

Genome-wide Association Analysis and Phenotypic Study of Nitrogen Use Efficiency in *Arabidopsis thaliana*

Xi Wei

Master Thesis , 2011



Supervisor: Nihal Erol-Öztolan, Mark Aarts

Plant Genetics, Wageningen University

6700HA Wageningen, The Netherlands

Abstract

Nitrogen, as one of the most important macronutrients for plants, is important for crop improvement. Better understanding of the genetic and physiological effects of Nitrogen Use Efficiency (NUE) for plants helps biologists and breeders select genotypes with a better capacity to grow crops under N-deficient conditions. To identify genes associated with NUE, genome wide association (GWA) mapping for HapMap population consisting of 353 natural *Arabidopsis thaliana* accessions was carried out. Extreme accessions with possible higher and lower NUE were selected according to their shoot biomass and the difference between control and deficient shoot biomass. In the confirmation experiment, Cha-0 and UKID22 were two extreme genotypes. Results showed several candidate genes for NUE within significant QTL on chromosome 2, 3, 5. With a haplotype test on three significant SNP markers, related to one of the candidate Sulfate Transporter 4;1, all accessions of the HapMap population were classified into three groups (Hap1, Hap2, Hap3). Hap1, Hap2, and Hap3 contained respectively 151, 170 and 18 accessions. Most of them belonged to the first two haplotypes. qRT-PCR for this gene did not show significant expression difference between Cha-0 (Hap1 with high NUE) and UKID22 (Hap2 with low NUE), but it was upregulated more than 3 folds in Col-0 (Hap 1 with middle NUE) under N-deficient condition. The reason could be that this gene involved in N metabolism but not NUE itself, or alleles function were interrupted, or random errors happened when biological replicates were not enough. Phenotypic analysis, comparing anthocyanin accumulation and root morphology among extreme genotypes, indicated plants tended to accumulate more anthocyanin and increase lateral root growth under N-deficient condition. The results of anthocyanin accumulation did not show direct correlation with NUE. For root analysis, big lateral root difference between deficiency and control plants seemed to be a good indicator for higher N uptake efficiency.

Key words: *Arabidopsis thaliana*, nitrogen, NUE, GWA, QTL, anthocyanin, root morphology.

Table of Contents

Abstract.....	2
1. Introduction	5
1.1 The importance of N for plants.....	5
1.2 Definition of NUE	5
1.3 N transport and homeostasis in plants.....	6
1.3.1 Function of roots in N uptake.....	6
1.3.2 N translocation in aerial parts of plant.....	7
1.3.3 N assimilation and remobilization	7
1.3.4 N metabolism and anthocyanin accumulation.....	9
1.4 Genetic mapping.....	9
1.4.1 QTL mapping	9
1.4.2 GWA mapping.....	10
1.4.3 Model plant for genetic mapping.....	11
1.5 Aim of the thesis	11
2. Material and Methods.....	11
2.1 Plant material and growth condition.....	11
2.2 Data collection and selection of extreme lines	12
2.2.1 Data collection.....	12
2.2.2 Selection of extreme lines.....	13
2.3 GWA mapping.....	13
2.4 Confirmation of candidate gene	13
2.5 Anthocyanin extraction and root architecture analysis on extreme lines.....	13
3. Results.....	15
3.1 Hapmap population screening.....	15
3.1.1 Growing performance of plants	15
3.1.2 Statistical analysis of DW.....	15
3.1.3 Heritability and correlations of N%.....	17
3.1.4 Statistical analysis of NUE.....	18
3.2 GWA mapping.....	19
3.2.1 GWA mapping on FW and DW	19
3.2.2 GWA mapping on N%	23

3.2.3 GWA mapping on NUE.....	24
3.3 Conformation experiments on extreme lines	26
3.3.1 FW and estimated NUE of extreme lines.....	26
3.3.2 Expression of SULTR4;1 in extreme lines.....	28
3.3.3 Anthocyanin comparison among extremes.....	29
3.3.4 Root architecture characterization on extremes.....	30
4. Discussion	34
4.1 Variation of biomass, N% and NUE in <i>Arabidopsis thaliana</i>	34
4.2 NUE determination.....	35
4.3 Expression of SULTR 4;1.....	36
4.4 Other possible candidate genes for NUE.....	36
4.5 NUE, anthocyanin and root morphology of extreme accessions	36
5. Conclusion.....	38
Acknowledgements	38
Reference	39
Appendix.....	45

1. Introduction

1.1 The importance of N for plants

Nitrogen (N) is an abundant element on earth and one of the macronutrients for plants. It is involved in many vital molecules of plants: nucleic acids, proteins and photosynthetic metabolites. Typical symptoms of N deficiency in plants are retarded shoot growth (Drew & Sisworo, 1977; Chapin, 1980), yellowing and chlorosis on leaves (Shaahan *et al.*, 1999), reduced grain yield (Papastylianou & Puckridge, 1983), and increased lateral/primary root ratio (Lemaître *et al.*, 2008) which can cause a serious loss in crop production. Demand for nitrogen fertilizer, therefore, increases in the last few decades (Roy & Hammond, 2004). Side effects of nitrogen fertilizer on environment make people realize that relying on chemical fertilizers might be full of risk (Frink, *et al.*, 1999). In addition, It is not economical for poor and nitrogen-deficient regions to apply high amount of nitrogen fertilizers. As a result, researchers and breeders start paying more attention on nitrogen use efficiency (NUE) in plants, searching for genotypes with a better capacity to grow under N-deficient conditions.

1.2 Definition of NUE

Moll *et al.* (1982) defined NUE as grain yield per unit of available N in soil:

$$\text{NUE} = \text{Gw}/\text{Ns} = (\text{Nt}/\text{Ns})(\text{Gw}/\text{Nt}),$$

where Gw is grain yield, Ns is available N in soil and Nt is the total N in plants when they are mature. This definition is divided into two components: uptake efficiency (Nt/Ns) and utilization efficiency (Gw/Nt).

Berendse and Aerts (1987) defined NUE as produced dry weight per unit of N taken up:

$$\text{NUE} = \text{A}/\text{Ln},$$

where A is N productivity ($\text{g dry weight g}^{-1} \text{N}$) and $1/\text{Ln}$ is mean residence time of N in plants.

Specifically, A is the dry matter production rate per unit of N in plants, where $(1/\text{Ln})$ represents the period when N is used for carbon fixation. Also, they claimed there is an evolutionary trade-off between a high A and a long $1/\text{Ln}$, because more N allocation to compounds or organs with longer life span would lead to a lower photosynthesis, when the N allocates to woody structures.

Good *et al.* (2004) defined NUE as the ratio of dry matter over N contents:

$$\text{NUE} = \text{DW}/\text{N}$$

Chardon *et al.* (2010) defined NUE as the ratio of dry matter over N concentration in *Arabidopsis thaliana*:

$$\text{NUE} = \text{DM}/\text{N}\%,$$

where DM is total shoot dry matter and N% is total N concentration available in plant.

In my thesis, NUE, in shoots of *Arabidopsis thaliana*, is calculated using the equation from Chardon *et al.* (2010):

$$\text{NUE} = \text{DM}/\text{N}\%,$$

and I would also check the differences in GWA mapping results, if using N instead of N% in the equation to recalculate NUE.

1.3 N transport and homeostasis in plants

N transport and homeostasis in plants are two key points to understand NUE. There are various forms of N, which can be classified generally into two categories: organic or inorganic N. Inorganic forms include N_2 , NO , NH_4^+ , NO_3^- , etc. Majority of N in soil is organic (e.g. protein-bound amino acid N), so it is mostly unavailable for plants to absorb unless soil microorganisms convert it into inorganic form (Dechorganat *et al.* 2011). Inorganic form N is available for plants to uptake and assimilate in N metabolism (Du & Peng, 2010). Many plants uptake NO_3^- to use in assimilation, resulting from an increase of the soil pH (Logan & Thomas, 1999). N transport in plants can be separated into several stages: uptake, translocation, assimilation and remobilization.

1.3.1 Function of roots in N uptake

Roots are fundamental organs related to NUE. To meet the N demand of plants, it develops mainly two N-uptake systems: low-affinity transport system under high nitrate supply and high-affinity transport system under low nitrate supply (< 0.1 mmol/L) (Mi & Zhang, 2007). Nitrate transporter 1, namely NRT1 family, is responsible for low-affinity transport system, whereas Nitrate transporter 2 (NRT2) regulates high-affinity system (Orsel *et al.*, 2002).

To take up N efficiently, root morphology is very important under N-deficient condition, because it needs to adapt itself for reaching insufficient and spatially distributed N. Root surface area, determined by root length, branching and diameter, is directly related to N uptake capacity (Noulas *et al.*, 2010). Also, the response of more lateral root growth plays a role in different N uptake efficiencies (Sattelmacher *et al.*, 1990; Mi & Zhang, 2007). In wheat, increasing root growth, at early growth stage, has been proven to be a main factor for N uptake efficiency (Liao *et al.*, 2004). In maize, increasing root

growth, however, has been found to be in relation with NUE at late growth stage. Hybrids with higher NUE had larger root system and less root senescence than inefficient ones in the field (Chun *et al.*, 2005). There are growing interests on improving root system to make plants uptake more N or use N more efficiently. However, there is still concern about root growth that would cause N to compete with C to produce grains after anthesis (Gooding *et al.*, 2005). To improve root system for higher NUE, it is necessary to research more on the root morphology of genotypes with NUE. Labor, time and cleaning of roots without destruction, however, are limiting factors for root analysis (Noulas *et al.*, 2010); also results might vary from controlled to field conditions (Bingham, 2001).

1.3.2 N translocation in aerial parts of plant

Roots absorb N from soil, load it into xylem through vascular tissues. After that, N is assimilated in above-ground parts of plant. In addition to N uptake system, NRT1 and NRT2 family genes take place also in N translocation system (Figure 1, Dechorganat *et al.*, 2011). There are also anion channels mediating nitrate homeostasis in plants. CLC stands for chloride channel well spreading in mammalian body cells. Whereas in *Arabidopsis thaliana* AtCLCa protein acts as a nitrate/H⁺ exchanger with the ability to regulate the nitrate transport to vacuole (De Angeli, 2006). Knockout of the *CLCa* gene in *Arabidopsis thaliana* led to 50% less nitrate accumulation in the whole plant compared to wild type counterparts (Angeli *et al.*, 2009).

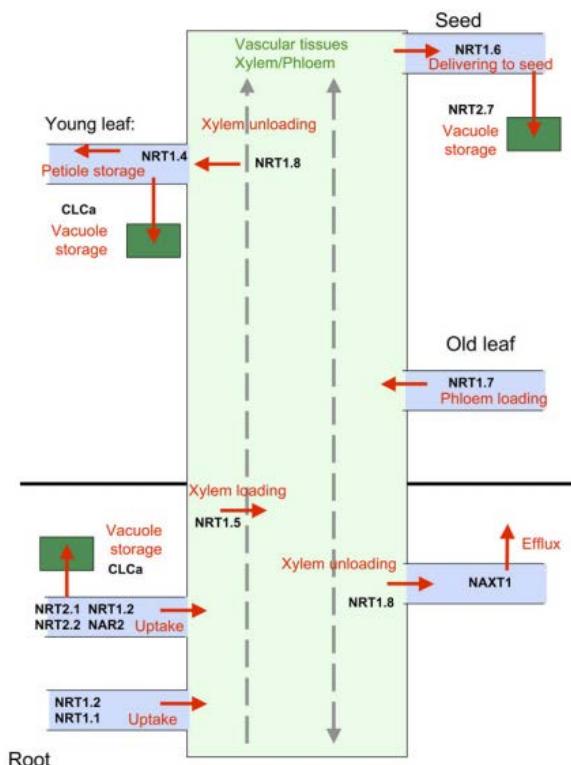


Figure 1. Schematic representation of nitrate routes within *Arabidopsis thaliana* (Dechorganat *et al.*, 2011). NRT1.1/1.2 and NRT2.1/2.2 uptake N from roots, and then NRT1.5 loads N to the xylem. After that, NRT1.8 unloads N from xylem to young leaf, while in old leaves, N is loaded by NRT1.7 to phloem. NRT2.7 and CLCa storage N in vacuoles.

1.3.3 N assimilation and remobilization

For assimilation in plants, nitrate needs to be reduced to ammonia ($\text{NO}_3^- + 8\text{H}^+ + 8\text{e}^- \rightarrow \text{NH}_3 + 2\text{H}_2\text{O} + \text{OH}^-$) (Logan & Thomas, 1999). Nitrate reductase and nitrite reductase (NR/NIA and NIR) coding genes are responsible for the two-steps reduction. (Jaiwal & Singh, 2003; Lemaître *et al.*, 2008). A study showed that a *nr* null mutant of

Arabidopsis thaliana lost the ability to use nitrate when it was the sole N source (Wang *et al.*, 2004). In another study, the expression of NRT1.1 and NRT2.1 increased in a NR-deficient mutant when they were induced by low nitrate (Filleur *et al.*, 1999). Loqué *et al.* (2003) detected upregulation of NRT1.1 and NIA1 genes in NR-deficient mutants with reduce NIA 2 activity when they were supplied with sufficient N, suggesting NR represses activities of both genes and affect the content of available N.

After nitrate has been reduced, ammonia is assimilated into amino acids, such as glutamine, glutamate, asparagine, which can be further used in protein synthesis to build up plant bodies or to catalyze other important biological pathways, like photosynthesis and transpiration (Kumagai *et al.*, 2001). Several enzymes are involved in assimilation of N into amino acids: glutamine synthetase (GS), glutamate synthase (GOGAT), glutamate dehydrogenase (GDH), etc (Lam *et al.*, 1996). They are, therefore, closely related to allocation of N for grain yield in crops. In maize, a *Gs1* locus on chromosome 5 was showed to be a common candidate that explains variations of yield under different N conditions, because in this region there were coincidences of QTL for grain yield, for nitrate content and also for GS, NR activity (Hirel *et al.*, 2001). Overexpression of a NADH-dependent GOGAT gene in rice increased grain weight, indicating this gene might play a key role in N use and yield (Yamaya *et al.*, 2002).

To further use N efficiently, plants recycle ammonia, and re-assimilate it into amino acids at specific stages. When plants are germinated, N moves from seed to growing seedlings to build up plant bodies. At vegetative stage, N is accumulated in vegetative parts to contribute to biomass; and at reproduction stage, it contributes to seed filling (Figure 2). In this way, N is distributed and redistributed within plants in different developmental stages (Bertrand *et al.*, 2007).

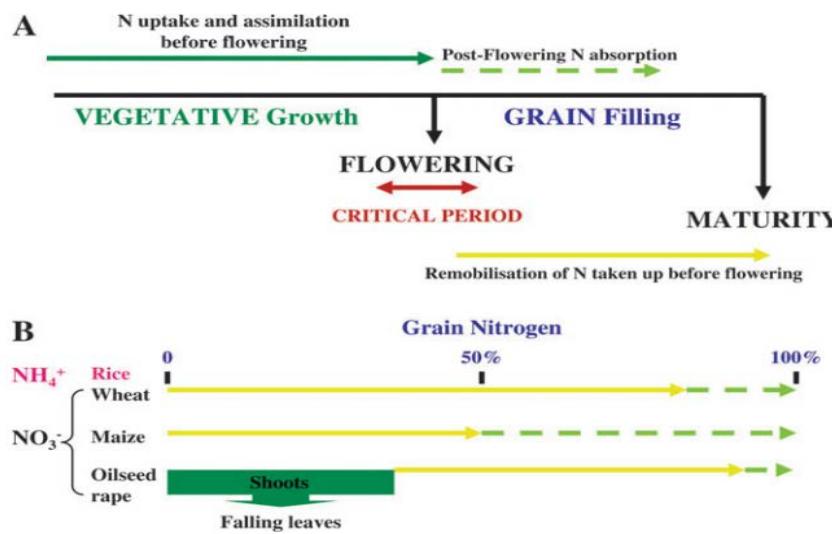


Figure 2. Schematic representation of nitrogen management in crops (Bertrand *et al.*, 2007). (A) Total N uptake, assimilation, and remobilization in plants all through from vegetative growth to grain filling; (B) Relative percentages of N in grain contributed by N remobilization using N, which is taken up before flowering (yellow arrow) or absorption after flowering (Green dash arrow) in rice, wheat, maize and oilseed rape. The grain N in rice, are mostly ammonium from the remobilization N, which is taken up before flowering. In maize, it needs more nitrate absorption (50% of the total N) after flowering. While in oilseed rape, although it also use mostly remobilization N from other plant parts, but much of it effluxes via falling leaves.

1.3.4 N metabolism and anthocyanin accumulation

The use of N in plants is also related to other secondary metabolite synthesis, like anthocyanin. Anthocyanin is usually accumulated in response to stress like light, temperature and nutrition depletion. Symptoms can be found in old leaves, showing a red-purple color. Being one part of the adaptive mechanism to stress, it plays a role in protecting leaves from light damage, and It also serves as a scavenger for reactive oxygen. It appears important for a better tolerance to N deficiency (Peng *et al.*, 2008). Anthocyanin protects N-deficient plants against light stress, therefore, avoids excessive C assimilation and photosynthesis (Paul & Driscoll, 1997). In *Arabidopsis thaliana*, MYB and bHLH transcription factors are induced in response to the stress of nitrogen depletion, regulating the synthesis of anthocyanin (Feyissa *et al.*, 2009). There are also reports about several N/NO₃⁻ related genes, like Lateral Organ Boundary Domain (LBD) family, acting as negative regulators of anthocyanin biosynthesis, when the N/NO₃⁻ amount is sufficient (Rubin *et al.*, 2009). Therefore, it may be possible to use anthocyanin accumulation as an indicator to detect the tolerance of plants with various NUE in response to N depletion.

1.4 Genetic mapping

1.4.1 QTL mapping

QTL is one of the techniques to identify candidate genes by analyzing quantitative traits. QTL mapping, or linkage mapping, is practiced on progeny from crosses between two different genotypes.

Advantages of this method are: 1) it is a well-tested method in scientific researches on many plant traits, like seed dormancy, disease resistance, etc. One of the successful examples is analysis of flowering time. A novel allele of CRY2 (blue light photoreceptor cryptochrome-2), was identified using this method (El-Din El-Assal *et al.*, 2001). It showed the effect of allelic variation on phenotype, caused by a single amino acid substitution. 2) Linkage mapping has bigger statistical power than association mapping to detect rare alleles; therefore, it is less likely to be influenced by subpopulation structures (Famoso *et al.*, 2011).

QTL mapping also has its limitations. One of the limitations of QTL mapping is that the process of crossing is longer than of association mapping, and finally it needs fine mapping to confirm genes of interest (Nordborg & Weigel, 2008). Another limitation is a poor resolution for fine-mapping. In each biparental cross, only two alleles for a specific locus can be studied at the same time (Flint-Garcia *et al.* 2005). Therefore recently, there is a surge of interest in another QTL approach — genome-wide association mapping (GWA mapping).

1.4.2 GWA mapping

GWA mapping or linkage disequilibrium (LD) mapping is a method searching for QTL by detecting associations between traits and genetic markers (Mackay & Powell, 2007). Single nucleotide polymorphisms (SNP) are used as marker to investigate genotypic variation of a trait in a population.

GWA mapping is based on the principle of LD. LD exists when alleles of two loci are inherited more frequently than expectation by random distributions (Pearson & Manolio, 2008), and it results from physical linkage, selection, etc. LD decays with recombination events along time in a random mating population. In an unstructured population, only the correlations between closely linked QTL and markers would remain after enough generations (Mackay & Powell, 2007). The extent of LD decay provides information for the required density of markers for GWA mapping.

Advantages of GWA mapping are: 1) It usually has greater resolution compared to QTL mapping because a rapid decay of LD needs more markers to ensure the success of GWA (Oraguzie & Wilcox, 2007). 2) Since the study applies on a natural population level, it is able to reflect some naturally occurring historical recombination events, providing a possibility of higher resolution (Hall *et al.*, 2010).

Apart from advantages of GWA mapping, there are also disadvantages. One of them is the likelihood to detect false positives, which is the false association of markers and alleles with target trait. One of the main factors inducing false positives is the population structure. Many plant species display complex population structures, because of same historical origin or artificial breeding strategy. When there is population structure, variations of allele frequency exist between different population subgroups. GWA mapping, therefore, identifies false positives that shows a false association between subgroups-related markers and target traits (Pearson & Manolio, 2008; Weigel, & Mott, 2009). Another disadvantage is possibility of false negatives, caused by exclusion of less significant SNP markers, which might be related to the target trait (Neale & Sham, 2004). In addition, GWA mapping has limited power to detect rare alleles (Famoso *et al.*, 2011).

To correct false positives, several statistical models have been developed. One of the promising models is mix-model, a statistical approach relating the pairwise relatedness matrix to the phenotype (Yu, *et al.*, 2006; Myles *et al.*, 2009; Hall *et al.*, 2010). In my thesis, I used efficient mixed mode association expedited (EMMAX), which uses a kinship matrix to estimate the phenotypic variance caused by population structures (Kang *et al.*, 2010).

1.4.3 Model plant for genetic mapping

To identify putative candidate genes involving in NUE with GWA mapping, many characteristics of *Arabidopsis thaliana* make the scientific research on NUE more advantageous. 1) It is an annual plant with short life cycle and it has worldwide distribution and variation; because it is diploid, mainly self-fertilized, homozygous lines are easily created. 2) It is a member of *Brassicaceae* family; therefore, findings on *Arabidopsis thaliana* can provide practical value for some important agricultural plants, like broccoli, cabbage, etc. 3) It contains only 5 chromosomes and fewer than 30,000 genes with dense genetic map available (Meinke & Cherry 1998; Weigel & Mott, 2009). With this map and the knowledge of well-known N-related genes in *Arabidopsis thaliana*, like GS, GOGAT coding genes, NRT and NR family genes, transformations and mutations of them (e.g. NR-deficient or null mutants), are easier to be conducted.

1.5 Aim of the thesis

Aims of this thesis was to find candidate genes and phenotypic indicators for NUE.

Objectives were: 1) to identify candidate genes associated with NUE; 2) To check whether anthocyanin accumulation could be used as a marker for NUE; 3) To find a correlation between root morphology and NUE.

2. Material and Methods

2.1 Plant material and growth condition

The used Hapmap population has 353 accessions of *Arabidopsis thaliana*, obtained from Borevitz' laboratory (University of Chicago, USA). They were collected worldwide. Hapmap population, therefore, includes variants from different parts of world. Details of the mapping population are available at <http://naturalvariation.org/hapmap>. Plants were cultivated under controlled conditions: 20°C/18°C day/night temperature, 10 hours day light, 200 $\mu\text{mol}/\text{m}^2/\text{s}$ light intensity, and 60% humidity. There were 2 treatments: N deficiency and control, with 3 replications per accession. A Standard Nutrition Solution

(NO₃⁻: NH₄⁺; 9:1) with 1 mM N was used to mimic N-deficient condition and 5 mM N for the control condition. The content of this solution can be found in the appendix. Seeds were sown on rock wool blocks, which were covered with black sheets. Rock wool blocks were placed on Ebb/Flood tables, where plants were automatically watered from the bottom. Plants were watered 3 times per week. Harvesting was carried out 30 days after germination.

2.2 Data collection and selection of extreme lines

2.2.1 Data collection

Fresh weights (FW) of shoots were recorded during harvesting, and dry weights (DW) were measured after drying in oven at 55-60 °C for three days. After shoots were dried, three replications of each accession were pooled and grinded. 10-15 mg powder, from homogenously grinded material, was weighed and closed in to a 1 centimeter diameter tin cup, before introducing in the sample tray.

For N measurement, a Flash EA 1112 N Analyzer (Thermo Fisher) was used. N percentage (N%) was determined from the produced gases after complete combustion of the sample. Maintenance should be done and leak test was necessary to carried out before each run. Two sample trays and 64 samples could be analyzed at the same time, and each sample took 265 seconds. Measurements started with a blank measurement (Cellulose, 10-15 mg) followed by five samples of standards (Methionine, 1-15 mg), which gave a calibration line. N % was automatically calculated by the EA 1112 software and showed in the result file. To make sure the accuracy of the measurements, one blank sample was analyzed after every 10 samples, and one standards was measured every 20 samples. For my thesis, Col-0 was additionally used as another standard to monitor the measurements. However in practical, it appeared that 64 samples each run may be too much to give reliable results. To avoid the errors which normally happened in the last 10-15 samples causing by the contamination of Methionine and other samples, sample size each run was decreased to 50. Also, samples were measured again if they were doubtful with contamination. It was also noticeable that except technical errors of machine, the homogeneity of the sample powder will also cause difference in results the because only a small volume (10-15mg) of samples was used to determine N%. After N% had been defined, NUE for each accession was calculated with the formula of DW/N%.

Statistical analysis on FW, DW and NUE was performed using two way ANOVA or t-test. Normality of these data were checked by SPSS, and their heritability was calculated by their genetic variance over the total variance by Microsoft Excel 2010. Genetic variance for FW and DW was calculated with variance of the averages of three replications for all accessions; whereas genetic variance for NUE was calculated

with variance for all accessions, and the environmental variance for NUE was calculated using 20 random accessions.

2.2.2 Selection of extreme lines

Accessions, which maintained high biomass both in N-deficient and control conditions, were defined as extremes with high NUE, whereas accessions, which showed high biomass in control but low in N-deficient condition, were defined as extremes with low NUE. 19 extreme lines showing low and high NUE were chosen. The names and ID numbers of these extremes were showed in the appendix. They were planted in climate chamber under controlled conditions. We applied three N treatments: 0.5mM as negative control, 1mM as N deficiency, and 5mM as standard control. Meanwhile, 8 extreme accessions were grown in silver sand covered with soil in greenhouse for root analysis. Hyponex solution was used to water the plants, and the content of it could be found in the appendix. Standard hyponex solution available in green house under control condition was 6 mM N instead of 5 mM N. Shoot and root of each plant were harvested 35 days after germination.

2.3 GWA mapping

GWA mapping was carried out with 214,000 SNPs as markers. Software R and EMMA were used to correct false positives (Kang *et al.*, 2010). The $-\log_{10}(P \text{ value})$ threshold was arbitrarily chosen at 3.8. Results from GWA were graphed and peaks above the threshold were used to identified candidate genes. Functions of the possible candidate gene were checked in the literature to see whether it was related to N metabolism.

2.4 Confirmation of candidate gene

Expression levels of candidate gene were checked with quantitative RealTime-Reverse Transcriptase - PCR (qRT-PCR). Total RNA of genes of interest was extracted from plant shoots using QIAGEN RNeasy Plant Mini Kit. Around 1 μ g of cDNA was synthesized from RNA using BIO-RAD iScript cDNA Synthesis Kit. The cDNA obtained in last step was five times diluted for qRT-PCR. qRT-PCR was performed using BIO-RAD CFX96 Real-Time PCR system instruments. The results of qRT-PCR for candidate gene were normalized by a housekeeping gene actin. For details of primers please check the Appendix.

2.5 Anthocyanin extraction and root architecture analysis on extreme lines

Anthocyanin was extracted in the 1% (v / v) HCl / methanol solution from around 100 mg fresh shoot powder, which was grinded from the whole shoots in the liquid N. Anthocyanin extracted from the powder was measured with Perkin Elmer Wallac 1420 Multipliable Counter. Relative anthocyanin levels

were checked and compared under the absorbance at 535 nm, normalized by grams of fresh weight of the shoot powder which was used for the extraction. Protocol for anthocyanin extraction was included in the Appendix. For roots analysis, after they had been washed by water, WinRhizo, an image system specially designed for the analysis of root morphology was used to measure the root lengths, surface area and average diameter (Himmelbauer *et al.*, 2004).

3. Results

3.1 Hapmap population screening

3.1.1 Growing performance of plants

Plants in control condition grew healthy, stayed green, and their biomass was greater than for N-deficient plants (Figure 3b). Plants in N deficiency by contrast, showed retarded growth after germination. Leaf senescence and chlorosis were observed on old leaves in N-deficient plants. Purple pigmentation was observed on stems and old leaves because of anthocyanin accumulation (Figure 3a).

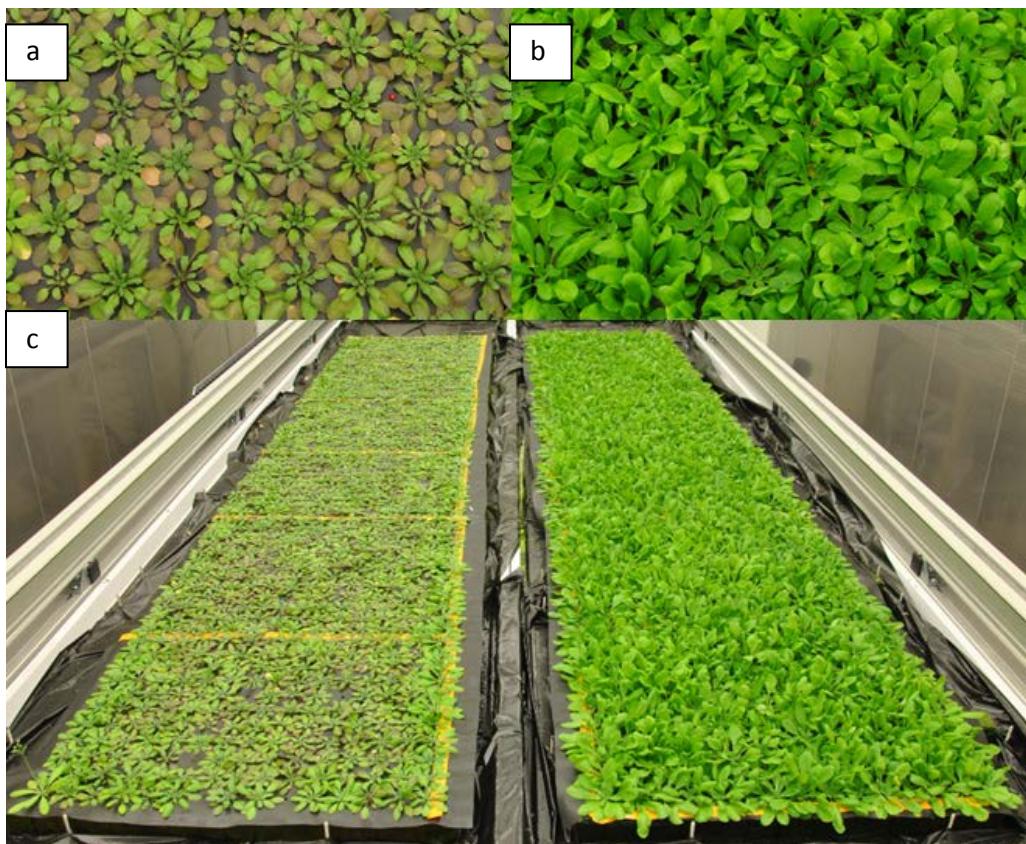


Figure 3. a. *Arabidopsis thaliana* grown under N-deficient condition (1mM N); b. *Arabidopsis thaliana* grown under control condition (5mM N); c. *Arabidopsis thaliana* grown on covered rock wool blocks. N-deficient treatment was applied on the left table, while control on the right.

3.1.2 Statistical analysis of DW

Normality distribution, correlation and ANOVA of DW

Statistical analysis of DW for HapMap population showed a bell shape curve of normal distribution in N-deficient and control conditions (Figure 4a, b). The dry weights and fresh weights under N-deficiency or control are positively correlated (Figure 5a, b). Control and N-deficient dry weights are also positively correlated. (Figure 5c). Average dry weight in control is 0.1065 g, more than 1.5 times of dry weight deficient. The broad sense heritability (H^2) of dry weight under N deficiency was 51.23%, less than that of control (65.22%) (Table 1). Results of ANOVA (Table 2) with F test showed that genotype (G) and N treatment (T) and GxT interactions significantly influenced the dry weight of shoots. Each genotypes responded differently to each treatment.

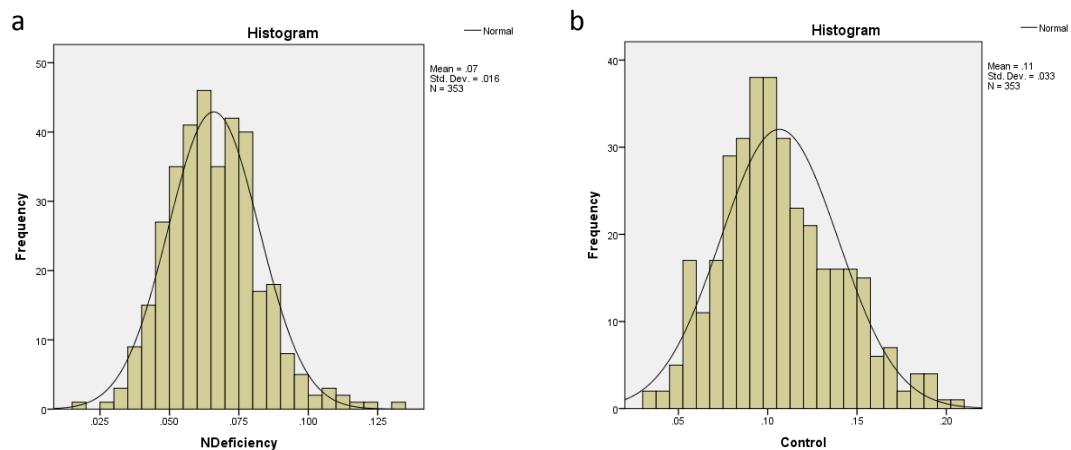


Figure 4. a. Normal distribution of dry weights in N deficiency; b. Normal distribution of dry weights in control.

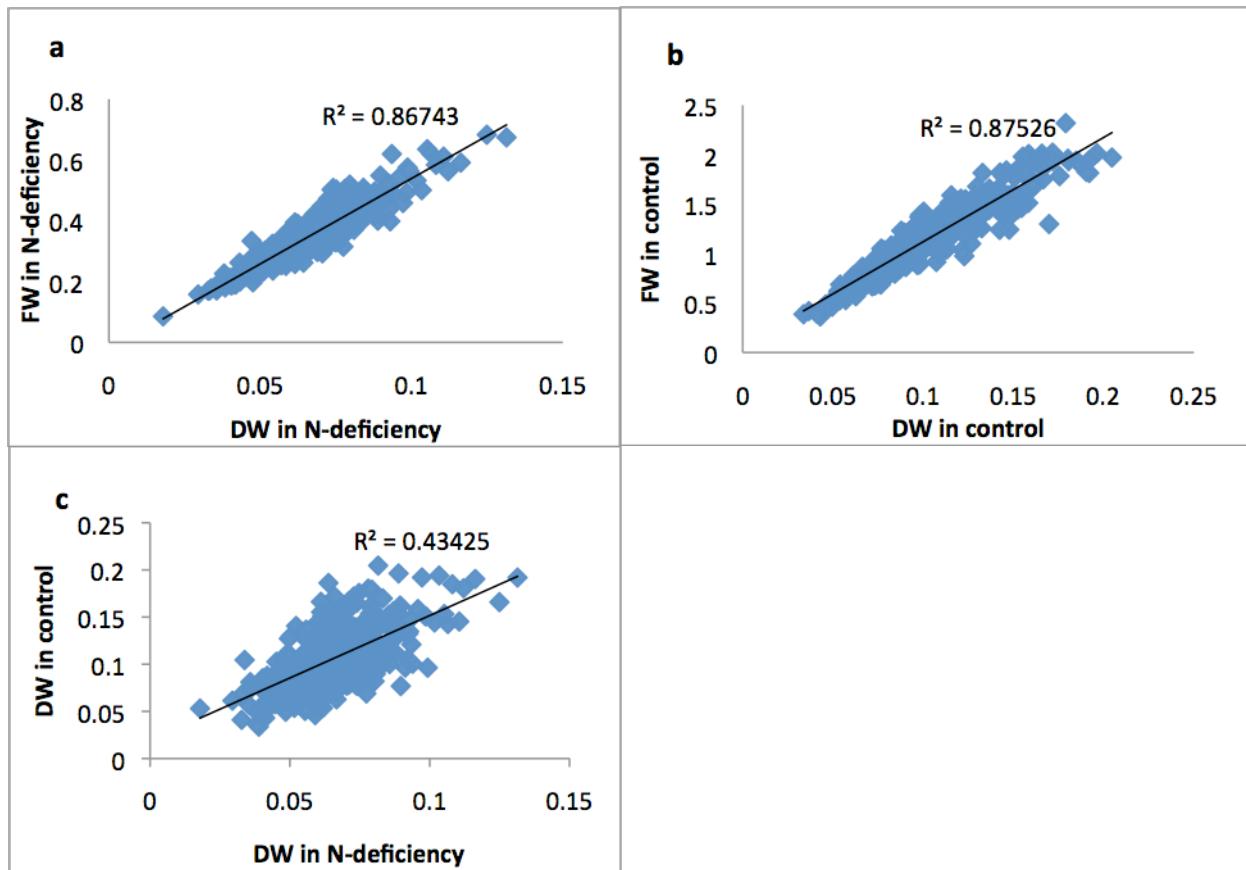


Figure 5. a. Correlation graph between fresh weights and dry weights under N-deficient condition;

b. Correlation graph between fresh weights and dry weights under control condition;

c. Correlation graph between control and N-deficient dry weights.

Table 1. Mean, broad sense heritability (H^2), and standard error of mean DW and NUE in N-deficient and control conditions.

Traits	DW		N%		NUE	
	Conditions	N-deficiency	control	N-deficiency	control	N-deficiency
Average	0.0659	0.1065	1.86	5.95	3.66	1.84
H^2 (%)	51.23	65.22	48.38	54.79	89.73	86.89
Std error	0.008	0.012	0.090	0.120	0.058	0.036

Table 2. Two way ANOVA for effects of genotype, treatment and GXT interaction on dry weight.

Dependent Variable: DW		
Source	F	Sig.
Genotype	6.127	.000
Treatment	1663.239	.000
Genotype * Treatment	2.208	.000

3.1.3 Heritability and correlations of N%

As expected, N% was higher under control than N-deficient condition. In the control, plant shoots contained 5.95% N, 3 times larger than that of plants under N-deficient condition (1.86% N). The heritabilities of N% were 48.38% under N-deficient condition, and 54.79% under control, respectively (Table 1). There was a very small positive correlation between N control and N-deficient condition regarding N% (Figure 6).

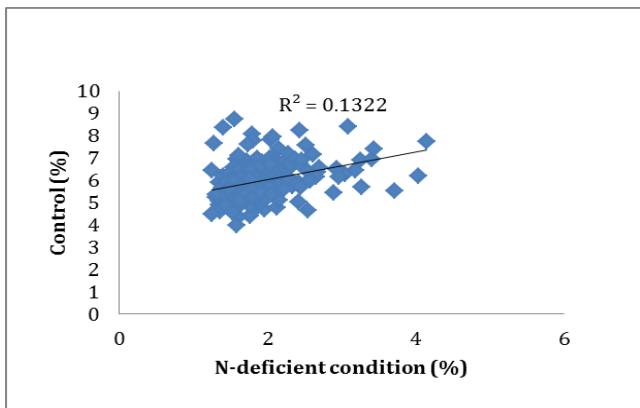


Figure 6. Correlation graph of N% between N control and N-deficient condition.

3.1.4 Statistical analysis of NUE

Normality distribution, correlation and T-test of NUE

Statistical analysis of NUE for HapMap population also showed a bell shape curve of normal distribution in N-deficient and control conditions (Figure 7a, b). NUE control and deficient conditions were positively correlated (Figure 8). Heritability of NUE under N-deficiency (89.73%) was higher than control condition (86.89%, Table 1). T-test of NUE showed accessions performed significantly differently within each treatment (Table 3). NUE of plants doubled when they were suffering from N deficiency. Average NUE under deficient and control condition was 3.66 and 1.84 respectively. The standard error of NUE under N-deficient condition (0.058) was bigger than that of control (0.036) (Table 1).

Table 3. Results of T-test for NUE under N-deficient and control conditions.

Source	t	Sig. (2-tailed)
NUE deficient	62.923	.000
NUE control	51.459	.000

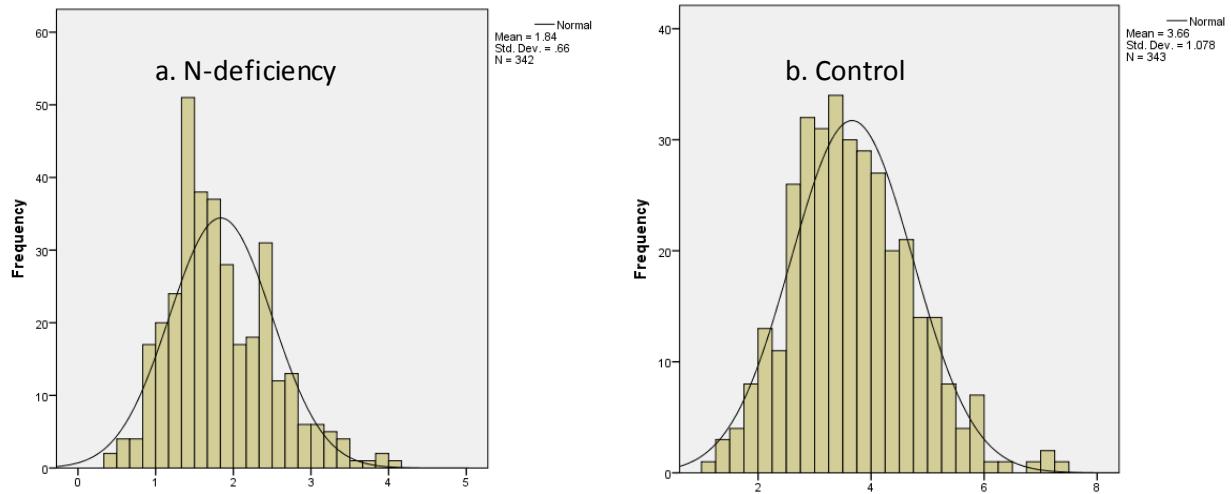


Figure 7. a. Normal distribution of NUE under N-deficient condition;
b. Normal distribution of NUE under control.

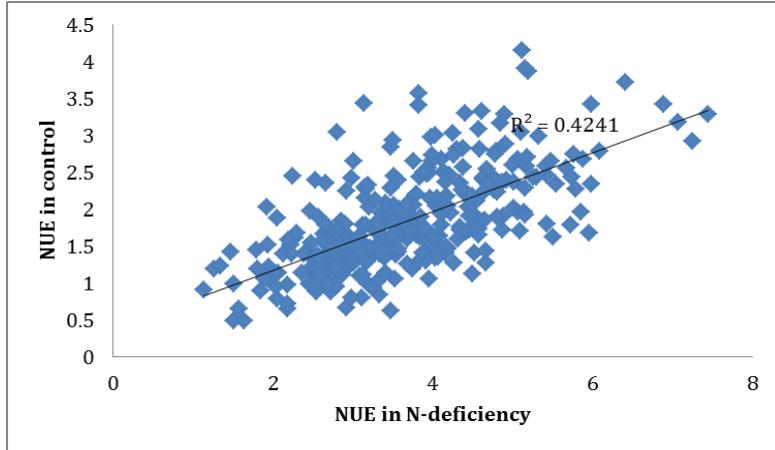


Figure 8. Correlation graph between control and N-deficient NUE.

3.2 GWA mapping

3.2.1 GWA mapping on FW and DW

To increase the reliability of results, I carried out GWA mapping both on fresh weight and dry weight. Results from two mapping, therefore, could be compared and used together to identify candidate genes.

GWA mapping for FW and DW both showed common highly significant peaks on chromosome 5 under N-deficient condition (Figure 9a, c, pink oval), which might correspond to QTL of N accumulation for biomass under N-deficiency. Highly significant SNPs (Table 4) in this region were analyzed to identify

candidate genes. AT5G13550, AT5G13920, and AT5G13760 might be possible candidates. Other genes like AT1G69560 and AT4G03000 on chromosome 1 and 4, could also be candidates.

In control condition, highly significant peaks were observed on chromosome 2 and 4 for FW. For DW, peaks were located on chromosome 1 and 4. There was a common region on chromosome 4 which had common highly significant associated SNPs in FW and GW mapping results (Table 5). Accelerate Cell Death 6 (ACD 6) was found in this region and it was more significant in FW mapping than that of DW (Figure 9b, d, blue oval). It was reasonable to observe it here, since ACD6 had marked impacts on growth and biomass in *Arabidopsis thaliana* (Todesco *et al.*, 2010). More details of this gene could be found in the discussion part (page 35).

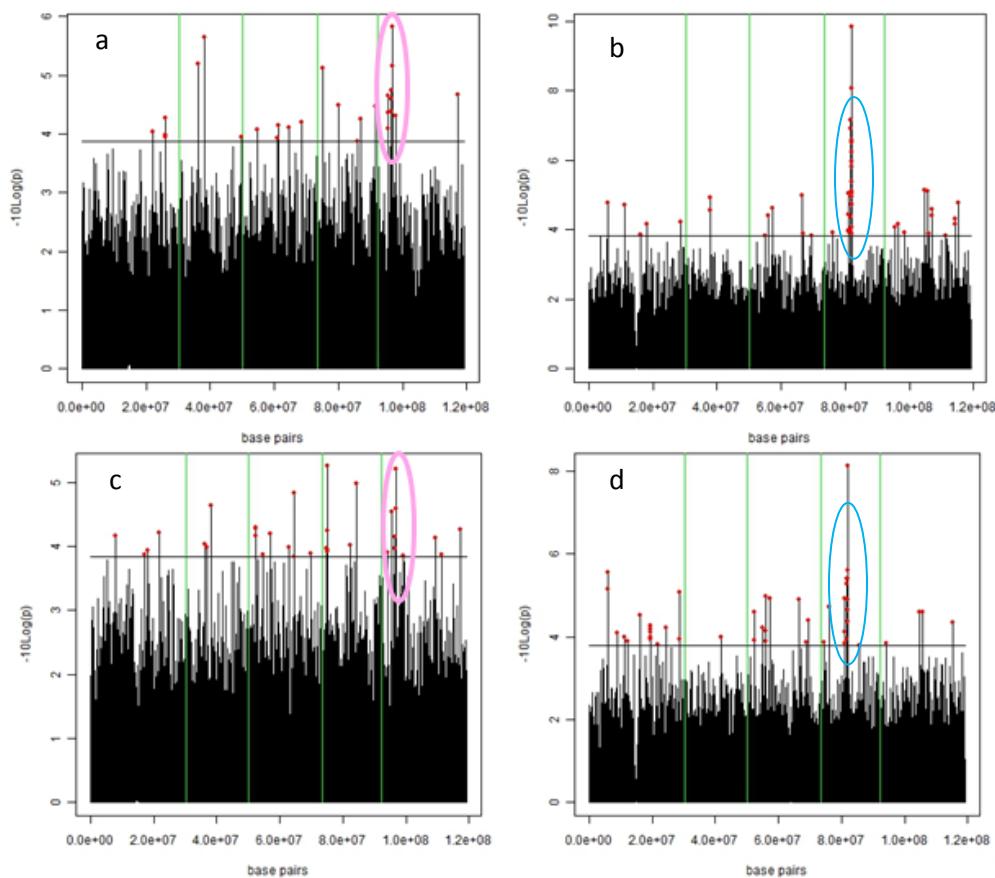


Figure 9. a & b : Association mapping of fresh weight under N-deficient and control condition; c & d : Association mapping of dry weight under N-deficient and control condition.

Five regions separated with green lines in the graph represented 5 chromosome from left to right. Threshold was chosen at 3.8 which was calculated as $-\log_{10}(P \text{ value})$. Peaks above the threshold with red points (SNPs) were used to identified candidate genes. The pink oval indicated the common region on chromosome 5 with highly significant peaks under N-deficient condition, and the blue oval indicated the common region with highly significant peaks on chromosome 4 under control.

Table 4. Significant SNPs from GWA mapping of fresh weight (a) and dry weight (b) under N-deficient condition.

a. Significant SNPs on chromosome 1 and 5 by FW mapping under N-deficient condition (threshold 3.8);

Marker	Chromosome	Position	P value	Gene	Description
m45148	1	26159252	0.000113	AT1G69560	Encodes LOF2 (Lateral Organ
m45152	1	26160608	0.000105	AT1G69560	Fusion1), a MYB-domain transcription factor expressed in organ boundaries;
m168919	5	4355492	1.85E-05	AT5G13550	Encodes a sulfate transporter;
m168920	5	4356592	4.13E-05	AT5G13550	
m168922	5	4357813	2.58E-05	AT5G13550	
m169063	5	4442912	0.000296	AT5G13760	Plasma-membrane choline transporter family protein;
m169130	5	4486785	0.000163	AT5G13920	GRF zinc finger / Zinc knuckle
m169131	5	4487408	0.000273	AT5G13920	protein, functions in zinc ion binding, nucleic acid binding.

b. Significant SNPs on chromosome 4 and 5 by DW mapping under N-deficient condition (threshold 3.8).

Marker	Chromosome	Position	P value	Gene	Description
m127205	4	1325316	0.000107	AT4G03000	RING/U-box superfamily protein, functions in zinc ion binding;
m127209	4	1327238	0.000107	AT4G03000	
m168919	5	4355492	0.000167	AT5G13550	Encodes a sulfate transporter;
m168920	5	4356592	7.00E-05	AT5G13550	
m168922	5	4357813	0.000161	AT5G13550	
m169062	5	4442875	0.000107	AT5G13760	Plasma-membrane choline transporter family protein;
m169063	5	4442912	0.000162	AT5G13760	
m169130	5	4486785	0.000233	AT5G13920	GRF zinc finger / Zinc knuckle protein.

Table 5. Significant SNPs from GWA mapping of fresh weight (a) and dry weight (b) under control condition.

a. Significant SNPs on chromosome 2 and 4 by FW mapping under control condition (threshold 3.8);

Marker	Chromosome	Position	P value	Gene	Description
m63169	2	7404016	2.87E-05	AT2G17036	Protein of unknown function;
m63170	2	7404505	1.21E-05	AT2G17036	
m139459	4	7532599	0.000117	AT4G12830	alpha/beta-Hydrolases
m139462	4	7532789	3.61E-05	AT4G12830	superfamily protein, functions in hydrolase
m139463	4	7533122	0.000107	AT4G12830	activity, catalytic activity;
m141612	4	8274339	1.07E-06	AT4G14368	Regulator of chromosome
m141613	4	8274503	1.48E-10	AT4G14368	condensation (RCC1) family
m141614	4	8274720	1.89E-05	AT4G14368	protein, functions in zinc ion
m141615	4	8274957	0.000131	AT4G14368	binding;
m141632	4	8278786	8.07E-06	AT4G14368	
m141722	4	8296951	2.74E-07	AT4G14400	ACD6, encodes a novel
m141723	4	8297194	9.01E-06	AT4G14400	protein with putative ankyrin
m141726	4	8297531	4.19E-06	AT4G14400	and transmembrane regions,
m141729	4	8297888	1.52E-06	AT4G14400	related to vegetative growth
					and pathogen resistance.

b. Significant SNPs on chromosome 1 and 4 by DW mapping under control condition (threshold 3.8);

Marker	Chromosome	Position	P value	Gene	Description
m31502	1	19311001	6.38E-05	AT1G51960	IQ-domain 26 (IQD26);
m31504	1	19311644	0.000116	AT1G51960	contains IQ calmodulin- binding region;
m31506	1	7532599	0.000117	AT1G51965	Encodes ABA Overly-
m31507	1	19312214	0.000101	AT1G51965	Sensitive5 (ABO5), involved in response to abscisic acid;
m31508	1	19313067	7.49E-05	AT1G51965	
m31512	1	19314517	5.33E-05	AT1G51965	
m141722	4	8296951	2.27E-05	AT4G14400	ACD6
m141729	4	8297888	4.28E-05	AT4G14400	

3.2.2 GWA mapping on N%

Fine mapping in the GWA map for N% (figure 10) showed there were two pairs of associated SNPs (Table 6) under N-deficient condition on chromosome 5, but these two pairs were far away from each other. On other chromosomes (1, 2, 3, 4), significant SNPs were single SNP without associations with each other. For the control condition, no clear peaks could be found (Figure 10).

Table 6. Significant SNPs in QTL on chromosome 5 from GWA mapping of N% under N-deficient condition (threshold 3.8).

Marker	Chromosome	Position	P-value	Gene	Description
m198277	5	18466256	7.69E-05	AT5G45560	Pleckstrin homology (PH) domain-containing protein / lipid-binding
m198287	5	18467607	9.18E-05	AT5G45560	START domain-containing protein
m199955	5	19153559	0.000105	AT5G47150	YDG/SRA domain-containing protein
m199958	5	19154378	0.000105	AT5G47150	

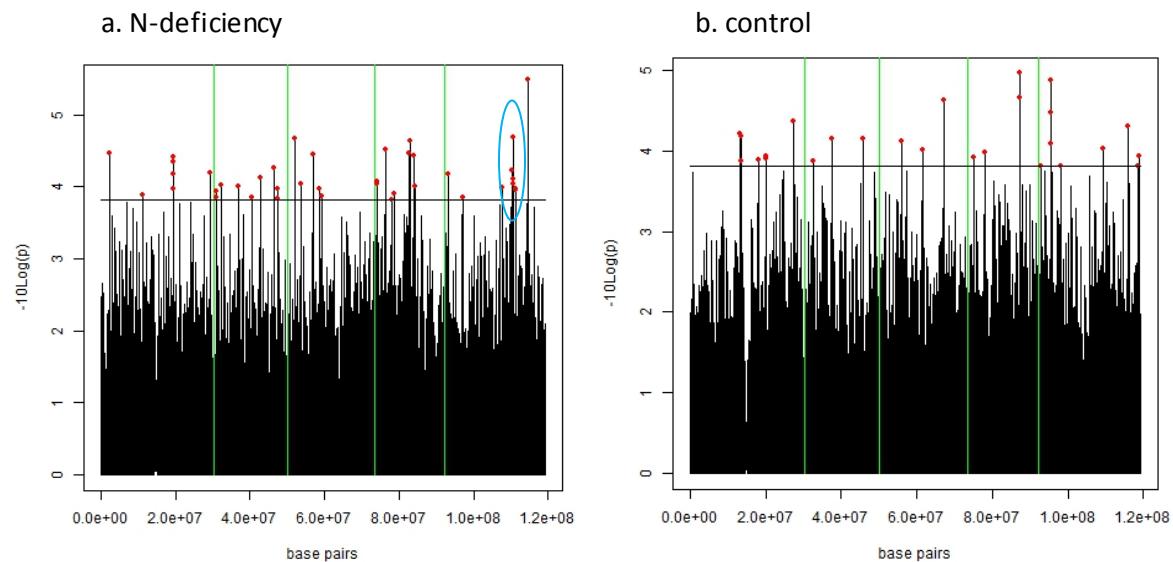


Figure 10. Association mapping results of N% of the HapMap population in N-deficient (a) and control (b) conditions. Five regions in the graph represented 5 chromosomes from left to right. Threshold was chosen at 3.8 which was calculated as $-\log_{10}(P \text{ value})$. Peaks above the threshold with red points(SNPs) were used to identified candidate genes. The blue oval on chromosome 5 contained peaks with associated highly significant SNPs.

3.2.3 GWA mapping on NUE

3.2.3.1 GWA mapping on NUE(DW/N%)

Results of GWA mapping on NUE(DW/N%) showed that chromosome 1, 2, 3 and 5, there were several significant SNPs (Table 7) appearing under N-deficient condition but not in the control. which might be helpful to identify candidate genes. It was noticeable that AT5G13550 (SULTR4;1) again gave one significant SNP under N-deficient condition. Although there was only one SNP here, SULTR4;1 was interesting and it was considered as a candidate, because it showed three times in GWA analysis: GWA mapping for FW, DW and also NUE under N-deficient condition. In addition, two previous researches suggested this gene may associate with N metabolism. In the first research, after inducing with 5 mM nitrate for 2 hours. microarray gene expressions were checked in shoots of *Arabidopsis thaliana* wide type and a NR null mutant (which could not use nitrate as sole N source). SULTR4;1 was one of 65 genes induced in both of wide type and mutant, suggesting it was nitrate responsive but NR-independent (Wang *et al.*, 2004). In another experiment in *Arabidopsis thaliana*, SULTR4;1 was shown to be upregulated under N-deficiency in a *Arabidopsis* nitrogen limitation adaptation (nla) mutant using responsive transcriptome (Peng *et al.*, 2007). Therefore, verification of the function of this gene under different N treatment might provide some information for NUE. AT2G22670 coding for IAA8 could be a candidate, considering it was auxin inducible, therefore it was possibly related to roots and N uptake efficiency. AT3G19610 on chromosome 3 with unknown function could also be a candidate (Figure 11).

Table 7. Significant SNPs in QTL on chromosome 1, 2, 3, 5 from GWA mapping of NUE under N-deficient condition.

Marker	Chromosome	Position	P value	Gene	Description
m13354	1	7827174	9.88E-05	AT1G22170	Phosphoglycerate mutase family protein
m66631	2	9637369	0.000119	AT2G22670	IAA8 (IAA8) gene is auxin inducible
m66632	2	9637537	9.83E-05	AT2G22670	
m92648	3	6814136	3.54E-05	AT3G19610	Unknown protein
m92650	3	6814289	2.45E-05	AT3G19610	
m92652	3	6814496	6.57E-05	AT3G19610	
m92653	3	6814511	4.21E-05	AT3G19610	
m168920	5	4356592	8.12E-05	AT5G13550	Encodes a sulfate transporter
m208184	5	23442402	1.84E-05	AT5G57870	Encodes a putative eukaryotic translation initiation factor

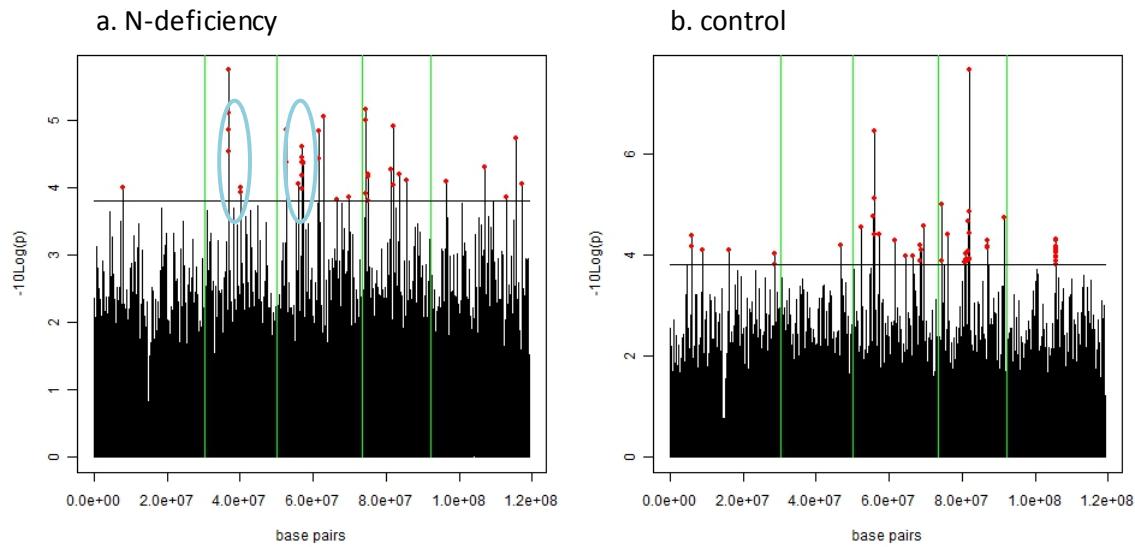


Figure 11. Association mapping results of NUE (DW/N%) of the HapMap population in N-deficient (a) and control (b) conditions. Five regions in the graph represented 5 chromosomes from left to right. Threshold was chosen at 3.8 which was calculated as $-\log_{10}(P \text{ value})$. Peaks above the threshold with red points(SNPs) were used to identified candidate genes. IAA8 was located within the blue oval on chromosome 2, and AT3G19610 with unknown function was located within the blue oval on chromosome 3.

3.2.3.2 GWA mapping on NUE(DW/N)

GWA mapping on NUE(DW/N) gave different results from NUE(DW/N%), with less significant associated SNPs under N-deficient condition (Table 8). Two peaks contained significant SNPs with associations, which might be helpful to identify candidates (Figure 12). One was on chromosome 5, AT5G47150, coding for the YDG/SRA domain-containing protein, and the other one with much less significance was AT2G02061, a Nucleotide-diphospho-sugar transferase family protein.

Table 8. Significant SNPs in QTL on chromosome 2 and 5 from GWA mapping of NUE (DW/N) under N-deficient condition (threshold 3.8).

Marker	Chromosome	Position	P value	Gene	Description
m52810	2	499143	0.000145414	AT2G02061	Nucleotide-diphospho-sugar transferase family protein
m52814	2	499863	0.000155444	AT2G02061	
m199955	5	19153559	0.00108147	AT5G47150	YDG/SRA domain-containing protein
m199958	5	19154378	0.000108147	AT5G47150	

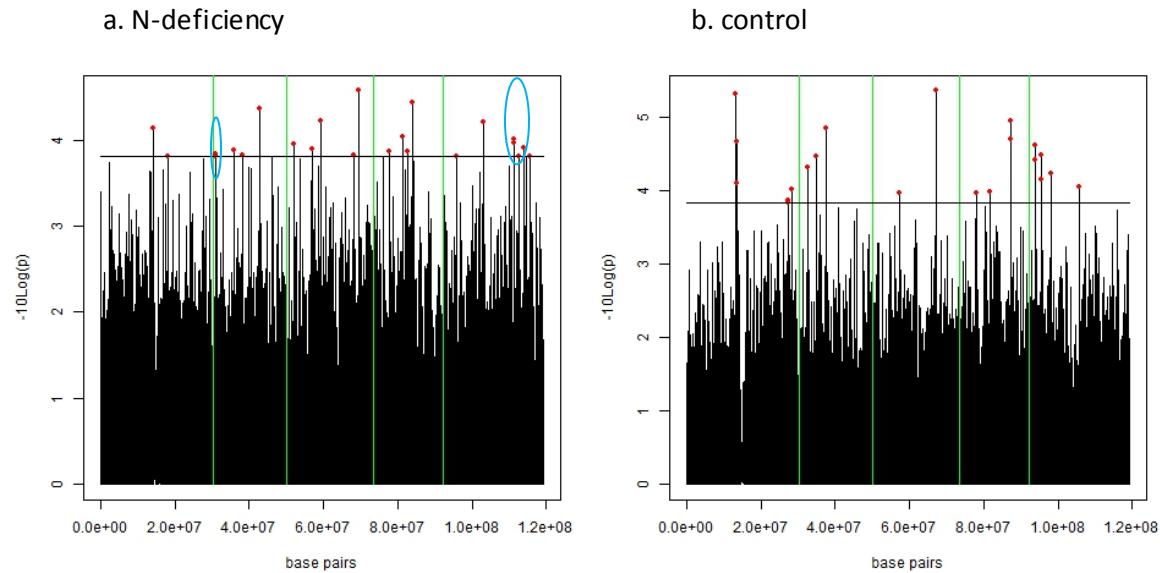


Figure 12. Association mapping results of NUE (DW/N) of the HapMap population in N-deficient (a) and control (b) conditions. Five regions in the graph represented 5 chromosomes from left to right. Threshold was chosen at 3.8 which was calculated as $-\log_{10}(P \text{ value})$. Peaks above the threshold with red points(SNPs) were used to identified candidate genes. AT5G47150 was indicated in blue oval on chromosome 5, while AT2G02061 was within the blue oval on chromosome 2.

3.3 Conformation experiments on extreme lines

3.3.1 FW and estimated NUE of extreme lines

In climate chamber, 19 extremes were grown to confirm their nitrogen use efficiency. They were selected from the Hapmap screening in my thesis and their correlations to the last screening from my supervisor. Accessions, which maintained high biomass both in N-deficient and control conditions, were defined as extremes with high NUE, whereas accessions, which showed high biomass in control but low in N-deficient condition, were defined as low NUE extremes. NUE abilities for extremes were estimated in this way and they were classified to three NUE class (Table 9). Under N-deficient condition (1mM N), Cha-0, Di-1, Tad01 used N efficiently, maintaining relatively high biomass; whereas LL-0, UKID22, Zu-1, Hovdala-2 had high biomass in control but low biomass under N deficiency (Figure13). Cha-0 and UKID22 were two most extreme genotypes. Cha-0 had both high NUE and high biomass under N-deficient condition. UKID22, on the contrary, performed worst. Zu-1 and Hovdala-2 could not use N efficiently under 1mM N deficiency, but their biomass almost did not further decreased under 0.5 mM N. Plants

under negative control of 0.5 mM N, performed slightly differently. Col-0, Tad01, Zu-1, Hovdala-2 and Cha-0 used N more efficiently; whilst UKID22 and Ba1-2 used N less efficiently.

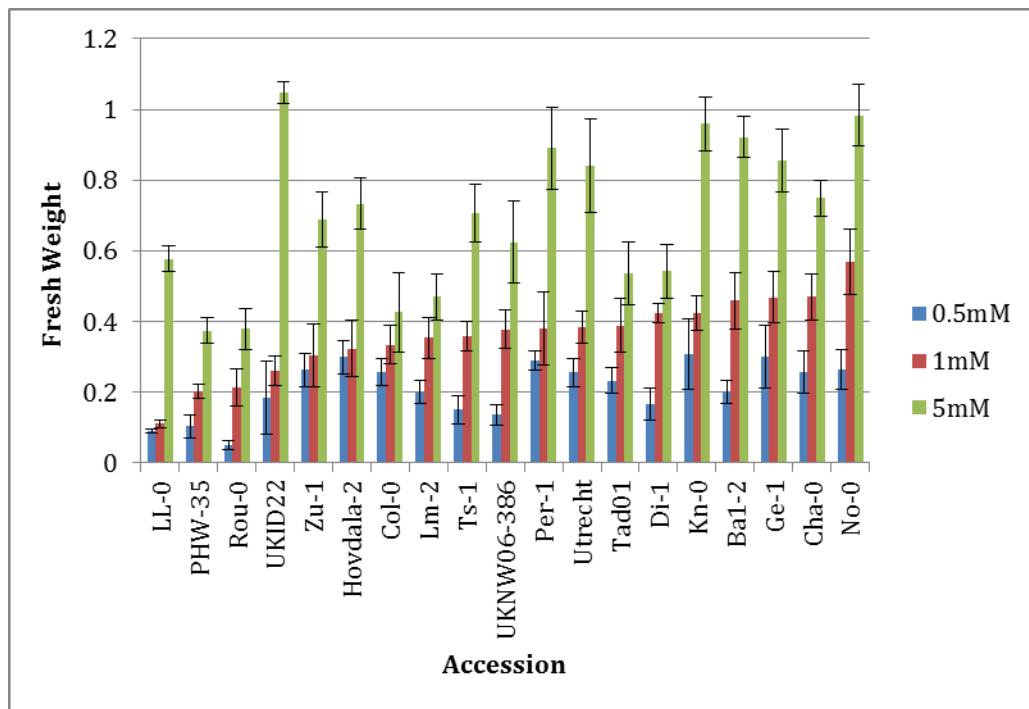


Figure 13. Histogram of fresh weights for 19 accessions in the confirmation experiment under three N treatments (0.5mM, 1mM and 5mM). It was sorted according to FW under 1 mM N.

Table 9. List of accessions classified by the NUE levels.

NUE class	ABRC Stock number	abbreviation
High NUE	CS28133	Cha-0
	CS28208	Di-1
	CS76243	Tad01
	CS28277	Ge-1
Moderate NUE	CS28692	Rou-0
	CS76113	Col-0
	CS76173	Lm-2
	CS76210	Per-1
	CS28795	Utrecht
	CS28395	Kn-0
	CS76093	Ba1-2
	CS28564	No-0
Low NUE	CS76277	UKNW06-386
	CS28635	PHW-35
	CS76271	UKID22
	CS76172	LL-0
	CS28847	Zu-1

CS76143	Hovdala-2
CS76268	Ts-1

3.3.2 Expression of SULTR4;1 in extreme lines

Haplotype test for three significant SNP markers (m168919, m168920 and m168922) related to candidate gene SULTR4;1 was carried out (Table 10). All the accessions were classified to three haplotype groups (Hap1, Hap2 and Hap3). Most accessions belong to Hap1 and 2, while only 18 accessions contained Hap3. Many high NUE extremes in the confirmation experiment belonged to Hap2, while some lower NUE extremes had Hap1. Haplotypes for all extreme accessions were showed in the appendix.

Table 10. Haplotype test for three significant markers related to SULTR4;1.

	m168919	m168920	m168922	accessions
Hap1	G	A	A	low NUE: UKID22, Zu-1, Ts-1, etc.
Hap2	T	G	G	high NUE: Cha-0, Ge-1, Tad01, etc.
Hap3	T	A	G	did not grow in confirmation experiments

To check the whether the difference of Hap1 and Hap2 related to SULTR4;1 would cause difference of NUE in *Arabidopsis thaliana*, three genotypes were used for gene verification. They were Cha-0 (high NUE, Hap2), UKID22 (low NUE, Hap1) and control accession Col-0 (moderate NUE, Hap1). qRT-PCR results (Figure 14) showed that for Cha-0 and Col-0, the expressions of SULTR4;1 were higher under N-deficient condition than under control. The expression difference between N-deficiency and control in Col-0 was the biggest, and there was no clear significant difference between Cha-0 (high NUE, Hap1) and UKID22 (low NUE, Hap2). Therefore, expression of different alleles were not shown to be different.

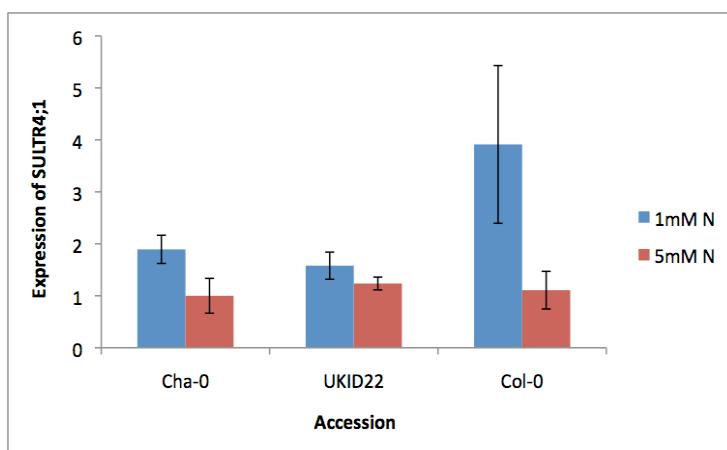


Figure 14. Comparative expressions of SULTR4;1 in three accessions with standard errors: Cha-0 (high NUE, Hap2), UKID22 (low NUE, Hap1) and control accession Col-0 (middle NUE, Hap1).

3.3.3 Anthocyanin comparison among extremes

As in Hapmap screening, anthocyanin pigmentation was also clearly visible on plants under N-deficient condition in the conformation experiments (Figure 15a); control plants were green and healthy (Figure 15b). Extraction of anthocyanin from the whole shoots also confirmed this. Many wells on the N-deficient plates showed purple-pink color, while plate wells for control plants showed light pink or transparent.

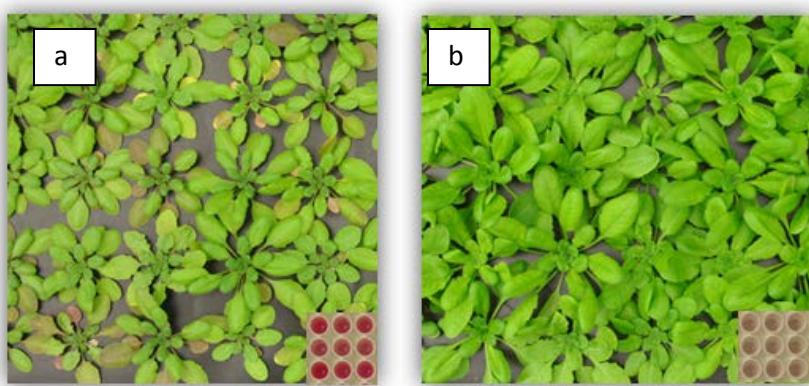


Figure 15. a. Anthocyanin pigmentation on N-deficient *Arabidopsis thaliana* and purple-pink plate containing anthocyanin extracted from their shoots; b. No anthocyanin pigmentation on control plants and low amount of anthocyanin in the wells.

Quantification of anthocyanin proved that there was obvious anthocyanin accumulation for all the genotypes when they faced N-deficiency (Figure 16); whereas, in control, plants accumulated fairly low anthocyanin. However, genotypes with similar NUE performance did not show similar anthocyanin accumulation response. Namely, there were no direct correlations between NUE and anthocyanin accumulation. High NUE extremes (Cha-0, Ge-1 and Di-1) responded differently in anthocyanin accumulation under N-deficient condition: Ge-1 and Di-1 seemed to accumulate much more anthocyanin than Cha-0. Genotypes with relatively low NUE (UKID22, Zu-1 and Col-0) also performed differently in terms of anthocyanin accumulation (Figure 16).

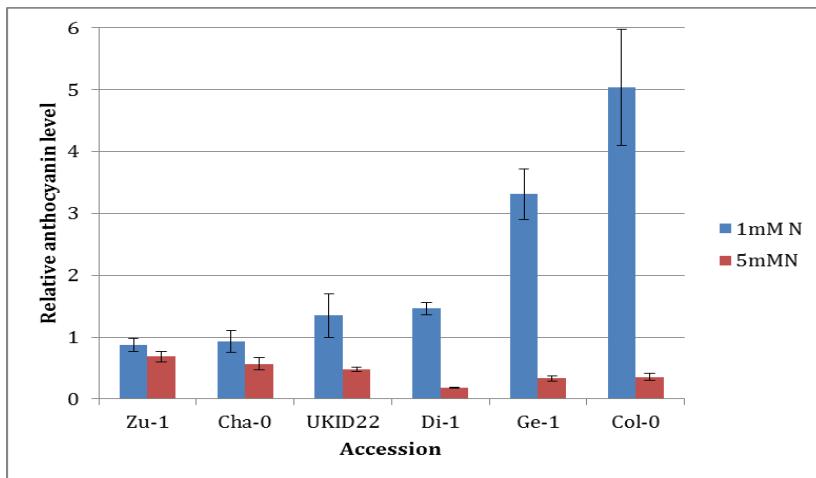


Figure 16. Relative anthocyanin accumulation of 6 accessions under N-deficient (1mM N) and control (5 mM N) conditions.

3.3.4 Root architecture characterization on extremes

There were positive correlations for of FW of shoots between climate chamber and greenhouse under N-deficient and control conditions (Figure 17a, b), but the correlations were not high. In sand experiment, Cha-0, Ge-1, Kn-0 and Di-1 maintained high biomass under deficient condition, suggesting higher NUE. By contrast, No-0, Hovdala-2, Ta-0 and Col-0 could not use N efficiently, with high biomass under control but low biomass under N deficiency (Figure 18).

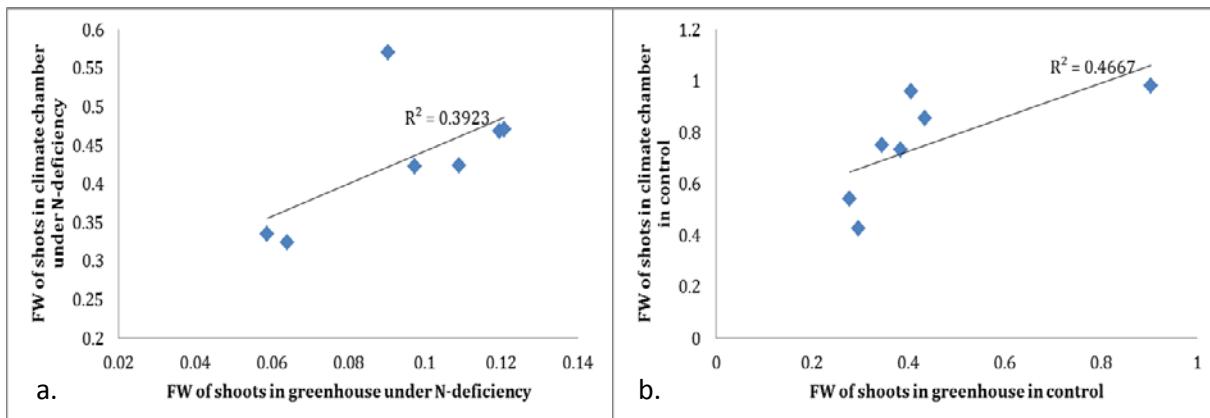


Figure 17. a. Correlation graph of FW of shoots between climate chamber and greenhouse under N-deficient condition; b. Correlation graph of FW of shoots between climate chamber and greenhouse under control condition.

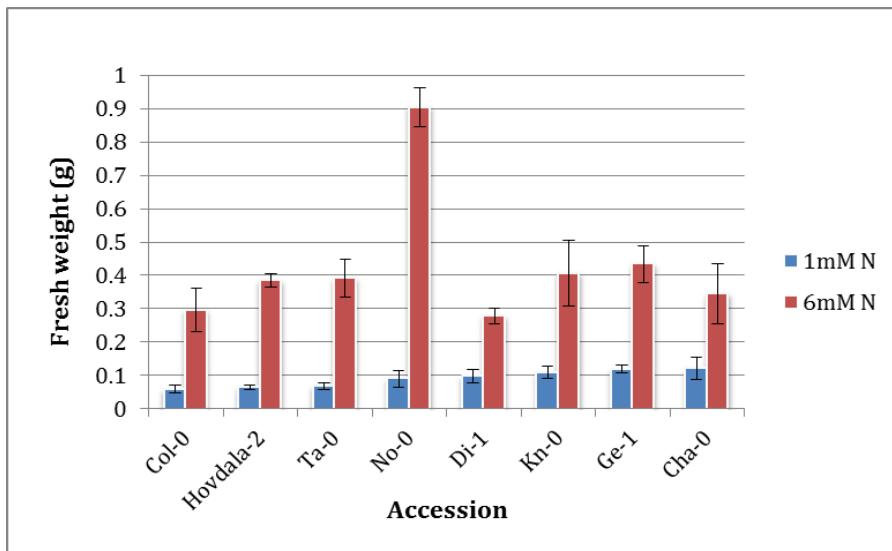


Figure 18. Shoot fresh weights of 8 accessions under N-deficient (1mM N) and control (6mM N) conditions in greenhouse.

8 accessions could be generally divided to two groups according to their NUE. No-0, Hovdala-2, Ta-0 and Col-0 were classified to high NUE group, whereas No-0, Hovdala-2, Ta-0 and Col-0 were put in lower NUE group. Root architecture differences between two NUE groups were showed in Figure 19 and 20, and three roots parameter (root lengths, surface area and average diameter) were measured and shown in Table 11).

High NUE genotypes developed distinct root architecture in N-deficiency compared to control. Cha-0, Ge-1 and Kn-0 shared the same natural root pattern under N-deficient condition. They develop strong root system with more lateral roots, stretching towards horizontal direction in the upper soil with a “umbrella” shape, and then expanded further to the deep (Figure 19). Whereas roots under control condition developed more smoothly towards deep soil, with less lateral roots and less twists and turns. Noticeably, Di-1 was different from other high NUE extremes. It had relative weak roots under N-deficiency.

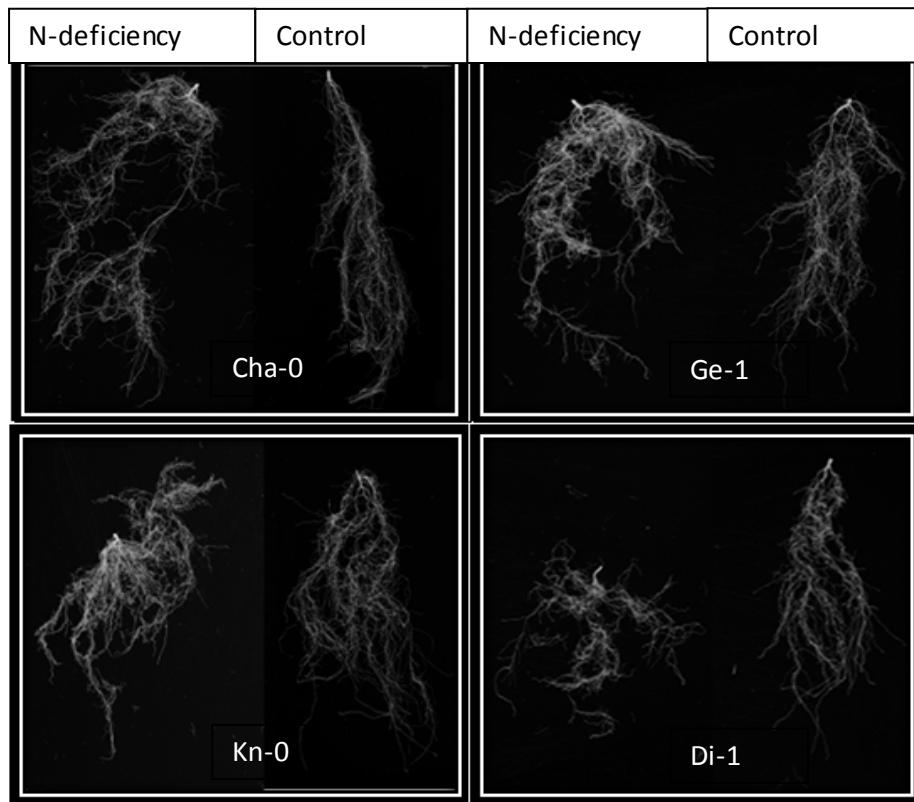


Figure 19. Root morphology of N-efficient accessions: Cha-0, Ge-1, Kn-0 and Di-1. Within each accession: the left root was under N-deficient condition, while the right one was under control condition.

By contrast, lower NUE genotypes contributed more to primary roots (Figure 20). In N-deficient condition, they either had less horizontal branches at the upper soil (Ta-0 and No-0), or branched at the upper soil without further extension (Hovdala-2). Roots of Col-0 with middle NUE is an exception. Although it formed a similar “umbrella” shape as high NUE genotypes under N-deficient condition, but still it had less branches compared to Cha-0, Ge-1 and Kn-0. Their lateral root differences were less than that of higher NUE accessions.

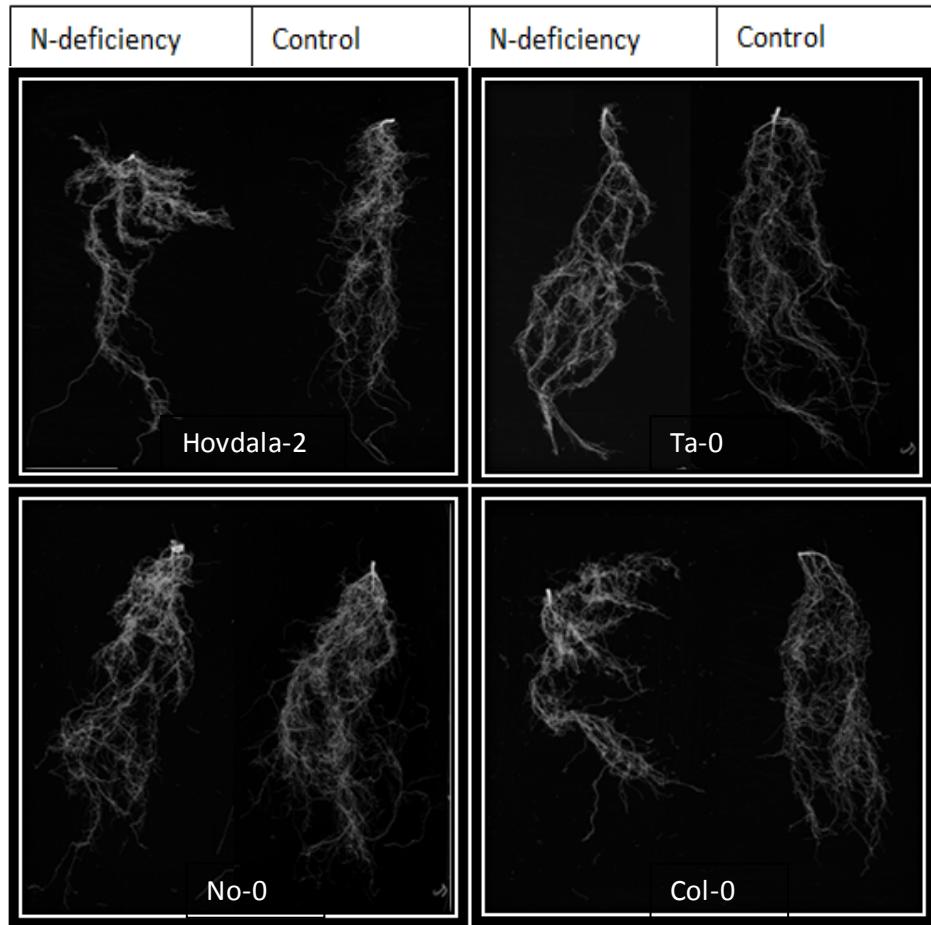


Figure 20. Root morphology of N-inefficient accessions: Hovdala-2, Ta-0, No-0 and Col-0. Within each accession: the left root was under N-deficient condition, while the right one was under control condition.

The root lengths, average diameters, and surface area were different in plants between N-deficient and control condition. 6 accessions had shorter root lengths under N-deficient condition, but not for the rest of two (Ge-1 and Di-1); The average diameters of roots were bigger for all accessions; The change of root surface which did not show a general trend in these accessions. It was aware that three lower NUE accession (Hovdala-2, Ta-0, No-0) had larger difference in root lengths between N-deficient and control condition (Table 12).

However, in N-deficient condition, accessions within the same NUE class (high or low) did not perform similarly. Within high NUE accessions, it was obvious that Cha-0 had longer root lengths, bigger surface area, and bigger average diameter than Di-1 (Table 11), which might suggest that Cha-0 could uptake N more efficiently than Di-0. For low NUE accession, it was also not a case that they had poor root systems with smaller results for three parameters. They sometimes could also have bigger root lengths (like No-0), bigger surface area and diameter (like Ta-0).

Table 11. The lengths, surface area and diameter of roots in 8 accessions: Cha-0, Ge-1, Kn-0, Di-1, Hovdala-2, Ta-0, No-0 and Col-0 under N-deficient and control condition.

Genotype	N-deficiency			Control		
	Length (cm)	SurfArea (cm ²)	Diameter (mm)	Length (cm)	SurfArea (cm ²)	Diameter (mm)
Cha-0	191.8a	23.83abc	0.41cd	272.9	24.56	0.29
Ge-1	134.2ab	19.00bcd	0.45c	131.4	15.43	0.39
Kn-0	135.6ab	24.01abc	0.56b	228.6	22.75	0.32
Di-1	108.3b	12.21de	0.35d	100.5	32.69	0.30
Hovdala-2	121.6b	16.81cd	0.44c	226.0	19.58	0.29
Ta-0	135.1ab	26.58ab	0.63a	272.8	22.43	0.26
No-0	145.3ab	18.37bcd	0.41cd	414.0	41.27	0.32
Col-0	121.9b	15.84cd	0.40cd	165.3	15.32	0.29

Means followed by different letters are different at $P < 0.05$ according to pair-wise LSD (least significant difference) comparisons.

Table 12. The difference of lengths, surface area and diameter of roots between N-deficient condition and control in 8 accessions: Cha-0, Ge-1, Kn-0, Di-1, Hovdala-2, Ta-0, No-0 and Col-0.

Difference between N-deficient and control			
Genotype	Length(cm)	SurfArea(cm ²)	Diameter(mm)
Cha-0	81.1	0.73	0.12
Ge-1	-2.8	-3.57	0.06
Kn-0	93.0	-1.26	0.24
Di-1	-7.8	20.48	0.06
Hovdala-2	104.4	2.77	0.15
Ta-0	137.7	-4.16	0.37
No-0	268.7	22.90	0.09
Col-0	43.4	-0.52	0.10

4. Discussion

4.1 Variation of biomass, N% and NUE in *Arabidopsis thaliana*

There were big genotypic variations in HapMap population of *Arabidopsis thaliana* in biomass and NUE. Observation of GxT interaction indicated that, different genotypes responded differently to each treatment. For biomass, plants had lower heritability in N deficiency (51.23%), indicating biomass in N-deficient condition was influenced more by the environment than in the control (65.22%) condition. NUE in control with lower heritability (86.89%) compared to N-deficient condition (89.73%), was affected more by environments. In climate chambers, light intensity, day length, temperature and

ventilation could be possible factors resulting in bigger environmental variance and lower heritability. Positive correlation, dose and high heritability of NUE under control and N-deficient conditions, indicated that it was a robust trait determined mainly on their genotype, which was consistent with the result from Chardon *et al.* (2010).

4.2 NUE determination

At the beginning of my thesis, NUE in extremes were once defined as DWdef/DWcon. However, this definition was found to be problematic. Firstly, small plants would easily be classified to high NUE genotypes. It was not favorable because for searching high NUE genotypes, plants with high biomass in control, which could also maintain high biomass under N-deficient condition, were our targets. This was an important factor breeders would take into account in selections, because bigger plants tended to give a bigger yield. Secondly, it would cause false positives in GWA mapping. The noise from control condition was big. For example, in mapping results of NUE by the definition of DWdef / DWcon, there was always a peak on chromosome 4, which was the same peak appearing in GWA mapping for DW and FW under control. Accelerate Cell Death 6 (ACD 6) was also found in this region. A previous research showed that the allelic diversity at this single locus was associated with vegetative growth and pathogen resistance in wide type *Arabidopsis thaliana* (Todesco *et al.*, 2010). A hyperactive ACD6 allele was shown with the ability to active defense system against pathogen, but meanwhile slowed the leaf initiation and reduced the biomass of mature leaves. It granted plants with environment-dependent fitness advantages in field if there were pathogens, despite its negative effects on growth and biomass.

The NUE I used in my thesis to run the GWA mapping later, was therefore NUE(DW/N%) (Chardon *et al.*, 2010). There were no correlations between DWdef/DWcon and DW/N%. Defining NUE from a different perspective, GWA mapping on DW/N% showed different peaks. However, the correlations of NUE between control and deficient conditions were still positively correlated ($R^2=0.42$), so even though with the correction by N%, NUE for each genotype is still dependent on the N-environment as in $NUE=DWdef / DWcon$. Recalculating NUE using N contents instead of N%, using $NUE = DW/N$, the mapping results were again totally different. Other peaks were found on chromosome 2 and 5. It is necessary to distinguish between these two calculation methods in the future research. Where in Good *et al.* (2004), they considered 'N usage efficiency', which was the efficiency with which the total plant N was used to produce grain or biomass. Nitrogen utilize efficiency was normally regarded as dry matter production from a given certain amount of N in plants, like what was also mentioned in Moll *et al.* (1982). From this

perspective, DW/N seems to be more reasonable. Nevertheless, Chardon *et al.* (2010) did not specify the reason for using N% clearly, therefore it is meaningful to explore and discuss further.

4.3 Expression of SULTR 4;1

Before the qRT-PCR for SULTR 4;1, haplotype test for three significant markers grouped all the accessions into three haplotype groups. qRT-PCR, therefore, was carried out based on the assumption that different haplotypes caused different NUE in plants. However, the results of qRT-PCR did not show clear significant expression difference between the highest and lowest NUE performance accessions with different haplotypes groups (Hap1 and Hap2). One explanation for this could be that the allele differences were caused by differences in protein sequence, contributing to expression difference. When the function of one of the alleles had been interrupted, there might be no difference for the expression of different alleles. Since there were only 3 accessions involved in qRT-PCR, the possibility of random error also increased. To better verify this gene, more accessions should be included in to test with qRT-PCR, and each accession should have more randomized biological replicates. Besides genetic analysis of SULTR4;1 on Hapmap population derived extremes, it might also be worthy of applying different N treatments on its mutants.

4.4 Other possible candidate genes for NUE

In the results of GWA mapping for NUE (DW/N%), Chromosome 2, 3 contained other possible candidate genes. For example, 2 markers targeted to AT2G22670, indoleacetic acid-induced protein 8 (IAA8), which is a gene related to N and C interaction, but previous research showed it was dependent on Carbon (Gutiérrez *et al.*, 2007). There was still no research about its direct relationship to NUE. 4 markers targeted to AT3G19610 with unknown function, which might be a promising candidate. There were also two SNPs which may possibly have relationship with NUE. Although they were single SNP, there were literatures suggesting their relationship with N metabolism. For example, AT1G22170 encodes phosphoglycerate mutase family protein, which was an associated loci involving in response to nitrate using *Arabidopsis* nitrate reductase null mutant (Wang *et al.*, 2004). AT5G57870, encodes eukaryotic translation initiation factor isoform 4G1. It is related to chlorophyll content in plants and it plays a role in the expression of chloroplast proteins. This was proven in *Arabidopsis* by i4g1/i4g2 double mutant, which had reducing amount of chlorophyll a and b (Lellis *et al.*, 2010). Therefore, it was possible that this gene was involved in photosynthetic nitrogen-use efficiency, when it contributed to the efficiency of C-fixation to leaf nitrogen content.

4.5 NUE, anthocyanin and root morphology of extreme accessions

Selection for high and low NUE extremes is important for QTL mapping. In my thesis, Cha-0 and UKID-22 were two relative stable accessions we selected with highest and lowest NUE in N-deficient condition. It might be helpful to further study NUE using them as parent to make a cross or create doubled haploids.

Anthocyanin extraction on extreme lines confirmed that anthocyanin synthesis would be induced in response to the stress of nitrogen depletion (Feyissa *et al.*, 2009); however, there was no positive correlation with N-deficiency tolerance (Peng *et al.*, 2008) or NUE. It is possibly only a signal of N-deficient stress. In some control plants, we could also see low amount of anthocyanin accumulation. So there might be other stresses possibly existed in the climate chambers. Although functionally anthocyanin protects N-deficient plants against light stress, it is still difficult to say whether it has a real correlation with NUE in *Arabidopsis thaliana*, because it could also be triggered by other stresses, like deficiency of other minerals. Another explanation is that anthocyanin accumulation is tissue specific (Steyn *et al.*, 2002). Anthocyanin was mainly located in stem and old leaves for all genotypes, fairly different from other metabolites spreading the whole-leaves. The NUE difference caused by anthocyanin accumulation between genotypes might be insignificant comparing to internal N regulation systems. For further anthocyanin analysis on its relation with NUE, it may be helpful to exclude accessions which are incapable of making anthocyanin. To test the anthocyanin accumulation, if the sample size is large, like in a mapping population, we could consider using image systems, which is more direct and efficient.

Three N-efficient genotypes in sand experiment generally showed increased lateral roots under N-deficiency. It suggested that developing more lateral roots might be a strategy they used to achieve a higher N uptake efficiency. Actually, the response of more lateral root growth was showed to be closely related to N uptake efficiencies in potato and maize (Sattelmacher *et al.*, 1990; Mi & Zhang, 2007). Di-1 was an exception. It was possible that although this genotype had less N uptake efficiency, it had high N assimilation and remobilization efficiency within shoots. To further understand the N uptake efficiency, it is advisable to measure different root parameters by Winrhizo. But there were still some problems need to be tackled if the plants were grown in sand. When the roots were isolated and washed by water in high pressure, they might be destroyed. Also, primary and lateral roots are sometimes very difficult to separate. These factors could also cause changes in root length, surface area, etc. The reliability of the data was undermined somewhat. Therefore, although measurements on the root parameters by image system could provide much more information for N uptake use efficiency in different accessions, the process itself needs to be improved.

5. Conclusion

In conclusion, GWA mapping of FW, DW and NUE in 354 *Arabidopsis thaliana* accessions identified several candidate genes for NUE. One of the candidates, SULTR 4;1, was studied with haplotype test and qRT-PCR in my thesis, results did not show significant expression difference between highest and lowest NUE extremes with two different haplotypes. It is obvious that more research has to be done for further verification of SULTR4;1 and other candidate genes. In phenotypic study for NUE extremes, plants tended to accumulate more anthocyanin and increase lateral roots under N-deficient condition. Although anthocyanin accumulation did not show direct correlation with NUE, big lateral root difference between deficiency and control might be a good indicator for higher N uptake efficiency.

Acknowledgements

I would like to thank my supervisors Nihal Erol-Öztolan and Mark Aarts for their attention and advice they gave to improve my thinking and my work. I also want to express my appreciation to other group members in Plant Genetics for their help in the lab. Many thanks to all of you who ever gave me support during my thesis.

Reference

Angeli, A., Monachello, D., Ephritikhine, G., Frachisse, J. M., Thomine, S., Gambale, F., & Barbier-Brygoo, H. (2009). CLC-mediated anion transport in plant cells. *Biological Sciences*, 364, 195-201.

Balasubramanian, V., Morales, A. C., Cruz, R. T., & Abdulrachman, S. (1999). On-farm adaptation of knowledge-intensive nitrogen management technologies for rice systems. *Nutrient Cycling in Agroecosystems*, 53, 59-69.

Berendse, F., & Aerts, R. (1987). Nitrogen-use-efficiency: a biologically meaningful definition? *Functional Ecology*, 1, 293-296.

Bingham, I. J. (2001). Soil-root-canopy interactions. *Annals of Applied Biology*, 138, 243-251.

Boussama, N., Ouariti, O., & Ghorbal, M. H. (1999). Changes in growth and nitrogen assimilation in barley seedlings under cadmium stress. *Journal of Plant Nutrition*, 22, 731-752.

Chapin, F. S., III. (1980). The mineral nutrition of wild plants. *Annual Review of Ecology and Systematics*, 11, 233-260.

Chardon, F., Barthélémy, J., Daniel-Vedele, F., & Masclaux-Daubresse, C. (2010). Natural variation of nitrate uptake and nitrogen use efficiency in *Arabidopsis thaliana* cultivated with limiting and ample nitrogen supply. *Journal of Experimental Botany*, 61, 2293-2302.

Chun L., Chen F., Zhang F., & Mi G.H. (2005). Root growth, nitrogen uptake and yield formation of hybrid maize with different N efficiency. *Plant Nutrition and Fertilizer Science*, 11, 615-619.

Coque, M., & Gallais, A. (2006). Genomic regions involved in response to grain yield selection at high and low nitrogen fertilization in maize. *Theoretical And Applied Genetics*. 112, 1205-1220.

Dechorganat, J., Chi Tam, N., Armengaud, P., Jossier, M., Diatloff, E., Filleur, S., & Daniel-Vedele, F. (2011). From the soil to the seeds: the long journey of nitrate in plants. *Journal of Experimental Botany*, 62, 1349-1359.

Drew, M. C., & Sisworo, E. J. (1977). Early Effects of Flooding on Nitrogen Deficiency and Leaf Chlorosis in Barley. *New Phytologist*, 79, 567-571.

Du, X., & Peng, F. (2010). Effect of inorganic nitrogen forms on growth and kinetics of ammonium and nitrate uptake in *Camellia sinensis* L. *Acta Agronomica Sinica*, 36, 327-334.

El-Din El-Assal, S., Alonso-Blanco, C., Peeters, A. J., Raz, V., & Koornneef, M. (2001). A QTL for flowering time in *Arabidopsis* reveals a novel allele of CRY2. *Nature Genetics*, 29(4), 435-440.

Famoso, A. N., Zhao, K. Y., Clark, R. T., Tung, C. W., Wright, M. H., Bustamante, C., & McCouch, S. R. (2011). Genetic architecture of aluminum tolerance in rice (*Oryza sativa*) determined through genome-wide association analysis and QTL mapping. *Plos Genetics*, 7, e1002221.

Feyissa, D. N., Løvdal, T., Olsen, K. M., Slimestad, R., & Lillo, C. (2009). The endogenous GL3, but not EGL3, gene is necessary for anthocyanin accumulation as induced by nitrogen depletion in *Arabidopsis* rosette stage leaves. *Planta*, 230, 747-754.

Filleur, S., & Daniel-Vedele, F. (1999). Expression analysis of a high-affinity nitrate transporter isolated from *Arabidopsis thaliana* by differential display. *Planta*, 207, 461-469.

Flint-Garcia, S. A., Thuillet, A. C., Yu, J. M., Pressoir, G., Romero, S. M., Mitchell, S. E., & Buckler, E. S. (2005). Maize association population: a high-resolution platform for quantitative trait locus dissection. *Plant Journal*, 44, 1054-1064.

Frink, C. R., Waggoner, P. E., & Ausubel, J. H. (1999). Nitrogen fertilizer: retrospect and prospect. *Proceedings of the National Academy of Sciences of the United States of America*, 96, 1175-1180.

Fulker, D. W., Cherny, S. S., Sham, P. C., & Hewitt, J. K. (1999). Combined linkage and association sib-pair analysis for quantitative traits. *American Journal of Human Genetics*, 64, 259-267.

Gallais, A., Coque, M., Quilléré, I., Prioul, J. L., & Hirel, B. (2006). Modelling postsilking nitrogen fluxes in maize (*Zea mays*) using ¹⁵N-labelling field experiments. *New Phytologist*, 172, 696-707.

Gallais, A. & Hirel, B. (2004). An approach to the genetics of nitrogen use efficiency in maize. *Journal of Experimental Botany*, 55, 295-306.

Good, A. G., A. K. Shrawat, et al. (2004). "Can less yield more? Is reducing nutrient input into the environment compatible with maintaining crop production?" *Trends in Plant Science*, 9, 597-605.

Gooding, M. J., Pepler, S., Ford, K., Gregory, E., & Cultivar, P. J. (2005). Fungicide and foliar urea effects on wheat grain yields of dry matter, nitrogen and sulphur: Associations with root distributions after ear emergence. *Aspects of Applied Biology*, 73: 179-186.

Gutiérrez, R. A., Lejay, L. V., Dean, A., Chiaromonte, F., Shasha, D. E., & Coruzzi, G. M. (2007). Qualitative network models and genome-wide expression data define carbon/nitrogen-responsive molecular machines in *Arabidopsis*. *Genome Biology*, 8, R7-R7.

Hall, D., Tegström, C., & Ingvarsson, P. K. (2010). Using association mapping to dissect the genetic basis of complex traits in plants. *Briefings in Functional Genomics & Proteomics*, 9, 157-165.

Himmelbauer, M. L., Loiskandl, W., & Kastanek, F. (2004). Estimating length, average diameter and surface area of roots using two different image analyses systems. *Plant and Soil*, 260, 111-120.

Hirel, B., Bertin, P., Quilleré, I., Bourdoncle, W., Attagnant, C., Dellay, C., & Gallais, A. (2001). Towards a better understanding of the genetic and physiological basis for nitrogen use efficiency in maize. *Plant Physiology*, 125, 1258-1270.

Hirel, B., Gouis, J. I., Ney, B., & Gallais, A. (2007). The challenge of improving nitrogen use efficiency in crop plants: towards a more central role for genetic variability and quantitative genetics within integrated approaches. *Journal of Experimental Botany*, 58, 2369-2387.

Jaiwal, P. K., & Singh, R. P. (2003). Improvement strategies of Leguminosae biotechnology. In P. K. Jaiwal & R. P. Singh (Eds.), *Improvement strategies of Leguminosae biotechnology*. Dordrecht; Netherlands: Kluwer Academic Publishers.

Kumagai, E., Araki, T., Hamaoka, N., & Ueno, O. (2011). Ammonia emission from rice leaves in relation to photorespiration and genotypic differences in glutamine synthetase activity. *Annals of Botany*, 108, 1381-1386.

Lam, H. M., Coschigano, K. T., Oliveira, I. C., Melo-Oliveira, R., & Coruzzi, G. M. (1996). The molecular genetics of nitrogen assimilation into amino acids in higher plants. *Annual Review of Plant Physiology and Plant Molecular Biology*, 47, 569-593.

Lellis, A. D., Allen, M. L., Aertker, A. W., Tran, J. K., Hillis, D. M., Harbin, C. R., & Browning, K. S. (2010). Deletion of the eIFiso4G subunit of the Arabidopsis eIFiso4F translation initiation complex impairs health and viability. *Plant Molecular Biology*, 74, 249-263.

Lemaître, T., Gaufichon, L., Boutet-Mercey, S., Christ, A., & Masclaux-Daubresse, C. (2008). Enzymatic and metabolic diagnostic of nitrogen deficiency in *Arabidopsis thaliana* Wassileskija accession. *Plant and Cell Physiology*, 49, 1056-1065.

Logan, K. A. B., & Thomas, R. J. (1999). Interactions between N supply and N uptake by perennial ryegrass, ¹⁵N recovery and soil pH for four acid Scottish soils. *Grass and Forage Science*, 54, 42-51.

Loqué, D., Tillard, P., Gojon, A., & Lepetit, M. (2003). Gene expression of the NO₃⁻ transporter NRT1.1 and the nitrate reductase NIA1 is repressed in *Arabidopsis* roots by NO₂⁻, the product of NO₃⁻ reduction. *Plant Physiology*, 132, 958-967.

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, 345-358.

Mackay, I., & Powell, W. (2007). Methods for linkage disequilibrium mapping in crops. *Trends In Plant Science*, 12, 57-63.

Myles, S., Peiffer, J., Brown, P. J., Ersoz, E. S., Zhang, Z., Costich, D. E., & Buckler, E. S. (2009). Association mapping: critical considerations shift from genotyping to experimental design. *The Plant Cell*, 21, 2194-2202.

Neale, B. M., & Sham, P. C. (2004). The Future of Association Studies: Gene-Based Analysis and Replication. *The American Journal of Human Genetics*, 75, 353-362.

Meinke, D. W., & Cherry, J. M. (1998). *Arabidopsis thaliana*: A Model Plant for Genome Analysis. *Science*, 282, 662.

Mi, G., Chen, F., & Zhang, F. (2007). Physiological and genetic mechanisms for nitrogen-use efficiency in maize. *Journal of Crop Science and Biotechnology*, 10, 57-63.

Moll, R. H., Kamprath, E. J., & Jackson, W. A. (1982). Analysis and interpretation of factors which contribute to efficiency of nitrogen utilization. *Agronomy Journal*, 74, 562-564.

Nordborg, M., & Weigel, D. (2008). Next-generation genetics in plants. *Nature*, 456, 720-723.

Noulas, C., Liedgens, M., Stamp, P., Alexiou, I., & Herrera, J. M. (2010). Subsoil root growth of field grown spring wheat genotypes (*Triticum aestivum* L.) differing in nitrogen use efficiency parameters. *Journal of Plant Nutrition*, 33, 1887-1903.

Novitskaya, L., Trevanion, S. J., Driscoll, S., Foyer, C. H., & Noctor, G. (2002). How does photorespiration modulate leaf amino acid contents? A dual approach through modelling and metabolite analysis. *Plant, Cell and Environment*, 25, 821-835.

Obara, M., Kajiura, M., Fukuta, Y., Yano, M., Hayashi, M., Yamaya, T., & Sato, T. (2001). Mapping of QTLs associated with cytosolic glutamine synthetase and NADH-glutamate synthase in rice (*Oryza sativa* L.). *Journal of Experimental Botany*, 52, 1209-1217.

Oraguzie, N. C., & Wilcox, P. L. (2007). An overview of association mapping. *Association Mapping in Plants*, 1-9.

Orsel, M., Filleur, S., Fraisier, V., & Daniel-Vedele, F. (2002). Nitrate transport in plants: which gene and which control? *Journal of Experimental Botany*, 53, 825-833.

Papastylianou, I., & Puckridge, D. W. (1983). Stem nitrate nitrogen and yield of wheat in a permanent rotation experiment. *Australian Journal of Agricultural Research*, 34, 599-606.

Paul, M. J., & Driscoll, S. P. (1997). Sugar repression of photosynthesis: the role of carbohydrates in signalling nitrogen deficiency through source:sink imbalance. *Plant, Cell and Environment*, 20, 110-116.

Pearson, T. A., & Manolio, T. A. (2008). How to interpret a genome-wide association study. *The Journal Of The American Medical Association*, 299, 1335-1344.

Peng, M., Bi, Y.-M., Zhu, T., & Rothstein, S. J. (2007). Genome-wide analysis of Arabidopsis responsive transcriptome to nitrogen limitation and its regulation by the ubiquitin ligase gene NLA. *Plant Molecular Biology*, 65, 775-797.

Peng M.S., Hudson D., Schofield A., Tsao R., Yang R., Gu H.L., Bi Y.M., & Rothstein S.J. (2008) Adaptation of Arabidopsis to nitrogen limitation involves induction of anthocyanin synthesis which is controlled by the NLA gene. *Journal of Experimental Botany*, 59, 2933-2944

Roy, A. H., & Hammond, L. L. (2004). Challenges and opportunities for the fertilizer industry. In A. R. Mosier, K. J. Syers & J. R. Freney (Eds.), *Agriculture and the nitrogen cycle: assessing the impacts of fertilizer use on food production and the environment*. Washington; USA: Island Press.

Rubin, G., Tohge, T., Matsuda, F., Saito, K., & Scheible, W. R. (2009). Members of the LBD family of transcription factors repress anthocyanin synthesis and affect additional nitrogen responses in Arabidopsis. *The Plant Cell*, 21, 3567-3584.

Sage, R. F., & Pearcy, R. W. (1987). The nitrogen use efficiency of C3 and C4 plants. I. Leaf nitrogen, growth, and biomass partitioning in *Chenopodium album* (L.) and *Amaranthus retroflexus* (L.). *Plant Physiology*, 84, 954-958.

Sattelmacher, B., Klotz, F., & Marschner, H. (1990). Influence of the nitrogen level on root growth and morphology of two potato varieties differing in nitrogen acquisition. *Plant and Soil*, 123, 131-137.

Shaahan, M. M., El-Sayed, A. A., & Abou El-Nour, E. A. A. (1999). Predicting nitrogen, magnesium and iron nutritional status in some perennial crops using a portable chlorophyll meter. *Scientia Horticulturae*, 82, 339-348.

Steyn, W. J., Wand, S. J. E., Holcroft, D. M., & Jacobs, G. (2002). Anthocyanins in vegetative tissues: a proposed unified function in photoprotection. *New Phytologist*, 155, 349-361.

Todesco, M., Balasubramanian, S., Hu, T. T., Traw, M. B., Horton, M., Epple, P., & Weigel, D. (2010). Natural allelic variation underlying a major fitness trade-off in *Arabidopsis thaliana*. *Nature*, 465, 632-636.

Wang, R., Tischner, R., Gutiérrez, R. A., Hoffman, M., Xing, X., Chen, M., & Crawford, N. M. (2004). Genomic analysis of the nitrate response using a nitrate reductase-null mutant of *Arabidopsis*. *Plant Physiology*, 136, 2512-2522.

Weigel, D., & Mott, R. (2009). The 1001 genomes project for *Arabidopsis thaliana*. *Genome Biology*, 10, 107-107.

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

Yu, J., Pressoir, G., Briggs, W. H., Vroh Bi, I., Yamasaki, M., Doebley, J. F., & Buckler, E. S. (2006). A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nature Genetics*, 38, 203-208.

Appendix

Anthocyanin extraction protocol:

Measure fresh weight of plant material;

The plant material is placed on mortar (preferably cooled to -20C) and frozen with liquid nitrogen;

Grind the whole shoots into even powder;

For each sample, measure approximately 100mg into 2ml tubes;

Extraction of pigments in 750ul of solution 1% (v / v) HCl / methanol;

Add 500ul of distilled water and Vortex;

Add 600ul of chloroform (to extract the chlorophylls) and vortex;

Separate the phases by centrifugation at 12000 rpm for 30 seconds;

Transfer supernatant to new tubes;

Anthocyanin quantification:

Quantificate (500ul of solution) in a 96 wells plate reader (auto-zero solution composed of water and HCl / Methanol, absorbance at 535 nm);

The level of anthocyanin per sample is expressed as the ratio between the absorbance value at 535 nm divided by grams of fresh weight.

Arabidopsis thaliana sulfate transporter 4.1 (SULTR4;1) mRNA, complete cds NCBI Reference Sequence: NM_121358.2

http://www.ncbi.nlm.nih.gov/nuccore/NM_121358.2

<http://www.ncbi.nlm.nih.gov/nuccore/30684459?report=fasta>

Primer/Hairpins, Self-dimers, Cross dimers

FW: CCTGACACCTCCAATGAAGC 0,0,0

RV: CCATCGGAAACAAGGAAAGA 0,0,0

Primer details:

Pair 109:					
<input type="checkbox"/> Left Primer 109:	Primer_108_F				
Sequence:	CCTGACACCTCCAATGAAGC				
Start: 189	Length: 20 bp	Tm: 60.7 °C	GC: 55.0 %	ANY: 5.0	SELF: 3.0
<input type="checkbox"/> Right Primer 109:	Primer_108_R				
Sequence:	CCATCGGAAACAAGGAAAGA				
Start: 317	Length: 20 bp	Tm: 60.0 °C	GC: 45.0 %	ANY: 3.0	SELF: 0.0
Product Size: 129 bp	Pair Any: 5.0	Pair End: 0.0	Product TM: 76.2 °C	Product - Tm: 16.2 °C	

Number and abbreviation of extreme lines used in climate chamber and greenhouse:

ABRC Stock number	abbreviation	Haplotype
CS76172	LL-0	Hap1
CS28635	PHW-35	Hap1
CS28692	Rou-0	Hap1
CS76271	UKID22	Hap1
CS28847	Zu-1	Hap1
CS28208	Di-1	Hap1
CS76113	Col-0	Hap1
CS76173	Lm-2	Hap1
CS76268	Ts-1	Hap1
CS76277	UKNW06-386	Hap1
CS76210	Per-1	Hap2
CS28795	Utrecht	Hap2
CS76243	Tad01	Hap2
CS76143	Hovdala-2	Hap2
CS28395	Kn-0	Hap2
CS76093	Ba1-2	Hap2
CS28277	Ge-1	Hap2
CS28133	Cha-0	Hap2
CS28564	No-0	Hap2
CS76242	Ta-0	Hap2

