

Patterns of symbiotic calcium oscillations

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

Anna Emma Ulrika Granqvist

John Innes Centre
Norwich, September 2012

© This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with the author and that use of any information derived there from must be in accordance with current UK Copyright Law. In addition, any quotation or extract must include full attribution.

To my family, with love

Abstract

Nuclear-localised calcium (Ca^{2+}) oscillations have been shown to be necessary for the establishment of both arbuscular mycorrhizal and rhizobial symbioses. This thesis focuses on the analysis of the nature of this Ca^{2+} signature and an investigation into how this signature might have changed during the evolution of arbuscular mycorrhizal and root nodule symbioses. The work combines mathematical and experimental biology.

A method for detecting frequencies using Bayesian inference, Bayesian Spectrum Analysis, has been tested and further developed for use in the analysis of Ca^{2+} oscillations. A mathematical model of the oscillatory system has been investigated with bifurcation analysis, focusing on how Ca^{2+} -binding proteins might account for the experimentally observed variation in the signal.

Data were collected for Ca^{2+} spiking induced by the diffusible signal molecules from the arbuscular mycorrhizal and rhizobial symbioses. No significant differences were found between the patterns of Ca^{2+} oscillations. This study was extended to a group of phylogenetically diverse plants in order to investigate whether there is evidence for changes in the patterns of Ca^{2+} oscillations during the evolution of the signalling pathway from the more ancient arbuscular symbiosis to the more recent root nodule symbiosis. Although some variation in Ca^{2+} patterns was found, this did not show a consistent correlation with phylogenetic relationships.

Taken together, the present study suggests possible sources of the observed variability of Ca^{2+} patterns. It also indicates that the system of Ca^{2+} oscillations is a highly conserved part of the symbiotic signalling pathway.

Contents

Abstract	iii
List of Figures	vii
List of Tables	ix
List of Abbreviations	x
Publications	xii
1 General introduction	1
1.1 Motivation	1
1.1.1 Nitrogen	1
1.1.2 Phosphorous	2
1.2 Plant-microbe root symbioses	3
1.2.1 Arbuscular mycorrhiza	3
1.2.2 Root nodule symbioses	4
1.3 Initial symbiotic events	6
1.3.1 Signal exchange in the soil	6
1.3.2 The common SYM pathway	8
1.4 The role of calcium signalling	12
1.4.1 Calcium oscillations in animal cells	13
1.4.2 Calcium oscillations in plant stomatal guard cells	13
1.4.3 Calcium oscillations in the SYM pathway	14
1.4.4 Calcium oscillations in defence pathways	16
1.5 Establishment of endosymbiosis	16
1.5.1 Mycorrhizal infection	17
1.5.2 Rhizobial infection	18
1.6 Hormone influences	21
1.7 Cross-talk with defence responses	24
1.8 Project objectives and thesis structure	25
2 Methods and materials	26
2.1 Experimental methods	26
2.1.1 Plant growth	26

2.1.2	Crude isolation of Nod factors	28
2.1.3	Root hair deformation assay	28
2.1.4	Calcium imaging	29
2.2	Experimental material	32
2.3	Computational methods	34
2.3.1	Fourier Analysis	34
2.3.2	Bayesian Spectrum Analysis	34
2.3.3	Modelling calcium oscillations	38
3	Employing Bayesian Spectrum Analysis to detect periodicity	39
3.1	Introduction	39
3.1.1	Fourier Analysis	40
3.1.2	Bayesian Spectrum Analysis	41
3.1.3	Model comparison	45
3.2	Results	49
3.2.1	Simulated test cases	49
3.2.2	Calcium spiking	56
3.2.3	Clock genes	61
3.3	Discussion	62
4	Bifurcation analysis of a calcium spiking model reveals sources of pattern variation	64
4.1	Introduction	64
4.1.1	Components needed for the symbiotic calcium oscillations	66
4.2	Description of the calcium model	68
4.2.1	Model components and equations	69
4.2.2	Model parameters	73
4.3	Results	75
4.3.1	The influence of calcium buffers on spiking patterns	76
4.3.2	Other sources of variation	81
4.4	Discussion	82
5	Comparing patterns of calcium oscillations in <i>Medicago truncatula</i>	84
5.1	Introduction	84
5.1.1	Rhizobial LCOs	84
5.1.2	Mycorrhizal LCOs	85
5.1.3	LCOs, COs and calcium spiking	85
5.1.4	The model legume <i>Medicago truncatula</i>	86
5.2	Results	86
5.2.1	Interspike intervals	87
5.2.2	Spike structures	89

5.2.3	Periodicity	91
5.3	Discussion	92
6	Defining the patterns of symbiotic calcium oscillations in diverse species in the nitrogen-fixing clade	96
6.1	Introduction	96
6.1.1	The legume family and symbiosis	97
6.1.2	The <i>Parasponia</i> case	98
6.1.3	Actinorhizal symbioses	98
6.2	Results	99
6.2.1	Choice of plant species to study	99
6.2.2	Induction of spiking with LCOs	101
6.2.3	Calcium oscillations in distantly related legumes	102
6.2.4	Calcium oscillations in the non-legume genera <i>Parasponia</i> and <i>Trema</i>	105
6.2.5	Data analysis of species-specific spiking	107
6.3	Discussion	111
7	General discussion	113
7.1	Patterns of symbiotic calcium oscillations	113
7.2	Similarities and differences in calcium signatures	114
7.3	The role of calcium-binding proteins	115
7.4	Evolution of the common SYM pathway	117
7.5	Prospects for biotechnology involving the common SYM pathway	118
	Appendices	120
.A	Mathematical background	120
.A.1	Interpretations of probability theory	120
.A.2	Dynamical systems theory	120
.A.3	Linear algebra	121
.A.4	Power series and polynomials	123
.B	Supplementary figures and tables	125
	Bibliography	132

List of Figures

1.1	Evolution of root symbioses	5
1.2	The SYM pathway	9
1.3	Ca ²⁺ spiking in the model legume <i>Medicago truncatula</i>	15
1.4	Infection of the plant root by arbuscular mycorrhiza	17
1.5	Different modes of rhizobial infection in symbiosis	19
1.6	The Nod factor signalling pathway in different cell layers	21
2.1	FRET with the Ca ²⁺ indicator cameleon YC2.1	29
2.2	The signal from the Ca ²⁺ reporter Cameleon YC2.1	30
2.3	Flowchart of the automated model development procedure for Bayesian Spectrum Analysis	36
3.1	An example of Fourier theory	41
3.2	Joint probability	42
3.3	Time series with a background trend	50
3.4	Short time series	51
3.5	Effects of noise on precision in FFT and BSA	52
3.6	Higher harmonics	53
3.7	Multiple close frequencies with noise	54
3.8	Change of frequency	55
3.9	Near-periodic oscillations	56
3.10	BSA and Ca ²⁺ spiking	59
3.11	Local BSA and Ca ²⁺ spiking	60
3.12	BSA and circadian genes	62
4.1	Initial signalling events in the SYM pathway	65
4.2	Phase space diagram of the Ca ²⁺ spiking model	70
4.3	The Ca ²⁺ spiking model	71
4.4	Simulated Ca ²⁺ spiking with the model including buffers	73
4.5	Simulations of different spike shapes	74
4.6	Examples of variability in Ca ²⁺ spiking	75
4.7	Unstable and stable steady states	76
4.8	Bifurcation analysis of buffer parameters	77
4.9	Bifurcation analysis of buffer parameters for Oregon Green	77
4.10	Two-parameter bifurcation diagrams of buffer parameters	79

4.11	Two-parameter bifurcation diagrams of buffer parameters	79
4.12	The effect of buffer concentrations on Ca^{2+} spiking period	80
4.13	Increasing buffer levels can inhibit Ca^{2+} oscillations.	80
4.14	Bifurcation analysis of other sensitive model parameters	81
4.15	The influence of pump rate on spike period and shape	82
5.1	Examples Ca^{2+} spiking induced by different signal molecules	88
5.2	Interspike intervals induced by the different signal molecules	89
5.3	Interspike intervals induced by higher LCO concentrations	90
5.4	Interspike intervals versus standard deviations	90
5.5	Spike shapes induced by the different signal molecules	91
5.6	Spike shapes induced by the different imaging methods	92
5.7	Bayesian Spectrum Analysis of the Ca^{2+} spiking	93
6.1	Ca^{2+} spiking in <i>Alnus glutinosa</i>	99
6.2	Phylogenetic tree of the nitrogen-fixing clade	100
6.3	Ca^{2+} spiking in legumes induced by Nod factors	103
6.4	Nuclear localised spiking in the four legumes species	104
6.5	Ca^{2+} oscillations in <i>Parasponia andersonii</i>	105
6.6	<i>Parasponia</i> and <i>Trema</i> responses to LCOs	106
6.7	Interspike intervals of Ca^{2+} oscillations	108
6.8	Comparison of Ca^{2+} spike structures	109
6.9	BSA of Ca^{2+} spiking in different plant species	110
1	The Legendre polynomials	124
2	NGR Nod factors inducing root hair deformation	127
3	Raw Ca^{2+} imaging data, example 1	128
4	Raw Ca^{2+} imaging data, example 2	129

List of Tables

2.1	List of plant material	32
2.2	List of media	33
2.3	Key BaSAR functions	37
3.1	BSA symbol definitions	46
3.2	Model ratio results	48
3.3	Automated model development	51
3.4	Comparison of BSA and FFT with simulated data	57
3.5	BSA on Ca^{2+} data	59
3.6	BSA on circadian data	61
4.1	Initial conditions and parameters used in the Ca^{2+} model	72
4.2	Buffer parameters for the simulations of Oregon Green	74
5.1	Descriptions of the signal molecules used to induce Ca^{2+} spiking	87
6.1	Number of wild-type cells showing Ca^{2+} spiking	102
6.2	Number of NupYC2.1 cells showing Ca^{2+} spiking	105
1	Sources of data in Chapter 5	125
2	Mann-Whitney-Wilcoxon test for interspike intervals	126
3	Kolmogorov-Smirnov test for BSA results	126
4	Kullback-Leibler divergence for BSA results	127
5	Mann-Whitney-Wilcoxon test for interspike intervals	130
6	Mann-Whitney-Wilcoxon test for spike structures	131

List of Abbreviations

ABA	Abscisic acid
AM	Arbuscular mycorrhiza
AON	Autoregulation of nodulation
AVG	L- α -(2-aminoethoxyvinyl) glycine
BSA	Bayesian Spectrum Analysis
Ca ²⁺	Calcium
CaM	Calmodulin
CFP	Cyan fluorescent protein
DFT	Discrete Fourier Transform
ECM	Ectomycorrhiza
ER	Endoplasmic reticulum
FFT	Fast Fourier Transform
FRET	Fluorescence resonance energy transfer
FT	Fourier Transform
GFP	Green fluorescent protein
INM	Inner nuclear membrane
IT	Infection thread
JA	Jasmonic acid
K ⁺	Potassium
LCO	Lipo-chito-oligosaccharide
M	Molar
MYA	Million years ago
Myc factor	Mycorrhizal factor
NE	Nuclear envelope
NFC	Nitrogen fixing clade

Nod factor	Nodulation factor
NPC	Nuclear pore complex
ODE	Ordinary differential equation
OG	Oregon Green
ONM	Outer nuclear membrane
PAMPs	Pathogen-associated molecular patterns
PPA	Prepenetration apparatus
RHD	Root hair deformation
RHI	Root hair infection
RNS	Root nodule symbiosis
ROS	Reactive oxygen species
SA	Salicylic acid
SERCA	Sarcoplasmic/endoplasmic reticulum calcium channel ATPases
TR	Texas Red
YFP	Yellow fluorescent protein

Publications

The manuscripts below include material from this thesis, and have been published or are in preparation. * Indicates that these authors have contributed equally.

Chapter 3

Granqvist, E., Oldroyd, G.E.D. & Morris, R. J. (2011) *Automated Bayesian model development for frequency detection in biological time series*. BMC Systems Biology 5:97.

Granqvist, E., Hartley, M., & Morris, R. J. (2012) *BaSAR - A tool in R for frequency detection*. BioSystems 110(1):60-63.

Chapter 4

Granqvist, E.,* Wysham, D.,* Hazledine, S.,* Kozłowski, W., Sun, J., Charpentier, M., Vaz Martins, T., Haleux, P., Tsaneva-Atanasova, K., Downie, J.A., Oldroyd, G.E.D. & Morris, R. J. (2012) *Buffering capacity explains signal variation in symbiotic calcium oscillations*. Plant Physiology 160(4):2300-2310.

Charpentier, M.,* Vaz Martins, T.,* **Granqvist, E.**, Oldroyd, G.E.D. & Morris, R. J. (2013) *The role of DMI1 in establishing calcium oscillations in legume symbioses*. Plant Signalling and Behaviour 8(2).

Chapter 5

Sun, J.,* Wang, E.,* Miller, J.B.,* **Granqvist, E.**, Gobbato, E., Maillet, F., Fort, S., Morris, R. J., Dénarié, J. & Oldroyd, G.E.D. *Discriminating between Nod factors and Myc factors in legumes and a non-legume*. **In preparation.**

Chapter 6

Granqvist, E., Sun, J., Op den Camp, R., Caopen, W., Normand, P., Geurts, R., Morris, R. J. & Oldroyd, G.E.D. *Symbiotic calcium oscillations have a conserved role in the nitrogen-fixing clade*. **In preparation.**

Acknowledgements

First of all I want to thank both my supervisors, Richard Morris and Giles Oldroyd. Richard has been enthusiastic and supportive throughout, both in my research and training. I am very fortunate to have had him as my supervisor. Giles has given indispensable input into the more experimental parts of my project, offering interesting discussions and highlighting the important questions. I really appreciate that both Richard and Giles have supported my efforts to keep the project interdisciplinary, which has made the project more challenging and interesting.

I also want to thank the people working with me in both of the groups. Their support have made it easier for me to feel at home in two places. Jongho Sun has taught me the microscopy, and has always been around to help. Saul Hazledine was there when I started learning programming, and has a never-ending patience. Derin Wysham has been part of the modelling project, and tried endlessly to make me understand the importance of limit cycles and steady states. Giulia Morieri helped me isolate the NGR Nod factors, and also helped with the microscopes. Myriam Charpentier, Ben Miller, Ward Capoen and Sarah Shailes have just been great to have around in the lab. And I want to thank Anne Edwards, who helps with all those questions that no one else can answer. Nick Pullen has been a valuable proof-reader for my early manuscripts, as well as good company in the “CSB boyz” room.

I also want to thank the other groups and group leaders involved in calcium research at the John Innes Centre, for making it a dynamic and interesting area to work in. These include the groups of Allan Downie, Jeremy Murray, and Stephen Bornemann. Allan has also been a great support, always up for a discussion and questioning what everyone else seems to have forgotten.

Thanks to John Doonan, who was my supervisor during my bachelor project at the JIC, and to Liam Dolan who supervised me during my rotation year. Krasimira Tsaneva-Atanasova taught me bifurcation analysis during a brief visit to Bristol University. Of course a warm thanks to Nick Brewin, for looking after the rotation students and always offering advice, as well as reading this thesis.

I acknowledge the John Innes Foundation that has funded me during my PhD. I am also grateful to several people for giving me plant seeds, bacterial symbionts and advice for my project. These include Kew Gardens Millennium Seed Bank, Milagros León Barrios, Ward Capoen, Katharina Pawlowski, Jean Dénarié, John Tjepkema, Tomasz Stepkowski, Maitrayee Das Gupta, Nick Brewin, Euan James, Janet Sprent

and René Geurts. A special thanks to Rik Op den Camp for a great collaboration and a few trips across the channel with plants. Thanks also to the staff in the JIC media kitchen.

My first and last periods in Norwich were spent in house of the Lamb family, a warm thanks to all of them for making me feel so welcome. I also want to thank the people that just made my PhD life nicer, like David Richards, Katerina Kalogeropoulou and fellow rotation students Evelyn Koerner and Dian Guan. Many thanks to Mike Croll for all the laughs and good times, and to Heather Edwards for helping me keep music in my life and for always brightening my week.

I owe much to my family. I have always felt encouraged by my relatives, which does mean a lot at challenging times. And the warmest thanks to my closest family: my mother Anna-Karin, my father Freddie and my brother Peter. You have always been there for me, and at times seemed much more convinced of my capability than I was. I cannot say how much you mean to me.

Finally, I want to thank Ruben, for all his support during this long and at times difficult period. We have had our fair share of trains, boats, and flights, but I have been lucky enough to finish my thesis in our joint home. Ruben, your presence makes me happier and stronger. The fact that we can now continue on our journey together is the most wonderful thing. Thank you for all your love.

Emma Granqvist

September 2012, Norwich

Chapter 1

General introduction

1.1 Motivation

Agriculture is facing one of the largest challenges for our modern society: How to feed the ever-growing human population. Vital for addressing this challenge is a sound understanding of plants and their requirements for optimal growth. With fertilisers being extensively used at a cost to both the economy and the environment, the acquisition of nutrients has become a research focus. Two major limiting nutrients, and therefore key components in fertilisers, are phosphate and nitrogen.

1.1.1 Nitrogen

Biologically available nitrogen is a limiting factor on primary production in both marine and terrestrial ecosystems, since N_2 cannot be assimilated directly by plants. Therefore, fertilisers are used to boost crop productivity, and a major breakthrough for fertilisers came with the Haber-Bosch process. This method was developed in the early 20th century, and it chemically reduces atmospheric N_2 to NH_3 , a process termed nitrogen fixation. Today it is an essential part of modern agriculture (Galloway and Cowling, 2002). Eutrophication, acidification, and large-scale energy waste are all consequences of the intense use of nitrogen fertilisers (Canfield et al., 2010). Furthermore, since geochemical cycles are closely connected, the pressing issue of global warming is also influenced by the nitrogen-cycle and our enforced changes upon it (Gruber and Galloway, 2008). Another man-made nitrogen addition to the environment, apart from nitrogen fertilisers, comes from fossil fuel combustion. This releases NO_x into the air, affecting chemical and physical pathways in the atmosphere. The combined effects of these nitrogen sources have risen sharply in the last 100 years, and are projected to keep increasing unless drastic changes are undertaken in a number of areas, such as agricultural practice, human diets and regulation of levels of pollution (Howarth et al., 2002). Increasing the capacity of crops to take up their own nitrogen would make a significant contribution towards these needed changes (Galloway et al., 2008).

Plants need nitrogen in an accessible form, and certain types of microbes are ca-

pable of delivering that. These microbes, collectively known as diazotrophs, use the enzyme nitrogenase to fix atmospheric nitrogen and produce ammonia. Some plants enter symbioses with diazotrophs that can provide them with extra nitrogen, for example the association between *Gunnera* and the cyanobacteria *Nostoc*, in which the plants' glandular cells are invaded (Bergman et al., 1992). But far more common is the intimate symbiosis between some plant groups and bacteria that have this nitrogen-fixing trait, and many legumes form symbiotic associations with bacteria collectively known as rhizobia. Small bumps called nodules are formed on the plant roots and these are infected by the bacteria. Within the nodule there is a microaerobic environment, allowing the rhizobia to fix N_2 efficiently by preventing nitrogenase from being inhibited by oxygen. The plant gains the reduced nitrogen, NH_3 , and in exchange the rhizobia are rewarded with a microaerobic safe-house and sugars (Long, 2001; Luciński et al., 2002).

This symbiosis greatly reduces the need for nitrogen fertiliser, and even after the crop is gone, nitrogen remains in the soil from decaying plant parts. The seeds of legumes are also very protein-rich, making them important for human and animal consumption. There is of course a carbohydrate cost to the plant which impacts the crude economic benefit of this symbiosis. However, in places where fertilisers are too expensive or in other ways inaccessible, or by sustainable considerations to the environment, enhancing the capacity of crops to nodulate is clearly beneficial.

1.1.2 Phosphorous

A second nutrient that is often limiting, and therefore also a key fertiliser component, is phosphorous. Natural cycling of phosphorous occurs slowly, but human impact is changing this and causing a series of environmental problems. The mineral is mined from rock, and the resources are found mainly in only a few countries, creating geopolitical influences on the availability (Cordell et al., 2009). Losing phosphorous from agricultural areas into the surrounding ecosystems causes eutrophication, erosion of the soils and depletion of mineral resources (Liu et al., 2008). Phosphorous is one of the many essential but non-renewable elements in the modern food system, and a predicted long-term decrease in available phosphorous to mine will affect the economics of fertiliser production (Vuuren et al., 2010). Thus, improving the ways of gaining phosphorous together with reducing the demand and waste will be essential in the coming decades. Food production plays a central role since 90% of global phosphorous fertilisers are used for food crops (Cordell et al., 2009).

Many groups of plants are naturally aided in their phosphate uptake by symbiosis with mycorrhizal fungi. This is an essential part of much of plant growth, and can, with good reason, be seen as a root extension mechanism, enhancing uptake of not only phosphorous but also water and some other required nutrients. It is estimated that up to 75% of phosphorous in plants is supplied by mycorrhiza (van der Heijden et al., 2008), and in some areas it is an absolute necessity to be connected to this

underground fungal network, for example in rainforest areas where the nutrient content in the soil is especially low. The increased water uptake is also a crucial contributor in some areas. Thus, these symbiotic fungi have an indisputable impact on the plant community, influencing both its structure and plant diversity. In the other direction, it should be noted that worldwide, plant photoassimilates go in huge quantities to mycorrhizal fungi (Bago et al., 2000).

In summary

The un-sustainability of modern agricultural practices requires us to reconsider the way food security might be achieved in the future. Many non-renewable resources will be depleted within the coming century, and environmental problems caused by chemical fertilisers, upsetting natural cycles, will only increase if the appropriate steps are not taken.

Some natural ways of aiding crops with nutrient uptake include exploiting our understanding of symbiotic processes such as nitrogen-fixation and mycorrhization. For nitrogen fixation, several approaches are open such as biofertilization by endophyte bacteria that associate with cereals and fix nitrogen, or to directly engineer the enzyme for nitrogen fixation into plants (Beatty and Good, 2011). A third option is the possibility of transferring the root nodule symbiosis between crop plants. This has received much attention and would be a huge leap much like the green revolution of the 20th century (Charpentier and Oldroyd, 2010). For any such step to be possible, more scientific knowledge of these plant-microbe relationships is required.

1.2 Plant-microbe root symbioses

1.2.1 Arbuscular mycorrhiza

The association with arbuscular mycorrhiza (AM) is formed by over 80% of all land plants with the phylum Glomeromycota (Humphreys et al., 2010). These fungi are thought to be asexual, and can have multiple nuclei in single cells, and even thousands of nuclei in a single spore (Bonfante and Genre, 2010). The fungi invade the plant root and connect it with the fungal hyphae in the soil, creating an extraordinary extension of the root system. Through this contact, the fungus supplies the plant with phosphate, water and other nutrients in exchange for carbon (Harrison, 2005). The ancient nature of this relationship is supported by the finding of fossil arbuscules, which are the fungal structures inside the plant cells, and they date back to the period of land colonisation by plants around 400-450 million years ago (MYA) (Remy et al., 1994; Redecker et al., 2000b,a). It is believed that this plant-fungus relationship was pivotal in land plant colonisation, and today it is significantly contributing to the productivity of plant ecosystems. Where there is already sufficient nutrients for the plants, this symbiosis is less common (Parniske, 2008). Plants that lack the ability to form mycorrhizal

associations are thought to have lost this characteristic, since it is an ancient trait even found in liverworts (Wang and Qiu, 2006).

AM fungi are obligate biotrophs, and thus live only for a short time as saprophytes in the soil and cannot be cultured in the laboratory without a plant host. At the start of symbiosis, the fungus penetrates into the plant root cortex and creates bush-like structures, arbuscules, inside host cells. However, there is no penetration of the plant plasma membrane, but instead a controlled nutrient exchange takes place between fungus and plant plasma membrane. AM is thus a form of endomycorrhiza, where the fungus enters plant cells. There are also mycorrhizal forms called ectomycorrhiza (ECM) in which the fungus never enters into the plant cells (Parniske, 2008). The ECM form has developed from AM independently several times (Wang and Qiu, 2006), and is found mainly in trees in ecological areas with seasonal growth, for example temperate forests (Smith and Douglas, 1987). When ECM fungi has located a suitable plant host the fungal hyphae create a sheath around the actively growing root tip and then colonise the intercellular space in the outer cortical root layers by forming branched hyphae that make up the so-called Hartig Net (Smith and Douglas, 1987).

Mycorrhizal fungus-plant specificity is very low, so fungal species are often capable of infecting a wide range of different plants. Benefits beyond direct nutrient acquisition have been documented, such as heightened biodiversity (van der Heijden et al., 1998) and resistance to pathogens (Hamel and Plenchette, 2007). The large mycorrhizal networks in soils worldwide most likely play important ecological roles as a carbon sink and possibly in connecting plants and transferring nutrients between them, influencing the very shape of plant communities (Smith and Douglas, 1987).

1.2.2 Root nodule symbioses

Root nodule symbiosis (RNS) takes place between some flowering plants and nitrogen-fixing soil bacteria. This symbiosis evolved much later than the mycorrhizal, estimated between 80-100 MYA. Nitrogen-fixation in the specialised structure of a nodule on the plant root is only found in some groups within the Eurosoid I Clade of the Angiosperms, collectively called the nitrogen-fixing clade. Since all nodulators are contained within this monophyletic group, a predisposition event enabling these plants to form nitrogen-fixation symbioses is thought to have occurred only once in their common ancestor (Soltis et al., 1995). What makes the root nodule symbiosis so successful are the optimal conditions within the nodule for fixing N_2 , as well as adaptations in the bacteria that shut down their glutamine synthetase and therefore prevents the bacteria from using the ammonia themselves. The ammonia that is the end result can be used directly by the plant (Oldroyd and Downie, 2004).

Within the nitrogen-fixing clade, there are different forms of root nodule symbioses. So-called actinorhizal plants form symbiosis with Gram-positive bacteria from the genus *Frankia* (Markmann and Parniske, 2009). Actinorhizal plants are found in many groups within the clade, but they are believed to have evolved independently because of several

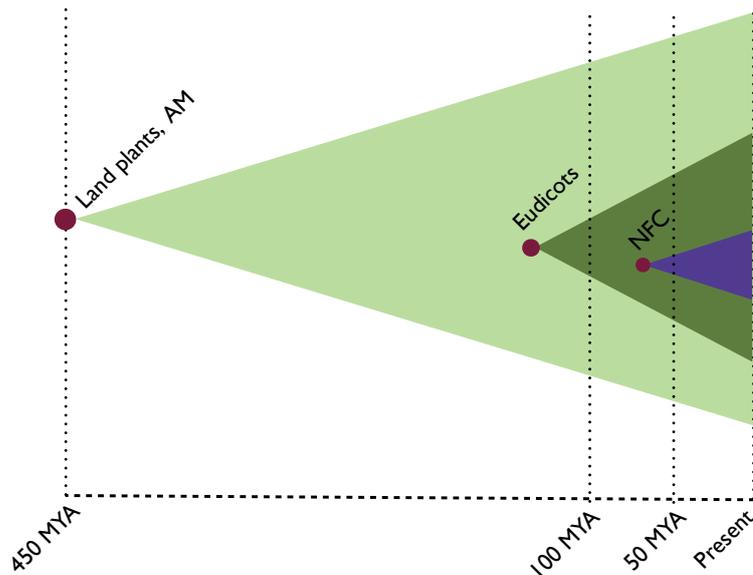


Figure 1.1: Evolution of root symbioses. Arbuscular mycorrhiza (AM) are believed to have aided plants to conquer land approximately 450 MYA. NFC stands for the nitrogen-fixing clade, containing all plants that form root nodule symbioses. After Kistner and Parniske (2002).

examples that have close non-nodulating neighbours. The most well-known N_2 -fixing symbioses, and the most interesting agriculturally, are found in the legume family *Fabaceae*. There, several groups associate with the diverse group of symbiotic Gram-negative bacteria collectively known as rhizobia. The only non-legume known to form a rhizobial symbiosis is the *Parasponia* genus, in the *Rosales* family (Markmann and Parniske, 2009).

Rhizobia-legume symbiosis

The large and varied *Fabaceae* family, of legumes, evolved around 60 MYA, which is during the period when many angiosperms are believed to have made their debut (Sprent, 2007). In legumes, specific signalling molecules from the bacterial symbionts induce a signalling pathway that leads to the establishment of the symbiosis. Many bacterial genes required for this symbiosis are shared between otherwise phylogenetically separate rhizobia, pointing to horizontal gene transfer. However, there is also a huge amount of variety in many aspects of the symbiosis, including the initiation stages and the final nodule structures. Several symbiotic traits are under the control of the host plant, which then exerts selective pressure on the microbial symbiont (Masson-Boivin et al., 2009). However, processes like the nitrogen-fixation are thought to be under rhizobial control (Sprent, 2007).

The symbiotic relationship between legumes and rhizobia is thought to have arisen multiple times. The legume family has three subfamilies and only in the most recent

one, *Papilionoideae*, are nodulators present at its base. Thus, nodulation should have been present from the start of this groups divergence (Doyle and Luckow, 2003).

The legumes *Medicago truncatula* and *Lotus japonicus* have become models for genetic research in legumes. The emergence of two models is down to differences in the symbiosis formed, for example *M. truncatula* and *L. japonicus* form different nodule types, indeterminate and determinate respectively. The determinate nodule type is also found in important agricultural plants such as bean and soybean. Indeterminate nodules are found in, for example, clover and pea. Therefore, studies into rhizobial infection of *L. japonicus* and *M. truncatula* in parallel can uncover important differences in these processes. Importantly, these plants also form AM symbiosis, and close genetic similarities in the initial stages of the rhizobial and AM symbioses have been found.

Connections between the symbioses are also present from a more ecological viewpoint. Nodulation has a requirement for phosphate, and has been shown to benefit from the presence of mycorrhizae (Smith and Douglas, 1987). *Lupinus* is the only known legume genus where the ability for AM symbiosis has been lost (Sprent and James, 2007), and in this special case so-called cluster roots help with phosphate acquisition. Other types of mycorrhiza, for example ECM, are sporadic in legumes and typically formed in some basal subfamilies. Interestingly, one branch which is completely non-nodulating is uniformly ECM (Lavin et al., 2005). However, ECM and nodulation are not mutually exclusive (Sprent and James, 2007).

1.3 Initial symbiotic events

1.3.1 Signal exchange in the soil

Plant-fungi dialogue

The key plant signal perceived by the fungal hyphae in the soil is strigolactone, which pushes the fungi from the initial growth stage into one of increased hyphal branching and metabolism (Akiyama et al., 2005; Besserer et al., 2006; Gutjahr et al., 2009a). Essentially the fungus is reaching out to the perceived nearby plant. Strigolactones are plant hormones that degrade rapidly in the soil and this characteristic makes them suitable as a signal for symbiotic location (Parniske, 2008).

In response to strigolactone, the fungi produce diffusible signal molecules of their own, that prepare the plant for fungal infection (Kosuta et al., 2008). Recently, it was shown that AM fungi release a mix of lipochitoooligosaccharides (LCOs), which have a chitin backbone with various decorations (Maillet et al., 2011). This is only the first identification of a mycorrhizal signal structure, and it is still unclear whether this is the only type of signal the fungi release or merely one type of signalling compound.

Plant-rhizobia dialogue

In the legume-rhizobia symbiosis, plant flavonoids (Zhang et al., 2009) are released from the plant roots and they are recognised by rhizobia in the soil (Perret et al., 2000). Flavonoids bind to and activate the bacterial NodD transcription factors, which in turn regulate the so-called nodulation (*nod*) genes that are required for infection (Fisher and Long, 1992). The NodD proteins are activated by different compounds in different rhizobia, and usually multiple NodD versions are present, thus contributing to host-bacterial specificity (Downie, 2010). Bypassing NodD activation using rhizobial mutants with constitutively active *nodD* promoters shows an increase in the rhizobia's host range (Cárdenas et al., 1995), demonstrating the role NodD can have in host specificity. Other key *nod* genes are the *nodABC* that are required for the synthesis of the Nod factor backbone, but there are many *nod* genes with known functions such as Nod factor secretion and modifications of the Nod factor structure (Masson-Boivin et al., 2009).

These Nod factors are diffusible signal molecules which initiate several developmental events in the plant root (Dénarié et al., 1996). They induce root hair deformation, which is required for root hair invasion in many legumes. Interestingly, Nod factors are also lipochitooligosaccharides (LCOs). Chitin is a typical fungal structure and LCOs also act in mycorrhizal symbiosis initiation where they are called Myc factors (Gough and Cullimore, 2011). Therefore, it is possible that rhizobia first started to produce LCOs by horizontal gene transfer, in this way mimicking the fungal symbiotic signal (Masson-Boivin et al., 2009). Nod factors can have a wide variety of modifications on top of this chitin backbone, and differences between strains play a key part in mediating host-specificity (Lerouge et al., 1990; D'Haeze and Holsters, 2002).

Other active compounds released by the rhizobia are polysaccharides, for example peptidoglycan which is a proteinacious polysaccharide from inside bacterial cell walls and has been implicated in the plant reception (Radutoiu et al., 2003). Polysaccharides have been shown to be required for rhizobial invasion (Pellock et al., 2000), showing that the Nod factors not sufficient for successful infection.

It should be noted that in the actinorhizal symbiosis, a *Frankia* factor has been hypothesised to play the same role as the Myc and Nod factors (Normand et al., 2007; Hocher et al., 2011). The characteristics of these different signal molecules and their responses will be dealt with more in depth in Chapter 5.

Plant signal receptors

Lysin motif (LysM) domains are key in perception of both Nod factors and chitin chains and are proposed to bind N-acetylglucosamine, which in turn forms the chain of chitin, the backbone of the LCO signal molecules (Buist et al., 2008). Detailed studies into the Nod factor structures and their effect on plant infection has led to the hypothesis that there are two key LysM-domain receptors perceiving Nod factors. There are also two modes of perception, which required different levels of specificity. The first is the

recognition of the basic structure of the signal, and this activates early responses. The second mode has a higher requirement on the exact signal structure such as decorations in the Nod factors, and this recognition controls further downstream events (Walker and Downie, 2000; Geurts and Franssen, 1996).

It is possible that these different levels come about through differential activation of one receptor or receptor complexes. In *M. truncatula* and *L. japonicus*, *MtNFP/LjNFR5* and *MtLYK3/LjNFR1* are putative Nod factor receptors (Radutoiu et al., 2003; Limpens et al., 2003; Madsen et al., 2003; Arrighi et al., 2006; Mulder et al., 2006). The mode of action could be a heterodimer at the plant plasma membrane, where *MtNFP1* has the predicted kinase domain and would transmit the perception downstream or interact with other proteins to form alternative dimers (Radutoiu et al., 2003). *MtLYK3* has been shown to be involved in early responses but the suggested function is to recognise structural differences in Nod factors (Smit et al., 2007).

MtNFP has recently been shown to be involved in Myc-LCO perception (Maillet et al., 2011), but surprisingly it is not essential for AM symbiosis (Amor et al., 2003). Other proteins in *M. truncatula* are also involved in LCO binding (Hogg et al., 2006), and *MtNFP*-independent signalling is found in the AM symbiosis (Chabaud et al., 2011; Mukherjee and Ané, 2011). Therefore, the receptors or complexes seem to have high redundancy (Kouchi et al., 2010). Also, small changes in CERK1, a chitin receptor in *A. thaliana*, allow it to complement an *nfr1* mutant of *L. japonicus* (Nakagawa et al., 2011), further showing how small differences can be crucial in perception and signal transduction.

Further components (beside the LysM receptor kinases) of the early recognition include lipid rafts, as demonstrated by the proteins linked to lipid rafts such as the flotillins FLOT2 and FLOT4 (Haney and Long, 2010), and SYMREM1 (Symbiotic Remorin 1) that interact with the receptors in root nodule symbiosis (Lefebvre et al., 2010). These proteins facilitate transport to and from membranes, and perhaps recruit important components of the signalling cascade in response to the rhizobial signal molecules.

1.3.2 The common SYM pathway

Through genetic studies in model legumes, several genes have been uncovered which are required for both the rhizobial and AM symbiosis. These are collectively called the common SYM genes, and are a part of the common symbiosis pathway that act in the early signalling events in the plant roots, leading to either rhizobial or fungal infection (Figure 1.2). Because of these shared elements, the bacterial symbiosis is thought to have recruited elements from the more ancient fungal pathway (Kistner and Parniske, 2002).

The common SYM genes are required after the fungal or rhizobial signal has been perceived by the appropriate receptor at the plant plasma membrane. Shared by both the rhizobial and AM signalling pathway is also the induction of calcium (Ca^{2+}) oscilla-

tions, also called Ca^{2+} spiking, in and around the nucleus, and these act as a secondary messenger to further relay the signal in the cell. Many of the common SYM genes either contribute to the production or the perception of these oscillations.

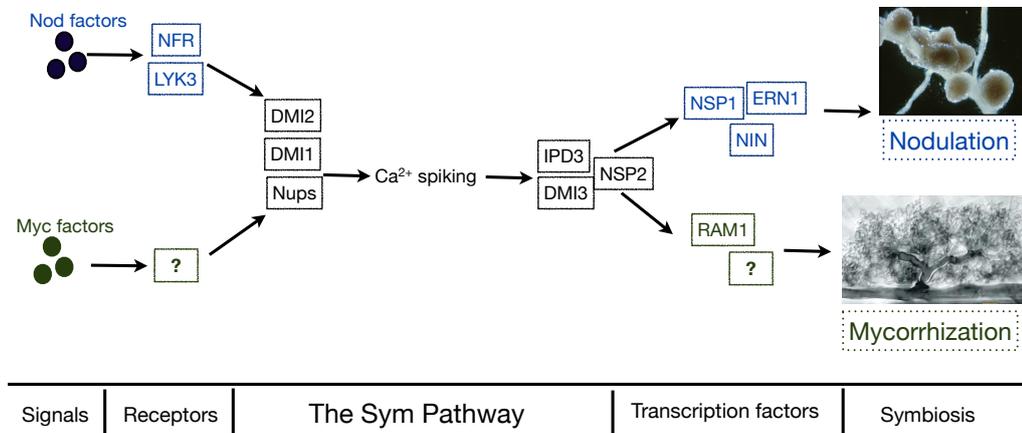


Figure 1.2: The SYM pathway. Shared components in early signalling events in the plant roots, leading to either rhizobial or arbuscular mycorrhizal infection.

DMI2

DMI2 is an early common SYM gene coding for a symbiosis receptor-like kinase, SymRK (Stracke et al., 2002). This is called nodulation receptor kinase, NOR, in *M. sativa* (Endre et al., 2002), and does-not-make-infection-2, *DMI2*, in *M. truncatula* (Bersoult et al., 2005). This is a receptor-like kinase gene with leucine-rich-repeats (LRR), an extracellular domain, a transmembrane domain and intracellular protein kinase domain. The mutants are impaired for Ca^{2+} spiking and downstream responses induced by Ca^{2+} spiking (Catoira et al., 2000; Wais et al., 2000). It has been shown that there is an enhanced sensitivity to touch in the *dmi2* mutant, and when avoiding this the plants show normal root hair deformation (Esseling et al., 2004). Therefore, the signal is perceived, but the signal inside the plant is not relayed. This mutation impacts on both rhizobial and AM symbioses, thus placing it firmly in the common SYM pathway (Gherbi et al., 2008; Markmann et al., 2008). There are mutations in the gene that cannot complement for nodulation, but that can for mycorrhization (Markmann et al., 2008). This raises questions of changes in *DMI2* (or its orthologues) can offer adaptation to new symbioses.

In the legume *Sesbania rostrata* (where the infection process does not necessarily start with root hair infection) the orthologue SymRK has been shown to have more roles than simply in the start of symbiosis; it is also involved later on during symbiosome formation in the cortex (Capoen et al., 2005). Furthermore, the mutant *symrk-14* uncouples the epidermal and cortical responses in *L. japonicus* (Kosuta et al., 2011).

The mutant plants have impaired infection thread formation, AM symbiosis, and NIN responses in the epidermis, but nodule formation and cortical infection was intact.

Secondary messengers

Many secondary messengers are induced in the signalling cascade of recognition, including Ca^{2+} , pH, reactive oxygen species (ROS) and phospholipid signals. Two Ca^{2+} responses are seen in root hair cells in response to Nod factors: the first is called Ca^{2+} flux and the second Ca^{2+} spiking. The Ca^{2+} flux is the earlier response in experimental conditions and precedes membrane depolarization (Felle et al., 1996). However, it does require higher concentrations of Nod factor than the induction of Ca^{2+} spiking (Shaw and Long, 2003). About one minute after Nod factor addition the cytoplasm becomes more alkaline (Felle et al., 1999, 1996, 1998). It has been shown that the Ca^{2+} concentration in a root hair cell is higher at the tip, and after Nod factor perception this gradient increases, and Ca^{2+} flows towards the nucleus (Cárdenas et al., 1999; Shaw and Long, 2003). Since this Ca^{2+} flux requires high concentrations of Nod factors, it could be that this occurs at a stage when bacteria are trapped close to the root hair and thus causes higher Nod factor concentrations locally (Miwa et al., 2006). Also, this flux seems to have a higher requirement of Nod factor structure specificity than other responses such as Ca^{2+} oscillations (Shaw and Long, 2003).

Another early response is ROS, and is produced in root hairs in response to Nod factors within minutes of recognition (Cárdenas and Quinto, 2008). They are also believed to play a role in the later infection process (Brewin, 2004).

Finally, phosphoinositide signalling through phospholipase C (PLC) and D (PLD) has been implicated by pharmacological studies to relay the signal before Ca^{2+} spiking (Pingret et al., 1998; Engstrom et al., 2002; Charron et al., 2004), and could play a role during the time from Nod factor addition to the start of Ca^{2+} spiking. Phosphoinositides are a type of phospholipids, small molecules playing major roles in eukaryotic secondary signal transduction and membrane compositions. For examples, phospholipids are precursors of the plant hormone jasmonic acid (JA) which is involved in defence responses. In animal cells, the phospholipids signal inositol-1,4,5-triphosphate (IP_3) is a key regulator which activates the release of Ca^{2+} inside cells, but no similar target receptors have been found in plants. However, inhibitors of PLC (Pingret et al., 1998; Engstrom et al., 2002; Charron et al., 2004) can still block reporter genes for the common SYM pathway, pointing to similar mechanisms of signal transduction (McAinsh and Pittman, 2009).

Encoding: the calcium spiking machinery

Several of the common SYM genes are components on the nuclear envelope, and known to be required for the production of the Ca^{2+} oscillations. One of these is DMII1, predicted to be a cation channel located preferentially in the inner nuclear envelope (Ané et al., 2004; Riely et al., 2007; Charpentier et al., 2008; Capoen et al., 2011).

Additional genes that are required for Ca^{2+} oscillations include components of the nuclear pore (Groth et al., 2010; Kanamori et al., 2006; Saito et al., 2007). These pores might be involved in trafficking secondary signals and/or ion channels to the inner nuclear membrane. There is also a Ca^{2+} pump, MCA8 (Capoen et al., 2011), which is thought to re-sequester Ca^{2+} into a store in the nuclear envelope and endoplasmic reticulum (ER) after each spike. A hypothesised Ca^{2+} channel for the release of Ca^{2+} is also necessary, but still unidentified.

These encoding components will be more closely examined in Chapter 4, where they form the backbone of a mathematical model simulating the Ca^{2+} spiking.

Decoding: the DMI3 complex

Downstream of Ca^{2+} spiking is a complex of proteins, in which a key player is a Ca^{2+} and calmodulin(CaM)-dependent kinase, CCaMK. The gene is known as *MtDMI3*, *LjC-CaMK* or *PsSYM9*, and is needed for downstream gene induction in both root nodule and AM symbiosis (Fearn and Larue, 1991; Mitra et al., 2004; Lévy et al., 2004). Interestingly, releasing autoinhibition for this kinase leads to spontaneous nodulation (Gleason et al., 2006; Tirichine et al., 2006), demonstrating the key role of this protein in nodule formation. CCaMK from *Oryza sativa* can complement *M. truncatula* for nodulation (Godfroy et al., 2006), and CCaMK from liverworts and hornworts can complement *M. truncatula* for mycorrhizal formation (Wang et al., 2010). This points to the highly conserved function of this protein, and is consistent with mycorrhization being essential for land plant colonisation. The protein is not needed for Ca^{2+} oscillations, but its absence leads to a lack of reporter genes for nodulation being switched on, the so-called *ENOD*-genes (Wais et al., 2000; Catoira et al., 2000). Some of these reporter genes, like *ENOD40*, are induced and involved in both symbioses (Charon et al., 1999).

One of the proteins that CCaMK closely interacts with is CYCLOPS in *L. japonicus*, or IPD3 (Interacting Protein of DMI3) in *M. truncatula* (Messinese et al., 2007). This is a nuclear coiled-coil protein which is needed for rhizobial infection to proceed normally, but is not required for nodule formation since spontaneous nodules can still be formed with a *cyclops* mutant (Yano et al., 2008). The mycorrhizal symbiosis is also affected, and that places IPD3 in the common SYM gene group (Gutjahr et al., 2008). The rice *IPD3* gene could complement mycorrhization in *M. truncatula*. IPD3 co-localises with CCaMK *in vivo*, and is phosphorylated by CCaMK *in vitro*, but the exact mechanism of interaction is not clear. Mutants are not completely lacking in symbiotic responses: nodule primordia are formed, infection initiated and some AM fungal colonisation occurs, i.e. more responses are present than in other common SYM gene mutants (Capoen and Oldroyd, 2008). Recently, even more roles for IPD3 have been shown, such as symbiosome formation and nodule development upstream of cytokinin signalling (Limpens et al., 2011). Since the *IPD3* gene is present in rice (Chen et al., 2008) it has been suggested to have a conserved role in the pathway (Wang et al.,

2010).

Even though there is clearly a complicated interaction taking place downstream of the Ca^{2+} oscillations, it should be noted that when it comes to nodule formation, when CCaMK is activated, the upstream genes of the common SYM pathway are not required (Gleason et al., 2006; Madsen et al., 2010; Hayashi et al., 2010).

In nodulation signalling, there is a pair of GRAS proteins that interact with the DMI3 complex: NSP1 (nodulation signalling pathway) (Smit et al., 2005) and NSP2 (Kaló et al., 2005; Oldroyd and Long, 2003). These GRAS proteins are binding to the promoters of early nodulin genes, promoting their transcription, and NSP2 interacts directly with CCaMK linking the complex together at the promoters (Hirsch et al., 2009b). NSP-like genes from tobacco and rice can complement legume *nsp* mutants, showing a conserved function over evolution of the SYM pathway (Heckmann et al., 2006; Yokota et al., 2010). Furthermore, NSP2 has recently also been implicated to have a role in the mycorrhizal pathway (Maillet et al., 2011). Two other transcription factors, ERN1 (ERF required for nodulation) (Middleton et al., 2007; Vernié et al., 2008) and NIN (nodule inception) (Schauser et al., 1999), also act downstream of CCaMK.

There is evidence supporting the presence of alternative pathways in parallel with CCaMK activation. This is suggested for example by experiments in which a gain-of-function *ccamk* cannot complement for infection (Madsen et al., 2010; Hayashi et al., 2010). What could be the controllers of such an infection pathway? In *M. truncatula*, LYK3, LIN and RPG are possible candidates (Limpens et al., 2003; Smit et al., 2007; Arrighi et al., 2006; Kiss et al., 2009). Such alternative pathways may also be present in the mycorrhizal symbiosis.

1.4 The role of calcium signalling

Increases in cellular Ca^{2+} levels play a role in many signalling pathways, and both abiotic and biotic stimuli gives rise to Ca^{2+} signals. There are well-known examples of where Ca^{2+} plays a role: in fertilisation, neural signals and cell death. Many proteins are targets of Ca^{2+} , including calmodulin (CaM), Ca^{2+} -dependent kinases and CaM-binding proteins, and Ca^{2+} often regulates downstream phosphorylation through these proteins (Kudla et al., 2010). Since Ca^{2+} plays a role in so many processes, a key question has been that of specificity i.e. how Ca^{2+} is capable of inducing such different downstream responses (Sanders et al., 2002). However, in some systems Ca^{2+} signals have been suggested to act only as a switch (Scrase-Field and Knight, 2003).

Many Ca^{2+} signals occur in the cytosol, in that Ca^{2+} is released from subcellular compartments such as the vacuole and cause an increased concentration of Ca^{2+} in the cytosol (Dodd et al., 2010). These Ca^{2+} transients can happen repetitively, and are then called Ca^{2+} oscillations or Ca^{2+} spiking. But a release of Ca^{2+} can also occur inside a compartment, such as a release from the nuclear envelope into the nucleus, and this is the case in the common SYM pathway (Oldroyd and Downie, 2004), where

the Ca^{2+} oscillations are perinuclear and the principal Ca^{2+} store is believed to be the nuclear envelope, contiguous with the ER (Oldroyd and Downie, 2006). Relatively little is known about nuclear Ca^{2+} signalling but it is a growing field (Bootman et al., 2009).

1.4.1 Calcium oscillations in animal cells

Ca^{2+} is an important signalling ion in animal cells as well as plant cells, and in many cases, the animal field has come further in elucidating the mechanisms of Ca^{2+} signalling than the plant counterpart.

Secondary messengers such as IP_3 often activate the machinery producing Ca^{2+} oscillations, via an IP_3 receptor (IP_3R) (Berridge, 2009). The IP_3R is located on the ER membrane, and release Ca^{2+} into the cytosol. STIM proteins on the ER sense when the ER Ca^{2+} store becomes depleted and then move to junctions between the ER and the plasma membrane. There the STIM proteins induce the assembly of CRAC channels, and their subunits consist of ORAI proteins (Hogan et al., 2010; Prakriya et al., 2006). The CRAC channels are highly Ca^{2+} -sensitive and let Ca^{2+} into the cytosol. Sarcoplasmic/endoplasmic reticulum calcium channel ATPases (SERCA) on the ER can now take up this Ca^{2+} into the depleted ER store. There are also pumps on the plasma membrane, PMCA, which can transport Ca^{2+} out of the cell when necessary (Parekh, 2011). Decoding of the Ca^{2+} oscillations in this system is for example done via Ca^{2+} /calmodulin-dependent protein kinase II (CaMKII) or protein kinase C (PKC) (Clapham, 2007). There are of course many other types of channels and pumps, and for a more complete review see Parekh (2011) or Clapham (2007).

In animal systems, both amplitude modulation and frequency modulation encoding by Ca^{2+} have been described (Parekh, 2011). Amplitude modulation encodes the information through the affinity of the Ca^{2+} sensor. In frequency modulation, the dynamics of the Ca^{2+} signal over time carries the information. These two options are further specialised by subcellular localization of the proteins, so the spatial information can separate sensors which may otherwise have been activated by the same signals.

1.4.2 Calcium oscillations in plant stomatal guard cells

In stomatal guard cells, many stimuli induce changes in cytosolic Ca^{2+} concentrations, leading to Ca^{2+} oscillations or Ca^{2+} transients. In *Arabidopsis thaliana*, these stimuli include ABA, cold, external Ca^{2+} and hydrogen peroxide, and downstream of the Ca^{2+} signals these stimuli lead to stomatal closure (Allen et al., 2001, 2000).

A correlation has been documented between strength of external Ca^{2+} stimuli, the pattern of the internal Ca^{2+} oscillations and the level of stomatal closure (McAinsh et al., 1995). This suggests that the Ca^{2+} oscillations could encode information on the strength and the type of the stimulus (McAinsh and Pittman, 2009). The Ca^{2+} oscillations are required to have a certain frequency, amplitude and duration to completely

close the stomata, but the pattern of the Ca^{2+} oscillations is only important in deciding the prolonged stomatal movement and not for the initial rapid closure after stimulus (Allen et al., 2001).

However, many things are still not understood about this pathway, such as how the stomatal movements are regulated in the long-term, nor what mechanism allow the signals to be stimulus-specific (Luan, 2011).

1.4.3 Calcium oscillations in the SYM pathway

In legumes, the Ca^{2+} oscillations also known as Ca^{2+} spiking are observed shortly after Nod factor perception by the plant, and the structure of the Nod factors has been shown to be crucial as to whether or not the response occurs (Ehrhardt et al., 1996). Nod factors can initiate Ca^{2+} oscillations at very low concentrations, from 10^{-13} M (Oldroyd et al., 2001b). The Ca^{2+} oscillations in root hair cells in response to Nod factors are observed after a lag of approximately 10-20 minutes. This lag probably involves changes in Ca^{2+} balances in the cell, which could be connected to the Ca^{2+} flux. However, the two responses can be separated (Shaw and Long, 2003; Miwa et al., 2006). It has also been demonstrated that the Ca^{2+} oscillations occur both on the inside and outside of the nucleus (Sieberer et al., 2009; Capoen et al., 2011), whereas the Ca^{2+} flux is found in the cytosol and not specifically in the nuclear area.

In a single Ca^{2+} transient during the oscillations, the initial upward phase is very rapid, and the subsequent fall in Ca^{2+} is slower. This is thought to be due to a rapid release of Ca^{2+} and a slower re-sequestering by Ca^{2+} pumps into the nuclear envelope, which is considered to be the Ca^{2+} store since the release is localised to the nuclear region (Oldroyd and Downie, 2004). The oscillations can go on for several hours, and both the length of this sustained spiking and the average period of spikes is variable between cells. The period, however, commonly lies between 60-120 seconds (Miwa et al., 2006). Ratiometric measurements of amplitudes have indicated that the concentration levels of Ca^{2+} released in *M. sativa* are in the vicinity of 500 nM (Ehrhardt et al., 1996). Measurements in pea (Walker et al., 2000), *M. truncatula* and *S. rostrata* (Capoen et al., 2009) have shown amplitudes around 200 nM. Example spiking from *M. truncatula* is shown in Figure 1.3.

The oscillations are perceived by DMI3 and IPD3. The Ca^{2+} signals are necessary for the activation of downstream genes, and it has been shown that about 36 spikes are required for the induction of reporter genes for nodulation (Miwa et al., 2006). Also, the position along the root influences the period of the spikes, with the cells closer to the root tip oscillating faster (Miwa et al., 2006).

Ca^{2+} spiking induced by Nod factors has so far been demonstrated in a variety of legume species, e.g. *M. truncatula* (Wais et al., 2000), *Pisum sativum* (Walker et al., 2000), *Phaseolus vulgaris* (Cárdenas et al., 1999) and *L. japonicus* (Harris et al., 2003) and *Sesbania rostrata* (Capoen et al., 2009), but it should be noted that all these plants are relatively closely related, and belong to the third subfamily of the legumes,

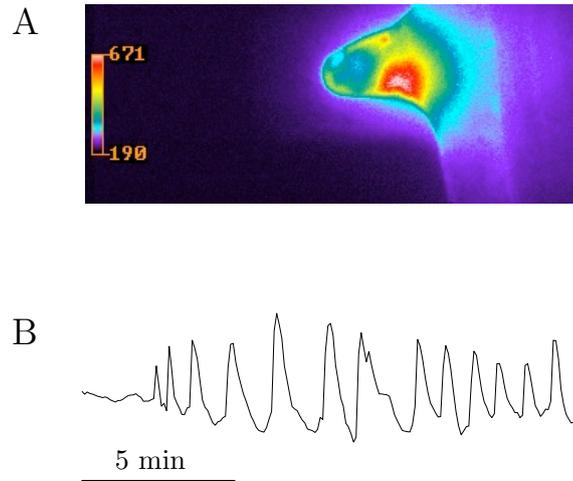


Figure 1.3: Ca²⁺ spiking in the model legume *Medicago truncatula*. A: A root hair cell injected with the Ca²⁺ indicator dyes Oregon Green and Texas Red. The highest fluorescence, and therefore Ca²⁺ concentration, can be seen in the nuclear area. B: A time series of Ca²⁺ levels in the root hair above, showing how the oscillations start. They were induced by Nod factors from *S. meliloti*, the rhizobial partner of *M. truncatula*.

the Papilionoids.

AM fungi also produce diffusible LCO signals and Ca²⁺ spiking has been recorded when legumes have been in close proximity to AM fungal hyphae (Kosuta et al., 2008). Interestingly, the AM and rhizobial signals reported in the study appeared different (Kosuta et al., 2008), and it was hypothesised that Ca²⁺ could be a conveyor of specificity in the common SYM pathway (Kosuta et al., 2008; Chabaud et al., 2011; Oldroyd et al., 2009)

It is clear that both AM and rhizobia activate the same signalling pathway, which diverges downstream resulting in either nodules or mycorrhizal symbiosis (Kosuta et al., 2008). The protein downstream of the spiking, CCaMK, has been predicted to be able to distinguish different Ca²⁺ signals (Lévy et al., 2004). However, the question remains of how such information would be encoded in the spiking (Oldroyd and Downie, 2006). It could be a question of spiking frequency, individual spike shape, and inter-spike period, but could also involve other elements like spatial patterning, considering that the spiking is localized around the nucleus of the cell. Essentially this question concerns the transfer and content of information. More research is clearly needed to understand this step in the initiation of symbiosis.

Using nuclear targeted constructs to monitor changes in Ca²⁺ has allowed further analysis (Sieberer et al., 2009) and a closer focus on similar cell types has also proved useful (Sieberer et al., 2011). This is because, in *M. truncatula*, mycorrhiza enter

through non-root hair cells and rhizobia through root hair cells: hence cell type-specific responses may confuse the comparison. It is noteworthy that similar Ca^{2+} spiking patterns were found in cells of the cortical layer during infection in both symbioses (Sieberer et al., 2011).

A non-LCO compound, mastoparan, also induces Ca^{2+} oscillations throughout the cell (Pingret et al., 1998), and with somewhat higher variability than those induced by Nod factors (Sun et al., 2007). Interestingly a synthetic analogue of mastoparan induced spiking in mutants *nfp*, *dmi1* and *dmi2* that are blocked for Nod factor-induced Ca^{2+} spiking (Sun et al., 2007). But the hypothesised decoder, CCaMK, was needed for further relay of the Ca^{2+} oscillations (Pingret et al., 1998; Sun et al., 2007). Therefore, a different machinery is probably involved in the production of the signals, but it seems that some SYM components are still needed.

Other roles for Ca^{2+} in symbiosis are even less understood than the LCO-induction of Ca^{2+} oscillations. For example, several Ca^{2+} -dependent protein kinases (CPKs) in rice are expressed in cortical cells during AM establishment, and are transcriptionally induced by signals secreted by the fungi (Campos-Soriano et al., 2011). There have also been reports of a rise in cytosolic Ca^{2+} for soybean, *Glycine max*, in the presence of AM fungal spores (Navazio et al., 2007). Perhaps with the isolation of the LCOs from the AM fungi *Glomus intraradices* (Maillet et al., 2011) such signals can be investigated further.

1.4.4 Calcium oscillations in defence pathways

Ca^{2+} signals in the cytosol are a key part of defence pathways in plants. Changes in the cytosolic Ca^{2+} concentration are activated in a range of plant-pathogen interactions, and can result in transients, oscillations, or sustained elevations of Ca^{2+} concentrations. For a review on Ca^{2+} signals in plant defences see Lecourieux et al. (2006).

Chitin oligomers are not only the backbone of Nod factors, they are also well-known defence-inducing molecules. They induce Ca^{2+} spiking in *P. sativum*, but with a less sustained response and a lower period than with Nod factors (Walker et al., 2000). Unlike the mastoparan-induced spikes, this response was lacking in common SYM mutants, which indicates that this relies on the same pathway as Nod factors. This is relevant since it could suggest an overlap between defence signalling and the common SYM pathway. Alternatively, it may be a consequence of a partial activation of Nod factor receptors by the Nod factor backbone.

1.5 Establishment of endosymbiosis

Downstream of the SYM pathway lie two routes to the establishment of endosymbiosis. Below is an overview of key events in both mycorrhizal and rhizobial symbiotic development.

1.5.1 Mycorrhizal infection

When the AM fungi have branched and reached the plant root, and the plant is a suitable host, the infection process begins. The fungus comes into contact with the epidermis, often at young lateral roots, and forms an hyphopodium, see Figure 1.4 (Bonfante and Genre, 2010). The plant cells underneath prepare for symbiosis. The epidermal nucleus migrates towards the contact area, and a structure called the prepenetration apparatus (PPA) start to form, guided by the nucleus and predicting the infection of the fungal hyphae. The PPA is a cytoplasmic bridge, like an empty road for the fungi, and it stretches across several cells. When it is completed the fungus is allowed to enter the plant cell (Parniske, 2008). Formation of the PPA is dependent on the common SYM pathway, and its formation is correlated with the induction of reporter genes such as *ENOD11* (Genre et al., 2005). The fungal path can also be predicted by the PPA all the way into the outer cortex, but the inner cortex gives rise to more complex patterns (Genre et al., 2008).

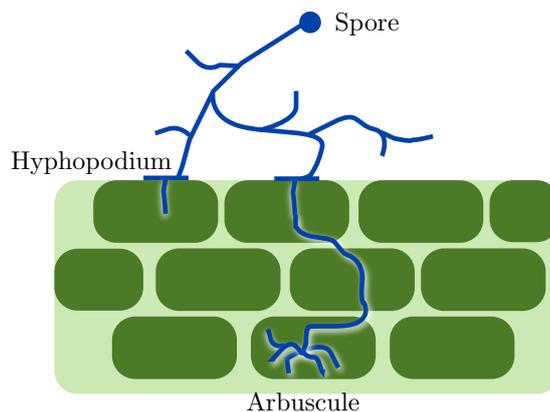


Figure 1.4: Overview of infection of the plant root by arbuscular mycorrhiza. After Bonfante and Genre (2010).

In the inner cortex, arbuscules are formed inside the plant cells. This is the bush-like structure that is shown in Figure 1.2 at the end of the mycorrhizal signalling pathway, and this stage of symbiosis is not considered to be dependent on the common SYM pathway (Bonfante and Genre, 2010). Genes such as *Vapyrin* are required for arbuscule formation, but the exact mechanism is unknown (Murray et al., 2011). Surrounding the arbuscule is a membrane derived from the plant, continuous with the plant plasma membrane but with a different composition. Nutrient exchange takes place across this membrane. The arbuscule lasts a limited time, a matter of days, which is probably related to the level of benefit that it provides to the plant. Phosphate transport over this membrane has proven essential for the arbuscule, and may be a way for the plant to control that the symbiosis is efficient (Oldroyd et al., 2009), and so phosphate signalling appears to be key in maintaining the symbiosis (Javot et al., 2007a,b; Maeda et al.,

2006). The plant makes sure that the fungal hyphae with the best access to nutrients at any given moment are preferred symbiotic partners, highlighting the dynamic process of this symbiosis.

1.5.2 Rhizobial infection

Intracellular entry

There are multiple ways for the rhizobia to enter their legume host. The most studied entry is intracellular, through the root hairs, and occurs in the model legumes such as *M. truncatula*. There, the induction of root hair deformation by Nod factors is a first step towards entrapping the bacteria fully. Once the rhizobia have attached to the plant root hair, root hair curling will start, and this will enclose bacteria and begin the process of infection (Figure 1.5). An ingrowth of the membrane and cell wall forms the infection thread (IT), which will grow through the root hair with rhizobia multiplying inside it forming a column that progresses down, through cell layers, towards the cortex. This IT is formed by the plant and controlled by the plant (Oldroyd et al., 2009), and a single rhizobial strain will usually remain in this IT even if there were more from the start. In this way, the host plant reduces the probability that any pathogenic (non-fixing) bacterium can co-infect the same nodule as the symbiotic strain of *Rhizobium* (Downie, 2010). The start of IT formation is preceded by cytoskeletal rearrangements in the microtubule and actin network, and this is a part of the change from normal root hair growth to the inward growth of the membrane; both actin and cytoskeleton dynamics continue to be important in IT growth (Yokota et al., 2009). Further, the nucleus is positioned close to the entrapped bacteria, and will guide the IT through the cell.

In cells that are about to be transversed by the IT, a so-called pre-infection thread is formed (van Brussel et al., 1992), and this bears similarities to the prepenetration apparatus (PPA) from the AM symbiosis. In non-symbiotic responses no such similar structure forms (Genre et al., 2009), and this is another common response that seems to have been conserved in the evolution of these root symbioses. The pre-infection thread appears in cortical cells and can to some extent predict where the IT is going to go, like the PPA in the AM symbiosis. Ca^{2+} is once again implicated at these stages, forming gradients that are thought to be necessary for both IT progression and root hair deformation (Murray, 2011). The signalling during the IT progression through the cells also involves bacterial surface polysaccharides (Gibson et al., 2008).

The IT finally reaches the cortex, where cell divisions have been initiated in a coordinated manner to the infection, forming the nodule. When the IT reaches the cortical cell, the cell wall is degraded (van Brussel et al., 1992). Then the bacteria are released into the cells and terminally differentiate into bacteroids, a form of rhizobia that resides inside the nodule and fixes nitrogen.

Intercellular entry

Another way for the rhizobia to enter the root is intercellularly. This takes place through cracks in the epidermis, for example at lateral root bases or wounding sites, but the exact mechanism varies depending on the plant species (Sprent, 2009). Sometimes ITs are formed, in other cases they are developed later or absent altogether.

The most studied example of intercellular entry, also called crack-entry, is found in the tropical legume *Sesbania rostrata*. Interestingly, it has been shown in *Sesbania* that Ca^{2+} spiking in the epidermis may not be necessary for this type of entry (Capoen et al., 2009). It is still, however, required for underlying nodule organogenesis. This points to a differential activation of DMI3 in epidermis and cortex.

Also, it has been shown that in several legumes, including *Lotus*, a basic infection intercellularly can occur without the need for Nod factors. This does, however, require SYMRK, and the bacteria reached the nodule directly between cells and without ITs. There was also the possibility of infection through cracks in the epidermis. Thus, several ways are available for the rhizobia to enter the plant, with different stringencies and different levels of efficiency (Karas et al., 2005; Madsen et al., 2010).

More information on natural variation of infection processes will be found in the introduction of Chapter 6.

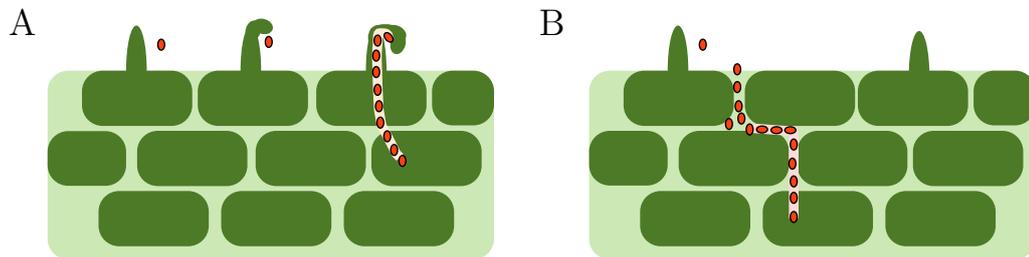


Figure 1.5: Different modes of rhizobial entry in symbiosis. A: Root hair invasion, intracellularly after root hair curling. B: Crack-entry or epidermal invasion, intercellularly. After Oldroyd and Downie (2008).

The nodule

Underneath the epidermis lies the cortex, and this is where the nodule forms. In response to Nod factors, plant nodulation genes are induced by the protein complex formed by DMI3. These nodulation genes play a part in starting the cell divisions that becomes the nodule primordia. It is crucial that the nodule is formed at a site close to where the IT is formed and progresses, so the processes must be synchronised (Oldroyd and Downie, 2008). However, the Nod factors are perceived at the epidermis, while the nodule primordium is in the cortex, so there should be a signal between cell layers that is still unidentified. A schematic overview of the signals in the different cell layers is shown in Figure 1.6.

The hormone cytokinin is required for nodule growth, as well as the transcription factor NIN (Marsh et al., 2007), which is up-regulated in response to both Nod factors and cytokinins. NIN has been suggested to play a part in the communication between epidermal and cortical cells, and *nin* mutants still show normal epidermal responses highlighting the active role required by NIN in the cortex (Crespi and Frugier, 2008). It has been suggested that cytokinin signalling is the underlying controller of nodule growth, and that cytokinin also controls NIN and its role in infection (Frugier et al., 2008).

There must of course be a way for the plant to ensure that it does not form more nodules than necessary. This is done partly through plant long-distance signalling known as autoregulation of nodulation (AON) (Mortier et al., 2011). An additional set of proteins involved in nodulation are the CLAVATA3/ endosperm-surrounding region-related (CLE) peptides (Mortier et al., 2010), and CYTOKININ RESPONSE1 (CRE1) (Plet et al., 2011; Gonzalez-Rizzo et al., 2006). It has been shown that CLE can regulate NIN via response regulators (RRs), linking AON with the cytokinin and common SYM pathway (Saur et al., 2011). Cytokinin and other hormone influences will be discussed further in the next section.

The nodule structure depends on the plant species, the two main types being determinate and indeterminate. A determinate nodule starts forming in the central cortex, resulting in a round nodule with a limited life span. An indeterminate nodule on the other hand comes from the inner cortex and has continuous apical meristematic growth, resulting in a more cylindrical shape (Sprent, 2009). There are other morphological and functional differences in these nodules, and there are also other variations such as lupinoid nodules.

Inside the host cells of the nodule, the symbiosome can be found. This organelle-like structure harbours one or several bacteroids enclosed in a plant membrane. The bacteroids can be similar in size to free rhizobia, or enlarged, depending on the legume species. In the “galeoid” tribe of legumes (which includes *Medicago*, *Lotus* and *Pisum*) the bacteroids are terminally differentiated. In *M. truncatula* they are enlarged by endoreduplication (Oldroyd et al., 2011). There is also an increase in the permeability of the membrane of the *M. truncatula* bacteroid, and it seems that cysteine-rich peptides found in bacteroids are connected to both the trait of membrane permeability and the endoreduplication (Van de Velde et al., 2010). Another protein that is key for bacteroids is BacA, an ATP-binding cassette superfamily (ABC) transporter that is thought to be involved in the transport of these cysteine peptides from the plant to the bacteroid (Oldroyd et al., 2011). Many traits are lost in the symbiotic bacteroid compared to the free-living form, and in some cases this leads to dependence on the plant to provide some amino acids (Prell et al., 2009). These observations, taken together, shows how the bacteria become adapted to living inside the plant, controlled by the plant. The bacteroids are also being monitored in terms of nitrogen production, and if they do not fix nitrogen the nodule will be shut down, showing once again a way for the plant

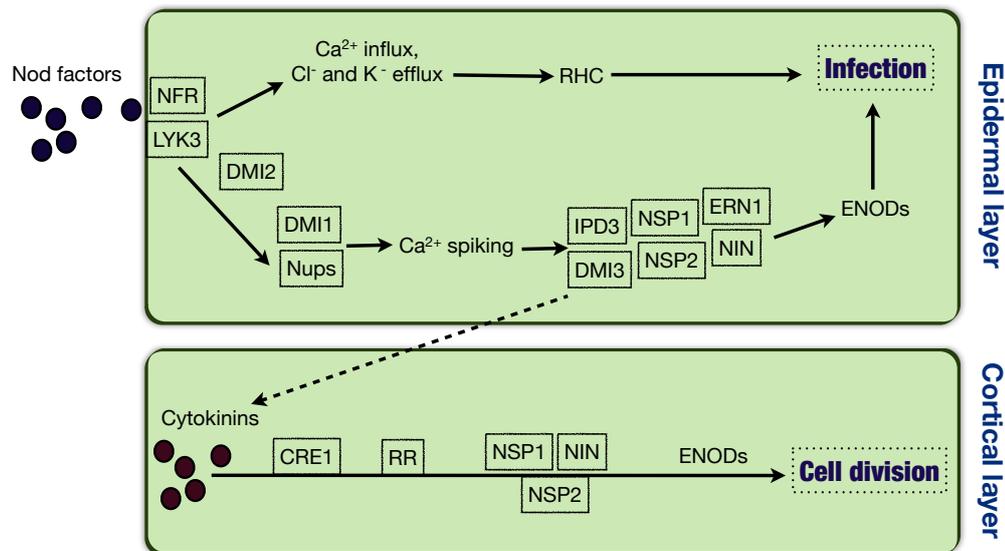


Figure 1.6: The Nod factor signalling pathway in different plant cell layers, after Oldroyd and Downie (2008).

to prevent impostors from infecting the root (Oldroyd et al., 2009).

1.6 Hormone influences

Endosymbioses involve changes in all major plant hormones, but the key players are presented here, especially those affecting Ca^{2+} signals.

Cytokinin

Cytokinin is an adenine-derived signal that is involved in many plant processes. It is a key player in the rhizobial symbiosis, and function downstream of Nod factor perception, dependent on the common SYM genes but upstream of the nodulation-specific transcription factors like NIN. It is crucial for nodule formation (Frugier et al., 2008).

The cytokinin receptor MtCRE1/LjLHK1 is required for nodule initiation by cell divisions, and to some extent also for infection thread progression (Murray et al., 2007; Frugier et al., 2008). Receptor homologues have been found in other legumes as well, such as *Lupinus albus* and *Medicago sativa* (Coba de la Peña et al., 2008a,b). The lack of LHK1 causes hyperinfection but absence of nodules (Murray et al., 2007), and furthermore, a gain-of-function *LHK1* mutant develops spontaneous nodules in the absence of rhizobia (Tirichine et al., 2007). This points to cytokinin as being an important regulator of nodule organogenesis, and these gain-of-function *LHK1* mutants

show similarity to the spontaneous nodulation of a *CCaMK* mutant (Gleason et al., 2006; Tirichine et al., 2006).

LHK1 is indicated to repress lateral root formation (Eckardt, 2006; Gonzalez-Rizzo et al., 2006), so could act as an important switch from one developmental program to another, namely nodule formation. The nodule originates in the cortex, and the cytokinin signalling could be a link between the epidermis, where Nod factor perception takes place, and the cortical cells (Oldroyd, 2007). Indeed, simple addition of exogenous cytokinin induces nodule primordia and nodulin genes in the cortex of *L. japonicus*, but the genes *NSP1*, *NSP2* and *NIN* were required for this to take place, in addition to *LHK1* (Heckmann et al., 2011).

The response regulators that act after *CRE1* but before the common SYM genes in the cortex are only active in the dividing cortex and pericycle tissues (Plet et al., 2011), and so this seems to limit the spatial role of cytokinins.

It should be noted that rhizobia produce cytokinins, and intriguingly, in Nod factor-independent symbioses cytokinin-precursors are thought to act as important signalling molecules, possibly circumventing the common SYM pathway (Newman et al., 1994; Giraud et al., 2007)

Thus, increases in cytokinin would be a key consequence of Nod factor perception, and trigger nodule formation. Furthermore, the roles of these cytokinin receptors are likely to be varied, and function throughout the lifetime of the nodule, as well as being involved in cross-talk with other pathways (Coba de la Peña et al., 2008c). It has been suggested that cytokinin is even integrating different inputs from other pathways, such as ethylene (Plet et al., 2011).

Auxin

Cytokinin seems to act together with auxin responses in the initiation of the nodule. A local accumulation of auxin occurs where the nodule primordia is formed, by an inhibition of polar auxin transport (Mathesius et al., 1998; van Noorden et al., 2006) through *MtCRE1* (Plet et al., 2011). Auxin and cytokinin work in opposition in the cells of a plant root: auxin is high at the tip, where cells divide, cytokinin is high at the crossroad between proximal meristem and elongation zone, so where cells differentiate. At both sites, the prevailing hormones suppress each other. Also, if considering the formation of a lateral root, auxin marks the site of initiation and cytokinin is suppressed at the start, but soon after cytokinin levels rise high (Benková and Bielach, 2010; Bishopp et al., 2011; Péret et al., 2009; Laplaze et al., 2007). This is opposite to the case for nodule initiation, where cytokinin suppresses auxin in the beginning. Manipulating auxin biosynthesis leads to changes in the number and size of nodules (Camerini et al., 2008). Auxin probably has a complicated role in nodulation with positive effects at some stages and negative effects at others (Ding and Oldroyd, 2009). Discussion on the exact role of the antagonists auxin and cytokinin are ongoing (Oldroyd et al., 2011).

Ethylene

Ethylene is a plant hormone that affects many processes, making it difficult to pinpoint its particular role(s) in symbiosis. It has been shown to negatively regulate the nodulation pathway, upstream of Ca^{2+} spiking (Oldroyd et al., 2001a), and represses both infection and nodule formation. In the AM symbiosis, tomato mutants that have an increased ethylene response showed a negative effect on the root colonisation (Torres de Los Santos et al., 2011). On the other hand, in the aquatic legume *S. rostrata*, ethylene seems to have a positive effect on lateral root nodulation (D’Haeze et al., 2003). This could be a consequence of the legume’s aquatic lifestyle, allowing it to alternate between rhizobial invasion strategies and nodule development (Fernandez-Lopez et al., 1998; Capoen et al., 2009). The fact that another legume, soybean, is insensitive to ethylene in its nodulation has added to this species-specific aspect of ethylene responses in symbiosis, and has highlighted the need for more studies on ethylene in a wider range of nodulating legumes (Guinel and Geil, 2002).

Jasmonic acid

Jasmonic acid (JA) works alongside ethylene in the suppression of nodulation. This is not surprising given its general role in plant defence signalling. The Ca^{2+} spiking on the other hand is affected differently by JA relative to ethylene: JA not only reduces the amplitude of Ca^{2+} spikes but also reduces the frequency of these spikes, an effect that is stronger in ethylene-insensitive plants (Sun et al., 2006). JA has also been shown to act in the AM symbiosis, (Garrido et al., 2010). In the case of *S. rostrata*, an increase in JA slowed Ca^{2+} spiking and thus had a similar effect to a decrease in ethylene (Capoen et al., 2009), showing once again the variation in responses that may exist between plant species.

Recent findings have indicated that JA may not be directly involved in the development of nodules, since biosynthesis of JA does not vary between nodulated and non-nodulated plants (Zdyb et al., 2011). There could still be another role for JA though, for example in the AON and movement between root and shoot.

Abscisic acid

The role of Ca^{2+} signals in guard cells is intimately related to abscisic acid (ABA) levels, a stress hormone regulating plant water status (Pei et al., 2000). ABA is an additional negative regulator of the nodulation pathway and its downstream responses, but interestingly act independently from the ethylene pathway. Furthermore, ABA affects both epidermal and cortical processes, pointing to a coordinating role in nodulation signalling (Ding et al., 2008).

Salicylic acid (SA) also negatively regulates infection and nodulation and thus joins the group of ethylene, ABA and JA. A threshold for response to Nod factors seems to lie in the balance between these hormones and Nod factor levels, suggesting the

accumulation of secondary messengers such as ROS could be the integrator of these pathways (Ding and Oldroyd, 2009).

1.7 Cross-talk with defence responses

When invading the plant, the microbes must overcome the plant's normal response to an attack, as the defence responses will automatically kick in when detecting signs of pathogens. Plants have evolved elaborate ways of recognising molecules from both fungi and bacteria, so-called pathogen-associated molecular patterns (PAMPs), and mount defences to them. However, those defences are not expressed, or only transiently expressed, during symbiotic entry of legume roots (Benedito et al., 2008). In the case of AM fungi, there must be specific strategies to avoid activation of defence pathways by chitin, which is a well-known PAMP.

Pathogens normally try to strike back by releasing effector proteins, inhibiting the plant responses to PAMPs. Do such proteins play a role in symbiosis, to avoid active defence pathways? Indeed some rhizobia have effector proteins (Oldroyd et al., 2009; Deakin and Broughton, 2009). AM fungi have recently been shown to secrete effectors interacting with and down-regulating defences (Kloppholz et al., 2011), and the same has been found in ECM fungi symbioses (Plett et al., 2011). There is probably a wide variety of molecules being released that target defences, and many such components and systems still remain to be discovered.

As mentioned earlier, chitin oligomers can elicit Ca^{2+} SYM-spiking (Oldroyd et al., 2001b; Walker et al., 2000), and can also activate downstream genes for nodulation (Nakagawa et al., 2011). These similarities between defence and symbiosis responses both in Ca^{2+} signals and receptors highlight the need to test these interactions for specificity, for example the chitin receptors.

Parasitic nematodes trigger similar responses to those induced by Nod factors in *L. japonicus*, and seem in part to make use of some common SYM components (Weerasinghe et al., 2005). Another similarity between parasitism and symbiosis is long distance transport of signals. In pathogen interactions, systemic acquired resistance allows the rest of the plant to be ready when attacked in one specific location, and in symbiosis CLE peptides have been proposed to play a similar role to regulate nodulation and mycorrhization systemically (Staelin et al., 2011).

The list of similarities between mutualism and parasitism is continuously growing, including signals released, signal perception mechanisms, transcriptional responses in the host and invasion mechanisms (Paszkowski, 2006), showing the continuum of plant-microbe interactions.

1.8 Project objectives and thesis structure

Nuclear-localised Ca^{2+} oscillations have been shown to be necessary in the establishment of both AM and rhizobial symbioses. This thesis focuses on an analysis of the nature of this Ca^{2+} signature and an investigation of how this signature might have changed during the evolution of AM and root nodule symbioses.

The work is interdisciplinary, combining experimental and mathematical biology to deepen the understanding of this signalling pathway. The approaches of the project can be divided into 1) experimental measurements of Ca^{2+} spiking, 2) data analysis to characterize different features of the oscillations, and 3) theoretical modelling to test hypotheses of how the Ca^{2+} oscillations are produced and what could influence their characteristics.

Chapter 2 is a general methods and materials outline for the project. Chapter 3 describes a method developed and used for the frequency analysis of the Ca^{2+} oscillations. Chapter 4 presents an analysis of a mathematical model of the oscillatory system, with focus on how Ca^{2+} -binding proteins could cause variation in the signal. Chapter 5 compares Ca^{2+} spiking from the AM and rhizobial symbioses, both with fungi-induced oscillations and induction with isolated LCO signals from AM and rhizobia. Chapter 6 presents measurements of Ca^{2+} spiking from different plant species and compares them with the developed data analysis methods, to look for correlations between phylogeny and variations in Ca^{2+} signatures. Chapter 7 is a summary and critical discussion.

Chapter 2

Methods and materials

2.1 Experimental methods

2.1.1 Plant growth

Unless stated otherwise, all plants were grown in a controlled environment at 20°C and with 16 h light and 8 h dark. The relative humidity was 32% and the light intensity was 300 $\mu\text{mol m}^{-2}\text{s}^{-1}$. Unless stated otherwise, the plates used were square with a side of 10 cm. When it states that AVG was added to the medium, it was always to a concentration of 0.1 μM . Specifics on experimental materials such as chemicals and antibiotics can be found in Section 2.2. An overview of the plant material can be found in Table 2.1. The media compositions are listed in Table 2.2.

Medicago truncatula: Seeds of *M. truncatula* were grown as previously described (Sun et al., 2007). They were scarified with sand paper to help germination, and subsequently left in bleach for two-three min. After washing with distilled water (dH₂O) approximately seven times, the seeds were allowed to imbibe while rotating for 4 h at room temperature, and then placed on DWA media in a Petri dish upside down overnight at room temperature. When the seeds had germinated, they were grown vertically on plates with BNM+AVG.

Sesbania rostrata: *S. rostrata* seeds were germinated and grown as described previously (Capoen et al., 2009). They were scarified in H₂SO₄ for 2 h, rinsed with H₂O six times, and sterilised for one minute in bleach. Then they were rinsed with dH₂O six times and moved to DWA Petri dishes sealed with parafilm and left at 28° in the dark, until germination. When germinated, the seedlings were transferred to Jensen growth tubes filled with liquid sterile FP.

Parasponia andersonii* and *Trema tomentosa: Plantlets (clones) were prepared by Rik Op den Camp from the laboratory of René Geurts in Wageningen University. The plantlets were kept vertically in a growth chamber of 28°, 80% relative

humidity and 16 h of light and 8 h of dark. The plates used for growth were square with a side of 15 cm. For the Ca^{2+} injection experiments, wild type *Parasponia* were initially cultivated on WPM media, but several days before the experiments the plantlets were transferred to EKM+AVG plates, between filter paper.

A group of plantlets of *P. andersonii* and *T. tomentosa* were also transformed by Rik Op den Camp with the cameleon construct NupYC2.1. These were grown on EM media throughout the Ca^{2+} experiments. The EM media was supplemented with the antibiotic Cefotaxime, 5 $\mu\text{g}/\text{litre}$.

***Cercis siliquastrum*:** The *C. siliquastrum* seeds were scarified for 20 min in H_2SO_4 , washed with H_2O six times, and left for 10 min in bleach. After this the seeds were washed with dH_2O six times and then kept under water for two-three days until imbibition. In sterile conditions, the seeds were cut open and the embryo taken out and placed on plates with WPM medium. When the plants were large enough to handle they were moved to EKM+AVG plates between filter papers.

***Cytisus proliiferus*:** The seeds were scarified for 30-40 min in H_2SO_4 , then left for 15 min in bleach and rinsed six times in dH_2O . Subsequently the seeds were placed on DWA Petri dishes sealed with parafilm, and left in the dark overnight, at room temperature. When germinated they were moved to BNM+AVG plates between filter paper.

***Lupinus pilosus*:** Seeds of *L. pilosus* were scarified with sand paper and surface sterilised with bleach for 10 min and washed with dH_2O six times. The seeds were then placed on DWA Petri dishes sealed with parafilm in 20°C , 16 h light and 8 hours dark, until the seeds germinated. After germination, the seedlings were grown on BNM+AVG plates between filter papers. These plates were square with a side of approx. 30 cm.

***Chamaecrista fasciculata*:** *C. fasciculata* seeds were scarified in H_2SO_4 for 20 min, washed six times in dH_2O and then in domestic bleach for two min and washed six times in dH_2O . After that, they were allowed to imbibe rotating overnight. Then the seeds were placed on Petri dishes with DWA, upside down, in the dark at room temperature. After germination the seedlings were grow on BNM+AVG plates between filter papers.

***Acacia retinoides*:** The seeds were scarified with H_2SO_4 for 2 h and washed six times in dH_2O . They were then left to sterilise in bleach for 2 min, washed six times in dH_2O and then soaked in dH_2O until imbibition. After this, the seeds were moved to DWA Petri dishes sealed with parafilm, and placed upside down in the dark at room temperature overnight. After germination, the seedlings were grown on BNM+AVG plates between filter papers.

***Vicia hirsuta*:** Seeds of *V. hirsuta* were scarified with H₂SO₄ for 20 min and was then washed three times in dH₂O, treated with bleach for 10 min and washed thoroughly. The seeds were allowed to imbibe for 3 h and then moved to DWA Petri dishes, sealed with parafilm. The plates were placed upside down in the dark at 4°C for four days or more, until the seeds germinated.

2.1.2 Crude isolation of Nod factors

Liquid culture of bacteria: The bacteria were grown on TY agar for 2 days at 28°C. To 100 ml liquid Y medium was added: 0.75 ml Vits, NaGlu 2.5 ml NaGlu and 0.5 ml Suc, and 100 µl of the flavanoid Apigenin (1 µM stock). One colony of bacteria was added to 100 ml of the solution prepared and left in a shaker at 28°C for two days. Then the cultures were centrifuged and tested for Nod factor activity with the root hair deformation (RHD) assay, see below. When the results were positive a larger culture was prepared: 2.1 litre of the liquid solution was prepared in three conical flasks (700 ml per flask) of same proportions as above (Y, Vits, NaGlu, Suc, Apigenin). 10 ml of bacteria was added from the earlier 100 µl cultures, per flask. This was left in a shaker at 28°C for two days.

Purification: The 2.1 litre culture was centrifuged at 5°C, 9000 rpm, 60 min. Using a peristaltic pump, the supernatant was then passed through a C18:1 reverse phase column (Sep Pak, Waters). Lastly 4 ml of MeOH, 100%, solution was also passed through the pump, and collected in eppendorf tubes. These eppendorfs contain the Nod factors, and were kept at -20°C. Their activity was tested with the RHD assay.

2.1.3 Root hair deformation assay

Preparations: The *V. hirsuta* roots were allowed to germinated to around 1 cm long (5 seedlings for each treatment plus positive control and 5-10 for negative). Cover slides were attached to sterile microscope glass slides with four grease (Dow Corning GMBH, Wiesbaden, USA) dots, to provide two layers of glass.

Assay: The optical density (OD) of the bacterial cultures was checked (at 600 nm) and if needed they were diluted to an OD between 0.01 and 0.05 for the assay. Solutions were prepared with dilutions of the Nod factor isolation, plus positive and negative controls (TY medium). Petri dishes with two small wooden sticks in the bottom were prepared, and on top one set of microscope slides was placed, as prepared. The roots were left for 1 h in FP medium between the glass layers, and after that the FP liquid was removed and the prepared solutions replaced the FP. The petri dishes were left at room temperature in the dark overnight, and then checked for root hair deformation.

Evaluation: Signs of root hair deformation include branching, curling, zigzag, and tip swelling on root hair cells that have been exposed to Nod factors for the longest

time, i.e. in the second half of the root but not in the tip itself. The examination was done on an inverted microscope.

2.1.4 Calcium imaging

Calcium indicator gene cameleon

Cameleons are GFP-based Ca^{2+} indicators that code for a construct that binds Ca^{2+} and also contains fluorophores (Rudolf et al., 2003). Yellow cameleon YC2.1 has YFP (yellow fluorescent protein) and CFP (cyan fluorescent protein) fluorophores, a CaM moiety binding, and a link domain between YFP and the CaM binding called M13 (Figure 2.1). As Ca^{2+} is bound at the CaM domain, the construct goes through a conformational change that brings YFP and CFP closer together. Fluorescence resonance energy transfer (FRET) between the fluorophores, from CFP to YFP, then allows the intensity of YFP to increase, indicating a higher Ca^{2+} concentration.

When this construct is monitored *in vivo* by fluorescent microscopy, the emissions from both CFP and YFP are measured and their ratio is a non-ratiometric measurement of the Ca^{2+} concentration. Non-ratiometric means that the FRET ratio is an relative indication of Ca^{2+} levels, not an absolute concentration measurement. A *M. truncatula* root hair expressing the YC2.1 construct is shown in Figure 2.2. The signal, and therefore the Ca^{2+} levels, are highest at the root tip and in/around the nucleus.

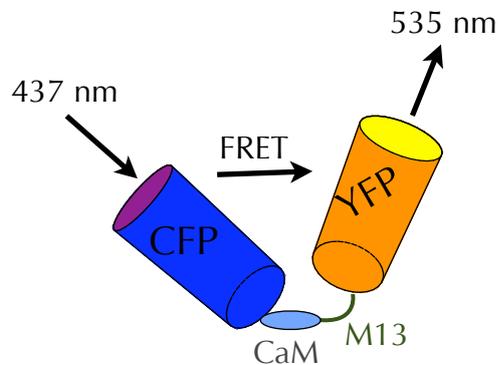


Figure 2.1: Fluorescence resonance energy transfer with the Ca^{2+} indicator cameleon YC2.1.

Cameleon YC2.1 imaging protocol

Seeds from a stable line of *M. truncatula* YC2.1 were germinated as outlined above, and 1-2 days after germination they can be used for Ca^{2+} imaging. The plant was placed on a microscope slide, with the root inside a small area restricted by vacuum

grease (Dow Corning GMBH, Wiesbaden, USA). The area inside the grease is filled up with liquid medium, varying depending on the size of the root.

The FRET measurements were taken using an epifluorescence microscope, with a Nikon 40x working lens for imaging. CFP is excited by 437 nm and a 11 nm bandpass with an Optoscan Monochromator (Cairn Research, Faversham, Kent, UK). An image splitter separated the fluorescence (dichroic mirror: 515 nm) and then an emission filter was used: 485 nm for CFP and 535 for YFP. The images were collected every 5 seconds and analysed with MetaFlour in Windows, and the data exported into R (R Development Core Team, 2008) for analysis. The YFP/CFP ratio is calculated, and the time series are detrended. For an overview of the process from raw data to presented data, see Appendix B, Figures 3-4. Amplitudes are disregarded since the YC2.1 used was very varying in signal strength, which impacted on noise levels and amplitudes. This can be seen by comparing Figures 3 and 4.

Cameleon NupYC2.1 imaging protocol

The nuclear-targeted cameleon NupYC2.1 was used for *Parasponia* and *Trema*. The plants were inspected using a GFP filter on a stereo microscope and roots expressing a strong signal were selected. These were cut off from the plant in sterile conditions and imaged as described in the previous section on Cameleon YC2.1.

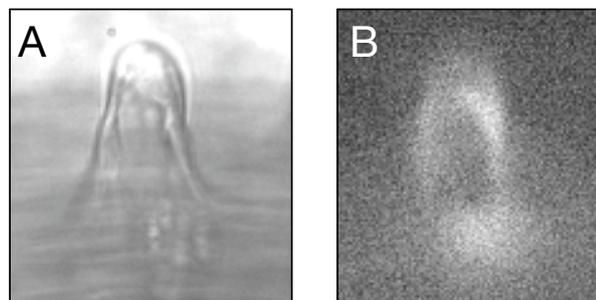


Figure 2.2: A: *M. truncatula* YC2.1 root hair. B: YC2.1 signal of same root hair, showing highest Ca^{2+} levels at the root tip and in/around nucleus.

Calcium indicator dyes

Ca^{2+} dyes are synthetic indicators for Ca^{2+} and are fluorescent polycarboxylate dyes. Here I will describe Oregon Green 488 BAPTA-1-dextran, 10,000 MW (OG). This indicator is based on BAPTA, which is similar in structure to EGTA, a well-known Ca^{2+} chelator. The important difference of BAPTA is that it is less pH-sensitive and binds Ca^{2+} faster, and it binds it in a 1:1 ratio (Putney, 2006). These dyes are often conjugated to dextran via a thiourea link, to avoid compartmentalization. The Ca^{2+}

affinity is therefore often reduced (Putney, 2006). When Ca^{2+} binds, a conformational change causes the chromophore of BAPTA to change emission. Another dye is often used together with Oregon Green: Texas Red-dextran, 10,000 MW (TR). This is a Ca^{2+} -insensitive dye which will not change emission when Ca^{2+} is present, and therefore the non-ratiometric measurement of Ca^{2+} levels is presented as a ratio of Oregon Green and Texas Red. This reduces the effects of for example cytoplasmic streaming. A figure showing a *M. truncatula* root hair injected with Oregon Green and Texas Red was shown in Chapter 1, Figure 1.3.

Microinjection protocol

This protocol follows Wais et al. (2000). The plants are prepared as described above, and growing root tips are used. The plant root is placed in the liquid-filled chamber that was described in the section on Cameleon YC2.1.

The Ca^{2+} dyes that were used were a mixture of Oregon Green (OG) and Texas Red (TR), and their starting concentration was 5 mM. The injection solution consists of 3.6 μl OG, 0.4 μl TR, and 1 μl of a dye salt (0.75 M KCl, 0.45 M HEPES, pH 7.0). This solution was centrifuged at 10,000 rpm for 2 min in a microcentrifuge. A glass needle is prepared using an electrode puller (model 773; Campden Instruments Ltd), and from borosilicate glass capillaries (1B120F-4; World precision Instruments Inc.). This needle is loaded with 0.5 μl of the dye solution with a Microloader pipette tip. The needle is then filled up with approximately 10 μl 1M KCl.

The glass slide with the plant is placed on the inverted microscope, and a reference electrode is placed in the liquid surrounding the plant. When the needle is in the liquid, a closed circuit is created which makes it possible to control the voltage in the needle (electro-manipulator, PatchMan NP2, Eppendorf). When the needle is in contact with a plant root hair cell, injection can start and the dyes are injected into the cell with a direct current set at 10 nA. After injection, only cells that stayed alive for 10 min were used for experiments (for example showed cytoplasmic streaming and correct localisation of the dyes in the cytoplasm). The dyes are of such a molecular weight that they also enter the nucleus, but no other cellular compartments.

The imaging was done with a Nikon TE2000U inverted microscope, and a Hamamatsu Photonics digital CC camera. OG is excited at 488 nm, giving a maximum 14-fold increase in intensity when binding to Ca^{2+} . Texas red is excited at 570 nm. OG emission was captured through a 545 nm emission filter, and TR with 620 nm. Every 5 seconds a data point was taken, one for OG and one for TR using an image splitter (Cairn Research, Faversham, Kent, UK) and a 1 second exposure. The images were analysed in MetaFluor in Windows, and the data exported into R (R Development Core Team, 2008) for analysis. The OG/TR ratio is calculated, and the time series are detrended. For an overview of the process from raw data to presented data, see Appendix B, Figures 3-4, which relate to cameleon YC2.1 but the process is the same for OG/TR. Amplitudes are disregarded since this is not a ratiometric method.

2.2 Experimental material

All plant materials are listed in Table 2.1. The media used for growing plants and bacteria are described in Table 2.2. All media, as well as NaGlu, vitamin solutions (Vits) and Succinate (Succ) solutions were prepared by the John Innes media kitchen. The bleach used for seed sterilization is commercially available (0.1% w/v sodium hypochlorite solution). The sulphuric acid referred to (H_2SO_4) is concentrated (98% v/v). L- α -(2-aminoethoxyvinyl) glycine (AVG) is an ethylene inhibitor, and is added to the plant media when the material is to be used for Ca^{2+} imaging. This increases the proportion of plants responding with Ca^{2+} spiking to Nod factors (Oldroyd et al., 2001a). The dyes for Ca^{2+} imaging, Oregon Green and Texas Red with a molecular weight of 10 000 kD, came from Molecular Probes (Eugene, USA). The antibiotic Cefotaxime was provided from the laboratory of René Geurts in Wageningen, the Netherlands. The flavonoid Apigenin used to induce the bacteria for Nod factor production came from Sigma-Aldrich Company Ltd (Poole, UK).

Table 2.1: List of plant material

Species	Seed source
<i>Medicago truncatula</i> A17	Oldroyd lab, JIC
<i>Sesbania rostrata</i> Brem	Oldroyd lab, JIC
<i>Chamaecrista fasciculata</i>	Easywildflowers Native Plant Nursery, USA
<i>Parasponia andersonii</i>	R. Geurts, Wageningen University
<i>Trema tomentosa</i>	R. Geurts, Wageningen University
<i>Lupinus pilosus</i>	Kew Gardens Millennium Seed Bank
<i>Cytisus proliferus</i>	M. León Barrios, Universidad de La Laguna
<i>Cercis siliquastrum</i>	R. Geurts, Wageningen University
<i>Acacia retinoides</i>	Kew Gardens Millennium Seed Bank

Table 2.2: Media composition

Medium	Abb.	Composition (per L)
EKM medium	EMK	1 ml 15 mM Fe-citrate, 1 ml Microelements stock (after autoclave), 0.25 g CaSO ₄ .2H ₂ O, 0.12 g KH ₂ PO ₄ , 0.36 g K ₂ HPO ₄ , 0.25 g MgSO ₄ , 0.1 g Na ₂ SO ₄ , pH 6.6, Daishin Agar 1% (w/v). After autoclave add 250ul 1M NH ₄ NO ₃ + microelements stock. Microelement stock mg/100ml: 100 MnSO ₄ , 25 ZnSO ₄ .7 H ₂ O, 25CuSO ₄ . 5 H ₂ O, 25 H ₃ BO ₃ , 0.5 NO ₂ MO ₄ .2H ₂ O (Becking, 1983)
Woody Plant medium	WPM	20 g sucrose, 2.4 g wpm salts (McCown's Woody Plant Basal Salt Mixture + Vitamins, from Duchefa) + vitamins, 8 g Daishin agar, pH 5.8, (Lloyd, 1980)
Emergence medium	EM	50 ml of 20x conc. SH-A, 50 ml of 20x conc. UM-C, 10 g of 1% sucrose, 3 ml of 1 M MES, pH 5.8, 0.9% agar, pH 5.8 (1M KOH), Stocks: 20x conc. SH-A: 50 g/l KNO ₃ , 8 g/l MgSO ₄ , 6 g/l NH ₄ H ₂ PO ₄ , 4 g/l CaCl ₂ .2H ₂ O, 0.2 g/l MnSO ₄ .4H ₂ O, 0.1 g/l H ₃ BO ₃ , 0.02 g/l ZnSO ₄ .7H ₂ O, 0.02 g/l KI, 0.004 g/l CuSO ₄ .5H ₂ O, 0.002 g/l NaMoO ₄ .2H ₂ O, 0.002 g/l CoCl ₂ .6H ₂ O, 0.3 g/l FeSO ₄ .7H ₂ O*, 0.4 g/l Na ₂ EDTA(TiTriplex III)*. * Dissolve in 100 ml H ₂ O 50°C before adding to the total solution. 20x conc. UM-C: 2 g/l Myoinositol, 0.1 g/l Nicotinic acid, 0.2 g/l Pyridoxine Hcl, 0.2 g/l Thiamine Hcl, 0.04 g/l Glycine.
Fahraeus plant medium	FP	0.1 g CaCl ₂ 2H ₂ O, 0.12 g MgSO ₄ 7H ₂ O, 0.1 g KHPO ₄ , 0.15 g Na ₂ HPO ₄ 12H ₂ O, 5 mg Ferric citrate, 2.86 mg H ₃ BO ₃ , 2.03 mg MnSO ₄ 4H ₂ O, 0.22 mg ZnSO ₄ 7H ₂ O, 0.08 mg CuSO ₄ 5H ₂ O, 0.08 mg H ₂ MoO ₄ H ₂ O, 0.5% Lab M agar (for solid), pH 6.3-6.7
Buffered nodulation medium	BNM	390 mg MES [2-(N-morpholino)-ethane sulphonic acid], 344 mg CaSO ₄ .2H ₂ O, 122 mg MgSO ₄ .7H ₂ O, 68 mg KH ₂ PO ₄ , 4.6 mg ZnSO ₄ 7H ₂ O, 3.1 mg H ₃ BO ₃ , 8.45 mg MnSO ₄ H ₂ O, 0.25 mg Na ₂ MoO ₄ 2H ₂ O, 0.016 mg CuSO ₄ 5H ₂ O, 0.025 mg CoCl ₂ 6H ₂ O, 18.65 mg Na ₂ EDTA, 13.9 mg Fe SO ₄ 7H ₂ O, for solid medium 11.5 g, pH 6.5
<i>Rhizobium</i> complete medium	TY	5 g Difco tryptone, 3 g Difco yeast extract, 1.325 g CaCl ₂ . Add 15 g agar for solid medium.
<i>Rhizobium</i> minimal medium	Y	0.1 mg MgSO ₄ 7H ₂ O, 0.22 CaCl ₂ 6H ₂ O, 0.22 K ₂ HPO ₄ , 0.02 FeCl ₃ , 1.1 g L-glutamate acid Na salt, 0.75 mg biotin, 0.75 mg aneurin HCl (Thiamine), 0.75 mg D,L-pantothenic acid Ca salt, 10 g agar, pH 6.8
Water agar	DWA	1.5 % Bacto agar, pH 5.7, adjusted with 1M KOH

2.3 Computational methods

2.3.1 Fourier Analysis

The frequency analysis in Chapter 3 consists of a comparison of Discrete Fourier Transforms (DFT) and Bayesian Spectrum Analysis (BSA).

The DFT was computed using the *fft* function in the programming language Octave (Eaton, 2002), where FFT stands for the Fast Fourier Transform, a more efficient way of calculating the DFT. The results are presented in a power spectrum to analyse which component carries the most power. This is also known as a periodogram,

$$S = \frac{1}{N} \sum_{k=1}^N |FFT(d)|^2, \quad (2.1)$$

i.e. the squared absolute value of the FFT of the data d , normalised over the number of data points N (Bracewell, 1978). For time series with strong trends, detrending was done before the FFT, using the moving average method (Brockwell and Davis, 2002).

Results from other frequency detection methods are also used in Chapter 3, for example the multitaper method (MTM), short-time Fourier transforms (STFT) and wavelet analysis. For the MTM, only the MTM spectrum is presented, but it should be noted that the Singular Spectrum Analysis - MultiTaper Method (SSA-MTM) toolkit provides additional features such as significance levels of the frequencies, relative to the estimated noise levels (Ghil et al., 2002). The STFT power spectrums were computed with the *specgram* function in Octave. The wavelet results were computed using software provided by Dr. C. Torrence and Dr. G. Compo, available online (<http://atoc.colorado.edu/research/wavelets/>). This wavelet software also provides additional tools such as significance levels (Torrence and Compo, 1998).

2.3.2 Bayesian Spectrum Analysis

The initial BSA package was coded in Octave by R. Morris, and was then extended and optimised by the author. Octave is freely available and compatible with the widely used Matlab.

An outline of the BSA process for a single time series is shown in Figure 2.3. The first step is to specify the appropriate model functions and the frequency domain of interest. This domain is then sampled with a chosen interval, and the posterior probability is computed at each frequency.

To account for background trends in the data, the modelled signal can have added trend functions $T(i)$ at each discrete time point i . We used the Legendre polynomials, see Appendix A.4 for more information. These are scaled to be orthogonal on the time interval of each time series, and expanded to an order which sufficiently represents the trends in the data. To avoid bias in the analysis by choosing the expansion order manually, model comparison was used to decide on the optimal number of trend

function.

To compare models, priors are assigned to each of the model parameters as normalised Gaussians. The parameters as well as their prior variances are then integrated out. It is important to take into account all parameters for the models being compared, since different models can have different number of parameters and must be normalised properly Bretthorst (1988). The model ratio in Chapter 3, Eq. 8, underlies the subsequent calculation, where j is a model including n background trends and k is a model with $n+1$ trends. When the ratio is below 1, the code stops automatically since enough background trends has been included.

Nested sampling

To calculate the evidence, the normalising component in Chapter 3, Eq. 3.7, an existing code using Nested Sampling was incorporated into the BSA code. Nested Sampling is a Bayesian variant of the Markov Chain Monte Carlo (MCMC) algorithm that concentrates its sampling efforts in high likelihood regions of parameter space (Sivia and Skilling, 2006). The algorithm focuses on the computation of the evidence, while at the same time generating samples from the posterior distribution. By transforming the problem to likelihood space, high-dimensional integration can be reduced to a sorting task. Random samples are taken from the prior and by rejecting the point with the worst likelihood, the algorithm iteratively contracts the spread of samples around high likelihood regions of parameter space. One of the remaining samples is chosen to generate a new sample by taking MCMC steps around it in search of a higher likelihood value. Then the new samples are again sorted and the worst point rejected. This process is iterated until a given number of posterior samples have been generated.

Downhill-Simplex

Since the values of angular frequency, ω , are sampled over the frequency domain of interest with a chosen interval, the most probable frequency from this set may have a close neighbour with even higher probability but which fell between sampling points. To avoid this, and confirm that the maximum peak has been found, the Nelder-Mead optimisation technique also known as the Downhill-Simplex (Press, 2002) is used in the BSA code. Subsequently, the area surrounding this peak is finely sampled, to achieve a better representation of the posterior probability distribution of ω .

BaSAR - BSA in R

The main drawback of Octave is a long computational time and therefore the whole BSA package in Octave was converted into an R package (R Development Core Team, 2008) by Matthew Hartley and the author. The package is named BaSAR, and is available online at CRAN - The Comprehensive R Archive Network: <http://cran.r-project.org/web/packages/BaSAR>.

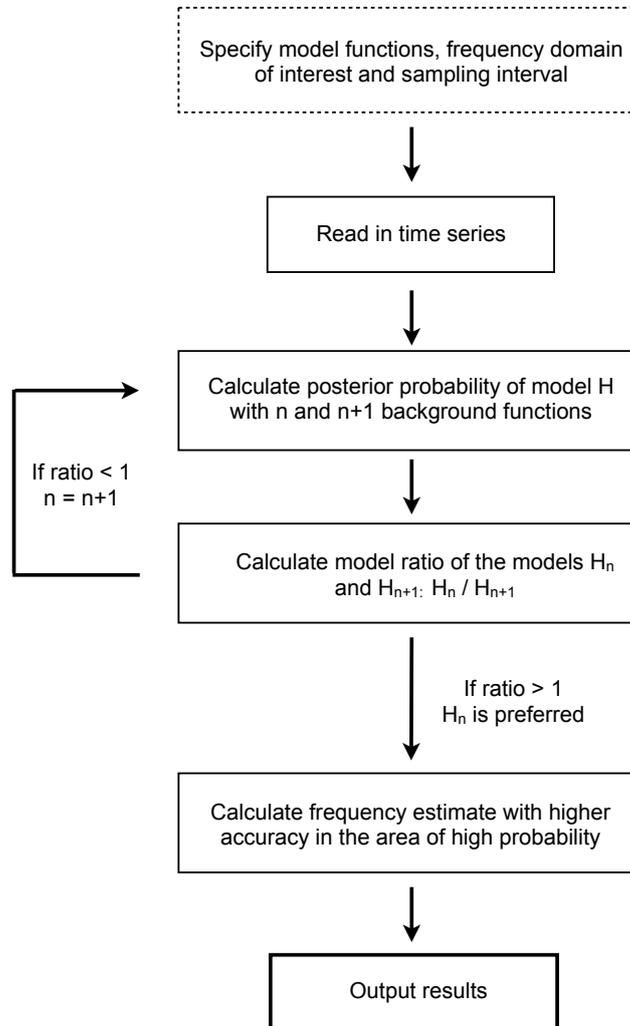


Figure 2.3: Flowchart of the automated model development procedure for Bayesian Spectrum Analysis. It is important to note that the proposed method for detecting the best number of background functions may give rise to local rather than global solutions for complex background trends and/or poor choices of background basis functions.

Table 2.3 lists the key functions. `BaSAR.post` returns a normalized posterior probability distribution over the chosen range of ω . This is invoked in the manner

```
BaSAR.post(data, start, stop, nsamples, nbackg, interval)
```

where `data` is the time series as a 1D vector, `start-stop` is the range of the period that is of interest (in seconds), `nsamples` is the number of samples that will be calculated from the posterior, and `interval` is the time between data points (in seconds). `BaSAR.nest` samples the posterior and calculates the evidence using nested sampling. Direct comparison of evidences makes it possible to evaluate models.

Table 2.3: Key functions in the BSAR package

Name	Description
<code>BaSAR.post</code>	normalized posterior probability distribution
<code>BaSAR.nest</code>	posterior and evidence using nested sampling
<code>BaSAR.fine</code>	high-resolution posterior of strongest frequency
<code>BaSAR.modelratio</code>	model comparison for background trends
<code>BaSAR.auto</code>	automated <code>BaSAR.modelratio</code>
<code>BaSAR.local</code>	2D posterior over time and ω by windowing

`BaSAR.modelratio` is an alternative model comparison method that uses model ratios to allow the user to compare two models with different background functions. This procedure has been automated in the function `BaSAR.auto`. For time series in which the dominant frequency changes over time, `BaSAR.local` can be used to calculate the local frequency by windowing.

2.3.3 Modelling calcium oscillations

The Ordinary Differential Equation (ODE) model used in Chapter 4 was originally set up by Saul Hazledine, and the initial buffer work was done by Wojciech Kozlowski. The ODEs were implemented in XPP (Ermentrout, 2002). The fourth-order Runge-Kutta method of solving ODEs was used, which solves for the next point, x_{n+1} , by calculating four numbers,

$$\begin{aligned}k_1 &= f(x_n)\Delta t \\k_2 &= f\left(x_n + \frac{1}{2}k_1\right)\Delta t \\k_3 &= f\left(x_n + \frac{1}{2}k_2\right)\Delta t \\k_4 &= f(x_n + k_3)\Delta t,\end{aligned}$$

where Δt is the stepsize in time between points. Then, the next point is given by

$$x_{n+1} = x_n + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4).$$

This method is very commonly used, and delivers good results without a extremely small stepsize, thus avoiding both computer cost and rounding off errors (Strogatz, 1994).

The bifurcation analysis was done in the bifurcation program AUTO, which is built into XPP (Ermentrout, 2002). To start the bifurcation analysis, the system needs to be in a stable steady state. This was achieved by using the value 0.1 for the parameter E_{ps} in combination with the parameters listed in Chapter 4, Table 4.1. This results in a steady state of the ODE model.

Chapter 3

Employing Bayesian Spectrum Analysis to detect periodicity

3.1 Introduction

The two symbiotic partners of legumes, rhizobia and arbuscular mycorrhizal fungi, both release a signalling compound into the soil which diffuses to the plant root cells. This triggers a signal cascade inside the cells, leading to either nodulation or mycorrhizal colonisation. Dissection of these signalling pathways has shown that they share several genetic components. This is called the common SYM pathway, see Chapter 1, Figure 1.2 for an overview. The common SYM pathway also specifies nuclear-associated Ca^{2+} oscillations, and several of the shared genetic elements are involved in the production and decoding of these oscillations, through, for example, cation channels, nuclear pores, and a Ca^{2+} and calmodulin protein kinase. Downstream of the common SYM pathway the signal paths diverge again, and it is unclear how the plant root cell distinguishes between the signals activated by the bacteria and the fungi.

One hypothesis is that the Ca^{2+} oscillations are conveying specificity in the system, allowing downstream components to respond differently to the two signals (Oldroyd and Downie, 2006). This is supported by evidence from other signalling systems, where Ca^{2+} oscillations have been shown to be more than a switch, and carry different messages in different circumstances (Dolmetsch et al., 1998; Allen et al., 2001; Evans et al., 2001). However, the strict dependence on specific patterns of Ca^{2+} oscillations for specific downstream events in the common SYM pathway has not been shown.

Ca^{2+} signalling has been studied in more detail in animal research, where frequency encoding plays a central role. A classic example of this is the case of neurons, where Ca^{2+} ions are involved in propagating the action potentials. The protein that decode pulses of Ca^{2+} at the cell membranes has been shown to distinguish different frequencies (Lisman et al., 2002), and the kinetics of these interactions have also recently been modelled (Pepke et al., 2010). The decoding protein in that study, CaMKII, is closely related to CCaMK, the hypothesised decoder in the common SYM pathway. The

structure of plant CCaMK is even more complex regarding its ability to bind Ca^{2+} , but its response to different Ca^{2+} pulses has not been studied. In plants, the most well-known system of Ca^{2+} signalling involve stomatal guard cells, where a defined range of frequencies in cytosolic Ca^{2+} oscillations deliver specific downstream responses (Allen et al., 2001). Considering such systems, the hypothesis of the Ca^{2+} oscillations in the root cells conveying specificity is plausible.

Could frequency encoding play a role in these symbiotic Ca^{2+} signals? If so, a significant frequency difference is to be expected. This is, however, not straightforward to address since the time series data contain background trends, sometimes have high noise levels, and change their pattern over time. In the following sections I will compare two approaches for frequency detection, Fourier Analysis and Bayesian Spectrum Analysis, and make further developments of the Bayesian method to aid in Ca^{2+} spiking analysis. These results and associated code have been published in Granqvist et al. (2011) and Granqvist et al. (2012a), respectively.

3.1.1 Fourier Analysis

Jean Joseph Fourier (1768-1830) developed a revolutionising theory which is now a widely used method in many scientific fields. He showed that it is possible to represent any integrable function as an infinite sum of sines and cosines. For a function $f(x)$ with period $P = 2L$, this can be written as

$$f(x) = a_0 + \sum_{n=1}^{\infty} \left(a_n \cos \frac{n\pi}{L}x + b_n \sin \frac{n\pi}{L}x \right). \quad (3.1)$$

This expression is called the Fourier series, and the constants a and b are the Fourier coefficients, where a is real and b is imaginary.

Fourier theory has given rise to a wide range of diverse developments and far-reaching applications, demonstrating the theory's undisputed importance and impact. A conventional usage for Fourier analysis is in frequency detection. The Fourier Transform (FT) can reveal the underlying frequency components that are present in a signal (Bracewell, 1978). An example of how several individual signals can make up the overall signal is shown in Figure 3.1. However, it is known that the FT works optimally for frequency detection only for uniformly sampled, long, stationary time series. Furthermore, common procedures of pre-processing the data can cause problems. Time series can contain low frequency background fluctuations or drift that are unrelated to the signal of interest. For the FT, it is then necessary to remove the trends using detrending techniques. As has been shown previously, this detrending leads to convolution of the signal that can both remove evidence for periodicity and add false patterns (Jaynes and Bretthorst, 2003). Another known problem is aliasing. If a signal containing high frequencies is recorded with a low sampling rate, peaks of high frequencies can fold back into the frequency spectrum, appearing as low frequencies (Bracewell, 1978). The Gibbs phenomenon provides another example where spurious peaks appear in a FT. It

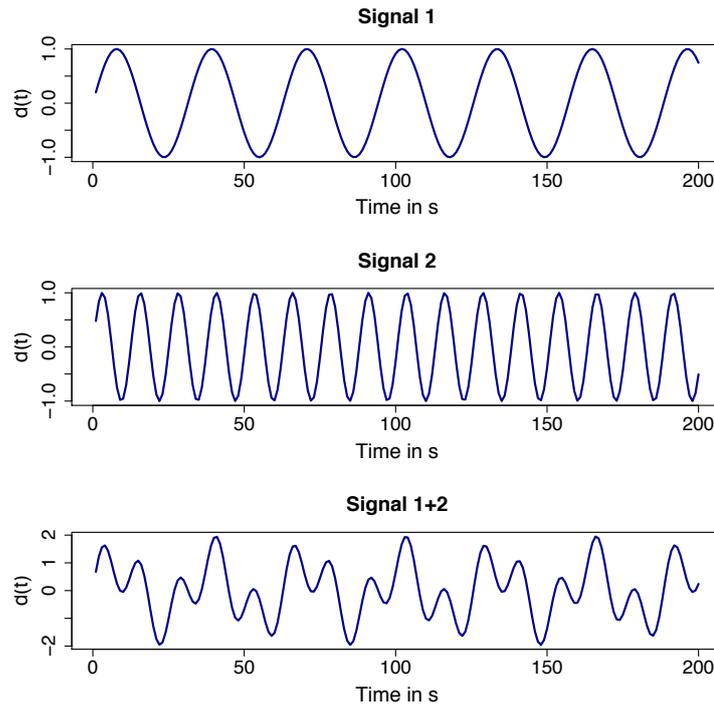


Figure 3.1: An example of Fourier theory: Periodic oscillations can be added together to form the final signal. Fourier Transformation can identify the two basic components.

occurs at points of discontinuity in a periodic function, and results in so-called ringing artifacts around the “true” frequency peak (Gibbs, 1899). As for the accuracy of the frequency estimate, no direct information of this is given by the output from a FT, since the sharpness of the peaks depends on a combination of factors such as sampling intervals, noise levels and the length of the time series. For further details, see the extensive FT literature (e.g. Bracewell (1978); Kammler (2007)).

Wavelet transforms (Mallat, 1999; Heil and Walnut, 2006; Daubechies, 1992; Prasad and Iyengar, 1997; Petrosian and Meyer, 2001) offer an attractive alternative to Fourier Transforms. The main difference is that they are localised in both the time and frequency domain. This property makes wavelets better adapted to problems with truncated data. Wavelets have found wide-ranging applications and have proven to be particularly powerful for image processing and data compression (Davis et al., 2007; Ricke et al., 1998; Lucier et al., 1994).

3.1.2 Bayesian Spectrum Analysis

Bayesian inference provides another approach for analysing data (for an introduction to Bayesian analysis, see Sivia and Skilling (2006)). It addresses additional aspects of the problem, such as the inherent uncertainty of the analysis and the effects of noise. Using this framework (Jaynes and Bretthorst, 2003), the method of Bayesian

Spectrum Analysis (BSA) was developed by Bretthorst (Bretthorst, 1988) and applied to Nuclear Magnetic Resonance (NMR) signals and parameter estimation with great success (Bretthorst et al., 1989; Neil and Bretthorst, 1993).

There are several advantages of the Bayesian approach, including an inherent mechanism for estimating the accuracy of the result and all parameters, as well as the ability to compare different hypotheses directly. Focus is shifted to the question of interest by integrating out other parameters and noise levels. Initial knowledge of the system can be incorporated in the analysis and expressed in the prior probability distributions. The Bayesian approach to time series analysis has proven its value in fields such as NMR and ion cyclotron resonance analysis (e.g. Meier and Marshall (1990) and Chylla and Markley (1993)).

Bayesian inference

Data is rarely available in sufficient quantity and quality to allow for exact scientific deduction. Instead we are forced to infer models from incomplete knowledge. Bayesian inference is based on Bayes' Rule, and is a method of assigning probabilities based on the current state of knowledge, allowing for subsequent re-evaluation as new data becomes available. Thomas Bayes (1702-1761) first described this expression, and it can be derived directly from conditional probabilities. Let us call two separate events A and B . The probability of A occurring given that B has occurred is the joint probability of A and B (Figure 3.2) divided by the probability of B , $P(A|B) = P(A \cap B)/P(B)$. The reverse is also true, $P(B|A) = P(A \cap B)/P(A)$, leading to $P(A|B)P(B) = P(B|A)P(A)$. Re-writing this expression results in $P(A|B) = P(B|A)P(A)/P(B)$. This is Bayes' rule.

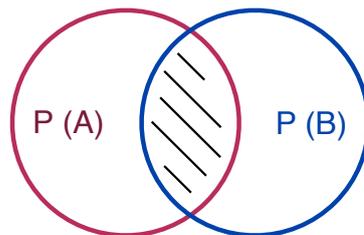


Figure 3.2: Joint probability of event A and B , $P(A \cap B)$, is the marked overlap of the two circles.

Bayes' rule allows us to evaluate a hypothesis, H , in light of some data, D , and some prior information, I . The goal is to determine $P(H|D, I)$, the posterior probability distribution of the hypothesis, given the data and the prior information. The posterior expresses the probability that the hypothesis is correct. With Bayes' rule, the posterior

can be expressed in terms of other probabilities as

$$P(H|D, I) = \frac{P(D|H, I) P(H|I)}{P(D|I)}, \quad (3.2)$$

where $P(D|H, I)$ is the probability of observing the data given the hypothesis and the prior, $P(H|I)$ is the prior probability of the hypothesis, and $P(D|I)$ is the probability of the data given the prior information. When the hypothesis is the variable and the data is held constant, $P(D|H, I)$ is called the likelihood function, and when the hypothesis is constant it is called the probability of obtaining a specified outcome (data). When evaluating only one hypothesis, $P(D|I)$ is a normalising constant, but when investigating more than one hypothesis this term plays a key role and is called the evidence (MacKay, 2003).

BSA theory

The presentation in this section follows closely that of Bretthorst (Bretthorst, 1988). The aim is to infer the most probable frequency (or frequencies), $\boldsymbol{\omega}$, from the given data. An overview of the symbols used can be found in Table 3.1. First, a model is defined (the hypothesis H) for the observed data, parameterised by angular frequency, $\boldsymbol{\omega}$, and amplitudes, \mathbf{c} , and then Bayes' rule is used to compute the posterior probability of the parameters, $P(\boldsymbol{\omega}, \mathbf{c}|D, H, I)$. By assigning priors to the model parameters \mathbf{c} and integrating over these, we arrive at the posterior probability for the parameter of interest, $\boldsymbol{\omega}$, $P(\boldsymbol{\omega}|D, H, I) = \int d\mathbf{c} P(\boldsymbol{\omega}, \mathbf{c}|D, H, I)$. This is referred to as the marginal posterior probability of $\boldsymbol{\omega}$. Note that $\boldsymbol{\omega}$ is an r -tuple, $\{\omega_1, \omega_2, \dots, \omega_r\}$, with as many elements as there are distinct frequencies in the data.

A general model for observed data that is used here is sampled at N discrete time points, $D = \{d(t_1), \dots, d(t_N)\}$, includes the signal of interest, $s(t_i)$, a possible background function, $g(t_i)$, and the noise present in the system, $e(t_i)$,

$$d(t_i) = s(t_i) + g(t_i) + e(t_i). \quad (3.3)$$

The signal function will usually be unknown and may be complicated, but can be approximated by a linear combination of m_s model functions, ψ_j , that we parameterise by the quantity of interest, $\boldsymbol{\omega}$:

$$s(t_i) = \sum_{j=1}^{m_s} a_j \psi_j(\boldsymbol{\omega}, t_i), \quad (3.4)$$

in which $\mathbf{a} = \{a_1, \dots, a_{m_s}\}$ are the expansion coefficients.

Similarly, the background function, $g(t_i)$, can be approximated by a set of m_g functions, ζ_j , that are independent of $\boldsymbol{\omega}$,

$$g(t_i) = \sum_{j=1}^{m_g} b_j \zeta_j(t_i), \quad (3.5)$$

where $\mathbf{b} = \{b_1, \dots, b_{m_g}\}$ are the background model function expansion coefficients.

Since \mathbf{a} and \mathbf{b} are not the main focus of the analysis, we will aim to integrate them out of the equations by marginalisation. Parameters that are treated in this manner as often referred to as nuisance parameters, which we denote here by $\mathbf{c} = \{a_1, \dots, a_{m_s}, b_1, \dots, b_{m_g}\}$. Although the signal function depends on $\boldsymbol{\omega}$, whereas the background function does not, for notational purposes we introduce the set of model functions, ϕ_i , which consists of both ψ_i and ζ_i . This allows us to condense the model equation into

$$f(t_i) = s(t_i) + g(t_i) = \sum_{j=1}^m c_j \phi_j(\boldsymbol{\omega}, t_i), \quad (3.6)$$

such that now $d(t_i) = f(t_i) + e(t_i)$ and m is now the total number of model functions, $m_s + m_g$. The model functions will typically not be orthogonal functions over the time series domain. This, however, can be achieved by Cholesky decomposition. In all subsequent calculations an orthogonal basis is used.

From Bayes' rule, the joint probability distribution of the model for the parameters $\boldsymbol{\omega}$ and \mathbf{c} is

$$P(\boldsymbol{\omega}, \mathbf{c} | D, H, I) = \frac{P(D | \boldsymbol{\omega}, \mathbf{c}, H, I) P(\boldsymbol{\omega}, \mathbf{c} | H, I)}{P(D | H, I)}. \quad (3.7)$$

The likelihood function, $P(D | \boldsymbol{\omega}, \mathbf{c}, H, I)$, is calculated by comparing data produced by the model signal, equation (3.6), to real experimental data. If the model perfectly captures the signal, the difference between the model data and the real data is simply the noise in the system. The model of the data in equation (3.3) includes noise, $e(t_i)$, which we assume to be time independent in the further developments. The true noise level is unknown, but for a given noise power, σ^2 , the principle of maximum entropy leads to the use of a normal distribution,

$$P(e | \sigma, I) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{e^2}{2\sigma^2}\right\}. \quad (3.8)$$

A noise model of this form ensures that the accuracy of the results is maximally conservative for a given noise power. We will later integrate over all possible noise levels to remove the dependence on σ . With the described signal and this noise model, the likelihood was calculated by Bretthorst (Bretthorst, 1988) to be

$$P(D | \boldsymbol{\omega}, \sigma, \mathbf{c}, H, I) \propto \sigma^{-N} \exp\left(\frac{-NQ}{2\sigma^2}\right), \quad (3.9)$$

where N is the number of data points, and

$$Q = \bar{d}^2 - \frac{2}{N} \sum_{j=1}^m \sum_{i=1}^N c_j d_i \phi_j(\boldsymbol{\omega}, t_i) + \frac{1}{N} \sum_{j=1}^m \sum_{k=1}^m \Phi_{jk} c_j c_k, \quad (3.10)$$

where \bar{d}^2 is the mean-square of the data, $\bar{d}^2 = (\sum_{i=1}^N d_i^2)/N$, and Φ_{jk} is the matrix of the model functions, $\Phi_{jk} = \sum_{i=1}^N \phi_j(\boldsymbol{\omega}, t_i) \phi_k(\boldsymbol{\omega}, t_i)$.

The goal of the analysis is to compute the posterior probability for frequencies in the data, i.e. to go from the joint probability distribution in equation (3.7) to a posterior probability of $\boldsymbol{\omega}$, independent of the other parameters. By integrating over all possible values of the parameters σ and \mathbf{c} , the remainder is the marginal posterior of the parameters of interest, $\boldsymbol{\omega}=\{\omega_1, \omega_2, \dots, \omega_r\}$. This is an essential advantage of the Bayesian framework, allowing the analysis to focus on estimating the parameters of interest, regardless of the values of the others. If necessary, the other parameters can be estimated at a later point.

To integrate over the σ and \mathbf{c} values, priors must first be assigned to them. Uniform priors are chosen for \mathbf{c} and $\boldsymbol{\omega}$, representing complete lack of knowledge. Since σ is continuous and must be positive, a Jeffreys prior is appropriate, $P(\sigma|I) = 1/\sigma$. Both the uniform distribution over continuous variables and Jeffreys prior are known as improper priors if bounds are not specified as they cannot be normalised. For more information on prior assignment see Bretthorst (1988); Jaynes and Bretthorst (2003).

Using the general model, equation (3.6), assigning the priors, calculating the likelihood function, equation (3.9), and integrating out the amplitudes and noise parameters, the posterior probability distribution of $\boldsymbol{\omega}$ is proportional to

$$P(\boldsymbol{\omega}|D, H, I) \propto \left[1 - \frac{m\bar{h}^2}{Nd^2} \right]^{\frac{m-N}{2}}, \quad (3.11)$$

where h is the projection of the data onto the orthonormal model functions, $h_j = \sum_{i=1}^N d_i \phi_j(\boldsymbol{\omega}, t_i)$, and \bar{h}^2 is the mean-square of the h_j , $\bar{h}^2 = (\sum_{j=1}^m h_j^2)/m$, (Bretthorst, 1988). This expression of the posterior allows us to identify the strongest frequencies present in the data. For a good model, there will be a high probability peak in the posterior distribution at that $\boldsymbol{\omega} = \{\omega_1, \omega_2, \dots, \omega_r\}$.

3.1.3 Model comparison

After evaluating the probability of parameters in light of a certain hypothesis, it is important to question the validity of that hypothesis. Thus, the next step in Bayesian inference is to compare the probability of different hypotheses. The hypothesis is now a particular model of the signal, H_i , out of a set of M models $\{H_1, \dots, H_M\}$, and using Bayes' Rule, the posterior probability of this model is

$$P(H_i|D, I) = \frac{P(H_i|I)P(D|H_i, I)}{P(D|I)}. \quad (3.12)$$

Then two different models, H_i and H_j , can be compared by taking their ratios,

$$\frac{P(H_i|D, I)}{P(H_j|D, I)} = \frac{P(H_i|I)P(D|H_i, I)}{P(H_j|I)P(D|H_j, I)}. \quad (3.13)$$

The probability of the data given our prior information, $P(D|H_i, I)$, which was a normalisation constant in equation (3.7), will now vary between models, and is called the

Table 3.1: BSA symbol definitions

Symbol	Description
s	signal function
g	background function
e	experimental noise
f	modelled signal, $s + g$
\mathbf{a}	signal model parameters, $\{a_1, \dots, a_{m_s}\}$
\mathbf{b}	background model parameters, $\{b_1, \dots, b_{m_g}\}$
\mathbf{c}	model parameters, $\{a_1, \dots, a_{m_s}, b_1, \dots, b_{m_g}\}$
N	number of data points
D	the data vector, $1 \leq d(t_i) \leq N$
t	vector of time points, $1 \leq t_i \leq N$
σ	noise standard deviation
σ^2	noise variance
ω	the angular frequency
$\boldsymbol{\omega}$	the vector of angular frequencies, $\{\omega_1, \omega_2, \dots, \omega_r\}$
$\{\hat{\omega}\}$	ω value that maximizes the likelihood function
$\bar{\omega}^2$	mean-squared ω at $\{\hat{\omega}\}$
ψ	set of signal model functions, dependent on ω
ζ	set of background model functions
ϕ	complete set of model functions, $\psi + \zeta$
Φ_{jk}	matrix of model functions ϕ , $1 \leq j, k \leq m$
m	total number of model functions, $m_s + m_g$
h_j	projection of the data onto the orthonormal model functions
\bar{h}^2	mean-square of the observed projections
$\bar{h}^2(\{\hat{\omega}\})$	mean-square h evaluated at $\hat{\omega}$
\bar{d}^2	observed mean-squared data value

evidence. It evaluates the fit of the data to the model, and for an equal fit it makes the model that include more parameters less likely. Each additional model parameter should be followed by a significant increase in fit, otherwise the simpler model is preferred. Thus, Bayesian model comparison naturally follows the principle of Occam's razor (MacKay, 2003; Toni and Stumpf, 2010).

Model development

It will often not be obvious which function to choose to model trends in the data, so an approach using basis functions and expanding these to different orders will be of advantage, as in equation (3.5). Each expansion represents a different model, H_i , and these can be compared using inference techniques. Likewise, different functions for capturing the signals in the data and modelling a different number of signals correspond to different models for data. Following Jaynes and Bretthorst (2003); MacKay (2003); Sivia and Skilling (2006), we use the posterior ratio to evaluate different models. This model ratio is used to determine the number of background model functions for each time series. Legendre polynomials are employed as background basis functions (ζ_j in equation (3.5)). Legendre functions are convenient as they form a basis that can be scaled to be orthogonal over the time domain and offer a level of detail that increases with expansion order. The software, however, will attempt to orthogonalise any given set of functions over the range determined by the data by Cholesky decomposition, so other functions can be employed.

The posterior probability ratio is calculated between model H_n and H_{n+1} , where H_n is a model including n background functions. To obtain the model ratio, priors are assigned to each of the models and their likelihood functions are calculated. Assigning equal prior probability to all models reduces this to the ratio of evidences. To compute the evidences we need to integrate the likelihood, $P(D|\boldsymbol{\omega}, \sigma, \mathbf{c}, H_n, I)$ from equation (3.9), over $\boldsymbol{\omega}$, σ , and \mathbf{c} for each model H_n . By assigning proper normalised priors to all model parameters it is possible to integrate over them around the maximum likelihood estimate. Following Bretthorst's derivation for location parameters (Bretthorst, 1988), we assign Gaussian priors to the amplitudes with hyperparameters for the variances. Since the variances are scale parameters, they are subsequently assigned Jeffreys priors with an upper and lower bound. This allows us to normalise them and integrate, leaving the defined bounds as parameters in the final equation. For models with the same bounds these terms cancel out in the model ratio. The evidence for a given model, H_n , was calculated by Bretthorst (Bretthorst, 1988) to be

$$P(D|H_n, I) = \frac{\Gamma(m/2)}{2 \log(R_\delta)} \left[\frac{m\bar{h}^2(\{\hat{\boldsymbol{\omega}}\})}{2} \right]^{-\frac{m}{2}} \frac{\Gamma(r/2)}{2 \log(R_\gamma)} \left[\frac{r\bar{\boldsymbol{\omega}}^2}{2} \right]^{-\frac{r}{2}} v_1^{-\frac{1}{2}} \dots v_r^{-\frac{1}{2}} \\ \times \frac{\Gamma(|N - m - r|/2)}{2 \log(R_\sigma)} \left[\frac{N\bar{d}^2 - m\bar{h}^2(\{\hat{\boldsymbol{\omega}}\})}{2} \right]^{\frac{m+r-N}{2}}, \quad (3.14)$$

where δ , γ and σ are the prior variances for amplitudes, frequencies and noise, respectively, R_δ , R_γ and R_σ are the ratios of the integral bounds for these variances, $\overline{h^2}(\{\hat{\omega}\})$ is the mean-square projection of the data onto the orthonormal model functions at the maximum likelihood point for model H_n , $\overline{\omega^2}$ is the mean-square of the ω value that maximises the likelihood, $\overline{\omega^2} = (\sum_{k=1}^r \hat{\omega}_k^2)/r$, and r is the number of ω parameters, $\omega = \{\omega_1, \dots, \omega_r\}$. The Jacobian $v_1^{-\frac{1}{2}} \dots v_r^{-\frac{1}{2}}$ is obtained by orthogonalising the Taylor-expansion of $\overline{h^2}$ around the maximum-likelihood point, $\hat{\omega}$. See Bretthorst for further details (Bretthorst, 1988). For cases in which the number of frequencies in the data exceeds the dimension of ω , for instance, multiple frequency data with a single frequency model, the above approximation for the evidence is ill-suited as the posterior will cease to be unimodal. For such scenarios, either multiple expansions or MCMC offer attractive solutions to marginalisation. For comparison we have included results from Nested Sampling (Sivia and Skilling, 2006; Skilling, 2006) as a means to perform the integration and compute the evidence. Nested Sampling is a variant of MCMC that employs a likelihood based sorting of sample points to efficiently guide the search strategy of the posterior distribution (Sivia and Skilling, 2006; Skilling, 2006).

When the model ratio, H_n/H_{n+1} becomes greater than 1, the simpler model, H_n , is favoured over H_{n+1} (MacKay, 2003). Adding more background functions than are justified by the data (based on the posterior model ratio) may lead to a lower probability for the frequency and in some cases possibly a location shift.

This model development approach used for the background functions above can also be used to decide on the number of underlying frequencies in the data. The model ratios of a time series containing one frequency (case A) and a time series containing two (case B) are presented in Table 3.2, analysed with both a one- and two-frequency model. The results show, as expected, a preference for the one-frequency model in case A, and for the two-frequency model in case B.

Table 3.2: Model ratio of number of frequencies, N_ω , in the data. Case A has only one frequency, and case B has two. The model ratio of a model including one frequency, $H_{1\omega}$, and a model with two, $H_{2\omega}$, is well above 1 for case A, showing that the simpler model should be chosen. For case B, the first ratio is close to 0, showing that the more complex model with two frequencies is far more likely. Values below 10^{-8} are listed as 0. The second ratio is above 1, preferring the model with two frequencies over the one with three.

Case	N_ω	Models	Model ratio
A	1	$H_{1\omega}/H_{2\omega}$	66936
B	2	$H_{1\omega}/H_{2\omega}$	0
B	2	$H_{2\omega}/H_{3\omega}$	686

Note that that the proposed method stops once the current best model has been found but is not guaranteed to find the global maximum from a predefined set of models. The procedure is thus part of model development rather than model selection. If the set of hypotheses are known in advance then the posterior ratios over the full set should be used to find the best model.

3.2 Results

3.2.1 Simulated test cases

The BSA approach is first tested on simulated data. In these tests, the goal was to recover known input parameters from the simulated data, to validate the BSA approach. Sines and cosines are the model functions (ψ_j in equation (3.4)). For comparison, Discrete Fourier Transforms were computed using Fast Fourier Transforms (FFT) (Cooley and Tukey, 1965). In the test cases, we varied key parameters such as noise levels, trace length, sampling intervals, amplitudes, frequencies, background trends and shape of oscillations.

Representative cases of noise levels and background trends are shown in Table 3.4, including FFT results on the same data set. A key observation is that the Bayesian approach extracts the correct answer from the data with extremely high precision. BSA also computes the signal-to-noise ratios which is a useful indication of how much of the data can not be accounted for in the model. Furthermore, the amplitudes do not impact the BSA results since they are integrated out.

BSA has a clear advantage over FFT when the data is non-uniformly sampled. FFT requires uniform sampling, whilst BSA is less stringent and delivers the correct result with high precision. Bretthorst also noted that non-uniformly sampled data removes aliases from the frequency domain, another significant advantage (Bretthorst, 1988). Five further distinct cases emerged from the tests in which BSA delivers superior results to FFT: time series which have background trends, few data points, high noise levels, multiple frequencies, and non-harmonic oscillations.

Background trends

One test consisted of a time series with a strong background trend, see Figure 3.3. In Table 3.3 the model ratios for different numbers of background functions are shown. The ratio is initially well below 1, but the ratio of models with expansion orders of two and three Legendre polynomials is above 1. Thus, background functions of Legendre polynomials to expansion order two is more likely, and should be used in the estimation of ω .

Example cases 8-10 in Table 3.4 also include background trends, and without pre-processing FFT cannot pick out the correct frequency. In contrast, BSA includes background functions in the model signal and delivers the desired result. Including

background functions, however, results in over-estimation of the signal-to-noise ratio.

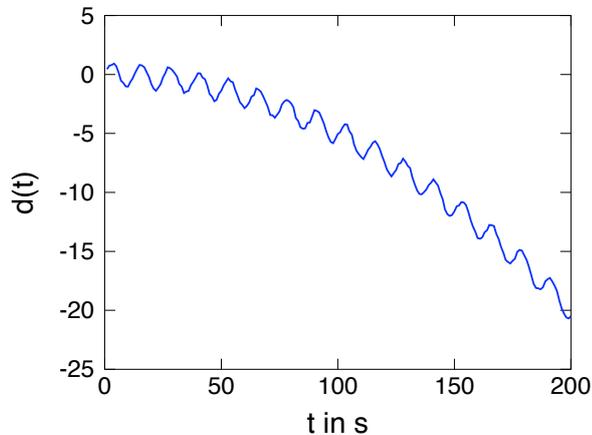


Figure 3.3: Time series with a background trend, simulated from $d(t) = \sin(\omega t) - 0.005t^2 + e$, with $\omega = 0.5$ rad/s, sampled with 1 s intervals to give 200 points. The noise level, e , is 0.1 which corresponds to 10 %.

Short time series

Figure 3.4 shows the results from analysing a short time series. The FFT power spectrum is very broad, see Figure 3.4B, which comes as no surprise given the FFT dependence on the number of data points. BSA estimates the correct frequency sharply, but the maximum probability drops compared to longer time series, see Figure 3.4C. This demonstrates the higher uncertainty of fewer time points.

High noise levels

BSA is also successful at handling high levels of noise, as highlighted in Examples 1-6 in Table 3.4. The frequency estimates are correctly reproduced by the FFT. In these simple test scenarios, the BSA posterior probability distribution estimates the frequency with a significantly higher precision than the FFT.

Whereas the estimated uncertainty of parameter expectation values is a built-in aspect of any probabilistic treatment such as BSA, FFT has no inherent mechanism for assessing the accuracy of the results. The FFT output is summarised by the average, $\hat{\omega}_{FFT}$, and the standard deviation, σ_{FFT} , over the transformed data set. The result show that different noise levels influence the σ_{FFT} more than the σ_{BSA} , see Figure 3.5.

Table 3.3: Automated Model Development. Posterior probability ratio of models including a different a number of background functions, ζ , in the analysis of a time series with a background trend. The first two ratios favour the more complex model (ratio below 1). The ratio between models with two and three background functions, $H_{2\zeta}/H_{3\zeta}$, is above 1 thus favouring the simpler model. The analysis would then normally automatically stop at two background functions, but for demonstration more functions are included here. The ratio stays above 1 thereby building a chain of decisions that always prefer the simpler model.

Models	Model ratio
$H_{0\zeta}/H_{1\zeta}$	1.9459e-06
$H_{1\zeta}/H_{2\zeta}$	1.0256e-167
$H_{2\zeta}/H_{3\zeta}$	622.5
$H_{3\zeta}/H_{4\zeta}$	566.3
$H_{4\zeta}/H_{5\zeta}$	501.8
$H_{5\zeta}/H_{6\zeta}$	99.2

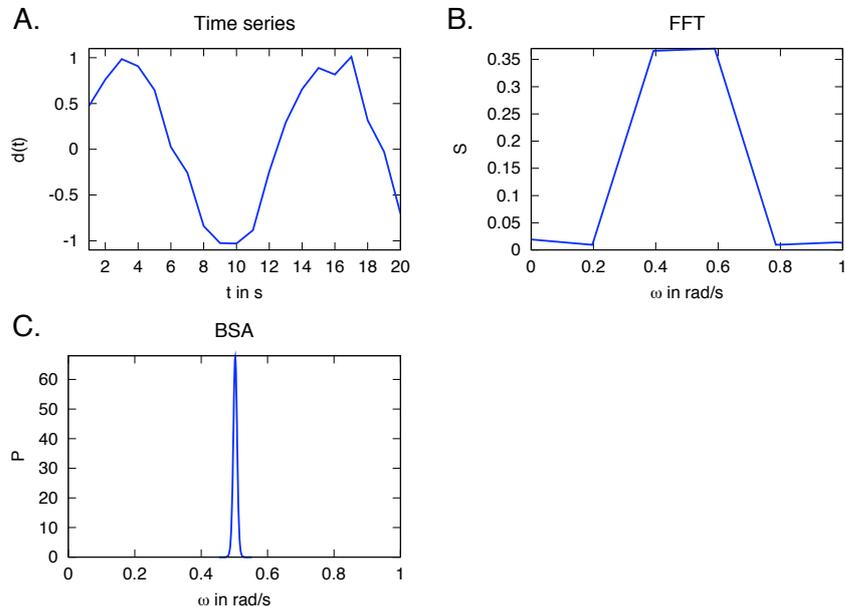


Figure 3.4: Short time series. A: A short time series simulated from $d(t) = \sin(\omega t) + e$, with $\omega = 0.5$ rad/s, and sampled with 1 s intervals to give 20 points. The noise level, e , is 0.1 which corresponds to 10 %. B: FFT results. The y-axis is the spectral power, S . C: BSA result. P denotes the posterior probability density. The BSA estimate of ω is correct, and has considerably less spread than the FFT estimate.

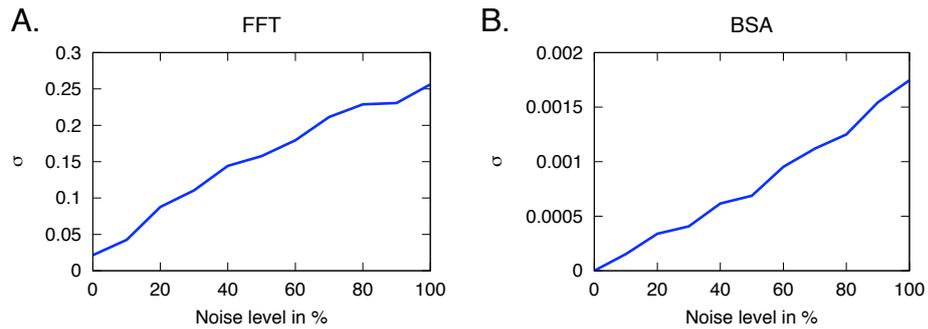


Figure 3.5: Effects of noise on precision. The effect of noise on σ for FFT (A) and BSA (B). The time series were simulated from $d(t) = \sin(\omega t) + e$, with $\omega = 0.5$ rad/s and sampled with 3 s intervals to give 100 points. The noise level, e , varies between 0 and 100 %. Although the qualitative behaviour of noise on the precision is the same for FFT and BSA in this example, BSA produces values of σ that are two orders of magnitude below the FFT values.

Multiple frequencies

Example 7 in Table 3.4 has two frequencies present in the data. Both BSA and FFT shows these two peaks in the resulting plots. Although BSA can be used in this manner with a one-dimensional ω to scan through frequency space and estimate the number of frequencies in the data and their location, if more than one frequency is present, the model should be extended to reflect this. Without this extension the integration procedure around a single point is not well suited, so we have employed Nested Sampling to compute the marginalisation in these cases. For the extension approach, when the posterior probability over $\omega = \{\omega_1\}$ reveals two strong frequencies, then a better model would be $\omega = \{\omega_1, \omega_2\}$. For example, Figure 3.6 shows BSA and FFT results for a test case that includes higher harmonics which give rise to multiple peaks in the log P plot. If more than one peak in the resulting posterior probability emerges, then the model can be extended further. One peak in the posterior probability over the number of modelled frequencies signifies that the correct number of frequencies has been captured.

As another example, Figure 3.7 shows the result of a two-frequency search. Figure 3.7A shows a time series with a high noise level and two very close frequencies of 0.498 and 0.505 rad/s. FFT cannot distinguish them and shows only one peak Figure 3.7B. BSA breaks the resolution and precision limitations inherent to FFT by introducing a continuous probability distribution instead of the fixed number of points and can therefore sample the posterior more finely in areas of high probability. This approach gives rise to a high resolution probability plot in which two distinct frequencies emerge, see Figure 3.7D. The peaks have a larger variance at this local level, but the qualitative information of two underlying frequencies is revealed.

To develop BSA further, windowing of the time series was used to compute the posterior probability distribution of ω at each time point. This is called BSA Local

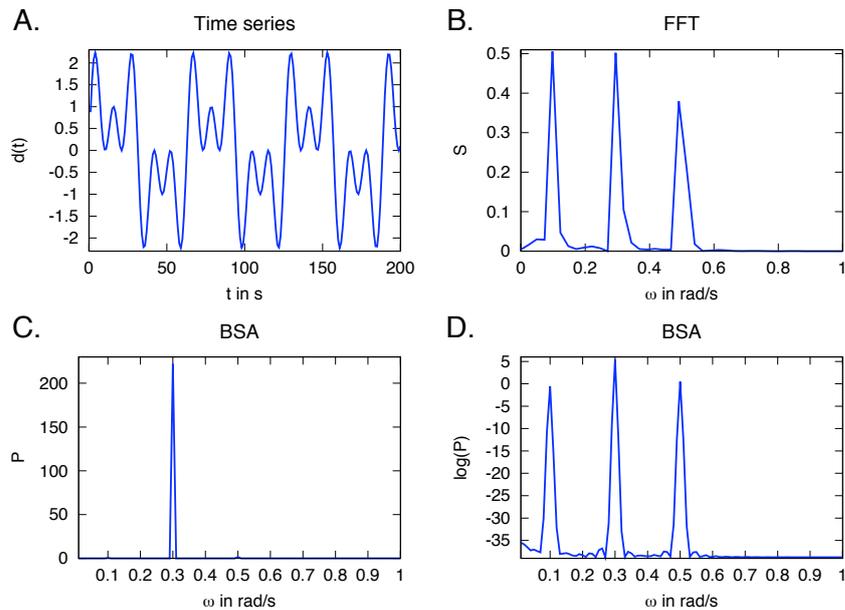


Figure 3.6: Higher harmonics. A: Time series with higher harmonic frequencies, simulated from $d(t) = \sin(\omega t) + \sin(3\omega t) + \sin(5\omega t)$, with $\omega = 0.1$ rad/s, sampled with 1 s intervals to give 200 points. B: FFT results. D: $\log(\text{Probability})$ plot of the BSA results. Both BSA and FFT show three strong peaks in ω . Depending on the length of the series, the truncation, and sampling interval not all peaks will result in an equal probability. C: Probability plot of the BSA results. In the current case, the question of which single frequency is the most probable results in the selection of the $\omega = 0.3$ rad/s frequency.

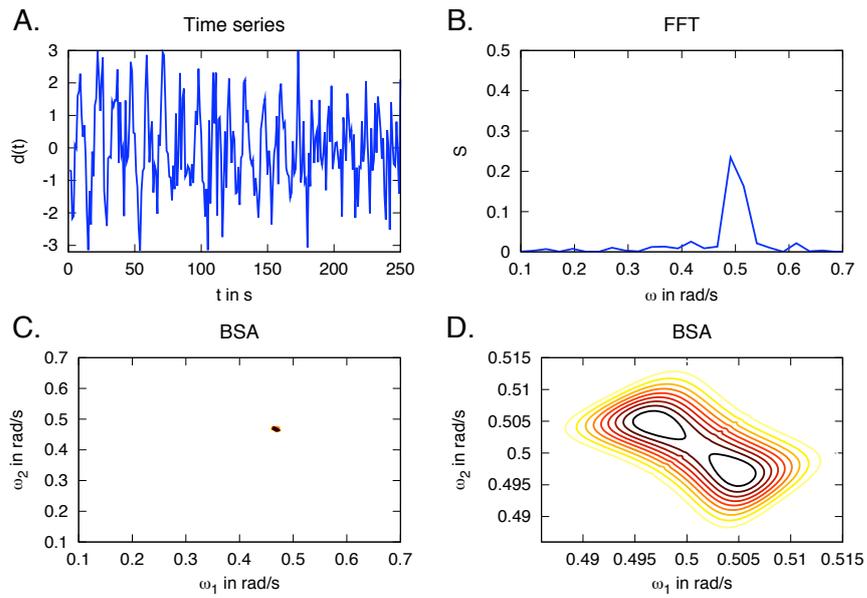


Figure 3.7: Multiple close frequencies with noise. A: A time series containing two close frequencies, $\omega_1 = 0.498$ rad/s and $\omega_2 = 0.505$ rad/s, simulated from $d(t) = \cos(\omega_1 t) + \cos(\omega_2 t) + e$, and sampled with 1 s intervals. The noise level, e , is 0.1, which corresponds to 10 %. B: The FFT result only shows one peak due to the sampling resolution that is determined by the time domain data. In the y-axis, S stands for spectral power. C: The BSA estimate using a two frequency model. Each point in this plot has two frequencies, so only off-diagonal elements correspond to two distinct frequencies and only if both are present will a high joint probability emerge. D: Sampling in the area around the peak of high probability show that two distinct frequencies emerge in strong off-diagonal peaks.

(BSAL). The robustness and negligible peak broadening of BSA with fewer time points allows for this windowing to proceed without the introduction of artifacts due to truncation. This local BSA captures changes in frequency, as shown in Figure 3.8B. The BSAL was compared to Short-Time Fourier Transform (STFT), Figure 3.8D, which is a windowed Fourier Transform, and to wavelets, Figure 3.8C. For the wavelet power spectrum a Morlet mother wavelet was used (Torrence and Compo, 1998). The advantages of BSAL are that it remains within the same BSA framework, has high accuracy, and does not require pre-processing of the data.

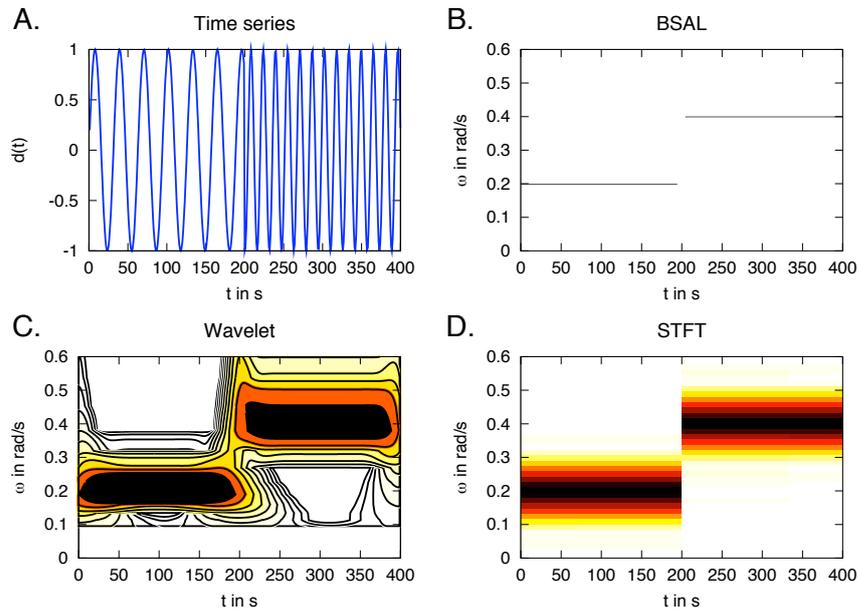


Figure 3.8: Frequency change. A: A time series with a sharp change in frequency half way through the observed time frame, simulated from $d(t) = \sin(\omega t)$ with $\omega_1 = 0.2$ rad/s and $\omega_2 = 0.4$ rad/s. B: The posterior probability distribution for the estimated frequency using a local BSA, BSAL, with a window size of 10, which is one third of the number of data points in a period of the first frequency. Nevertheless, the resulting distribution is so narrow that it resembles a sharp line. C: The wavelet power spectrum using a Morlet mother wavelet (Torrence and Compo, 1998), with a lower cut-off at $\omega = 0.1$ rad/s in the spectrum. D: The STFT power spectrum, window size of 100 and overlap of 5. Both the wavelet and STFT results gives the correct answer but with high variance.

Non-harmonic oscillations

BSA results for oscillations with a non-harmonic shape are superior to the FFT. It highlights an essential difference in the two methods since biological data is often repetitive, but with a wide range of oscillatory patterns. To demonstrate this further, Figure 3.9A shows a time series simulated from an ODE model of cellular Ca^{2+} signals (Kummer

et al., 2000). Such time series presents two potential problems: the time series is chaotic and thus not perfectly periodic, and the signal shape is non-harmonic. The calculation of interspike intervals (ISI) of the time series show that multiple intervals are present, see Figure 3.9D. The highest peak of the FFT plot, in Figure 3.9B, suggests that the entire time series is one period, while BSA suggests a strong angular frequency around 1.2 rad/s, Figure 3.9C . The BSA suggestion is similar to the second FFT peak .

This highlights the differences between frequencies in the data and spike intervals. ISIs are a common way of characterising spike data, however, multiple ISIs need not correspond to multiple frequencies in the data. Of the four strong ISIs shown here, only one of these is identified as a regular period by both BSA and FFT.

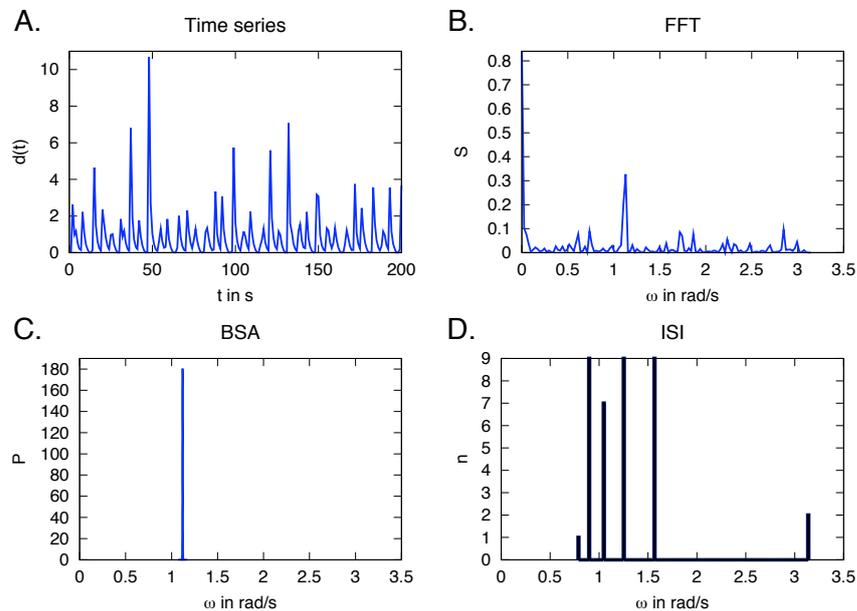


Figure 3.9: Near-periodic oscillations. A: A near-periodic time series simulated from a set of ODEs describing Ca^{2+} oscillations in animal cells (Kummer et al., 2000). In the y-axes, n stands for the number of interspike intervals (ISI) in the ISI plot, P for probability in the BSA plot and S for spectral power in the FFT plot. The ISI plot show several intervals present (D), but both FFT (B) and BSA (C) identify a frequency around $\omega = 1.2$ rad/s.

3.2.2 Calcium spiking

The first biological data set comes from intracellular signalling in plant-microbe interactions. Symbiotic bacteria induce Ca^{2+} oscillations, called Ca^{2+} spiking, in legume root cells (see Chapter 1). These are non-stationary and often noisy time series, causing problems in identifying periodicity. One hypothesis for signal transduction in this system is via frequency encoding (Oldroyd and Downie, 2006), so concluding whether

Table 3.4: BSA and FFT results from simulated harmonic data with noise and background trends. Each time series was generated with a sine function of angular frequency, ω , of 0.5 rad/s with a level of noise in amplitude, e_a , and phase, e_p . In some time series a background trend (b) was included, and in case number 7 an additional sine function of 0.3 rad/s is present. The resulting function was sampled 200 times at an interval of 1 s. Results from FFT are presented in the form of the angular frequency with the highest power, $\hat{\omega}_{FFT}$, and the estimated standard deviation, σ_{FFT} . The Bayesian frequency estimate at the maximum posterior point is denoted by $\hat{\omega}_{BSA}$ and its standard deviation by σ_{BSA} . For comparison, the expectation value of ω and its standard deviation computed using BSA and Nested Sampling (BSA-NS) are denoted by $\bar{\omega}_{BSA-NS}$ and σ_{BSA-NS} . Values of σ below 10^{-8} are listed as 0. The estimated signal-to-noise ratio (s-n) from the Bayesian analysis is given in the last column. The BSA and BSA-NS approaches deliver the same results, other than for the case of multiple frequencies in a 1D search of ω (see case No. 7) which leads an intermediate estimate between the frequencies with a higher standard deviation.

No.	ω	e_a (%)	e_p (%)	b	$\hat{\omega}_{FFT}$	σ_{FFT}	$\hat{\omega}_{BSA}$	σ_{BSA}	$\bar{\omega}_{BSA-NS}$	σ_{BSA-NS}	s-n
1	0.5	1	-	-	0.49	0.06	0.5	0	0.5	0.0002	70
2	0.5	10	-	-	0.49	0.20	0.5	0.0002	0.5	0.0004	6.5
3	0.5	40	-	-	0.49	0.54	0.5	0.0005	0.5	0.0011	1.9
4	0.5	10	10	-	0.49	0.27	0.5	0	0.5	0.0003	4.2
5	0.5	10	40	-	0.49	0.57	0.5	0.0002	0.5	0.0007	2.2
6	0.5	100	40	-	0.49	0.89	0.5	0.0006	0.5	0.0020	0.7
7	0.3, 0.5	10	10	-	0.29, 0.51	0.14	0.3, 0.5	0.0003	0.34	0.0832	1
8	0.5	10	-	$-t$	0	0.15	0.5	0.0002	0.5	0.0002	110
9	0.5	10	-	$-t^2$	0	0.19	0.5	0.0002	0.5	0.0002	90
10	0.5	10	-	$-t^3$	0.02	0.24	0.5	0.0003	0.5	0.0002	35

there is indeed underlying periodicity is of great interest.

The Ca^{2+} spiking has background trends present due to fluorescence bleaching and cell movements, which are assumed to be unrelated to the underlying signal in the cell. Therefore, accounting for the background functions plays a key role in the analysis. Example time series are shown in Figure 3.10A. Nine spiking cells were analysed for an underlying period. The data is obtained by microinjecting a root hair cell with the Ca^{2+} indicator dyes Oregon Green (responds to Ca^{2+}) and Texas Red (non-responsive), after exposing the plant to the bacterial signal molecule that induces the oscillations. The data is a ratio of the fluorescence of the two dyes, showing changes in Ca^{2+} concentrations. The data has been published in Kosuta et al. (2008).

The FFT of Ca^{2+} data results in a very broad periodogram, due to multiple frequencies and high noise levels, Figure 3.10D. Also, the spiking produces a non-harmonic signal which might be another problem for the FFT. For comparison, we also present results from the multitaper method (MTM). The MTM is a non-parametric method of spectral analysis that uses tapers to minimize the variance in the power estimate, and is targeted at short and noisy time series (Ghil et al., 2002). The MTM results were very similar to the FFT, see Figure 3.10B. These periodograms do not address the question of interest: Is there a key period in the Ca^{2+} signal?

In the BSA analysis in Figure 3.10C, the Ca^{2+} spiking data used the Legendre background functions to an expansion order of 1-2, depending on the individual trace. Nested Sampling was used to compute the evidences. Frequencies with high probabilities were picked out, but varied in the interval of approximately 50-120 s, see Table 3.5. However, the strongest periods were in the interval of 75-100 s. If periodicity plays a role in the signal transduction of this system, then the key period should be in this interval. Those with a probability between 0.95 and 1 also have a very low standard deviation, in contrast to the other traces where additional peaks cause a spread. The signal-to-noise ratios were relatively high, between 100-200, possibly as a consequence of including several background functions.

Since the oscillations often start with a rapid period (of approximately one minute) and often slow down over time as the signal continues, BSAL was also tested on Ca^{2+} spiking data. An example time series is shown in Figure 3.11A. In Figure 3.11B, the result of analysing this time series with BSAL is shown, demonstrating that this package copes with the background trend as well as shows the signal period changes over time, without need for pre-processing such as detrending. Good results can be achieved by a windowed Fourier spectra, see (Figure 3.11C), but only after detrending, which was done using a moving average (Brockwell and Davis, 2002). BSA delivers superior results and without the need for data pre-processing.

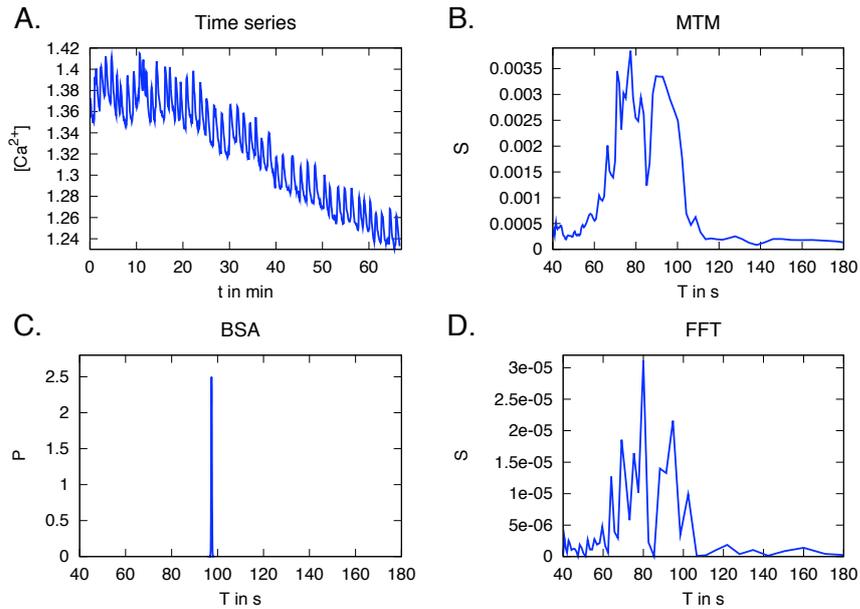


Figure 3.10: Example results of data from Ca^{2+} spiking. A: Time series of Ca^{2+} oscillations measured in a *Medicago truncatula* root hair cell in response to a bacterial symbiont. The data is a relative ratio of the fluorescence of the Ca^{2+} indicator dyes Oregon Green and Texas Red, showing changes in concentration, $[\text{Ca}^{2+}]$. The measurements are taken with 5 s intervals. BSA identifies one strong frequency at a period, T , of about 100 s (C), while MTM (B) and FFT (D) deliver a broad spectrum.

Table 3.5: BSA on Ca^{2+} data. Analysed Ca^{2+} concentrations in *Medicago truncatula* root hair cells in response to bacterial symbionts. The BSA shows strong underlying signals in the data, but the cell-to-cell variability is high. BSA and the BSA-NS disagree when there are multiple peaks present, as can be seen by the large standard deviation of the BSA-NS average period estimate.

Cell	BSA Period \pm std (s)	BSA-NS Period \pm std (s)
1	97.4 \pm 0.23	97.3 \pm 0.15
2	80.9 \pm 0.63	75.2 \pm 10.1
3	74.6 \pm 0.19	74.6 \pm 0.85
4	123.8 \pm 0.16	124.2 \pm 1.18
5	88.9 \pm 0.22	123.9 \pm 0.61
6	74.6 \pm 0.21	113.7 \pm 16.16
7	121.9 \pm 0.22	146.1 \pm 21.53
8	74.4 \pm 0.92	75.2 \pm 2.69
9	48.2 \pm 0.3	64.5 \pm 13.94

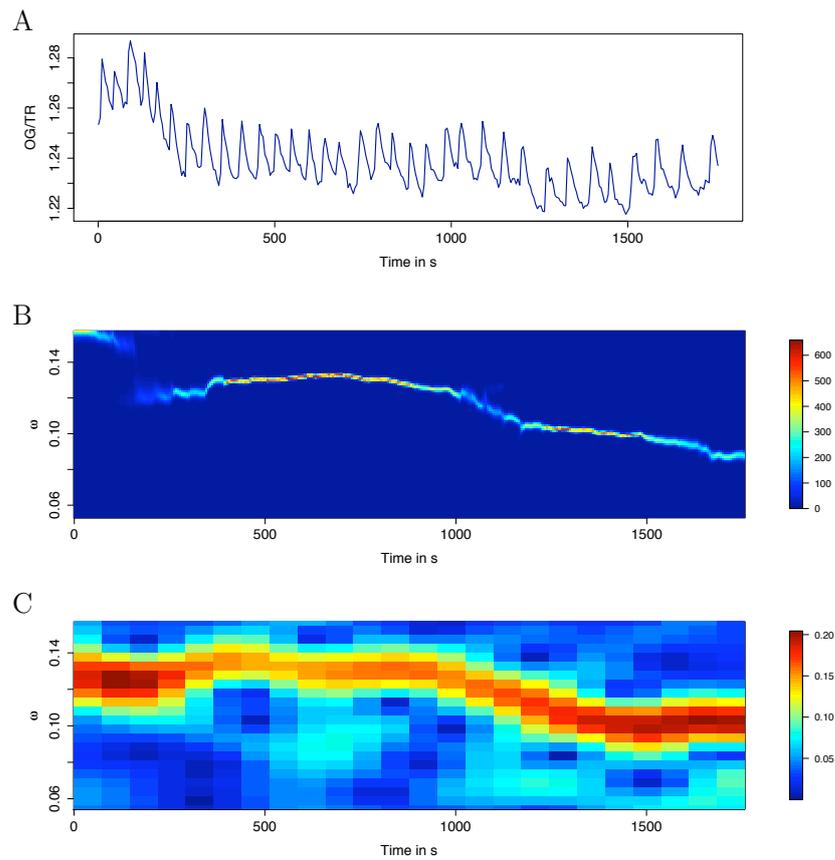


Figure 3.11: Example results from BSAL. A: Calcium oscillations measured in *M. truncatula* root hair cells. B: The 2-dimensional PDF over time and angular frequency (ω), showing that oscillation frequency varies over time. C: The 2-dimensional Fourier spectra from a windowed Fourier transform.

3.2.3 Clock genes

The second biological data set shows gene expression of so-called clock genes. Many processes in plants follow a circadian rhythm (for reviews see e.g. Pruneda-Paz and Kay (2010) or Gardner et al. (2006)). A number of genes in *Arabidopsis thaliana* have been shown to regulate circadian rhythms, and time series of RNA levels show how these clock genes are expressed in cycles (Edwards et al., 2006). Time series with only a couple of cycles are common in biology and provide another suitable test case.

For these circadian rhythms, we chose to analyse RT-PCR data from four clock genes in two genotypes of *Arabidopsis thaliana*. The plants are either wild type, *FRI;FLC*, or mutants, *fri;flc*, of the genes *FRI* and *FLC*. The RNA was extracted from seedlings, and each time series is an average of two biological replicates. An example of the RNA levels of a clock gene is shown in Figure 3.12A. (The data has been published in Edwards et al. (2006).)

FFT on the RNA levels of our clock genes did not give any clear periods, either having only a vague peak or none at all, see Figure 3.12D. This is caused by the FFT's dependence on the length of the time series, and these only had 1-2 cycles. The MTM method had more of a peak in the 20-25 h period, but still lacking in precision, see Figure 3.12B. BSA on the other hand provides a clear peak close to 23 h, see Figure 3.12C, consistently for all 8 time series, see Table 3.6. Nested Sampling was used to compute the evidences. The assigned probabilities are relatively low, but similar probabilities were obtained using simulated data with few data points and high noise levels. The period peaks are stable over all time series, and suggests a probable period which is unaffected by the mutations, in agreement with the conclusions of the original study (Edwards et al., 2006).

Table 3.6: BSA on circadian data. The BSA results of RNA levels of four so-called clock-genes in *A. thaliana*, measured in two different genotypes. The plants are either wild type, *FRI;FLC*, or mutants, *fri;flc*.

Gene	Genotype	BSA Period \pm std (h)	BSA-NS Period \pm std (h)
<i>TOC1</i>	<i>fri;flc</i>	22.75 \pm 0.18	22.58 \pm 0.43
<i>TOC1</i>	<i>FRI;FLC</i>	23.36 \pm 0.20	23.26 \pm 0.42
<i>CCA1</i>	<i>fri;flc</i>	23.58 \pm 0.15	23.67 \pm 0.94
<i>CCA1</i>	<i>FRI;FLC</i>	23.98 \pm 0.16	24.23 \pm 0.72
<i>GI</i>	<i>fri;flc</i>	22.39 \pm 0.14	22.54 \pm 0.86
<i>GI</i>	<i>FRI;FLC</i>	23.41 \pm 0.16	23.61 \pm 0.83
<i>LHY</i>	<i>fri;flc</i>	23.84 \pm 0.16	23.82 \pm 1.54
<i>LHY</i>	<i>FRI;FLC</i>	25.74 \pm 0.19	24.03 \pm 1.23

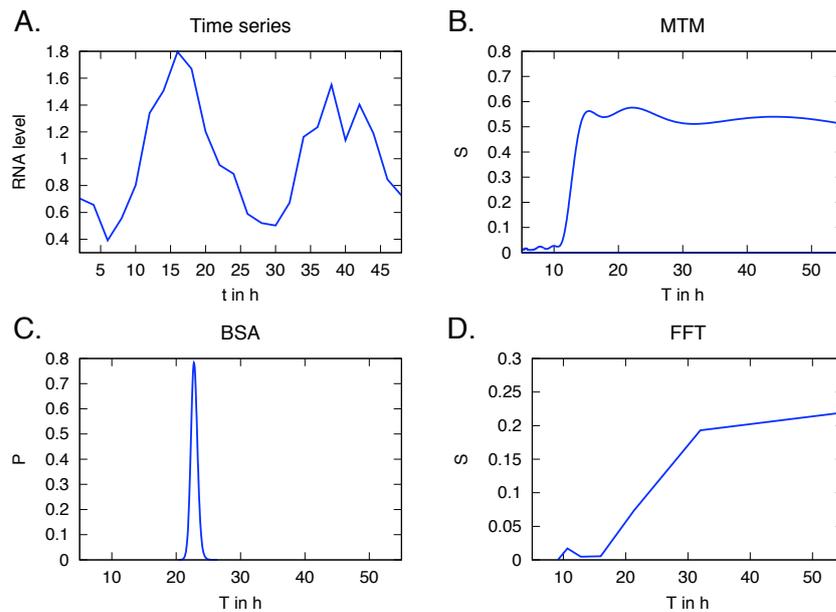


Figure 3.12: Example results of data from a clock gene’s RNA levels. A: RNA levels of a clock gene, *CCA1*, in a wild type background of *A. thaliana*. The data points have 2 h intervals and consist of two averaged biological replicates. The values are normalised to the average of the housekeeping gene *ACTIN2* in wild type.

3.3 Discussion

The tests have shown that BSA delivers superior results in cases where the FFT assumptions are too constraining, most notably in the five cases of background trends, short time series, high noise levels, multiple frequencies and non-harmonic oscillations. BSA is a flexible method allowing the underlying hypothesis to be changed depending on the focus of the analysis, and to directly compare the validity of different hypotheses. It can handle non-uniformly sampled data and has no need for pre-processing procedures. The price of these superior results comes at a computational cost that ranged from tens of seconds to hundreds for the examples shown here.

Bayesian inference offers a powerful way of analysing biological time series. Despite the undisputed value of Fourier theory, there are cases when the necessary requirements for its optimality for time series analysis are not met. This is a consequence of the underlying assumptions of a Fourier Transform, causing it to work optimally only for uniformly sampled, long, stationary, harmonic signals that have either no or white noise. In biology these requirements are rarely fulfilled, requiring pre-processing of the data, such as noise reduction and detrending techniques, with the risk of convoluting the signal and losing valuable information.

By placing the problem of frequency extraction in the framework of Bayesian inference, the known and well-documented problems of Fourier analysis can be overcome. This approach also breaks the resolution and precision limitations inherent to the FFT

by introducing a continuous probability distribution instead of the fixed number of points maintained by the discrete Fourier Transform.

The suggested automated model development worked well in our hands but must be used with caution in practice as the approach is not guaranteed to find a global optimum in model space. Alternate models should be explored and compared using posterior probability ratios or approximations thereof. Conservative priors has been used (uniform, Jeffreys, Gaussian) that make an analytical treatment tractable but in some cases more information could warrant a different choice of prior that might require substantial alternations to our approach to handle the numerics of marginalisation.

The further developments of the BSA method presented here, with automated model development and a local version of BSA, will make the method usable in more situations. Given that the recorded Ca^{2+} oscillations in symbiosis have several of the data characteristics studied here, with high noise levels and strong background trends being the most prominent, the BSA method will be a useful tool in the frequency characterisation of the Ca^{2+} spiking data.

Chapter 4

Bifurcation analysis of a calcium spiking model reveals sources of pattern variation

4.1 Introduction

The oscillations of Ca^{2+} required for the initiation of rhizobial and arbuscular mycorrhizal symbioses occur in the nucleus and in the cytosol associated with the nucleus (Capoen et al., 2011). It is not known how the signal is transduced from the plasma membrane once the receptors there have perceived the symbiotic signals in the soil (see a schematic overview of the pathway in Figure 4.1). Nor is the complete complex of components producing the Ca^{2+} oscillations identified. However, there are several genetic components that are necessary for the Ca^{2+} oscillations, and the proteins that they code for have been used to build a mathematical model to simulate the Ca^{2+} spiking (Hazledine, 2010). This model could recapitulate the symbiotic spikes well, but it was limited in the variability of spiking behaviour that it could produce.

Experimentally, differences in spike period and shape are seen, as well as behaviour such as rapid spiking at the start of some time series, and therefore we sought to use modelling to understand these observations. A natural extension was to add buffers, Ca^{2+} -binding proteins, to the nucleoplasm. Such buffers are present in the plant cell nucleoplasm at high concentrations, and they are a likely candidate to impact on Ca^{2+} signals.

In this chapter, I will use bifurcation analysis to investigate how the Ca^{2+} buffers in the model affect spiking behaviour. This will provide important insights into what mechanisms could cause the observed variations in Ca^{2+} patterns.

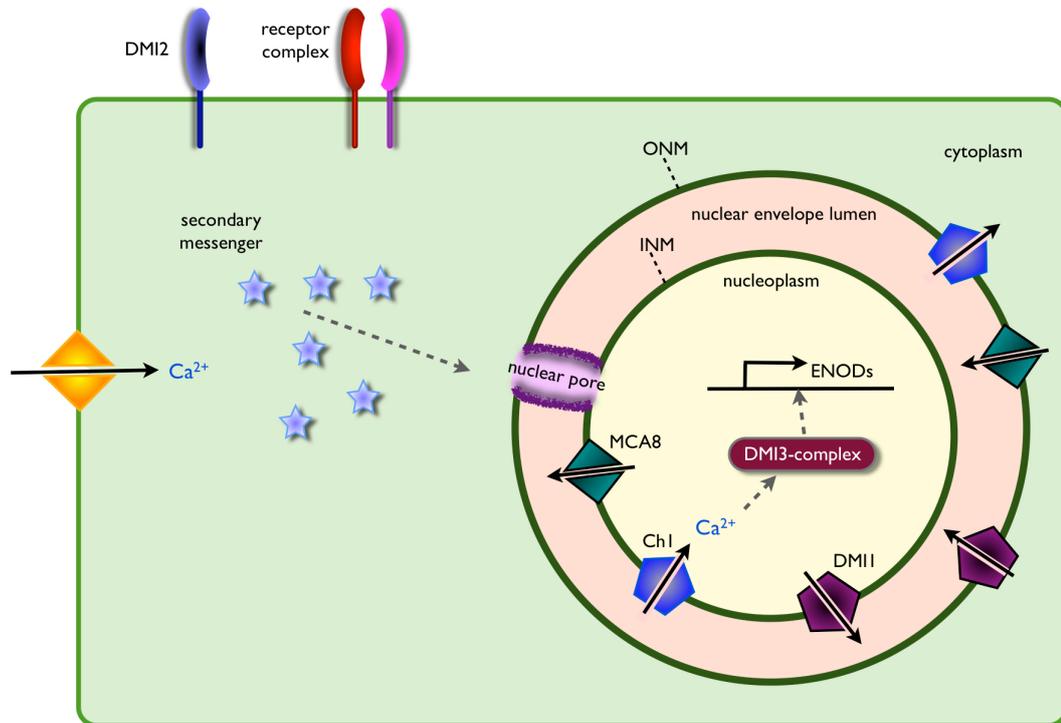


Figure 4.1: Initial signalling events in the SYM pathway, adapted from Oldroyd and Downie (2008). The receptors on the plant epidermal cell perceive signal molecules released from the symbionts (AM fungi and rhizobia). This initiates a signalling cascade inside the cell, involving DMI2 and an unknown secondary messenger in the cytosol. In the rhizobial symbiosis, there is also an early influx of Ca^{2+} into the cytosol. The signalling cascade reaches the nucleus, where the Ca^{2+} oscillations are seen both on the inside and outside of the nuclear envelope. On the nuclear membrane, these oscillations require the cation channel DMI1, a Ca^{2+} channel for release (Ch1), Ca^{2+} pumps such as MCA8, and several nucleoporins. Inside the nucleus, the complex of proteins formed by DMI3 perceive the Ca^{2+} oscillations and regulate downstream gene expressions of for example *ENOD* genes. ONM: outer nuclear membrane, INM: inner nuclear membrane.

4.1.1 Components needed for the symbiotic calcium oscillations

The nuclear envelope and the ER: a Ca^{2+} store

The nuclear envelope (NE) that surrounds the nucleus is enclosed by an outer and inner membrane, where the outer nuclear membrane (ONM) is contiguous with the endoplasmic reticulum (ER) (Petersen et al., 1998). The inner nuclear membrane (INM) on the other hand is isolated, and has discrete characteristics (Graumann and Evans, 2010). The lumen of the NE is directly connected to the ER lumen, but it is unclear whether diffusion and transport is uniform between them.

The NE and the ER are believed to share a resting Ca^{2+} concentration at approximately $100 \mu\text{M}$ (Petersen et al., 1998), making the NE and ER one large Ca^{2+} store. Ion channels on both the INM and ONM control the release of Ca^{2+} . In the SYM pathway where Ca^{2+} is observed in and around the nucleus, this store is the likely source for the released Ca^{2+} .

The nuclear pore complex

Three nucleoporins, which are components of the nuclear pore complex (NPC), have been identified in genetic screens as being necessary for the Ca^{2+} oscillations in the SYM pathway, and are located upstream of the oscillations (Saito et al., 2007; Kanamori et al., 2006; Groth et al., 2010), but their physiological role has not been demonstrated. One hypothesis could be that Ca^{2+} needs to pass between the outside and inside of the nucleus during spiking. However, conflicting results as to whether or not Ca^{2+} can pass through the NPC has led to a long-standing controversy, but some convincing results point to the allowed passage for Ca^{2+} in animal systems, - reviewed in Bootman et al. (2009); Petersen et al. (1998). It has been shown that the plant NE is similar to the animal counterparts, as are the plant NPCs (Fiserova et al., 2009). Also, the size of the NPC would suggest that Ca^{2+} can pass through freely. However, recent studies in plants argue that Ca^{2+} is not passively diffusing, but that there is instead an autonomous role for the nucleus versus the cytosol in the generation of Ca^{2+} signals. This is based partly on observed delays between nucleus and cytoplasmic Ca^{2+} signals in response to several stimuli (Mazars et al., 2011). In the SYM pathway it has been argued that regardless of how freely Ca^{2+} can move through these pores, the oscillations seen on the outside and inside of the nucleus need to be generated separately, and therefore are not an artifact of diffusion (Capoen et al., 2011).

There could also be roles for the NPC in the SYM pathway that do not directly involve Ca^{2+} . Perhaps the unknown secondary messenger from the plasma membrane needs to pass through the NPC, or possibly the NPC is involved in protein trafficking, for example moving ion channels between the inner and outer nuclear membranes.

The cation channel DMI1

The cation channel DMI1 (in *Medicago truncatula*) seems to play a key role in the Ca^{2+} oscillations. DMI1 is a predicted cation channel located at the nuclear rim (Ané et al., 2004; Riely et al., 2007). The orthologue genes of *MtDMI1* has been found in *Lotus japonicus*, named *POLLUX* (Charpentier et al., 2008), and in *Pisum sativum*, named *SYM8* (Edwards et al., 2007). Furthermore in *L. japonicus* *POLLUX* has a twin, *CASTOR*, and the proteins that they specify have both been identified as cation channels localised to the nuclear envelope. Results indicate that the *CASTOR* channel might be found only on one side of the nuclear envelope, while *POLLUX* could be placed on both the inner and outer membranes (Charpentier et al., 2008), but preferentially on the inner membrane (Capoen et al., 2011).

The mode of action of DMI1 could be to allow cations into the nuclear envelope and in that way to counterbalance the release of Ca^{2+} into the nucleoplasm and cytoplasm. Alternatively, the cation channel could change the electrical potential across the nuclear membranes, triggering the opening of voltage-activated Ca^{2+} channels, or indeed a combination of these roles. There have been studies supporting a membrane potential over the nuclear envelope in plants (Matzke and Matzke, 1986), which would allow membrane hyperpolarization to happen.

POLLUX has a weak similarity to potassium (K^+) channels in yeast (Charpentier et al., 2008). K^+ channels are not as well characterised in plants as in animals. They are multimeric proteins with several transmembrane domains, so-called α -subunits, forming one or two pores (Lebaudy et al., 2007). The regulation of K^+ channels varies widely, sometimes including Ca^{2+} or voltage.

The calcium channel

Ca^{2+} channels are needed to release Ca^{2+} from the store, but the Ca^{2+} channel acting in the SYM oscillations is still unidentified. However, active research is ongoing and it is likely that more information on this channel will be available soon (M. Charpentier, personal communication). So far, only indirect studies through pharmacological inhibitors have pointed to the SYM Ca^{2+} channel (Engstrom et al., 2002). Therefore it is also not known what the mechanism of activating this channel might be. One hypothesis is that it could be triggered to open through voltage-gating (Edwards et al., 2007; Oldroyd and Downie, 2008). This means that the channel is modulated in its activity by the membrane potential. Hyperpolarization of the membrane is when the membrane potential becomes more negative, by cations exiting the compartment or anions entering. When the reverse happens, it is termed membrane depolarization. Voltage-sensitive ion channels respond to either depolarization or hyperpolarization.

Other activators of channel opening or closure include small ligands acting as secondary messengers, and examples of such are IP_3 , cADPR, NAADP, Ca^{2+} ions, cAMP, cGMP, glutamate and glycine. Cyclic nucleotides such as cAMP and cGMP are common regulators of cation channels, and have been identified in that role both in animal

and plant cells (McAinsh and Pittman, 2009).

The ATP-driven Ca^{2+} pumps

Pharmacological evidence as well as the characteristics of the SYM Ca^{2+} oscillations support the involvement of plant type IIA Ca^{2+} pumps in these Ca^{2+} signals (Engstrom et al., 2002). The pumps are needed to re-sequester Ca^{2+} after each release event, actively transporting Ca^{2+} against the concentration gradient into the nuclear envelope. Recently, a study confirming this has found a SERCA-type IIB Ca^{2+} ATPase, MCA8, that is located on the inner and outer nuclear envelope of *M. truncatula*, and lies upstream of symbiotic Ca^{2+} oscillations (Capoen et al., 2011). MCAs are also found on animal plasma membranes, and have central roles in the Ca^{2+} signalling of those cells not only by the removal of the toxic Ca^{2+} ions from the cytoplasm but also by interacting with Ca^{2+} -sensitive proteins and modulating the signals (Holton et al., 2010). In plants, type IIB SERCA-pumps are widely distributed on plant membranes and the variation in their structure points to them being differentially regulated (Sze et al., 2000).

Ca^{2+} buffers: Ca^{2+} -binding proteins

After release of Ca^{2+} into the cytosol and nucleoplasm, Ca^{2+} -binding proteins, or Ca^{2+} buffers, will quickly bind these ions. These buffers are often kinases that require bound Ca^{2+} for their activity (Sanders et al., 2002). Buffering of Ca^{2+} contributes to keeping the concentration of free Ca^{2+} inside the cell low, which is necessary due to the toxicity of Ca^{2+} .

Ca^{2+} buffers may play an important role in determining the nonlinear behaviour of the Ca^{2+} signalling system, and indeed, in animal systems such relationships have been found (Falcke, 2003a). As Ca^{2+} buffers can diffuse at different speeds, and with widely different affinities for binding Ca^{2+} , their effect on the system is highly complex.

4.2 Description of the calcium model

As a consequence of the many unknown mechanisms in the production of the Ca^{2+} oscillations in the SYM pathway, mathematical modelling was initiated in the groups of Morris and Oldroyd in an attempt to understand the underlying system. Saul Hazledine, a PhD student in the Morris group, built the initial model of symbiotic Ca^{2+} oscillations (Hazledine, 2010). The group later added buffers to the simple model, which was done by Wojciech Kozłowski. The results of this chapter consist of bifurcation analyses using this extended model. The extended model and results thereof was published in Granqvist et al. (2012b).

4.2.1 Model components and equations

The Ca^{2+} spiking model is a set of four first-order ordinary differential equations (ODEs) that describe how current, voltage, and the amount of bound buffer (binding to Ca^{2+}) change over time. For a brief background into dynamical systems theory and ODEs, see Appendix A.2.

The simple model, which has no buffer species present, is based on the three key proteins known to be required for the Ca^{2+} oscillations: DMI1, MCA8, and the Ca^{2+} channel (labelled Ch1 in Figure 4.3).

The structure of DMI1 predicts partial gating by calcium (Edwards et al., 2007), and this allows the formulation of an oscillatory mechanism: this cation channel, upon activation by a secondary messenger from the plant plasma membrane, causes hyperpolarization of the nuclear envelope. Based on the hypothesised mechanism of the Ca^{2+} release events, the unknown Ca^{2+} channel was assumed to be voltage dependent. Through this change in voltage, the Ca^{2+} channel opens and Ca^{2+} feeds back onto the cation channel. The SERCA-type Ca^{2+} ATPase re-sequesters the Ca^{2+} released into the Ca^{2+} store, against the concentration gradient. This mathematical model is placed on the inner nuclear envelope, based on the preferential localisation of DMI1 as well as the results showing the protein decoding the spikes to reside inside the nucleus (Oldroyd et al., 2011).

Therefore, the model oscillations arise as a combination of a voltage-activation of the Ca^{2+} channel, and a Ca^{2+} -activation of the cation channel DMI1. A phase space diagram of the oscillation behaviour as a consequence of voltage (v) and Ca^{2+} (c) in the simple model is shown in Figure 4.2. The oscillation is a stable limit cycle with a period of approximately 85 seconds that oscillates between 130 and 630 nM Ca^{2+} . The system has three points where it does not oscillate, so-called fixed points, in this parameter range at $(v, c) = \{(55, 0), (54.975, 0.03185), (-2.64, 0.158)\}$, each of which are unstable except for the point at $(55, 0)$. If the concentration of Ca^{2+} at the start of the simulation is below a threshold of approximately 22 nM Ca^{2+} , then the system decays to this stable fixed point. The fixed points appear as circles in Figure 4.2., while the nullclines for voltage and Ca^{2+} appear in blue and green, respectively. The nullclines are the values for which the voltage or Ca^{2+} do not change over time. Where nullclines intersect neither of the systems variables change, in our case voltage and Ca^{2+} , and thus the system is at rest.

Buffers were added as a natural extension of the model, as the Ca^{2+} data measured comes from Ca^{2+} indicators that themselves act as buffers. Therefore the ideal situation would be to match experimental data with simulated Ca^{2+} buffers. Furthermore, buffers were included in the model to try to capture the variation in system dynamics. Given that DMI1 binds Ca^{2+} in the model, leading to changes in membrane voltage and the subsequent release of more Ca^{2+} , it is reasonable to consider that buffers binding to the released Ca^{2+} might affect the oscillations. When Ca^{2+} levels oscillate in the nucleoplasm, the proportion of bound and unbound buffers will also oscillate, since

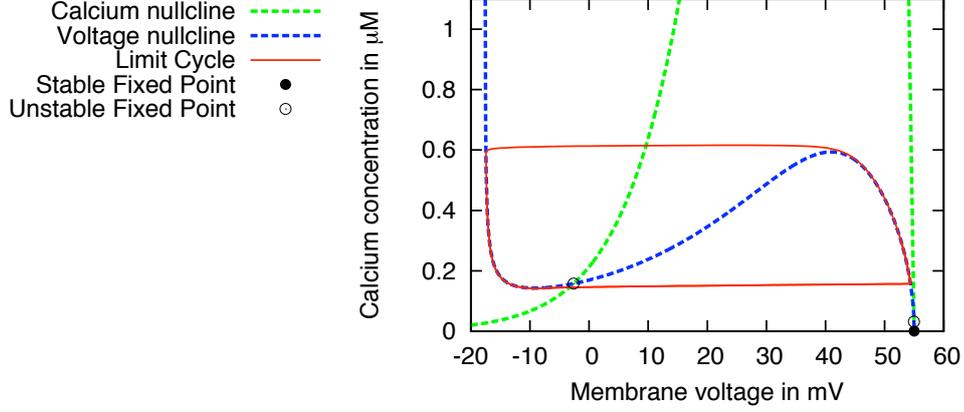


Figure 4.2: Phase space diagram of Ca^{2+} concentration and voltage for the basic ODE model without buffers. The system trajectory from an arbitrary initial condition is shown making one full oscillation in red. The nullclines $\dot{c} = 0$ and $\dot{v} = 0$ are shown in green and blue, respectively. The stable fixed point at $(0, E_{Ca}) = (0, 55)$ is shown as a full black circle, and the two unstable fixed points as empty black circles. Model parameters and initial conditions can be found in Table 4.1.

more of the buffer will bind Ca^{2+} when more Ca^{2+} ions are available. A schematic overview of the model, including buffers, is shown in Figure 4.3.

The model equations contains terms representing a voltage over the inner nuclear membrane, v , the Ca^{2+} concentration in the nucleus, c , and the concentration of buffers that are bound to Ca^{2+} , $\mathbf{p} = \{p_1, p_2\}$. This is expressed as

$$\dot{v} = \frac{1}{C_m}(I_{ca} + I_k) \quad (4.1)$$

$$\dot{c} = E_{ps}(\alpha I_{ca} - \mu c) + \sum_{i=1}^N R_i \quad (4.2)$$

$$\dot{p}_i = -R_i \quad (4.3)$$

$$R_i = k_i^- p_i - k_i^+ (p_i^{\text{tot}} - p_i) c \quad (4.4)$$

$$k_i^+ = \frac{k_i^-}{K_i^d}, \quad (4.5)$$

where $N = 2$, k_i^+ and k_i^- are the on-rate and off-rate for the i^{th} buffer species ($i = 1, \dots, N$), K_i^d is the dissociation constant, and p_i^{tot} is the total concentration of the i^{th} buffer. This model assumes that binding to Ca^{2+} occurs in a 1:1 ratio to the Ca^{2+} buffers. All initial conditions and parameter values are given in Table 4.1.

The model is electrical in nature and has two currents. I_{ca} is the Ca^{2+} current

through the voltage-gated channel, and I_k is the K^+ current through the cation channel DMI1, which is modelled here as a K^+ channel. These currents are described by

$$I_{ca} = G_{ca} \left(1 + \exp \left(\frac{V_{ml} - v}{K_{ml}} \right) \right)^{-2} (E_{ca} - v) \quad (4.6)$$

$$I_k = G_k \frac{c^2}{c^2 + K^2} (E_k - v), \quad (4.7)$$

and all parameters are found in Table 4.1.

The expressions for I_{ca} and I_k imply that the Ca^{2+} channel has two voltage-dependent, Hodgkin-Huxley type gates (Izhikevich, 2007) and that the conductance of the K^+ channels is Ca^{2+} -dependent. In animal systems Ca^{2+} -gated K^+ channels have been shown to possess a voltage dependency; however, investigations into DMI1 have shown that changes in voltage have no discernible effect on the conductance of the channels. The Ca^{2+} pump is modelled as electrically neutral, with a constant pump rate, as many Ca^{2+} pumps are electroneutral (Thomas, 2009).

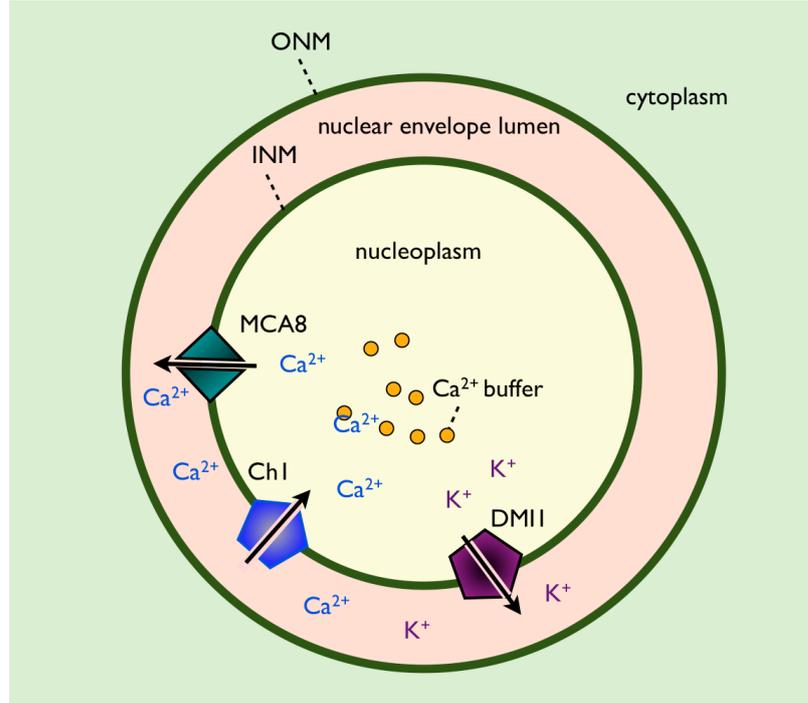


Figure 4.3: The Ca^{2+} spiking model including buffers. The model simulates the Ca^{2+} oscillations observed inside of the nucleus. The three proteins included are sufficient for simulated Ca^{2+} spiking: a Ca^{2+} channel, here called Ch1, the cation channel DMI1, here assumed to be a K^+ channel, and the Ca^{2+} pump MCA8. Ca^{2+} -binding proteins, buffers, are found at high levels inside the nucleus and were included in the Ca^{2+} spiking model to capture observed variation in oscillation regularity, period and spike shapes.

Table 4.1: Definitions and values for initial conditions and parameters used in the Ca^{2+} model. The source indicates whether the value was fitted (Fit) or taken from literature (Lit.). The top four are initial conditions and the rest are model parameters.

Symbol	Description	Value	Unit	Source
v	initial voltage	-27.0	mV	-
c	initial calcium	0.23	μM	-
p_1	initial bound buffer 1	0.0	μM	-
p_2	initial bound buffer 2	0.0	μM	-
V_n	Volume of nucleus	160	μm^3	Lit.
C_m	Capacitance of nuclear envelope	5.1	pF	Lit.
E_{ca}	Equilibrium potential of Ca^{2+}	55	mV	Lit.
F	Faraday constant	10^{14}	10^{14}	$\text{fC} \cdot \mu\text{mol}^{-1}$
E_{ps}	Scaling factor relating total Ca^{2+} changes to changes in free Ca^{2+}	0.001	-	Lit.
α	Conversion of Ca^{2+} current to Ca^{2+} flux	0.03239	$\mu\text{M} \cdot \text{fC}^{-1}$	$\frac{1}{2FV_n}$
G_c	Total max conductance of voltage gated Ca^{2+} channels	2864	pS	Fit
V_{ml}	Half maximal activation of voltage gated Ca^{2+} channel	50.0	mV	Fit
K_{ml}	Constant in scaling function for voltage-gated Ca^{2+} channel	14.7	mV	Fit
G_k	Total max conductance of Ca^{2+} activated K^+ channels	302	pS	Fit
K	Constant in Hill function for Ca^{2+} activated K^+ channel	0.953	μM	Fit
μ	Pump rate into the nuclear envelope	24.9	s^{-1}	Fit
E_k	Equilibrium potential of K^+	-17.7	mV	Fit
k_1^-	Off-rate for buffer 1	5.04	s^{-1}	Fit
K_1^d	Disassociation constant for buffer 1	5.4	μM	Lit.
p_1^{tot}	Total concentration of buffer 1	0.1	μM	Fit
k_2^-	Off-rate for buffer 2	4.5	s^{-1}	Fit
K_2^d	Disassociation constant for buffer 2	5.4	μM	Fit
p_2^{tot}	Total concentration of buffer 2	0.1	μM	Fit

4.2.2 Model parameters

Initial parameters for the simple model, excluding buffers, were fitted by Saul Hazledine using a swarm optimiser (He et al., 2007). The buffer parameters were later optimised to produce different spike shapes, which are observed with two different experimental methods. The first experimental technique is to inject fluorescent Ca^{2+} -sensitive dyes into the cell, which increase their fluorescence when binding Ca^{2+} (Rudolf et al., 2003). The second method is to transform the plants with Ca^{2+} reporter genes that encode proteins that indicates increases in Ca^{2+} through fluorescent resonant energy transfer (FRET) (Miyawaki et al., 1999). The differences in spike shapes seen is a result of different binding characteristics of such Ca^{2+} indicators, and this was discovered together with Jongho Sun and Saul Hazledine.

To allow the model to reproduce these differences, the buffer parameters were optimised by fitting model-generated spikes to different experimentally observed spikes. The dissociation constant of one of the buffers was chosen to reflect the Ca^{2+} indicators typically used in experiments, which is either a Ca^{2+} dye or a Ca^{2+} reporter. The Ca^{2+} -sensitive dye Oregon Green has a K^d of $0.17 \mu\text{M}$ (Molecular Probes) and the Ca^{2+} reporter Cameleon YC2.1 has a K^d of $5.4 \mu\text{M}$ (Miyawaki et al., 1999). Unless when stated otherwise, in this chapter K^d values matching those of the Cameleon system will be used, see example oscillations in Figure 4.4. The buffer oscillation shown is the concentration of buffer that is bound to Ca^{2+} .

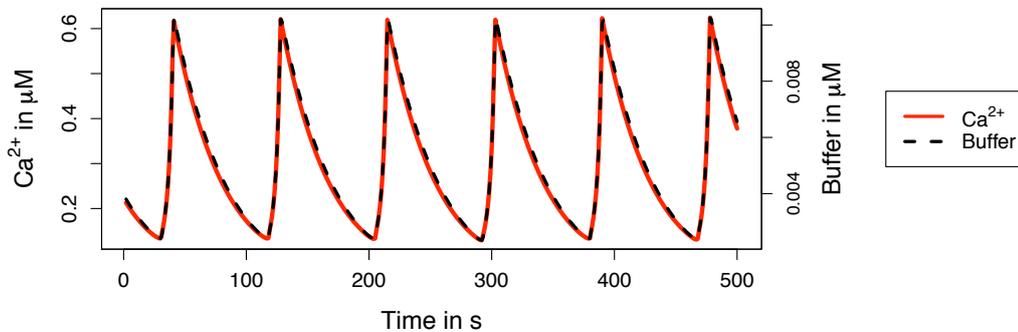


Figure 4.4: Simulated Ca^{2+} spiking with the Ca^{2+} model including buffers. The left y-axis shows the Ca^{2+} concentration, and the right shows the concentration of buffer bound to Ca^{2+} . The bound buffer oscillation follows the Ca^{2+} oscillation.

By changing the buffer characteristics to the known values of OG and Cameleon YC2.1, it was possible to match both Oregon Green and Cameleon YC2.1 spike shapes, see Figure 4.5. For the buffer parameters of the OG simulations see Table 4.2.

Even when comparing SYM oscillations from the same method and treatments,

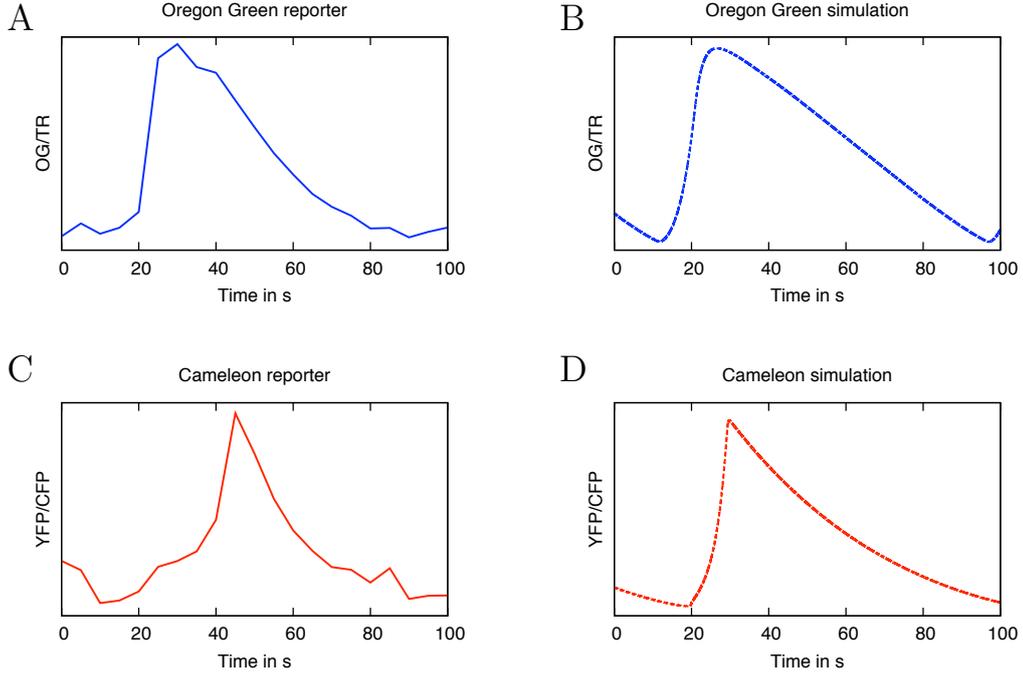


Figure 4.5: The model parameters can be fitted to different spikes shapes, as demonstrated here by matching experimental data from *M. truncatula* root hairs. A: An experimental Ca^{2+} transient measured using the dyes Oregon Green (OG) and Texas Red (TR). The y-axis is the ratio OG/TR after the trace has been de-trended with the moving average method. B: A simulated Ca^{2+} transient with parameters optimized to replicate the shape measured with OG. C: An experimental Ca^{2+} transient recorded using cameleon (YC2.1) as a reporter. The y-axis is the ratio of YFP/CFP, after the trace has been de-trended with the moving average method. D: A simulated Ca^{2+} transient with parameters optimized to replicate the shape measured with cameleon. The y-axes of the simulated data represents the Ca^{2+} bound to the fluorescent dye/cameleon complex, and thus represents what would be measured in the experiments.

Table 4.2: Buffer parameters for the simulations Oregon Green, the fluorescent Ca^{2+} indicator used in many Ca^{2+} imaging experiments.

Symbol	Description	Value	Unit	Source
k_1^-	Off-rate for buffer 1	0.1	s^{-1}	Fit
K_1^d	Disassociation constant for buffer 1	0.17	μM	Lit.
p_1^{tot}	Total concentration of buffer 1	0.005	μM	Fit
k_2^-	Off-rate for buffer 2	1.0	s^{-1}	Fit
K_2^d	Disassociation constant for buffer 2	0.17	μM	Fit
p_2^{tot}	Total concentration of buffer 2	0.005	μM	Fit

variability is common in the experimentally observed Ca^{2+} oscillations. Examples are shown in Figure 4.6, where three time series of the Ca^{2+} spiking are displayed. They are all recorded in root hair cells of the transgenic *M. truncatula* cameleon YC2.1, and the oscillations are induced identically. Can the Ca^{2+} model reproduce this variability?

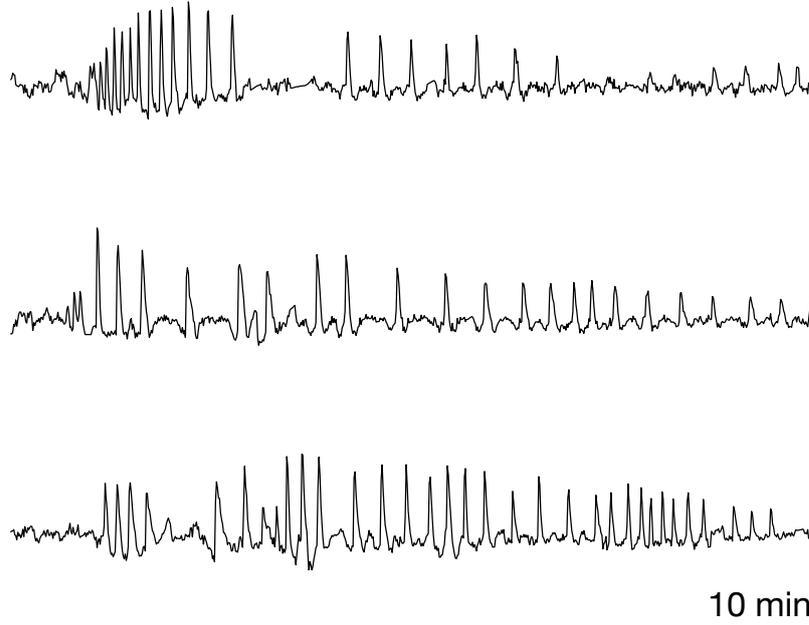


Figure 4.6: Examples of variability in Ca^{2+} spiking. I collected the data in *M. truncatula* with the Ca^{2+} reporter YC2.1, and the data have been detrended. The traces are examples of the variability in period both between and within individual time series. The y-axis is not shown, since the data is non-ratiometric and therefore has no known concentrations. After detrending the relative y-value of the data has also been removed. For an overview of the process from raw data to presented data, see Appendix B, Figures 3-4.

4.3 Results

The Ca^{2+} model described above reproduces the observed Ca^{2+} oscillations induced both by rhizobial and mycorrhizal signals. As will be described below, it also offers a plausible explanation for other observations such as the often observed rapid spiking at the beginning of oscillations, the period and spike shape variability, and the initiation and termination of spiking. The parameter influences on the spiking patterns will be further investigated here using bifurcation analysis.

Bifurcation analysis investigates the stability of a system and how it depends on the parameters, and the name comes from the Latin word for branching, *bifurcus*. For example, in the Ca^{2+} oscillation example, a bifurcation point is when the oscillations start or stop. Changes in the oscillation behaviour, for example in period length, is

not a qualitative change, only quantitative, and thus does not involve bifurcations. Analysing where, and which, bifurcations occur can give insight into the system and its parameter sensitivity. For more information of the theory of bifurcation analysis, see Appendix A.2-3.

The main behaviours found were stable limit cycles, stable steady states and unstable steady states. On a few occasions, unstable limit cycles were also found. Unstable steady states does not attract the trajectories. As an example, Figure 4.7 shows how a ball might move over a two-dimensional landscape. It will never settle on the unstable points unless placed there undisturbed, which will not happen *in vivo*. Rather, if the ball moves over this landscape, it will go to stable states, such as the valley in our example figure, representing a stable steady state. The presence of an unstable steady state does, however, indicate that at that point, there is an alternative direction for the ball to go.

In this model, the points of interest in this system are the transitions between the stable steady states and the stable limit cycles, which is the transition between resting state and a stable oscillation. For simplicity, the unstable states are not shown in this analysis.

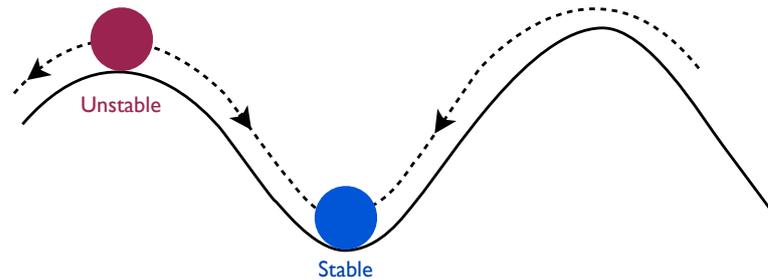


Figure 4.7: Unstable and stable steady states. The coloured balls represent the location of the dynamical system, in a two-dimensional landscape. Red indicates that the ball is resting in an unstable steady state, and blue is resting in a stable steady state. The arrows indicate where the balls would roll if given a small push.

4.3.1 The influence of calcium buffers on spiking patterns

The buffer parameters are the focus of the bifurcations analysis: the concentration of buffers and their binding characteristics. Also, the parameter E_{ps} is linked to buffer concentrations as it controls the proportions of free Ca^{2+} in the nucleoplasm. More free Ca^{2+} can also be achieved by lower buffer concentrations.

Bifurcation diagrams are shown in Figure 4.8 for the buffer parameters E_{ps} and total buffer concentration (p_1^{tot}). E_{ps} is the proportion of Ca^{2+} that is unbound in the nucleoplasm and p_1^{tot} is the total concentration of buffer 1 in the model. In red, the stable limit cycles are shown, which is where the Ca^{2+} model oscillates. The point of

transition between the stable steady state, in black, and the red limit cycles, is known as a supercritical Hopf bifurcation. Both parameters have this Hopf bifurcation, and the results show that when there is low amount of buffer present, or when there is a lot of Ca^{2+} free in the nucleoplasm, the model does not oscillate. Perhaps this might be because the difference between the level of Ca^{2+} in the nuclear envelope and the nucleoplasm is smaller, so the hyperpolarization created by DMI1 is not large enough to induce oscillations.

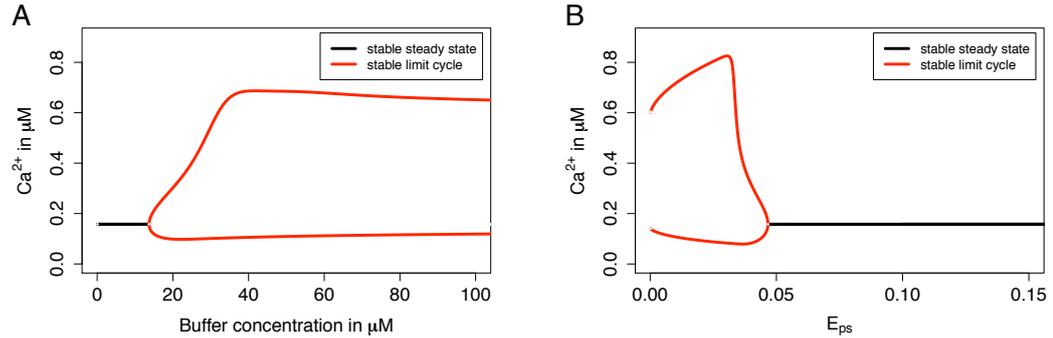


Figure 4.8: Bifurcation analysis of the buffer parameters p_1^{tot} and E_{ps} . At low total buffer concentrations, p_1^{tot} , there are no oscillations and the system is in a stable steady state (A). Conversely, at high values of the parameter E_{ps} , which means that more Ca^{2+} is free in the nucleus, there are also no oscillations.

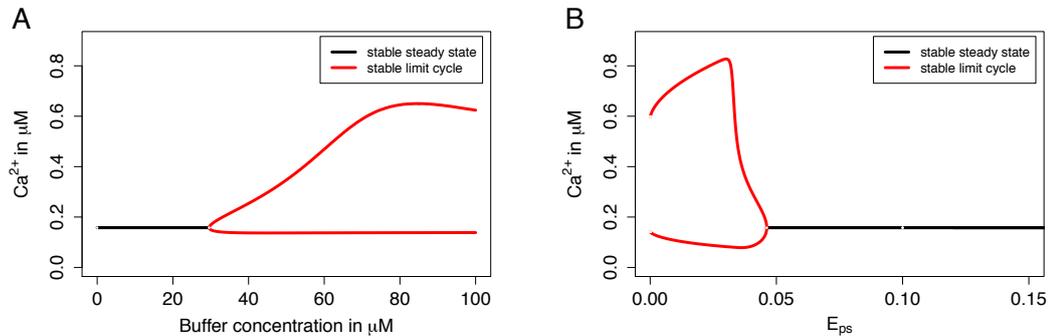


Figure 4.9: Bifurcation analysis of the buffer parameters p_1^{tot} and E_{ps} with OG buffer parameters (see Table 4.2). At low buffer concentrations, p_1^{tot} , there are no oscillations and the system is in a stable steady state (A). Conversely, at high values of the parameter E_{ps} , which means that more Ca^{2+} is free in the nucleus, there are also no oscillations.

The Cameleon reporter complex produces sharper Ca^{2+} spikes than the Oregon Green complex due to its higher K^d . However, using different K^d and k^- values only changes the bifurcation diagrams quantitatively, not qualitatively. The bifurcation point shown in for example Figure 4.8A may move, but does not disappear or

change type. As an example, the bifurcation diagrams of p_1^{tot} and E_{ps} with OG buffer parameters (see Table 4.2) are shown in Figure 4.9.

A two-parameter bifurcation diagram of these two parameters, p_1^{tot} and E_{ps} , is shown in Figure 4.10A. As either of the two parameters are varied, the Hopf point moves so that the point of onset or termination of oscillations changes. The parameter ranges where oscillations or a steady state exist are separated by the trajectory of the Hopf point. At very low values of E_{ps} , it ceases to matter what the buffer concentration is: the model will oscillate. However, at high levels of E_{ps} , which means that more Ca^{2+} is free in the nucleoplasm, high concentrations of buffers are also required for the oscillations. These two parameters act on the same mechanism in the model, buffer saturation, but in different ways. One is controlling the buffer concentration and one is controlling the free Ca^{2+} concentration, and they can therefore balance each other.

Note that with the two-parameter bifurcation diagrams, only the trajectory of the chosen Hopf point is monitored. There may be parameter changes that affect the system behaviour (oscillation or rest) in addition to the movement of the Hopf. For example, later result will show that high levels of buffer concentrations terminate oscillations, but they do that by pushing the system down to the stable steady state shown in Figure 4.2. Such dynamics are not visible in the plot showing the Hopf point trajectory.

Varying the buffer characteristics for one of the buffers, K_1^d , causes a twist in the trajectory of the Hopf bifurcation at low values of K_1^d (4.10B). This is to be expected since a location move is also present when comparing Figure 4.8B and Figure 4.9B, even though it is too small to be seen clearly from the plot. Furthermore, the Hopf bifurcation location over E_{ps} is not moved by high levels of the dissociation rate, k^- , see Figure 4.11A. However, at low values of k^- , a slight twist in the Hopf point trajectory can be seen, see Figure 4.11B.

Taken together, both buffer concentrations and to a smaller extent buffer characteristics such as K^d and k^- values, can shift the location of the Hopf bifurcation, as has been shown here in relation to the E_{ps} parameter.

Changing the buffer concentration (p_1^{tot}) has a strong influence on oscillatory period, see Figure 4.12. A higher buffer concentration gives longer periods of oscillations. This can be understood as the Ca^{2+} released being bound up by buffer proteins, causing a longer time until the Ca^{2+} channel closes in response to a high electrochemical gradient created by the released Ca^{2+} . It then also takes longer for the cation channel DMI1 to be activated again by the feedback of Ca^{2+} .

The E_{ps} parameter also has a marked influence on the periodicity of oscillations, although opposite to the buffer concentrations. At low values of E_{ps} , less free Ca^{2+} is available which causes the oscillations to be slower. As the E_{ps} value approach zero, the period increases exponentially.

At very high buffer concentrations, over 1000 μM , the oscillations are inhibited, see Figure 4.13. This occurs because there is not enough free Ca^{2+} to sustain the oscillations, and provide the necessary feedback to the cation channel DMI1.

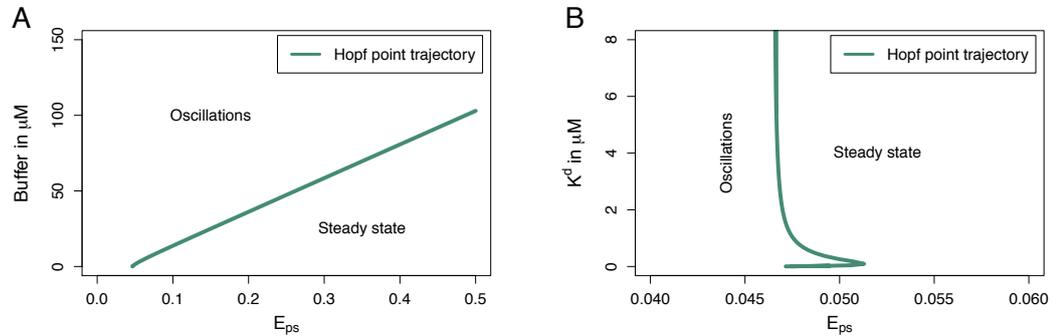


Figure 4.10: Two-parameter bifurcation analysis of the buffer concentration p_1^{tot} and E_{ps} (A) and the dissociation constant for the buffer, K_1^d , and E_{ps} (B). The line shows the trajectory of the Hopf bifurcation that separates steady state from oscillations, and therefore the plots show how the parameter area where oscillations can be found changes by varying parameters.

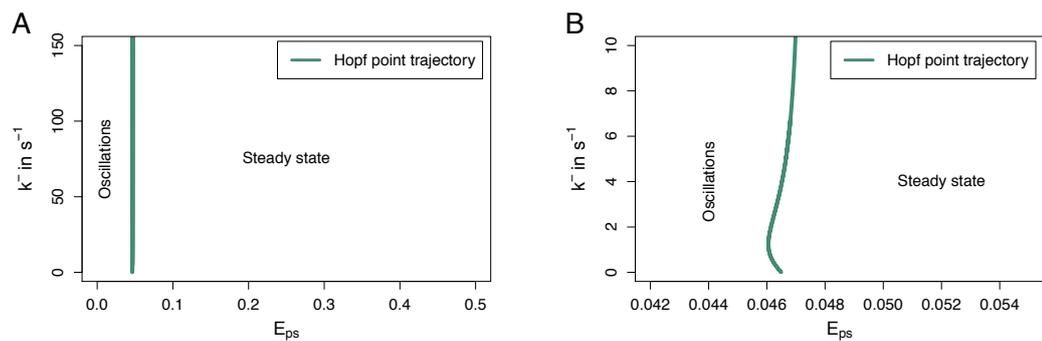


Figure 4.11: Two-parameter bifurcation analysis of E_{ps} and the dissociation rate of buffer 1, k_1^- . At high values of k_1^- , the Hopf point does not move (A), but at low value there is a slight shift (B). The line shows the trajectory of the Hopf bifurcation that separates steady state from oscillations, and therefore the plots show how the parameter area where oscillations can be found changes by varying parameters.

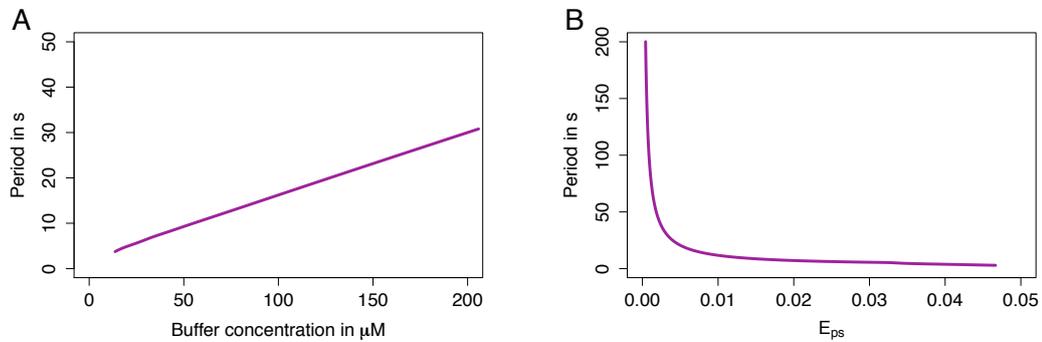


Figure 4.12: The effect of changes in total buffer concentrations on Ca^{2+} spiking period. The buffer parameters in A and B work in opposite directions, one increasing the period of oscillations (A) and the other decreasing it (B).

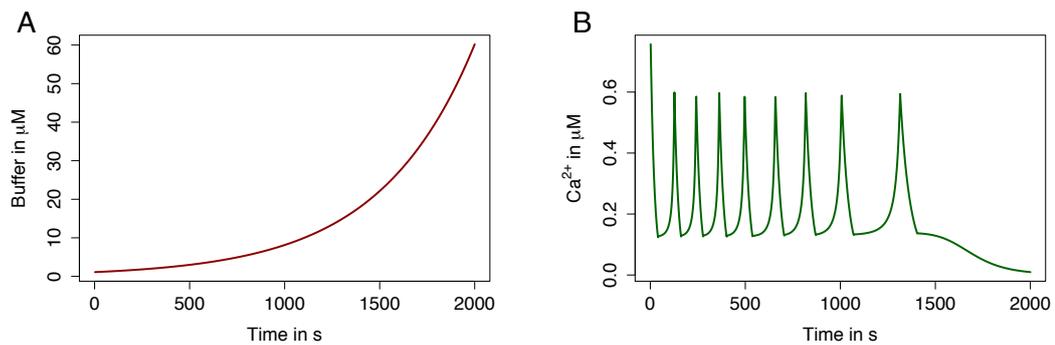


Figure 4.13: Simulation demonstrating how increasing buffer levels can inhibit Ca^{2+} oscillations. The model includes two buffers, one that is held constant and a second that is increasing over time. The total buffer concentration of buffer species 1 increases with every time step in the simulation, starting at 0.1 μM (A). The Ca^{2+} oscillations terminate as the buffer concentration increases (B).

4.3.2 Other sources of variation

There are other model parameters that could also be candidates for producing the observed spike variability, beside the buffer parameters. One of them is the parameter designated α , and it is related to the nucleoplasmic volume (see Table 4.1). The bifurcation diagram of α is shown in Figure 4.14A, and shows the limited parameter range where oscillations arise. For some other parameter sets, the bifurcation analysis of α show unstable limit cycles and possibly the start of a period doubling cascade in the region between the two stable limit cycle areas in Figure 4.14A. This is interesting since period doubling is a well-known feature that leads to chaos in some systems, and the SYM oscillations have been investigated for chaos (Hazledine et al., 2009). However, the small range of parameters in which this was found makes it unlikely to be biologically relevant (K. Tsaneva-Atanasova, personal communication).

Another sensitive parameter is the pump rate, μ , which could be a key controller of spike shape since the downwards slope of the Ca^{2+} is when the pump is the key player. This hypothesis is confirmed in the model, but the pump rate range in which the model oscillates is relatively small. The bifurcation diagram of the pump rate is shown in Figure 4.14B, showing two Hopf bifurcations enclosing the region of oscillations.

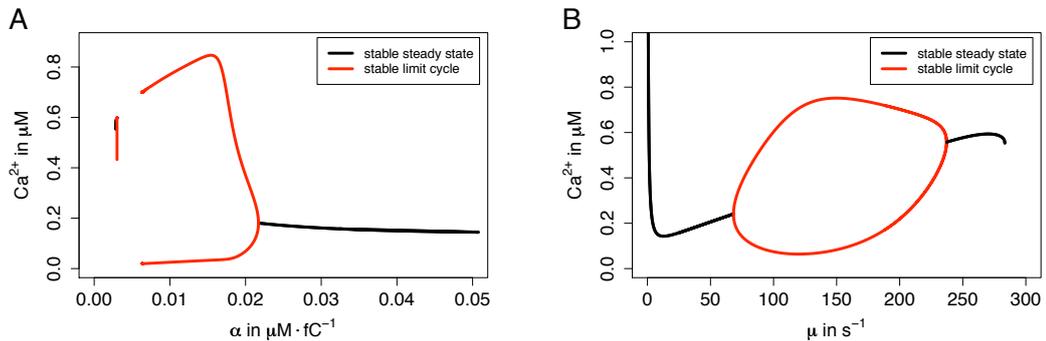


Figure 4.14: Bifurcation analysis of other sensitive model parameters. α is related to nuclear volume, and shows a very small area in which oscillations occur (A). The pump rate, μ , also restricts oscillations to a specific parameter range (B).

An example of the effect on spike shape and period due to pump rate is shown in Figure 4.15. With a lower pump rate the period is longer and the spike shapes are more drawn out (Figure 4.15A). A higher pump rate results in a shorter period and a more symmetrical spike (Figure 4.15B). However, this variability is confined to the oscillating interval of the pump rate in Figure 4.14B.

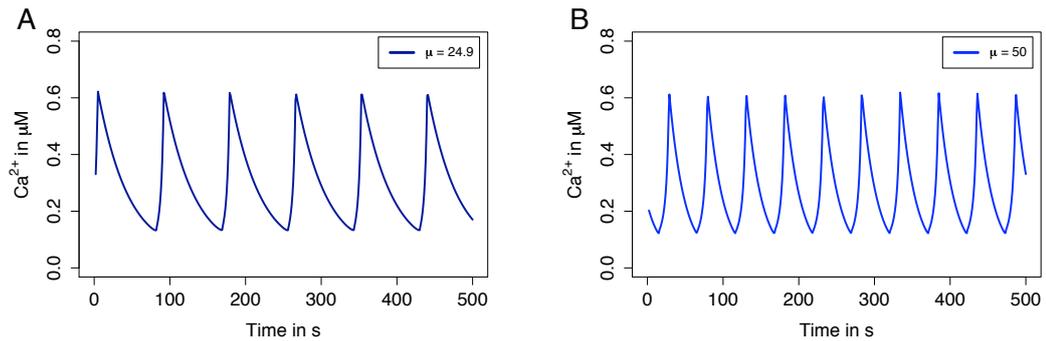


Figure 4.15: The influence of pump rate, μ , on spike period and shape. With lower pump rates (A), the oscillations have a longer period and a broader spike shape than at higher pump rates (B).

4.4 Discussion

The model of these Ca^{2+} oscillations recapitulates the oscillatory behaviour seen in the real system, and including Ca^{2+} buffers allows the biological variability to be reproduced. This effect of buffers is consistent with previous observations (Marhl et al., 1997; Falcke, 2004). During the oscillations, concentration fluctuations of proteins binding to Ca^{2+} may give rise to the variability with a single Ca^{2+} trace, such as the rapid spiking sometime observed at the beginning of spiking, and at other times later in the oscillations. Furthermore, cell-to-cell differences may be caused by concentration differences of these buffers. However, other causes cannot be ruled out with the current model, since many parts of the mechanisms creating Ca^{2+} oscillations are unknown.

Pump rate is another possible source of spike shape variation, since the downwards slope represents uptake into the Ca^{2+} store, and with a faster pump this might go faster. The pump rate in this model does have an influence, but the parameter range in which oscillations can be found is smaller than for example the buffer parameters. Thus, the pump rate is likely to contribute to the oscillatory behaviour, but not be the major determining factor for the observed variation. The buffers are perhaps also more likely to vary during a single time series, since Ca^{2+} buffers may be more movable than pumps in the nuclear envelope. There are also reports of buffers like calmodulin moving into the nucleus in response to some signalling pathways (Deisseroth et al., 1998).

Differences in Ca^{2+} oscillation signatures have been seen not only between genetically identical cells, but also between legume species. The aquatic legume *Sesbania rostrata* has been shown to have faster oscillations when nodulated at lateral root bases, compared to the oscillations associated with root hair invasion in *M. truncatula* (Capoen et al., 2009). Such differences could be caused by buffers, but could also be caused by other differences, for example the pump rate.

This model hypothesized a purely voltage-controlled Ca^{2+} channel. Voltage changes in the nuclear membrane of neural cells have been shown to be involved in periodic

release of Ca^{2+} from the nuclear envelope (Yamashita, 2011). However, there are other mechanisms compatible with the data such as the more complex option of voltage and ligand gating or purely ligand gating (Dodd et al., 2010).

The model presented here could be modified as more biological information becomes available on the proteins producing the Ca^{2+} oscillations. Other aspects could also benefit from further work, for example the high sensitivity of the parameter α . This parameter is related to nuclear volume, and the bifurcation analysis found this parameter to have a variable behaviour in a narrow region. This result is interesting in light of observations that the nucleus can grow in volume during spiking (Sieberer et al., 2011). It could be that nuclear volume changes contribute to variability in oscillation patterns, but this would require further studies.

Ca^{2+} spiking does not only occur on the inside of the nucleus, but also in the cytoplasm surrounding the nucleus. The current model does not address this, as it is located on the inside of the nuclear envelope and uses parameters such as nuclear volume. This was a choice based on the knowledge that the main Ca^{2+} decoder in *M. truncatula*, CCaMK, is located in the nucleoplasm. However, future studies could also investigate the cytoplasmic spiking, and spatial modelling of the Ca^{2+} oscillations have already started addressing aspects of this (Capoen et al., 2011).

The model (Hazledine, 2010) demonstrates a possible mechanism for Ca^{2+} spiking in the common SYM pathway. By including buffers, it also addresses the effects that buffer characteristics could have on the spiking. I have shown that the main bifurcation in the system is a supercritical Hopf bifurcation at the point between rest and oscillations. Furthermore I have shown that concentrations of Ca^{2+} buffer can drastically change the period of spiking, and that they also move the Hopf point. At very high levels of buffers the spiking terminates altogether. Other buffer characteristics, such as disassociation constants, can move the Hopf point as well but to a lesser degree. They can also change the spike shapes and reproduce differences seen experimentally. Given these insights, it is a plausible hypothesis that changes in Ca^{2+} buffer concentrations and/or characteristics could give rise to the experimentally observed variability in Ca^{2+} spiking.

Chapter 5

Comparing patterns of calcium oscillations in *Medicago truncatula*

5.1 Introduction

The existence of shared components of the SYM pathway between AM and rhizobial symbioses (see Figure 1.2 in Chapter 1) has led to questions about how the plant cell distinguishes the two signalling pathways. One hypothesis is that differences in the patterns of Ca^{2+} oscillation could allow the downstream components to separate AM and rhizobial signals (Oldroyd and Downie, 2006). Subsequently, when Ca^{2+} oscillations induced by the AM fungi and the rhizobial Nod factors were compared, differences were found in both spike structures and periodicity (Kosuta et al., 2008).

Since then, a group of diffusible signal molecules released by the AM fungi *Glomus intraradices* has been identified (Maillet et al., 2011). These signal molecules have been termed “Myc factors”, analogous to rhizobial Nod factors. Using Myc factors in the Ca^{2+} experiments allows for a higher control and a better comparison to the experiments with isolated Nod factors. This chapter will consist of a rigorous comparison of the Ca^{2+} spiking induced by mycorrhizal and rhizobial signal molecules.

5.1.1 Rhizobial LCOs

During the establishment of symbiosis, legume flavonoids or isoflavonoids in the soil are perceived by the bacterial symbionts (Zhang et al., 2009). As described in Chapter 1, this activates the release of lipochitooligosaccharides (LCOs) called Nodulation (Nod) factors from rhizobia. The backbone of LCOs consist of chitin, a polysaccharide that is made up of β 1-4-linked *N*-acetylglucosamine residues. Chitin is a major constituent of the arthropod exoskeleton and of fungal cell walls. Several Nod factors have been studied in detail, and their structures and genetics are well understood. Small differ-

ences in the decorations of the chitin backbone between rhizobial strains contribute to host-specificity. There is an acyl chain which can hold differences, and also substitutions at the reducing and non-reducing end of the glucosamine residues (Dénarié et al., 1996; Lerouge et al., 1990; D’Haeze and Holsters, 2002).

5.1.2 Mycorrhizal LCOs

Several studies have demonstrated that AM fungi release diffusible signals, which are perceived by the plant (Weidmann et al., 2004; Oláh et al., 2005; Gutjahr et al., 2009b; Kosuta et al., 2008; Chabaud et al., 2011). The delay in identifying these signals is probably due to the difficulty of working with AM fungi, which are obligate, multinucleate biotrophs (Parniske, 2008). It had been speculated that these signal molecules might be similar to Nod factors, given the fungal nature of chitin. Also, because the AM symbiosis developed far earlier than the rhizobial symbiosis, it seemed a reasonable hypothesis that the rhizobia are imitating fungal signals (Masson-Boivin et al., 2009). Indeed, recently LCOs have been shown to be released from AM fungi (Maillet et al., 2011). The LCOs are tetrameric and pentameric, with a mix of sulphated and non-sulphated varieties, and with the acyl chains that are common in fungi. These LCOs, sometimes referred to as Myc factors, increased mycorrhization in *Medicago truncatula*, *Tagetes patula* and *Daucus carota*, as well as stimulating the root branching at concentrations similar to those at which Nod factors act (Maillet et al., 2011).

Interestingly, the mycorrhizal LCOs have oleic acid chains while rhizobial Nod factors have never been shown to have an oleic acid chain, and this may be a key difference between fungal and bacterial LCOs (Maillet et al., 2011; Gough and Cullimore, 2011). The acyl group might also be important, since small changes could change the biologically active structure and affect perception mechanism and binding affinities to the plant receptor.

A mixture of sulphated and non-sulphated LCOs are also found in broad-host range rhizobia, which perhaps indicates a similar strategy in specifying a broad host range. AM fungi can infect an enormous variety of plants, as opposed to many highly host-specialised rhizobial strains (Groves et al., 2005; Morando et al., 2011).

5.1.3 LCOs, COs and calcium spiking

It was shown that AM fungi and exudates derived from them induce Ca^{2+} spiking (Kosuta et al., 2008; Chabaud et al., 2011). However, the concentrations of the diffusible signal molecules in these experiments are unknown, and may vary widely. Purified Nod factors have long been used in isolation from the rhizobia, since it was demonstrated that they induce the same Ca^{2+} response as the bacteria themselves (Wais et al., 2002). Purified Nod factors can activate Ca^{2+} spiking at concentration above 10^{-14} M, with a dose-response curve showing an increasing number of cells responding with higher concentrations (Oldroyd et al., 2001a).

Chitoooligosaccharide elicitors (COs) come from fungal cell walls, and they activate defence responses in plants. These COs are structurally very similar to Nod factors, consisting of a chitin backbone (Hamel and Beaudoin, 2010). LCOs induce Ca^{2+} spiking, but so do undecorated chitin chains. This chitin-induced spiking has been observed in pea (Walker et al., 2000) but not in bean (Cárdenas et al., 1999) nor in alfalfa (Ehrhardt et al., 1996), which requires a specific Nod factor decoration: an o-linked sulphate group. The chitin-induced Ca^{2+} spiking in pea was of a lower frequency than normal spiking from Nod factors, and the concentrations needed for the response were higher.

5.1.4 The model legume *Medicago truncatula*

Medicago truncatula has been used in the present study to compare different regimes for LCO induction of Ca^{2+} oscillations. *M. truncatula*, or barrel medic, is a model legume used extensively for genetic studies. It is a Mediterranean herb that can self-fertilise and has a short generation time, and currently good transformation protocols with *Agrobacterium* have been achieved (Cook, 1999). This plant forms symbioses with *Sinorhizobium meliloti*, a highly studied rhizobial species, and with arbuscular mycorrhizal fungi such as *Glomus intraradices*. The genome of *M. truncatula* is small, diploid and is currently being sequenced (Young et al., 2011).

5.2 Results

The mycorrhizal LCOs were prepared by the group of Jean Dénarié, at INRA in France (Maillet et al., 2011). The pattern of Ca^{2+} spiking was compared for the synthetic Nod factor NodL (henceforth abbreviated as Nod factor); synthetic sulphated mycorrhizal LCOs (Syn S-LCO); synthetic non-sulphated mycorrhizal LCOs (Syn NS-LCO); a mixture of unpurified biological mycorrhizal LCOs (M-LCOs); the AM fungi *Glomus intraradices* (*G. intraradices*); and chitin tetramers as a control (CT4). An overview of these Ca^{2+} inducers and their abbreviations can be found in Table 5.1. Data were collected using a stable line of *M. truncatula* YC2.1 seedlings, which expressed the Ca^{2+} reporter cameleon YC2.1 (Miyawaki et al., 1999). Root hair cells in zone 1 were monitored.

Syn NS-LCO and Syn S-LCO were initially tested for spiking by Jongho Sun, who recorded differences in the concentration level required for the same proportion of cells to respond. Therefore, the Ca^{2+} spiking in the present study were compared at levels which activate the same proportion of cells. The focus was on two levels, 50% and 80%, since they correspond to half-way up the sigmoidal curve from the titration series and to the point where the curve levels off towards saturation. The 50% level is reached by 100 pM Nod factor, 1 nM Syn S-LCO, and 1 μM Syn NS-LCO, respectively. The 80% level is reached by 1 nM Nod factor, 10 nM Syn S-LCO, and 10 μM Syn NS-LCO,

respectively. From each of these treatments, a minimum of ten traces were collected for the analysis.

In addition to this core data set, data was included from previous fungal experiments performed by Jongho Sun and Sonia Kosuta, as well as data for M-LCOs and CT4, also from Jongho Sun. An overview of all the data groups collected and who performed the experiments is found in Appendix B.1, Table 1. The main analysis will be in relation to the 50% concentration level, because this seemed to be a more biologically relevant level. Some additional results will be shown at the 80% level.

Table 5.1: Descriptions of the signal molecules used to induce Ca^{2+} spiking.

Description	Abbreviation
Synthesised Nod factor, Nod factor L	Nod factor
Synthesised fungal LCO A, sulphated	Syn S-LCO
Synthesised fungal LCO B, non-sulphated	Syn NS-LCO
Biological mix of fungal LCOs A and B	M-LCOs
Chitin tetramer	CT4

Representative traces of the 50% concentration level of Nod factor, Syn S-LCO and Syn NS-LCO, as well as the other groups, can be seen in Figure 5.1. For M-LCOs, no titration curve was established since the amount of the M-LCOs was very limited. Here, 100 times dilution of the original concentration was used in comparison to the lower concentration level of the other LCOs. Undiluted M-LCOs were used for the higher concentration level.

5.2.1 Interspike intervals

The interspike interval distributions are defined as the times between all the individual Ca^{2+} spikes within a group. The distributions for the 50% level are shown as boxplots in Figure 5.2. The distributions are all very similar, with a median in the vicinity of 60-80 seconds in all groups apart from CT4. The CT4-induced time series had longer interspike intervals, with a median around 120 seconds.

The similarity of the distributions were tested with a Mann-Whitney-Wilcoxon test, and the results of this can be found in Table 2 in Appendix B. At this concentration level, with 50% of cells spiking, the only groups showing a significant difference to the others (with a p-value < 0.05) was Syn NS-LCO and CT4. When going to the 80% level the difference for Syn NS-LCO disappeared (see Figure 5.3). It is possible that the difference at lower concentrations was due to less initial rapid spiking, something often seen in the start of traces at high concentrations of Nod factors. For CT4, data from a higher concentration level was not collected, but these results are consistent

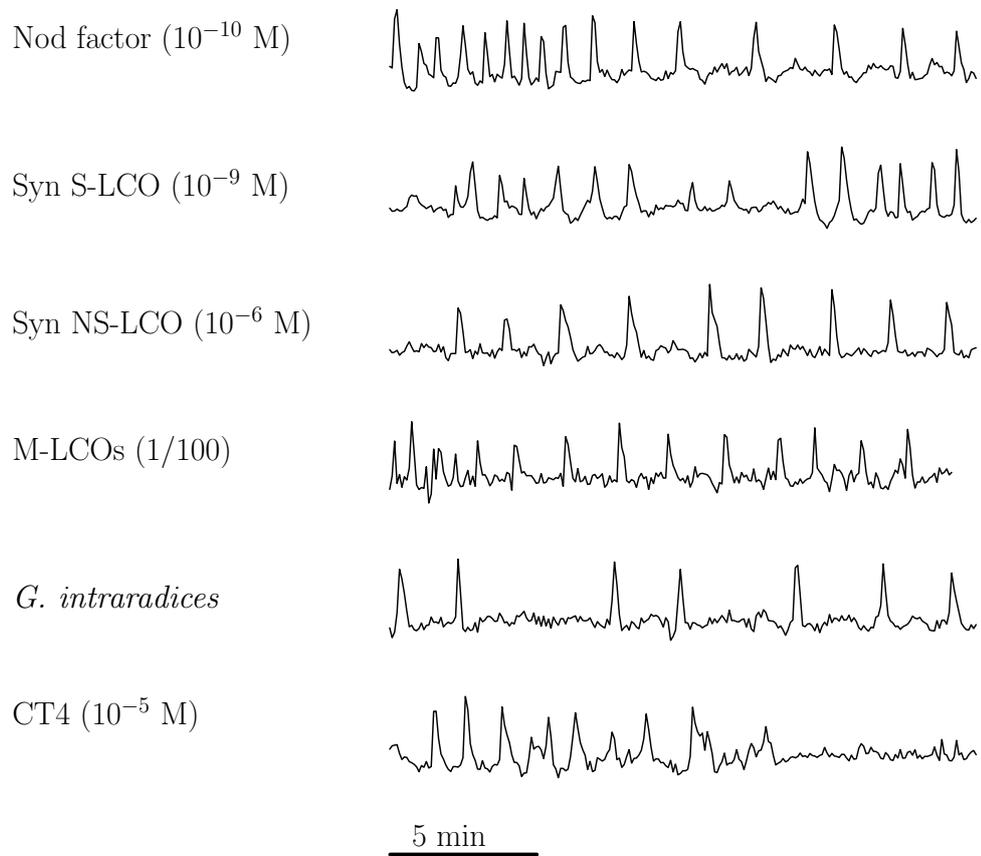


Figure 5.1: Examples of Ca^{2+} spiking induced by different signal molecules.

with previous reports that chitin oligomers induce weaker Ca^{2+} spiking (Walker et al., 2000).

The mean interspike intervals were compared to the standard deviations in the individual traces, and this is shown in Figure 5.4 for the higher concentration level of Nod factor, Syn S-LCO and Syn NS-LCO. This plot shows a correlation, where time series with a higher mean interspike interval are also more likely to have a higher standard deviation. This is interesting because it has been suggested in the animal field of Ca^{2+} oscillations that when this correlation is found, it can be taken as a sign that the spikes are stochastic in their timing (Falcke, 2003b). The symbiotic Ca^{2+} oscillations on the other hand have been found to be deterministic (Hazledine et al., 2009). However, the symbiotic Ca^{2+} spiking data is an integration over the whole nuclear area, and it is not clear at this point how individual ion channels behave in plants.

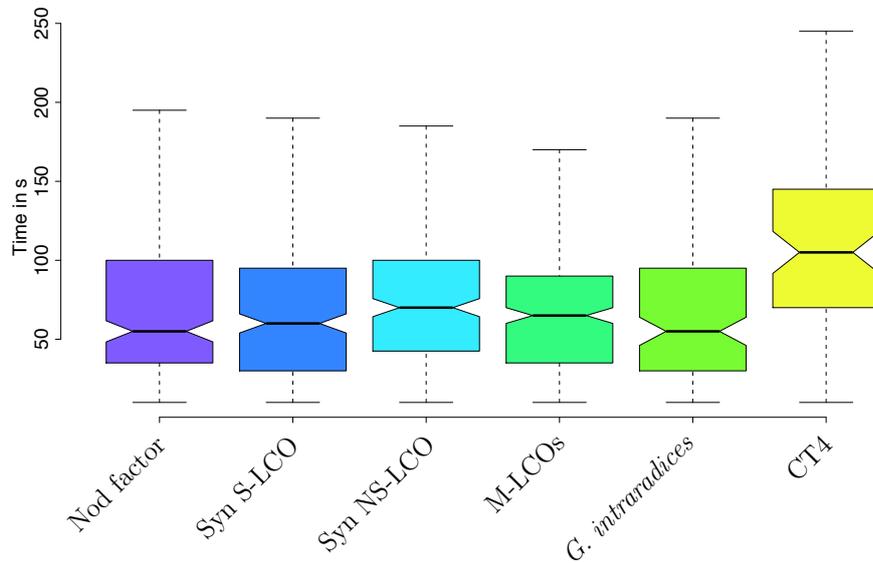


Figure 5.2: Interspike intervals induced by the different treatments, at the concentration level where 50% of cells respond in the case of Nod factor, Syn S-LCO and Syn NS-LCO.

5.2.2 Spike structures

The structures of the Ca^{2+} spikes were also investigated in an effort to approximate their shape. This method is sensitive to noise levels, but it is the best option when working with non-ratiometric data, since we cannot estimate “true” spike shapes. Therefore, and as the time series are detrended with a moving average before analysis starts, the y-axis has no absolute value attached to it. For an overview of the process from raw data to presented data, see Appendix B, Figures 3-4. If the real amplitude value was

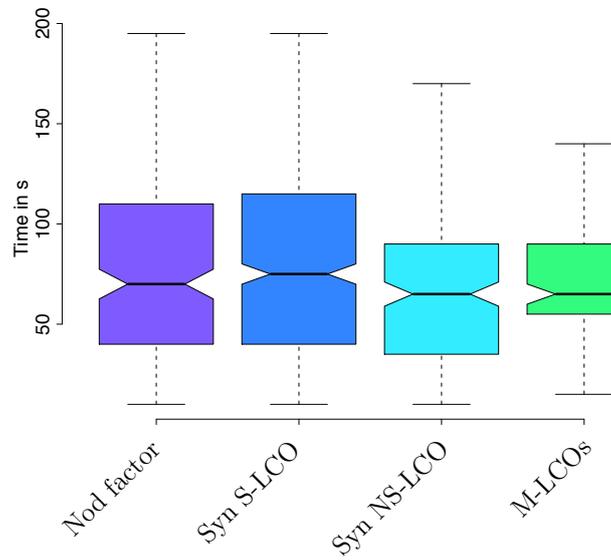


Figure 5.3: Interspike intervals induced by LCOs at the concentration level where 80% of cells respond.

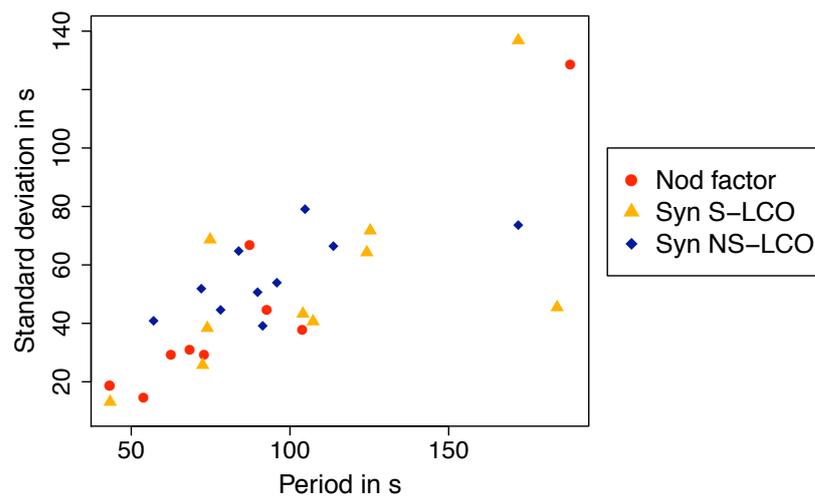


Figure 5.4: The mean interspike intervals for individual time series, versus their standard deviations, at the concentration level where 80% of cells respond.

known, the slope of the upward and downward spike phases would be the natural way of measuring individual spike shapes.

The structure of the spike is measured by the time it takes to reach its maximal height after leaving the baseline, and the time it takes for it to return to baseline from that maximum. The baseline of a Ca^{2+} time series consist of background fluctuations, so a Ca^{2+} spike needs to rise above the standard deviation of the baseline to count as a spike. The results of the spike structure analysis is shown in Figure 5.5. There are very small differences between all groups. The large standard deviations can be explained mainly by two points: 1) the differences in noise levels between traces have an impact on when the spike leaves and comes back to the "baseline", and 2) all the time series have an interval of 5 seconds between data points, which makes the difference between one and two data points rather large when one spike only lasts for approximately 20-25 seconds.

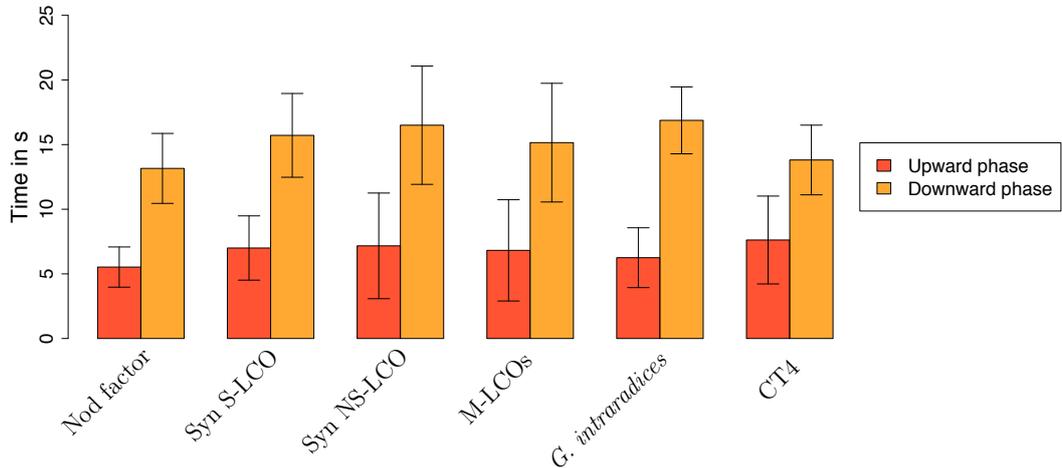


Figure 5.5: Spike shapes induced by the different signal molecules, at the concentration level where 50% of cells respond.

The result that there are no differences in spike structures is contrary to what was found by Kosuta et al. (2008). The difference presented in Kosuta et al. (2008) arose when comparing data from different imaging methods: microinjection of fluorescent dyes compared to the cameleon YC2.1 reporter. This is illustrated in Figure 5.6. The reason for this discrepancy is because different Ca^{2+} reporters have different Ca^{2+} buffer characteristics. The impact of buffers on Ca^{2+} spiking was discussed in Chapter 4, and a similar figure to Figure 5.5 is shown in Figure 4.5 in Chapter 4.

5.2.3 Periodicity

Bayesian Spectrum Analysis (BSA), which was described in detail in Chapter 3, was used to assess the periodicity of the different Ca^{2+} spiking groups. This focuses on

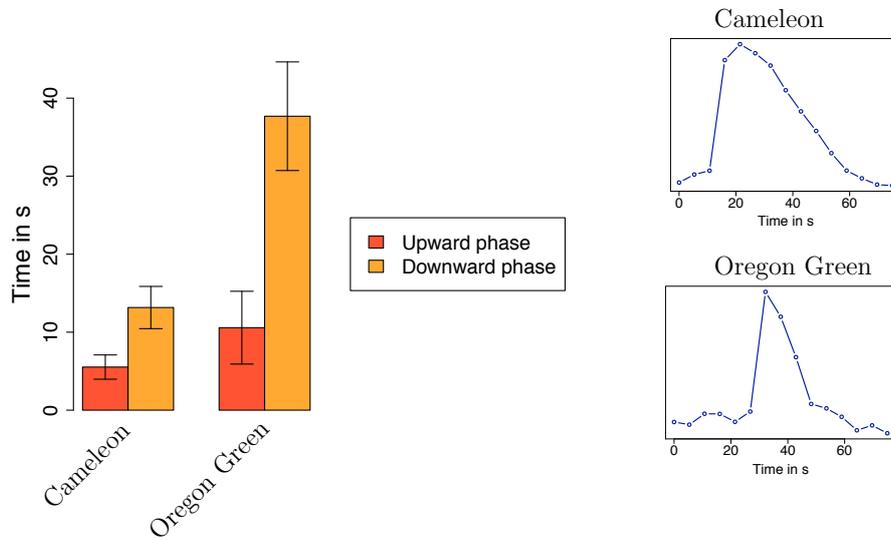


Figure 5.6: Spike shapes induced by the different imaging methods, namely microinjection of Oregon Green and use of a transgenic cameleon reporter, YC2.1.

Ca^{2+} spiking periods that re-occur, while interspike intervals aggregates all periods within a time series. The probability distributions are shown in Figure 5.7, where it can be seen that all isolated LCOs produce distributions of periods with very similar probability. Treatments with the less controlled M-LCOs and fungal induced material show a less strong preference for specific periods. The CT4 treatments induce, as with the interspike intervals, longer periods as well a less strong preference compared to the isolated LCOs. The statistical comparisons confirm these observations using as measures the Kullback-Leibler divergence, and the Kolmogorov-Smirnov test, and can be found in Appendix B, Table 3-4.

5.3 Discussion

Overall, no significant differences could be seen in the Ca^{2+} spiking behaviour between the synthetic mycorrhizal LCOs and Nod factor. The main difference appears to be in the activity of the signal molecule, rather than in the spiking profile itself: Syn NS-LCO required much higher concentrations to activate Ca^{2+} spiking than Nod factor or Syn S-LCO in *M. truncatula*. This could be a result of the specificity of the LCO receptors and their preferred LCO ligand, which will vary between plant species. This specificity can be seen clearly in the case of Nod factors, where a specific Nod factor structure is often recognised, and the AM fungi probably produces a variety of LCOs. It is known

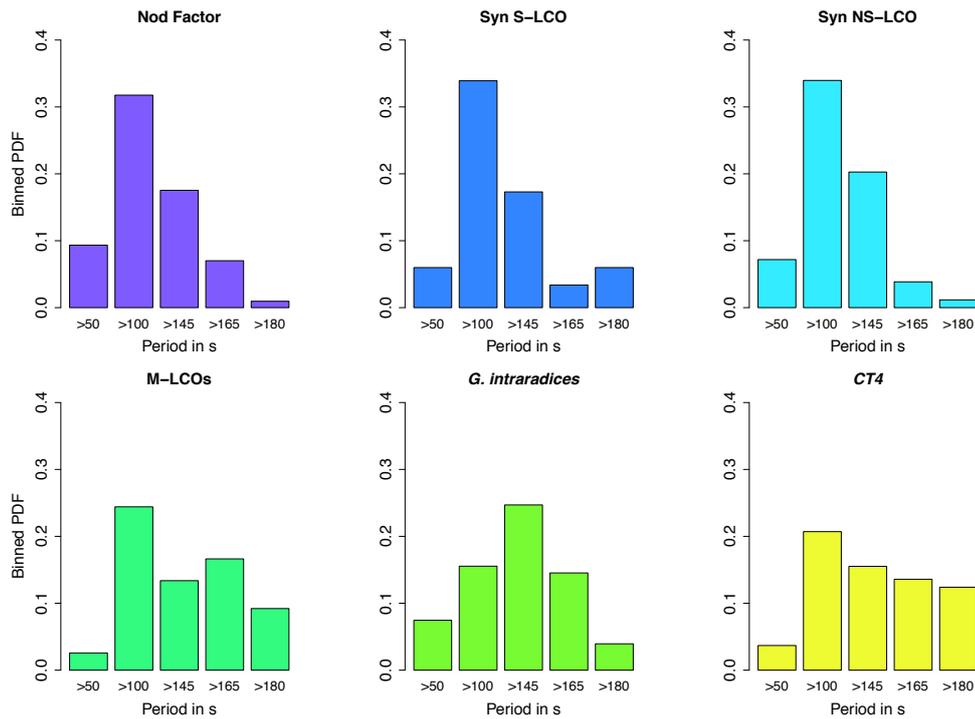


Figure 5.7: Bayesian Spectrum Analysis of the Ca^{2+} spiking, at the concentration level where 50% of cells respond. The y-axis gives the binned probability density function (PDF).

that there is a dual receptor complex on the epidermal surface of *M. truncatula*, and one theory could be that many receptors might work co-operatively on the cell surface, generating multiple signals that are transduced into the cell. Ca^{2+} spiking would then be one of several distinct signal transduction pathways arising from stimulation of LCO receptors.

The fungal data are difficult to compare to the data for synthetic factors, because they come from much less controlled experiments. However, no difference could be seen in spike structure. A difference was detected in overall periodicity (with Bayesian Spectrum Analysis), but this may be a result of the experimental differences. There is a much higher level of uncertainty in the fungal data because of the technical problems associated with growing the fungi in association with the plants, and one can never be certain of the start time of the oscillations.

There may also be other diffusible fungal signals. The method of identifying the Myc-LCOs in Maillet et al. (2011) used an approach specifically targeted towards the isolation of LCOs. This was based on the hypothesis that the rhizobia might have mimicked the original fungal symbiotic signal. However, there may be additional fungal signals, which, together with LCOs, prepare the plant root for AM infection.

Interestingly, the LCOs are still independently recognised, based on data from the Oldroyd laboratory on gene expression profiles. There is little overlap shown between gene expression profiles induced by Nod factor, Syn NS-LCO and Syn S-LCO (Sun et

al., under review). But this information does not seem to come from different Ca^{2+} signals, based on the data presented in this chapter. Furthermore, the gene expression data does show that Syn S-LCO is activating *NIN*, as does Nod factor. This fits with the more similar structure of these two signal molecules, as well as their similarity in Ca^{2+} spiking induction. *MSBP1* is the AM specific marker, and is induced by NS-LCO. Both *NIN* and *MSBP1* induction by the respective symbionts require *DMI3*. SYM signalling mutants were also tested for Ca^{2+} spiking by Jongho Sun, showing a dependence of all LCOs on the genes *NFP*, *DMI1* and *DMI2*, but not *DMI3* which is downstream of the Ca^{2+} spiking.

An aspect of this LCO study is that it only looks at trichoblasts. In *M. truncatula*, rhizobia enter through trichoblasts but AM infection goes through epidermal cells. Perhaps these signals are perceived in a very cell-type specific manner. A recent study has used a nuclear-localisedameleon to monitor even cortical cells, with confocal microscopy (Sieberer et al., 2011). Thus, cell-specific information is becoming more and more available, and it will be interesting to see whether any cellular differences can be seen.

It is clear that the results on the fungal-induced Ca^{2+} spiking data obtained in the present study do not agree with earlier studies (Kosuta et al., 2008), and there are several possible reasons for this discrepancy. Firstly, the fungal experiments and the comparison to LCO-induced spiking must be taken with caution. The fungal experiments were less controlled and therefore it is unknown when the Ca^{2+} spiking starts, which can significantly impact the analysis of periodicity and interspike intervals. Secondly, microinjection traces were compared to camelon YC2.1 traces in the previous studies. Spike shape and the connection to buffers and binding kinetics was discussed in Chapter 4.

This study has not investigated Ca^{2+} influx, another and separate Ca^{2+} response to Nod factors. Ca^{2+} influx has been shown to be unnecessary for nodule organogenesis, but it is apparently required for infection. Compared to Ca^{2+} spiking, a more stringent Nod factor structure is needed to activate Ca^{2+} influx, as well as higher concentrations of Nod factors. The high requirement for a specific Nod factor structure is consistent with chitin oligomers not being able to induce this response. In the light of this, it would be important to test whether mycorrhizal LCOs can induce Ca^{2+} influx, and at what concentration levels.

It is important to note the high level of variability within and between Ca^{2+} traces. There are, as we have seen in Chapter 4, many possible reasons for this variability, for example cell size, buffer concentrations and binding kinetics. Experiments have also confirmed an intrinsic axis of variability along the root, with cells closer to the tip oscillating faster (Miwa et al., 2006). This was suggested to be connected to the developmental status or size of the cell. Indeed, based on the hormone influences known to impact on Ca^{2+} spiking, hormone differences could potentially cause this result.

Could Ca^{2+} signalling still be a conveyor of specificity in the SYM pathway despite

all the variability between cells and the observed similarity in the responses to different LCOs? Perhaps a more likely explanation may be that there are parallel active pathways that modulate the cell to prepare for the inducing symbiont and that these parallel pathways may differ for rhizobial and AM infection.

Chapter 6

Defining the patterns of symbiotic calcium oscillations in diverse species in the nitrogen-fixing clade

6.1 Introduction

Arbuscular mycorrhiza are believed to have associated with very early land plants (Remy et al., 1994; Redecker et al., 2000a): this is therefore a much older symbiosis than that between legumes and rhizobia. In the previous chapter, the patterns of Ca^{2+} oscillations induced by the signal molecules during these two symbioses were compared, and found to be very similar. However, this comparison was performed with only one plant species, the model legume *M. truncatula*. According to the hypothesis that the pattern of Ca^{2+} oscillations encodes information regarding the specificity of the endosymbiont, the question arises whether the evolution of the root nodule symbioses could have been associated with changes in Ca^{2+} patterns as new plant species evolved. Addressing this question requires a broader comparison of Ca^{2+} spiking patterns in symbiotic plant species, and this will be the focus of the current chapter.

As angiosperms diversified, the symbiosis with nitrogen-fixing soil bacteria appeared: this was associated with the development of specialised structures termed nodules on the plant roots. Today this root nodule symbiosis is found in many plant families, but they are all confined to the nitrogen-fixing clade, within the Eurosid I clade (Soltis et al., 1995). This has led to a hypothesis that the most recent common ancestor of the nitrogen-fixing clade harboured traits that predisposed this group of plants for root nodule symbiosis, facilitating the emergence of nodulation in later species (Soltis et al., 1995; Doyle, 2011).

Within the nitrogen-fixing clade, two different types of root nodule symbiosis can be distinguished. The first is an association between plants and filamentous gram-positive

Frankia bacteria, and is known as the actinorhizal symbiosis (Stacey et al., 1992). The second type of symbiosis is that between the legume family, *Fabaceae*, and single-celled gram-negative rhizobia (Sprent, 2009). The only known non-legume to have a rhizobial symbiosis is the *Parasponia* genus in the *Rosales* (Trinick, 1973, 1979).

Ca^{2+} oscillations have been measured in different legume plant species in response to Nod factors, for example *Medicago truncatula* (Wais et al., 2000), *Pisum sativum* (Walker et al., 2000), *Phaseolus vulgaris* (Cárdenas et al., 1999) and *Lotus japonicus* (Harris et al., 2003). The oscillations have also been recorded in the non-legume *Daucus carota* in response to AM exudates (Chabaud et al., 2011). Furthermore, in Chapter 5 of this thesis, the Ca^{2+} oscillations were presented after induction by the recently identified AM equivalent of Nod factors (Maillet et al., 2011). With the exception of carrot, these plants are all legumes and, moreover, they are all relatively closely related, belonging to the same legume subfamily. There may be several disadvantages of focusing so closely on a few species, especially because of the many dissimilarities among the model legumes and other legumes that have been studied (Sprent and James, 2008).

Also, most studies have involved legume species with a specific type of rhizobial infection strategy, namely root hair infection (RHI). This was shown schematically in Chapter 1, Figure 1.5. Other infection processes in legumes appear simpler than the root hair infection process, and are therefore possibly more ancient in evolutionary terms (Sprent, 2009). Additionally, it has been shown that the aquatic legume *Sesbania rostrata* has two modes of infection, and these correlated with two different signatures of Ca^{2+} spiking (Capoen et al., 2009). This provides further motivation to extend the study of Ca^{2+} spiking patterns to species with different infection strategies.

Here I will present a study of Ca^{2+} oscillations covering more distantly related legumes and non-legumes compared to what has been previously examined, as well as an investigation of Ca^{2+} responses in *Parasponia andersonii* and its close relative *Trema tomentosa*, which is a non-nodulator. The chapter will investigate whether Ca^{2+} spiking is a conserved part of the common SYM pathway in all the symbiotic species, and whether any variations of Ca^{2+} signatures could be correlated to evolutionary relationships or infection strategies.

6.1.1 The legume family and symbiosis

The Fabaceae is the third largest family of flowering plants, and probably evolved in the late Cretaceous period (60-70 MYA). It belongs to the Fabales order in the nitrogen-fixing clade. It is unclear where legumes originated geographically, and this is subject to much debate, but today legumes are found on all continents apart from Antarctica (Sprent, 2009). There are three subfamilies, the most basal is the Caesalpinioideae, then the Mimosoideae and finally the Papilionoideae. Non-nodulating plant species form the majority in the Caesalpinioideae and the symbiosis with rhizobia is far more common in the other two families (Sprent, 2007). The subfamily Papilionoideae has

nodulators at its base, and thus nodulation was probably present from the start of this group's divergence (Doyle and Luckow, 2003).

6.1.2 The *Parasponia* case

The *Parasponia* genus consists of small tropical trees from Papua New Guinea, and they are often found as pioneers in ash-covered soil. This genus is the only non-legume found that forms a symbiosis with rhizobia leading to nodule formation (Stacey et al., 1992). The *Rhizobium* strain and the ability of the *Parasponia* nodules to fix nitrogen was first found in the 1970s (Trinick, 1973), and since then this relationship has been studied sporadically (see for example Trinick (1979); Marvel et al. (1987); Davey et al. (1993); Lafay et al. (2006)). *Parasponia* is placed in the monophyletic family Cannabaceae (Sytsma et al., 2002) and its close sister-genera, *Trema*, does not have the root nodule symbiosis (Akkermans et al., 1978).

Parasponia nodules closely resemble legume nodules (Trinick, 1979), with rhizobia entering intercellularly through the epidermis to the cortex, and the infection threads growing intracellularly at a later stage of infection (Lancelle and Torrey, 1984). Both *Rhizobium* and *Bradyrhizobium* nodulates *Parasponia*, but efficiency can vary widely depending on the symbiotic partner (Op den Camp et al., 2012).

6.1.3 Actinorhizal symbioses

Some plants are nodulated by *Frankia* bacteria, and these are referred to as actinorhizal species, including 24 plant genera in eight families (Stacey et al., 1992). The nitrogen-fixing clade contains four orders: the Fabales, Fagales, Rosales and Cucurbitales. Actinorhizal plants are found in all of these orders and these symbioses are therefore more distantly related to each other than to the legume-rhizobial symbiosis, which is confined to a single family. It is believed that actinorhizal nodulation has arisen several times within the nitrogen-fixing clade (Swensen, 1996; Swensen and Mullin, 1997). Many actinorhizal plants are found in nutrient-poor soil around the world, and many have been utilised for agriculture and forestry.

The *Frankia* genus consist of multicellular, filamentous bacteria that belong to the order Actinomycetales. Three major strains of *Frankia* are often recognised, based on their ability to infect different groups of plants (Baker, 1987), and sometimes these are further divided into two groups representing those that enter plants via inter- or intracellular infection. Most *Frankia* can survive outside the plant as vegetative mycelia present in the soil where they possess vesicles involved in N₂-fixation (Stacey et al., 1992; Pawlowski and Bisseling, 1996).

The tree *Alnus glutinosa* belongs to the *Betulaceae* in the order Fagales, and is an important coloniser in many ecosystems because of its capacity to form nitrogen-fixing symbioses. *Frankia* invade *Alnus* species through a root hair infection process (Wall, 2000). It was recently shown that Ca²⁺ spiking occurs in *Alnus glutinosa* after

application of an unidentified *Frankia* signalling molecule present in bacterial exudates (J. Sun, unpublished data), which are referred to as “Frankia” factors in line with Nod factors. This data will be included in my Ca^{2+} spiking study, as a representative of an actinorhizal plant response.

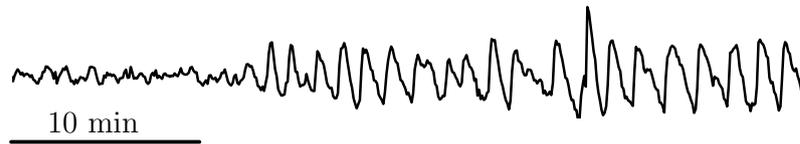


Figure 6.1: Ca^{2+} spiking in *Alnus glutinosa* in response to an unidentified Frankia factor present in extracts from *Frankia*. The experiments were performed by Jongho Sun in the Oldroyd laboratory. Root hair cells of *A. glutinosa* were imaged following microinjection with Oregon Green.

6.2 Results

Five legume species that have not been tested for symbiotic Ca^{2+} spiking were chosen for this study. The selection of species focused on species that could inform us of the evolution of Ca^{2+} spiking, as well as species with distinct infection strategies that may further a previous hypothesis that infection strategy may be associated with the modality of Ca^{2+} spiking (Capoen et al., 2009). New data was also collected on *P. andersonii* and its close relative *T. tomentosa*. These are compared to *M. truncatula*, *S. rostrata* (Capoen et al., 2009), and *A. glutinosa* Ca^{2+} spiking (J. Sun, unpublished work). For a phylogenetic overview of the plant species, see Figure 6.2.

6.2.1 Choice of plant species to study

The five legume species were chosen not only because they represent different groups of legumes, but also because they have a variety of different rhizobial infection strategies. In previous studied species such as the model legumes *M. truncatula* and *L. japonicus* the infection occurs intracellularly through the root hair cell, referred to as root hair infection (RHI). In contrast, the aquatic legume *S. rostrata* can undergo intercellular infection through cracks in the root epidermis or via root hairs according to the environment. These three examples all belong to the most recently evolved legume subfamily, Papilionoideae.

Caesalpinioideae is the most basal legume subfamily, originating around 60 MYA at a point very close to the origin of legumes. Most members of this group do not nodulate (Sprent, 2009). The tree *Cercis siliquastrum* is investigated as a representative of a

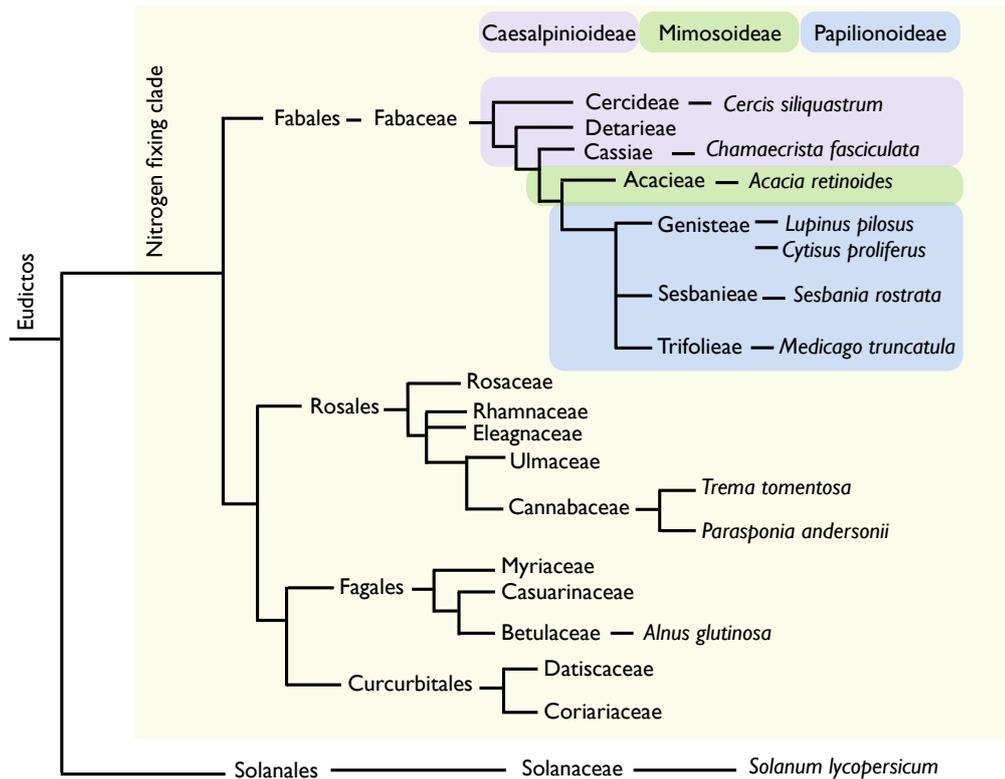


Figure 6.2: Phylogenetic tree of the nitrogen-fixing clade. The tree is a simplification of the clade, and provides an overview of the relationships between the plant species included in the present study. The three subfamilies of the legume family Fabaceae are indicated with coloured boxes. The tree is unscaled, and tomato, *S. lycopersicum*, is included as an outgroup which has also shown Ca^{2+} spiking (Sun et al., under review). *S. lycopersicum* belongs to the Asterids, while the nitrogen-fixing clade is found among the Rosids.

non-nodulator, and belongs to a basal branch in the Caesalpinioideae. In comparison *Chamaecrista fasciculata* is included in the study as a nodulating caesalpinoid. *C. fasciculata* is herbaceous with easy growth conditions that is rapidly emerging as a model basal legume (Singer et al., 2009; Naisbitt et al., 1992). The infection process for *C. fasciculata* has been suggested to be an intermediate between RHI and crack infection (Hirsch et al., 2009a).

In the Mimosoideae subfamily, *Acacia retinoides* was tested for Ca^{2+} spiking. *Acacia* is a large, paraphyletic group that can be found worldwide and harbours both nodulators and non-nodulators. However, some of the non-nodulators are thought to have lost the capacity relatively recently in their evolutionary history (Sprent, 2009). This group also has extensive mycorrhizal associations, with many species forming both AM and ectomycorrhizal (ECM) symbioses (Sprent, 2009). Previously investigated *Acacia* species have shown RHI (Räsänen et al., 2001).

In the more well-studied subfamily Papilionoideae two legumes were tested, *Cytisus proliferus* and *Lupinus pilosus*. *C. proliferus* is a woody plant that forms symbioses with *Bradyrhizobia*. The infection process starts with RHI, which then aborts and the bacteria finally enter via crack-entry, with no infection threads formed (Vega-Hernandez et al., 2001).

Lupinus forms a unique type of nodule after infection by *Bradyrhizobia*, and the infection process leading up to it is also unusual. The rhizobia enter intercellularly, but not via the classical crack-infection since entry points are uncorrelated with lateral roots. Instead, the bacteria seem to enter between the base of root hair cells and epidermal cells (González-Sama et al., 2004). Infection threads are only found in young nodules (Novikova and Gordienko, 2001), and the nodule is a mix between the two classical morphologies of determinate and indeterminate nodules (González-Sama et al., 2004).

6.2.2 Induction of spiking with LCOs

Sinorhizobium NGR234

To induce Ca^{2+} spiking with Nod factors in several plant species the *Sinorhizobium* strain NGR234 (Saldaña et al., 2003) was chosen. It is a broad host-range rhizobia that can nodulate over 70 legume genera and, in addition, *Parasponia* (Pueppke and Broughton, 1999). A crude isolation of the NGR234 Nod factors was purified using published methods (Price et al., 1992), and the activity was tested in a root hair deformation assay on the legume *Vicia hirsuta*, which is known to perceive a variety of Nod factors (Schultze et al., 1992). A root hair cell that responds to Nod factors changes growth direction and starts curling around the group of rhizobia in the soil (Oldroyd and Downie, 2008). This behaviour can be seen by branching, zig-zagging and curling of root hairs 6-12 hours after addition of Nod factors or bacteria to the surrounding media. The assay showed that the isolated NGR234 Nod factors were very

active at 1/100 dilutions of the mixture, see Figure 2 in Appendix B. Therefore, this is the dilution that was used in the experiments, and it will be referred to from now on as NGR NFs.

Mycorrhizal LCOs

To complement the Nod factor experiments, some plants were also tested for their response to mycorrhizal LCOs. For this, some of the recently described Myc factors from *Glomus intraradices* were used, namely the non-sulphated species (Maillet et al., 2011). These were used at a concentration of 10^{-5} M, which has been shown to give a high proportion of Ca^{2+} spiking in both legumes and non-legumes (Sun et al., under review) and was described in Chapter 5. This Myc factor solution will be referred to as NS-LCO in the rest of this chapter.

6.2.3 Calcium oscillations in distantly related legumes

The five legumes *C. siliquastrum*, *C. fasciculata*, *A. retinoides*, *C. proliferus* and *L. pilosus* were challenged with NGR NF after microinjection of growing root hairs. All of them apart from *C. siliquastrum* have been reported to form nodules in response to NGR234 (Pueppke and Broughton, 1999). All four symbiotic species exhibited Ca^{2+} oscillations in response to NGR NF, while the non-nodulator *C. siliquastrum* did not. This highlights a correlation with the ability to nodulate and the ability to activate the signalling pathway with LCOs. Representative Ca^{2+} responses are shown in Figure 6.3 and the results are listed in Table 6.1. The nature of the response was equal to what has been reported in other legumes such as *M. truncatula* and *L. japonicus*, in that the responses were activated 10-20 min after Nod factor addition and the oscillations can continue for at least an hour (Ehrhardt et al., 1996). The oscillations were also localised to the nuclear area of the root hair cell, see Figure 6.4, which is also the same for the previous studies on other legumes (Capoen et al., 2009, 2011).

Table 6.1: Cells showing Ca^{2+} spiking after microinjection

Plant species	Spiking cells
<i>Parasponia andersonii</i>	2/3
<i>Cercis siliquastrum</i>	0/15
<i>Chamaecrista fasciculata</i>	13/20
<i>Acacia retinoides</i>	13/17
<i>Lupinus pilosus</i>	11/12
<i>Cytisus proliferus</i>	12/17

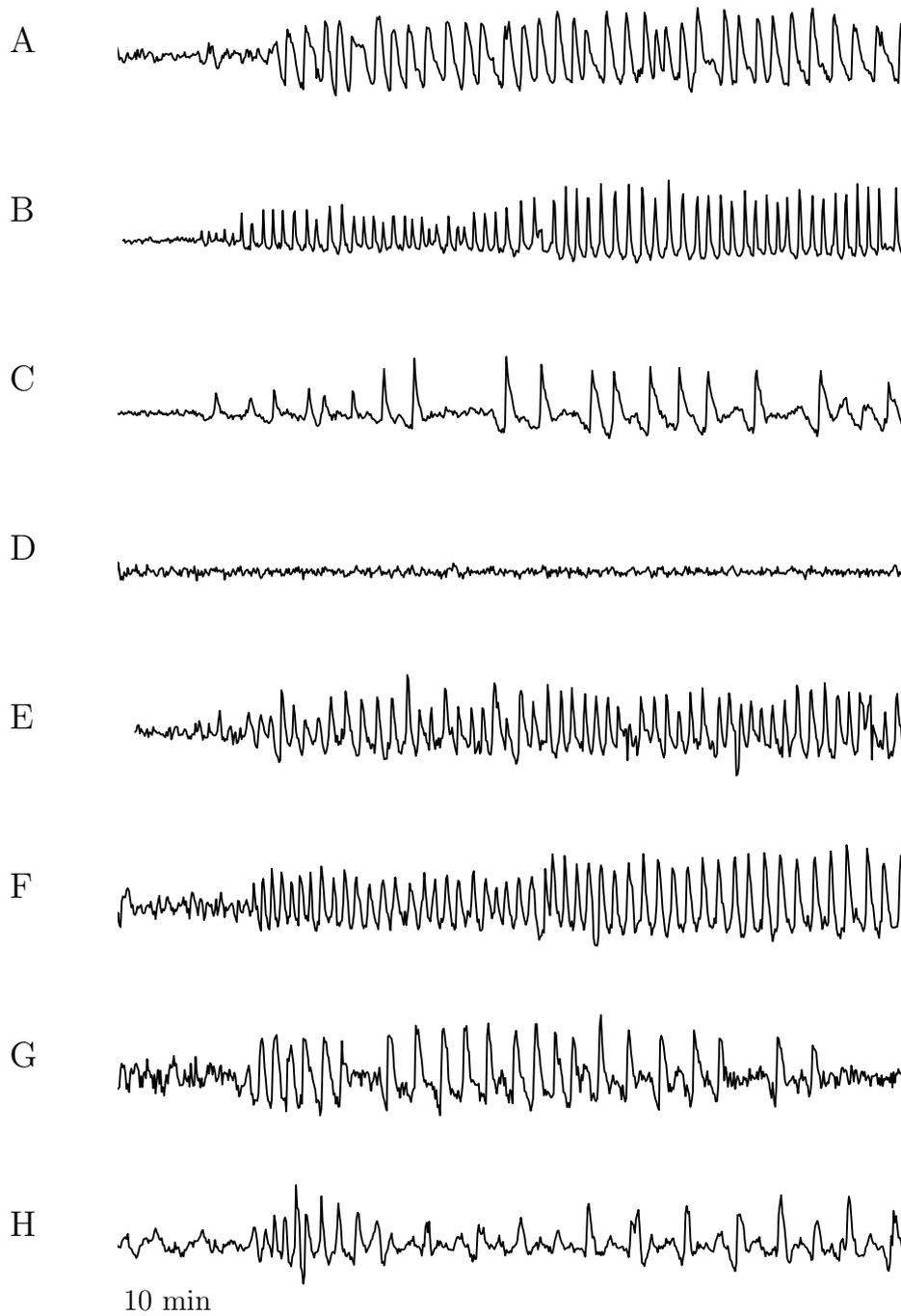


Figure 6.3: Ca^{2+} spiking in legumes induced by Nod factors **A:** *Medicago truncatula* and *S. meliloti* NFs. **B:** *Sesbania rostrata* crack infection and *A. caulinodans* NFs. **C:** *Sesbania rostrata* root hair infection and *A. caulinodans* NFs. **D:** *Cercis siliquastrum* and NGR NFs. **E:** *Chamaecrista fasciculata* and NGR NFs. **F:** *Acacia retinoides* and NGR NFs. **G:** *Lupinus pilosus* and NGR NFs. **H:** *Cytisus proliferus* and NGR NFs. See Table 6.1 for an overview of the proportions of cells responding.

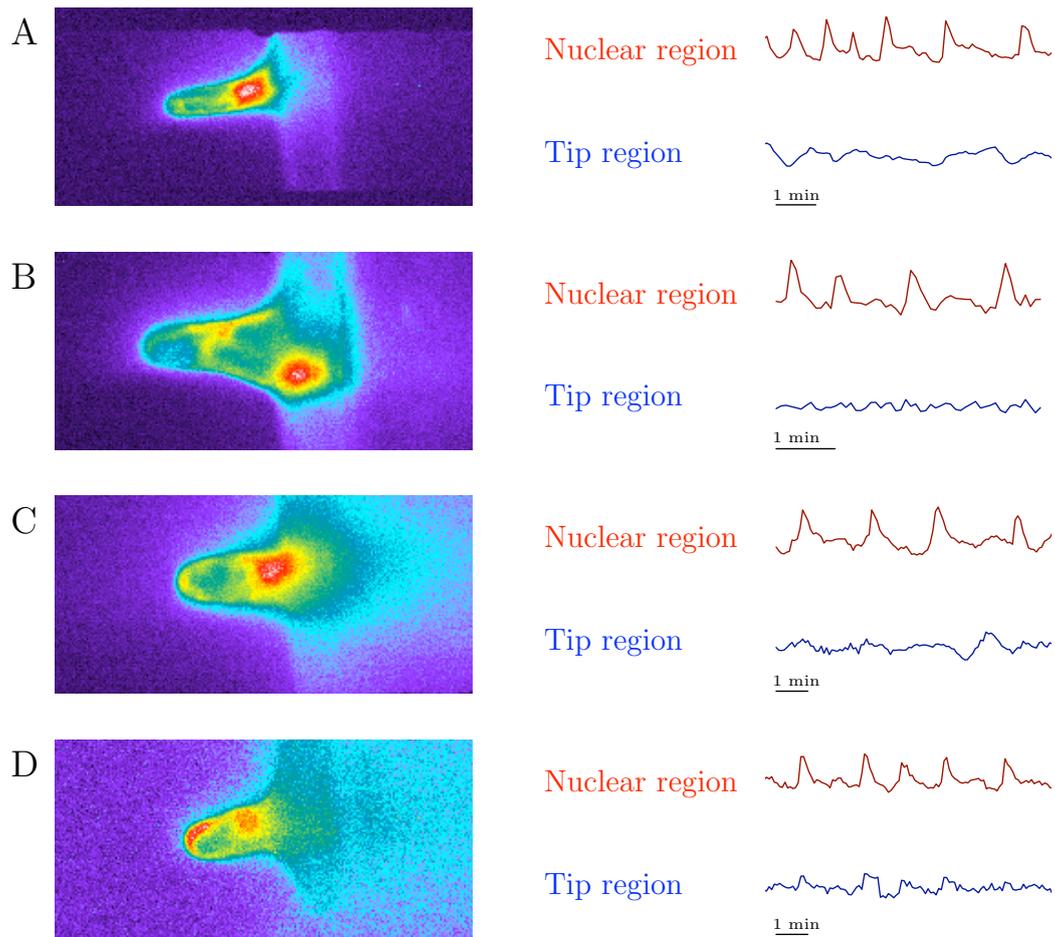


Figure 6.4: Nuclear localised spiking in the four legumes species: **A:** *Lupinus pilosus*. **B:** *Chamaecrista fasciculata*. **C:** *Cytisus proliferus*. **D:** *Acacia retinoides*.

6.2.4 Calcium oscillations in the non-legume genera *Parasponia* and *Trema*

P. andersonii is a tropical tree which poses some difficulties in cultivation, and therefore plantlets from vegetative propagation were used in this study. This was done in collaboration with René Geurts' group at Wageningen University.

P. andersonii showed Ca^{2+} oscillations in root zone 1 in response to NGR NF after injection of root hairs, see Figure 6.5. These responses appear equivalent to the Ca^{2+} oscillations observed in legumes (Figure 6.3). It was technically challenging to inject the *P. andersonii* roots, and in order to gather more data, *P. andersonii* and *T. tomentosa* plantlets were transformed with the nuclear localisedameleon Ca^{2+} reporter NupYC2.1 (Sieberer et al., 2009). The successful nuclear localisation of the construct is shown in Figure 6.6A-B. When challenged with NGR NF, the *P. andersonii* epidermal cells showed Ca^{2+} oscillations in 15/23 cells (Figure 6.6C, Table 6.2). We also tested NS-LCO, and this signal induced Ca^{2+} oscillations in *P. andersonii* in 11/18 cells (Figure 6.6D, Table 6.2).

When testing *T. tomentosa* for Ca^{2+} spiking, there were no observed Ca^{2+} oscillations in 0/35 cells imaged after challenging them with NGR NF (Figure 6.6E, Table 6.2). However, when challenged with NS-LCO at the same concentration as *P. andersonii* 7/20 cells responded (Figure 6.6F, Table 6.2). This reveals that although *T. tomentosa* does not respond to NGR NF, it has the capability to recognise LCOs and activate Ca^{2+} oscillations.



Figure 6.5: *Parasponia andersonii* oscillations in a root hair cell after induction by NGR NFs.

Table 6.2: Epidermal root cells showing Ca^{2+} spiking in *Parasponia* and *Trema* NupYC2.1 plants in response to LCOs.

Plant species	NGR NF	NS-LCO
<i>Parasponia andersonii</i>	15/23	11/18
<i>Trema tomentosa</i>	0/35	7/20

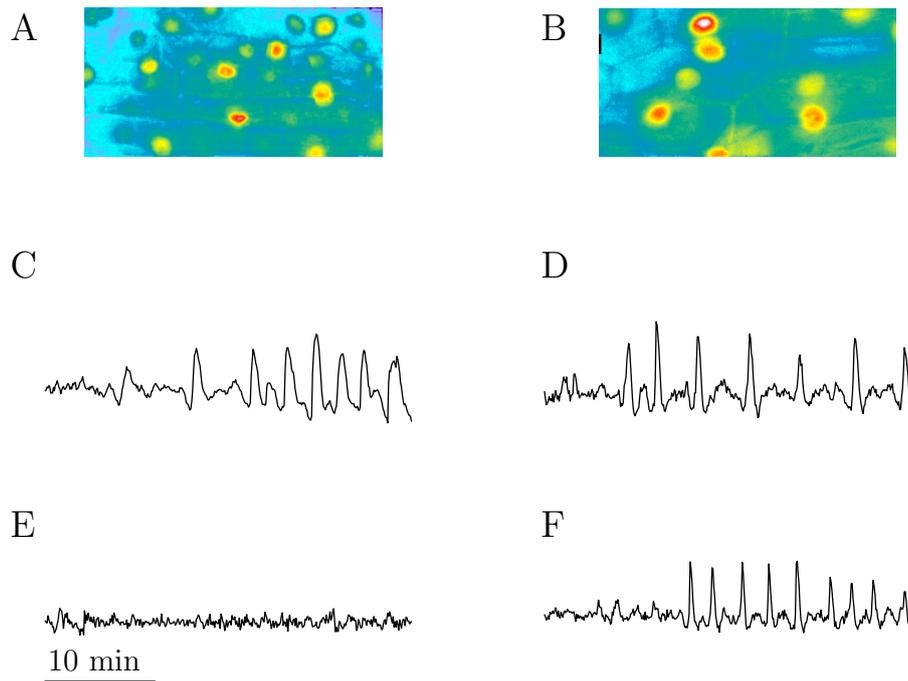


Figure 6.6: Response of epidermal cells of *Parasponia* and *Trema* to LCOs. **A:** *Parasponia andersonii* with nuclear CFP signal. **B:** *Trema tomentosa* with nuclear CFP signal. **C:** *Parasponia andersonii* and NGR Nod factors. **D:** *Parasponia andersonii* and mycorrhizal NS-LCO. **E:** *Trema tomentosa* and NGR Nod factors. **F:** *Trema tomentosa* and mycorrhizal NS-LCO. For an overview of the proportions of cells spiking, see Table 6.2.

6.2.5 Data analysis of species-specific spiking

Ca²⁺ spiking recorded with microinjection of Ca²⁺-responsive dyes was analysed for comparability between the different plant species. Ca²⁺ traces from *S. rostrata* published previously (Capoen et al., 2009) were also included in this analysis. This plant can exhibit both intercellular crack infection by rhizobia, or the more well-studied RHI. The latter occurs when *Sesbania* is grown in aerated conditions, where it develops root hairs. Under water-logged conditions, the bacteria instead enter through cracks in the epidermis at lateral root emergence sites (Goormachtig et al., 2004). The crack infection was shown to be correlated with faster and sharper Ca²⁺ spikes than the RHI (Capoen et al., 2009).

Data on *M. truncatula* oscillations are also included as well as unpublished data on *A. glutinosa* (J. Sun, unpublished data). For each plant species, 5-10 high-quality time series were used, with the only exception being *P. andersonii*, where it has only been possible to obtain one high-quality trace. For interspike intervals and spike structures, the time series were detrended before analysis. The Bayesian Spectrum Analysis (BSA, presented in Chapter 3) did not require any detrending or other data processing before the analysis.

Interspike intervals, which is the time between two individual spikes, showed some variation between all groups (Figure 6.7). The median is marked as a black line in the boxplots, and these values differ between all groups. The statistical Mann-Whitney-Wilcoxon tests also conclude that the distributions of interspike intervals are significantly different between almost all groups (See Table 5 in Appendix B). However, it should be noted that many of the distributions largely fall within the already well-known spiking interval for legumes: 1-2 minutes. Also, all plant species had one or two other plants that were not significantly different, creating a chain of overlaps in similarities. The exception from this was *S. rostrata* grown in a crack-entry environment, which was significantly different from all other groups. In this environment this plant has a shorter interspike interval than, for example, *M. truncatula* (Capoen et al., 2009). During root hair infection (RHI) Ca²⁺ oscillations in *S. rostrata* showed longer intervals between spikes, similar to the other species. In contrast, *L. pilosus* had very long interspike intervals, often with 1-2 minute interspike intervals at the start of the oscillations, when other species generally show very rapid spiking, and slower interspike intervals over 2 minutes between spikes after that (Figure 6.3G).

The structures of individual Ca²⁺ oscillations were also investigated. This was done by measuring the time for the spike to reach its maximal height after leaving the baseline, and the time it takes for it to return to baseline from that maximum. This method is sensitive to noise levels, but it is the best option when working with non-ratiometric data, since we cannot estimate “true” spike shapes. The results of the spike structure analysis are shown in Figure 6.8. The bars show the mean and standard deviation for each plant species. Here the two outliers *S. rostrata* and *L. pilosus* are confirmed, each having short- or long-lasting spikes, respectively. *A. retinoides* also showed short spike

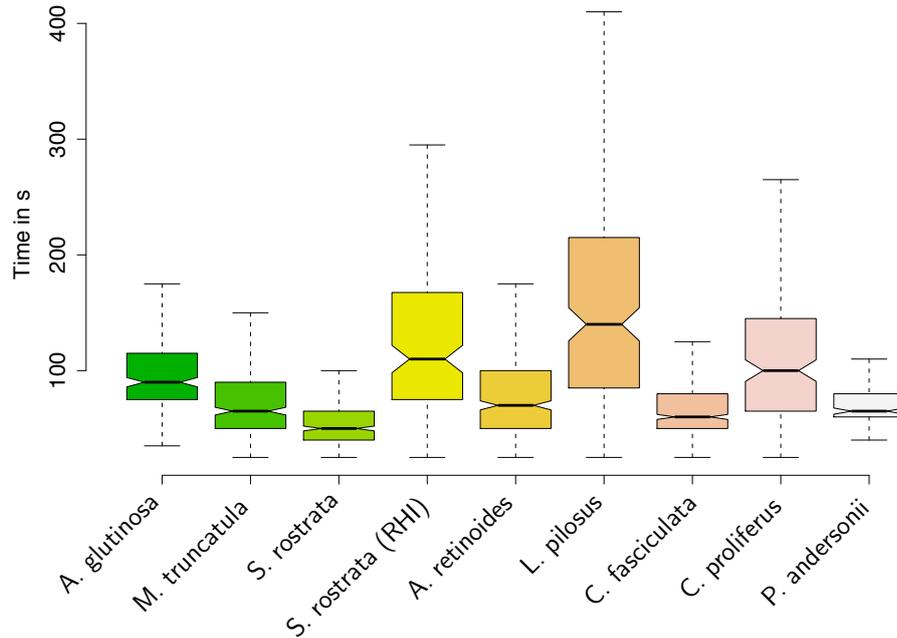


Figure 6.7: Interspike intervals of Ca^{2+} oscillations. The boxplots show the median value with the black line. The box contains 50% of the distribution, and the whiskers show the maximum and minimum value of the distribution, excluding outliers.

durations. Among the other plant species, the spike structures were largely similar. This is confirmed by the Mann-Whitney-Wilcoxon test (Table 6 in Appendix B). There it can be seen that *A. retinoides* and *S. rostrata* are significantly different from all groups except each other on the initial period of an individual transient. *L. pilosus* was also significantly different from all other groups on the initial period of the transients. On the second phase of a Ca^{2+} transient there are more mixed results, but the conclusion that *S. rostrata* and *A. retinoides* have shorter spike durations hold.

The Bayesian Spectrum Analysis is shown in Figure 6.9, where binned posterior density functions are presented for each group. The analysis confirms that the most common period lies in the interval of 1-2 minutes. Some plant species have periodicities of over 2 minutes, which usually occurs as the oscillations are slowing down towards the end of the Ca^{2+} response. This is consistent with previous period analyses of Ca^{2+} traces in *M. truncatula*, also showing a group of interspike intervals in the region around 170 seconds (Miwa et al., 2006). *L. pilosus* have a spread distribution in the BSA analysis, showing a less strong preference for a certain frequency than the other plants. This is consistent with the result presented in Chapter 5: that longer interspike intervals are correlated with a larger standard deviation. *S. rostrata* displayed a shorter period, with intervals of around 60 seconds while other plant species had between 80-100 seconds.

It should be noted that only one trace from *P. andersonii* was used in this analysis.

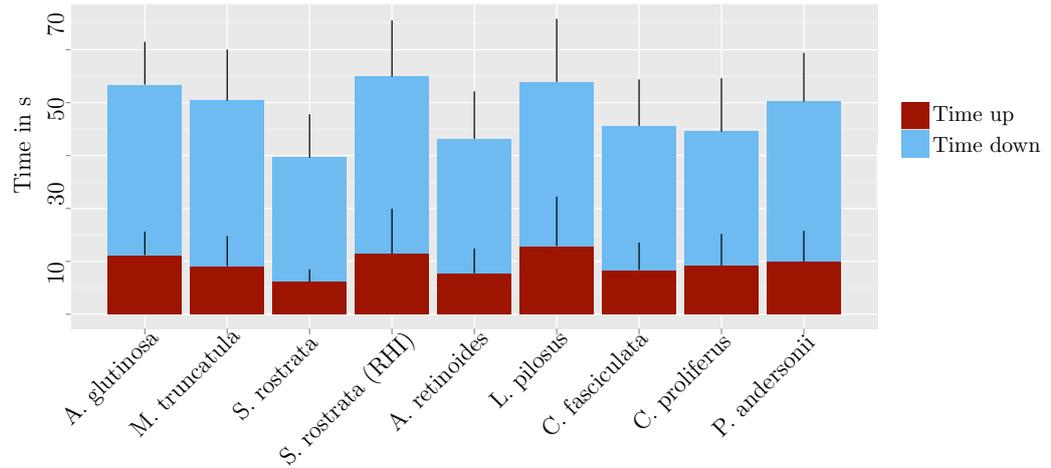


Figure 6.8: Comparison of Ca^{2+} spike structures in the different plant species. The red part shows the time for the spike to reach its maximum height, after leaving the base level. The blue part shows the time for the spike to return to the base level from the maximum height. The bars mark the mean value, and one standard deviation is represented by the black vertical line on the bar.

However, this trace showed no special features compared to the other plant species, and therefore *P. andersonii* is not likely to have a special pattern of Ca^{2+} oscillations.

In summary, there are some differences in the patterns of Ca^{2+} oscillations between the plant species investigated. The most extreme differences are found in *L. pilosus* and *S. rostrata*. But strikingly, plant species that are very distantly related in the nitrogen-fixing clade, such as *M. truncatula* (model legume), *C. fasciculata* (basal legume), *P. andersonii* (non-legume), and *A. glutinosa* (actinorhizal plant), all have very similar Ca^{2+} oscillations in response to symbiotic signals. The differences therefore appear unrelated to the evolution of the nitrogen-fixing clade; no pattern of changes in the Ca^{2+} oscillations correlates with phylogeny. The infection process, which was correlated to different Ca^{2+} spiking signatures in *S. rostrata*, is not correlated to different Ca^{2+} spiking signatures when other species were analysed.

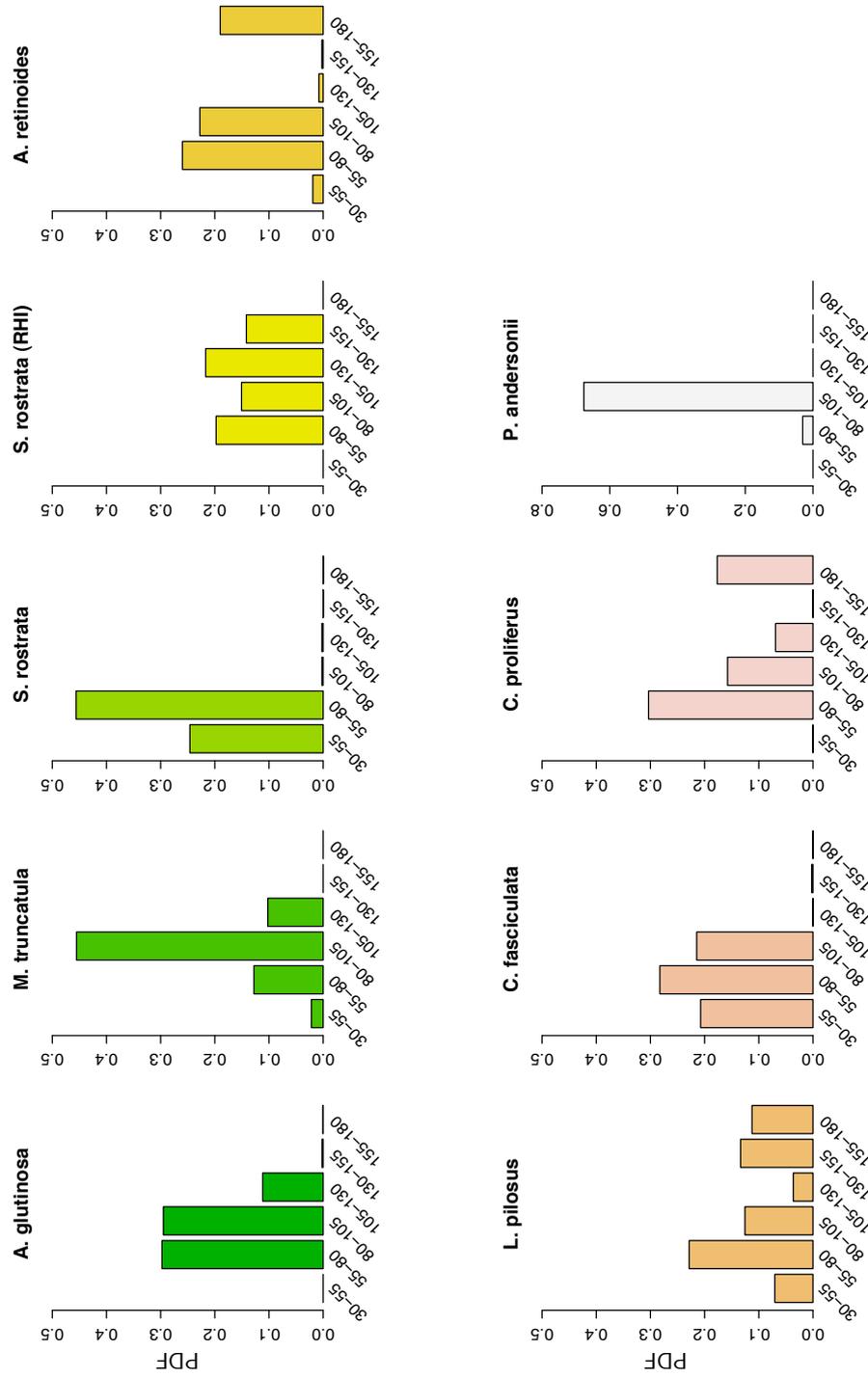


Figure 6.9: Bayesian Spectrum Analysis of Ca^{2+} spiking in different plant species. The plots are binned posterior density functions (PDFs) over the interval 30-180 seconds. Note that the y-axis of *P. andersonii*, at the bottom right, is different from the rest. This is caused by this group containing only one time series, as explained in the main text, so the PDF is less spread.

6.3 Discussion

All plant species that form root nodule symbioses are confined to a monophyletic clade (Soltis et al., 1995). However, the majority of species within this clade do not nodulate. It has been hypothesised that some genetic event predisposed this group of plants to develop root nodule symbioses, and that the symbiosis subsequently evolved independently on several occasions as the group diverged. An alternative view could be that the root nodule symbiosis evolved once in the clade, and that many losses of the symbiotic capacity has occurred as the group diverged. But such a hypothesis seems unlikely since the majority of species in this clade do not nodulate. In either case, the results presented in this chapter show that symbiotic Ca^{2+} oscillations are a conserved part of the symbiosis in the nitrogen-fixing clade present in all groups with the capacity to nodulate, and that no clear pattern of evolution affecting the Ca^{2+} signature has been found.

This is consistent with the fact that many common SYM genes, among them those directly involved in the production of the Ca^{2+} spiking, have been shown to be highly conserved both inside the nitrogen-fixing clade and outside (see for example Banba et al. (2008), Chen et al. (2009), Yokota et al. (2010), Hocher et al. (2011)). As the nodulation signalling pathway probably adapted from an arbuscular mycorrhizal signalling pathway, it is logical that the Ca^{2+} oscillations formed an indispensable part of the nodulation pathway.

Even though SYM components have been conserved, there must be a difference within the nitrogen-fixing clade, since other plants can only form symbioses with AM. Indeed, orthologues of common SYM genes from rice can sometimes complement for nodule formation in *M. truncatula* mutants, but not for rhizobial infection (Chen et al., 2009).

Mechanisms of rhizobial infection have clearly changed over evolution. It has been suggested that the ancient mode of rhizobial invasion was intercellular (Madsen et al., 2010), which may have evolved into crack infection between epidermal cells, and only later perhaps the more controlled root hair infection (Sprent, 2008).

Since crack infection was correlated with a faster Ca^{2+} spiking signature in *S. rostrata*, it was hypothesised that this would represent a more ancient Ca^{2+} response. However, by extending the Ca^{2+} study to a wider range of plants in this chapter, that also undergo crack invasion, this hypothesis does not hold: There is no correlation between infection strategy and Ca^{2+} signatures in the plants investigated.

The tropical legume *Sesbania rostrata* was a clear outlier in the analysis. This plant form symbioses with *Azorhizobium*, and both nodule type and infection process differ depending on environmental conditions. Root nodules of determinate type develop under high levels of ethylene associated with flooding (Fernandez-Lopez et al., 1998). Ethylene is a plant hormone known to be involved in several responses such as stress and phenotypic plasticity. However, stem nodules are also formed and are always indeterminate, as are nodules that develop on the root during dry conditions. The

different infection processes are also linked to ethylene. In laboratory conditions, a long period in the presence of the ethylene-inhibitor was necessary for the RHI and the induction of slower Ca^{2+} oscillations were more comparable to those observed in *M. truncatula* (Goormachtig et al., 2004; Capoen et al., 2009). Therefore, it is possible that these very different levels of ethylene may cause the different Ca^{2+} signatures. Indeed, it has previously been found that ethylene plays a positive role in *S. rostrata* nodulation (D’Haeze et al., 2003), opposite to the role that has been reported for the non-aquatic legume *M. truncatula* (Oldroyd et al., 2001a).

The genus *Lupinus* is the only legume group that does not form an AM symbiosis (Oba et al., 2001). In this light it is interesting that *L. pilosus* appeared to have differential Ca^{2+} oscillations to the other species. Hormones may also be playing a role in defining this differential structure of *L. pilosus* oscillations, but little is known of how hormones regulate nodulation in this species.

Parasponia species are the only non-legumes reported to form root nodule symbioses with rhizobia (Trinick, 1973). Recently, molecular studies revealed that the same receptor for the symbiotic signal molecules act in both AM and rhizobial symbioses in *Parasponia* (Op den Camp et al., 2011). This might indicate that the *Parasponia*-rhizobia symbiosis has evolved relatively recently, and it is interesting to note that its close sister-genera *Trema* does not perceive Nod factors and does not nodulate. However, even if the only difference between *Trema* and *Parasponia* lies in a small change in the ligand receptor, this is unlikely to be connected to the predisposition event since all plants that form AM symbioses are likely to perceive LCOs. If a small adaptation of the LCO receptor was all that was required, there should be many observations of root nodule symbioses spread among higher plants.

In summary this work reveals that Ca^{2+} oscillations induced by diffusible signals are a common feature of plants that enter nitrogen-fixing symbioses with resultant nodule formation. This suggests that the recruitment of the symbiosis signalling pathway for recognition of rhizobial bacteria may be associated with the predisposition to nodulate. However, the present study found that non-nodulating species within this group lacked the ability to perceive Nod factors, but were able to respond to LCOs from mycorrhizal fungi. It was previously suggested that the modality of Ca^{2+} oscillations may be associated with the mechanism of bacterial invasion, but this correlation is not sustained in this larger analysis. I propose that a single mechanism of Ca^{2+} oscillation induction by diffusible signals from nitrogen fixing bacteria is likely ubiquitous among all nodulating plants.

Chapter 7

General discussion

7.1 Patterns of symbiotic calcium oscillations

The objective of this research project was to investigate the structure of Ca^{2+} oscillations associated with the onset of symbioses in legumes using a combination of computational and experimental approaches. The work focused on two hypotheses: 1) do mycorrhizal and rhizobial induced Ca^{2+} oscillations differ sufficiently to explain specificity in symbiosis signalling and 2) are different modes of Ca^{2+} oscillations associated with the different modalities of rhizobial infection. These questions were embedded in a broader assessment of the recruitment of the symbiosis signalling pathway during the evolution of nitrogen fixing symbioses in the Eurosid I clade.

In order to assess the different Ca^{2+} oscillations regimes I have used and further developed Bayesian Spectrum Analysis to quantify the frequencies of Ca^{2+} spiking. This method has proven especially useful when handling data with background trends and varying noise levels, both of which applies to the Ca^{2+} spiking data. I have also tried to assess what may explain differences in the structure of Ca^{2+} oscillations between cells and between species. For this I used a model of Ca^{2+} oscillations to analyse the impact of Ca^{2+} -binding proteins, Ca^{2+} buffers, and I have shown that buffer concentrations and characteristics are predicted to have a large impact on the observed oscillations.

The discovery that mycorrhizal fungi produce LCOs (Maillet et al., 2011) provided an invaluable tool for the comparison of mycorrhizal and rhizobial induced Ca^{2+} oscillations. I showed that Ca^{2+} oscillations induced by the different LCOs had no significant differences in their structure. The comparison of Ca^{2+} spiking was extended to a group of phylogenetically diverse plant species in the nitrogen-fixing clade, to investigate whether the Ca^{2+} spiking signatures could have changed as these symbioses evolved. Some differences in Ca^{2+} patterns were found, but these were not correlated with phylogenetic relationships.

In summary, the present study has deepened the understanding of how Ca^{2+} buffers influence Ca^{2+} patterns in symbiosis. It has also provided data showing that Ca^{2+} spiking patterns are highly similar throughout the nitrogen-fixing clade, as well as

between AM and rhizobial symbioses.

7.2 Similarities and differences in calcium signatures

The results presented in Chapter 5 and 6 demonstrate the remarkable similarity of Ca^{2+} oscillations embedded in the common SYM pathway for the wide range of plant signalling systems that I have studied. This is true both for the comparison of oscillations induced by Myc and Nod factors, but also in the broader comparison of plants responding to LCOs. However, species-specific differences were found, such as the previously studied Ca^{2+} signature of *S. rostrata* (Capoen et al., 2009), but also new variations were described such as the oscillations with slower frequency in *L. pilosus*.

The effects of plant growth regulators on the symbiotic Ca^{2+} oscillations are well-known, for example the inhibitory effect of ABA (Ding et al., 2008) and the species-specific influences of ethylene (Fernandez-Lopez et al., 1998; Oldroyd et al., 2001a). Perhaps a plausible explanation for plant species-specific variability is to be found in differences in physiological state as reflected in local hormone levels. This is suggested on the basis that the machinery of the common SYM pathway seems highly conserved (Gutjahr et al., 2008; Chen et al., 2009; Yokota et al., 2010) and thus physiological differences seem a more likely explanation for the differences in the Ca^{2+} oscillations than genetic variation in the components producing the Ca^{2+} oscillations. Hormones are highly influenced by the environment and the diversity of environmental conditions in which the different species prosper will undoubtedly lead to differences in hormone levels. A good example is *S. rostrata* that can grow under waterlogged conditions under which ethylene accumulates (Goormachtig et al., 2004). Hence, the differences in the structure of Ca^{2+} oscillations may simply reflect unavoidable consequences arising from the growth habits. Alternatively, it may be advantageous for the plant to modulate Ca^{2+} spiking in response to the environment, much like ABA modulates plant guard cells in response to water limitations (Allen et al., 2001). Thus the alternate Ca^{2+} oscillation patterns observed in *S. rostrata* (Capoen et al., 2009) may control the different modes of rhizobial infection according to the environmental conditions, but if this is the case then such innovation must be a recent evolutionary phenomenon since I found no correlation between the structure of the Ca^{2+} oscillations and the mode of rhizobial infection in the diversity of species I tested.

Previously it has been reported that different concentrations of Nod factors do not induce different patterns of Ca^{2+} oscillations. However, in a recent study that investigated later stages of signalling, when the microbes were entering the plant root cells, different frequencies of Ca^{2+} spiking were observed (Sieberer et al., 2011). These correlated with the proximity of the cell to invading microbe, rhizobia or mycorrhiza (Sieberer et al., 2011), which could suggest that diffusible signals were reaching the cells at different concentrations and this could lead to the induction of different modes of

Ca^{2+} oscillations. A concentration effect was observed in the present study (in Chapter 5), following treatment with the biological mix of mycorrhizal LCOs. This points to the possibility of two levels of Ca^{2+} oscillations, with differing frequencies and this diversity of response may be dependent on the mix of LCOs present in the microbial exudates. The weakness of this study is that since the authors measured Ca^{2+} oscillations in cortical cells with confocal microscopy, they were only able to capture short time series of the oscillations. It is difficult to compare Ca^{2+} traces with such short observations to the much longer traces induced by isolated signalling molecules. Perhaps working with synthetic factors is too artificial to simulate the biological situation, since the low-concentration response was mostly evident when working with the microbes or their un-purified extracts. This work suggests that the mix of LCOs present in the microbial exudates may well be important in coordinating the precise nature of the Ca^{2+} oscillations induced.

An unexplained difference was observed in Chapter 6 between the NupYC2.1 transformants and the other plants. The *Parasponia* and *Trema* oscillations were measured in the atrichoblasts of the plantlets transformed with the nuclear-localised NupYC2.1. These oscillations were not directly compared to the data collected with injection, because of the previous problems involved with comparing data from different methods. However, the Ca^{2+} spikes in the *Parasponia* and *Trema* epidermal cells were much broader than Ca^{2+} spikes in any other species, and had long periods of several minutes. This could be an effect of physiological and/or hormone changes due to the transformation, since the NupYC2.1 transformed roots were generated using *A. rhizogenes* that maintained the disease-causing Ri plasmid. It could also be that this is revealing cell-specific differences, since most Ca^{2+} oscillations are measured in root hair cells. Such differences have not been reported in studies of *Medicago* or *Lotus* when using NupYC2.1 (see for example Sieberer et al. (2011) and Kosuta et al. (2011)), but perhaps the cell types have not been consistently compared. Jongho Sun has observed a large variation in spike shapes and period in Ca^{2+} oscillations in atrichoblasts in *M. truncatula* (personal communication), although this has not been rigorously quantified.

Such cell-specific differences could be due to hormone levels and developmental status (Miwa et al., 2006), but the cell size may also play a role as mentioned in Chapter 4. These questions require further investigation.

7.3 The role of calcium-binding proteins

The mathematical modelling simulating the effect of Ca^{2+} buffers on oscillations (as described in Chapter 4) showed how both period and spike shape might be strongly influenced by Ca^{2+} buffers. This has been demonstrated in animal systems (Marhl et al., 1997; Falcke, 2004), but not in plants.

There are a myriad of natural Ca^{2+} buffers in plant cells, and indeed the target proteins themselves act as buffers when they bind Ca^{2+} . Therefore, perhaps the ad-

dition of a new Ca^{2+} buffer, in the shape of the Ca^{2+} reporters, will not have a large impact on the physiological response. However, the Ca^{2+} reporters do certainly have a large impact on the nature of the fluorescence data that is collected, and that is why knowledge of their buffering characteristics is important and should be taken into consideration when choosing a Ca^{2+} reporter.

A high-affinity reporter, with a dissociation constant close to the basal Ca^{2+} concentration, will in theory give the most accurate measurement. However, since such reporters bind Ca^{2+} tightly, they are more likely to buffer, and possibly perturb the signal. At the opposite end of the spectrum, low-affinity reporters are better able to capture the speed of the response, but may be less accurate in other characteristics (Putney, 1998). To further complicate the issue, the inside of the cell can change the action spectra and kinetics of these reporters (Putney, 1998), making the *in vitro* characteristics less relevant to the *in vivo* situation.

The main conclusions from this study are twofold. Firstly, it is a plausible hypothesis that Ca^{2+} buffers in the cell can play a role in shaping the nature of the Ca^{2+} patterns that are observed. Secondly, the experiments must be carried out with the clear understanding that the data may not reveal the actual Ca^{2+} response, since the act of measuring the Ca^{2+} oscillations may impact their observed structure.

The protein complex decoding the Ca^{2+} oscillations can act as Ca^{2+} buffers leading to the question: what is the nature of the signal perceived? In order to address this question, we should perhaps strive to make the binding kinetics of our Ca^{2+} indicators as similar as possible to the binding kinetics of the protein decoding the Ca^{2+} signals.

The protein that is required for decoding the symbiotic Ca^{2+} oscillations is a calcium/calmodulin-dependent protein kinase (CCaMK) that is only found in plants. It has structural similarities to proteins in animals such as CaMKII, that has been shown to distinguish different frequencies of Ca^{2+} oscillations and activate different downstream responses (De Koninck and Schulman, 1998). The fact that CCaMK is structurally more complex than its animal counterpart in the way that it binds Ca^{2+} and calmodulin (CaM), made an attractive hypothesis that CCaMK could distinguish different patterns of Ca^{2+} oscillations and that this could be a means for specificity in symbiosis signalling (Oldroyd and Downie, 2006). This hypothesis was also supported by other cases of Ca^{2+} signals that conferred specificity of downstream responses, such as plant guard cells that have been shown to respond to different Ca^{2+} oscillations with different strengths of guard cell closure (McAinsh and Pittman, 2009).

Taken together with comparisons of data using different methods, this led to the hypothesis of specificity in Ca^{2+} oscillations being suggested in Kosuta et al. (2008). This hypothesis does not seem to be consistent with the new experimental evidence obtained in the present study. However, it has also been shown that during infection rhizobia and mycorrhiza activate different Ca^{2+} signatures according to their distance from cell (Sieberer et al., 2011). Could it then still be the case that CCaMK can decode different Ca^{2+} oscillations?

The mechanism of decoding the Ca^{2+} signal is largely unknown for CCaMK. However, the protein can bind Ca^{2+} in two ways: via three EF-hands or via binding CaM, which in turns bind Ca^{2+} (Oldroyd and Downie, 2006). Several studies have recently suggested ways in which CCaMK could be regulated (Shimoda et al., 2012; Takeda et al., 2012; Liao et al., 2012), as well as studies of the Ca^{2+} -binding properties of CCaMK (Swainsbury et al., 2012). However, several results seem contradictory at this point, and the current state was reviewed in Singh and Parniske (2012).

Perhaps the complex structure of CCaMK does allow it to identify different Ca^{2+} signatures, for example those that were associated with infection (Sieberer et al., 2011). This could perhaps serve to separate cells in close proximity to the symbionts and those that will actually be infected. In this case it may be that CCaMK acts similarly in both arbuscular mycorrhizal and rhizobial symbiosis, since the PPA and the pre-IT are similar structures. Another hypothesis could be that CCaMK only responds to one Ca^{2+} signature, and that the complex structure and regulation of this protein guarantees that it is only activated under appropriate conditions.

7.4 Evolution of the common SYM pathway

The common SYM pathway consists of components that are required for the establishment of both AM and rhizobial symbioses. Several genes related to the common SYM genes have been found in non-legumes such as rice (Chen et al., 2009; Gutjahr et al., 2008), liverworts and hornworts (Wang et al., 2010), highlighting the importance of this signalling pathway in the mycorrhizal association with these non-legumes. Furthermore, CCaMK from rice could complement nodulation in *M. truncatula ccamk* mutants (Godfroy et al., 2006). This points to the common SYM components being highly conserved (Yokota and Hayashi, 2011). The results presented in Chapter 5 which show equivalent Ca^{2+} oscillations induced by both AM and rhizobial signal, agrees with this conservation. The predisposition event which is believed to have allowed plants to acquire the capacity to form root-nodule symbioses appears unlikely to be associated with a change in the structure of symbiotic Ca^{2+} oscillations, but may still be associated with the recruitment of this signalling pathway for recognition of nitrogen-fixing bacteria.

An alternative hypothesis is that the predisposition event lies outside of the common SYM pathway and this is supported by the observation that some rhizobia are able to bypass the SYM pathway, but invade the plant root and form nodules. In the *Aeschynomene* genus, several plant species form root and stem nodules with photosynthetic *Bradyrhizobia*. One strain of these rhizobia has been shown to lack *Nod* genes, and so does not produce Nod factors (Giraud et al., 2007). These *Bradyrhizobia* are believed to produce a cytokinin-related signal (Giraud et al., 2007) which could then activate parts of the signalling pathway downstream of the Ca^{2+} spiking components.

Analyses in *Lotus japonicus* mutants showed that rhizobia can infect the root in-

tercellularly and this appeared to be independent of Nod factor (Madsen et al., 2010), since this mode of infection did not require Nod factor receptors in the plant nor the Nod factors from the rhizobia.

The evidence for rhizobia entering into root nodule symbiosis with legumes without an absolute requirement of the common SYM pathway shows how versatile and adaptable the rhizobial-legume symbiosis is. There is a broad spectrum of infection modes by rhizobia, as can be seen, for example, in the legumes *Sesbania* and *Aeschynomene*. The results from the present study (Chapter 6) also indicate that the evolution of these different infection modes did not involve a change in the Ca^{2+} spiking signatures. Therefore, I propose that the acquisition of the SYM pathway from the AM symbiosis did not involve a change in the modality of signalling. The predisposition event which facilitated the emergence of nodulation in the Eurosid 1 clade (Soltis et al., 1995) may be associated with the capability for rhizobial bacteria to activate the SYM pathway, but could equally lie outside the common SYM pathway.

7.5 Prospects for biotechnology involving the common SYM pathway

Understanding of arbuscular mycorrhizal and rhizobial symbioses is of great interest to fundamental plant science, since these symbioses represent highly complex plant-microbe interactions. Both symbioses are also important to understand in the context of agriculture. Phosphorous and nitrogen are the two main ingredients of inorganic fertilisers. The AM and rhizobial symbioses provide phosphorous and nitrogen respectively, and increasing the knowledge about them may enable the optimisation of these associations in agricultural production. With the global human population projected to increase dramatically in the coming decades, new ways of increasing food production are needed. Because of this, the possibility of transferring the nitrogen-fixing symbiosis to crops such as cereals has received much attention (Charpentier and Oldroyd, 2010).

Since cereals form AM symbioses they are assumed to contain the components for the SYM pathway, and in rice this has been confirmed for several common SYM genes (Chen et al., 2009; Gutjahr et al., 2008). Enabling the rhizobia to infect cereals would require biotechnological engineering, imitating the evolution that occurred in the nitrogen-fixing clade. The genetic components that will need to be transferred will be involved in recognition of the rhizobial signals. Considering that AM fungi use signals that are structurally very similar to the rhizobial Nod factors makes this step look very possible to achieve, which would allow rhizobia to activate the SYM pathway in cereals.

Other traits may be more difficult to transfer, including components regulating nodule organogenesis. Furthermore, if the predisposition event is indeed outside the common SYM pathway, more components will need to be identified. But given the fast pace that research in the plant symbiosis field is progressing, and the strong motivations for nitrogen-fixing cereals, the prospects are promising.

Appendices

.A Mathematical background

.A.1 Interpretations of probability theory

Probability is straightforward in its mathematical theory, but the practical interpretation opens up for different opinions. There are several approaches to probability theory, but the two main are the frequentist approach and the Bayesian approach (Wilkinson, 2006). The frequentist defines probability as the proportion of times a certain outcome is observed, out of the total number of observations, under identical conditions. Limiting points to this approach is that probability is not defined unless there is the possibility of identical repeats of the observation (Wilkinson, 2006).

The alternative is the Bayesian approach, and allows for a prior belief in the event before it is measured. This is the view of probability that underlies much of the work in Chapter 3. Bayesian methods have received criticism because it allows one to incorporate assumptions into the probability evaluation. However, Bayesians argue that assumptions, or prior knowledge as it is called, are always present and are better stated clearly to be used to the advantage of the inference problem at hand (Jaynes and Bretthorst, 2003).

Both approaches have merits as well as disadvantages. However, a key advantage of the Bayesian approach is that it is consistent in its treatment of inference problems regardless of the details of the questions. The frequentist approach often utilises tailor-made methods to cope with different situations, something that gives rise to much confusion (MacKay, 2003).

.A.2 Dynamical systems theory

Dynamical systems can be represented by the state of its variables in so-called phase space. There, each set of variable values is a single point, with the time variable excluded. Depending on the starting point, the initial conditions, the system will develop in a different way, or follow different *trajectories* in phase space.

To describe dynamical systems ordinary differential equations (ODEs) are often used, describing the changes in the variables over time. A minimum of two ODEs are required to produce oscillations, and for chaos there must be a minimum of three ODEs (Schuster et al., 2002). The system will be drawn to its steady solutions, the so-called attractors in phase space. Two fundamental non-chaotic attractors are the steady states (or fixed points by another terminology) and limit cycles. A steady state, or a point attractor, is a state which draws the system towards it and, once reached, does not change over time. A simple example of this is a bouncing ball which eventually comes to rest. Limit cycles are similar, in that once the system enters a limit cycle it will stay there. Essentially they are periodic attractors, for example the sine waves discussed in Chapter 3. A beautiful example of a limit cycle is your own beating heart, which will return to its rhythm as soon as possible after a disturbance. However, there

is both global and local stability, where local stability only stands firm towards small perturbations. If the perturbations are large enough, the system may flow towards another attractor. Global steady states on the other hand will always draw the system toward them, regardless of the size of perturbations.

How the system reaches these attractors is often a point of interest. For more information on dynamical systems, see the excellent basic textbook by Strogatz (1994).

Bifurcation Analysis

Analysing the stability of a system and how it depends on the system parameters is called *bifurcation* analysis. This was done using the Ca^{2+} spiking model described in Chapter 4, with focus on Ca^{2+} buffer parameters. Bifurcations can be global or local, and the global types are more complicated to identify. Local bifurcations on the other hand can be found through stability analysis of the fixed points. A bifurcation with respect to some parameter, λ , is found at the value of λ where the number of solutions to the system change. There are different types of bifurcations, but we will focus the types occurring in continuous systems, since those are described by ODEs.

In neurons and excitable spiking, there are an enormous amount of underlying mechanisms that can give rise to Ca^{2+} spiking (oscillations), but only four types of bifurcations: Saddle-node, saddle node on invariant circle, subcritical Andronov-Hopf bifurcation and supercritical Andronov-Hopf bifurcation (Izhikevich, 2007). The two latter are often simply known as Hopf bifurcations. Thus, without knowing if our underlying model is correct, by studying the bifurcations that give rise to the desired data fit, we can gain insight into paths to spiking. We should also note that different initial conditions will result in different solutions in some systems, and also the threshold value of some parameter at which a change occurs (Seydel, 2010).

An important type of bifurcation in our oscillatory set of ODEs is the Hopf bifurcation. These connect stable solutions to periodic ones (oscillations). As mentioned above, there are two types of Hopf bifurcations, subcritical and supercritical. Subcritical occurs when the steady state transitions into an unstable periodic cycle. The supercritical on the other hand is when a stable periodic cycle is born (Seydel, 2010). Such Hopf bifurcations are found in many famous systems, for example the Hodgkin-Huxley model.

To find bifurcation points, the state of the system is determined by calculating the Jacobian. When the eigenvalues of the Jacobian are passing the imaginary axis in the complex plane, a bifurcation takes place (Strogatz, 1994). See the next section for details on ODEs, the Jacobian and eigenvalues.

.A.3 Linear algebra

Systems of equations

A system of ordinary differential equations (ODEs),

$$\begin{cases} y_1 = f_1(x_1, x_2, \dots, x_n) \\ \vdots \\ y_n = f_n(x_1, x_2, \dots, x_n) \end{cases} \quad (1)$$

is often used to describe a dynamical system. There are three key ingredients in such systems: time, state and functions governing the system's evolution. In an autonomous system, the evolution does not depend on time, but simply on the current state of the system and the function f defining how it moves from there (Seydel, 2010). As an example, for a two-dimensional system we can visualise the phase space for the two variables, with one axis for each, and at each point in this diagram the system has a state and a direction, a *trajectory*. Time will simply move the system along these trajectories. There are three types of trajectories for ODE systems, based on the uniqueness theorems, namely stationary solutions, periodic solutions, and one-to-one solutions (Seydel, 2010).

Importantly, there are some points in phase space where the system stays, so called *fixed* or *stationary* points. These are solutions to the ODE, i.e. for our 2-dimensional example $\dot{y}_1 = 0$ and $\dot{y}_2 = 0$, the system is at rest.

The stability of the other points can of course also be evaluated. The vector field of the system, ϕ , consisting of the system's trajectories is described by

$$\phi = Ax + b, \quad (2)$$

where A is a matrix of values, b is a vector, and x is the position vector. When b is zero, the the vector field will describe if the system is moving towards or from a fixed point, the equilibrium. For this, we need to examine the eigenvalues and eigenvectors of the vector field.

Eigenvalues and eigenvectors

In dynamical systems described by systems of differential equations, eigenvectors and eigenvalues are essential for analysing phase space trajectories. Eigenvalues can be both real and complex, and together with their sign this carries important information of the system's stability.

For any square matrix, A , vectors that do not change direction after multiplication with the matrix are called the eigenvectors, and their length after multiplication is the corresponding eigenvalue. Thus, the eigenvectors form a basis for the matrix. The number of eigenvectors equals the number of variables of the system. Denoting the square matrix A , an eigenvector \mathbf{v} and an eigenvalue λ , then

$$A\mathbf{v} = \lambda\mathbf{v}, \quad (3)$$

where $\mathbf{v} \neq 0$. When this is true, it will also be true that

$$\det(A - \lambda I) = 0, \tag{4}$$

where $\det(A - \lambda I)$ denotes the determinant of the matrix $A - \lambda I$, and I is the identity matrix of A . This equation is also called the characteristic equation, and through it the eigenvalues can be solved for, and subsequently the eigenvectors.

The Jacobian and stability analysis

The Jacobian is often the short name of the determinant of the Jacobian matrix, a matrix consisting of the first-order partial derivatives of a function with respect to some other function. So for a system of ODEs such as the one in Equation 1, the Jacobian matrix is

$$J(x_1, \dots, x_n) = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \dots & \frac{\partial y_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_n}{\partial x_1} & \dots & \frac{\partial y_n}{\partial x_n} \end{bmatrix}. \tag{5}$$

For dynamical systems, the determinant for the Jacobian matrix is used to determine the stability at fixed points. When all the real parts of the eigenvalues are negative, the point is stationary, in equilibrium. The interpretation of this is straightforward in that if the system is slightly disturbed from this point, all trajectories are negative i.e. carry the system back to the equilibrium. If one or more of the eigenvalues are positive, the point is unstable. Furthermore, if there is no real part to the eigenvalues, only an imaginary, then the system is in periodic orbit. For more details on linear algebra and stability of systems, see for example Strogatz (1994).

.A.4 Power series and polynomials

Taylor series

A Taylor series, $T(x)$, is a power series of a infinitely differential function, $f(x)$, at some point, x_0 . It requires all derivatives at that point to be known, and then the Taylor series is created around the point using the power series

$$T(x) = \sum_{n=0}^{\infty} a_n (x - x_0)^n, \tag{6}$$

where a are the coefficients which must be of the form

$$a_n = \frac{f^n(x_0)}{n!} \tag{7}$$

for $f(x) = T(x)$ to be true. This can be used to approximate the function $f(x)$ with a finite number of terms of the series. For more details on Taylor series see for example

Edelstein-Keshet (2005) or background texts on power series.

Legendre polynomials

The Legendre polynomials are solutions to the Legendre differential equation,

$$\frac{d}{dx} \left[(1 - x^2) \frac{d}{dx} P_n(x) \right] + n(n + 1)P_n(x) = 0. \quad (8)$$

If n is an integer the Legendre functions are polynomials, and orthogonal on the interval $-1 \leq x \leq 1$. The Legendre polynomials up to order 10 on the interval $0 \leq x \leq 1$ are shown in Figure 1.

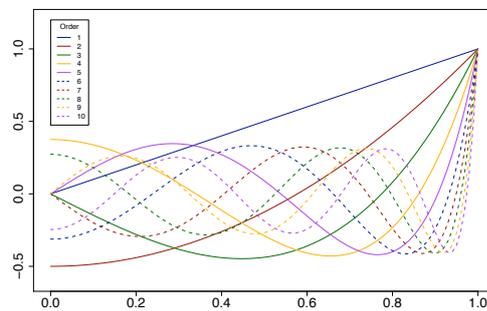


Figure 1: The Legendre polynomials

.B Supplementary figures and tables

Table 1: List of data used in Chapter 5 from *M. truncatula* YC2.1. Source indicates person performing experiments, EG: author, JS: Jongho Sun, SK: Sonia Kosuta. Where no absolute concentration is defined, the number indicate the strength relative to the original exudate. In the fungal experiments, with *Glomus intraradices*, the whole fungi is used in close proximity to the plant.

Induction	Conc.	Source
NodL	1 nM	EG
NodL	100 pM	EG
NodL	10 pM	EG
S-LCO	1/100	EG
Syn S-LCO	10 nM	EG
Syn S-LCO	1 nM	EG
NS-LCO	1	JS
Syn NS-LCO	100 μ M	JS
Syn NS-LCO	10 μ M	EG, JS
Syn NS-LCO	1 μ M	EG
M-LCOs	1	JS
M-LCOs	1/10	JS
M-LCOs	1/100	JS
M-LCOs	1/1000	JS
<i>G. intraradices</i>	-	SK, JS
CT4	10 μ M	JS

Table 2: Results from the statistical Mann-Whitney-Wilcoxon test for interspike intervals induced by different LCOs. The values are p-values, and a p-value below 0.05 shows a significant difference in the compared distributions. The results have been rounded off to three decimal places. Since the pair-wise comparison is symmetric, to avoid redundancy only half the table is filled.

	Nod factor	Syn S-LCO	Syn NS-LCO	M-LCOs	<i>G. intraradices</i>	CT4
Nod factor	1					
Syn S-LCO	0.151	1				
Syn NS-LCO	0.178	0.01	1			
M-LCOs	0.23	0.87	0.009	1		
<i>G. intraradices</i>	0.086	0.585	0.012	0.655	1	
CT4	0	0	0	0	0	1

Table 3: Kolmogorov-Smirnov test for BSA results on Ca^{2+} spiking induced by different LCOs. The results have been rounded off to three decimal places. Since the pair-wise comparison is symmetric, to avoid redundancy only half the table is filled.

	Nod factor	Syn S-LCO	Syn NS-LCO	M-LCOs	<i>G. intraradices</i>	CT4
Nod factor	1					
Syn S-LCO	0.111	1				
Syn NS-LCO	0.004	0.054	1			
M-LCOs	0	0	0.01	1		
<i>G. intraradices</i>	0	0	0	0	1	
CT4	0	0	0	0	0.468	1

Table 4: Kullback-Leibler divergence for BSA results on Ca^{2+} spiking induced by different LCOs. The results have been rounded off to three decimal places. Since the pair-wise comparison is symmetric, to avoid redundancy only half the table is filled.

	Nod factor	Syn S-LCO	Syn NS-LCO	M-LCOs	<i>G. intraradices</i>	CT4
Nod factor	0					
Syn S-LCO	0.203	0				
Syn NS-LCO	0.046	0.13	0			
M-LCOs	0.585	0.444	0.695	0		
<i>G. intraradices</i>	0.364	0.518	0.495	0.316	0	
CT4	0.664	0.403	0.741	0.044	0.275	1

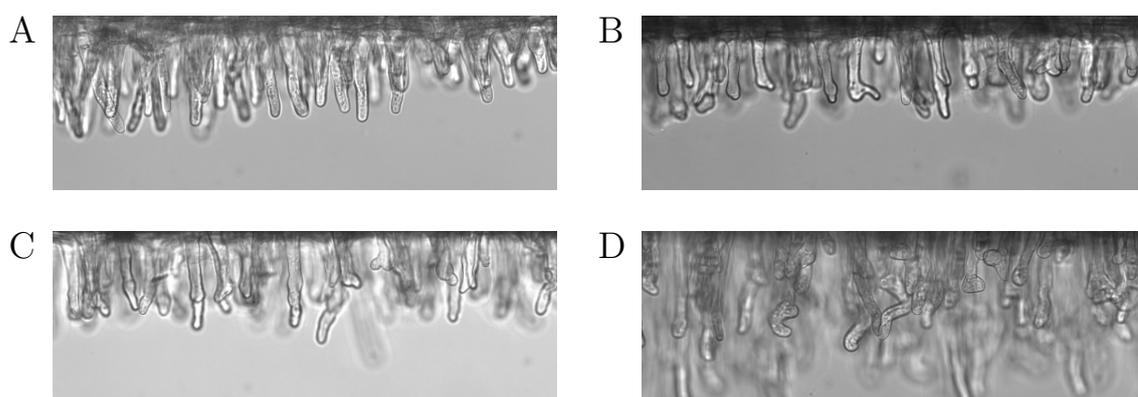


Figure 2: NGR Nod factors inducing root hair deformation in *Vicia hirsuta*. **A:** Negative control **B:** Supernatant of NGR234 liquid culture **C:** NGR234 Nod factors, diluted 1/1000 **D:** NGR234 Nod factors diluted 1/100

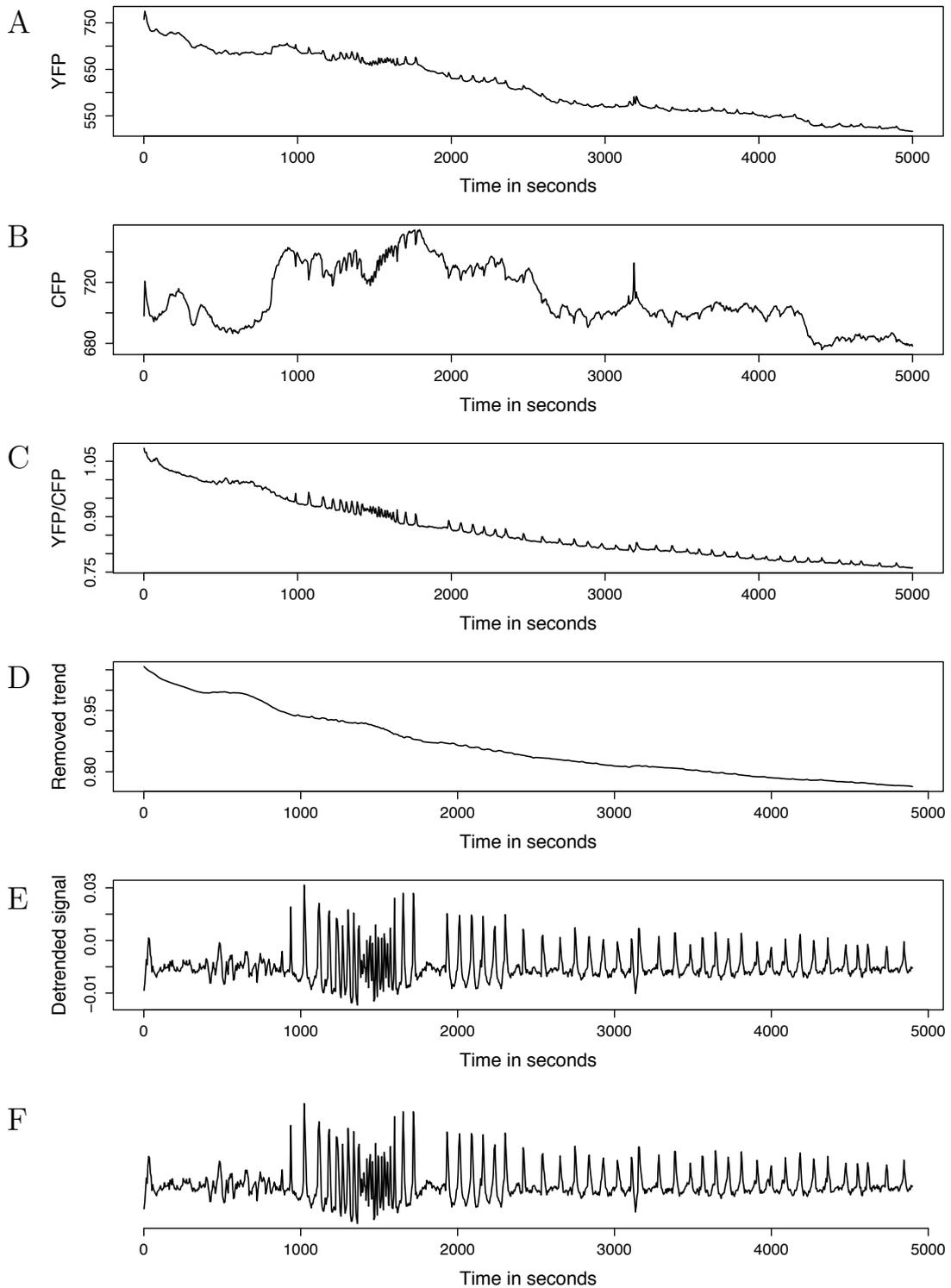


Figure 3: The process from raw YC2.1 Ca^{2+} imaging data to the presented data. Example 1. **A:** YFP channel. **B:** CFP channel. **C:** Ratio YFP/CFP. **D:** Removed background trend, using moving average. **E:** Remaining signal after detrending. **F:** Presented signal, excluding amplitude. For more details on the methods see Chapter 2, Section 2.1.4.

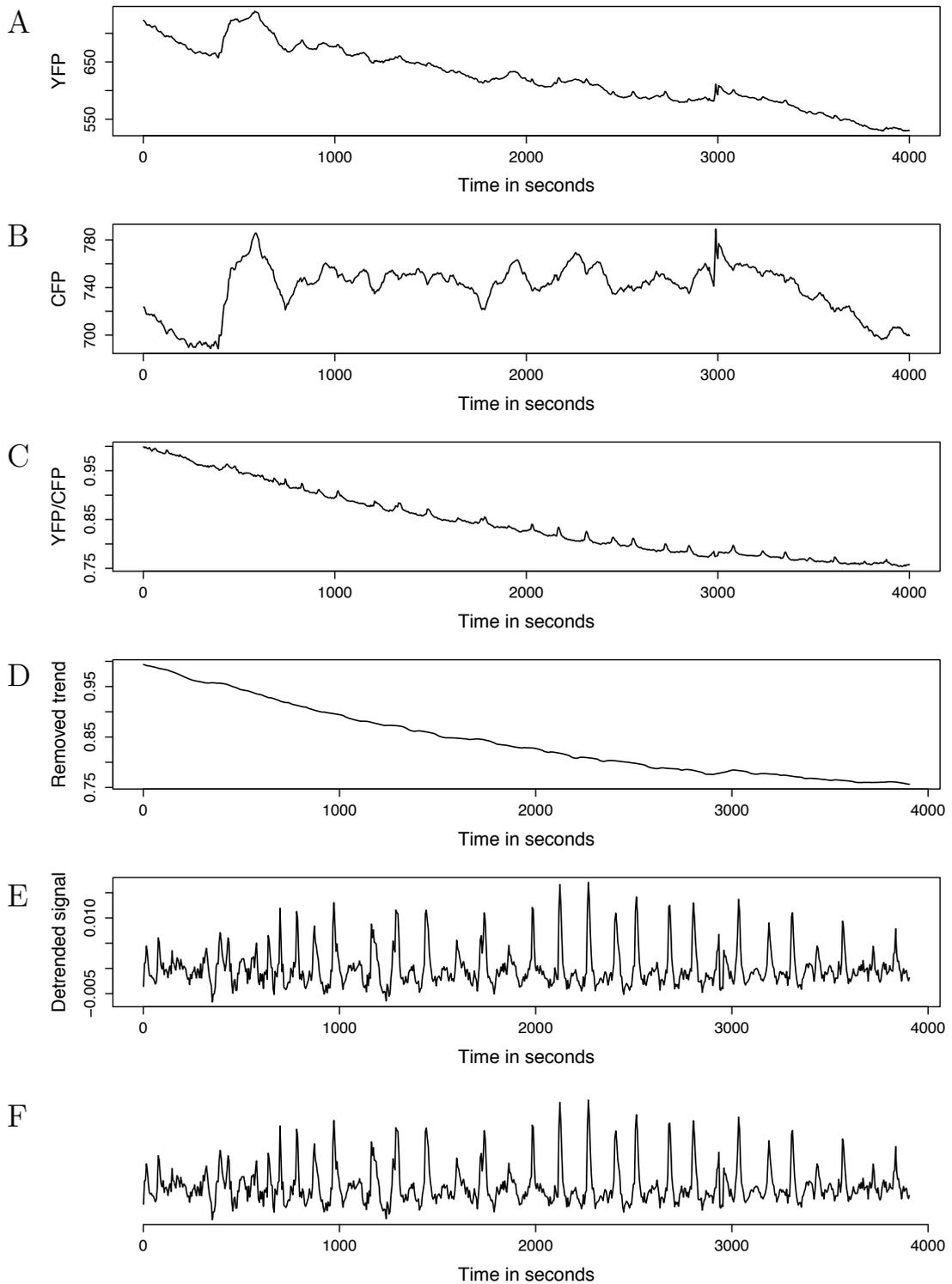


Figure 4: The process from raw YC2.1 Ca^{2+} imaging data to the presented data. Example 2. **A:** YFP channel. **B:** CFP channel. **C:** Ratio YFP/CFP. **D:** Removed background trend, using moving average. **E:** Remaining signal after detrending. **F:** Presented signal, excluding amplitude. For more details on the methods see Chapter 2, Section 2.1.4.

Table 5: Results from the statistical Mann-Whitney-Wilcoxon test for interspike intervals in the nitrogen-fixing clade. The values are p-values, and a p-value below 0.05 shows a significant difference in the compared distributions. The results have been rounded off to three decimal places. Since the pair-wise comparison is symmetric, to avoid redundancy only half the table is filled.

	<i>A. glutinosa</i>	<i>M. truncatula</i>	<i>S. rostrata</i>	<i>S. rostrata</i> (RHI)	<i>A. retinoides</i>	<i>L. pilosus</i>	<i>C. fasciculata</i>	<i>C. proliferus</i>	<i>P. andersonii</i>
<i>A. glutinosa</i>	1								
<i>M. truncatula</i>	0	1							
<i>S. rostrata</i>	0	0	1						
<i>S. rostrata</i> (RHI)	0.003	0	0	1					
<i>A. retinoides</i>	0	0.021	0	0	1				
<i>L. pilosus</i>	0	0	0	0.004	0	1			
<i>C. fasciculata</i>	0	0.225	0	0	0	0	1		
<i>C. proliferus</i>	0.17	0	0	0.151	0	0	0	1	
<i>P. andersonii</i>	0	0.58	0	0	0.275	0	0.01	0	1

Table 6: Results from the statistical Mann-Whitney-Wilcoxon test for spike structure during up-phase and down-phase, in the nitrogen-fixing clade. The values are p-values, and a p-value below 0.05 shows a significant difference in the compared distributions. The results have been rounded off to three decimal places. Since the pair-wise comparison is symmetric, to avoid redundancy only half the table is filled.

(a) Mann-Whitney-Wilcoxon test for spike structure of up-phase

	<i>A. glutinosa</i>	<i>M. truncatula</i>	<i>S. rostrata</i>	<i>S. rostrata</i> (RHI)	<i>A. retinoides</i>	<i>L. pilosus</i>	<i>C. fasciculata</i>	<i>C. proliferus</i>	<i>P. andersonii</i>
<i>A. glutinosa</i>	1								
<i>M. truncatula</i>	0	1							
<i>S. rostrata</i>	0	0	1						
<i>S. rostrata</i> (RHI)	0.001	0.007	0	1					
<i>A. retinoides</i>	0	0	0	0	1				
<i>L. pilosus</i>	0.5	0	0	0.056	0	1			
<i>C. fasciculata</i>	0	0.012	0	0	0.148	0	1		
<i>C. proliferus</i>	0	0.459	0	0.002	0.007	0	0.181	1	
<i>P. andersonii</i>	0	0.076	0	0.461	0	0.006	0	0.035	1

(b) Mann-Whitney-Wilcoxon test for spike structure of down-phase

	<i>A. glutinosa</i>	<i>M. truncatula</i>	<i>S. rostrata</i>	<i>S. rostrata</i> (RHI)	<i>A. retinoides</i>	<i>L. pilosus</i>	<i>C. fasciculata</i>	<i>C. proliferus</i>	<i>P. andersonii</i>
<i>A. glutinosa</i>	1								
<i>M. truncatula</i>	0.143	1							
<i>S. rostrata</i>	0	0	1						
<i>S. rostrata</i> (RHI)	0.024	0.002	0	1					
<i>A. retinoides</i>	0	0	0	0	1				
<i>L. pilosus</i>	0.08	0.551	0	0.004	0	1			
<i>C. fasciculata</i>	0	0	0	0	0.003	0	1		
<i>C. proliferus</i>	0	0	0.005	0	0.841	0	0.009	1	
<i>P. andersonii</i>	0.014	0.279	0	0.001	0	0.643	0	0	1

Bibliography

- Akiyama, K., Matsuzaki, K.-i., and Hayashi, H. (2005). Plant sesquiterpenes induce hyphal branching in arbuscular mycorrhizal fungi. *Nature*, 435(7043):824–7.
- Akkermans, A., Abdulkadir, S., and Trinick, M. (1978). N₂-fixing root nodules in Ulmaceae: *Parasponia* or (and) *Trema* spp.? *Plant and Soil*, 49:711–715. 10.1007/BF02183301.
- Allen, G. J., Chu, S. P., Harrington, C. L., Schumacher, K., Hoffmann, T., Tang, Y. Y., Grill, E., and Schroeder, J. I. (2001). A defined range of guard cell calcium oscillation parameters encodes stomatal movements. *Nature*, 411(6841):1053–7.
- Allen, G. J., Chu, S. P., Schumacher, K., Shimazaki, C. T., Vafeados, D., Kemper, A., Hawke, S. D., Tallman, G., Tsien, R. Y., Harper, J. F., Chory, J., and Schroeder, J. I. (2000). Alteration of stimulus-specific guard cell calcium oscillations and stomatal closing in *Arabidopsis det3* mutant. *Science*, 289(5488):2338–42.
- Amor, B. B., Shaw, S. L., Oldroyd, G. E. D., Maillet, F., Penmetsa, R. V., Cook, D., Long, S. R., Dénarié, J., and Gough, C. (2003). The NFP locus of *Medicago truncatula* controls an early step of nod factor signal transduction upstream of a rapid calcium flux and root hair deformation. *Plant J*, 34(4):495–506.
- Ané, J.-M., Kiss, G. B., Riely, B. K., Penmetsa, R. V., Oldroyd, G. E. D., Ajax, C., Lévy, J., Debelle, F., Baek, J.-M., Kalo, P., Rosenberg, C., Roe, B. A., Long, S. R., Dénarié, J., and Cook, D. R. (2004). *Medicago truncatula DMI1* required for bacterial and fungal symbioses in legumes. *Science*, 303(5662):1364–7.
- Arrighi, J.-F., Barre, A., Ben Amor, B., Bersoult, A., Soriano, L. C., Mirabella, R., de Carvalho-Niebel, F., Journet, E.-P., Ghérardi, M., Huguet, T., Geurts, R., Dénarié, J., Rougé, P., and Gough, C. (2006). The *Medicago truncatula* lysin motif-receptor-like kinase gene family includes *NFP* and new nodule-expressed genes. *Plant Physiol*, 142(1):265–79.
- Bago, B., Pfeffer, P. E., and Shachar-Hill, Y. (2000). Carbon metabolism and transport in arbuscular mycorrhizas. *Plant Physiol*, 124(3):949–58.
- Baker, D. D. (1987). Relationships among pure cultured strains of *Frankia* based on host specificity. *Physiologia Plantarum*, 70(2):245–248.
- Banba, M., Gutjahr, C., Miyao, A., Hirochika, H., Paszkowski, U., Kouchi, H., and Imaizumi-Anraku, H. (2008). Divergence of evolutionary ways among common *sym* genes: CASTOR and CCaMK show functional conservation between two symbiosis systems and constitute the root of a common signaling pathway. *Plant Cell Physiol*, 49(11):1659–71.

-
- Beatty, P. H. and Good, A. G. (2011). Future prospects for cereals that fix nitrogen. *Science*, 333(6041):416–7.
- Becking, J. (1983). The *Parasponia parviflora*—Rhizobium symbiosis. Host specificity, growth and nitrogen fixation under various conditions. *Plant and Soil*, 75:309–342.
- Benedito, V. A., Torres-Jerez, I., Murray, J. D., Andriankaja, A., Allen, S., Kakar, K., Wandrey, M., Verdier, J., Zuber, H., Ott, T., Moreau, S., Niebel, A., Frickey, T., Weiller, G., He, J., Dai, X., Zhao, P. X., Tang, Y., and Udvardi, M. K. (2008). A gene expression atlas of the model legume *Medicago truncatula*. *Plant J*, 55(3):504–13.
- Benková, E. and Bielach, A. (2010). Lateral root organogenesis - from cell to organ. *Curr Opin Plant Biol*, 13(6):677–83.
- Bergman, B., Johansson, C., and Söderbäck., E. (1992). The *Nostoc*–*Gunnera* symbiosis. *New Phytol*, 122:379–400.
- Berridge, M. J. (2009). Inositol trisphosphate and calcium signalling mechanisms. *Biochim Biophys Acta*, 1793(6):933–40.
- Bersoult, A., Camut, S., Perhald, A., Kereszt, A., Kiss, G., and Cullimore, J. (2005). Expression of the *Medicago truncatula* *DMI2* gene suggests roles of the symbiotic nodulation receptor kinase in nodules and during early nodule development. *Mol Plant Microbe Interact*, 18(8):869–876.
- Besserer, A., Puech-Pagès, V., Kiefer, P., Gomez-Roldan, V., Jauneau, A., Roy, S., Portais, J.-C., Roux, C., Bécard, G., and Séjalon-Delmas, N. (2006). Strigolactones stimulate arbuscular mycorrhizal fungi by activating mitochondria. *PLoS Biol*, 4(7):e226.
- Bishopp, A., Benková, E., and Helariutta, Y. (2011). Sending mixed messages: auxin-cytokinin crosstalk in roots. *Curr Opin Plant Biol*, 14(1):10–6.
- Bonfante, P. and Genre, A. (2010). Mechanisms underlying beneficial plant–fungus interactions in mycorrhizal symbiosis. *Nat Commun*, 1:48.
- Bootman, M. D., Fearnley, C., Smyrniak, I., MacDonald, F., and Roderick, H. L. (2009). An update on nuclear calcium signalling. *J Cell Sci*, 122(Pt 14):2337–50.
- Bracewell, R. N. (1978). *The Fourier transform and its applications*. McGraw-Hill, New York, 2nd edition.
- Bretthorst, G. L. (1988). *Bayesian spectrum analysis and parameter estimation*. Lecture notes in statistics. Springer-Verlag, New York.
- Bretthorst, G. L., Kotyk, J. J., and Ackerman, J. J. (1989). 31P NMR bayesian spectral analysis of rat brain *in vivo*. *Magn Reson Med*, 9(2):282–7.
- Brewin, N. J. (2004). Plant cell wall remodelling in the Rhizobium-legume symbiosis. *Critical Reviews in Plant Sciences*, 23:293–316.
- Brockwell, P. J. and Davis, R. A. (2002). *Introduction to time series and forecasting*. Springer, New York, 2nd edition.

-
- Buist, G., Steen, A., Kok, J., and Kuipers, O. P. (2008). LysM, a widely distributed protein motif for binding to (peptido)glycans. *Mol Microbiol*, 68(4):838–47.
- Camerini, S., Senatore, B., Lonardo, E., Imperlini, E., Bianco, C., Moschetti, G., Rotino, G. L., Campion, B., and Defez, R. (2008). Introduction of a novel pathway for IAA biosynthesis to rhizobia alters vetch root nodule development. *Arch Microbiol*, 190(1):67–77.
- Campos-Soriano, L., Gomez-Ariza, J., Bonfante, P., and San Segundo, B. (2011). A rice calcium-dependent protein kinase is expressed in cortical root cells during the presymbiotic phase of the arbuscular mycorrhizal symbiosis. *BMC Plant Biol*, 11(1):90.
- Canfield, D. E., Glazer, A. N., and Falkowski, P. G. (2010). The evolution and future of earth’s nitrogen cycle. *Science*, 330(6001):192–6.
- Capoen, W., Den Herder, J., Sun, J., Verplancke, C., De Keyser, A., De Rycke, R., Goormachtig, S., Oldroyd, G., and Holsters, M. (2009). Calcium spiking patterns and the role of the calcium/calmodulin-dependent kinase CCaMK in lateral root base nodulation of *Sesbania rostrata*. *Plant Cell*, 21(5):1526–40.
- Capoen, W., Goormachtig, S., De Rycke, R., Schroeyers, K., and Holsters, M. (2005). SrSymRK, a plant receptor essential for symbiosome formation. *Proc Natl Acad Sci U S A*, 102(29):10369–10374.
- Capoen, W. and Oldroyd, G. (2008). How CYCLOPS keeps an eye on plant symbiosis. *Proc Natl Acad Sci U S A*, 105(51):20053–4.
- Capoen, W., Sun, J., Wysham, D., Otegui, M. S., Venkateshwaran, M., Hirsch, S., Miwa, H., Downie, J. A., Morris, R. J., Ané, J.-M., and Oldroyd, G. E. D. (2011). Nuclear membranes control symbiotic calcium signaling of legumes. *Proc Natl Acad Sci U S A*.
- Cárdenas, L., Domínguez, J., Quinto, C., López-Lara, I. M., Lugtenberg, B. J., Spaink, H. P., Rademaker, G. J., Haverkamp, J., and Thomas-Oates, J. E. (1995). Isolation, chemical structures and biological activity of the lipo-chitin oligosaccharide nodulation signals from *Rhizobium etli*. *Plant Mol Biol*, 29(3):453–64.
- Cárdenas, L., Feijó, J. A., Kunkel, J. G., Sánchez, F., Holdaway-Clarke, T., Hepler, P. K., and Quinto, C. (1999). *Rhizobium* Nod factors induce increases in intracellular free calcium and extracellular calcium influxes in bean root hairs. *Plant J*, 19(3):347–52.
- Cárdenas, L. and Quinto, C. (2008). Reactive oxygen species (ROS) as early signals in root hair cells responding to rhizobial nodulation factors. *Plant Signal Behav*, 3(12):1101–2.
- Catoira, R., Galera, C., de Billy, F., Penmetsa, R. V., Journet, E. P., Maillet, F., Rosenberg, C., Cook, D., Gough, C., and Dénarié, J. (2000). Four genes of *Medicago truncatula* controlling components of a Nod factor transduction pathway. *Plant Cell*, 12(9):1647–66.
- Chabaud, M., Genre, A., Sieberer, B. J., Faccio, A., Fournier, J., Novero, M., Barker, D. G., and Bonfante, P. (2011). Arbuscular mycorrhizal hyphopodia and germinated spore exudates trigger Ca²⁺ spiking in the legume and nonlegume root epidermis. *New Phytol*, 189(1):347–55.

-
- Charon, C., Sousa, C., Crespi, M., and Kondorosi, A. (1999). Alteration of *enod40* expression modifies *Medicago truncatula* root nodule development induced by *Sinorhizobium meliloti*. *Plant Cell*, 11(10):1953–66.
- Charpentier, M., Bredemeier, R., Wanner, G., Takeda, N., Schleiff, E., and Parniske, M. (2008). *Lotus japonicus* CASTOR and POLLUX are ion channels essential for perinuclear calcium spiking in legume root endosymbiosis. *Plant Cell*, 20(12):3467–79.
- Charpentier, M. and Oldroyd, G. (2010). How close are we to nitrogen-fixing cereals? *Curr Opin Plant Biol*, 13(5):556–64.
- Charron, D., Pingret, J.-L., Chabaud, M., Journet, E.-P., and Barker, D. G. (2004). Pharmacological evidence that multiple phospholipid signaling pathways link rhizobium nodulation factor perception in *Medicago truncatula* root hairs to intracellular responses, including Ca^{2+} spiking and specific *ENOD* gene expression. *Plant Physiol*, 136(3):3582–3593.
- Chen, C., Ané, J.-M., and Zhu, H. (2008). OsIPD3, an ortholog of the *Medicago truncatula* DMI3 interacting protein IPD3, is required for mycorrhizal symbiosis in rice. *New Phytol*, 180(2):311–5.
- Chen, C., Fan, C., Gao, M., and Zhu, H. (2009). Antiquity and function of *CASTOR* and *POLLUX*, the twin ion channel-encoding genes key to the evolution of root symbioses in plants. *Plant Physiol*, 149(1):306–17.
- Chylla, R. A. and Markley, J. L. (1993). Improved frequency resolution in multidimensional constant-time experiments by multidimensional Bayesian analysis. *J Biomol NMR*, 3(5):515–33.
- Clapham, D. E. (2007). Calcium signaling. *Cell*, 131(6):1047–58.
- Coba de la Peña, T., Cárcamo, C. B., Almonacid, L., Zaballos, A., Lucas, M. M., Balomenos, D., and Pueyo, J. J. (2008a). A cytokinin receptor homologue is induced during root nodule organogenesis and senescence in *Lupinus albus* L. *Plant Physiol Biochem*, 46(2):219–25.
- Coba de la Peña, T., Cárcamo, C. B., Almonacid, L., Zaballos, A., Lucas, M. M., Balomenos, D., and Pueyo, J. J. (2008b). A salt stress-responsive cytokinin receptor homologue isolated from *Medicago sativa* nodules. *Planta*, 227(4):769–79.
- Coba de la Peña, T., Cárcamo, C. B., Lucas, M. M., and Pueyo, J. J. (2008c). Multiple roles for cytokinin receptors and cross-talk of signaling pathways. *Plant Signal Behav*, 3(10):791–4.
- Cook, D. R. (1999). *Medicago truncatula*—a model in the making! *Curr Opin Plant Biol*, 2(4):301–4.
- Cooley, J. and Tukey, J. (1965). An algorithm for machine calculation of complex Fourier series. *Mathematics of Computation*, 19(90):297–301.
- Cordell, D., Drangert, J.-O., and White, S. (2009). The story of phosphorus: Global food security and food for thought. *Global Environmental Change*, 19(2):292–305.
- Crespi, M. and Frugier, F. (2008). De novo organ formation from differentiated cells: root nodule organogenesis. *Sci Signal*, 1(49):re11.

-
- Daubechies, I. (1992). *Ten lectures on wavelets*, volume 61. Society for Industrial and Applied Mathematics, Philadelphia, Pa.
- Davey, M., Webster, G., Manders, G., Ringrose, F., Power, J., and Cocking, E. (1993). Effective nodulation of micro-propagated shoots of the non-legume *Parasponia andersonii* by *Bradyrhizobium*. *J Exp Bot*, 44(262):863–867.
- Davis, R. A., Charlton, A. J., Godward, J., Jones, S. A., Harrison, M., and Wilson, J. C. (2007). Adaptive binning: An improved binning method for metabolomics data using the undecimated wavelet transform. *Chemometrics and Intelligent Laboratory Systems*, 85(1):144–154.
- De Koninck, P. and Schulman, H. (1998). Sensitivity of CaM kinase II to the frequency of Ca^{2+} oscillations. *Science*, 279(5348):227–30.
- Deakin, W. J. and Broughton, W. J. (2009). Symbiotic use of pathogenic strategies: rhizobial protein secretion systems. *Nat Rev Microbiol*, 7(4):312–20.
- Deisseroth, K., Heist, E. K., and Tsien, R. W. (1998). Translocation of calmodulin to the nucleus supports CREB phosphorylation in hippocampal neurons. *Nature*, 392(6672):198–202.
- Dénarié, J., Debelle, F., and Promé, J. C. (1996). Rhizobium lipo-chitooligosaccharide nodulation factors: signaling molecules mediating recognition and morphogenesis. *Annu Rev Biochem*, 65:503–35.
- D’Haeze, W., De Rycke, R., Mathis, R., Goormachtig, S., Pagnotta, S., Verplancke, C., Capoen, W., and Holsters, M. (2003). Reactive oxygen species and ethylene play a positive role in lateral root base nodulation of a semiaquatic legume. *Proc Natl Acad Sci U S A*, 100(20):11789–94.
- D’Haeze, W. and Holsters, M. (2002). Nod factor structures, responses, and perception during initiation of nodule development. *Glycobiology*, 12(6):79R–105R.
- Ding, Y., Kalo, P., Yendrek, C., Sun, J., Liang, Y., Marsh, J. F., Harris, J. M., and Oldroyd, G. E. D. (2008). Abscisic acid coordinates Nod factor and cytokinin signaling during the regulation of nodulation in *Medicago truncatula*. *Plant Cell*, 20(10):2681–2695.
- Ding, Y. and Oldroyd, G. E. (2009). Positioning the nodule, the hormone dictum. *Plant Signal Behav*, 4(2):89–93.
- Dodd, A. N., Kudla, J., and Sanders, D. (2010). The language of calcium signalling. *Annu Rev Plant Biol*, 61:593–620.
- Dolmetsch, R. E., Xu, K., and Lewis, R. S. (1998). Calcium oscillations increase the efficiency and specificity of gene expression. *Nature*, 392(6679):933–6.
- Downie, J. A. (2010). The roles of extracellular proteins, polysaccharides and signals in the interactions of rhizobia with legume roots. *FEMS Microbiol Rev*, 34(2):150–70.
- Doyle, J. J. (2011). Phylogenetic perspectives on the origins of nodulation. *Mol Plant Microbe Interact*, 24(11):1289–95.

-
- Doyle, J. J. and Luckow, M. A. (2003). The rest of the iceberg. Legume diversity and evolution in a phylogenetic context. *Plant Physiol*, 131(3):900–10.
- Eaton, J. W. (2002). *GNU Octave Manual*. Network Theory Limited.
- Eckardt, N. A. (2006). *Medicago truncatula* CRE1 cytokinin receptor regulates nodulation and lateral root development. *Plant Cell*, 18(10):2419.
- Edelstein-Keshet, L. (2005). *Mathematical models in biology*, volume 46. Society for Industrial and Applied Mathematics, Philadelphia.
- Edwards, A., Heckmann, A. B., Yousafzai, F., Duc, G., and Downie, J. A. (2007). Structural implications of mutations in the pea *SYM8* symbiosis gene, the *DMI1* ortholog, encoding a predicted ion channel. *Mol Plant Microbe Interact*, 20(10):1183–91.
- Edwards, K. D., Anderson, P. E., Hall, A., Salathia, N. S., Locke, J. C. W., Lynn, J. R., Straume, M., Smith, J. Q., and Millar, A. J. (2006). FLOWERING LOCUS C mediates natural variation in the high-temperature response of the *Arabidopsis* circadian clock. *Plant Cell*, 18(3):639–50.
- Ehrhardt, D. W., Wais, R., and Long, S. R. (1996). Calcium spiking in plant root hairs responding to Rhizobium nodulation signals. *Cell*, 85(5):673–81.
- Endre, G., Kereszt, A., Kevei, Z., Mihacea, S., Kaló, P., and Kiss, G. B. (2002). A receptor kinase gene regulating symbiotic nodule development. *Nature*, 417(6892):962–6.
- Engstrom, E. M., Ehrhardt, D. W., Mitra, R. M., and Long, S. R. (2002). Pharmacological analysis of Nod factor-induced calcium spiking in *Medicago truncatula*. Evidence for the requirement of type IIA calcium pumps and phosphoinositide signaling. *Plant Physiol*, 128(4):1390–401.
- Ermentrout, B. (2002). *Simulating, analyzing, and animating dynamical systems: a guide to XPPAUT for researchers and students*. Society for Industrial and Applied Mathematics, Philadelphia.
- Esseling, J. J., Lhuissier, F. G. P., and Emons, A. M. C. (2004). A nonsymbiotic root hair tip growth phenotype in NORK-mutated legumes: implications for nodulation factor-induced signaling and formation of a multifaceted root hair pocket for bacteria. *Plant Cell*, 16(4):933–44.
- Evans, N. H., McAinsh, M. R., and Hetherington, A. M. (2001). Calcium oscillations in higher plants. *Curr Opin Plant Biol*, 4(5):415 – 420.
- Falcke, M. (2003a). Buffers and oscillations in intracellular Ca^{2+} dynamics. *Biophys J*, 84(1):28–41.
- Falcke, M. (2003b). On the role of stochastic channel behavior in intracellular Ca^{2+} dynamics. *Biophys J*, 84(1):42–56.
- Falcke, M. (2004). Reading the patterns in living cells —the physics of Ca^{2+} signaling. *Advances in Physics*, 53:255–440.

-
- Fearn, J. C. and Larue, T. A. (1991). Ethylene inhibitors restore nodulation to *sym* 5 mutants of *Pisum sativum* L. cv Sparkle. *Plant Physiol*, 96(1):239–44.
- Felle, H. H., Kondorosi, É., Kondorosi, A., and Schultze, M. (1996). Rapid alkalization in alfalfa root hairs in response to rhizobial lipochitooligosaccharide signals. *Plant J*, 10(2):295–301.
- Felle, H. H., Kondorosi, É., Kondorosi, Á., and Schultze, M. (1998). The role of ion fluxes in Nod factor signalling in *Medicago sativa*. *Plant J*, 13(4):455–463.
- Felle, H. H., Kondorosi, E., Kondorosi, A., and Schultze, M. (1999). Elevation of the cytosolic free $[Ca^{2+}]$ is indispensable for the transduction of the Nod factor signal in alfalfa. *Plant Physiol*, 121(1):273–80.
- Fernandez-Lopez, M., Goormachtig, S., Gao, M., D’Haeze, W., M, V. M., and Holsters, M. (1998). Ethylene-mediated phenotypic plasticity in root nodule development on *Sesbania rostrata*. *Proc Natl Acad Sci U S A*, 95(21):12724–8.
- Fiserova, J., Kiseleva, E., and Goldberg, M. W. (2009). Nuclear envelope and nuclear pore complex structure and organization in tobacco BY-2 cells. *Plant J*, 59(2):243–55.
- Fisher, R. F. and Long, S. R. (1992). Rhizobium–plant signal exchange. *Nature*, 357(6380):655–60.
- Frugier, F., Kosuta, S., Murray, J. D., Crespi, M., and Szczyglowski, K. (2008). Cytokinin: secret agent of symbiosis. *Trends Plant Sci*, 13(3):115–20.
- Galloway, J. N. and Cowling, E. B. (2002). Reactive nitrogen and the world: 200 years of change. *Ambio*, 31(2):64–71.
- Galloway, J. N., Townsend, A. R., Erisman, J. W., Bekunda, M., Cai, Z., Freney, J. R., Martinelli, L. A., Seitzinger, S. P., and Sutton, M. A. (2008). Transformation of the nitrogen cycle: recent trends, questions, and potential solutions. *Science*, 320(5878):889–92.
- Gardner, M. J., Hubbard, K. E., Hotta, C. T., Dodd, A. N., and Webb, A. A. R. (2006). How plants tell the time. *Biochem J*, 397(1):15–24.
- Garrido, J. M. G., Morcillo, R. J. L., Rodríguez, J. A. M., and Bote, J. A. O. (2010). Variations in the mycorrhization characteristics in roots of wild-type and ABA-deficient tomato are accompanied by specific transcriptomic alterations. *Mol Plant Microbe Interact*, 23(5):651–64.
- Genre, A., Chabaud, M., Faccio, A., Barker, D. G., and Bonfante, P. (2008). Prepenetration apparatus assembly precedes and predicts the colonization patterns of arbuscular mycorrhizal fungi within the root cortex of both *Medicago truncatula* and *Daucus carota*. *Plant Cell*, 20(5):1407–20.
- Genre, A., Chabaud, M., Timmers, T., Bonfante, P., and Barker, D. G. (2005). Arbuscular mycorrhizal fungi elicit a novel intracellular apparatus in *Medicago truncatula* root epidermal cells before infection. *Plant Cell*, 17(12):3489–99.

-
- Genre, A., Ortu, G., Bertoldo, C., Martino, E., and Bonfante, P. (2009). Biotic and abiotic stimulation of root epidermal cells reveals common and specific responses to arbuscular mycorrhizal fungi. *Plant Physiol*, 149(3):1424–34.
- Geurts, R. and Franssen, H. (1996). Signal transduction in Rhizobium-induced nodule formation. *Plant Physiol*, 112(2):447–53.
- Gherbi, H., Markmann, K., Svistoonoff, S., Estevan, J., Autran, D., Giczey, G., Auguy, F., Péret, B., Laplaze, L., Franche, C., Parniske, M., and Bogusz, D. (2008). SymRK defines a common genetic basis for plant root endosymbioses with arbuscular mycorrhiza fungi, rhizobia, and *Frankia* bacteria. *Proc Natl Acad Sci U S A*, 105(12):4928–32.
- Ghil, M., Allen, M. R., Dettinger, M. D., Ide, K., Kondrashov, D., Mann, M. E., Robertson, A. W., Saunders, A., Tian, Y., Varadi, F., and Yiou, P. (2002). Advanced spectral methods for climatic time series. *Rev. Geophys.*, 40(1):1003–1.
- Gibbs, J. W. (1899). Fourier’s series. *Nature*, 606.
- Gibson, K. E., Kobayashi, H., and Walker, G. C. (2008). Molecular determinants of a symbiotic chronic infection. *Annu Rev Genet*, 42:413–41.
- Giraud, E., Moulin, L., Vallenet, D., Barbe, V., Cytryn, E., Avarre, J.-C., Jaubert, M., Simon, D., Cartieaux, F., Prin, Y., Bena, G., Hannibal, L., Fardoux, J., Kojadinovic, M., Vuillet, L., Lajus, A., Cruveiller, S., Rouy, Z., Mangenot, S., Segurens, B., Dossat, C., Franck, W. L., Chang, W.-S., Saunders, E., Bruce, D., Richardson, P., Normand, P., Dreyfus, B., Pignol, D., Stacey, G., Emerich, D., Verméglio, A., Médigue, C., and Sadowsky, M. (2007). Legumes symbioses: absence of *Nod* genes in photosynthetic bradyrhizobia. *Science*, 316(5829):1307–12.
- Gleason, C., Chaudhuri, S., Yang, T., Muñoz, A., Poovaiah, B. W., and Oldroyd, G. E. D. (2006). Nodulation independent of rhizobia induced by a calcium-activated kinase lacking autoinhibition. *Nature*, 441(7097):1149–52.
- Godfroy, O., Debelle, F., Timmers, T., and Rosenberg, C. (2006). A rice calcium- and calmodulin-dependent protein kinase restores nodulation to a legume mutant. *Mol Plant Microbe Interact*, 19(5):495–501.
- Gonzalez-Rizzo, S., Crespi, M., and Frugier, F. (2006). The *Medicago truncatula* CRE1 cytokinin receptor regulates lateral root development and early symbiotic interaction with *Sinorhizobium meliloti*. *Plant Cell*, 18(10):2680–93.
- González-Sama, A., Lucas, M. M., de Felipe, M. R., and Pueyo, J. J. (2004). An unusual infection mechanism and nodule morphogenesis in white lupin (*Lupinus albus*). *New Phytol*, 163(2):371–380.
- Goormachtig, S., Capoen, W., James, E. K., and Holsters, M. (2004). Switch from intracellular to intercellular invasion during water stress-tolerant legume nodulation. *Proc Natl Acad Sci U S A*, 101(16):6303–6308.
- Gough, C. and Cullimore, J. (2011). Lipo-chitooligosaccharide signalling in endosymbiotic plant-microbe interactions. *Mol Plant Microbe Interact*.

-
- Granqvist, E., Hartley, M., and Morris, R. J. (2012a). BaSAR - A tool in R for frequency detection. *Biosystems*, 110(1):60–3.
- Granqvist, E., Oldroyd, G. E., and Morris, R. J. (2011). Automated bayesian model development for frequency detection in biological time series. *BMC Syst Biol*, 5(1):97.
- Granqvist, E., Wysham, D., Hazledine, S., Kozłowski, W., Sun, J., Charpentier, M., Martins, T. V., Haleux, P., Tsaneva-Atanasova, K., Downie, J. A., Oldroyd, G. E., and Morris, R. J. (2012b). Buffering capacity explains signal variation in symbiotic calcium oscillations. *Plant Physiology*, 160(4):2300–2310.
- Graumann, K. and Evans, D. E. (2010). The plant nuclear envelope in focus. *Biochem Soc Trans*, 38(Pt 1):307–11.
- Groth, M., Takeda, N., Perry, J., Uchida, H., Dräxl, S., Brachmann, A., Sato, S., Tabata, S., Kawaguchi, M., Wang, T. L., and Parniske, M. (2010). *NENA*, a *Lotus japonicus* homolog of *Sec13*, is required for rhizodermal infection by arbuscular mycorrhiza fungi and rhizobia but dispensable for cortical endosymbiotic development. *Plant Cell*, 22(7):2509–26.
- Groves, P., Offermann, S., Rasmussen, M. O., Canada, F. J., Bono, J.-J., Driguez, H., Imbert, A., and Jimenez-Barbero, J. (2005). The relative orientation of the lipid and carbohydrate moieties of lipochitooligosaccharides related to nodulation factors depends on lipid chain saturation. *Org Biomol Chem*, 3(8):1381–6.
- Gruber, N. and Galloway, J. N. (2008). An earth-system perspective of the global nitrogen cycle. *Nature*, 451(7176):293–6.
- Guinel, F. and Geil, R. (2002). A model for the development of the rhizobial and arbuscular mycorrhizal symbioses in legumes and its use to understand the roles of ethylene in the establishment of these two symbioses. *Can J Bot*, 80:695–720.
- Gutjahr, C., Banba, M., Croset, V., An, K., Miyao, A., An, G., Hirochika, H., Imaizumi-Anraku, H., and Paszkowski, U. (2008). Arbuscular mycorrhiza-specific signaling in rice transcends the common symbiosis signaling pathway. *Plant Cell*, 20(11):2989–3005.
- Gutjahr, C., Casieri, L., and Paszkowski, U. (2009a). *Glomus intraradices* induces changes in root system architecture of rice independently of common symbiosis signaling. *New Phytol*, 182(4):829–37.
- Gutjahr, C., Novero, M., Guether, M., Montanari, O., Udvardi, M., and Bonfante, P. (2009b). Presymbiotic factors released by the arbuscular mycorrhizal fungus *Gigaspora margarita* induce starch accumulation in *Lotus japonicus* roots. *New Phytol*, 183(1):53–61.
- Hamel, C. and Plenchette, C. (2007). *Mycorrhizae in crop production*. Food Products Press, Binghamton, NY.
- Hamel, L.-P. and Beaudoin, N. (2010). Chitooligosaccharide sensing and downstream signaling: contrasted outcomes in pathogenic and beneficial plant-microbe interactions. *Planta*, 232(4):787–806.
- Haney, C. H. and Long, S. R. (2010). Plant flotillins are required for infection by nitrogen-fixing bacteria. *Proc Natl Acad Sci U S A*, 107(1):478–83.

-
- Harris, J. M., Wais, R., and Long, S. R. (2003). Rhizobium-induced calcium spiking in *Lotus japonicus*. *Mol Plant Microbe Interact*, 16(4):335–41.
- Harrison, M. J. (2005). Signaling in the arbuscular mycorrhizal symbiosis. *Annu Rev Microbiol*, 59:19–42.
- Hayashi, T., Banba, M., Shimoda, Y., Kouchi, H., Hayashi, M., and Imaizumi-Anraku, H. (2010). A dominant function of CCaMK in intracellular accommodation of bacterial and fungal endosymbionts. *Plant J*, 63(1):141–154.
- Hazledine, S. (2010). *The analysis and modelling of chaotic calcium oscillations*. PhD thesis, University of East Anglia.
- Hazledine, S., Sun, J., Wysham, D., Downie, J. A., Oldroyd, G. E. D., and Morris, R. J. (2009). Nonlinear time series analysis of nodulation factor induced calcium oscillations: evidence for deterministic chaos? *PLoS One*, 4(8):e6637.
- He, Q., Wang, L., and Liu, B. (2007). Parameter estimation for chaotic systems by particle swarm optimization. *Chaos, Solitons and Fractals*, 34(2):654 – 661.
- Heckmann, A. B., Lombardo, F., Miwa, H., Perry, J. A., Bunnewell, S., Parniske, M., Wang, T. L., and Downie, J. A. (2006). *Lotus japonicus* nodulation requires two GRAS domain regulators, one of which is functionally conserved in a non-legume. *Plant Physiol*, 142(4):1739–50.
- Heckmann, A. B., Sandal, N., Bek, A. S., Madsen, L. H., Jurkiewicz, A., Nielsen, M. W., Tirichine, L., and Stougaard, J. (2011). Cytokinin induction of root nodule primordia in *Lotus japonicus* is regulated by a mechanism operating in the root cortex. *Mol Plant Microbe Interact*.
- Heil, C. and Walnut, D. F. (2006). *Fundamental papers in wavelet theory*. Princeton University Press, Princeton, N.J.
- Hirsch, A. M., Yu, N., Ma, N., Schwartz, A. R., and De Hoff, P. (2009a). The *Chamaecrista fasciculata* nitrogen-fixing symbiosis. In *Botany and Mycology, abstract 446*.
- Hirsch, S., Kim, J., Muñoz, A., Heckmann, A. B., Downie, J. A., and Oldroyd, G. E. D. (2009b). GRAS proteins form a DNA binding complex to induce gene expression during nodulation signaling in *Medicago truncatula*. *Plant Cell*, 21(2):545–57.
- Hoche, V., Alloisio, N., Auguy, F., Fournier, P., Dumas, P., Pujic, P., Gherbi, H., Queiroux, C., Da Silva, C., Wincker, P., Normand, P., and Bogusz, D. (2011). Transcriptomics of actinorhizal symbioses reveals homologs of the whole common symbiotic signaling cascade. *Plant Physiol*.
- Hogan, P. G., Lewis, R. S., and Rao, A. (2010). Molecular basis of calcium signaling in lymphocytes: STIM and ORAI. *Annu Rev Immunol*, 28:491–533.
- Hogg, B. V., Cullimore, J. V., Ranjeva, R., and Bono, J.-J. (2006). The *DMI1* and *DMI2* early symbiotic genes of *Medicago truncatula* are required for a high-affinity nodulation factor-binding site associated to a particulate fraction of roots. *Plant Physiol*, 140(1):365–73.

-
- Holton, M. L., Wang, W., Emerson, M., Neyses, L., and Armesilla, A. L. (2010). Plasma membrane calcium ATPase proteins as novel regulators of signal transduction pathways. *World J Biol Chem*, 1(6):201–8.
- Howarth, R. W., Boyer, E. W., Pabich, W. J., and Galloway, J. N. (2002). Nitrogen use in the united states from 1961-2000 and potential future trends. *Ambio*, 31(2):88–96.
- Humphreys, C. P., Franks, P. J., Rees, M., Bidartondo, M. I., Leake, J. R., and Beerling, D. J. (2010). Mutualistic mycorrhiza-like symbiosis in the most ancient group of land plants. *Nat Commun*, 1:103.
- Izhikevich, E. M. (2007). *Dynamical systems in neuroscience: the geometry of excitability and bursting*. MIT Press, Cambridge, Mass.
- Javot, H., Penmetsa, R. V., Terzaghi, N., Cook, D. R., and Harrison, M. J. (2007a). A *Medicago truncatula* phosphate transporter indispensable for the arbuscular mycorrhizal symbiosis. *Proc Natl Acad Sci U S A*, 104(5):1720–5.
- Javot, H., Pumplin, N., and Harrison, M. J. (2007b). Phosphate in the arbuscular mycorrhizal symbiosis: transport properties and regulatory roles. *Plant Cell Environ*, 30(3):310–22.
- Jaynes, E. T. and Bretthorst, G. L. (2003). *Probability theory: the logic of science*. Cambridge University Press, Cambridge, UK.
- Kaló, P., Gleason, C., Edwards, A., Marsh, J., Mitra, R. M., Hirsch, S., Jakab, J., Sims, S., Long, S. R., Rogers, J., Kiss, G. B., Downie, J. A., and Oldroyd, G. E. D. (2005). Nodulation signaling in legumes requires NSP2, a member of the GRAS family of transcriptional regulators. *Science*, 308(5729):1786–9.
- Kammler, D. W. (2007). *A first course in fourier analysis*. Cambridge University Press, Cambridge, 2nd edition.
- Kanamori, N., Madsen, L. H., Radutoiu, S., Frantescu, M., Quistgaard, E. M. H., Miwa, H., Downie, J. A., James, E. K., Felle, H. H., Haaning, L. L., Jensen, T. H., Sato, S., Nakamura, Y., Tabata, S., Sandal, N., and Stougaard, J. (2006). A nucleoporin is required for induction of Ca²⁺ spiking in legume nodule development and essential for rhizobial and fungal symbiosis. *Proc Natl Acad Sci U S A*, 103(2):359–64.
- Karas, B., Murray, J., Gorzelak, M., Smith, A., Sato, S., Tabata, S., and Szczyglowski, K. (2005). Invasion of *Lotus japonicus* root hairless 1 by *Mesorhizobium loti* involves the nodulation factor-dependent induction of root hairs. *Plant Physiol*, 137(4):1331–44.
- Kiss, E., Oláh, B., Kaló, P., Morales, M., Heckmann, A. B., Borbola, A., Lózsa, A., Kontár, K., Middleton, P., Downie, J. A., Oldroyd, G. E. D., and Endre, G. (2009). LIN, a novel type of U-box/WD40 protein, controls early infection by rhizobia in legumes. *Plant Physiol*, 151(3):1239–49.
- Kistner, C. and Parniske, M. (2002). Evolution of signal transduction in intracellular symbiosis. *Trends Plant Sci*, 7(11):511–8.
- Kloppholz, S., Kuhn, H., and Requena, N. (2011). A secreted fungal effector of *Glomus intraradices* promotes symbiotic biotrophy. *Curr Biol*, 21(14):1204–9.

-
- Kosuta, S., Hazledine, S., Sun, J., Miwa, H., Morris, R. J., Downie, J. A., and Oldroyd, G. E. D. (2008). Differential and chaotic calcium signatures in the symbiosis signaling pathway of legumes. *Proc Natl Acad Sci U S A*, 105(28):9823–8.
- Kosuta, S., Held, M., Hossain, M. S., Morieri, G., Macgillivray, A., Johansen, C., Antolín-Llovera, M., Parniske, M., Oldroyd, G. E. D., Allan Downie, J., Karas, B., and Szczyglowski, K. (2011). *Lotus japonicus symRK-14* uncouples the cortical and epidermal symbiotic program. *Plant J*.
- Kouchi, H., Imaizumi-Anraku, H., Hayashi, M., Hakoyama, T., Nakagawa, T., Umehara, Y., Suganuma, N., and Kawaguchi, M. (2010). How many peas in a pod? Legume genes responsible for mutualistic symbioses underground. *Plant Cell Physiol*, 51(9):1381–97.
- Kudla, J., Batistic, O., and Hashimoto, K. (2010). Calcium signals: The lead currency of plant information processing. *Plant Cell*.
- Kummer, U., Olsen, L. F., Dixon, C. J., Green, A. K., Bornberg-Bauer, E., and Baier, G. (2000). Switching from simple to complex oscillations in calcium signaling. *Biophys J*, 79(3):1188–95.
- Lafay, B., Bullier, E., and Burdon, J. J. (2006). Bradyrhizobia isolated from root nodules of *Parasponia* (Ulmaceae) do not constitute a separate coherent lineage. *Int J Syst Evol Microbiol*, 56(Pt 5):1013–8.
- Lancelle, S. A. and Torrey, J. G. (1984). Early development of Rhizobium-induced root nodules of *Parasponia rigida*. I. Infection and early nodule initiation. *Protoplasma*, 123:26–37.
- Laplaze, L., Benkova, E., Casimiro, I., Maes, L., Vanneste, S., Swarup, R., Weijers, D., Calvo, V., Parizot, B., Herrera-Rodriguez, M. B., Offringa, R., Graham, N., Doumas, P., Friml, J., Bogusz, D., Beekman, T., and Bennett, M. (2007). Cytokinins act directly on lateral root founder cells to inhibit root initiation. *Plant Cell*, 19(12):3889–900.
- Lavin, M., Herendeen, P. S., and Wojciechowski, M. F. (2005). Evolutionary rates analysis of Leguminosae implicates a rapid diversification of lineages during the tertiary. *Syst Biol*, 54(4):575–94.
- Lebaudy, A., Véry, A.-A., and Sentenac, H. (2007). K⁺ channel activity in plants: Genes, regulations and functions. *FEBS Lett*, 581(12):2357–66.
- Lecourieux, D., Ranjeva, R., and Pugin, A. (2006). Calcium in plant defence-signalling pathways. *New Phytol*, 171(2):249–69.
- Lefebvre, B., Timmers, T., Mbengue, M., Moreau, S., Hervé, C., Tóth, K., Bittencourt-Silvestre, J., Klaus, D., Deslandes, L., Godiard, L., Murray, J. D., Udvardi, M. K., Raffaele, S., Mongrand, S., Cullimore, J., Gamas, P., Niebel, A., and Ott, T. (2010). A remorin protein interacts with symbiotic receptors and regulates bacterial infection. *Proc Natl Acad Sci U S A*, 107(5):2343–8.
- Lerouge, P., Roche, P., Faucher, C., Maillet, F., Truchet, G., Promé, J. C., and Dénarié, J. (1990). Symbiotic host-specificity of *Rhizobium meliloti* is determined by a sulphated and acylated glucosamine oligosaccharide signal. *Nature*, 344(6268):781–4.

-
- Lévy, J., Bres, C., Geurts, R., Chalhoub, B., Kulikova, O., Duc, G., Journet, E.-P., Ané, J.-M., Lauber, E., Bisseling, T., Dénarié, J., Rosenberg, C., and Debellé, F. (2004). A putative Ca^{2+} and calmodulin-dependent protein kinase required for bacterial and fungal symbioses. *Science*, 303(5662):1361–4.
- Liao, J., Singh, S., Hossain, M. S., Andersen, S. U., Ross, L., Bonetta, D., Zhou, Y., Sato, S., Tabata, S., Stougaard, J., Szczyglowski, K., and Parniske, M. (2012). Negative regulation of CCaMK is essential for symbiotic infection. *Plant J.*
- Limpens, E., Franken, C., Smit, P., Willemse, J., Bisseling, T., and Geurts, R. (2003). LysM domain receptor kinases regulating rhizobial Nod factor-induced infection. *Science*, 302(5645):630–3.
- Limpens, E., Ovchinnikova, E., Journet, E.-P., Chabaud, M., Cosson, V., Ratet, P., Duc, G., Fedorova, E., Liu, W., Op den Camp, R., Zhukov, V., Tikhonovich, I., Borisov, A., and Bisseling, T. (2011). IPD3 controls the formation of nitrogen-fixing symbiosomes in pea and *Medicago*. *Mol Plant Microbe Interact.*
- Lisman, J., Schulman, H., and Cline, H. (2002). The molecular basis of CaMKII function in synaptic and behavioural memory. *Nat Rev Neurosci*, 3(3):175–90.
- Liu, Y., Villalba, G., Ayres, R., and Schroder, H. (2008). Global phosphorus flows and environmental impacts from a consumption perspective. *Journal of Industrial Ecology*, 12(2):229–247.
- Lloyd, G.; McCown, B. (1980). Commercially-feasible micropropagation of mountain laurel, *Kalmia latifolia*, by use of shoot-tip culture. *Combined Proceedings, International Plant Propagators' Society*, 30:421–427.
- Long, S. R. (2001). Genes and signals in the rhizobium-legume symbiosis. *Plant Physiol*, 125(1):69–72.
- Luan, S. (2011). *Coding and decoding of calcium signals in plants*. Signaling and communication in plants. Springer, Heidelberg.
- Lucier, B. J., Kallergi, M., Qian, W., DeVore, R. A., Clark, R. A., Saff, E. B., and Clarke, L. P. (1994). Wavelet compression and segmentation of digital mammograms. *J Digit Imaging*, 7(1):27–38.
- Luciński, R., Polcyn, W., and Ratajczak, L. (2002). Nitrate reduction and nitrogen fixation in symbiotic association Rhizobium-legumes. *Acta Biochim Pol*, 49(2):537–46.
- MacKay, D. J. C. (2003). *Information theory, inference, and learning algorithms*. Cambridge University Press, Cambridge, UK.
- Madsen, E. B., Madsen, L. H., Radutoiu, S., Olbryt, M., Rakwalska, M., Szczyglowski, K., Sato, S., Kaneko, T., Tabata, S., Sandal, N., and Stougaard, J. (2003). A receptor kinase gene of the LysM type is involved in legume perception of rhizobial signals. *Nature*, 425(6958):637–40.

-
- Madsen, L. H., Tirichine, L., Jurkiewicz, A., Sullivan, J. T., Heckmann, A. B., Bek, A. S., Ronson, C. W., James, E. K., and Stougaard, J. (2010). The molecular network governing nodule organogenesis and infection in the model legume *Lotus japonicus*. *Nat Commun*, 1:10.
- Maeda, D., Ashida, K., Iguchi, K., Chechetka, S. A., Hijikata, A., Okusako, Y., Deguchi, Y., Izui, K., and Hata, S. (2006). Knockdown of an arbuscular mycorrhiza-inducible phosphate transporter gene of *Lotus japonicus* suppresses mutualistic symbiosis. *Plant Cell Physiol*, 47(7):807–17.
- Maillet, F., Poinso, V., André, O., Puech-Pagès, V., Haouy, A., Gueunier, M., Cromer, L., Giraudet, D., Formey, D., Niebel, A., Martinez, E. A., Driguez, H., Bécard, G., and Dénarié, J. (2011). Fungal lipochitooligosaccharide symbiotic signals in arbuscular mycorrhiza. *Nature*, 469(7328):58–63.
- Mallat, S. G. (1999). *A wavelet tour of signal processing*. Academic Press, San Diego, 2nd edition.
- Marhl, M., Schuster, S., Brumen, M., and Heinrich, R. (1997). Modeling the interrelations between the calcium oscillations and ER membrane potential oscillations. *Biophys Chem*, 63(2-3):221–39.
- Markmann, K., Giczey, G., and Parniske, M. (2008). Functional adaptation of a plant receptor-kinase paved the way for the evolution of intracellular root symbioses with bacteria. *PLoS Biol*, 6(3):e68.
- Markmann, K. and Parniske, M. (2009). Evolution of root endosymbiosis with bacteria: How novel are nodules? *Trends Plant Sci*, 14(2):77–86.
- Marsh, J. F., Rakocevic, A., Mitra, R. M., Brocard, L., Sun, J., Eschstruth, A., Long, S. R., Schultze, M., Ratet, P., and Oldroyd, G. E. D. (2007). *Medicago truncatula* *NIN* is essential for rhizobial-independent nodule organogenesis induced by autoactive calcium/calmodulin-dependent protein kinase. *Plant Physiol*, 144(1):324–35.
- Marvel, D. J., Torrey, J. G., and Ausubel, F. M. (1987). Rhizobium symbiotic genes required for nodulation of legume and nonlegume hosts. *Proc Natl Acad Sci U S A*, 84(5):1319–23.
- Masson-Boivin, C., Giraud, E., Perret, X., and Batut, J. (2009). Establishing nitrogen-fixing symbiosis with legumes: how many rhizobium recipes? *Trends Microbiol*, 17(10):458–66.
- Mathesius, U., Schlaman, H. R., Spaink, H. P., Of Sautter, C., Rolfe, B. G., and Djordjevic, M. A. (1998). Auxin transport inhibition precedes root nodule formation in white clover roots and is regulated by flavonoids and derivatives of chitin oligosaccharides. *Plant J*, 14(1):23–34.
- Matzke, M. A. and Matzke, A. J. M. (1986). Visualization of mitochondria and nuclei in living plant cells by the use of a potential-sensitive fluorescent dye. *Plant, Cell and Environment*, 9(1):73–77.
- Mazars, C., Brière, C., Bourque, S., and Thuleau, P. (2011). Nuclear calcium signaling: An emerging topic in plants. *Biochimie*.
- McAinsh, M. R. and Pittman, J. K. (2009). Shaping the calcium signature. *New Phytol*, 181(2):275–94.

-
- McAinsh, M. R., Webb, A., Taylor, J. E., and Hetherington, A. M. (1995). Stimulus-induced oscillations in guard cell cytosolic free calcium. *Plant Cell*, 7(8):1207–1219.
- Meier, J. E. and Marshall, A. G. (1990). Bayesian versus Fourier spectral analysis of ion cyclotron resonance time-domain signals. *Anal Chem*, 62(2):201–8.
- Messinese, E., Mun, J.-H., Yeun, L. H., Jayaraman, D., Rougé, P., Barre, A., Loughon, G., Schornack, S., Bono, J.-J., Cook, D. R., and Ané, J.-M. (2007). A novel nuclear protein interacts with the symbiotic DMI3 calcium- and calmodulin-dependent protein kinase of *Medicago truncatula*. *Mol Plant Microbe Interact*, 20(8):912–21.
- Middleton, P. H., Jakab, J., Penmetsa, R. V., Starker, C. G., Doll, J., Kalo, P., Prabhu, R., Marsh, J. F., Mitra, R. M., Kereszt, A., Dudas, B., VandenBosch, K., Long, S. R., Cook, D. R., Kiss, G. B., and Oldroyd, G. E. D. (2007). An ERF transcription factor in *Medicago truncatula* that is essential for Nod factor signal transduction. *Plant Cell*, 19(4):1221–1234.
- Mitra, R. M., Gleason, C. A., Edwards, A., Hadfield, J., Downie, J. A., Oldroyd, G. E. D., and Long, S. R. (2004). A Ca^{2+} /calmodulin-dependent protein kinase required for symbiotic nodule development: gene identification by transcript-based cloning. *Proc Natl Acad Sci U S A*, 101(13):4701–4705.
- Miwa, H., Sun, J., Oldroyd, G. E. D., and Downie, J. A. (2006). Analysis of calcium spiking using a cameleon calcium sensor reveals that nodulation gene expression is regulated by calcium spike number and the developmental status of the cell. *Plant J*, 48(6):883–894.
- Miyawaki, A., Griesbeck, O., Heim, R., and Tsien, R. Y. (1999). Dynamic and quantitative Ca^{2+} measurements using improved cameleons. *Proc Natl Acad Sci U S A*, 96(5):2135–40.
- Morando, M. A., Nurisso, A., Grenouillat, N., Vauzeilles, B., Beau, J.-M., Cañada, F. J., Jiménez-Barbero, J., and Imberty, A. (2011). NMR and molecular modelling reveal key structural features of synthetic nodulation factors. *Glycobiology*.
- Mortier, V., Den Herder, G., Whitford, R., Van de Velde, W., Rombauts, S., D’Haeseleer, K., Holsters, M., and Goormachtig, S. (2010). CLE peptides control *Medicago truncatula* nodulation locally and systemically. *Plant Physiol*, 153(1):222–37.
- Mortier, V., Holsters, M., and Goormachtig, S. (2011). Never too many? How legumes control nodule numbers. *Plant Cell Environ*.
- Mukherjee, A. and Ané, J.-M. (2011). Germinating spore exudates from arbuscular mycorrhizal fungi: molecular and developmental responses in plants and their regulation by ethylene. *Mol Plant Microbe Interact*, 24(2):260–70.
- Mulder, L., Lefebvre, B., Cullimore, J., and Imberty, A. (2006). LysM domains of *Medicago truncatula* NFP protein involved in Nod factor perception. Glycosylation state, molecular modeling and docking of chitoooligosaccharides and Nod factors. *Glycobiology*, 16(9):801–9.
- Murray, J. D. (2011). Invasion by invitation: rhizobial infection in legumes. *Mol Plant Microbe Interact*, 24(6):631–9.

-
- Murray, J. D., Karas, B. J., Sato, S., Tabata, S., Amyot, L., and Szczyglowski, K. (2007). A cytokinin perception mutant colonized by *Rhizobium* in the absence of nodule organogenesis. *Science*, 315(5808):101–4.
- Murray, J. D., Muni, R. R. D., Torres-Jerez, I., Tang, Y., Allen, S., Andriankaja, M., Li, G., Laxmi, A., Cheng, X., Wen, J., Vaughan, D., Schultze, M., Sun, J., Charpentier, M., Oldroyd, G., Tadege, M., Ratet, P., Mysore, K. S., Chen, R., and Udvardi, M. K. (2011). *Vapyrin*, a gene essential for intracellular progression of arbuscular mycorrhizal symbiosis, is also essential for infection by rhizobia in the nodule symbiosis of *Medicago truncatula*. *Plant J*, 65(2):244–52.
- Naisbitt, T., James, E. K., and Sprent, J. I. (1992). The evolutionary significance of the legume genus *Chamaecrista*, as determined by nodule structure. *New Phytol*, 122(3):487–492.
- Nakagawa, T., Kaku, H., Shimoda, Y., Sugiyama, A., Shimamura, M., Takashi, K., Yazaki, K., Aoki, T., Shibuya, N., and Kouchi, H. (2011). From defense to symbiosis: limited alterations in the kinase domain of LysM receptor-like kinases are crucial for evolution of legume-*Rhizobium* symbiosis. *Plant J*, 65(2):169–80.
- Navazio, L., Moscatiello, R., Genre, A., Novero, M., Baldan, B., Bonfante, P., and Mariani, P. (2007). A diffusible signal from arbuscular mycorrhizal fungi elicits a transient cytosolic calcium elevation in host plant cells. *Plant Physiol*, 144(2):673–81.
- Neil, J. J. and Bretthorst, G. L. (1993). On the use of Bayesian probability theory for analysis of exponential decay data: An example taken from intravoxel incoherent motion experiments. *Magn Reson Med*, 29(5):642–7.
- Newman, J. D., Diebold, R. J., Schultz, B. W., and Noel, K. D. (1994). Infection of soybean and pea nodules by *Rhizobium* spp. purine auxotrophs in the presence of 5-aminoimidazole-4-carboxamide riboside. *J Bacteriol*, 176(11):3286–94.
- Normand, P., Lapierre, P., Tisa, L. S., Gogarten, J. P., Alloisio, N., Bagnarol, E., Bassi, C. A., Berry, A. M., Bickhart, D. M., Choisine, N., Couloux, A., Cournoyer, B., Cruveiller, S., Daubin, V., Demange, N., Francino, M. P., Goltsman, E., Huang, Y., Kopp, O. R., Labarre, L., Lapidus, A., Lavire, C., Marechal, J., Martinez, M., Mastrorunzio, J. E., Mullin, B. C., Niemann, J., Pujic, P., Rawnsley, T., Rouy, Z., Schenowitz, C., Sellstedt, A., Tavares, F., Tomkins, J. P., Vallenet, D., Valverde, C., Wall, L. G., Wang, Y., Medigue, C., and Benson, D. R. (2007). Genome characteristics of facultatively symbiotic *Frankia* sp. strains reflect host range and host plant biogeography. *Genome Res*, 17(1):7–15.
- Novikova, T. I. and Gordienko, N. I. (2001). Study of the root nodules in some species of the Papilionaceae subfamily by scanning electron microscopy. *Tsitologiya*, 43(2):107–13.
- Oba, H., Tawaray, K., and Wagatsuma, T. (2001). Arbuscular mycorrhizal colonization in *Lupinus* and related genera. *Soil Science and Plant Nutrition*, 47(4):685–694.
- Oláh, B., Brière, C., Bécard, G., Dénarié, J., and Gough, C. (2005). Nod factors and a diffusible factor from arbuscular mycorrhizal fungi stimulate lateral root formation in *Medicago truncatula* via the DMI1/DMI2 signalling pathway. *Plant J*, 44(2):195–207.

-
- Oldroyd, G. E., Engstrom, E. M., and Long, S. R. (2001a). Ethylene inhibits the Nod factor signal transduction pathway of *Medicago truncatula*. *Plant Cell*, 13(8):1835–1849.
- Oldroyd, G. E., Mitra, R. M., Wais, R. J., and Long, S. R. (2001b). Evidence for structurally specific negative feedback in the Nod factor signal transduction pathway. *Plant J*, 28(2):191–199.
- Oldroyd, G. E. D. (2007). Nodules and hormones. *Science*, 315(5808):52–3.
- Oldroyd, G. E. D. and Downie, J. A. (2004). Calcium, kinases and nodulation signalling in legumes. *Nat Rev Mol Cell Biol*, 5(7):566–76.
- Oldroyd, G. E. D. and Downie, J. A. (2006). Nuclear calcium changes at the core of symbiosis signalling. *Curr Opin Plant Biol*, 9(4):351–357.
- Oldroyd, G. E. D. and Downie, J. A. (2008). Coordinating nodule morphogenesis with rhizobial infection in legumes. *Annu Rev Plant Biol*, 59:519–46.
- Oldroyd, G. E. D., Harrison, M. J., and Paszkowski, U. (2009). Reprogramming plant cells for endosymbiosis. *Science*, 324(5928):753–4.
- Oldroyd, G. E. D. and Long, S. R. (2003). Identification and characterization of *nodulation-signaling pathway 2*, a gene of *Medicago truncatula* involved in Nod factor signaling. *Plant Physiol*, 131(3):1027–32.
- Oldroyd, G. E. D., Murray, J. D., Poole, P. S., and Downie, J. A. (2011). The rules of engagement in the legume-rhizobial symbiosis. *Annu Rev Genet*, 45:119–44.
- Op den Camp, R., Streng, A., De Mita, S., Cao, Q., Polone, E., Liu, W., Ammiraju, J. S. S., Kudrna, D., Wing, R., Untergasser, A., Bisseling, T., and Geurts, R. (2011). LysM-type mycorrhizal receptor recruited for rhizobium symbiosis in nonlegume *Parasponia*. *Science*, 331(6019):909–12.
- Op den Camp, R. H. M., Polone, E., Fedorova, E., Roelofsen, W., Squartini, A., Op den Camp, H. J. M., Bisseling, T., and Geurts, R. (2012). Nonlegume *Parasponia andersonii* deploys a broad rhizobium host range strategy resulting in largely variable symbiotic effectiveness. *Mol Plant Microbe Interact*, 25(7):954–63.
- Parekh, A. B. (2011). Decoding cytosolic Ca²⁺ oscillations. *Trends Biochem Sci*, 36(2):78–87.
- Parniske, M. (2008). Arbuscular mycorrhiza: the mother of plant root endosymbioses. *Nat Rev Microbiol*, 6(10):763–75.
- Paszkowski, U. (2006). Mutualism and parasitism: the yin and yang of plant symbioses. *Curr Opin Plant Biol*, 9(4):364–70.
- Pawlowski, K. and Bisseling, T. (1996). Rhizobial and actinorhizal symbioses: What are the shared features? *Plant Cell*, 8(10):1899–1913.
- Pei, Z. M., Murata, Y., Benning, G., Thomine, S., Klüsener, B., Allen, G. J., Grill, E., and Schroeder, J. I. (2000). Calcium channels activated by hydrogen peroxide mediate abscisic acid signalling in guard cells. *Nature*, 406(6797):731–4.

-
- Pellock, B. J., Cheng, H. P., and Walker, G. C. (2000). Alfalfa root nodule invasion efficiency is dependent on *Sinorhizobium meliloti* pepske polysaccharides. *J Bacteriol*, 182(15):4310–8.
- Pepke, S., Kinzer-Ursem, T., Mihalas, S., and Kennedy, M. B. (2010). A dynamic model of interactions of Ca^{2+} , calmodulin, and catalytic subunits of Ca^{2+} /calmodulin-dependent protein kinase II. *PLoS Comput Biol*, 6(2):e1000675.
- Péret, B., De Rybel, B., Casimiro, I., Benková, E., Swarup, R., Laplaze, L., Beeckman, T., and Bennett, M. J. (2009). Arabidopsis lateral root development: an emerging story. *Trends Plant Sci*, 14(7):399–408.
- Perret, X., Staehelin, C., and Broughton, W. J. (2000). Molecular basis of symbiotic promiscuity. *Microbiol Mol Biol Rev*, 64(1):180–201.
- Petersen, O. H., Gerasimenko, O. V., Gerasimenko, J. V., Mogami, H., and Tepikin, A. V. (1998). The calcium store in the nuclear envelope. *Cell Calcium*, 23(2-3):87–90.
- Petrosian, A. A. and Meyer, F. G. (2001). *Wavelets in signal and image analysis: from theory to practice*, volume 19. Kluwer Academic, Dordrecht.
- Pingret, J.-L., Journet, E.-P., and Barker, D. G. (1998). Rhizobium Nod factor signaling: Evidence for a G protein-mediated transduction mechanism. *Plant Cell*, 10(5):659–671.
- Plet, J., Wasson, A., Ariel, F., Le Signor, C., Baker, D., Mathesius, U., Crespi, M., and Frugier, F. (2011). MtCRE1-dependent cytokinin signaling integrates bacterial and plant cues to coordinate symbiotic nodule organogenesis in *Medicago truncatula*. *Plant J*, 65(4):622–33.
- Plett, J. M., Kemppainen, M., Kale, S. D., Kohler, A., Legué, V., Brun, A., Tyler, B. M., Pardo, A. G., and Martin, F. (2011). A secreted effector protein of *Laccaria bicolor* is required for symbiosis development. *Curr Biol*, 21(14):1197–203.
- Prakriya, M., Feske, S., Gwack, Y., Srikanth, S., Rao, A., and Hogan, P. G. (2006). Orai1 is an essential pore subunit of the CRAC channel. *Nature*, 443(7108):230–3.
- Prasad, L. and Iyengar, S. S. (1997). *Wavelet analysis with applications to image processing*. CRC Press, Boca Raton.
- Prell, J., White, J. P., Bourdes, A., Bunnewell, S., Bongaerts, R. J., and Poole, P. S. (2009). Legumes regulate Rhizobium bacteroid development and persistence by the supply of branched-chain amino acids. *Proc Natl Acad Sci U S A*, 106(30):12477–82.
- Press, W. H. (2002). *Numerical recipes in C++: the art of scientific computing*. Cambridge University Press, Cambridge, UK, 2nd edition.
- Price, N. P., Relić, B., Talmont, F., Lewin, A., Promé, D., Pueppke, S. G., Maillet, F., Dénarié, J., Promé, J. C., and Broughton, W. J. (1992). Broad-host-range *Rhizobium* species strain NGR234 secretes a family of carbamoylated, and fucosylated, nodulation signals that are o-acetylated or sulphated. *Mol Microbiol*, 6(23):3575–84.
- Pruneda-Paz, J. L. and Kay, S. A. (2010). An expanding universe of circadian networks in higher plants. *Trends Plant Sci*, 15(5):259–65.

-
- Pueppke, S. G. and Broughton, W. J. (1999). *Rhizobium* sp. strain NGR234 and *R. fredii* USDA257 share exceptionally broad, nested host ranges. *Mol Plant Microbe Interact*, 12(4):293–318.
- Putney, J. W. (2006). *Calcium signaling*. CRC Taylor and Francis, Boca Raton, 2nd edition.
- Putney, Jr, J. W. (1998). Calcium signaling: up, down, up, down...what's the point? *Science*, 279(5348):191–2.
- R Development Core Team (2008). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.
- Radutoiu, S., Madsen, L. H., Madsen, E. B., Felle, H. H., Umehara, Y., Grønlund, M., Sato, S., Nakamura, Y., Tabata, S., Sandal, N., and Stougaard, J. (2003). Plant recognition of symbiotic bacteria requires two LysM receptor-like kinases. *Nature*, 425(6958):585–92.
- Räsänen, L. A., Sprent, J. I., and Lindström, K. (2001). Symbiotic properties of sinorhizobia isolated from *Acacia* and *Prosopis* nodules in sudan and senegal. *Plant and Soil*, 235:193–210.
- Redecker, D., Kodner, R., and Graham, L. E. (2000a). Glomalean fungi from the ordovician. *Science*, 289(5486):1920–1.
- Redecker, D., Morton, J. B., and Bruns, T. D. (2000b). Ancestral lineages of arbuscular mycorrhizal fungi (Glomales). *Mol Phylogenet Evol*, 14(2):276–84.
- Remy, W., Taylor, T. N., Hass, H., and Kerp, H. (1994). Four hundred-million-year-old vesicular arbuscular mycorrhizae. *Proc Natl Acad Sci U S A*, 91(25):11841–3.
- Ricke, J., Maass, P., Lopez Hänninen, E., Liebig, T., Amthauer, H., Stroszczyński, C., Schauer, W., Boskamp, T., and Wolf, M. (1998). Wavelet versus JPEG (Joint Photographic Expert Group) and fractal compression. Impact on the detection of low-contrast details in computed radiographs. *Invest Radiol*, 33(8):456–63.
- Riely, B. K., Loughnon, G., Ané, J.-M., and Cook, D. R. (2007). The symbiotic ion channel homolog DMI1 is localized in the nuclear membrane of *Medicago truncatula* roots. *Plant J*, 49(2):208–16.
- Rudolf, R., Mongillo, M., Rizzuto, R., and Pozzan, T. (2003). Looking forward to seeing calcium. *Nat Rev Mol Cell Biol*, 4(7):579–86.
- Saito, K., Yoshikawa, M., Yano, K., Miwa, H., Uchida, H., Asamizu, E., Sato, S., Tabata, S., Imaizumi-Anraku, H., Umehara, Y., Kouchi, H., Murooka, Y., Szczygłowski, K., Downie, J. A., Parniske, M., Hayashi, M., and Kawaguchi, M. (2007). NUCLEOPORIN85 is required for calcium spiking, fungal and bacterial symbioses, and seed production in *Lotus japonicus*. *Plant Cell*, 19(2):610–24.
- Saldaña, G., Martínez-Alcántara, V., Vinardell, J. M., Bellogín, R., Ruíz-Sainz, J. E., and Balatti, P. A. (2003). Genetic diversity of fast-growing rhizobia that nodulate soybean (*Glycine max* L. Merr). *Arch Microbiol*, 180(1):45–52.
- Sanders, D., Pelloux, J., Brownlee, C., and Harper, J. F. (2002). Calcium at the crossroads of signaling. *Plant Cell*, 14 Suppl:S401–17.

-
- Saur, I. M. L., Oakes, M., Djordjevic, M. A., and Imin, N. (2011). Crosstalk between the nodulation signaling pathway and the autoregulation of nodulation in *Medicago truncatula*. *New Phytol.*
- Schauser, L., Roussis, A., Stiller, J., and Stougaard, J. (1999). A plant regulator controlling development of symbiotic root nodules. *Nature*, 402(6758):191–5.
- Schultze, M., Quiclet-Sire, B., Kondorosi, E., Virelizer, H., Glushka, J. N., Endre, G., Géro, S. D., and Kondorosi, A. (1992). *Rhizobium meliloti* produces a family of sulfated lipooligosaccharides exhibiting different degrees of plant host specificity. *Proc Natl Acad Sci U S A*, 89(1):192–6.
- Schuster, S., Marhl, M., and Hofer, T. (2002). Modelling of simple and complex calcium oscillations. From single-cell responses to intercellular signalling. *Eur J Biochem*, 269(5):1333–1355.
- Scrase-Field, S. A. M. G. and Knight, M. R. (2003). Calcium: just a chemical switch? *Curr Opin Plant Biol*, 6(5):500–6.
- Seydel, R. (2010). *Practical bifurcation and stability analysis*, volume 5 of *Interdisciplinary applied mathematics*. Springer, New York, 3rd edition.
- Shaw, S. L. and Long, S. R. (2003). Nod factor elicits two separable calcium responses in *Medicago truncatula* root hair cells. *Plant Physiol*, 131(3):976–84.
- Shimoda, Y., Han, L., Yamazaki, T., Suzuki, R., Hayashi, M., and Imaizumi-Anraku, H. (2012). Rhizobial and fungal symbioses show different requirements for calmodulin binding to calcium calmodulin-dependent protein kinase in *Lotus japonicus*. *Plant Cell*.
- Sieberer, B. J., Chabaud, M., Fournier, J., Timmers, A. C. J., and Barker, D. G. (2011). A switch in Ca^{2+} spiking signature is concomitant with endosymbiotic microbe entry into cortical root cells of *Medicago truncatula*. *Plant J*.
- Sieberer, B. J., Chabaud, M., Timmers, A. C., Monin, A., Fournier, J., and Barker, D. G. (2009). A nuclear-targeted cameleon demonstrates intranuclear Ca^{2+} spiking in *Medicago truncatula* root hairs in response to rhizobial nodulation factors. *Plant Physiol*, 151(3):1197–206.
- Singer, S. R., Maki, S. L., Farmer, A. D., Ilut, D., May, G. D., Cannon, S. B., and Doyle, J. J. (2009). Venturing beyond beans and peas: what can we learn from *Chamaecrista*? *Plant Physiol*, 151(3):1041–7.
- Singh, S. and Parniske, M. (2012). Activation of calcium- and calmodulin-dependent protein kinase (CCaMK), the central regulator of plant root endosymbiosis. *Curr Opin Plant Biol*, 15(4):444–53.
- Sivia, D. S. and Skilling, J. (2006). *Data analysis: a Bayesian tutorial*. Oxford science publications. Oxford University Press, Oxford, 2nd edition.
- Skilling, J. (2006). Nested sampling for general Bayesian computation. *Bayesian Analysis*, 1(4):833–860.

-
- Smit, P., Limpens, E., Geurts, R., Fedorova, E., Dolgikh, E., Gough, C., and Bisseling, T. (2007). Medicago LYK3, an entry receptor in rhizobial nodulation factor signaling. *Plant Physiol*, 145(1):183–91.
- Smit, P., Raedts, J., Portyanko, V., Debellé, F., Gough, C., Bisseling, T., and Geurts, R. (2005). NSP1 of the GRAS protein family is essential for rhizobial Nod factor-induced transcription. *Science*, 308(5729):1789–91.
- Smith, D. C. and Douglas, A. E. (1987). *The biology of symbiosis*. E. Arnold, London.
- Soltis, D. E., Soltis, P. S., Morgan, D. R., Swensen, S. M., Mullin, B. C., Dowd, J. M., and Martin, P. G. (1995). Chloroplast gene sequence data suggest a single origin of the predisposition for symbiotic nitrogen fixation in angiosperms. *Proc Natl Acad Sci U S A*, 92(7):2647–51.
- Sprent, J. I. (2007). Evolving ideas of legume evolution and diversity: a taxonomic perspective on the occurrence of nodulation. *New Phytol*, 174(1):11–25.
- Sprent, J. I. (2008). 60Ma of legume nodulation. What’s new? What’s changing? *J Exp Bot*, 59(5):1081–4.
- Sprent, J. I. (2009). *Legume nodulation: a global perspective*. Wiley-Blackwell, Chichester, U.K.
- Sprent, J. I. and James, E. K. (2007). Legume evolution: where do nodules and mycorrhizas fit in? *Plant Physiol*, 144(2):575–81.
- Sprent, J. I. and James, E. K. (2008). Legume-rhizobial symbiosis: an anorexic model? *New Phytol*, 179(1):3–5.
- Stacey, G. S., Burris, R. H., and Evans, H. J. (1992). *Biological nitrogen fixation*. Chapman and Hall, New York.
- Staehelin, C., Xie, Z.-P., Illana, A., and Vierheilig, H. (2011). Long-distance transport of signals during symbiosis: are nodule formation and mycorrhization autoregulated in a similar way? *Plant Signal Behav*, 6(3).
- Stracke, S., Kistner, C., Yoshida, S., Mulder, L., Sato, S., Kaneko, T., Tabata, S., Sandal, N., Stougaard, J., Szczyglowski, K., and Parniske, M. (2002). A plant receptor-like kinase required for both bacterial and fungal symbiosis. *Nature*, 417(6892):959–62.
- Strogatz, S. H. (1994). *Nonlinear dynamics and Chaos: with applications to physics, biology, chemistry, and engineering*. Addison-Wesley Pub., Reading, Mass.
- Sun, J., Cardoza, V., Mitchell, D. M., Bright, L., Oldroyd, G., and Harris, J. M. (2006). Crosstalk between jasmonic acid, ethylene and Nod factor signaling allows integration of diverse inputs for regulation of nodulation. *Plant J*, 46(6):961–970.
- Sun, J., Miwa, H., Downie, J. A., and Oldroyd, G. E. D. (2007). Mastoparan activates calcium spiking analogous to Nod factor-induced responses in *Medicago truncatula* root hair cells. *Plant Physiol*, 144(2):695–702.
- Swainsbury, D. J. K., Zhou, L., Oldroyd, G. E. D., and Bornemann, S. (2012). Calcium ion binding properties of *Medicago truncatula* calcium/calmodulin-dependent protein kinase. *Biochemistry*.

-
- Swensen, S. M. (1996). The evolution of actinorhizal symbioses: Evidence for multiple origins of the symbiotic association. *Am. J. Bot.*, 83:1503–1512.
- Swensen, S. M. and Mullin, B. C. (1997). Phylogenetic relationships among actinorhizal plants. The impact of molecular systematics and implications for the evolution of actinorhizal symbioses. *Physiologia Plantarum*, 99(4):565–573.
- Sytsma, K. J., Morawetz, J., Pires, J. C., Nepokroeff, M., Conti, E., Zjhra, M., Hall, J. C., and Chase, M. W. (2002). Urticalean rosids: circumscription, rosid ancestry, and phylogenetics based on *rbcL*, *trnL-F*, and *ndhF* sequences. *Am J Bot*, 89(9):1531–46.
- Sze, H., Liang, F., Hwang, I., Curran, A. C., and Harper, J. F. (2000). Diversity and regulation of plant Ca^{2+} pumps: insights from expression in yeast. *Annu Rev Plant Physiol Plant Mol Biol*, 51:433–62.
- Takeda, N., Maekawa, T., and Hayashi, M. (2012). Nuclear-localized and deregulated calcium- and calmodulin-dependent protein kinase activates rhizobial and mycorrhizal responses in *Lotus japonicus*. *Plant Cell*, 24(2):810–22.
- Thomas, R. C. (2009). The plasma membrane calcium ATPase (PMCA) of neurones is electroneutral and exchanges 2 H^+ for each Ca^{2+} or Ba^{2+} ion extruded. *J Physiol*, 587(Pt 2):315–27.
- Tirichine, L., Imaizumi-Anraku, H., Yoshida, S., Murakami, Y., Madsen, L. H., Miwa, H., Nakagawa, T., Sandal, N., Albrektsen, A. S., Kawaguchi, M., Downie, A., Sato, S., Tabata, S., Kouchi, H., Parniske, M., Kawasaki, S., and Stougaard, J. (2006). Deregulation of a Ca^{2+} /calmodulin-dependent kinase leads to spontaneous nodule development. *Nature*, 441(7097):1153–6.
- Tirichine, L., Sandal, N., Madsen, L. H., Radutoiu, S., Albrektsen, A. S., Sato, S., Asamizu, E., Tabata, S., and Stougaard, J. (2007). A gain-of-function mutation in a cytokinin receptor triggers spontaneous root nodule organogenesis. *Science*, 315(5808):104–7.
- Toni, T. and Stumpf, M. P. H. (2010). Simulation-based model selection for dynamical systems in systems and population biology. *Bioinformatics*, 26(1):104–110.
- Torrence, C. and Compo, G. (1998). A practical guide to wavelet analysis. *Bulletin of the American Meteorological Society*, 79(1):61–78.
- Torres de Los Santos, R., Vierheilig, H., Ocampo, J. A., and García Garrido, J. M. (2011). Altered pattern of arbuscular mycorrhizal formation in tomato ethylene mutants. *Plant Signal Behav*, 6(5).
- Trinick, M. J. (1973). Symbiosis between *Rhizobium* and the non-legume, *Trema aspera*. *Nature*, 244:459–460.
- Trinick, M. J. (1979). Structure of nitrogen-fixing nodules formed by *Rhizobium* on roots of *Parasponia andersonii* Planch. *Can J Microbiol*, 25(5):565–78.
- van Brussel, A. A. N., Bakhuizen, R., van Spronsen, P. C., Spaink, H. P., Tak, T., Lugtenberg, B. J. J., and Kijne, J. W. (1992). Induction of pre-infection thread structures in the leguminous host plant by mitogenic lipo-oligosaccharides of *Rhizobium*. *Science*, 257(5066):70–72.

-
- Van de Velde, W., Zehirov, G., Szatmari, A., Debreczeny, M., Ishihara, H., Kevei, Z., Farkas, A., Mikulass, K., Nagy, A., Tiricz, H., Satiat-Jeunemaitre, B., Alunni, B., Bourge, M., Kucho, K.-i., Abe, M., Kereszt, A., Maroti, G., Uchiumi, T., Kondorosi, E., and Mergaert, P. (2010). Plant peptides govern terminal differentiation of bacteria in symbiosis. *Science*, 327(5969):1122–6.
- van der Heijden, M., Klironomos, J., Ursic, M., Moutoglis, P., Streitwolf-Engel, R., Boller, T., Wiemken, A., and Sanders, I. (1998). Mycorrhizal fungal diversity determines plant biodiversity, ecosystem variability and productivity. *Nature*, 396(6706):69–72.
- van der Heijden, M. G. A., Bardgett, R. D., and van Straalen, N. M. (2008). The unseen majority: soil microbes as drivers of plant diversity and productivity in terrestrial ecosystems. *Ecol Lett*, 11(3):296–310.
- van Noorden, G. E., Ross, J. J., Reid, J. B., Rolfe, B. G., and Mathesius, U. (2006). Defective long-distance auxin transport regulation in the *Medicago truncatula* super numeric nodules mutant. *Plant Physiol*, 140(4):1494–506.
- Vega-Hernandez, M., Perez-Galdona, R., Dazzo, F., Jarabo-Lorenzo, A., Alfayate, M., and Leon-Barrios, M. (2001). Novel infection process in the indeterminate root nodule symbiosis between *Chamaecyctisus proliferus* (tagasaste) and *Bradyrhizobium* sp. *New Phytol*, 150(3):707–721.
- Vernié, T., Moreau, S., de Billy, F., Plet, J., Combier, J.-P., Rogers, C., Oldroyd, G., Frugier, F., Niebel, A., and Gamas, P. (2008). EFD is an ERF transcription factor involved in the control of nodule number and differentiation in *Medicago truncatula*. *Plant Cell*, 20(10):2696–713.
- Vuuren, D. V., Bouwman, A., and Beusen, A. (2010). Phosphorus demand for the 1970–2100 period: A scenario analysis of resource depletion. *Global Environmental Change*.
- Wais, R. J., Galera, C., Oldroyd, G., Catoira, R., Penmetsa, R. V., Cook, D., Gough, C., Denarie, J., and Long, S. R. (2000). Genetic analysis of calcium spiking responses in nodulation mutants of *Medicago truncatula*. *Proc Natl Acad Sci U S A*, 97(24):13407–13412.
- Wais, R. J., Wells, D. H., and Long, S. R. (2002). Analysis of differences between *Sinorhizobium meliloti* 1021 and 2011 strains using the host calcium spiking response. *Mol Plant Microbe Interact*, 15(12):1245–52.
- Walker, S. A. and Downie, J. A. (2000). Entry of *Rhizobium leguminosarum* bv. *viciae* into root hairs requires minimal Nod factor specificity, but subsequent infection thread growth requires *nodO* or *nodE*. *Mol Plant Microbe Interact*, 13(7):754–62.
- Walker, S. A., Viprey, V., and Downie, A. J. (2000). Dissection of nodulation signaling using pea mutants defective for calcium spiking induced by Nod factors and chitin oligomers. *Proc Natl Acad Sci U S A*, 97:13413–13418.
- Wall, L. G. (2000). The actinorhizal symbiosis. *Journal of Plant Growth Regulation*, 19:167–182.
- Wang, B. and Qiu, Y.-L. (2006). Phylogenetic distribution and evolution of mycorrhizas in land plants. *Mycorrhiza*, 16(5):299–363.

-
- Wang, B., Yeun, L. H., Xue, J.-Y., Liu, Y., Ané, J.-M., and Qiu, Y.-L. (2010). Presence of three mycorrhizal genes in the common ancestor of land plants suggests a key role of mycorrhizas in the colonization of land by plants. *New Phytol*, 186(2):514–25.
- Weerasinghe, R. R., Bird, D. M., and Allen, N. S. (2005). Root-knot nematodes and bacterial Nod factors elicit common signal transduction events in *Lotus japonicus*. *Proc Natl Acad Sci U S A*, 102(8):3147–52.
- Weidmann, S., Sanchez, L., Descombin, J., Chatagnier, O., Gianinazzi, S., and Gianinazzi-Pearson, V. (2004). Fungal elicitation of signal transduction-related plant genes precedes mycorrhiza establishment and requires the *dmi3* gene in *Medicago truncatula*. *Mol Plant Microbe Interact*, 17(12):1385–93.
- Wilkinson, D. J. (2006). *Stochastic modelling for systems biology*. Chapman & Hall/CRC mathematical and computational biology series. Taylor & Francis, Boca Raton.
- Yamashita, M. (2011). Fluctuations in nuclear envelope’s potential mediate synchronization of early neural activity. *Biochem Biophys Res Commun*, 406(1):107–11.
- Yano, K., Yoshida, S., Muller, J., Singh, S., Banba, M., Vickers, K., Markmann, K., White, C., Schuller, B., Sato, S., Asamizu, E., Tabata, S., Murooka, Y., Perry, J., Wang, T. L., Kawaguchi, M., Imaizumi-Anraku, H., Hayashi, M., and Parniske, M. (2008). CYCLOPS, a mediator of symbiotic intracellular accommodation. *Proc Natl Acad Sci U S A*, 105(51):20540–20545.
- Yokota, K., Fukai, E., Madsen, L. H., Jurkiewicz, A., Rueda, P., Radutoiu, S., Held, M., Hosain, M. S., Szczyglowski, K., Morieri, G., Oldroyd, G. E. D., Downie, J. A., Nielsen, M. W., Rusek, A. M., Sato, S., Tabata, S., James, E. K., Oyaizu, H., Sandal, N., and Stougaard, J. (2009). Rearrangement of actin cytoskeleton mediates invasion of *Lotus japonicus* roots by *Mesorhizobium loti*. *Plant Cell*, 21(1):267–284.
- Yokota, K. and Hayashi, M. (2011). Function and evolution of nodulation genes in legumes. *Cell Mol Life Sci*, 68(8):1341–51.
- Yokota, K., Soyano, T., Kouchi, H., and Hayashi, M. (2010). Function of GRAS proteins in root nodule symbiosis is retained in homologs of a non-legume, rice. *Plant Cell Physiol*, 51(9):1436–42.
- Young, N. D., Debellé, F., Oldroyd, G. E. D., Geurts, R., Cannon, S. B., Udvardi, M. K., Benedito, V. A., Mayer, K. F. X., Gouzy, J., Schoof, H., Van de Peer, Y., Proost, S., Cook, D. R., Meyers, B. C., Spannagl, M., Cheung, F., De Mita, S., Krishnakumar, V., Gundlach, H., Zhou, S., Mudge, J., Bharti, A. K., Murray, J. D., Naoumkina, M. A., Rosen, B., Silverstein, K. A. T., Tang, H., Rombauts, S., Zhao, P. X., Zhou, P., Barbe, V., Bardou, P., Bechner, M., Bellec, A., Berger, A., Bergès, H., Bidwell, S., Bisseling, T., Choisine, N., Couloux, A., Denny, R., Deshpande, S., Dai, X., Doyle, J. J., Duzde, A.-M., Farmer, A. D., Fouteau, S., Franken, C., Gibelin, C., Gish, J., Goldstein, S., González, A. J., Green, P. J., Hallab, A., Hartog, M., Hua, A., Humphray, S. J., Jeong, D.-H., Jing, Y., Jöcker, A., Kenton, S. M., Kim, D.-J., Klee, K., Lai, H., Lang, C., Lin, S., Macmil, S. L., Magdelenat, G., Matthews, L., McCarrison, J., Monaghan, E. L., Mun, J.-H., Najar, F. Z., Nicholson, C., Noirot, C., O’Bleness, M., Paule, C. R., Poulain, J., Prion, F., Qin, B., Qu, C., Retzel, E. F.,

-
- Riddle, C., Sallet, E., Samain, S., Samson, N., Sanders, I., Saurat, O., Scarpelli, C., Schiex, T., Segurens, B., Severin, A. J., Sherrier, D. J., Shi, R., Sims, S., Singer, S. R., Sinharoy, S., Sterck, L., Viollet, A., Wang, B.-B., Wang, K., Wang, M., Wang, X., Warfsmann, J., Weissenbach, J., White, D. D., White, J. D., Wiley, G. B., Wincker, P., Xing, Y., Yang, L., Yao, Z., Ying, F., Zhai, J., Zhou, L., Zuber, A., Dénarié, J., Dixon, R. A., May, G. D., Schwartz, D. C., Rogers, J., Quétier, F., Town, C. D., and Roe, B. A. (2011). The *Medicago* genome provides insight into the evolution of rhizobial symbioses. *Nature*, 480(7378):520–4.
- Zdyb, A., Demchenko, K., Heumann, J., Mrosk, C., Grzeganeck, P., Göbel, C., Feussner, I., Pawlowski, K., and Hause, B. (2011). Jasmonate biosynthesis in legume and actinorhizal nodules. *New Phytol*, 189(2):568–79.
- Zhang, J., Subramanian, S., Stacey, G., and Yu, O. (2009). Flavones and flavonols play distinct critical roles during nodulation of *Medicago truncatula* by *Sinorhizobium meliloti*. *Plant J*, 57(1):171–83.