-DNA lines with insertions in the three selected candidate***

Gene	Line name	NASC line number	Comment
AT5G03430	PAPS (Line 34)	SALK_141801C	primary line
	PAPS1_1	SALK_031937 (AE)	heterozygous
	PAPS1_2	SALK_043060	didn't germinate
	PAPS1_3	SALK_128101	no insertion
	PAPS1_4	SALK_140720	no insertion
	PAPS1_5	SALK_144629C	no insertion
AT1G03450	CAF1 (Line12)	SALK_033906C	primary line
	CAF1_1	SALK_101782 (BO)	didn't germinate
	CAF1_2	SAIL_533_G10	homozygous
	CAF1_3	SAIL_1306_A08	homozygous
	CAF1_4	SALK_009021C	no insertion
AT1G03470	ST (line 15)	SALK_124733C	primary line
	ST1_1	SALK_088510 (CI)	no insertion
	ST1_2	SAIL_156_B12	no insertion
	ST1_3	SALK_126893C	homozygous

The column "Line name" corresponds to working names of the lines

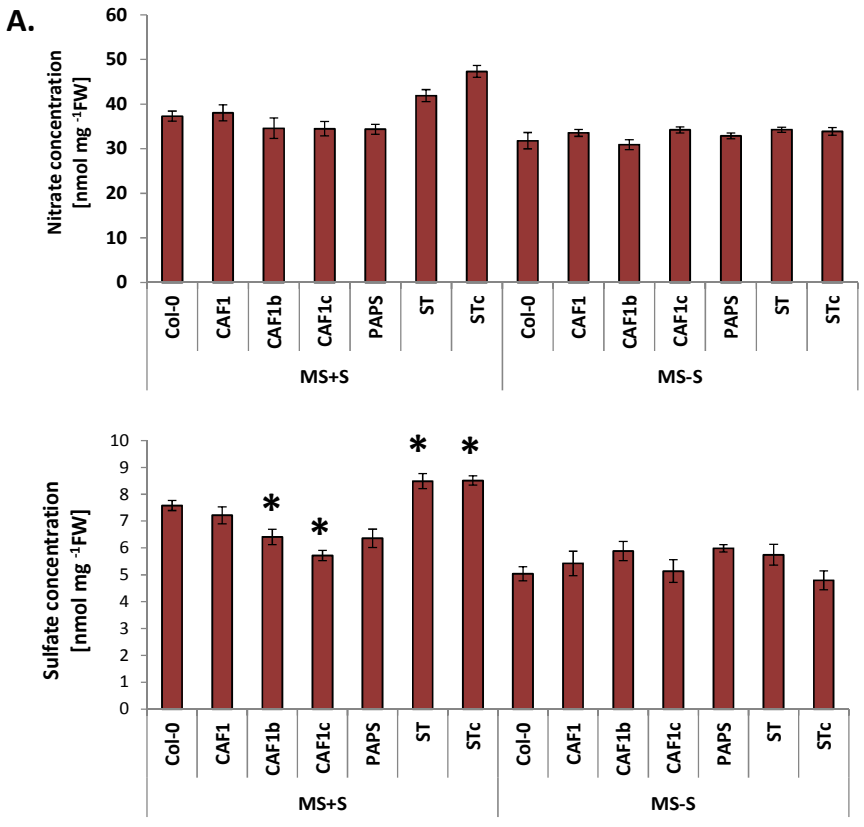


Figure 5.10 ***Anion analyses in additional T-DNA lines***

The analysis of nitrate (A) and sulfate (B) concentration in the additional T-DNA lines with insertions in the three candidate genes; the bars represent mean values of six biological replicates ± standard error; Asterisks indicate lines significantly different from Col-0 (*t*-test P-value ≤ 0.05); MS+S indicates Murashige-Skoog medium supplemented with 0.8% sucrose; MS-S indicates Murashige-Skoog with no sucrose.

5.4 Discussion

Many traits of agronomic importance are quantitative and as such required development of specific tools that would allow analysis of their complex genetic architecture (Koornneef et al. 2004). In fact, many tools for quantitative genetics were developed by breeders. *Arabidopsis thaliana* was adopted as a model organism to study the genetic architecture of quantitative traits after the molecular markers for mapping became available (Chang et al. 1988, Nam et al. 1989). Currently, around 1,500 accessions have been genotyped for 250,000 SNPs (Horton et al. 2012) and the data are publically available and can be used for the study of natural variation. Additionally, a number of statistical and computational tools have been developed to overcome various limitations and increase the efficiency of such analyses. Therefore, there is an increasing interest in the studies of natural variation. To date, quantitative genetic studies of arabidopsis natural variation have been successfully used to examine genetic variation underlying a number of agronomically important traits such as drought tolerance (Thudi et al. 2014, Varshney et al. 2012), salt tolerance (Baxter et al. 2010), shade avoidance (Filiault & Maloof 2012) and more.

The GWAS on 317 arabidopsis accessions described in this chapter was performed to get new insight into the regulation of anion accumulation in this model plant. Rather than a few peaks with large effects, as seen in some previously published arabidopsis GWAS (Atwell et al. 2010, Chao et al. 2012), this analysis revealed many peaks with small effects. These results suggest that variation in homeostasis of the three nutrients tested is a complex trait controlled by many different loci as might be expected for an environmentally sensitive trait (Atwell et al. 2010, Filiault & Maloof 2012). Through the analysis of the most significant associations and the follow up analysis of T-DNA insertion lines it was possible to identify genes that might be involved in control of nutrient accumulation. The analysis of the gene variants in 1001 Genomes database revealed various haplotypes for each of the genes tested. Subsequent cross reference with anion accumulation datasets showed differences in anion concentration between the accessions with different variants of each of the genes tested. This indicates further the possible involvement of these genes in the control of nutrient homeostasis. Future work should be focused on the verification of these associations by complementary methods such as QTL mapping that minimize the confounding due to effects of population structure. Secondly, the polymorphisms in these genes should be functionally characterised and their involvement in nutrient homeostasis described in more detail.

For two of the candidates revealed from the analysis of T-DNA insertion lines, CAF1 and phosphate/sugar transporter, I investigated only the information available online to get better understanding of their functions and possible involvement in sulfate homeostasis. Briefly, CAF1 (AT1G06450) is an integral subunit of carbon catabolite repressor 4 – CCR4 associated factor 1 (CCR4-CAF1) complex. CCR4-CAF1 is the major enzyme complex that catalyses mRNA deadenylation (shortening of the poly(A) tail which is the initial step of mRNA degradation; Sarowar et al. 2007). This process is well studied in yeast (Tucker et al. 2001), but relatively less is known about this complex in plants. CAF1 shows deadenylase activity *in vitro* (Daugeron et al. 2001) but the *in vivo* role remains unclear. The arabidopsis CAF1 protein family consists of 11 homologs among which CAF1a and CAF1b are best described (Walley et al. 2010). In the study of Walley et al. (2010) AT1G06450 is referred to as CAF1f. In this study they analysed the expression of all eleven members of family in response to mechanical wounding. The expression of CAF1f was undetectable under these conditions (Walley et al. 2010). To date, it is the only report of this gene in the literature. However, CAF1 proteins are known to be involved in mediating plant development and response to biotic and abiotic stress (Chou et al. 2014, Liang et al. 2009). Additionally, according to Genevestigator expression data AT1G06450 is induced by sulfate and nitrate depletion. These results, together with the analysis of the two haplotype groups in this gene described above suggest that it is likely to be involved in the control of sulfate homeostasis in arabidopsis. Plant nutrition is often regulated on the post-transcriptional level (Yoshimoto et al. 2007, Yuan et al. 2007) and these regulation processes often involve mRNA turnover, e.g. after sulfate resupply to sulfur-starved plants (Smith et al. 1997, Tavares et al. 2008, Yoshimoto et al. 2007). Further characterisation of the involvement of CAF1f in the regulation of sulfate metabolism could provide an interesting, novel insight into sulfate homeostasis and perhaps sulfate use efficiency.

Phosphate/sugar transporter (AT1G06470) is the next gene revealed as a candidate in GWAS, based on leaf sulfate accumulation data. The S63V amino acid substitution leads to increased sulfate concentration in the accessions with the V allele. The two T-DNA lines, one with an insertion at the beginning of the gene and the second with the insertion in the last (11) exon (according to T-DNA Express database) showed high sulfate concentration. Increased sulfate concentration in these lines is more pronounced after addition of sucrose to the growing medium. The only report of this gene in the literature concerns its localisation in the tonoplast membrane (Schmidt et al. 2007). Therefore, higher sulfate accumulation in the presence of sucrose in the T-DNA lines with insertion in this gene is particularly interesting. The analysis of regulation of arabidopsis ion transporters showed induction of nitrate and

sulfate transporters by sucrose (Lejay et al. 2003). Therefore it might be hypothesized that this transporter might be involved in transport of sulfate across vacuolar membranes. However in that case rather low sulfate accumulation would be expected. Vacuolar sulfate transporters are key regulators of the distribution of internal sulfate in arabidopsis. They have been shown to respond to sulfate limitation with an increased transcript accumulation (Kataoka et al. 2004b). To date no sulfate vacuolar importer has been identified at a molecular level. Because the sulfate accumulation phenotype of T-DNA lines with insertions in this gene was not consistent in independent experiments more detailed analysis is necessary to examine the involvement of this protein in sulfate transport. Nevertheless, its further examination could reveal an important regulatory factor involved in sulfate distribution within the cell.

Due to limited length of my PhD project I could focus on detailed analysis of only one gene. I chose to analyse in more detail AT5G03430 which is annotated in TAIR database as a putative PAPS reductase family protein. In sulfate assimilation APS (activated form of sulfate) can be either directly reduced to sulfite by APS reductase or phosphorylated by APS kinase to form PAPS (Leustek et al. 2000). In bacteria PAPS can be reduced to sulfite by PAPS reductase (Leustek et al. 2000, Kopriva 2006). The existence of PAPS reductase activity in plants is controversial. Up to date no PAPS reductases homologous to those from *E. coli* other than APR have been identified in the arabidopsis or rice genomes (Kopriva et al. 2007). It has been shown that algae and higher plants reduce APS (Tsang et al. 1971) whereas the PAPS reductase activity is identified mainly in yeast and enteric bacteria (Schmidt & Jager 1992). However, it could be that such an enzyme in plants has a completely divergent structure from the one found in bacteria (see Chapter 6). This controversy made that gene particularly interesting for me. The AT5G03430 came up in the GWAS on both leaf nitrate and sulfate accumulation data. However, the T-DNA line with an insertion in the gene differed only in concentration of nitrate. Sulfate concentration in that line was comparable with Col-0 and other lines (Figure 5.6C). Additionally, this phenotype was the most consistent among all the lines tested in a number of independent experiments. More detailed analysis of this gene is described in Chapter 6.

Due to time limitation a number of potentially interesting genes selected as candidates were not verified. The 17 heterozygous T-DNA lines were immediately excluded from further analysis. Among these, one had an insertion in AT5G03730 which was linked to SNP *Chr5:955628*. This is the major peak of linked SNPs associated with leaf phosphate accumulation revealed by both GWAS methods used (Figure 5.4B). The AT5G03730 is a Constitutive Triple-Response 1 CTR1 kinase which was identified by mutants that displayed the

triple response morphology in the absence of exogenously added ethylene (Huang et al. 2003, Kieber et al. 1993). This protein kinase negatively regulates ethylene signalling in arabidopsis (Huang et al. 2003, Yang et al. 2013). Ethylene is a gaseous hormone involved in the regulation of number of cellular processes and stress responses. It has been shown that ethylene is involved in phosphate-dependent root modifications and systemic phosphate signalling pathways (Nagarajan & Smith 2012, Roldan et al. 2013). Therefore, it is possible that CTR1 is directly or indirectly involved in the regulation of phosphate accumulation in arabidopsis leaves. However, it was not selected for further analysis because the T-DNA line was a heterozygote and none of the heterozygous lines was analysed. Additionally, this gene is already described in detail and the aim of the study was to identify novel candidate genes involved in the regulation of nutrient concentration. Since phosphate is a common limiter of plant growth, a better understanding of the regulation of phosphate metabolism would help to improve breeding strategies towards more efficient use of phosphate available in the soil. Therefore, further investigation of this and other candidate genes revealed by this GWAS analysis should be considered in the near future.

Similarly, none of the candidates derived from the ratio data were analysed in more detail even though a number of interesting candidates were revealed. Here again the datasets including phosphate accumulation and especially the S/P ratio dataset yielded candidate genes that are likely to be involved in the control of nutrient use efficiency. The GWAS on the S/P ratio dataset revealed four phosphate transporters. Phosphate transporters are the major factors in the response to phosphate depletion. They are strongly induced in response to phosphate deficiency (Karthikeyan et al. 2002). Additionally, both sulfate and phosphate transporters are regulated by miRNAs. MiR395 is involved in the regulation of sulfate assimilation (Matthewman et al. 2012), whereas miR399 is known to be involved in regulation of phosphate uptake and translocation (Chiou et al. 2006, Fujii et al. 2005, Lin et al. 2008). It was also suggested that miR395 might be suppressed in phosphate deficient plants (Hsieh et al. 2009). Together with the Phosphate Response 1 (PHR1) transcription factor miRNAs are known to be involved in the interconnection of sulfate and phosphate assimilation pathways (Rouached 2011). Therefore, more careful analysis of the genes revealed from the GWAS on ratio data may provide new insights into interconnection of the control of homeostasis of various nutrients and should be considered as a future goal.

In general, GWAS on leaf phosphate concentration data produced the highest amount of significant associations (Figure 5.4). Moreover, it was the only dataset that apart from few significant associations spread across the genome showed a peak of SNPs associated with leaf

phosphate accumulation on chromosome 5 discussed above. Taking into account the low significance of the GWAS outcome, highly significant associations from well-defined peaks with several adjacent SNPs having high score are more likely to contain real information. In contrast, a single isolated SNP is probably a false positive even if it is above the threshold (Atwell et al. 2010, Korte & Farlow 2013). However, none of the T-DNA lines tested showed significant differences compared to other lines or Col-0 (Figures 5.6 to 5.8).

Contrary to GWAS on leaf phosphate concentration data, the strongest associations from GWAS on leaf nitrate and sulfate concentration data were relatively weak and below significance thresholds commonly used (Figure 5.4). This might be due to a complex architecture of the analysed traits (Korte & Farlow 2013). The power of GWAS to identify the association between the SNP and the variation in the trait depends on the phenotypic variance within the population. The phenotypic variance depends on the strength of the phenotypic effect of the different alleles and their frequency in the sample. Therefore, there are two cases where GWAS is likely to produce a large number of false positive associations. First, the analysed trait is controlled by many common variants, each with only a small phenotypic effect. Second, the analysed trait is controlled by many rare variants, each having a large effect on phenotype (Korte & Farlow 2013). Moreover, computer simulations which were subsequently coupled with available data showed that the rare variations are very often strongly correlated with other, non-causative rare variants within the genome, irrespective of LD decay. Therefore, a single rare causative locus may drag with it many synthetic associations (Dickson et al. 2010).

As described previously, the analysis of natural variation in sulfate concentration between two wild arabidopsis accessions Bay-0 and Shahdara, performed to identify the gene(s) controlling sulfate concentration, revealed APR2 as a key control step in the sulfate reduction pathway (Loudet et al. 2007). The analysis of Bay-0 x Sha recombinant inbred lines (RIL) led to the identification of a single nucleotide polymorphism in the APR2 isoform of APR. The substitution of alanine with glutamate in a conserved domain of the protein resulted in significant differences in enzyme activity leading to sulfate accumulation (Loudet et al. 2007). However, APR2 did not come up in the GWAS on leaf sulfate accumulation data. As described in Chapter 3 and by Chao et al. (2014) further analysis of natural variation in APR2 gene architecture led to the identification of a number of very rare loci that severely reduced the activity of the protein. This complex architecture of the gene was probably the reason why APR2 presented difficulties for GWAS and did not produce a significant association. The analysis of APR2 and ATPS1 described in Chapter 3 as well as the results described in this

chapter are excellent examples of how combining both GWAS and QTL mapping may improve the overall understanding of such a complex process as nutrient use efficiency.

Although GWAS was proved to be successful in investigation of the genetic architecture of a number of quantitative traits related to flowering time (Aranzana et al. 2005), drought tolerance (Bouchabke et al. 2008), or nutrient accumulation (Koprivova et al. 2014) it has its limitations. Traits with complex genetic architecture are known to present difficulties for GWAS (Korte & Farlow 2013). There are two possible scenarios: the traits confounding for GWAS are either controlled by many common variants with small effect on phenotype, or many rare variants each having a large effect on phenotype (Korte & Farlow 2013). Additionally, there is number of other issues such as sample size, mapping panel composition or missing heritability that can affect the performance of GWAS (Brachi et al. 2011, Korte & Farlow 2013). All these limitations are difficult to overcome with the currently used techniques especially when environmentally sensitive traits are analysed (Korte & Farlow 2012, Weigel 2012).

The first GWAS on anion accumulation was described by Atwell et al. (2010) in a proof of concept studies on 107 arabidopsis accessions. In this analysis no significant associations were identified for sulfate concentration data (Atwell et al. 2010). It was concluded that increasing the population to 192 accessions can double the power of GWAS (Atwell et al. 2010). Therefore, in GWAS described in this chapter the data from more than 200 accessions were used to compose each of the three datasets. Nevertheless, it did not improve the outcome of the analysis. GWAS was also not successful in the analysis of total sulfur (and selenium) concentration data in 349 accessions from Borevitz collection published recently (Chao et al. 2014). In the study of Chao et al. (2014) the causal locus for total sulfur was identified using the extreme array mapping (XAM) method which combines bulk segregant analysis with SNP microarray genotyping (Chao et al. 2014, Becker et al. 2011). The results described in Chapter 3 of this thesis and by Chao et al. (2014) indicate that sulfate accumulation is controlled by a number of singleton loci with a large phenotypic effect. The same might be true for nitrate concentration and most likely is the reason of low significance of the studies described here.

An inadequate experimental design might also be a reason for low significance of GWAS described in this chapter (Figure 5.2). The seeds of various accessions were provided by David E. Salt (University of Aberdeen) and Caroline Dean (John Innes Centre, Norwich, UK). Therefore the performance of the plants obtained from these seeds might have been affected

by the place of origin of parental plant. This phenomenon is known as a genomic imprinting (Costa et al. 2012, Kohler et al. 2005, Baroux et al. 2002). To avoid the changes in plant performance due to genomic imprinting the seeds should be cycled at the John Innes Centre (Norwich, UK) at least twice before the actual analysis. However, due to time limitation of my PhD project it was not possible and the plants originated from seeds obtained from different locations were used directly for the analysis of anion concentration. Additionally, more randomization should be introduced during plant growth and harvesting. Instead of only one leaf which was used for the analysis described here, entire rosette should be harvested and an aliquot of this should be analysed. The consistency of the data could be improved by inducing more reference accessions (Chao et al. 2012). These inaccuracies in the experimental design as well as lack of proper normalisation of the data might have affected the outcome of the studies. It can be speculated that introducing all these improvements in experimental design would improve the outcome of the studies. However, taking into account the issues described in previous paragraph it also should be stated that traits with such complex architecture as ion accumulation still present difficulties for GWAS with currently used statistical methods for identification of meaningful associations (Chao et al. 2014, Korte & Farlow 2013, Weigel 2012). There is no certainty that the changes in experimental design would improve the outcome. Therefore, it can be concluded that further improvement of existing methods for GWAS and development of new statistical tools is required to overcome the difficulties described in this paragraph.

Despite the low significance of the GWAS results, it revealed a number of genes which are known to be involved in the control of nutrient homeostasis. An excellent example of such known associations are all the transporter genes: mitochondrial phosphate transporter MPT1 derived from the GWAS on leaf phosphate accumulation data, SULTR3;3 derived from the GWAS on leaf sulfate accumulation data, and the four phosphate transporters derived from GWAS on S/P ratio data (Table 5.7). Additionally, ATPS1 was revealed in the GWAS on leaf sulfate accumulation data. As described in Chapter 3, the natural variation in the transcript abundance of ATPS1 contributes to the control of sulfate metabolism. These “known” associations indicate that there is real information in the data, but the low significance will increase the number of false-positive associations. Therefore, it was necessary to verify the primary candidate genes by the analysis of T-DNA insertion lines. On the other hand, the inconsistent phenotypes of the T-DNA lines analysed here (Figure 5.6 to 5.8) indicate that some of the selected candidates might have been false-positive. Due to time limitation of my PhD project a number of the experiments described in this chapter were not repeated and the

selection of candidates was rough. Therefore, the results described here should be interpreted with caution and the phenotyping should be repeated before possible future analyses. However, it should be stressed that the phenotype of the T-DNA line with the insertion in the promoter of the putative PAPS reductase – a candidate gene characterisation of which is described in the next chapter – was strongly consistent over a number of experiments performed by me and by Anna Koprivova (JIC, Norwich) who used this line for other purposes and the results are not shown here.

An additional confirmation of the real associations in the data was the fact that a number of the strongest associations were common for both methods used (Tables 5.3 to 5.5). However, despite some obvious similarities of the results derived from the two GWAS methods (such as the peak of associations on the chromosome 5 revealed in the GWAS on phosphate accumulation data), differences between significance and location of the strongest associations were clearly visible. Similar phenomenon was observed in the proof of concept studies described by Atwell et al. (2010) who also used two different methods of the correction for population structure – non-parametric Wilcoxon rank-sum test and EMMA method. In the course of their studies, Atwell et al. (2010) have concluded that the P-values produced by EMMA are not always well estimated and should be interpreted with caution. This was also observed in the analysis described here – the EMMA method produced slightly elevated P-values compared to GAPIT, especially in phosphate concentration data. Because of the differences in the results delivered by two different methods, the selection of candidates for follow up studies should be supported by additional information such as the current functional knowledge and publically available information such as membership of genes in specific regulatory networks (Atwell et al. 2010, Verslues et al. 2014, Weigel 2012) or by microarray data, as was presented here.

5.5 Conclusion

GWAS methodology has developed significantly in recent years and can be successfully used as a powerful tool to discover genes underlying traits with simple architecture. However, factors such as genetic heterogeneity, unexpected LD, small effect size, low allele frequency or complex genetic architectures still remain a challenge (Korte & Farlow 2013). Nevertheless, these difficulties can be overcome by complementing and/or verifying GWAS outcomes by additional methods such as QTL mapping or follow up analysis of primary candidate genes. Moreover, current functional knowledge and publically available databases are helpful in identification of meaningful associations.

Despite the low significance of the results, the GWAS analysis described in this chapter revealed a number of candidate genes. The functions of these genes are quite diverse. In general, genes which were selected as candidates fall into four categories: transcription factors, regulatory factors, transporters and enzymes. Each of these genes might be involved in the control of accumulation of one or more macronutrients tested. The follow up analysis of T-DNA lines with insertions in selected candidate genes was conducted to eliminate possible false-positive associations. It led to selection of three final candidates. Each of them showed differences in anion accumulation capacity in the T-DNA insertion lines compared to other T-DNA lines tested and Col-0. Additionally, the analysis of the genetic variation in these genes revealed natural haplotypes that also differed in anion accumulation capacity. These results indicate that selected candidates may be involved in the control of nutrient homeostasis. However, more detailed analysis is required to investigate the exact function of these genes in this process. Nevertheless, GWAS led to identification of known genes that have not previously been related to plant nutrition as well as a previously undescribed gene: AT5G03430. The detailed analysis of this novel gene is the subject of the following chapter of this dissertation.

Chapter 6:

Functional Characterisation of AT5G03430 Gene Revealed by GWAS

6.1 Introduction

The analysis described so far in this thesis has provided new information concerning plant mineral nutrition. Through various collaborations the regulatory functions of two enzymes of the sulfate reduction pathway were characterised and their contribution to natural variation in sulfate accumulation in arabidopsis accessions was investigated. Subsequently, plant adaptation strategies to different sulfate and nitrate availability were investigated and marker traits that define plant growth during nutrient limitation were identified. Finally, GWAS was performed to identify new genes that were not previously known to be involved in plant nutrition, and show that disruptions in these genes are associated with differences in anion accumulation among natural arabidopsis accessions. In this chapter I describe my contribution to the characterisation of a previously undescribed gene identified by GWAS analysis and its involvement in the variation in anion accumulation in different accessions.

The gene identified by GWAS – AT5G03430 – was annotated in TAIR as a putative PAPS reductase family protein (see Chapter 5). In plant cells sulfate is first activated by ATPS to form APS (see Chapter 1). Subsequently, APS is directly reduced to sulfite by APS reductase. However, in bacteria, where the PAPS reductase activity was first identified, APS is further activated by APS kinase to form 3' –phosphoadenosine 5' –phosphosulfate (PAPS). In this case, PAPS can serve as a starting point for the synthesis of secondary sulfate metabolites or can be reduced by thioredoxin-dependent PAPS reductase (PAPR) to sulfite which is subsequently reduced to sulfide and as such, incorporated into O-acetylserine to form cysteine (for details see chapter 1 and (Kopriva & Koprivova 2004, Takahashi et al. 2011).

A single phospho group in PAPS is the only difference between APS and PAPS. APR and PAPR are homologous proteins which share about 20% amino acid sequence identity (Kopriva & Koprivova 2004). The main difference between the two proteins is a lack of an iron-sulfur cluster in PAPS reductase. Additionally, APS reductases possess a C-terminal domain (thioredoxin-like domain) that acts as a glutaredoxin during the transfer of electrons from glutathione (Kopriva et al. 2007). Otherwise, the comparison of APS reductase crystallised from *Pseudomonas aeruginosa* and PAPS reductase crystallised from *Escherichia coli* revealed that the two enzymes have similar structure (Chartron et al. 2006, Savage et al. 1997). The active site of both enzymes is defined by three elements: the P-loop motif, the LDTG motif, and the Arg-loop. Both of these enzymes recognise adenosine as a substrate. Because the adenosine recognition residues are highly conserved between the two proteins it was proposed that PAPS binds to the enzyme in the same manner as APS (Chartron et al. 2006).

The PAPS reductase activity is identified mainly in yeast and enteric bacteria (Schmidt & Jager 1992) whereas APS is utilised by photosynthetic organisms (Schmidt 1975, Schmidt & Trüper 1977). Despite the fact that it has long been known that in plants and algae APS rather than PAPS is the main form of activated sulfur used for further reduction (Tsang et al. 1971) the existence of a sulfate assimilation pathway dependent on PAPS was never excluded. Indeed, PAPS reductase activity has been reported in spinach (Schwenn 1989). Up to date, apart from genes encoding APR no other genes homologous to the PAPS reductase gene identified in *Escherichia coli* have been found in the arabidopsis and rice genome. The first plant species from which a putative gene for PAPS reductase was cloned was the moss *Physcomitrella patens* (Koprivova et al. 2002). It was suggested that in this species PAPS- and APS-dependent sulfate assimilation co-exist. However, further analysis of the two enzymes revealed that even though the PAPS reductase enzyme from *Physcomitrella patens* is structurally more similar to bacterial PAPS reductases (it does not contain the iron-sulfur cluster and the thioredoxin-like domain) it preferentially binds APS (Koprivova et al. 2002). Therefore, the enzyme was renamed as APR-B (Kopriva et al. 2007). These results were remarkable since they revealed that the reduction of APS without an iron-sulfur cluster is possible and that APS-dependent sulfate reduction dominates in plants.

Because no genes encoding PAPS reductase were known in arabidopsis, the identification of AT5G03430 was potentially very interesting. However, the analysis of the sequence of this gene revealed that it does not contain the active site necessary for the PAPS reductase activity (see the results section). Instead, it contains a cinA domain which belongs to the MoCF_BD superfamily (according to NCBI Conserved Domain Database). Domains included in this superfamily such as MoeA or MogA_MoaB are found in arabidopsis Cnx1, and bacterial MoeA – proteins involved in biosynthesis of molybdenum cofactor (MoCo) which possess residues that allow binding of molybdopterin (Rizzi & Schindelin 2002). The present state of knowledge about MoCo synthesis and related aspects of molybdenum metabolism as well as the link between molybdenum and nitrate and sulfate metabolism in plants is described in the following paragraphs.

The specific redox chemistry of molybdenum is used by number of enzymes such as nitrate reductase and sulfite oxidase to catalyse diverse redox reactions (Hille et al. 2011). However, the molybdenum atom cannot be directly introduced into the protein, but needs to be first attached to a specific cofactor scaffold to be able to execute its catalytic function. This molecule is a tricyclic pterin called molybdopterin or metal-containing pterin (MPT), to account for the fact that in some bacteria tungsten is also coordinated by this pterin (Mendel 2013).

MoCo is synthesised in all higher organisms (including humans) by a conserved four-step biosynthetic pathway. In general, six proteins are involved in MoCo biosynthesis in plants (Mendel & Hansch 2002), fungi (Millar et al. 2001), and humans (Reiss et al. 1998, Stallmeyer et al. 1999). Homologs of these genes are found in bacteria. Moreover, bacterial mutants deficient in some MoCo biosynthetic genes can be complemented by their eukaryotic homologs (Mendel 2013). A mutation in MoCo biosynthetic genes leading to inhibition of the pathway has dramatic consequences. The complete loss of MoCo, as observed in *cnx* (**c**ofactor for **n**itrate reductase and **x**anthine dehydrogenase) mutants missing one or more Cnx proteins (see *below*), is lethal for plants grown in soil and has severe consequences for humans because the activity of all MoCo enzymes is altered (Mendel & Kruse 2012).

Conversion of 5'-GTP into cyclic pyranopterin monophosphate (cPMP), known previously as precursor Z, is the initial step of MoCo biosynthesis pathway (Figure 6.1). This reaction is catalysed by two proteins. Cnx2 initiates the transformation of 5'-GTP (Hoff et al. 1995) through the C-terminal [4Fe-4S] cluster. The function of second protein, Cnx3, is not yet understood. However it is believed to be involved in the release of pyrophosphate after the rearrangement reaction (Mendel 2013). cPMP is the first and most stable intermediate of the pathway (Wuebbens & Rajagopalan 1993). The first step of MoCo biosynthesis occurs in mitochondria, since both Cnx2 and Cnx3 were demonstrated to be localised in this cell compartment (Teschner et al. 2010). All subsequent steps were demonstrated to be localised in the cytosol (Kaufholdt et al. 2013). Therefore, cPMP is transported through the mitochondrial inner membrane via a recently identified ABC-transporter: ATM3 (Teschner et al. 2010). In the cytosol MPT is generated by transfer of sulfur to cPMP in a two-step reaction catalysed by MPT synthase (Figure 6.1; Leimkuhler et al. 2011). Subsequently, the molybdenum atom is incorporated into the chemical backbone of the MPT moiety. This step is catalysed by molybdenum insertase – Cnx1 (Schwarz et al. 2000). The two domains of the enzyme are named E-domain and G-domain and each has a different mechanistic function. Cnx1G activates the MPT before insertion of the metal and Cnx1E catalases subsequent insertion of molybdenum (Llamas et al. 2004, Llamas et al. 2006). As a result a physiologically active MoCo is synthesised.

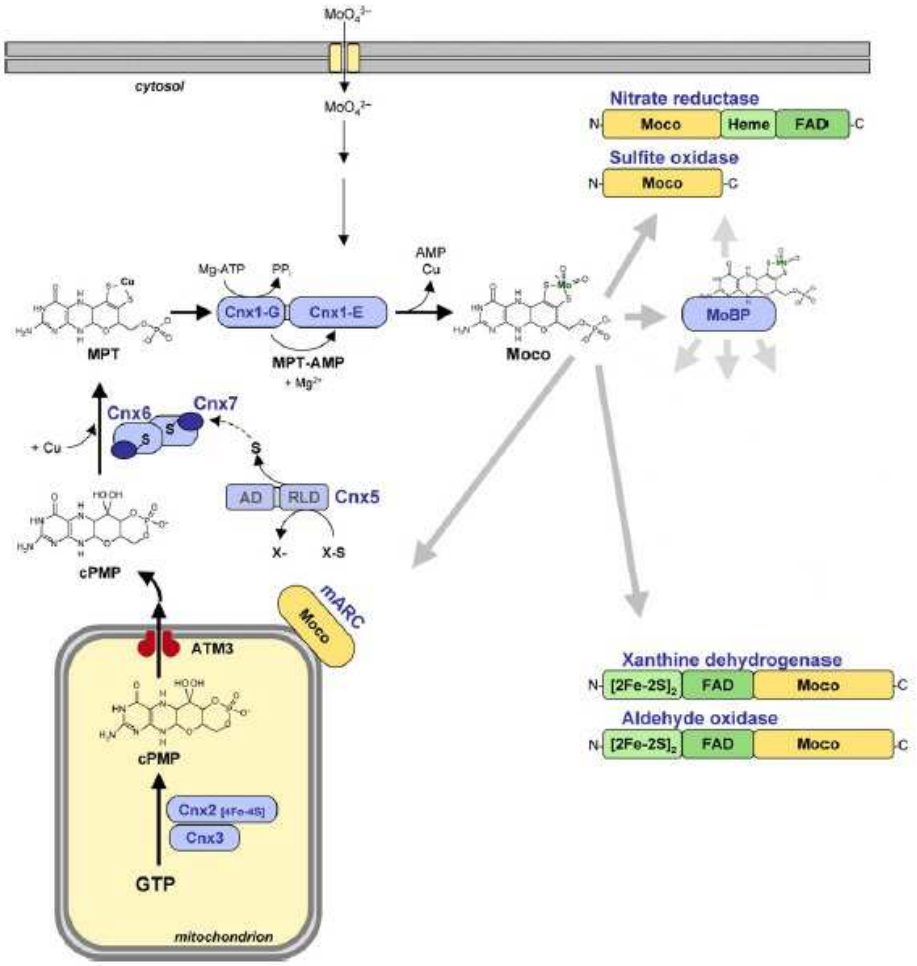


Figure 6.1 ***Biosynthesis of MoCo in plant cell***
MoCo biosynthesis pathway starts from the conversion of GTP to cPMP in the mitochondria. The enzymes involved in MoCo biosynthesis are shown in blue boxes. The five enzymes which use MoCo as a cofactor are indicated with full names in blue and their schematic structure is shown as yellow and green boxes. Abbreviations: (AD) adenylation domain of Cnx5, (RLD) rhodanese-like domain of Cnx5, (MoBP) MoCo binding proteins. The figure is modified after Mendel 2011.

Since mature MoCo is extremely sensitive to oxidation it is known to be permanently-bound to protein in the cell (Rajagopalan & Johnson 1992). It needs to be transported immediately to the target enzymes to minimize its degradation. Both of these requirements are fulfilled by MoCo-binding proteins, which bind newly synthesised MoCo allowing its storage and providing it according to demand (Mendel 2013). The first MoCo-binding protein was identified in *Chlamydomonas reinhardtii* (Witte et al. 1998) and named MoCo carrier protein (MCP). Recently, a family of eight MCP-related proteins was identified in arabidopsis (Kruse et al. 2010). However, they seem to be involved in cellular distribution of MoCo rather than storage since they undergo protein-protein interactions with both Cnx1 and the MoCo

acceptor protein nitrate reductase (NR; Kruse et al. 2010). To date, the exact mechanism of MoCo insertion into the acceptor proteins is not understood. The crystallographic analysis of MoCo enzymes revealed that the cofactor is usually localised deep inside the holoenzymes (Dobbek 2011). Since all eukaryotic MoCo enzymes are dimers, it has been suggested that MoCo needs to be incorporated into the targeted protein during the folding and dimerization of the apoprotein monomers, before the mature protein is synthesised (Mendel 2013).

The eukaryotic molybdenum enzymes fall into two groups (Hille et al. 2011). The first, called the sulfite oxidase (SO) family, also includes nitrate reductase (NR) and the mitochondrial amidoxime-reducing component (mARC; Figure 6.1). SO and NR are links connecting molybdenum with sulfate and nitrate metabolism. Therefore, it is possible that AT5G03430 which was identified from GWAS from both nitrate and sulfate accumulation data (see Chapter 5) can be involved in molybdenum metabolism. Enzymes from this family are activated by insertion of MoCo. The second group is called the xanthine oxidase (XO) family and includes xanthine dehydrogenase (XDH) and aldehyde oxidase (AO; Figure 6.1). Enzymes from this family require the addition of a terminal sulfido group to the MoCo during or after its insertion which is the final maturation step necessary to gain the enzymatic activity (Hille et al. 2011). In addition to these proteins which contain the pterin type of cofactor, there is another type of Mo-containing cofactor which is known in only one type of enzyme in nature. Bacterial nitrogenase, which is required for biological nitrogen fixation, contains the so-called iron-molybdenum cofactor, FeMoCo (Hu & Ribbe 2013).

Sulfite oxidase catalyses the oxidation of sulfite to sulfate (Eilers et al. 2001). It has a sulfite-detoxifying function, removing excess sulfite from the cell (Brychkova et al. 2007). In plants the reaction catalysed by SO results in production of hydrogen peroxide. Therefore, SO is localised in peroxisomes where the hydrogen peroxide can be easily eliminated by catalase (Nowak et al. 2004). Similarly to mARC it possesses only MoCo as a redox centre and is the simplest molybdenum enzyme found in plants (Figure 6.1). However, the animal SO is much more complex. It consists of an iron-heme N-terminal domain containing cytochrome b_5 and the C-terminal domain which is responsible for MoCo binding and dimerization (Kisker et al. 1997). It is localised in the intermembrane space of mitochondria and a loss of its activity results in severe neurological disability and early death (Johnson & Rajagopalan 1979). The second member of the SO family is NR, which is the main enzyme of nitrate assimilation. In the cytosol it catalyses the reduction of nitrate to nitrite providing essential nitrogen metabolites to the plant (Eckardt 2005). The MoCo domains of SO and NR are very similar (Fischer et al. 2005, Schrader et al. 2003). mARC is the last and the least known (at least in plants) member of

the SO family. In eukaryotes mARCs form a small protein family (Wahl et al. 2010). They were predicted to be localised in mitochondria and it is assumed that they play a detoxifying role (Plitzko et al. 2013).

The second group of Mo-enzymes includes molybdo-flavoenzymes that are known to catalyse the oxidative hydroxylation of a number of aldehydes and aromatic heterocycles. Xanthine dehydrogenase (XDH) is the first member of the family and a key enzyme of purine degradation. It was shown to oxidize hypoxanthine to xanthine and further to uric acid (Zarepour et al. 2010). The second member of the family AO converts aromatic and non-aromatic heterocycles and aldehydes to the respective carboxylic acids via oxidation (Yesbergenova et al. 2005). AO is essential for the biosynthesis of abscisic acid (ABA) – a plant hormone which plays an important role in seed development and dormancy and plant response to various environmental stresses (Seo & Koshiba 2002). AO and XDH show a high level of sequence similarity and the main difference between the two enzymes concerns the substrate binding at the MoCo centre as well as binding of the physiological acceptor of electrons. Since AO are strict oxidases they are unable to bind NAD^+ and they use molecular oxygen as electron acceptor (Hille 2005).

In this chapter I describe my attempt to functionally characterise the novel protein revealed by GWAS. This analysis is still at an early stage; however some interesting results can be already presented. First of all, I conducted a bioinformatics analysis of the gene and the protein using publicly available bioinformatics tools to get an idea about the possible function of the protein. In the characterisation of AT5G03430, with advice from Vasilios Andriotis (Metabolic Biology, John Innes Centre), I performed an experiment which revealed that a mutation in this gene is embryo-lethal. I also prepared GFP-fusion constructs to verify the subcellular localisation of the protein. Moreover, I undertook some steps to explain the involvement of the protein in the metabolism of nitrate and sulfate as well as metabolism of molybdenum. I analysed a number of arabidopsis accessions with different versions of the gene to investigate how the changes in the gene affect anion accumulation as well as activity of enzymes possibly related to/regulated by the product of AT5G03430.

6.2 Materials and Methods

6.2.1 Growth conditions in the soil

Accessions with different architecture of AT5G03430 gene used for the analysis of anion accumulation and mature plants for the expression analysis of AT5G03430 (Figure 6.6C) were grown for five weeks in the CER as described in Chapter 5 (see section 5.2.1) Plants for the analysis of anion accumulation were supplied with 1 l of ¼Hoagland solution per tray twice a week in order to provide sufficient nutrient supply over the growth period (see Chapter 5).

Plants used to study the time course of embryo development and for transformation were sown directly in the soil (Levington Horticulture soil mix, Ipswich, UK) in single pots (9 cm diameter) or in a 40-cell tray. They were first stored at 4°C for three days to break the seed dormancy and subsequently moved to a glasshouse where they were grown until seed production.

6.2.2 Growth conditions on plates

Young plants for the analysis of expression of AT5G03430 (Figure 6.6A, B) were grown on plates with MS medium without sucrose for three weeks as described in Chapter 5 (see section 5.2.3).

6.2.3 Bioinformatic tools and software

The arabidopsis Information Resource (TAIR) database was used for the functional annotation of the genes (http://www.arabidopsis.org/); National Centre for Biotechnology Information (NCBI; http://www.ncbi.nlm.nih.gov/) was used to obtain the gene and protein sequences; the Conserved Domain Database (NCBI CDD) was used to investigate the conserved domains in the protein sequences (Marchler-Bauer et al. 2011); Blast was used to look for homologous sequences (Altschul et al. 1990); ClustalW was used for multiple alignment of gene sequences (Larkin et al. 2007); 1001 Genomes database was used for the investigation of the natural variation in the architecture of candidate genes (http://signal.salk.edu/atg1001/3.0/gebrowser.php); the "Stress Response" tool within Genevestigator (Zimmermann et al. 2004; https://www.genevestigator.com/gv/plant.jsp) was used to investigate the changes in gene expression in different conditions; Mfold software was used for the analysis of DNA secondary structure (Zuker 2003); the molybdenum accumulation data were obtained from publically available ionome database iHUB (Baxter et al. 2007; www.ionomicshub.org), Vector NTI Advance Suite 11 (Invitrogen) software was used for the analysis of sequencing data; the T-DNA Express search tool was used to obtain T-DNA lines with insertions in genes of interest

(Alonso et al. 2003); Arabidopsis electronic Fluorescent Pictograph (eFP) browser was used for expression data of gene of interest (Winter et al. 2007); Subcellular Proteomic Database (SUBA3) was used for the prediction of subcellular location of gene of interest (Tanz et al. 2013);

6.2.4 Genotyping of the T-DNA lines

The DNA from plant tissue was isolated as described in Chapter 2. The genotyping of T-DNA lines was carried out on genomic DNA by PCR amplification as described in Chapter 5 using gene specific primers listed in Table 6.1M and the left-border SALK primer for Salk lines.

Table 6.1M ***Gene specific primers used for genotyping***

Name	T-DNA line	Gene	Sequence 5' -> 3'
34Fw	SALK_141801C	AT5G03430	CAGATTCTTCTTCTTCTCG
34Rw			cgatgatcaattggcttcg
PAPS1_1Fw	SALK_031937	AT5G03430	AGTTGCAGATTGTAACTGCCG
PAPS1_1Rw			gcccattattagtgtccgagac
PAPS1_2Fw	SALK_043060	AT5G03430	GCTCAAAATTTTCTGCACGTC
PAPS1_2Rw			cgccatcttcgtcattaaaag
PAPS1_3Fw	SALK_128101	AT5G03430	GAATTGCTCTTGACCAACCTG
PAPS1_3Rw			tgcagatactccgaaatttgg
PAPS1_4Fw	SALK_140720	AT5G03430	AGGCACCTGATGAAGAGTTTG
PAPS1_4Rw			tgtgaccggtacatagaagcc
PAPS1_5Fw	SALK_144629C	AT5G03430	GAGGTCGTGTGGGATATGTTG
PAPS1_5Rw			acagaaacgagagcagagcag
PAPS1_6FW	SALK_140626	AT5G03430	GTGTTGCTAAGGCATTTGGAG
PAPS1_6RF			tgtgaccggtacatagaagcc
SALK_LB1	Salk lines	insertion	GCGTGGACCGCTTGCTGCAACT

One of the primers from each pair which is underlined and bolded was used in combination with the Salk left-border primer to identify homozygous lines.

6.2.5 Time course of arabidopsis embryo development

The siliques from the main flowering stem of three plants were analysed after the flowering of the plant. Every second silique (starting from the youngest one) was opened under a dissecting microscope. Developing seeds were cleared on microscope slides according to Andriotis et al. (2010) overnight at room temperature using Hoyer's solution (100 g chloral hydrate, 5 ml glycerol, 30 ml water; Liu & Meinke 1998). Whole-mount preparations of cleared seeds were viewed under differential interference contrast (DIC) optics with a Leica DM6000 microscope operated with the Leica LAS AF7000 software. For the quantification of abnormal seeds up to six siliques from three heterozygous plants were opened under the

dissecting microscope and the number of phenotypically normal and abnormal seeds was calculated.

Table 6.2M *Primers used for qPCR and RT-PCR of AT5G03435*

Name	Gene	Sequence 5' -> 3'
PAPSqFw	AT5G03430	TGCACAGGTGACAGGAATGAGATG
PAPSqRw		ttcccactctttctcgagctcctc
03435qFw	AT5G03435	GTGGGATATGTTGAAGCCGAAGAC
03435qRw		aaagcaagcgcagcaactttcc
03435qFwa	AT5G03435	GTGGGATATGTTGAAGCCGAAGAC
03435qRwa		gcgcagcaactttccgcatatc
03435qFwb	AT5G03435	TGGGATATGTTGAAGCCGAAGACG
03435qRwb		gcgcagcaactttccgcatatc
03435rtFw	AT5G03435	ATGGCTGCCAATAAAGATGAATTCTCCGTC
03435rtRw		gaacaatacttcgtttgtgggcaatcttcg
S18F	S18	GGTACGTGCTACTCGGATAACC
S18R		tctccggaatcgaacccta
ACT2F	actin	GCACCCTGTTCTTCTTACCG
ACT2R		aaccctcgtagattggcaca

The primers for qPCR were generated using the Quant Prime program

6.2.6 Expression analysis

RNA was isolated from frozen plant tissue using the RNeasy Plant Mini Kit (Qiagen) following the manufacturer's instructions. RNA was eluted from the RNeasy spin column using 30 µl of RNase-free dH2O and its quality and quantity were assessed by measurements of absorbance at 260 nm and 280 nm using the NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, http://www.nanodrop.com/). Subsequently the RNA samples were DNase treated using the DNase Ambion kit (Life Technologies™) according to the manufacturer's instructions. The quality and quantity of RNA were assessed using the NanoDrop spectrophotometer as described above. Highly concentrated samples with good quality RNA were used for cDNA synthesis. For the reaction 1.5 µl of Oligo(dT) and 1.5 µl of dNTP were added to 1 µg of RNA in the final reaction volume of 30 µl. Samples were incubated at 65°C for 5 min and cooled down on ice for 1 min. For reverse transcription the SuperScriptIII Reverse Transcriptase (RTase) kit (Invitrogen) was used. To each sample, 6 µl of 5x First Strand Buffer, 1.5 µl of RNAsin RNAse Inhib, 1.5 µl of 0.1M DTT, and 1.5 µl of SuperScriptIII RTase were added. The samples were gently mixed and incubated at 50°C for 50 min and subsequently 70°C for 15 min to inactivate the RTase. After that 30 µl of water were added to each sample and they were tested for genomic contamination via PCR reaction with primers specific for actin (ACT2 – Table 6.2M) against the genomic DNA as a control. The transcript

abundance was measured via real-time quantitative RT-PCR (qPCR) using 1 µl of cDNA and the fluorescent intercalating dye SYBR Green (Applied Biosystems) in a DNA engine OPTICON2 continuous fluorescence detector (Bio-Rad). The qPCR was performed using gene specific primers and the results were normalised to the *S18* gene (Table 6.2M). The qPCR was performed in triplicate for each of three independent samples.

6.2.7 Complementation of heterozygous T-DNA line with wild-type Col-0 DNA

The cDNA obtained for the expression analysis (see section 6.2.6) was PCR amplified using Platinum™ High-Fidelity DNA Polymerase (Invitrogen) which leaves 3'-A overhangs on the PCR products and gene specific primers (Table 6.3M). PCR products were purified using the QIAquick®PCR-purification kit (Qiagen). Cloning of the PCR fragments was performed using a reaction mix containing: 0.5 µl PCR product, 1 µl salt solution provided with the vector, 0.5 µl vector (pCR8/GW/TOPO TA) and dH_2O in a final volume of 3 µl. Subsequently, samples were incubated for 10 min at room temperature and transformed into chemically competent *E. coli* TOP10 or HD5α cells via a heat shock (see Chapter 2). Transformants were selected on LB medium supplemented with spectinomycin. Subsequently, the plasmid DNA was isolated from overnight liquid bacterial cultures using the QIAprep Spin Miniprep Kit (Qiagen) according to the manufacturer's instruction. Successful cloning was confirmed by restriction digestion of plasmid DNA with EcoRI (Roche; see Chapter 2). Successful transformants were sequenced (section 6.2.9) with plasmid specific primers by the sequencing service at Eurofins Genomics (www.eurofins.com). The constructs with sequences identical to template were then transformed from pCR8 entry vector to the destination vector pGWB2 via overnight LR reaction (see Chapter 2). Subsequently the constructs were transformed to chemically competent *E. coli* TOP10 or DH5α cells via a heat shock (see Chapter 2). Transformants were selected on LB medium supplemented with kanamycin and hygromycin. Successful cloning was confirmed by colony PCR. Subsequently, the successful constructs were transformed to electrocompetent *Agrobacterium tumefaciens* via electroporation (see Chapter 2). Successful transformants were selected on LB medium supplied with carbenicilin, rifampicin, kanamycin, and hygromycin. The successful transformants were used to prepare the cell suspension and the plants were transformed via the floral dip method (see Chapter 2). Successful plant transformants were selected on MS medium supplemented with hygromycin.

Table 6.3M *Primers used for high-fidelity PCR amplification of AT5G03430 for cloning*

Name	Sequence 5' -> 3'	Comment
PAPSFw	CCGGACTGGAAAACTTGATAGACCACC	With START codon
PAPSRw	CTTGATTTCTACGAACACATCTTTCTTGAATTTCTTGCG	Without STOP codon

The 1.2kb fragments cloned for sequencing of the fragment indicated in the 1001 Genomes database as a deletion were amplified with primers PAPS1_1Fw and PAPS1_1Rw from Table 6.1M.

6.2.8 Subcellular localisation of GFP fused AT5G03430

The cDNA obtained for the expression analysis (section 6.2.6) was PCR amplified using Platinum™ High-Fidelity DNA Polymerase, cloned to pCR8 vector and successful transformants were selected and verified as described before (section 6.2.7). Since the destination vector pK7FWG2,0 has resistance to the same antibiotic as the entry vector the constructs with sequences identical to template were first digested with BglI (New England Biolabs) to disable the resistance gene of the entry vector. Subsequently, the plasmid DNA was purified using the QIAquick®PCR-purification kit (Qiagen) and introduced to the destination vector pK7FWG2,0 via overnight LR reaction (see Chapter 2). The constructs were then transformed to chemically competent *E. coli* TOP10 or DH5α cells via a heat shock (see Chapter 2). Transformants were selected on LB medium supplemented with spectinomycin. Successful cloning was confirmed by restriction digestion with ClaI and BspHI (Roche, New England Biolabs respectively) and colony PCR. Subsequently, the successful constructs were transformed to electrocompetent *Agrobacterium tumefaciens* via electroporation (see Chapter 2). Successful transformants were selected on LB medium supplied with spectinomycin, carbenicilin, and rifampicin. The successful transformants were used to prepare the cell suspension used to infiltrate the *Nicotiana benthamiana* plants (see Chapter 2). After three days the tobacco leaves were analysed under a confocal microscope Leica SP5 (II).

6.2.9 Sequencing of plasmid DNA

The DNA for sequencing was PCR amplified using the Platinum™ High-Fidelity DNA Polymerase and cloned to pCR8 vector. Subsequently the plasmid DNA from the successful transformants was isolated using the QIAprep Spin Minipre Kit (Qiagen) according to the manufacturer's instructions. The plasmid DNA was eluted from the spin column using 30 µl of dH_2O and its quality and quantity was assessed using the NanoDrop spectrophotometer (NanoDrop Technologies). Subsequently, the appropriate dilutions of DNA were prepared and sequenced with plasmid and gene specific primers (Table 6.4M) by the sequencing service at Eurofins Genomics.

Table 6.4M *Primers used for sequencing*

Name	Sequence 5' -> 3'	Comment
M13 rev (-29)	CAG GAA ACA GCT ATG ACC	Supplied by Eurofins, specific for pCR8 vector
M13 uni (-21)	TGT AAA ACG ACG GCC AGT	
Seq960F	TATCATTCGCCAGGATTTCAAATCC	Designed by me to sequence genomic DNA of AT5G03430 with the promoter
Seq1910R	CTTACAGAATCTCATCGCCGACAGC	
Seq1730F	GGTTCCCAACTCATTATTGTCTG	
Seq2610R	CACATTCATTGAGATGTTAAGGCAC	
Seq2550F	GGAGTTCGCCTGGTACGTATC	
Seq3130R	CAGACATTCCCACTCTTTCTCG	
SeqProm300F	GGTGGGTTTTAGATAAATTTGGTGTGG	Designed by me to sequence genomic DNA of AT5G03430 with the promoter
SeqProm1570R	GCATCTGGATCTTGAGATTAAAGATTTGAAG	
SeqProm130F	CTCTCTCTCGCCATGAGTAGTAACACGCGG	
SeqProm600R	CTGAAGAATGAGAGAGAGAGAGAGCTCACGAG	

.

6.3 Results

6.3.1 Bioinformatics analysis of AT5G03430 sequence

The GWAS on nitrate, phosphate, and sulfate accumulation data and subsequent analysis of T-DNA insertion lines described in Chapter 5 revealed a number of candidate genes that could potentially be linked to the regulation of anion accumulation. Moreover, data supporting the hypothesis that disruptions in selected genes may affect anion accumulation in these lines were provided (Figure 5.9). Among genes selected as candidates AT5G03430, annotated in the TAIR database as a putative PAPS reductase family protein, seemed particularly interesting, since PAPS reductase activity had not been previously identified in arabidopsis. AT5G03430 was revealed in GWAS on both nitrate and sulfate concentration data. The T-DNA line with insertion in the gene promoter was characterised by a low nitrate concentration, but there was no difference in sulfate concentration compared to Col-0 (see Chapter 5). The low nitrate phenotype was consistent over a number of experiments performed by myself and Anna Koprivova who used this line for other purposes. Additional information on AT5G03430 was delivered from the investigation of its architecture in the 1001 Genome database which revealed a number of accessions with various deletions and amino acid changes in the gene body. Further analysis showed that these accessions accumulate in general less nitrate compared to accessions with gene sequence similar to Col-0 (Figure 5.9B). There was no difference in sulfate concentration in these accessions when compared to accessions with the gene coding region as in Col-0 (see Chapter 5). These results indicated that the variation in AT5G03430 gene sequence is likely to affect nitrate accumulation. However the exact mechanism was not known at this stage.

First, I investigated the information available online in various databases to get a better understanding of the function of the protein encoded at AT5G03430 and its possible involvement in nitrate and/or sulfate metabolism. Analysis of the protein sequence using the NCBI conserved domain search tool (Marchler-Bauer et al. 2011) revealed that the product of AT5G03430 is a 497 amino acid protein containing two conserved domains: a PAPS reductase domain specific for PAPS reductase enzymes and a cinA domain similar to domains found in proteins involved in MoCo biosynthesis which bind MPT (Figure 6.2A). However, further analysis of the protein sequence revealed that AT5G03430 does not have the conserved residues which are crucial for PAPS reductase activity (Figure 6.2B). These residues are conserved among species and found in both plant APS reductases and bacterial PAPS reductases.

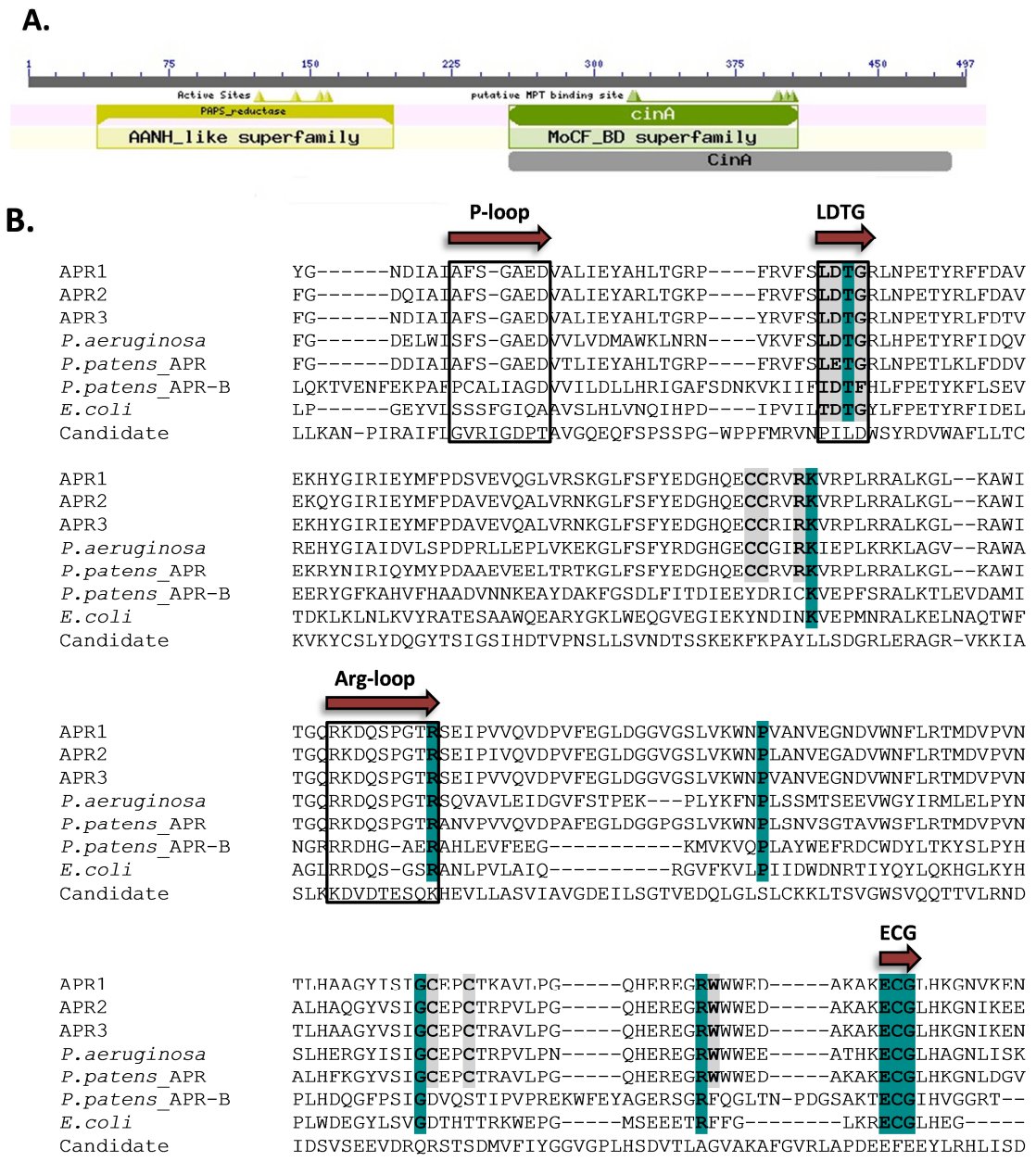


Figure 6.2 ***Conserved domains in AT5G03430***

(A) The conserved domains found in the sequence of the entire protein encoded by AT5G03430 using the NCBI Conserved Domain Database (Marchler-Bauer et al. 2011). The result indicates the presence of two domains in the submitted protein sequence which was obtained from NCBI database: PAPS reductase domain (yellow box) and cinA domain (green box). (B) A fragment of the alignment of APS/PAPS reductase domains from *Arabidopsis thaliana* (APR1, APR2, APR3), *Pseudomonas aeruginosa* (APS reductase), *Physcomitrella patens* (APS reductase and APR-B), and *Escherichia coli* (PAPS reductase) with first 225 amino acids of AT5G03430 protein sequence indicated by Conserved Domain Database as PAPS reductase domain (Candidate). The red arrows and black frames indicate conserved motifs defining the active site of the enzymes. Strictly conserved residues are marked in green, semi conserved residues or residues common for APS reductases only are marked in grey. The multiple alignment was made by using the ClustalW software. The numbers are created by the software and do not correspond to actual positions of particular residues in the protein sequence since only a fragment of the sequence was used for the alignment.

Figure 6.2B presents the alignment of the putative PAPS reductase domain of AT5G03430 with different isoforms of arabidopsis APS reductase and bacterial APS and PAPS reductases. It revealed that the fragment of AT5G03430 indicated by the conserved domain search tool as PAPS reductase domain (first 225 amino acids of the protein) does not contain any of the residues recognised before as crucial for this enzymatic activity. Among the motifs that define the active site of the enzyme, the P-loop plays an important role in substrate recognition and specificity (Chartron et al. 2006). It consists of FS-GAED sequence which is highly conserved in APS reductases and related motifs in other ATPases (Mougous et al. 2006). It aligns with SSSFGIQA motif from the PAPS reductase from *E. coli* which contains the consensus SXG motif that is involved in the interaction with phosphate of AMP (Savage et al. 1997). It has been proposed that the two last amino acids found in P-loop motif in APS reductases – Gly and Asp – which are not present in this motif in PAPS reductases are mimicking the interaction of a negatively charged 3'-phosphate group in PAPS reductase (Chartron et al. 2006). None of the residues of the P-loop motif are found in the AT5G03430 protein.

Correspondingly, none of the residues of LDTG motif or Arg-loop are found in the PAPS reductase domain of AT5G03430. The LDTG motif plays an important role in the proper conformation of the protein as well as substrate binding (Chartron et al. 2006). Threonine in this motif is strongly conserved among a number of species in both APS and PAPS reductases and is involved in stabilizing the protein. Glycine from this motif is also highly conserved and it has been recognized to be involved in the proper conformation of the protein (Chartron et al. 2006, Savage et al. 1997). Only APR-B from *Physcomitralla patens* does not share this residue on the alignment shown in the Figure 6.2B. Similarly, the residues building Arg-loop, especially the last arginine from the motif which plays a crucial role in alternate conformation of Arg-loop (dependent on the presence of the substrate), are conserved among both APS and PAPS reductases but they are not present in the AT5G03430 protein (Figure 6.2B; Chartron et al. 2006, Savage et al. 1997).

Most importantly, the ECG motif that contains the catalytic cysteine is not present in the PAPS reductase domain from AT5G03430 (Figure 6.2B). To date this motif has been found in all identified APS/PAPS reductases and is crucial for the catalytic function of the enzyme. Moreover, the cysteine from this motif which is the only conserved cysteine in both APS and PAPS reductases has been suggested to be involved in the formation of a disulphide bond that joins the two monomers of the enzyme (Kopriva & Koprivova 2004).

PAPS/APS reductases are not the only enzymes that contain the PAPS reductase domain. Some bacteria such as *Pseudomonas syringae* possess a specific sulfate assimilation gene cluster in which genes encoding ATPS are associated with genes encoding GTP hydrolases (Mougous et al. 2006). Because the formation of APS is thermodynamically unfavourable bacteria evolved mechanisms that shift the reaction equilibrium towards production of APS (Liu et al. 1994). In bacteria such as *Pseudomonas syringae* the ATPS activity is coupled with GTP hydrolysis activity in one protein. In that protein, the monomer that is able to bind GTP possesses the PAPS reductase domain. The GTPase activity is gained after the two monomers are coupled as a heterodimer (Mougous et al. 2006).

Another group of proteins which possess the PAPS reductase domain are FAD synthases. FAD is a cofactor necessary for the functioning of large variety of dehydrogenases, reductases, and oxidases involved in a number of vital processes in the cell (Joosten & van Berkel 2007). It is synthesised from riboflavin in a two step reaction catalysed by riboflavin kinase and FAD synthase (Massey 2000). The first eukaryotic gene for FAD synthase was identified in *Saccharomyces cerevisiae* and named *FAD1*. In yeast this protein contains a PAPS reductase domain and is not similar to the bacterial FAD synthases (Leulliot et al. 2010). The sequence of yeast FAD synthase was used to identify two human isoforms of that protein (Brizio et al. 2006). In contrast to yeast FAD synthase, the enzyme found in humans is organised in two domains: the PAPS reductase domain at the C-terminus and the cinA domain at the N-terminus (Figure 6.3A). It has been shown that in the human FAD synthase the PAPS reductase domain itself is able to catalyse FAD synthesis and its cleavage (Miccolis et al. 2012). The role of the cinA domain has not been described in detail. However, the analysis of recombinant PAPS domain (Miccolis et al. 2012) revealed differences in the response to Mg^{2+}, Co^{2+}, and Ca^{2+}, and GTP compared to the response observed in the analysis of the entire protein (Torchetti et al. 2011). Therefore, it was suggested that this domain may play some role in the regulation of hFADS contributing to binding of GTP in a site different from the substrate sites (Miccolis et al. 2012). It is worth noticing that the FAD synthetases identified in the arabidopsis genome – RibF1 and RibF2 – have a completely different structure from the human enzymes (Sandoval et al. 2008).

The Blast search against the sequence of the PAPS reductase domain from AT5G03430 protein (amino acids from 1 to 224) revealed 90% similarity and 42% identity between it and FAD synthase isoform 1 and 2, 83% similarity and 40% identity with FAD synthase isoform 3 and 79% similarity and 32% identity with yeast FAD synthase. The alignment of the sequences from these proteins revealed that PAPS reductase domain from AT5G03430 shares all the

residues involved in the conserved PP-loop, ARG1, flavin and γ-phosphate motifs commonly found in FAD synthases from different species (Figure 6.3B). In FAD synthases the PP-loop and ARG1 motifs are known to be involved in the binding of ATP, the flavin motif is involved in FMN binding, and the function of γ-phosphate motif is not clear yet (Huerta et al. 2009, Miccolis et al. 2012). The high similarity of the PAPS reductase domains from the two proteins might have been a reason why the AT5G03430 is annotated as a putative PAPS reductase family protein.

Likewise, the Blast search against the sequence of the cinA domain from AT5G03430 protein (amino acids from 225 to 497) revealed 86% similarity and 21% identity between it and FAD synthase isoform 1 and 2, and 78% similarity and 20% identity with FAD synthase isoform 4. In contrast to other isoforms of human FAD synthase, the isoform 4 consists of a cinA domain only. According to Conserved Domain Database all of these proteins share the sequence features important for binding of molybdopterin (Figure 6.3C). Blast search also revealed that the cinA domain is similar to Cnx1G domain of Cnx1 from *Arabidopsis thaliana* (56% similarity, 15% identity), mammalian gephyrin isoform 1 and 2 (62% similarity, 12% identity), and *Drosophila melanogaster* cinnamon protein (37% similarity, 15% identity). There was also a high similarity to three monofunctional proteins from *Escherichia coil* – MoeA (47% similarity, 19% identity), MogA (40% similarity, 16% identity), and MoaB (38% similarity, 16% identity). All these proteins (except FAD synthases) are functional homologues and are known to be involved in the final step of the MoCo biosynthesis i.e. incorporation of molybdenum into molybdopterin which yields the active cofactor (Schwarz & Mendel 2006). Moreover, human gephyrin and cinnamon protein for Drosophila have a number of assigned functions in addition to MoCo biosynthesis (Nawrotzki et al. 2012, Wittle et al. 1999).

In summary, there is no similarity between AT5G03430 and the known APS and PAPS reductases (Figure 6.2). Because no homologues to PAPS reductase from *E. coli* other than APR have been identified in arabidopsis or rice it could be that such an enzyme does not exist or has a structure completely divergent from the bacterial enzyme. Nevertheless, it seems unlikely that AT5G03430 has this activity, especially when taking into account the high similarity of the protein to FAD synthases and proteins involved in MoCo biosynthesis (Figure 6.3). MoCo biosynthesis as well as molybdenum metabolism are linked to the metabolism of nitrate and sulfate (see discussion of this Chapter). The T-DNA line with insertion in AT5G03430 showed low nitrate concentration and no changes in sulfate accumulation (see Chapter 5; Figure 5.6). Moreover, the role of the cinA domain and the structural relationship between the two domains in FAD synthases are not known. Therefore, I decided to focus

further analysis of the protein encoded by AT5G03430 on the function of the cinA domain and the involvement of this protein in nitrate and/or sulfate metabolism.

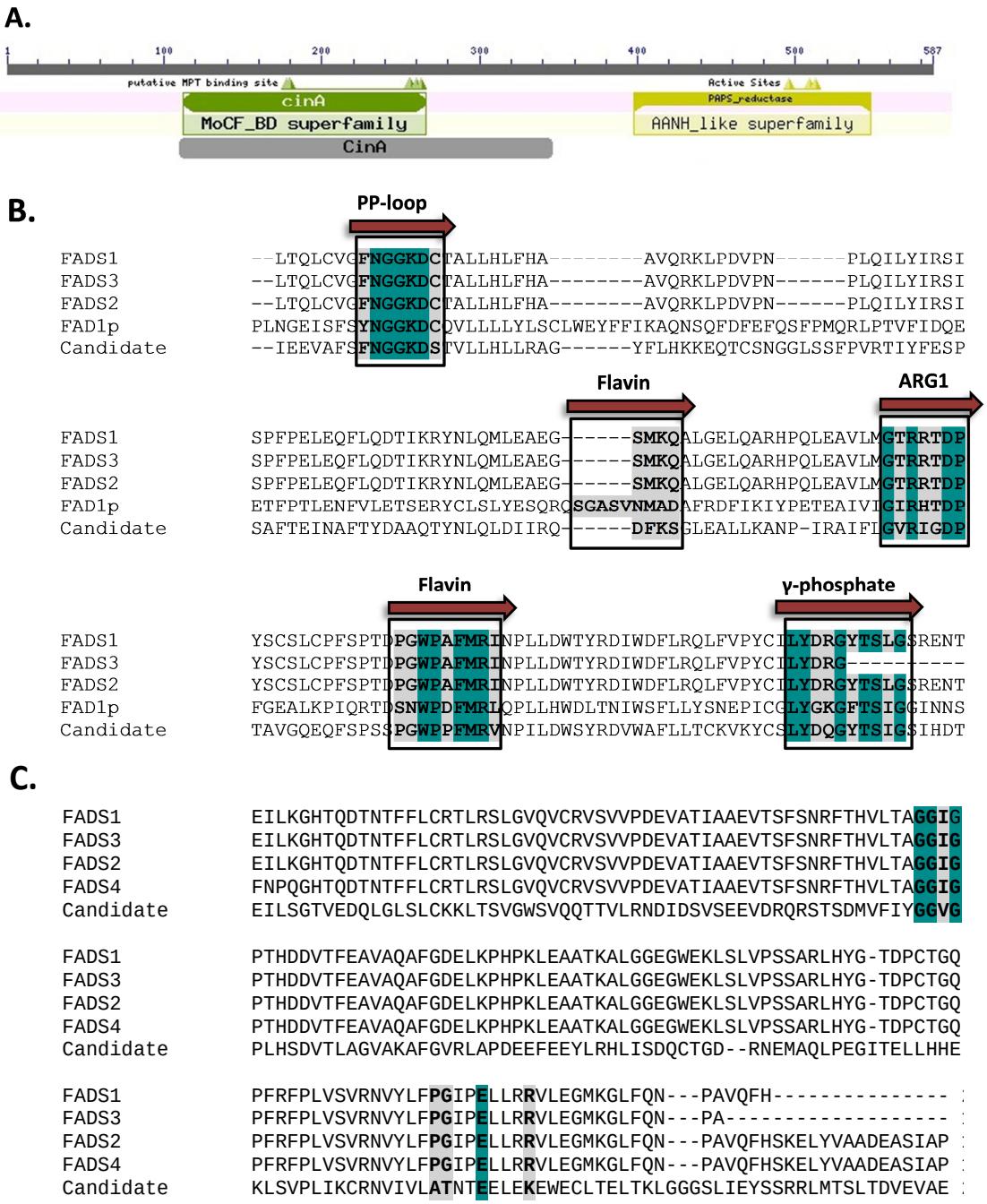


Figure 6.3 ***Comparison of conserved domains from AT5G03430 with other proteins***

(A) The conserved domains found in human FAD synthase isoform 1 protein sequence by using the NCBI Conserved Domain Database (Marchler-Bauer et al. 2011). Two domains are indicated: cinA domain (green box) and PAPS reductase domain (yellow box). (B) Fragment of the alignment of PAPS reductase domains from different isoforms of human FAD synthase (FADS1, FADS2, FADS3), *Saccharomyces cerevisiae* FAD synthase (FAD1p), with first 225 amino acids of AT5G03430 protein sequence indicated by Conserved Domain Database as PAPS reductase domain (Candidate). The red arrows and black frames indicate conserved motifs defining the active site of the enzymes. Strictly conserved residues within these motifs are marked in green, semi conserved residues

within these motifs are marked in grey. (C) Fragment of the alignment of cinA domains from different isoforms of human FAD synthase (FADS1, FADS2, FADS3, FADS4) with the part of AT5G03430 sequence indicated by Conserved Domain Database (CDD) as a cinA domain (amino acids 225 to 497). Strictly conserved residues that are involved in MPT binding (according to CDD search tool) are marked in green, semi conserved residues that are involved in MPT binding are marked in grey. The multiple alignment was made by using the ClustalW software. The numbers are created by the software and do not correspond to actual positions of particular residues in the protein sequence since only a fragment of the sequence was used for the alignment.

6.3.2 The embryo-lethality of AT5G03430 mutant

The T-DNA line with insertion in the promoter showed low nitrate concentration. Therefore, I obtained additional T-DNA lines with insertions in this gene to investigate the effect of disruption of the gene on anion accumulation and eventually, the exact function of the gene product (Table 6.1). I ordered six additional T-DNA lines from NASC and genotyped them via PCR. The seeds for one line did not germinate. The genotyping of the other five lines revealed that four of them had no insertion and one was heterozygous (Table 6.1). Therefore, I grew the heterozygous line again in order to genotype the next generation of the plants. However, in this case I also did not identify any homozygous mutants in PCR genotyping screens. I then examined the dry seeds obtained from WT Col-0 and the heterozygous line under the microscope. The seeds from WT Col-0 had a proper egg-shape with comb-like rise at the side and not many aborted seeds. In contrast a large number of seeds obtained from the heterozygous line were aborted. This observation suggests that the insertion in AT5G03430 might alter the development of plant embryos.

Table 6.1 ***T-DNA lines with insertions in AT5G03430***

Line name	NASC line number	Localisation according to T-DNA Express	Comment
PAPS (Line 34)	SALK_141801C	promoter	homozygous
PAPS1_1	SALK_031937 (AE)	beginning of the gene	heterozygous
PAPS1_2	SALK_043060	beginning of the gene	didn't germinate
PAPS1_3	SALK_128101	beginning of the gene	no insertion
PAPS1_4	SALK_140720	end of the gene	no insertion
PAPS1_5	SALK_144629C	promoter	no insertion
PAPS1_6	SALK_140626 (BY)	end of the gene	no insertion

Line names correspond to working names of the lines

In arabidopsis the phenotype of the immature seeds changes from white to green, and then to brown before seed release (Andriotis et al. 2010). In normal conditions the immature seeds within the silique develop at the same rate. Very often seeds containing defective

embryos may appear white or brown and shrivelled when seeds containing normal embryos have turned green (Andriotis et al. 2010, Patton et al. 1998, Sparkes et al. 2003, Yu et al. 2004). Therefore, to examine whether disruption in AT5G03430 could affect the development of plant embryos, I performed a time-course experiment to investigate the development of the embryos from heterozygous plants. I had advice and assistance from Vasilios Andriotis (JIC), who is experienced in work with embryo lethal mutants. All seeds from every second silique starting from the top of the stem which had 20 siliques at the time of the analysis were investigated (Figure 6.4). In the first silique (the youngest) all seeds were at the early globular stage. They all looked identical and they were properly developed according to the literature (Figure 6.4A; Andriotis et al. 2010, Patton et al. 1998, Sparkes et al. 2003). Similarly, in seeds from the third silique most of the embryos were at the succeeding, transition stage with properly developed suspensors and seed tissues (Figure 6.4B). In the fifth silique some embryos were at the transition stage, whereas some others were at the early heart stage (Figure 6.4C). No abnormalities were observed in the development of these embryos. Likewise, in silique seven the embryos were at the early or late heart stage (Figure 6.4D). However, the morphology of some embryos in silique nine differed between seeds from this silique. Most of the embryos were at the early torpedo stage (Figure 6.4E), but some were significantly lagging behind. They were at the early or late heart stage with short suspensors containing abnormally enlarged cells (Figure 6.4F).

Silique 11 was the first where, among green seeds, a number of white seeds was detected. The embryos from green seeds from silique 11 were at the late torpedo stage and did not show signs of abnormal development (Figure 6.4G). However, the embryos from white seeds were at the heart stage with short suspensor containing abnormally enlarged cells (Figure 6.4H). Similarly, in silique 13 the embryos from green seeds were maturing (Figure 6.4) whereas the embryos from white seeds were still arrested at the heart stage (Figure 6.4J). In silique 15 apart from the white seeds, pale-yellow seeds were detected. Embryos from these seeds developed slightly further than the embryos from the white seeds. They were arrested at the early torpedo stage but their shape was severely altered (Figure 6.4L). The cells of these embryos were abnormally large and they developed some kind of tumours suggesting disorganised cell division. The cotyledons in these embryos were growing apart (Figure 6.4L). It is likely that these embryos would abort eventually, in the same way as the embryos from white seeds. Indeed, in silique 17 and 19 where the embryos from green seeds were fully developed (Figure 6.4N, R) the embryos from white and pale seeds were arrested at the torpedo stage and presented abnormal

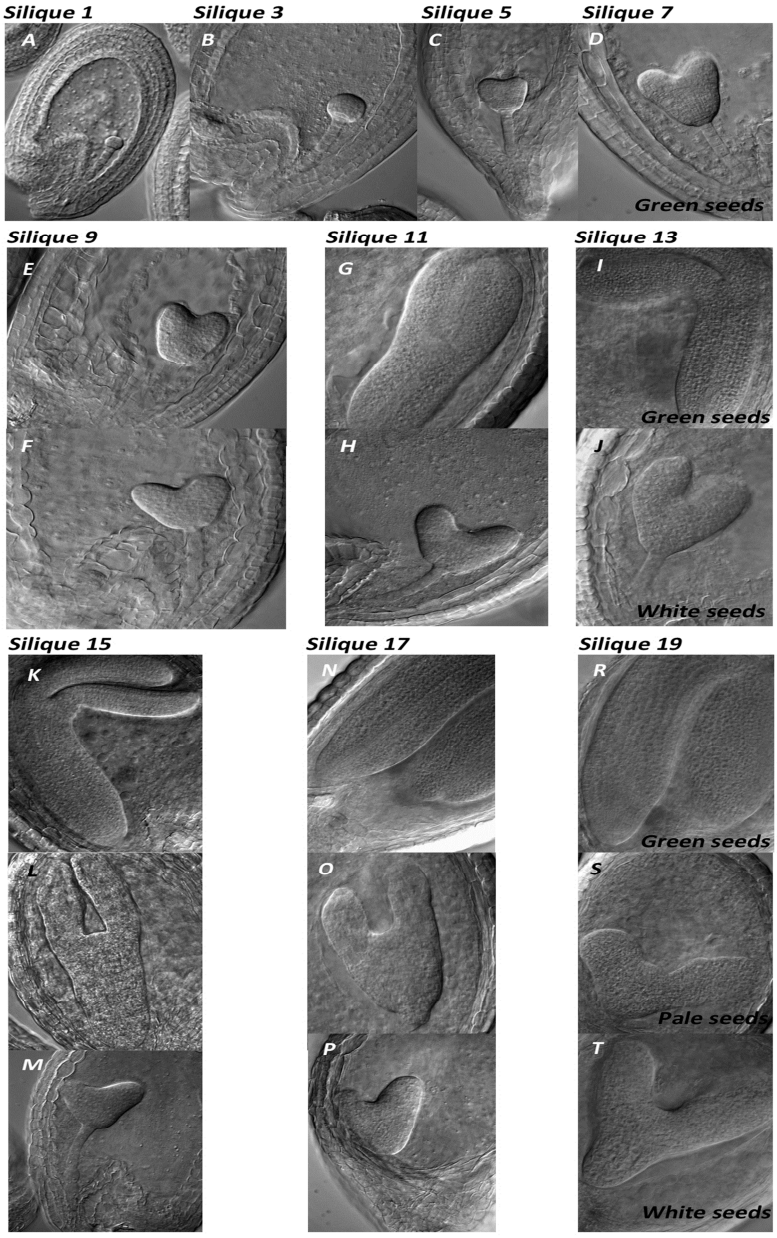


Figure 6.4 ***Developmental arrest of embryos with disruption in AT5G03430 gene***

The siliques of a heterozygous T-DNA line with insertion in AT5G03430 were examined under DIC optics. Particular pictures correspond to representative embryos at different developmental stages: (A) early globular, (B) globular, (C) early heart, (D) heart, (E) early torpedo, (F) globular stage from the same silique as E, (G) late torpedo stage from green seed, (H) heart stage from white seed in the same silique as G, (I) bent torpedo from green seed, (J) heart stage from white seed in the same silique as I, (K) mid torpedo, (L) misshaped torpedo stage from a pale seed in the same silique as K, (M) heart stage from white seed in the same silique as K, (N) mature embryo, (O) misshaped torpedo in pale seed in the same silique as N, (P) heart stage from white seed in the same silique as N, (R) post-mature embryo, (S) late heart stage from pale seed in the same silique as R, (T) early torpedo from the white seed in the same silique as R.

morphology (Figure 6.4O, P, S, T). These seeds would most certainly fail to germinate at maturity. Silique 19 was the last analysed and the oldest from this stem. The results described here are for a representative plant. The phenotypes and the time points, at which the

significant differences in the embryo development were observed, were common for three independent plants analysed.

To quantify the percentage of abnormal seeds I counted the numbers of green, white, pale, and dry (aborted) seeds from up to six siliques (about 150 seeds per plant) from three independent plants (Figure 6.5). The percent of abnormal seeds in each plant varied between 20% and 35% with a mean close to 25% when dry and white seeds were treated as abnormal (excluding pale seeds). The percentage was higher when dry, white, and pale seeds were treated as abnormal (Figure 6.5). Nevertheless, these quantification data, together with the microscopic analysis of the embryos indicate that the homozygous mutant is embryo-lethal and that the mutation is recessive.

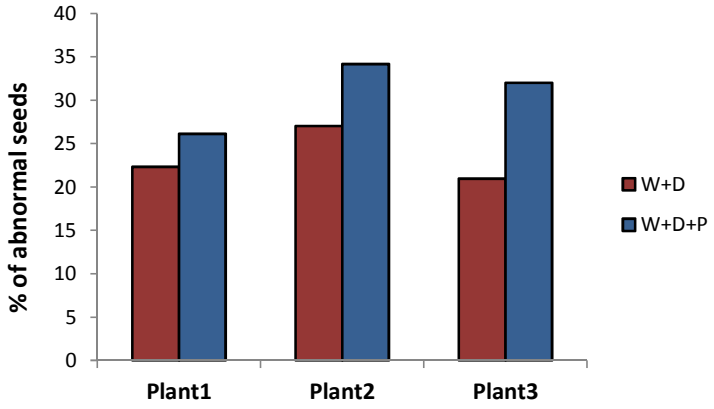


Figure 6.5 ***Quantification of abnormal*** **see*ds***

The bars indicate the percentage of abnormal seeds calculated as average from up to six siliques from a plant. The red bars correspond to percentage of white and dry seeds (W+D) among all the seeds in the silique. The blue bars correspond to white, dry, and pale seeds (W+D+P) among all the seeds in the silique.

To verify whether it is the mutation in AT5G03430 that results in the embryo-lethality I complemented the heterozygous line with the wild-type cDNA of AT5G03430 from Col-0 under 35S promoter. The arabidopsis heterozygous line with T-DNA insertion in AT5G03430 (SALK_031937) was transformed by dipping the flowering plants in a suspension of *Agrobacterium tumefaciens* transformed with appropriate gene constructs (see methods section specific for this Chapter). The transformants were selected on hygromycin B. The transgenic lines were not analysed in time before the thesis submission.

It is important to consider how it was possible to obtain the homozygous insertion line in this gene (line 34 described in Chapter 5), if the mutation is embryo-lethal. According to T-

DNA Express search tool (Alonso et al. 2003) this T-DNA line has the insertion in the promoter and the gene sequence itself should not be disrupted. One possible explanation is that the insertion in the homozygous line did not knock out the gene expression completely. Therefore, I analysed the AT5G03430 transcript abundance in the homozygous line. The investigation of the electronic fluorescent pictographic tool (eFP browser; Winter et al. 2007) revealed that AT5G03430 is expressed continuously at all developmental stages (Figure S6.1). In young seedlings it is expressed in both shoots and roots, whereas in mature plants at the stage of flowering it is expressed mainly in entire rosette and in cauline leaves. It is not expressed in the mature pollen. Therefore, for the analysis of AT5G03430 transcript abundance in the homozygous T-DNA line I first used whole three week old plants (Figure 6.6A). I measured the gene transcript abundance by qPCR using the gene specific primers (see methods section specific for this chapter). I used the arabidopsis *S18* gene as a standard for all measurements described in this chapter. The analysis of three week old plants did not show any difference in AT5G03430 transcript accumulation between the insertion line and the wild type Col-0. Subsequently, I measured the transcript accumulation in roots and shoots separately (Figure 6.6B). However, similarly to the previous analysis I did not observe any difference in AT5G03430 transcript accumulation in homozygous line compared to WT Col-0. There was higher accumulation of the transcript in shoots than in roots in both the insertion line and WT Col-0 in three week old plants. In the next step, I examined the leaves of mature plants for transcript accumulation to finally resolve that issue (Figure 6.6C). This analysis also did not reveal any significant differences between transcript accumulation in the homozygous line and the WT Col-0. These results indicate that the T-DNA insertion in the promoter did not affect the accumulation of the gene transcript in this line. This observation is surprising. The question arises, what exactly affects the nitrate accumulation in this line if the expression of the gene remains unaltered.

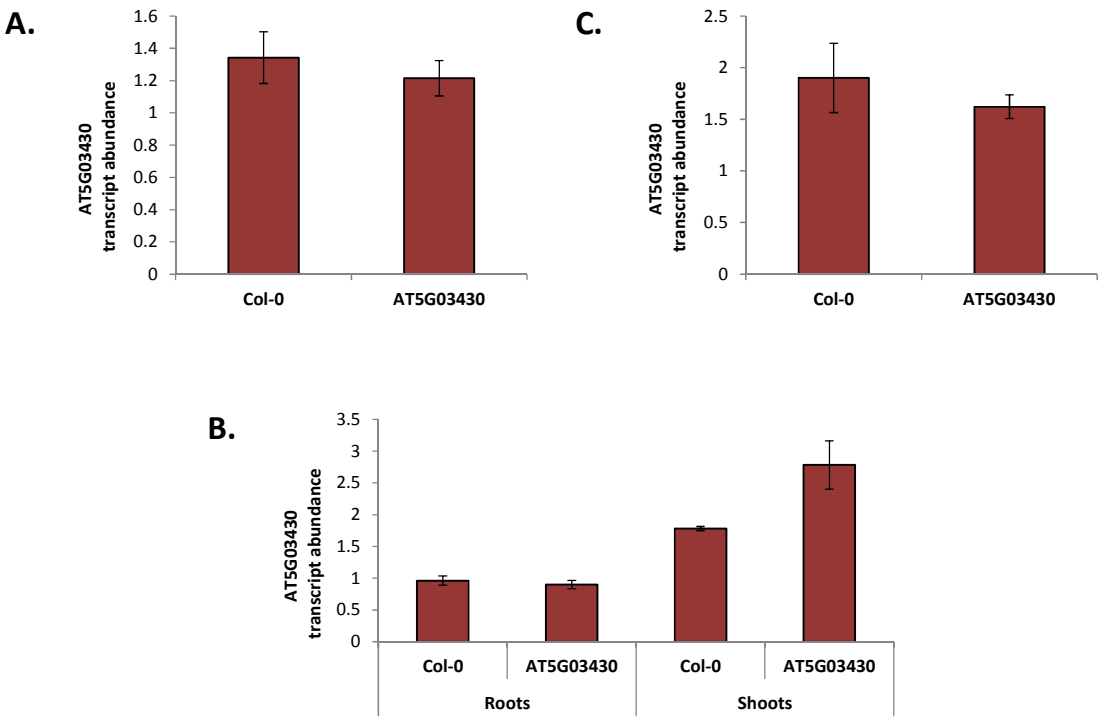


Figure 6.6 ***Transcript abundance of AT5G03430 in the homozygous T-DNA insertion line***

(A) The transcript abundance in whole, three week old seedlings, (B) the transcript abundance in shoots and roots from three week old seedlings, (C) the transcript abundance in leaves of mature plants; the bars indicate average value of three biological replicates measured in triplicate each ± standard error.

One possible explanation of this phenomenon is that the insertion in the promoter might affect adjacent genes. In that case altered expression of AT5G03435 (the closest to the gene of my interest) rather than AT5G03430, might give rise to the changes in nitrate accumulation in the T-DNA line. The AT5G03435 gene is annotated in TAIR as a calcium dependent plant phosphoribosyl transferase family protein, but nothing is known about the function of the gene product. It was not mentioned in the literature before and no data are available in the gene co-expression database ATTEDII or in eFP browser. I analysed the accumulation of its transcript by qPCR with three pairs of different primers as described above using *S18* as a standard. I tested three and five week old plants using shoots and roots separately. The expression of AT5G03435 was undetectable in each of these experiments. These results indicate that AT5G03435 has a very low expression or is expressed only in specific conditions or at a certain developmental stage. Therefore, it seemed unlikely that the changes in its expression are responsible for changes in nitrate accumulation which were consistent in a number of experiments on plants at different developmental stages.

Since the T-DNA insertion in the gene promoter does not alter its transcript accumulation it might be that the low nitrate phenotype in this line is not affected by the insertion itself but some other changes in the gene. One of the possibilities is alternative splicing of the gene. In that case the function of the gene product would be altered, rather than its expression. To see whether there are any differences in the gene transcript between WT Col-0 and the AT5G03430 line I amplified the cDNA of AT5G03430 with gene specific primers. However, there was no difference in the size of PCR products between the insertion line and WT Col-0 in both shoots and roots. Subsequently, I sequenced the genomic DNA of the entire gene isolated from WT Col-0 and the homozygous line to investigate whether there are any other changes in the gene sequence that could result in a low nitrate phenotype. The sequencing of AT5G03430 from the insertion line did not reveal any changes in the sequence compared to WT Col-0 and the genomic sequence of that gene obtained from the TAIR database (data not shown).

An experiment which could provide the explanation for low nitrate phenotype in the homozygous line is the analysis of the regulation of gene expression by different nitrogen sources and nitrate limitation. However, according to Genevestigator there was no difference in expression of this gene in response to nitrate or sulfate deficiency. Therefore, I first examined the other possibilities. Finally, due to time limitation I did not manage to perform these experiments.

A.

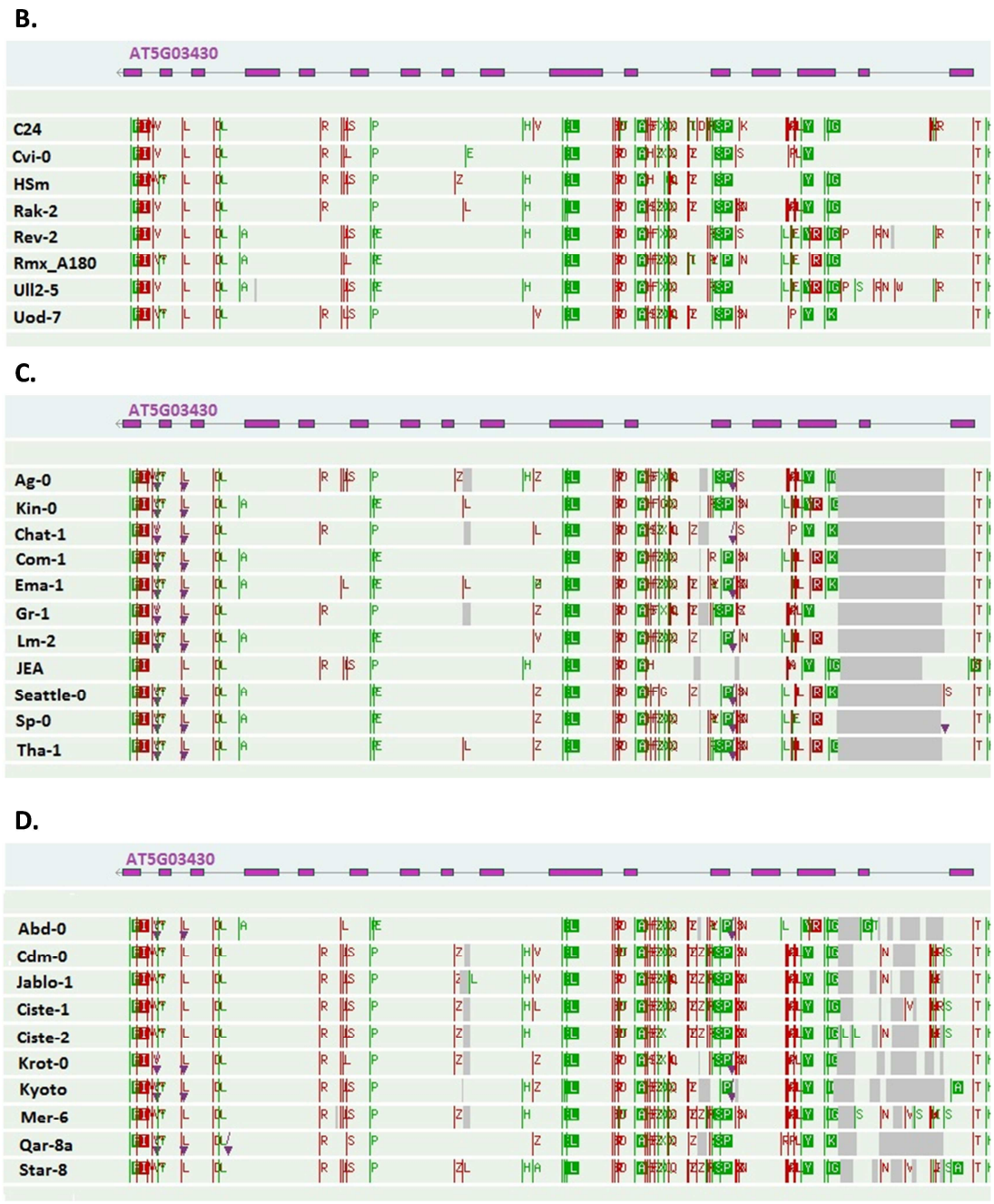


Figure 6.7 ***Schematic representation of different haplotype groups of AT5G03430***

(A) Accessions with no deletion or SNPs in the coding region of the gene, (B) accessions with various SNPs in the gene sequence; (C) accessions with deletion including second exon, (D) accessions with deletion excluding second exon; the deletions are indicated as a grey boxes, codons with synonymous substitutions are marked in green (in coding region in green box), codons with non-synonymous substitutions are marked in red (in coding region in red box), the arrow indicates the left border primer used for sequencing; the right border primer is not indicated on the figure.

6.3.3 The analysis of AT5G03430 haplotypes

To investigate further the involvement of AT5G03430 in the variation of nitrate and/or sulfate accumulation I examined the genetic variation of this gene in arabidopsis using the 1001 Genomes database. This analysis revealed four different haplotypes. In some accessions there were various amino acid changes in the protein sequence compared to Col-0 (Figure 6.7B). There was also a group of accessions with a long deletion including the second exon (Figure 6.7C) and accessions with a deletion excluding the second exon (Figure 6.7D). The group which included Col-0 did not have any of the changes found in other haplotypes and was treated as a control group (Figure 6.7A). Since removing one exon from the gene can have a significant effect on the gene product (e.g. alterations of protein folding or protein function) the difference between the two haplotype groups seemed substantial. Therefore I analysed the available data to investigate the effect of this deletion on anion accumulation.

Such an analysis for AT5G03430 was already described in Chapter 5 where the anion accumulation datasets generated for GWAS were used to investigate the differences in anion accumulation between accessions with different changes in the gene sequence compared to Col-0 (Figure 5.9A, B, and C). In this analysis the three groups of accessions were treated as one and compared to accessions with no changes in the gene sequence. This analysis revealed that the accessions with various changes in AT5G03430 sequence accumulate less nitrate compared to accessions with gene sequence similar to that found in Col-0 (Figure 5.9B). There was no difference in sulfate accumulation between the two groups of accessions (Figure 5.9C).

In the analysis shown in Figure 6.8A and B the same data as above are classified into three different groups (instead of two as was shown in Chapter 5). The first group includes accessions with no changes in the gene sequence compared to Col-0 (as shown in Figure 6.7A). The second group includes accessions with deletion including second exon (as shown in Figure 6.7C). The last group includes the accessions with various predicted amino acid changes and deletion excluding second exon (as shown on Figure 6.7B and D). The data for molybdenum concentration were obtained from a publicly available database of ionomic data iHUB (Baxter et al. 2007; www.ionomicshub.org). This analysis revealed that the accessions with various predicted amino acid changes and deletion excluding second exon have significantly lower nitrate concentration and significantly higher molybdenum concentration compared to accessions with no changes in the gene sequence as found in Col-0. There was no difference in sulfate concentration which is consistent with the results from Chapter 5 (Figure 5.9C).

Surprisingly, these data suggest that the deletion of second exon does not have an effect on the accumulation of anions.

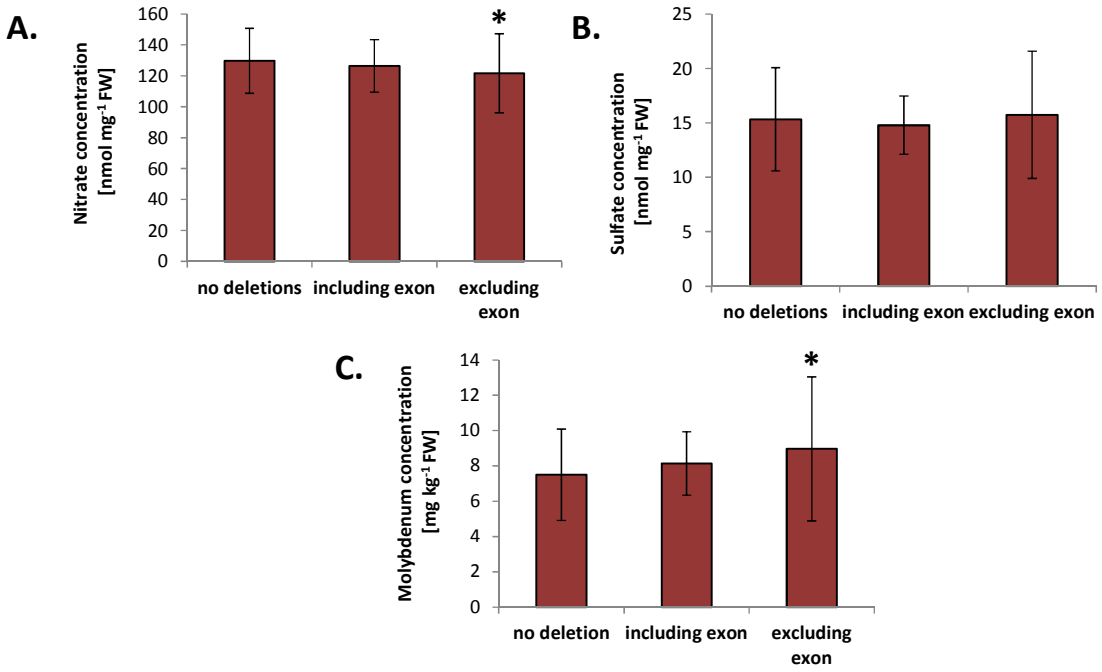


Figure 6.8 ***Analysis of AT5G03430 haplotypes based on available data***

Nitrate (A) and sulfate (B) accumulation data from the datasets generated for GWAS analysis described in Chapter 5. (C) Molybdenum accumulation data were obtained from the publically available database iHUB (www.ionomicshub.org). The bars correspond to average value of anion accumulation for accessions with gene sequence similar to Col-0 (no deletion), accessions with the deletion including second exon (including deletion), and accessions with various amino acid changes and deletions excluding second exon (excluding exon) ± standard deviation. Asterisks indicate significantly different values compared to accessions with no changes in the gene sequence (*t*-test P-value ≤ 0.05).

To investigate further the effect of the variation in AT5G03430 I selected 10-15 accessions representing different groups. I grew these plants for five weeks in soil and analysed them for anion accumulation. This analysis revealed that the nitrate concentration in accessions with the deletion including the second exon was significantly higher (*t*-test P-value ≤ 0.05) compared to accessions with no changes in gene sequence and accessions with the deletion excluding the exon (Figure 6.9A). This is in contrast to the results shown in Figure 6.8 where the accessions with deletion including exon did not differ from the accessions with no changes in the gene sequence as seen in Col-0 and the accession with the deletion excluding exon showed lower nitrate concentration. Moreover, the analysis shown in Figure 6.9B revealed that the accessions with the deletion excluding exon showed lower sulfate

concentration compared to the accessions with no changes in the gene sequence. In the previous analysis (Figure 6.8B) there was no difference in sulfate accumulation between the accessions. Additionally, the analysis shown in Figure 6.9 revealed higher phosphate concentration in accessions with deletion including and excluding second exon compared to accessions with no deletion in the gene sequence as seen in Col-0 (Figure 6.9C). The anion accumulation data for particular accessions are shown in Figure S6.2.

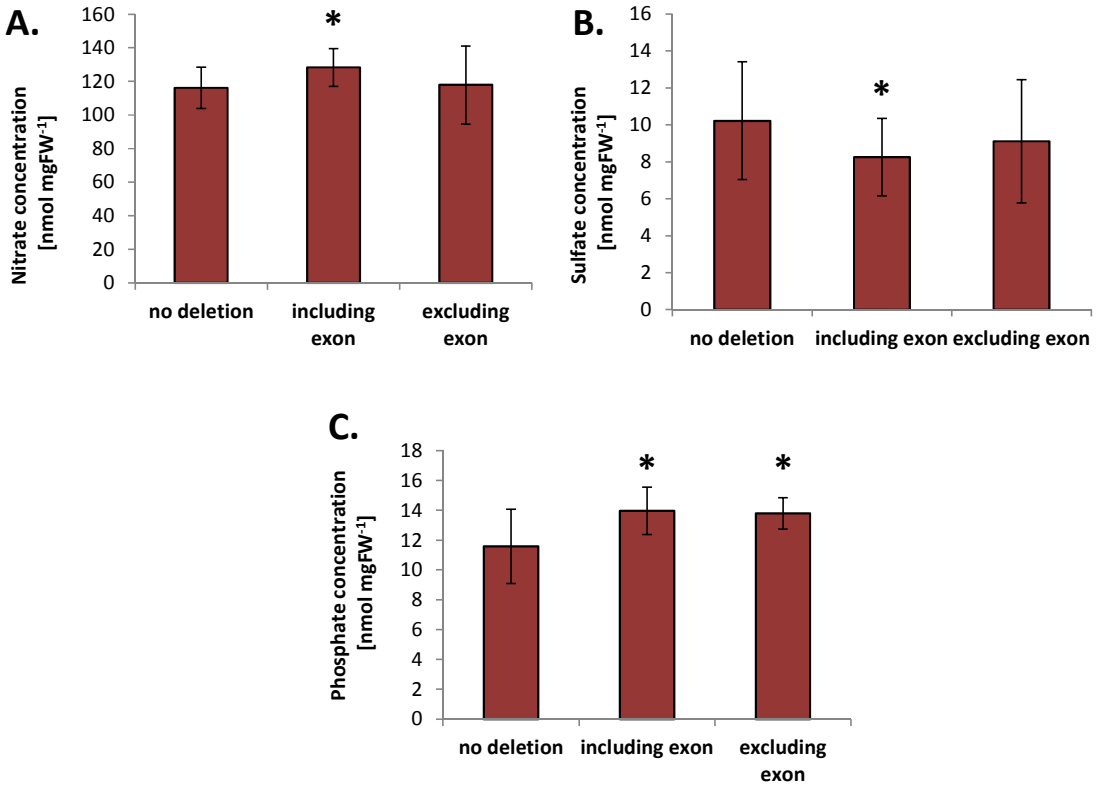
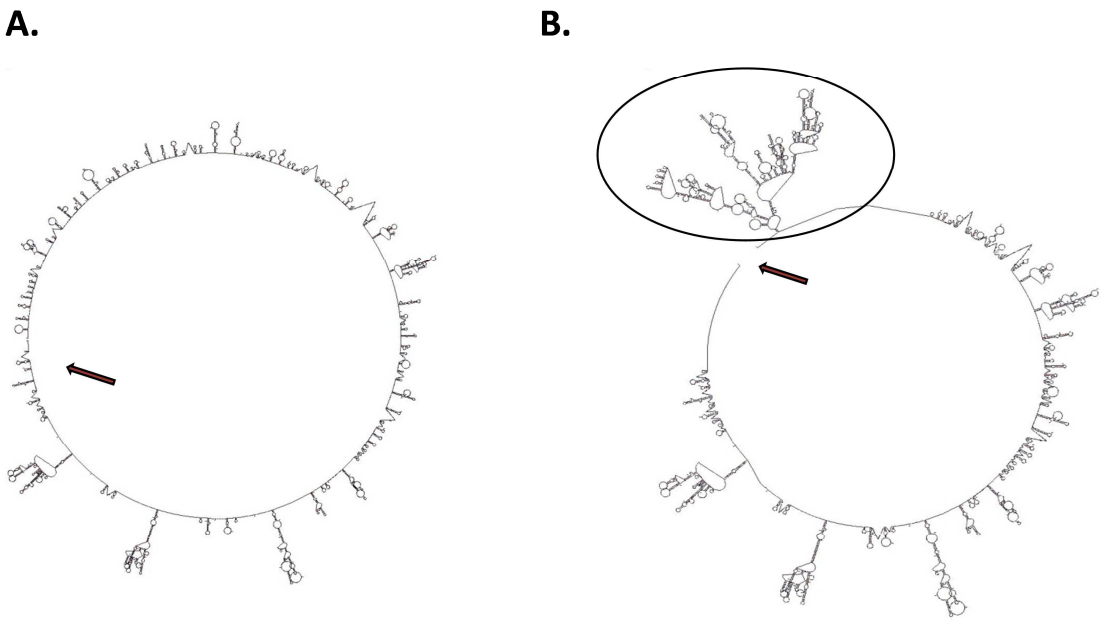


Figure 6.9 ***Analysis of anion concentration in three different AT5G03430 haplotype groups***

The bars correspond to average value of concentration of nitrate (A), sulfate (B), and phosphate (C) in 10-15 accessions selected as representatives of each haplotype group shown on figure 6.5 ± standard deviation. Asterisks indicate values significantly different compared to control group with no deletions in the gene coding region (*t*-test P-value ≤ 0.05).

To investigate this discrepancy of the results more analyses were needed. First, I investigated the presence of the deletion. I isolated the genomic DNA from all the accessions and PCR amplified a 1.2kb fragment of AT5G03430 which was indicated by 1001 Genomes database to include the deletions. Subsequently, I cloned these fragments into the pCR8 vector and sequenced the plasmid DNA. The sequencing showed that there is no deletion in the 1.2 kb fragment in any of the accessions analysed. Moreover, the PCR of the genomic

DNA amplifying entire gene did not show any difference in the PCR product size between accessions. These results indicate not only that the deletion is not in the place indicated by the 1001 Genomes Database, but it is not there at all. One of the possible reasons why this region could be indicated as a deletion in the database is the presence of secondary structure in that region which would disrupt the sequencing. To examine that hypothesis I used the Mfold Web Server (Zuker 2003) that is an online tool to predict nucleic acid secondary structure. First I folded the genomic sequence of AT5G03430 obtained from TAIR (Figure 6.10A, B). The analysis revealed five alternative structures (two are shown in the figure), four of which had a very complicated folding in the region indicated by 1001 Genomes database as a deletion. Subsequently, I folded the 1.2kb fragment of AT5G03430 obtained from sequencing of accessions with no changes in gene sequence as in Col-0 and accessions which were indicated to have a deletion (Figure 6.10C, D). This analysis revealed that the folding of DNA is more complicated in the accessions which were indicated as accessions with deletion in 1001 Genomes database compared to accessions similar to Col-0.



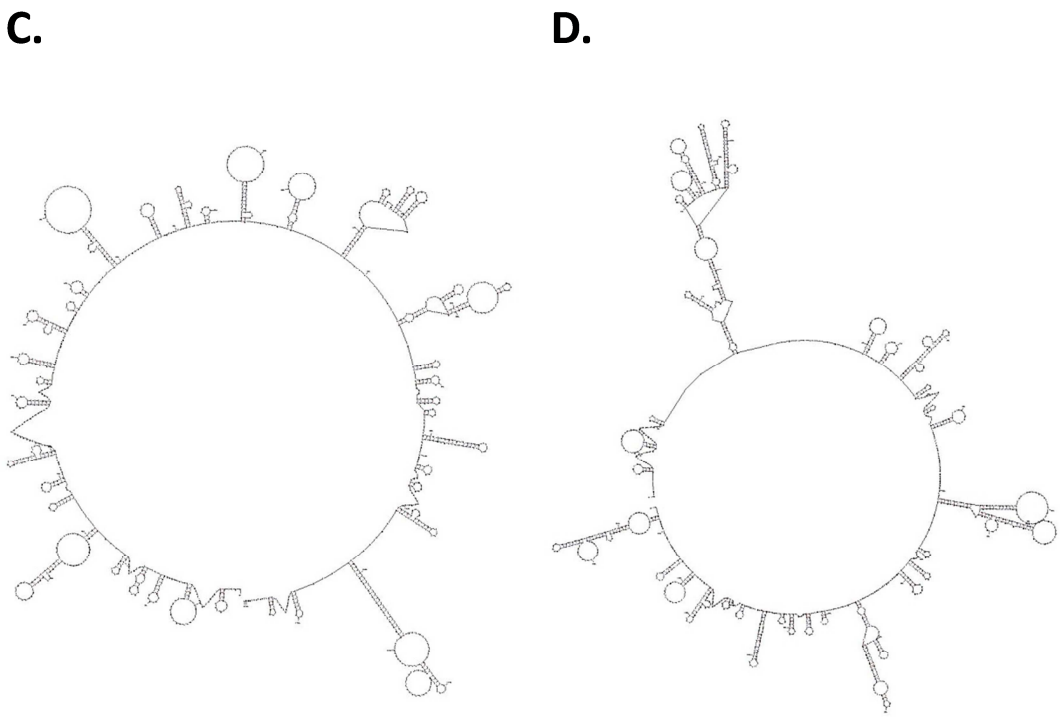


Figure 6.10 ***Secondary structure of AT5G03430 genomic sequence***

(A) Secondary structure with the highest ΔG obtained from the AT5G03430 genomic sequence from TAIR database; (B) Secondary structure with the lowest ΔG obtained from the AT5G03430 genomic sequence from TAIR database; The region indicated as a deletion is circled. The three other structures had similar complicated fold as this one; (C) Secondary structure obtained from 1.2kb fragment of AT5G03430 sequenced from Col-0; (D) Secondary structure obtained from 1.2kb fragment of AT5G03430 sequenced from Com-1 indicated in 1001 Genomes database as deletion; The prediction was made by using the Mfold Web Server. The folding temperature was set to 37°C. The arrow indicates the start codon.

The sequence analysis revealed a lack of deletion, but the differences in anion concentration between the accessions described before were noticable (Figure 6.9). Therefore it was of interest to discover whether there are any other changes in the gene sequence that could be responsible for these differences. The sequenced fragment of the gene included three first exons. Within that fragment I found only one non-synonymous nucleic acid change which led to substitution of proline for arginine at position 86 (third exon). Since only a short fragment of the gene was sequenced it was impossible to verify whether this substitution is responsible for the differences between accessions. The multiple alignment of the fragments obtained from sequencing revealed three clearly distinguished groups, one of which included

only accessions with the Pro/Arg substitution (Figure 6.11B). In the alignment based on the sequences obtained from the 1001 Genomes database the groups of accessions were not that clearly pronounced (Figure 6.11A). This could result in less pronounced differences in anion accumulation between groups of accessions and the discrepancies in the results seen before and described above.

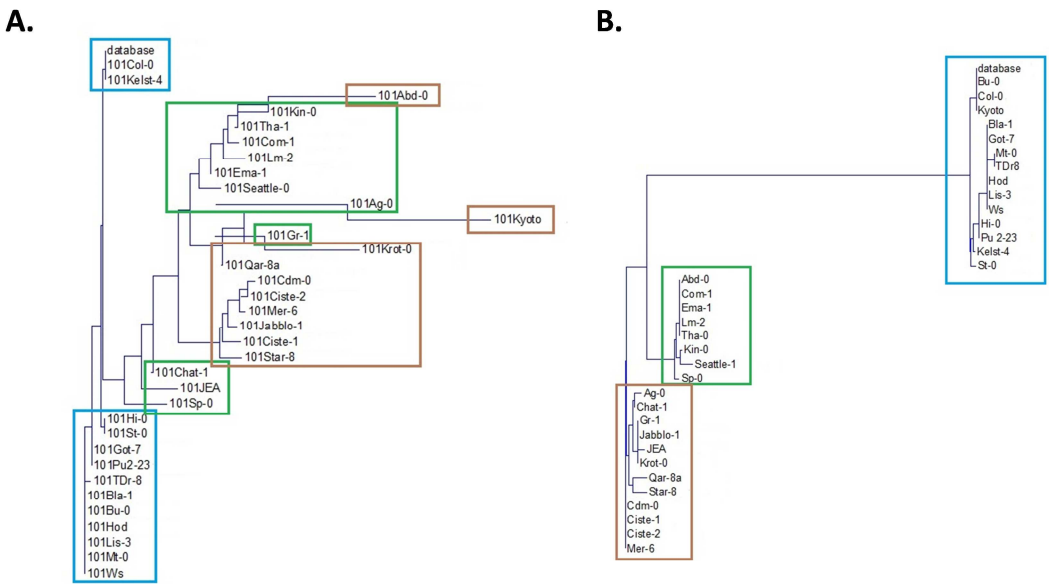


Figure 6.11 ***Changes in haplotype grouping revealed by sequencing***

(A) The guide tree created based on the alignment of 1.2 kb fragment of AT5G03430 sequence obtained from the 1001 Genomes database. According to the database this fragment was expected to have two different types of deletion. (B) The guide tree created based on the alignment of 1.2 kb fragment of AT5G03430 obtained after sequencing. The trees were generated by Vector NTI software as a part of multiple sequence alignment analysis. Different colour boxes indicated accessions belonging to different groups before (A) and after (B) the sequencing.

Following the results of sequence alignment I regrouped the anion accumulation data so that I could compare the effect of the P86R amino acid substitution. Most of the accessions with arginine in position 86 instead of proline were previously selected as accessions with deletion including exon (Figure 6.11). The only exception was Abd-0 which, based on the 1001 Genomes database selection, belonged before to the group of accessions with deletion excluding exon. Moreover, four accessions - Ag-0, Gr-1, Chat-1, and JEA - which before sequencing belonged to the group of accessions with deletion including exon did not have the amino acid substitution. In the multiple alignment obtained based on the sequencing results these accessions were included in the group selected before sequencing as accessions with

deletion excluding exon. The control group remained unchanged with the exception of Kyoto which before sequencing was selected as accession with deletion excluding exon (Figure 6.11).

After regrouping of the accessions the results of anion accumulation were in agreement with what was seen before (Figure 6.9) but the differences between groups were much more pronounced and had higher significance (Figure 6.12). The nitrate concentration was significantly higher in the accessions with the P86R amino acid substitution compared to both the control group (including Col-0) and the third group distinguished from the sequence alignment (Figure 6.12A). I did not observe any non-synonymous nucleic acid substitutions in the coding region in accessions from this group. However, there were a number of SNPs in introns common for these accessions which most likely resulted in differentiating this group from the accessions with no changes in the gene sequence as seen in Col-0. Moreover, sulfate concentration was significantly lower in accessions with the P86R substitution compared to control group (Figure 6.12B). There was no significant difference between the accessions with no changes in the gene coding region and the third group of accessions. In contrast, phosphate concentration was higher in accessions with P86R substitution as well as in the accessions from the third group (Figure 6.12C). This indicates that there might be additional amino acid substitutions or other changes in this gene or related genes in the accessions from this group which were not included in the sequenced fragment.

At this stage it is impossible to infer the exact reason for the differences in anion accumulation between the accessions. However, the results described here indicate that the P86R amino acid substitution in AT5G03430 might be involved in the natural variation in anion accumulation. Given the discrepancy between the sequence provided in the database and the results of sequencing thus far, it would be necessary to re-sequence the entire gene in these accessions to investigate whether it is the only difference in the gene sequence among these accessions. Moreover, additional experiments such as analysis of transcript accumulation would be necessary to investigate the mechanism of that variation (similar to the results described in Chapter 3).

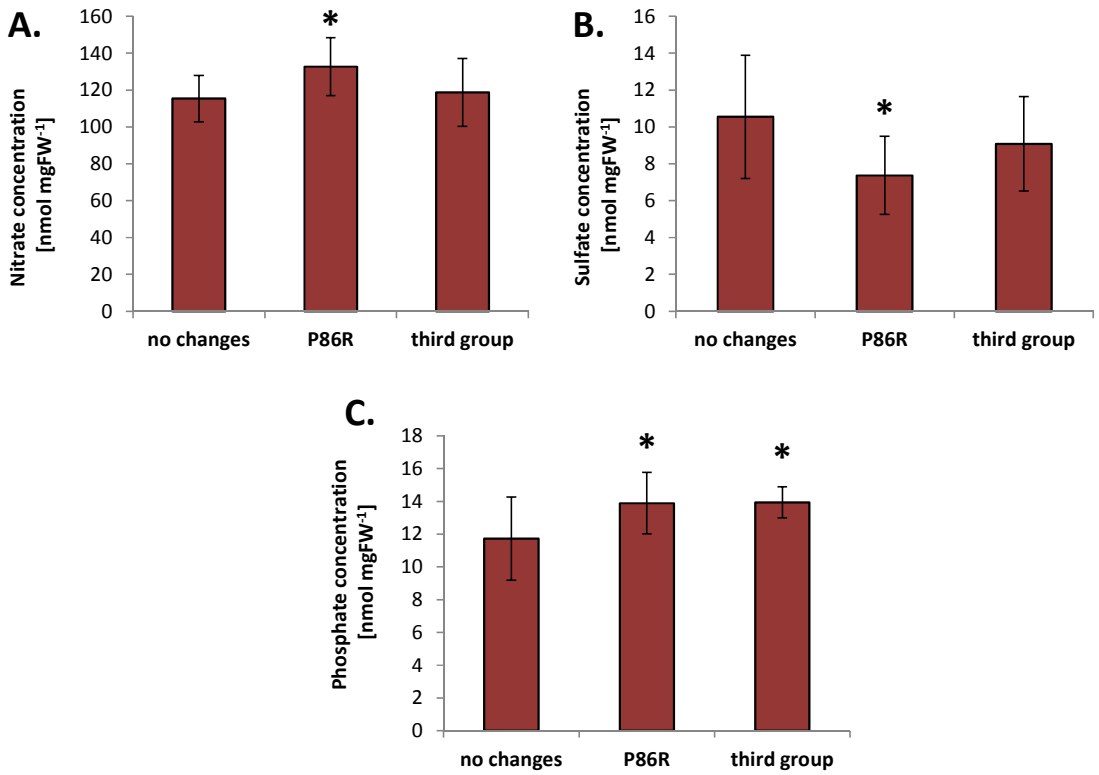


Figure 6.12 ***Anion concentration in different AT5G03430 haplotype groups after***

The bars correspond to average value of concentration of nitrate (A), sulfate (B), and phosphate (C) in 10-15 accessions selected as representatives of each haplotype group shown on figure 7.9B ± standard deviation. Asterisks indicate values significantly different compared to control group with no changes in the gene coding region (*t*-test P-value ≤ 0.05).

6.3.4 Subcellular localisation of the AT5G03430 gene product

The prediction of subcellular localisation of AT5G03430 in the SubCellular Proteomic Database (SUBA3; Tanz et al. 2013) was not clear. According to this database AT5G03430 gene product was predicted to be localised in number of cellular compartments such as cytosol, nucleus, mitochondria, and Golgi. The SLPFA tool also showed an extracellular localisation of this protein. Therefore, to discover the location of the AT5G03430 gene product I fused the coding region of AT5G03430 to the C-terminal GFP vector under the control of a 35S promoter. I then transferred the gene constructs to *Agrobacterium tumefaciens* by electroporation and infiltrated *Nicotiana benthamiana* leaves. I detected the GFP fluorescence by confocal microscopy two days after infiltration. This analysis revealed that the product of AT5G03430 is localised in cytoplasm and nucleus (Figure 6.13A). This is in agreement with the MoCo biosynthesis pathway which, except the first step (the conversion of 5'-GTP into cPMP occurs in mitochondria), occurs in the cytosol.

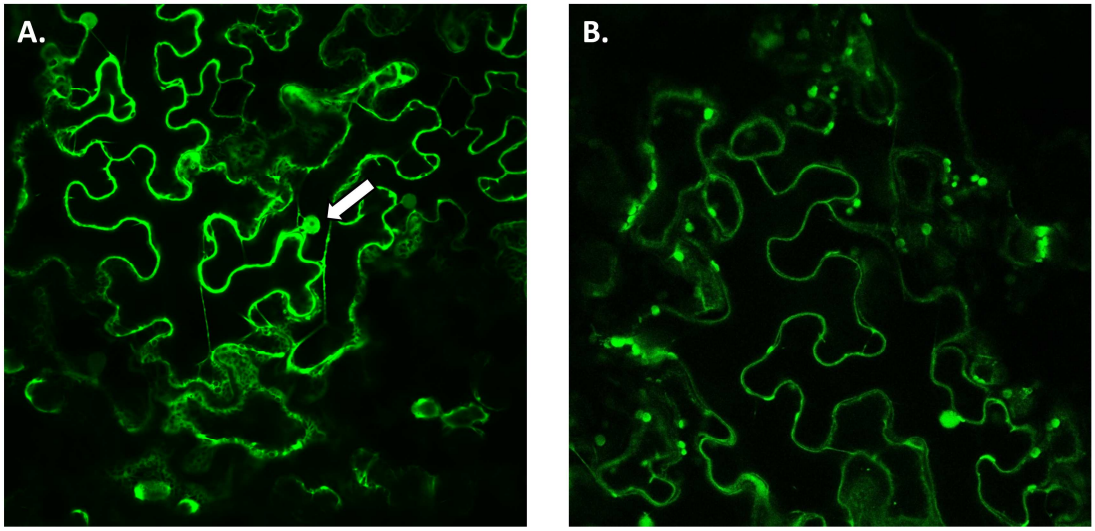


Figure 6.13 ***Subcellular localization of the GFP::AT5G03430 in tabacco leaves***

Confocal microscopy images of leaf epidermis of *Nicotiana benthamiana* infiltrated with *Agrobacterium tumefaciens* transformed with (A) AT5G03430 coding region fused to GFP under the 35S promoter. The image indicates protein location in cytoplasm and nucleus (white arrow); (B) the expression of GFP under the control of 35S promoter as a positive control.

6.4 Discussion

The gene described in this chapter was detected by GWAS conducted on nitrate and sulfate concentration data. Based on the analysis of the protein sequence I hypothesised that it might be involved in the molybdenum cofactor biosynthesis pathway. Since the metabolic pathways of molybdenum, nitrate and sulfate are closely linked it is entirely possible that this novel gene could affect the accumulation of nitrate and sulfate. Two of the molybdenum enzymes known in eukaryotes catalyse basic reactions in the metabolism of nitrogen and sulfur compounds (Hille 1996, Mendel & Hansch 2002). Moreover, bacterial molybdenum nitrogenase links molybdenum with nitrate metabolism in bacteria (Seefeldt et al. 2009). The crucial biochemical role of molybdenum in nitrogen fixation and nitrate reduction is long and well recognised in plants (Nicholas & Nason 1955, Spencer & Wood 1954, Yang et al. 2011). Nitrate reductase catalyses the first step in the nitrate assimilation pathway and is a key regulatory step in nitrate metabolism (Masclaux-Daubresse et al. 2010). Nitrate assimilation is regulated by a complex network responding to a number of internal and external signals e.g. light, phytohormones, nitrate, and carbon (Crawford & Forde 2002). The complete loss of enzyme activity e.g. in NR mutants or during molybdenum deficiency is lethal confirming further the importance of this enzyme for plant fitness.

Sulfur is embedded in molybdenum metabolism at several steps. First, the uptake system of molybdate and sulfate are closely related. The first information about the similarity between molybdenum and sulfate transport system was delivered by Stout et al. (1951) who showed that sulfate is a potential inhibitor of molybdenum uptake. This finding was later confirmed in research on rice (Kannan & Ramani 1978). Recently it was also shown that molybdate uptake increases under sulfate deficiency (Shinmachi et al. 2010). Since variation in sulfate availability may affect molybdenum accumulation and distribution within the plant this relationship has a practical significance in crop production (Balik et al. 2006, Macleod et al. 1997). Extensive use of sulfate fertilizers can result in molybdenum deficiency in soil (Macleod et al. 1997). On the other hand, high sulfate concentration can interfere with the capacity of the plant to remove inorganic forms of molybdenum during phytoremediation (Schiavon et al. 2012).

Molybdenum and sulfate share a high degree of structure similarity: they both have a double negative charge and tetrahedral structure. They are also similar in size (Bittner 2014). Therefore, it is not surprising that the two molecules are taken up via the same transport system. Indeed, the first molybdenum transporter identified in arabidopsis MOT1 was

previously classified as a sulfate transporter SULTR5.2 (Baxter et al. 2008, Tomatsu et al. 2007). This class of sulfate transporters lacks the STAS domain commonly found in the sulfate transporters (see Chapter 1). However, the analysis of the *mot1* mutant did not show any changes in the shoot sulfate concentration (Baxter et al. 2008). The two reports describing MOT1 as a molybdate transporter provided different subcellular localisation of the protein: either endomembrane (Tomatsu et al. 2007) or in mitochondria (Baxter et al. 2008). Therefore the exact function of MOT1 in molybdate transport remains to be investigated. Molybdenum was also shown to affect not only sulfate uptake but also sulfate assimilation. It was shown that molybdate can serve as a substrate for the ATP sulfurylase (Reuveny 1977, Wangeline et al. 2004). However, no stable products are formed in this reaction (Reuveny 1977). The overexpression of ATPS in *Brassica juncea* resulted in lower tolerance of molybdenum toxicity and higher accumulation of molybdenum (Wangeline et al. 2004). It was suggested that this is due to increased loss of ATP which was bound by ATPS to molybdenum yielding an unstable product (Wangeline et al. 2004). It was shown recently that plants treated with high concentration of molybdate accumulate less cysteine and glutathione in sulfate sufficient plants (Schiavon et al. 2012). The concentration of thiols in these plants was similar to that observed in plants grown under sulfate limitation (Schiavon et al. 2012). It was suggested that the decreased thiol synthesis is an effect of competition between sulfate and molybdate for access to the sulfate reduction pathway as was observed before (Reuveny 1977, Wangeline et al. 2004).

As mentioned before, sulfite oxidase, the molybdenum enzyme that oxidises sulfite to sulfate, is another link between sulfate and molybdenum metabolism (see introduction to this Chapter). In arabidopsis, sulfite oxidase together with APR is involved in the co-regulation of the sulfate assimilation pathway (Randewig et al. 2012). In contrast to nitrate reductase, the loss of sulfite oxidase is not lethal unless the plant is exposed to high sulfur dioxide in the atmosphere (Lang et al. 2007). The availability of sulfate also has an important role during MoCo biosynthesis since the sulfur atom is necessary to generate MPT, and molybdenum enzymes of xanthine oxidase family require the addition of a terminal inorganic sulfur to acquire enzymatic activity (Wahl et al. 1984). Therefore it is likely that fluctuations in sulfate accumulation could affect molybdenum metabolism and *vice versa*.

The high sequence identity between the cinA domain of AT5G03430 and the proteins involved in molybdenum cofactor biosynthesis in *E. coli* (MoeA and MoaB), arabidopsis (Cnx1G) and human (gephyrin) is another clue suggesting that this novel gene might be involved in MoCo metabolism. Some of the proteins in MoCo biosynthesis in different species

are known to be multifunctional proteins. The Cnx5 in *Nicotiana plumbaginifolia* was shown to be a two domain protein. One of the domains is homologous to *E. coli* MoeB and the other to Ntdin, a protein for which the exact involvement in MoCo biosynthesis in tobacco is still not clarified (Yang et al. 2003). However, it was shown that transgenic tobacco mutants with reduced expression of *Ntdin* showed low activities of NR and xanthine dehydrogenase and very low MoCo concentration (Yang et al. 2003). Similarly, Cnx1 and human gephyrin consist of two domains which are homologous to *E. coli* MogA and MoeA proteins (Stallmeyer et al. 1999, Stallmeyer et al. 1995). In plants the E-domain is localised at the N-terminus of the protein and the G-domain is localised at the C-terminus. Mammals and fungi have the G-domain at the N terminus (Mendel & Schwarz 2011). This indicates that the fusion of the two bacterial proteins occurred at least twice during evolution and points out the functional benefit of having the two enzymatic functions coupled into one protein (Belaidi & Schwarz 2013).

Given that human FAD synthases have the same two domains as found in AT5G03430, but in reverse order, it is possible to speculate that a similar event in evolution would lead to the two-domain structure of AT5G03430 gene product. FAD synthase acts in the second step of FAD synthesis by adenylation of FMN yielding FAD (Barile et al. 2013). Some of the proteins involved in MoCo biosynthesis which show high sequence similarity to cinA domain of AT5G03430 have similar function (Mendel & Kruse 2012). They are involved in the adenylation of the MPT to activate it before the insertion of molybdenum atom (Mendel & Kruse 2012). This could suggest that AT5G03430 may have a similar function to human FAD synthases. However, it remains to be investigated if and how the human FAD could adenylate the molybdopterin and what is the relationship between the two domains in the FAD synthase protein (Miccolis et al. 2012). Therefore, the investigation of the function of AT5G03430 gene product could lead to the better understanding of the function of human FAD synthases.

The crucial function of AT5G03430 for plant fitness was confirmed by the embryo-lethality of homozygous mutant. The embryo-lethal phenotype was observed in three independent heterozygous plants. Since I could not obtain homozygous lines from several generations and the embryos from about 25% of seeds from each plant were arrested it is likely that they were the homozygous mutants. The embryo development in these plants was arrested at the early heart stage. The embryos from the pale seeds developed further, however their shape was abnormal with the suspensors and cotyledons enlarged by disorganised cell division and expansion. The embryo arrest at the early developmental stage resulted in embryo death and seed abortion. It is not possible to evaluate the role of the AT5G03430 gene product in embryo development at this stage when the function of the

protein is not known. However, following the hypothesis that it might be involved in the MoCo biosynthesis, this phenotype would not be surprising. In plants, loss of any of the proteins involved in MoCo biosynthesis results in the loss of the enzymatic function of all MoCo enzymes (Mendel 2011). Mutants in MoCo biosynthetic genes need to be supplemented with ammonium since they don't have NR activity which is crucial for nitrate assimilation (Mendel & Kruse 2012). In humans, MoCo deficiency results in severe neurological disorders in new-born babies and early childhood death (Mendel & Kruse 2012). However, it can be treated by repetitive injection of cPMP overproduced and purified from *E. coli* (Schwarz et al. 2004).

The understanding of molybdenum metabolism and the biosynthesis of MoCo has progressed rapidly over past years. The molybdo-enzymes have been characterised and the enzymes of MoCo synthesis have been cloned. However, many aspects still remain poorly understood. The currently open questions concern mainly molybdate transporters including the import and export routes in plants as well as its uptake from the soil. A family of MoCo binding proteins was recently discovered in arabidopsis, however their exact function is still not clear (Kruse et al. 2010). Moreover, no proteins that could store molybdate or transport it between organelles and tissues are known. Very little is also known about the control of the gene expression and interaction of proteins involved in molybdenum metabolism. Interestingly, the crystallisation of Cnx1G revealed the existence of copper bound to the MPT dithiolate sulfur (Kuper et al. 2004). However, very little is known about the interaction between molybdenum and copper metabolism and the mechanism of copper insertion into MPT. Moreover, the proteins that donate copper to MPT or that accept it after its release during MoCo formation have not been identified to date.

6.5 Future work

The analysis described in this chapter is still at an early stage and not many conclusions can be drawn. In order to verify the hypothesis of the involvement of AT5G03430 in MoCo biosynthesis and its possible function in that process, further analyses have been planned. In collaboration with the group of Ralf Mendel from the Technical University of Braunschweig, who is a specialist in molybdenum metabolism, I prepared appropriate gene constructs to over express and purify the AT5G03430 gene product. The recombinant protein will be used to quantify the MoCo, MPT and molybdenum via HPLC. This information will show whether there is any connection of AT5G03430 gene product with molybdenum metabolism. Since loss of function of this gene is embryo lethal the characterisation of its function by a classic method including detailed analysis of loss-of-function alleles is problematic. However, there are other methods which can be used to uncover the function of embryo-lethal genes. Approaches such as rescue of mutant embryos using special culture conditions or partial complementation of mutant phenotypes during embryogenesis have been successfully used to dissect the essential gene functions (Baus et al. 1986, Candela et al. 2011, Despres et al. 2001). Direct gene silencing in post-embryonic tissues has also been very effective in uncovering gene function (Schwab et al. 2006, Schwab et al. 2010).

Additionally, the analysis of complemented heterozygous lines is necessary to confirm that it is disruption in the AT5G03430 gene which causes early stage embryo arrest in the homozygous lines. The analysis of Col-0 lines overexpressing AT5G03430 will be helpful to investigate further the low nitrate accumulation phenotype in the homozygous T-DNA insertion line which did not show changes in gene transcript abundance. It can also be helpful in dissecting the function of the gene product. I have already obtained these transgenic lines, but due to time limitation I was not able to analyse them before the thesis submission deadline.

In order to examine the effect of natural variation in AT5G03430 on the accumulation of nitrate, phosphate, sulfate and possibly molybdenum more detailed analysis of haplotypes would be necessary. Since the sequencing of the gene fragment revealed some discrepancies between the database and re-sequencing results, the entire gene would need to be re-sequenced in all the accessions used for such an analysis. Furthermore, the transcript abundance in the accessions with different alleles of AT5G03430 should be measured to investigate the exact effect of the P86R and possible other substitutions revealed by the gene sequencing. The analysis of metabolite accumulation and enzymatic activity of key enzymes

from sulfate and nitrate metabolic pathways would aid discovery of the nature of the effect of these substitutions. Directed mutagenesis would be useful to confirm whether the observed variation in measured traits is caused by the variation in AT5G03430. Due to the limited length of my PhD project I was not able to perform any of these analyses. However, they are necessary to characterise the exact function of AT5G03430 gene product.

Chapter 7:

General Discussion

7.1 Summary

The work described in this thesis concerned the investigation of genetic variation in arabidopsis response to sulfate, nitrate, and phosphate availability. The main aims of this work were to investigate the natural variation in plant response to low nutrient availability as well as identify and functionally characterise genes involved in the control of nitrate, phosphate, and sulfate metabolism using forward genetic screens and biochemical and molecular biology techniques.

The investigation of natural variation in the gene architecture of APR2 and ATPS1 described in Chapter 3 led to remarkable results. First, it revealed that two consecutive enzymes of sulfate assimilation pathway are nearly equally involved in its regulation (Chao et al. 2014, Koprivova et al. 2013). This is in contrast with previous results suggesting almost exclusive control of APR2 over the pathway (Vauclare et al. 2002). Detailed analysis of the natural variation in the gene sequence of these two enzymes revealed different mechanisms by which they control the pathway (Chao et al. 2014, Koprivova et al. 2013). Moreover, it led to the identification of individuals with extremely rare alleles and a large phenotypic effect which would not have been identified using the classic mutant approaches (Chao et al. 2014). Further characterisation of these individuals can lead to better understanding of the regulation of the entire pathway. Additionally, further characterisation of identified individuals may provide new insights in the relation between enzyme structure and its function (Herrmann et al. 2014). In general, this analysis complemented current understanding of the regulation of sulfate metabolism, set the background for further analyses and highlighted the benefits of the analysis of natural variation.

Subsequently, natural variation was used to characterise a general response profile of arabidopsis to low sulfate supply and characterise specific patterns of response among accessions, followed by identification of accessions which were well adapted to the nutrition regimes tested in this experiment (see Chapter 4). Additionally, the natural variation in a worldwide population of arabidopsis was used to identify novel genes involved in the control of nutrient homeostasis using GWAS (see Chapter 5). The analysis described in Chapter 3 revealed that sulfate accumulation is a complex trait controlled by a number of rare loci with strong phenotypic effects (see Chapter 3). Since such traits present difficulties for GWAS (Nordborg & Weigel 2008) this might be the reason for the low significance of the results. However, a follow up analysis of insertion mutants in candidate genes delivered by GWAS led to identification of genes potentially involved in the control of nutrient accumulation: CCR-

4NOT associated factor 1 (CAF1), phosphate/sugar transporter, and putative PAPS reductase. The analysis of sequences of these candidate genes revealed single nucleotide polymorphisms which might be associated with variation in anion accumulation (see Chapter 5, Figure 5.9). Further characterisation of individuals with different variations in each of these genes might be helpful in dissection of their exact functions in control of anion homeostasis. Additionally, these individuals can also be used for creation of specific populations with positive variations in the traits of interest (Baxter & Dilkes 2012).

Further investigation of a candidate gene delivered by GWAS and partially described in Chapter 6 – AT5G03430 – may lead to an improvement of our knowledge about molybdenum metabolism. Given the high similarity of the novel gene delivered by GWAS to genes found in humans (based on the computational analysis of the sequence) the analysis of plant homologue may result in an improvement of knowledge of an important human protein (FAD synthase) the exact function of which has not yet been resolved. Here again, the analysis of natural variation in gene sequence revealed specific groups of individuals with different phenotypes which may be used in future investigation of function of the gene product as was described in Chapter 3.

7.2 Practical application of the results

My PhD project was part of Marie Skłodowska-Curie International Training Network, Bionut ITN (http://bionutitneu2.fatcow.com/). The main objective of this network was to provide state-of-the-art training for young researchers through international research programmes, workshops, and transfer of knowledge. The scientific focus of the network partners was based around biochemical and genetic dissection of the control of plant mineral nutrition. Given the constantly growing global population and the increasing pressure on land resources, the food security became a major challenge of our society (Mba et al. 2012). It has been projected that the human population will exceed nine billion by the year 2050 (Mba et al. 2012). Therefore, food production must increase around 70% over next decades to nourish adequately the human population (Tomlinson 2013). Moreover, food quality, particularly nutrient concentration, needs to be improved and the agricultural input needs to be reduced (Tester & Langridge 2010). Achieving these goals is even more challenging in the face of global environmental changes (Beddington et al. 2011). Therefore, it requires fundamental modifications in breeding strategies and crop improvement (Moose & Mumm 2008). Thus, the major scientific questions of Bionut ITN concerned the control of nutrient use efficiency in plants and its practical application in generation of crop plants with optimized yields without

increasing inputs. The implementation of the main objectives of the network fell into three main areas: investigation of genetic control of nutrient use efficiency, functional analysis of new genes affecting plant mineral nutrition, and evaluation of candidate genes in crop plants.

In order to develop new crop varieties ensuring high yield at low inputs in the changing environment a better understanding of plant response to environmental fluctuations is necessary. This was achieved by the analyses described in Chapter 4. First of all, this analysis revealed the traits that can be targeted in the process of development of new crop varieties. The general response profile of 25 arabidopsis accessions as well as the detailed characterisation of four different classes of accessions representing different patterns of plant response to sulfate availability highlighted the importance of increased root development under limited sulfate. In that case the possible targets for genetic engineering could be focused on improving resource capture from the soil (Hawkesford 2000). It could be achieved by modulation of transport system or root structure and proliferation. It has been shown previously that the expression of sulfate transporter genes increases significantly under sulfate limitation, but decreases when sulfate is abundant (Takahashi et al. 1997). Overriding this control might be achieved by expressing transporter genes under the control of an appropriate constitutive promoter. However, the control mechanisms would have to be removed only for sulfate transport and not for other steps of the pathway in order to prevent e.g. accumulation of sulfide which is toxic for plants in high concentration (Lamers et al. 2013).

Alternatively, root structure and proliferation could be targeted. The mechanism of increased root proliferation under sulfate deficiency was already described in the discussion section of Chapter 4. Briefly, sulfate limitation results in an induction of *NIT3* gene leading to increased production of auxin which subsequently results in increased root growth and branching (Kutz et al. 2002). Moreover, root growth is also modulated by alterations in auxin homeostasis caused by depletion of GSH (Koprivova et al. 2010). In fact, the results described in Chapter 4 also revealed significant correlation between GSH concentration in roots and root thickness. All these three factors: expression of *NIT3*, GSH concentration in roots, and auxin homeostasis could be manipulated in order to develop crop varieties which will be more resistant to nutrient depletion. Moreover, manipulation of GSH concentration could increase the resistance of plants to biotic and abiotic stress.

An increased biomass allocation towards the roots was a common plant response to sulfate deficiency observed in the analyses described in Chapter 4. Moreover, the primary metabolite profiling revealed accumulation of sugars in all accessions analysed in response to

sulfate limitation and/or starvation. A shift of biomass production towards the greater root biomass is a sensible adaptation to maximise the acquisition of minerals. However, it needs to be well balanced in order to avoid a decrease in shoot biomass (Hawkesford 2000). An increased accumulation of carbohydrate and carbon allocation towards the roots has been observed in nitrate and phosphate deficient plants (Ciereszko et al. 2001, Linkohr et al. 2002, Lopez-Bucio et al. 2002, Paul & Driscoll 1997). Based on the analysis of microarray data obtained from phosphate and nitrate deficient plants a mechanism leading to optimisation of root architecture has been proposed (Hermans et al. 2006). Changes in gene expression under phosphate and nitrate deficiency result in changes in primary metabolism leading to increased transport of sugars to the roots and simultaneous fluctuations in the hormone concentration, which results in modification of the root architecture (Hermans et al. 2006). Given the results described in Chapter 4 it might be hypothesised that a similar mechanism could apply to sulfate deficient plants. However, these results concerned the investigation of changes in primary metabolite composition in shoots only. The analysis of primary metabolites in roots as well as the analysis of gene expressions in sulfur starved plants (perhaps using the available data) would be necessary to answer that question.

Another important feature revealed in the analysis described in this thesis is the importance of the interconnection between elements in nutrient homeostasis. It is long been known that the lack of one element can limit the use efficiency of other elements (Ahmad et al. 2005, Anjum et al. 2012, Rouached 2011). Therefore, the investigation of such interconnections is crucial for developing new breeding strategies. The need of simultaneous analysis of changes in plant mineral composition was recognised a decade ago when the term “ionome” was introduced to describe all metals, metalloids, and non-metallic elements in an organism (Hirschi 2003, Salt 2004). It has been shown that one gene can control the metabolism of more than one element and about 2-4% of all genes are involved in the control of ion uptake (Eide et al. 2005, Lahner et al. 2003). Since then complex studies of the ionome in different plant species and its connection with the genome have been carried out to expand our knowledge (Buescher et al. 2010, Chao et al. 2011, Chen et al. 2009, Parida et al. 2004). Therefore, one of the main goals in the studies described here was to identify and characterise regulatory components that might be involved in the control of concentration of one or more elements in the plant cells. In Chapter 4 an involvement of a mechanism of coordination of long distance transport of minerals in response to nutrient limitation involving miRNA and transcription factors was suggested. A signalling role of GSH, sugars and other molecules was also revealed in these studies. Moreover, genes potentially involved in the coordination of

nutrient homeostasis were revealed in the GWAS described in Chapter 5. Evidence was provided that the natural variation in the architecture of these genes may result in changes in patterns of anion accumulation. These results, together with further analyses and functional characterisation of particular genes could provide more targets for improving crop yield and quality.

In the course of the studies described here accessions best adapted to sulfate limitation and accessions with various polymorphisms in genes of interest were identified. The genetic modification (GM) of organisms is a common method of implementing novel genetic variation to create new crop varieties (Mba et al. 2012). The development of genetically modified crops and its evaluation remains a very attractive area of research (Tester & Langridge 2010). However, currently the access to transgenic technologies is highly restricted, especially in Europe, mainly by political and bioethical issues (Moose & Mumm 2008). Although it is believed that deployment of GM technologies is a matter of time, currently non-GM technologies such as molecular plant breeding are in favour, perhaps to ease public concerns (Moose & Mumm 2008). In general, plant breeding is a three step process including creation of populations with a genetic variation of commercial value, identification of individuals with superior phenotypes, and development of improved cultivars (Moose & Mumm 2008). Therefore, the identification of genetic variation in traits of commercial interest as well as identification of individuals with specific phenotypes is of great importance. Further investigation of best adapted accessions identified in the analysis described in Chapter 4 including specific patterns of gene expression can help to reveal particular changes in the metabolism of plants that foster specific adaptations. Further characterisation of individuals with specific polymorphisms described in different chapters of this dissertation will set the background for further implementation of these specific changes in the process of crop improvement.

7.3 Future research directions

The understanding of sulfur metabolism, its regulation and interconnection with other elements has progressed rapidly over the last two decades. However, it is still lagging behind the understanding of metabolism of other elements such as nitrate and phosphate. A detailed analysis of accessions with interesting phenotypes (concerning the link between APR activity and sulfate concentration) revealed in the analysis described in Chapter 3 such as Tfa-08, ICE61, or Qar-8a may bring new insights into regulation of sulfate assimilation. These accessions may also be used for creation of new segregating populations which subsequently

could be used for the discovery of novel genes and regulatory elements involved in the control of sulfate reduction pathway. Moreover, the single amino acid substitutions revealed in the course of this project could be used for the investigation of the relation between protein structure and enzymatic activity of enzymes for which the crystal structure is available.

Further characterisation of different patterns of plant response to sulfate and nitrate deficiency including global analysis of gene expression may improve our understanding of plant response to the limitation of these two elements and eventually characterise different adaptation strategies specific for particular groups of accessions. Moreover, detailed analysis of different patterns of plant response to nutrient deficiency may lead to a development of "catalogue" of specific changes in metabolism important for improving particular traits: for example plant biomass, or increased accumulation of a specific metabolite. More in depth analysis of the accessions from the best adapted group would reveal the changes in metabolism important for response to sulfate and nitrate deficiency.

A large number of candidate genes derived by GWAS were not analysed due to time limitation. However, the evidence was provided that the disruption in some of these genes may have a real effect on the accumulation of related anions. Further characterisation of these genes as well as investigation of the T-DNA insertion lines which were omitted in this analysis could result in the discovery of new genes and regulatory factors involved in the control of nutrient homeostasis. Similarly, the detailed analysis of the AT5G03430 – a gene identified by GWAS and subjected to subsequent functional characterisation – may bring new insight into molybdenum metabolism in plants and humans and perhaps better understanding of the function of human FAD synthases.

7.4 Conclusion

In conclusion, the analysis described in this dissertation revealed different patterns of arabidopsis response to sulfate and/or nitrate availability. Individuals well adapted to sulfate limitation were selected and the specific mechanisms involved in plant response to sulfate limitation were described. Additionally, number of genes which, when disrupted, affect the nutrient homeostasis were identified and partially characterised. Functions of some of these genes were already known, however they were not related to nutrition previously (e.g. CAF1). Moreover, a novel gene was identified, which seems to be involved in the metabolism of molybdenum. Taken together, the main aims of this project listed on the beginning of this chapter were met in the course of the research described here. Moreover, the results delivered from a number of experiments also meet the main objectives of the Bionut network

of which this project was a part. Some of the results described here such as the marker traits for sulfate availability can be used immediately in developing new breeding strategies. Some others require more in depth analysis and further investigation to be applied prior to the commercial use. However, this project undoubtedly improved our current understanding of sulfate metabolism and set the background for future research.

Supplement

```
ATPS1_Bay GAGTCAAATCTTTGAGTTAAGACACTAAATGATTTTACAAAACTCGAATAATATAATAAG
ATPS1_Sha GAGTCAAATCTTTGAGTTAAGACACTAAATGATTTTACAAAACTCGAATAATATAATAAG
ATPS1_Col TAGTCAAATCTTTGAGTTAAGACACCAAATGATTTTACAAAACTCGAATAATATAAGAAG

ATPS1_Bay AGTATACTTGTTGTAAGTTGTAACCTTGGAGTACTACGAATTC---AATTTGATGAAGTG
ATPS1_Sha AGTATACTTGTTGTAAGTTGTAACCTTGGAGTACTACGAATTC---AATTTGATGAAGTG
ATPS1_Col AGTATACTTGTTGTAA-------CCTTGGAGTACTACGAATTCAACAATTTGATGAAGTG

ATPS1_Bay AAAACATAGTTAAGAAATGATTGGGAGATTATTCATCTAAAGCTACCAGTCTACCAACTT
ATPS1_Sha AAAACATAGTTAAGAAATGATTGGGAGATTATTCATCTAAAGCTACCAGTCTACCAACTT
ATPS1_Col AAAACATAGTTAAGAAATGATTGGGAGATTATTCATCTAAAGCTACCAGTCTACCAACTT

ATPS1_Bay TCCAGTTTTGACAATATTTACCACGTGGCCCTGAGGAACATCATTGTCAGTTTATCAAGC
ATPS1_Sha TCCAGTTTTGACAATATTTACCACGTGGCCCTGAGGAACATCATTGTCAGTTTATCAAAC
ATPS1_Col TCCAGTTTTGACAATATTTACCACGTGGCCCTGAGGAACATCATTGTCAGTTTATCAAAC

ATPS1_Bay ACTCCATATTTCGTGGCAGCTTCGGGTCAAGAATCCAAATTGGTAATGGTATGGTCACTT
ATPS1_Sha ACTCCATATTTCGTGGCAGCTTCGGGTCAAGAATCCAAATTGGTAATGGTATGGTCACTT
ATPS1_Col ACTCCATATTTCGTGGCAGCTTCGGGTCAAGAATCCAAATT------GGTATTGTCACTT

ATPS1_Bay GGTCAGGTCAAAGTTGAGGAGCTTTTATCTATAGTATACTCAATGAGTCGGCTAAAATGT
ATPS1_Sha GGTCAGGTCAAAGTTGAGGAGCTTTTATCTATAGTATACTCAATGAGTCGGCTAAAATGT
ATPS1_Col GGTCAGGTCAAAGTTGAGGAGCTTTTATATATAGTATAATCAATGAGTCGGCTAAAATGT

ATPS1_Bay AAAATAAAAAACATTTTCTAAATTAAACGTAATAAATTACTATAATCGTATCAAATTTAT
ATPS1_Sha AAAATAAAAAACATTTTCTAAATTAAACGTAATAAATTACTATAATCGTATCAAATTTAT
ATPS1_Col AAAATAAAAAACATGTTCTAAATTAAACGTAATAAATTACTATAATCGTATCAAATTTAT

ATPS1_Bay CCCTCTTATCAAATTAGTACTTTCGAGTTTCGATATTGTTTTTCTTTCTCTGGGAAACTA
ATPS1_Sha CCCTCTTATCAAATTAGTACTTTCGAGTTTCGATATTGTTTTTCTTTCTCTGGGAAACTA
ATPS1_Col CCCTCTTATCAAATTAGTACTTTCGAGTTTCGATATTGTTTTTCTTTCTCTGGGAAACTA

ATPS1_Bay TATATATTTACATTTTTAGAAAGAAAGAAAAAAACTATATATTACATTTTATACGTGCTA
ATPS1_Sha TATATATTTACATTTTTAGAAAGAAAGAAAAAAACTATATATTACATTTTATACGTGCTA
ATPS1_Col TATATATTTACATTTCTAGAAAGAAAGAAAAAAACTATATATTACATTTTATACGTGCTA

ATPS1_Bay AGCCTAGAAAAGATTAATTACAAAGAAATTATACATTTTATTGATCAAGTGGTGCTTAAA
ATPS1_Sha AGCCTAGAAAAGATTAATTACAAAGAAATTATACATTTTATTGATCAAGTGGTGCTTAAA
ATPS1_Col AGCCTAGAAAAGATTAATTACAAAGAAATTATACATTTTATTGATCAAGTGGTGCTTAAA

ATPS1_Bay GCAGTAAAAACAATTTGAACTCATAAATCGTAAATACGTCGACATATTTCGTGATCTCCA
ATPS1_Sha GCAGTAAAAACAATTTGAACTCATAAATCGTAAATACGTCGATATATTTCGTGATCTCCA
ATPS1_Col GCAGTAAAAACAATTTGAACTCATAAATCGTAAATACGTCGACATATTTCGTGATCTCCA

ATPS1_Bay TTAATTTTTCTTTTTTTAAAGATTGACGCAAAATAATATTCTGAAAATGAAAAAGTAAAA
ATPS1_Sha TTAATTTTTCTTTTCTTAAGGATTGACGCAAAATAATATTCTGAAAATGAAAAAGTAAAA
ATPS1_Col TTAATTTTTCTTTTTTTAAAGATTGACGCAAAATAATATTCTGAAAATGAAAAAGTAAAA

ATPS1_Bay TAAGAGGGGACAAATATTCGAGATGTGACGTGGCAGATCGAGTGGTTTAAATATTCTATT
ATPS1_Sha TAAGAGGGGACAAATATTCGAGATGTGACGTGGCAGATCGAGTGGTTTAAATATTCTATT
ATPS1_Col TAAGAGGGGACAAATATTCGAGATGTGACGTGGCAGATCGAGTGGTTTAAATATTCTATT

ATPS1_Bay AGCAAGTGGTTTGTGTAATAAGCAAATGGGTGGTCGAACCTGACCGTATTCTTGGATCTA
ATPS1_Sha AGCAAGTGGTTTGTGTAATAAGCAAATGGGTGGTCGAACCTGACCGTATTCTTGGATCTA
ATPS1_Col AGCAAGTGGTTTGTGTAATAAGCAAATGGGTGGTCGAACCTGACCGTATTCTTGGATCTA

ATPS1_Bay TTCAACTGTAGCATCAGTCCACCTTCCTTACCTCATCTTTCCTAACTTTTTAATACCTTT
ATPS1_Sha TTCAACTGTAGCATCAGTCCACCTTCCTTACCTCATCTTTCCTAACTTTTTAATACCTTT
ATPS1_Col TTCAACTGTAGCATCAGTCCACCTTCCTTACCTCATCTTTCCTAACTTTTTAATACCTTT

ATPS1_Bay TTAATTTGCTAAAAACATCACTACTATTTATATTTGATCTCTAATTACTGTTTCAACTCT
ATPS1_Sha TTAATTTGCTAAAAACATCACTACTATTTATATTTGATCTCTAATTACTGTTTCAACTCT
ATPS1_Col TTAATTTGCTAAAAACATCACTACTATTTATATTTGATCTCTAATTACTGTTTCAACTCT

ATPS1_Bay GAATATTCCGTAATCCTTATATTAATATGTCCAAAAATATGTAATTTCTGGACTTTCCCT
ATPS1_Sha GAATATTCCGTAATCCTTATATTAATATGTCCAAAAATATGTAATTTCTGGACTTTCCCT
ATPS1_Col GAATATTCCGTAATCCTTATATTAATATGTCCAAAAATATATAATTTCTGGACTTTCCCT
```

```
ATPS1_Bay TGAGATGGGAATATGAAAGAATTGCATCATTTACCTAATCATATGAACACAAATAGATAT
ATPS1_Sha TGAGATGGGAATATGAAAGAATTGCATCATTTACCTAATCATATGAACACAAATAGATAT
ATPS1_Col TGAGATGGGAATATGAAAGAATTGCATCATTTACCTAATCATATGAACACAAATAGATAT

ATPS1_Bay TGGAAAAAATGTGGTTTTATTTTTCATGTTTTGTTCGATTATCTTTATCTTTATCCCAAA
ATPS1_Sha TGGAAAAAATGTGGTTTTATTTTTCATGTTTTGTTCGATTATCTTTATCTTTATCCCAAA
ATPS1_Col TGGAAAAAATGTGGTTTTATTTTTCATGTTTTGTTCGATTATCTTTATCTTTATCCCAAA

ATPS1_Bay AAAAAAAAAATCATTCGATTATATCTAAAATTCAAAATAGATTACAGAAATATAATTTAT
ATPS1_Sha AAAAAAAAAATCATTCGATTATATCTAAAATTCAAAATAGATTACAGAAATATAATTTAT
ATPS1_Col AAAAAAAAAATCATTCGATTATATCTAAAATTCAAAATAGATTACAGAAATATAATTTAT

ATPS1_Bay TAGCAAAACGGCATGTTTAAAGCGTTTGAAGTTATTAAATCATTAGTAAGAATATATAAG
ATPS1_Sha TAGCAAAACGGCATGTTTAAAGCGTTTGAAGTTATTAAATCATTAGTAAGAATATATAAG
ATPS1_Col TAGCAAAACGGCATGTTTAAAGCGTTTGAAGTTATTAAATCATTAGTAAGAATATATAAG

ATPS1_Bay AATATAATTAGTGGAAATACATAGTAAGTATCATTGGTTTTTGCCACATATGGTGAGCAA
ATPS1_Sha AATGTAATTAGTGGAAATACATAGTAAGTATCATTGGTTTTTGCCACATATGGTGAGCAA
ATPS1_Col AATGTAATTAGTGGAAATACATAGTAAGTATCATTGGTTTTTGCCACATATGGTGAGCAA

ATPS1_Bay TTTTTTATTTTAAGAAGGGAAAATCAATTTGTACATAGATTTATGTCACTTATTCAATTG
ATPS1_Sha TTTTTTATTTTAAGAAGGGAAAATCAATTTGTACATAGATTTATGTCACTTATTCAATTG
ATPS1_Col TTTTTTATTTTAAGAAGGGAAAATCAATTTGTACATAGATTTATGTCACTTATTCAATTG

ATPS1_Bay AATAATACAGAAGGATTTAAAGTCTAAAGTAAAAACAGGCAAAATAATAATATGTTTTTT
ATPS1_Sha AATAATACAGAAGGATTTAAAGTCTAAAGTAAAAACAGGCAAAATAATAATATGTTTTTT
ATPS1_Col AATAATACAGAAGGATTTAAAGTCTAAAGTAAAAACAGGCAAAATAATAATATGTTTTTT

ATPS1_Bay TCTTTGATCGCTCAGATTATCGTATTAAAATTTGGATTATGACATAACAACGATAATAAT
ATPS1_Sha TCTTTGATCGCTCAGATTATCGTATTAAAATTTGGATTATGACATAACAACGATAATAAT
ATPS1_Col TCTTTGATCGCTCAGATTATCGTATTAAAATTTGGATTATGACATAACAACGATAATAAT

ATPS1_Bay ACAAACTAGTTGGTTATGAACTCTGAATAAATTATTTTAAAGAAAGAATACTACTATTTA
ATPS1_Sha ACAAACTAGTTGGTTATGAACTCTGAATAAATTATTTTAAAGAAAGAATACTACTATTTA
ATPS1_Col ACAAACTAGTTGGTTATGAACTCTGAATAAATTATTTTAAAGAAAGAATACTACTATTTA

ATPS1_Bay ATTATAAAATGACTCTGCATCATATCAATAAGGTAACCTCGTTATTATAAACGTCACACT
ATPS1_Sha ATTATAAAATGACTCTGCATCATATCAATAAGGTAACCTCGTTATTATAAACGTCACACT
ATPS1_Col ATTATAAAATGACTCTGCATCATATCAATAAGGTAACCTCGTTATTATAAACGTCACACT

ATPS1_Bay AACACACTGTATTAGTATTTTAATTACACAGTGAAAAAATTTAATTAATTACTAATCTCT
ATPS1_Sha AACACACTGTATTAGTATTTTAATTACACAGTGAAAAAATTTAATTAATTACTAATCTCT
ATPS1_Col AACACACTGTATTAGTATTTTAATTACACAGTGAAAAAATTTAATTAATTACTAATCTCT

ATPS1_Bay GTCCAGGTACATAATATTATTCCAAGATACGGTCCTTCGTTACTATAAACTCTATAAAAA
ATPS1_Sha GTCCAGGTACATAATATTATTCCAAGATACGGTCCTTCGTTACTATAAACTCTATAAAAA
ATPS1_Col GTCCAGGTACATAATATTATTCCAAGATACGGTCCTTCGTTACTATAAACTCTATAAAAA

ATPS1_Bay CCAATTTTCACTTCCAATTGAATTGGGAACAAACCAAATCTCTATCTCTCTCCATTAGAG
ATPS1_Sha CCAATTTTCACTTCCAATTGAATTGGGAACAAACCAAATCTCTATCTCTCTCCATTAGAG
ATPS1_Col CCAATTTTCACTTCCAATTGAATTGGGAACAAACCAAATCTCTATCTCTCTCCATTAGAG

ATPS1_Bay CTTGAAGCAGCCATAGCCTGAGAAAACCTTCAACAATGGCTTCAATGGCTGCCGTCTTAA
ATPS1_Sha CTTGAAGCAGCCATAGCCTGAGAAAACCTTCAACAATGGCTTCAATGGCTGCCGTCTTAA
ATPS1_Col CTTGAAGCAGCCATAGCCTAACAAAACCTTCAACAATGGCTTCAATGGCTGCCGTCTTAA
                                                  **************************
ATPS1_Bay GCAAAACTCCATTCCTCTCTCAACCACTAACCAAATCATCTCCAAACTCCGATCTCCCCT
ATPS1_Sha GCAAAACTCCATTCCTCTCTCAACCACTAACCAAATCATCTCCAAACTCCGATCTCCCCT
ATPS1_Col GCAAAACTCCATTCCTCTCTCAACCACTAACCAAATCATCTCCAAACTCCGATCTCCCCT
          ************************************************************

ATPS1_Bay TCGCCGCGGTTTCCTTCCCTTCCAAATCCCTACGCCGCCGCGTAGGATCAATCCGAGCCG
ATPS1_Sha TCGCCGCGGTTTCCTTCCCTTCCAAATCCCTACGCCGCCGCGTAGGATCAATCCGAGCCG
ATPS1_Col TCGCCGCGGTTTCCTTCCCTTCCAAATCCCTACGCCGCCGCGTAGGATCAATCCGAGCCG
          ************************************************************
```

```
ATPS1_Bay GATTAATCGCTCCCGACGGTGGTAAGCTTGTAGAGCTCATCGTGGAAGAGCCAAAGCGGC
ATPS1_Sha GATTAATCGCTCCCGACGGTGGTAAGCTTGTAGAGCTCATCGTGGAAGAGCCAAAGCGGC
ATPS1_Col GATTAATCGCTCCCGACGGTGGTAAGCTTGTAGAGCTCATCGTGGAAGAGCCAAAGCGGC
          ************************************************************

ATPS1_Bay GAGAGAAGAAACACGAGGCGGCGGATTTGCCACGTGTTGAGCTGACGGCGATTGACTTGC
ATPS1_Sha GAGAGAAGAAACACGAGGCGGCGGATTTGCCACGTGTTGAGCTGACGGCGATTGACTTGC
ATPS1_Col GAGAGAAGAAACACGAGGCGGCGGATTTGCCACGTGTTGAGCTGACGGCGATTGACTTGC
          ************************************************************

ATPS1_Bay AATGGATGCATGTATTAAGCGAAGGCTGGGCAAGTCCACTCGGAGGTTTCATGAGAGAAT
ATPS1_Sha AATGGATGCATGTATTAAGCGAAGGCTGGGCAAGTCCACTCGGAGGTTTCATGAGAGAAT
ATPS1_Col AATGGATGCATGTATTAAGCGAAGGCTGGGCAAGTCCACTCGGAGGTTTCATGAGAGAAT
          ************************************************************

ATPS1_Bay CCGAGTTCCTCCAAACTCTTCATTTTAACTCGCTACGTCTTGACGACGGCTCCGTCGTTA
ATPS1_Sha CCGAGTTCCTCCAAACTCTTCATTTTAACTCGCTACGTCTTGACGACGGCTCCGTCGTTA
ATPS1_Col CCGAGTTCCTCCAAACTCTTCATTTTAACTCGCTACGTCTTGACGACGGCTCCGTCGTTA
          ************************************************************

ATPS1_Bay ACATGTCCGTGCCTATTGTTCTCGCTATTGACGATGAACAAAAAGCACGTATCGGCGAGT
ATPS1_Sha ACATGTCCGTGCCTATTGTTCTCGCTATTGACGATGAACAAAAAGCACGTATCGGCGAGT
ATPS1_Col ACATGTCCGTGCCTATTGTTCTCGCTATTGACGATGAACAAAAAGCACGTATCGGCGAGT
          ************************************************************

ATPS1_Bay CTACACGTGTCGCTCTTTTTAATTCCGATGGTAACCCCGTCGCTATCCTCAGCGAGTAAG
ATPS1_Sha CTACACGTGTCGCTCTTTTCAATTCCGATGGTAACCCCGTCGCTATCCTCAGCGAGTAAG
ATPS1_Col CTACACGTGTCGCTCTTTTCAATTCCGATGGTAACCCCGTCGCTATCCTCAGCGAGTAAG
                ********************************************************

ATPS1_Bay TCCTCCTTCAACATACTCAGATTCAAATTCAGATCATAGATTGTTCGAGAAGACGAAAAC
ATPS1_Sha TCCTTCTTCAACATACTCAGATTCAAATTCAGATCATAGATTGTTGGAGAAGACGAAAAC
ATPS1_Col TCCTTCTTCAACATACTCAGATTCAAATTCAGATCATAGATTGTTGGAGAAGACGAAAAC

ATPS1_Bay TGATATTTGAATTTGATTTTGAACT------------GATTAGTTAAATACAGTAGAAT
ATPS1_Sha TGAGATTTGAATTTGATTTTGAACTTATCTCTCTTTGAGATTAGTTAAATACAGTATAAT
ATPS1_Col TGAGATTTGAATTTGATTTTGAACTTATCTCTCTTTGAGATTAGTTAAATACAGTATAAT

ATPS1_Bay CTCCATAAATTAATACTCGATAAATTAAGAA--ACTTAGACTAGATCTGATTAGTATTGA
ATPS1_Sha CTCGATAAAGTAATACTCGATAAATTAATAACTATGTATTCACGATTAGACTTAGACTAG
ATPS1_Col CTCGATAAAGTAATACTCGATAAATTAATAACTATGTATTCACGATTAGACTTAGACTAG

ATPS1_Bay GATTAGTCAAATATGT-ATT-GAACTTGGCTATATGTGTAATTGATTACTGAAATCTGTT
ATPS1_Sha ATCTGATTAATCTTGAGATTAGAATTTGGCTATATGTGTAATTGGTTACTGAAATCTGTT
ATPS1_Col ATCTGATTAATCTTGAGATTAGAATTTGGCTATATGTGTAATTGGTTACTGAAATCTGTT

ATPS1_Bay TGTTTTTATGTTTGATCAGTATTGAGATTTATAAGCATCCAAAGGAAGAAAGGATAGCTA
ATPS1_Sha TGTTTTTATGTTTGATCAGTATTGAGATTTATAAGCATCCAAAGGAAGAAAGGATAGCTA
ATPS1_Col TGTTTTTATGTTTGATCAGTATTGAGATTTATAAGCATCCAAAGGAAGAAAGGATAGCTA
                                  ***************************************

ATPS1_Bay GAACATGGGGTACGACGGCTCCAGGTTTGCCTTACGTAGACGAGGCGATAACTAATGCTG
ATPS1_Sha GAACATGGGGTACGACGGCTCCAGGTTTGCCTTACGTAGACGAGGCGATAACTAATGCTG
ATPS1_Col GAACATGGGGTACGACGGCTCCAGGTTTGCCTTACGTAGACGAGGCGATAACTAATGCTG
          ************************************************************

ATPS1_Bay GAAACTGGCTCATTGGGGGTGATCTTGAGGTTCTTGAGCCAGTGAAGTACAATGATGGGC
ATPS1_Sha GAAACTGGCTCATTGGGGGTGATCTTGAGGTTCTTGAGCCAGTGAAGTACAATGATGGGC
ATPS1_Col GAAACTGGCTCATTGGGGGTGATCTTGAGGTTCTTGAGCCAGTGAAGTACAATGATGGGC
          ************************************************************

ATPS1_Bay TTGATCGTTTCAGGCTTTCGCCTGCTGAGTTACGTAAAGAGTTGGAGAAGCGTAATGCGG
ATPS1_Sha TTGATCGTTTCAGGCTTTCGCCTGCTGAGTTACGTAAAGAGTTGGAGAAGCGTAATGCGG
ATPS1_Col TTGATCGTTTCAGGCTTTCGCCTGCTGAGTTACGTAAAGAGTTGGAGAAGCGTAATGCGG
          ************************************************************
```

```
ATPS1_Bay ATGCGGTGTTTGCTTTCCAGCTGAGGAATCCTGTTCATAATGGTCATGCTCTTCTTATGA
ATPS1_Sha ATGCGGTGTTTGCTTTCCAGCTGAGGAATCCTGTTCATAATGGTCATGCTCTTCTTATGA
ATPS1_Col ATGCGGTGTTTGCTTTCCAGCTGAGGAATCCTGTTCATAATGGTCATGCTCTTCTTATGA
          ************************************************************

ATPS1_Bay CTGATACTCGTAGGAGACTTCTTGAGATGGGTTACAAAAACCCTATTCTTTTGCTTCATC
ATPS1_Sha CTGATACTCGTAGGAGACTTCTTGAGATGGGTTACAAAAACCCTATTCTTTTGCTTCATC
ATPS1_Col CTGATACTCGTAGGAGACTTCTTGAGATGGGTTACAAAAACCCTATTCTTTTGCTTCATC
          ************************************************************

ATPS1_Bay CGTTAGGTGGGTTTACAAAGGCTGATGATGTTCCTTTAGATTGGAGGATGAAGCAACACG
ATPS1_Sha CGTTAGGTGGGTTTACAAAGGCTGATGATGTTCCTTTAGATTGGAGGATGAAGCAACACG
ATPS1_Col CGTTAGGTGGGTTTACAAAGGCTGATGATGTTCCTTTAGATTGGAGGATGAAGCAACACG
          ************************************************************

ATPS1_Bay AGAAGGTAAAAAAAGTTATTGATTGTTGTGTGTCTCTTGAGTGTGAGTGTGTGTATCTTG
ATPS1_Sha AGAAGGTAAAAAAAGTTATTGATTGTTGTGTGTCTCTTGAGTGTGAGTGTGTGTATCTTG
ATPS1_Col AGAAGGTAAAAAAAGTTATTGATTGTTGTGTGTCTCTTGAGTGTGAGTGTGTGTATCTTG
          *****

ATPS1_Bay AAATGCTAAATGTGTTTATGAATGTAAAACAGGTTCTAGAGGATGGTGTTCTCGATCCGG
ATPS1_Sha AAATGCTAAATGTGTTTATGAATGTAAAACAGGTTCTAGAGGATGGTGTTCTCGATCCGG
ATPS1_Col AAATGCTAAATGTGTTTATGAATGTAAAACAGGTTCTAGAGGATGGTGTTCTCGATCCGG
                                          ****************************

ATPS1_Bay AGACTACAGTGGTTTCGATATTCCCGTCACCTATGCATTACGCTGGTCCAACCGAAGTGC
ATPS1_Sha AGACTACAGTGGTTTCGATATTCCCGTCACCTATGCATTACGCTGGTCCAACCGAAGTGC
ATPS1_Col AGACTACAGTGGTTTCGATATTCCCGTCACCTATGCATTACGCTGGTCCAACCGAAGTGC
          ************************************************************

ATPS1_Bay AGTGGCACGCAAAGGCTAGAATCAATGCTGGTGCTAACTTTTACATTGTGGGTCGTGATC
ATPS1_Sha AGTGGCACGCAAAGGCTAGAATCAATGCTGGTGCTAACTTTTACATTGTGGGTCGTGATC
ATPS1_Col AGTGGCACGCAAAGGCTAGAATCAATGCTGGTGCTAACTTTTACATTGTGGGTCGTGATC
          ************************************************************

ATPS1_Bay CTGCTGGGATGGGTCATCCAGTAGAGAAACGTGATCTTTACGATGCTGATCATGGAAAGA
ATPS1_Sha CTGCTGGGATGGGTCATCCAGTAGAGAAACGTGATCTTTACGATGCTGATCATGGAAAGA
ATPS1_Col CTGCTGGGATGGGTCATCCAGTAGAGAAACGTGATCTTTACGATGCTGATCATGGAAAGA
          ************************************************************
ATPS1_Bay AAGTACTAAGCATGGCACCAGGACTCGAACGACTCAACATCCTTCCTTTCAGGGTATATA
ATPS1_Sha AAGTACTAAGCATGGCACCAGGACTCGAACGACTCAACATCCTTCCTTTCAGGGTATATA
ATPS1_Col AAGTACTAAGCATGGCACCAGGACTCGAACGACTCAACATCCTTCCTTTCAGGGTATATA
          ********************************************************

ATPS1_Bay CAATTCGAAAAGATTCACACTTTTTGTTTGACAATGTAGAGATCTAATTCCTTGGTGTAA
ATPS1_Sha CAATTCGAAAAGATTCACACTTTTTGTTTGACAATGTAGAGATCTAATTCCTTGGTGTAA
ATPS1_Col CAATTCGAAAAGATTCACACTTTTTGTTTGACAATGTAGAGATCTAATTCCTTGGTGTAA

ATPS1_Bay CTGCAGGTTGCTGCATATGACAAGACGCAAGGCAAGATGGCTTTCTTCGATCCCTCGAGG
ATPS1_Sha CTGCAGGTTGCTGCATATGACAAGACGCAAGGCAAGATGGCTTTCTTCGATCCCTCGAGG
ATPS1_Col CTGCAGGTTGCTGCATATGACAAGACGCAAGGCAAGATGGCTTTCTTCGATCCCTCGAGG
                ************************************************

ATPS1_Bay CCTCAAGATTTCTTGTTCATCTCCGGCACTAAGGTAATATACCAGTCCTACATTGTTAAA
ATPS1_Sha CCTCAAGATTTCTTGTTCATCTCCGGCACTAAGGTAATATACCAGTCCTACATTGTTAAA
ATPS1_Col CCTCAAGATTTCTTGTTCATCTCCGGCACTAAGGTAATATACCAGTCCTACATTGTTAAA
          ********************************

ATPS1_Bay ATTCTTCATAGTTTGTTTATAAAACAAACCTCTAAATGTTTTTCGATTATTCTAGATGCG
ATPS1_Sha ATTCTTCATAGTTTGTTTATAAAACAAACCTCTAAATGTTTTTCGATTATTCTAGATGCG
ATPS1_Col ATTCTTCATAGTTTGTTTATAAAACAAACCTCTAAATGTTTTTCGATTATTCTAGATGCG
                                                    *****

ATPS1_Bay CACATTGGCAAAGAACAACGAAAACCCGCCAGACGGTTTTATGTGCCCAGGTGGATGGAA
ATPS1_Sha CACATTGGCAAAGAACAACGAAAACCCGCCAGACGGTTTTATGTGCCCAGGTGGATGGAA
ATPS1_Col CACATTGGCAAAGAACAACGAAAACCCGCCAGACGGTTTTATGTGCCCAGGTGGATGGAA
          ************************************************************
```

```
ATPS1_Bay AGTTCTGGTGGATTACTATGAGAGCTTGACTCCGGCGGGTAATGGTAGACTACCAGAAGT
ATPS1_Sha AGTTCTGGTGGATTACTATGAGAGCTTGACTCCGGCGGGTAATGGTAGACTACCAGAAGT
ATPS1_Col AGTTCTGGTGGATTACTATGAGAGCTTGACTCCGGCGGGTAATGGTAGACTACCAGAAGT
          ************************************************************

ATPS1_Bay GGTTCCGGTGTAAGACAAAACTGTTCGTTTCAAATTGTAACGTTTGTGTTGTGAAGCCTT
ATPS1_Sha GGTTCCGGTGTAAGACAAAACTGTTCGTTTCAAATTGTAACGTTTGTGTTGTGAAGCCTT
ATPS1_Col GGTTCCGGTGTAAGACAAAACTGTTCGTTTCAAATTGTAACGTTTGTGTTGTGAAGCCTT
          ************

ATPS1_Bay GTAGCAACAATCATTGTTGTATTGGGAGAGAAGCCTATGTATAATCTGGCTTGACCTTTT
ATPS1_Sha GTAGCAACAATCATTGTTGTATTGGGAGAGAAGCCTATGTATAATCTGGCTTGACCTTTT
ATPS1_Col GTAGCAACAATCATTGTTGTATTGGGAGAGAAGCCTATGTATAATCTGGCTTGACCTTTT

ATPS1_Bay TCCAAATAAAATACAGAAGAAAAAAAGACTGTTTTTCGTTTGCAAGATAATTTACGAAAC
ATPS1_Sha TCCAAATAAAATACAGAAGAAAAAAAGACTGTTTTTCGTTTGCAAGATAATTTACGAAAC
ATPS1_Col TCCAAATAAAATACAGAAGAAAAAAAGACTGTTTTTCGTTTGCAAGATAATTTACGAAAC

ATPS1_Bay TTGTAATATTTGGGCCTCAAACTTTGTACCATATTAATGAAACGATTGTGTTTACATATA
ATPS1_Sha TTGTAATATTTGGGCCTCAAACTTTGTACCATATTAATGAAACGATTGTGTTTACATATA
ATPS1_Col TTGTAATATTTGGGCCTCAAACTTTGTACCATATTAATGAAACGATTGTTTTTACATATA

ATPS1_Bay AGGATTTACGTACTTTTGACTT---------------------------------------
ATPS1_Sha AGGATTTACGTATTTTTGACTTTTGACTACATCTTTTTCTATTTAATCACTCACTTTTTT
ATPS1_Col AGGATTTACGTATTTTTGACTTTTGACTACATCTTTTTCTATTTAATCACTCACTTTTTT

ATPS1_Bay ----------------------------------------CATAAAATCTAGTGAT
ATPS1_Sha TTTTTTTTTTTTCGCCATTTAGTCACTCACTTTTGACTTCATATATGTATTCTAGTAGTC
ATPS1_Col TTTTTTTTTTTTCGCCATTTAATCACTCACTTTTGACTTCATATATGTATTCTAGTAGTC

ATPS1_Bay G-ATCCAATTGTAGTAATGGGCCTTCAAAGGCATCTTGTGGTAATGGGCCTATTAATTTG
ATPS1_Sha GTACCCAATTGTTGTAATGGGCCTTCAAAGGCCTCTTGTAGTAATGGGCCTATTAATTTG
ATPS1_Col GTACCCAATTGTTGTAATGGGCCTTCAAAGGCCTCTTGTAGTAATGGGCCTATTAATTTG

ATPS1_Bay TAAACTAACCAAACCCGAAGTTTCGTCTCCTGAGTCATTTCACGACTGATGTCTTTCTTC
ATPS1_Sha TAAACTAACCAAACCCGAAGTTTCGTCTCCTGAGTCATTTCACGACTGATGTTTTTCTTC
ATPS1_Col TAAACTAACCAAACCCGAAGTTTCGTCTCCTGAGTCATTTCACGACTGATGTTTTTCTTC

ATPS1_Bay TCCGACGCTACAATCGGTGTCTGTTTCCGAAGTTGTTCGTTCCATTAGAGCTTCAGATTT
ATPS1_Sha TCCGACGCTACAATCGGTGTCTGTTTCCGAAGTTGTTCGTTCCATTAGAGCTTCAGATTT
ATPS1_Col TCCGACGCTACAATCGGTGTCTGTTTCCGAAGTTGTTCGTTCCATTAGAGCTTCAGATTT

ATPS1_Bay CGTCGTTTGGTAATTTCAAGTTAGAGCTACTTAATCTTTTGATCTGTCAAACCAATTTCT
ATPS1_Sha CGTCGTTTGGTAATTTCAAGTTAGAGCTACTTAATCTTTTGATCTGTCAAACCAATTTCT
ATPS1_Col CGTCGTTTGGTAATTTCAAGTTAGAGCTACTTAATCTTTTGATCTGTCAAACCAATTTCT

ATPS1_Bay CAGATAAATTCACTCTAGGGTTTGTGAAGAACGATGATATGATCACATTTCAGAATGCAA
ATPS1_Sha CAGATAAATTCACTCTAGGGTTTGTGAAGAACGATGATATGATCACATTTCACAATGCAA
ATPS1_Col CAGATAAATTCACTCTAGGGTTTGTGAAGAACGATGATATGATCACATTTCACAATGCAA
```

Figure S3.1 ***Comparison of ATPS1 gene sequence from Bay-0, Sha and Col-0***

Identical residues are in blue, SNPs and indels are in red. Coding region is marked with asterisks. The results were obtained by Anna Koprivova and the figure comes from Koprivova et al. (2013).

```
Col-0      CACGTGTCGCTCTTTTCAATTCCGATGGTAACCCCGTCGCTATCCTCAGCGAGTAAGTCC
Sha        CACGTGTCGCTCTTTTCAATTCCGATGGTAACCCCGTCGCTATCCTCAGCGAGTAAGTCC
Blh-1_1    CACGTGTCGCTCTTTTTAATTCCGATGGTAACCCCGTCGCTATCCTCAGCGAGTAAGTCC
Aa-0_1     CACGTGTCGCTCTTTTCAATTCCGATGGTAACCCCGTCGCTATCCTCAGCGAGTAAGTCC
Hel-3_1    CACGTGTCGCTCTTTTTAATTCCGATGGTAACCCCGTCGCTATCCTCAGCGAGTAAGTCC
Ak-1_1     CACGTGTCGCTCTTTTTAATTCCGATGGTAACCCCGTCGCTATCCTCAGCGAGTAAGTCC
Krot-0_1   CACGTGTCGCTCTTTTTAATTCCGATGGTAACCCCGTCGCTATCCTCAGCGAGTAAGTCC
Bor-1_1    CACGTGTCGCTCTTTTTAATTCCGATGGTAACCCCGTCGCTATCCTCAGCGAGTAAGTCC
Bay-0      CACGTGTCGCTCTTTTTAATTCCGATGGTAACCCCGTCGCTATCCTCAGCGAGTAAGTCC
           ********************************** ***********************************

Col-0      TTCTTCAACATACTCAGATTCAAATTCAGATCATAGATTGTTGGAGAAGACGAAAACTGA
Sha        TTCTTCAACATACTCAGATTCAAATTCAGATCATAGATTGTTGGAGAAGACGAAAACTGA
Blh-1_1    TCCTTCAACATACTCAGATTCAAATTCAGATCATAGATTGTTGGAGAAGACGAAAACTGA
Aa-0_1     TTCTTCAACATACTCAGATTCAAATTCAGATCATAGATTATTGGAGAAGACGAAAACTGA
Hel-3_1    TCCTTCAACATACTCAGATTCAAATTCAGATCATAGATTGTTCGAGAAGACGAAAACTGA
Ak-1_1     TCCTTCAACATACTCAGATTCAAATTCAGATCATAGATTGTTGGAGAAGACGAAAACTGA
Krot-0_1   TCCTTCAACATACTCAGATTCAAATTCAGATCATAGATTGTTGGAGAAGACGAAAACTGA
Bor-1_1    TCCTTCAACATACTCAGATTCAAATTCAGATCATAGATTGTTCGAGAAGACGAAAACTGA
Bay-0      TCCTTCAACATACTCAGATTCAAATTCAGATCATAGATTGTTCGAGAAGACGAAAACTGA

Col-0      GATTTGAATTTGATTTTGAACTTATCTCTCTTTGAGATTAGTTAAATACAGTATAATCTC
Sha        GATTTGAATTTGATTTTGAACTTATCTCTCTTTGAGATTAGTTAAATACAGTATAATCTC
Blh-1_1    GATTTGAATTTGATTTTGAACTTATCTCTCTTTGAGATTAGTTAAATACAGTATAATCTC
Aa-0_1     GATTTGAATTTGATTTTGAACTTATCTCTCTTTGAGATTAGTTAAATACAGTATAATCTC
Hel-3_1    TATTTGAATTTGATTTTGAACTTATCTCTCTTTGAGATTAGTTAAATACAGTATAATCTC
Ak-1_1     GATTTGAATTTGATTTTGAACT-------------GATTAGTTAAATACAGTAGAATCTC
Krot-0_1   GATTTGAATTTGATTTTGAACT-------------GATTAGTTAAATACAGTAGAATCTC
Bor-1_1    GATTTGAATTTGATTTTGAACT-------------GATTAGTTAAATACAGTAGAATCTC
Bay-0      TATTTGAATTTGATTTTGAACT-------------GATTAGTTAAATACAGTAGAATCTC

Col-0      GATAAAGTAATACTCGATAAATTAATAACTATGTATTCACGATTAGACTTAGACTAGATC
Sha        GATAAAGTAATACTCGATAAATTAATAACTATGTATTCACGATTAGACTTAGACTAGATC
Blh-1_1    GATAAAGTAATACTCGATAAATTAATAACTATGTATTCACGATTAGACTTAGACTAGATC
Aa-0_1     GATAAATTAATACTCCATAAATTAATAACTATGTACTCACGATTAGACTTAGACTAGATC
Hel-3_1    GATAAATTAATACTCGATAAATTAATAACTATGTACTCACGATTAGACTTAGACTAGATC
Ak-1_1     CATAAAGTAATACTCGATAAATTAAGAAACTTAGACTA--GATCTGATTAGTATTGAGAT
Krot-0_1   CATAAAGTAATACTCGATAAATTAAGAAACTTAGACTA--GATCTGATTAGTATTGAGAT
Bor-1_1    CATAAAGTAATACTCGATAAATTAAGAAACTTAGACTA--GATCTGATTAGTATTGAGAT
Bay-0      CATAAAGTAATACTCGATAAATTAAGAAACTTAGACTA--GATCTGATTAGTATTGAGAT

Col-0      TGATTAATCTTGAGATTAGAATTTGGCTATATGTGTAATTGGTTACTGAAATCTGTTTGT
Sha        TGATTAATCTTGAGATTAGAATTTGGCTATATGTGTAATTGGTTACTGAAATCTGTTTGT
Blh-1_1    TGATTAATCTTGAGATTAGAATTTGGCTATATGTGTAATTGATTACTGAAATCTGTTTGT
Aa-0_1     TGATTAATCTTGAGATTAGAATTTGGCTATATGTGTAATTGGTTACTGAAATCTGTTTGT
Hel-3_1    TGATTAATCTTGAGATTAGAATTTGGCTATATGTGTAATTGATTACTGAAATCTGTTTGT
Ak-1_1     TAGTCAAATATGTATT--GAACTTGGCTATATGTGTAATTGATTACTGAAATCTGTTTGT
Krot-0_1   TAGTCAAATATGTATT--GAACTTGGCTATATGTGTAATTGGTTACTGAAATCTGTTTGT
Bor-1_1    TAGTCAAATATGTATT--GAACTTGGCTATATGTGTAATTGATTACTGAAATCTGTTTGT
Bay-0      TAGTCAAATATGTATT--GAACTTGGCTATATGTGTAATTGATTACTGAAATCTGTTTGT

Col-0      TTTTATGTTTGATCAGTATTGAGATTTATAAGCATCCAAAGGAAGAAAGGATAGCTAGAA
Sha        TTTTATGTTTGATCAGTATTGAGATTTATAAGCATCCAAAGGAAGAAAGGATAGCTAGAA
Blh-1_1    TTTTATGTTTGATCAGTATTGAGATTTATAAGCATCCAAAGGAAGAAAGGATAGCTAGAA
Aa-0_1     TTTTATGTTTGATCAGTATTGAGATTTATAAGCATCCAAAGGAAGAAAGGATAGCTAGAA
Hel-3_1    TTTTATGTTTGATCAGTATTGAGATTTATAAGCATCCAAAGGAAGAAAGGATAGCTAGAA
Ak-1_1     TTTTATGTTTGATCAGTATTGAGATTTATAAGCATCCAAAGGAAGAAAGGATAGCTAGAA
Krot-0_1   TTTTATGTTTGATCAGTATTGAGATTTATAAGCATCCAAAGGAAGAAAGGATAGCTAGAA
Bor-1_1    TTTTATGTTTGATCAGTATTGAGATTTATAAGCATCCAAAGGAAGAAAGGATAGCTAGAA
Bay-0      TTTTATGTTTGATCAGTATTGAGATTTATAAGCATCCAAAGGAAGAAAGGATAGCTAGAA
                                           ****************************************
```

```
Col-0     AATATTTGGGCCTCAAACTTTGTACCATATTAATGAAACGATTGTTTTTACATATAAGGA
Sha       AATATTTGGGCCTCAAACTTTGTACCATATTAATGAAACGATTGTGTTTACATATAAGGA
Aa-0_2    AATATTTGGGCCTCAAACTTTGTACCATATTAATGAAACGATTGTGTTTACATATAAGGA
Blh-1_2   AATATTTGGGCCTCAAACTTTGTACCATATTAATGAAACGATTGTGTTTACATATAAGGA
Hel-3_2   AATATTTGGGCCTCAAACTTTGTACCATATTAATGAAACGATTGTGTTTACATATAAGGA
Ak-1_2    AATATTTGGGCCTCAAACTTTGTACCATATTAATGAAACGATTGTGTTTACATATAAGGA
Krot-0_2  AATATTTGGGCCTCAAACTTTGTACCATATTAATGAAACGATTGTGTTTACTATAAGGA
Bor-1_2   AATATTTGGGCCTCAAACTTTGTACCATATTAATGAAACGATTGTGTTTACTATAAGGA
Bay-0     AATATTTGGGCCTCAAACTTTGTACCATATTAATGAAACGATTGTGTTTACATATAAGGA

Col-0     TTTACGTATTTTTGACTTTTGACTACATCTTTTTCTATTTAATCACTCACTTTTTTTTTT
Sha       TTTACGTATTTTTGACTTTTGACTACATCTTTTTCTATTTAATCACTCACTTTTTTTTTT
Blh-1_2   TTTACGTATTTTTGACTTTTGACTACACCTTTTTCTATTTAATCACTCACTTTTTTTTTT
Aa-0_2    TTTACGTATTTTTTACTTTTGACTACATCTTTTTCTATTTAATCACTCACTTTT------
Hel-3_2   TTTACGTACTTTTGACTTCA----------------------------------------
Ak-1_2    TTTACGTACTTTTGACTTCA----------------------------------------
Krot-0_2  TTTACGTACTTTTGACTTCA----------------------------------------
Bor-1_2   TTTACGTATTTTTGACTTCA----------------------------------------
Bay-0     TTTACGTACTTTTGACTTCA----------------------------------------

Col-0     TTTTTTTTCGCCATTTAATCACTCACTTTTGACTTCATATATGTATTCTAGTAGTCGTAC
Sha       TTTTTTTTCGCCATTTAGTCACTCACTTTTGACTTCATATATGTATTCTAGTAGTCGTAC
Blh-1_2   TTTTTTTTCGCCATTTAATCACTCACTTCTGACTTCATATATGTATTCTAGTAGTCGTAC
Aa-0_2    ----------------------------GACTTCATATATGTATTCTAGTAGTCGTAC
Hel-3_2   ----------------------------T-----------AAAATCTAGTGATGAT--
Ak-1_2    ----------------------------T-----------AAAATCTAGTGATGAT--
Bor-1_2   ----------------------------T-----------AAAATCTAGTGATGAT--
Krot-0_2  ----------------------------T-----------AAAATCTAGTGATGAT--
Bay-0     ----------------------------T-----------AAAATCTAGTGATGAT--

Col-0     CCAATTGTTGTAATGGGCCTTCAAAGGCCTCTTGTAGTAATGGGCCTATTAATTTGTAAA
Sha       CCAATTGTTGTAATGGGCCTTCAAAGGCCTCTTGTAGTAATGGGCCTATTAATTTGTAAA
Blh-1_2   CCAATTGTTGTAATGGGCCTTCAAAGGCCTCTTGTAGTAATGGGCCTATTAATTTGTAAA
Aa-0_2    CCAATTGTTGTAATGGGCCTTCAAAGGCCTCTTGTAGTAATGGGCCTATTAATTTGTAAA
Hel-3_2   CCAATTGTAGTAATGGGCCTTCAAAGGCATCTTGTGGTAATGGGCCTATTAATTTGTAAA
Ak-1_2    CCAATTGTAGTAATGGGCCTTCAAAGGCATCTTGTGGTAATGGGCCTATTAATTTGTAAA
Krot-0_2  CCAATTGTTGTAATGGGCCTTCAAAGGCCTCTTGTAGTAATGGGCCTATTAATTTGTAAA
Bor-1_2   CCAATTGTTGTAATGGGCCTTCAAAGGCCTCTTGTAGTAATGGGCCTATTAATTTGTAAA
Bay-0     CCAATTGTAGTAATGGGCCTTCAAAGGCATCTTGTGGTAATGGGCCTATTAATTTGTAAA

Col-0     -TAACCAAACCCGAAGTTTCGTCTCCTGAGTCATTTCACGACTGATGTTTTTCTTCTCCG
Sha       -TAACCAAACCCGAAGTTTCGTCTCCTGAGTCATTTCACGACTGATGTTTTTCTTCTCCG
Blh-1_2   -TAACCAAACCCGAAGTTTCGTCTCCTGAGTCATTTCACGACTGATGTTTTTCTTCTCCG
Aa-0_2    -TAACCAAACCCGAAGTTTCGTCTCCTGAGTCATTTCACGACTGATGTTTTTCTTCTCCG
Hel-3_2   CTAACCAAACCCGAAGTTTCGTCTCCTGAGTCATTTCACGACTGATGTCTTTCTTCTCCG
Ak-1_2    CTAACCAAACCCGAAGTTTCGTCTCCTGAGTCATTTCACGACTGATGTCTTTCTTCTCCG
Krot-0_2  CTAACCAAACCCGAAGTTTCGTCTCCTGAGTCATTTCACGACTGATGTCTTTCTTCTCCG
Bor-1_2   CTAACCAAACCCGAAGTTTCGTCTCCTGAGTCATTTCACGACTGATGTCTTTCTTCTCCG
Bay-0     CTAACCAAACCCGAAGTTTCGTCTCCTGAGTCATTTCACGACTGATGTCTTTCTTCTCCG
```

Figure S3.2 ***Sequence comparison of the two deletions in ATPS1 gene***

Identical residues are in blue, coding region is marked with asterisks. The results were obtained by Anna Koprivova and the figure comes from Koprivova et al. (2013).

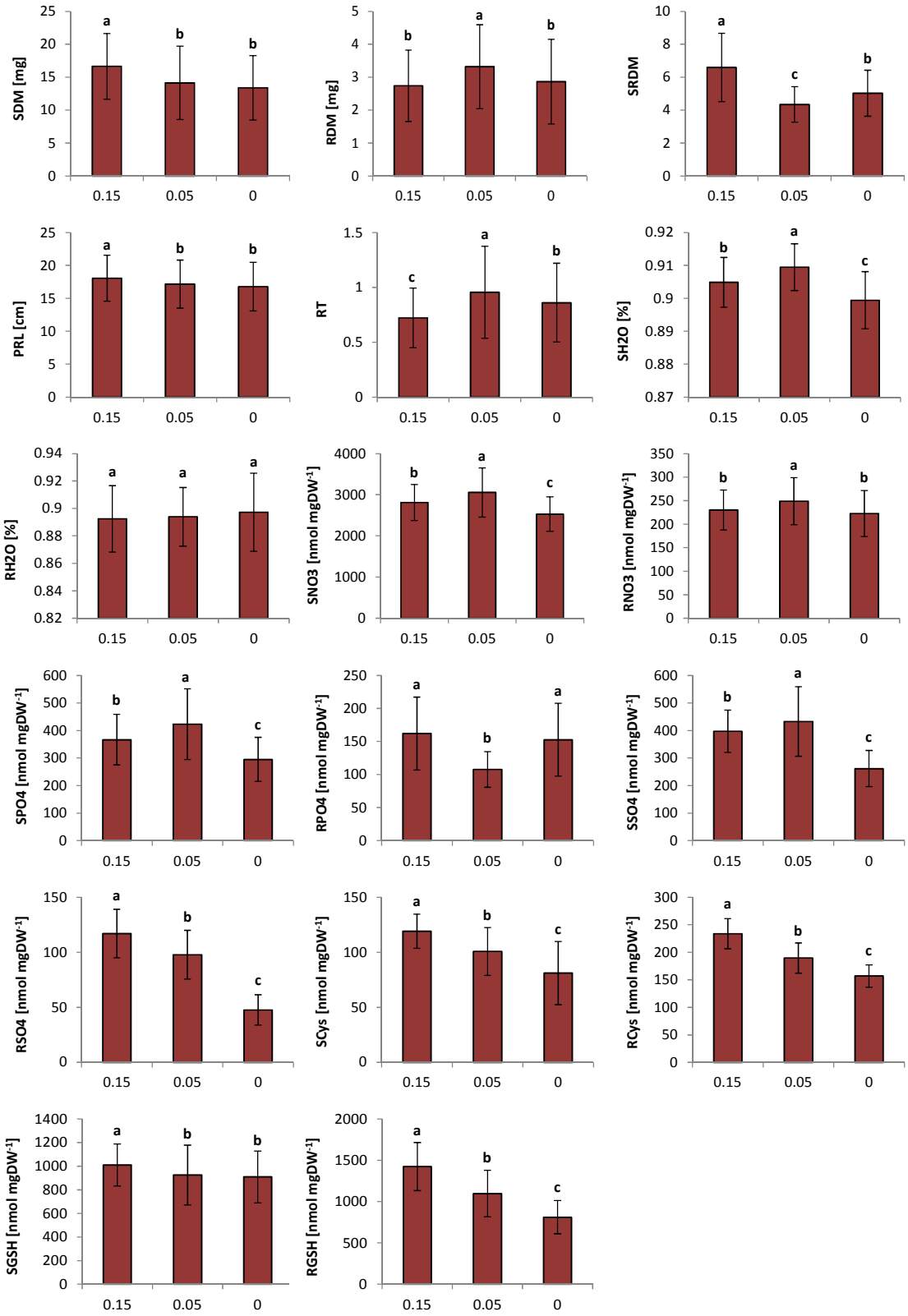
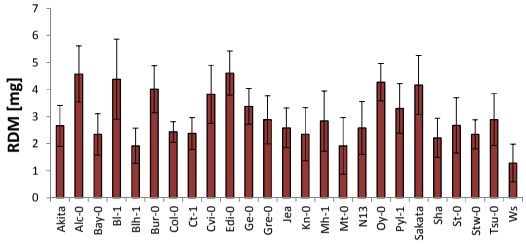
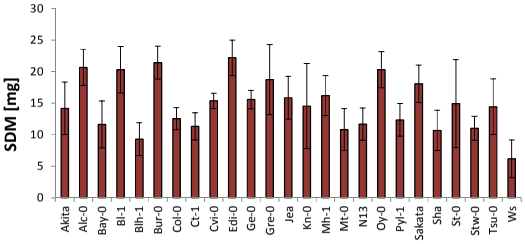


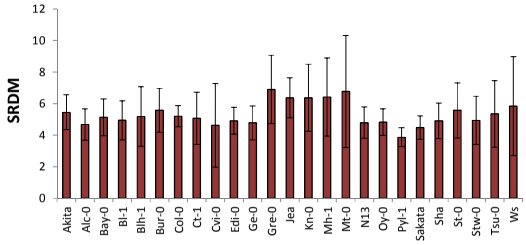
Figure S4.1 ***Global nutrition effect***

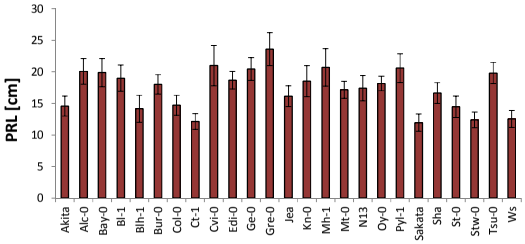
Arabidopsis response to different nutrition regimes shown as average of all 25 accessions for each trait in normal sulfate condition (0.15), sulfate limitation (0.05), and sulfate starvation (0). Different letters indicate values significantly different at P-value < 0.05 obtained from ANOVA with the Newman-Keuls (SNK) multiple comparison grouping method. The error bars correspond to standard deviation. Abbreviations: shoot dry matter (SDM), root dry matter (RDM), shoot to root

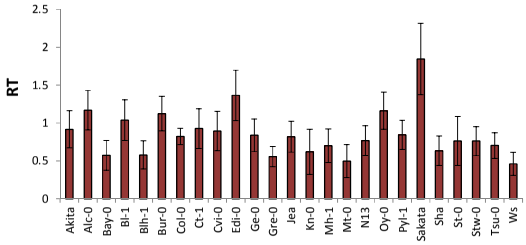
ratio of dry matter (SRDM), primary root length (PRL), root thickness (RT), shoot water concentration (SH2O), and root water concentration (RH2O), shoot nitrate concentration (SNO3), root nitrate concentration (RNO3), shoot phosphate concentration (SPO4) root phosphate concentration (RPO4), shoot sulfate concentration (SSO4), root sulfate concentration (RSO4), shoot cysteine concentration (SCys), root cysteine concentration (RCys), shoot glutathione concentration (SGSH), root glutathione concentration (RGSH).

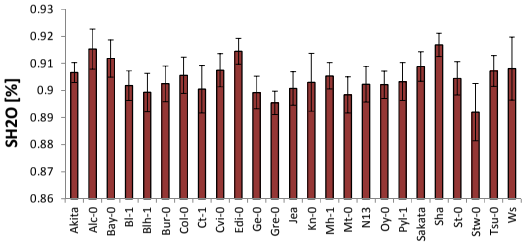

RDM [mg]
0 1 2 3 4 5 6 7
Akita Alc-0 Bay-0 Bl-1 Blh-1 Bur-0 Col-0 Ct-1 Cvi-0 Edi-0 Ge-0 Gre-0 Jea Kn-0 Mh-1 Mt-0 N13 Oy-0 Pyl-1 Sakata Sha St-0 Stw-0 Tsu-0 Ws

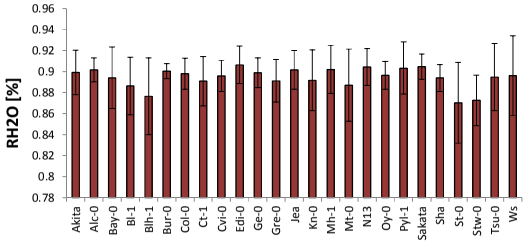

SDM [mg]
0 5 10 15 20 25 30
Akita Alc-0 Bay-0 Bl-1 Blh-1 Bur-0 Col-0 Ct-1 Cvi-0 Edi-0 Ge-0 Gre-0 Jea Kn-0 Mh-1 Mt-0 N13 Oy-0 Pyl-1 Sakata Sha St-0 Stw-0 Tsu-0 Ws

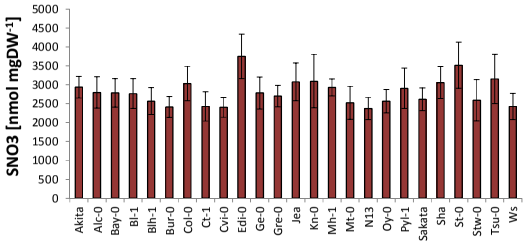

SRDM
0 2 4 6 8 10 12
Akita Alc-0 Bay-0 Bl-1 Blh-1 Bur-0 Col-0 Ct-1 Cvi-0 Edi-0 Ge-0 Gre-0 Jea Kn-0 Mh-1 Mt-0 N13 Oy-0 Pyl-1 Sakata Sha St-0 Stw-0 Tsu-0 Ws

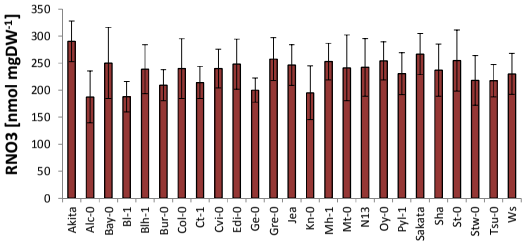

PRL [cm]
0 5 10 15 20 25 30
Akita Alc-0 Bay-0 Bl-1 Blh-1 Bur-0 Col-0 Ct-1 Cvi-0 Edi-0 Ge-0 Gre-0 Jea Kn-0 Mh-1 Mt-0 N13 Oy-0 Pyl-1 Sakata Sha St-0 Stw-0 Tsu-0 Ws

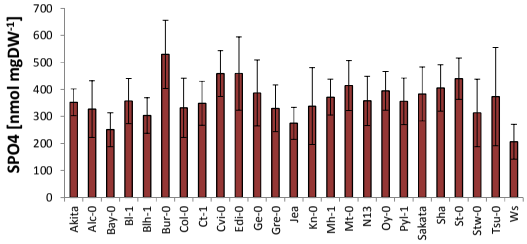

RT
0 0.5 1 1.5 2 2.5
Akita Alc-0 Bay-0 Bl-1 Blh-1 Bur-0 Col-0 Ct-1 Cvi-0 Edi-0 Ge-0 Gre-0 Jea Kn-0 Mh-1 Mt-0 N13 Oy-0 Pyl-1 Sakata Sha St-0 Stw-0 Tsu-0 Ws

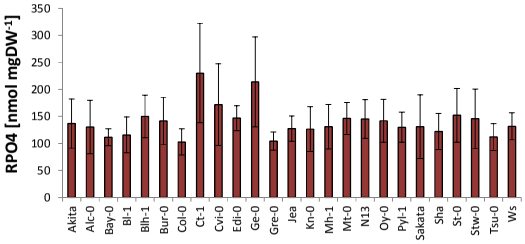

SH2O [%]
0.86 0.87 0.88 0.89 0.9 0.91 0.92 0.93
Akita Alc-0 Bay-0 Bl-1 Blh-1 Bur-0 Col-0 Ct-1 Cvi-0 Edi-0 Ge-0 Gre-0 Jea Kn-0 Mh-1 Mt-0 N13 Oy-0 Pyl-1 Sakata Sha St-0 Stw-0 Tsu-0 Ws


RH2O [%]
0.78 0.8 0.82 0.84 0.86 0.88 0.9 0.92 0.94 0.96
Akita Alc-0 Bay-0 Bl-1 Blh-1 Bur-0 Col-0 Ct-1 Cvi-0 Edi-0 Ge-0 Gre-0 Jea Kn-0 Mh-1 Mt-0 N13 Oy-0 Pyl-1 Sakata Sha St-0 Stw-0 Tsu-0 Ws


SNO3 [nmol mgDW-1]
0 500 1000 1500 2000 2500 3000 3500 4000 4500 5000
Akita Alc-0 Bay-0 Bl-1 Blh-1 Bur-0 Col-0 Ct-1 Cvi-0 Edi-0 Ge-0 Gre-0 Jea Kn-0 Mh-1 Mt-0 N13 Oy-0 Pyl-1 Sakata Sha St-0 Stw-0 Tsu-0 Ws


RNO3 [nmol mgDW-1]
0 50 100 150 200 250 300 350
Akita Alc-0 Bay-0 Bl-1 Blh-1 Bur-0 Col-0 Ct-1 Cvi-0 Edi-0 Ge-0 Gre-0 Jea Kn-0 Mh-1 Mt-0 N13 Oy-0 Pyl-1 Sakata Sha St-0 Stw-0 Tsu-0 Ws


SPO4 [nmol mgDW-1]
0 100 200 300 400 500 600 700
Akita Alc-0 Bay-0 Bl-1 Blh-1 Bur-0 Col-0 Ct-1 Cvi-0 Edi-0 Ge-0 Gre-0 Jea Kn-0 Mh-1 Mt-0 N13 Oy-0 Pyl-1 Sakata Sha St-0 Stw-0 Tsu-0 Ws


RPO4 [nmol mgDW-1]
0 50 100 150 200 250 300 350
Akita Alc-0 Bay-0 Bl-1 Blh-1 Bur-0 Col-0 Ct-1 Cvi-0 Edi-0 Ge-0 Gre-0 Jea Kn-0 Mh-1 Mt-0 N13 Oy-0 Pyl-1 Sakata Sha St-0 Stw-0 Tsu-0 Ws

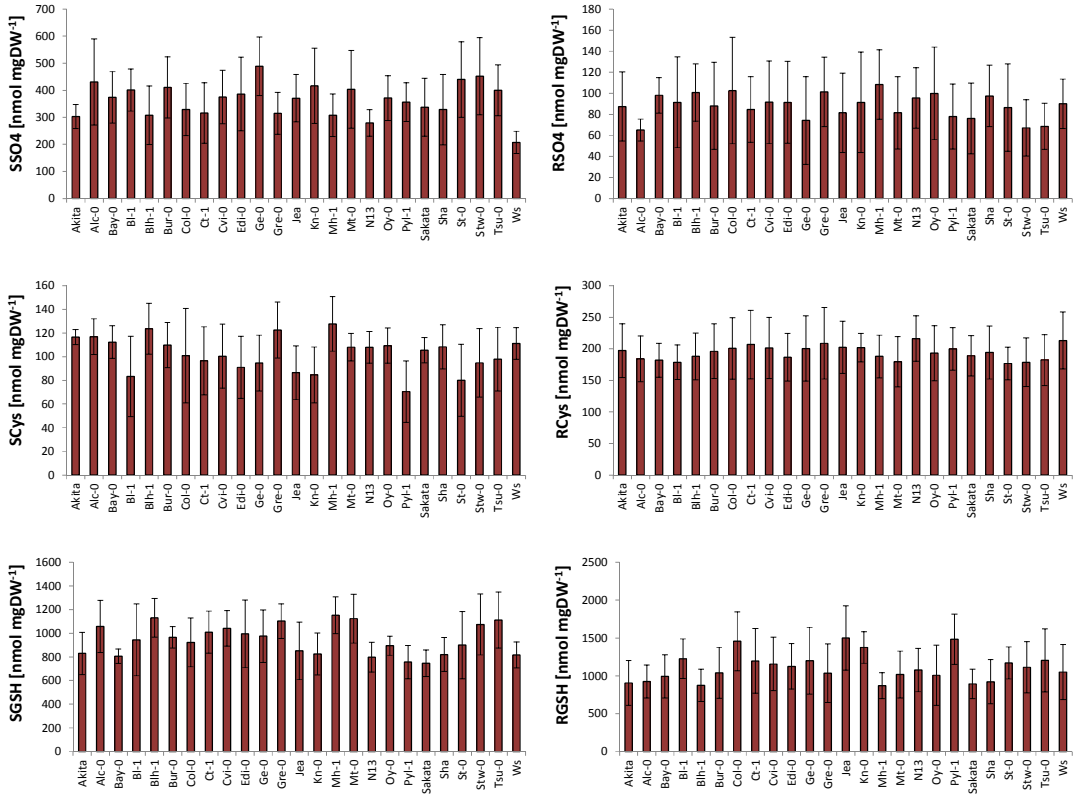


Figure S4.2 ***Global genetic effect***

The difference between arabidopsis accessions is shown as average of the values for all three nutrition regimes for each trait. The error bars correspond to standard deviation. Abbreviations: shoot dry matter (SDM), root dry matter (RDM), shoot to root ratio of dry matter (SRDM), primary root length (PRL), root thickness (RT), shoot water concentration (SH2O), and root water concentration (RH2O), shoot nitrate concentration (SNO3), root nitrate concentration (RNO3), shoot phosphate concentration (SPO4) root phosphate concentration (RPO4), shoot sulfate concentration (SSO4), root sulfate concentration (RSO4), shoot cysteine concentration (SCys), root cysteine concentration (RCys), shoot glutathione concentration (SGSH), root glutathione concentration (RGSH).

Table S4.1 ***Percentage of variation among 25 arabidopsis accessions***

	Percentage of variation			
Trait	**Nutrition**	**Genotype**	**Nutrition x Genotype**	**Residual**
SFM	8.48	59.77	10.80	20.96
P-value	< 0.0001	< 0.0001	< 0.0001	
RFM	4.08	57.43	10.11	28.38
P-value	< 0.0001	< 0.0001	0.027	
SRFM	41.89	22.40	ns	35.67
P-value	< 0.0001	< 0.0001	0.086	
SDM	6.89	58.61	11.77	22.73
P-value	< 0.0001	< 0.0001	< 0.0001	
RDM	4.15	53.52	14.45	27.84
P-value	< 0.0001	< 0.0001	< 0.0001	
SRDM	25.97	15.61	19.27	39.13
P-value	< 0.0001	< 0.0001	0.0002	
RL	1.19	73.38	6.66	18.77
P-value	0.002	< 0.0001	0.028	
RT	6.75	59.42	9.50	23.36
P-value	< 0.0001	< 0.0001	0.005	
SH20	24.04	47.83	ns	26.22
P-value	< 0.0001	< 0.0001	0.057	
RH20	ns	14.56	31.45	53.90
P-value	0.262	0.001	< 0.0001	
SNO3	16.14	38.01	27.74	18.11
P-value	< 0.0001	< 0.0001	< 0.0001	
RNO3	5.43	26.58	41.12	26.87
P-value	< 0.0001	< 0.0001	< 0.0001	
SPO4	21.65	33.30	25.08	19.96
P-value	< 0.0001	< 0.0001	< 0.0001	
RPO4	18.81	31.15	33.42	16.62
P-value	< 0.0001	< 0.0001	< 0.0001	
SSO4	37.56	24.98	21.99	15.46
P-value	< 0.0001	< 0.0001	< 0.0001	
RSO4	67.75	10.28	11.82	10.15
P-value	< 0.0001	< 0.0001	< 0.0001	
SCys	27.66	31.41	20.45	20.48
P-value	< 0.0001	< 0.0001	< 0.0001	
RCys	60.59	7.42	17.53	14.46
P-value	< 0.0001	< 0.0001	< 0.0001	
SGSH	3.39	29.74	47.29	19.58
P-value	< 0.0001	< 0.0001	< 0.0001	
RGSH	46.86	26.13	12.07	14.94
P-value	< 0.0001	< 0.0001	< 0.0001	

The percentage variation explained by each factor among 2 arabidopsis accessions was computed from the sum of squares obtained from ANOVA which was also used for the significance testing. Traits with P-value<0.05 were considered as significantly altered.

Table S4.2 ***Percentage of variation among five arabidopsis accessions***

	Percantage of variation						
Trait	**Genotype**	**N nutrition**	**S nutrition**	**N nutrition x S nutrition**	**Genotype x N nutrition**	**Genotype x S nutrition**	**Residual**
NO3	23.72	41.64	5.44	ns	16.54	2.99	9.67
P-value	<.0001	<.0001	<.0001	0.2164	<.0001	0.0106	
PO4	78.76	4.11	ns	0.78	1.91	5.82	8.62
P-value	<.0001	<.0001	0.1602	0.0144	0.0068	<.0001	
SO4	ns	2.84	44.98	1.76	14.78	11.41	24.24
P-value	0.7637	0.0059	<.0001	0.0285	<.0001	0.0005	
Cys	34.88	19.05	3.35	ns	8.40	9.91	24.41
P-value	<.0001	<.0001	0.0097	0.1548	0.0002	0.0013	
GSH	0.33	92.46	2.41	2.86	0.50	0.41	1.04
P-value	0.0004	<.0001	<.0001	<.0001	<.0001	0.0011	
ITh	11.45	18.69	32.65	2.80	ns	ns	34.41
P-value	0.0005	<.0001	<.0001	0.0206	0.5669	0.3129	
IP	3.44	63.18	ns	ns	ns	ns	33.37
P-value	0.0034	<.0001	0.2688	0.3443	0.1064	0.7019	
TTS	25.43	8.88	38.24	ns	ns	ns	27.46
P-value	<.0001	<.0001	<.0001	0.2169	0.187	0.1407	
ASF	8.17	35.27	25.06	ns	ns	ns	31.51
P-value	0.0032	<.0001	<.0001	0.0852	0.418	0.4682	
RSF	15.75	40.80	8.68	8.21	ns	ns	26.56
P-value	<.0001	<.0001	<.0001	<.0001	0.712	0.3505	
TU	17.10	13.19	9.75	ns	11.35	ns	48.61
P-value	0.0002	<.0001	0.0013	0.3322	0.0037	0.1092	

The percentage variation explained by each factor among five arabidopsis accessions was computed from the sum of squares obtained from ANOVA which was also used for the significance testing. Traits with P-value<0.05 were considered as significantly altered.

Table S5.1 ***Borevitz collection accessions***

Accession	Alternative name	Country	Comment
Ag.0	n.a.	FRA	
Alc.0	n.a.	ESP	
ALL1.2	n.a.	FRA	
ALL1.3	n.a.	FRA	
An.1	n.a.	BEL	
App1.16	n.a.	SWE	
Ba1.2	n.a.	SWE	
Bay.0	n.a.	GER	didn't grow
Belmonte.4.94	n.a.	ITA	
Bg.2	n.a.	USA	
Bla.1	n.a.	ESP	
Blh.1	n.a.	CZE	
Bor.1	n.a.	CZE	
Bor.4	n.a.	CZE	
Br.0	n.a.	CZE	
Bro1.6	n.a.	SWE	
Bu.0	n.a.	GER	
BUI	n.a.	FRA	
Bur.0	n.a.	IRL	
C24	n.a.	POR	
CAM.16	n.a.	FRA	
CAM.61	n.a.	FRA	
Can.0	n.a.	ESP	
Cen.0	n.a.	FRA	
CIBC.17	n.a.	UK	
CLE.6	n.a.	FRA	
Col.0	n.a.	USA	
CS28007	Aa-0	GER	
CS28013	Alst-1	UK	
CS28014	Amel-1	NED	
CS28017	An-2	BEL	
CS28018	Ang-0	BEL	
CS28049	Ann-1	FRA	
CS28053	Ba-1	UK	
CS28054	Baa-1	NED	
CS28063	Be-1	GER	
CS28064	Benk-1	NED	
CS28090	Blh-2	CZE	
CS28091	Boot-1	UK	
CS28097	Bs-2	SUI	
CS28099	Bsch-0	GER	
CS28108	Bu-8	GER	
CS28128	Ca-0	GER	
CS28133	Cha-0	SUI	
CS28135	Chat-1	FRA	
CS28140	CIBC-2	UK	
CS28141	CIBC-4	UK	
CS28142	CIBC-5	UK	
CS28158	Cit-0	FRA	
CS28163	Co-2	POR	

Accession	Alternative name	Country	Comment
CS28165	Co-4	POR	
CS28181	CSHL-5	USA	
CS28193	Com-1	FRA	
CS28200	Da-0	GER	
CS28201	Da(1)-12	CZE	
CS28202	Db-0	GER	
CS28208	Di-1	FRA	
CS28210	Do-0	GER	
CS28214	Dra-2	CZE	
CS28217	Ede-1	NED	
CS28236	Ep-0	GER	
CS28241	Es-0	FIN	
CS28243	Est-0	RUS	
CS28252	Fi-1	GER	
CS28268	Fr-4	GER	
CS28274	Ga-2	GER	
CS28277	Ge-1	SUI	
CS28279	Gel-1	NED	
CS28280	Gie-0	GER	
CS28282	Go-0	GER	
CS28326	Gr-5	AUT	
CS28332	Gu-1	GER	
CS28336	Ha-0	GER	
CS28344	Hey-1	NED	
CS28345	Hh-0	GER	
CS28350	Hn-0	GER	
CS28364	Je-0	GER	
CS28369	Jl-3	CZE	
CS28373	Jm-1	CZE	
CS28382	Kelsterbach-2	GER	
CS28394	Kl-5	GER	
CS28395	Kn-0	LTU	
CS28407	KNO-11	USA	
CS28419	Kr-0	GER	
CS28420	Kro-0	GER	
CS28423	unknown	GER	
CS28454	Li-3	GER	
CS28457	Li-5:2	GER	
CS28459	Li-6	GER	
CS28461	Li-7	GER	
CS28490	Mc-0	UK	
CS28492	Mh-0	POL	
CS28495	Mnz-0	GER	didn't germinate
CS28510	N4	RUS	
CS28513	N7	RUS	
CS28527	Nc-1	FRA	
CS28550	NFC-20	UK	
CS28564	No-0	GER	
CS28568	Nok-1	NED	
CS28573	Nw-0	GER	

Accession	Alternative name	Country	Comment
CS28575	Nw-2	GER	
CS28578	Nz1	NZL	
CS28580	Ob-1	GER	
CS28583	Old-1	GER	
CS28587	Or-0	GER	
CS28595	Pa-2	ITA	
CS28610	PHW-10	UK	
CS28613	PHW-13	UK	
CS28614	PHW-14	UK	
CS28620	PHW-20	UK	
CS28622	PHW-22	UK	
CS28628	PHW-28	UK	
CS28640	Pla-0	ESP	
CS28645	Pn-0	FRA	
CS28650	Pog-0	CAN	
CS28651	Pr-0	GER	
CS28663	Pu2-24	CZE	
CS28685	Rhen-1	NED	
CS28692	Rou-0	FRA	
CS28713	RRS-7	USA	
CS28720	S96	Unknown	
CS28724	Sapporo-0	JPN	
CS28725	Sav-0	CZE	
CS28729	Sei-0	ITA	
CS28734	Sh-0	GER	
CS28739	Si-0	GER	
CS28743	Sp-0	GER	
CS28750	Ste-0	GER	
CS28759	Ting-1	SWE	
CS28760	Tiv-1	ITA	
CS28779	Tscha-1	AUT	
CS28780	Tsu-0	JPN	
CS28786	Ty-0	UK	
CS28787	Uk-1	GER	
CS28795	Utrecht	NED	
CS28800	Ven-1	NED	
CS28804	Wa-1	POL	
CS28808	Wag-3	NED	
CS28809	Wag-4	NED	
CS28810	Wag-5	NED	
CS28812	WAR	USA	
CS28814	Wc-2	GER	
CS28822	Wl-0	GER	
CS28823	Ws	RUS	
CS28833	Wt-3	GER	
CS28847	Zu-1	SUI	
CS28849	Ors-2	ROU	
Ct.1	n.a.	ITA	didn't germinate
CUR.3	n.a.	FRA	
Cvi.0	n.a.	CPV	

Accession	Alternative name	Country	Comment
Dra3.1	n.a.	SWE	seeds not available
DraII.1	n.a.	CZE	
DraIV.1.14	n.a.	CZE	
DraIV.1.5	n.a.	CZE	
DraIV.1.7	n.a.	CZE	
DraIV.6.16	n.a.	CZE	
DraIV.6.35	n.a.	CZE	
Duk	n.a.	CZE	
Eden.2	n.a.	SWE	
Edi.0	n.a.	UK	
Est.1	n.a.	RUS	
Fab.2	n.a.	SWE	
Fei.0	n.a.	POR	
Fja1.1	n.a.	SWE	didn't germinate
Fja1.2	n.a.	SWE	
Fja1.5	n.a.	SWE	
Ga.0	n.a.	GER	
Gd.1	n.a.	GER	
Hi.0	n.a.	NED	
Hod	n.a.	CZE	
Hov4.1	n.a.	SWE	
Hovdala.2	n.a.	SWE	
HR.5	n.a.	UK	
Hs.0	n.a.	GER	
HSm	n.a.	CZE	
In.0	n.a.	AUT	
JEA	n.a.	FRA	
Ka.0	n.a.	AUT	
Kas.2	n.a.	IND	
KBS.Mac.8	n.a.	USA	
Kelsterbach.4	n.a.	GER	
Kin.0	n.a.	USA	
Kno.18	n.a.	USA	
Koln	n.a.	GER	
Kulturen.1	n.a.	SWE	
LAC.3	n.a.	FRA	didn't germinate
LAC.5	n.a.	FRA	
Lc.0	n.a.	UK	
LDV.25	n.a.	FRA	didn't germinate
LDV.34	n.a.	FRA	
LDV.58	n.a.	FRA	
Ler.1	n.a.	GER	
LI.OF.095	n.a.	USA	
Liarum	n.a.	SWE	
Lillo.1	n.a.	SWE	didn't germinate
Lip.0	n.a.	POL	
Lis.1	n.a.	SWE	
Lis.2	n.a.	SWE	
Lisse	n.a.	NED	
LL.0	n.a.	ESP	

Accession	Alternative name	Country	Comment
Lm.2	n.a.	FRA	
Lom1.1	n.a.	SWE	
Lov.5	n.a.	SWE	
Lp2.2	n.a.	CZE	
Lp2.6	n.a.	CZE	
Lund	n.a.	SWE	
Lz.0	n.a.	FRA	
Map.42	n.a.	USA	
MIB.15	n.a.	FRA	
MIB.22	n.a.	FRA	
MIB.28	n.a.	FRA	
MIB.84	n.a.	FRA	
MNF.Che.2	n.a.	USA	
MNF.Jac.32	n.a.	USA	
MNF.Pot.48	n.a.	USA	
MNF.Pot.68	n.a.	USA	
MOG.37	n.a.	FRA	
Mr.0	n.a.	ITA	
Mrk.0	n.a.	GER	
Mt.0	n.a.	LIB	
Mz.0	n.a.	GER	
N13	n.a.	RUS	
Na.1	n.a.	FRA	
NC.6	n.a.	USA	
Ost.0	n.a.	SWE	
Oy.0	n.a.	NOR	
Pa.1	n.a.	ITA	
PAR.3	n.a.	FRA	
PAR.4	n.a.	FRA	
PAR.5	n.a.	FRA	
Paw.3	n.a.	USA	
Pent.1	n.a.	USA	
Per.1	n.a.	RUS	
Petergof	n.a.	RUS	
PHW.34	n.a.	FRA	didn't germinate
Pna.17	n.a.	USA	
Pro.0	n.a.	ESP	
Pu2.23	n.a.	CZE	
Ra.0	n.a.	FRA	
Rak.2	n.a.	CZE	
Ren.1	n.a.	FRA	
Rev.2	n.a.	SWE	
Rmx.A180	n.a.	USA	
ROM.1	n.a.	FRA	
RRS.10	n.a.	USA	
Rsch.4	n.a.	RUS	
Sanna.2	n.a.	SWE	
Sap.0	n.a.	CZE	
Sav.0	n.a.	CZE	
Se.0	n.a.	ESP	

Accession	Alternative name	Country	Comment
Shahdara	n.a.	TJK	
SLSP.30	n.a.	USA	
Sparta.1	n.a.	SWE	
Sq.8	n.a.	UK	didn't germinate
St.0	n.a.	SWE	
Ste.3	n.a.	USA	
T1040	n.a.	SWE	
T1060	n.a.	SWE	
T1080	n.a.	SWE	
T1110	n.a.	SWE	
T510	n.a.	SWE	
T540	n.a.	SWE	
T620	n.a.	SWE	
T690	n.a.	SWE	
Ta.0	n.a.	CZE	
TAD.01	n.a.	SWE	
Tamm.2	n.a.	FIN	
TDr.18	n.a.	SWE	
TDr.3	n.a.	SWE	
TDr.8	n.a.	SWE	
Tomegap.2	n.a.	SWE	
Tottarp.2	n.a.	SWE	
TOU.A1.115	n.a.	FRA	
TOU.A1.116	n.a.	FRA	
TOU.A1.12	n.a.	FRA	
TOU.A1.43	n.a.	FRA	didn't germinate
TOU.A1.62	n.a.	FRA	
TOU.A1.67	n.a.	FRA	didn't germinate
TOU.A1.96	n.a.	FRA	
TOU.C.3	n.a.	FRA	
TOU.I.6	n.a.	FRA	
TOU.J.3	n.a.	FRA	
TOU.K.3	n.a.	FRA	didn't germinate
Ts.1	n.a.	ESP	
UduI.1.34	n.a.	CZE	
UKID101	n.a.	UK	didn't germinate
UKID22	n.a.	UK	
UKID37	n.a.	UK	
UKID48	n.a.	UK	
UKID80	n.a.	UK	didn't germinate
UKNW06.059	n.a.	UK	
UKNW06.060	n.a.	UK	didn't germinate
UKNW06.386	n.a.	UK	
UKNW06.436	n.a.	UK	didn't germinate
UKNW06.460	n.a.	UK	
UKSE06.062	n.a.	UK	
UKSE06.192	n.a.	UK	didn't germinate
UKSE06.272	n.a.	UK	didn't germinate
UKSE06.278	n.a.	UK	
UKSE06.349	n.a.	UK	didn't germinate

Accession	Alternative name	Country	Comment
UKSE06.351	n.a.	UK	
UKSE06.414	n.a.	UK	
UKSE06.429	n.a.	UK	
UKSE06.466	n.a.	UK	
UKSE06.482	n.a.	UK	didn't germinate
UKSE06.520	n.a.	UK	didn't germinate
UKSE06.628	n.a.	UK	
UKSW06.202	n.a.	UK	
Ull2.5	n.a.	SWE	
Ull3.4	n.a.	SWE	
Uod.7	n.a.	AUT	
Var2.1	n.a.	SWE	
VOU.1	n.a.	FRA	didn't germinate
VOU.2	n.a.	FRA	didn't germinate
Wil.1	n.a.	LTU	
Ws.0	n.a.	RUS	
Wt.5	n.a.	GER	
X11ME1.32	n.a.	USA	didn't germinate
X11PNA4.101	n.a.	USA	didn't germinate
X328PNA054	n.a.	USA	
X627ME.4Y1	n.a.	USA	
Yo.0	n.a.	USA	didn't germinate
Zdr.6	n.a.	CZE	
ZdrI.2.24	n.a.	CZE	
ZdrI.2.25	n.a.	CZE	

Table S5.2 ***Anion concentration***

Accession	Anion accumulation [nmol mg^{-1} FW]					
	NO_3^{1-}	SD	PO_4^{3-}	SD	SO_4^{2-}	SD
Ag.0	130.2	11.5	13.3	1.2	-	-
Alc.0	114.5	11.4	13.2	1.3	-	-
ALL1.2	159.2	10.0	16.7	3.7	12.7	1.8
ALL1.3	136.0	10.4	12.4	2.3	15.3	1.3
An.1	119.2	11.7	14.3	1.3	-	-
App1.16	125.3	15.6	27.4	2.4	10.7	1.6
Ba-1.2	147.0	1.9	14.6	1.4	13.7	1.6
Belmonte.4.94	127.4	10.1	13.5	4.5	-	-
Bg.2	116.0	14.8	-	-	-	-
Bla.1	138.8	13.6	24.1	2.1	20.5	2.0
Blh.1	144.9	7.1	15.0	2.4	18.7	2.5
Bor.1	131.4	14.7	16.2	1.7	16.1	0.8
Bor.4	103.5	8.6	-	-	17.3	2.4
Br.0	112.5	15.7	-	-	17.1	3.4
Brö1-6	128.4	16.9	9.7	0.9	10.0	1.6
Bu.0	124.9	8.8	10.8	1.0	10.6	1.6
Bur.0	-	-	-	-	13.6	3.2
C24	98.2	4.1	6.7	1.3	11.9	2.4
CAM.16	115.2	3.0	14.1	2.0	9.9	0.9
CAM.61	118.4	2.9	10.9	1.5	9.0	2.0
Can.0	126.9	22.0	-	-	32.8	6.6
Cen.0	-	-	-	-	15.1	1.8
CIBC.17	124.0	10.1	17.8	1.6	16.1	0.9
Col.0	137.2	21.6	14.3	4.4	13.7	3.7
Aa-0	139.0	8.6	15.5	1.9	18.3	1.0
Alst-1	133.4	12.0	13.1	3.0	8.1	1.5
Amel-1	119.8	5.6	21.9	0.9	10.4	2.0
An-2	135.3	3.5	18.1	2.6	-	-
Ang-0	122.4	9.2	15.4	1.6	8.3	0.5
Ann-1	138.7	12.2	18.3	1.4	9.7	0.5
Ba-1	145.1	1.9	15.4	1.4	6.7	1.6
Baa-1	175.4	15.9	12.8	1.5	7.5	0.9
Benk-1	143.4	11.6	20.8	3.0	11.3	0.9
Blh-2	143.5	12.4	16.1	1.5	11.9	0.8
Boot-1	131.2	9.3	21.3	1.1	10.2	0.7
Bs-2	144.7	9.9	-	-	12.1	0.9
Bsch-0	120.8	1.4	16.7	2.2	16.2	2.3
Bu-8	148.7	16.7	14.0	1.8	-	-
Ca-0	136.5	4.2	15.4	0.9	10.1	0.4
Cha-0	116.7	8.4	31.2	1.2	22.7	0.5
Chat-1	153.0	13.3	20.0	0.7	13.3	1.1
CIBC-2	164.5	2.6	-	-	10.3	0.8
CIBC-4	145.0	6.4	24.1	1.9	11.9	1.1
CIBC-5	163.4	12.2	19.0	1.8	10.7	1.4
Cit-0	164.9	8.3	18.0	1.6	-	-
Co-2	157.0	4.1	20.4	1.7	-	-

Accession	Anion accumulation [nmol mg^{-1} FW]					
	NO_3^{1-}	SD	PO_4^{3-}	SD	SO_4^{2-}	SD
Co-4	134.6	6.1	14.0	0.5	13.4	1.2
CSHL-5	143.4	10.5	11.5	1.2	12.0	1.0
Com-1	131.6	4.1	17.2	1.4	15.5	1.0
Da-0	148.1	2.2	17.3	1.5	-	-
Da(1)-12	140.1	11.1	12.8	1.7	15.0	1.7
Db-0	140.4	14.1	13.7	1.0	14.6	1.6
Di-1	99.7	10.8	-	-	21.1	1.8
Do-0	166.4	5.2	15.7	1.7	8.8	0.5
Dra-2	145.5	11.8	17.6	2.1	10.8	1.6
Ede-1	161.4	6.1	16.0	3.2	9.9	0.5
Ep-0	154.1	7.9	15.0	2.4	11.2	0.7
Es-0	125.8	3.6	14.5	2.3	12.4	2.6
Est-0	122.8	4.3	12.3	1.4	-	1.2
Fi-1	152.4	11.4	11.1	0.8	11.2	1.3
Fr-4	124.2	5.8	15.0	1.2	-	-
Ga-2	120.9	6.9	17.5	1.0	-	-
Ge-1	114.7	3.4	21.4	1.3	18.4	1.4
Gel-1	112.2	7.0	20.9	0.9	19.8	2.1
Gie-0	116.1	10.6	14.1	0.8	14.1	1.6
Go-0	122.7	8.6	-	-	-	-
Gr-5	135.1	12.5	29.4	2.3	11.6	1.0
Gu-1	119.4	11.1	14.4	2.0	14.7	1.3
Ha-0	147.2	10.4	-	-	7.9	1.2
Hey-1	130.2	14.7	9.8	1.0	10.9	1.8
Hh-0	134.4	6.7	9.8	1.2	8.9	0.8
Hn-0	117.1	10.5	-	-	-	-
Je-0	134.7	12.7	20.5	1.0	17.3	0.8
Jl-3	120.2	11.7	18.9	1.3	14.5	0.5
Jm-1	-	-	8.8	0.2	14.8	2.5
Kelsterbach-2	151.7	9.1	14.3	0.2	9.6	0.8
Kl-5	143.4	9.5	-	-	13.7	0.3
Kn-0	137.7	14.0	13.7	1.3	12.7	0.9
KNO-11	226.0	18.4	-	-	25.6	3.2
Kr-0	201.6	8.0	41.8	3.8	31.8	2.0
Kro-0	188.0	14.8	-	-	27.7	1.0
Li-3	177.7	9.4	22.7	0.9	24.2	1.6
Li-5:2	135.9	15.5	20.1	1.6	10.4	1.0
Li-6	130.6	16.6	15.3	3.2	13.6	1.8
Li-7	155.3	10.3	12.9	1.6	15.1	2.5
Mc-0	129.0	14.3	20.9	0.9	-	-
Mh-0	125.2	7.5	9.0	2.5	13.0	1.5
Mnz-0	119.4	4.5	13.3	1.9	-	-
N4	124.7	10.9	-	-	15.7	1.5
N7	-	-	8.4	1.6	9.5	0.7
Nc-1	-	-	-	-	9.7	1.3
NFC-20	129.2	12.1	-	-	11.1	2.0

Accession	Anion accumulation [nmol mg^{-1} FW]					
	NO_3^{1-}	SD	PO_4^{3-}	SD	SO_4^{2-}	SD
Ob-1	138.2	8.9	-	-	-	-
Old-1	-	-	18.0	2.4	12.2	1.3
Or-0	155.7	13.3	12.3	0.7	13.1	2.0
Pa-2	172.9	15.6	-	-	-	-
PHW-10	160.2	14.8	-	-	-	-
PHW-13	147.3	12.0	12.5	2.2	9.9	1.4
PHW-14	156.6	9.8	17.3	1.8	9.8	1.1
PHW-20	101.1	8.9	21.4	3.3	17.9	1.3
PHW-22	157.8	8.8	14.7	2.9	10.6	1.6
PHW-28	125.3	11.1	35.1	4.9	-	-
PHW-31	150.4	9.5	-	-	-	-
PHW-35	132.2	7.2	24.6	3.2	14.8	2.7
PHW-36	152.8	12.1	22.1	2.0	11.3	1.4
PHW-37	150.2	10.9	22.8	2.6	13.7	0.7
Pla-0	147.5	5.0	18.7	0.5	15.2	1.5
Pn-0	130.8	5.3	21.5	1.7	17.5	0.9
Pr-0	163.0	17.1	16.7	0.5	9.5	0.8
Pu2-24	141.3	5.8	21.3	3.5	12.5	2.1
Rhen-1	135.8	13.5	-	-	14.4	4.0
Rou-0	130.3	9.6	22.5	1.8	19.2	2.7
RRS-7	125.9	12.6	19.3	3.1	13.7	1.3
S96	130.5	6.5	-	-	16.5	1.3
Sapporo-0	97.9	9.2	-	-	28.2	1.1
Sav-0	120.0	8.8	7.8	1.7	13.3	1.3
Sei-0	119.2	13.0	10.0	1.3	-	-
Sh-0	120.8	9.4	9.2	1.0	18.1	2.1
Si-0	-	-	11.7	1.6	12.7	0.7
Sp-0	138.7	12.1	9.1	2.0	14.9	0.4
Ste-0	123.9	5.9	33.4	1.8	14.0	1.0
Ting-1	119.1	15.0	19.5	1.2	15.8	1.3
Tiv-1	140.8	4.0	25.1	1.2	12.7	1.6
Tscha-1	-	-	16.7	0.9	15.5	1.9
Tsu-0	138.6	17.7	-	-	14.5	1.0
Ty-0	128.1	9.1	-	-	16.4	0.9
Uk-1	134.4	8.4	22.2	1.6	18.3	1.6
Utrecht	131.7	6.0	16.1	1.5	11.9	0.7
Ven-1	132.1	13.4	15.2	1.2	11.2	1.2
Wa-1	116.3	15.6	-	-	12.0	1.4
Wag-3	125.1	10.1	17.5	1.7	16.5	1.8
Wag-4	144.7	14.6	14.8	1.7	11.8	1.3
Wag-5	128.6	12.3	19.4	0.4	18.5	1.7
WAR	-	-	-	-	13.6	1.8
Wc-2	108.0	3.4	16.7	0.3	14.3	0.5
Wl-0	114.0	6.9	14.6	1.8	11.1	0.8
Ws	145.6	17.4	-	-	15.2	5.5
Wt-3	137.1	5.9	19.5	1.2	16.6	0.5

Accession	Anion accumulation [nmol mg^{-1} FW]					
	NO_3^{1-}	SD	PO_4^{3-}	SD	SO_4^{2-}	SD
DraIV.1.5	129.4	5.7	27.2	1.1	17.6	0.9
DraIV.6.16	138.4	15.9	15.5	2.0	10.2	1.3
DraIV.6.35	116.3	17.3	22.3	3.2	17.8	2.0
DraIV 1-14	157.5	10.7	25.5	1.6	19.6	1.4
Duk	143.9	13.9	16.6	1.8	14.7	1.2
Eden.2	112.0	7.7	30.7	2.3	26.7	1.8
Edi.0	152.0	13.5	-	-	12.6	1.4
Est.1	129.1	13.8	27.2	3.4	20.7	1.2
Fäb-2	116.4	14.2	23.4	3.4	25.5	6.6
Fja1.2	179.9	15.6	26.8	4.5	17.0	2.0
Fja1.5	158.4	17.3	-	-	13.9	2.4
Ga.0	134.5	18.5	32.8	4.2	-	-
Gd.1	119.1	16.9	16.9	1.0	12.6	1.4
Ge.0	170.8	8.8	20.6	1.1	22.6	2.6
Got.7	163.2	9.5	20.8	0.9	-	-
Gr.1	133.5	18.4	19.1	0.3	16.9	1.3
Gy.0	92.5	6.9	21.4	0.7	18.3	1.8
Hi.0	115.5	9.4	22.4	2.4	13.6	0.7
Hod	-	-	-	-	32.5	8.2
Hov4.1	114.7	14.5	-	-	-	-
Hovdala.2	118.9	14.1	19.4	0.7	19.5	1.9
HR.5	119.4	10.2	17.8	2.3	-	-
Hs.0	120.7	18.0	17.2	0.8	18.6	2.6
HSm	121.7	9.5	19.3	1.3	16.2	0.7
In.0	166.6	24.5	16.6	0.9	15.1	1.0
JEA	124.3	6.3	17.0	0.6	16.8	1.1
Ka.0	120.9	13.3	14.8	0.5	14.6	1.5
Kas.2	112.2	9.6	21.3	1.2	13.4	1.4
KBS.Mac.8	129.7	11.2	16.5	1.2	18.9	0.9
Kelsterbach.4	118.9	9.3	16.1	1.4	11.3	0.3
Kin.0	137.9	9.0	15.7	1.8	12.2	1.4
Kno.18	109.1	9.6	12.0	0.3	-	-
Koln	119.4	7.7	16.8	2.0	18.2	1.0
Kulturen.1	94.7	8.4	29.7	1.7	19.3	2.9
LAC.5	107.0	14.1	11.8	0.5	21.6	1.1
Lc.0	103.4	7.8	14.6	1.0	-	-
LDV.34	135.4	15.2	18.1	2.3	13.4	0.6
LDV.58	105.4	4.7	27.7	0.7	21.2	0.8
Ler.1	129.5	9.2	15.2	1.1	16.8	0.2
LI.OF.095	-	-	14.1	0.9	19.7	0.3
Liarum	-	-	16.3	1.1	27.4	1.6
Lip.0	104.7	4.6	15.5	1.5	23.0	1.0
Lis.1	165.2	26.0	19.9	1.6	20.5	1.7
Lisse	125.3	5.3	14.5	0.5	15.9	1.1
LL.0	134.1	14.0	20.7	1.5	20.6	1.0
Lm.2	114.1	17.8	18.5	1.8	18.2	0.8

Accession	Anion accumulation [nmol mg^{-1} FW]					
	NO_3^{1-}	SD	PO_4^{3-}	SD	SO_4^{2-}	SD
Lz.0	127.2	5.6	21.0	3.3	17.9	1.3
Map.42	141.6	5.6	-	-	-	-
MIB.15	152.0	9.2	17.6	1.9	-	-
MIB.22	107.9	7.1	-	-	-	-
MIB.28	133.3	2.6	18.3	1.2	13.9	1.7
MIB.84	127.1	13.1	16.3	1.5	18.6	1.1
MNF.Che.2	111.0	10.3	16.8	1.0	15.7	1.3
MNF.Jac.32	164.8	22.6	14.9	0.2	17.2	0.6
MNF.Pot.48	134.3	4.0	18.3	1.7	12.7	0.8
MNF.Pot.68	135.1	11.7	19.8	1.9	13.1	1.0
MOG.37	129.3	12.7	17.0	0.6	13.1	1.2
Mr.0	130.0	1.3	-	-	-	-
Mrk.0	106.1	7.4	15.5	0.4	18.4	1.3
Mt.0	103.3	6.1	18.7	1.4	15.5	0.7
Mz.0	119.2	9.5	17.8	1.5	15.8	1.1
N13	123.6	3.9	15.5	1.7	9.3	0.5
Na.1	125.4	8.7	16.2	1.5	12.7	0.6
Nd-1	-	-	-	-	20.8	0.8
NFA-10	166.0	11.2	22.6	2.3	24.6	2.1
NFA-8	182.7	26.0	-	6.6	-	-
Omo2.1	118.1	10.1	15.8	1.7	16.3	3.6
Or-1	126.5	13.3	-	-	12.6	2.0
Ost.0	94.9	5.2	14.3	2.1	18.5	1.5
Oy.0	112.2	13.5	14.6	0.8	13.4	1.4
Pa.1	82.3	9.7	8.4	0.9	9.4	0.9
PAR.3	113.6	8.8	12.9	0.7	10.3	0.4
PAR.4	107.3	10.2	13.8	1.9	7.8	1.9
PAR.5	117.2	13.7	25.0	0.9	16.0	1.4
Paw.3	126.7	12.1	18.7	2.6	9.8	1.3
Per.1	110.2	19.2	-	-	9.9	1.5
Petergof	125.8	14.1	-	-	12.0	2.2
Pna.17	124.5	12.7	-	-	-	-
Pro.0	136.9	19.7	15.7	3.2	12.2	1.6
Pu2.23	94.5	4.5	16.3	1.5	18.7	1.5
Ra.0	120.0	12.9	10.2	1.6	14.0	0.4
Rak.2	94.8	5.9	16.1	2.4	13.8	1.3
Ren.1	124.4	13.5	12.1	1.6	-	-
Rev.2	106.2	16.3	12.8	3.2	18.3	2.7
Rmx.A180	123.7	16.6	12.4	1.2	10.2	1.6
ROM.1	114.8	13.4	11.0	1.9	12.5	0.9
RRS.10	131.4	13.2	15.2	1.0	9.5	0.5
Rsch.4	-	-	13.3	2.1	20.0	0.9
Sanna.2	157.7	20.0	15.2	2.4	17.0	1.6
Sap.0	-	-	15.0	1.3	16.9	1.6
Sav.0	128.4	8.8	11.7	1.7	11.8	1.3
Se.0	-	-	12.5	0.4	14.8	2.9

Accession	Anion accumulation [nmol mg^{-1} FW]					
	NO_3^{1-}	SD	PO_4^{3-}	SD	SO_4^{2-}	SD
T1040	175.0	16.1	13.6	0.1	11.6	0.8
T1060	151.6	2.1	19.5	1.7	11.7	1.6
T1080	168.8	7.2	25.2	1.3	10.9	0.7
T1110	152.6	12.2	19.3	0.5	13.6	1.4
T510	142.8	8.5	18.1	1.0	11.4	0.8
T540	144.3	6.0	25.8	1.0	18.8	2.7
T620	139.4	10.8	24.9	1.3	12.6	1.9
T690	175.0	5.4	17.1	2.4	9.2	2.3
Ta.0	109.7	11.5	18.7	1.4	14.5	1.9
TAD.01	133.9	4.6	17.8	1.3	11.9	0.7
Tamm.2	-	-	14.7	2.1	18.7	1.3
TDr.3	108.2	1.9	19.6	1.6	12.2	1.1
TDr.8	111.7	9.4	16.6	0.6	16.1	2.1
TDr.18	112.1	10.2	17.4	2.2	13.6	1.3
Tomegap.2	-	-	-	-	13.4	2.2
Tottarp.2	142.9	7.2	15.1	0.8	13.7	0.8
TOU.A1.115	131.4	8.3	19.1	1.3	15.6	1.0
TOU.A1.116	146.4	7.7	16.2	0.9	16.9	1.4
TOU.A1.62	247.1	25.7	37.0	3.5	27.1	2.5
TOU.A1.96	242.1	23.4	33.0	3.1	23.3	0.7
TOU.C.3	229.0	7.3	38.9	1.3	30.2	1.5
TOU.E.11	94.2	5.0	-	-	17.0	0.4
TOU.I.17	108.2	10.6	-	-	12.7	1.6
TOU.I.6	106.0	7.8	-	-	12.7	1.1
TOU.J.3	108.2	11.6	-	-	11.3	1.0
Ts.1	107.4	12.6	13.1	1.2	16.7	0.8
UduI.1.34	106.2	10.7	14.5	0.9	18.0	0.7
UKID22	103.3	5.6	73.1	1.1	12.7	0.7
UKID37	134.8	4.9	16.8	0.2	17.7	0.8
UKID48	129.4	8.2	13.2	1.0	15.7	0.8
UKNW06.059	123.8	5.5	13.7	0.7	15.2	0.3
UKNW06.386	121.6	12.9	18.7	1.4	14.3	1.5
UKNW06.460	127.2	9.9	18.3	1.1	12.6	0.4
UKSE06.062	133.1	7.5	20.6	1.5	17.3	1.6
UKSE06.278	118.0	13.7	21.3	2.5	19.8	0.6
UKSE06.351	124.8	14.1	13.1	2.0	12.5	1.0
UKSE06.414	104.9	6.1	15.1	1.2	16.6	3.4
UKSE06.429	114.8	19.4	13.9	2.3	12.6	2.0
UKSE06.466	101.9	9.2	-	-	13.5	0.6
UKSE06.628	107.6	2.9	15.9	0.8	13.6	1.5
UKSW06.202	101.0	7.9	14.6	2.8	13.7	1.4
Ull2.5	99.8	4.4	13.1	1.3	15.2	1.2
Ull3.4	140.7	13.1	15.6	1.9	11.3	2.5
Uod.7	-	-	18.2	2.5	17.3	1.2
Var2.1	115.7	10.0	15.8	0.8	10.1	1.3
Wil.1	96.6	5.1	-	-	14.5	1.3

Accession	Anion accumulation [nmol mg^{-1} FW]					
	NO_3^{1-}	SD	PO_4^{3-}	SD	SO_4^{2-}	SD
Wt.5	111.6	7.7	13.3	1.9	18.3	1.4
X328PNA054	126.2	15.6	13.0	1.3	6.7	1.4
Zdrl.2.24	124.4	5.9	18.5	1.5	12.0	1.9
Zdrl.2.25	140.5	7.0	22.6	3.9	13.6	0.9
Zdr.6	-	-	14.5	1.6	12.7	1.8

The dataset generated for GWAS (see Chapter 5) includes the concentration of nitrate, phosphate and sulfate in arabidopsis natural accessions. The anion concentration was measured in five week old plants grown in the soil in 10 hours light conditions (see section 5.2.2).

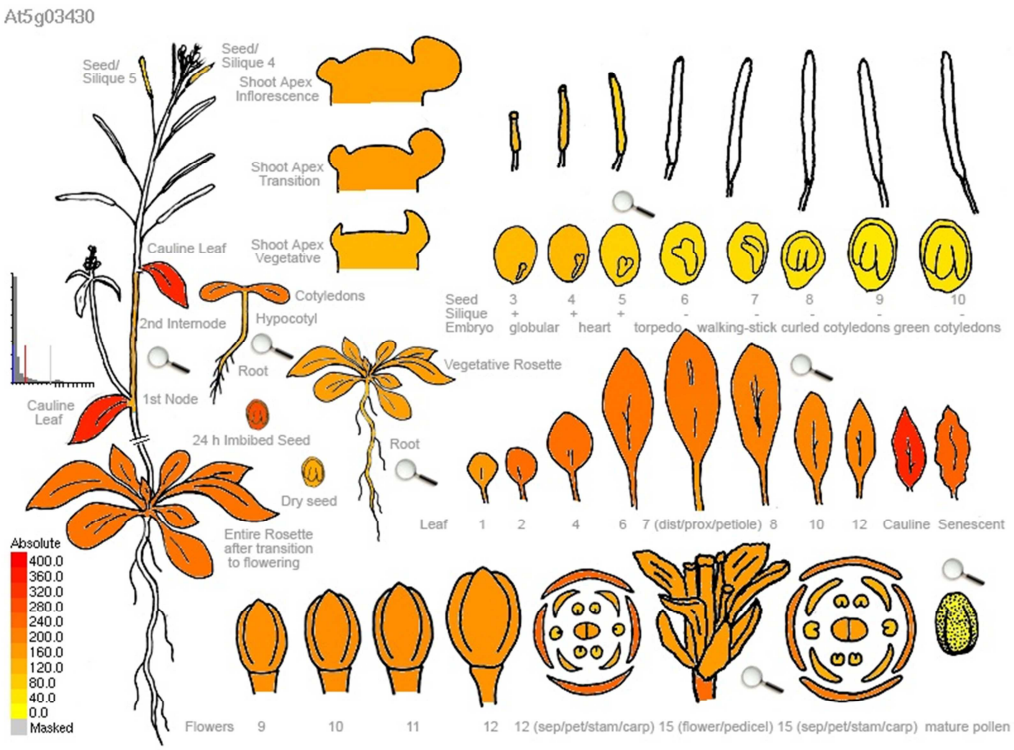


Figure S6.1 ***Expression of AT5G03430 at different developmental stages***

The image of transcript abundance of AT5G03430 at different developmental stages in different plant organs was obtained from eFP browser (Winter et al. 2007).

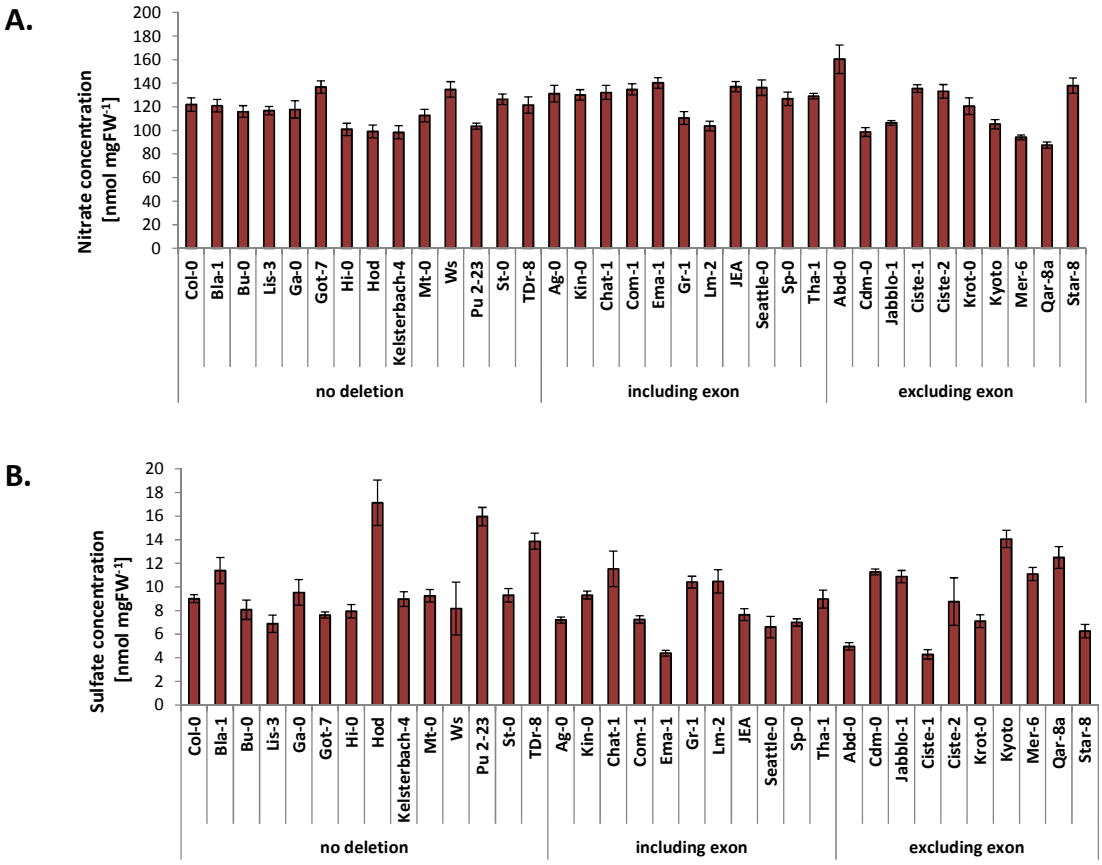


Figure S6.2 A*nion concentration in particular accessions from three AT5G03430 haplotype groups*

(A) Nitrate concentration, (B) sulfate concentration, (C) phosphate concentration in the representative accessions with no amino acid changes in the protein sequence, with deletion including second exon, and with deletion excluding second exon. The bars indicate average value of five independent plants ± standard error.

Bibliography

Abdallah M, Dubousset L, Meuriot F, Etienne P, Avice JC, Ourry A. 2010. Effect of mineral sulphur availability on nitrogen and sulphur uptake and remobilization during the vegetative growth of *Brassica napus* L. *Journal of Experimental Botany* 61: 2635-46

Ahmad A, Khan I, Anjum NAM, Diva I, Abdin MZ, Iqbal M. 2005. Effect of timing of sulfur fertilizer application on growth and yield of rapeseed. *Journal of Plant Nutrition* 28: 1049-59

Allwood JW, Erban A, de Koning S, Dunn WB, Luedemann A, et al. 2009. Inter-laboratory reproducibility of fast gas chromatography-electron impact-time of flight mass spectrometry (GC-EI-TOF/MS) based plant metabolomics. *Metabolomics* 5: 479-96

Alonso-Blanco C, Blankestijn-de Vries H, Hanhart CJ, Koornneef M. 1999. Natural allelic variation at seed size loci in relation to other life history traits of *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences USA* 96: 4710-7

Alonso JM, Stepanova AN, Leisse TJ, Kim CJ, Chen H, et al. 2003. Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science* 301: 653-57

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *Journal of Molecular Biology* 215: 403-10

Álvarez C, García I, Romero LC, Gotor C. 2012. Mitochondrial sulfide detoxification requires a functional isoform of O-acetylserine(thiol)lyase C in *Arabidopsis thaliana*. *Molecular Plant* 5: 1217-26

Amir R, Hacham Y, Galili G. 2002. Cystathionine γ-synthase and threonine synthase operate in concert to regulate carbon flow towards methionine in plants. *Trends in Plant Science* 7: 153-56

Anandacoomaraswamy A, De Costa WAJM, Tennakoon PLK, Van Der Werf A. 2002. The physiological basis of increased biomass partitioning to roots upon nitrogen deprivation in young clonal tea (*Camellia sinensis* (L.) O. Kuntz). *Plant and Soil* 238: 1-9

Andriotis VM, Pike MJ, Bunnewell S, Hills MJ, Smith AM. 2010. The plastidial glucose-6-phosphate/phosphate antiporter GPT1 is essential for morphogenesis in arabidopsis embryos. *Plant Journal* 64: 128-39

Anjum NA, Gill SS, Umar S, Ahmad I, Duarte AC, Pereira E. 2012. Improving growth and productivity of oleiferous Brassicas under changing environment: significance of nitrogen and sulphur nutrition, and underlying mechanisms. *The Scientific World Journal* 2012: 657808

Aranzana MJ, Kim S, Zhao K, Bakker E, Horton M, et al. 2005. Genome-wide association mapping in arabidopsis identifies previously known flowering time and pathogen resistance genes. *PLoS Genetics* 1: e60

Aravind L, Koonin EV. 2000. The STAS domain — a link between anion transporters and antisigma-factor antagonists. *Current Biology* 10: R53-R55

Atwell S, Huang YS, Vilhjalmsson BJ, Willems G, Horton M, et al. 2010. Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines. *Nature* 465: 627-31

Awazuhara M, Kim H, Goto DB, Matsui A, Hayashi H, et al. 2002. A 235-bp region from a nutritionally regulated soybean seed-specific gene promoter can confer its sulfur and

nitrogen response to a constitutive promoter in aerial tissues of *Arabidopsis thaliana*. *Plant Science* 163: 75-82

Balasubramanian S, Schwartz C, Singh A, Warthmann N, Kim MC, et al. 2009. QTL mapping in new *Arabidopsis thaliana* advanced intercross-recombinant inbred lines. *PLoS ONE* 4: e4318

Balik J, Pavlikova D, Tlustos P, Sykora K, Cerny J. 2006. The fluctuation of molybdenum concentration in oilseed rape plants after the application of nitrogen and sulphur fertilizers. *Plant Soil and Environment* 52: 301-07

Barberon M, Berthomieu P, Clairotte M, Shibagaki N, Davidian JC, Gosti F. 2008. Unequal functional redundancy between the two *Arabidopsis thaliana* high-affinity sulphate transporters SULTR1;1 and SULTR1;2. *New Phytologist* 180: 608-19

Bari R, Pant BD, Stitt M, Scheible W-R. 2006. PHO2, microRNA399, and PHR1 define a phosphate-signaling pathway in plants. *Plant Physiology* 141: 988-99

Barile M, Giancaspero TA, Brizio C, Panebianco C, Indiveri C, et al. 2013. Biosynthesis of flavin cofactors in man: implications in health and disease. *Current Pharmaceutical Design* 19: 2649-75

Baroux C, Spillane C, Grossniklaus U. 2002. Genomic imprinting during seed development. *Advances in Genetics* 46: 165-214

Bartel DP. 2004. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116: 281-97

Baus AD, Franzmann L, Meinke DW. 1986. Growth in vitro of arrested embryos from lethal mutants of *Arabidopsis thaliana*. *Theoretical and Applied Genetics.* 72: 577-86

Baxter I, Brazelton JN, Yu D, Huang YS, Lahner B, et al. 2010. A coastal cline in sodium accumulation in *Arabidopsis thaliana* is driven by natural variation of the sodium transporter AtHKT1;1. *PLoS Genetics* 6: e1001193

Baxter I, Dilkes BP. 2012. Elemental profiles reflect plant adaptations to the environment. *Science* 336: 1661-3

Baxter I, Muthukumar B, Park HC, Buchner P, Lahner B, et al. 2008. Variation in molybdenum concentration across broadly distributed populations of *Arabidopsis thaliana* is controlled by a mitochondrial molybdenum transporter (MOT1). *PLoS Genetics* 10: e1000004

Baxter I, Ouzzani M, Orcun S, Kennedy B, Jandhyala SS, Salt DE. 2007. Purdue ionomics information management system. An integrated functional genomics platform. *Plant Physiology* 143: 600-11

Beddington J, Mohammed A, Clark M, Fernandez A, Guillou M, et al. 2011. Achieving food security in the face of climate change: summary for policy makers from the Commission on Sustainable Agriculture and Climate Change.

Belaidi AA, Schwarz G. 2013. Metal insertion into the molybdenum cofactor: product-substrate channelling demonstrates the functional origin of domain fusion in gephyrin. *The Biochemical Journal* 450: 149-57

Benjamini Y, 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* 57: 289-300

Bergelson J, Roux F. 2010. Towards identifying genes underlying ecologically relevant traits in *Arabidopsis thaliana*. *Nature Reviews Genetics* 11: 867-79

Bermúdez MÁ, Galmés J, Moreno I, Mullineaux PM, Gotor C, Romero LC. 2012. Photosynthetic adaptation to length of day is dependent on S-sulfocysteine synthase activity in the thylakoid lumen. *Plant Physiology* 160: 274-88

Bick JA, Aslund F, Chen Y, Leustek T. 1998. Glutaredoxin function for the carboxyl-terminal domain of the plant-type 5'-adenylylsulfate reductase. *Proceedings of the National Academy of Sciences USA* 95: 8404-9

Bick JA, Setterdahl AT, Knaff DB, Chen Y, Pitcher LH, et al. 2001. Regulation of the plant-type 5'-adenylyl sulfate reductase by oxidative stress. *Biochemistry* 40: 9040-8

Bittner F. 2014. Molybdenum metabolism in plants and crosstalk to iron. *Frontiers in Plant Science* 5: 28

Blake-Kalff MM, Harrison KR, Hawkesford MJ, Zhao FJ, McGrath SP. 1998. Distribution of sulfur within oilseed rape leaves in response to sulfur deficiency during vegetative growth. *Plant Physiology* 118: 1337-44

Blake-Kalff MMA, Hawkesford MJ, Zhao FJ, McGrath SP. 2000. Diagnosing sulfur deficiency in field-grown oilseed rape (*Brassica napus* L.) and wheat (*Triticum aestivum* L.). *Plant and Soil* 225: 95-107

Blake-Kalff MMA, Zhao FJ, Hawkesford MJ, McGrath SP. 2001. Using plant analysis to predict yield losses caused by sulphur deficiency. *Annals of Applied Biology* 138: 123-27

Bolchi A, Petrucco S, Tenca PL, Foroni C, Ottonello S. 1999. Coordinate modulation of maize sulfate permease and ATP sulfurylase mRNAs in response to variations in sulfur nutritional status: stereospecific down-regulation by L-cysteine. *Plant Molecular Biology* 39: 527-37

Borevitz JO, Maloof JN, Lutes J, Dabi T, Redfern JL, et al. 2002. Quantitative trait loci controlling light and hormone response in two accessions of *Arabidopsis thaliana*. *Genetics* 160: 683-96

Botto JF, Alonso-Blanco C, Garzaron I, Sanchez RA, Casal JJ. 2003. The Cape Verde Islands allele of cryptochrome 2 enhances cotyledon unfolding in the absence of blue light in arabidopsis. *Plant Physiology* 133: 1547-56

Brachi B, Aime C, Glorieux C, Cuguen J, Roux F. 2012. Adaptive value of phenological traits in stressful environments: predictions based on seed production and laboratory natural selection. *PLoS ONE* 7: 504555411

Brachi B, Faure N, Horton M, Flahauw E, Vazquez A, et al. 2010. Linkage and association mapping of *Arabidopsis thaliana* flowering time in nature. *PLoS Genetics* 6: e1000940

Brader G, Tas E, Palva ET. 2001. Jasmonate-dependent induction of indole glucosinolates in arabidopsis by culture filtrates of the nonspecific pathogen *Erwinia carotovora*. *Plant Physiology* 126: 849-60

Bradford MM. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry* 72: 248-54

Brizio C, Galluccio M, Wait R, Torchetti EM, Bafunno V, et al. 2006. Over-expression in *Escherichia coli* and characterization of two recombinant isoforms of human FAD synthetase. *Biochemical and Biophysical Research Communications* 344: 1008-16

Brown PD, Tokuhisa JG, Reichelt M, Gershenzon J. 2003. Variation of glucosinolate accumulation among different organs and developmental stages of *Arabidopsis thaliana*. *Phytochemistry* 62: 471-81

Brunold C, Schiff JA. 1976. Studies of sulfate utilization of algae: 15. Enzymes of assimilatory sulfate reduction in Euglena and their cellular localization. *Plant Physiology* 57: 430-6

Brunold C, Suter M. 1989. Localization of enzymes of assimilatory sulfate reduction in pea roots. *Planta* 179: 228-34

Brunold C, Suter M. 1991. Adenosine 5'-phosphosulfate sulfotransferase from spruce. *Plant Physiology* 96: 155

Brychkova G, Xia Z, Yang G, Yesbergenova Z, Zhang Z, et al. 2007. Sulfite oxidase protects plants against sulfur dioxide toxicity. *Plant Journal* 50: 696-709

Buchner P, Prosser IM, Hawkesford MJ. 2004a. Phylogeny and expression of paralogous and orthologous sulphate transporter genes in diploid and hexaploid wheats. *Genome* 47: 526-34

Buchner P, Takahashi H, Hawkesford MJ. 2004b. Plant sulphate transporters: co-ordination of uptake, intracellular and long-distance transport. *Journal of Experimental Botany* 55: 1765-73

Buescher E, Achberger T, Amusan I, Giannini A, Ochsenfeld C, et al. 2010. Natural genetic variation in selected populations of *Arabidopsis thalianaIs* associated with ionomic differences. *PLoS ONE* 5: e11081

Burstenbinder K, Rzewuski G, Wirtz M, Hell R, Sauter M. 2007. The role of methionine recycling for ethylene synthesis in arabidopsis. *Plant Journal* 49: 238-49

Cakmak I, Hengeler C, Marschner H. 1994. Partitioning of shoot and root dry matter and carbohydrates in bean plants suffering from phosphorus, potassium, and magnesium deficiency. *Journal of Experimental Botany* 45: 1245-50

Candela H, Perez-Perez JM, Micol JL. 2011. Uncovering the post-embryonic functions of gametophytic- and embryonic-lethal genes. *Trends in Plant Science* 16: 336-45

Cao J, Schneeberger K, Ossowski S, Gunther T, Bender S, et al. 2011. Whole-genome sequencing of multiple *Arabidopsis thaliana* populations. *Nature Genetics* 43: 956-63

Cao MJ, Wang Z, Wirtz M, Hell R, Oliver DJ, Xiang CB. 2013. SULTR3;1 is a chloroplast-localized sulfate transporter in *Arabidopsis thaliana*. *Plant Journal* 73: 607-16

Ceccotti SP. 1995. Plant nutrient sulphur-a review of nutrient balance, environmental impact and fertilizers. *Fertilizer Research* 43: 117-25

Celenza JL, Quiel JA, Smolen GA, Merrikh H, Silvestro AR, et al. 2005. The arabidopsis ATR1 MYB transcription factor controls indolic glucosinolate homeostasis. *Plant Physiology* 137: 253-62

Chang C, Bowman JL, DeJohn AW, Lander ES, Meyerowitz EM. 1988. Restriction fragment length polymorphism linkage map for *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences USA* 85: 6856-60

Chao D-Y, Baraniecka P, Danku J, Koprivova A, Lahner B, et al. 2014. Variation in sulfur and selenium accumulation is controlled by naturally occurring isoforms of the key sulfur assimilation enzyme APR2 across the *Arabidopsis thaliana* species range. *Plant Physiology*

Chao DY, Gable K, Chen M, Baxter I, Dietrich CR, et al. 2011. Sphingolipids in the root play an important role in regulating the leaf ionome in *Arabidopsis thaliana*. *Plant Cell* 23: 1061-81

Chao DY, Silva A, Baxter I, Huang YS, Nordborg M, et al. 2012. Genome-wide association studies identify heavy metal ATPase3 as the primary determinant of natural variation in leaf cadmium in *Arabidopsis thaliana*. *PLoS Genetics* 8: e1002923

Chartron J, Carroll KS, Shiau C, Gao H, Leary JA, et al. 2006. Substrate recognition, protein dynamics, and iron-sulfur cluster in *Pseudomonas aeruginosa* adenosine 5'-phosphosulfate reductase. *Journal of Molecular Biology* 364: 152-69

Chen YF, Matsubayashi Y, Sakagami Y. 2000. Peptide growth factor phytosulfokine-alpha contributes to the pollen population effect. *Planta* 211: 752-5

Chen Z, Watanabe T, Shinano T, Ezawa T, Wasaki J, et al. 2009. Element interconnections in *Lotus japonicus*: a systematic study of the effects of element additions on different natural variants. *Soil Science and Plant Nutrition* 55: 91-101

Chevalier F, Pata M, Nacry P, Doumas P, Rossignol M. 2003. Effects of phosphate availability on the root system architecture: large-scale analysis of the natural variation between arabidopsis accessions. *Plant, Cell & Environment* 26: 1839-50

Chiba Y, Ishikawa M, Kijima F, Tyson RH, Kim J, et al. 1999. Evidence for autoregulation of cystathionine gamma-synthase mRNA stability in arabidopsis. *Science* 286: 1371-4

Chiba Y, Sakurai R, Yoshino M, Ominato K, Ishikawa M, et al. 2003. S-adenosyl-L-methionine is an effector in the posttranscriptional autoregulation of the cystathionine gamma-synthase gene in arabidopsis. *Proceedings of the National Academy of Sciences USA* 100: 10225-30

Childs LH, Lisec J, Walther D. 2012. Matapax: an online high-throughput genome-wide association study pipeline. *Plant Physiology* 158: 1534-41

Chiou TJ, Aung K, Lin SI, Wu CC, Chiang SF, Su CL. 2006. Regulation of phosphate homeostasis by microRNA in arabidopsis. *Plant Cell* 18: 412-21

Chou WL, Huang LF, Fang JC, Yeh CH, Hong CY, et al. 2014. Divergence of the expression and subcellular localization of CCR4-associated factor 1 (CAF1) deadenylase proteins in *Oryza sativa*. *Plant Molecular Biology* 85: 443-58

Chung LY. 2006. The antioxidant properties of garlic compounds: allyl cysteine, alliin, allicin, and allyl disulfide. *Journal of Medicinal Food* 9: 205-13

Ciereszko I, Johansson H, Hurry V, Kleczkowski LA. 2001. Phosphate status affects the gene expression, protein concentration and enzymatic activity of UDP-glucose pyrophosphorylase in wild-type and pho mutants of arabidopsis. *Planta* 212: 598-605

Clarkson DT, Smith FW, Berg PJV. 1983. Regulation of sulphate transport in a tropical legume, *Macroptilium atropurpureum*, *cv. siratro*. *Journal of Experimental Botany* 34: 1463-83

Costa Liliana M, Yuan J, Rouster J, Paul W, Dickinson H, Gutierrez-Marcos Jose F. 2012. Maternal Control of Nutrient Allocation in Plant Seeds by Genomic Imprinting. *Current Biology* 22: 160-65

Coupland G. 1995. Genetic and environmental control of flowering time in arabidopsis. *Trends in Genetics* 11: 393-7

Crawford NM, Forde BG. 2002. Molecular and developmental biology of inorganic nitrogen nutrition. *The arabidopsis book / American Society of Plant Biologists* 1: e0011

Cumming M, Leung S, McCallum J, McManus MT. 2007. Complex formation between recombinant ATP sulfurylase and APS reductase of *Allium cepa* (L.). *FEBS Letters* 581: 4139-47

Curtis TY, Powers SJ, Balagiannis D, Elmore JS, Mottram DS, et al. 2010. Free amino acids and sugars in rye grain: implications for acrylamide formation. *Journal of Agricultural and Food Chemistry* 58: 1959-69

Datko AH, Mudd SH. 1984. Sulfate uptake and its regulation in *Lemna paucicostata Hegelm*. 6746. *Plant Physiology* 75: 466-73

Daugeron MC, Mauxion F, Seraphin B. 2001. The yeast POP2 gene encodes a nuclease involved in mRNA deadenylation. *Nucleic Acids Research* 29: 2448-55

Davey MW, Montagu MV, Inzé D, Sanmartin M, Kanellis A, et al. 2000. Plant L-ascorbic acid: chemistry, function, metabolism, bioavailability and effects of processing. *Journal of the Science of Food and Agriculture* 80: 825-60

Davidian J-C, Kopriva S. 2010. Regulation of sulfate uptake and assimilation—the same or not the same? *Molecular Plant* 3: 314-25

De Pessemier J, Chardon F, Juraniec M, Delaplace P, Hermans C. 2013. Natural variation of the root morphological response to nitrate supply in *Arabidopsis thaliana*. *Mechanisms of Development* 130: 45-53

Despres B, Delseny M, Devic M. 2001. Partial complementation of embryo defective mutations: a general strategy to elucidate gene function. *Plant Journal* 27: 149-59

Dethloff F, Erban A, Orf I, Alpers J, Fehrle I, et al. 2014. Profiling methods to identify cold-regulated primary metabolites using gas chromatography coupled to mass spectrometry. *Methods in Molecular Biology* 1166: 171-97

Dickson SP, Wang K, Krantz I, Hakonarson H, Goldstein DB. 2010. Rare variants create synthetic genome-wide associations. *PLoS Biology* 8: e1000294

Dietz KJ. 1989. Leaf and chloroplast development in relation to nutrient availability. *Journal of Plant Physiology* 134: 544-50

Dijkshoorn W, van Wijk AL. 1967. The sulphur requirements of plants as evidenced by the sulphur-nitrogen ratio in the organic matter – a review of published data. *Plant and Soil* 26: 129-57

Dinkova-Kostova AT, Kostov RV. 2012. Glucosinolates and isothiocyanates in health and disease. *Trensd in Molecular Medicine* 18: 337-347

Dobbek H. 2011. Structural aspects of mononuclear Mo/W-enzymes. *Coordination Chemistry Reviews* 255: 1104-16

Dorner AJ, Kaufman RJ. 1990. Analysis of synthesis, processing, and secretion of proteins expressed in mammalian cells. *Methods in Enzymology* 185: 577-96

Doughty KJ, Kiddle GA, Pye BJ, Wallsgrove RM, Pickett JA. 1995. Selective induction of glucosinolates in oilseed rape leaves by methyl jasmonate. *Phytochemistry* 38: 347-50

Droux M, Ruffet ML, Douce R, Job D. 1998. Interactions between serine acetyltransferase and O-acetylserine (thiol) lyase in higher plants--structural and kinetic properties of the free and bound enzymes. *European Journal of Biochemistry* 255: 235-45

Durenkamp M, De Kok LJ. 2004. Impact of pedospheric and atmospheric sulphur nutrition on sulphur metabolism of *Allium cepa* L., a species with a potential sink capacity for secondary sulphur compounds. *Journal of Experimental Botany* 55: 1821-30

Durenkamp M, De Kok LJ, Kopriva S. 2007. Adenosine 5'-phosphosulphate reductase is regulated differently in *Allium cepa* L. and *Brassica oleracea* L. upon exposure to H_2S. *Journal of Experimental Botany* 58: 1571-9

Eckardt NA. 2005. Moco mojo: Crystal structure reveals essential features of eukaryotic assimilatory nitrate reduction. *Plant Cell* 17: 1029-31

Eide D, Clark S, Nair TM, Gehl M, Gribskov M, et al. 2005. Characterization of the yeast ionome: a genome-wide analysis of nutrient mineral and trace element homeostasis in *Saccharomyces cerevisiae*. *Genome Biology* 6: R77

Eilers T, Schwarz G, Brinkmann H, Witt C, Richter T, et al. 2001. Identification and biochemical characterization of *Arabidopsis thaliana* sulfite oxidase. A new player in plant sulfur metabolism. *The Journal of Biological Chemistry* 276: 46989-94

El-Soda M, Kruijer W, Malosetti M, Koornneef M, Aarts MG. 2014. Quantitative trait loci and candidate genes underlying genotype by environment interaction in the response of *Arabidopsis thaliana* to drought. *Plant, Cell & Environment* 10.1111; 12418

Erban A, Schauer N, Fernie AR, Kopka J. 2007. Nonsupervised construction and application of mass spectral and retention time index libraries from time-of-flight gas chromatography-mass spectrometry metabolite profiles. *Methods in Molecular Biology* 358: 19-38

Essigmann B, Guler S, Narang RA, Linke D, Benning C. 1998. Phosphate availability affects the thylakoid lipid composition and the expression of SQD1, a gene required for sulfolipid biosynthesis in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences USA* 95: 1950-5

Fahey JW, Zalcmann AT, Talalay P. 2001. The chemical diversity and distribution of glucosinolates and isothiocyanates among plants. *Phytochemistry* 56: 5-51

Fazili IS, Jamal A, Ahmad S, Masoodi M, Khan JS, Abdin MZ. 2008. Interactive effect of sulfur and nitrogen on nitrogen accumulation and harvest in oilseed crops differing in nitrogen assimilation potential. *Journal of Plant Nutrition* 31: 1203-20

Fiehn O, Kopka J, Dormann P, Altmann T, Trethewey RN, Willmitzer L. 2000. Metabolite profiling for plant functional genomics. *Nature Biotechnology* 18: 1157-61

Filiault DL, Maloof JN. 2012. A genome-wide association study identifies variants underlying the *Arabidopsis thaliana* shade avoidance response. *PLoS Genetics* 8: e1002589

Fischer K, Barbier GG, Hecht HJ, Mendel RR, Campbell WH, Schwarz G. 2005. Structural basis of eukaryotic nitrate reduction: crystal structures of the nitrate reductase active site. *Plant Cell* 17: 1167-79

Fismes J, Vong PC, Guckert A, Frossard E. 2000. Influence of sulfur on apparent N-use efficiency, yield and quality of oilseed rape (*Brassica napus* L.) grown on a calcareous soil. *European Journal of Agronomy* 12: 127-41

Flæte NES, Hollung K, Ruud L, Sogn T, Færgestad EM, et al. 2005. Combined nitrogen and sulphur fertilisation and its effect on wheat quality and protein composition measured by SE-FPLC and proteomics. *Journal of Cereal Science* 41: 357-69

Flint J, Eskin E. 2012. Genome-wide association studies in mice. *Nature Review Genetics* 13: 807-17

Francois JA, Kumaran S, Jez JM. 2006. Structural basis for interaction of O-acetylserine sulfhydrylase and serine acetyltransferase in the arabidopsis cysteine synthase complex. *Plant Cell* 18: 3647-55

Friedman M. 2003. Chemistry, biochemistry, and safety of acrylamide. A Review. *Journal of Agricultural and Food Chemistry* 51: 4504-26

Friedrich JW, Schrader LE. 1978. Sulfur deprivation and nitrogen metabolism in maize seedlings. *Plant Physiol*ogy 61: 900-3

Fritz C, Palacios-Rojas N, Feil R, Stitt M. 2006. Regulation of secondary metabolism by the carbon-nitrogen status in tobacco: nitrate inhibits large sectors of phenylpropanoid metabolism. *Plant Journal* 46: 533-48

Fujii H, Chiou T-J, Lin S-I, Aung K, Zhu J-K. 2005. A miRNA involved in phosphate-starvation response in arabidopsis. *Current Biology* 15: 2038-43

Galant A, Preuss ML, Cameron JC, Jez JM. 2011. Plant glutathione biosynthesis: diversity in biochemical regulation and reaction products. *Frontiers in Plant Science* 2

Gallie DR. 2013. L-Ascorbic Acid: A multifunctional molecule supporting plant growth and development. *Scientifica* 2013: 24

Gasber A, Klaumann S, Trentmann O, Trampczynska A, Clemens S, et al. 2011. Identification of an arabidopsis solute carrier critical for intracellular transport and inter-organ allocation of molybdate. *Plant Biology* 13: 710-18

Gigolashvili T, Berger B, Mock HP, Muller C, Weisshaar B, Flugge UI. 2007. The transcription factor HIG1/MYB51 regulates indolic glucosinolate biosynthesis in *Arabidopsis thaliana*. *Plant Journal* 50: 886-901

Gigolashvili T, Engqvist M, Yatusevich R, Muller C, Flugge UI. 2008. HAG2/MYB76 and HAG3/MYB29 exert a specific and coordinated control on the regulation of aliphatic glucosinolate biosynthesis in *Arabidopsis thaliana*. *New Phytology* 177: 627-42

Gilbert SM, Clarkson DT, Cambridge M, Lambers H, Hawkesford MJ. 1997. SO_4^{2-} deprivation has an early effect on the concentration of ribulose - 1,5 - bisphosphate carboxylase / oxygenase and photosynthesis in young leaves of wheat. *Plant Physiology* 115: 1231-39

Gill SS, Tuteja N. 2010. Reactive oxygen species and antioxidant machinery in abiotic stress tolerance in crop plants. *Plant Physiology and Biochemistry* 48: 909-30

Gillaspy GE. 2011. The cellular language of myo-inositol signaling. *New Phytologist* 192: 823-39

Giovanelli J, Mudd SH, Datko AH. 1985. Quantitative analysis of pathways of methionine metabolism and their regulation in *Lemna*. *Plant Physiology* 78: 555-60

Gombert J, Etienne P, Ourry A, Le Dily F. 2006. The expression patterns of SAG12/Cab genes reveal the spatial and temporal progression of leaf senescence in *Brassica napus* L. with sensitivity to the environment. *Journal of Experimental Botany* 57: 1949-56

Haas FH, Heeg C, Queiroz R, Bauer A, Wirtz M, Hell R. 2008. Mitochondrial serine acetyltransferase functions as a pacemaker of cysteine synthesis in plant cells. *Plant Physiology* 148: 1055-67

Habtegebrial K, Singh BR. 2006. Effects of timing of nitrogen and sulphur fertilizers on yield, nitrogen, and sulphur concentrations of Tef (*Eragrostis tef* (Zucc.) Trotter). *Nutrient Cycling in Agroecosystems* 75: 213-22

Hacham Y, Schuster G, Amir R. 2006. An in vivo internal deletion in the N-terminus region of arabidopsis cystathionine gamma-synthase results in CGS expression that is insensitive to methionine. *Plant Journal* 45: 955-67

Hagen G, Guilfoyle T. 2002. Auxin-responsive gene expression: genes, promoters and regulatory factors. *Plant Molecular Biology* 49: 373-85

Halford NG, Curtis TY, Muttucumaru N, Postles J, Elmore JS, Mottram DS. 2012. The acrylamide problem: a plant and agronomic science issue. *Journal of Experimental Botany* 63(8):2841-51

Halkier BA, Gershenzon J. 2006. Biology and biochemistry of glucosinolates. *Annual Reviews Plant Biology* 57: 303-33

Hanahan D. 1983. Studies on transformation of *Escherichia coli* with plasmids. *Journal of Molecular Biology* 166: 557-80

Hanai H, Nakayama D, Yang H, Matsubayashi Y, Hirota Y, Sakagami Y. 2000. Existence of a plant tyrosylprotein sulfotransferase: novel plant enzyme catalyzing tyrosine O-sulfation of preprophytosulfokine variants in vitro. *FEBS Letters* 470: 97-101

Hancock AM, Brachi B, Faure N, Horton MW, Jarymowycz LB, et al. 2011. Adaptation to climate across the *Arabidopsis thaliana* genome. *Science* 334: 83-86

Hannah MA, Wiese D, Freund S, Fiehn O, Heyer AG, Hincha DK. 2006. Natural genetic variation of freezing tolerance in arabidopsis. *Plant Physiology* 142: 98-112

Harada E, Kusano T, Sano H. 2000. Differential expression of genes encoding enzymes involved in sulfur assimilation pathways in response to wounding and jasmonate in *Arabidopsis thaliana*. *Journal of Plant Physiology* 156: 272-76

Harmon M. 2011. Normality testing in Excel - the Excel statistical master. *Excel Master Series*.

Hatzfeld Y, Lee S, Lee M, Leustek T, Saito K. 2000. Functional characterization of a gene encoding a fourth ATP sulfurylase isoform from *Arabidopsis thaliana*. *Genetics* 248: 51-58

Hawkesford M, Davidian J-C, Grignon C. 1993. Sulphate/proton cotransport in plasma-membrane vesicles isolated from roots of *Brassica napus* L.: increased transport in membranes isolated from sulphur-starved plants. *Planta* 190: 297-304

Hawkesford MJ. 2000. Plant responses to sulphur deficiency and the genetic manipulation of sulphate transporters to improve S-utilization efficiency. *Journal of Experimental Botany* 51: 131-38

Hawkesford MJ. 2003. Transporter gene families in plants: the sulphate transporter gene family — redundancy or specialization? *Physiologia Plantarum* 117: 155-63

Hawkesford MJ, De Kok LJ. 2006. Managing sulphur metabolism in plants. *Plant, Cell, and Environment* 29: 382-95

Hell R, Hillebrand H. 2001. Plant concepts for mineral acquisition and allocation. *Current Opinion in Biotechnology* 12: 161-8

Hell R, Jost R, Berkowitz O, Wirtz M. 2002. Molecular and biochemical analysis of the enzymes of cysteine biosynthesis in the plant *Arabidopsis thaliana*. *Amino Acids* 22: 245-57

Hell R, Wirtz M. 2008. Metabolism of cysteine in plants and phototrophic bacteria in *Sulfur Metabolism in Phototrophic Organisms*, ed. R Hell, C Dahl, D Knaff, T Leustek, pp. 59-91: Springer Netherlands

Hell R, Wirtz M. 2011. Molecular biology, biochemistry and cellular physiology of cysteine metabolism in *Arabidopsis thaliana*. *The arabidopsis Book / American Society of Plant Biologists* 9: e0154

Hermans C, Hammond JP, White PJ, Verbruggen N. 2006. How do plants respond to nutrient shortage by biomass allocation? *Trends in Plant Science* 11: 610-7

Herrmann J, Ravilious GE, McKinney SE, Westfall CS, Lee SG, et al. 2014. Structure and mechanism of soybean ATP sulfurylase and the committed step in plant sulfur assimilation. *Journal of Biological Chemistry*

Herschbach C, Rennenberg H. 1991. Influence of glutathione (GSH) on sulphate influx, xylem loading and exudation in excised tobacco roots. *Journal of Experimental Botany* 42: 1021-29

Herschbach C, Rennenberg H. 1994. Influence of glutathione (GSH) on net uptake of sulphate and sulphate transport in tobacco plants. *Journal of Experimental Botany* 45: 1069-76

Herschbach C, van Der Zalm E, Schneider A, Jouanin L, De Kok LJ, Rennenberg H. 2000. Regulation of sulfur nutrition in wild-type and transgenic poplar over-expressing gamma-glutamylcysteine synthetase in the cytosol as affected by atmospheric H_2S. *Plant Physiology* 124: 461-73

Hesse H, Kreft O, Maimann S, Zeh M, Hoefgen R. 2004a. Current understanding of the regulation of methionine biosynthesis in plants. *Journal of Experimental Botany* 55: 1799-808

Hesse H, Nikiforova V, Gakiere B, Hoefgen R. 2004b. Molecular analysis and control of cysteine biosynthesis: integration of nitrogen and sulphur metabolism. *Journal of Experimental Botany* 55: 1283-92

Hesse H, Trachsel N, Suter M, Kopriva S, von Ballmoos P, et al. 2003. Effect of glucose on assimilatory sulphate reduction in *Arabidopsis thaliana* roots. *Journal of Experimental Botany* 54: 1701-9

Hille R. 1996. The mononuclear molybdenum enzymes. *Chemical Reviews* 96: 2757-816

Hille R. 2005. Molybdenum-containing hydroxylases. *Archives of Biochemistry and Biophysics* 433: 107-16

Hille R, Nishino T, Bittner F. 2011. Molybdenum enzymes in higher organisms. *Coordination Chemistry Reviews* 255: 1179-205

Hirai MY, Sugiyama K, Sawada Y, Tohge T, Obayashi T, et al. 2007. Omics-based identification of arabidopsis MYB transcription factors regulating aliphatic glucosinolate biosynthesis. *Proceedings of the National Academy of Sciences USA* 104: 6478-83

Hirai MY, Yano M, Goodenowe DB, Kanaya S, Kimura T, et al. 2004. Integration of transcriptomics and metabolomics for understanding of global responses to nutritional stresses in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences USA* 101: 10205-10

Hirschhorn JN, Daly MJ. 2005. Genome-wide association studies for common diseases and complex traits. *Nature Reviews Genetics* 6: 95-108

Hirschi KD. 2003. Strike while the ionome is hot: making the most of plant genomic advances. *Trends in Biotechnology* 21: 520-1

Hoff T, Schnorr KM, Meyer C, Caboche M. 1995. Isolation of two arabidopsis cDNAs involved in early steps of molybdenum cofactor biosynthesis by functional complementation of *Escherichia coli* mutants. *The Journal of Biological Chemistry* 270: 6100-7

Hoffmann MH. 2002. Biogeography of *Arabidopsis thaliana* (L.) Heynh. (Brassicaceae). *Journal of Biogeography* 29: 125-34

Honsel A, Kojima M, Haas R, Frank W, Sakakibara H, et al. 2012. Sulphur limitation and early sulphur deficiency responses in poplar: significance of gene expression, metabolites, and plant hormones. *Journal of Experimental Botany* 63: 1873-93

Hopkins L, Parmar S, Blaszczyk A, Hesse H, Hoefgen R, Hawkesford MJ. 2005. O-acetylserine and the regulation of expression of genes encoding components for sulfate uptake and assimilation in potato. *Plant Physiology* 138: 433-40

Hopkins L, Parmar S, Bouranis DL, Howarth JR, Hawkesford MJ. 2004. Coordinated expression of sulfate uptake and components of the sulfate assimilatory pathway in maize. *Plant Biology* 6: 408-14

Horton MW, Hancock AM, Huang YS, Toomajian C, Atwell S, et al. 2012. Genome-wide patterns of genetic variation in worldwide *Arabidopsis thaliana* accessions from the RegMap panel. *Nature Genetics* 44: 212-16

Howarth JR, Fourcroy P, Davidian JC, Smith FW, Hawkesford MJ. 2003. Cloning of two contrasting high-affinity sulfate transporters from tomato induced by low sulfate and infection by the vascular pathogen *Verticillium dahliae*. *Planta* 218: 58-64

Hsieh LC, Lin SI, Shih AC, Chen JW, Lin WY, et al. 2009. Uncovering small RNA-mediated responses to phosphate deficiency in arabidopsis by deep sequencing. *Plant Physiology* 151: 2120-32

Hu Y, Ribbe MW. 2013. Biosynthesis of the iron-molybdenum cofactor of nitrogenase. *The Journal of Biological Chemistry* 288: 13173-7

Huang X, Effgen S, Meyer RC, Theres K, Koornneef M. 2012a. Epistatic natural allelic variation reveals a function of AGAMOUS-LIKE6 in axillary bud formation in arabidopsis. *Plant Cell* 24: 2364-79

Huang X, Paulo MJ, Boer M, Effgen S, Keizer P, et al. 2011. Analysis of natural allelic variation in arabidopsis using a multiparent recombinant inbred line population. *Proceedings of the National Academy of Sciences USA* 108: 4488-93

Huang X, Zhao Y, Wei X, Li C, Wang A, et al. 2012b. Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. *Nature Genetics* 44: 32-39

Huang Y, Li H, Hutchison CE, Laskey J, Kieber JJ. 2003. Biochemical and functional analysis of CTR1, a protein kinase that negatively regulates ethylene signaling in arabidopsis. *Plant Journal* 33: 221-33

Hubberten HM, Drozd A, Tran BV, Hesse H, Hoefgen R. 2012a. Local and systemic regulation of sulfur homeostasis in roots of *Arabidopsis thaliana*. *Plant Journal* 72: 625-35

Hubberten HM, Klie S, Caldana C, Degenkolbe T, Willmitzer L, Hoefgen R. 2012b. Additional role of O-acetylserine as a sulfur status-independent regulator during plant growth. *Plant Journal* 70: 666-77

Huerta C, Borek D, Machius M, Grishin NV, Zhang H. 2009. Structure and mechanism of a eukaryotic FMN adenylyltransferase. *Journal of Molecular Biology* 389: 388-400

Hummel J, Strehmel N, Selbig J, Walther D, Kopka J. 2010. Decision tree supported substructure prediction of metabolites from GC-MS profiles. *Metabolomics* 6: 322-33

Hussain A, Larsson H, Kuktaite R, Prieto-Linde ML, Johansson E. 2012. Towards the understanding of bread-making quality in organically grown wheat: Dough mixing behaviour, protein polymerisation and structural properties. *Journal of Cereal Science* 56: 659-66

Ikram S, Bedu M, Daniel-Vedele F, Chaillou S, Chardon F. 2012. Natural variation of arabidopsis response to nitrogen availability. *Journal of Experimental Botany* 63: 91-105

Jez J. 2008. Sulfur: a missing link between soils, crops and nutrition. *Madison: American Society of Agronomy*. 323 pp.

Johnson JL, Rajagopalan KV. 1979. The oxidation of sulphite in animal systems. *CIBA Foundation Symposium*: 119-33

Jones-Rhoades MW, Bartel DP. 2004. Computational identification of plant microRNAs and their targets, including a stress-induced miRNA. *Molecular Cell* 14: 787-99

Joosten V, van Berkel WJ. 2007. Flavoenzymes. *Current Opinions in Chemical Biology* 11: 195-202

Jost R, Altschmied L, Bloem E, Bogs J, Gershenzon J, et al. 2005. Expression profiling of metabolic genes in response to methyl jasmonate reveals regulation of genes of primary and secondary sulfur-related pathways in *Arabidopsis thaliana*. *Photosynthesis Research* 86: 491-508

Juenger T, Purugganan M, Mackay TF. 2000. Quantitative trait loci for floral morphology in *Arabidopsis thaliana*. *Genetics* 156: 1379-92

Kang HM, Zaitlen NA, Wade CM, Kirby A, Heckerman D, et al. 2008. Efficient control of population structure in model organism association mapping. *Genetics* 178: 1709-23

Kannan S, Ramani S. 1978. Studies on molybdenum absorption and transport in bean and rice. *Plant Physiology* 62: 179-81

Kant S, Bi YM, Rothstein SJ. 2011. Understanding plant response to nitrogen limitation for the improvement of crop nitrogen use efficiency. *Journal of Experimental Botany* 62: 1499-509

Karmoker J, Clarkson D, Saker L, Rooney J, Purves J. 1991. Sulphate deprivation depresses the transport of nitrogen to the xylem and the hydraulic conductivity of barley (*Hordeum vulgare* L.) roots. *Planta* 185: 269-78

Karthikeyan AS, Varadarajan DK, Mukatira UT, D'Urzo MP, Damsz B, Raghothama KG. 2002. Regulated expression of arabidopsis phosphate transporters. *Plant Physiology* 130: 221-33

Kataoka T, Hayashi N, Yamaya T, Takahashi H. 2004a. Root-to-shoot transport of sulfate in arabidopsis. Evidence for the role of SULTR3;5 as a component of low-affinity sulfate transport system in the root vasculature. *Plant Physiology* 136: 4198-204

Kataoka T, Watanabe-Takahashi A, Hayashi N, Ohnishi M, Mimura T, et al. 2004b. Vacuolar sulfate transporters are essential determinants controlling internal distribution of sulfate in arabidopsis. *Plant Cell* 16: 2693-704

Kaufholdt D, Gehl C, Geisler M, Jeske O, Voedisch S, et al. 2013. Visualization and quantification of protein interactions in the biosynthetic pathway of molybdenum cofactor in *Arabidopsis thaliana*. *Journal of Experimental Botany* 64: 2005-16

Kawashima CG, Berkowitz O, Hell R, Noji M, Saito K. 2005. Characterization and expression analysis of a serine acetyltransferase gene family involved in a key step of the sulfur assimilation pathway in arabidopsis. *Plant Physiology* 137: 220-30

Kawashima CG, Matthewman CA, Huang S, Lee B-R, Yoshimoto N, et al. 2011. Interplay of SLIM1 and miR395 in the regulation of sulfate assimilation in arabidopsis. *Plant Journal* 66: 863-76

Kawashima CG, Yoshimoto N, Maruyama-Nakashita A, Tsuchiya YN, Saito K, et al. 2009. Sulphur starvation induces the expression of microRNA-395 and one of its target genes but in different cell types. *Plant Journal* 57: 313-21

Kellermeier F, Chardon F, Amtmann A. 2013. Natural variation of arabidopsis root architecture reveals complementing adaptive strategies to potassium starvation. *Plant Physiol*ogy 161: 1421-32

Khan MS, Haas FH, Allboje Samami A, Moghaddas Gholami A, Bauer A, et al. 2010. Sulfite reductase defines a newly discovered bottleneck for assimilatory sulfate reduction and is essential for growth and development in *Arabidopsis thaliana*. *The Plant Cell* 22: 1216-31

Kieber JJ, Rothenberg M, Roman G, Feldmann KA, Ecker JR. 1993. CTR1, a negative regulator of the ethylene response pathway in arabidopsis, encodes a member of the raf family of protein kinases. *Cell* 72: 427-41

Kim S, Plagnol V, Hu TT, Toomajian C, Clark RM, et al. 2007. Recombination and linkage disequilibrium in *Arabidopsis thaliana*. *Nature Genetics* 39: 1151-55

Kisker C, Schindelin H, Pacheco A, Wehbi WA, Garrett RM, et al. 1997. Molecular basis of sulfite oxidase deficiency from the structure of sulfite oxidase. *Cell* 91: 973-83

Kobayashi T, Eun C-H, Hanai H, Matsubayashi Y, Sakagami Y, Kamada H. 1999. Phytosulphokine-α, a peptidyl plant growth factor, stimulates somatic embryogenesis in carrot. *Journal of Experimental Botany* 50: 1123-28

Koch K. 2004. Sucrose metabolism: regulatory mechanisms and pivotal roles in sugar sensing and plant development. *Current Opinions in Plant Biology* 7: 235-46

Kohler C, Grossniklaus U. 2005. Seed development and genomic imprinting in plants. *Progress in Molecular and Subcellular Biology* 38: 237-62

Koornneef M, Alonso-Blanco C, Vreugdenhil D. 2004. Naturally occurring genetic variation in *arabidopsis thaliana*. *Annual Review of Plant Biology* 55: 141-72

Kopka J, Schauer N, Krueger S, Birkemeyer C, Usadel B, et al. 2005. GMD@CSB.DB: the Golm metabolome database. *Bioinformatics* 21: 1635-8

Kopriva S. 2006. Regulation of sulfate assimilation in arabidopsis and beyond. *Annals of Botany* 97: 479-95

Kopriva S, Buchert T, Fritz G, Suter M, Weber M, et al. 2001. Plant adenosine 5'-phosphosulfate reductase is a novel iron-sulfur protein. *The Journal of Biological Chemistry* 276: 42881-6

Kopriva S, Fritzemeier K, Wiedemann G, Reski R. 2007. The putative moss 3'-phosphoadenosine-5'-phosphosulfate reductase is a novel form of adenosine-5'-phosphosulfate reductase without an iron-sulfur cluster. *The Journal of Biological Chemistry* 282: 22930-8

Kopriva S, Koprivova A. 2004. Plant adenosine 5 '-phosphosulphate reductase: the past, the present, and the future. *Journal of Experimental Botany* 55: 1775-83

Kopriva S, Mugford S, Matthewman C, Koprivova A. 2009. Plant sulfate assimilation genes: redundancy versus specialization. *Plant Cell Reports* 28: 1769-80

Kopriva S, Mugford SG, Baraniecka P, Lee B-R, Matthewman CA, Koprivova A. 2012. Control of sulfur partitioning between primary and secondary metabolism in arabidopsis. *Frontiers in Plant Science* 19;3:163

Kopriva S, Muheim R, Koprivova A, Trachsel N, Catalano C, et al. 1999. Light regulation of assimilatory sulphate reduction in *Arabidopsis thaliana*. *The Plant Journal* 20: 37-44

Kopriva S, Rennenberg H. 2004. Control of sulphate assimilation and glutathione synthesis: interaction with N and C metabolism. *Journal of Experimental Botany* 55: 1831-42

Kopriva S, Suter M, von Ballmoos P, Hesse H, Krähenbühl U, et al. 2002. Interaction of sulfate assimilation with carbon and nitrogen metabolism in *Lemna minor*. *Plant Physiology* 130: 1406-13

Koprivova A, Giovannetti M, Baraniecka P, Lee BR, Grondin C, et al. 2013. Natural variation in the ATPS1 isoform of ATP sulfurylase contributes to the control of sulfate levels in arabidopsis. *Plant Physiology* 163: 1133-41

Koprivova A, Meyer AJ, Schween G, Herschbach C, Reski R, Kopriva S. 2002. Functional knockout of the adenosine 5 '-phosphosulfate reductase gene in Physcomitrella patens revives an old route of sulfate assimilation. *Journal of Biological Chemistry* 277: 32195-201

Koprivova A, Mugford ST, Kopriva S. 2010. arabidopsis root growth dependence on glutathione is linked to auxin transport. *Plant Cell Reports* 29: 1157-67

Koprivova A, North KA, Kopriva S. 2008. Complex signaling network in regulation of adenosine 5'-phosphosulfate reductase by salt stress in arabidopsis roots. *Plant Physiology* 146: 1408-20

Koprivova A, Suter M, Op den Camp R, Brunold C, Kopriva S. 2000. Regulation of sulfate assimilation by nitrogen in arabidopsis. *Plant Physiology* 122: 737-46

Koralewska A, Posthumus FS, Stuiver CE, Buchner P, Hawkesford MJ, De Kok LJ. 2007. The characteristic high sulfate concentration in *Brassica oleracea* is controlled by the expression and activity of sulfate transporters. *Plant Biology* 9: 654-61

Korte A, Farlow A. 2013. The advantages and limitations of trait analysis with GWAS: a review. *Plant Methods* 9: 29

Kover PX, Valdar W, Trakalo J, Scarcelli N, Ehrenreich IM, et al. 2009. A multiparent advanced generation inter-cross to fine-map quantitative traits in *Arabidopsis thaliana*. *PLoS Genetics* 5: e1000551

Krouk G, Lacombe B, Bielach A, Perrine-Walker F, Malinska K, et al. 2010. Nitrate-regulated auxin transport by NRT1.1 defines a mechanism for nutrient sensing in plants. *Developemental Cell* 18: 927-37

Krueger S, Niehl A, Lopez Martin MC, Steinhauser D, Donath A, et al. 2009. Analysis of cytosolic and plastidic serine acetyltransferase mutants and subcellular metabolite distributions suggests interplay of the cellular compartments for cysteine biosynthesis in arabidopsis. *Plant, Cell, and Environment* 32: 349-67

Kruse T, Gehl C, Geisler M, Lehrke M, Ringel P, et al. 2010. Identification and biochemical characterization of molybdenum cofactor-binding proteins from *Arabidopsis thaliana*. *The Journal of Biological Chemistry* 285: 6623-35

Kuktaite R, Larsson H, Johansson E. 2004. Variation in protein composition of wheat flour and its relationship to dough mixing behaviour. *Journal of Cereal Science* 40: 31-39

Kuper J, Llamas A, Hecht HJ, Mendel RR, Schwarz G. 2004. Structure of the molybdopterin-bound Cnx1G domain links molybdenum and copper metabolism. *Nature* 430: 803-6

Kutz A, Muller A, Hennig P, Kaiser WM, Piotrowski M, Weiler EW. 2002. A role for nitrilase 3 in the regulation of root morphology in sulphur-starving *Arabidopsis thaliana*. *Plant Journal* 30: 95-106

Kvalheim OM, Brakstad F, Liang Y. 1994. Preprocessing of analytical profiles in the presence of homoscedastic or heteroscedastic noise. *Analytical Chemistry* 66: 43-51

Lahner B, Gong J, Mahmoudian M, Smith EL, Abid KB, et al. 2003. Genomic scale profiling of nutrient and trace elements in *Arabidopsis thaliana*. *Nature Biotechnology* 21: 1215-21

Lambein I, Chiba Y, Onouchi H, Naito S. 2003. Decay kinetics of autogenously regulated CGS1 mRNA that codes for cystathionine gamma-synthase in *Arabidopsis thaliana*. *Plant Cell Physiology* 44: 893-900

Lamers LP, Govers LL, Janssen IC, Geurts JJ, Van der Welle ME, et al. 2013. Sulfide as a soil phytotoxin-a review. *Frontiers in Plant Science* 4: 268

Lang C, Popko J, Wirtz M, Hell R, Herschbach C, et al. 2007. Sulphite oxidase as key enzyme for protecting plants against sulphur dioxide. *Plant, Cell, and Environment* 30: 447-55

Lappartient AG, Touraine B. 1996. Demand-driven control of root ATP sulfurylase activity and sulfate uptake in intact canola (the role of phloem-translocated glutathione). *Plant Physiology* 111: 147-57

Lappartient AG, Vidmar JJ, Leustek T, Glass ADM, Touraine B. 1999. Inter-organ signaling in plants: regulation of ATP sulfurylase and sulfate transporter genes expression in roots mediated by phloem-translocated compound. *Plant Journal* 18: 89-95

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, et al. 2007. Clustal W and Clustal X version 2.0. *Bioinformatics* 23: 2947-8

Lasky JR, Des Marais DL, McKay JK, Richards JH, Juenger TE, Keitt TH. 2012. Characterizing genomic variation of *Arabidopsis thaliana*: the roles of geography and climate. *Molecular Ecology* 21: 5512-29

Lass B, Ullrich-Eberius CI. 1984. Evidence for proton/sulfate cotransport and its kinetics in *Lemna gibba* G1. *Planta* 161: 53-60

Lee B-R, Huseby S, Koprivova A, Chételat A, Wirtz M, et al. 2012. Effects of *fou8/fry1* mutation on sulfur metabolism: is decreased internal sulfate the trigger of sulfate starvation response? *PLoS ONE* 7: e39425

Lee B-R, Koprivova A, Kopriva S. 2011. The key enzyme of sulfate assimilation, adenosine 5 '-phosphosulfate reductase, is regulated by HY5 in arabidopsis. *Plant Journal* 67: 1042-54

Leimkuhler S, Wuebbens MM, Rajagopalan KV. 2011. The history of the discovery of the molybdenum cofactor and novel aspects of its biosynthesis in bacteria. *Coordination Chemistry Reviews* 255: 1129-44

Lejay L, Gansel X, Cerezo M, Tillard P, Muller C, et al. 2003. Regulation of root ion transporters by photosynthesis: functional importance and relation with hexokinase. *Plant Cell* 15: 2218-32

Leulliot N, Blondeau K, Keller J, Ulryck N, Quevillon-Cheruel S, van Tilbeurgh H. 2010. Crystal structure of yeast FAD synthetase (Fad1) in complex with FAD. *Journal of Molecular Biology* 398: 641-6

Leustek T, Martin MN, Bick JA, Davies JP. 2000. Pathways and regulation of sulfur metabolism revealed through molecular and genetic studies. *Annual Reviews of Plant Biology* 51: 141-65

Lewandowska M, Sirko A. 2008. Recent advances in understanding plant response to sulfur-deficiency stress. *Acta Biochimica Polonica* 55: 457-71

Li X, Yan W, Agrama H, Jia L, Jackson A, et al. 2012. Unraveling the complex trait of harvest index with association mapping in rice (*Oryza sativa* L.). *PLoS ONE* 7: e29350

Li Y, Dankher OP, Carreira L, Smith AP, Meagher RB. 2006. The shoot-specific expression of gamma-glutamylcysteine synthetase directs the long-distance transport of thiol-peptides to roots conferring tolerance to mercury and arsenic. *Plant Physiology* 141: 288-98

Li Y, Huang Y, Bergelson J, Nordborg M, Borevitz JO. 2010. Association mapping of local climate-sensitive quantitative trait loci in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences USA* 107: 21199-204

Liang G, Yang F, Yu D. 2010. MicroRNA395 mediates regulation of sulfate accumulation and allocation in *Arabidopsis thaliana*. *Plant Journal* 62: 1046-57

Liang W, Li C, Liu F, Jiang H, Li S, et al. 2009. The arabidopsis homologs of CCR4-associated factor 1 show mRNA deadenylation activity and play a role in plant defence responses. *Cell Research* 19: 307-16

Lin S-I, Chiang S-F, Lin W-Y, Chen J-W, Tseng C-Y, et al. 2008. Regulatory network of microRNA399 and PHO2 by systemic signaling. *Plant Physiology* 147: 732-46

Linkohr BI, Williamson LC, Fitter AH, Leyser HM. 2002. Nitrate and phosphate availability and distribution have different effects on root system architecture of arabidopsis. *Plant Journal* 29: 751-60

Lipka AE, Tian F, Wang Q, Peiffer J, Li M, et al. 2012. GAPIT: genome association and prediction integrated tool. *Bioinformatics* 28: 2397-9

Lisec J, Schauer N, Kopka J, Willmitzer L, Fernie AR. 2006. Gas chromatography mass spectrometry-based metabolite profiling in plants. *Nature Protocols* 1: 387-96

Liu C, Martin E, Leyh TS. 1994. GTPase activation of ATP sulfurylase: the mechanism. *Biochemistry* 33: 2042-7

Liu CM, Meinke DW. 1998. The titan mutants of arabidopsis are disrupted in mitosis and cell cycle control during seed development. *Plant Journal* 16: 21-31

Llamas A, Mendel RR, Schwarz G. 2004. Synthesis of adenylated molybdopterin: an essential step for molybdenum insertion. *The Journal of Biological Chemistry* 279: 55241-6

Llamas A, Otte T, Multhaup G, Mendel RR, Schwarz G. 2006. The mechanism of nucleotide-assisted molybdenum insertion into molybdopterin. A novel route toward metal cofactor assembly. *The Journal of Biological Chemistry* 281: 18343-50

Lloyd JC, Zakhleniuk OV. 2004. Responses of primary and secondary metabolism to sugar accumulation revealed by microarray expression analysis of the arabidopsis mutant, pho3. *Journal of Experimental Botany* 55: 1221-30

Logan HM, Cathala N, Grignon C, Davidian JC. 1996. Cloning of a cDNA encoded by a member of the *Arabidopsis thaliana* ATP sulfurylase multigene family. Expression studies in yeast and in relation to plant sulfur nutrition. *Journal of Biological Chemistry* 271: 12227-33

Lopez-Bucio J, Cruz-Ramirez A, Herrera-Estrella L. 2003. The role of nutrient availability in regulating root architecture. *Current Opinions in Plant Biology* 6: 280-7

Lopez-Bucio J, Hernandez-Abreu E, Sanchez-Calderon L, Nieto-Jacobo MF, Simpson J, Herrera-Estrella L. 2002. Phosphate availability alters architecture and causes changes in hormone sensitivity in the arabidopsis root system. *Plant Physiology* 129: 244-56

López-Martín MC, Becana M, Romero LC, Gotor C. 2008. Knocking out cytosolic cysteine synthesis compromises the antioxidant capacity of the cytosol to maintain discrete concentrations of hydrogen peroxide in arabidopsis. *Plant Physiology* 147: 562-72

Loudet O, Chaillou S, Merigout P, Talbotec J, Daniel-Vedele F. 2003. Quantitative trait loci analysis of nitrogen use efficiency in arabidopsis. *Plant Physiology* 131: 345-58

Loudet O, Saliba-Colombani V, Camilleri C, Calenge F, Gaudon V, et al. 2007. Natural variation for sulfate concentration in *Arabidopsis thaliana* is highly controlled by APR2. *Nature Genetics* 39: 896-900

Luedemann A, Strassburg K, Erban A, Kopka J. 2008. TagFinder for the quantitative analysis of gas chromatography--mass spectrometry (GC-MS)-based metabolite profiling experiments. *Bioinformatics* 24: 732-7

Lunde C, Zygadlo A, Simonsen HT, Nielsen PL, Blennow A, Haldrup A. 2008. Sulfur starvation in rice: the effect on photosynthesis, carbohydrate metabolism, and oxidative stress protective pathways. *Physiology of Plant* 134: 508-21

Lunn JE, Droux M, Martin J, Douce R. 1990. Localization of ATP sulfurylase and O-acetylserine(thiol)lyase in spinach leaves. *Plant Physiology* 94: 1345-52

Macleod JA, Gupta UC, Stanfield B. 1997. Molybdenum and sulfur relationships in plants. *Cambridge University Press*. 229-44 pp.

Malhi SS, Gan Y, Raney JP. 2007. Yield, seed quality, and sulfur uptake of Brassica oilseed crops in response to sulfur fertilization. *Agronomy Journal* 99: 570-77

Malitsky S, Blum E, Less H, Venger I, Elbaz M, et al. 2008. The transcript and metabolite networks affected by the two clades of arabidopsis glucosinolate biosynthesis regulators. *Plant Physiology* 148: 2021-49

Marchler-Bauer A, Lu S, Anderson JB, Chitsaz F, Derbyshire MK, et al. 2011. CDD: a Conserved domain database for the functional annotation of proteins. *Nucleic Acids Research* 39: D225-9

Martin MN, Tarczynski MC, Shen B, Leustek T. 2005. The role of 5'-adenylylsulfate reductase in controlling sulfate reduction in plants. *Photosynthesis Research* 86: 309-23

Martinoia E, Massonneau A, Frangne N. 2000. Transport processes of solutes across the vacuolar membrane of higher plants. *Plant & Cell Physiology* 41: 1175-86

Maruyama-Nakashita A, Inoue E, Watanabe-Takahashi A, Yamaya T, Takahashi H. 2003. Transcriptome profiling of sulfur-responsive genes in arabidopsis reveals global effects of sulfur nutrition on multiple metabolic pathways. *Plant Physiology* 132: 597-605

Maruyama-Nakashita A, Nakamura Y, Tohge T, Saito K, Takahashi H. 2006. arabidopsis SLIM1 is a central transcriptional regulator of plant sulfur response and metabolism. *The Plant Cell* 18: 3235-51

Maruyama-Nakashita A, Nakamura Y, Watanabe-Takahashi A, Inoue E, Yamaya T, Takahashi H. 2005. Identification of a novel cis-acting element conferring sulfur deficiency response in arabidopsis roots. *Plant Journal* 42: 305-14

Maruyama-Nakashita A, Nakamura Y, Watanabe-Takahashi A, Yamaya T, Takahashi H. 2004a. Induction of SULTR1;1 sulfate transporter in arabidopsis roots involves protein phosphorylation/dephosphorylation circuit for transcriptional regulation. *Plant Cell Physiology* 45: 340-5

Maruyama-Nakashita A, Nakamura Y, Yamaya T, Takahashi H. 2004b. A novel regulatory pathway of sulfate uptake in arabidopsis roots: implication of CRE1/WOL/AHK4-mediated cytokinin-dependent regulation. *Plant Journal* 38: 779-89

Masclaux-Daubresse C, Daniel-Vedele F, Dechorgnat J, Chardon F, Gaufichon L, Suzuki A. 2010. Nitrogen uptake, assimilation and remobilization in plants: challenges for sustainable and productive agriculture. *Annals of Botany* 105: 1141-57

Massey V. 2000. The chemical and biological versatility of riboflavin. *Biochemical Society Transactions* 28: 283-96

Matsubayashi Y, Sakagami Y. 1996. Phytosulfokine, sulfated peptides that induce the proliferation of single mesophyll cells of *Asparagus officinalis* L. *Proceedings of the National Academy of Sciences USA* 93: 7623-7

Matsubayashi Y, Sakagami Y. 2006. Peptide hormones in plants. *Annual Reviews in Plant Biology* 57: 649-74

Matsubayashi Y, Takagi L, Sakagami Y. 1997. Phytosulfokine-alpha, a sulfated pentapeptide, stimulates the proliferation of rice cells by means of specific high- and low-affinity binding sites. *Proceedings of the National Academy of Sciences USA* 94: 13357-62

Matthewman CA, Kawashima CG, Huska D, Csorba T, Dalmay T, Kopriva S. 2012. miR395 is a general component of the sulfate assimilation regulatory network in arabidopsis. *FEBS Letters* 586: 3242-48

Mba C, Guimaraes EP, Ghosh K. 2012. Re-orienting crop improvement for the changing climatic conditions of the 21st century. *Agriculture and Food Security* 1:7

McGrath SP, Zhao FJ. 1996. Sulphur uptake, yield responses and the interactions between nitrogen and sulphur in winter oilseed rape (*Brassica napus*). *The Journal of Agricultural Science* 126: 53-62

McGrath SP, Zhao FJ, Withers PJA. 1996. Development of sulphur deficiency in crops and its treatment. *Plants* 233(4):649-60

McKay JK, Richards JH, Mitchell-Olds T. 2003. Genetics of drought adaptation in *Arabidopsis thaliana*: I. Pleiotropy contributes to genetic correlations among ecological traits. *Molecular Ecology* 12: 1137-51

McKhann HI, Camilleri C, Bérard A, Bataillon T, David JL, et al. 2004. Nested core collections maximizing genetic diversity in *Arabidopsis thaliana*. *Plant Journal* 38: 193-202

Mendel RR. 2011. Cell biology of molybdenum in plants. *Plant Cell Reports* 30: 1787-97

Mendel RR. 2013. The molybdenum cofactor. *Journal of Biological Chemistry* 288: 13165-72

Mendel RR, Hansch R. 2002. Molybdoenzymes and molybdenum cofactor in plants. *Journal of Experimental Botany* 53: 1689-98

Mendel RR, Kruse T. 2012. Cell biology of molybdenum in plants and humans. *Biochimica et Biophysica Acta* 1823: 1568-79

Mendel RR, Schwarz G. 2011. Molybdenum cofactor biosynthesis in plants and humans. *Coordination Chemistry Reviews* 255: 1145-58

Miccolis A, Galluccio M, Giancaspero TA, Indiveri C, Barile M. 2012. Bacterial over-expression and purification of the 3'phosphoadenosine 5'phosphosulfate (PAPS) reductase domain of human FAD synthase: functional characterization and homology modeling. *International Journal of Molecular Sciences* 13: 16880-98

Millar LJ, Heck IS, Sloan J, Kana'n GJ, Kinghorn JR, Unkles SE. 2001. Deletion of the cnxE gene encoding the gephyrin-like protein involved in the final stages of molybdenum cofactor biosynthesis in *Aspergillus nidulans*. *Molecular Genetics and Genomics* 266: 445-53

Misson J, Raghothama KG, Jain A, Jouhet J, Block MA, et al. 2005. A genome-wide transcriptional analysis using *Arabidopsis thaliana* Affymetrix gene chips determined plant responses to phosphate deprivation. *Proceedings of the National Academy of Sciences USA* 102: 11934-9

Mithen R, Faulkner K, Magrath R, Rose P, Williamson G, Marquez J. 2003. Development of isothiocyanate-enriched broccoli, and its enhanced ability to induce phase 2 detoxification enzymes in mammalian cells. *Theoretical Applied Genetics* 106: 727-34

Montesinos A, Tonsor SJ, Alonso-Blanco C, Pico FX. 2009. Demographic and genetic patterns of variation among populations of *Arabidopsis thaliana* from contrasting native environments. *PLoS ONE* 4: e7213

Moose SP, Mumm RH. 2008. Molecular plant breeding as the foundation for 21st century crop improvement. *Plant Physiology* 147: 969-77

Moss H, Wrigley C, MacRichie R, Randall P. 1981. Sulfur and nitrogen fertilizer effects on wheat. II. Influence on grain quality. *Australian Journal of Agricultural Research* 32: 213-26

Mougous JD, Lee DH, Hubbard SC, Schelle MW, Vocadlo DJ, et al. 2006. Molecular basis for G protein control of the prokaryotic ATP sulfurylase. *Molecular Cell* 21: 109-22

Mugford SG, Lee B-R, Koprivova A, Matthewman C, Kopriva S. 2011. Control of sulfur partitioning between primary and secondary metabolism. *Plant Journal* 65: 96-105

Mugford SG, Yoshimoto N, Reichelt M, Wirtz M, Hill L, et al. 2009. Disruption of adenosine-5'-phosphosulfate kinase in arabidopsis reduces levels of sulfated secondary metabolites. *Plant Cell* 21: 910-27

Murillo M, Leustek T. 1995. Adenosine-5'-triphosphate-sulfurylase from *Arabidopsis thaliana* and *Escherichia coli* are functionally equivalent but structurally and kinetically divergent: nucleotide sequence of two adenosine-5'-triphosphate-sulfurylase cDNAs from *Arabidopsis thaliana* and analysis of a recombinant enzyme. *Archives of Biochemistry and Biophysics* 323: 195-204

Myles S, Peiffer J, Brown PJ, Ersoz ES, Zhang Z, et al. 2009. Association mapping: critical considerations shift from genotyping to experimental design. *Plant Cell* 21: 2194-202

Nagarajan VK, Smith AP. 2012. Ethylene's role in phosphate starvation signaling: more than just a root growth regulator. *Plant & Cell Physiology* 53: 277-86

Nakayama M, Akashi T, Hase T. 2000. Plant sulfite reductase: molecular structure, catalytic function and interaction with ferredoxin. *Journal of Inorganic Biochemistry* 82: 27-32

Nam HG, Giraudat J, Den Boer B, Moonan F, Loos W, et al. 1989. Restriction fragment length polymorphism linkage map of *Arabidopsis thaliana*. *Plant Cell* 1: 699-705

Nawrotzki R, Islinger M, Vogel I, Volkl A, Kirsch J. 2012. Expression and subcellular distribution of gephyrin in non-neuronal tissues and cells. *Histochemistry and Cell Biology* 137: 471-82

Nemri A, Atwell S, Tarone AM, Huang YS, Zhao K, et al. 2010. Genome-wide survey of arabidopsis natural variation in downy mildew resistance using combined association and linkage mapping. *Proceedings of the National Academy of Sciences USA* 1;107(22):10302-7

Nicholas DJ, Nason A. 1955. Role of molybdenum as a constituent of nitrate reductase from soybean leaves. *Plant Physiology* 30: 135-8

Nikiforova V, Freitag J, Kempa S, Adamik M, Hesse H, Hoefgen R. 2003. Transcriptome analysis of sulfur depletion in *Arabidopsis thaliana*: interlacing of biosynthetic pathways provides response specificity. *Plant Journal* 33: 633-50

Nikiforova VJ, Kopka J, Tolstikov V, Fiehn O, Hopkins L, et al. 2005. Systems rebalancing of metabolism in response to sulfur deprivation, as revealed by metabolome analysis of arabidopsis plants. *Plant Physiology* 138: 304-18

Nimni M, Han B, Cordoba F. 2007. Are we getting enough sulfur in our diet? *Nutrition & Metabolism* 4: 24

Noble WS. 2009. How does multiple testing correction work? *Nature Biotechnology* 27: 1135-37

Noctor G, Arisi A-CM, Jouanin L, Kunert KJ, Rennenberg H, Foyer CH. 1998. Glutathione: biosynthesis, metabolism and relationship to stress tolerance explored in transformed plants. *Journal of Experimental Botany* 49: 623-47

Noctor G, Gomez L, Vanacker H, Foyer CH. 2002. Interactions between biosynthesis, compartmentation and transport in the control of glutathione homeostasis and signalling. *Journal of Experimental Botany* 53: 1283-304

Noji M, Inoue K, Kimura N, Gouda A, Saito K. 1998. Isoform-dependent differences in feedback regulation and subcellular localization of serine acetyltransferase involved in cysteine biosynthesis from *Arabidopsis thaliana*. *Journal of Biological Chemistry* 273: 32739-45

Nordborg M, Weigel D. 2008. Next-generation genetics in plants. *Nature* 456: 720-3

North KA, Ehlting B, Koprivova A, Rennenberg H, Kopriva S. 2009. Natural variation in arabidopsis adaptation to growth at low nitrogen conditions. *Plant Physiology and Biochemistry* 47: 912-8

Nowak K, Luniak N, Witt C, Wustefeld Y, Wachter A, et al. 2004. Peroxisomal localization of sulfite oxidase separates it from chloroplast-based sulfur assimilation. *Plant & Cell Physiology* 45: 1889-94

Ohkama-Ohtsu N, Oikawa A, Zhao P, Xiang C, Saito K, Oliver DJ. 2008. A gamma-glutamyl transpeptidase-independent pathway of glutathione catabolism to glutamate via 5-oxoproline in arabidopsis. *Plant Physiology* 148: 1603-13

Ohkama N, Takei K, Sakakibara H, Hayashi H, Yoneyama T, Fujiwara T. 2002. Regulation of sulfur-responsive gene expression by exogenously applied cytokinins in *Arabidopsis thaliana*. *Plant & Cell Physiology* 43: 1493-501

Ohto M, Onai K, Furukawa Y, Aoki E, Araki T, Nakamura K. 2001. Effects of sugar on vegetative development and floral transition in arabidopsis. *Plant Physiology* 127: 252-61

Olsen HG, Hayes BJ, Kent MP, Nome T, Svendsen M, et al. 2011. Genome-wide association mapping in Norwegian Red cattle identifies quantitative trait loci for fertility and milk production on BTA12. *Animal Genetics* 42: 466-74

Olsen LR, Huang B, Vetting MW, Roderick SL. 2004. Structure of serine acetyltransferase in complexes with CoA and its cysteine feedback inhibitor. *Biochemistry* 43: 6013-9

Ouyang Y, Lane WS, Moore KL. 1998. Tyrosylprotein sulfotransferase: purification and molecular cloning of an enzyme that catalyzes tyrosine O-sulfation, a common posttranslational modification of eukaryotic proteins. *Proceedings of the National Academy of Sciences USA* 95: 2896-901

Pant BD, Buhtz A, Kehr J, Scheible W-R. 2008. MicroRNA399 is a long-distance signal for the regulation of plant phosphate homeostasis. *Plant Journal* 53: 731-38

Parida A, Das AB, Mittra B. 2004. Effects of salt on growth, ion accumulation, photosynthesis and leaf anatomy of the mangrove, *Bruguiera parviflora*. *Trees* 18: 167-74

Patton DA, Schetter AL., Franzmann LH, Nelson K, Ward ER, Meinke DW. 1998. An embryo-defective mutant of arabidopsis disrupted in the final step of biotin synthesis. *Plant Physiology* 116: 935-46

Paul MJ, Driscoll SP. 1997. Sugar repression of photosynthesis: the role of carbohydrates in signalling nitrogen deficiency through source:sink imbalance. *Plant, Cell & Environment* 20: 110-16

Pico FX, Mendez-Vigo B, Martinez-Zapater JM, Alonso-Blanco C. 2008. Natural genetic variation of *Arabidopsis thaliana* is geographically structured in the Iberian peninsula. *Genetics* 180: 1009-21

Pinkerton A. 1998. Critical sulfur concentrations in oilseed rape (*Brassica napus*) in relation to nitrogen supply and to plant age. *Australian Journal of Experimental Agriculture* 38: 511-22

Platt A, Horton M, Huang YS, Li Y, Anastasio AE, et al. 2010. The scale of population structure in *Arabidopsis thaliana*. *PLoS Genetics* 6: e1000843

Plitzko B, Ott G, Reichmann D, Henderson CJ, Wolf CR, et al. 2013. The involvement of mitochondrial amidoxime reducing components 1 and 2 and mitochondrial cytochrome b5 in N-reductive metabolism in human cells. *The Journal of Biological Chemistry* 288: 20228-37

Prosser IM, Purves JV, Saker LR, Clarkson DT. 2001. Rapid disruption of nitrogen metabolism and nitrate transport in spinach plants deprived of sulphate. *Journal of Experimental Botany* 52: 113-21

Rae A, Smith F. 2002. Localisation of expression of a high-affinity sulfate transporter in barley roots. *Planta* 215: 565-68

Rajagopalan KV, Johnson JL. 1992. The pterin molybdenum cofactors. *The Journal of Biological Chemistry* 267: 10199-202

Randewig D, Hamisch D, Herschbach C, Eiblmeier M, Gehl C, et al. 2012. Sulfite oxidase controls sulfur metabolism under SO_2 exposure in *Arabidopsis thaliana*. *Plant, Cell, and Environment* 35: 100-15

Ravanel S, Block MA, Rippert P, Jabrin S, Curien G, et al. 2004. Methionine metabolism in plants: chloroplasts are autonomous for *de novo* methionine synthesis and can import S-adenosylmethionine from the cytosol. *Journal of Biological Chemistry* 279: 22548-57

Ravanel S, Gakière B, Job D, Douce R. 1998. The specific features of methionine biosynthesis and metabolism in plants. *Proceedings of the National Academy of Sciences USA* 95: 7805-12

Reiss J, Cohen N, Dorche C, Mandel H, Mendel RR, et al. 1998. Mutations in a polycistronic nuclear gene associated with molybdenum cofactor deficiency. *Nature Genetics* 20: 51-3

Rennenberg H. 1984. The fate of excess sulfur in higher plants. *Annual Reviews in Plant Biology* 35: 121-53

Rennenberg H, Schmitz K, Bergmann L. 1979. Long-distance transport of sulfur in *Nicotiana tabacum*. *Planta* 147: 57-62

Renosto F, Patel HC, Martin RL, Thomassian C, Zimmerman G, Segel IH. 1993. ATP sulfurylase from higher plants: kinetic and structural characterization of the chloroplast and cytosol enzymes from spinach leaf. *Archievs of Biochemistry and Biophysics* 307: 272-85

Renwick JA. 2001. Variable diets and changing taste in plant-insect relationships. *Journal of Chemical Ecology* 27: 1063-76

Reuveny Z. 1977. Derepression of ATP sulfurylase by the sulfate analogs molybdate and selenate in cultured tobacco cells. *Proceedings of the National Academy of Sciences USA* 74: 619-22

Richard-Molard C, Krapp A, Brun F, Ney B, Daniel-Vedele F, Chaillou S. 2008. Plant response to nitrate starvation is determined by N storage capacity matched by nitrate uptake capacity in two arabidopsis genotypes. *Journal of Experimental Botany* 59: 779-91

Rizzi M, Schindelin H. 2002. Structural biology of enzymes involved in NAD and molybdenum cofactor biosynthesis. *Current Opinion in Structural Biology* 12: 709-20

Roessner U, Wagner C, Kopka J, Trethewey RN, Willmitzer L. 2000. Technical advance: simultaneous analysis of metabolites in potato tuber by gas chromatography-mass spectrometry. *Plant Journal* 23: 131-42

Roldan M, Dinh P, Leung S, McManus MT. 2013. Ethylene and the responses of plants to phosphate deficiency. *AoB Plants* 5: plt013

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

Rossato L, Lainé P, Ourry A. 2001. Nitrogen storage and remobilization in *Brassica napus* L. during the growth cycle: nitrogen fluxes within the plant and changes in soluble protein patterns. *Journal of Experimental Botany* 52: 1655-63

Rothstein SJ. 2007. Returning to our roots: making plant biology research relevant to future challenges in agriculture. *The Plant Cell* 19: 2695-99

Rotte C, Leustek T. 2000. Differential subcellular localization and expression of ATP sulfurylase and 5'-adenylylsulfate reductase during ontogenesis of arabidopsis leaves indicates that cytosolic and plastid forms of ATP sulfurylase may have specialized functions. *Plant Physiology* 124: 715-24

Rouached H. 2011. Multilevel coordination of phosphate and sulfate homeostasis in plants. *Plant Signalling & Behavior* 6: 952-5

Rouached H, Berthomieu P, El Kassis E, Cathala N, Catherinot V, et al. 2005. Structural and functional analysis of the C-terminal STAS (Sulfate Transporter and Anti-sigma Antagonist) domain of the *Arabidopsis thaliana* sulfate transporter SULTR1.2. *Journal of Biological Chemistry* 280: 15976-83

Rouached H, Wirtz M, Alary R, Hell R, Arpat AB, et al. 2008. Differential regulation of the expression of two high-affinity sulfate transporters, SULTR1.1 and SULTR1.2, in arabidopsis. *Plant Physiology* 147: 897-911

Rubio V, Linhares F, Solano R, Martin AC, Iglesias J, et al. 2001. A conserved MYB transcription factor involved in phosphate starvation signaling both in vascular plants and in unicellular algae. *Genes & Development* 15: 2122-33

Saito K. 2000. Regulation of sulfate transport and synthesis of sulfur-containing amino acids. *Current Opinion in Plant Biology* 3: 188-95

Saito K. 2004. Sulfur assimilatory metabolism. The long and smelling road. *Plant Physiology* 136: 2443-50

Sakakibara H, Takei K, Hirose N. 2006. Interactions between nitrogen and cytokinin in the regulation of metabolism and development. *Trends in Plant Science* 11: 440-48

Salt DE. 2004. Update on plant ionomics. *Plant Physiology* 136: 2451-56

Sambrook J, Fritsch EF, Maniatis T. 1989. Molecular Cloning: A Laboratory Manual. *Cold Spring Harbor Laboratory Press*.

Sandoval FJ, Zhang Y, Roje S. 2008. Flavin nucleotide metabolism in plants: monofunctional enzymes synthesize FAD in plastids. *The Journal of Biological Chemistry* 283: 30890-900

Sarowar S, Oh HW, Cho HS, Baek KH, Seong ES, et al. 2007. *Capsicum annuum* CCR4-associated factor CaCAF1 is necessary for plant development and defence response. *Plant Journal* 51: 792-802

Savage H, Montoya G, Svensson C, Schwenn JD, Sinning I. 1997. Crystal structure of phosphoadenylyl sulphate (PAPS) reductase: a new family of adenine nucleotide alpha hydrolases. *Structure* 5: 895-906

Scaife A, Burns IG. 1986. The sulfate-S total S-ratio in plants as an index of their sulfur status. *Plant and Soil* 91: 61-71

Schauer N, Steinhauser D, Strelkov S, Schomburg D, Allison G, et al. 2005. GC-MS libraries for the rapid identification of metabolites in complex biological samples. *FEBS Letters* 579: 1332-7

Scheerer U, Haensch R, Mendel RR, Kopriva S, Rennenberg H, Herschbach C. 2010. Sulphur flux through the sulphate assimilation pathway is differently controlled by adenosine 5'-phosphosulphate reductase under stress and in transgenic poplar plants overexpressing γ-ECS, SO, or APR. *Journal of Experimental Botany* 61: 609-22

Scherer HW. 2001. Sulphur in crop production. *European Journal of Agronomy* 14: 81-111

Schiavon M, Pittarello M, Pilon-Smits EAH, Wirtz M, Hell R, Malagoli M. 2012. Selenate and molybdate alter sulfate transport and assimilation in *Brassica juncea* L. Czern.: Implications for phytoremediation. *Environmental and Experimental Botany* 75: 41-51

Schmidt A. 1975. Distribution of APS-sulfotransferase activity among higher plants. *Plant Science Letters* 5: 407-15

Schmidt A, Jager K. 1992. Open questions about sulfur metabolism in plants. *Annual Reviews in Plant Biology* 43: 325-49

Schmidt A, Trüper HG. 1977. Reduction of adenylylsulfate and 3'-phosphoadenylylsulfate in phototrophic bacteria. *Experientia* 33: 1008-10

Schmidt UG, Endler A, Schelbert S, Brunner A, Schnell M, et al. 2007. Novel tonoplast transporters identified using a proteomic approach with vacuoles isolated from cauliflower buds. *Plant Physiology* 145: 216-29

Schrader N, Fischer K, Theis K, Mendel RR, Schwarz G, Kisker C. 2003. The crystal structure of plant sulfite oxidase provides insights into sulfite oxidation in plants and animals. *Structure* 11: 1251-63

Schwab R, Ossowski S, Riester M, Warthmann N, Weigel D. 2006. Highly specific gene silencing by artificial microRNAs in arabidopsis. *Plant Cell* 18: 1121-33

Schwab R, Ossowski S, Warthmann N, Weigel D. 2010. Directed gene silencing with artificial microRNAs. *Methods in Molecular Biology* 592: 71-88

Schwarz G, Mendel RR. 2006. Molybdenum cofactor biosynthesis and molybdenum enzymes. *Annual Review of Plant Biology* 57: 623-47

Schwarz G, Santamaria-Araujo JA, Wolf S, Lee HJ, Adham IM, et al. 2004. Rescue of lethal molybdenum cofactor deficiency by a biosynthetic precursor from *Escherichia coli*. *Human Molecular Genetics* 13: 1249-55

Schwarz G, Schulze J, Bittner F, Eilers T, Kuper J, et al. 2000. The molybdenum cofactor biosynthetic protein Cnx1 complements molybdate-repairable mutants, transfers molybdenum to the metal binding pterin, and is associated with the cytoskeleton. *Plant Cell* 12: 2455-72

Schwenn JD. 1989. Sulfate assimilation in higher plants – a thioredoxin-dependent PAPS reductase from spinach leaves. *Journal of Biosciences* 44: 504-08

Seefeldt LC, Hoffman BM, Dean DR. 2009. Mechanism of Mo-dependent nitrogenase. *Annual Review of Biochemistry* 78: 701-22

Sengupta S, Mukherjee S, Goswami L, Sangma S, Mukherjee A, et al. 2012. Manipulation of inositol metabolism for improved plant survival under stress: a "network engineering approach". *Journal of Plant Biochemistry and Biotechnology* 21: S15-S23

Seo M, Koshiba T. 2002. Complex regulation of ABA biosynthesis in plants. *Trends Plant Science* 7: 41-8

Shahsavani S, Gholami A. 2008. Effect of sulphur fertilization on breadmaking quality of three winter wheat varieties. *Pakistan Journal of Biological Sciences* 11: 2134-8

Shibagaki N, Grossman AR. 2004. Probing the function of STAS domains of the arabidopsis sulfate transporters. *Journal of Biological Chemistry* 279: 30791-99

Shibagaki N, Rose A, McDermott JP, Fujiwara T, Hayashi H, et al. 2002. Selenate-resistant mutants of *Arabidopsis thaliana* identify SULTR1;2, a sulfate transporter required for efficient transport of sulfate into roots. *Plant Journal* 29: 475-86

Shinmachi F, Buchner P, Stroud JL, Parmar S, Zhao FJ, et al. 2010. Influence of sulfur deficiency on the expression of specific sulfate transporters and the distribution of sulfur, selenium, and molybdenum in wheat. *Plant Physiology* 153: 327-36

Smeekens S, Hellmann HA. 2014. Sugar sensing and signaling in plants. *Frontiers in Plant Science* 5: 113

Smith FW, Ealing PM, Hawkesford MJ, Clarkson DT. 1995a. Plant members of a family of sulfate transporters reveal functional subtypes. *Proceedings of the National Academy of Sciences USA* 92: 9373-77

Smith FW, Hawkesford MJ, Ealing PM, Clarkson DT, van den Berg PJ, et al. 1997. Regulation of expression of a cDNA from barley roots encoding a high affinity sulphate transporter. *Plant Journal* 12: 875-84

Smith FW, Hawkesford MJ, Prosser IM, Clarkson DT. 1995b. Isolation of a cDNA from *Saccharomyces cerevisiae* that encodes a high affinity sulphate transporter at the plasma membrane. *Molecular Genetics and Genomics* 247: 709-15

Smith IK. 1980. Regulation of sulfate assimilation in tobacco cells: effect of nitrogen and sulfur nutrition onn sulfate permease and O-acetylserine sulfhydrylase. *Plant Physiology* 66: 877-83

Smolders E, Merckx R. 1992. Growth and shoot:root partitioning of spinach plants as affected by nitrogen supply. *Plant, Cell & Environment* 15: 795-807

Sønderby IE, Hansen BG, Bjarnholt N, Ticconi C, Halkier BA, Kliebenstein DJ. 2007. A systems biology approach identifies a R2R3 MYB gene subfamily with distinct and overlapping functions in regulation of aliphatic glucosinolates. *PLoS ONE* 2: e1322

Sparkes IA, Brandizzi F, Slocombe SP, El-Shami M, Hawes C, Baker A. 2003. An arabidopsis pex10 null mutant is embryo lethal, implicating peroxisomes in an essential role during plant embryogenesis. *Plant Physiology* 133: 1809-19

Spencer D, Wood JG. 1954. The role of molybdenum in nitrate reduction in higher plants. *Australian Journal of Biological Sciences* 7: 425-34

Stallmeyer B, Drugeon G, Reiss J, Haenni AL., Mendel RR. 1999. Human molybdopterin synthase gene: identification of a bicistronic transcript with overlapping reading frames. *American Journal of Human Genetics* 64: 698-705

Stallmeyer B, Nerlich A, Schiemann J, Brinkmann H, Mendel RR. 1995. Molybdenum co-factor biosynthesis: the *Arabidopsis thaliana* cDNA cnx1 encodes a multifunctional two-domain protein homologous to a mammalian neuroprotein, the insect protein Cinnamon and three *Escherichia coli* proteins. *Plant Journal* 8: 751-62

Stout PR, Meagher WR, Pearson GA, Johnson CM. 1951. Molybdenum nutrition of crop plants. 1. The influence of phosphate and sulfate on the absorption of molybdenum from soil and solution cultures. *Plant and Soil* 3: 51-87

Strehmel N, Hummel J, Erban A, Strassburg K, Kopka J. 2008. Retention index thresholds for compound matching in GC-MS metabolite profiling. *Journal of Chromatography.* 871: 182-90

Strohm M, Jouanin L, Kunert KJ, Pruvost C, Polle A, et al. 1995. Regulation of glutathione synthesis in leaves of transgenic poplar (*Populus tremula* X *P. alba*) overexpressing glutathione synthetase. *Plant Journal* 7: 141-45

Stuiver CEE, DeKok LJ, Westerman S. 1997. Sulfur deficiency in *Brassica oleracea* L: development, biochemical characterization, and sulfur/nitrogen interactions. *Russian Journal of Plant Physiology* 44: 505-13

Sugimoto K, Sato N, Tsuzuki M. 2007. Utilization of a chloroplast membrane sulfolipid as a major internal sulfur source for protein synthesis in the early phase of sulfur starvation in *Chlamydomonas reinhardtii*. *FEBS Letters* 581: 4519-22

Sun J, Gibson KM, Kiirats O, Okita TW, Edwards GE. 2002. Interactions of nitrate and CO_2 enrichment on growth, carbohydrates, and Rubisco in arabidopsis starch mutants. Significance of starch and hexose. *Plant Physiology* 130: 1573-83

Suzuki A, Shirata Y, Ishida H, Chiba Y, Onouchi H, Naito S. 2001. The first exon coding region of cystathionine gamma-synthase gene is necessary and sufficient for downregulation of its own mRNA accumulation in transgenic *Arabidopsis thaliana*. *Plant & Cell Physiology* 42: 1174-80

Takahashi H, Kopriva S, Giordano M, Saito K, Hell R. 2011. Sulfur assimilation in photosynthetic organisms: molecular functions and regulations of transporters and assimilatory enzymes. *Annual Review of Plant Biology 62:* 157-84

Takahashi H, Saito K. 2008. Molecular biology and functional genomics for identification of regulatory networks of plant sulfate uptake and assimilatory metabolism. *Sulfur Metabolism in Phototrophic Organisms*, ed. R Hell, C Dahl, D Knaff, T Leustek, pp. 149-59: Springer Netherlands

Takahashi H, Watanabe-Takahashi A, Smith FW, Blake-Kalff M, Hawkesford MJ, Saito K. 2000. The roles of three functional sulphate transporters involved in uptake and translocation of sulphate in *Arabidopsis thaliana*. *Plant Journal* 23: 171-82

Takahashi H, Yamazaki M, Sasakura N, Watanabe A, Leustek T, et al. 1997. Regulation of sulfur assimilation in higher plants: A sulfate transporter induced in sulfate-starved roots plays a central role in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences USA* 94: 11102-07

Tanz SK, Castleden I, Hooper CM, Vacher M, Small I, Millar HA. 2013. SUBA3: a database for integrating experimentation and prediction to define the SUBcellular location of proteins in arabidopsis. *Nucleic Acids Research* 41: D1185-91

Tareke E, Rydberg P, Karlsson P, Eriksson S, Tornqvist M. 2002. Analysis of acrylamide, a carcinogen formed in heated foodstuffs. *Journal of Agricultural and Food Chemistry* 50: 4998-5006

Tavares S, Sousa C, Carvalho LC, Amancio S. 2008. Derepressed sulfate transporters are strongly and rapidly repressed after sulfate addition to sulfur-depleted *Vittis* cells. *International Journal of Plant Sciences* 169: 987-97

Teschner J, Lachmann N, Schulze J, Geisler M, Selbach K, et al. 2010. A novel role for arabidopsis mitochondrial ABC transporter ATM3 in molybdenum cofactor biosynthesis. *Plant Cell* 22: 468-80

Tester M, Langridge P. 2010. Breeding technologies to increase crop production in a changing world. *Science* 327: 818-22

Thudi M, Upadhyaya HD, Rathore A, Gaur PM, Krishnamurthy L, et al. 2014. Genetic dissection of drought and heat tolerance in chickpea through genome-wide and candidate gene-based association mapping approaches. *PLoS ONE* 9: e96758

Tian F, Bradbury PJ, Brown PJ, Hung H, Sun Q, et al. 2011. Genome-wide association study of leaf architecture in the maize nested association mapping population. *Nature Genetics* 43: 159-62

Todesco M, Balasubramanian S, Hu TT, Traw MB, Horton M, et al. 2010. Natural allelic variation underlying a major fitness trade-off in *Arabidopsis thaliana*. *Nature* 465: 632-36

Tomatsu H, Takano J, Takahashi H, Watanabe-Takahashi A, Shibagaki N, Fujiwara T. 2007. An *Arabidopsis thaliana* high-affinity molybdate transporter required for efficient uptake of molybdate from soil. *Proceedings of the National Academy of Sciences USA* 104: 18807-12

Tomlinson I. 2013. Doubling food production to feed the 9 billion: A critical perspective on a key discourse of food security in the UK. *Journal of Rural Studies* 29: 81-90

Torchetti EM, Bonomi F, Galluccio M, Gianazza E, Giancaspero TA, et al. 2011. Human FAD synthase (isoform 2): a component of the machinery that delivers FAD to apo-flavoproteins. *FEBS Journal* 278: 4434-49

Tsang MLS, Goldschm EE, Schiff JA. 1971. Adenosine 5'-phosphosulfate (APS35) as an intermediate in conversion of adenosine 3'-phosphate 5'-phosphate (PAPS35) to acid volatile radioactivity. *Plant Physiology* 47: 20

Tucker M, Valencia-Sanchez MA, Staples RR, Chen J, Denis CL, Parker R. 2001. The transcription factor associated Ccr4 and Caf1 proteins are components of the major cytoplasmic mRNA deadenylase in *Saccharomyces cerevisiae*. *Cell* 104: 377-86

Valluru R, Van den Ende W. 2011. Myo-inositol and beyond - emerging networks under stress. *Plant Science* 181: 387-400

van den Berg RA, Hoefsloot HC, Westerhuis JA, Smilde AK, van der Werf MJ. 2006. Centering, scaling, and transformations: improving the biological information concentration of metabolomics data. *BMC Genomics* 7: 142

Van Larebeke N, Zaenen I, Teuchy H, Schell J. 1973. Circular DNA plasmids in Agrobacterium strains. Investigation of their role in the induction of crown-gall tumors. *Archives Internationales de Physiologie et de Biochimie* 81: 986

Varshney RK, Paulo MJ, Grando S, van Eeuwijk FA, Keizer LCP, et al. 2012. Genome wide association analyses for drought tolerance related traits in barley (*Hordeum vulgare* L.). *Field Crops Research* 126: 171-80

Vauclare P, Kopriva S, Fell D, Suter M, Sticher L, et al. 2002. Flux control of sulphate assimilation in *Arabidopsis thaliana*: adenosine 5 '-phosphosulphate reductase is more susceptible than ATP sulphurylase to negative control by thiols. *Plant Journal* 31: 729-40

Vernoux T, Wilson RC, Seeley KA, Reichheld JP, Muroy S, et al. 2000. The ROOT MERISTEMLESS1/CADMIUM SENSITIVE2 gene defines a glutathione-dependent pathway involved in initiation and maintenance of cell division during postembryonic root development. *Plant Cell* 12: 97-110

Verslues PE, Lasky JR, Juenger TE, Liu TW, Kumar MN. 2014. Genome-wide association mapping combined with reverse genetics identifies new effectors of low water potential-induced proline accumulation in arabidopsis. *Plant Physiology* 164: 144-59

Veyel D, Erban A, Fehrle I, Kopka J, Schroda M. 2014. Rationales and approaches for studying metabolism in eukaryotic microalgae. *Metabolites* 4: 184-217

Vidal EA, Tamayo KP, Gutierrez RA. 2010. Gene networks for nitrogen sensing, signaling, and response in *Arabidopsis thaliana*. *Wiley Interdisciplinary Reviews. Systems Biology and Medicine* 2: 683-93

Vidmar JJ, Tagmount A, Cathala N, Touraine B, Davidian JE. 2000. Cloning and characterization of a root specific high-affinity sulfate transporter from *Arabidopsis thaliana*. *FEBS Letters* 475: 65-9

Wagner C, Sefkow M, Kopka J. 2003. Construction and application of a mass spectral and retention time index database generated from plant GC/EI-TOF-MS metabolite profiles. *Phytochemistry* 62: 887-900

Wahl B, Reichmann D, Niks D, Krompholz N, Havemeyer A, et al. 2010. Biochemical and spectroscopic characterization of the human mitochondrial amidoxime reducing components hmARC-1 and hmARC-2 suggests the existence of a new molybdenum enzyme family in eukaryotes. *The Journal of Biological Chemistry* 285: 37847-59

Wahl RC, Hageman RV, Rajagopalan KV. 1984. The relationship of Mo, molybdopterin, and the cyanolyzable sulfur in the Mo cofactor. *Archives of Biochemistry and Biophysics* 230: 264-73

Walker RL, Burns IG, Moorby J. 2001. Responses of plant growth rate to nitrogen supply: a comparison of relative addition and N interruption treatments. *Journal of Experimental Botany* 52: 309-17

Walley JW, Kelley DR, Nestorova G, Hirschberg DL, Dehesh K. 2010. arabidopsis deadenylases AtCAF1a and AtCAF1b play overlapping and distinct roles in mediating environmental stress responses. *Plant Physiology* 152: 866-75

Wangeline AL., Burkhead JL, Hale KL, Lindblom SD, Terry N, et al. 2004. Overexpression of ATP sulfurylase in Indian mustard: effects on tolerance and accumulation of twelve metals. *Journal of Environmental Quality* 33: 54-60

Waraich EA, Ahmad R, Halim A, Aziz T. 2012. Alleviation of temperature stress by nutrient management in crop plants: a review. *Journal of Soil Science and Plant Nutrition* 12: 221-44

Watanabe M, Mochida K, Kato T, Tabata S, Yoshimoto N, et al. 2008. Comparative genomics and reverse genetics analysis reveal indispensable functions of the serine acetyltransferase gene family in arabidopsis. *Plant Cell* 20: 2484-96

Weigel D. 2012. Natural variation in arabidopsis: from molecular genetics to ecological genomics. *Plant Physiology* 158: 2-22

Westerman S, Stulen I, Suter M, Brunold C, De Kok LJ. 2001. Atmospheric H_2S as sulphur source for *Brassica oleracea*: consequences for the activity of the enzymes of the assimilatory sulphate reduction pathway. *Plant Physiology and Biochemistry* 39: 425-32

White PJ, Brown PH. 2010. Plant nutrition for sustainable development and global health. *Annals of Botany* 105: 1073-80

Winter D, Vinegar B, Nahal H, Ammar R, Wilson GV, Provart NJ. 2007. An "Electronic Fluorescent Pictograph" browser for exploring and analyzing large-scale biological data sets. *PLoS ONE* 2: e718

Wirtz M, Berkowitz O, Droux M, Hell R. 2001. The cysteine synthase complex from plants. Mitochondrial serine acetyltransferase from *Arabidopsis thaliana* carries a bifunctional domain for catalysis and protein-protein interaction. *European Journal of Biochemistry* 268: 686-93

Wirtz M, Droux M, Hell R. 2004. O-acetylserine (thiol) lyase: an enigmatic enzyme of plant cysteine biosynthesis revisited in *Arabidopsis thaliana*. *Journal of Experimental Botany* 55: 1785-98

Wirtz M, Hell R. 2006. Functional analysis of the cysteine synthase protein complex from plants: structural, biochemical and regulatory properties. *Journal of Plant Physiology* 163: 273-86

Witte CP, Igeno MI, Mendel R, Schwarz G, Fernandez E. 1998. The *Chlamydomonas reinhardtii* MoCo carrier protein is multimeric and stabilizes molybdopterin cofactor in a molybdate charged form. *FEBS Letters* 431: 205-9

Wittle AE, Kamdar KP, Finnerty V. 1999. The Drosophila cinnamon gene is functionally homologous to arabidopsis cnx1 and has a similar expression pattern to the mammalian gephyrin gene. *Molecular Genomics and Genetics* 261: 672-80

Wu X, Zhou Q-h, Xu K. 2009. Are isothiocyanates potential anti-cancer drugs? *Acta Pharmacologica Sinica* 30: 501-12

Wuebbens MM, Rajagopalan KV. 1993. Structural characterization of a molybdopterin precursor. *The Journal of Biological Chemistry* 268: 13493-8

Xiang C, Oliver DJ. 1998. Glutathione metabolic genes coordinately respond to heavy metals and jasmonic acid in arabidopsis. *The Plant Cell* 10: 1539-50

Yaeno T, Iba K. 2008. BAH1/NLA, a RING-type ubiquitin E3 ligase regulates the accumulation of salicylic acid and immune responses to *Pseudomonas syringae* DC3000. *Plant Physiology* 148: 1032-41

Yamaguchi Y, Nakamura T, Harada E, Koizumi N, Sano H. 1999. Differential accumulation of transcripts encoding sulfur assimilation enzymes upon sulfur and/or nitrogen deprivation in *Arabidopsis thaliana*. *Bioscience, Biotechnology, and Biochemistry* 63: 762-6

Yamakawa S, Sakuta C, Matsubayashi Y, Sakagami Y, Kamada H, Satoh S. 1998. The promotive effects of a peptidyl plant growth factor, phytosulfokine-α, on the formation of adventitious roots and expression of a gene for a root-specific cystatin in cucumber hypocotyls. *Journal of Plant Research* 111: 453-58

Yang H, Matsubayashi Y, Hanai H, Sakagami Y. 2000. Phytosulfokine-alpha, a peptide growth factor found in higher plants: its structure, functions, precursor and receptors. *Plant & Cell Physiology* 41: 825-30

Yang H, Matsubayashi Y, Nakamura K, Sakagami Y. 1999. *Oryza sativa* PSK gene encodes a precursor of phytosulfokine-alpha, a sulfated peptide growth factor found in plants. *Proceedings of the National Academy of Sciences USA* 96: 13560-5

Yang H, Matsubayashi Y, Nakamura K, Sakagami Y. 2001. Diversity of arabidopsis genes encoding precursors for phytosulfokine, a peptide growth factor. *Plant Physiology* 127: 842-51

Yang SH, Berberich T, Miyazaki A, Sano H, Kusano T. 2003. Ntdin, a tobacco senescence-associated gene, is involved in molybdenum cofactor biosynthesis. *Plant & Cell Physiology* 44: 1037-44

Yang Z, Guo G, Zhang M, Liu CY, Hu Q, et al. 2013. Stable isotope metabolic labeling-based quantitative phosphoproteomic analysis of arabidopsis mutants reveals ethylene-regulated time-dependent phosphoproteins and putative substrates of constitutive triple response 1 kinase. *Molecular & Cellular Proteomics : MCP* 12: 3559-82

Yang ZY, Danyal K, Seefeldt LC. 2011. Mechanism of Mo-dependent nitrogenase. *Methods in Molecular Biology* 766: 9-29

Yatusevich R, Mugford SG, Matthewman C, Gigolashvili T, Frerigmann H, et al. 2010. Genes of primary sulfate assimilation are part of the glucosinolate biosynthetic network in *Arabidopsis thaliana*. *Plant Journal* 62: 1-11

Yesbergenova Z, Yang G, Oron E, Soffer D, Fluhr R, Sagi M. 2005. The plant Mo-hydroxylases aldehyde oxidase and xanthine dehydrogenase have distinct reactive oxygen species signatures and are induced by drought and abscisic acid. *Plant Journal* 42: 862-76

Yoshimoto N, Inoue E, Saito K, Yamaya T, Takahashi H. 2003. Phloem-localizing sulfate transporter, SULTR1;3, mediates re-distribution of sulfur from source to sink organs in arabidopsis. *Plant Physiology* 131: 1511-7

Yoshimoto N, Inoue E, Watanabe-Takahashi A, Saito K, Takahashi H. 2007. Posttranscriptional regulation of high-affinity sulfate transporters in arabidopsis by sulfur nutrition. *Plant Physiology* 145: 378-88

Yoshimoto N, Takahashi H, Smith FW, Yamaya T, Saito K. 2002. Two distinct high-affinity sulfate transporters with different inducibilities mediate uptake of sulfate in arabidopsis roots. *Plant Journal* 29: 465-73

Yu B, Wakao S, Fan J, Benning C. 2004. Loss of plastidic lysophosphatidic acid acyltransferase causes embryo-lethality in arabidopsis. *Plant & Cell Physiology* 45: 503-10

Yu J, Pressoir G, Briggs WH, Vroh Bi I, Yamasaki M, et al. 2006. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nature Genetics* 38: 203-8

Yuan L, Loqué D, Ye F, Frommer WB, von Wirén N. 2007. Nitrogen-dependent posttranscriptional regulation of the ammonium transporter AtAMT1;1. *Plant Physiology* 143: 732-44

Zarepour M, Kaspari K, Stagge S, Rethmeier R, Mendel RR, Bittner F. 2010. Xanthine dehydrogenase AtXDH1 from *arabidopsis thaliana* is a potent producer of superoxide anions via its NADH oxidase activity. *Plant Molecular Biology* 72: 301-10

Zhang B, Pasini R, Dan H, Joshi N, Zhao Y, et al. 2014. Aberrant gene expression in the arabidopsis SULTR1;2 mutants suggests a possible regulatory role for this sulfate transporter in response to sulfur nutrient status. *Plant Journal* 77: 185-97

Zhang H, Forde BG. 2000. Regulation of arabidopsis root development by nitrate availability. *Journal of Experimental Botany* 51: 51-9

Zhao F, McGrath SP. 1994. Extractable sulphate and organic sulphur in soils and their availability to plants. *Plant and Soil* 164: 243-50

Zhao FJ, Hawkesford MJ, McGrath SP. 1999. Sulphur assimilation and effects on yield and quality of wheat. *Journal of Cereal Science* 30: 1-17

Zhao FJ, Hawkesford MJ, Warrilow AGS, McGrath SP, Clarkson DT. 1996. Responses of two wheat varieties to sulphur addition and diagnosis of sulphur deficiency. *Plant and Soil* 181: 317-27

Zhao K, Aranzana MJ, Kim S, Lister C, Shindo C, et al. 2007. An arabidopsis example of association mapping in structured samples. *PLoS Genetics* 3: e4

Zimmermann P, Hirsch-Hoffmann M, Hennig L, Gruissem W. 2004. GENEVESTIGATOR. arabidopsis microarray database and analysis toolbox. *Plant Physiology* 136: 2621-32

Zuker M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research* 31: 3406-15