GENETIC ANALYSIS OF NITROGEN USE EFFICIENCY IN ARABIDOPSIS THALIANA

NİHAL ÖZTOLAN EROL

GENETIC ANALYSIS OF NITROGEN USE EFFICIENCY IN ARABIDOPSIS THALIANA N. ÖZTOLAN EROL 2019

Propositions

1. Genome-wide association studies outperform conventional bi-parental linkage mapping studies in determining the genetic architecture of nitrogen use efficiency in *Arabidopsis thaliana*.
   (this thesis)

2. Nitrogen use efficient *Arabidopsis thaliana* accessions exhibit higher nitrogen utilization than uptake efficiency during the vegetative phase.
   (this thesis)

3. Machine learning with big data can contribute to sustainable agriculture through the implementation of prediction algorithms for crop management.

4. CRISPR gene editing is a promising new technology and its use should not be restricted by law if the genetic outcome is indistinguishable from existing genetic variation.

5. Scientific institutions that wish to attract innovative and progressive talent should not allow politics to influence their decisions and will need to respect freedom of expression.

6. Restrictions of the global exchange and flow of scientific talent, capital and know-how as a result of populist protectionist policies will impede scientific progress.

Propositions belonging to the thesis, entitled

"Genetic analysis of nitrogen use efficiency in *Arabidopsis thaliana*"

Nihal Öztolan Erol
Wageningen, 14 May 2019

Genetic Analysis of Nitrogen Use Efficiency in *Arabidopsis thaliana*

Nihal Öztolan Erol

Thesis committee

Promotors

prof. dr M.G.M. Aarts
Personal chair at the Laboratory of Genetics
Wageningen University & Research

prof. dr J.J.B. Keurentjes
Associate professor, Laboratory of Genetics
Wageningen University & Research

Co-promotor

prof. dr M. Koornneef
Emeritus professor, Laboratory of Genetics
Wageningen University & Research

Other members

dr L.M. Trindade, Wageningen University & Research
dr C.R.M. Hermans, Université Libre de Bruxelles, Belgium
dr E.J.W. Visser, Radboud University
dr L. Bentsink, Wageningen University & Research

This research was conducted under the auspices of the Graduate School Experimental Plant Sciences

Genetic Analysis of Nitrogen Use Efficiency in *Arabidopsis thaliana*

Nihal Öztolan Erol

Thesis
submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus,
Prof. Dr A.P.J. Mol,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Tuesday 14 May 2019
at 4 p.m. in the Aula.

Nihal Öztolan Erol

Genetic Analysis of Nitrogen Use Efficiency in *Arabidopsis thaliana*,

126 pages.

PhD thesis, Wageningen University, Wageningen, the Netherlands (2019)

With summary in English

ISBN 978-94-6343-440-9

DOI: https://doi.org/10.18174/471718

To Emre, Deniz, and Cevriye

Content:

Chapter 1

Introduction

Nitrogen (N) is an abundant element on earth; around 80% of the atmosphere consists of N_2. As it is vital for all living organisms, plants use N as a macronutrient in the composition of photosynthetic metabolites (Seeman 1987; Sage 1987), amino acids (Lam 1996) and also genetic information carriers. N is considered to be one of the most important growth limiting factors on crop productivity and it is eventually one of the crucial supplements of food production for feeding an increasing human population.

Although N exists in a very stable formation in the atmosphere, symbiotic microorganisms are able to make it available for other living systems and this creates a N cycle in nature. After N is fixed by microorganisms, it is released in the form of ammonium (NH_4^+) and nitrate (NO_3^-) to soil. Plants can only uptake N from soil in the form of NO_3^- and NH_4^+. The absorbable form of nitrogen depends on the pH of the soil and the type of plant. In low pH conditions, plants, which are adapted to that condition, take up N in the form of NH4+. However, at high pH, plants generally absorb N in the form of NO_3^- (Maathuis 2009). NO_3^- is readily available in aerated soils and it is rather mobile in the xylem. Once it is taken up it is either reduced to NH_4^+ to assimilate into the organic compounds and/or stored in the vacuoles. Whereas, NH_4^+ is toxic even at very low concentrations and it is very detrimental when stored in the vacuoles. Therefore, the incorporation of NH_4^+ into the organic compounds is a way to detoxify NH_4^+ taken up by roots and derived from NO_3^- reduction (Marschner 1995).

N fertilizers and their effect on different ecosystems

The green revolution in 1960's has left a booming effect on the agricultural production with the increased application of chemical disease treatments and synthetic nitrogen (N) fertilizers (Tilman 1998). Although it has met the food needs for the increasing human population, it has also remained behind a controversial fact about environmental pollution. N fertilizers, which lead to great anthropogenic impact by habitat destruction, account for approximately 85 to 90 million metric tons every year (Good et al. 2004). Unfortunately, only 33% of this amount can be taken up and assimilated by crops (Raun and Johnson 1999) due to leaching in groundwater, volatilization to the atmosphere, surface runoff or microbial denitrification. Indeed, many studies have revealed the negative impact of nitrogen fertilization on bacterial, animal and plant ecosystems, which include eutrophication of terrestrial and aquatic systems (Gruber and Galloway 2008). A striking example of eutrophication was the dramatic effect of hypoxia in the Gulf of Mexico caused by proliferation of phytoplankton in the seabed fed by over-loaded

N fertilizer flowing from Mississippi River, yearly reported by the Louisiana Universities Marine Consortium (Hypoxia).

The green revolution has also come along with the increased emission of greenhouse gases and loss of crop genetic diversity due to using fossil fuel to produce synthetic fertilizers and deforesting for croplands (Tilman 1998; Chen et al. 2014; Garnett et al. 2009). Novel restoring strategies have been investigated to establish an improved crop-environment interaction. N use efficiency (NUE) is a remedy which protects ecosystems with less N input and also boosts crop productivity with efficient N uptake and utilization. Changing fertilization practices in agriculture greatly affects crop improvement in NUE (Lassaletta et al. 2014). Netherlands is a good example in terms of increasing yield production and NUE by decreasing fertilization. A regular increase was recorded in both yield and fertilization between 1960-1975; however, since 1980's this trend was shifted towards yield increase without any fertilizing activity. Thanks to strict regulations and policies in Europe, many other European countries increase their crop productivity without further fertilization and even with decreasing fertilization (Lassaletta et al. 2014). Apart from changing practices in terms of N fertilizers, plant genetics techniques, together with novel biotechnological tools, facilitate to improve crop-environment interaction for a sustainable agriculture.

Nitrogen use efficiency

Crop breeders have so far developed their varieties under high N fertilized soils for an optimal yield, whereas crop breeding under low N fertilized soils still needs attention. Plants are very plastic in their response to environmental changes and can also adapt to N deficient conditions. They are able to assimilate limited amounts of N to complete their life cycle, starting from germination to seed production by maximizing NUE.

NUE has been interpreted in different ways in the literature. Chapin (1980) argued that the inverse of tissue nutrient concentration does not always explain nutrient use efficiency in all plants. Slowly growing species, adapted to marginal habitats, tend to accumulate more N than rapidly growing species, so by definition slowly growing species should be less efficient in nutrient use. However, rapidly growing species show deficiency symptoms like impaired photosynthesis, reduced root absorption capacity and respiration under nutrient limited conditions, whereas slowly growing species show none of these symptoms. Therefore Chapin concludes that nutrient use efficiency must include respiration, photosynthesis, and net nutrient assimilation per gram of input (Chapin 1980). Moll et al. (1982) stated that NUE is associated with efficient N absorption, translocation, assimilation and redistribution. However, direct

measurements of these features are not feasible in large fields, so they suggested using estimates of N accumulation in certain plant parts at certain growth stages for the definition of NUE (Moll R.H. 1982). Birk (1986) and Vitousek (1982) argued that these definitions are applicable for short-lived plants but not for long-living plants like forest stands. Therefore, they define NUE as the amount of organic matter lost from a plant or permanently stored in wood, divided by the amount of nitrogen lost or permanently stored (Vitousek 1982; Birk 1986). In turn Berendse and Aerts (1987) suggested a NUE definition, which would be applicable for any plant species. According to them, the calculation of NUE must be the product of two parameters in plants: the first one is the N productivity, which is simply the production of dry matter per unit of N in the plants; the second one is the mean residence time of N in the plant, which is the mean period that the plant uses N for carbon fixation (Berendse 1987). Finally, Good et al. (2004) suggested to make estimations on the basis of the crop of interest, its harvestable material, and specific physiological measurements related to NUE (Good et al. 2004). For the effect of NUE on grain as a harvestable product, it is calculated as the grain weight divided by the N supply. For vegetative harvest products Good et al. (2004) calculated NUE as the amount of harvestable dry matter per N content (note that N content is dry matter times N concentration (N%), so NUE is 1/N%) and this calculation is exempt from the residual dry biomass.

Only a few genetic studies on NUE in *Arabidopsis thaliana* have been reported and the approaches usually confuse N usage index (NUI) and N utilization efficiency (NUtE) with NUE. NUI is the total plant biomass multiplied by the ratio of the total plant biomass to the total N content, which is the dry matter times NUE (Han et al. 2015; Good et al. 2004). NUtE is the efficiency of assimilation and remobilization of plant N to produce the dry matter and the grain (Han et al. 2015; Good et al. 2004; Chardon et al. 2010). Chardon et al. (2010) used shoot dry weight (SDW) per N% to estimate NUE; but this calculation takes total biomass increase into account, which in fact is defined as the NUI by Good et al. (2004) (Chardon et al. 2010; Good et al. 2004). A recent study estimated NUE as SDW per N content in the pot; but for the calculation of NUtE they calculated SDW per shoot N content, similar to what Good et al. (2004) used for NUE calculations (Menz et al. 2018; Good et al. 2004). I used NUE and NUI definitions for *A. thaliana* as Good et al. (2004) suggested. Shoot dry matter per N content was used to estimate NUE and shoot dry matter times the ratio of the shoot dry matter to N content was used to estimate NUI in *A. thaliana.*

Two components of NUE: N uptake and utilization efficiencies for an improved crop production

N uptake efficiency (NUpE) and NUtE are two components of NUE. NUpE relies on the plant nutrient transportation system from source to sink organs, which is controlled by efficient plant nutrient uptake by roots and translocation. NUtE determines the increase in the shoot dry matter and the grain production, which is controlled by efficient nutrient assimilation into essential proteins and remobilization (Chardon et al. 2010). Therefore, it is very important to address plant adaptation mechanisms in terms of physiological changes in roots and shoots, such as lateral root elongation, prolonged leaf greenness, photosynthetic activity, shoot growth and seed formation under N deficient growth conditions.

The root to shoot ratio is evidently increasing with root length and the density of lateral roots. The root to shoot ratio, however, decreases under nutrient-sufficient conditions, in which plants invest in shoot growth rather than root growth (Tian et al. 2008). In Arabidopsis, lateral root growth was inhibited under increasing NO_3^- levels (Zhang and Forde 2000). In contrast, increased lateral root growth provides a larger root surface area, which facilitates increased nutrient uptake and reduced carbon allocation in root systems when N is deficient in soils (Marschner 1995; Peng et al. 2007; Bi et al. 2007; Wang et al. 2003; Wang et al. 2004). Lateral root growth is induced when the external N availability is low and homogeneous. However, when the external N availability is heterogeneous, lateral root growth is also induced in patches of high N content in the growth media. This explains the adaptation of root system architectures to forage N under stress conditions (Kant 2018; Chardon et al. 2010). The level of external N availability also regulates the activity of N transporters. A low affinity transport system is active when the concentration of soil-N is above 0.5 mM, whereas a high affinity transport system is active when the external N concentration is below 0.5 mM (Kant 2018; Krapp 2015). An efficient N uptake and transport activity enhances increased N accumulation in plants. The increase in NUpE may increase NUE without the contribution of NUtE (Garnett et al. 2009). However, such a system does not provide an improvement in yield capacity (Gao et al. 2018). Therefore, an efficient N remobilization and assimilation should secure increased yield capacity when orchestrating with NUpE.

N remobilization and assimilation efficiencies are components of NUtE (Yang and Udvardi 2018). Reduced photosynthesis and shoot growth, in addition to the accumulation of anthocyanins, are subsequent effects as opposed to increased root growth under N limited conditions (Bi et al. 2007; Loudet et al. 2003b; Peng et al. 2007; Wang et al. 2003; Wang et al. 2004). Anthocyanin production is one of the protection strategies of plants against photo-damage in poorly photosynthesizing conditions (Kant et al. 2011; Peng et al. 2007; Bi et al. 2007). The stress physiology in plants further triggers N remobilization, which initiates amino

acid translocation from source to sink organs, followed by assimilation in sink organs (Bi et al. 2007; Loudet et al. 2003b). Leaf senescence is eventually observed in older leaves during N limited conditions. During this process, leaf senescence is accompanied with chlorophyll, protein, lipid, and nucleic acid degradation and mineral translocation from source to sink organs and finally the senesced leaf is discarded from the plant (Buchanan-Wollaston et al., 2003). Nevertheless, leaf senescence is not a favorable trait among crop breeders. Delayed leaf senescence, or the stay-green trait, is an economically valuable trait to invest for many seed companies. The stay-green trait provides a prolonged shelf life especially for green crops such as lettuce, broccoli and/or high NUtE for all crops. The stay-green trait is investigated for two objectives: cosmetic stay-green and functional stay-green. Cosmetic stay-green considers only chlorophyll degradation such that, independent of the environmental condition, a plant does not initiate chlorophyll degradation even though a photosynthetic reduction can still be observed (Grassl et al., 2012). However, functional stay-green considers whole leaf delayed chlorophyll degradation mechanisms coupled to N remobilization and assimilation efficiencies, in which a plant remains green and extends photosynthesis for an increased yield production (Yang and Udvardi 2018; Han et al. 2015).

An efficient N assimilation, under N deficient conditions, enhances photosynthetic capacity. N is a rate limiting mineral for photosynthesis, due to a high amount of leaf-N fixed in the enzyme ribulose-1,5-biphosphate carboxylase/oxygenase (Rubisco) in the chloroplasts (Poorter and Evans 1998). Therefore, an increased photosynthetic capacity depends on the N assimilation in Rubisco and carbon fixation (Hikosaka and Shigeno 2009). In C3 plants, Rubisco acts as a key factor in CO_2 fixation in ribulose-1,5-biphosphate (RuBP) in the Calvin cycle. However, under high temperature, high irradiance and low CO_2 conditions, this enzyme has increased affinity to O_2 leading to photorespiration and fixating O_2 by RuBP. Photorespiration is an unfavorable pathway, which consumes energy and resources to produce glucose molecules. NUE has become an important aspect for C4 plants as morphological and biochemical differences between C3 and C4 plants result in large differences in their photosynthetic capacities. C4 plants developed biochemical and morphological modifications to cope with this inefficient photosynthetic trade-off. The main characteristic of C4 photosynthesis is to concentrate CO_2 around Rubisco in the bundle sheats, where diffusion of O_2 is almost impossible. This makes C4 photosynthesis more efficient in terms of biomass production and resource use under low CO_2 conditions and high temperatures. Under these conditions, reduction in N consumption enhances NUE in C4 plants because an efficient activity of the Rubisco enzyme requires only 4-21% of the leaf N content whereas C3 plants consume more than 40% of the leaf N content

for sufficient Rubisco activity (Williams et al. 2012; Hikosaka and Hirose 2000; Ripley et al. 2008).

During seed formation, N remobilization and assimilation mediates N partitioning into the seeds. In this process, vegetative organs become source and reproductive organs become sink organs. To provide a good nutrient-supply to the growing reproductive tissues, plants maintain their biochemical and photosynthetic machinery for growth, architecture and development under N limited conditions (Masclaux-Daubresse et al., 2010). Upholding maximum yield-capacity under N limited conditions is a challenge for crop breeders. It is, therefore, important to understand the physiological mechanisms of varieties that obtain a better N uptake and utilization-capacity.

In this thesis, Good et al. was taken as a reference for making rational estimations for NUE in *A. thaliana* (Good et al. 2004). If an accession can still produce high yield by accumulating less N, it does not need further N fertilization. Therefore, by definition, improving NUtE and N translocation-efficiency in crops have more potential for crop improvement than N accumulation-efficiency. In my research, N utilization related traits such as shoot growth, carbon concentration per unit biomass and photosynthetic activity were recorded in *A. thaliana.* Causal genetic variants, controlling these traits, were investigated to assign candidate genes involved in sustainable crop-production under low N conditions.

Genetic analysis of NUE in *Arabidopsis thaliana*

A. thaliana, as a member of the mustard family (Brassicaceae), is a good reference for many plant species (David W. Meinke and Koornneef 1998; Rauh et al. 2002; Koornneef et al. 2004). Extensive genomic and physical maps of its small genome size (120 megabase) with 5 chromosomes, self-fertilization, and rapid life cycle are advantageous for scientific research (Weigel and Mott 2009; TAIR). A large number of resources, such as T-DNA libraries and silenced and overexpressed gene sets, make *A. thaliana* an unequivocal resource for other higher plants (Koornneef et al. 2004).

A. thaliana has a wide geographic distribution occupying Eurasia, North America and Africa with a large latitudinal range (Durvasula et al. 2017). Its range of habitats differs from disturbed to cultivated soils and from sea level to up to the mountains (Koornneef et al. 2004). This wide geographic distribution makes *A. thaliana* a suitable model for studying the genetics of adaptation mechanisms by using natural variation. The natural variation in *A. thaliana* in NUE was previously studied to identify which accessions are adapted to N deficiency, how it is phenotypically expressed and which genes are up- or down-regulated under the stress

conditions (Gifford et al. 2013; Chardon et al. 2012; Chardon et al. 2010; Menz et al. 2018). Along with its easy growth, its small genome size, the availability of several functional genetic analysis tools, and the existing natural genetic variation make *A. thaliana* an advantageous plant to investigate NUE.

Quantitative trait loci (QTL) mapping has been used to identify putative genes that might control traits under N deficient conditions. To date, family-based populations have been solely used for the genetic mapping of NUE and related traits. Rauh et. al. (2002) used recombinant inbred lines (RILs) derived from a cross of Columbia (Col-0) and Landsberg *erecta* (L*er*), to identify loci controlling aerial mass, root mass and root length of plants growing in different N sources (nitrate, ammonium nitrate, or ammonium) and also in limited N supply (Rauh et al. 2002). Loudet et al. contributed to QTL mapping of NUE in *A. thaliana* by studying a Bayreuth-0 x Shahdara RIL population in two different N supply conditions and mapping shoot growth, total N, nitrate, amino acid, anions and also water content (Loudet et al. 2003b; Loudet et al. 2003a). These studies associated the amino acid transporter gene *AAP5*, the cytosolic GS gene *GLN1.2*, and a high-affinity nitrate transporter gene *NRT2.6* with shoot growth, total N, nitrate and amino acid content; aquaporin, chloride, and phosphate channel/transporters with water and anion contents. In another study Diaz et al. investigated the genetic co-regulation of NUE, flowering time, anthocyanin accumulation and leaf yellowing in the Bayreuth-0 x Sha RIL population (Diaz et al. 2006). The Col-0 and Br-0 accessions were screened under low and normal N conditions showing contrasting behavior under low N conditions in this thesis. Later, a novel RIL population was used to find out QTLs contributing to differing NUE phenotypes.

Using population of natural accessions of *A. thaliana* (HapMap population) has not yet been done for the genetic mapping of NUE and related traits. In this thesis, a genome-wide association study (GWAS) was conducted by utilizing genetic variation of the HapMap population for the identification of functional alleles associated with NUE and other related traits. GWAS is the analysis of statistical association of genotypes and phenotypes in a set of individuals (Rafalski 2010). It is based on linkage disequilibrium mapping, which decays after approximately 10 kb in *A. thaliana* (Brachi et al. 2010; Kim et al. 2007; Atwell et al. 2010). A dense and high resolution physical map, created with 250K single nucleotide polymorphism (SNP) markers (Weigel and Mott 2009), provides a great advantage to pinpoint naturally occurring single nucleotide mutations within the population used for GWAS (Nordborg and Weigel 2008; Korte and Farlow 2013). The advantages of using the HapMap population are the presence of several recombinant individuals because of random mating in nature and homozygosity because of the self-inbreeding property of *A. thaliana* (Mackay 2009). The

detection power in GWAS increases with the magnitude of effect and also the frequency of an allele in a natural population (Rafalski 2010).

In this thesis, GWAS was first conducted using the HapMap population to find out putative loci on *A. thaliana* genome that might associate with NUE and related traits. Two extreme accessions from this population were later used for bi-parental linkage mapping. As both mapping techniques yielded different results, co-localized QTLs might shed a light to the genetics of NUE.

Scope, objectives and outline of this thesis

The scope of this thesis is to phenotype shoot traits of *Arabidopsis thaliana* to understand how causal alleles regulate the adaptation to N deficiency stress.

The objectives of this thesis are (i) shedding light on the genetic mechanisms that explain NUE; (ii) understanding what NUE means for plants by using *A. thaliana* as a model organism; (iii) assessing the applicability of NUE for crop improvement.

Chapter 1 introduces the importance of N for plants and NUE for a sustainable agriculture. Different NUE definitions are presented and discussed to assess the best approach for *A. thaliana*. How the natural variation in *A. thaliana* was exploited to get candidate genetic variants is also explained here.

In Chapter 2 a GWAS was conducted in the HapMap population of *A. thaliana*, growing under sub-optimal and optimal N conditions. This is the first step to identify causal alleles associated with shoot dry weight (SDW), shoot fresh weight (SFW), shoot N concentration (N%), NUI, NUE, total leaf area (TLA), relative growth rate (RGR), Photosystem II efficiency (ΦPSII), and water content (WC). Phenotypic plasticity of SDW, SFW, N%, NUI, NUE, ΦPSII and WC was also investigated to detect alleles associated with the change of phenotypic expression across different conditions.

In Chapter 3, accessions with high shoot dry weight but contrasting NUE, called N efficient (NE) and N inefficient (NI) accessions, were selected and reassessed, as a follow-up for understanding how SDW, SFW, WC, N%, NUE, NUI, ΦPSII, TLA and RGR were represented for NE and NI accessions under sub-optimal N conditions. Phenotypic plasticity of these traits, except for TLA and RGR, was investigated to see which accessions show plasticity across the sub-optimal and optimal N supplies.

In Chapter 4, an F_8 RIL population, for which parents were selected from each of the NE and NI groups, was analyzed for QTLs associated with SDW, SFW, N%, NUI, NUE, C concentration (C%), and TLA.

In Chapter 5, an overview of examples of the genetic research of NUE in crop improvement is presented. The importance of NUE is discussed for the present and the future of the agriculture and nature.

References

Atwell S, Huang YS, Vilhjalmsson BJ, Willems G, Horton M, Li Y, Meng D, Platt A, Tarone AM, Hu TT, Jiang R, Muliyati NW, Zhang X, Amer MA, Baxter I, Brachi B, Chory J, Dean C, Debieu M, de Meaux J, Ecker JR, Faure N, Kniskern JM, Jones JD, Michael T, Nemri A, Roux F, Salt DE, Tang C, Todesco M, Traw MB, Weigel D, Marjoram P, Borevitz JO, Bergelson J, Nordborg M (2010) Genome-wide association study of 107 phenotypes in Arabidopsis thaliana inbred lines. Nature 465:627-631.

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

Bi YM, Wang RL, Zhu T, Rothstein SJ (2007) Global transcription profiling reveals differential responses to chronic nitrogen stress and putative nitrogen regulatory components in Arabidopsis. BMC genomics 8:1-17.

Birk EM, Vitousek PM (1986) Nitrogen availability and nitrogen use efficiency in loblolly pine stands. Ecology 67:69-79.

Brachi B, Faure N, Horton M, Flahauw E, Vazquez A, Nordborg M, Bergelson J, Cuguen J, Roux F (2010) Linkage and association mapping of Arabidopsis thaliana flowering time in nature. PLoS Genetics 6:1-17.

Chapin FS (1980) The mineral nutrition of wild plants. Annual Review of Ecology and Systematics 11:233-260.

Chardon F, Barthelemy 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.

Chardon F, Noel V, Masclaux-Daubresse C (2012) Exploring NUE in crops and in Arabidopsis ideotypes to improve yield and seed quality. Journal of Experimental Botany 63:3401-3412.

Chen XP, Cui ZL, Fan MS, Vitousek P, Zhao M, Ma WQ, Wang ZL, Zhang WJ, Yan XY, Yang JC, Deng XP, Gao Q, Zhang Q, Guo SW, Ren J, Li SQ, Ye YL, Wang ZH, Huang JL, Tang QY, Sun YX, Peng XL, Zhang JW, He MR, Zhu YJ, Xue JQ, Wang GL, Wu L, An N, Wu LQ, Ma L, Zhang WF, Zhang FS (2014) Producing more grain with lower environmental costs. Nature 514:486-489.

Meinke DW, Cherry JM, Dean C, Rounsley SD, Koornneef M (1998) Arabidopsis thaliana: A model plant for genome analysis. Science 282:662+679-682.

Diaz C, Saliba-Colombani V, Loudet O, Belluomo P, Moreau L, Daniel-Vedele F, Morot-Gaudry JF, Masclaux-Daubresse C (2006) Leaf yellowing and anthocyanin accumulation are two genetically independent strategies in response to nitrogen limitation in Arabidopsis thaliana. Plant & Cell Physiology 47:74-83.

Durvasula A, Fulgione A, Gutaker RM, Alacakaptan SI, Flood PJ, Neto C, Tsuchimatsu T, Burbano HA, Pico FX, Alonso-Blanco C, Hancock AM (2017) African genomes illuminate the early history and transition to selfing in Arabidopsis thaliana.

Proceedings of the National Academy of Sciences of the United States of America 114:5213-5218.
Gao R, Guo G, Fang C, Huang S, Chen J, Lu R, Huang J, Fan X, Liu C (2018) Rapid generation of barley mutant lines with high nitrogen uptake efficiency by microspore mutagenesis and field screening. Frontiers in Plant Science 9:450.
Garnett T, Conn V, Kaiser BN (2009) Root based approaches to improving nitrogen use efficiency in plants. Plant, Cell & Environment 32:1272-1283.
Gifford ML, Banta JA, Katari MS, Hulsmans J, Chen L, Ristova D, Tranchina D, Purugganan MD, Coruzzi GM, Birnbaum KD (2013) Plasticity regulators modulate specific root traits in discrete nitrogen environments. PLoS Genetics 9:1-13.
Good AG, Shrawat AK, Muench DG (2004) Can less yield more? Is reducing nutrient input into the environment compatible with maintaining crop production? Trends in Plant Science 9:597-605.
Gruber N, Galloway JN (2008) An Earth-system perspective of the global nitrogen cycle. Nature 451:293-296.
Han M, Okamoto M, Beatty PH, Rothstein SJ, Good AG (2015) The Genetics of nitrogen use efficiency in crop plants. Annual Review of Genetics 49:269-289.
Hikosaka K, Hirose T (2000) Photosynthetic nitrogen-use efficiency in evergreen broad-leaved woody species coexisting in a warm-temperate forest. Tree Physiology 20:1249-1254.
Hikosaka K, Shigeno A (2009) The role of Rubisco and cell walls in the interspecific variation in photosynthetic capacity. Oecologia 160:443-451.
Hypoxia GoM https://gulfhypoxia.net/research/shelfwide-cruise/?y=2013 ,on https://gulfhypoxia.net.
Kant S (2018) Understanding nitrate uptake, signaling and remobilisation for improving plant nitrogen use efficiency. Seminars in Cell Development Biology 74:89-96.
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 62:1499-1509.
Kim S, Plagnol V, Hu TT, Toomajian C, Clark RM, Ossowski S, Ecker JR, Weigel D, Nordborg M (2007) Recombination and linkage disequilibrium in Arabidopsis thaliana. Nature Genetics 39:1151-1155.
Koornneef M, Alonso-Blanco C, Vreugdenhil D (2004) Naturally occurring genetic variation in Arabidopsis thaliana. Annual Review of Plant Biology 55:141-172.
Korte A, Farlow A (2013) The advantages and limitations of trait analysis with GWAS: a review. Plant Methods 9:29.
Krapp A (2015) Plant nitrogen assimilation and its regulation: a complex puzzle with missing pieces. Current Opinion in Plant Biology 25:115-122.
Lam HM, Coschigano KT, Oliveira IC, Melo-Oliveira R, Coruzzi GM (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.
Lassaletta L, Billen G, Grizzetti B, Anglade J, Garnier J (2014) 50 year trends in nitrogen use efficiency of world cropping systems: the relationship between yield and nitrogen input to cropland. Environmental Research Letters 9:1-9.
Loudet O, Chaillou S, Krapp A, Daniel-Vedele F (2003a) Quantitative trait loci analysis of water and anion contents in interaction with nitrogen availability in Arabidopsis thaliana. Genetics 163:711-722.
Loudet O, Chaillou S, Merigout P, Talbotec J, Daniel-Vedele F (2003b) Quantitative trait loci analysis of nitrogen use efficiency in Arabidopsis. Plant physiology 131:345-358.

Maathuis FJM (2009) Physiological functions of mineral macronutrients. Current Opinion in Plant Biology 12:250-258.
Mackay TF (2009) Q&A: Genetic analysis of quantitative traits. Journal of Biology 8:23.
Marschner H (1995) Mineral nutrition of higher plants. 2nd edn. Academic Press, London etc.
Menz J, Range T, Trini J, Ludewig U, Neuhauser B (2018) Molecular basis of differential nitrogen use efficiencies and nitrogen source preferences in contrasting Arabidopsis accessions. Scientific Reports 8:1-11.
Moll RH, Kamprath EJ, Jackson WA (1982) Analysis and interpretation of factors which contribute to efficiency of nitrogen utilizatilon. Agronomy Journal 74:562-564.
Nordborg M, Weigel D (2008) Next-generation genetics in plants. Nature 456:720-723.
Peng M, Bi YM, Zhu T, Rothstein SJ (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.
Poorter H, Evans JR (1998) Photosynthetic nitrogen-use efficiency of species that differ inherently in specific leaf area. Oecologia 116:26-37.
Rafalski JA (2010) Association genetics in crop improvement. Current Opinion in Plant Biology 13:174-180.
Rauh L, Basten C, Buckler St (2002) Quantitative trait loci analysis of growth response to varying nitrogen sources in Arabidopsis thaliana. Theoretical and Applied Genetics 104:743-750.
Raun WR, Johnson GV (1999) Improving nitrogen use efficiency for cereal production. Agronomy Journal 91:357-363.
Ripley BS, Abraham TI, Osborne CP (2008) Consequences of C4 photosynthesis for the partitioning of growth: a test using C3 and C4 subspecies of Alloteropsis semialata under nitrogen-limitation. Journal of Experimental Botany 59 (7):1705-1714.
Sage RF, Pearcy RW, Seemann JR (1987) The nitrogen use efficiency of C3 and C4 plants. Plant Physiology 85:355-359.
Seeman JR, Sharkey TD, Wang JL, Osmond CB (1987) Environmental effects on photosynthesis, nitrogen-use efficiency, and metabolite pools in leaves of sun and shade plants. Plant Physiology 84:796-802.
TAIR https://www.arabidopsis.org/portals/education/aboutarabidopsis.jsp, on www.arabidopsis.org.
Tian Q, Chen F, Liu J, Zhang F, Mi G (2008) Inhibition of maize root growth by high nitrate supply is correlated with reduced IAA levels in roots. Journal of Plant Physiology 165:942-951.
Tilman D (1998) The greening of the green revolution. Nature 396:211-212.
Vitousek P (1982) Nutrient cycling and nutrient use efficiency. American Naturalist 119:553-572.
Wang R, Okamoto M, Xing X, Crawford NM (2003) Microarray analysis of the nitrate response in Arabidopsis roots and shoots reveals over 1,000 rapidly responding genes and new linkages to glucose, trehalose-6-phosphate, iron, and sulfate metabolism. Plant Physiology 132:556-567.
Wang R, Tischner R, Gutierrez RA, Hoffman M, Xing X, Chen M, Coruzzi G, Crawford NM (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.
Williams BP, Aubry S, Hibberd JM (2012) Molecular evolution of genes recruited into C(4) photosynthesis. Trends in Plant Science 17:213-220.

Yang J, Udvardi M (2018) Senescence and nitrogen use efficiency in perennial grasses for forage and biofuel production. Journal of Experimental Botany 69:855-865.
Zhang H, Forde BG (2000) Regulation of Arabidopsis root development by nitrate availability. Journal of Experimental Botany 51:51-59.

Chapter 2

Genome-wide association study of *Arabidopsis thaliana* growth and response to sub-optimal nitrogen supply

Nihal Oztolan Erol[1], Xi Wei[1], Padraic Flood[1], Joost JB Keurentjes[1], Mark GM Aarts[1]

[1]Laboratory of Genetics, Wageningen University, Wageningen, the Netherlands

Abstract:

Nitrogen use efficiency (NUE) in plants is an important trait in order to decrease the agricultural production costs and the detrimental effect of excessive use of synthetic N fertilization on environment. It is therefore imperative to identify the genes regulating nitrogen use efficiency (NUE) in plants. A population of *Arabidopsis thaliana*, hereafter referred to as the HapMap population, was used in this chapter to identify candidate genes related to NUE. Shoot dry and fresh weights (SDW and SFW), total leaf area (TLA), shoot N concentration (N%), Photosystem II (PSII) efficiency (ΦPSII) were measured and water content (WC), NUE, N usage index (NUI), and relative growth rate (RGR) were calculated for both sub-optimal (1 mM) and optimal (5 mM) N supply. Genome-wide association study (GWAS) was used to identify trait-associated markers and allelic variation in candidate genes underlying the traits. Overall, 28 QTLs were identified with marker trait associations above an arbitrary threshold of -log10(*p*-value) ≥ 4. These correspond to 157 candidate genes, as determined based on linkage disequilibrium association with the most significant SNP. Candidate genes are enriched for functions in hormone signalling; transmembrane transport; and carbon metabolism and photosynthesis. Among the candidate genes are several that previously have been implicated in plant response to a variation in N supply, but also several novel candidate genes, with proposed functions in plant growth and developmental, are identified, which may be involved in the regulation of N homeostasis under sub-optimal N supply.

Introduction

Nitrogen (N) is one of the crucial growth-limiting macro-elements for plants. It has a significant role in many plant development processes as a component of amino acids, signaling molecules and also genetic coding molecules. It is, therefore, one of the crucial supplements for food production for the increasing human population. To overcome N shortage in soil, N fertilizers have been applied in large quantities for many years. However, only approximately one-third of the applied N in fertilizers can be taken up and assimilated by crops (Raun W.R. 1999) due to leaching to groundwater, volatilization to the atmosphere, surface runoff or microbial denitrification. Many studies have demonstrated the negative impact of N fertilization on bacterial, animal and plant ecosystems, including eutrophication of terrestrial and aquatic systems (Gruber and Galloway 2008). Apart from the negative ecological effects, the economic costs of N fertilizers increase with an increasing demand for fossil fuels needed to produce them (Garnett et al. 2009). Thus, from an ecological and an economic perspective, the use of N fertilizers for agriculture comes at a high cost. Improving N use efficiency (NUE) of crops not only secures high yields at low inputs but also provides an ecologically sustainable and economically viable agriculture for an increasing human population. NUE was previously defined as grain and biomass production per unit of N content in plant (Good et al. 2004). Plants evolved to survive on N deficient soils and to assimilate N efficiently to complete their vegetative and reproductive growth. Adaptation to deficient conditions modifies a whole suite of plant features, ranging from morphology to metabolism. Genes controlling these modifications, might be functionally altered or transcriptionally up- or down-regulated in the course of evolution. However, adaptive responses of plants may vary according to their ecological conditions, so different strategies of coping with limited N supplies might have evolved in different accessions.

Plants perform different strategies to cope with the changing environmental conditions. They can either be stable or modify their phenotypes in order to adapt to the surrounding conditions (El-Soda et al. 2014). The adaptational change of phenotypes regarding the environmental conditions is called phenotypic plasticity. Different *A. thaliana* accessions perform differing phenotypic plasticity, which is called as genotype by environment interaction (El-Soda et al. 2014). For instance, under given conditions, (I) two different genotypes may perform similar reactions, resulting in parallel reaction norms, the ranking of genotypes does not change; (II) they may perform different reactions, where there is no intersection of reaction norms, so the ranking still does not change; (III) they may perform contrasting reactions, where there is a cross-over intersection of reaction norms which could be explained as one genotype performs

a decreasing reaction, whereas the other performs an increasing reaction. The ranking of genotypes, thereby, changes under the given circumstances (El-Soda et al. 2014; Lacaze et al. 2009). The natural variation in phenotypic plasticity in *A. thaliana* grown under changing N supply was previously discussed all these reaction forms in different genotypes under given N conditions (Chardon et al. 2010; Menz et al. 2018; Meyer et al. 2019; North et al. 2009). They often compared the ratio of phenotypic responses to increasing N supply. Residual analysis is also suggested to quantify phenotypic plasticity in plants (Anderson et al. 2012). Residual analysis measures the difference between the observed value of dependent variable (y) and the expected value of dependent variable ($\hat{y}$). This method could be easily adapted for changing N supply conditions to dissect the genetic mechanisms behind the adaptation and to understand general adaptation behavior of a population.

A. thaliana has been previously used in several genetic studies into NUE. Genome-wide microarray and transcriptome analysis revealed many N-responsive genes. Mild and severe stress conditions, long term and short term N applications cause differential up- or down-regulation of many N-responsive genes (Wang et al. 2003; Bi et al. 2007; Peng et al. 2007; Meyer et al. 2019). Overall, significant changes in plant physiology under low N availability are associated with arrested growth and development, impaired photosynthesis, and increased anthocyanin accumulation.

Limited N activates protein degradation mechanisms by inducing ubiquitination and amino acid catabolism, whereas cytokinin down-regulation is one of the reasons for arrested growth and development. Furthermore, lateral root elongation, caused by auxin activity, is triggered by N scavenging as an adaptive response to limited N conditions. Carbon metabolism is also affected by limited N in addition to impaired anabolism of nitrogenous macromolecules. Accordingly, genes involved in photosynthesis, such as those responsible for chlorophyll synthesis, chlorophyll a/b binding proteins, subunits of PS I and PS II, plastocyanin proteins, and Calvin-Benson cycle enzymes are down-regulated. Additionally, genes encoding enzymes active in the tricarboxylic acid cycle and pentose phosphate pathway are also down-regulated. Genes involved in starch accumulation and cell wall biosynthesis are also N supply responsive and can be up- or down-regulated depending on the adaptation strategy (Peng et al. 2007; Bi et al. 2007; Wang et al. 2003; Wang et al. 2004).

Inevitable degradation of nitrogenous macromolecules activates N translocation and re-assimilation regulatory genes. Genes involved in nitrate and ammonium transport and re-assimilation are up-regulated under N limited conditions (Bi et al. 2007; Loudet et al. 2003b). Stress-induced accumulation of toxic compounds, released during the cell death process, is also

under control of genes involved in the detoxification (Peng et al. 2007; Bi et al. 2007). Anthocyanin accumulation takes place in order to protect plant cells against photodamage, which might be realized due to N limited conditions. Therefore, genes regulating anthocyanin and phenylpropanoid synthesis are up-regulated (Peng et al. 2007; Bi et al. 2007).

These studies show that a lot of NUE regulating genes has been uncovered, however there are still many unknown genes involved in the NUE mechanism. A genome-wide association study (GWAS) can identify novel genes based on genetic analysis, rather than expression differences, making us of the allelic variation expected to be present in large collections of natural germplasm.

The HapMap population, consisting of 354 natural inbred lines showing genetic diversity based on differences in geography and ecology (Zhu et al. 2008), has previously been analyzed by GWAS for a wide variety of traits, including development-related phenotypes (Atwell et al. 2009; Baxter et al. 2010; Chao et al. 2012; Platt et al. 2010; Kooke et al. 2016; Bac-Molenaar et al. 2016). Considering the decay of linkage disequilibrium (LD) within approximately 10 kb in *A. thaliana* (Atwell et al. 2009; Kim et al. 2007), a high mapping resolution can be obtained when using around 250,000 SNP markers, approximating genome-wide gene density. In this chapter, the HapMap population has been analyzed for variation in NUE and NUE related agronomic traits (NRAT) (shoot fresh and shoot dry weights (SFW and SDW), water content (WC), shoot N concentration (N%), N usage index (NUI), PSII efficiency (ΦPSII), total leaf area (TLA), relative growth rate (RGR)) in both sub-optimal and optimal N supply. In order to understand how *A. thaliana* reacts to changing N supply conditions and how genetic variants control phenotypic plasticity, residual values of SDW (SDW_Res), SFW (SFW_Res), N% (N%_Res), NUE (NUE_Res), NUI (NUI_Res), WC (WC_Res) and Photosystem II (ΦPSII_Res) have been analyzed through GWAS as well. These analyses revealed a number of known and novel candidate genes, putatively involved in the regulation of NUE and NRAT.

Materials and Methods

Plant material and experimental setup

The HapMap population, containing 354 diverse accessions of natural genotypes of *A. thaliana,* was previously described in (Li et al. 2010). These accessions, collected from the northern hemisphere and genotyped with 250k SNP bi-allelic markers, were selected to maximize genetic diversity and reduce population structure as much as possible.

The entire HapMap population was grown in a climate cell at short day length (10 hours light) conditions. Light intensity was set at 200 µmol m^{-2} s^{-1}; humidity at 60%; and day and night temperatures were 20˚C and 18˚C, respectively. The experiment was carried out using a randomized design with three blocks per treatment and one replication per block. Plants were grown on rock-wool blocks placed on flooding tables in order to sub-irrigate and drain rock-wool blocks automatically (Kooke et al. 2016) Two treatments were applied, consisting of an optimal and sub-optimal N-supply. Standard nutrient solution contained basal ions (2.94 mM K, 1.31 mM Ca, 0.402 mM Mg, 0.938 mM SO_4, 0.536 mM P, 7.5 µM Fe, 3 µM Mn, 1.5 µM Zn, 5 µM B, 0.25 µM Cu, 0.25 µM Mo) supplemented with either 0.1 mM NH_4 / 0.904 mM NO_3 for the sub-optimal nitrogen condition or 0.5 mM NH_4 / 4.521 mM NO_3 for the optimal condition. Plants were watered with nutrient solution three times per week.

Plant phenotyping

During plant growth, ΦPSII was recorded automatically using a high-throughput phenotyping platform to determine ΦPSII, according to (Baker 2008). ΦPSII was recorded starting from the second week after germination, at three times per day (10:14, 13:14, and 16:14), until the end of the fourth week. TLA was monitored at eight time points on every day (00:12, 03:12, 06:12, 09:12, 12:12, 15:12, 18:12, and 21:12). RGR was estimated using the formula $(lnA_x\text{-}lnA_y)/dt_{(x\text{-}y)}$, which indicates the difference in the leaf area value $(lnA_x\text{-}lnA_{y)}$ of a certain time period $(dt_{(x\text{-}y)})$ (El-Lithy et al. 2004). ΦPSII, TLA and RGR were determined for the sub-optimal and optimal N supply conditions separately.

Plants were harvested 32 days after sowing upon which SFW was determined. SDW was determined after drying at 60˚C for three days. N% was measured using the micro-combustor DUMAS method (Matejovic 1995). Subsequently, NUE and NUI were calculated as *SDW/N* and *SDW*(SDW/N)* respectively (where N is shoot N content) (Good et al. 2004). Water content (WC) was calculated as *(SFW-SDW)/SDW*.

Statistical Analysis

Descriptive analysis (maximum & minimum values, mean, standard deviation, and coefficient of variance (CV)) and analysis of variance (ANOVA) for testing of differences between genotypes, treatments and genotype by treatment interaction, was performed on raw data using

IBM SPSS Statistics 19 (New York: IBM). Broad-sense heritability was calculated using the formula:

$$H^2 = \sigma_g^2 / \sigma_g^2 + \sigma_e^2 ,$$

where σ_g^2 represents genetic variance and σ_e^2 represents residual variance.

In order to test phenotypic plasticity of genotypes, linear and fitted regressions between the control and sub-optimal nitrogen conditions were investigated for all traits (except for TLA and RGR) using IBM SPSS Statistics 19 (New York: IBM) (Chu et al. 2008; Maloof et al. 2001).
The linear regression model was created as:

$$y = \beta_0 + \beta_1 \, x + \varepsilon$$

where y is the observed values of the optimal N treatment and x is that of the sub-optimal N treatment; β_0 and β_1 refer to the regression coefficients explaining the intercept and the slope of the regression line respectively; ε refers to the random error.
The fitted regression model was created as:

$$\hat{y} = \beta_0 + \beta_1 \, x$$

where $\hat{y}$ refers to the estimated values of the optimal N treatment for the given x, the observed values of the sub-optimal N treatment.

The residual values (e) were calculated as the difference between the estimated ($\hat{y}$) and the observed (y) values of the optimal N treatment:

$$e = y - \hat{y}$$

Genome wide association study (GWAS)

GWAS was performed on all phenotypic data using an efficient mixed-model association expedited (EMMAX) to correct population structure, for which an R-package is available (http://www.r-project.org/) (Kang et al. 2010). The mixed model was built supposing marker effects are zero, so the hypothesis was H_0: $\beta_j = 0$;

$$Y = \mu I_n + x\beta_j + g + e$$

where Y is the vector of phenotype value for every genotype $Y = (Y_{1;\ldots\ldots;}Y_n)$' and n is the number of genotypes = 354 ; μ is the population mean, 1_n is the vector of ones $1_n = (1;\ldots;1)$'; x is the vector of bi-allelic (Col-0 vs non-Col-0) SNP marker score $x = (x_1;\ldots;x_n)' \in \{0,1\}$, β_j is the effect of SNP marker $j = (1;\ldots;p)$' and p is the number of markers = 214051; g is the vector of genotypic effect $g = (g_1;\ldots.;g_n)$' which is normally distributed with covariance matrix $\sigma^2_g K$, where K is the kinship matrix on all available SNPs; e is the vector of independent environmental noise $e = (e_1;\ldots..;e_n)$'. The minor allele frequency threshold was set to 0.05, and the significance threshold was arbitrarily set to $-\log_{10}(p\text{-value}) = 4$ (Li et al. 2010; Atwell et al. 2009; El-Soda et al. 2015).

Candidate gene selection

QTLs were selected based on the appearance of two or more SNPs within the same LD region above the arbitrary threshold $-\log_{10}(p\text{-value}) \geq 4$ (El-Soda et al. 2015). Candidate genes were selected within 10 kb proximity of the most significant SNPs in QTLs with an LD R^2 value 0.5. Their expression profile and function information were determined using gene ontology annotations, from The Arabidopsis Information Resource 10 (TAIR 10: www.arabidopsis.org) were used during confirmation process.

Results

Descriptive analyses of phenotypes

The *A. thaliana* HapMap population of 354 accessions was grown in sub-optimal (1 mM) and optimal (5 mM) N supply conditions. Plants grown under sub-optimal nitrogen conditions showed significantly different responses compared to plants grown under optimal conditions (Figure 1 and Table 1), most obvious being symptoms of nitrogen deficiency. Chlorosis, anthocyanin accumulation, enhanced senescence of older leaves, retarded shoot growth and early flowering were observed.

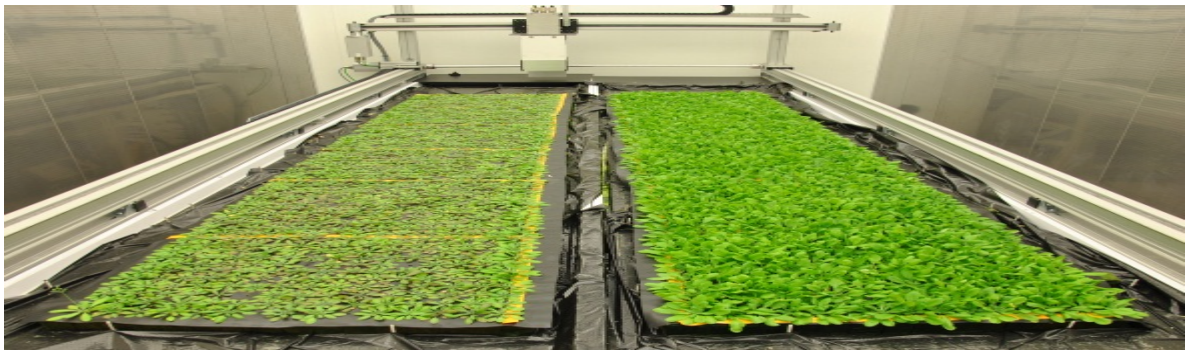


Figure 1: The HapMap population growing on flooding benches. **Left:** under sub-optimal N conditions (1 mM). **Right:** under optimal conditions (5 mM).

Plants growing under optimal condition had higher SFW, SDW, N%, and WC than those growing under sub-optimal N supply. However, a better estimate of NUE and NUI could only be acquired under sub-optimal conditions as NUE and NUI were higher in plants grown in sub-optimal compared to those grown in optimal conditions.

ANOVA showed significant differences between genotypes and between treatments, as well as genotype by environment interactions. Differences between replications for ΦPSII, SDW, N%, NUE, NUI, and TLA per treatment were not significant; however, replicates of SFW, WC, and RGR were significantly different which might explain experimental errors on these traits (Table 1).

Frequency distributions indicated a normal distribution for all traits and treatments and CV explained the presence of phenotypic variation in the population. SDW, SFW, NUI, and TLA showed larger phenotypic variation than ΦPSII, N%, NUE, WC, and RGR. H^2 estimates ranged between 38% to 80%. 80%, 75%, and 69% of phenotypic variations in N%, NUE, and NUI, respectively, were controlled by genes, not by environment, under sub-optimal N condition.

Table 1: Minimum and maximum values, mean, standard deviation, coefficient of variation (CV), genetic variation (σ_g^2), residual variation (σ_e^2), broad sense heritability (H^2) and significance of variation among genotypes (G), replications (R), treatments (T), and of genotype-by-environment interactions (G x E) of Photosystem II efficiency (Φ PSII), shoot dry weight (SDW), shoot fresh weight (SFW), nitrogen concentration (N%), nitrogen use efficiency (NUE), nitrogen usage index (NUI), water content (WC), total leaf area (TLA), and relative growth rate (RGR) under sub-optimal N (S) and optimal (O) conditions. s refers to significant (p-value<<0,05) and ns refers to non-significant (p-value>0,05).

Traits	Descriptive Statistics					σ_g^2	σ_e^2	H^2	ANOVA			
	Minimum	Maximum	Mean	Std. Deviation	CV				G	R	T	G x E
Chl Fluo O	0.597	0.692	0.639	0.013	0.02	0.0002	0.0001	0.58	s	ns	s	s
Chl Fluo S	0.600	0.709	0.671	0.015	0.02	0.0003	0.0004	0.45				
SDW O	0.033	0.204	0.105	0.032	0.30	0.0010	0.0005	0.65	s	ns	s	s
SDW S	0.017	0.131	0.065	0.016	0.25	0.0002	0.0002	0.52				
SFW O	0.364	2.319	1.177	0.368	0.31	0.1357	0.0743	0.64	s	s	s	s
SFW S	0.085	0.682	0.344	0.098	0.28	0.0096	0.0094	0.50				
N% O	3.979	8.720	5.931	0.670	0.11	0.4493	0.7293	0.38	s	ns	s	s
N% S	1.246	3.178	1.822	0.316	0.17	0.1635	0.0397	0.80				
NUE O	11.466	25.130	17.070	1.913	0.11	3.6605	2.8699	0.56	s	ns	s	s
NUE S	31.463	80.218	56.327	8.607	0.15	88.7668	29.4272	0.75				
NUI O	0.004	0.041	0.018	0.006	0.33	0.00004	0.00007	0.36	s	ns	s	s
NUI S	0.011	0.074	0.036	0.010	0.28	0.0001	0.00004	0.69				
WC O	6.65	13.193	10.159	1.193	0.12	1.6442	2.5268	0.39	s	s	s	s
WC S	3.061	6.064	4.227	0.522	0.12	0.2789	0.2899	0.49				
TLA S	131.5	2055.5	1381.386	297.763	0.22	88663	83978	0.51	s	ns	-	-
RGR S	0.173	0.332	0.234	0.020	0.09	0.0004	0.0006	0.41	s	s	-	-

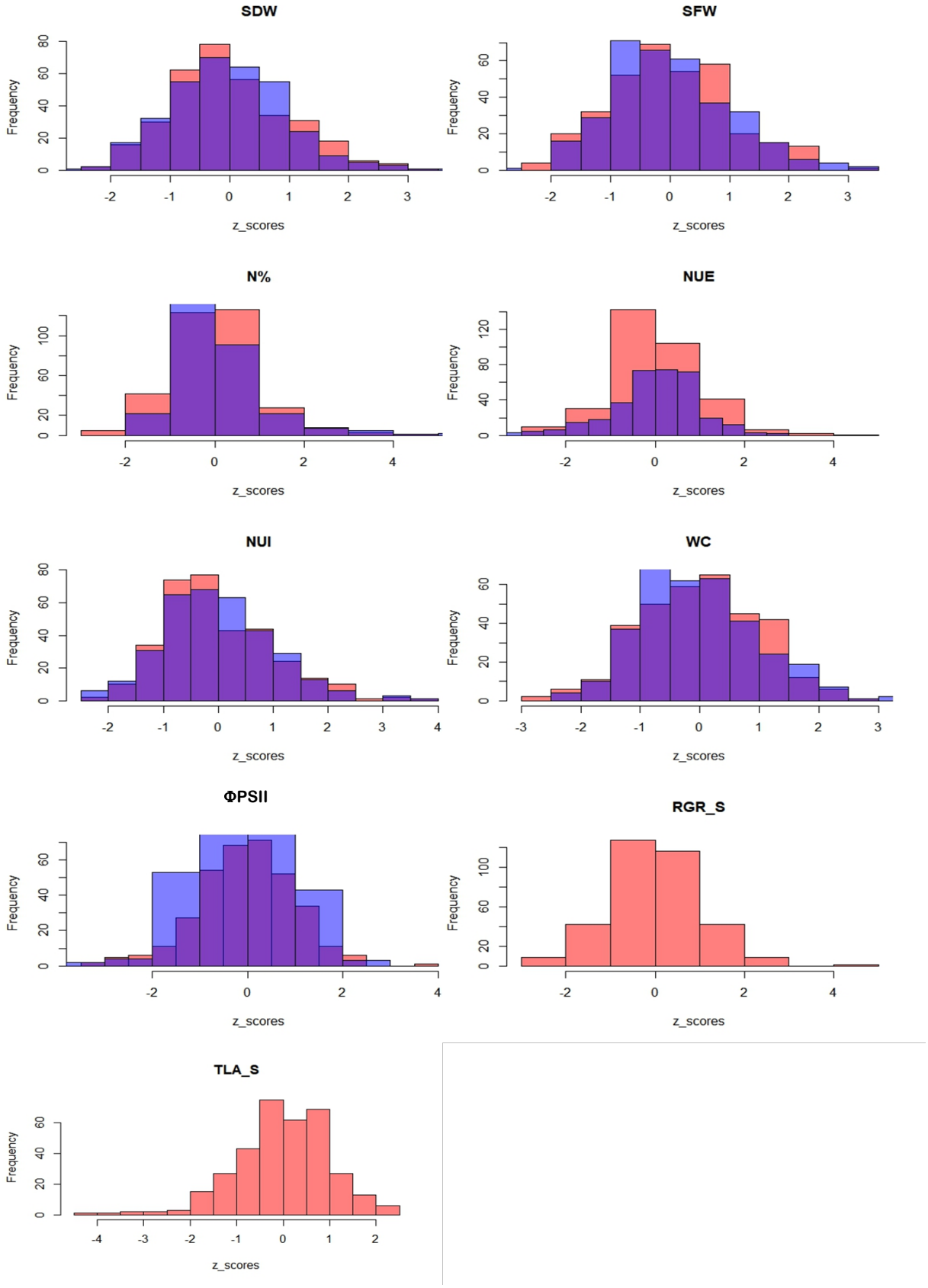


Figure 2: Frequency distributions of z_scores of shoot dry weight (SDW), shoot fresh weight (SFW), nitrogen concentration (N%), nitrogen use efficiency (NUE), nitrogen usage index (NUI), water content (WC), Photosystem II efficiency (ΦPSII) under sub-optimal (red) and optimal (blue) N conditions; relative growth rate (RGR) and total leaf area (TLA) under sub-optimal (S) N condition.

A Pearson correlation matrix showed the relation between traits and treatments (Table 2). Results of the sub-optimal and optimal N treatments were significantly correlated in each trait (Table 2, Figure 3). In addition, SFW, SDW, NUI, WC, and TLA were significantly correlated to each other under both treatments (Table 2). However, N% was negatively correlated with NUI and NUE. Likewise, WC and N% were positively correlated; whereas the correlation between WC and NUE was negative (Table 2). Finally, RGR, only observed under the sub-optimal N supply, was negatively correlated with SFW, N%, ΦPSII, WC, and TLA, but positively correlated with NUE.

Table 2: Pearson correlations (2-tailed) of shoot dry weight (SDW), shoot fresh weight (SFW), nitrogen concentration (N%), nitrogen use efficiency (NUE), nitrogen usage index (NUI), water content (WC), Photosystem II efficiency (ΦPSII), relative growth rate (RGR), and total leaf area (TLA) under sub-optimal nitrogen (S) and optimal (O) conditions. NS means not significant.

* $p \leq 0.01$, * $p \leq 0.05$.

	SDW S	SDW O	SFW S	SFW O	N% S	N% O	NUE S	NUE O	NUI S	NUI O	WC S	WC O	PSII S	PSII O	TLA S	RGR S
SDW S	1															
SDW O	.660**	1														
SFW S	.933**	.563**	1													
SFW O	.677**	.935**	.655**	1												
N% S	NS	NS	.143**	NS	1											
N% O	-.123*	-.258**	NS	NS	.333**	1										
NUE S	NS	.141*	-.171**	NS	-.969**	-.343**	1									
NUE O	NS	.264**	NS	NS	-.348**	-.982**	.365**	1								
NUI S	.841**	.636**	.699**	.602**	-.496**	-.314**	.500**	.307**	1							
NUI O	.600**	.947**	.477**	.821**	-.209**	-.529**	.249**	.546**	.652**	1						
WC S	.191**	NS	.523**	.196**	.333**	.303**	-.378**	-.308**	NS	NS	1					
WC O	.144**	NS	.339**	.294**	.174**	.419**	-.205**	-.453**	NS	-.212**	.600**	1				
PSII S	NS	NS	.125*	NS	NS	NS	NS	NS	NS	NS	NS	.162**	1			
PSII O	NS	-.154**	NS	-.130*	NS	NS	NS	NS	NS	-.114*	NS	NS	.344**	1		
TLA S	.453**	.494**	.448**	.522**	NS	NS	NS	NS	.382**	.449**	.165**	.159**	.191**	NS	1	
RGR S	NS	NS	-.174**	NS	-.187**	NS	.209**	NS	NS	NS	-.243**	-.153**	-.267**	NS	-.525**	1

**. Correlation is significant at the 0.01 level (2-tailed). *. Correlation is significant at the 0.05 level (2-tailed).

Phenotypic plasticity

In order to understand the adaptive responses to the change of N supply from sub-optimal to optimal, residual analysis was conducted for all genotypes. The residual analysis was the deviation of the observed values from the estimated results. This estimated adaptation response was calculated using the fitted regression modelling. If the difference between estimated and observed values is very small, then the residual should be close to 0 and the squared regression coefficient (R^2) should be close to 1. However, as shown on Figure 3, genotypes were deviated from the regression line and this pointed out the effect of phenotypic plasticity. N%, NUE and ΦPSII showed a high deviation from the regression line compared to SDW, SFW, WC, and NUI (Figure 3). This might explain that genotype by environment interaction greatly affected N%, NUE and ΦPSII under two different N supplies.

The positive correlation between WC_Res, N%_Res, SFW_Res and SDW_Res and the negative correlation between WC_Res and NUE_Res (Table 3) corresponded well with the correlations observed between the raw trait values (Table 2) as it could be expected from their interdependency. Frequency distributions showed the phenotypic variation of phenotypic plasticity in the population (Figure 4). The phenotypic variation in N%_Res, NUE_Res and ΦPSII_Res was not as high as in SDW_Res, SFW_Res, NUI_Res, and WC_Res. This wide range of variation in SDW_Res, SFW_Res, WC_Res and NUI_Res might be controlled by the genetic variation in the HapMap population. It was, therefore, possible to identify of candidate genes associated with phenotypic plasticity of *A. thaliana* under changing N supply conditions.

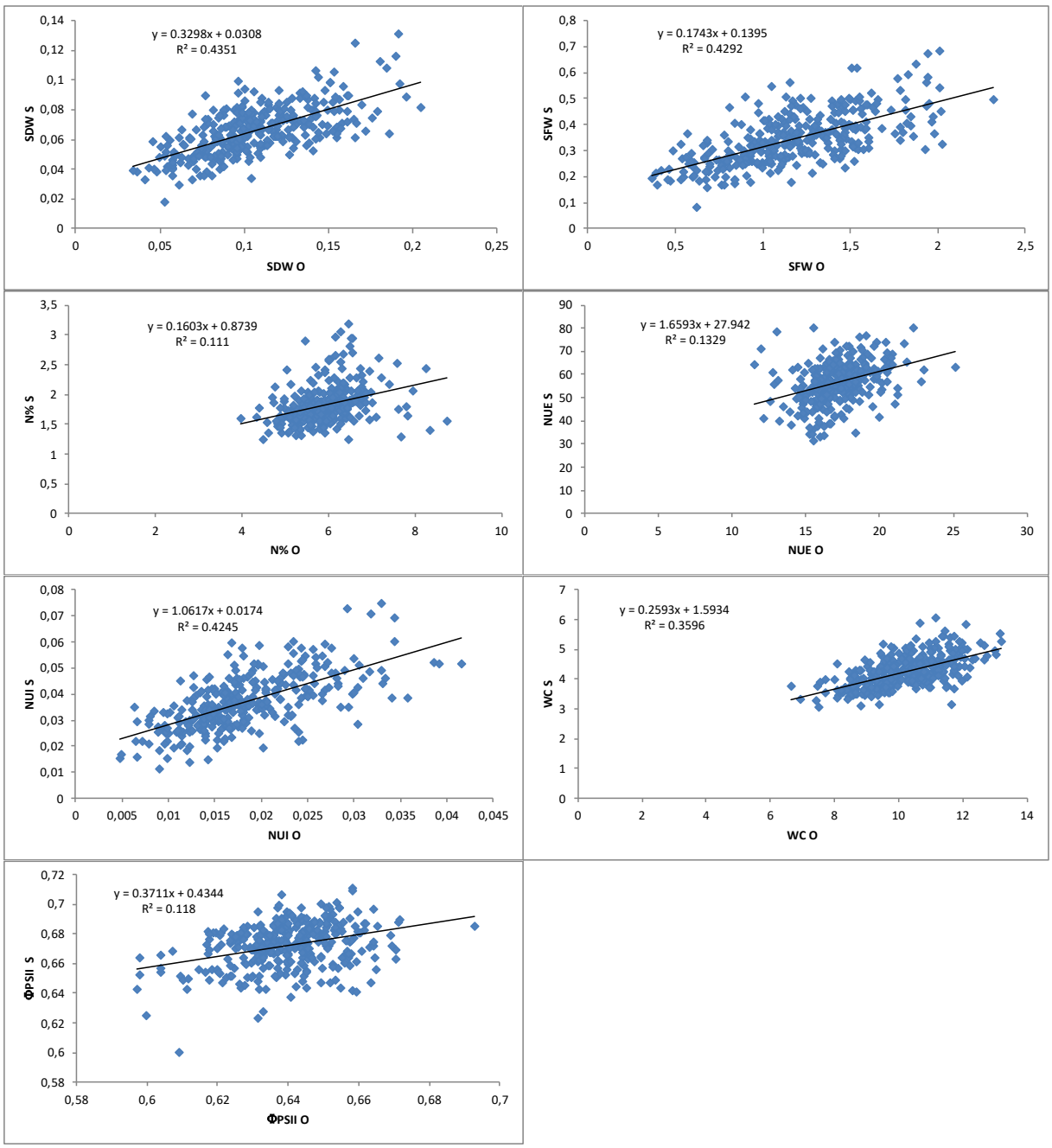


Figure 3: Correlation plots with scattered dots and regression lines comparing data from sub-optimal (S) and optimal (O) N supplies for shoot dry weight (SDW), shoot fresh weight (SFW), shoot N concentration (N%), N use efficiency (NUE), N usage index (NUI), water content (WC), Photosystem II efficiency (ΦPSII).

Table 3: Pearson correlation analysis (2-tailed) of residuals (Res) of shoot dry weight (SDW), shoot fresh weight (SFW), shoot nitrogen concentration (N%), nitrogen use efficiency (NUE), nitrogen usage index (NUI), water content (WC), and Photosystem II efficiency (ΦPSII). NS means not significant.

	SDW Res	SFW Res	N% Res	NUE Res	NUI Res	WC Res	PSII Res
SDW Res	1						
SFW Res	.923**	1					
N% Res	.171**	.214**	1				
NUE Res	NS	-.232**	-.645**	1			
NUI Res	.770**	.604**	-.366**	.542**	1		
WC Res	.243**	.497**	.383**	-.114*	NS	1	
PSII Res	.151**	.117*	.109*	NS	NS	NS	1

**. Correlation is significant at the 0.01 level (2-tailed). *. Correlation is significant at the 0.05 level (2-tailed).

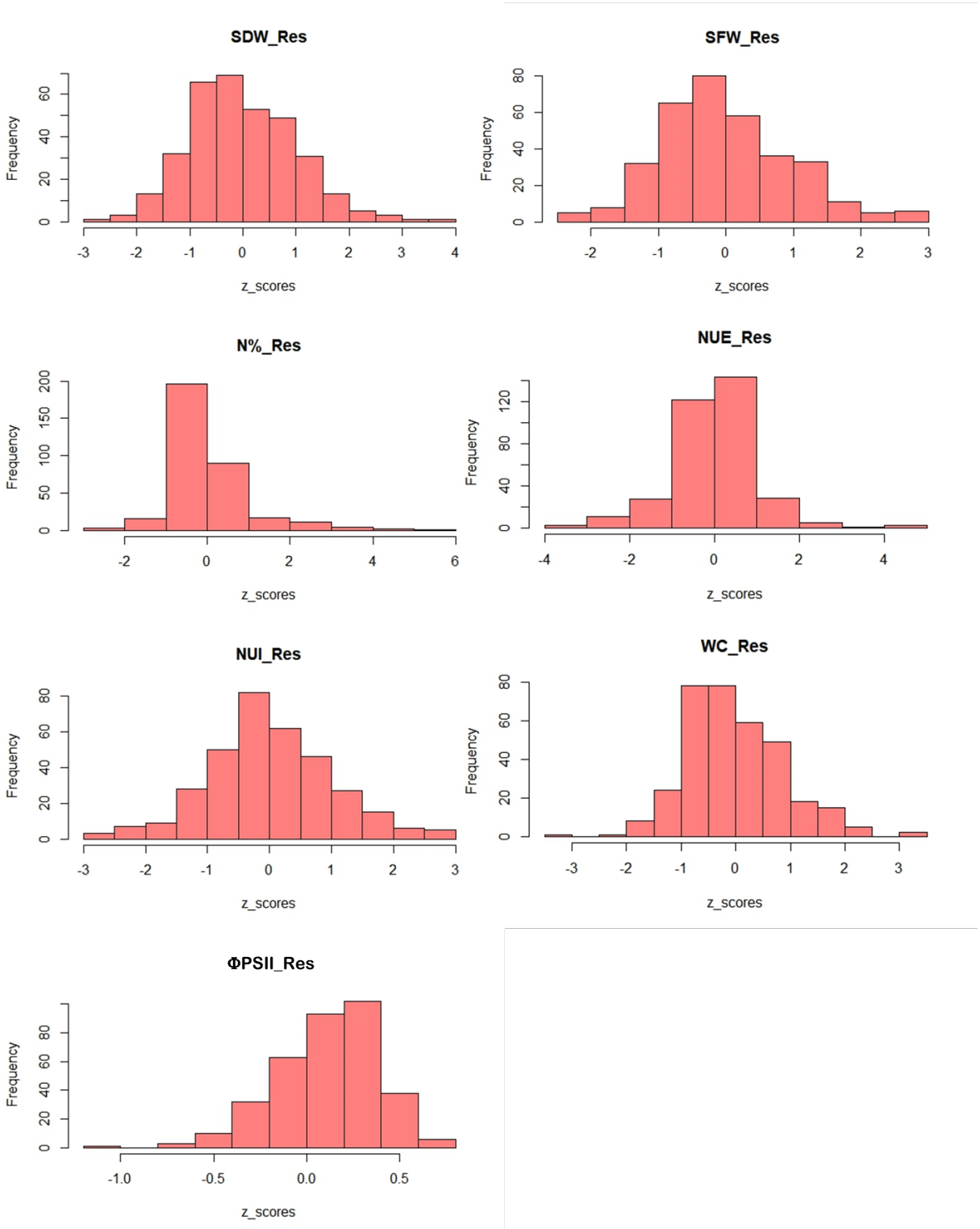


Figure 4: Frequency distributions of z_scores of residuals (Res) of shoot dry weight (SDW), shoot fresh weight (SFW), shoot nitrogen concentration (N%), nitrogen use efficiency (NUE), nitrogen usage index (NUI), water content (WC), and Photosystem II (ΦPSII).

Mapping quantitative traits and identifying associated candidate genes

GWAS was used to find causal genetic variants associated with traits of interest. GWAS is based on the LD mapping. LD decays in approximately 10 kb in *A. thaliana* (Kim et al. 2007). Candidate QTLs were selected based on the appearance of multiple SNPs, having a -log10(*p*-

value)≥4. GWAS detected 28 QTLs associated with the afformentioned NUE and NRAT, including phenotypic plasticity (Figure 5). GWAS detected four QTLs for SDW; one QTL for ΦPSII, two QTL for N% and NUE; four QTLs for NUI; three QTLs for WC, three QTLs for TLA and finally one QTL for RGR above the arbitrary threshold (-log10(*p*-value)≥4) (Figure 5 and Table 4).

QTLs associated with SDW_Res, N%_Res, NUE_Res, NUI_Res, ΦPSII _Res, WC_Res were mapped with GWAS (Figure 6). GWAS detected three QTLs for SDW_Res; three QTLs for NUI_Res; four QTLs for WC_Res; three QTLs for NUE_Res; two QTLs for N%_Res; and four QTLs for ΦPSII_Res above the arbitrary threshold (-log10(*p*-value)≥4).

Candidate genes were selected within 10 kb around the most significant marker of the candidate locus (Table 4 and 5). SDW, SFW, NUI, TLA, and NUI_Res shared a QTL on chromosome 5 (Figure 5a, b, c, d and Figure 6b; Table 4 and 5). The explained genetic component of phenotypic variation of this locus varied between 9% and 36%. This locus was co-locating with AT5G13550 gene. SDW and SDW_Res shared a QTL on chromosome 1. The explained genetic component of phenotypic variations were as 6% and 21% respectively. This locus was co-localized with AT1G22180 gene. NUI and SDW_Res shared a QTL on chromosome 5 with the explained genetic component of phenotypic variations as 10% and 27% respectively. This locus was co-localized with AT5G13380 gene. SFW and SDW shared a QTL on chromosome 5 with 7% of the explained genetic component of phenotypic variation. This locus was co-located with AT5G10290 gene. N% and NUE shared two QTLs with the explained genetic variations as 12%/44% and 13%/42% respectively. These loci were co-located withAT2G28800 and AT5G27870 genes. Two loci associated with N% (44% and 42%) and one locus associated with WC_Res (49.86%) showed the highest explained genetic component of phenotypic variation, where they might provide a major effect for these traits.

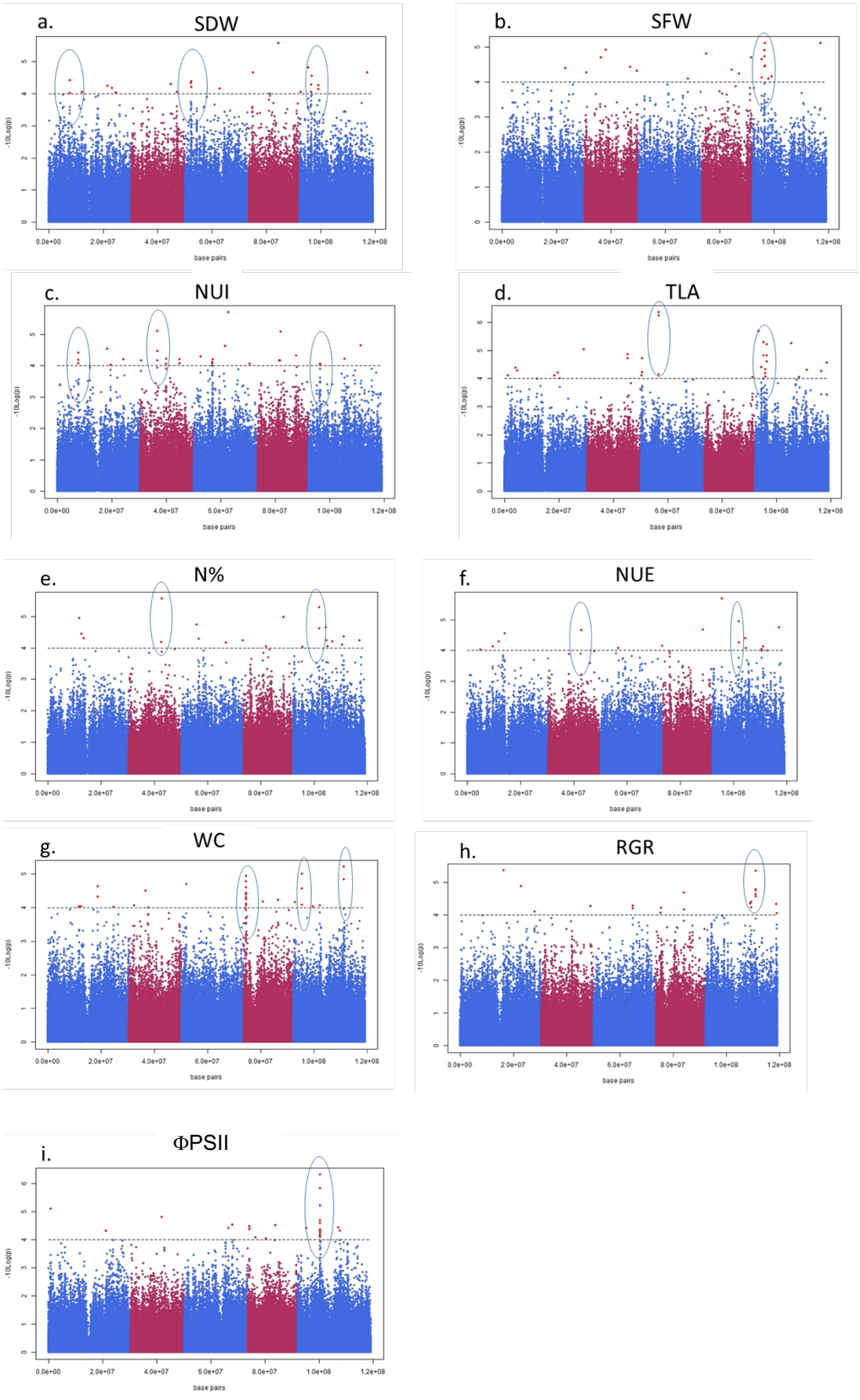


Figure 5: Manhattan plots of shoot dry weight (SDW) (a), shoot fresh weight (SFW) (b), nitrogen usage index (NUI) (c), total leaf area (TLA) (d), shoot nitrogen concentration (N%)

(e), nitrogen use efficiency (NUE) (f), water content (WC) (g), relative growth rate (RGR) (h), and Photosystem II efficiency (ΦPSII) (i) in the sub-optimal N condition. Every dot indicates a SNP with its corresponding $-\log_{10}(p\text{-value})$. -Potential trait-associated SNPs, displayed as $-\log_{10}(p\text{-value}) \geq 4$, are shown as red dots above the threshold indicated as a dashed line. The alternating blue and purple colours refer to five chromosomes of *A. thaliana*. Ovals show the associated loci containing promising candidate genes.

Table 4: List of gene IDs corresponding with significant SNP markers with a -$\log_{10}$(*p*-value)≥4, associated with shoot dry weight (SDW), total leaf area (TLA), water content (WC), nitrogen usage index (NUI), nitrogen use efficiency (NUE), shoot nitrogen concentration (N%), Photosystem II efficiency (ΦPSII), and relative growth rate (RGR) under sub-optimal N supply. Indicated are chromosome numbers (Chr.), SNP positions as basepairs (bp), estimated marker effect (β), and % explained genetic component of phenotypic variation (%genetic var.), for every significant SNP marker. "LD" shows additional candidates within linkage disequilibrium (10 kb) of the significant marker. Common loci found in more than one trait are highlighted in green.

	Chr.	Gene ID	-log10(p-value)	SNP position (bp)	β	%genetic var.	LD
SDW	1	AT1G22150					within 10 kb
		AT1G22160					within 10 kb
		AT1G22170					within 10 kb
		AT1G22180	4.41	7831029	-0.007	6	
		AT1G22190					within 10 kb
		AT1G22200					within 10 kb
	3	AT3G07060					within 10 kb
		AT3G07070					within 10 kb
		AT3G07080	4.21	2241776	0.012	7	
		AT3G07090					within 10 kb
		AT3G07100					within 10 kb
	5	AT5G10280					within 10 kb
		AT5G10290	4.81	3235839	0.014	7	
		AT5G10300					within 10 kb
		AT5G10310					within 10 kb
		AT5G10320					within 10 kb
		AT5G10330					within 10 kb
		AT5G13530					within 10 kb
		AT5G13550	4.28	4356589	-0.008	9	
		AT5G13560					within 10 kb
		AT5G13565					within 10 kb
		AT5G13570					within 10 kb
SFW	5	AT5G10280					within 10 kb
		AT5G10290	4.65	3235839	0.082	7	
		AT5G10300					within 10 kb
		AT5G10310					within 10 kb
		AT5G10320					within 10 kb
		AT5G10330					within 10 kb
		AT5G13530					within 10 kb
		AT5G13550	4.90	4355489	-0.055	10	
		AT5G13560					within 10 kb
		AT5G13565					within 10 kb
		AT5G13570					within 10 kb
PSII	5	AT5G24310					within 10 kb
		AT5G24313					within 10 kb
		AT5G24314					within 10 kb
		AT5G24316					within 10 kb
		AT5G24318					within 10 kb
		AT5G24320	6.31	8285674	-0.016	10	
		AT5G24330					within 10 kb
		AT5G24340					within 10 kb

NUE & N%	2	AT2G28790					within 10 kb
		AT2G28800	4.66/5.56	12354308	-4.5/0.18	12/44	
		AT2G28810					within 10 kb
		AT2G28815					within 10 kb
		AT2G28830					within 10 kb
	5	AT5G27850					within 10 kb
		AT5G27860					within 10 kb
		AT5G27870	4.95/5.29	9880146	-4.66/0.17	13/42	
		AT5G27880					within 10 kb
NUI	1	AT1G22150					within 10 kb
		AT1G22160					within 10 kb
		AT1G22170	4.17	7827370	0.005	8	
		AT1G22180					within 10 kb
		AT1G22190					within 10 kb
		AT1G22200					within 10 kb
	2	AT2G14800					within 10 kb
		AT2G14810					within 10 kb
		AT2G14820					within 10 kb
		AT2G14825					within 10 kb
		AT2G14830					within 10 kb
		AT2G14835	5.1	6369482	0.006	10	
	5	AT5G13350					within 10 kb
		AT5G13360					within 10 kb
		AT5G13370					within 10 kb
		AT5G13380	4.04	4291018	-0.005	10	
		AT5G13390					within 10 kb
		AT5G13400					within 10 kb
		AT5G13530					within 10 kb
		AT5G13550	4.06	4356589	-0.005	11	
		AT5G13560					within 10 kb
		AT5G13565					within 10 kb
		AT5G13570					within 10 kb
WC	4	AT4G02010					within 10 kb
		AT4G02020					within 10 kb
		AT4G02030	4.94	891994	0.5	11	
		AT4G02040					within 10 kb
		AT4G02050					within 10 kb
		AT4G02060					within 10 kb
		AT4G02070					within 10 kb
	5	AT5G10710					within 10 kb
		AT5G10720	4.57	3390523	0.27	11	
		AT5G10730	5.01	3391648	0.27	12	
		AT5G10740					within 10 kb
		AT5G10745					within 10 kb
		AT5G10750					within 10 kb
	5	AT5G46825					within 10 kb
		AT5G46840					within 10 kb
		AT5G46845	4.85	19010585	0.46	10	
		AT5G46850					within 10 kb
		AT5G46860	5.21	19012414	0.55	11	
		AT5G46870					within 10 kb
		AT5G46871					within 10 kb
		AT5G46873					within 10 kb
TLA	3	AT3G19280					within 10 kb
		AT3G19290					within 10 kb
		AT3G19300					within 10 kb
		AT3G19310	6.34	6694727	-311.48	17	
		AT3G19320					within 10 kb
		AT3G19330					within 10 kb
		AT3G19340					within 10 kb
		AT3G19350					within 10 kb
	5	AT5G10140					within 10 kb
		AT5G10150	4.81	3184569	-192.41	14	
		AT5G10160					within 10 kb
		AT5G10170	5.27	3188327	-200.22	16	
		AT5G10180					within 10 kb
		AT5G10190					within 10 kb
		AT5G13530					within 10 kb
		AT5G13550	5.21	4355489	-172.77	18	
		AT5G13560					within 10 kb
		AT5G13565					within 10 kb
		AT5G13570					within 10 kb
RGR	5	AT5G46260					within 10 kb
		AT5G46270					within 10 kb
		AT5G46280					within 10 kb
		AT5G46290	5.35	18776780	-0.01	11	
		AT5G46295					within 10 kb
		AT5G46297					within 10 kb
		AT5G46230					within 10 kb

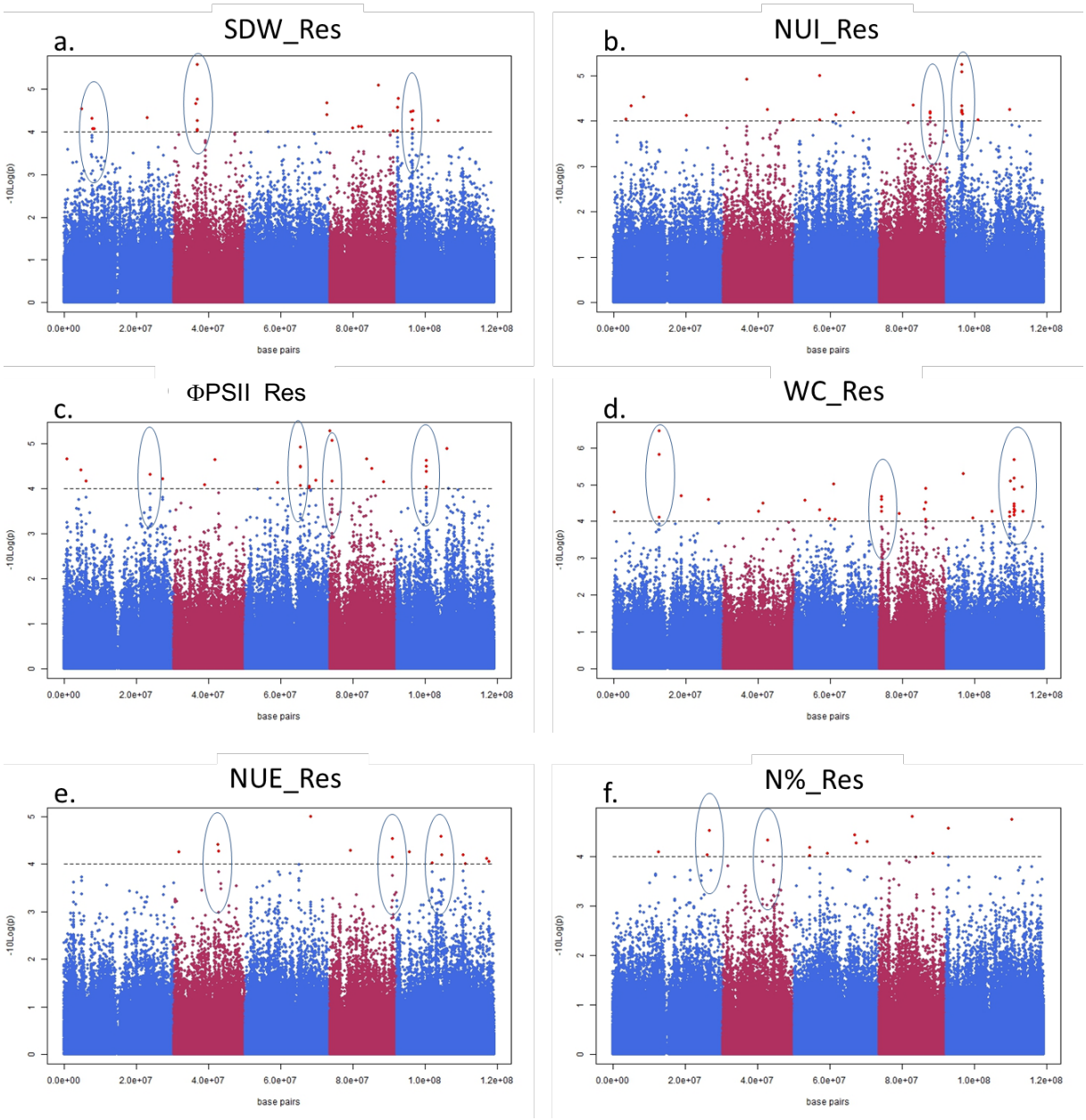


Figure 6: Manhattan plots of GWAS of residuals of shoot dry weight (SDW_Res), N usage index (NUI_Res), Photosystem II efficiency (ΦPSII_Res), water content (WC_Res), N use efficiency (NUE_Res), and shoot N concentration (N%_Res). Significant SNPs, displayed as $-\log_{10}(p\text{-value})$, are shown with red dots above the threshold, which was arbitrarily set at 4. Blue and red areas refer to the 5 chromosomes of *A. thaliana*. Ovals show the associated loci corresponding to promising candidate genes.

Table 5: Chromosome numbers (Chr.) and gene IDs consisting significant SNP marker above the threshold (-$\log_{10}$(*p*-value)≥4) value associated with residuals (Res) of shoot dry weight (SDW), water content (WC), nitrogen usage index (NUI), nitrogen use efficiency (NUE), shoot nitrogen concentration (N%), and Photosystem II efficiency (ΦPSII). SNP positions as base pairs (bp), estimated marker effects (β), and % explainedgenetic component of phenotypic variation (%genetic var.) are displayed for every significant SNP marker and LD shows additional candidates within 10 kb of significant marker. Common loci were highlighted in green colour.

	Chr.	Gene ID	-log10(p-value)	SNP position (bp)	β	%genetic var.	LD
NUE_Res	2	AT2G28690					within 10 kb
		AT2G28700	4.41	12317854	-5.11	33.79	
		AT2G28710					within 10 kb
		AT2G28720					within 10 kb
	4	AT4G36960					within 10 kb
		AT4G36970					within 10 kb
		AT4G36980	4.53	17436233	7.04	32.55	
		AT4G37000					within 10 kb
		AT4G37010					within 10 kb
	5	AT5G33290					within 10 kb
		AT5G33300	4.2	12568616	-6.5	32.93	
N%_Res	1	AT1G71100					within 10 kb
		AT1G71110					within 10 kb
		AT1G71120					within 10 kb
		AT1G71130					within 10 kb
		AT1G71140	4.52	26825107	0.28	35.75	
		AT1G71150					within 10 kb
		AT1G71160					within 10 kb
		AT1G71170					within 10 kb
		AT1G71180					within 10 kb
		AT1G71190					within 10 kb
	2	AT2G28790	4.34	12354308	0.15	34.15	
		AT2G28800					within 10 kb
		AT2G28810					within 10 kb
NUI_Res	4	AT4G28670					within 10 kb
		AT4G28680					within 10 kb
		AT4G28690					within 10 kb
		AT4G28700	4.07	14161592	0.003	34.7	
		AT4G28703					within 10 kb
		AT4G28706					within 10 kb
	5	AT5G13350					within 10 kb
		AT5G13360					within 10 kb
		AT5G13370					within 10 kb
		AT5G13380	5.23	4290632	-0.004	44.9	
		AT5G13390					within 10 kb
		AT5G13400					within 10 kb
		AT5G13530					within 10 kb
		AT5G13550	4.18	4356589	0.003	36.4	
		AT5G13560					within 10 kb
		AT5G13565					within 10 kb
		AT5G13570					within 10 kb

		AT1G34400					within 10 kb
	1	AT1G34403					within 10 kb
		AT1G34405	6.44	12571916	0.55	49.86	
		AT1G34410					within 10 kb
		AT4G01720					within 10 kb
		AT4G01730	4.68	751568	0.33	38.42	
	4	AT4G01735					within 10 kb
		AT4G01740					within 10 kb
		AT4G01750					within 10 kb
		AT4G01760					within 10 kb
WC_Res		AT5G44010					within 10 kb
		AT5G44020					within 10 kb
		AT5G44030	5.09	17719004	0.43	38.14	
		AT5G44040					within 10 kb
		AT5G44050					within 10 kb
	5	AT5G46360					within 10 kb
		AT5G46370					within 10 kb
		AT5G46380					within 10 kb
		AT5G46390	5.17	18817297	0.44	38.39	
		AT5G46400	5.68	18823192	0.49	42.66	
		AT5G46410					within 10 kb
		AT5G46420					within 10 kb
		AT1G22150					within 10 kb
		AT1G22160					within 10 kb
	1	AT1G22170					within 10 kb
		AT1G22180	4.31	7831463	0.006	21.85	
		AT1G22190					within 10 kb
		AT1G22200					within 10 kb
		AT2G15220					within 10 kb
SDW_Res	2	AT2G15230	5.57	6614233	-0.006	28.57	
		AT2G15240					within 10 kb
		AT5G13350					within 10 kb
		AT5G13360					within 10 kb
	5	AT5G13370					within 10 kb
		AT5G13380	4.28	4291018	0.006	27.28	
		AT5G13390					within 10 kb
		AT5G13400					within 10 kb
		AT1G64295					within 10 kb
		AT1G64300					within 10 kb
	1	AT1G64310					within 10 kb
		AT1G64320	4.31	23868751	-0.01	7.56	
		AT1G64330					within 10 kb
		AT1G64340					within 10 kb
	3	AT3G43444	4.49	15378752	0.006	7.62	
		AT4G01800					within 10 kb
		AT4G01810					within 10 kb
PSII_Res	4	AT4G01820	5.06	781252	0.007	10.23	
		AT4G01830					within 10 kb
		AT4G01840					within 10 kb
		AT5G24310					within 10 kb
		AT5G24313					within 10 kb
		AT5G24314					within 10 kb
		AT5G24316					within 10 kb
		AT5G24318					within 10 kb
	5	AT5G24320	4.49	8285955	-0.009	8.06	
		AT5G24330					within 10 kb

Discussion

Shoot phenotyping and phenotypic plasticity

N is a macronutrient required for plant growth and development. Plants develop different adaptation mechanisms to cope with limited N supply. In this chapter, I phenotyped shoot traits in *A. thaliana* related to growth, photosynthesis, N uptake and utilization. I also checked for phenotypic plasticity to understand the responsiveness of the population. Symptoms of sub-

optimal N supply were detected as the reduction in shoot biomass and leaf yellowing in leaves. ΦPSII, TLA, and RGR was measured during the course of plant growth. Harvested plants were measured for their SDW, SFW, N%, NUE, NUI, and WC. Correlation matrix showed that growth related traits like SDW, SFW, and TLA were found to be positively correlating as a matter of course. However, shoot phenotyping revealed interesting relations between traits. Phenotypic analysis of HapMap population showed positive correlation between N% and WC under sub-optimal and optimal conditions (Table 2). Indeed, N% and WC plasticity were positively correlated in the population, which explained the parallel response to the changing N supply (Table 3). Water and N are two important determinants of photosynthesis and therefore plant growth, which infers that shoot N% changes simultaneously with shoot WC (Easlon and Bloom 2013; Loudet et al. 2003a). N supply reduction triggered differed N% and WC for diverse genotypes, which might explain a differing communication system between root and shoot in order to perform an adaptive response to N limitation. A differing root/shoot communication might account for a differing root architecture among diverse genotypes. Further root analyses would benefit to understand this communication between root and shoot.

Phenotypic plasticity of SDW, SFW, N%, NUE, NUI, and WC largely varied in the population. Only ΦPSII plasticity did not show a large phenotypic variation. It is important to note that HapMap population constitutes genotypes with varying N responsiveness to changing N supply. Natural variation in the responsiveness to changing N supply was previously studied (Chardon et al. 2010; Meyer et al. 2019; Menz et al. 2018; North et al. 2009; De Pessemier et al. 2013). Since their application procedures considerably differ from each other, they recorded varying responses to changing N supply, which explains the strong effect of genotype by environment interaction. They compared either severe N stress (0.4 mM, 0.2 mM, or 0.01 mM) to optimal N condition (9 mM, 4 mM, or 10 mM respectively) (Menz et al. 2018; Meyer et al. 2019; De Pessemier et al. 2013) or mild N stress (2 mM or 1 mM) to optimal N condition (10 mM or 4 mM, respectively) (Chardon et al. 2010; North et al. 2009). They applied either long (16 hours light) (Meyer et al. 2019; North et al. 2009; De Pessemier et al. 2013) or short days (8 and 10 hours light) (Menz et al. 2018; Chardon et al. 2010), high (160 or 200 μM m^{-2} s^{-1}) (Chardon et al. 2010; North et al. 2009; Menz et al. 2018) or low (75 or 100 μM m^{-2} s^{-1}) (Meyer et al. 2019; De Pessemier et al. 2013) light intensities. They also differed in types of N supplies. Some studies applied nitrate as a sole N source (Chardon et al. 2010; Meyer et al. 2019; North et al. 2009; De Pessemier et al. 2013), whereas some applied ammonium nitrate (Menz et al. 2018; Meyer et al. 2019) to find the plasticity of genotypes. Moreover, sand (Chardon et al. 2010), agar (North et al. 2009; Meyer et al. 2019; De Pessemier et al. 2013), and hydroponics

(Menz et al. 2018) were three different growth media where plants were tested in these studies. The variability of experiments greatly affects the ranking of genotypes for a certain trait under every unique environmental condition. Therefore, the comparability of environmental conditions is very important to evaluate the phenotypic plasticity of traits.

QTL detection through GWAS

This research presented the first QTL mapping analysis in NUE, NRAT, and phenotypic plasticity in HapMap population of *A. thaliana* and revealed 28 QTLs in total. HapMap population is known as the collection of diverse natural ecotypes (Platt et al. 2010). After the elimination of population structure, a core set of 354 genetically diverse *A. thaliana* accessions was also utilized in other QTL mapping researches (Kooke et al. 2016; van Rooijen et al. 2017; Li et al. 2010; El-Soda et al. 2015). The results showed a large phenotypic variation for NUE, NRAT, and phenotypic plasticity controlled by varying genetic mechanisms in the population (Meyer et al. 2019). Indeed, high broad sense heritabilities estimated that the majority of phenotypic variation was supposedly explained by many genetic variants. However, the number of SNPs surpassing the threshold was not in agreement with high heritabilities of NUE and NRAT (Table 1, Table 4, Figure 5). This discrepancy between high heritability and few detected SNPs could be due to (I) rare alleles, (II) allelic heterogeneity, (III) single marker approach, (IV) epistatic interactions, (V) epigenetics, (VI) genotype by environment interaction, and (VII) complex traits controlled by small effect QTLs (Brachi et al. 2011; Kooke et al. 2016). Although, some confounding factors were eliminated from GWAS (the population structure and rare alleles (minor allele frequency (0.05))), the missing heritability was still observed. The reason might be the complexity of these traits, which are controlled by several small effect QTLs.

It was well established that NUE and NRAT are complex traits controlled by many small effect QTLs (Han et al. 2015). For instance, Loudet et al. (2003a and 2003b) and Diaz et al. (2006) conducted QTL analysis to dissect the genetic architecture of NUE in *A. thaliana* (Loudet et al. 2003a; Loudet et al. 2003b; Diaz et al. 2006). They identified many small effect QTLs in bi-parental RIL population. Likewise, GWAS revealed many small effect QTLs in NUE and NRAT (Figure 5 and Table 4). Further analysis could be done using multiple parental RIL populations to increase the power of QTL detection since GWAS might not represent the whole picture of genetic mechanism underlying the phenotypic variation of the traits.

QTLs and candidate gene prioritization

GWAS identified 28 candidate loci. In order to confirm their association with NUE, NRAT, and phenotypic plasticity, they were first of all checked to find if they were co-located with previously identified QTLs in literature. Rauh et al. (2002), Loudet et al. (2003a and 2003b), and Diaz et al. (2006), published four different QTL mapping results, where Loudet et al. (2003a and 2003b) and Diaz et al. (2006) used the same RIL population and the same growth conditions (Rauh et al. 2002; Loudet et al. 2003a; Loudet et al. 2003b; Diaz et al. 2006). Loudet et al. (2003a and 2003b) and Diaz et al. (2006) grew Bay-0 and Sha RIL population under mild N stress (3 mM) and used nitrate as a sole N source. Rauh et al. (2006) grew Col-0 and Ler RIL population under very severe N stress (only tap water, no N fertilizer). Loudet et al. (2003b) shared some QTLs, whereas Rauh et al. (2002) did not. The reason why Loudet et al. (2003b) and Diaz et al. (2006) showed co-localizing QTLs was because dry matter, measured under low N by Loudet et al. (2003b), was correlated with leaf yellowing, measured by Diaz et al. (2006) in Bay-0 and Sha RIL population. Since they used exactly the same genetic pool under the same environmental conditions, they apparently identified co-localizing QTLs. On the other hand, Rauh et al. (2006) did not share any common loci with them, since their growth conditions and RIL population were completely different from each other. A similar reasoning would be applied to GWAS results of HapMap population. First of all, GWAS results did not show any co-localization with QTL results of Col-0 and Ler RIL population since HapMap population was subjected to a mild N stress, unlike Col-0 and Ler RIL population. Although, Bay-0 and Sha RIL population was also applied to a mild N stress, GWAS results showed different QTLs associated with NUE, NRAT, and phenotypic plasticity. The genetic variation in HapMap population offers diverse allele combinations, so this diversity controls varying genetic mechanisms under N stress conditions (Meyer et al. 2019). As a result, it was not unexpected to find lack of co-localization between different QTL analysis of NUE and NRAT.

Bonferroni correction method is often used to eliminate false positive SNPs in GWAS (El-Soda et al. 2015; Atwell et al. 2009). As the detection of small effect QTLs was precluded through GWAS, Bonferroni correction offers a highly stringent -log10(*p*-value) to indicate significant SNPs above the threshold. Therefore, an arbitrary threshold was set to indicate significant SNPs above the threshold (El-Soda et al. 2015; Atwell et al. 2009; van Rooijen et al. 2017). Strong peaks (the appearance of multiple significant SNPs near the most significant SNP within the LD) were identified as candidate QTLs (Figure 5 and 6). Thereafter, candidate genes were

selected from these QTLs (Table 4 and 5). They were prioritized regarding their relation with NUE by using evidences in literature and *A. thaliana* gene ontology annotation database. Prioritized ones were later categorized according to their biological functions: plant growth, development, transporter activity, carbon metabolism, and photosynthesis.

Plant growth and development factors under sub-optimal N condition

Reduced growth and retarded development in plants are prevalent responses for sub-optimal N nutrition. Growth and development regulating plant phytohormones play a very crucial role under stress conditions. Cytokinin, auxin, and abscisic acid are three common growth and development regulating hormones also utilized as signalling mediators for N homeostasis along with amino acids, sugars, nitrate and other phytohormones (Kiba et al. 2011). Cytokinin is a regulator of cell proliferation and differentiation in plants. Its expression is highly correlated with N nutrition (Sakakibara et al. 2006; Brenner et al. 2005). GWAS revealed one candidate gene involved in lipid metabolism and a cytokinin regulated pathway associated with RGR: AT5G46290 (3-Ketoacyl-acyl carrier protein synthase 1, *KAS1*) (Figure 5h, Table 4) (Liu et al. 2013). A study about the stay-green trait in broccoli revealed that in vitro cytokinin induction down-regulated KAS1 protein production, while showing a prolonged stay-green phenotype after harvest and dark treatment (Liu et al. 2013). Sub-optimal N supply limited shoot growth and development in *A. thaliana*, so it might down-regulate the production of cytokinin hormone. Consequently, the down-regulation of cytokinin might induce the production of KAS1 protein. The role of auxin is to trigger lateral root growth towards N rich environment (Krouk et al. 2010; Walch-Liu et al. 2006; Mounier et al. 2014). Nitrate transporter genes (*NRT1.1*) mediate auxin trafficking to the root primordia and facilitate auxin accumulation in the root tips (Mounier et al. 2014; Kiba et al. 2011). As a result of GWAS, AT5G13380 (Auxin responsive GH3 family protein), and AT1G34410 (*Auxin responsive 21: ARF21*), which was previously identified as an auxin activator in the lateral roots induced by N (Gifford et al. 2008), were listed as auxin regulated candidate genes associated with NUI (Figure 5c and Table 4) and SDW_Res, NUI_Res, and WC_Res (Figure 6a, b, d and Table 5). AT5G46845 (*MIR160*) is a micro RNA encoding gene and was detected by GWAS of WC under sub-optimal N condition (Figure 5g and Table 4). This gene is induced to produce small RNA fragments for negative regulation of *ARF10/16/17* genes during N starvation, which affects vegetative growth in plants (Liang et al. 2012; Nguyen et al. 2015). *MIR160* was previously identified in *A. thaliana* as a lateral root development factor under N deficiency (Liang et al. 2012). Lateral root growth is

an adaptation response for *A. thaliana* to cope with the limited N supply by scavenging towards nutrient rich environment. Under sub-optimal N condition, auxin and *MIR160* might be up-regulated to adapt to stressful environments.

Absicisic acid is known as a stress induced plant hormone which down-regulates root production (Kiba et al. 2011). Abscisic acid dependent protein producing gene AT5G10720 (Histidine Kinase 5*: AHK5*) was found as a result of GWAS of WC (Figure 5g and Table 4). This gene is involved in root elongation, where its AHK5 protein defective mutants type *ahk5-1, ahk5-2* and *cki2-1* have reduced effects through the control of abscisic acid and ethylene (Iwama et al. 2007). The *AHK5* gene might negatively regulate the inhibitory effect of abscisic acid on root elongation, which could be the case for induced lateral root formation under sub-optimal N condition.

The phenotyping results obviously showed that N limitation evidently affected the shoot growth and development in *A. thaliana* accessions. Our GWAS results would suggest an adaptive fine-tuning in response to the (in)activity of cytokinin, auxin, and abscisic acid mediated proteins controlling plant development under sub-optimal N condition.

Mineral and amino acid transporters and transporter facilitator proteins

Reduced N supply induces the production of high affinity N transporters (Krapp 2015; Forde 2000; Wang et al. 2012; Dechorgnat et al. 2011). Several other proteins are also induced as mineral and amino acid transporters or transporter facilitators along with the activation of N transporter proteins. GWAS of NUE_Res identified the ammonium transporter gene *Ammonium transporter 1;4* (AT4G28700) on the chromosome 4 (Figure 6e and Table 5). Moreover, the QTL on the chromosome 1 was co-located with Sec14p-like phosphatidylinositol transfer family protein encoding gene (AT1G22180) associated with SDW, NUI, and SDW_Res (Figure5a, c Figure6a, Table4 and Table5), and one on the chromosome 5 was co-located with Major facilitator superfamily protein encoding gene (AT5G10190) associated with TLA (Figure5d and Table4). These proteins were referred as nitrate transporter controlling or facilitating proteins (Heyndrickx and Vandepoele 2012; Gaufichon et al. 2010).

Two QTLs on the chromosome 5 were co-located with *SULTR4;1* gene (AT5G13550), associated with SDW, SFW, NUI, TLA, and NUI_Res (Figure5a, b, c, d, Figure6b, Table4 and Table5), and *SULTR2;1* (AT5G10180) gene, associated with TLA (Figure5d and Table4).. SULTR4;1 and SULTR2;1 are two transporter proteins, not only responsible for sulphur transportation but also responsive to N limitation (Zuber et al. 2010; Kataoka et al. 2004). A study conducted by Peng et al. (2007) found that Sec14p-like phosphatidylinositol transfer

family, SULTR2;1, Major facilitator superfamily protein and SULTR4;1 encoding genes were differentially expressed under changing N conditions. *SULTR2;1* was 2-fold down-regulated in the rosette leaves of the *nla* (*nitrogen limitation adaptation*) mutant in 3mM nitrate compared to 10mM nitrate, whereas the expression of *SULTR2;1* was not affected in the wild type Col-0 upon N limitation. On the other hand, *SULTR4;1* was up-regulated in the same mutant grown in 3mM nitrate compared to 10mM nitrate. Sec14p-like phosphatidylinositol transfer family encoding gene and Major facilitator superfamily protein encoding gene were also found to be up-regulated in the *nla* mutant grown in 3mM nitrate compared to 10mM nitrate (Peng et al. 2007). Therefore, GWAS, together with transcriptome analysis, would suggest that Sec14p-like phosphatidylinositol transfer family, Major facilitator superfamily protein, SULTR2;1 and SULTR4;1 proteins function for keeping up ion homeostasis and amino acid mobilization during sub-optimal N supply. However, this study could not show any further evidence to prove this hypothesis. In order to prove the crosstalk between N deficiency and ion balance in *A. thaliana*, a mineral analysis could be suggested for knock-out mutants of the identified genes.

Carbon metabolism and photosynthesis

Carbon metabolism in plants is identified highly N responsive. N limitation reduces photosynthetic activity and induces cell wall modification in plants (Peng et al. 2007). GWAS revealed two genes involved in cell wall modification. AT5G27870 (Plant invertase/pectin methylesterase inhibitor superfamily protein) associated with N%, NUE (Figure 5 e, f, Table 4) and AT5G44030 (*CESA4*) associated with WC_Res (Figure 6d, Table 5) (Bellec et al. 2002; Fernandes et al. 2013; Hermans et al. 2011). Cell wall structure regulating gene, *CESA4* was previously described in the second group of cellulose synthases including *CESA7* and *CESA8* (Ramirez et al. 2009). These genes were found to be involved in the resistance to the soil borne bacterium *Ralstonia solanacearum* and the necrotrophic fungus *Plectosphaerella cucumerina* (Hernandez-Blanco et al. 2007). The second important gene involving in the cell wall structure was Plant invertase/pectin methylesterase inhibitor superfamily protein encoding gene. In a research on the cell wall composition of grapevine (*Vitis vinifera* L.) it was observed that N deficiency decreased the methyl esterification of pectin in the plant in order to protect cell wall integrity (Fernandes et al. 2013). Apparently, cell wall modification involving genes are one of the regulatory factors of a plant growth and development under sub-optimal N condition.

Limited N nutrition causes catabolism of carbon molecules. GWAS revealed AT1G22170 (*Phosphoglycerate mutase*) gene associated with SDW, NUI, and SDW_Res under sub-optimal N condition (Figure5a, c, Figure6a, Table4 and Table5). *Phosphoglycerate mutase* gene is

identified as a glycolytic enzyme involved in energy metabolism (Wang et al. 2003). Up-regulation of this gene was previously detected by a micro-array assay, conducted by Wang et. al. (2003), showing its rapid response to the added nitrate to the nutrient solution. Its expression takes place dominantly in roots (Wang et al. 2003). However, limited N condition down-regulates *Phosphoglycerate mutase* gene (Peng et al. 2007). This might indicate that this gene could be activated for producing energy for up-taken N assimilation (Wang et al. 2004), so sub-optimal N condition might down-regulate it to decrease glycolysis activity.

Production of photosynthetic proteins are also negatively affected by limited N nutrition due to reduced photosynthesis. GWAS of N%, NUE (Figure 5e, f, Table 4) and NUE_Res (Figure6e, Table 5) revealed candidate genes involved in light induced reactions in plants: AT2G28800 (*Albino3*: *ALB3*) and AT4G37000 (*ACD2*). *ALB3* has a role as membrane invertase activity on chloroplast, which integrates light harvesting chlorophyll proteins onto chloroplast membrane (Asakura et al. 2008; Schneider et al. 2014). The knock-out mutant shows deadly phenotypes: reduced growth and leaf yellowing. The *A. thaliana ACD2* gene is induced by light and provides a defence mechanism against biotic stresses (Mach et al. 2001). Over-expression of *ACD2* shows reduced leaf yellowing and delayed cell death. As leaf yellowing is a very distinctive symptom for sub-optimal N condition, *ALB3* and *ACD2* might be down-regulated under sub-optimal N conditions.

This thesis did not show any confirmation of prioritized candidate genes through functional analysis. Further studies would identify the association of these genes with NUE, NRAT, and phenotypic plasticity using knock-out mutants and differential gene expression analysis.

Conclusion

GWAS revealed 28 QTLs and proposed 157 candidate genes. Genes involved in cytokinin and auxin pathways *KAS1, ARF21,* and *MIR160*; in mineral and amino acid transporting systems *SULTR4;1, SULTR2;1,* Sec14p-like phosphatidylinositol transfer family and Major facilitator superfamily proteins; in carbon metabolism Plant invertase/pectin methylesterase inhibitor superfamily protein and *Phosphoglycerate mutase* were previously shown to be responsive to limited N conditions. Although, there are some evidences in literature, these candidate genes still need to be confirmed through further functional analysis.

SDW, SFW, NUE, NUI, WC, RGR, TLA, and Chl Fluo showed small effect QTLs with the explained genetic component of phenotypic variations ranged between 2% to 18%. However, N% showed two loci co-locating with AT2G28800 and AT5G27870 with the explained genetic

component of phenotypic variations 44% and 42%. These candidate genes are involved in carbon metabolism and photosynthesis. Plant invertase/methylesterase inhibitor superfamily protein has previously identified to be responsive to limited N conditions; however, *ALB3* has remained to be a strong candidate that its interaction with changing N supply conditions has not yet been resolved. GWAS of phenotypic plasticity revealed bigger effect QTLs compared to traits of sub-optimal N condition. AT1G34405 showed the highest explained genetic component of phenotypic variation for WC_Res. This candidate gene might be the major QTL controlling phenotypic plasticity of WC, but its interaction with changing N supply conditions has also been unresolved. Therefore, a future experiment on AT2G28800 and AT1G34405 may provide a novel achievement for genetic analysis of NUE.

Moreover, GWAS proposed candidate genes which are expressed specifically in roots: *ARF21, MIR160, AHK5,* and *Phosphoglycerate mutase.* Their involvement in the root system, especially for lateral root elongation, has previously been studied in order to investigate root adaptation to limited N conditions. Therefore, a future experiment on *A. thaliana* root systems would provide a complete phenotypic screening for NUE and NRAT under sub-optimal N condition.

References

Anderson JT, Inouye DW, McKinney AM, Colautti RI, Mitchell-Olds T (2012) Phenotypic plasticity and adaptive evolution contribute to advancing flowering phenology in response to climate change. Proceedings of Royal Society B:-Biological Sciences 279:3843-3852.

Asakura Y, Kikuchi S, Nakai M (2008) Non-identical contributions of two membrane-bound cpSRP components, cpFtsY and Alb3, to thylakoid biogenesis. The Plant journal: For Cell and Molecular Biology 56:1007-1017.

Atwell S, Huang YS, Vilhj´almsson BJ, Willems G, Horton M, Li Y, Meng D, Platt A, Tarone AM, Hu TT, Jiang R, Muliyati NW, Zhang X, Amer MA, Baxter I, Brachi B, Chory J, Dean C, Debieu M, de Meaux J, Ecker JR, Faure N, Kniskern JM, Jones JDG, Michael T, Nemri A, Roux F, Salt DE, Tang C, Todesco M, Traw MB, Weigel D, Marjoram P, Borevitz JO, Bergelson J, Nordborg M (2009) Genome-wide association study of 107 phenotypes in a common set of Arabidopsis thaliana inbred lines. Nature 465:627–631

Bac-Molenaar JA, Granier C, Keurentjes JJ, Vreugdenhil D (2016) Genome-wide association mapping of time-dependent growth responses to moderate drought stress in Arabidopsis. Plant, Cell & Environment 39:88-102.

Baker NR (2008) Chlorophyll fluorescence: a probe of photosynthesis in vivo. Annual Review of Plant Biology 59:89-113.

Baxter I, Brazelton JN, Yu D, Huang YS, Lahner B, Yakubova E, Li Y, Bergelson J, Borevitz JO, Nordborg M, Vitek O, Salt DE (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.

Bellec Y, Harrar Y, Butaeye C, Darnety S, Bellini C, Faure JD (2002) Pasticcino2 is a protein tyrosine phosphatase-like involved in cell proliferation and differentiation in Arabidopsis. The Plant Journal 32:713-722.

Bi YM, Wang RL, Zhu T, Rothstein SJ (2007) Global transcription profiling reveals differential responses to chronic nitrogen stress and putative nitrogen regulatory components in Arabidopsis. BMC Genomics 8:281.

Brachi B, Morris GP, Borevitz JO (2011) Genome-wide association studies in plants: the missing heritability is in the field. Genome Biology 12:232.

Brenner WG, Romanov GA, Kollmer I, Burkle L, Schmulling T (2005) Immediate-early and delayed cytokinin response genes of Arabidopsis thaliana identified by genome-wide expression profiling reveal novel cytokinin-sensitive processes and suggest cytokinin action through transcriptional cascades. The Plant journal: For Cell and Molecular Biology 44:314-333.

Chao DY, Silva A, Baxter I, Huang YS, Nordborg M, Danku J, Lahner B, Yakubova E, Salt DE (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.

Chardon F, Barthelemy 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.

Chu VT, Gottardo R, Raftery AE, Bumgarner RE, Yeung KY (2008) MeV+R: using MeV as a graphical user interface for Bioconductor applications in microarray analysis. Genome Biology 9:R118.

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. Mechanical Development 130:45-53.

Dechorgnat J, Nguyen CT, 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.

Diaz C, Saliba-Colombani V, Loudet O, Belluomo P, Moreau L, Daniel-Vedele F, Morot-Gaudry JF, Masclaux-Daubresse C (2006) Leaf yellowing and anthocyanin accumulation are two genetically independent strategies in response to nitrogen limitation in Arabidopsis thaliana. Plant & Cell Physiology 47:74-83.

Easlon HM, Bloom AJ (2013) The effects of rising atmospheric carbon dioxide on shoot-root nitrogen and water signaling. Frontiers in plant science 4:304.

El-Lithy ME, Clerkx EJ, Ruys GJ, Koornneef M, Vreugdenhil D (2004) Quantitative trait locus analysis of growth-related traits in a new Arabidopsis recombinant inbred population. Plant Physiology 135:444-458.

El-Soda M, Kruijer W, Malosetti M, Koornneef M, Aarts MG (2015) Quantitative trait loci and candidate genes underlying genotype by environment interaction in the response of Arabidopsis thaliana to drought. Plant, Cell & Environment 38:585-599.

El-Soda M, Malosetti M, Zwaan BJ, Koornneef M, Aarts MG (2014) Genotypexenvironment interaction QTL mapping in plants: lessons from Arabidopsis. Trends in Plant Science 19:390-398.

Fernandes JC, Garcia-Angulo P, Goulao LF, Acebes JL, Amancio S (2013) Mineral stress affects the cell wall composition of grapevine (Vitis vinifera L.) callus. Plant Science : an International Journal of Experimental Plant Biology 205-206:111-120.

Forde BG (2000) Nitrate transporters in plants: structure, function and regulation. Biochimica et Biophysica Acta 1465:219-235.

Garnett T, Conn V, Kaiser BN (2009) Root based approaches to improving nitrogen use efficiency in plants. Plant, Cell & Environment 32:1272-1283.

Gaufichon L, Reisdorf-Cren M, Rothstein SJ, Chardon F, Suzuki A (2010) Biological functions of asparagine synthetase in plants. Plant Science 179:141-153.

Gifford ML, Dean A, Gutierrez RA, Coruzzi GM, Birnbaum KD (2008) Cell-specific nitrogen responses mediate developmental plasticity. Proceedings of the National Academy of Sciences of the United States of America 105:803-808.

Good AG, Shrawat AK, Muench DG (2004) Can less yield more? Is reducing nutrient input into the environment compatible with maintaining crop production? Trends in Plant Science 9:597-605.

Gruber N, Galloway JN (2008) An Earth-system perspective of the global nitrogen cycle. Nature 451:293-296.

Han M, Okamoto M, Beatty PH, Rothstein SJ, Good AG (2015) The Genetics of Nitrogen Use Efficiency in Crop Plants. Annual Review of Genetics 49:269-289.

Hermans C, Porco S, Vandenbussche F, Gille S, De Pessemier J, Van Der Straeten D, Verbruggen N, Bush DR (2011) Dissecting the Role of CHITINASE-LIKE1 in Nitrate-Dependent Changes in Root Architecture. Plant Physiology 157 (3):1313-1326. doi:10.1104/pp.111.181461

Hernandez-Blanco C, Feng DX, Hu J, Sanchez-Vallet A, Deslandes L, Llorente F, Berrocal-Lobo M, Keller H, Barlet X, Sanchez-Rodriguez C, Anderson LK, Somerville S, Marco Y, Molina A (2007) Impairment of cellulose synthases required for Arabidopsis secondary cell wall formation enhances disease resistance. The Plant Cell 19:890-903.

Heyndrickx KS, Vandepoele K (2012) Systematic identification of functional plant modules through the integration of complementary data sources. Plant Physiology 159:884-901.

Iwama A, Yamashino T, Tanaka Y, Sakakibara H, Kakimoto T, Sato S, Kato T, Tabata S, Nagatani A, Mizuno T (2007) AHK5 histidine kinase regulates root elongation through

an ETR1-dependent abscisic acid and ethylene signaling pathway in Arabidopsis thaliana. Plant & Cell Physiology 48:375-380.
Kang HM, Sul JH, Service SK, Zaitlen NA, Kong SY, Freimer NB, Sabatti C, Eskin E (2010) Variance component model to account for sample structure in genome-wide association studies. Nature Genetics 42:348-354.
Kataoka T, Watanabe-Takahashi A, Hayashi N, Ohnishi M, Mimura T, Buchner P, Hawkesford MJ, Yamaya T, Takahashi H (2004) Vacuolar sulfate transporters are essential determinants controlling internal distribution of sulfate in Arabidopsis. The Plant Cell 16:2693-2704.
Kiba T, Kudo T, Kojima M, Sakakibara H (2011) Hormonal control of nitrogen acquisition: roles of auxin, abscisic acid, and cytokinin. Journal of Experimental Botany 62:1399-1409.
Kim S, Plagnol V, Hu TT, Toomajian C, Clark RM, Ossowski S, Ecker JR, Weigel D, Nordborg M (2007) Recombination and linkage disequilibrium in Arabidopsis thaliana. Nature Genetics 39:1151-1155.
Kooke R, Kruijer W, Bours R, Becker F, Kuhn A, van de Geest H, Buntjer J, Doeswijk T, Guerra J, Bouwmeester H, Vreugdenhil D, Keurentjes JJ (2016) Genome-Wide Association Mapping and Genomic Prediction Elucidate the Genetic Architecture of Morphological Traits in Arabidopsis. Plant Physiology 170:2187-2203.
Krapp A (2015) Plant nitrogen assimilation and its regulation: a complex puzzle with missing pieces. Current Opinion in Plant Biology 25:115-122.
Krouk G, Lacombe B, Bielach A, Perrine-Walker F, Malinska K, Mounier E, Hoyerova K, Tillard P, Leon S, Ljung K, Zazimalova E, Benkova E, Nacry P, Gojon A (2010) Nitrate-regulated auxin transport by NRT1.1 defines a mechanism for nutrient sensing in plants. Developmental Cell 18:927-937.
Lacaze X, Hayes PM, Korol A (2009) Genetics of phenotypic plasticity: QTL analysis in barley, Hordeum vulgare. Heredity 102:163-173.
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 of the United States of America 107:21199-21204.
Liang G, He H, Yu D (2012) Identification of nitrogen starvation-responsive microRNAs in Arabidopsis thaliana. PloS One 7:e48951.
Liu MS, Li HC, Lai YM, Lo HF, Chen LF (2013) Proteomics and transcriptomics of broccoli subjected to exogenously supplied and transgenic senescence-induced cytokinin for amelioration of postharvest yellowing. Journal of Proteomics 93:133-144.
Loudet O, Chaillou S, Krapp A, Daniel-Vedele F (2003a) Quantitative Trait Loci Analysis of Water and Anion Contents in Interaction With Nitrogen Availability in Arabidopsis thaliana. Genetics 163:711-722
Loudet O, Chaillou S, Merigout P, Talbotec J, Daniel-Vedele F (2003b) Quantitative trait loci analysis of nitrogen use efficiency in Arabidopsis. Plant Physiology 131:345-358.
Mach JM, Castillo AR, Hoogstraten R, Greenberg JT (2001) The Arabidopsis-accelerated cell death gene ACD2 encodes red chlorophyll catabolite reductase and suppresses the spread of disease symptoms. Proceedings of the National Academy of Sciences of the United States of America 98:771-776.
Maloof JN, Borevitz JO, Dabi T, Lutes J, Nehring RB, Redfern JL, Trainer GT, Wilson JM, Asami T, Berry CC, Weigel D, Chory J (2001) Natural variation in light sensitivity of Arabidopsis. Nature Genetics 29:441-446.
Matejovic I (1995) Total nitrogen in plant material determinated by means of dry combustion: A possible alternative to determination by Kjeldahl digestion. Communications in Soil Science and Plant Analysis 26:2217-2229.

Menz J, Range T, Trini J, Ludewig U, Neuhauser B (2018) Molecular basis of differential nitrogen use efficiencies and nitrogen source preferences in contrasting Arabidopsis accessions. Scientific Reports 8.

Meyer RC, Gryczka C, Neitsch C, Muller M, Brautigam A, Schlereth A, Schon H, Weigelt-Fischer K, Altmann T (2019) Genetic diversity for nitrogen use efficiency in Arabidopsis thaliana. Planta.

Mounier E, Pervent M, Ljung K, Gojon A, Nacry P (2014) Auxin-mediated nitrate signalling by NRT1.1 participates in the adaptive response of Arabidopsis root architecture to the spatial heterogeneity of nitrate availability. Plant, Cell & Environment 37:162-174.

Nguyen GN, Rothstein SJ, Spangenberg G, Kant S (2015) Role of microRNAs involved in plant response to nitrogen and phosphorous limiting conditions. Frontiers in Plant Science 6:629.

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 47:912-918.

Peng M, Bi YM, Zhu T, Rothstein SJ (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.

Platt A, Horton M, Huang YS, Li Y, Anastasio AE, Mulyati NW, Agren J, Bossdorf O, Byers D, Donohue K, Dunning M, Holub EB, Hudson A, Le Corre V, Loudet O, Roux F, Warthmann N, Weigel D, Rivero L, Scholl R, Nordborg M, Bergelson J, Borevitz JO (2010) The scale of population structure in Arabidopsis thaliana. PLoS Genetics 6:e1000843.

Ramirez V, Coego A, Lopez A, Agorio A, Flors V, Vera P (2009) Drought tolerance in Arabidopsis is controlled by the OCP3 disease resistance regulator. The Plant journal : For Cell and Molecular Biology 58:578-591.

Rauh L, Basten C, Buckler St (2002) Quantitative trait loci analysis of growth response to varying nitrogen sources in Arabidopsis thaliana. Theoretical and Applied Genetics 10:743-750.

Raun WR, Johnson GV (1999) Improving nitrogen use efficiency for cereal production. Agronomy Journal 91:357-363

Sakakibara H, Takei K, Hirose N (2006) Interactions between nitrogen and cytokinin in the regulation of metabolism and development. Trends in Plant Science 11:440-448.

Schneider A, Steinberger I, Strissel H, Kunz HH, Manavski N, Meurer J, Burkhard G, Jarzombski S, Schunemann D, Geimer S, Flugge UI, Leister D (2014) The Arabidopsis Tellurite resistance C protein together with ALB3 is involved in photosystem II protein synthesis. The Plant Journal: For Cell and Molecular Biology 78:344-356.

van Rooijen R, Kruijer W, Boesten R, van Eeuwijk FA, Harbinson J, Aarts MGM (2017) Natural variation of YELLOW SEEDLING1 affects photosynthetic acclimation of Arabidopsis thaliana. Nature Communications 8.

Walch-Liu P, Ivanov, II, Filleur S, Gan Y, Remans T, Forde BG (2006) Nitrogen regulation of root branching. Annals of Botany 97:875-881.

Wang R, Okamoto M, Xing X, Crawford NM (2003) Microarray analysis of the nitrate response in Arabidopsis roots and shoots reveals over 1,000 rapidly responding genes and new linkages to glucose, trehalose-6-phosphate, iron, and sulfate metabolism. Plant Physiology 132:556-567.

Wang R, Tischner R, Gutierrez RA, Hoffman M, Xing X, Chen M, Coruzzi G, Crawford NM (2004) Genomic analysis of the nitrate response using a nitrate reductase-null mutant of Arabidopsis. Plant Physiology 136:2512-2522.

Wang YY, Hsu PK, Tsay YF (2012) Uptake, allocation and signaling of nitrate. Trends in Plant Science 17:458-467.

Zhu C, Gore M, Buckler ES, Yu J (2008) Status and Prospects of Association Mapping in Plants. The Plant Genome Journal 1:5.

Zuber H, Davidian JC, Wirtz M, Hell R, Belghazi M, Thompson R, Gallardo K (2010) Sultr4;1 mutant seeds of Arabidopsis have an enhanced sulphate content and modified proteome suggesting metabolic adaptations to altered sulphate compartmentalization. BMC Plant Biology 10:78.

Chapter 3

A phenotypic investigation of 20 accessions of *Arabidopsis thaliana* differing in nitrogen use efficiency

Nihal Oztolan Erol[1], Joost JB Keurentjes[1], Mark GM Aarts[1]

[1]Laboratory of Genetics, Wageningen University, Wageningen, the Netherlands

Abstract:

Plants evolved to survive under limited N supply by using N efficiently. Improving N use efficiency (NUE) has become a sustainable strategy for wild species to grow in low N environments. In order to clarify physiological responses in relation with NUE and related traits. 20 accessions of *Arabidopsis thaliana* were analysed under sub-optimal N condition. These accessions were equal in shoot dry weight (SDW), however, different in NUE. N-use-efficient (NE) and N-use-inefficient (NI) accessions were subjected to detailed investigations on phenotypic differences of shoot fresh weight (SFW), SDW, water content (WC), relative growth rate (RGR), total leaf area (TLA), shoot N concentration (N%), NUE, N usage index (NUI) and Photosystem II efficiency (ΦPSII); and phenotypic plasticity determined with the residuals of SDW (SDW_Res), SFW (SFW_Res), N% (N%_Res), NUE (NUE_Res), NUI (NUI_Res), WC (WC_Res), and ΦPSII (ΦPSII _Res). Results showed that NE accessions had higher NUE, NUI, and RGR, but lower SFW, WC, and N% values than NI accessions under sub-optimal N condition. NE accessions could efficiently assimilate and use N with less N uptake, because they produce similar amount of SDW like NI accessions. WC and SFW were higher in NI accessions than NE accessions, this might explain simultaneous increase in water and N uptake in NI accessions under sub-optimal N condition. Cytokinin, auxin, and abscisic acid levels would explain this antagonistic behaviour between NE and NI accessions. Phenotypic plasticity was observed separately in two different clusters, namely NE and NI accessions. Highly plastic and stable accessions were identified within each cluster. CIBC17, Uk-1, Bur-0, and Col-0 belonged to the plastic and Ca-0, Ag-0, Kelsterbach-2, Aa-0, CIBC2, and Sei-0 belonged to the stable NI accessions; where, Ven-1, Paw-2, NFA-8, and Br-0 belonged to the plastic and TOU-A1-67, PHW-36, Wt-3, WAR, TOU-I-6, and Sav-0 belonged to the stable NE accessions. These results would support a further investigation about the genetics of phenotypic stability and genotype by environment interaction, which would be helpful for agricultural strategies developed by crop breeders.

Introduction

Nitrogen (N) is an essential macronutrient for plants. It plays an important role during plant development. N uptake, translocation and utilization ensure a healthy growth and reproduction for plants under an optimal level of N availability in soil. Plants, growing under limited N supply, show impaired physiological responses. Therefore, investigating different physiological responses in accordance with changing N supplies would help us to understand how traits are affected and correlated.

Research on natural variation in plant species shows that heterogeneous environments lead to varieties evolving in accordance with their surroundings (Weigel 2012). Limited N-environments have driven plants into an adaptation process of a sustainable life cycle by using N more efficiently. The broad distribution range in the northern hemisphere and the availability of naturally occurring inbred accessions make *A. thaliana* very convenient for investigation of natural intraspecific variation for many adaptive traits, including N-use-efficiency (NUE) and related agronomic traits (NRAT) (Chardon et al., 2010; Gifford et al. 2013; Loudet et al., 2003; Loudet et al., 2003). Diverse varieties of *Arabidopsis thaliana* were used here in order to test the physiological response under sub-optimal N-supply.

Biomass accumulation, nutrient uptake and usage, and also photosynthesis are three mechanisms that work in harmony in plants (Krouk et al. 2011; Rauh et al. 2002; Chapin et al. 1988; D.W. 2002; Balasubramanian et al. 2006; Ripley et al. 2008). Growth and photosynthesis impairments in plants are observed simultaneously due to the limited N supply (Overman AR 2011).In chapter 2, we analysed shoot biomass, chlorophyll fluorescence, N-concentration and related traits to estimate the performance of 354 accessions, the so-called HapMap population, under optimal and sub-optimal N-supplies. In this chapter, we re-examined the HapMap population data in order to focus on extreme accessions in the population which are grouped as N-use-efficient (NE) and N-use-inefficient (NI) accessions. We investigated differences in shoot dry weight (SDW), shoot fresh weight (SFW), water content (WC), N concentration (N%), NUE, N usage index (NUI), chlorophyll fluorescence (ΦPSII), total leaf area (TLA), and relative growth rate (RGR) between these two groups. We also investigated the phenotypic plasticity of SDW, SFW, WC, N%, NUE, NUI, and ΦPSII to understand which accessions are prone to change its phenotypic expressions under changing environmental conditions.

Materials and Methods

GWAS data re-assessment

In chapter 2, a population of 354 diverse natural accessions of *A. thaliana* (the so-called HapMap population (Weigel and Mott 2009; Li et al. 2010)) was phenotyped for SDW, SFW, WC, N%, NUE, NUI, ΦPSII, TLA, and RGR. Phenotypic results were re-assessed in order to find extreme groups of accessions. Initially, 100 accessions displaying the highest SDW under sub-optimal N-supply were selected. A further selection was conducted among these 100 accessions regarding their SDW, NUE, NUI and N% values under sub-optimal N-supply using heatmap and dendrogram methods to eliminate small plants, and to compare big plants for their NUE. Out of 100 accessions, 10 accessions showing low NUE and 10 showing high NUE were selected and referred to as NI and NE, respectively.

Experimental setup

Three subsequent experiments were carried on under the same climate conditions. The twenty selected accessions were grown in a climate-controlled growth chamber at short day conditions (10 hours light). Light intensity was set at 200 µmol m^{-2} s^{-1}; relative humidity at 60%; and day and night temperatures were 20˚C and 18˚C, respectively. Plants were grown on rock-wool blocks placed on tables prepared for irrigation by flooding and draining rock-wool blocks automatically (Gent and McAvoy 2011). 5 replications per accession were grown in every treatment with random block design. Plants were grown with two treatments, optimal N-supply and sub-optimal N-supply. Standard nutrient solutions contained 2.94 mM K, 1.31 mM Ca, 0.402 mM Mg, 0.938 mM SO_4, 0.536 mM P, 7.5 µM Fe, 3 µM Mn, 1.5 µM Zn, 5 µM B, 0.25 µM Cu, 0.25 µM Mo supplemented with 0.5 mM NH_4 and 4.521 mM NO_3 (optimal N-supply) or 0.1 mM NH_4 and 0.904 mM NO_3 (sub-optimal N-supply) in the form of ammonium-nitrate. Nutrient solutions were applied three times per week in the morning. Plants were harvested 30 days after germination.

Plant phenotyping

For all plants SFW was measured at the time of harvesting. SDW of collected rosettes was measured after drying at 60˚C for three days. N% was measured using the micro-combustor DUMAS method with a Flash EA 1112 N/Protein Analyzer (Thermo Scientific, US) (Matejovic 1995). Subsequently, NUE and NUI were calculated as *SDW/N* and *SDW*(SDW/N)* respectively (where *N* is the shoot N-content expressed as *N%*SDW*) (Good et al. 2004). WC was calculated as *(SFW-SDW)/SDW*.

ΦPSII, TLA, and RGR data for the extreme accessions were retrieved from the HapMap population screening in chapter 2.

Statistical analysis

In order to understand the correlations between traits SDW, SFW, NUE, N%, NUI, TLA, RGR, WC, ΦPSII under two different N-supplies, the bivariate Pearson correlation module was used in R commander. Averages, standard errors, and T-test calculations of both NE and NI accessions were done using Microsoft Excel 2016, to test the significances of the differences between NE and NI.

For phenotypic plasticity, residuals were calculated using linear and fitted regression models for sub-optimal and optimal N treatments of SDW, SFW, NUE, N%, NUI, WC, and ΦPSII using R commander (Chu et al. 2008; Maloof et al. 2001; Fox 2005).

The linear model was created as:

$$y = \beta_0 + \beta_1 \, x + \varepsilon$$

where y is the observed values of the optimal N treatment and x is that of the sub-optimal N treatment; β_0 and β_1 refer to the regression coefficients explaining the intercept and the slope of the regression line respectively; ε refers to the random error.

The fitted regression model was created as:

$$\hat{y} = \beta_0 + \beta_1 \, x$$

where $\hat{y}$ refers to the estimated values of the optimal N treatment for the given x, the observed values of the sub-optimal N treatment.

The residual values (e) were calculated as the difference between the estimated ($\hat{y}$) and the observed (y) values of the optimal N treatment:

$$e = y - \hat{y}$$

In order to show NE and NI accessions under two separate clusters, dendrograms and heatmaps for SFW, N%, NUI, NUE, RGR, and WC of the sub-optimal N condition and for the residuals of SFW, N%, NUI, NUE, and WC were constructed using R project (R 2016).

Z-scores of SFW, N%, NUI, NUE, RGR, and WC of the sub-optimal N condition and the residuals of SFW, N%, NUI, NUE, and WC were calculated to standardize heatmap scales as:

$$z\text{-score} = (x - \mu) / \sigma$$

where x is the average trait values of each accession, μ refers to the mean value of NE or NI clusters, σ refers to standard deviation of NE or NI clusters.

Results

Re-assessing 100 accessions and selection of 20 extreme accessions

The HapMap population of 354 diverse *A. thaliana* natural accessions was re-assessed to select the extreme accessions. The investigation process excluded plants with small biomass under the sub-optimal condition. Therefore, top 100 *A. thaliana* accessions of the SDW scale were selected from the HapMap population and re-screened to confirm their stability across the subsequent experiments. SDW, SFW and N% were measured and NUE and NUI were calculated for the 100 accessions. Heatmap clustering of z-scores of SDW, NUE, NUI and N% revealed the phenotypic distance between accessions (Figure 1). The dendrogram clustering distinguished accessions in two clusters. The heatmap for SDW did not separate 100 accessions into clusters, whereas for NUE, N%, and NUI, it showed two separate clusters. Additionally, any correlation was not observed between SDW and NUE, N%, or NUI, on the other hand NUE, N%, and NUI were highly correlated with each other. Heatmap and dendrogram proved that the clustering of 100 accessions was independent of SDW, but it depended on NUE, N%, and NUI. 10 accessions from each cluster were chosen for further investigation (Table 1). The accessions Ca-0, Col-0, Sei-0, Uk-1, CIBC17, Bur-0, Aa-0, Ag-0, Kelsterbach-2, CIBC2 at the lower extreme of the NUE distribution were selected as NI representatives. The accessions WAR, Sav-0, TOU-I-6, Wt-3, TOU-A1-67, NFA-8, PHW-36, Paw-3, Br-0, and Ven-1 at the higher extreme of the NUE distribution were selected as NE representatives.

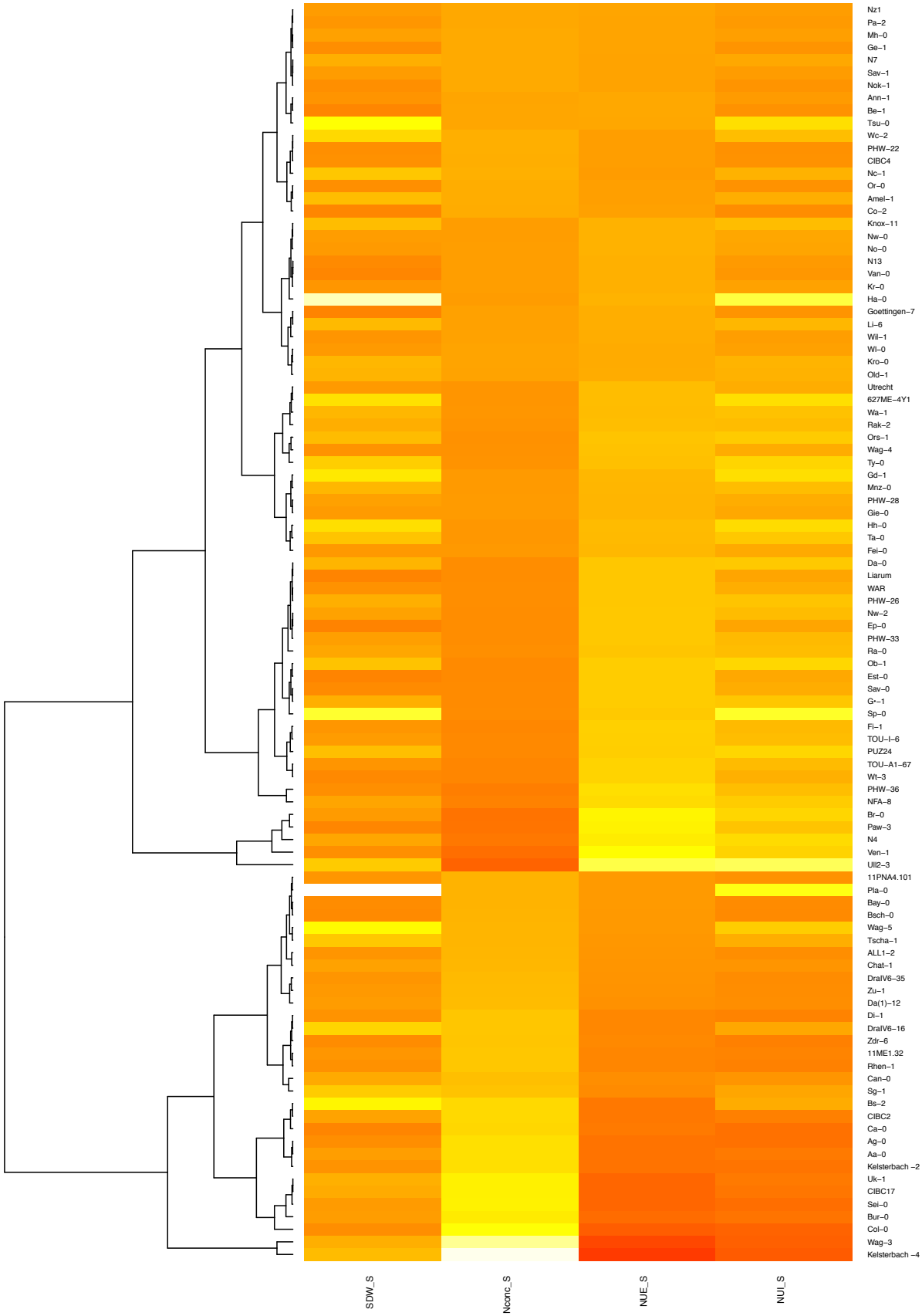


Figure 1: Heatmap and dendrogram of shoot dry weight (SDW), shoot nitrogen concentration (Nconc), nitrogen use efficiency (NUE), and nitrogen usage index (NUI) of 100 accessions collected under sub-optimal N supply (S). Color scale indicates the z-score of the traits.

Table 1: a. Average shoot dry weight (SDW), shoot fresh weight (SFW), nitrogen concentration (N%), nitrogen use efficiency (NUE), nitrogen usage index (NUI), water content (WC), Photosystem II efficiency (ΦPSII), total leaf area (TLA), and relative growth rate (RGR) of nitrogen-use-efficient (NE) and nitrogen-use-inefficient (NI) accessions grown under the sub-optimal (S) condition. **b.** Residuals of SDW (SDW_Res), SFW (SFW_Res), N% (N%_Res), NUE (NUE_Res), NUI (NUI_Res), WC (WC_Res), ΦPSII (ΦPSII _Res)

a.

NE	SDW_S	SFW_S	N%_S	NUE_S	NUI_S	WC_S	PSII_S	TLA_S	RGR_S
Ven-1	0,077	0,342	1,351	74,041	0,057	3,464	0,647	1249,750	0,244
Br-0	0,080	0,383	1,382	72,364	0,058	3,815	0,661	1746,889	0,227
Paw-3	0,074	0,331	1,393	71,791	0,053	3,480	0,672	1306,667	0,263
PHW-36	0,077	0,359	1,477	67,719	0,052	3,684	0,668	1159,333	0,262
NFA-8	0,083	0,408	1,501	66,635	0,055	3,948	0,666	1535,333	0,230
TOU-A1-67	0,078	0,401	1,539	64,986	0,051	4,116	0,650	1639,250	0,236
Wt-3	0,075	0,339	1,540	64,927	0,049	3,531	0,675	1234,000	0,253
TOU-I-6	0,080	0,396	1,557	64,243	0,052	3,938	0,666	1069,667	0,240
Sav-0	0,075	0,398	1,578	63,379	0,048	4,291	0,663	1308,500	0,229
WAR	0,077	0,370	1,604	62,360	0,048	3,803	0,690	1620,250	0,228
NI									
CIBC2	0,082	0,433	2,224	44,958	0,037	4,285	0,678	873,750	0,242
Kelsterbach -2	0,077	0,497	2,271	44,039	0,034	5,420	0,669	1547,500	0,214
Ag-0	0,076	0,477	2,280	43,864	0,033	5,261	0,695	1543,667	0,221
Aa-0	0,081	0,502	2,282	43,829	0,035	5,226	0,651	1579,333	0,217
Bur-0	0,081	0,490	2,362	42,330	0,034	5,088	0,685	1540,000	0,208
CIBC17	0,084	0,508	2,421	41,312	0,035	5,042	0,675	1474,500	0,225
Uk-1	0,086	0,422	2,421	41,312	0,035	3,921	0,673	1233,000	0,227
Sei-0	0,080	0,411	2,429	41,164	0,033	4,161	0,673	1571,500	0,217
Col-0	0,076	0,490	2,549	39,234	0,030	5,444	0,682	1349,182	0,223
Ca-0	0,074	0,502	2,205	45,362	0,033	5,822	0,683	1345,250	0,208

b.

NE	SDW_Res	SFW_Res	N%_Res	NUE_Res	NUI_Res	WC_Res	PSII_Res
Ven-1	0,003	-0,015	-0,326	13,264	0,013	-0,369	-0,026
Br-0	0,002	-0,020	-0,300	11,770	0,012	-0,323	-0,011
Paw-3	0,000	-0,032	-0,340	12,880	0,011	-0,405	0,008
PHW-36	-0,002	-0,029	-0,249	8,568	0,007	-0,221	0,000
NFA-8	0,019	0,073	-0,382	11,867	0,021	-0,342	-0,006
TOU-A1-67	0,015	0,075	-0,288	8,780	0,016	-0,031	-0,011
Wt-3	-0,014	-0,112	-0,240	7,438	-0,002	-0,434	0,004
TOU-I-6	0,015	0,059	-0,198	6,008	0,014	-0,208	-0,006
Sav-0	0,003	-0,001	-0,178	5,154	0,007	-0,107	0,007
WAR	0,010	0,021	-0,236	6,503	0,011	-0,356	0,014
NI							
CIBC2	0,019	0,105	0,347	-9,939	0,003	0,029	0,003
Kelsterbach -2	0,006	0,096	0,259	-7,864	-0,002	0,879	-0,002
Ag-0	0,012	0,089	0,151	-5,852	0,002	0,251	0,027
Aa-0	0,020	0,152	0,302	-8,736	0,004	0,386	-0,017
Bur-0	0,009	0,080	0,506	-13,090	-0,005	0,501	-0,006
CIBC17	0,006	0,099	0,754	-19,820	-0,012	0,898	0,004
Uk-1	0,022	0,081	0,567	-14,178	0,001	-0,390	-0,001
Sei-0	0,009	0,019	0,401	-10,398	-0,003	-0,272	0,000
Col-0	0,008	0,108	0,614	-14,304	-0,006	0,908	0,002
Ca-0	0,011	0,141	0,164	-5,965	0,001	1,094	0,009

Trait correlations, average values and genotype clusters

To investigate the relationship between traits and environments Pearson correlations were calculated (Table 2). Traits, which were recorded under both N-supply conditions, did not show

very strong correlation between the two N-conditions. In fact, WC was the only trait showing a strong and positive correlation between the two environments. The lack of correlation for SDW and SFW between two N-treatments indicates that the 20 extreme accessions were differently affected by environmental changes.

A clustering of normalised trait values in sub-optimal conditions clearly illustrates the differences between the NE and NI accessions (Figure 2). While SFW, N% and WC displayed significantly higher values in the NI accessions, NUI, NUE and to a lesser extent RGR showed higher values in the NE accessions (Table 1 and 3). From all accessions Col-0 was the least efficient, whereas Ven-1 was the most efficient genotype in terms of N-use. Furthermore, Ca-0 had the highest WC among NI accessions. Paw-3 and PHW-36 had the highest RGR among NE accessions; whereas Bur-0 and Ca-0 had the lowest RGR among NI accessions. Paw-3 and Wt-3 had the lowest SFW among NE accessions. The traits TLA, SDW and ΦPSII measured for the sub-optimal N-supply, together with all measured and calculated traits collected under the optimal N-supply did not exhibit a clear separation between the NE and NI accessions (Table 1 and 3) and were not included in the clustering.

The correlation matrix and the heatmap of the 20 extreme accessions reveals interesting interactions between traits. For example, high values of RGR, NUE, and NUI in NE accessions and their strong and positive correlations might explain that NE accessions would accelerate the shoot growth in order to achieve equal SDW as NI accessions under sub-optimal N-conditions. On the other hand, high SFW, WC and N% in NI accessions and their strong and positive correlation might signify that high water intake couples with high N intake. This might raise questions about the effect of sub-optimal N-supply on stomatal closure and primary root growth in NI accessions.

Table 2: Pearson correlation analysis of Photosystem II efficiency (ΦPSII), shoot nitrogen concentration (N%), nitrogen use efficiency (NUE), nitrogen usage index (NUI), relative growth rate (RGR), shoot dry weight (SDW), shoot fresh weight (SFW), total leaf area (TLA), and water content (WC) of 20 extreme accessions grown in optimal (O) and sub-optimal (S) nitrogen treatments. Regression coefficient (R^2) values in blue

	PSII_O	PSII_S	N%_O	N%_S	NUE_O	NUE_S	NUI_O	NUI_S	RGR_S	SDW_O	SDW_S	SFW_O	SFW_S	TLA_S	WC_O	WC_S
PSII_O	1															
PSII_S	0,4692*	1														
N%_O	0,1632	0,4253	1													
N%_S	0,439*	0,459*	0,6316*	1												
NUE_O	-0,1832	-0,4205	-0,9947***	-0,627**	1											
NUE_S	-0,4142	-0,4794*	-0,6634***	-0,9945***	0,662**	1										
NUI_O	-0,0944	-0,1847	-0,7948***	-0,4829*	0,8154***	0,5183**	1									
NUI_S	-0,3803	-0,5224**	-0,6968***	-0,9789***	0,6918***	0,9854***	0,4943*	1								
RGR_S	-0,3757	-0,2628	-0,6262**	-0,683***	0,6169**	0,696***	0,5581**	0,6696***	1							
SDW_O	-0,0206	-0,0249	-0,5628**	-0,3347	0,5848**	0,3642	0,9449***	0,3139	0,4423*	1						
SDW_S	0,2621	-0,1233	-0,1077	0,3492	0,0877	-0,3404	-0,1661	-0,1796	-0,2354	-0,2835	1					
SFW_O	0,1193	0,249	-0,0862	0,1107	0,1209	-0,0945	0,6264**	-0,1601	-0,0758	0,8059***	-0,2979	1				
SFW_S	0,3176	0,3013	0,5917**	0,8387***	-0,5772**	-0,8507***	-0,5266**	-0,8302***	-0,802***	-0,4351*	0,2723	0,0639	1			
TLA_S	0,0051	-0,0455	0,2355	0,0615	-0,2139	-0,0524	-0,1135	-0,0396	-0,5347**	-0,0469	0,0153	0,1999	0,2448	1		
WC_O	0,167	0,3737	0,8511***	0,6988***	-0,836***	-0,7276***	-0,7238***	-0,7407***	-0,8071***	-0,5644**	0,0247	0,0256	0,8084***	0,3332	1	
WC_S	0,2498	0,3556	0,6614***	0,7626***	-0,6405**	-0,7783***	-0,514*	-0,8063***	-0,7683***	-0,3821	-0,0284	0,1423	0,9534***	0,2448	0,8394***	1

Table 3: Average and standard error (Stderr) values of shoot dry weight (SDW), shoot fresh weight (SFW), shoot nitrogen concentration (N%), nitrogen use efficiency (NUE), nitrogen usage index (NUI), water content (WC), Photosystem II efficiency (ΦPSII), total leaf area (TLA), and relative growth rate (RGR) of 10 nitrogen-use-efficient (NE) and 10 nitrogen-use-inefficient (NI) accessions, measured under sub-optimal (S) and optimal (O) N-conditions. Paired-end t-tests were calculated to show if the average trait value differences between NE and NI were significant or not (*p*-value<0,05).

	NE		NI		
	Average	Stderr	Average	Stderr	ttest
SDW_S	0,077	0,001	0,08	0,001	1,73E-01
SDW_O	0,126	0,008	0,11	0,005	1,21E-01
SFW_S	0,373	0,009	0,473	0,012	3,00E-06
SFW_O	1,327	0,07	1,359	0,053	7,22E-01
N%_S	1,492	0,028	2,344	0,035	4,89E-13
N%_O	5,486	0,101	6,302	0,196	2,00E-03
NUE_S	67,244	1,299	42,74	0,623	3,15E-10
NUE_O	18,283	0,332	16,017	0,541	3,00E-03
NUI_S	0,052	0,001	0,034	0,001	1,20E-09
NUI_O	0,023	0,002	0,018	0,001	2,40E-02
WC_S	3,807	0,087	4,967	0,199	1,60E-04
WC_O	9,628	0,225	11,372	0,328	4,70E-04
PSII_S	0,666	0,004	0,676	0,004	6,30E-02
PSII_O	0,632	0,005	0,647	0,006	5,90E-02
TLA_S	1386,964	72,798	1405,768	69,857	8,54E-01
RGR_S	0,241	0,004	0,22	0,003	1,00E-03

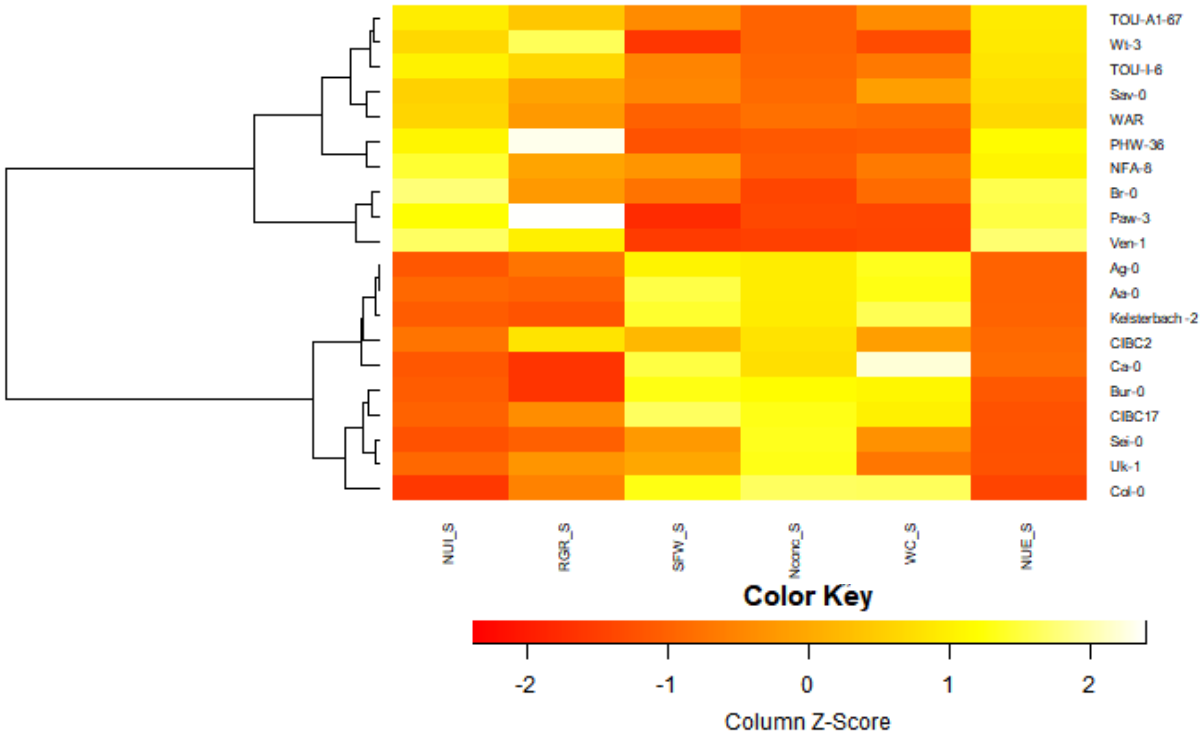


Figure 2: Heatmap and dendrogram of nitrogen usage index (NUI), relative growth rate (RGR), shoot fresh weight (SFW), shoot nitrogen concentration (Nconc), water content (WC) and nitrogen use efficiency (NUE) of nitrogen-use-efficient (NE) (TOU-A1-67, Wt-3, TOU-1-6, Sav-0, WAR, PHW-36, NFA-8, Br-0, Paw-3, Ven-1) and nitrogen-use-inefficient (NI) (Ag-0, Aa-0, Kelsterbach-2, CIBC2, Ca-0, Bur-0, CIBC17, Sei-0, Uk-1, Col-0) accessions collected under sub-optimal N supply (S). Color scale indicates the z-score of the traits.

Phenotypic plasticity of extreme accessions

In nature, plants adapt to their changing environment by changing their phenotypic expression, which is called phenotypic plasticity. Here, extreme accessions were tested to find out if their plasticity values put them in the pre-defined segregated clusters and how plastic they are in terms of environmental adaptation. Residual value of a trait shows the difference between observed and estimated value. Observed values were obtained via linear regression analysis, estimated values via fitted linear regression analysis between the sub-optimal and the optimal N conditions. Fitted regression analysis excludes the environmental error factor out of the model, so fewer environmental constraint would result in a smaller difference between the observed and estimated values and a residual closer to 0. However, greater the environmental constraint would cause greater the difference, hence greater the deviation in residuals.

Heatmap and pearson correlation analysis supported the significant and positive correlation between SFW_Res, WC_Res, and N%_Res, but significant and negative correlation between NUE_Res and NUI_Res (Figure 3 and Table 4) as previously mentioned. T-test results showed significant difference between NE and NI accessions for SFW_Res, N%_Res, NUE_Res, NUI_Res, and WC_Res and dendrogram showed a clear separation of NE and NI accessions into two clusters based on their phenotypes (Figure 3 and Table 5). This suggested that the phenotypic plasticity would be interpreted separately under two pre-defined clusters, namely NE and NI accessions.

Deviation from estimated value depends on the adaptation nature of every accession. The clustering based on residuals clearly created two sub-groups under both clusters. This sub-grouping separated plastic from stable accessions (Figure 3). For NUE and N%, CIBC17, Uk-1, Bur-0, and Col-0 belonged to the plastic and Ca-0, Ag-0, Kelsterbach-2, Aa-0, CIBC2, and Sei-0 belonged to the stable NI accessions; where, Ven-1, Paw-2, NFA-8, and Br-0 belonged to the plastic and TOU-A1-67, PHW-36, Wt-3, WAR, TOU-I-6, and Sav-0 belonged to the stable NE accessions. This explained that some accessions were tremendously affected by the genotype by environment interactions, so they underwent physiological changes.

Table 4: Pearson correlation analysis of residuals of Photosystem II efficiency (ΦPSII _Res), shoot nitrogen concentration (N%_Res), nitrogen use efficiency (NUE_Res), nitrogen usage index (NUI_Res), shoot dry weight (SDW_Res), shoot fresh weight (SFW_Res), water content

(WC_Res) of 20 extreme accessions. *** Correlation is significant at the 0, 001 level and ** at the 0.01 level (red).

	PSII_Res	N%_Res	NUE_Res	NUI_Res	SDW_Res	SFW_Res	WC_Res
PSII_Res	1						
N%_Res	0,1571	1					
NUE_Res	-0,2233	-0,9903***	1				
NUI_Res	-0,3073	-0,8539***	0,8473***	1			
SDW_Res	-0,1045	0,3367	-0,3693	0,1786	1		
SFW_Res	0,0256	0,625**	-0,6668**	-0,2466	0,8006***	1	
WC_Res	0,1627	0,652**	-0,6791***	-0,625**	0,1478	0,7072***	1

Table 5: Average and standard errors of residuals of Photosystem II efficiency (ΦPSII _Res), shoot nitrogen concentration (N%_Res), nitrogen use efficiency (NUE_Res), nitrogen usage index (NUI_Res), shoot dry weight (SDW_Res), shoot fresh weight (SFW_Res), water content (WC_Res) for 10 nitrogen-use-efficient (NE) and 10 nitrogen-use-inefficient (NI) accessions. Paired-end t-tests were calculated to show if the average trait value differences between NE and NI were significant or not (*p*-value<0,05).

	NE		NI		
	Average	Stderr	Average	Stderr	ttest
SDW_Res	0,005	0,003	0,012	0,002	7,01E-02
SFW_Res	0,002	0,018	0,097	0,011	4,81E-04
N%_Res	-0,274	0,020	0,406	0,063	6,51E-07
NUE_Res	9,223	0,950	-11,015	1,380	1,92E-09
NUI_Res	0,011	0,002	-0,002	0,002	9,19E-05
WC_Res	-0,280	0,042	0,429	0,165	1,90E-03
PSII_Res	-0,003	0,004	0,002	0,003	3,73E-01

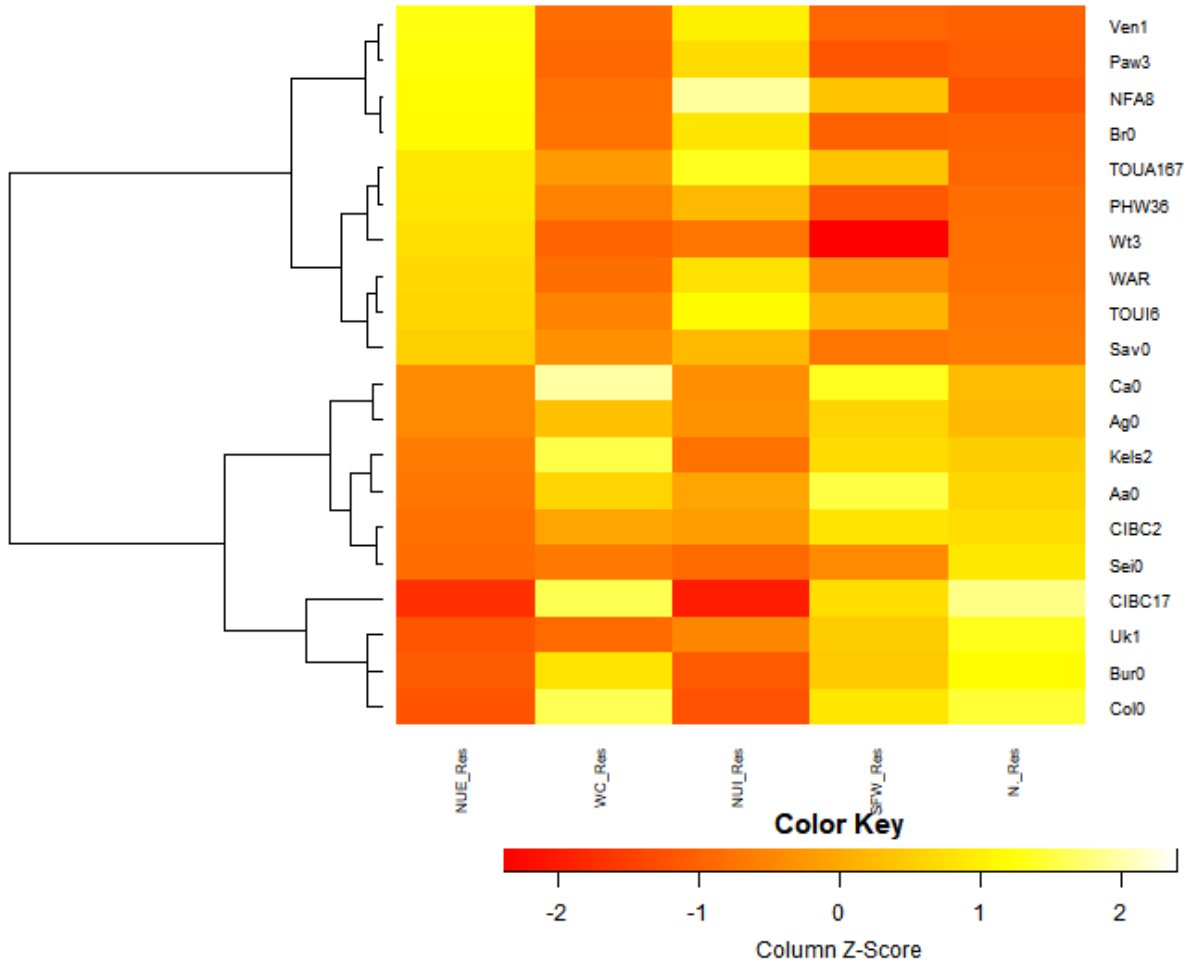


Figure 3: Dendrogram and heatmap of residuals of nitrogen use efficiency (NUE_Res), water content (WC_Res), nitrogen usage index (NUI_Res), shoot fresh weight (SFW_Res), and shoot nitrogen concentration (N%_Res) of nitrogen-use-efficient (NE) and nitrogen-use-inefficient (NI) accessions. Color scale indicates the z-scores of phenotypic plasticity across the changing environments. If $x \leq -2$ and $x \geq 2$, then those were considered highly unstable; if $-2 \leq x \leq -1$ and $1 \leq x \leq 2$, then those were considered moderately stable; if $-1 \leq x \leq 1$, then those were considered stable.

Discussion

The identification of extreme accessions is an important step for understanding the physiology of N-deficiency. NE and NI accessions were selected out of 100 accessions having very high SDW but contrasting NUE levels. While keeping one trait stable and doing the selection based

on NUE, we observed that not only NUE but also RGR, N%, WC, NUI, and SFW were different between NE and NI accessions.

Relationship between WC, RGR and NUE

N-deficiency triggers concentration changes in phytohormones such as cytokinin, auxin and abscisic acid levels in plant parts (Kiba et al. 2011; Marschner 1995). Under the suboptimal or N-deficient conditions, antagonistic effects of cytokinin and abscisic acid regulate adaptation to environmental changes; for instance, altered root, shoot, and leaf elongation rate, and rapid responsiveness of stomatal closure are well-observed responses against N-deficiency, governed by these hormones (Marschner 1995). Physiological alterations across NE and NI accessions under sub-optimal N-conditions could be well explained with these hormonal cross-talks in the plants.

SFW, WC and N% were on average lower, while NUE, NUI and RGR were higher in NE accessions than NI accessions under sub-optimal N-supply conditions. The explanation for the relationship between N% and WC might be that NI accessions developed a longer root system than NE accessions so that they could take up more water and nutrients and/or close their stomata more so that they would prevent water loss (Marschner 1995). The positive correlation between water and nitrate content was explained by the hypothesis of osmosis created by nitrate to absorb water (Loudet et al. 2003). This might be also the reason why NI accessions had higher SFW levels compared to NE accessions. On the other hand, we observed high NUE, NUI and RGR values in NE accessions. As these plants take up and allocate fewer N, their growth was accelerated under sub-optimal N-supply. This would explain the equal SDW values across NE and NI accessions. NE accessions with relatively low N-input could produce as much dry matter as NI accessions. Under these circumstances, auxin, cytokinin and abscisic acid levels of NE and NI accessions should be measured separately in the roots and the shoot. One would argue that the adjusted hormone levels might likely cause these physiological differences.

These results show that N-deficit-adapted accessions might be also adapted to water deficiency, since those ones could still produce equal amounts of shoot dry matter as high N and water accumulating accessions. Our findings would be very beneficial for a sustainable agriculture, when both water and N are deficient in soil.

Phenotypic plasticity

Gene, environment, and genotype by environment interaction are the basic determinants of a plant phenotype. The effects of these components can vary in every genotype under a certain environment, so plants are able to show phenotypic plasticity that they express different phenotypes under different environmental conditions. The reaction norms measure phenotypic plasticity of a genotype. A stable variety was at first described as a variety with stable yield capacity which is not affected by changing environmental conditions (Adugna and Labuschagne 2002; Eberhart 1966; Becker 1981). Even though phenotypic stability suggests highly predictable and controllable characters in plants, it does not respond to improved growth conditions. This approach was revisited and a stable variety was identified as a variety which produce expected amount of yield under given environmental conditions by neglecting genotype by environment interaction (Becker 1981; El-Soda et al. 2014). Our results indicated that NUE and N% traits were proportionally changed across the sub-optimal and the optimal N conditions for TOU-A1-67, PHW-36, Wt-3, WAR, TOU-I-6, and Sav-0, which belong to NE accessions, and for Ca-0, Ag-0, Kelsterbach-2, Aa-0, CIBC2, and Sei-0, which belong to NI accessions. In another words, genotype by environment interaction did not show a great effect on these genotypes. On the other hand, CIBC17, Uk-1, Bur-0, and Col-0, which belong to NI accessions, and Ven-1, Paw-2, NFA-8, and Br-0, which belong to the NE accessions, clearly showed the effect of genotype by environment interaction by expressing unexpected NUE and N% values across the two environments.

Investigating the determinants of genotype by environment interaction at first requires a significant confirmation analysis to understand plants behaviour under given environmental condition. Our results were confirmed with three subsequent experiments with five replications within every experiment, so this supports the validity of the phenotypic plasticity analysis across the sub-optimal and the optimal N conditions. Five genetic factors affecting genotype by environment interaction were identified as: overdominance, pleiotropy, epistasis, genetic linkage, and epigenetics (El-Soda et al. 2014). However, which of these factors explain the genetics of the significant genotype by environment interaction of the highly plastic accessions is yet needed to be investigated.

Conclusion

Our findings show that decreased N-uptake and allocation to above-ground organs might couple with decreased water uptake and allocation. This might mean that the adaptation to low N-input might also compromise water uptake of a plant. Under changing climate conditions, nutrient-deficiency and drought have become a rising environmental stress in agriculture. Excessive fertilizer input is causing environmental pollution and it is also an expensive application for farmers. Given results would lead plant geneticists and breeders to a combinatorial stress study to develop sustainable agricultural strategies.

The genetics of phenotypic plasticity would be a great added value to understand trade-offs and benefits of the plasticity. QTL analysis of two plastic NI and NE accessions might unravel genes associated with NUE adaptation across the optimal and the sub-optimal N conditions.

References

Adugna W, Labuschagne MT (2002) Genotype-environment interactions and phenotypic stability analyses of linseed in Ethiopia. Plant Breeding 121:66-71.

Balasubramanian S, Sureshkumar S, Agrawal M, Michael TP, Wessinger C, Maloof JN, Clark R, Warthmann N, Chory J, Weigel D (2006) The PHYTOCHROME C photoreceptor gene mediates natural variation in flowering and growth responses of Arabidopsis thaliana. Nature Genetics 38:711-715.

Becker HC (1981) Correlations among some statistical measures of phenotypic stability. Euphytica 30:835-840.

Chapin FS, Walter CHS, Clarkson DT (1988) Growth response of barley and tomato to nitrogen stress and its control by abscisic acid, water relations and photosynthesis. Planta 173:352-366.

Chu VT, Gottardo R, Raftery AE, Bumgarner RE, Yeung KY (2008) MeV+R: using MeV as a graphical user interface for Bioconductor applications in microarray analysis. Genome Biology 9:R118.

Lawlor DW (2002) Carbon and nitrogen assimilation in relation to yield mechanisms are the key to understanding production systems. Journal of Experimental Botany 53:773-787.

Eberhart SA, Russel WA (1966) Stability parameters for compairing varieties. Crop Science 6:36-40.

El-Soda M, Malosetti M, Zwaan BJ, Koornneef M, Aarts MGM (2014) Genotype x environment interaction QTL mapping in plants: lessons from Arabidopsis. Trends in Plant Science 19:390-398.

Fox J (2005) The R commander: A basic-statistics graphical user interface to R. Journal of Statistical Software 14:1-42.

Gent M, McAvoy R (2011) Water and nutrient uptake and use efficiency with partial saturation ebb and flow watering. HortScience 46:791-798.

Gifford ML, Banta JA, Katari MS, Hulsmans J, Chen L, Ristova D, Tranchina D, Purugganan MD, Coruzzi GM, Birnbaum KD (2013) Plasticity regulators modulate specific root traits in discrete nitrogen environments. PLoS Genetics 9:1-13.

Good AG, Shrawat AK, Muench DG (2004) Can less yield more? Is reducing nutrient input into the environment compatible with maintaining crop production? Trends in Plant Science 9:597-605.

Kiba T, Kudo T, Kojima M, Sakakibara H (2011) Hormonal control of nitrogen acquisition: roles of auxin, abscisic acid, and cytokinin. Journal of Experimental Botany 62:1399-1409.

Krouk G, Ruffel S, Gutierrez RA, Gojon A, Crawford NM, Coruzzi GM, Lacombe B (2011) A framework integrating plant growth with hormones and nutrients. Trends in Plant Science 16:178-182.

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 of the United States of America 107:21199-21204.

Loudet O, Chaillou S, Krapp A, Daniel-Vedele F (2003) Quantitative trait loci analysis of water and anion contents in interaction with nitrogen availability in Arabidopsis thaliana. Genetics 163:711-722.

Maloof JN, Borevitz JO, Dabi T, Lutes J, Nehring RB, Redfern JL, Trainer GT, Wilson JM, Asami T, Berry CC, Weigel D, Chory J (2001) Natural variation in light sensitivity of Arabidopsis. Nature Genetics 29:441-446.

Marschner H (1995) Mineral nutrition of higher plants. 2nd edn. Academic Press, London etc.

Matejovic I (1995) Total nitrogen in plant material determinated by means of dry combustion: A possible alternative to determination by Kjeldahl digestion. Communications in Soil Science and Plant Analysis 26:2217-2229.
Overman AR, Scholtz RV (2011) Accumulation of biomass and mineral elements with calendar time by corn: Application of the expanded growth model. PlosOne 6:1-5.

R CT (2016) R: A language and environment for statistical computing.
Rauh L, Basten C, Buckler St (2002) Quantitative trait loci analysis of growth response to varying nitrogen sources in Arabidopsis thaliana. Theoretical and applied genetics 104:743-750.
Ripley BS, Abraham TI, Osborne CP (2008) Consequences of C4 photosynthesis for the partitioning of growth: a test using C3 and C4 subspecies of Alloteropsis semialata under nitrogen-limitation. Journal of Experimental Botany 59:1705-1714.
Weigel D (2012) Natural variation in Arabidopsis: from molecular genetics to ecological genomics. Plant Physiology 158:2-22.
Weigel D, Mott R (2009) The 1001 genomes project for Arabidopsis thaliana. Genome Biology 10:107.

Chapter 4

Quantitative trait loci mapping of nitrogen use efficiency using the Br-0 x Col-0 recombinant inbred lines population of *Arabidopsis thaliana*

Nihal Oztolan Erol[1], Joost JB Keurentjes[1], Mark GM Aarts[1]

[1]Laboratory of Genetics, Wageningen University, Wageningen, the Netherlands

Abstract

Classical agricultural techniques increased the usage of N fertilizers and this caused the accumulation of N derived pollution in marine and soil ecosystems. As a solution, N use efficiency (NUE), the efficiency to take up and utilize biologically reactive N from the growth medium, has been employed in crop breeding. In this chapter an F8 recombinant inbred line (RIL) population of Br-0 and Col-0, differing in NUE, was exploited in search for causal genes involved in NUE and NUE related traits (NRAT) in *Arabidopsis thaliana* by using linkage and composite interval mapping. Shoot fresh weight (SFW), shoot dry weight (SDW), shoot N concentration (N%), N usage index (NUI), N use efficiency (NUE), C concentration (C%), and total leaf area (TLA) were measured and used for the analysis. Residual values were calculated using linear regression analysis in order to find phenotypic plasticity across the environments. Linkage and composite interval mapping showed seven loci in different genomic regions: Q1_1 (chromosome 2), Q1_2 and Q1_3 (chromosome 5) from SFW under sub-optimal N condition, Q2_1 (chromosome 5) from SFW under optimal N condition, Q3_1 and Q3_2 (chromosome 1) from NUI under sub-optimal N condition, Q4_1 (chromosome 5) from SFW residuals. Non-synonymous amino acid changes were detected in Q1_2, Q3_1 and Q4_1 using Genome Express Browser 3.0. Although the support intervals of these QTLs consisted of several candidate genes, *Arabidopsis thaliana* genome annotation database was used to select out promising candidate genes, in addition to genes co-locating with Q1_2, Q3_1, and Q4_1. One locus was identified as co-locating in both bi-parental linkage and genome-wide association mapping results. In conclusion, studying a novel bi-parental population would contribute to identify the genetics of a complex trait; therefore, Br-0 and Col-0 RIL population might shed a light on the genetics of NUE and NRAT in *Arabidopsis thaliana*.

Introduction

Nitrogen (N) is a growth limiting macro-nutrient for plants. It has various crucial roles in vital mechanisms in plants. It is involved in macro and micro molecules, secondary and signalling compounds such as nucleic acids, hormones, vitamins, amino acids and proteins (Krapp 2015). N loss to the environment causes severe problems in agriculture. In order to cope with the inadequate N nutrition in plants, fertilizers have been utilized; yet their usage has negative impacts on marine and terrestrial ecosystems (Vitousek 1997; D.W. 2002; Hirel et al. 2007). Another strategy, which has been applied by many crop breeders, is to develop new varieties which are improved in the efficiency of N usage. Therefore, an elaborate comprehension on the N uptake and assimilation and its genetic variation, is necessary to enhance agricultural strategies along with protecting natural habitats.

Reduced shoot biomass and seed production are frequently seen under limited N conditions (Schulte auf'm Erley et al. 2009). However, during the course of evolution different species have developed various strategies to be able to grow under N limited conditions (Gallais 2005; Chardon et al. 2012). Plants, with high N use efficiency (NUE), are able to grow and reproduce efficiently also under low N supply. This important characteristic has become a promising research subject for plant geneticists, physiologists and breeders. Working on NUE in plants reduces the need for N fertilizers, which are costly for farmers and also detrimental for ecosystems. Therefore, unravelling genes controlling NUE and related quantitative traits enhances the development of N-efficient crop varieties.

NUE and related phenotypes, such as shoot and root biomass, photosynthesis, carbon and nitrogen metabolism are quantitative (complex) traits and they show continuous variation within a population (Han et al. 2016). Quantitative traits are typically controlled by more than one gene, which can be revealed in genetic screens as quantitative trait loci (QTL) (Koornneef et al. 2004). QTLs can be distributed all over the genome and, in the absence of epistasis, show independent assortment in a Mendelian fashion (Griffiths AJF 2008). Another important feature of QTLs is the scale of their effect (Moreau et al. 2012). QTLs of a quantitative trait have often small-effect sizes and they can be sensitive to genotype by environment interaction. Additionally, their relative additive effect and QTL by QTL interaction (epistasis) affect phenotypic variation (Lian et al. 2005). Evidently, it may not be immediately clear to see a direct effect of genotype on phenotype for quantitative traits, unlike Mendelian variants in

qualitative traits where a large-effect gene corresponds with a specific phenotype (Mackay 2009).

Mapping strategies facilitate the investigation of evolutionary and functional mechanisms of genes. Improvements in genome sequencing have enhanced map building of many plant species. Single nucleotide polymorphisms (SNPs) are commonly utilized molecular markers to link polymorphisms on a chromosome map. The genotypic difference between different individuals is caused among others by single nucleotide differences between their alleles and its effect is detectable by genome-wide association study (GWAS) in natural populations (Korte and Farlow 2013; Chardon et al. 2014) and linkage mapping in experimental bi-parental populations (Nordborg and Weigel 2008). GWAS is based on haplotype mapping of linkage disequilibrium while linkage mapping is based on recombinant-frequency mapping. Even though the principle of both relies on finding closely linked loci of causal alleles, both techniques yield different outcomes due to utilization of different population types. The difference in population types determines what kind of alleles are expected to be discovered. If one would like to identify rare mutations, then conducting linkage mapping is more relevant than conducting GWAS. Because the detection power in GWAS increases with the magnitude of effect and also the frequency of an allele in a natural population rare alleles are difficult to detect (Rafalski 2010). Another feature of linkage mapping is the low resolution genetic map, but high QTL detection power. Fortunately, increasing resolution is possible by using additional bi-parental populations or increasing population size (Boer et al. 2007; Rafalski 2010; Chardon et al. 2014; Quraishi et al. 2011).

Genetic analysis of NUE has long been studied to understand the correlation between genetic variants and NUE and NUE related traits (NRAT). QTL mapping of RIL populations of *Arabidopsis thaliana* was performed to reveal putative genes controlling quantitative traits of interest (Alonso-Blanco et al. 2009; Weigel 2012). Rauh et. al. (2002) used RILs derived from a cross of Columbia (Col-0) and Landsberg *erecta* (Ler), to identify loci controlling aerial mass, root mass and root length of plants growing in different N sources (nitrate, ammonium nitrate, or ammonium) and also in limited N supply (Rauh et al. 2002). Loudet et al. contributed to QTL mapping of NUE in *A. thaliana* by studying the Bay-0 x Shahdara RIL population under two different N supplied conditions and mapping this population for shoot growth, total N, nitrate, amino acid, anions and also water contents (Loudet et al. 2003b; Loudet et al. 2003a). They identified co-localisations with the amino acid transporter gene *AAP5*, the cytosolic GS

gene *GLN1.2*, and the high-affinity nitrate transporter gene *NRT2.6* when they mapped shoot growth, total N, nitrate and amino acid contents. In addition, an aquaporin, chloride, and phosphate channel/transporter were identified when they mapped water and anion contents. In yet another study Diaz et al. reported a QTL co-localisation for NUE, flowering time, anthocyanin accumulation and leaf yellowing with the Bay-0 x Shahdara RIL population (Diaz et al. 2006). N regulatory genes were also investigated through RNA-seq techniques and novel microRNAs were found to be related with nitrogen-carbon crosstalk (Vidal et al. 2013). Regarding the nitrogen metabolism from soil to roots, and from roots to shoots, a group of genes have been frequently associated: ammonium transporters to uptake ammonium as a source of N (AMTs), nitrate transporters (*NRT1.1, NRT2.1, and NRT2.2*), nitrate reductases to convert nitrate to nitrite (NRs) and nitrite reductase (NiRs), ATP-dependent glutamine synthase and asparagine synthase isozymes to assimilate ammonium into amino acids (GS, AS), GS activity catalysers and their isoforms (GS1 and GS2), and finally, glutamate-oxoglutarate amino transferase (GOGAT) (Liu et al. 2015; Krapp 2015; Chardon et al. 2012; Quraishi et al. 2011; Vidal et al. 2015).

So far, a comparative analysis connecting QTLs from GWAS and bi-parental linkage mapping has not been done in *A. thaliana*. In this chapter, I aimed to increase the power of QTL detection by doing such a comparative analysis. Therefore, Br-0 and Col-0, two distinct genotypes, were selected as two parents for a RIL population, out of the natural population that we have reported on in Chapter 2 and 3. Br-0 and Col-0 were selected because these two accessions exhibited contrasting physiological NUE responses in both sub-optimal and optimal N conditions, as described in Chapter 3. This new RIL population may provide novel loci controlling NUE and the related quantitative traits. The parameters that we measured and used for QTL mapping were shoot fresh (SFW) and dry weights (SDW), carbon (C%) and shoot N concentration (N%) and also total leaf area (TLA). Later NUE and N usage index (NUI) were calculated using formulas described in Good et.al (2004). Residual analysis was studied in order to identify the genetic effect of phenotypic plasticity across sub-optimal and optimal N conditions.

Materials and Methods

Plant material, experimental setup and phenotyping

The RIL population (F8 generation), derived from the parents Br-0 (Brno from Czech Republic) and Col-0 (Poland), was obtained by single seed descent from an F2 population derived from

the cross between both parents (Zhang et al. 2014) and were kindly provided by Jan A.L. van Kan. 143 RILs were grown in a climate cell at short day (10 hours light) conditions. Light intensity was set at 200 µmol m^{-2} s^{-1}; humidity at 60%; and day and night temperatures were 20˚C and 18˚C, respectively. Three subsequent experiments were carried out in the same climate cell. In each of those experiments different measurements were reported from plants grown under optimal and/or sub-optimal nitrogen conditions. SDW, N%, and C%, from sub-optimal and optimal N conditions, were measured in the first experiment; SFW, from sub-optimal and optimal N conditions was measured in the second; and finally TLA, from only sub-optimal N condition, was measured in the third one. Every experiment was carried out in a randomized design with three blocks per treatment and one replication per block. Plants were grown on rock-wool blocks placed on flooding tables in order to sub-irrigate and drain rock-wool blocks automatically (Gent and McAvoy 2011). Standard nutrient solutions contained basal nutrients (2.94 mM K, 1.31 mM Ca, 0.402 mM Mg, 0.938 mM SO_4, 0.536 mM P, 7.5 µM Fe, 3 µM Mn, 1.5 µM Zn, 5 µM B, 0.25 µM Cu, 0.25 µM Mo) supplemented with 0.904 mM NO_3^- / 0.1 mM NH_4 for the sub-optimal nitrogen condition and 4.521 mM NO_3^- / 0.5 mM NH_4^+ for the optimal condition. Nutrient solutions were applied three times per week. Plants were harvested 32 days after sowing. SDW was measured after drying at 60˚C for three days. Mineral measurements were carried out with the support of Ismail Cakmak and Levent Ozturk (Sabanci University, Turkey). 100 mg of dried and ground shoot material were combusted and released gas was analysed to determine N and C concentrations using LECO TruSpec C/N analyser (Leco, Michigan, US). Subsequently, NUE and NUI were calculated as *SDW/N* and *SDW*(SDW/N)* respectively (where N is shoot N content: N%*SDW) (Good et al. 2004). In the second experiment, freshly harvested plants were weighed for SFW. And finally in third experiment, TLA was recorded using cameras in the climate cell and images were analysed using ImageJ.

Statistical analysis

Rcmdr, a package in R project, was used to calculate mean, coefficient of variance, one-way and two-way ANOVA, residual and fitted values (R 2016; Fox 2005). Residual values were calculated for the phenotypic plasticity analysis. The linear model was created as:

$$y = \beta_0 + \beta_1 \, x + \varepsilon$$

where y is the observed values of the optimal N treatment and x is that of the sub-optimal N treatment; β_0 and β_1 refer to the regression coefficients explaining the intercept and the slope of the regression line respectively; ε refers to the random error.
The fitted regression model was created as:

$$\hat{y} = \beta_0 + \beta_1 \text{ x}$$

where $\hat{y}$ refers to the estimated values of the optimal N treatment for the given x, the observed values of the sub-optimal N treatment.

The residual values (e) were calculated as the difference between the estimated ($\hat{y}$) and the observed (y) values of the optimal N treatment:

$$e = y - \hat{y}$$

Histograms of trait values were also drawn by Rcmdr. Broad sense heritability (H^2) was calculated using the following formula:

$$H^2 = V_G/(V_G+V_E),$$

where V_G stands for the genetic variance (the variation among phenotypic responses of individuals) and V_E stands for the environmental variance (the average of variation between replications of every individual).

DNA extraction and genotyping

DNA extracts were collected from fresh F8 flower buds of individuals of the Br-0 x Col-0 RIL population using 96-Well format DNA extraction with CTAB protocol designed by Maloof Lab (Maloof Lab). Homogenization procedure was carried out by simply grinding in a mortar using liquid nitrogen. Ground samples were mixed with 300 µl of 2x CTAB buffer (1 lt of 2x CTAB buffer contains 20 g of cetyltrimethyl or hexyltrimethyl ammonium bromide, 81,82 g of NaCl, 100 ml of 1 M Tris (pH8), 40 ml of 0,5 M EDTA). Samples were incubated in water bath at 65°C for 30 minutes. After they were cooled down to room temperature, they were centrifuged and mixed with 300 µl of chloroform and vortexed. Samples were centrifuged at 3250 rpm for 15 minutes. The supernatant was transferred into new tubes and 200 µl very cold isopropyl alcohol was added. Tubes were centrifuged at 3250 rpm for 15 minutes. the liquid was poured off and the pellet was washed with 200 µl of 70% ethanol and centrifuged again at 3250 rpm for 10 minutes. The washing step was repeated after pouring off the liquid. The pellet was dried

out through overnight and suspended within 50 µl of TE buffer. 15 µg of purified DNA samples from 143 lines were sent on dry ice to the Gregor Mendel Institute of Molecular Plant Biology in Vienna in order to conduct genotyping using a 384 Illumina SNP Assay (https://www.illumina.com/Documents/products/workflows/workflow_goldengate_assay.pdf)
.

Construction of the genetic map and QTL mapping

The genetic map was constructed using 138 polymorphic SNP markers by JoinMap 4.0 (Van Ooijen 2006). The average distance between markers was 3,29 cM with a minimum distance of 0,14 cM and a maximum distance of 10,05 cM. Interval mapping was conducted with R/qtl (Broman et al. 2003) to search for putative QTLs, afterwards composite interval mapping was performed by choosing the most significant marker as a cofactor to search for new QTLs within the proximity of 10 cM (Jiang and Zeng 1995; Zeng 1994). Permutation test was iterated for 1000 times in order to establish LOD scores for every single trait. Explained and additive variation were calculated using a linear model analysis:

For explained variation: *observed values ~ marker*
For additive variation: *observed values ~marker1 + marker2 + marker3*

In consideration with permutation test results, candidate genes were selected from the interval identified by significant SNP markers near the most significant SNP marker, which had LOD scores above the threshold within the 95% confidence interval. In order to estimate the effect of genetic variants on Br-0 and Col-0, non-synonymous amino acid change for the most significant marker was detected using *Arabidopsis* 1001 Genome Express Browser 3.0 at http://signal.salk.edu/atg1001/3.0/gebrowser.php. The Gene Ontology Annotations at www.arabidopsis.org were used as an evidence from published data to support association of candidate genes to N limitation.

Results

Physiological analysis of F8 population and the parental accessions Br-0 and Col-0

Br-0 and Col-0 (Figure 1) are two distinct *A. thaliana* accessions showing different physiological responses under sub-optimal and optimal N environments (Chapter 3).

Figure 1: Col-0 and Br-0 accessions grown in the sub-optimal N supply.

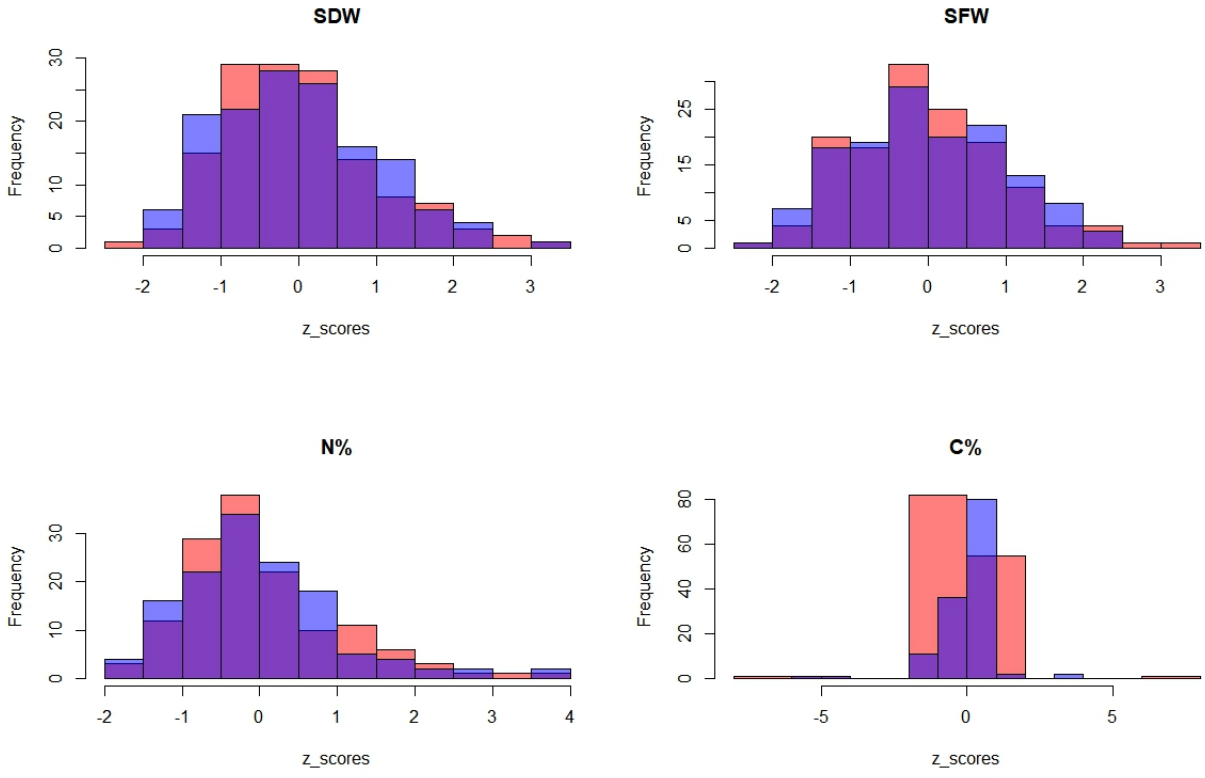
The F8 RIL population was screened in three subsequent experiments in the same climate cell. SFW, SDW, N%, C% were recorded under sub-optimal and optimal N conditions; NUI and NUE were computed as SDW(SDW/N) and SDW/N, respectively (where N is the shoot N content and calculated as N%*SDW). TLA was only measured under sub-optimal N condition. A Linear model analysis of the population trait values yielded residual and fitted values per trait, which was used to remove the outliers for further analysis. All traits, except for NUE, showed smaller mean values in the sub-optimal N supply than in the optimal N supply. Br-0 and Col-0 had equal values for SDW under sub-optimal N condition. Br-0 had lower N% but higher NUE than Col-0 under sub-optimal N condition. These results supported the results in Chapter 3. TLA in Br-0 is higher than Col-0, which indicates that leaf expansion in Br-0 might be bigger than Col-0 under sub-optimal N condition. Br-0 might trigger an increased photosynthetic rate for the purpose of boosting production of dry matter.

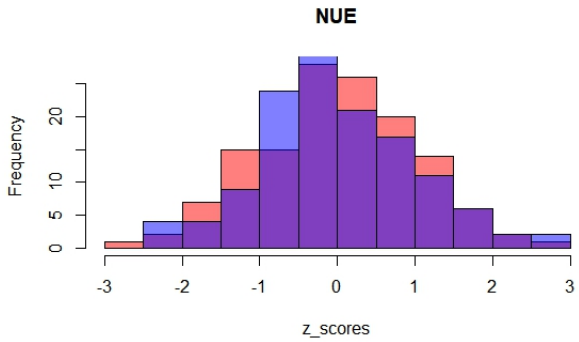
Coefficient of variation (CV) was calculated in order to understand the phenotypic variation in RIL population. CV ranged from 4,2 to 37,22, and H^2 ranged from 0,71 to 0,34 (Table 1). C%, under both N conditions, showed relatively smaller CV (9,608 for the sub-optimal and 4,2 for the optimal N condition) than others which symbolized a narrow phenotypic variation of the trait. Regarding H^2 values, around half of the phenotypic variation in RIL population for C% was regulated by genes under both N conditions (0,47 for the sub-optimal and 0,52 for the optimal N condition). TLA and NUI of sub-optimal N conditions and SDW of optimal N condition showed a higher range of H^2 (0,71, 0,60, and 0,59 respectively). More than half of the phenotypic variation for TLA, NUI and SDW was controlled by genes in RIL population. These traits also showed higher CV values, which explained the large phenotypic variation in the population.

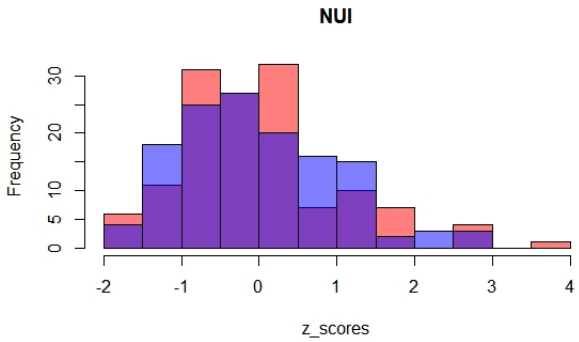
Table 1: Means, coefficient of variation (CV) and broad-range heritability (H^2) of total leaf area (TLA), shoot fresh weight (SFW), shoot dry weight (SDW), shoot nitrogen concentration (N%), carbon concentration (C%), nitrogen use efficiency (NUE), nitrogen usage index (NUI) in RILs and parental lines (Br-0 and Col-0) grown under sub-optimal (S) and optimal (C) nitrogen condition.

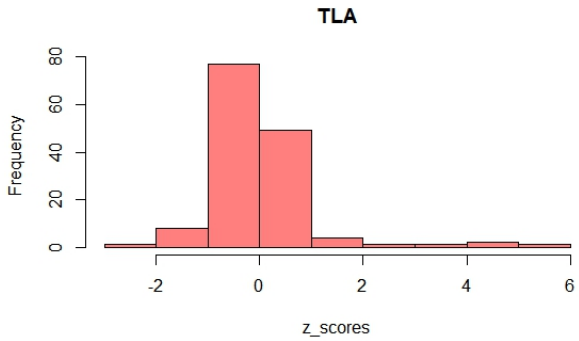
	RIL			Br-0	Col-0
	Mean	CV	H^2	Mean	Mean
SDW_S	0.039	19.785	0.359	0.036	0.036
SDW_O	0.111	27.85	0.591	0.13	0.07
SFW_S	0.496	16.828	0.43	0.38	0.49
SFW_O	1.105	26.739	0.34	1.51	1.39
N%_S	2.405	17.913	0.45	1.96	3.06
N%_O	3.271	17.668	0.45	2.82	4.35
C%_S	39.035	9.608	0.47	38.93	37.95
C%_O	38.634	4.2	0.52	41.47	36.32
NUE_S	43.346	24.69	0.5	50.96	32.67
NUE_O	31.484	17.067	0.47	35.34	22.94
NUI_S	1.698	37.222	0.6	1.87	1.18
NUI_O	3.450	26.597	0.5	4.93	1.77
TLA_S	27098	32.876	0.71	17510	15456

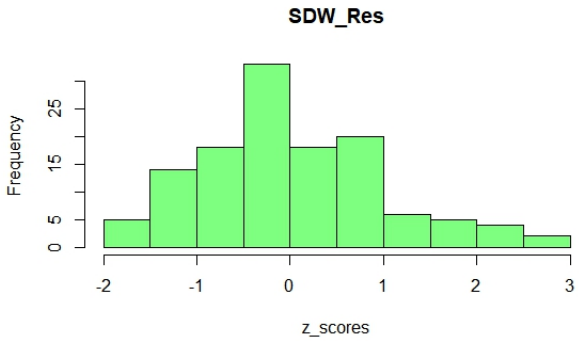
SDW, C%, TLA (under sub-optimal N), and SFW (under both N treatments) showed transgressive segregation (Table 1). SDW, C%, and TLA mean values of RIL population were higher than both parents. Frequency distribution graph of TLA, however, showed some outliers that might be the cause of the increase in mean value. On the other hand, SFW mean values of RIL population were smaller than parents.




NUE
Frequency
0
5
10
20
-3
-2
-1
0
1
2
3
z_scores


NUI
Frequency
0
5
10
20
30
-2
-1
0
1
2
3
4
z_scores


TLA
Frequency
0
20
40
60
80
-2
0
2
4
6
z_scores


SDW_Res
Frequency
0
5
15
25
-2
-1
0
1
2
3
z_scores

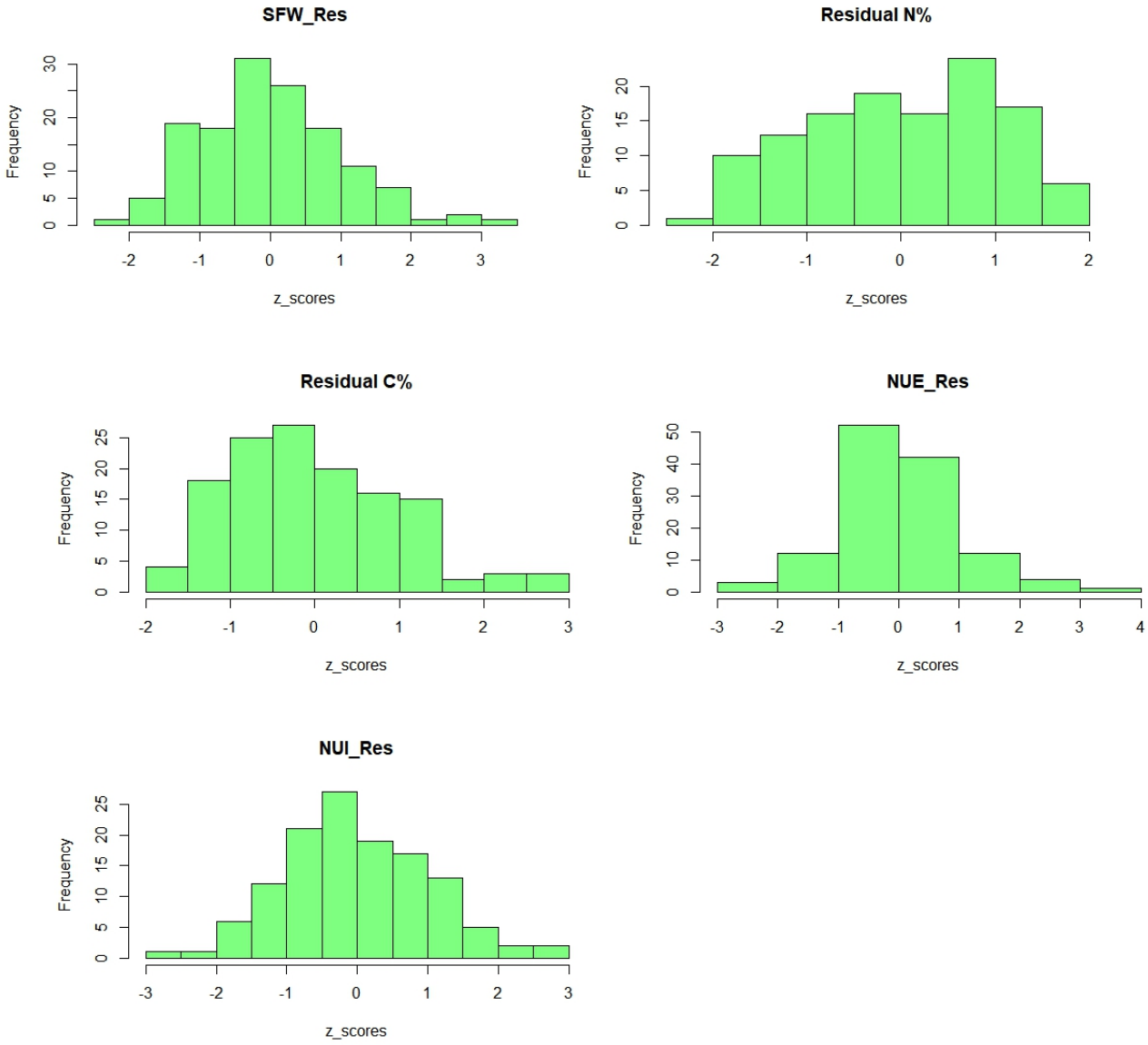


Figure 2: Frequency distribution of z-scores of shoot dry weight (SDW), shoot fresh weight (SFW), carbon concentration (C%), shoot nitrogen concentration (N%), nitrogen use efficiency (NUE), nitrogen usage index (NUI), and total leaf area (TLA) of the RIL population grown in the sub-optimal (red bars) and the optimal (blue bars) N environments and residuals of SDW, SFW, C%, N%, NUE, NUI (green bars).

Pearson correlation matrix showed the relation between traits and treatments. The correlation matrix displayed that N%, NUI, and NUE were strongly correlated with each other under both N conditions, because of interdependency of these traits. SDW and N% were positively correlated with each other under optimal N condition. However, a strong independency of traits was observed for the rest of the correlations. The missing correlation between SFW-TLA and SDW-TLA might be due to environmental effects, since they were measured in three subsequent experiments. On the other hand, the lack of positive correlation between two N

treatments within each experiment might be due to strong genotype by environment interaction (Table 2).

Marker segregation and the genetic map

Br-0 and Col-0 RIL population was not sequenced and mapped before. In this novel population, 138 out of 384 SNP markers were polymorphic between the two parents (Figure 3). Markers were positioned over 5 linkage groups corresponding to the 5 chromosomes of *A. thaliana* where it showed a correlated alignment in the physical and genetic maps. Marker segregation in the population yielded 49% Col-0 allele (AA) and 49% Br-0 allele (BB) along with a 2% ratio of heterozygous allele (AB). Homozygosity was 98% in the F8 generation of Br-0 and Col-0 RIL population.

QTL analysis

Trait CV values showed that there was a large phenotypic variation by which around half of the variation might be explained by genetic variants (Table 1). A QTL analysis was performed for SDW, SFW, N%, C%, NUE, NUI in the sub-optimal and optimal N conditions; TLA in the sub-optimal N condition; and phenotypic plasticity across the two N supplies to unravel those genetic variants. Interval mapping and subsequent composite interval mapping resulted in three QTLs on chromosomes 2 and 5 explaining variation in SFW under sub-optimal N supply; one QTL for SFW under optimal N supply on chromosome 5; one QTL for SFW residuals on chromosome 5; and two QTLs for NUI under sub-optimal N supply on chromosome 1 (Figure 4). Explained variance showed total phenotypic variance explained by every QTL (Table 3). QTL1_1, QTL1_2 and QTL1_3 individually explained 8%, QTL2_1 explained 9%, QTL3_1 and QTL3_2 explained 7.5% and 7.3% respectively, and finally QTL4_1 explained 8.5% of explained variation. Cumulative additive variation of Q1_1, Q1_2, and Q1_3 is 14% in SFW; and that of Q3_1 and Q3_2 is 7% in NUI under sub-optimal conditions. The found QTLs partially represented of genetic variation. It might be due to the nature of complex traits which are regulated by multiple small effect QTLs.

Average trait values of Br-0 or Col-0 haplotypes for each QTL were calculated to determine the allelic effects on phenotypes (Table 3). For each trait and QTL, Col-0 and Br-0 allele groups were slightly different from each other. In general, Br-0 alleles resulted in higher phenotypic values than Col-0 alleles. This small phenotypic difference might complicate the process of finding large effect QTLs.

Table 2: Pearson correlation matrix including shoot dry weight (SDW), shoot fresh weight (SFW), carbon concentration (C%), shoot shoot nitrogen concentration (N%), nitrogen use efficiency (NUE), nitrogen usage ndex (NUI), and total leaf area (TLA) under sub-optimal (Ndef) and optimal (Control) growth conditions.

	C%_O	C%_S	N%_O	N%_S	NUE_O	NUE_S	NUI_O	NUI_S	SDW_O	SDW_S	SFW_O	SFW_S	TLA_S
C%_O	1												
C%_S	0.2893**	1											
N%_O	-0.176	0.0812	1										
N%_S	0.2093*	-0.1206	0.0487	1									
NUE_O	0.146	-0.1127	-0.986***	-0.0182	1								
NUE_S	-0.2357*	0.1067	-0.044	-0.9845***	0.0172	1							
NUI_O	-0.1225	-0,.99	-0.2838**	-0.0866	0.2916**	0.0611	1						
NUI_S	-0.2954**	-0.0342	-0.0239	-0.6675***	0.0235	0.6828***	0.0666	1					
SDW_O	-0.2337*	-0.143	0.3083**	-0.078	-0.2968**	0.0539	0.814***	0.0655	1				
SDW_S	-0.2079	-0.1655	0.0152	-0.0042	0.0182	0.0076	0.0469	0.7276**	0.0553	1			
SFW_O	0.0629	0.0466	0.0055	0.0703	-0.0065	-0.0664	0.0796	-0.0912	0.0831	-0.061	1		
SFW_S	-0.0263	0.0043	-0.0738	-0.0668	0.0487	0.0589	0.1067	0.2055	0.0729	0.2302*	0.2114*	1	
TLA_S	0.1561	-0.0326	-0.0457	-0.0937	0.0281	0.0615	0.0138	0.0323	0.0074	0.0018	-0.0454	0.1022	1

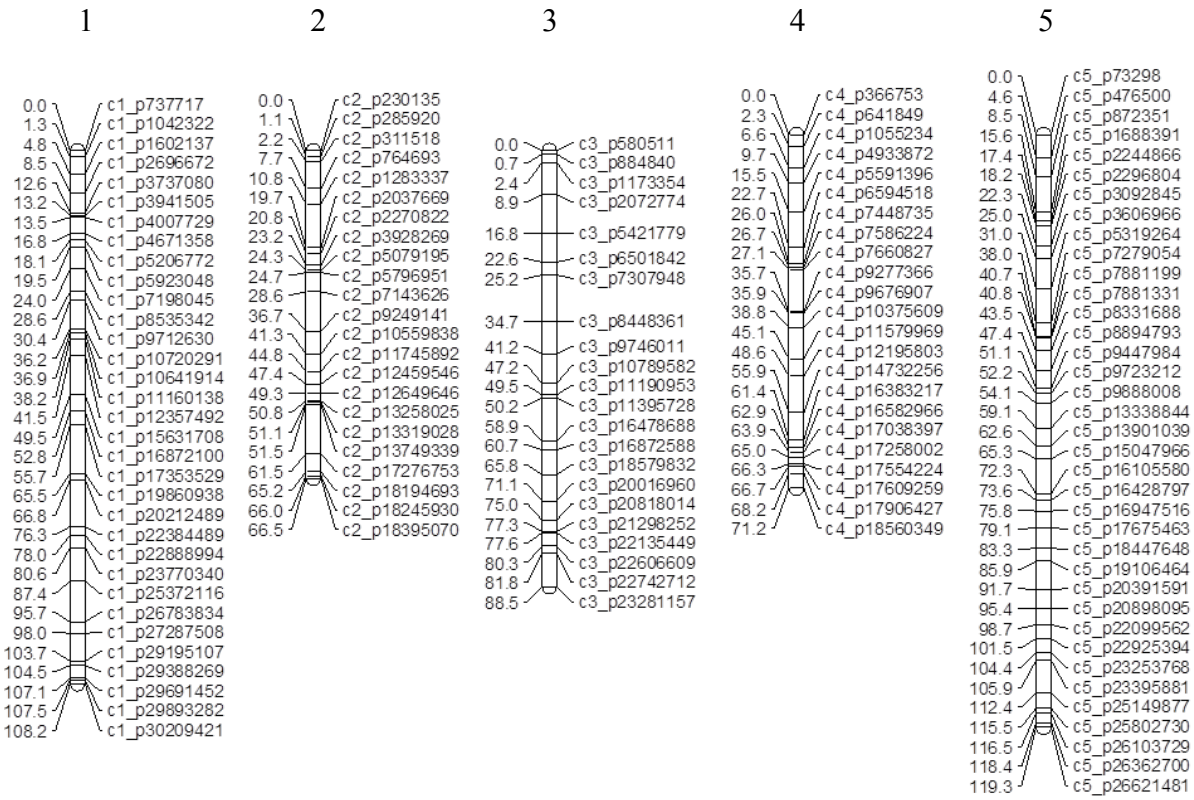


Figure 3: Genetic map of the *A. thaliana* Br-0 x Col-0 RIL population. 138 SNP markers were positioned across the genome over the five linkage groups corresponding to the five chromosomes of *A. thaliana*.

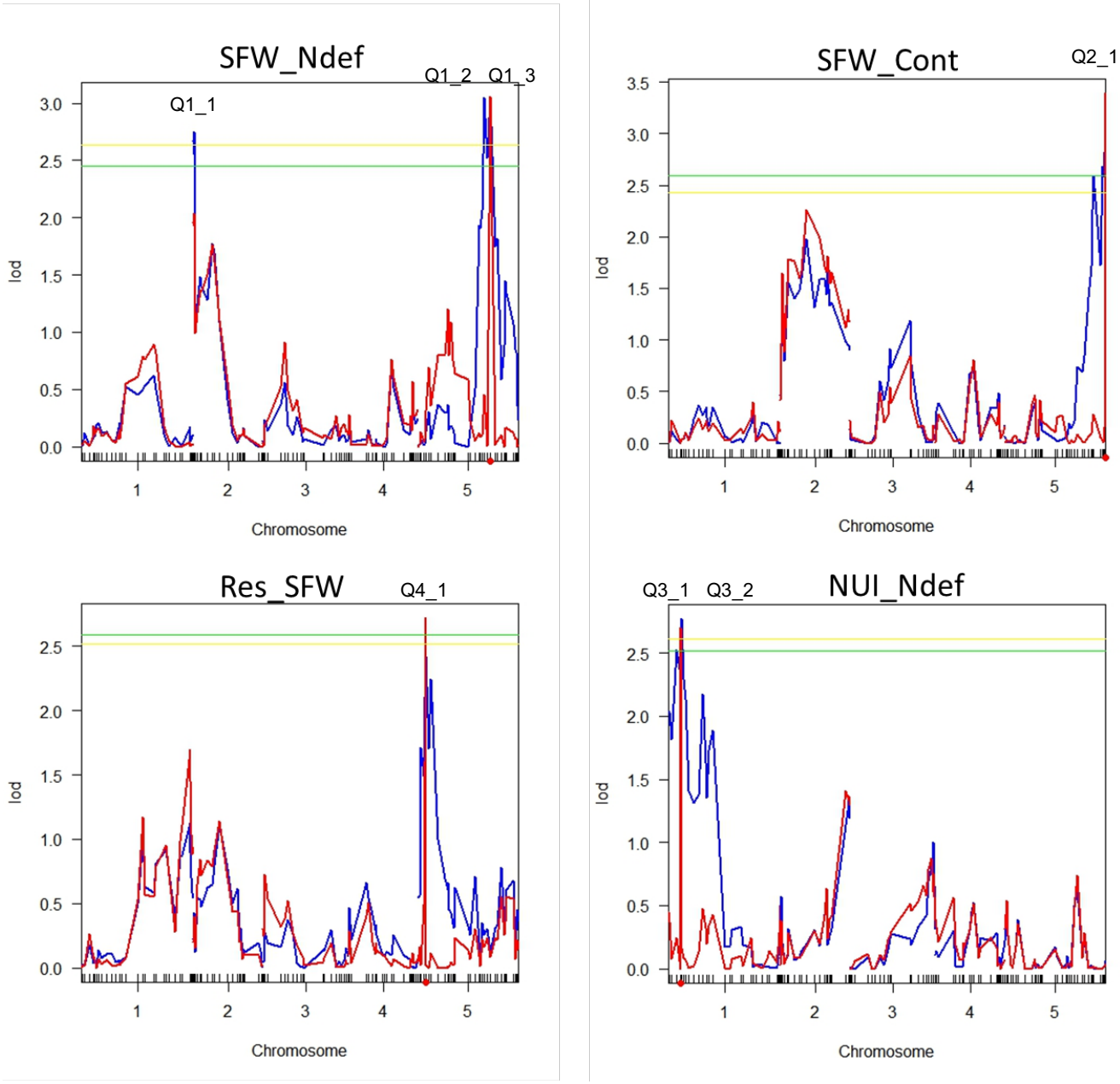


Figure 4: Interval mapping (IM) (blue) and composite interval mapping (CIM) (red) graphs of shoot fresh weight (SFW) under sub-optimal (Ndef) and optimal (Cont) nitrogen (N) conditions, residuals of SFW, and N usage index (NUI) under sub-optimal N condition. The most significant markers selected as co-factor for CIM were shown as red dots on the x-axis. LOD score thresholds were calculated using permutation test iterated 1000 times with *p-values* ≤ 0,05. Markers passing this threshold in interval mapping (green) and composite interval mapping (yellow) were selected as significant QTLs. Candidate QTL IDs are shown on top of the strongest peaks.

Table 3: Marker IDs, physical positions (Phy_Pos), chromosome number (Chr), LOD scores (LOD), explained (Exp. Var.), additive variations (Add. Var.) of QTLs, candidate gene position interval in kb (Cand. gen. int. (kb)) and LOD cut-off values (LOD cut-off) , averages (Col-0

and Br-0), standard deviations (Stdev Col-0 and Stdev Br-0), and percentages of Col-0 and Br-0 alleles (%Col-0 and %Br-0) of QTLs found in shoot fresh weight (SFW) in sub-optimal (Ndef) and optimal (Cont) nitrogen (N) conditions, N usage index (NUI) in sub-optimal N condition, and residuals of SFW. LOD scores were calculated using permutation test with 1000 times iteration (*p-values* ≤ 0,05). Candidate gene position interval is within 95% confidence interval. Explained and additive variations were calculated using linear model analysis.

Trait	QTL	Marker_ID	Phy_Pos	Chr	LOD	Exp. Var.	Add. Var.	Cand.gen.int. (kb)	LOD cut-off	Col-0	Br-0	Stdev Col-0	Stdev Br-0	% Col-0	% Br-0
SFW_Ndef	Q1_1	c2_p311518	311518	2	2.74	8%	0.146	230-311	2.56	0.47	0.5	0.06	0.08	43	57
	Q1_2	c5_p19106464	19106464	5	3.057	8%		17675-19106	2.52	0.47	0.5	0.08	0.08	51	49
	Q1_3	c5_p17675463	17675463	5	3.052	8%				0.47	0.5	0.08	0.08	52	47
SFW_Cont	Q2_1	c5_p26621481	26621481	5	3.392	9%		23395-26621	2.59	1.02	1.2	0.25	0.25	42	58
NUI_Ndef	Q3_1	c1_p4007729	4007729	1	2.771	7.5%	0.067	2696-4007	2.52	1.5	1.8	0.44	0.43	41	56
	Q3_2	c1_p3941505	3941505	1	2.699	7.3%				1.5	1.8	0.44	0.43	42	55
SFW_Res	Q4_1	c5_p2244866	2244866	5	2.722	8.5%		2244	2.72	-0.01	0.009	0.31	0.03	45	51

Bi-parental linkage mapping and GWAS co-localization and candidate genes

QTL mapping showed the putative causal loci associated with NUE, NRAT, and phenotypic plasticity. Candidate genes, associated with SFW, NUI under sub-optimal N condition and SFW_Res, were selected within a significant interval near the most significant SNP marker. The interval near Q1_1 included 37 candidate genes; Q1_2 & Q1_3 and Q3_1 & Q3_2 constituted a large significant interval including 697 and 801 candidate genes. Q4_1 was not surrounded with any significant markers so no interval was identified for this QTL. The prioritized candidate genes were selected based on the presence of non-synonymous amino acid change in Br-0. Q1_2 was co-located with Transducin family protein producing gene (AT5G43920). Q3_1 was co-located with Branchless trichome-like protein producing gene (AT1G11690). Q4_1 was co-located with *GA20OX3* (AT5G07200). The most significant markers were checked to find non-synonymous amino acid alterations between Br-0 and Col-0 via *A. thaliana* Genome Express Browser 3.0 (http://signal.salk.edu/atg1001/3.0/gebrowser.php) (Figure 5). Q1_2, on the chromosome 5 at the position 17675463, Q3_1, on the chromosome 1 at the position 3941505, and Q4_1, on the chromosome 5 at the position 2244866, pointed out non-synonymous amino acid change on coding regions in Br-0. The ones with non-synonymous amino acid alteration might modify the encoding protein and cause functional difference between Br-0 and Col-0.

Q1_2 and Q1_3 loci were common in GWAS results of WC_Res (Chapter 2) (Table 5). These loci were associated with SFW and WC_Res, which were two related traits phenotyped from two different populations. In Chapter 2, *CESA4* (AT5G44030) was detected within the significant proximity of the most significant SNP marker on chromosome 4 and consisted of

two non-synonymous amino acid change on coding region (Figure 5). This gene has a role in cellulose synthesis for cell wall (Guevara et al. 2014). This candidate gene would be also involved in the differing genetic mechanisms of Br-0 and Col-0 to sub-optimal N condition. Consequently, all candidate genes were listed in Table 5:

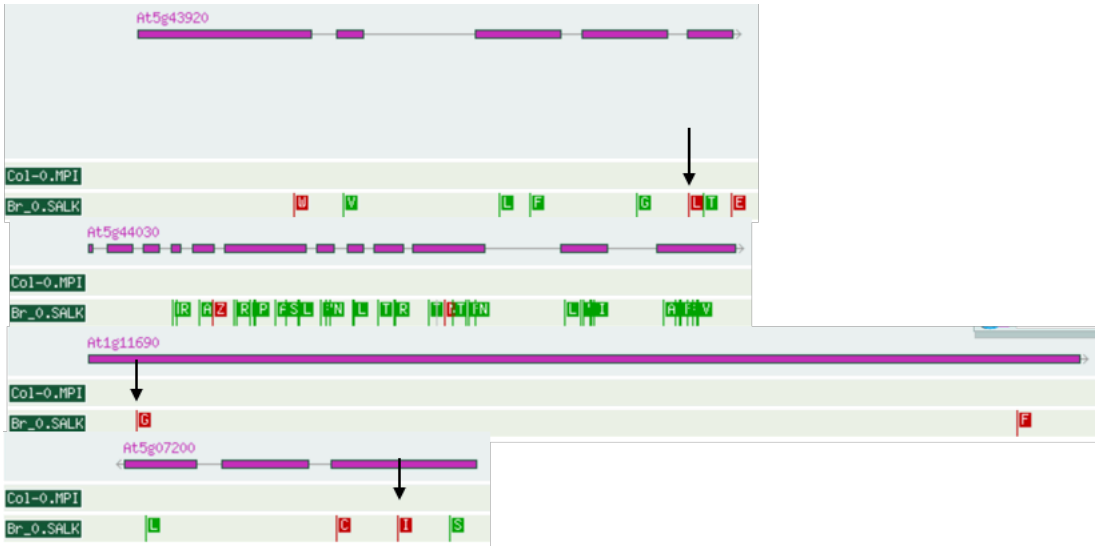


Figure 5: Synonymous (green lines) and non-synonymous (red lines) amino acid alterations in the candidate genes between Br-0 and Col-0. Standard amino acid codes are used. Blue arrows point significant alleles popping out through linkage mapping. Purple triangle indicates 1-3 bp insertions.

Table 5: Prioritized candidate genes and their associated traits, QTL IDs, chromosome (Chr) numbers, locus IDs (Gene) and names. Co-localized locus was indicated under GWAS co-localization. Q1_3 includes candidate genes proposed in a co-locating locus associated with WC_Res in Chapter 2.

Trait	QTL	Chr	Gene	Name	GWAS co-localization
SFW_S	Q1_2	5	AT5G43920	Transducin family protein	WC_Res
	Q1_3	5	AT5G44010 AT5G44020 AT5G44030 AT5G44040 AT5G44050	CESA4	
NUI_S	Q3_1	1	AT1G11690	Branchless trichome-like protein	
SFW_Res	Q4_1	5	AT5G07200	GA20OX3	

Discussion

Br-0, Col-0 and RIL population

Br-0 and Col-0 are two accessions that were selected out of a large natural population of *A. thaliana.* The most important characteristic was the difference between their NUE, TLA, NUI, and N% in a sub-optimal N supply (Table 1). A RIL population from these parents could be useful to investigate novel genes linked to NUE and NRAT which has never been done before.

North et al. 2009 compared a number of *A. thaliana* accessions including Br-0 and Col-0 (North et al. 2009). They grew *A. thaliana* accessions on agar plate consisting of 1 mM nitrate for sub-optimal and 4mM nitrate for optimal N conditions. They collected seedlings after 17 days and measured SFW. They found that SFW of Br-0 was always higher than Col-0 under both N supplies. On the contrary, in Chapter 3 and 4, results showed that Col-0 had higher SFW and N% than Br-0. Additionally, the positive correlation between SFW, WC, and N% under sub-optimal N condition (Chapter 3) explained an improved N and water uptake capacity of Col-0 under sub-optimal N condition to scavenge for mineral and water. These non-matching observations might be due to the differences in the source type of N applied to plants, in the cultivation of seedlings in different developmental stage, and/or in the growth medium. The lack of correlation between agar and soil was well established before, so comparing plant phenotypes grown in different media would give varying results (Meyer et al. 2019). Besides, Menz et al. (2018) showed that Col-0 was very responsive to ammonium nitrate than solely nitrate supply (Menz et al. 2018). Therefore, higher SFW in Col-0 than Br-0 might indicate that Col-0 was very sensitive to the type N source and favour ammonium nitrate as a source of N supply.

North et al. (2009) found that SFW reduction from optimal to sub-optimal N condition was quite significant in Br-0 than Col-0, which explains higher phenotypic plasticity for Br-0 (North et al. 2009). In agreement with North et al. (2009), SFW plasticity in Br-0 was more pronounced than Col-0 as explained in Chapter 3. Likewise, Chardon et al. (2010) supported the tolerant behaviour of Col-0 under changing N conditions (Chardon et al. 2012)). This made Col-0 more tolerant than Br-0 to changing N conditions in terms of SFW.

Another difference between our studies was that North et. al. did not show dry weight and/or N content differences between these two accessions under low N condition. Calculation of NUE and NUI are crucial measures for a better comprehension of the contribution of absorbed and

allocated N on biomass. Even though it is the only publication, which tested Br-0 and Col-0 accessions under limited N supply, it did not provide a deeper phenotypic investigation due to missing SDW, NUE and NUI values.

We screened a bigger population involving distinct individuals (Chapter 2) and conducted several tests to identify two phenotypically distant genotypes (Chapter 3). Br-0 and Col-0 were selected to produce high biomass using contrasting levels of N concentrations, where Br-0 absorbs and allocates less N compared to Col-0. The question then raises; how could Br-0 produce an equal amount of biomass as Col-0? The answer might lay in the differences of carbon metabolism between Br-0 and Col-0. Larger total leaf area would increase light harvesting capacity of Br-0 under a given photosynthetic nitrogen use efficiency (amount of carbon assimilated into sugars per unit leaf N content). As a result, starch accumulation might be increased to compromise biomass production under decreased N supply. Our results indicate that increased photosynthetic nitrogen use efficiency helped Br-0 to produce carbon bodies and accumulate them rather than use them for amino acid production. QTL mapping of RIL population revealed Q1_3 locus, which was co-located with *CESA4*. This gene was found to be associated with carbohydrate assimilation to cell wall (Guevara et al. 2014) and could be the reason for the biomass production in Br-0 under low leaf N content. These characteristics of Br-0 make it even more interesting to continue with elaborative research about the relationship between photosynthesis and nitrogen use efficiency.

QTL mapping

QTL mapping of Br-0 and Col-0 RIL population was expected to show QTLs in association with NUE, NRAT, and phenotypic plasticity across two N conditions. Rauh et al. (2002), Loudet et al. (2003), and Diaz et al. (2006) previously studied QTL analysis in *A. thaliana* RIL populations under changing N supplies (Loudet et al. 2003b; Loudet et al. 2003a; Diaz et al. 2006; Rauh et al. 2002). Rauh et al. (2002) used Col-0 and Ler RIL population in order to unravel genetic mechanisms associated with different N sources and N limitation. Loudet et al. (2003a and 2003b) and Diaz et al. (2006) used the same Bay-0 and Sha RIL population to detect QTLs associated with SDW, amino acid, nitrate, chloride, phosphate, and water content, and also leaf yellowing and anthocyanin accumulation under sub-optimal N condition (Loudet et al. 2003a; Loudet et al. 2003b; Diaz et al. 2006). Both Loudet et al. (2003a and 2003b) and Diaz et al. (2006) showed common QTLs since they used the same RIL population under similar growth conditions. Also they identified correlating traits such as dry mass and leaf yellowing, so their results very likely matched with each other. Rauh et al. (2002) and Loudet et al.(2003b)

identified one common QTL on the chromosome 5 associated with shoot biomass under optimal N condition. On the other hand, no common loci were identified between their and my QTL study. The reason might be due to using different parents exhibiting varying genetic mechanisms to N limitation. The natural variation in *A. thaliana* consists of a large genetic variation (Meyer et al. 2019). This causes a large phenotypic variation among the *A. thaliana* accessions. Bay-0, Sha, Ler, Col-0, and Br-0 might undergo different strategies to cope with limited N conditions.

Likewise the results presented in this chapter, Rauh et al. (2002), Loudet et al. (2003a and 2003b), and Diaz et al. (2006) showed small effect QTLs. It was well-known that NUE, NRAT, and phenotypic plasticity are complex traits that are controlled by multiple QTLs. In order to overcome this problem, a mixed RIL population could be used to detect more powerful QTL results instead of bi-parental population. Further QTL analysis would be done using a complex cross RIL population generated from extreme *A. thaliana* accessions explained in Chapter 3.

Candidate genes

Bi-parental linkage and composite interval mapping revealed small effect QTLs associated with SFW, NUI and Res_SFW. Supporting QTL intervals consisted of several candidate genes yet to be discovered. Non-synonymous amino acid change in Transducin family protein encoding gene, Branchless trichome-like protein encoding gene, *CESA4,* and *GA20OX3* might be associated with phenotypic differences in NUI, SFW (under sub-optimal N condition) and SFW plasticity between Br-0 and Col-0. In order to confirm these results, a detailed research was done in literature and *A. thaliana* annotation resources. Consequently, the literature search provided more information about *CESA4* gene in relation with N stress conditions.

CESA4 is co-located within Q1_3 locus on the chromosome 5, associated with SFW under sub-optimal N condition. It was also identified in GWAS result of WC_Res. This gene was identified in defence mechanism against soil borne bacterium *Ralstonia solanacearum* and the necrotrophic fungus *Plectosphaerella cucumerina* (Hernandez-Blanco et al. 2007). Another study showed that N limitation could down-regulate *CESA4* gene in *A. thaliana* and rice which would cause reduction in carbohydrate deposition to cell wall (Guevara et al. 2014). The discovery of *CESA4,* as a common locus in GWAS and bi-parental linkage mapping, might answer the question about Br-0's high biomass production absorbing fewer N than Col-0. Plants undergo starch accumulation and alter carbon metabolism under N stress conditions. They

rather use carbon bodies for growth instead of using them in energy metabolism for amino acid production (Robinson 1997). *CESA4* could have an important role in carbohydrate accumulation in Br-0. Further analysis would be to make a *cesa4* knock-out mutant Br-0 and check for the effect of the candidate gene in comparison to wild type and Col-0 under sub-optimal N condition.

Conclusion

My initial motive to choose Br-0 and Col-0 as two parents was to narrow down the investigation towards N%, NUE and NUI, when holding other parameters stable. By this way, I could unravel the exact genetic identity of NUE in *A. thaliana.* The QTL analysis of a Br-0 x Col-0 RIL population for NUE and NRAT yielded small effect QTLs segregating in the RIL population. However, due to the complexity of traits, I could only identify seven small effect QTLs from SFW, under both sub-optimal and optimal N conditions, NUI, under sub-optimal N condition, and SFW_Res. Among them *CESA4* was found to be associated with SFW. My other interest in this research was to detect co-locating QTLs resulting from the HapMap and bi-parental RIL populations. One locus (Q1_3), consisting of *CESA4* gene, was identified in both populations supporting candidate gene validation process.

Genetic analysis of NUE will provide strong impacts not only on the agriculture but also on the marine and terrestrial ecosystems. Diminishing usage of N fertilizers will be economical for farmers, besides its negative effects on the environment will be vanished. Therefore, it is very important to keep on studying this trait by analysing mutant types of identified candidate genes to see the effect of sub-optimal N nutrition on plant growth and development and also by investigating natural genetic variants shaping the large phenotypic variation among *A. thaliana* accessions.

References

Alonso-Blanco C, Aarts MG, Bentsink L, Keurentjes JJ, Reymond M, Vreugdenhil D, Koornneef M (2009) What has natural variation taught us about plant development, physiology, and adaptation? The Plant Cell 21:1877-1896.

Boer MP, Wright D, Feng LZ, Podlich DW, Luo L, Cooper M, van Eeuwijk FA (2007) A mixed-model quantitative trait loci (QTL) analysis for multiple-environment trial data using environmental covariables for QTL-by-environment interactions, with an example in maize. Genetics 177:1801-1813.

Broman KW, Wu H, Sen S, Churchill GA (2003) R/qtl: QTL mapping in experimental crosses. Bioinformatics 19:889-890

Chardon F, Jasinski S, Durandet M, Lecureuil A, Soulay F, Bedu M, Guerche P, Masclaux-Daubresse C (2014) QTL meta-analysis in Arabidopsis reveals an interaction between leaf senescence and resource allocation to seeds. Journal of Experimental Botany 65:3949-3962.

Chardon F, Noel V, Masclaux-Daubresse C (2012) Exploring NUE in crops and in Arabidopsis ideotypes to improve yield and seed quality. Journal of Experimental Botany 63:3401-3412.

Lawlor DW (2002) Carbon and nitrogen assimilation in relation to yield mechanisms are the key to understanding production systems. Journal of Experimental Botany 53:773-787.

Diaz C, Saliba-Colombani V, Loudet O, Belluomo P, Moreau L, Daniel-Vedele F, Morot-Gaudry JF, Masclaux-Daubresse C (2006) Leaf yellowing and anthocyanin accumulation are two genetically independent strategies in response to nitrogen limitation in Arabidopsis thaliana. Plant & Cell Physiology 47:74-83.

Fox J (2005) The R commander: A basic-statistics graphical user interface to R. Journal of Statistical Software 14.

Gallais A, Coque M (2005) Genetic variation and selection for nitrogen use efficiency in maize: A synthesis. Maydica 50:531-547

Gent M, McAvoy R (2011) Water and nutrient uptake and use efficiency with partial saturation ebb and flow watering. HortScience 46:791-798.

Good AG, Shrawat AK, Muench DG (2004) Can less yield more? Is reducing nutrient input into the environment compatible with maintaining crop production? Trends in Plant Science 9:597-605.

Griffiths AJF, Wessler S, Lewontin RC, Carroll SB (2008) Introduction to Genetic Analysis. 9th edn. W. H. Freeman and Company.

Guevara DR, El-Kereamy A, Yaish MW, Mei-Bi Y, Rothstein SJ (2014) Functional characterization of the rice UDP-glucose 4-epimerase 1, OsUGE1: A potential role in cell wall carbohydrate partitioning during limiting nitrogen conditions. PloS One 9:e96158.

Han M, Wong JL, Su T, Beatty PH, Good AG (2016) Identification of nitrogen use efficiency genes in barley: Searching for QTLs controlling complex physiological traits. Frontiers in Plant Science 7.

Hernandez-Blanco C, Feng DX, Hu J, Sanchez-Vallet A, Deslandes L, Llorente F, Berrocal-Lobo M, Keller H, Barlet X, Sanchez-Rodriguez C, Anderson LK, Somerville S, Marco Y, Molina A (2007) Impairment of cellulose synthases required for Arabidopsis secondary cell wall formation enhances disease resistance. The Plant Cell 19:890-903.

Hirel B, Le Gouis J, 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.
Jiang C, Zeng ZB (1995) Multiple trait analysis of genetic mapping for quantitative trait loci. Genetics 140:1111-1127.
Koornneef M, Alonso-Blanco C, Vreugdenhil D (2004) Naturally occurring genetic variation in Arabidopsis thaliana. Annual Review of Plant Biology 55:141-172.
Korte A, Farlow A (2013) The advantages and limitations of trait analysis with GWAS: a review. Plant Methods 9:29.
Krapp A (2015) Plant nitrogen assimilation and its regulation: a complex puzzle with missing pieces. Current Opinion in Plant Biology 25:115-122.
Lian XM, Xing YZ, Yan H, Xu CG, Li XH, Zhang QF (2005) QTLs for low nitrogen tolerance at seedling stage identified using a recombinant inbred line population derived from an elite rice hybrid. Theoretical and Applied Genetics 112:85-96.
Liu Q, Chen X, Wu K, Fu X (2015) Nitrogen signaling and use efficiency in plants: what's new? Current Opinion Plant Biology 27:192-198.
Loudet O, Chaillou S, Krapp A, Daniel-Vedele F (2003a) Quantitative trait loci analysis of water and anion contents in interaction with nitrogen availability in Arabidopsis thaliana. Genetics 163:711-722.
Loudet O, Chaillou S, Merigout P, Talbotec J, Daniel-Vedele F (2003b) Quantitative trait loci analysis of nitrogen use efficiency in Arabidopsis. Plant Physiology 131:345-358.
Mackay TF (2009) Q&A: Genetic analysis of quantitative traits. J Biology 8:23.
Maloof Lab 96-Well Format DNA Extraction with CTAB. http://malooflab.openwetware.org/96well_CTAB.html. Accessed 18 Nov 2016
Menz J, Range T, Trini J, Ludewig U, Neuhauser B (2018) Molecular basis of differential nitrogen use efficiencies and nitrogen source preferences in contrasting Arabidopsis accessions. Scientific Reports 8.

Meyer RC, Gryczka C, Neitsch C, Muller M, Brautigam A, Schlereth A, Schon H, Weigelt-Fischer K, Altmann T (2019) Genetic diversity for nitrogen use efficiency in Arabidopsis thaliana. Planta.
Moreau D, Burstin J, Aubert G, Huguet T, Ben C, Prosperi JM, Salon C, Munier-Jolain N (2012) Using a physiological framework for improving the detection of quantitative trait loci related to nitrogen nutrition in Medicago truncatula. Theoretical and Applied Genetics 124:755-768.
Nordborg M, Weigel D (2008) Next-generation genetics in plants. Nature 456:720-723.
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 47:912-918.
Quraishi UM, Abrouk M, Murat F, Pont C, Foucrier S, Desmaizieres G, Confolent C, Riviere N, Charmet G, Paux E, Murigneux A, Guerreiro L, Lafarge S, Le Gouis J, Feuillet C, Salse J (2011) Cross-genome map based dissection of a nitrogen use efficiency ortho-metaQTL in bread wheat unravels concerted cereal genome evolution. The Plant Journal : for Cell and Molecular Biology 65:745-756.
R CT (2016) R: A language and environment for statistical computing.
Rafalski JA (2010) Association genetics in crop improvement. Current Opinion in Plant Biology 13:174-180.
Rauh L, Basten C, Buckler St (2002) Quantitative trait loci analysis of growth response to varying nitrogen sources in Arabidopsis thaliana. Theoretical and Applied Genetics 104:743-750.

Robinson JM (1997) Nitrogen limitation of spinach plants causes a simultaneous rise in foliar levels of orthophosphate, sucrose, and starch. International Journal of Plant Sciences 158:432-441.

Schulte auf'm Erley G, Dewi ER, Nikus O, Horst WJ (2009) Genotypic differences in nitrogen efficiency of white cabbage (Brassica oleracea L.). Plant and Soil 328:313-325.

Van Ooijen JW (2006) JoinMap 4, software for the calculation of genetic linkage maps in experimental populations. Wageningen, Netherlands Patent.

Vidal EA, Alvarez JM, Moyano TC, Gutierrez RA (2015) Transcriptional networks in the nitrate response of Arabidopsis thaliana. Current Opinion in Plant Biology 27:125-132.

Vidal EA, Moyano TC, Krouk G, Katari MS, Tanurdzic M, McCombie WR, Coruzzi GM, Gutierrez RA (2013) Integrated RNA-seq and sRNA-seq analysis identifies novel nitrate-responsive genes in Arabidopsis thaliana roots. BMC Genomics 14:701.

Vitousek PMA, J.D.; Howarth, R.W.; Likens, G.E.; Matson, P.A.; Schindler, D.W.; Schlesinger, W.H.; Tilman, D.G. (1997) Human alteration of the global nitrogen cycle: sources and consequences. Ecological Applications 7:737-750

Weigel D (2012) Natural variation in Arabidopsis: from molecular genetics to ecological genomics. Plant Physiology 158:2-22.

Zeng ZB (1994) Precision mapping of quantitative trait loci. Genetics 136:1457-1468.

Zhang L, Kars I, Essenstam B, Liebrand TW, Wagemakers L, Elberse J, Tagkalaki P, Tjoitang D, van den Ackerveken G, van Kan JA (2014) Fungal endopolygalacturonases are recognized as microbe-associated molecular patterns by the arabidopsis receptor-like protein RESPONSIVENESS TO BOTRYTIS POLYGALACTURONASES1. Plant Physiology 164:352-364.

Chapter 5

General Discussion

Improving nitrogen (N) use efficiency (NUE) in crops is one of the crop management strategies for a sustainable agriculture. This thesis focuses on the genetic structure of NUE and investigates it using the knowledge of natural and genetic variation in *Arabidopsis thaliana. A. thaliana* is responsive to changing N conditions (Masclaux-Daubresse and Chardon 2011; Chardon et al. 2010). Retarded shoot growth, leaf senescence, anthocyanin accumulation, and reduced N accumulation and water content were observed in my screenings as the strategies for *A. thaliana* to cope with the sub-optimal N condition. Phenotyping shoot traits under sub-optimal and optimal N conditions, investigating the phenotypic plasticity across the changing N conditions, genetically dissecting NUE, NUE related agronomic traits (NRAT) (shoot dry & fresh weights (SDW & SFW), N usage index (NUI), N concentration (N%), carbon concentration (C%), Photosystem II (ΦPSII), water content (WC), relative growth rate (RGR), and total leaf area (TLA)), and phenotypic plasticity (Res) were addressed throughout the study to understand the mechanism(s) underlying genetics of NUE in *A. thaliana*.

Natural genetic variation and phenotypic plasticity in *A. thaliana*

Knowledge of natural genetic variation is necessary to understand the genetics of NUE in plants (Stahl et al. 2016; Bouchet et al. 2016). However, domestication practices in agriculture have selected cultivars for high yield production under adequate N conditions. Alleles associated with traits adapted to N limited conditions might, therefore, be missing in domesticated crops (Hawkesford 2017; Ren et al. 2018). Therefore, investigating genetic mechanism(s) of NUE in crops initially needs a wide-range of genetic variation, including landraces and wild relatives that covers all alleles characterizing NUE and NRAT within a species.

A. thaliana is a dicot plant growing in differing habitats in Eurasia, North America and Africa (Durvasula et al. 2017; Koornneef et al. 2004). In Chapter 2, a population of natural accessions (HapMap population) consisting of 354 genotypes of *A. thaliana* (Li et al. 2010) was screened and this screen showed a wide range of phenotypic variation for NUE and NRAT, controlled by genetic variation in the population. It provided (I) phenotypic data to study the genetics of NUE and NRAT through a genome-wide association study (GWAS) (Chapter 2); (II) a sub-set of HapMap population including extreme *A. thaliana* accessions to deeply understand correlations among NUE, NRAT, and phenotypic plasticity (Chapter 3); (III) phenotypic data

of a RIL population generated from two distinct accessions (Br-0 and Col-0) to dissect genetic mechanism of NUE, NRAT, and phenotypic plasticity through bi-parental linkage mapping.

Natural variation in *A. thaliana* was previously detected and studied for NUE (Chardon et al. 2010; Menz et al. 2018). Chardon et al. (2010) and Menz et al. (2018) both agreed that more N supply generally increases plant biomass production. The data from this thesis on average showed a similar phenotypic response by changing N supply (Chapter 2, 3, and 4). However, high N supply was not a good indicator for the adaptive behavior of *A. thaliana* to the sub-optimal N condition. Under N limited conditions a higher genetic variation in NUE was observed compared to normal conditions (Chardon et al. 2010). In agreement with Chardon et al. (2010), I observed a higher genetic variation in N% and NUE under the sub-optimal N conditions than under the optimal conditions in Chapter 2 and 4. Heritability values of N% and NUE (Chapter 2: 0.80 and 0.75; Chapter 4: 0.45 and 0.50, respectively) were higher under sub-optimal N condition than the optimal condition. This indicated that around half or more than half of the phenotypic variation in N% and NUE was controlled by genes under sub-optimal N condition.

In Chapter 3, phenotypic responses were investigated in a subset of the HapMap population consisting of 20 extreme *A. thaliana* accessions. Ven-1, Br-0, Paw-3, PHW-36, NFA-8, TOU-A1-67, Wt-3, TOU-I-6, Sav-0, and WAR were high in NUE; CIBC2, Kelsterbach-2, Ag-0, Aa-0, Bur-0, CIBC17, Uk-1, Sei-0, Ca-0, and Col-0 were low in NUE. These accessions had very similar SDW values under sub-optimal N condition, although they differed in NUE. The purpose was to confirm extreme accessions in HapMap population and characterize their behavior under sub-optimal N condition. Results revealed very interesting trait correlations. SFW, WC, and N% were positively correlated under sub-optimal N conditions. As N is an important determinant of water uptake, this positive relation between WC and N uptake might explain a good water and mineral absorbing capability of a developed root system under sub-optimal N conditions (Easlon and Bloom 2013; Ding et al. 2018). Although water and N are two growth limiting factors for plants, accessions with high NUE produced a similar SDW as those with low NUE. Starch accumulation in high NUE plants might provide the same SDW production. Plants, under prolonged N limitation, are adapted to produce more glucose, sucrose, and starch, instead of amino acid and protein synthesis, than those growing under optimal conditions (Robinson 1997). In order to identify the underlying genes responsible for increasing WC along with N uptake and increasing starch accumulation along with NUE, in Chapters 2 and 4, I examined both the HapMap and a RIL population generated from two extreme

accessions (Col-0 (low NUE) and Br-0 (high NUE)). I will discuss these results further under the "Candidate genes" segment of this chapter.

Plants adapt to their changing environments by modifying their phenotypes (El-Soda et al. 2014). Genes regulating phenotypic plasticity can sustain the adaptive responses to changing environmental conditions. Phenotypic plasticity was studied in Chapters 2, 3, and 4 by the analysis of the residuals between sub-optimal and optimal N conditions (Becker 1981; Eberhart and Russell 1966; Rauh et al. 2002). The HapMap population showed a larger phenotypic variation in SDW_Res, SFW_Res, WC_Res, and NUI_Res than in NUE_Res and N%_Res (Chapter 2). Therefore, the HapMap population was highly responsive to changing N conditions. A deep phenotyping was conducted on phenotypic plasticity using a subset of the HapMap population in Chapter 3. Phenotypic plasticity assessment of 20 extreme accessions resulted in eight contrasting but highly plastic accessions. Br-0, Ven-1, Paw-3, NFA-8, Col-0, Bur, Uk-1, and CIBC17 showed high phenotypic plasticity for N%_Res, NUE_Res, and NUI_Res. Natural variation in phenotypic plasticity was studied before (Chardon et al. 2010; Menz et al. 2018; North et al. 2009; Meyer et al. 2019). North et al. (2009) and Chardon et al. (2010) applied sub-optimal nitrate supply (1mM nitrate on agar and 2mM nitrate on soil, respectively) to different *A. thaliana* accessions, whereas Menz et al. (2018) and Meyer et al. (2019) applied very severe N deficiency (0.4mM ammonium nitrate and 0.4mM nitrate) on agar plate. They all included Col-0 in their set of different accessions and all concluded that Col-0 was not responsive in terms of biomass gain to increased N supply. In agreement with them, Col-0 in this thesis did not show a high plasticity in SFW and SDW (Chapter 3). Chardon et al. (2010) also found that Col-0 did not show any phenotypic plasticity in N% under two N supplies, whereas North et al. (2009) showed a significant decrease in total N content in Col-0 under sub-optimal N condition. Menz et al. (2018) stated that Col-0 was more responsive to ammonium supply than nitrate as a sole N source. Shoot N content in Col-0 increased more with increased ammonium nitrate than only nitrate supply. In this thesis, N was always applied in the form of ammonium nitrate and shoot N% of Col-0 was found very responsive to changing environment (Chapter 3). The varying results for Col-0 might be due to different growth conditions. These results infer that ranking *A. thaliana* accessions might vary depending on the growth condition due to strong genotype by environment interaction. The phenotypic predictability of plastic traits could, therefore, be improved with subsequent trials under different environmental conditions. In order to substantiate the root plasticity to changing environmental conditions, root visualization and quantification should be conducted in these

trials. It was well established that adaptation to different environmental conditions is regulated by different inter- and intra-specific genetic mechanisms (Meyer et al. 2019; Bi et al. 2007; Rauh et al. 2002; Loudet et al. 2003b). Therefore, a genome-wide transcriptome analysis or QTL mapping may provide the causal genetic variants involved in adaptational responses to changing N conditions. Consequently, a deeper comprehension would be achieved for accessions of *A. thaliana*.

Experimental design and mapping strategies

NUE and NRAT are known as complex traits regulated by multiple genes each explaining only part of the phenotypic variation. The genetic dissection of complex traits requires accurate phenotyping methodologies (Ingvarsson and Street 2011; Han et al. 2015). Experimental designs greatly affect the accuracy of phenotyping, therefore I grew multiple replicates per genotype in a randomized block design within every treatment to account for any heterogeneity. Innovations in agricultural bio-mechatronics improve high-throughput phenotyping methodologies in order to speed up experiments and to increase accuracy and precision of phenotyping for complex traits. Therefore, we used a high-throughput phenotyping platform and digital phenotyping methodologies (Flood et al. 2016) to detect ΦPSII , TLA, and RGR in Chapter 2 and 4.

Selecting the correct mapping methodology was as crucial as accurate phenotyping in order to detect small effect quantitative trait loci (QTLs). Researchers have identified a number of QTLs associated with NUE using either population-based association mapping or family-based linkage mapping for maize, rice and wheat (Morosini et al. 2017; Abdel-Ghani et al. 2015; Zhou et al. 2017; Liu et al. 2016; Fan et al. 2018; Ren et al. 2018; Monostori et al. 2017). These studies compared QTLs with formerly identified loci or genes associated with NUE and NRAT. In this thesis, I performed two different QTL mapping approaches: GWAS, where the genetic dissection of NUE and NRAT of the HapMap population was performed, and conventional bi-parental population linkage mapping, where a Br-0 x Col-0 recombinant inbred line (RIL) population was used for the genetic analysis.

GWAS

GWAS is based on linkage disequilibrium (LD) mapping and LD distances decay approximately after 10kb in *A. thaliana* (Brachi et al. 2010; Kim et al. 2007; Atwell et al. 2010). Along with the high number of recombination events in the HapMap population, GWAS provided me a high-resolution mapping using 250K SNP markers (Li et al. 2010). The population structure might, however, result in false positive SNPs due to shared ancestry of population members. False positives are caused by inflation of significance of SNPs which are not actually associated with the trait of interest (Mitchell-Olds 2010; Atwell et al. 2010). Therefore, the efficient mixed model association expedited (EMMAX), developed by Kang et al. (2010), was used to correct for population structure in Chapter 2 (Kang et al. 2010). EMMAX is a time efficient implementation in GWAS to correct pairwise relatedness by using high density molecular markers.

Adjusting the significance threshold was another correction method applied to eliminate false positives in GWAS results. Bonferroni correction is often used to eliminate false positives in GWAS. However, the Bonferroni significance threshold (-log(*p*-value)=6.5) is quite stringent for complex traits like NUE and NRAT that are regulated by multiple small effect QTLs (Atwell et al. 2010; El-Soda et al. 2015). An enrichment of *a priori* candidate genes was previously reported with a significance threshold arbitrarily set to -log(*p*-value)=4 (Li et al. 2010; Atwell et al. 2010; Kooke et al. 2016). In my study, a number of SNPs showed a significance of -log(*p*-value)>4 for association with variation in response to sub-optimal N condition. These SNPs were found to be in linkage with several candidate genes, of which most are likely to be false positives and additional confirmation methods must be performed to verify the true positives. The linkage drag indicates linked SNP markers within a LD window. First of all, I selected putative QTLs based on the appearance of multiple SNP markers above the arbitrary threshold. Then, I listed all candidate genes within a proximity of 10kb near the most significant SNP marker. The promising candidate genes were then checked according to the evidences in relevant literature to see if they play an important role in the N responsive regulatory pathways.

As mentioned before, the HapMap population provided phenotypic data for GWAS (Chapter 2). The large phenotypic variation, however, could confound the power of GWAS due to multigenic regulation, low minor allele frequencies and allelic heterogeneity (Meyer et al. 2019). Bi-parental populations, with lower genetic variation, might improve the power of QTL detection of such complex traits. The combination of both mapping approaches, i.e. high resolution from GWAS and high detection power from bi-parental linkage mapping might increase the accuracy of finding causal alleles. Therefore, another confirmation method of

GWAS results was the combination with bi-parental linkage mapping outputs to identify co-localization of putative loci in different populations (Morosini et al. 2017). This approach was not before applied in the genetic analysis of NUE in *A. thaliana*. Whereas, QTLs were more often verified through finding a co-localization with a previously confirmed gene (Rauh et al. 2002; Loudet et al. 2003a; Loudet et al. 2003b; Diaz et al. 2006).

Bi-parental Linkage mapping

Br-0 and Col-0 were selected from the HapMap population because of their contrasting NUE characteristics. In a number of subsequent experiments in Chapter 3, Br-0 expressed high NUE and Col-0 expressed low NUE even though they yielded equal amounts of SDW under sub-optimal N condition. The linkage mapping of the Br-0 and Col-0 RIL population was conducted to identify co-localizations of QTLs for GWAS and bi-parental linkage mapping and novel genetic mechanisms associated with NUE and NRAT in Chapter 4. However, in a linkage mapping, the genetic variation of pedigrees is always limited to the genetic variation of parents, so fewer molecular markers are required for mapping the bi-parental populations due to limited recombination (makes higher LD) (Mitchell-Olds 2010). This reduces the genetic resolution and accuracy of mapping positions in the Br-0 and Col-0 RIL.

Both GWAS and the linkage mapping confirmed the complexity of NUE, NRAT, and phenotypic plasticity by revealing many putative small-effect loci. I identified 28 loci in GWAS associated with SDW, SFW, NUE, N%, NUI, TLA, ΦPSII, RGR, and WC under sub-optimal N condition and phenotypic plasticity (SDW_Res, NUI_Res, N%_Res, NUE_Res, WC_Res, and ΦPSII_Res). The fraction of explained phenotypic variation ranged between 6% and 49.86% for all QTLs. Five QTLs were identified from the linkage mapping of SFW and NUI under sub-optimal N condition, where proportions of explained phenotypic variation ranged around 8%. This showed that NUE, NRAT, and phenotypic plasticity were regulated by small effect QTLs, which explained only a part of the phenotypic variation in both populations.

QTL mapping were performed by Rauh et al. (2002), Loudet et al. (2003), Diaz et al. (2006) to find QTLs associated with adaptive responses in *A. thaliana* to low N supplies (Rauh et al. 2002; Loudet et al. 2003a; Loudet et al. 2003b; Diaz et al. 2006). Rauh et al. (2002) used Col-0 and Ler RIL population for detecting QTLs associated with the adaptation to differing N sources (0mM N (control), 16mM ammonium sulfate, nitrate, and ammonium nitrate) (Rauh et al. 2002). Loudet et al (2003a and 2003b) used Bay-0 and Sha RIL population for QTLs

associated with shoot dry mass, N%, amino acid, water, and chloride content under low and optimal N conditions (3mM and 10mM nitrate, respectively). They also looked for QTLs associated with nitrate content under the optimal and phosphate content under the low N conditions (Loudet et al. 2003a; Loudet et al. 2003b). Diaz et al. (2006) used the same Bay-0 and Sha RIL population under similar N supply conditions and checked for QTLs associated with leaf yellowing and anthocyanin accumulation (Diaz et al. 2006). Co-localizations of QTLs were evidently detected between Loudet et al. (2003b) and Diaz et al. (2006), since they assessed highly correlating traits (shoot dry biomass, N%, amino acid content, and leaf yellowing) in the same RIL population grown in similar conditions. However, no co-localization with QTLs of Rauh et al. (2002) was detected in these studies, and neither in this thesis (Chapter 2 and 4). The lack of co-localized QTLs might be due to the large phenotypic variation in NUE and NRAT in different *A. thaliana* accessions that is regulated by varying genetic mechanisms (Meyer et al. 2019). Therefore, it would be beneficial using mixed populations created by inter-crossing parents to increase the power of QTLs for such complex traits.

Candidate genes

Throughout this thesis, I focused on the effect of N limitation on plant growth, N accumulation, NUE, and photosynthesis related traits. Applying GWAS and linkage mapping, I identified plant phytohormones, transporters, transporter facilitators, cell wall structure, carbon metabolism and photosynthesis related candidate genes. Some candidates were already identified in *A. thaliana*, but most of them were novel. For instance, *MIR160* (AT5G46845) (Liang et al. 2012; Nguyen et al. 2015), Sec14p-like phosphatidylinositol transfer family protein encoding gene (AT1G22180), Major facilitator superfamily protein encoding gene (AT5G10190), *SULTR4;1* (AT5G13550), *SULTR2;1* (AT5G10180) (Peng et al. 2007), *PGM* (AT1G22170) (Peng et al. 2007; Wang et al. 2003; Wang et al. 2004) and *CESA4* (AT5G44030) (Guevara et al. 2014) were identified through mutant and/or genome-wide transcriptomic analysis and very responsive to N limitation.

GWAS and linkage mapping revealed three key candidate genes, which might regulate WC and SDW under sub-optimal N condition. GWAS showed a QTL on chromosome 5, associated with variation in WC under sub-optimal N conditions, which was in LD with *MIR160* (Chapter 2). This gene was previously identified in *A. thaliana.* Studies showed that *MIR160* encoded small RNA fragments (microRNAs) to down-regulate expression of *Auxin Response Factors (ARF)*

10, 16, and *17*. Therefore, lateral root growth was increased upon the expression of *MIR160* under N-limited conditions (Liang et al. 2012; Nguyen et al. 2015). Its association with variation in WC seems to be based on an up-regulation of adaptation mechanisms to the sub-optimal N condition in order to absorb water and N through elongated lateral root system.

Another aforementioned interesting phenotype was high SDW production with minimum water and N-uptake in accessions with high NUE. Both mapping approaches indicated the association of this phenotype with candidate genes involved in carbon metabolism. GWAS revealed a QTL on chromosome 1, associated with SDW, NUI, and SDW_Res. This QTL co-located with a gene encoding phosphoglucomutase (PGM), a glycolytic enzyme in energy metabolism to finally produce amino acids and proteins for a healthy plant growth under optimal environmental conditions (Wang et al. 2003; Robinson 1997). Under N-limited conditions, carbon skeletons produced by photosynthesis are stored as starch and sucrose in plant leaves instead of consumed for amino acid production (Robinson 1997). Regarding the results of Chapter 2 and 3 *PGM* might be the causal gene for high SDW production of accessions with high NUE. Finally, the outcomes of GWAS and linkage mapping revealed a common locus co-locating with the cell wall structure regulating gene *CESA4* (Ramirez et al. 2009; Hernandez-Blanco et al. 2007; Guevara et al. 2014). *CESA4* was previously reported in a study on *A. thaliana* and rice, of which the down-regulation impaired the carbohydrate deposition onto the cell wall under N limitation (Guevara et al. 2014). This candidate gene was associated to both SFW under sub-optimal N condition and WC_Res and it might affect the adaptational response of plants WC under sub-optimal N condition (Chapter 2 and 4).

Further experiments would be to test the function of candidate genes using allelic variants and validate their effect through differential gene expression analysis using the extreme 20 accessions of *A. thaliana* identified in Chapter 3.

The genetic analysis of NUE in *Arabidopsis thaliana*

The genetic analysis of NUE in *A. thaliana* was studied before by others using populations based on Columbia (Col-0) x Landsberg *erecta* (Ler), by Rauh et al. (2002), and Bay-0 x Shahdara by Loudet et al. (2003a, 2003b) and Diaz et al. (2006) (Rauh et al. 2002; Loudet et al. 2003a; Loudet et al. 2003b; Diaz et al. 2006). Transcriptome analysis in *Arabidopsis thaliana* showed that several genes were up/down-regulated under limited N conditions. A number of genes involved in N-uptake, translocation, assimilation, and remobilization have been characterized (Wang et al. 2003; Bi et al. 2007; Peng et al. 2007; Li et al. 2017; Meyer et al.

2019). This thesis introduces two novel populations to the field and investigated novel genetic mechanisms of NUE, NRAT, and the phenotypic plasticity in *A. thaliana* by combining two mapping strategies. These findings will provide new alleles for the plant breeding, especially for *Brassica* breeding, because *A. thaliana* is a member of Brassicaceae family and closely related to *Brassica* species. An identified NUE or NRAT associated locus in *A. thaliana* will indicate orthologous loci within *Brassica* genome using comparative genome analysis (Schranz et al. 2006). This provides a great opportunity for *Brassica* crop species to improve NUE.

Conclusion and recommendations

The results presented in this thesis explain the great potential of the exploration of natural genetic variation in *A. thaliana* for NUE improvement. For future experiments, root visualization and quantification, quantifying differential gene expression, recording the effect of promising alleles through genetic transformation to a null-allele background, and testing the function of candidate genes through transgenic *A. thaliana* will be additional steps to confirm the results mentioned in this thesis. Finally, the collected knowledge, resulting from genetic analysis and phenotyping, will further help agronomists to use the available information for a sustainable NUE improvement in crops.

References

Abdel-Ghani AH, Kumar B, Pace J, Jansen C, Gonzalez-Portilla PJ, Reyes-Matamoros J, San Martin JP, Lee M, Lubberstedt T (2015) Association analysis of genes involved in maize (Zea mays L.) root development with seedling and agronomic traits under contrasting nitrogen levels. Plant Molecular Biology 88:133-147.

Atwell S, Huang YS, Vilhjalmsson BJ, Willems G, Horton M, Li Y, Meng D, Platt A, Tarone AM, Hu TT, Jiang R, Muliyati NW, Zhang X, Amer MA, Baxter I, Brachi B, Chory J, Dean C, Debieu M, de Meaux J, Ecker JR, Faure N, Kniskern JM, Jones JD, Michael T, Nemri A, Roux F, Salt DE, Tang C, Todesco M, Traw MB, Weigel D, Marjoram P, Borevitz JO, Bergelson J, Nordborg M (2010) Genome-wide association study of 107 phenotypes in Arabidopsis thaliana inbred lines. Nature 465:627-631.

Becker HC (1981) Correlations among some statistical measures of phenotypic stability. Euphytica 30:835-840.

Bi YM, Wang RL, Zhu T, Rothstein SJ (2007) Global transcription profiling reveals differential responses to chronic nitrogen stress and putative nitrogen regulatory components in Arabidopsis. BMC genomics 8:281.

Bouchet AS, Laperche A, Bissuel-Belaygue C, Snowdon R, Nesi N, Stahl A (2016) Nitrogen use efficiency in rapeseed. A review. Agronomy for Sustainable Development 36.

Brachi B, Faure N, Horton M, Flahauw E, Vazquez A, Nordborg M, Bergelson J, Cuguen J, Roux F (2010) Linkage and association mapping of Arabidopsis thaliana flowering time in nature. PLoS Genetics 6:e1000940.

Chardon F, Barthelemy 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.

Diaz C, Saliba-Colombani V, Loudet O, Belluomo P, Moreau L, Daniel-Vedele F, Morot-Gaudry JF, Masclaux-Daubresse C (2006) Leaf yellowing and anthocyanin accumulation are two genetically independent strategies in response to nitrogen limitation in Arabidopsis thaliana. Plant & Cell Physiology 47:74-83.

Ding L, Lu ZF, Gao LM, Guo SW, Shen QR (2018) Is Nitrogen a Key Determinant of Water Transport and Photosynthesis in Higher Plants Upon Drought Stress? Frontiers in Plant Sciences 9.

Durvasula A, Fulgione A, Gutaker RM, Alacakaptan SI, Flood PJ, Neto C, Tsuchimatsu T, Burbano HA, Pico FX, Alonso-Blanco C, Hancock AM (2017) African genomes illuminate the early history and transition to selfing in Arabidopsis thaliana. Proceedings of the National Academy of Sciences of the United States of America 114:5213-5218.

Easlon HM, Bloom AJ (2013) The effects of rising atmospheric carbon dioxide on shoot-root nitrogen and water signaling. Frontiers Plant Sciences 4.

Eberhart AS, Russell WA (1966) Stability parameters for comparing varieties. Crop Science 6:36-40.

El-Soda M, Kruijer W, Malosetti M, Koornneef M, Aarts MG (2015) Quantitative trait loci and candidate genes underlying genotype by environment interaction in the response of Arabidopsis thaliana to drought. Plant, Cell & Environment 38:585-599.

El-Soda M, Malosetti M, Zwaan BJ, Koornneef M, Aarts MG (2014) Genotypexenvironment interaction QTL mapping in plants: lessons from Arabidopsis. Trends in Plant Science 19:390-398.

Fan X, Zhang W, Zhang N, Chen M, Zheng S, Zhao C, Han J, Liu J, Zhang X, Song L, Ji J, Liu X, Ling H, Tong Y, Cui F, Wang T, Li J (2018) Identification of QTL regions for seedling root traits and their effect on nitrogen use efficiency in wheat (Triticum aestivum L.). Theoretical and Applied Genetics 131:2677-2698.

Flood PJ, Kruijer W, Schnabel SK, van der Schoor R, Jalink H, Snel JFH, Harbinson J, Aarts MGM (2016) Phenomics for photosynthesis, growth and reflectance in Arabidopsis thaliana reveals circadian and long-term fluctuations in heritability. Plant Methods 12.

Guevara DR, El-Kereamy A, Yaish MW, Mei-Bi Y, Rothstein SJ (2014) Functional characterization of the rice UDP-glucose 4-epimerase 1, OsUGE1: A potential role in cell wall carbohydrate partitioning during limiting nitrogen conditions. PloS one 9:e96158.

Han M, Okamoto M, Beatty PH, Rothstein SJ, Good AG (2015) The Genetics of nitrogen use efficiency in crop plants. Annual Review of Genetics 49:269-289.

Hawkesford MJ (2017) Genetic variation in traits for nitrogen use efficiency in wheat. Journal of Experimental Botany 68:2627-2632.

Hernandez-Blanco C, Feng DX, Hu J, Sanchez-Vallet A, Deslandes L, Llorente F, Berrocal-Lobo M, Keller H, Barlet X, Sanchez-Rodriguez C, Anderson LK, Somerville S, Marco Y, Molina A (2007) Impairment of cellulose synthases required for Arabidopsis secondary cell wall formation enhances disease resistance. The Plant Cell 19:890-903.

Ingvarsson PK, Street NR (2011) Association genetics of complex traits in plants. The New Phytologist 189:909-922.

Kang HM, Sul JH, Service SK, Zaitlen NA, Kong SY, Freimer NB, Sabatti C, Eskin E (2010) Variance component model to account for sample structure in genome-wide association studies. Nature genetics 42 (4):348-354. doi:10.1038/ng.548

Kim S, Plagnol V, Hu TT, Toomajian C, Clark RM, Ossowski S, Ecker JR, Weigel D, Nordborg M (2007) Recombination and linkage disequilibrium in Arabidopsis thaliana. Nature Genetics 39:1151-1155.

Kooke R, Kruijer W, Bours R, Becker F, Kuhn A, van de Geest H, Buntjer J, Doeswijk T, Guerra J, Bouwmeester H, Vreugdenhil D, Keurentjes JJ (2016) Genome-Wide Association Mapping and Genomic Prediction Elucidate the Genetic Architecture of Morphological Traits in Arabidopsis. Plant Physiology 170:2187-2203.

Koornneef M, Alonso-Blanco C, Vreugdenhil D (2004) Naturally occurring genetic variation in Arabidopsis thaliana. Annual Review of Plant Biology 55:141-172.

Li H, Hu B, Chu CC (2017) Nitrogen use efficiency in crops: lessons from Arabidopsis and rice. Journal of Experimental Botany 68:2477-2488.

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 of the United States of America 107:21199-21204.

Liu ZY, Zhu CS, Jiang Y, Tian YL, Yu J, An HZ, Tang WJ, Sun J, Tang JP, Chen GM, Zhai HQ, Wang CM, Wan JM (2016) Association mapping and genetic dissection of nitrogen use efficiency-related traits in rice (Oryza sativa L.). Functional and Integrative Genomics 16:323-333.

Loudet O, Chaillou S, Krapp A, Daniel-Vedele F (2003a) Quantitative trait loci analysis of water and anion contents in interaction with nitrogen availability in Arabidopsis thaliana. Genetics 163:711-722.

Loudet O, Chaillou S, Merigout P, Talbotec J, Daniel-Vedele F (2003b) Quantitative trait loci analysis of nitrogen use efficiency in Arabidopsis. Plant Physiology 131:345-358.

Masclaux-Daubresse C, Chardon F (2011) Exploring nitrogen remobilization for seed filling using natural variation in Arabidopsis thaliana. Journal of Experimental Botany 62:2131-2142.

Menz J, Range T, Trini J, Ludewig U, Neuhauser B (2018) Molecular basis of differential nitrogen use efficiencies and nitrogen source preferences in contrasting Arabidopsis accessions. Scientific Reports 8.

Meyer RC, Gryczka C, Neitsch C, Muller M, Brautigam A, Schlereth A, Schon H, Weigelt-Fischer K, Altmann T (2019) Genetic diversity for nitrogen use efficiency in Arabidopsis thaliana. Planta.

Mitchell-Olds T (2010) Complex-trait analysis in plants. Genome biology 11:113.

Monostori I, Szira F, Tondelli A, Arendas T, Gierczik K, Cattivelli L, Galiba G, Vagujfalvi A (2017) Genome-wide association study and genetic diversity analysis on nitrogen use efficiency in a Central European winter wheat (Triticum aestivum L.) collection. PloS One 12.

Morosini JS, Mendonca LD, Lyra DH, Galli G, Vidotti MS, Fritsche-Neto R (2017) Association mapping for traits related to nitrogen use efficiency in tropical maize lines under field conditions. Plant and Soil 421:453-463.

North KA, Ehlting B, Koprivova A, Rennenberg H, Kopriva S (2009) Natural variation in Arabidopsis adaptation to growth at low nitrogen conditions. Plant Physiol Bioch 47:912-918.

Peng M, Bi YM, Zhu T, Rothstein SJ (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.

Ramirez V, Coego A, Lopez A, Agorio A, Flors V, Vera P (2009) Drought tolerance in Arabidopsis is controlled by the OCP3 disease resistance regulator. The Plant journal : for cell and molecular biology 58:578-591.

Rauh L, Basten C, Buckler St (2002) Quantitative trait loci analysis of growth response to varying nitrogen sources in Arabidopsis thaliana. Theoretical and Applied Genetics 104:743-750.

Ren DQ, Fang XJ, Jiang P, Zhang GX, Hu JM, Wang XQ, Meng Q, Cui WA, Lan SJ, Ma X, Wang HW, Kong LR (2018) Genetic architecture of nitrogen-deficiency tolerance in wheat seedlings based on a nested association mapping (NAM) population. Frontiers in plant science 9.

Robinson JM (1997) Nitrogen limitation of spinach plants causes a simultaneous rise in foliar levels of orthophosphate, sucrose, and starch. Internationa Journal of Plant Sciences 158:432-441.

Schranz ME, Lysak MA, Mitchell-Olds T (2006) The ABC's of comparative genomics in the Brassicaceae: building blocks of crucifer genomes. Trends in Plant Science 11:535-542.

Stahl A, Friedt W, Wittkop B, Snowdon RJ (2016) Complementary diversity for nitrogen uptake and utilisation efficiency reveals broad potential for increased sustainability of oilseed rape production. Plant and Soil 400:245-262.

Wang R, Okamoto M, Xing X, Crawford NM (2003) Microarray analysis of the nitrate response in Arabidopsis roots and shoots reveals over 1,000 rapidly responding genes and new linkages to glucose, trehalose-6-phosphate, iron, and sulfate metabolism. Plant Physiology 132:556-567.

Wang R, Tischner R, Gutierrez RA, Hoffman M, Xing X, Chen M, Coruzzi G, Crawford NM (2004) Genomic analysis of the nitrate response using a nitrate reductase-null mutant of Arabidopsis. Plant Physiology 136:2512-2522.

Zhou Y, Tao YJ, Tang DN, Wang J, Zhong J, Wang Y, Yuan QM, Yu XF, Zhang Y, Wang YL, Liang GH, Dong GC (2017) Identification of QTL associated with nitrogen uptake and nitrogen use efficiency using high throughput genotyped CSSLs in rice (Oryza sativa L.). Frontiers in Plant Science 8.

Summary in English:

Nitrogen (N) is an essential macro-nutrient for plant to sustain a healthy growth and reproduction. Its deficiency generally causes retarded shoot growth, increased lateral root, impaired photosynthesis, chlorosis, anthocyanin accumulation, and early flowering in plants. N use efficiency (NUE) is the ability of a plant to grow and to reproduce under sub-optimal N conditions. NUE is defined as the amount of yield produced per N content in the plant. NUE, growth and photosynthesis related traits are highly responsive to sub-optimal N conditions. These are complex traits controlled by multiple genes. The natural genetic variation in *Arabidopsis thaliana* consists of allelic variants associated with these traits. Quantitative trait locus (QTL) mapping is an appropriate strategy for genetically dissecting NUE, growth, and photosynthesis related traits under sub-optimal N condition. The main objectives of this thesis are: (i) shedding light on the genetic mechanisms that explain NUE; (ii) understanding what NUE means for plants by using *A. thaliana* as a model organism; (iii) assessing the applicability of NUE for crop improvement. To achieve these objectives, shoot traits of *A. thaliana* were phenotyped, the collected phenotyping data were used to map causal alleles associated with NUE.

In Chapter 2, I screened a population of natural accessions (HapMap population) consisting of 354 diverse genotypes of *A. thaliana* under sub-optimal and optimal N condition. I phenotyped shoot dry and fresh weight (SDW and SFW), shoot N concentration (N%), total leaf area (TLA), and Photosystem II efficiency (ΦPSII) and calculated NUE, N usage index, water content (WC), and relative growth rate. To understand the phenotypic plasticity across the two conditions, I also calculated residual values of every genotype for every trait. Genome-wide association study (GWAS) was conducted to unravel single nucleotide polymorphisms (SNPs) associated with these traits. 28 QTLs were identified which were enriched for functions in hormone signalling, transmembrane transport, and carbon metabolism. These results proposed novel candidate genes involved in plant growth and developmental mechanisms under sub-optimal N condition.

In Chapter 3, I selected 20 extreme accessions of *Arabidopsis thaliana* from HapMap population and phenotypically analysed them under sub-optimal N condition. These accessions had equal SDW, however, 10 had very low NUE and 10 had very high NUE. These two clusters were named as N-use-efficient (NE) and N-use-inefficient (NI) accessions. They were screened under sub-optimal and optimal N conditions and phenotyped for SFW, SDW, WC, RGR, TLA, N%, NUE, NUI, and ΦPSII. The residual values exhibited the phenotypic plasticity of every

genotype across the two environments. High NUE, NUI, and RGR make NE accessions more efficient in N assimilation and utilization with less N uptake. Br-0 and Col-0 were identified as two contrasting genotypes belonging to NE and NI accessions, respectively. These results suggested that the genetic variation between Br-0 and Col-0 might explain the genetic structure of NUE using a bi-parental population linkage mapping.

In Chapter 4, referring to the Chapter 3, an F8 recombinant inbred line (RIL) population of Br-0 and Col-0 was genotyped and screened under sub-optimal and optimal N conditions. Bi-parental population linkage and composite interval mapping were applied to find significant QTLs associated with SFW, SDW, N%, NUI, NUE, C%, TLA, and their residuals. Seven loci in different genetic locations were identified. Non-synonymous amino acid changes were detected in three loci. One QTL co-locating *CESA4* was also detected significant in GWAS results. Br-0 and Col-0 is a novel bi-parental population in the study of the genetic analysis of NUE in *A. thaliana*, which might suggest novel alleles associated with NUE.

The study presented in this thesis explores the genetic mechanisms of NUE in *A. thaliana* using novel approaches in the field. The findings of experimental chapters will contribute to crop NUE improvement for a sustainable agriculture.

Acknowledgements

The decision to come to the Netherlands for a PhD research was a monumental turning point in my life. Looking back to this journey makes me realize how it transformed me into a more mature, responsible and dedicated scientist. It was a transformative personal experience. I experienced both success and challenges throughout this journey. I always felt to be loved and supported, especially by my husband, Emre Erol, my family, and my friends. In bringing this thesis into completion, I am thankful to all of them for always being with me.

I would like to express my gratitude to my promoters Professor Mark GM Aarts and Professor Joost JB Keurentjes and my co-promoter Professor Maarten Koornneef. Their caring guidance and helpful suggestions were essential in developing this dissertation. Our meetings and subsequent discussions were very useful to me. Like most researchers, I encountered challenges during my PhD study in Wageningen University, and I must thank both my promoters and co-promoter for consistently encouraging and assisting me in the research directions without allowing me to lose my focus.

My research experience in Wageningen University would not have been nearly as enjoyable without the company of Dr. Mohamed El-Soda, Dr. Pingping Huang, Dr. Christina May, and Dr. Aina E Prinzenberg. I owe them a heartful of thanks for supporting me and cheering me up. Our relationship has transpired beyond sharing a lab bench and transformed into lasting friendship.

Throughout my scientific journey, I was lucky to have friends who helped me in various ways by sharing my stress, uplifting my spirit and encouraging my intellectual curiosities. Alaettin Çarıkçı, Maurits Nikkels, Fleur Van Odylle, Başak Barış, Başak Tektemur Altay, Tsolin Nalbantian, Eftychia Mylona, Gaye Eksen, Berk Çallı, Yunus Sarıçay, and Uğur Aktay, your friendship has been my greatest source of motivation when I needed to cope with difficulties of work, life and the notoriously cloudy and dark days in the Netherlands. Nalan Liv and Caner Hamarat, especially you guys... You became my family. I cannot imagine a life without you. I feel so lucky to have you. Thanks for being with me, Emre and Deniz Meriç in good times and bad times!

My family, Funda, Ethem, Didem Öztolan (Arslan), Selcan and Necati Erol always cared about my progress, fostered my interest, and gave me their full backing. I am indebted to them for all

their care and support. I would like to specially thank my father Ethem Öztolan for his progressive thoughts about my education life which started with Saint Joseph and continued with Sabancı University. You gave a lot to me against all odds. Thank you!

Finally, my beloved Emre Erol! You are the reason of all of these my dear! You influence me in so many ways, your high energy and motivation always give me the will, the patience, and the endurance in my life. You and our son Deniz Meriç are my biggest source of motivation that helps me to succeed in this endeavor. Sizi çok seviyorum, hep benimle kalın!

About the author

Nihal Öztolan Erol was born on 11th of November, 1983 in Izmir (Turkey). She studied a major of Biological Sciences and Bioengineering in Sabancı University, Istanbul (Turkey) between 2002 and 2007. She completed her master thesis in Sabancı University under the supervision of Professor İsmail Çakmak in 2009. The topic was the investigation of anti-oxidant capacity of selenium enriched wheat extract on mammalian cells. Afterwards, she moved to the Netherlands and applied for a PhD position at the laboratory of Genetics of Wageningen University, supervised by Professor Mark GM Aarts and Joost JB Keurentjes. She began to work on the genetic analysis of nitrogen use efficiency in *Arabidopsis thaliana* as a PhD candidate between 2009 and 2013. After the completion of her research in Wageningen University, she worked in Enza Zaden as a biotech breeder with open field crops group in Enkhuizen, between 2014 and 2016. After the relocation of her family to Turkey, she worked on the genetic diversity of Turkish hazelnut cultivars as a post-graduate researcher in Sabancı University between 2016-2017. She still works as a post-graduate in the molecular genetics laboratory in Sabancı University to unravel genetic mechanisms of crops for a sustainable agriculture. This thesis presents the outcome of her PhD research.

List of publications:

Published:

Hiroshi Tsugawa, Ryo Nakabayashi, Tetsuya Mori, Yutaka Yamada, Mikiko Takahashi, Amit Rai, Ryosuke Sugiyama, Hiroyuki Yamamoto, Taiki Nakaya, Mami Yamazaki, Rik Kooke, Johanna A. Bac-Molenaar, **Nihal Oztolan-Erol**, Joost J. B. Keurentjes, Masanori Arita & Kazuki Saito (2019) A cheminformatics approach to characterize metabolomes in stable-isotope-labeled organisms. Nature Methods 16: 295-298.

In preparation:

Nihal Oztolan-Erol, Xi Wei, Padraic Flood, Joost JB Keurentjes, Mark GM Aarts.Genome-wide association study of *Arabidopsis thaliana* growth and response to sub-optimal nitrogen supply

Nihal Oztolan-Erol, Joost JB Keurentjes, Mark GM Aarts. A phenotypic investigation of 20 accessions of *Arabidopsis thaliana* differing in nitrogen use efficiency

Nihal Oztolan-Erol, Joost JB Keurentjes, Mark GM Aarts. Quantitative trait loci mapping of nitrogen use efficiency using the Br-0 x Col-0 recombinant inbred lines population of *Arabidopsis thaliana*

Education Statement of the Graduate School

Experimental Plant Sciences

Issued to: Nihal Oztolan Erol
Date: 14 May 2019
Group: Laboratory of Genetics
University: Wageningen University & Research

1) Start-Up Phase	date	cp
► **First presentation of your project**		
Genetic Analysis of Nitrogen Use Efficiency in Arabidopsis thaliana, Plant Physiology seminar	29 Mar 2010	1,5
► **Writing or rewriting a project proposal**		
Genetic Analysis of Nitrogen Use Efficiency in Arabidopsis thaliana	30 Jun 2010	6,0
► **Writing a review or book chapter**		
► **MSc courses**		
Subtotal Start-Up Phase		7,5

2) Scientific Exposure	date	cp
► **EPS PhD student days**		
EPS PhD students Day Utrecht, the Netherlands	1.Haz.10	0,3
EPS PhD students Day Wageningen, the Netherlands	20 May 2011	0,3
► **EPS theme symposia**		
Theme 3 Symposium "Metabolism and Adaptation", the Netherlands	10 Feb 2011	0,3
Theme 4 Symposium "Genome Plasticity"	14 Dec 2012	0,3
Theme 3 Symposium "Metabolism and Adaptation", the Netherlands	22 Mar 2013	0,3
► **Lunteren Days and other national platforms**		
Annual meeting "Experimental Plant Sciences", Lunteren, the Netherlands	19-20 Apr 2010	0,6
Annual meeting "Experimental Plant Sciences", Lunteren, the Netherlands	04-05 Apr 2011	0,6
Annual meeting "Experimental Plant Sciences", Lunteren, the Netherlands	02-03 Apr 2012	0,6
► **Seminars (series), workshops and symposia**		
Workshop: Farmer participatory approaches for variety selection/breeding and nutrient/environmental management, NUE-Crops training workshop, Crete, Greece	09-10 Jun 2012	0,6
Minisymposium: How to write a worldclass paper, Wageningen, the Netherlands	19 Apr 2011	0,2
Seminars: K. Joeri van der Velde, Danny Arends, Ritsert Jansen, Morris Swertz, Learning From Nature, Wageningen, the Netherlands	23 Feb 2012	0,3
Meeting: between Nigde University (Turkey) and Wageningen University	22 Jan 2014	0,2
► **Seminar plus**		
► **International symposia and congresses**		
NUE-Crops General Assembly, New Castle, UK	12-13 Oct 2010	0,6
NUE-Crops General Assembly, Crete, Greece	11-12 Jun 2012	0,6
International Plant Nutrition Colloquium (IPNC), Istanbul, Turkey	19-22 Aug 2013	1,2
► **Presentations**		
Talk: Genetic Analysis of Nitrogen Use Efficiency in Arabidopsis thaliana in INRA, Versailles,	30 Mar 2010	1,0
Poster: Genetic Analysis of Nitrogen Use Efficiency in Arabidopsis thaliana EPS PhD Students day, Utrecht, the Netherlands	01 Jun 2010	1,0
Poster: Towards Genetic Analysis of Nitrogen Use Efficiency in Arabidopsis thaliana, Hannover, Germany	30 Sep - 02 Oct 2010	1,0
Talk: Genetic Analysis of Nitrogen Use Efficiency in Arabidopsis thaliana NUE-Crops General Assembly Meeting New Castle, UK	12-13 Oct 2010	1,0
Poster: Genetic Analysis of Nitrogen Use Efficiency in Natural Population of Arabidopsis thaliana Using Association Mapping, Lunteren, The Netherlands	04-05 Apr 2011	1,0
Poster: Genetic Analysis of Nitrogen Use Efficiency in Arabidopsis thaliana Using Association Mapping, EPS PhD Students Day	20 May 2011	1,0
Poster: Association Mapping in Arabidopsis thaliana Unravels Candidate Genes Associated with Nitrogen Use Efficiency, Vienna, Austria	08-11 Feb 2012	1,0
Talk: Genetic Analysis of Nitrogen Use Efficiency in Arabidopsis thaliana, NUE-Crops General Assembly Meeting, Crete, Greece	11-12 June 2012	1,0
Poster: Association Mapping in Arabidopsis thaliana Unravels Candidate Genes Associated with Nitrogen Use Efficiency, Natural variation in plants summer school, Wageningen, the	21-24 Aug 2012	1,0
Talk: Natural variation for Arabidopsis adaptation to low N environments unveiled in the progeny of two extreme accessions, EPS Theme 3 Symposium, Amsterdam, the Netherlands	22 Mar 2013	1,0
Talk: Natural Variation of Nitrogen Use Efficiency in Response to Suboptimal Nitrogen Nutrition in Arabidopsis thaliana, IPNC, Istanbul, Turkey	20 Aug 2013	1,0
Talk: Genetic analysis of Nitrogen Use Efficiency in Arabidopsis thaliana, meeting with Nigde University (Turkey)	22 Jan 2014	1,0
► **IAB interview**		
Interview with Prof.dr. Ted Farmer	15 Nov 2012	0,7
► **Excursions**		
Visit Mineral and Metal Analysing Laboratory Sabanci University, Istanbul, Turkey	03-07 Sep 2012	1,5
Subtotal Scientific Exposure		21,2

CONTINUED ON NEXT PAGE

3) In-Depth Studies	date	cp
► **Advanced scientific courses & workshops**		
Bioinformatics- a User's Approach, Wageningen, the Netherlands	04-08 Mar 2013	1,5
Natural Variation of Plants, Wageningen, the Netherlands	21-24 Aug 2012	1,2
► **Journal club**		
Literature discussion Genetics and Plant Physiology 2010 -2013	2010-2013	3,0
► **Individual research training**		
Subtotal In-Depth Studies		5,7

4) Personal Development	date	cp
► **General skill training courses**		
Techniques for Writing and Presenting a Scientific Paper, Wageningen, the Netherlands	13-16 Dec 2011	1,2
Project and Time Management, Wageningen, the Netherlands	11 Jan - 22 Feb 2011	1,5
► **Organisation of meetings, PhD courses or outreach activities**		
Participation in Plant Science Slam during Annual meeting "Experimental Plant Sciences"	02 Apr 2012	0,3
► **Membership of EPS PhD Council**		
Subtotal Personal Development		3,0

TOTAL NUMBER OF CREDIT POINTS*	**37,4**

Herewith the Graduate School declares that the PhD candidate has complied with the educational requirements set by the Educational Committee of EPS with a minimum total of 30 ECTS credits.

** A credit represents a normative study load of 28 hours of study.*

The research described in this thesis was conducted at the laboratory of Genetics of Wageningen University. The research was supported by the European Commission through the NUE-CROPS research project (FP7-KBBE-2007-2A) on *Improving nutrient efficiency in major European food, feed and biofuel crops to reduce the negative environmental impact of crop production*.

Cover design: Yasemin Yılmaz Üsküplü, yasoyilmaz@gmail.com
Thesis layout: Nihal Öztolan Erol
Printed by Digiforce: || Proefschriftmaken.nl ||