Radboud University Nijmegen

Faculty of Science

Nitrogen Signalling in *Arabidopsis thaliana* roots

Thesis MSc Computer Science

Daily supervisor:
Bas van den Herik

Author:
Lili Mészáros

First reader:
Tom Claassen

Second reader:
Kirsten ten Tusscher

June 2022

Abstract

Through the plasticity of their root system architecture, plants adapt to varying environmental conditions, including the availability of nutrients such as nitrogen. Nitrate is both a nutrient and a signal and it is one of the major sources of nitrogen for plants. Under very low nitrate, root growth almost stops, whilst under low nitrate conditions, the root exhibits a foraging response, aiming to find more nitrate. Under high nitrate, root growth slows down, as other nutrients become limited. When a single plant perceives different amounts of nitrate in a split root experiment, the low nitrate side is suppressed more than its counterpart growing in a homogeneous environment, and the high side grows quicker than the side in a uniformly high nitrate condition. This indicates that there are not only local, but systemic nitrate signals determining root growth. In this thesis we implemented a functional structural plant model (FSPM) simulating these split root experiments. The FSPM was based on the model created by Boer et al. [1], and contains nitrate (N), C-terminally encoded peptide (CEP), and cytokinin (CK) signalling pathways. The FSPM explicitly represents the growing branching architecture of plant roots, and the calculation of the growth rate is an improved version of the one taken from Boer et al. Additionally, transcriptomics data from Ruffel et al. [2] was analysed where genes that are part of N, CEP, and CK signalling pathways were found. This included transporters of the NRT1 and the NRT2 family, auxin and cytokinin biosynthesis genes, ROXYs, and nitrate assimilation genes. In addition, sugar transporters were identified. However, they did not show strong signals, and could not be integrated into the FSPM model to further improve the accuracy of the growth of the root system architecture.

Contents

Chapter 1

Introduction

The global population is growing rapidly, and the current rate of crop production is insufficient to meet demands [3]. Climate change is severely damaging crop growth through extreme droughts, floods, and heat [4]. Therefore there is a necessity to find more efficient ways of growing crops to ensure food security. Industrial agriculture suffers from numerous problems and current practices have to be rethought. The high yields have come at the cost of ecological and environmental damages. A monoculture produces greater yields and has lower costs on a short term, but the risk of diseases and pests is larger. Pesticides are used excessively to tackle the problem which in return causes chemical pollution and biodiversity loss [5]. The machines on the fields cause soil compaction, there is not enough space for water and air, the minerals are pressed together, therefore the plants are less able to grow roots. Since the soil is not able to absorb the water, soil erosion intensifies as well [6]. The lacking nutrients in the soil, such as nitrogen are compensated by using large amount of nitrogenous fertilisers to meet the crop demand. As a result nitrous oxide emissions increase, which causes eutrophication in fresh water and marine ecosystems [7]. Plants take up nitrogen as ammonium or nitrate from the soil. In this thesis we focused on how to limit the use of nitrogenous fertilisers by finding how a plant can efficiently take up nitrate. We aimed to find nitrate signalling routes; how nitrate is taken up, transported, and stored; and how these affect growth.

We tried to find out how plants interact with their environment, which is very different from how members of the animal kingdom do. Mammals react very quickly to the changes in their surroundings, for example they sense if they touch something too hot and immediately move their hand. They are capable of doing this because they have a developed central nervous system, the nerves on the hand send a signal to the brain, which in return tells the mammal to pull away the hand. This is not the case for all animals, for example the nervous system of a cephalopod, such as an octopus is not centralised. Each tentacle is able to perform complex actions on its own and combined, these navigate the octopus in the right direction. A plant's physiology and morphology is very different from that of an animal's. Plants are not able to adapt as quickly, however they have a "nervous system" of their own which sends signals in their "body" through which they move in a direction where they have the highest possibility of survival, and possibly reproduction. Above the soil the leaves are oriented so that they get the maximum amount of sun, below the soil the roots navigate to patches where nutrients are optimal.

The mechanisms governing root formation and growth are complicated, even causing surprises, as a larger amount of nitrogen does not necessarily mean that roots will become bigger. As illustrated on figure 1.1 when external nitrate is scarce, – hence the internal

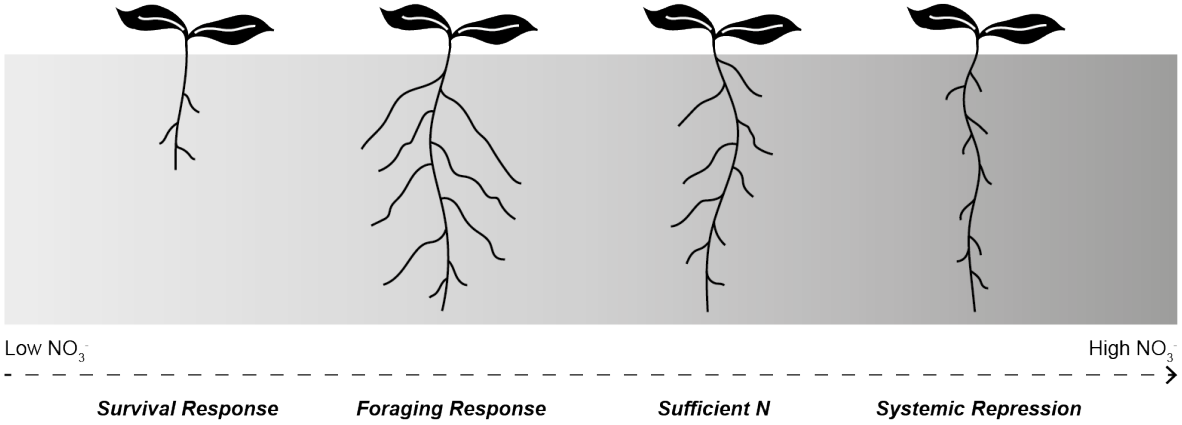


Figure 1.1: Responses of roots under varying nitrate concentrations in a homogeneous environment [1].

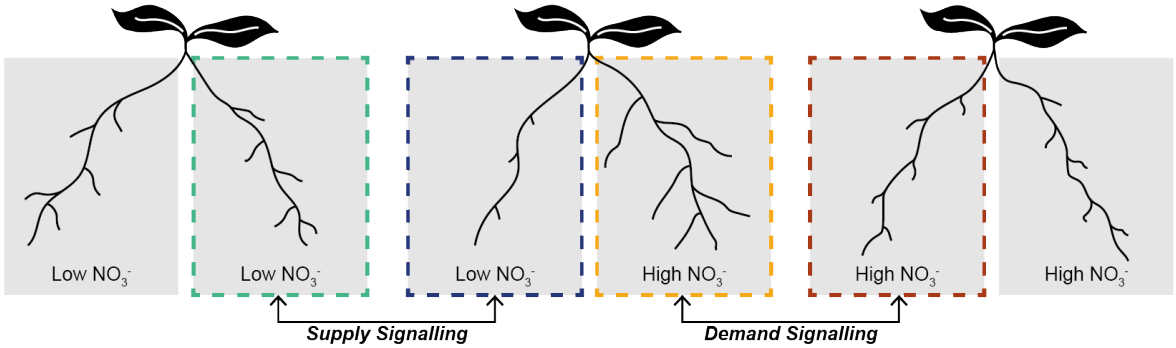


Figure 1.2: Responses of roots under varying nitrate concentrations in a heterogeneous environment [1].

levels too – the root shows a survival response where the root growth slows down. As the levels increase, the response is a foraging one and the root starts searching for patches where nitrate availability is higher. However if internal nitrate levels are very high, the response is systemic repression, where root growth is repressed as other nutrients' availability is limited.

To find out more about these underlying processes, researchers conduct split root experiments, where the taproot is separated into two compartments (figure 1.2). In the separate compartments, the nitrogen concentration can be adjusted to different levels, creating heterogeneous environments. In a setting of combined low and high environments the two roots are asymmetrical, but what is even more interesting is that the root from the high nitrogen side exceeds the size of roots grown with the same N concentration that grew in a homogeneous environment, whereas the heterogeneous low side is significantly smaller than the homogeneous low. This shows the "intelligence" of plants as they do not only respond locally to different signals, but are able to respond systemically as well. The heterogeneous high side shows a systemic demand signal as growth is promoted, whilst the low is sending systemic supply signals and growth is repressed [1].

Arabidopsis thaliana, or by its common name mouse-ear cress shown in figure 1.3 has been used as a model organism in plant biology, therefore this was our model plant. The *A. thaliana* plant has no agricultural purpose but shares similarities with other flowering

Figure 1.3: The flowering plant *Arabidopsis thaliana*.

plants, thus the findings can be extended to crops [8]. During the thesis we modelled the growth of the root system architecture and analysed transcriptomics data on nitrate dependent root growth. The model was based on the model of Boer et al. where known molecular pathways were incrementally added in order to describe preferential root foraging [1]. The data was taken from Ruffel et al. where split root experiments were investigated using microaarray technology [2]. The analysis on the transcriptomics data was conducted with limma (Linear Models for MicroArray data), an R package created for the analysis of gene expression data. The package provides the possibility of comparing multiple RNA targets in parallel, and produces reliable results even when the amount of arrays available are small [9, 10].

1.1 Outline of thesis

In chapter 2 we introduce the biological background of the thesis, the terminology used, how a plant is built up, the processes governing it, the molecular pathways, and the experiments.

In chapter 3 the modelling of plant root growth is discussed, starting with the model of Boer et al. [1] in section 3.1, the programming language used for it, and the outcome of the experiments.

The focus of chapter 4 is the data analysis of transcriptomics data of split root experiments, where we identify key players of nitrate signalling.

Finally, in chapter 5 we summarise our findings, draw our conclusion, and propose future research.

The project is available at `https://github.com/meszlili96/NitrateSplitRoots`.

Chapter 2

Introduction to vascular plants

Plants are able to adapt to their environment through their physiology. This includes both chemical and physical processes such as producing proteins, hormones and other chemicals. For example chlorophyll enables a plant to create its own energy from light; with toxins they defend themselves from diseases, animals, and other plants; and ethylene governs leaf loss. Another important aspect is the morphology, which determines the shape and structure of plants. In particular leaves catch light to produce sugar and the plasticity of the root system architecture (RSA) is crucial for a plant to be able to take up numerous nutrients such as nitrate, phosphorus, and potassium.

2.1 Vascular plants

In our research we work with a vascular plant, *Arabidopsis thaliana*. Here we will discuss the tissues and organs defining these plants. As the name suggests a prominent feature of vascular plants is the vascular tissue, which is responsible for the transport of food and water [11].

2.1.1 Morphology

A vascular plant has reproductive organs, and three basic vegetative organs: the roots, the stems, and the leaves. The former constitutes the root system, the latter two the shoot system. The leaves are responsible for photosynthesis, and the stem connects the leaves to the roots [11].

The root system of an Arabidopsis plant consists of a primary root, also known as the taproot, and lateral roots (LRs) that are responsible for the foraging behaviour. The taproot is connected to the shoot, the sink of the nutrients. The primary root has different zones, the differentiation zone (DZ), the elongation zone (EZ), the meristematic zone (MZ) which can be seen on figure 2.1. The LRs form from the primary root in three steps, priming, initiation and elongation. Priming occurs periodically as a result of oscillating genes, – genes that are expressed in periodic cycles – in the region of the EZ and MZ of the taproot. If a site has primed, only then can a LR initiate, and if priming occurred, initiation always follows. The development of a lateral root primordium takes approximately 40 hours unless interrupted, and generally only about 2/3 of the LRPs develop based on the patterning dynamics of the root tip. Further lateral roots develop later which are dependent on the systemic control of the shoot [12].

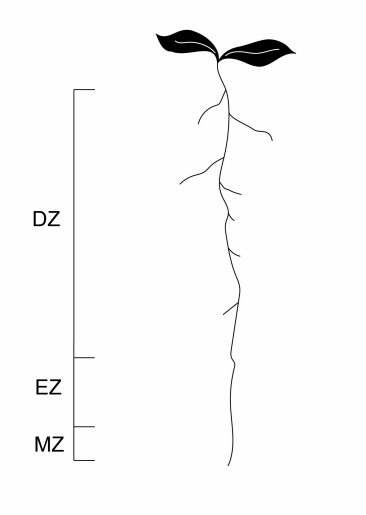
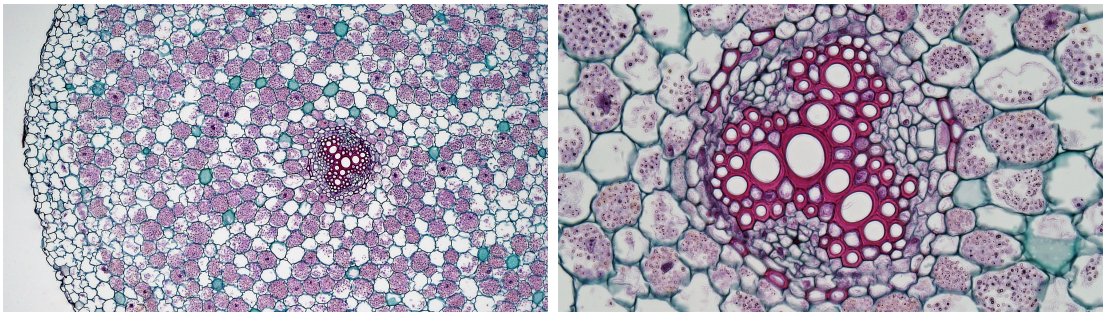


Figure 2.1: Developmental zones of a root. Most of the root is made up of the differentiation zone (DZ), followed by the elongation zone (EZ), and finally the meristematic zone (MZ).

2.1.2 Anatomy

Multiple kinds of tissues occur in the body of a plant, the dermal tissue covers and protects the outside of the plant, the vascular tissue transports water and nutrients, and the ground tissue fills the gaps between the dermal and vascular tissues. Roots, stems, and leaves are all built from these three kinds of tissues. On figure 2.2 there is a cross section of a root on the left, where the different type of tissue cells can clearly be distinguished with the dermal on the white outer layer, the ground tissue the middle light purple, and the vascular tissue at the center in magenta. The dermal tissue in the root is made up of root hairs and the epidermis. The ground tissue consists of the cortex, the endodermis and the pericycle.



Cross section of a root Phloem and xylem tissue

Figure 2.2: Cross sections displaying buttercup root tissues[1]. On the left, the outmost layer in white is the dermal tissue. Only the epidermis is visible here, the root hairs are not. The middle layer with blue and light purple is the cortex which is part of the ground tissue. On the right, there is a closer view on the center of the tissue. The two circles of small cells are the endodermis and the pericycle, which are also part of the ground tissue. The vascular tissue is at the center and consists of the xylem in magenta, and the phloem surrounding the xylem in purple [13].

On the right, there is a closer view of the vascular tissue, which consists of the phloem

and the xylem. These are crucial for the development of the root architecture as the nutrients needed for survival, and the signalling molecules are transported through these tissues as well. Nutrients (e.g. nitrogen, phosphorus, calcium) are taken up by the roots, and sugar and oxygen are produced in the leaves during photosynthesis, all of which need to be transported from one part of the plant to the other [11]. The details of water and food transportation are shown in figure 2.3. Where the food is being produced is the source (leaf), and where it is stored is the sink (root, stem). Water moves upwards from the sink to the source in the xylem through transpiration, and during translocation sucrose is pushed down in the phloem. Water moves into the phloem through osmosis as the phloem has a higher sucrose concentration from the source cells. When the sink cells take up the sucrose, the sugar concentration in the phloem decreases, allowing the water to pass back to the xylem through osmosis. However in young *Arabidopsis* plants there is very little storage, almost all of the sugar is used for growth.

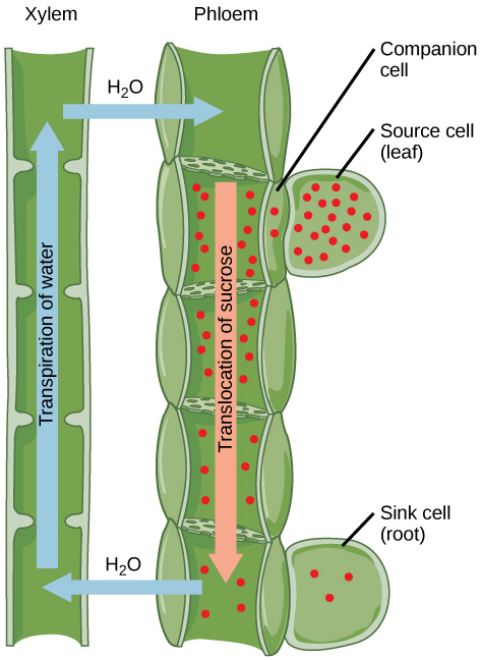


Figure 2.3: The transpiration of water and translocation of sugar in a plant. Water is shown in blue, and the red dots represent sucrose molecules. The phloem compartments are separated by sieve plates, which have holes in them [14].

2.2 Conditions of root growth

Nitrogen is one of the nutrients that a plant needs in large quantities, and as such roots develop in the direction of high nitrate patches, and where nitrate is scarcer, roots will grow slower. However, if nitrate is abundant everywhere in the soil, the root growth almost halts to a stop. As illustrated on figure 1.1, the roots are showing different responses based on the soil nitrate availability, under low NO_3^- the plant shows a survival response, and as the availability increases it starts a foraging behaviour, until NO_3^- levels become too high, and systemic repression occurs [7, 1]. These responses are governed by phytohormones, such as auxins (AUX) which influence cell enlargement and root initiation;

[1]Buttercup (*Ranunculus*) roots are similar to *Arabidopsis* roots.

cytokinins (CK) which control cell division and shoot formation. AUX and CK together are responsible for most of the growth in plants, including root growth. C-terminally encoded peptide (CEP), a kind of peptide hormone also plays an important role in lateral root formation, and acts as a root-to-shoot signalling molecule. The signalling and transportation of these hormones occurs through different gene pathways, dependent upon the external soil nitrate. Some of these pathways are depicted on figure 2.4.

Uptake depends on the nitrate concentration, and different transport systems are activated based on the availability. The low affinity transport system (LATS) is activated if the nitrate levels are high, and the high affinity transport system (HATS) is responsible when nitrate levels are low. The nitrate transporter family NRT1 is involved in the LATS, and NRT2 in HATS. The exceptional member is NPF6.3/NRT1.1, which is a dual affinity transporter, meaning it is active under both LATS and HATS. This is visible on figure 2.4, NRT1.1 appears on both the left and right root. Next to NRT1.1, NRT1.2 is a well-known low affinity transporter responsible for nitrate influx. From the NRT2 gene family NRT2.1, NRT2.2, NRT2.4 and NRT2.5 are the main contributors to nitrate uptake under low nitrogen availability. NRT2.1 and NRT2.2 import nitrate through the cortex and the endodermis, whilst NRT2.4 and NRT2.5 take it up through the root hairs [7]. NRT2.1 can be seen on the left side where it is indirectly activated under high local N.

After uptake, the nitrate is translocated via the xylem with the aid of NRT1.5, and reduced to nitrite by nitrate reductase (e.g. NIA1, NIA2), and to ammonium by nitrite reductase (e.g. NIR1). Ammonium is then incorporated into amino acids, catalysed primarily by glutamine synthetase (e.g. GSR1, GLN1;4) and glutamate synthase (e.g. GS2) [7, 15].

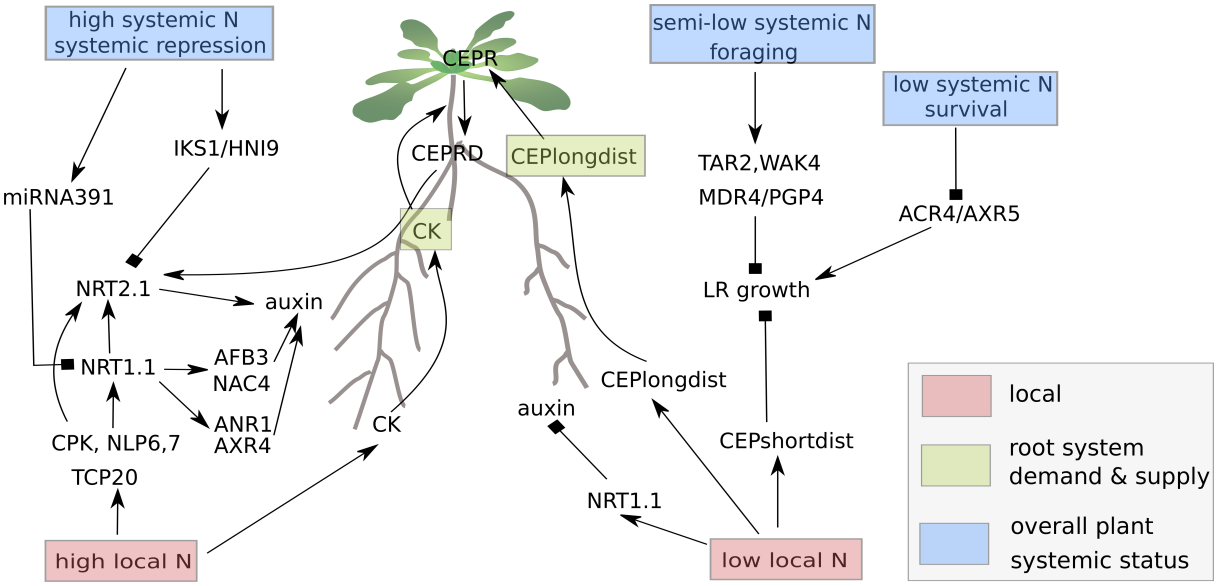


Figure 2.4: Genes and their functions. High systemic nitrogen causes systemic repression through miRNA391 and IKS1/HNI9, which block the nitrate transporters NRT1.1 and NRT2.1 subsequently. These transporters are also affected by high local nitrogen. CPK, NLP and TCP20 influence NRT1.1 which in return influences AFB3, NAC4, ANR1, AXR4 and NRT2.1. All of these have an impact on auxin. High local N increases cytokinin levels as well. The combination of increased CK and AUX leads to lateral root growth.

When nitrate is scarce, plants exhibit a survival response, and genes such as ACR4

and AXR5 are down-regulated [1]. They can be seen on figure 2.4 on the right side being blocked by low systemic N survival. In response to slightly higher levels of nitrate (semi-low systemic N foraging), roots exhibit a foraging response. This is due to TAR2 which causes auxin biosynthesis, the cell wall–associated protein kinase WAK4, and the auxin transporters MDR4 and PGP4 [16]. At medium levels of nitrate, NRT1.1 increases auxin signalling through the AFB3, NAC4, OBP4 pathway, resulting in lateral root formation (left root). However, for very high levels of nitrate, systematic repression occurs, which is further repressed by HNI9. HNI9 is visible on the upper left side and blocks NRT2.1.

CEP pathways also play an important role in foraging behaviour. Under low nitrate, roots produce CEPs locally from where they are translocated via the xylem to receptors in the shoot, called CEP receptors (CEPR) [17]. These produce CEP DOWNSTREAM, CEPD1 and CEPD2, which are transported back through the phloem to roots where nitrate is higher and upregulate NRT2.1, hence promoting LR growth [18].

Category	Genes	Function
AUX	ACR4, AXR5	Causes LR formation, nitrogen deficiency down-regulates them [16].
	AFB3, NAC4 and OBP4	This pathway represses auxin sensing when nitrate is very high, and enhances auxin through NRT1.1 when nitrate is sufficient [19].
	NRT1.1/NPF6.3	Auxin importer at low, reducing laterals. At medium it does not transport, increasing LRs [1].
	NRT2.1	Transports nitrate under low nitrate conditions [7].
	HNI9	Represses nitrate transport [20].
	TAR2	Foraging under low nitrate which results in increased auxin biosynthesis [21].
	WAK4	Regulates cell elongation, expressed under low nitrate [16].
	MDR4, PGP4	Downstream auxin transporter, foraging [16].
CK	IPT3, IPT5, CYP735A2, LOG5, and CKX4	Cytokinin metabolism genes [22].
CLE-CV1	CLE1, CLE2, CLE3, CLE4, CLE5, CLE7	When nitrogen is very limited, LR number decreases [23].
	CLV1	When nitrogen is scarce it mediates CLE signalling pathways, thus contributes to inhibited LR growth [23].
CEP	CEP1, CEP2, CEP3, CEP5, CEP9, CEP13, CEP14, CEP15	Roots produce them under low nitrate conditions [17].
	CEPR1, CEPR2	CEP receptors are part of the CEP demand signaling pathway, CEPS bind to them [17].
	CEPD1, CEPD2	CEP downstreams are produced in the shoot [18] and up-regulate NRT2.1 in roots that are in high N [17].
	miRNA391 miRNA393	Targets AFB3 [19].

Table 2.1: Genes involved in root growth. Phytohormones are in blue, peptides in red, and micro RNA in yellow.

Chapter 3

Functional structural plant model

Computational models have an important role in understanding the multiple processes ongoing in living organisms. Here we used a functional structural plant model to explain how root growth is affected under different nitrogen conditions. The FSPM is based on the model of Boer et al. introduced in section 3.1, where the combined signalling pathways describe lateral root growth.

3.1 Model of Boer

The model introduced by Boer et al. [1] simulates split root experiments by creating two individual root compartments and measuring the root length for them separately. The root length is assumed to be proportional to the shoot area. The lengths are also dependent on N, CEP, and CK signalling pathways. Altogether these lead to the central equation of the model which measures root growth using an ordinary differential equation (ODE)[1]:

$$\frac{\mathrm{d}L_x}{\mathrm{d}t} = \frac{1}{n} \cdot conv \cdot r_x \cdot \sum_{y=1}^{n} L_y \tag{3.1}$$

where L is the accumulated root length, x is the root compartment being modelled, and n is the number of compartments (in our case $n = 2$). The maximum rate of root length increase per unit of shoot area is given by the conversion factor $conv$, and r_x is the growth rate dependent on different signals. $conv$ and n are constants defined preliminarily. The signalling pathways are added incrementally to the model by changing the parameter r in every step. The signalling pathways are numeric factors which are multiplied together, giving the product

$$r = f_{\text{basic}} \cdot f_{\text{local}} \cdot f_{\text{CEP}} \cdot f_{\text{systrepr}} \cdot f_{\text{systfor}} \tag{3.2}$$

The basic signalling (f_{basic}) is dependent on the systemic nitrate levels. Local nitrate levels, therefore local signalling (f_{local}) affects growth through the nitrate transporter NRT1.1 which transports auxin depending on whether external nitrate levels are high or not. In a heterogeneous environment, the root on the high nitrate side is significantly bigger than the one on the low side which suggests systemic demand signalling. This signalling is conducted through CEP pathways (f_{CEP}) and the regulation of NRT2.1. In a homogeneous environment under very high nitrate levels the roots become smaller which

[1] All the ODEs shown here differentiate over functions that are dependent on t but for simplicity are abbreviated, e.g. $L_x(t)$ to L_x.

occurs because of systemic repression (f_{systrepr}). In extreme conditions plants exhibit either a survival response or systemic repression. In between, there is another response that has been observed which is referred to as foraging (f_{systfor}). A schematic drawing of nitrate, C-terminally coded peptide and cytokinin transport is shown on figure 3.1, together with the signalling pathways.

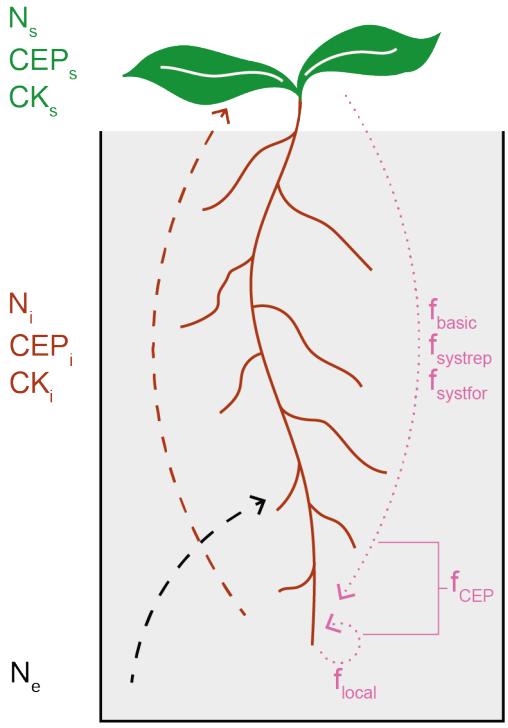


Figure 3.1: The root senses and takes up external nitrate (N_e, black), from which the internal N, CEP, and CK is calculated (N_i, CEP_i, CK_i, red) in equations 3.3, 3.5, and 3.7. The locally produced N, CEP, and CK are transported to their systemic pools (N_s, CEP_s, CK_s, green), calculated in 3.4, 3.6, and 3.8. The growth rate (r, pink) calculated in equation 3.2 depends on these signals. f_{basic} is dependent on the systemic nitrate N_s; f_{local} on the external nitrate N_e; and f_{CEP} on N_e, CEP_s, and CK_s. f_{systrepr} and f_{systfor} depend on N_s.

These signals depend on external nitrate, which was defined per root compartment as $N_{e,x}$. Using the external nitrate, the internal nitrate $N_{i,x}$ can be calculated:

$$\frac{\mathrm{d}N_{i,x}}{\mathrm{d}t} = \left(up_1 \frac{N_{e,x}}{N_{e,x} + K_{up}} + up_2 N_e \right) L_x - T_{up} N_{i,x} \tag{3.3}$$

where the parameters $up_1, up_2, K_{up}, T_{up}$ respectively are the uptake rates, the concentration, and the rate of transport. The systemic nitrate N_s is given by

$$\frac{\mathrm{d}N_s}{\mathrm{d}t} = T_{up} \sum_{x=1}^{n} N_{i,x} - u_m \sum_{x=1}^{n} L_x - eN_s \tag{3.4}$$

where u_m and e are rates of nitrate loss.

To model the demand signalling, the produced CEP needs to be calculated as well:

$$\frac{\mathrm{dCEP}_x}{\mathrm{d}t} = p_{\text{CEP}} \frac{K_{\text{CEP}}^2}{K_{\text{CEP}}^2 + N_{e,x}^2} + L_x - T_{\text{CEP}} \text{CEP}_x \tag{3.5}$$

$$\frac{\mathrm{dCEP}_s}{\mathrm{d}t} = T_{\mathrm{CEP}} \sum_{x=1}^{n} \mathrm{CEP}_x - d_{\mathrm{CEP}} \mathrm{CEP}_s \tag{3.6}$$

where p_{CEP} is the maximum production rate of CEP, K_{CEP} is the concentration, T_{CEP} the rate of transport, and d_{CEP} the degradation rate.

For the supply signalling, CK dynamics are added:

$$\frac{\mathrm{dCK}_x}{\mathrm{d}t} = p_{\mathrm{CK}} \frac{N_{e,x}^2}{K_{\mathrm{CK}}^2 + N_{e,x}^2} + L_x - T_{\mathrm{CK}} \mathrm{CK}_x \tag{3.7}$$

$$\frac{\mathrm{dCK}_s}{\mathrm{d}t} = T_{\mathrm{CK}} \sum_{x=1}^{n} \mathrm{CK}_x - d_{\mathrm{CK}} \mathrm{CK}_s \tag{3.8}$$

where p_{CK} is the maximum production rate, K_{CK} is the concentration, T_{CK} is the transport rate, and d_{CK} is the degradation rate of CK.

When all of these signals are added up, the modelled root lengths matched what happened in experimental studies except for the low side of the plant root in the heterogeneous nitrate environment. The suspicion of Boer et al. on why this root compartment was longer than what was seen in available experimental data was that carbon allocation was missing from the model. Growth in all the plant organs heavily depends on carbon, and the allocation of carbon depends on an organ's sink strength. Quicker growth enhances sink strength, hence more carbon is allocated, which in return increases growth further. Therefore equation 3.1 was modified to

$$\frac{\mathrm{d}L_x}{\mathrm{dt}} = g_{carbon,x} \cdot conv \cdot r_x \cdot \sum_{y=1}^{n} L_y, \quad g_{carbon,x} = \frac{r_x \frac{L_x}{\sum_{y=1}^{n} L_y}}{\sum_{y=1}^{n} r_y \frac{L_y}{\sum_{z=1}^{n} L_z}} = \frac{r_x L_x}{\sum_{y=1}^{n} r_y L_y} \tag{3.9}$$

Before, root compartments received equal amounts of carbon through $\frac{1}{n}$. With the modification, the compartments' growth is weighted based on their relative size $\frac{L_x}{\sum_{y=1}^{n} L_y}$ which simulates the carbon sink strength of a compartment. The model showed correct results after the modification of the growth equation [1], which can be seen on figure 3.3f in the next section.

3.2 Implementation of the Boer model as an FSPM

The model implemented by Boer et al. [1] has limitations as it is an abstract representation, only the cumulative length L of the laterals is calculated. There is neither branching in the RSA, nor a distinction made between the main and the lateral roots in terms of growth dynamics. Therefore as a first step in our modelling we re-created the Boer model as a functional structural plant model (FSPM). The aim of a FSPM is to accurately describe the 3D architecture of a plant and the processes governing its development over time, taking the environmental factors into account. This can be done on different scales, from a cell to a forest. All FSPMs follow the paradigm of object oriented programming, where the plant is built of elementary units, like roots, stems, nodes, buds, etc., each with its own properties and functions. These components are parameterised and instantiated separately. The explicit representation allows feedback between the structure and the functions, making them particularly suitable for analysing problems where the spatial structure is relevant to the behaviour of the plant. There are numerous approaches

```
RU(180)
ShootOrg(true)

[
    RU(45)
    for((1:2))(RU(-1.5) RootSegment(1, Nleft))
    EZSegment(1, 0, 0, Nleft, 0)
]

[
    RU(-45)
    for((1:2))(RU(1.5) RootSegment(2, Nright))
    EZSegment(2, 0, 0, Nright, 0)
]
```

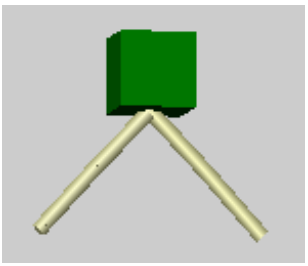


Figure 3.2: The definition and the snapshot of the initial root.

to FSPMs: morphological models focus on the plant morphology, some models focus on the physical and biological processes occurring in a plant, and integrated models were created for both single plants and plant communities. Most implementations are based on L-systems [24]. A Lindenmayer system is a type of formal grammar very similar to a Chomsky grammar. The key difference between the two is how the productions are applied to a word. In a Chomsky grammar the productions are applied sequentially, whilst in a Lindenmayer system they are applied in parallel. The aim of a L-system is to represent the behaviour of living organisms. Replacing all letters in a word in parallel reflects the biological motive behind L-systems as cell divisions in organisms happen at the same time [25, 26]. The programming language XL is an implementation of L-systems. More specifically XL is a relational growth grammar with built-in ODE solver [27] and is discussed in more detail in appendix A.

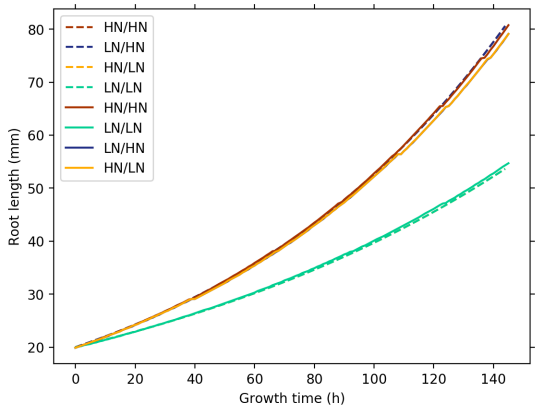
In our model we heavily relied on the Boer model and kept the equations the same. The improvement was creating a physical representation and the accompanying functions necessary for it. We did not create the entire plant as some of the organs, like reproductive ones are of no interest for root growth. Leaves were left out, adding them was not within the scopes of the thesis. The basic model that we created is made of a shoot and two roots, as described on figure 3.2.

The shoot is defined as `ShootOrg`, and both the taproots consist of two root segments (`RootSegment`) and one elongation zone segment (`EZSegment`) each. The `RL(a)`, `RU(a)`, `RH(a)` commands rotate these components by `a` degrees around the x/y/z-axis, respectively [27].

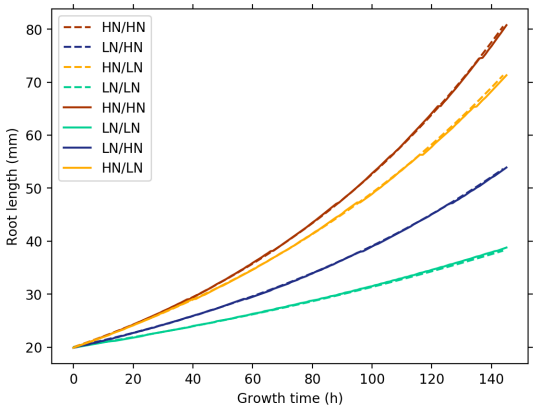
All plant organs are based on the `Organ` abstract class which has properties defining the external and internal amount of nitrate, as well as the amount of CEP and CK, and the age. In addition to these parameters, the `RootSegment` is initialised with the length, and the order of the root, e.g. the taproot is 0^{th}, the lateral 1^{st} order. The parameter that differs for a root segment is the external nitrate concentration. The `EZSegment` has similar properties, and is extended with the growth related functions (figure A.1, A.2) as this is the part of the plant that is growing, the rest of the root is "constant". These functions are implementations of the equations 3.1 and 3.2.

To verify the accuracy of the FSPM, after implementing the classes with their functions, we proceeded to compare the root lengths of the FSPM and the Boer model incrementally adding the different signals. On figure 3.3a we compare the baseline models. The roots grow with similar speed on both plots, however the FSPM is not as smooth as the Boer version. The incremental increases in the FSPM are a simulation artefact. These can be explained by the discrete addition of new root segments and increasing the spatial resolution in the simulations would resolve these. Indicating the simplicity of the baseline model, all the roots, except for the homogeneous low ones have almost the exact same length, therefore those lines overlap on the plots.

Adding the local signalling distinguishes the various conditions, the roots with a high nitrate environment are growing quicker than the roots on the low side as seen on figure 3.3b.



(a) $r = f_{\text{basic}}$
The baseline model only focuses on systemic and not on local internal nitrate levels, resulting in the separation of the roots in only low nitrate (green) from the rest.



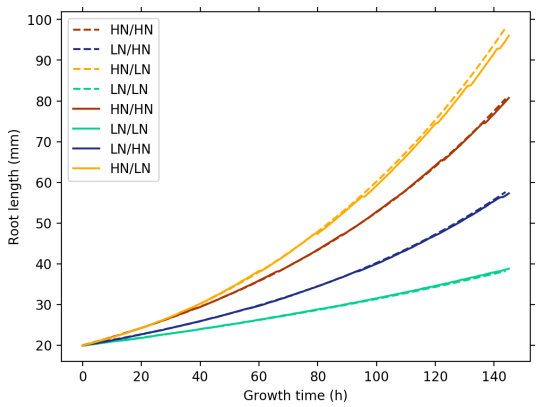
(b) $r = f_{\text{basic}} \cdot f_{\text{local}}$
Local external nitrate separated roots experiencing high (red, yellow) from roots experiencing low (blue, green) conditions. However, the homogeneous high (red) is larger than the heterogeneous (yellow).

Figure 3.3: Growth dynamics for models. The dashed lines represent the Boer model, while the solid lines represent the FSPM version.

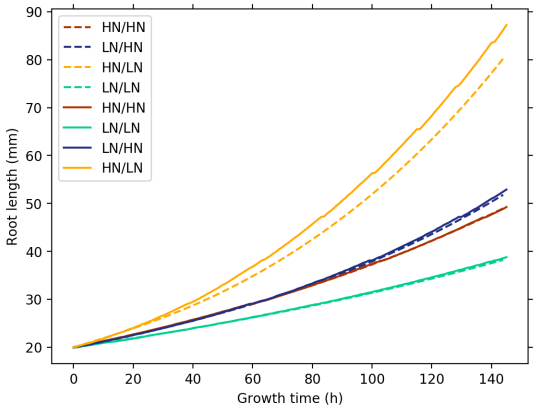
The results seen on figure 3.3b do not match experimental results where roots grown on homogeneous soil compared to their heterogeneous counterparts are larger when nitrate is lacking, and smaller when nitrate is plentiful. When systemic demand signalling is included, the plots on figure 3.3c show a more realistic image as the HN/LN line increases the most rapidly.

After adding the systemic demand signalling caused by CEPs the roots in the heterogeneous high nitrate side are the fastest growing ones, however the ones on the homogeneous side do not show the systemic repression that would be expected on homogeneous nitrate rich soil. This response is included in the next step and is shown on figure 3.3d. The HN/HN growth is "flattened" and closely follows the line of LN/HN.

All signalling for preferential foraging on the high nitrate side is now added, but the low nitrate side has been ignored so far. Roots in homogeneous nitrate deprived conditions exhibit a foraging behaviour and grow more quickly than ones in a heterogeneous environment and roots grown in nitrate rich conditions. After adding systemic foraging the root lengths are in the right order on figure 3.3e with HN/LN the highest, followed



(c) $r = f_{\text{basic}} \cdot f_{\text{local}} \cdot f_{\text{CEP}}$
Systemic demand signalling increases the size of the root that is experiencing high nitrate on one side (yellow) from the one experiencing it on both (red).



(d) $r = f_{\text{basic}} \cdot f_{\text{local}} \cdot f_{\text{CEP}} \cdot f_{\text{systrepr}}$
Roots growing in only high nitrate (red) are smaller than roots in high nitrate on one side (yellow), but they are significantly smaller as a result of systemic repression.

Figure 3.3: Growth dynamics for models. The dashed lines represent the Boer model, while the solid lines represent the FSPM version.

by LN/LN, LN/HN, and finally HN/HN.

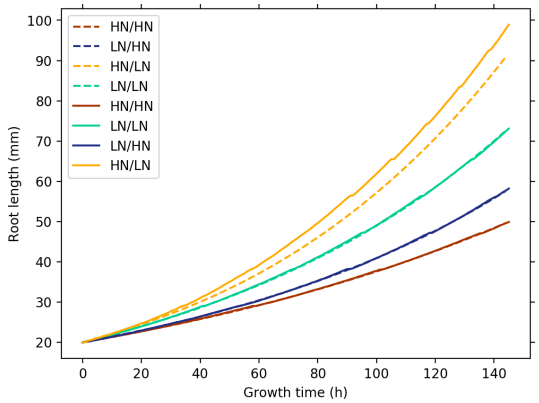
Figure 3.3f shows the model after the modification of the growth equation to a version with carbon allocation. The results match the data available from experimental studies. The fastest growing root is on the heterogeneous high side because of extensive foraging, which is followed by the homogeneous low side where foraging still occurs, only with weaker foraging signals and survival becoming stronger. The homogeneous high side comes next where systemic repression is occurring, and finally the heterogeneous low side where the root is showing a survival response.

Adding the signalling pathways incrementally verified the accuracy of the FSPM allowing for improvements to be included in the following steps.

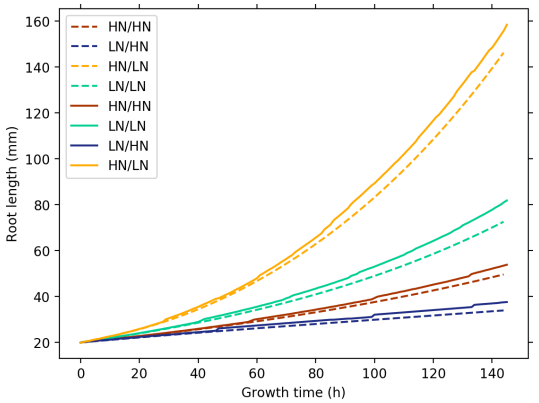
3.3 Extending the model

The Boer model consisted of eight "lateral" components but with no architecture, whereas the basic FSPM had more potential as it resembled a plant more by having an explicit shoot, and two roots made up of root segments which ended in elongation zones. However, we observed that the rate of taproot growth is similar, independent of the nitrate levels [28, 29]. During root growth, lateral roots are the roots that show a foraging behaviour and this is where differences in root sizes are visible. If a plant senses more nitrate, more laterals will start emerging, and they will grow more quickly than one with less nitrate availability [28]. Therefore the calculation of the root length had to be reconsidered to separate the lateral and primary roots to an equation where signalling plays a role, and to one that has constant growth.

We made a distinction between the behaviour of the primary and the lateral roots by giving them orders. The primary was given the 0^{th}, the lateral the 1^{st} order. The amount of nitrogen in the soil does not affect the rate of taproot growth [28, 29], therefore they



(e) $r = f_{\text{basic}} \cdot f_{\text{local}} \cdot f_{\text{CEP}} \cdot f_{\text{systrepr}} \cdot f_{\text{systfor}}$ With systemic foraging included, the homogeneous low side (green) becomes the second biggest, instead of smallest root length.



(f) $r = f_{\text{basic}} \cdot f_{\text{local}} \cdot f_{\text{CEP}} \cdot f_{\text{systrepr}} \cdot f_{\text{systfor}}$ Carbon allocation increases the length of the homogeneous high side (red) to become larger than the heterogeneous low (blue).

Figure 3.3: Growth dynamics for models. The dashed lines represent the Boer model, while the solid lines represent the FSPM version.

grow with a constant rate l:

$$\frac{\mathrm{d}L_x^{(0)}}{\mathrm{d}t} = l \tag{3.10}$$

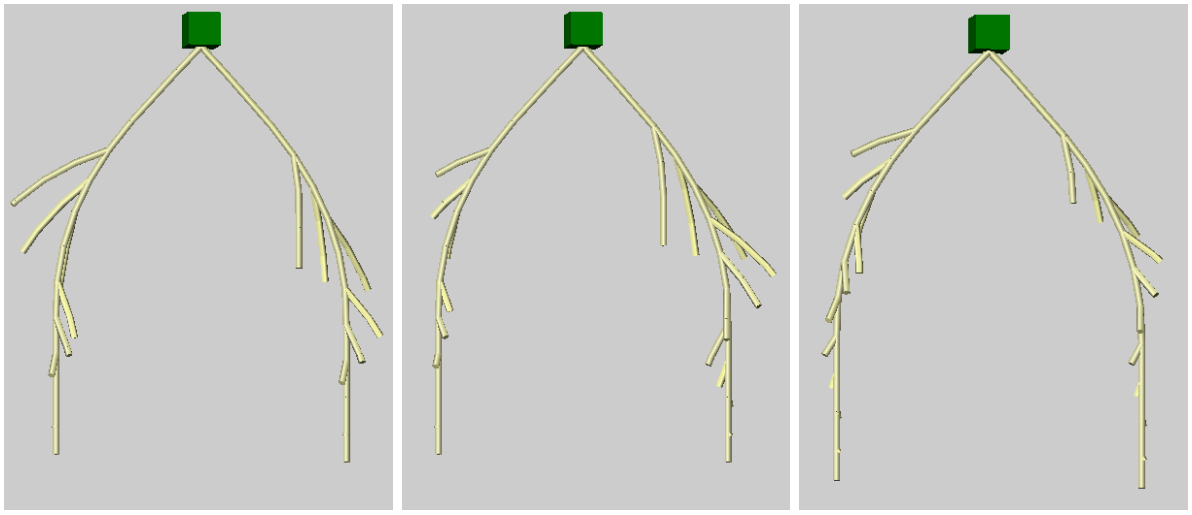
The growth rate for the laterals (1^{st} order) is

$$\frac{\mathrm{d}L_x^{(1)}}{\mathrm{d}t} = \frac{1}{n_0 + n_1} \cdot conv \cdot r_x \cdot \sum_{y=1}^{n} L_y, \quad n_0 + n_1 = n \tag{3.11}$$

The is equation is the same as equation 3.1 as all parts of the root system contribute to the growth of the laterals. However there is a difference, the differential equations given in equations 3.3, 3.4, 3.5, 3.6, 3.7 and 3.8 do not calculate the signals for 0^{th} order roots, only for 1^{st} order ones. Essentially, the taproots are just a structure for the laterals to grow, the length and the signalling of them does not matter. Calculating the growth rate r remained the same as in the Boer model, given on equation 3.2.

In the model of Boer et al. a carbon-sink function was used [1], and the growth equation 3.1 was replaced by equation 3.9. Originally all segments received the same amount of carbon, which was then replaced by a factor for carbon competition using the relative sizes of the compartments. This meant that if a root initially started growing more quickly it would be able to allocate more carbon, increasing its size more rapidly than a slower growing one. Whereas we wanted the roots to emerge from the growth dynamics. Laterals form more often if they are in a high nitrate condition, there are about 20-25% more roots than in a nitrate deprived condition, therefore more roots appear on those plants [28] which in return enhances the dominance. This does not mean that they are also long, repression is blocking growth in a completely nitrate rich environment, and the roots remain small. The ability to replace the carbon-sink function strongly suggests that there are signalling pathways underlying LR formation. We do not know yet what signals play a role, but this can be investigated in future work. Including this new signal would mean that we could control lateral root emergence.

Another aspect of designing the model was its physical appearance. This involved the implementation of gravitational tropism in both the taproots and laterals, and the angle and periodicity of the lateral roots. Gravitropism is a plant's response to gravity where the root changes its orientation and bends downwards. It is a growth response during which cells on the lower side of the elongation zone expand less compared to the cells on the upper side [30]. Laterals develop following the pattern of a corkscrew, meaning they rotate around the main root. Moreover, laterals do not form immediately, only after a certain age of their parent root do they start developing, and if the parent root becomes too old, they stop forming. On figure 3.4 it is visible that at the top, the taproots do not have any laterals growing on them, they missed the developmental window. From top to bottom, the laterals decrease in length as the roots on the lower part started developing later.



LN/LN LN/HN HN/HN

Figure 3.4: *Arabidopsis* roots after developing for 8 days under different nitrate concentrations. There is substantial asymmetry between the lateral roots in differing environments.

Adding the laterals changed the growth dynamics of the roots for each set up which can be seen on figure 3.5. Since we are only measuring the lateral root lengths, in the beginning all the lengths are zero, and afterwards they start growing more rapidly. The LN/LN and HN/HN lines cross each other which can be attributed to more laterals forming in HN/HN initially, but because of the systemic repression the growth slows down, whilst LN/LN starts foraging. The results of the new model do not overlap with the initial model. They are not completely comparable because the calculations are slightly different, therefore the order in which the roots appear was of significance.

3.4 Conclusion

The changes made to the FSPM represent root growth more accurately than in previous models. This involves a 3D physical representation using gravitropism and the patterns of LR emergence. The taproots and lateral roots are clearly separated, allowing for accurate measurements of lateral root growth. At the same time the initial growth function is kept with the aid of controlling LR emergence depending on nitrate availability. However

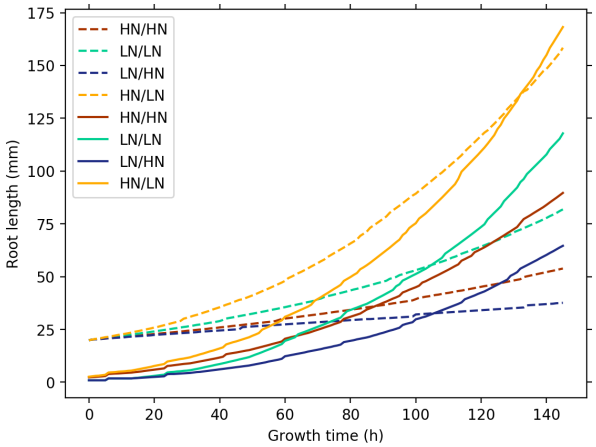


Figure 3.5: The growth rate of the LRs shown of figure 3.4. The smallest laterals form on the heterogeneous low side (blue) indicating a systemic supply signal. The homogeneous high environment (red) is the second smallest. This is followed by the homogeneous low side (green), and then the heterogeneous high side (yellow) showing systemic demand signalling. The new model is shown with a continuous, the initial model with a dashed line.

it remains unknown what signalling pathways determine that more laterals appear in a nitrate rich environment. Additionally, realistic carbon transport and allocation is missing, therefore in the next chapter transcriptomics data is analysed in order to find genes that could play a role.

Chapter 4

Analysis of split root data

Underlying a plant's morphology and physiology are molecular pathways, and through our data analysis we aimed to find certain genes and modules that have already been identified to play a key role in root growth adaptation, and other genes that could be crucial to the development of the root system architecture (RSA) in response to nitrate.

The data analysis was conducted on a single-channel microarray data set of split root experiments on *Arabidopsis thaliana* by Ruffel et al. [2]. Microarrays employ probes, also known as oligos that are bound to a solid surface, – such as glass [31]. A probe is a short section of a gene and as such, a gene can have many different probes [32]. There are two types of microarrays, two- and single-channel. In a two-channel experiment, probes are obtained from two samples and labelled with two fluorescent dyes, Cy3 and Cy5. The two samples are hybridised to the same microarray chip and the relative mRNA levels are determined from the Cy3/Cy5 signal [33]. In single-channel arrays, genes are usually represented by 16–20 pairs of probe sets. A probe pair is made up of a perfect match (PM) and a mismatch (MM) probe. One sample is hybridised at a time, and signal intensities are compared from different chips [32, 33]. Both types of microarrays allow gene expression levels to be compared across several RNA samples. Examples of raw microarray images taken from the data set of Ruffel et al. are shown on 4.1.

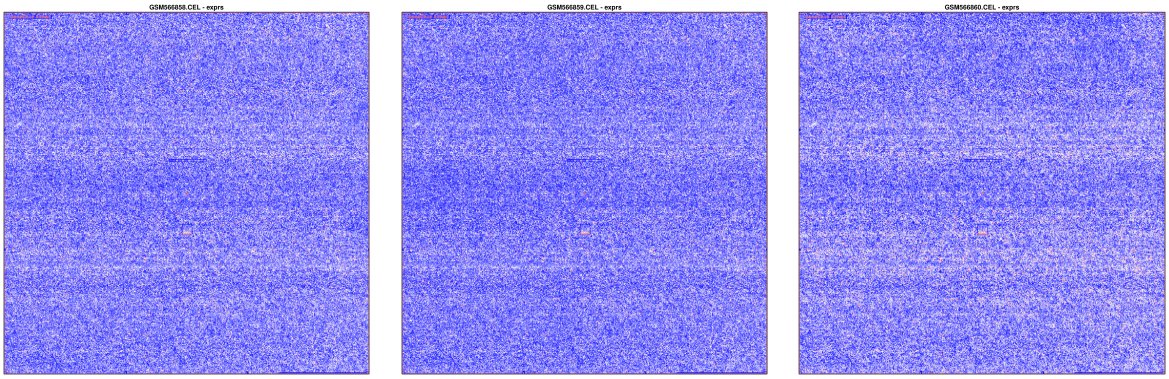


Figure 4.1: Three raw images of microarray data from the Ruffel data set. Each represents a sample of KNO_3 treated roots after 2 hours.

The data set is based on roots being grown in varying nitrate conditions, and is made up of 36 experiments on 22810 probes. The 36 experiments consist of 3 samples in 3 time points (2 hours, 8 hours, 2 days) and 4 environmental conditions (KNO_3 and KCl in homogeneous soil, KNO_3 and KCl in heterogeneous soil). Ruffel et al. found that there are

genetically distinct systemic signalling pathways, a long distance nitrogen supply signal and the importance of CK in N demand signalling [2]. Several articles applied the results in crops, such as the role of cytokinin in barley [34] and rice [35], or systemic N signalling in graminaceous plant species [36]. Others combined it with other nutrient deficiencies and stresses [37].

Microarrays compare gene expression levels across several RNA samples. Identifying these genes is a complicated task as there are thousands of genes, therefore a large number of tests are needed. There are various conditions, resulting in differing expression levels and the distributions are often not normally distributed. There are multiple methods available to find differentially expressed genes, however each one finds a different gene list resulting in further complications in deciding whether a gene is indeed differentially expressed or it is only a result of a test statistic. It is essential that the right feature selection method is used based on the number of samples and genes, and the noise in the data set to avoid wrong conclusions. The simplest approach, hence one of the most popular approaches used are t-statistic methods. These are based on the t-test which assumes that the means are equal in two groups. It has been shown that on small samples it performs poorly as the false discovery rate is high, the statistical power is small [38, 39] and the p-values are unreliable [40]. Microarray data faces two issues, the level of noise is high and the sample size is often small. Combined, these falsely predict genes to be differentially expressed. The variance estimate can be skewed by the genes that have a low variance, and because there are a large number of genes, there are always some that have a low variance by coincidence. Analysis of variance (ANOVA) is a more advanced t-statistic method, and was originally used on the data set we reanalysed. Unsurprisingly, in many aspects ANOVA does not show any improvement compared to the basic t-test as it calculates the standard t-statistic [38, 39].

Empirical Bayes methods, such as Linear Models for MicroArray data (limma) performed the best on multiple data sets, including simulated and real-life, across varying sample sizes [38, 39, 40]. Empirical Bayes methods are useful if the number of observations are small as they offer a more complex model for calculating the gene variance. Empirical Bayes estimates the prior distribution from the data. The moderated variance is a weighted average of the prior's variance and the observed variance for each gene. Essentially, the observed variances are shrinked towards the mean variance. Using the moderated variance gives a better estimate of the standard deviation compared to the t-test that uses the sample variance [10]. The advantage of the t-test and ANOVA is that they execute very quickly, though limma is almost as efficient [39]. Therefore we used limma in our analysis.

`limma` uses a general linear model to analyse experiments on gene expression data and is implemented as an R-Bioconductor package [9]. In the following section some of theory behind the package is presented together with our implementation.

4.1 Linear modelling

The central idea of the package is to fit a linear model to the expression data for each gene. We have $n = 36$ microarrays as there are 3 samples in 3 time points and 4 environmental conditions. The response vector for the gth gene is $\mathbf{y}_g^T = (y_{g1}, \ldots, y_{gn})$. However, based on the time points this was split into separate models, resulting in $(y_{g1}, \ldots, y_{g12}), (y_{g13}, \ldots, y_{g24})$, and $(y_{g25}, \ldots, y_{g36})$. The responses in our data are log-intensities as we have single channel data. Then we can assume that the expectation is

$$E[\mathbf{y}_g] = X\boldsymbol{\alpha}_g \tag{4.1}$$

where X is the design matrix and $\boldsymbol{\alpha}_g$ is a vector of coefficients. There are 4 columns, one for each scenario, KCl.Control, KCl.Split, KNO3.Control, KNO3.Split. There were 3 repeats for each scenario, meaning $4 \times 3 = 12$ rows. The design matrix in our case is

$$X = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \tag{4.2}$$

α_{gj} represents the expression level at the jth experiment. The variance is assumed to be

$$Var(\mathbf{y}_g) = W_g\sigma_g^2 \tag{4.3}$$

with W_g a non-negative definite weight matrix.

Our data is of a single channel type, therefore it can be analysed similarly to an ordinary univariate linear model, with the additional step of extracting the contrasts that are of interest. These are given by $\boldsymbol{\beta}_g = C^T\boldsymbol{\alpha}_g$ and we want to test whether contrast values β_{gj} are equal to 0. In this model we make four pairwise comparisons, that reflect the demand and supply, for both local and systemic signalling. The contrast matrix C contains the comparisons in its columns in the order of KNO3.Control - KCl.Control, KNO3.Split - KCl.Split, KNO3.Control - KNO3.Split, and KCl.Control - KCl.Split. The rows are KCl.Control, KCl.Split, KNO3.Control, and KNO3.Split respectively. Then the contrast matrix is

$$C = \begin{pmatrix} -1 & 0 & 0 & 1 \\ 0 & -1 & 0 & -1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 \end{pmatrix} \tag{4.4}$$

The linear model is fitted for all genes g in order to get the coefficient estimators $\hat{\boldsymbol{\alpha}}_g$, s_g^2 of σ_g^2, and the estimated covariance matrices

$$Var(\hat{\boldsymbol{\alpha}}_g) = V_g s_g^2 \tag{4.5}$$

where V_g is a positive definite matrix that is independent from s_g^2. The contrast estimators are $\hat{\boldsymbol{\beta}}_g = C^T\hat{\boldsymbol{\alpha}}_g$ with covariance matrices

$$Var(\hat{\boldsymbol{\beta}}) = C^T V_g C s_g^2 \tag{4.6}$$

The jth diagonal element of $C^T V_g C$ is v_{gj}. Finally, the ordinary t-statistic can be given

$$t_{gj} = \frac{\hat{\beta}_{gj}}{s_g\sqrt{v_{gj}}} \tag{4.7}$$

We want to test the null hypotheses $H_0 : \beta_{gj} = 0$ [10]. In this case what is of interest to us is $H_0 : \beta_{g3} = 0$ and $H_0 : \beta_{g4} = 0$ individually because the systemic signalling is in the last two columns, the third column for high and the fourth for low nitrate.

Further theoretical details of limma are described in the paper of Smyth, including the moderated t-statistic [10], and Ritchie et al. describe how to implement it within R [9].

4.2 Quality check

Microarrays can contain many anomalies, the chips can be damaged, the experiment could have been conducted wrong, the equipment could be malfunctioning. If a sample is much lighter or darker than the others, or some unexpected artifacts appear, that could be an indication of bad quality data which needs to be removed. To avoid using faulty data, we carried out a quality check.

Pre-processing and visualising the data was done with the `oligo` package [41]. Microarray images are very noisy and no patterns or artifacts should appear. Therefore the quality of the data was first checked by observing the raw images of the data. The image plots did not show any deformation, hence all samples were kept.

Observing the box and density plots of the probe-level data can also show abnormalities. The box plots should appear in similar ranges, where the interquartile ranges and boxes closely align. The density plots should overlap and not show bimodality. A distribution that is far from the rest indicates a chip with abnormally high or low values. A bimodal distribution is usually a sign of a spatial artifact in the data [42, 43]. The density plots on the left largely overlap and there are no signs of bimodality. On the right, the box plots and the interquartile ranges also indicate good quality data. They are between 4-14 on the $\log_2$ scale and the boxes closely align too [32, 42].

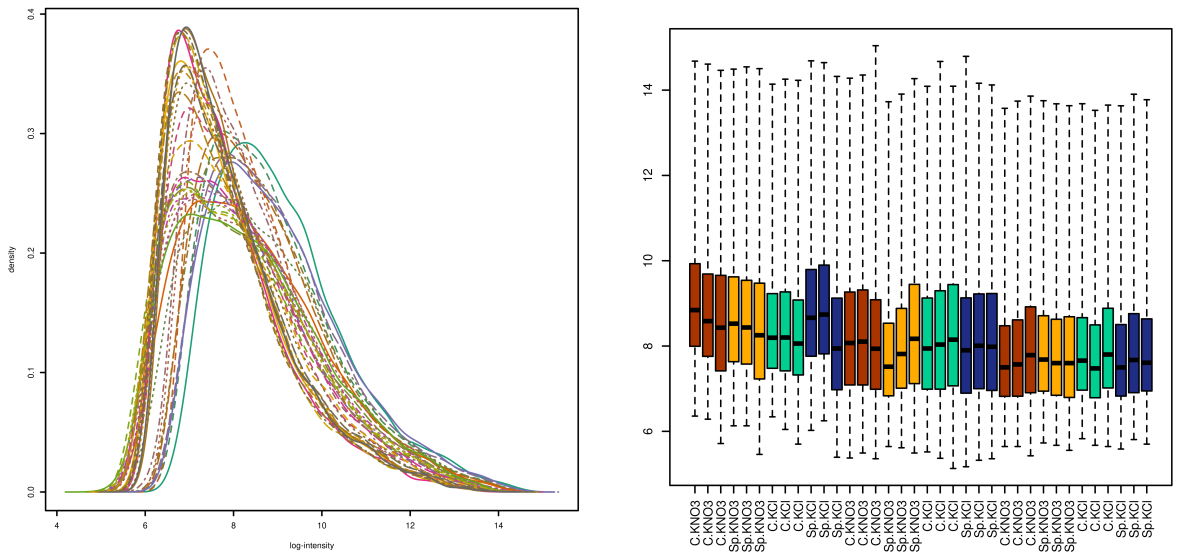


Figure 4.2: Histogram and boxplot of raw data. The density plots on the left do not show anomalies. The boxplots on the right appear comparable ranges.

In a microarray data set samples are highly correlated, as the same set of genes are used across all repeats. In biomedicine, the Bland–Altman plot [44], also known as the

Tukey mean-difference plot [45] is widely used to measure the agreement between two samples by plotting the differences against the means. The aim of the mean-difference plot is easier visual assessment, as it is easier to observe deviations from a horizontal line than from a slope. Instead of direct comparison of two samples with the linear line $y = x$ as the perfect alignment, the coordinate system is transformed, and the horizontal line $y = 0$ becomes the perfect alignment [45]. Bland and Altman argue that high correlation does not imply agreement, therefore they advise against using the correlation coefficient, and plotting two samples against each other is not informative enough as the data points will be closely clustered. Thus they suggest to plot the differences against their mean [44]. The MA-plot is an application of the Bland-Altman plot, and was initially implemented for two channel data. A plot is created from one array, where the log intensities of the red and green channel are compared. The difference in log expression values $M = \log_2 R/G$ is plotted against the average of the log expression values $A = \log_2 \sqrt{RG}$ [46]. The data used here is single channel, so two arrays are compared at the same time. For two arrays i, j and intensities x_{ki}, x_{kj} where $k = 1, \ldots, p$ are the probes, the intensity ratio is

$$M_k = \log_2 \left(\frac{x_{ki}}{x_{kj}} \right), \tag{4.8}$$

and the average is

$$A_k = \frac{1}{2} \log_2 \left(x_{ki} x_{kj} \right). \tag{4.9}$$

A normalisation curve is fitted to the data points using the loess regression. The assumption with these plots is that the experiments are similar, therefore the loess curve and most of the data points should appear close to $M = 0$. The data points that are outliers are the genes that are potentially differentially expressed [47]. In order to create MA-plots for a set of single-channel arrays, an artificial array is created as a reference by averaging over all the other arrays [9]. On figure 4.3 as examples, the first three arrays are plotted against the reference. The loess curve is typically not affected by the small fraction of differential genes, and it should be a smooth curve. An oscillating loess curve would suggest bad quality data [42, 43], and the plots here do not show a sign of that. A line above or below suggests that there is a bias towards a positive or a negative log fold change in an array compared to the rest. A decreasing line suggests that genes with lower intensities are more likely to be up-regulated, and the ones with higher intensities are more likely to be down-regulated. For an increasing line, it is the opposite. The differences between the loess lines of the samples and the fact that they do not align with $M = 0$ mean that normalisation is necessary [47].

Based on the histogram, the image, bar and MA-plots, all the samples were good quality, therefore none of the samples have to be discarded and we could proceed to the normalisation of the data.

4.3 RMA algorithm

The data cannot be compared against each other in its raw form as there is noise and variation in the measurements which could be due to laboratory conditions, labelling, physical problems. Not taking into account the noise and using the raw data could result in incorporation of batch effects or even wrong results. Therefore, the data was transformed using the RMA algorithm [47, 48, 32]. This involved background correction,

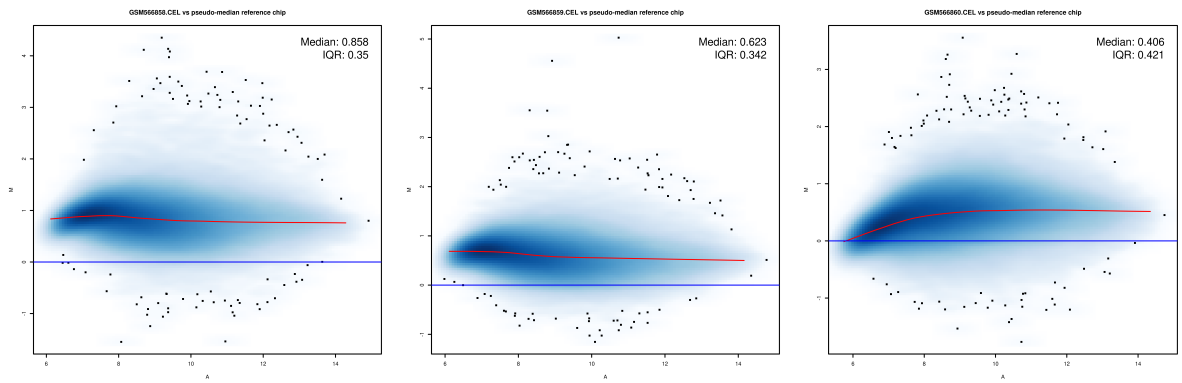


Figure 4.3: MA plots of KNO_3 treated roots after 2 hours against a pseudo-median reference chip. The data points that are non-differentially expressed appear in blue and the outliers are black dots. $M = 0$ is the loess regression line indicated with red. It does not oscillate suggesting the data is good quality.

normalisation and summarising expressions. For the normalisation, each array needs to be normalised against all of the others which is achieved through quantile normalisation [32]. Plotting the cleaned and normalised data, the boxplot and histogram on figure 4.4 and the MA plots on figure 4.5 show that the RMA algorithm was successful [42].

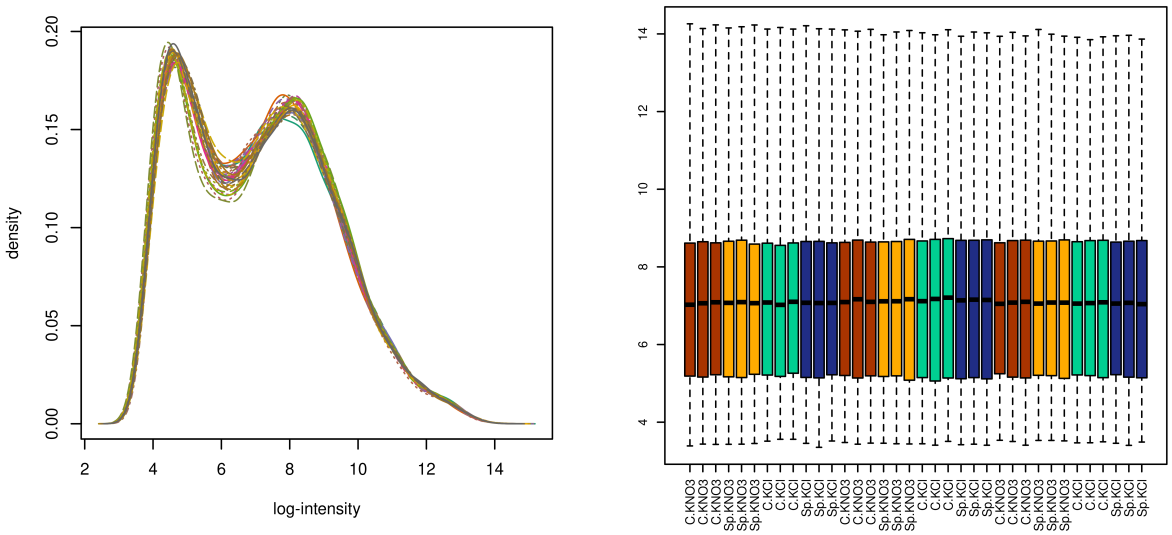


Figure 4.4: Density and box plot of normalised data. The density plot is bimodal, which is the expected result, the higher peak corresponds to the true signals, the lower to the background noise. The box plots are aligned, the normalisation was successful, therefore we can proceed with the analysis.

4.4 Result of the statistical analysis

The data is dependent on three variables, time, nitrogen content and environment (split or control). Genes can exhibit large differentiation whether they belong to 2 hours, 8 hours or 2 days, thus we created a multidimensional scaling (MDS) plot on the experiments

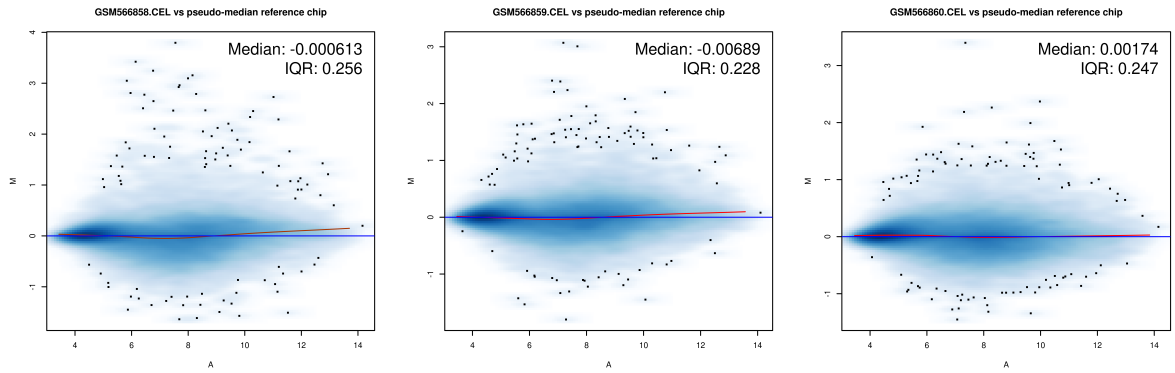


Figure 4.5: KNO_3 treated roots after 2 hours against a pseudo-median reference chip (normalised). The red loess regression line almost completely aligns with the blue line, further supporting the success of the RMA algorithm.

which resulted in figure 4.6. As visible on this plot, the different time points are indeed closely clustered which is expected, however the treatments do not. This makes it difficult to find the changes that are happening due to the different treatments. Therefore we split up the data into 3 separate experiments for each time point. In each time point the split root condition was compared against the control for each plant, and the high and low nitrate sides.

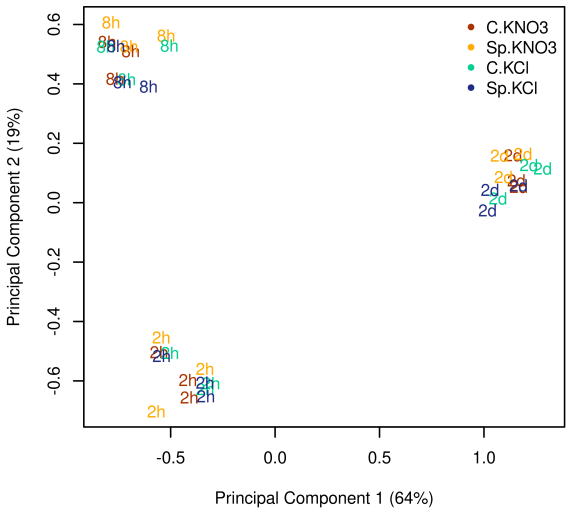


Figure 4.6: MDS plot on all the experiments. Multidimensional scaling creates a matrix of Euclidean distances based on the expression levels of the samples. All the genes are averaged into one point, and the distances represent the $\log_2$-fold changes between the experiments. The data points cluster based on time, not on treatment.

We checked if any clusters appeared for the separate time points on figure 4.7. We were focusing on systemic, not on local signalling, therefore we looked for differences recorded between Sp.KCl and C.KCl, and Sp.KNO3 and C.KNO3 samples. The previous is indicative of a N-supply, the latter of a N-demand long-distance signal. Initially, roots respond to local signalling, visible on the MDS plot too. After 2 and 8 hours the data

points can only be separated based on whether the nitrate amount was high or low. After 2 days, the experiments clustered more clearly, therefore this is where systemic signals were likely to be found. However, the small log-fold changes made picking up signals more challenging. This partly corresponds to the results of Ruffel et al. [2] where after 2 hours only local signalling occurred, however using dendograms they noticed patterns appearing after 8 hours. After conducting the linear modelling we were not able to find the same as the results after 8 hours had high p-values thus statistically unreliable, proving our assumptions based on the MDS plot true.

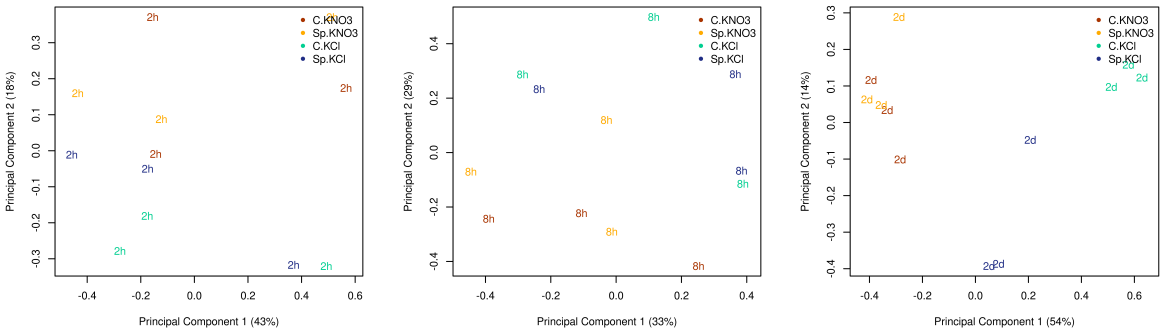
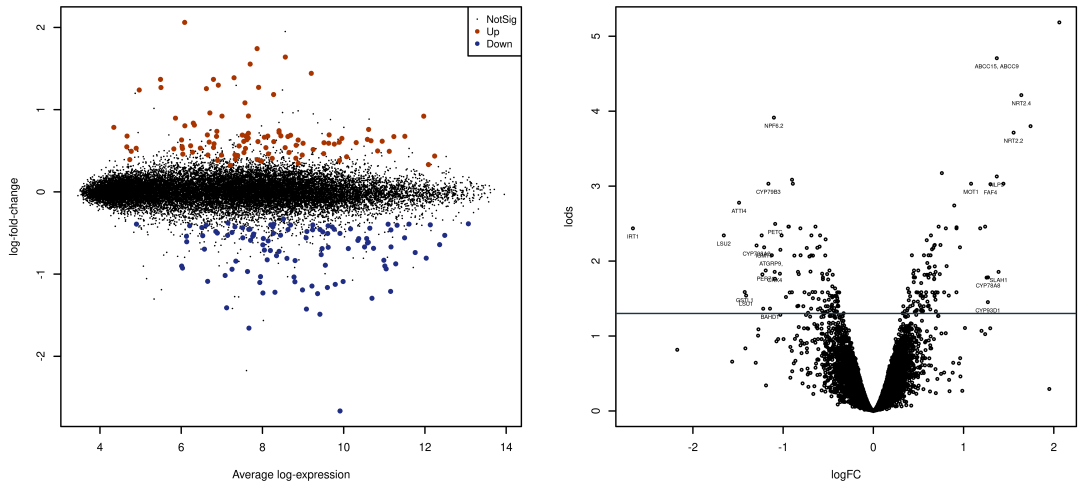


Figure 4.7: MDS plot in the three different time points. After 2 hours, the data points are scattered, however they can be separated on nitrate content. After 8 hours, the separation is still based on nitrate availability. After 2 days, the experiments cluster based on the condition too, the control clearly separates from the split for the low nitrate side.

After 2 days we found several genes that were differentially expressed in nitrate deprived conditions which is presented on figure 4.8a. From here on, the genes that we focused on were the ones of the volcano plot on figure 4.8b that appear above the horizontal line. A volcano plot is a scatter plot that plots the strength of a gene's strength ($\log_2$-fold change) against their statistical significance ($-\log_{10}$ of the p-value). The most interesting genes therefore are the ones on the top left and right corner as these show the biggest log-fold changes whilst having the smallest p-values.

In the list of potential genes not all were relevant as they were not amongst the target genes, nor were they related to root growth through the molecular pathways that we were searching for. We made a selection of 18 genes that play a role in N uptake or signalling, CK biosynthesis, and sugar transport. As noticeable on the heatmap on figure 4.9a there is a stark contrast between the control low nitrate condition compared to the rest of the experiments, even more so on figure 4.9b where the means of the experiments are taken. The dendograms on the sides order the genes based on complete-linkage hierarchical clustering using the Pearson distance metric as Gibbons and Roth showed that the complete-linkage method combined with Pearson correlation performed the best amongst hierarchical clustering methods on single-channel microarray data [49]. The clusters reflect the groups of genes discussed in the following sections. NRT2 genes appear in a big cluster, together with ROXYs which are known to regulate NRT2.1 [50]. The pairs NIA2 and HHO1, CKX4 and IPT5, CYP79B2 and CYP79B3, and SWEET16 and SWEET17 are clustered together, of which all perform similar roles.

Ruffel et al. [2] selected some sentinel genes by looking at the intersection of C.KNO3 vs. Sp.KNO3 and Sp.KCl vs. C.KCl then clustered them using dendograms. We aimed

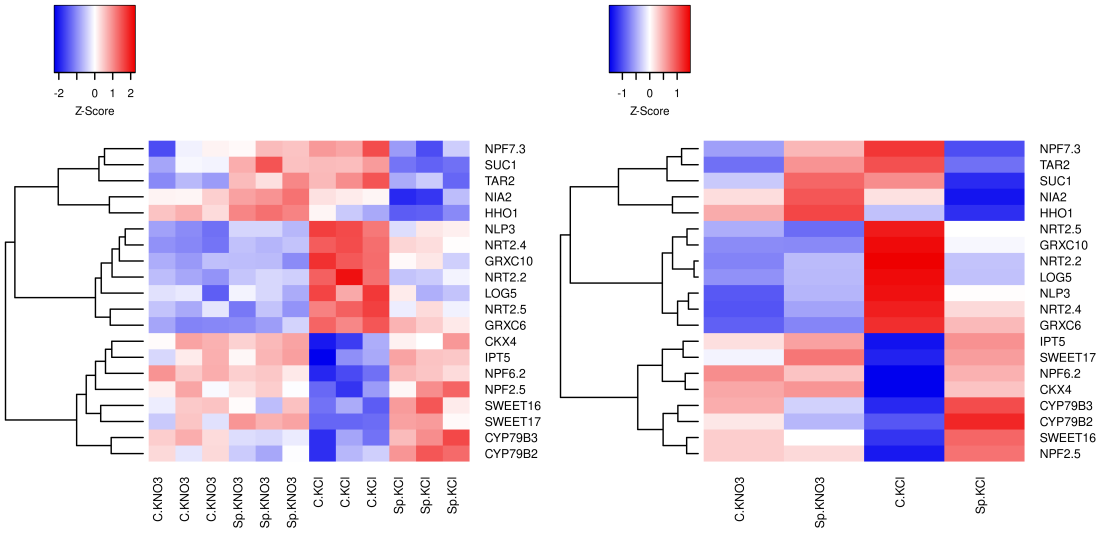


(a) Up and down-regulated genes after 2 days. The red points are up, the blue ones are down-regulated, and the black ones are not significant. The further a point is from zero, the higher the $\log_2$-fold change.

(b) Volcano plot highlighting the genes that showed the highest $\log_2$-fold change. The horizontal line indicates $p = 0.05$, the genes below it have $p > 0.05$. The genes that are statistically significant and have the highest log-fold changes have their names depicted.

Figure 4.8: Log-fold changes of the genes after 2 days.

to recreate the same heatmap with the dendogram based on the expression levels on the microarrays, but the results on figure 4.10a do not resemble the one seen in the paper. Nevertheless, IAA17, G6PD3, NIR1, and UPM1 do show a noticeable difference between Sp.KCl and C.KCl after 2 days. The authors also selected NIA2, though it does not show differences in the expression levels between experiments. NRT3.1 and NRT2.1 were added as well, though the latter was not found by their microarray analysis.



(a) Selected genes after 2 days with all repeats.

(b) Means of the experiments on selected genes after 2 days.

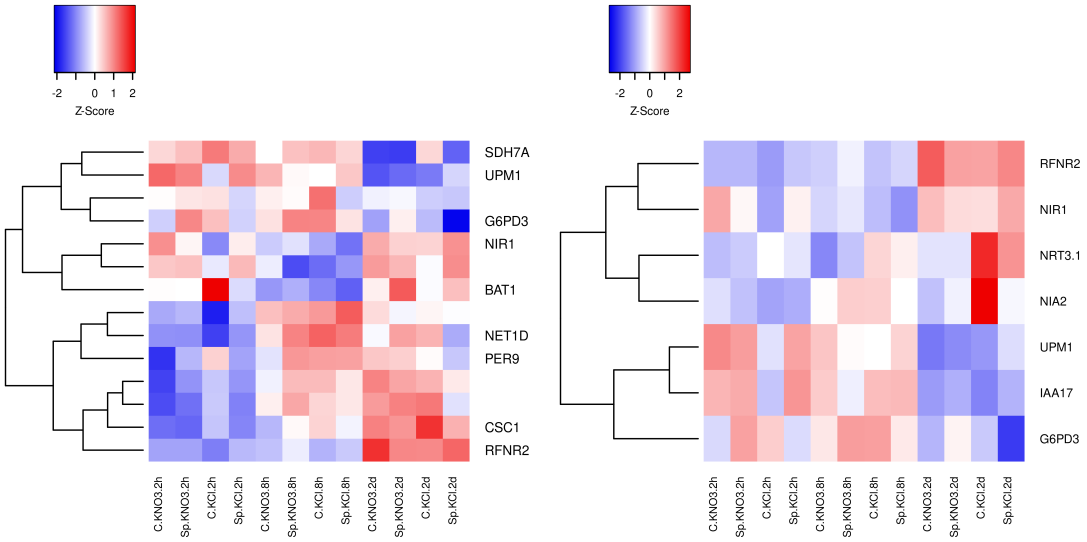
Figure 4.9: The heatmaps show the log-expression values of the selected genes across experiments after two days. They are clustered by relative changes in the expression, using complete-linkage hierarchical clustering combined with Pearson correlation. The Z-score represents the standard deviation of a gene from the mean, which is derived from the set of genes expressed.

4.4.1 NRT1 family

There are two families of nitrate transporters NRT1 and NRT2. Most NRT1s are low-affinity nitrate transporters which are activated if the nitrate concentration is high, with the exception of NRT1.1, which is a dual-affinity transporter responsible for both low- and high-affinity nitrate uptake. In total Arabidopsis has 53 NRT1 transporters [51, 52] of which we found NRT1.5/NPF7.3 (AT1G32450) and NRT1.4/NPF6.2 (AT2G26690) (figure 4.11).

NRT1.4 is required for petiole nitrate storage [52]. NRT1.4 is down-regulated under both limiting and stressful conditions, and up-regulated when N levels increase [15], which was shown during our analysis as well.

NRT1.5 is a low-affinity, pH-dependent bidirectional nitrate transporter. NRT1.5 is located in the plasma membrane and is expressed in root pericycle cells close to the xylem. Transport from the root to the shoot is achieved by NRT1.5 exporting nitrate out of the pericycle cells and loading it into the xylem vessels depending on the pH level in the xylem [51]. Surprisingly under limited nitrate, NRT1.5 was up-regulated. However NRT1.5 is known to respond slowly to changes, it only starts responding to high nitrate after 8 hours. In low nitrate even after 16 hours, the expression level remains unchanged [51]. On the right plot of figure 4.11, the expression levels are almost the same across all conditions after 8 hours. After 2 days the control low experiment is still unchanged, but the split root and the control high nitrate experiment show the expected response. The split high is bigger than the control high experiment, and the expression level of the split low experiment is smaller than the split high.



(a) Differentially expressed genes of Ruffel et al. [2].

(b) Selected genes of Ruffel et al. based on microarray data analysis [2].

Figure 4.10: The heatmaps show the log-expression values of the selected genes across experiments through each time point. They are clustered by relative changes in the expression, using complete-linkage hierarchical clustering combined with Pearson correlation. The Z-score represents the standard deviation of a gene from the mean, which is derived from the set of genes expressed.

4.4.2 NRT2 family

NRT2s are high-affinity nitrate transporters which are involved when there has been a prolonged nitrogen deprivation. The genes we found are NRT2.2 (AT1G08100), NRT2.4 (AT5G60770), and NRT2.5 (AT1G12940) which are involved in inducible high-affinity nitrate uptake (figure 4.12). The results show that as expected, NRT2.2 and NRT2.4 were induced by NO_3 while NRT2.5 was repressed [53]. NRT2.2 and NRT2.4 were among the genes that showed the highest log-fold change.

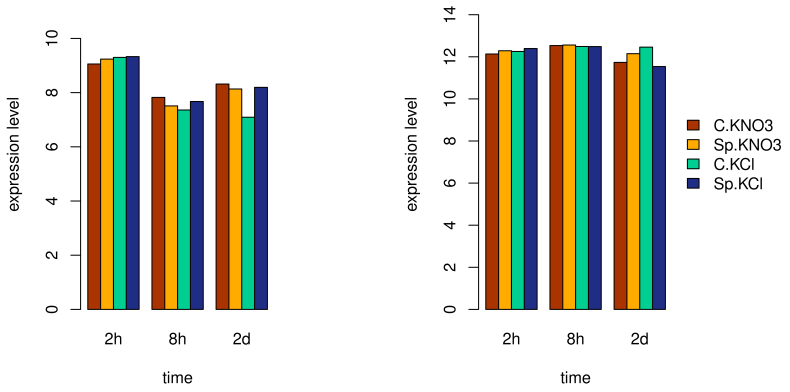
4.4.3 Auxin biosynthesis

Nitrate is able to regulate the expressions of genes that belong to auxin biosynthetic pathways, of which all affect lateral root development. TAR2 (AT4G24670) [21], CYP79B2, and CYP79B3 [54] are the auxin biosynthesis genes that appeared in our search. TAR2 plays a role in low nitrogen conditions [21], which was the case in our analysis as well as TAR2 was up-regulated under nitrogen deprived conditions. Plants overexpressing CYP79B2 and CYP79B3 boost auxin levels [54], aligning with our results where both were down-regulated under nitrate deficiency (figure 4.13).

4.4.4 Other nitrate related genes

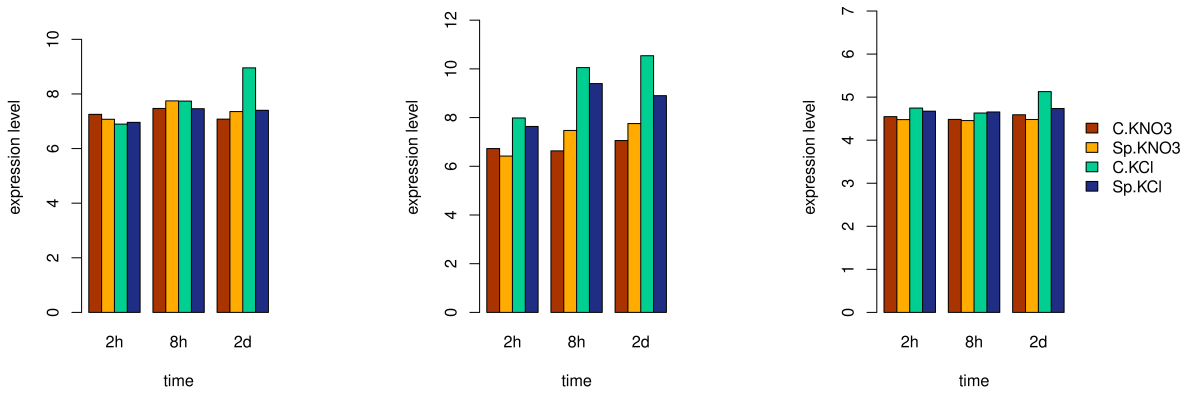
HHOs such as HHO1 directly repress NRT2.4 and NRT2.5, thus HHO1 is an important player in the signalling pathway in response to nitrogen starvation [55]. HHO1 was slightly up-regulated.

NIN-like proteins (NLP) regulate many nitrate signaling and assimilation genes, such



NRT1.4 NRT1.5

Figure 4.11: Expression levels of NRT1 family members through time.



NRT2.2 NRT2.4 NRT2.5

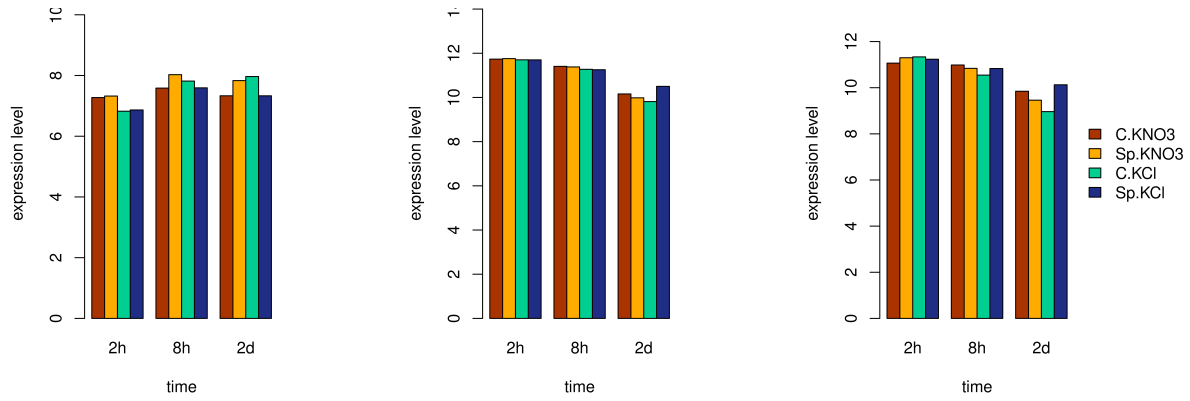
Figure 4.12: Expression levels of NRT2 family members through time.

as NLP3 (AT4G38340) which activates genes involved in nitrate assimilation. Enhancement of NLPs promote, and suppression of NLPs downregulate nitrate-inducible gene expression, including the nitrate reductase gene NIA2 (AT1G37130) [56]. Both NIA2 and NLP3 were up-regulated, with NLP3 showing one of the biggest positive log-fold changes amongst the genes. NIA2 is activated by NLPs, therefore it was less significant than NLP3 (figure 4.14).

4.4.5 CEP

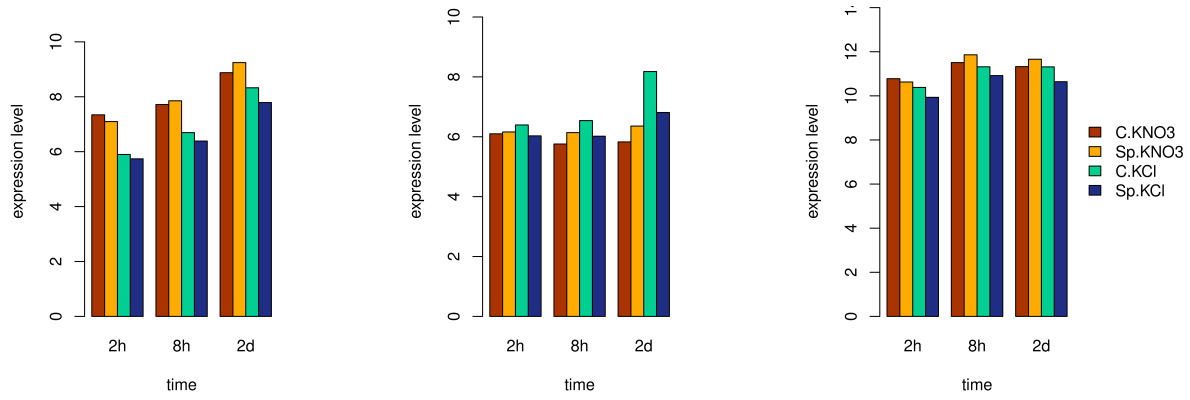
N deprived roots trigger CEP production, which are then translocated to the shoot and bind to CEP receptors [17]. CEPRs cause the production of CEPD1/ROXY6 and CEPD2/ROXY9. These downstream signals up-regulate the expression of the nitrate transporter gene NRT2.1 [18].

The other CC-type glutaredoxins (ROXY) are also differentially expressed under N starvation. The expression of six ROXYs (ROXY6,8,9,19–21) are significantly up-regulated by nitrate deficiency. CEPD1 and CEPD2 are not the only ROXYs that up-regulate the transcription of NRT2.1 and promote root hair elongation, but so do their paralogs, ROXY8,9,19-21 [50]. ROXY20/GRXC10 (AT5G11930) and ROXY21/GRXC6 (AT4G33040) appeared in our search. Both were up-regulated, consistent with the NRT2s. ROXY20 showed strong signals, whereas ROXY21 had a lower log-fold change (figure 4.15).



TAR2 CYP79B2 CYP79B3

Figure 4.13: Expression levels of auxin biosynthesis genes through time.



HHO1 NLP3 NIA2

Figure 4.14: Expression levels of HHO1, NLP3, and NIA2 through time.

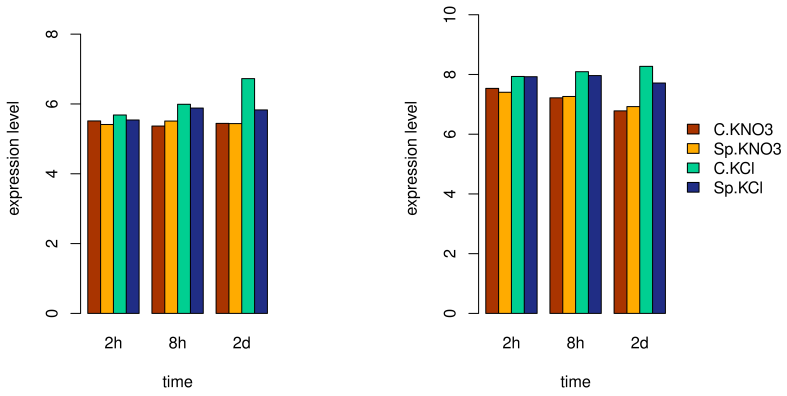
4.4.6 Cytokinin

Cytokinins are hormones that play a crucial role as negative regulators in lateral root formation and growth. The steps of cytokinin biosynthesis are the following. First the isopentenyl transferases (IPT) catalyse the biosynthesis and isopentenyl adenine (iP) metabolites are formed. Then part of iP metabolites break down to trans-zeatin (tZ) metabolites with the aid of hydroxylase enzymes CYP735A. iP and tZ cytokinins release active cytokinins with the aid of genes called LONELY GUY (LOG). Finally, from iP and tZ cytokinins CK oxidase/dehydrogenase (CKX) enzymes catalyse cytokinin breakdown [57]. The steps are also presented on figure 4.16.

The genes that we found are IPT5 (AT5G19040), LOG5 (AT4G35190) and CKX4 (AT4G29740) which are all very responsive cytokinin metabolism genes. As expected high nitrate down-regulated LOG5, while up-regulated IPT5 and CKX4 [22] (figure 4.17). IPT5 plays a central role LR formation, it suppresses the initiation of new LRs [57].

4.4.7 Unexpected genes

Some genes that were found had relations to chlorophyll and photosynthesis which one does not expect in the root. This could be in relation to how the experiment was conducted, such as exposing the root to light. However, there has also been proof that a high number of mRNAs travel between tissues, such as the one produced by the gene LHCA3



ROXY20 ROXY21

Figure 4.15: Expression levels of genes playing a role in CEP signalling.

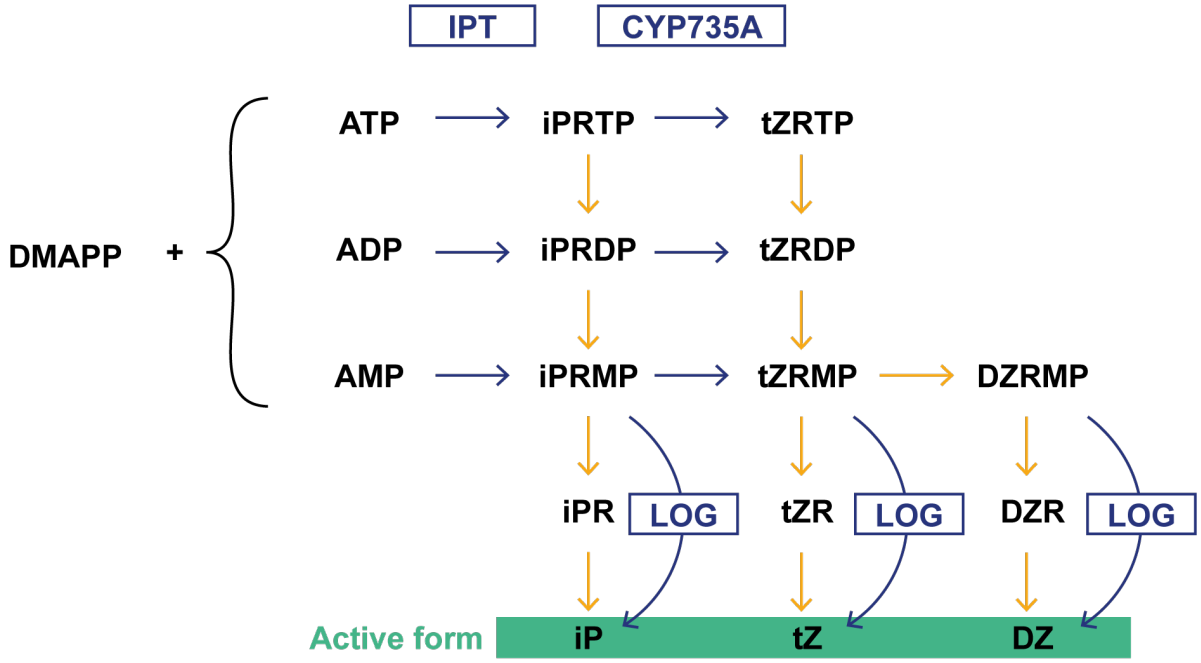
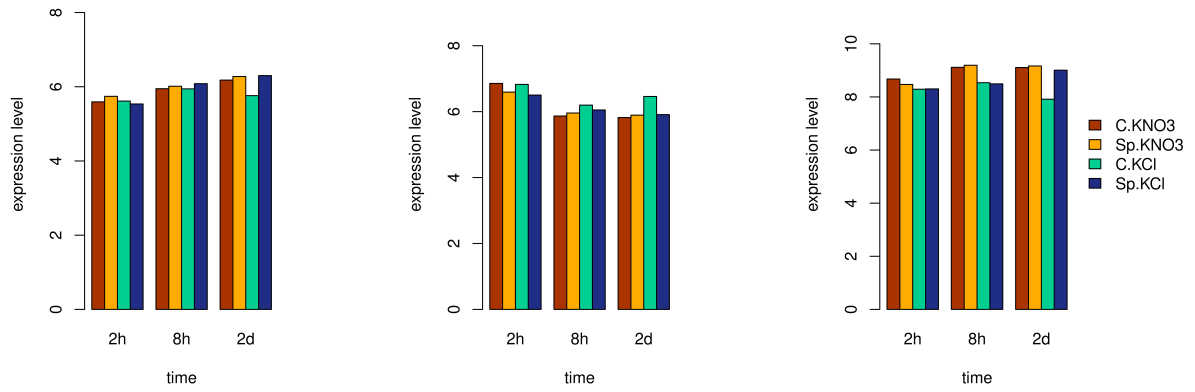


Figure 4.16: Cytokinin pathway.

(AT1G61520). LHCA3 is expressed in root tissue and can be transported from both root to shoot and shoot to root [58].[1]

We also found some genes related to sugar transport such as SWEET16 (AT3G16690) and SWEET17 (AT4G15920) belong to clade IV of SWEETs and are expressed in root cortical cells. (There are 4 clades. All except for III mainly transport monosaccharides. Clade III members transport disaccharides.) SWEET17 can export and import fructose, during normal conditions it acts as an importer, while in unfavourable conditions it becomes an exporter. SWEET16 is redundant as it behaves similarly [59]. Both of these transporters were down-regulated under N deprivation, suggesting that the plant is exporting the nutrients to its organs where the sugar is needed, instead of importing it to an organ where growth will not be possible. In addition we found SUC1 (AT1G71880) which is responsible for sucrose unloading in roots and is one of the main transporters of

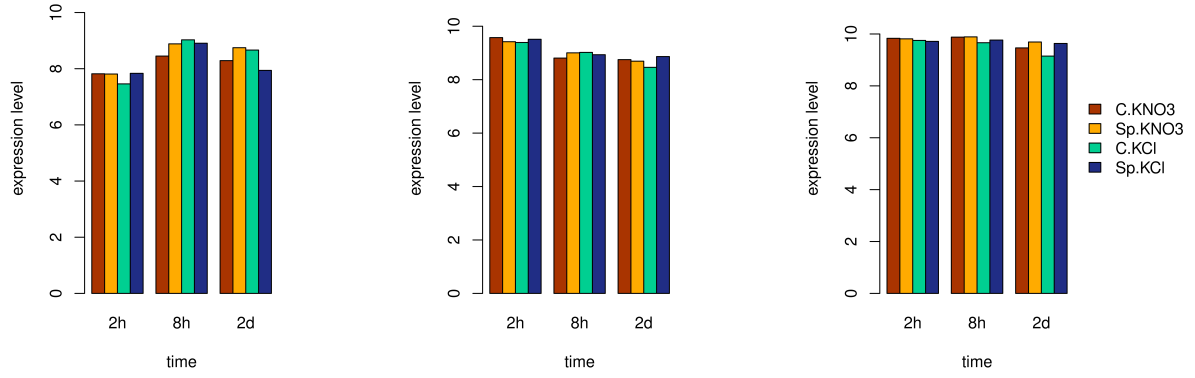
[1]The Lhca proteins are associated with the light-harvesting complexes of photosystem I (PSI) and the Lhcb proteins are associated with those of photosystem II (PSII).



IPT5 LOG5 CKX4

Figure 4.17: Expression levels of cytokinin genes through time.

the SUC family in roots. It is expressed in the root vasculature in the elongation zone of the root and sites of lateral root formation [60, 61]. Interestingly, SUC1 was upregulated in split root compared to control conditions, when N was scarce. However, this is a systemic signal, and the local signalling matched the expectations and SUC1 was upregulated under nitrogen rich conditions (figure 4.18).



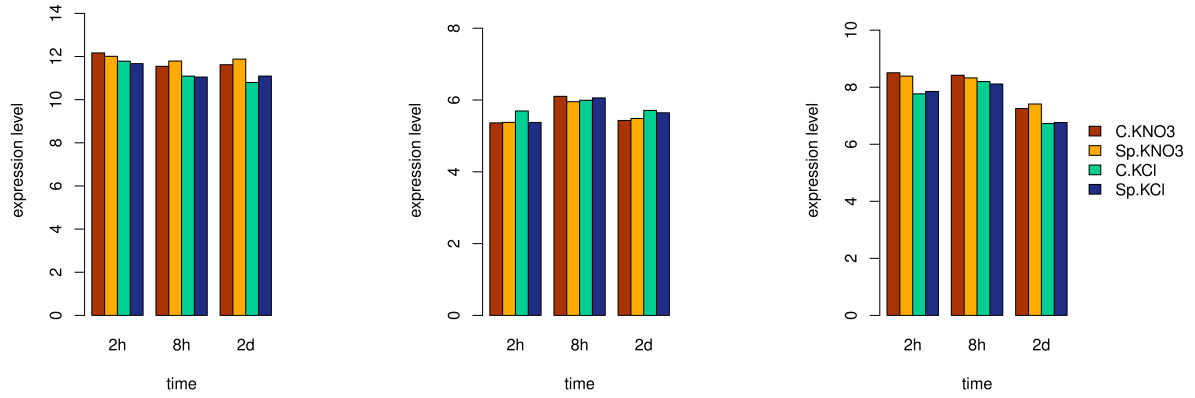
SUC1 SWEET16 SWEET17

Figure 4.18: Expression levels of sugar transport genes through time.

4.5 Discussion

We found numerous genes, that are key players in nitrogen transport, to be differentially expressed, such as low (NRT1.4, NRT1.5, NPF2.5) and high (NRT2.2, NRT2.4, NRT2.5) affinity nitrogen transporters, cytokinin related genes (IPT5, CKX4, LOG5) and others regulating nitrogen signalling. However, several of the most well-known genes related to N-signalling did not show significant difference between control and split root environments. NRT1.1, the best known nitrate transporter and the CEPs we aimed to identify only appeared in local signalling. Some of these are shown on figure 4.19. NRT1.1 and CLE2 clearly had local signals, the high nitrate roots' expression levels are bigger than the nitrate deprived one. On the other hand, CEP1 barely changed through time.

Our findings are not unlike Ruffel et al. [2] who also only found a few genes that are directly involved in lateral root development and growth. They did not identify NRT1.1 or



NRT1.1 CEP1 CLE2

Figure 4.19: Expression levels of NRT1.1, CEP1, CLE2.

any genes from the NRT1 family, nor NRT2.1, however they found NRT3.1 the partner of NRT2.1 which is also responsible for N uptake. N assimilation genes like NIR1 were found, and G6PD3 and RFNR2 which also play a role in assimilation. Though minimal, there was some overlap between the genes of ours and Ruffel et al.'s, most notably NRT2.2, NRT2.4, HHO1, NIA2, NLP3, and SUC1. We considered these to be important, but they did not mention them in their analysis. In our analysis we saw a significant correlation based on time, for which Ruffel et al. did not account for, especially the control KCl stood out from the other experiments. It would be expected that one of KNO_3 experiments would show a stark contrast as well. In the paper they also did not find significant changes after 2 hours, as did we. However they found that after 8 hours genes were differentially expressed. This contradicts our findings, though their clustering was not based on nitrate availability, but root length. Ruffel et al. used the transcriptomics data set, but their results rely heavily on q-PCR tests. NRT2.1 was not found in the microaarray data analysis, and was added on the basis of sharing similarities with NRT3.1.

In order to find out the reliability of Ruffel et al., we searched for papers that investigate the data set, but only few papers citing the results of Ruffel et al. do so. Araya et al. found CLE3 to be induced by N deprivation [23], however it is unclear what their methodology was in analysing the data set. No description was given, only an image with the relative mRNA levels of CLE3 suggesting that they did not use any statistical tests, and only relied on measuring log-fold changes. Canales et al. combined numerous data sets, including the one from Ruffel et al. and found more than 2000 genes. They did not describe in detail which genes were found in which data set and they used a different statistical method, hence a direct comparison between our results is not possible. Amongst the genes that they highlighted were ones inducing auxin biosynthesis, such as TAR2, CYP79B3, and CYP79B2 [62]. In conclusion, the results found in this thesis are considered more reliable as time was accounted for and a better statistical method was used [38, 39, 40]. Furthermore, the experimental setup of measuring after 2, 8, and 48 hours raises questions of why there was such a big gap left between 8 and 48 hours. Most likely there were changes happening during that time frame and it would have improved with finding more details on how gene expressions change.

There are different explanations why many important players did not appear in our search. Looking at the MDS plot on figure 4.6, we can see that the ranges are quite small and the experiments are clustered based on time. Catching signals from data like this is hard, even if differentially expressed genes are found, they often have to be discarded

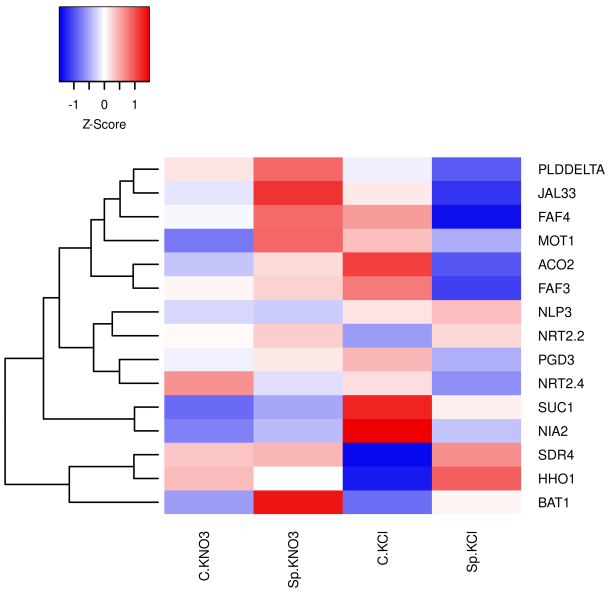


Figure 4.20: The heatmap shows the log-expression values of the genes that were found by both Ruffel et al. and us after two days. They are clustered by relative changes in the expression, using complete-linkage hierarchical clustering combined with Pearson correlation. The Z-score represents the standard deviation of a gene from the mean, which is derived from the set of genes expressed.

because the results are unreliable, i.e. the p-values are high. It is a possibility that microarrays are not the best way to find all of the signals, the signals found using molecular studies are not strong enough, and they only appear when they are knocked out or when they are up-regulated using genetic techniques.

Chapter 5

Conclusion and Discussion

The root system architecture of plants allows them to adapt to their environmental conditions, such as to the availability of nutrients including nitrate. There has been extensive research on the RSA's response to nitrate availability which has shown that roots can exhibit various responses, survival when nitrate is very low, foraging when low, and repression when high. Split root experiments showed that roots grow significantly quicker on the high nitrate side, and slower on the low side much more than their counterparts grown in homogeneous conditions. Several molecular pathways related to root growth are already known, and Boer et al. [1] successfully combined signals in their model. They included local nitrate, CEP, and CK signalling, systemic nitrate dependent foraging and repression. However the model was simplified and there were possible signalling pathways left to be included, therefore our research was two-fold, consisting of a modelling and a data analysis part.

During the modelling we created a FSPM based on the XL programming language that realistically models the initial development of the roots. The model incorporated all of the signalling mentioned in the Boer model [1]. The initial model consisted of two taproots made up of compartments, and a shoot, and the growth of these roots followed the results of the Boer model. However a root architecture does not only consist of taproots, but lateral roots as well. In experiments it is visible that the laterals are the part of the root architecture that perform the preferential root foraging behaviour, whilst the taproots grow at a similar rate. Thus the root lengths reported in papers are based on lateral roots. We wanted to calculate lengths the same way and extended the FSPM with laterals. The signalling pathways stayed the same, but the taproots became "transit areas" where they served nothing more than to have a place for the laterals to grow on and to connect them with the shoot. Furthermore we added a pattern of stopping laterals to emerge on low nitrate sides as in experiments we saw that there were 20-25% more LRs on high nitrate than on low nitrate sides [28]. This was a critical addition as it allowed us to use the growth function instead of the carbon-sink function that Boer et al. used to compensate for the root lengths not completely appearing in the right order. Without the carbon-sink function, the heterogeneous high side of the plant was the longest, followed by the homogeneous low side, however the heterogeneous low side was longer than the plant experiencing homogeneous high nitrate conditions. This suggests that there is indeed another signalling pathway determining lateral root development in response to nitrate. NRT1.1 was found to play a key role in negatively regulating the auxin carrier LAX3 and the auxin biosynthesis gene TAR2, of which both are necessary for optimal LR development. However the exact mechanisms are still poorly understood

[63].

We experienced difficulties with the modelling software, GroIMP produced different results when run from the command line than from the IDE. After extensive research we were still not able to find where this problem originated from. Until the software problems are fixed a solution would be to create a field of plants with different growing conditions.

Transcriptomics offer a more efficient way of finding important genes in terms of as they allow hundreds of genes to be inspected at the same time, contrarily to mutant studies where genes are targeted one by one. We reanalysed transcriptomics data from Ruffel et al. [2] in the hopes of finding major nitrate signalling routes that were already incorporated in the model. This was in order to prove that this type of data indeed offers the possibility of relying less on mutant studies. The genes related to the molecular pathways of the FSPM were not found, however some genes that share similarities were discovered which are summarised in table 5.1. NRT1.1 did not appear in our analysis, however we found NRT1.4 and NRT1.5 which are both low-affinity nitrate transporters [52], with NRT1.5 being a bidirectional one [51]. We found numerous members of the NRT2 family (NRT2.2, NRT2.4, and NRT2.5) which are part of the high-affinity transport system just as NRT2.1 [53]. Though CEP/CEPR/CEPD1/CEPD2 were not observed, the ROXYs ROXY20 and ROXY21 did, which also play a role in CEP signalling [50]. The cytokinin genes IPT5, LOG5 and CKX4 [57] were discovered. Our other goal was to find more genes that play a role in nitrate uptake, transport, and storage, and the genes listed before met this criteria. NRT1.5 participates in nitrate uptake in LATS [51] and NRT1.4 stores nitrate [52]. There are four NRT2 transporters that nitrate uptake depends on in HATS, namely NRT2.1, NRT2.2, NRT2.4, and NRT2.5 of which three were found [53].

Category	Genes
NRT1 family	NRT1.4, NRT1.5
NRT2 family	NRT2.2, NRT2.4, NRT2.5
AUX biosynthesis	TAR2, CYP79B2, CYP79B3
Other nitrate related	HHO1, NLP3, NIA2
CEP	ROXY20, ROXY21
CK	IPT5, LOG5, CKX4
Sugar transport	SWEET16, SWEET17, SUC1

Table 5.1: Genes found during the data analysis.

Future research would involve improving the FSPM as it is still a simplified version of a root system architecture. The surface of the root should be taken into account whereas a larger surface means more uptake. The ageing of the root should be considered in view of the roots becoming woodier as they grow, thus they take up less water and consequently nitrate. A drier environment disturbs transport as well, because of the lack of water the nitrate remains in the soil. The lengths are proportionally accurate, but some fine-tuning is required with the parameters. We still need to identify what signals cause less laterals to emerge in low nitrate conditions which would make the model more realistic. Adding carbon allocation and transportation would also improve the model as carbon is one the nutrients determining root growth. The FSPM could also be extended with varying nitrate concentrations as plants often experience temporal changes which affects foraging. Roots develop quicker in environments where nitrate concentrations are increasing, even when their other roots are growing in richer patches [64]. Transport is slower in longer roots which could be reflected by including delays in the transport.

We aimed to combine the modelling and data analysis where significant genes could have led to adding more signalling pathways. However the two parts remained separate. It was challenging to identify differentially expressed genes in the microarray data because the differences in expression levels were small between experiments. With other data sets it could be possible to pick up more genes that play an important role in preferential root foraging. Additionally, other data sets could involve shoot data to analyse root-shoot transport. Microarray data could be replaced with RNA-Seq in view of the costs of the technology dropping [65]. RNA-Seq could show new insights considering that they can produce a higher resolution and sensitivity of differentially expressed genes than microarrays [66].

Acknowledgements

I would like to thank Bea and Christian Bachem for finding me this topic and for always welcoming me in their home.

Thank you to my supervisors Bas van den Herik, Kirsten ten Tusscher, and Tom Claassen for their useful inputs and having the patience to bear through all the difficulties that I had this year. Thank you Bas for the long hours of discussions, and for always being so encouraging, your reassurance meant so much whenever I was stressed. Thank you Kirsten for asking the tough questions and staying critical. And thank you Tom for joining a project that neither of us had almost any clue about and for keeping an eye on the computer science side of it.

Lastly, thanks to Julian Besems for helping with the images and for being a harsh proof reader, even when I did not want to hear it.

Bibliography

[1] M. D. Boer, J. Santos Teixeira, and K. H. Ten Tusscher, "Modeling of root nitrate responses suggests preferential foraging arises from the integration of demand, supply and local presence signals," *Frontiers in Plant Science*, vol. 11, p. 708, 2020.

[2] S. Ruffel, G. Krouk, D. Ristova, D. Shasha, K. D. Birnbaum, and G. M. Coruzzi, "Nitrogen economics of root foraging: Transitive closure of the nitrate–cytokinin relay and distinct systemic signaling for n supply vs. demand," *Proceedings of the National Academy of Sciences*, vol. 108, no. 45, pp. 18524–18529, 2011.

[3] D. K. Ray, N. D. Mueller, P. C. West, and J. A. Foley, "Yield trends are insufficient to double global crop production by 2050," *PloS one*, vol. 8, no. 6, p. e66428, 2013.

[4] M. A. Altieri and C. I. Nicholls, "The adaptation and mitigation potential of traditional agriculture in a changing climate," *Climatic Change*, vol. 140, no. 1, pp. 33–45, 2017.

[5] A. Lithourgidis, C. Dordas, C. A. Damalas, and D. Vlachostergios, "Annual intercrops: an alternative pathway for sustainable agriculture," *Australian journal of crop science*, vol. 5, no. 4, pp. 396–410, 2011.

[6] M. F. Nawaz, G. Bourrie, and F. Trolard, "Soil compaction impact and modelling. a review," *Agronomy for sustainable development*, vol. 33, no. 2, pp. 291–309, 2013.

[7] A. Iqbal, D. Qiang, M. Alamzeb, W. Xiangru, G. Huiping, Z. Hengheng, P. Nianchang, Z. Xiling, and S. Meizhen, "Untangling the molecular mechanisms and functions of nitrate to improve nitrogen use efficiency," *Journal of the Science of Food and Agriculture*, vol. 100, no. 3, pp. 904–914, 2020.

[8] H. Sauquet, M. Von Balthazar, S. Magallón, J. A. Doyle, P. K. Endress, E. J. Bailes, E. Barroso de Morais, K. Bull-Hereñu, L. Carrive, M. Chartier, *et al.*, "The ancestral flower of angiosperms and its early diversification," *Nature communications*, vol. 8, no. 1, pp. 1–10, 2017.

[9] M. E. Ritchie, B. Phipson, D. Wu, Y. Hu, C. W. Law, W. Shi, and G. K. Smyth, "limma powers differential expression analyses for RNA-sequencing and microarray studies," *Nucleic Acids Research*, vol. 43, pp. e47–e47, 01 2015.

[10] G. K. Smyth, "Linear Models and Empirical Bayes Methods for Assessing Differential Expression in Microarray Experiments," *Statistical Applications in Genetics and Molecular Biology*, vol. 3, no. 1, 2004.

[11] K. Esau *et al.*, "Plant anatomy.," *Plant Anatomy.*, no. 2nd Edition, 1965.

[12] W. T. McCleery, N. A. Mohd-Radzman, and V. A. Grieneisen, "Root branching plasticity: collective decision-making results from local and global signalling," *Current Opinion in Cell Biology*, vol. 44, pp. 51–58, 2017. Cell Architecture.

[13] F. A. Reynolds, "Introduction to Vascular Plant Structure." `https://blogs.berkshirecc.edu/bccoer/`, 02 2014.

[14] C. Rye, R. Wise, V. Jurukovski, J. DeSaix, J. Choi, and Y. Avissar, "Biology." `https://openstax.org/books/biology/pages/30-5-transport-of-water-and-solutes-in-plants`, 10 2016.

[15] Y.-M. Bi, R.-L. Wang, T. Zhu, and S. J. Rothstein, "Global transcription profiling reveals differential responses to chronic nitrogen stress and putative nitrogen regulatory components in arabidopsis," *BMC genomics*, vol. 8, no. 1, pp. 1–17, 2007.

[16] R. F. H. Giehl, B. D. Gruber, and N. von Wirén, "It's time to make changes: modulation of root system architecture by nutrient signals," *Journal of Experimental Botany*, vol. 65, pp. 769–778, 12 2013.

[17] R. Tabata, K. Sumida, T. Yoshii, K. Ohyama, H. Shinohara, and Y. Matsubayashi, "Perception of root-derived peptides by shoot lrr-rks mediates systemic n-demand signaling," *Science*, vol. 346, no. 6207, pp. 343–346, 2014.

[18] Y. Ohkubo, M. Tanaka, R. Tabata, M. Ogawa-Ohnishi, and Y. Matsubayashi, "Shoot-to-root mobile polypeptides involved in systemic regulation of nitrogen acquisition," *Nature plants*, vol. 3, no. 4, pp. 1–6, 2017.

[19] E. A. Vidal, V. Araus, C. Lu, G. Parry, P. J. Green, G. M. Coruzzi, and R. A. Gutiérrez, "Nitrate-responsive mir393/afb3 regulatory module controls root system architecture in arabidopsis thaliana," *Proceedings of the National Academy of Sciences*, vol. 107, no. 9, pp. 4477–4482, 2010.

[20] T. Girin, E.-S. El-Kafafi, T. Widiez, A. Erban, H.-M. Hubberten, J. Kopka, R. Hoefgen, A. Gojon, and M. Lepetit, "Identification of arabidopsis mutants impaired in the systemic regulation of root nitrate uptake by the nitrogen status of the plant," *Plant Physiology*, vol. 153, no. 3, pp. 1250–1260, 2010.

[21] W. Ma, J. Li, B. Qu, X. He, X. Zhao, B. Li, X. Fu, and Y. Tong, "Auxin biosynthetic gene TAR2 is involved in low nitrogen-mediated reprogramming of root architecture in Arabidopsis," *The Plant Journal*, vol. 78, no. 1, pp. 70–79, 2014.

[22] E. Ramireddy, L. Chang, and T. Schmülling, "Cytokinin as a mediator for regulating root system architecture in response to environmental cues," *Plant signaling & behavior*, vol. 9, no. 1, pp. 5021–32, 2014.

[23] T. Araya, M. Miyamoto, J. Wibowo, A. Suzuki, S. Kojima, Y. N. Tsuchiya, S. Sawa, H. Fukuda, N. Von Wirén, and H. Takahashi, "CLE-CLAVATA1 peptide-receptor signaling module regulates the expansion of plant root systems in a nitrogen-dependent manner," *Proceedings of the National Academy of Sciences*, vol. 111, no. 5, pp. 2029–2034, 2014.

[24] R. Sievänen, C. Godin, T. M. DeJong, and E. Nikinmaa, "Functional–structural plant models: a growing paradigm for plant studies," *Annals of Botany*, vol. 114, pp. 599–603, 09 2014.

[25] A. Lindenmayer, "Mathematical models for cellular interactions in development i. filaments with one-sided inputs," *Journal of Theoretical Biology*, vol. 18, no. 3, pp. 280–299, 1968.

[26] A. Lindenmayer, "Mathematical models for cellular interactions in development ii. simple and branching filaments with two-sided inputs," *Journal of Theoretical Biology*, vol. 18, no. 3, pp. 300–315, 1968.

[27] O. Kniemeyer, *Design and implementation of a graph grammar based language for functional-structural plant modelling*. PhD thesis, BTU Cottbus-Senftenberg, 2008.

[28] E. MOUNIER, M. PERVENT, K. LJUNG, A. GOJON, and P. NACRY, "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*, vol. 37, no. 1, pp. 162–174, 2014.

[29] P. Guan, R. Wang, P. Nacry, G. Breton, S. A. Kay, J. L. Pruneda-Paz, A. Davani, and N. M. Crawford, "Nitrate foraging by arabidopsis roots is mediated by the transcription factor tcp20 through the systemic signaling pathway," *Proceedings of the National Academy of Sciences*, vol. 111, no. 42, pp. 15267–15272, 2014.

[30] J. L. Mullen, H. Ishikawa, and M. L. Evans, "Analysis of changes in relative elemental growth rate patterns in the elongation zone of arabidopsis roots upon gravistimulation," *Planta*, vol. 206, no. 4, pp. 598–603, 1998.

[31] M. Schena, D. Shalon, R. W. Davis, and P. O. Brown, "Quantitative monitoring of gene expression patterns with a complementary dna microarray," *Science*, vol. 270, no. 5235, pp. 467–470, 1995.

[32] R. A. Irizarry, B. Hobbs, F. Collin, Y. D. Beazer-Barclay, K. J. Antonellis, U. Scherf, and T. P. Speed, "Exploration, normalization, and summaries of high density oligonucleotide array probe level data," *Biostatistics*, vol. 4, no. 2, pp. 249–264, 2003.

[33] J. C. Mills, K. A. Roth, R. L. Cagan, and J. I. Gordon, "Dna microarrays and beyond: completing the journey from tissue to cell," *Nature Cell Biology*, vol. 3, no. 8, pp. E175–E178, 2001.

[34] L. Vysotskaya, L. Timergalina, G. Akhiyarova, A. Korobova, V. Fedyaev, I. Ivanov, G. Kudoyarova, and D. Veselov, "Association of barley root elongation with aba-dependent transport of cytokinins from roots and shoots under supra-optimal concentrations of nitrates and phosphates," *Cells*, vol. 10, no. 11, p. 3110, 2021.

[35] T. Kamada-Nobusada, N. Makita, M. Kojima, and H. Sakakibara, "Nitrogen-dependent regulation of de novo cytokinin biosynthesis in rice: the role of glutamine metabolism as an additional signal," *Plant and Cell Physiology*, vol. 54, no. 11, pp. 1881–1893, 2013.

[36] Z. Jia and N. von Wirén, "Signaling pathways underlying nitrogen-dependent changes in root system architecture: from model to crop species," *Journal of Experimental Botany*, vol. 71, no. 15, pp. 4393–4404, 2020.

[37] R. Tabata, T. Kamiya, S. Imoto, H. Tamura, K. Ikuta, M. Tabata, T. Hirayama, H. Tsukagoshi, K. Tanoi, T. Suzuki, *et al.*, "Systemic regulation of iron acquisition by arabidopsis in environments with heterogeneous iron distributions," *Plant and Cell Physiology*, 2022.

[38] I. B. Jeffery, D. G. Higgins, and A. C. Culhane, "Comparison and evaluation of methods for generating differentially expressed gene lists from microarray data," *BMC bioinformatics*, vol. 7, no. 1, pp. 1–16, 2006.

[39] M. Jeanmougin, A. De Reynies, L. Marisa, C. Paccard, G. Nuel, and M. Guedj, "Should we abandon the t-test in the analysis of gene expression microarray data: a comparison of variance modeling strategies," *PloS one*, vol. 5, no. 9, p. e12336, 2010.

[40] C. Kooperberg, A. Aragaki, A. D. Strand, and J. M. Olson, "Significance testing for small microarray experiments," *Statistics in medicine*, vol. 24, no. 15, pp. 2281–2298, 2005.

[41] B. S. Carvalho and R. A. Irizarry, "A framework for oligonucleotide microarray preprocessing," *Bioinformatics*, vol. 26, no. 19, pp. 2363–2367, 2010.

[42] W. Gregory Alvord, J. A. Roayaei, O. A. Quiñones, and K. T. Schneider, "A microarray analysis for differential gene expression in the soybean genome using Bioconductor and R," *Briefings in Bioinformatics*, vol. 8, pp. 415–431, 09 2007.

[43] R. Gentleman, V. J. Carey, W. Huber, R. A. Irizarry, S. Dudoit, *et al.*, *Bioinformatics and computational biology solutions using R and Bioconductor*, vol. 1. Springer, 2005.

[44] J. M. Bland and D. Altman, "Statistical methods for assessing agreement between two methods of clinical measurement," *The lancet*, vol. 327, no. 8476, pp. 307–310, 1986.

[45] W. S. Cleveland, *Visualizing data*. Hobart press, 1993.

[46] S. Dudoit, Y. H. Yang, M. J. Callow, and T. P. Speed, "Statistical methods for identifying differentially expressed genes in replicated cdna microarray experiments," *Statistica sinica*, pp. 111–139, 2002.

[47] B. M. Bolstad, R. A. Irizarry, M. Åstrand, and T. P. Speed, "A comparison of normalization methods for high density oligonucleotide array data based on variance and bias," *Bioinformatics*, vol. 19, no. 2, pp. 185–193, 2003.

[48] R. A. Irizarry, B. M. Bolstad, F. Collin, L. M. Cope, B. Hobbs, and T. P. Speed, "Summaries of affymetrix genechip probe level data," *Nucleic acids research*, vol. 31, no. 4, pp. e15–e15, 2003.

[49] F. D. Gibbons and F. P. Roth, "Judging the quality of gene expression-based clustering methods using gene annotation," *Genome research*, vol. 12, no. 10, pp. 1574–1581, 2002.

[50] J.-Y. Jung, J. H. Ahn, and D. P. Schachtman, "CC-type glutaredoxins mediate plant response and signaling under nitrate starvation in Arabidopsis," *BMC plant biology*, vol. 18, no. 1, pp. 1–13, 2018.

[51] S.-H. Lin, H.-F. Kuo, G. Canivenc, C.-S. Lin, M. Lepetit, P.-K. Hsu, P. Tillard, H.-L. Lin, Y.-Y. Wang, C.-B. Tsai, A. Gojon, and Y.-F. Tsay, "Mutation of the Arabidopsis NRT1.5 Nitrate Transporter Causes Defective Root-to-Shoot Nitrate Transport," *The Plant Cell*, vol. 20, pp. 2514–2528, 09 2008.

[52] Y.-F. Tsay, C.-C. Chiu, C.-B. Tsai, C.-H. Ho, and P.-K. Hsu, "Nitrate transporters and peptide transporters," *FEBS letters*, vol. 581, no. 12, pp. 2290–2300, 2007.

[53] S. Ruffel, V. Chaput, J. Przybyla-Toscano, I. Fayos, C. Ibarra, T. Moyano, C. Fizames, P. Tillard, J. A. O'Brien, R. A. Gutiérrez, A. Gojon, and L. Lejay, "Genome-wide analysis in response to nitrogen and carbon identifies regulators for root AtNRT2 transporters," *Plant Physiology*, vol. 186, pp. 696–714, 02 2021.

[54] Y. Zhao, A. K. Hull, N. R. Gupta, K. A. Goss, J. Alonso, J. R. Ecker, J. Normanly, J. Chory, and J. L. Celenza, "Trp-dependent auxin biosynthesis in arabidopsis: involvement of cytochrome p450s cyp79b2 and cyp79b3," *Genes & development*, vol. 16, no. 23, pp. 3100–3112, 2002.

[55] A. Safi, A. Medici, W. Szponarski, A. Marshall-Colon, S. Ruffel, F. Gaymard, G. Coruzzi, B. Lacombe, and G. Krouk, "HRS1/HHOs GARP transcription factors and reactive oxygen species are regulators of Arabidopsis nitrogen starvation response. Affiliation," 2018.

[56] M. Konishi and S. Yanagisawa, "Arabidopsis NIN-like transcription factors have a central role in nitrate signalling," *Nature communications*, vol. 4, no. 1, pp. 1–9, 2013.

[57] L. Chang, E. Ramireddy, and T. Schmülling, "Cytokinin as a positional cue regulating lateral root spacing in Arabidopsis," *Journal of Experimental Botany*, vol. 66, pp. 4759–4768, 05 2015.

[58] C. J. Thieme, M. Rojas-Triana, E. Stecyk, C. Schudoma, W. Zhang, L. Yang, M. Miñambres, D. Walther, W. X. Schulze, J. Paz-Ares, *et al.*, "Endogenous arabidopsis messenger rnas transported to distant tissues," *Nature Plants*, vol. 1, no. 4, pp. 1–9, 2015.

[59] D. Chandran, "Co-option of developmentally regulated plant sweet transporters for pathogen nutrition and abiotic stress tolerance," *IUBMB life*, vol. 67, no. 7, pp. 461–471, 2015.

[60] A. B. Sivitz, A. Reinders, and J. M. Ward, "Arabidopsis sucrose transporter AtSUC1 is important for pollen germination and sucrose-induced anthocyanin accumulation," *Plant Physiology*, vol. 147, no. 1, pp. 92–100, 2008.

[61] M. Durand, D. Mainson, B. Porcheron, L. Maurousset, R. Lemoine, and N. Pourtau, "Carbon source–sink relationship in Arabidopsis thaliana: the role of sucrose transporters," *Planta*, vol. 247, no. 3, pp. 587–611, 2018.

[62] J. Canales, T. C. Moyano, E. Villarroel, and R. A. Gutiérrez, "Systems analysis of transcriptome data provides new hypotheses about arabidopsis root response to nitrate treatments," *Frontiers in plant science*, vol. 5, p. 22, 2014.

[63] A. Maghiaoui, E. Bouguyon, C. Cuesta, F. Perrine-Walker, C. Alcon, G. Krouk, E. Benková, P. Nacry, A. Gojon, and L. Bach, "The Arabidopsis NRT1.1 transceptor coordinately controls auxin biosynthesis and transport to regulate root branching in response to nitrate," *Journal of Experimental Botany*, vol. 71, pp. 4480–4494, 05 2020.

[64] H. Shemesh, A. Arbiv, M. Gersani, O. Ovadia, and A. Novoplansky, "The effects of nutrient dynamics on root patch choice," *PLOS ONE*, vol. 5, pp. 1–6, 05 2010.

[65] K. J. Mantione, R. M. Kream, H. Kuzelova, R. Ptacek, J. Raboch, J. M. Samuel, and G. B. Stefano, "Comparing bioinformatic gene expression profiling methods: microarray and rna-seq," *Medical science monitor basic research*, vol. 20, p. 138, 2014.

[66] Z. Wang, M. Gerstein, and M. Snyder, "Rna-seq: a revolutionary tool for transcriptomics," *Nature reviews genetics*, vol. 10, no. 1, pp. 57–63, 2009.

Appendix A

The XL Programming Language

The XL programming language is an implementation of relational growth grammars (RGG). It is an extension of Java, therefore all relevant libraries of Java can be imported. However the more interesting topic is what an RGG is which will be the topic of the next section, followed by the design of XL with some examples.

A.1 Theoretical background of XL

A.1.1 L-systems

Lindenmayer-systems are a type of formal grammar widely used in structural plant modelling. They operate on an alphabet from which a string of symbols can be constructed, and using production rules the symbols can be expanded to larger symbols. The productions are executed in parallel, as in living organisms growth processes happen in parallel.

The simplest kind of L-system is a context-free grammar, given in definition A.1.

Definition A.1 (context-free L-system)**.** *A context-free L-system (0L-system)* $\mathcal{G} = (V, \alpha, P)$ *consists of an alphabet* V*, an initial empty word* $\alpha \in V^+$*, and a set of productions* $P \subset V \times V^*$*. A rule* $(a, \chi) \in P$ *is written as* $a \to \chi$*, where* a *is the predecessor and* χ *is the successor.*

From the grammar, a context-free language can be generated as in definition A.2.

Definition A.2 (generated language)**.** *Let* $\mathcal{G}$ *be a context-free L-system and* $\mu = a_1 \dots a_n$ *be a word of* n *symbols. If* $\exists \chi_1, \dots, \chi_n \in V^*$ *such that* $a_i \to \chi_i \in P$*, then there is a direct derivation* $\mu \overset{\mathcal{G}}{\Rightarrow} \nu$ *from* μ *to* $\nu = \chi_1 \dots \chi_n$ *in* $\mathcal{G}$*. The language generated by* $\mathcal{G}$ *is* $L(\mathcal{G}) = \left\{ \nu \in V^* \middle| \alpha \overset{\mathcal{G}}{\Rightarrow}{}^* \nu \right\}$.

The words generated like this are abstract, therefore it is necessary to add an interpretation. We aim to create plant structures with L-systems, so a turtle interpretation is used to create 3-dimensional structures.

$$\alpha = \texttt{F Right Right F Right Right F},$$
$$\texttt{F} \to \texttt{F Left F Right Right F Left F}$$

If a non-deterministic L-system is constructed - which is almost always the case for plant growth -, then stochasticity is ensured by assigning probabilities to the rules:

Definition A.3 (stochastic 0L-system)**.** *A stochastic 0L-system* $(\mathcal{G}, \pi)$ *is a context-free L-system* $\mathcal{G}$ *with a probability function* $\pi : P \to (0, 1]$ *such that* $\forall a \in V : \sum_{\chi \in V^*} \pi(a \to \chi) = 1$.

Definition A.4 (stochastic derivation)**.** *Let* $(\mathcal{G}, \pi)$ *be a stochastic 0L-system. A stochastic derivation* $\mu \overset{\mathcal{G}}{\Rightarrow} \nu$ *is a derivation such that* $\forall a \in \mu$ *the probability of applying* $a \to \chi \in P$ *is* $\pi(a \to \chi)$.

A.2 The design of XL

A.2.1 Generator Expressions

Generator methods return numerous values one after the other and were defined similarly to Python.

1. The range operator `a:b` returns the values $a, a + 1, \ldots, b$.
2. The array generator `a[:]` returns all the values of array a.
3. The guard operator `a::b` returns a if b is true, and no value if b is false.

A.2.2 Aggregate Expressions

Aggregate methods are the counterpart to generator expressions, they take multiple values and give one result back. These were also defined similarly to Python's syntax.

- The containment operator `a in b` returns true if a equals one of the values in b, otherwise returns false.
- Standard aggregate methods, such as `count`, `max`, `min`, `mean` are available in the `de.grogra.xl.util.Operators` class.

A.2.3 Queries

Queries are the reason that XL stands out from other programming languages. They are also the most complex part of the language. In a query a pattern is given that is evaluated against the graph. For a pattern P the query is given by

```
(* P *).
```

A pattern can be simple such as

```
(* f:F & (f.diameter > 0.01) *)
```

where all nodes f of type F that have bigger diameters than 0.01 are returned. We can make use of the L-system syntax as in

```
(* x (<-successor-)* Basis *)
```

where all `Basis` nodes are returned, that can be reached from `x` through successor edges backwards. Here, `successor` is a predefined constant which stands for successor edges, the arrows indicate the direction, and `(...)*` is a regex which says that an arbitrary number of edges can be used.

Patterns

Queries are built from elementary patterns, the simplest being node patterns which are separated by white spaces. If `T` is a type, then the term `T` can already be used as a query which matches instances of type `T`. If a reference is necessary, we can declare a variable as `var:T`. A special case is `var:.` where `var` matches all nodes.

L-systems allow parameterisation, hence XL allows this too, and it is defined similarly to the mathematical notation of a function and parameter, for example `F(a)`. It is possible to not give any values with writing `F(.)`, or to omit parameters: `G(a,,b)`.

Nodes are connected through paths, therefore different kinds of path patterns were created too. Explicit path patterns are defined as

```
-r-> <-r- -r- <-r->
```

where r is the relation, which can be a type, a pattern, or an expression.

Regular expressions are crucial to express patterns in a short way, the quantifiers are

```
+ * ? {n} {n, m} {n,}
```

which are added to the parenthesis of the pattern, as in the example earlier

```
(<-successor-)*
```

A.2.4 Production statements

Production statements form the right-hand side of rules which are made up of expressions. The simplest form is a node expression, and they share multiple similarities with node patterns. The nodes here too are separated by white spaces, and they may be prefixed by an identifier and a colon. They can be prefixed by operators

```
> < <-> --- >> << >>> in ::
+> <+ <+> -+- >= <= <=> ++ --
/> </ </> -/- + * / % **
--> <-- <--> | || & &&
```

which are further explained on table A.1, and by edge expressions `e`

```
-e-> n  <-e- n  -e- n  <-e-> n
```

Code blocks can be added as well using { indicating the beginning and } the end, such as

```
{float f = x/3;}
```

A.2.5 Rules

Rules consist of queries and production statements presented in the previous two subsections. The two parts are connected by the arrows `==>>` and `==>`. For example the rule of the snowflake is given by

```
[
    Axiom ==> F(1) RU(120) F(1) RU(120) F(1);
    F(x) ==> F(x/3) RU(-60) F(x/3) RU(120) F(x/3) RU(-60) F(x/3);
]
```

These are so-called structural rules, and a third arrow is used `::>` for execution rules. Execution rules differ from structural rules in production statements, the right-hand side of these are single imperative statements. These do not change the structure of the graph, they are good for printing values, or changing values of nodes for example.

Rules are specified in rule blocks, they are defined between square brackets `[]`. Expressions in node blocks are executed sequentially.

A.2.6 Properties

Properties are similar to instance variables of Java, though they have a different syntax

```
e[n]
```

where `e` is an expression of type `T` and `n` is the name of the property.

Assignments

Values can be assigned to variables with the usual equation symbol `=`, and compound assignments like `+=` are part of the XL language too. In the case of properties, the assignment operators are

```
:= :**= :*= :/= :%= :+= :-=
:<<= :>>= :>>>= :&= :^= :|=
```

These are deferred assignments, that do not execute immediately, only after the parallel production has.

A.2.7 Other extensions

for statements

The usual for loops of the Java, including

```
for (T i : e) b
```

where `i` is the iterator of type `T`, `e` the expression, and `b` the body. XL has an even shorter notation:

```
for (e) b
```

Expression lists

Expression lists are taken from the programming language C which defines the comma operator `a, b`. This expression evaluates `a` and passes it to `b`, discards the result of `a`, and then evaluates `b` and returns the result. The idea is similar here, for example calculating the squares from 0 to 100 is given by

```
(int x = 0 :  100, x * x)
```

Comparison operator

Next to the traditional comparison operators like `==`, `<`, and `<=`, a binary one is part of the XL language defined as

```
a <=> b
```

which returns 0 if a and b are equal, 1 if $a > b$, and -1 otherwise.

The summary of all operators is presented on table A.1.

Operator	Description
`‘a‘`	quote
`a[b]`	array index
`a(b)`	invocation
`a[:]`	array generator
`a->b`	right arrow
`a<-b`	left arrow
`a++`	postfix increment
`a--`	postfix decrement
`a**b`	exponentiation
`++a`	prefix increment
`--a`	prefix decrement
`+a`	unary plus
`-a`	negation
`~a`	bitwise complement
`!a`	logical complement
`a*b`	multiplication
`a/b`	division
`a%b`	remainder
`a+b`	addition
`a-b`	subtraction
`a<<b`	left shift
`a>>b`	right shift
`a>>>b`	unsigned right shift
`a instanceof T`	type comparison
`a<b`	less than
`a>b`	greater than
`a<=b`	less than or equal
`a>=b`	greater than or equal
`a<=>b`	comparison
`a in b`	containment
`a<->b`	left-right arrow
`a-->b`	long right arrow
`a<--b`	long left arrow
`a<-->b`	long left-right arrow
`a---b`	line
`a+>b`	plus right arrow
`a<+b`	plus left arrow
`a<+>b`	plus left-right arrow

`a-+-b`	plus line
`a/>b`	slash right arrow
`a</b`	slash left arrow
`a <\/> b`	slash left-right arrow
`a -/- b`	slash line
`a==b`	equality
`a!=b`	inequality
`a&b`	bitwise and
`a^b`	bitwise exclusive or
`a|b`	bitwise inclusive or
`a&&b`	conditional and
`a||b`	conditional or
`a::b`	guard
`a?b:c`	conditional
`a:b`	range
`a = b`	assignment
`a := b`	deferred assignment
`a` *op*`= b`	compound assignment
`a :`*op*`= b`	compound deferred assignment
	op $\in$ {`**`, `*`, `/`, `%`, `+`, `-`, `<<`, `>>`, `>>>`, `&`, `^`, `|`}

Table A.1: Operators decreasing by precedence.

```
double growthFunction(){
  if(side == 1) {
    Nex = Nleft;
  } else if(side == 2){
    Nex = Nright;
  }

  double size = sum((*r:RootOrgan, r.order == 1)*)[length]);
  double N = sum((*ShootOrg*)[Ns]) / size;
  double CEP = sum((*ShootOrg*)[CEPs]) / size;
  double CK = sum((*ShootOrg*)[CKs]) / size;

  double gNE = Nex**2 / (Nex**2 + KCEPNE**2);
  double gCK = aCK + (1 - aCK) * (CK**2 / (CK**2 + KCEPCK**2));

  double fbasic = abasic + (1 - abasic) * (N**2 / (N**2 + Kbasic**2));
  double flocal = (1 - alocal) + alocal * (Nex**2 / (Nex**2 + Klocal**2));

  double fsystrep = Ksystrep**4 / (Ksystrep**4 + N**4);
  double fsystfor = 1 + asystfor * (Ksystfor**4 / (N**4 + Ksystfor**4));

  double fCEP = 1 + aCEP * (CEP**2 / (CEP**2 + KNRT21**2)) * gNE;
  double fCEPCK = 1 + aCEP * (CEP**2 / (CEP**2 + KNRT21**2)) * gNE * gCK;

  return fbasic * flocal * fCEP * fsystrep * fsystfor;
}
```

Figure A.1: The equations from the Boer model that calculate all the signalling occurring in the roots [1].

```
void calcGrowth() {
  double size = sum((*r:RootOrgan*)[length]);
  double growth = growthFunction() * conv * size / count((*ez:EZSegment*));
  length += growth;
}
```

Figure A.2: The implementation of equation 3.1.