Characterisation of grain legume rhizobia for the potential development of inoculants with an improved shelf-life

Francesc Ferrando Molina

Thesis submitted for the degree of Doctor of Philosophy

Department of Biological and Environmental Sciences

Faculty of Natural Sciences

University of Stirling

December 2021

To Araceli
and Alba

Statement of originality

I hereby declare that this PhD thesis is an original piece of work that embodies the results of my own research. All work contained herein has not been submitted for any other degree.

All research material has been duly acknowledged and cited.

Signature of candidate

Francesc Ferrando Molina

Acknowledgements

I would like to start this section thanking the University of Stirling, The James Hutton Institute, PGRO and Legume Technology Ltd. for their financial support over this 4-year project and making it possible. I would also like to thank them for providing the means and training needed to be able to successfully accomplish it.

I am enormously thankful to my lead supervisor Richard Qulliam and JHI co-supervisors Euan James, Pete Iannetta, Adrian Newton and Marta Maluk for their exceptional support and guidance and for being there at any time and always giving the support required when and as it was needed.

I would like to thank landowners Roberto Torró, Ricardo Paredes and CF Folgado for lending soil for isolating rhizobia from, this project could not have started without their selfless participation at the beginning of it. I would like to express my sincere gratitude to Marta Maluk for the patience she has had over the past four years when teaching me all the different techniques and her ways of doing things that have saved me so much wasted time in the lab and on the computer. I would also like to thank Ashley Murdoch for cheerfully giving a hand in the lab at any time I needed. Furthermore, it is essential to acknowledge other JHI lab co-workers, Susan Mitchell, David Roberts, Carmen Escudero and Maddy Giles for the fun moments in the lab and their help around it and the interesting discussions on the different methods and techniques used in the lab. I would also like to thank all farm, media kitchen and agroecology department staff from JHI that have helped on meeting this goal. Furthermore, I would like to thank Ronnie from University of Stirling for his help on the intricacies of ordering lab supplies through the university system and always assisting on chasing up deliveries.

During the last four years, I have shared many moments with many people and some of them helped me developing as a scientist and naming them all here would make this a never-ending section. Nevertheless, it is essential to thank Javier, my PhD colleague at the university, to whom I am truly grateful for giving a hand when needed from the beginning of the project discussing statistics, methods and having some fun after work.

Family is an important factor in everybody’s life, and I am grateful for the support that my parents and sister have given me, not only during these past four years but for the whole of my life, always encouraging me to pursue my goals and giving support and advice when needed. Lastly but not last, I am enormously grateful to my wife Araceli who has been there every step of the way of this project and with whom I have shared the last 10 years of my life. Moving to a foreign country seeking to pursue a career in science eight years ago was a tough experience but, it certainly would have been much worse without the emotional and professional support from Araceli. Undertaking this project has been a long and sometimes challenging endeavour and I thank her for always believing in me and for being the light that kept my focus and sanity along this process. Finally, I would like to thank our little one, Alba, who recently came to our lives bringing joy and amusement to it and during the final 6 months of this project.

Abstract

Nitrogen (N) is a limiting element for plants; however, the use of synthetic N fertilisers in agriculture has increased crop production and yield. Importantly, a significant proportion of chemical fertilisers applied to soils will not be taken up by the roots of crops, but lost to the environment via run-off into waterways, or denitrification by soil bacteria. Legumes are plants that can transform atmospheric di-N into ammonia through a symbiotic association with rhizobia, a group of N-fixing bacteria, in root organs called nodules. Natural populations of rhizobia often exhibit below optimal N-fixation or nodulation, although so-called 'elite' strains with optimal abilities can be applied as inoculants. Although inoculants can be formulated with crop-compatible elite strains of rhizobia, their shelf-life is often compromised by high rates of cell die-off caused mainly by desiccation, which is an environmental stress that rhizobia are not good at withstanding. Therefore, there is a need to identify novel rhizobial strains that are able to tolerate desiccation stress. Recent evidence has suggested that strains isolated from areas with higher water deficit can better tolerate desiccation than those from wetter locations. Therefore, the overarching aim of this project was to isolate and characterise novel rhizobia strains from a semi-arid environment and assess their tolerance to desiccation for their potential use in inoculants for grain legumes. In addition, this project also evaluated the impact of agricultural land management on natural soil populations of rhizobia. Over 80 strains were isolated from soil from a semi-arid area of Spain using pea as a trapping plant. After a series of glasshouse and growth room experiments two strains were tested in field trials during two consecutive seasons where they showed a similar performance to strains from commercially available inoculants (used as positive control strains). Desiccation tolerance of strains isolated from Spain was tested and compared *in vitro* against strains from a wetter environment. The strains isolated from the semi-arid region showed 1.55-fold increased tolerance to desiccation. The genomes of 70 strains were sequenced and characterised, and a genome-wide association study on desiccation tolerance revealed that genes involved with regulating the concentration of solutes in the cytoplasm, and the protection and stabilisation of genetic material, were involved in the tolerance to this environmental stress. Finally, it was found that a change in land management and the presence of legumes in the crop rotation increased nodulating rhizobia in soil by 15 and 30 %

respectively over a period of 4 years. This project has successfully isolated strains with comparable symbiotic performance to standard commercial strains that show improved tolerance to desiccation, which makes them potentially superior for use in commercial inoculants with longer shelf-lives. Furthermore, this project has demonstrated that the reintroduction of a legume host after long absences produces an at least 4-year lasting effect that increases the proportion of nodulating rhizobia in soil year-on-year.

Table of contents

Chapter 2 | Isolation and symbiotic characterisation of pea-nodulating rhizobia from a semi-arid environment for their use in commercial inoculants 57

List of figures

List of tables

List of abbreviations

ANI	Average Nucleotide Identity
ATP	Adenosine triphosphate
A_w	Water activity
Bgmf	Below ground mass fraction
BH	Benjamini and Hochberg 1995 method of p-value adjustment
BLAST	Basic Local Alignment Search Tool
BNF	Biological Nitrogen Fixation
bv	Biovar
C	Carbon
$CaCl_2$	Calciul chloride
CFU	Colony forming units
CSC	Center for Sustainable Cropping
CSV	Comma-Separated Values file
cv	Cultivar
DAS	Days After Sowing
DF	Desiccation factor
DNA	Deoxyribonucleic Acid
dNTP	Deoxyribonucleic single nucleotides (Adenine, Timine, Citosine and Guanine)
DT	Desiccation Tolerance
EDTA	Ethylenediaminetetraacetic acid
EPS	Exopolysaccharides
FAO	Food and Agriculture Organisation of the United Nations

FAOSTAT	Food and Agriculture Organisation Corporate Statistical Database
FDA	Food and Drug Administration
$FeCl_3$	Iron(III) chloride
FGP	Final Germination Percentage
GI	Germination Index
GIC	Corrected Germination index
GLMM	Generalised Linear Mixed Effects Model
GPS	Global Positioning System
gs	Genospecies
GWAS	Genome Wide Association Study
H, S and B	In image analysis, each of the three channels of an HSB image which correspond to Hue, Saturation and Brightness channels respectivelly
H_2O	Water
HCl	Hydroclhoride acid
ISO	In photograpy stands for International Organisation of Standardization. It is a sensor sensitivity setting
JHI	The James Hutton Institute
K	Potassium
K_2HPO_4	Dipotassium phosphate
LB	Luria-Bertani broth (liquid)
LCO	Lipo-chito-oligosaccharides
LMM	Linear Mixed Effects Model
LN	Liquid Nitrogen
LPWG	The Legume Phylogeny Working Group
LSD	Least Significant Difference

Mg	Magnesium
$MgSO_4$	Magnesium sulphate
MGT	Mean Germination Time
MPN	Most probable number
N	Nitrogen
N_2	Di-nitrogen
NaCl	Sodium chloride
NC	Negative Control
NC1	For screening experiments, un-inoculated (SDW) NC
NC2	For screening experiments, inoculated NC with autoclaved culture
NCBI	National Center for Biotechnology Information
NH_3	Ammonia
NJ	Neighbor-Joining method
Nmf	Nodule mass fraction
NO_3^-	Nitrate
OD	Optical Density
P	Phosphorus
PC	Positive Control
PCR	Polymerase Chain Reaction
QC	Quality control
qPCR	Quantitative PCR
R, G and B	In image analysis, each of the three channels of an RGB image which correspond to the red, green and blue channels respectivelly
Rlc	Rhizobium leguminosarum species complex
Rleg	Rhizobium leguminosarum

Rlp	Rhizobium leguminosarum bv phaseoli
Rlt	Rhizobium leguminosarum bv trifolii
Rlv	Rhizobium leguminosarum bv vicia
ROS	Reactive oxygen species
rRNA	Ribosomal Ribonucleic Acid
SDS	Sodium dodecyl sulfate
SDW	Sterile Distilled Water
SNP	Single Nucleotide Polymorphism
SOC	Soil Organic Carbon
T_{50}	Time to 50% germination/emergence
TSW	Thousand Seed Weight
Tukey HSD	Tukey Honestly Significant difference
TY	Tryptone Yeast broth medium (liquid)
UK	United Kingdom
UN	United Nations
UPGMA	Unweighted Pair Group Method Average
USA	United States of America
UV	Ultraviolet
YMA-CR	Yeast Mannitol Agar with Congo Red medium (solid)
YMB	Yeast Mannitol Broth medium (liquid)

Chapter 1 | General introduction

1.1 The global context of nitrogen in food production

Nitrogen (N) is a limiting element for all organisms despite being a fundamental component of important biomolecules such as proteins and nucleic acids (Bernhard, 2010). Although N is always present in the environment, it is usually in a non-reactive form and therefore not available for most organisms. With the development of the Haber-Bosch process in the 20th century, it became possible to transform atmospheric N into reactive forms, which provides an important component of synthetic fertilisers. It has been calculated that 40 % of the current global population are dependent on this process (Smil, 2000), but it is estimated that by 2050 this will increase to 5.5 billion people (Crews and Peoples, 2004), which is more than half of the estimated population for that decade (United Nations, 2019). Furthermore, the Food and Agriculture Organisation of the United Nations (FAO) has projected that to meet the food demand of this population will require an increase in food production by at least 60 % (Rockström *et al.*, 2017). Consequently, to meet this demand the amount of artificially fixed N fertiliser will also need to be significantly increased under current cropping practices.

Some estimates suggest that human processes convert between 120 and 160 Tg of atmospheric di-nitrogen (N_2) per year (Galloway and Cowling, 2002; Galloway *et al.*, 2008; Gruber and Galloway, 2008; Rockström *et al.*, 2009; Steffen *et al.*, 2015). A significant part of this artificially fixed N is applied to the soil as a fertiliser; however, it is estimated that between 30 and 50 % of fertilisers applied to soil are lost due to leaching, much of which will eventually be discharged into water bodies thus causing eutrophication (Graham and Vance, 2003). Furthermore, it has been calculated that rivers discharge about 48 Tg N yr^{-1} into coastal zones (Statham, 2012). Human-mediated reactivated N is one of the main sources of reactive N in the land N cycle and agriculture and is the major contributor to perturbations of this cycle (Carpenter, 2005; Gruber and Galloway, 2008; Rockström *et al.*, 2017).

This artificially reactivated N has the potential to destabilise the biogeochemical N cycle and with it compromise the stability of the geological Era in which we live, the Holocene (Rockström *et al.*, 2009). It has been proposed that the amount of artificially reactivated N that the Earth system can tolerate is between 35 and 62 Tg N yr^{-1}, which is close to an estimate of pre-industrial levels of N reactivation (Rockström *et al.*, 2009; Steffen *et al.*, 2015). Thus, with estimates of human N-fixation being at least twice the amount of the upper threshold (Galloway and Cowling, 2002; Galloway *et al.*, 2008; Gruber and Galloway, 2008; Steffen *et al.*, 2015), it is evident that action needs to be taken in order to reduce current N reactivation rates to a more sustainable levels.

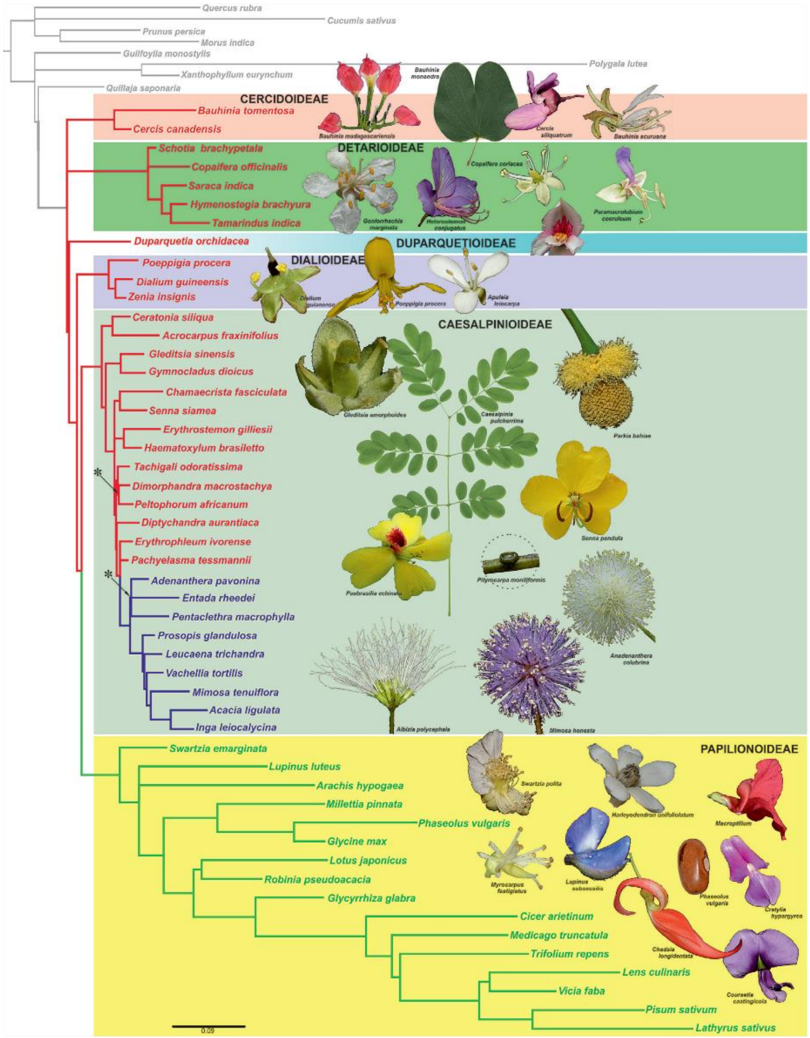
In terrestrial and aquatic ecosystems, some microorganisms are capable of reactivating, or fixing, N (Gruber and Galloway, 2008; de Bruijn, 2015). This natural N reactivation can be exploited to reduce or even replace the use of chemical N fertilisers in arable systems. N-fixing organisms are capable of reducing N_2 from the atmosphere into ammonia through a process called biological nitrogen fixation (BNF) and this can add up to 110 and 140 Tg N yr^{-1} to terrestrial and oceanic ecosystems respectively (Burris, 1980; Gruber and Galloway, 2008; Herridge *et al.*, 2008). A group of these N-fixing organisms include terrestrial bacteria, commonly called rhizobia, have developed a symbiotic interaction with legume plants, whereby sugars fixed in photosynthesis are exchanged for inorganic ammonia in plant root organs called nodules (Kiers *et al.*, 2008; Sprent *et al.*, 2017). It has been estimated that the total amount of fixed N of various legume crops and pastures could be 200 to 300 Kg of N ha^{-1} (Peoples *et al.*, 1995; Sessitsch *et al.*, 2002; Iannetta *et al.*, 2016; Maluk *et al.*, 2022). For instance, a yield of 4 Mg ha^{-1} of faba beans (*Vicia faba* L.) with a 4.5 % content of N accounts for about 180 Kg of fixed N ha^{-1} (Köpke and Nemecek, 2010). After yield offtake, most of the remaining fixed N remains as crop residues which are mineralised and the fixed N is potentially available for use by the next crop (Iannetta *et al.*, 2016; Maluk *et al.*, 2022). Thus, there is significant potential to reduce the amount of chemical N-fertiliser input and improve soil functions for the following crop, and these provisions may provide an additional yield increase for the following crop (Preissel *et al.*, 2015; Iannetta *et al.*, 2016). This can have an impact also on global warming alleviation as the same amount of N generated by industrial processes would release 480 Kg of CO_2, and, also the denitrification of excess N fertiliser in soil releases nitrous oxide (N_2O) where each released molecule has more than 200-fold potential to contribute to global warming compared with a molecule of CO_2 (Crews and Peoples, 2004; Jensen *et al.*, 2012; Barłóg *et al.*, 2018). The application of N fertilisers in agriculture is considered the largest single source of greenhouse gas emissions in this sector (Smith *et al.*, 2008; White and Brown, 2010). Thus, optimising or enhancing the interaction between legumes and rhizobia could present a sustainable solution for our dependence on chemical N fertilisers, and without compromising crop yields.

1.2 The nitrogen fixing symbiosis partners: legumes and their rhizobia

The term ‘diazotrophic’ describes organisms that can reduce atmospheric N_2 into ammonia and includes members of different phyla in the Eubacteria and Archaea (Young, 2000; James, 2017). Such organisms are found in the environment in either a free-living form or in mutualistic symbioses with other organisms (James, 2017). Symbiotic BNF has been described in many different organisms such as wood-boring molluscs (Altamia *et al.*, 2020), termites (Yamada *et al.*, 2006), tortoises (Montes-Grajales *et al.*, 2019), ferns (Raja *et al.*, 2012), and angiosperms where, in the Fabaceae family, symbiotic BNF is a common feature in most of the clade (Soltis *et al.*, 1995; Doyle, 2011).

The Fabaceae (commonly known as legumes) are the second most cultivated plant family (after the grasses) and include globally important food and forage crops, such as soybean (*Glycine max* (L.) Merr.), common beans (*Phaseolus vulgaris* L.), faba beans (*Vicia faba* L.), peas (*Pisum sativum* L.) and lucerne (*Medicago sativa* L.) (Sprent *et al.*, 2017; Koenen *et al.*, 2020). Legumes play a significant role in arable and natural soil ecology due to their mutualistic symbioses with soil microbiota (Graham and Vance, 2003). The Fabaceae are located on a branch of the Rosid I clade (Soltis *et al.*, 2000; Sprent *et al.*, 2017) and include a range of plant forms (trees, herbs and shrubs) with a wide global distribution (LPWG, 2013, 2017; Andrews and Andrews, 2017). The legume family has more than 19,000 species, over 750 genera (LPWG, 2013, 2017; Andrews and Andrews, 2017) and has demonstrated a higher than average diversification rate in the last 60 million years in the angiosperm clade (Magallon and Sanderson, 2001; LPWG, 2013). It is divided into six subfamilies (Figure 1.1) from which only some Caesalpinioideae (mainly in the Mimosoid clade) and most of the Papilionoideae are able to fix N symbiotically through nodulation (LPWG, 2017; Sprent *et al.*, 2017; van Velzen *et al.*, 2019). The subfamily Papilionoideae is the most diversified subfamily of the Fabaceae with around 14,000 species (LPWG, 2017) and is the group that hosts all the major cultivated grain legumes such as peas, faba beans, common beans and soybean.

Legumes capable of BNF interact symbiotically with rhizobia, a polyphyletic group of bacteria that describes those diazotrophic bacteria that fix atmospheric N inside specialised plant organs called nodules (de Lajudie and Young, 2017; Sprent *et al.*, 2017). Rhizobia belong to the bacterial classes Alpha- and Beta-proteobacteria and comprise 18 genera with over 200 described species (de Lajudie *et al.*, 2019). However, not all species belonging to these 18 genera are rhizobia (O'Hara *et al.*, 2016).



	No genera	No species
Absent nodulation	12	ca 335
	84	ca 760
	1	1
	17	ca 85
Some nodulating spp (indeterminate nodules)	184	ca 4,400
Mainly nodulating spp (indeterminate and determinate nodules)	503	ca 14,000

Figure 1.1. Fabaceae family phylogenetic tree. The diagram has been taken from The Legume Phylogeny Working Group (LPWG, 2017) with data added from the same source.

To be classed as a rhizobium, a species undergoes a nodulation test where it needs to show the ability to form nodules, and then following Koch's postulates, an identical strain needs to be re-isolated from the nodules formed in this test (Hungria *et al.*, 2016; Yates *et al.*, 2016; de Lajudie *et al.*, 2019).

1.2.1 Peas and faba beans, and their rhizobia

Peas and faba beans are two widely cultivated grain legume crops in the Papilionoideae subfamily (Cousin, 1997; Rubiales *et al.*, 2016; Peoples *et al.*, 2021). These grain legumes are of economic importance due to their use as human food, feed for animals, or as green manure (Cousin, 1997; Jensen *et al.*, 2010; Rubiales *et al.*, 2016). Their seeds are highly nutritious and offer high levels of both proteins, carbohydrates, minerals including essential minerals, and many other health promoting non-nutritionals too like fibre (Hall *et al.*, 2017; Ferreira *et al.*, 2021). In arable systems, they can represent an important source of income for farmers, and may elicit the transition towards a more sustainable and diversified agricultural systems and downstream value chains (Iannetta *et al.*, 2016, 2021).

The roots of peas and faba beans can be successfully nodulated with several species of rhizobia belonging to the genus *Rhizobium* (Saïdi *et al.*, 2014; Andrews and Andrews, 2017; Jorrin *et al.*, 2020). Pea- and faba bean-nodulating rhizobia are members of the *Rhizobium leguminosarum* species complex (Rlc) a group of Gram-negative bacteria formed by (to date) 18 genetically differentiated genospecies that share a common set of nodulation and N fixation genes (Kumar *et al.*, 2015; Boivin *et al.*, 2020; Young *et al.*, 2021). The nodulation and N fixation genes of these rhizobia are located in their accessory genome in what is called the Symbiotic plasmid (Sym-plasmid) (Laguerre *et al.*, 2001, 2003; Young *et al.*, 2006).

Despite the ancient and widespread use of these legumes, the area used for cropping faba bean has been declining since the 1960s as yield is often considered by growers to be unreliable (Figure 1.2)(Jensen *et al.*, 2010) due to pests and disease (Sillero *et al.*, 2010), and also to the increased use of synthetic N fertilisers, which reduce the benefits of including a legume in crop rotations (Crews and Peoples, 2004; Iannetta *et al.*, 2016).

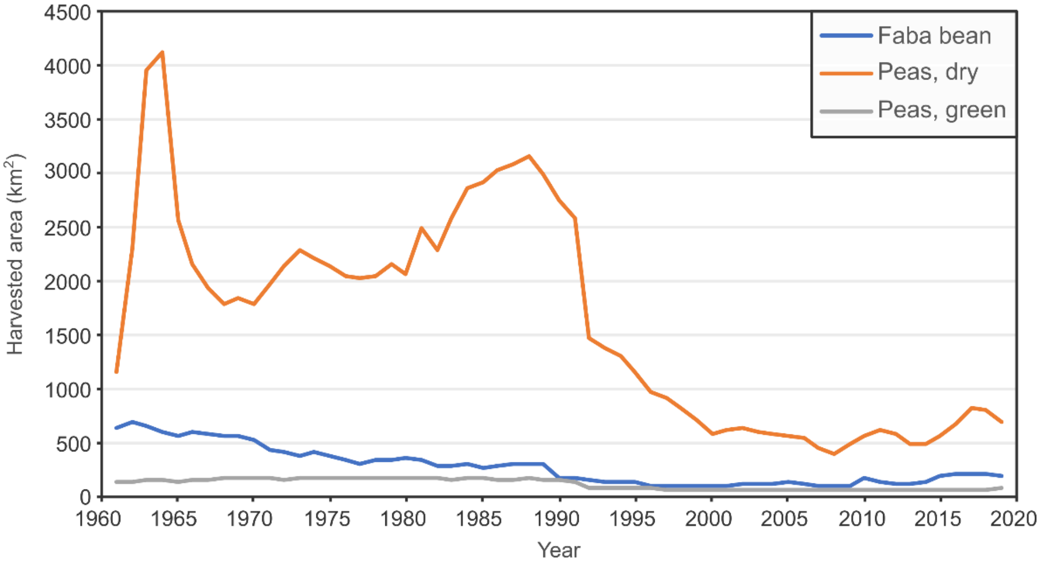


Figure 1.2. Area harvested of peas and faba beans in Europe (FAOSTAT, 2021).

1.3 The nodule and the nodulation process

1.3.1 On the origin of nodulation

It is widely accepted that nodules provide a source of N for the host plant, thus giving an advantage in N poor soils for these plants. Nodulation is a feature shown by members of the N-fixing clade, the group of plants within the Rosids where legumes are located (Soltis *et al.*, 1995, 2000). There are two main hypothesis on how nodulation appeared in this clade, one which proposes that the clade ancestor was a nodulator, and another which proposes that the ancestor was predisposed to nodulation but did not nodulate (Soltis *et al.*, 1995, 2000; Doyle, 2011, 2016; Werner *et al.*, 2014; Li *et al.*, 2015) .

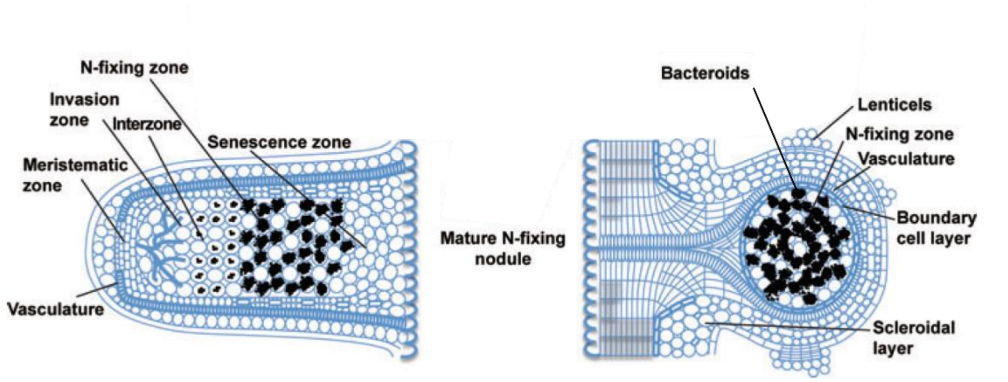
The monophyly of the N-fixing clade and the dispersion of nodulation within it are the main supporting arguments for the nodulating ancestor hypothesis and would have implied multiple losses of this ability to nodulate (Soltis *et al.*, 1995, 2000). When improved phylogenetic analyses of the Rosids resolved the uncertainty regarding some of its branches (Wang *et al.*, 2009; Bell *et al.*, 2010) the predisposed to nodulation ancestor hypothesis was brought forward because of the structural and anatomical diversity of nodules (Doyle, 2011). In this scenario, nodulation would have to have evolved several times within the N-fixing clade (Doyle, 2011; Werner *et al.*, 2014; Li *et al.*, 2015).

Recently, the nodulating ancestor hypothesis of Soltis *et al.* (1995) has seen a resurgence. This hypothesis argues that the predisposed to nodulation ancestor hypothesis does not explain (i) why nodulation only appeared in the N-fixing clade when parallel evolution could occur in any taxa; (ii) that these parallel origins drawn on phylogenetic trees are in conflict with structural and developmental data; and, (iii) that the predisposition to nodulate ancestor hypothesis would imply that the development of nodules would be more likely than losing them - whereas a more parsimonious approach would indicate the opposite given the complexity of this trait (Griesmann *et al.*, 2018; van Velzen *et al.*, 2019). van Velzen *et al.* (2019) hypothesise that in the nodulating ancestor hypothesis, nodules would have evolved from *Frankia* species capable of interacting with plants rather than rhizobia species; following this, a horizontal transfer of the genes needed for this interaction from *Frankia* to proteobacteria allowed the latter to compete for nodule occupancy of the legume ancestors.

1.3.2 Nodule types

Nodulation generally occurs on the roots although nodules can also develop on stems, for example, in the genera *Sesbania, Aeschynomene* or *Discolobium* (Boivin *et al.*, 1997; Sprent, 2009; James, 2017). When on roots, nodulation may occur either on the root hair, on a lateral root emergence or on the root epidermis depending on the legume host (Sprent *et al.*, 2013, 2017).

Based on their cellular components and primordium formation, nodules can be classified as either ‘determinate’ or ‘indeterminate’, which can be further divided into eight different nodule types depending on their morphological features (Maunoury *et al.*, 2008; Sprent *et al.*, 2013, 2017; Andrews and Andrews, 2017). Determinate nodules (Figure 1.3) are those nodules where the primordium is formed in the outer or middle cortex, is comprised of a finite number of plant cells, and its growth depends on bacteroid growth and division (Maunoury *et al.*, 2008).



Indeterminate Determinate

Figure 1.3. General histological composition of indeterminate (left) and determinate (right) nodules in peas and soybean respectively (modified from Ferguson *et al.*, 2010).

In contrast, indeterminate nodules (Figure 1.3) are those nodules where the primordium is originated in the inner cortex, has one or two meristems and growth depends on both continuous plant cell division and bacteroid growth (Maunoury *et al.*, 2008). Consequently, two different symbiotic models were developed for studying nodulation, i.e., determinate nodulation is studied via the *Lotus japonicus* (legume) and *Mesorhizobium loti* (rhizobium) interaction, and indeterminate nodulation studies use *Medicago truncatula* (legume) and *Ensifer (Sinorhizobium) meliloti* (rhizobium) (VandenBosch and Stacey, 2003).

1.3.3 Nodule organogenesis

Nodule formation is dependent on several biotic and abiotic factors including both the density and diversity of rhizobia populations living in soil saprophytically, and the plant-available N in the soil (Fonouni-Farde *et al.*, 2017; Reid *et al.*, 2018). The process of nodulation is initiated by the legume host, and the number of nodulation events, the number of nodules formed and which nodules go on to develop successfully is controlled by the legume (Ferguson *et al.*, 2010, 2019; Downie, 2014).

A legume with a N requirement will release flavonoids to the rhizosphere in root exudates (Figure 1.4A) (Kobayashi and Broughton, 2008; Oldroyd, 2013; Downie, 2014). Flavonoids are a group of plant secondary metabolites that can accumulate in the rhizosphere (Downie, 2014). The flavonoid-specific NodD protein in the cellular wall of compatible rhizobium will bind compatible flavonoid(s) and will activate a signal transduction cascade in the rhizobium which will result in the synthesis of lipo-chito-oligosaccharides (LCOs), or nodulation (nod-)factors, that are host specific (Kobayashi and Broughton, 2008). These LCOs are excreted into the rhizosphere and will bind to nod-factor-specific receptor-like kinases (NFR1/NFR5) (Figure 1.4B), which generate calcium oscillations in the epidermal cell of the legume root that will activate the expression of genes involved in the nodulation process - such as ERN1 and NIN (Oldroyd *et al.*, 2011; Cerri *et al.*, 2012; Oldroyd, 2013; Ferguson *et al.*, 2019).

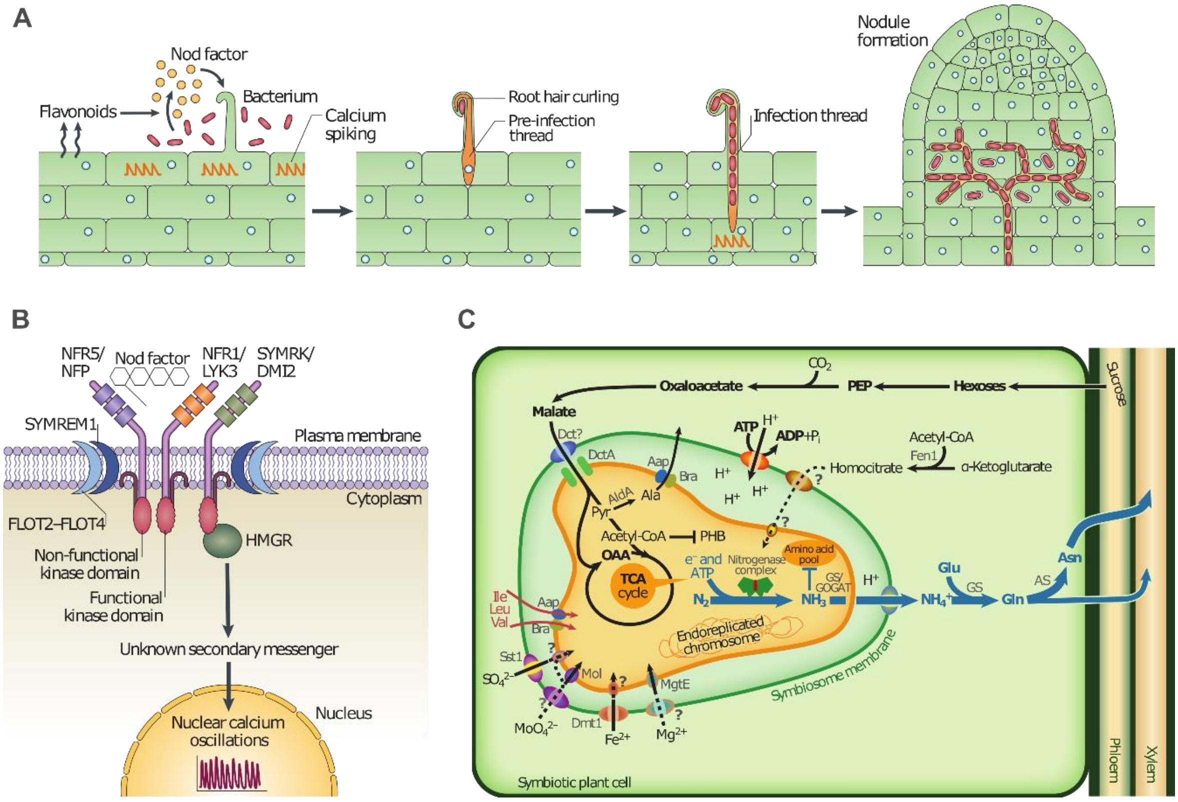


Figure 1.4. Nodule formation process, plus the main signalling and N-fixing pathways. Root colonisation until root nodule formation (A), nod factor recognition by receptor kinases NFR1/NFR5 (B) and transport and metabolism of an infected nodule cell (C). The figure is a collage of figures from Udvardi and Poole (2013) and Oldroyd (2013) with some modifications.

When the mutual recognition is completed, this triggers the root hair to curl around the rhizobium cell and a tubular infection thread (an invagination of the cell wall) will open and guide the rhizobium into the inner cortex (or outer/middle cortex for determinate nodules) of the root where a further wave of calcium oscillations will be triggered in the receptor cortical cells (Figure 1.4A) (Maunoury *et al.*, 2008;

Oldroyd, 2013). The rhizobium multiplies in the infection thread which guides it to the cytoplasm of the host cortical cell surrounded by the plant membrane (Oldroyd *et al.*, 2011; Udvardi and Poole, 2013). During this process, some of the duplicated rhizobia will be differentiated into bacteroids, a modified bacterial cell that cannot multiply any longer, though it may undergo endoreduplications, a process by which the cell undergoes duplication without mitosis, although this is limited to the Inverted Repeat Lacking Clade (IRLC) of the Papilionideae subfamily (Maunoury *et al.*, 2008; Downie, 2014; Sprent *et al.*, 2017). Bacteroids are thus located in the cytoplasm of the cortical cell inside a cellular membrane called the symbiosome, an organelle-like structure, where N is fixed and exchanged with the host for carbohydrates (Figure 1.4C) (Udvardi and Poole, 2013; Downie, 2014).

1.4 Characteristics of an optimal legume inoculant

In agriculture, legumes play a crucial role by providing N fixed through their nodule rhizobia which reduces the need for using chemical N fertilisers. They are therefore capable of alleviating some of the environmental issues of this sector and represent a key for transition to more sustainable agricultural production (Rees *et al.*, 2013; Iannetta *et al.*, 2021; Oliveira *et al.*, 2021; Peoples *et al.*, 2021; Udvardi *et al.*, 2021). Faba beans and peas obtain respectively 74 and 62 % of the N they need from N fixation (Peoples *et al.*, 2021), and have the capacity to fix up to 300 kg N ha^{-1} (Maluk *et al.*, 2022). Inoculation is the process by which one or several rhizobial strains with proven enhanced N-fixing and plant growth promoting capabilities (i.e., ‘elite’ strains) are applied to soil, seed, or seedlings before or at sowing, in the form of an inoculant. Autochthonous legume crops usually have compatible rhizobia present in soil, though their symbiotic and N-fixing abilities may be under optimal conditions (Mutch and Young, 2004). When foreign legume crops, like soybean in America or faba bean and peas in Australia, are used they often do not have compatible rhizobia in their new soil environment to nodulate with (Alves *et al.*, 2003; Denton *et al.*, 2013; Chibeba *et al.*, 2018). Thus, it is in these two cases when inoculation is recommended (Giller *et al.*, 2016).

Since the discovery of the BNF process and its causative agents in the late 19th century (reviewed on Burris, 1994), inoculation of legumes by the addition of specific rhizobium strains has become an established agricultural practice in many countries (Catroux *et al.*, 2001; Herridge, 2008). The production of legume inoculants involves the growth of the potentially elite strain in an optimal liquid medium and its subsequent mixing with a carrier substance tailored to sustain the rhizobium until their application (Deaker, 2004; Bashan *et al.*, 2014). Each component and step of inoculant production provides an opportunity to improve the final product.

1.4.1 The strain

The rhizobium strain, or strains, need to fulfil certain criteria to ensure inoculant success and commercial viability (Keyser and Li, 1992; Brockwell *et al.*, 1995; Herridge, 2008). Desirable rhizobia strains need to be a genetically stable, with an ability to form nodules and fix N, with not only the target crop, but desirably also with a wider range of genotypes/hosts. Its effectiveness needs to be stable across a range of environmental conditions including in the presence of soil nitrate and local rhizobia competitors (e.g. native soil rhizobia). It needs to be able to grow in different environments such as artificial media, the carrier, and in soil, and it must be able to not only colonize the soil and rhizosphere, but also persist over several seasons and migrate from the inoculation area to the roots of the target crop. Finally, an ideal strain must demonstrate low mortality when applied on seed and show compatibility with agrochemicals.

The use of locally sourced strains will increase ‘resilience potential’ against local competitors and predators (Paau, 1989; Bashan, 1998; Bashan *et al.*, 2014; Pastor-Bueis *et al.*, 2019). Once isolated, the strain needs to undergo several tests or screenings to ensure host compatibility and range, as well as N-fixing ability (Paau, 1989). This requires considerable amounts of time, which might make a specific field- and/or crop-tailored strategy commercially non-viable. Furthermore, there are soils with low or non-existent compatible rhizobia, for example, in Australia. Such soils are often cropped with European legumes like pea, faba beans or clovers (*Trifolium spp.*) which have to be inoculated with rhizobia sourced from other continents (Herridge, 2008).

The most commonly used growth media for rhizobia is yeast mannitol agar/broth (YMA or YMB) (Vincent, 1970), amended with different substances, e.g. carboxymethyl cellulose, glycerol, or sucrose (Manikandan *et al.*, 2010; Taurian *et al.*, 2010; Jha and Saraf, 2012), depending on the bacterial isolate. However, there are a range of alternative substrates that have also been used as media for growing rhizobia, such as industrial by-products with amendments like cheese whey, malt sprouts, industrial-grade yeast extracts, pea husk, molasses, and water hyacinth extract, a brown Indian sugar made from palm sap called ‘jaggery’ (all reviewed by Ben Rebah *et al.*, 2007), or wastewater sludge (Ben Rebah *et al.*, 2002).

1.4.2 The carrier

The carrier is the substrate that is used during the formulation to sustain the rhizobium following their growth in culture media until their use as an inoculant, and the carrier is the major component (by weight or volume) of the final inoculant (Bashan, 1998; Bashan *et al.*, 2014).

The carrier must be able to sustain the growth of the formulated rhizobium and maintain its population over time to allow transportation, storage and use (Stephens and Rask, 2000; Bashan *et al.*, 2014). To achieve this, it is important that the carrier is able to contain high levels of water, regulate the pH within a range suitable for the rhizobium, and be non-toxic for both these bacteria and the environment (Smith, 1992; Albareda *et al.*, 2008; Deaker *et al.*, 2016). From a manufacturing perspective, the carrier also needs to be a substance that is widely available in different forms, inexpensive, easy to sterilise, have chemical and physical uniformity, and to be amendable (Smith, 1992; Brockwell and Bottomley, 1995; Stephens and Rask, 2000; Albareda *et al.*, 2008). Other important characteristics of an ideal carrier are that it must not have heat of wetting, i.e., the release of heat that some colloidal materials have upon wetting (Smith, 1992), as this rise in temperature may have a detrimental effect on the formulated rhizobium.

There are five different categories of inoculant carriers: soils (e.g. peat or clay), plant material (e.g. cellulose), inert materials (e.g. vermiculite), plain lyophilised microbial cultures, and liquid inoculants (Smith, 1992; Bashan, 1998; Bashan *et al.*, 2014; Deaker *et al.*, 2016). The most commonly used carrier is peat (Smith, 1992; Mahmood *et al.*, 2016) with new carriers usually tested against it as a standard (Date, 2001; Bashan *et al.*, 2014). Peat possess many of the characteristics of an ideal carrier, e.g., the capacity to support the growth of rhizobia and regulate the pH, but its lack of widespread availability (as a limited bioresource) and the variability of its composition depending on the plant material from which it is formed (Chao and Alexander, 1984; Mahmood *et al.*, 2016), has led to a search for alternative carriers (Ben Rebah *et al.*, 2002; Albareda *et al.*, 2008). Despite decades of research directed at discovering an ideal carrier, little progress has been made with synthesising a carrier with all the ideal properties (Smith, 1992; Bashan *et al.*, 2014). It has been proposed to incorporate polymers from other industries, such as nanotechnology or pharmaceuticals, to create a carrier that meets all these properties; however, there are as yet no commercial products available on the market (John *et al.*, 2011; Schoebitz *et al.*, 2013; Bashan *et al.*, 2014).

1.4.3 The formulation process

Formulation is the process of mixing the bacterial isolate with the carrier, and is the most critical step during inoculant production (Bashan, 1998; Bashan *et al.*, 2014). Despite the high level of potentially efficient strains reported in the literature, not all of them reach the commercial market possibly due to an inefficient formulation (Bashan, 1998; Bashan *et al.*, 2014). Both the bacterial isolate and the carrier must reach the formulation stage a near optimum state to produce an effective inoculant i.e. a product capable of having a repeatedly positive effect on plant biomass, fitness, and/or yield. Thus, the bacterial isolate must be at a suitable growth phase at the time of mixing (Bashan *et al.*, 2014; Deaker *et al.*, 2016) and the carrier must be adjusted to best accommodate the specific bacterial isolate. For example, peat is acidic and, therefore, the pH needs to be adjusted and held to near neutral before adding the bacterial culture (Smith, 1992). Sterile peat can hold larger populations of bacterial isolates than non-sterile peat (Stephens and Rask, 2000), but involves the costly process of sterilisation prior to formulation by autoclaving, heating or gamma irradiation (Deaker *et al.*, 2016).

The inoculant must support the growth of the bacterial strain, maintain a high concentration of viable cells over a prolonged period (i.e., during the ‘shelf-life’), and deliver enough viable bacteria at the time of inoculation (Stephens and Rask, 2000; Date, 2001; Bashan *et al.*, 2014). The characteristics of the ideal inoculant are slightly different depending on the various needs of manufacturers and farmers, but it is necessary to take these into consideration to ensure inoculant success (Bashan *et al.*, 2014). For farmers, the ideal inoculant must be; (1) easy to use and compatible with seed drilling equipment and other growers’ practices; (2) facilitate improved yield; (3) be effective in different field conditions; (4) tolerate storage conditions whilst maintaining high concentrations of rhizobia; and (5) facilitate the survival of microorganism during and after inoculation until the host plant requires the rhizobia (Catroux *et al.*, 2001; Bashan *et al.*, 2014). From the manufacturers' point of view, the ideal inoculant must have an extended shelf-life, preferably spanning different growing seasons, and must have a consistent effect on yield, biomass and/or fitness with no adverse environmental effects (Catroux *et al.*, 2001; Bashan *et al.*, 2014).

Inoculants have been classified into different categories depending on their physical conditions (solid or liquid) and application method (to seed or soil) (see Smith, 1992; Brockwell *et al.*, 1995; Deaker, 2004; Bashan *et al.*, 2014; Mahmood *et al.*, 2016). Encapsulation of bacteria in polymers is a current experimental formulation technique in the field of agriculture (Schoebitz *et al.*, 2013; Bashan *et al.*, 2014). It was first introduced by Bashan (1986) and consists of entrapping cells in beads of a polymer matrix, with alginate being the most common polymer (Berninger *et al.*, 2016). Bashan *et al.* (2014) divides liquid inoculants as either 'primitive' (i.e. those used at research facilities which are basically composed of unaltered growth media) or, liquid or 'non-primitive' inoculants, which are amended growth media with enhanced properties; solid inoculants can be classified as either organic inoculants, with peat being the most common, or inorganic inoculants.

1.4.4 Soil inoculant application

The technique used for applying the inoculant will also determine the survival and efficiency of the bacteria in the soil or on the seed (Mahmood *et al.*, 2016). The general methods of application are seed and soil inoculation (Smith, 1992; Bashan *et al.*, 2014; Mahmood *et al.*, 2016) yet there are other methods less frequently used like foliar spray, or application by irrigation (as mentioned in Mahmood *et al.*, 2016). Direct seed inoculation has the advantage of delivering a high density of bacterial cells in close proximity to the developing seedling (Graham, 2008), and it has been suggested that soil inoculation also enhances the competitiveness of the inoculant strain against native soil rhizobia (López-García *et al.*, 2002). However, soil inoculation requires larger volumes of inoculum, which raises the price of inoculating soil (Bashan *et al.*, 2014). Either way, once in the soil, the survival of the rhizobium strain will be determined by biotic and abiotic factors such as soil temperature, moisture, presence of nutrients, pH, predation and niche availability (Mahmood *et* *al.*, 2016). Importantly, survival of the inoculated strain can be compromised if it is not capable of competing with the autochthonous microbial communities (Bashan *et al.*, 2014).

The effectiveness of any inoculant depends on the quality of the rhizobial strain used, its competitiveness with other soil bacteria, its resilience against predation and its survival during the formulation and storage of the inoculant until its application in soil (Paau, 1989; Bashan, 1998; Bashan *et al.*, 2014). Generally, rhizobial strains are isolated from soil using trapping experiments with the target host-plant and soil of interest (Requena *et al.*, 1997; Khalid *et al.*, 2004; Diez-Mendez *et al.*, 2015). This method ensures the isolation of legume-compatible rhizobia. However, the isolated strains also requires further screening for their N-fixation and nodulation ability (Mutch and Young, 2004; Maluk *et al.*, 2022), and their competitiveness for nodule occupancy (Boivin *et al.*, 2020; Mendoza-Suárez *et al.*, 2020, 2021), as well as assessment of their tolerance to environmental stresses or formulation processes that may affect their use as a viable commercial product (Bashan *et al.*, 2002; Deaker *et al.*, 2007; Albareda *et al.*, 2008; Casteriano *et al.*, 2013; Atieno *et al.*, 2018).

1.5 Desiccation tolerance in rhizobia, and its importance in the production of inoculants

At the formulation stage, the bacterial strain will be transferred from an optimal growth environment where it has access to all of its required nutrients, optimal growth temperature, and humidity, to an environment where it will experience some degree of desiccation stress (Casteriano *et al.*, 2013) or anhydrobiosis, a state where the organism reduces or even stops its vital functions (Berninger *et al.*, 2018). Low tolerance to desiccation is considered one of the main causes of lack of effectiveness of inoculation (Deaker, 2004; Casteriano *et al.*, 2013; Casteriano, 2014), and due to it being responsible for high cell die-off is often a major constraint for the development of efficacious inoculants (Berninger *et al.*, 2018). Desiccation causes cellular damage via three different mechanisms: (i) oxidative damage caused by the formation of reactive oxygen species (ROS); (ii) phase transition after rehydration which affects cell envelope permeability; and (iii) browning or Maillard reactions (García, 2011). For non-spore forming bacteria like

Rhizobium, damage caused by anhydrobiosis is very likely to be deleterious (Berninger *et al.*, 2018). The usual pattern of bacterial survival after formulation is a bi- or tri-phasic curve (Date, 2001) with the most acute die-off during the first phase (1-7 days) (Figure 1.5). Hence, tolerance to desiccation is a desirable phenotype for rhizobial inoculants as it can reduce the inoculant curation period (i.e. the time under controlled environment which allows bacteria to acclimatise to their new conditions and to multiply, thus alleviating viable cell loss caused by die-off), expand its shelf-life during storage, and increase rhizobial survival at sowing and subsequent saprophytic stages in the soil (van Ham *et al.*, 2016; Molina-Romero *et al.*, 2017).

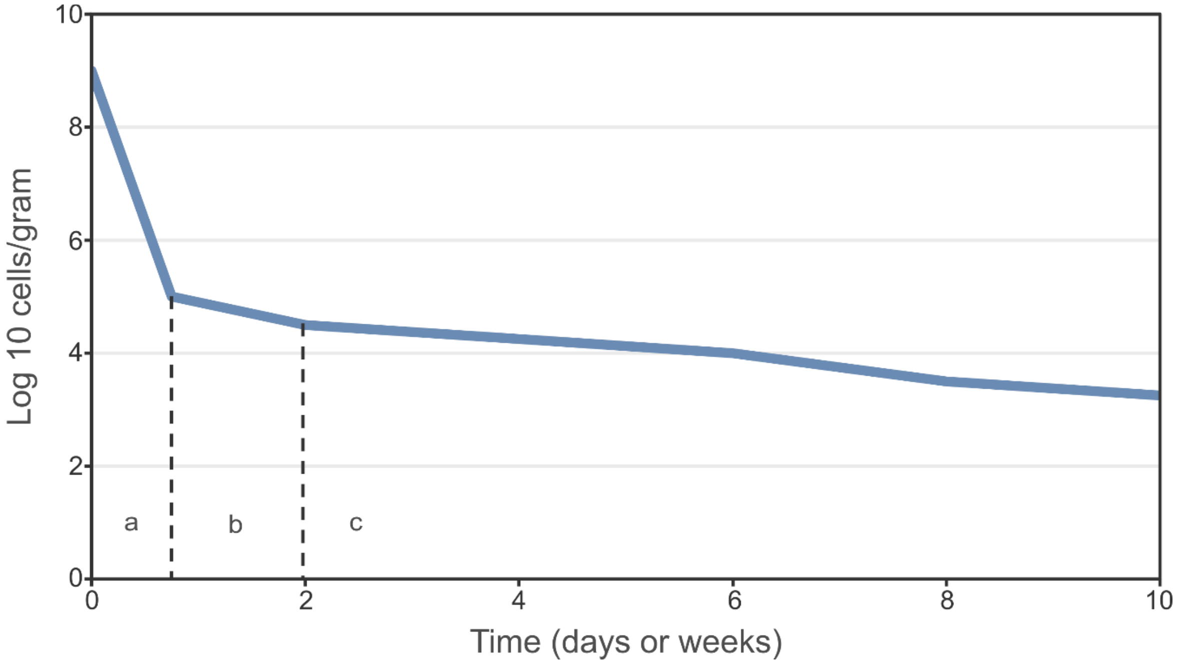


Figure 1.5. Adapted from Date (2001). Survival of bacteria after inoculant formulation showing the characteristic thee-phasic (a, b and c) curve with differential rate of rhizobia die-off.

Desiccation tolerance in bacteria has previously been studied using a diverse range of techniques, e.g., micro-droplet assays (Slininger and Schisler, 2013; van Ham *et al.*, 2016), desiccation assays with and without substrate (Molina-Romero *et al.*, 2017), desiccation experiments on glass beads (McIntyre *et al.*, 2007), or directly on-seed (Streeter, 2003), yet, there has not been a predilection for one method over any of the others.

Desiccation tolerance in rhizobia is determined by several factors such as accumulation of trehalose, the synthesis of proteins that can protect the cell envelope, repair DNA or stabilise other proteins, and responses to oxidative stress damage caused by desiccation stress (Cytryn *et al.*, 2007; Casteriano *et al.*, 2013). The synthesis of the disaccharide trehalose is a strategy adopted by several bacteria when exposed to an anhydrobiotic environment, for example, Zhang and Yan (2012) observed that when exposing *Escherichia coli* to desiccation all the strains they tested synthesised trehalose, with those isolated from soil producing significantly more trehalose than the reference strains. In rhizobia, trehalose biosynthesis has been observed in *R. leguminosarum* sv *trifolii* when exposed to an anhydrobiotic environment (McIntyre *et al.*, 2007). Nevertheless, transcriptional and physiological studies showed that for *Bradyrhizobium japonicum* the desiccation tolerance mechanism comprised not only increases in cytoplasmic trehalose but also the synthesis of proteins that protected the cell membrane or repaired DNA, although the presence of trehalose had a significant role in tolerance to this stress (Cytryn *et al.*, 2007; Zhu *et al.*, 2021). Anhydrobiosis also generates oxidative stress caused by ROS (García, 2011), e.g. under oxidative stress, *B. japonicum* has a wide range of response mechanisms, such as increased motility or the production of exopolysaccharides (EPS) (Donati *et al.*, 2011).

In addition to the intrinsic capability of desiccation tolerance, largely driven by bacterial genotype and phenotype, survival during anhydrobiosis can be enhanced by elicitors which are molecules that induce some degree of tolerance to this environmental stress. For example, the addition of trehalose to the growth medium of *B. japonicum* increased bacterial survival after a period of anhydrobiosis (Streeter, 2003). Furthermore, growing rhizobia in aqueous peat extracts can induce tolerance to desiccation through the upregulation of amino acid metabolism, inorganic ion transport or cell wall biogenesis, yet the elicitor(s) in the peat extract for this enhanced tolerance to desiccation remains unknown (Casteriano *et al.*, 2013; Casteriano, 2014; Atieno *et al.*, 2018). Other substances may confer a protective environment against desiccation stress during the formulation process, such as the synthetic polymer polyvinyl alcohol (Deaker *et al.*, 2007). Therefore, it is possible to enhance survival during inoculant formulation by the addition of protectants to the growth medium and/or carrier, which can trigger internal anhydrobiotic metabolic processes in the formulated organisms or by the co-formulation with other bacteria which will provide indirect protection. These three methods of enhancing cell viability in non-sporulating bacteria have recently been reviewed by Berninger *et al.* (2018).

1.6 Water activity: a key parameter for bacterial survival but rarely used in inoculant technology

Selecting the appropriate strain, understanding its physiology, and optimising its growth medium and formulation are key processes for producing more effective and longer-lasting inoculants. However, other abiotic factors can also play important roles in the survival of rhizobia. A characteristic that is often overlooked during inoculant development is water activity (A_W). Water activity is described as the ratio between the water pressure of the sample at a given temperature and that of pure water at the same temperature (Food and Drug Administration (FDA), 2015), and indicates the availability of water for metabolic processes (Deaker *et al.*, 2012; Tadapaneni *et al.*, 2018). Water activity changes with temperature (Liu *et al.*, 2018) and the physical and chemical properties of the substrate (Chirife and Fontana, 2007). Consequently, matrices with the same moisture content show an increase in water activity when the temperature is increased. Table 1.1 shows water activities above which certain biological processes and growth of different microorganisms are possible.

Table 1.1. Water activity thresholds above which some biological processes can develop (extracted from Bell and Labuza (2000)).

Water activity	Biological processes
0.1	Oxidation occurs
0.3	Liquid chemical activity
0.4	Enzyme activity
0.6	Osmophilic fungal growth
0.7	Xerophilic mould growth
0.8	Halophilic bacteria and most moulds growth
0.9	Bacterial growth

In the food and pharmaceutical industries, water activity is widely used as an indicator of conditions that might support growth of spoilage or pathogenic microorganisms, which can reduce shelf-life or quality of a product. Low A_W reduces microbial activity, induces dormancy and/or kills cells (Stapelfeldt *et al.*, 1997; Food and Drug Administration (FDA), 2015; Laranjo *et al.*, 2017; Angamuthu *et al.*, 2018;

Liu *et al.*, 2018; Patel *et al.*, 2018; Shi *et al.*, 2018). Therefore, quantifying water activity could also be of utility for the inoculant industry to optimise the specific environmental conditions that will help optimise the survival of formulated rhizobia. Vriezen *et al*. (2007) suggests that at water activities below 0.53 the monolayer of water molecules stabilizing proteins and other molecules is removed, causing irreversible damage to the cell membrane and thus reducing rhizobia survival. Despite the essential role of A_W in bacterial survival, there are only a few studies that have characterised water activity during inoculant development (Mugnier and Jung, 1985; Kosanke *et al.*, 1992; Paul *et al.*, 1993; Goss *et al.*, 2003; Friesen *et al.*, 2004, 2005, 2006; Deaker *et al.*, 2007, 2012). Therefore, it is important to take A_W into account when designing novel solid or seed coating formulations, and attention must be paid to the threshold at which the formulated rhizobia are able to survive, but also the inherent characteristics of the carrier and additives, plus the storage conditions of the final product, as all these factors will ultimately modify the available water in the product and thus affect the shelf-life of the inoculant.

1.7 Research rationale, aims and objectives

The production and use of N fertilisers generates and aggravates global environmental problems through reactive N loss such as via eutrophication of waterways (N-leaching), and global warming (via greenhouse gasses) (Graham and Vance, 2003; Barłóg *et al.*, 2018). Legumes, however, remain an under-developed important asset for agriculture, as they can enhance soil functions, and provide a renewable source of fertiliser-N through BNF reducing the need for artificial (mineral) N-fertiliser use in support to more sustainable production systems (Iannetta *et al.*, 2016, 2021; Peoples *et al.*, 2021; Udvardi *et al.*, 2021; Maluk *et al.*, 2022).

Legume nodulation can be improved by using inoculant formulations that can support and maintain the growth and survival of rhizobia until they are applied to the soil (Stephens and Rask, 2000; Date, 2001; Bashan *et al.*, 2014). However, one of the main constraints of the inoculant industry is the shelf-life of the inoculant, which is not yet comparable to that of chemical fertilisers (Catroux *et al.*, 2001; Bashan *et al.*, 2014). Shelf-life is affected by several factors with the most relevant being the elevated die-off of formulated rhizobia mainly due to a lack of desiccation tolerance (Deaker *et al.*, 2012; Casteriano *et al.*, 2013). Thus, the isolation of nodulating rhizobia with improved tolerance to desiccation can help alleviate this problem. Recent evidence suggests that rhizobia isolated from soils with less available water are inherently more tolerant to desiccation due to their habitual exposure to this event in their natural habitat (van Ham *et al.*, 2016).

Therefore, using novel pea- and faba bean-compatible strains of rhizobia, the aim of this project was to characterise desiccation tolerance and the capacity for sustaining viable rhizobia with a longer shelf-life than currently available commercial inoculants.

Specifically, this thesis aimed to address the following objectives:

1. The isolation and characterisation of elite strains of pea- and faba bean-compatible rhizobia from a semi-arid environment. (Chapter 2).
2. To quantify desiccation tolerance in a group of rhizobia from two contrasting environments, i.e., the UK and Spain. (Chapter 3).
3. By using a genome-wide approach, to determine which genes are involved in desiccation tolerance of Rlc rhizobia. (Chapter 4).
4. To evaluate the impact of two different agricultural management regimes on natural populations of Rlc rhizobia. (Chapter 5).

Chapter 2 | Isolation and symbiotic characterisation of pea-nodulating rhizobia from a semi-arid environment for their use in commercial inoculants

Abstract

Legumes play an important role in natural and arable systems due to their symbiotic interaction with some soil bacteria collectively known as rhizobia. These bacteria are capable of fixing atmospheric nitrogen (N) inside specialised root organs (nodules) through a process called biological nitrogen fixation (BNF). In the UK, pea and faba bean are the two main grain legume crops, both of which are often characterised by their yield instability. This may, in part, be due to the presence of inefficient rhizobia in soil. Thus, supplying the crop with effective rhizobia through the application of inoculants may contribute to reducing variability in yield. Pea and faba bean rhizobia do not withstand desiccation very well, which is one of the main stresses that reduces the effectiveness of inoculants in both the field and during storage. Therefore, the main aim of this study was to isolate rhizobia from a semi-arid environment and assess their symbiotic performance in field and glasshouse experiments for future potential use in commercial inoculants. Soil from five fields with low N input were sampled from eastern Spain. More than 80 strains of pea-nodulating rhizobia were isolated, using pea as a trapping plant. All strains were genetically characterised and a selection of 40 strains were subsequently tested for nodulation and N fixation in glasshouse experiments. The top six performers were further screened in different legume crops, and two were selected for field testing based on their capacity to nodulate and increase pea and faba bean biomass in a N-free environment. In two successive field trials, the candidate strains successfully improved crop establishment when applied to seeds before sowing during a dry season, yet this was not reflected in any change of yield.

Keywords

Rhizobia, legumes, field trial, *Rhizobium leguminosarum*, pea, faba bean

2.1 Introduction

Legumes have important ecological and economic value (Graham and Vance, 2003), as many of these plants interact symbiotically with soil bacteria called rhizobia that form nodules on the root of the plant (Sprent *et al.*, 2013). Inside the nodule, rhizobia are provided with a suitable environment where they can transform atmospheric di-nitrogen (N_2) into ammonia (NH_3) by a process called biological nitrogen fixation (BNF) (Howieson and Dilworth, 2016). The nitrogen (N) fixed by BNF can account globally for more than 200 Tg N y^{-1} and is thus the main natural source of assimilable nitrogen (N) (Vitousek *et al.*, 2002; Gruber and Galloway, 2008). Legume crops in pastures are important for their capacity of fixing N, which is estimated between 200 and 300 kg N ha^{-1} dependent on the region (Peoples *et al.*, 1995; Sessitsch *et al.*, 2002; Alves *et al.*, 2003; Iannetta *et al.*, 2016). In the UK, the two main leguminous crops are peas (*Pisum sativum* L.) and faba bean (*Vicia faba* L.), which on average are grown on 3.1% of arable land annually (FAOSTAT, 2021). The yields of these crops have traditionally been characterised as unstable, defined by the significant yield variation from one season to the next; which has led to their decreased use in modern agriculture in recent decades (Jensen *et al.*, 2010). However, legumes can also provide food and feed with a high protein and carbohydrate contents, in addition to a range of other direct or indirect services such as improved soil structure, reduction of the use of pesticides (by diversifying the crops used in a rotation), and the reduction of greenhouse gas from the avoided manufacture, transport and decomposition of synthetic N-fertilisers produced by the Haber-Bosch process (Jensen *et al.*, 2010).

For nodulation to occur, both symbiotic partners need to recognise each other and be compatible (Maunoury *et al.*, 2008; Andrews and Andrews, 2017). Both peas and faba bean nodulate with *Rhizobium leguminosarum* bv *viciae* (Rlv), although this bacterium is also capable of nodulating with other legume crops and wild legumes such as vetches (*Vicia* spp.), lentil (*Lens culinaris* L) and vetchlings (*Lathyrus* spp.) (Howieson and Dilworth, 2016). In UK soils, there are native populations of compatible rhizobia for pea and faba bean (ca.10^4 bacteria per gram of dry soil) (Hirsch, 1996; Macdonald *et al.*, 2011; Mauchline *et al.*, 2018), although not all populations are equally infective or effective at fixing N. However, nodulation does not always directly correlate with the rate of N fixation and does not necessarily translate into improved plant growth. This has been shown for a range of rhizobia isolated from *Vicia, Lathyrus* and *Pisum* species which, when re-inoculated with the same or different host, some were capable of inducing nodulation without fixing N, while others could not nodulate at all (Mutch and Young, 2004). Therefore, it is likely that some of the Rlv interactions with pea and faba bean roots will not produce nodules, or some of them will generate nodules without benefiting the host (e.g., due to the increased sink strength of the nodule), which may contribute to the yield instability of these two crops.

Delivering known strains of rhizobia (that are both infective and effective at fixing N) directly to the crop could contribute to reducing instability in yield. Compatible bacteria can be added to a crop at the time of sowing, in the form of either a solid or liquid inoculant, i.e. a mixture of bacteria in a carrier substance such as peat (Deaker, 2004). Inoculation is a common practice in Australia and the Americas, and is essential when the soil contains no compatible bacteria (Howieson and Dilworth, 2016). A critical step during inoculant production is the transfer of the bacteria from the optimal growth conditions in the fermenter to being mixed with the carrier (Bashan, 1998; Bashan *et al.*, 2014). This environmental change produces the highest die-off of bacterial cells (Date, 2001), mainly due to desiccation stress which is the main cause of inoculants losing their effectiveness (Deaker, 2004; Casteriano *et al.*, 2013; Casteriano, 2014). In common with all Gram-negative bacteria, *Rhizobium leguminosarum* (Rleg) has two cell envelopes and does not form spores (an environmental stress resistant structure) and therefore desiccation stress is often lethal (Berninger *et al.*, 2018).

Rhizobial strains with some degree of desiccation tolerance are therefore desirable, and could increase the final shelf-life of inoculant products (Bashan *et al.*, 2014; Molina-Romero *et al.*, 2017). Rhizobia isolated from arid areas or from soil that has been previously dried are more likely to tolerate desiccation compared to rhizobia from wetter environments or the same soil prior to drying (van Ham *et al.*, 2016). Thus, it is supposed that strains isolated from a semi-arid environment will tolerate desiccation better than strains from wetter environments providing them with an industry-desired phenotype that will ultimately extend inoculant shelf-life.

The east of Spain has a characteristic semi-arid Mediterranean climate with dry and hot summers, followed by mild winters (Peel et al 2007, Perez-Cuevas, 1994). Rhizobium is a bacterial genus native to Spanish soils and legumes such as faba bean and pea are widely grown in this country with an average annual production in the first two decades of the 21st century of 38,372 and 124,712 tonnes per year respectively (FAOSTAT, 2021). Therefore, the main aim of this chapter was to isolate and select pea-nodulating rhizobia from soil from eastern Spain and assess their symbiotic performance in glasshouse and field experiments and their potential use in commercial inoculants. We hypothesise that pea, used as a trap plant, will spontaneously nodulate with free-living compatible rhizobia from the *Rhizobium leguminosarum* species complex (Rlc) (Young *et al.*, 2021) and that some of these will present similar symbiotic performance in terms of plant biomass and increased yield to that of commercially used strains. To address this aim, the study had the following objectives: (I) the isolation of rhizobia from soil sampled from a semi-arid environment; (II) to genetically characterise and identify pure lines of isolated rhizobia; and (III) quantify the symbiotic and N-fixing ability of these rhizobia in glasshouse and field experiments.

2.2 Materials and methods

2.2.1 Soil sampling and chemical composition analysis

Soil was sampled from five different low N input fields in two localities, Ontinyent and Valencia, in the eastern region of Spain of the Valencian Province. Valencia (39°28'13.29"N, 0°22'33.87"W) is the most northern of the two with an elevation of 20 m above sea level while Ontinyent (38°49'19.57"N, 0°36'23.27"W) is situated 75 km south from Valencia at 350 m above sea level in a limestone valley called La Vall d'Albaida. Both areas have a semi-arid Mediterranean climate (Peel *et al.*, 2007) characterised by a dry and hot summer followed by a mild winter (Pérez-Cuevas, 1994).

Soil was collected from a random 400 cm^2 area within each field where the top 20 cm of soil had been mixed homogeneously. From this mixture, ca 16 g of soil were sampled in a plastic tube. Between 1 and 3 soil samples were taken from each of the five fields giving a total of twelve soil samples in total, which was used for the subsequent trapping of rhizobia. An additional soil sample from each field was sampled for chemical analysis. For this, the field was subdivided into six square areas and in each one of them a 'W' walk was carried out sampling ca 400 g of the top 20 cm of soil at each vertex and twice between vertexes, the orientation of the 'W' was different in each area to produce a random sampling of the field. The resulting 2.4 kg of topsoil was mixed well in a bucket and ca 1 kg was sampled and taken to the laboratory. All samples were kept at 4°C until arrival at the laboratory where a sub-sample from each sample (ca. 1 g) was snap frozen in liquid N and stored at -80°C. The remaining sample was stored in a cold room at 4°C. From the second set of soil samples, a sub-sample of 400 g was shipped to Lancrop Laboratories (York, UK) for chemical analyses following the company's usual standard operating procedures. The soil was processed to measure pH and quantify phosphorus (P), potassium (K), magnesium (Mg) concentration (Analysis code S1); and soil mineral N (Analysis code SA10) which measures nitrate N (NO_3^-) and ammonia N (NH_3) concentration.

2.2.2 Trapping *Rhizobium* spp. using pea

Seeds of pea cv. Corus, one of the most frequent vining pea cultivars used in Scotland (Wardlaw *et al.*, 2019), were aseptically surface sterilised in a laminar flow hood. Briefly, seeds were immersed in a 95 % v/v ethanol solution for 30 seconds with continuous mixing followed by a rinse with sterile distilled water (SDW). Seeds were subsequently immersed in a 3 % sodium hypochlorite solution with a drop of TWEEN 20 (Sigma-Aldrich, Germany) for 5 minutes followed by five rinses with SDW. After the fifth rinse, the seeds were aseptically transferred to sterile Petri dishes containing 0.5 % w/v water agar and incubated at 28°C until germinated.

A mix of perlite and vermiculite 50:50 v/v was added to 1 L pots and moistened with 400 mL of distilled water. All pots were autoclaved inside autoclavable bags at 121°C for a minimum of 15 minutes. Once the seeds were germinated, each of the twelve soil samples were suspended in 300 mL of SDW. Five germinated seeds were sown in each sterile pot and 100 mL of the suspended soil was added to each one of the three replicate pots per soil sample, together with enough SDW to provide a moist environment for the seedling. Following this, all pots were transferred to a glasshouse with a controlled environment of 16 h / 8 h (light/darkness) photoperiod and 20°C / 14°C day/night temperatures respectively. The growth of the plants was checked every three days and they were watered when needed with SDW. After the emergence of the third seedling any further seedlings were removed to allow enough space for these plants to grow without stress. After 6 – 8 weeks, each plant was removed from the pot, and all the substrate was carefully removed from the root system under running tap water. Following this, all nodules on the root of each plant were counted, and all those N-fixing active nodules showing Leghaemoglobin characteristic red/pink colour were harvested. From these, a random two nodules per plant were selected for isolation of rhizobia, and the rest were snap frozen in liquid N and stored at -80 °C.

2.2.3 Isolation, identification, and characterisation of rhizobia

2.2.3.1 Nodule bacteria isolation

Nodules were surface sterilized with a 3 % sodium hypochlorite solution, rinsed with SDW three times and homogenised in 500 µL of SDW in a 1.5 mL tube with a sterile pestle. The homogenate was aliquoted and streaked onto two sterile Petri dishes containing yeast mannitol agar and Congo red (YMA-CR) (10 g mannitol, 0.5 g glutamate, 0.5 g K_2HPO_4, 0.1 g $MgSO_4 \cdot 7H_2O$, 0.05 g NaCl, 1 mL (40 g L^{-1}) $CaCl_2$, 1 mL (4 g L^{-1}) $FeCl_3$, 1 g yeast extract (Fermtech, Merck), 15 g technical grade agar (Difco), 10 mL 0.25 % w/v Congo Red, 1 L SDW and pH 6.8), and incubated at 28 °C. Plates were checked every other day and when growth was visible, single colonies were passaged onto Petri dishes with fresh YMA-CR medium. Once the cultures were 'clean' (i.e. growth of only one colony type), a single colony was picked off and inoculated in tubes of tryptone yeast (TY) broth (5 g tryptone, 3 g yeast extract, 1 L SDW and pH 6.8) and incubated at 30 ºC and 120 rpm for 48 hours or until the media culture was cloudy or reached log phase growth (OD_{600} 0.2 – 0.8). After incubation, 1 mL of the solution was mixed with the same volume of 50 % v/v glycerol and incubated at room temperature for 20 minutes after which they were snap frozen with liquid N and stored at -80 °C to generate a glycerol stock of each strain.

2.2.3.2 Bacterial genomic DNA isolation

The remaining culture from each isolate was centrifuged at 1,900 RCF for 15 minutes. The supernatant was removed, and the pellet re-suspended in a lysis buffer (10 mM Tris-HCl, 1 mM EDTA, 0.5 % SDS and 19.05 units mL^{-1} Proteinase K) and incubated for at least one hour at 37 °C. After incubation, the same volume of phenol:chloroform:isoamilalcohol (Sigma-Aldrich) was added, vortexed for a few seconds and centrifuged at 13,200 RCF for 12 minutes. After centrifugation, the top layer of the tube content was recovered and mixed with a solution of 78 mM sodium acetate in 100 % propanol, and incubated overnight in a freezer at -20°C. After incubation, the sample was centrifuged at 13,200 RCF for 12 minutes, the supernatant discarded, and the pellet resuspended in 70 % ethanol and vortexed. The sample was centrifuged again at 13,200 RCF for 12 minutes followed by the removal of the supernatant and the sample was stored in an incubator at 28 °C until all ethanol had evaporated. Once dry, the pellet was resuspended in SDW and the DNA concentration assessed with a NanoDrop ND-1000 (NanoDrop Technologies, Inc., Wilmington, USA).

2.2.3.3 PCR and PCR product sequencing

For bacterial identification, the primer pair 8-27F and rD1 (Weisburg *et al.*, 1991) were used for the partial amplification of the 16S rRNA region of the core genome (Table 2.1). The master mix for the PCR reaction contained GoTaq® G2 DNA Polymerase (Promega, USA) used at manufacturer recommendations with 10 mM of dNTP (Invitrogen, USA) and 10 µM of both forward and reverse primers, together with 1 µL of isolated DNA. The PCR was run with an initial denaturalisation step at 95 °C for 2 minutes followed by 35 PCR cycles (denaturalisation at 95 °C for 1.5 minutes, annealing step at 58 °C for 1 minute and elongation at 72 °C for 1.5 minutes) and a final elongation step at 72 °C for 15 minutes in a G-Storm GS1 thermal cycler (GRI Ltd, Braintree, UK).

Table 2.1. Primers used for 16S- and BOX-PCR procedures. Y = T or C, R = A or G, K = T or G, M = A or C, W = A or T, N = any base.

Usage	Primer	Sequence	Reference
PCR	fD1 (8-27F)	5' - AGA GTT TGA TCC TGG CTC AG - 3'	Weisburg *et al.* (1991)
	rD1	5' - AAG GAG GTG ATC CAG CC - 3'	
BOX PCR	BOXA1R	5' - CTA CGG CAA GGC GAC GCT GAC G - 3'	Versalovic *et al.* (1994)

PCR products were subsequently run in a 1% w/v agarose gel in 1x TBE with SYBR Safe (Invitrogen, UK) and visualised under UV light in a UVP BioDoc-It^2 Imager (Analytik Jena GmbH, Germany). The PCR product size was estimated by comparison with a 1 kb DNA ladder (Promega). When the PCR product showed a single band, a 6 µL aliquot of the remaining PCR product was purified with illustra™

ExoProStar™ 1-Step (GE, USA) following manufacturer recommendations. The purified PCR product was processed for sequencing using a Big Dye Terminator v3.1 kit (Applied Biosystems, USA) and a Bio-Rad Tetrad 2 thermal cycler (Bio-Rad Laboratories Inc, USA) at manufacturer's recommendations. The sequencing reactions were then precipitated, cleaned and analysed in an ABI3730 DNA analyser (Applied Biosystems, USA) by Sanger sequencing.

2.2.3.4 Phylogenetic analysis

Raw data from the sequencing analyser was processed with BioEdit v 7.0.5.3 (Hall, 1999) for reversing and complementing the reverse primer sequence. All forward and reverse sequences were trimmed and contrasted with the National Center for Biotechnology Information (NCBI) database through the nucleotide basic local alignment tool (BLASTn) on the NCBI website. Subsequently, both forward and reverse sequences were aligned with Mega-7 v 7.0.26 (Kumar *et al.*, 2016) by Clustal-W and the full partial sequence was queried on the NCBI BLASTn suite for confirming the previous match, and 16S rRNA sequences from type species of these matches (when available) were downloaded from NCBI for the phylogenetic analysis.

All sequences were split between two datasets, one with Rhizobiales isolates and NCBI sequences (plus an outgroup Betaproteobacteria), and a second one with all other strains and the corresponding NCBI sequences (outgroup genus *Bacillus*).

The sequences were aligned by Muscle (Edgar, 2004) on Mega-X version 10.0.5 (Kumar *et al.*, 2018) with default gap opening and gap extension penalties for DNA (-400 and 0 respectively), a maximum number of iterations of the algorithm of 16, clustering method UPGMA and a minimum diagonal length of 24. After the alignment, all positions with gaps were removed using BioEdit, and the pairwise distances were computed for removing duplicated sequences by the 'number of differences' method. The overall nucleotide identity was calculated to assess the quality of the alignment by the p-distances method, and the suitability of the dataset for running the Neighbour-Joining phylogenetic analysis was calculated by estimating the Jukes-Cantor distance. Following this, the gamma parameter was calculated to adjust the rates among sites of the phylogenetic analysis.

Finally, the phylogeny was estimated by Neighbour-Joining by the Maximum Composite Likelihood model and with the calculated gamma parameter. The phylogeny was tested by the Bootstrap method with 1000 replications (Felsenstein, 1985; Saitoh, 1987; Tamura *et al.*, 2004).

2.2.3.5 BOX PCR run and band analysis

Strain diversity was assessed by BOX PCR for all confirmed *Rhizobium leguminosarum* strains from the phylogenetic analysis. First, the concentration of DNA for all samples was adjusted to 12.5 ng µL^{-1}, and a PCR master mix was prepared following the same protocol as described in Section 2.2.3.3 with two differences: only the palindromic primer BOXA1R (Table 2.1)(Versalovic *et al.*, 1994) was used at double volume per reaction to achieve the same primer concentration and the volume of template DNA was doubled. The PCR was run with an initial denaturalisation step at 94 °C for 30 seconds followed by 35 PCR cycles (denaturalisation at 94 °C for 10 seconds, annealing step at 50 °C for 30 seconds and elongation at 72 °C for 30 seconds) and a final elongation step at 72 °C for 10 minutes in a G-Storm GS1 thermal cycler (GRI Ltd, Braintree, UK).

The PCR product was run on a 2 % w/v agarose gel in 1x TBE with SYBR Safe at manufacturer recommendations, run in 1x TBE and a 1 kb DNA ladder (Promega) as used in Section 2.2.3.3 for band length estimation. The gels were read under UV light as described above.

All images taken were analysed on a VisionWorks Acquisition and Analysis software v 8.20.17096.9551. The software settings were setup in the ‘Find Lanes and Bands’ menu with lane and band sensitivity set to 20 and 92 respectively and the boxes ‘constant lane width’ and ‘force all lanes straight’ ticked. For standardisation of bands generated in different gels, all images were calibrated with the 1 kb DNA ladder (Promega) included in each individual gel by selecting the manufacturer’s ladder details downloaded from their website. After the software automatically detected lanes and bands, all bands wrongly identified were removed and bands not identified were added.

A dendrogram was constructed with the calibrated band lengths using the tool in the same software selecting the Jaccard similarity coefficient (Jaccard, 1901) as a lane formula and Unweighted Pair-Group Method Average (UPGMA) (Sokal and Michener, 1958) as a linkage formula.

2.2.4 Screening of bacterial isolates on plant

Initially, 40 strains were selected based on their BOX PCR band pattern and were screened on pea cv. Corus in a glasshouse during a preliminary experiment (Appendix 2.1 | Screening and selection of rhizobia). From this initial group of 40 strains, six were selected based on their capacity of increasing biomass on pea and were assessed in more detail with increased replication and in a controlled environment growth room in sterile conditions.

Table 2.2. Positive and negative control treatments used in screening experiments.

Controls	Strain	Isolated from	Country of origin	References
Positive (PC)	rcr1045	*Pisum sativum* L.	Ireland	(Dye, 1978; Bitanyi, 1983)
	JHI388	*Pisum sativum* L.	Scotland	(Maluk *et al.*, 2022)
	USDA2364	*Pisum sativum* L.	Virginia, USA	(van Berkum *et al.*, 1995)
Negative (NC1)	SDW	**-**	-	-
Negative (NC2)	Autoclaved rcr1045	-	-	-

The experimental setup for screening rhizobia was similar to the trapping of rhizobia from soil (Section 2.2.2). In this case, the 1 L pots were prepared and sterilised as described above. Rhizobia were screened with two pea cultivars, Corus and Kareni, and one faba bean cultivar, Fuego, in separate experiments. Seeds of these cultivars were surface sterilised and incubated at 28 °C until germination. Three germinated seeds were sown in each pot, watered with sterile N-free rooting solution (Burchill *et al.*, 2014) under sterile conditions and moved to a controlled environment growth room (16 h / 8 h light/dark photoperiod, 23 °C / 15 °C day/night temperature and 70% relative humidity). The pots were checked daily until the seedlings had emerged at which point each pot was thinned to one plant per pot.

Individual cultures of the six rhizobial isolates plus a range of known effective nodulating *R. leguminosarum* strains which were used as positive controls (Table 2.2), were grown in TY broth until they reached log phase growth (OD_{600} 0.2 - 0.8) and adjusted to the same absorbance before inoculating 1 mL of the adjusted culture next to the emergence area of the seedling. Each experiment included at least one negative control, either an uninoculated treatment where SDW was added instead of a grown culture (NC1) or an autoclaved PC culture treatment (NC2) (Table 2.2). Five replicates per treatment were prepared for each crop cultivar which were randomly placed in the growth room. The pots were checked every other day and watered with sterile N-free rooting solution (Burchill *et al.*, 2014) when needed and harvested at 33 days after sowing.

At harvest, each plant was removed from the pot, any adhering substrate was carefully removed, and the root systems cleaned with running tap water. Shoot, root, and nodule fresh weight were measured. Shoots and roots were placed in individual paper bags and dried for 72 h at 70 °C, and the dry weights measured. The freshly harvested nodules were carefully arranged on a flat surface and an image taken with a Canon EOS 1200D camera with a Canon EF-S 18-55 mm f/3.5-5.6 III lens or a Sigma 105 mm f/2.8 EX DG OS HSM macro lens depending on availability. When using the Sigma lens, pictures were taken at 105 mm focal length, F-stop f/14, 1/200 s of exposure time and ISO-1600, and when the Canon lens was used, pictures were taken at 55 mm focal length, 1/250 s exposure time, F-stop f/25-36 and ISO-400-3200 depending on the lighting conditions (Appendix 2.2 | Optimisation of a macro for automated counting and measuring of nodules using digital images). After the image was captured, all nodules from the same plant were put in a plastic tube and dried at 70 °C for at least 12 h and the dry weight measured. The nodule images were analysed on FIJI ImageJ v1.52n (Schindelin *et al.*, 2012; Rueden *et al.*, 2017) using two custom scripts for measuring the nodule number, length, perimeter, area, and the shape parameters, circularity and solidity (Appendix 2.2 | Optimisation of a macro for automated counting and measuring of nodules using digital images). Furthermore, the below-ground mass fraction (Bgm) (i.e. the combined dry weight of nodules and roots divided by the total biomass dry weight) and nodule mass fraction (Nmf) (dry weight of nodules divided by the combined dry weight of root and nodules) was calculated.

2.2.5 Field trial

The two strains that showed the highest biomass increase from the growth room experiments were selected to be tested in field conditions. These two strains were tested against a commercial standard strain (rcr1045), the best performing isolate in terms of pea biomass increase from the James Hutton Institute collection (JHI388) and a non-inoculated control in two field trials carried out in 2019 and 2020 with pea cvs Corus and Kareni during the first year and cvs Zero4 and Kareni during the second year.

2.2.5.1 *Inoculant preparation*

For the preparation of inoculant, all strains were grown in YMB at 30°C and 150 rpm for 48 h. After the incubation, cultures were adjusted to the same absorbance, and 2.61 mL of this solution was diluted in 150 mL of YMB in triplicate. The resulting dilution was injected into a small sterile peat bag supplied by Legume Technology Ltd. (Nottingham, UK) and mixed well by massaging the bag. The sealed peat bag was then placed in an incubator at 28 °C for 7 days.

Following the incubation, two of the peat bags were stored at 4°C until they were used, whilst the third bag was opened in sterile conditions and mixed well. A sub-sample of the peat inoculant was diluted 1:1 w/v in SDW and mixed well. The resulting slurry was serially diluted 1:10 in 0.85 % w/v saline and a 50 µL aliquot from each dilution was inoculated in YMA+CR Petri dishes and incubated at 28°C. Once the colonies were visible, the concentration of rhizobia in each bag was calculated.

2.2.5.2 *Seed weighing and inoculation*

For the 2019 field trial, non-dressed pea seeds cv. Corus (92.9 g thousand seed weight (TSW) and 97 % germination) were purchased from Syngenta (France), and cv. Kareni (294 g TSW and 97 % germination) purchased from Senova (UK). For the second year, the same Kareni seed was used but cv. Corus was substituted with cv. Zero4 (190 g TSW and 99 % germination) purchased from Limagrain (UK). These cultivars were selected based on their frequent choice by farmers in the area where the trials were run and being in the recommended lists of the Processors and Growers Research Organisation (PGRO, 2013, 2017; Wardlaw *et al.*, 2019). Seeds were stored at 4°C and 14 % RH upon arrival until use. Sowing densities targeted

70 plants m^{-2} for cvs Corus and Kareni and 110 plants m^{-2} for cv. Zero4 (PGRO, 2013, 2017). The actual seed weight sown into each plot accounted for between 1 and 3 % germination loss depending on the cultivar germination and 13 % field loss. Thus, for each plot of cv. Corus, 73 g of seed were placed in individual paper bags, for cv. Kareni, 232 g of seed and for cv. Zero4, 240 g of seed. All weighed seed was stored in the seed store at 4 °C in dark until inoculation and sowing.

In the 24 h before sowing the field trials, the seeds were treated with the inoculants as described in Section 2.2.5.1. When inoculating the seeds, the content of each paper bag was moistened with SDW at a ratio of 2 mL kg^{-1} and mixed well. After this, the peat inoculant was added at a ratio of 4 g of inoculant per 1 kg of seed and mixed well. Once applied, the inoculated seed were returned to the paper bags and stored at 4 °C in a cold room until sown.

2.2.5.3 Experimental design and data collection

In both years, each pea cultivar was sown in 8 rows at 3 cm depth in three blocks formed by twenty 6.25 x 1.55 m plots divided in four rows and five columns (Figure 2.1 A). Each block had four replicates of each of the five seed treatments (four strains plus un-inoculated control). The plots in each column were sown with the same treatment and the columns within blocks were organised randomly (using a random number generator) so that each block had the treatment columns in random order. To minimise the edge effect, all plots and blocks were surrounded by ‘guard plots’ of the same area. The crop was managed with conventional legume arable practices common for this part of the country, like ploughing and harrowing prior cultivation, pre-emergence herbicide application (Stomp Aqua 2.9 L ha^{-1}, BASF) and no addition of N fertiliser.

Seed emergence was monitored at 2-3 day intervals until 23 days after sowing (DAS) in the same 1 x 1.55 m area within each plot (Figure 2.1 B and C). From these data the following germination and emergence parameters were calculated:

The time to 50 % germination (T_{50}) was calculated according to (Farooq *et al.*, 2005) equation (Eq. 2.1):

$$T_{50} = t_i + \frac{\left(\frac{N}{2} - n_i\right) \times (t_j - t_i)}{(n_j - n_i)} \qquad \text{(Eq. 2.1)}$$

where N is the maximum number of emerged seeds and n_i and n_j are the accumulated number of emerged seedlings by adjacent counts at times t_i and t_j when the condition $n_i < \frac{N}{2} < n_j$ is given.

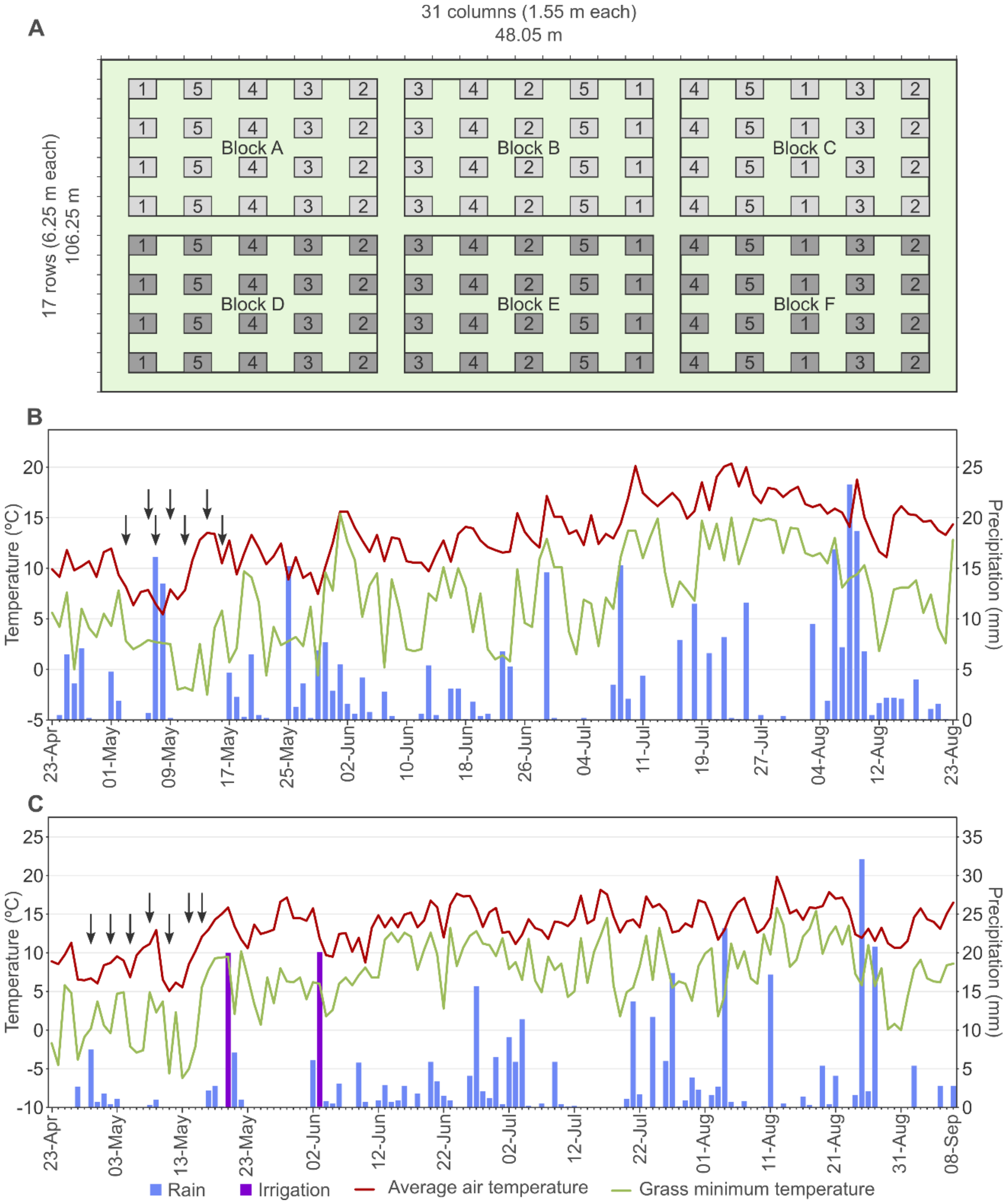


Figure 2.1. Example of field trial experimental layout (A) and temperature and precipitation plots for 2019 (B) and 2020 (C). The experimental blocks are delimited by the black border, light grey squares are plots with pea cv. Corus, dark grey for pea cv. Kareni and the light green area are barley guard plots. The numbers

inside each plot are the inoculant treatment. The arrows in plots B and C mark emergence count events. Both B and C x-axis start at sowing and end the day the t was harvested.

The final percentage of germination (FGP) was calculated as (Eq. 2.2):

$$FGP = \frac{N}{N_{exp}} \times 100 \qquad \text{(Eq. 2.2)}$$

where, as above, N is the maximum number of emerged seeds and N_{exp} is the number of expected seeds in the monitored area after accounting for germination and field loss.

The mean germination time (MGT) was calculated with the Ellis and Roberts (1980) equation (Eq. 2.3):

$$MGT = \frac{\sum_{i=1}^{k} n_i \times t_i}{\sum_{i=1}^{k} n_i} \qquad \text{(Eq. 2.3)}$$

where in this case n_i is the number of non-accumulated emerged seedlings counted at the ith time, t_i is the time in days at which they were counted and k is the last counting time.

In the second year, six plants per plot were removed at early- to mid-pod fill stage (growth stage 206 on Knott, 1987) and dried for 72 h at 70 °C for assessing shoot dry weight.

Once the peas had reached growth stage 301 (Knott, 1987), desiccant was applied to achieve a uniform and synchronised drying of the seeds and after two weeks they were harvested. Following this, seeds were placed in a drying area to reduce the moisture content to a suitable level for storage. After drying until about 13 % moisture content, the seeds were cleaned of any debris from pods and total yield was weight, TSW and seed size distribution were measured with a Marvin Seed Analyser 176 (Marvitech, Germany) and moisture and protein content were measured using an Infratec 1241 Grain Analyzer (Foss, Denmark).

2.2.6 Statistical analyses

All statistical analyses were carried out using R software v 3.5.1 (R Core Team, 2019) implemented on RStudio v1.1.456 (RStudio Team, 2020). R package dplyr v 0.8.3 (Wickham *et al.*, 2019) was used for data handling and processing.

One- and two-way ANOVA and generalised linear models were performed using R base package v 3.5.1(R Core Team, 2019). Linear mixed effects models (LMM) and generalised linear mixed effects models (GLMM) were performed with package lme4 v 1.1.21 (Bates *et al.*, 2015). If treatment effects were significant, the Tukey HSD test was used for treatment pairwise comparison with the package multcomp v 1.4.10 (Hothorn *et al.*, 2008).

For field trial data, treatment pairwise comparison was implemented by Least Significant Difference (LSD) and P-values adjusted for false discovery rates by the BH method (Benjamini and Hochberg, 1995) both implemented on package predictmeans v 1.0.4 (Luo *et al.*, 2020).

For the LMM and GLMM models used on the screenings, 'treatment' was used as a fixed effect and 'plant' as a random effect (i.e., Biomass~Treatment + (1|Plant)). For the models used on the field data, treatment and cultivar were used as fixed effects and column and block as a nested random effect (i.e., Yield~Cultivar*Treatment + (1|Block/Column)).

Percentage data such as those of the FGP parameter were transformed by the arcsine of the square root before analysing the data by LMM.

For data visualisation, R package ggplot2 v 3.3.0 (Wickham, 2016), R base v 3.5.1 (R Core Team, 2019) and dotwhisker v 0.5.0 (Solt and Hu, 2018) were used for sketching the plots and were later optimised for publication on Inkscape v 0.92.

2.3 Results

2.3.1 Trapping isolation and selection of pea rhizobia

The chemical analysis of the soil samples from Spain showed that all soil samples had an alkaline pH and low nitrate content (Table 2.3). All plants grown in these soils had large root nodules, with red (or pink) nodule-cores. A total of 103 bacterial lines were isolated from these nodules. Nodules were also present on the roots of negative control plants; however, these nodules were small and white suggesting late contamination of the negative control pots.

The NCBI BLASTn of the partial 16S rRNA sequences of the isolated bacteria revealed that 78.6 % showed high similarity to sequences of the Rhizobium genus (Table S2.2). Both the Rhizobiales and non-Rhizobiales sequence datasets met the standard thresholds for proceeding with the Neighbour-Joining (NJ) phylogenetic analysis (Nei and Kumar, 2000; Kumar and Filipski, 2007). The nucleotide identity was greater than 66 % (93 % for the Rhizobiales dataset and 85 % for the non-Rhizobiales) and the Jukes-Cantor distance was lower than 1 (0.07 for Rhizobiales dataset and 0.17 for non-Rhizobiales). The phylogenetic analysis of all Rhizobiales lines involved 67 nucleotide sequences with 1151 positions where all gaps were removed and ambiguous positions deleted. The phylogeny was inferred by NJ and tested by the bootstrap test with 1000 replicates. The resulting tree had a total branch length of 1.396 substitutions. The results of the phylogenetic analysis confirmed the preliminary result where 90 % of Rhizobiales strains were clustered within the Rhizobium branch next to the *R. leguminosarum* or *R. laguerreae* type strains (Figure 2.2).

Table 2.3. Chemical analysis and the legume cropping history of the field at time of the second sampling event. Symbols correspond to the ADAS soil index (numbers) and its interpretation: (↓) index 0 to 2 or very low to medium; (−) index 3 or medium to high; (↑) index 4 or high; (↑↑) index 5 to 9 or very high.

Field	Location	pH	P (mg L^{-1})	K (mg L^{-1})	Mg (mg L^{-1})	N (mg kg^{-1})	Time since last legume crop (y)
F1	Valencia	7.8	59 ↑	746 ↑↑	343 ↑↑	36.7 ↓	>1
F2	Valencia	7.9	58 ↑	929 ↑↑	246 ↑	35.8 ↓	>1
F3	Valencia	8.1	104 ↑↑	689 ↑↑	250 ↑	12.6 ↓	<1
F4	Ontinyent	8.3	20 ↓	853 ↑↑	284 ↑↑	29.3 ↓	>2
F5	Ontinyent	8.3	31 −	862 ↑↑	143 −	10.5 ↓	>5

The dendrogram generated using the BOX PCR band patterns from all *Rhizobium* isolates showed some level of grouping dependent on the strain field origin (Figure 2.3). All field F3 isolates grouped on adjacent branches and were distinct from the isolates from other fields. Some F3 isolates also clustered with one group of the negative control plant isolates. Isolates from the negative control pots showed high resemblance between the band patterns and grouped together in two main groups. One group contained most of the isolates from the negative control pots and two F3 strains: 52A12 and 52B11. The second group contained mainly F3 isolates and two negative control strains: NC1A21 and NC1A11. Isolates from the other fields were dispersed over the other branches of the dendrogram, yet field F4 was the only field that was present in most of them.

Forty-three representative *Rhizobium* strains from all fields and from all BOX PCR dendrogram groups were selected to assess their plant growth-promoting potential on pea cv. Corus during preliminary experiments (data shown on Appendix 2.1 | Screening and selection of rhizobia). From these experiments the four strains with the highest biomass increase and two strains with performance comparable to *R. leguminosarum* bv. *viciae* 3841 were selected for further screening on different pea and faba bean cultivars.

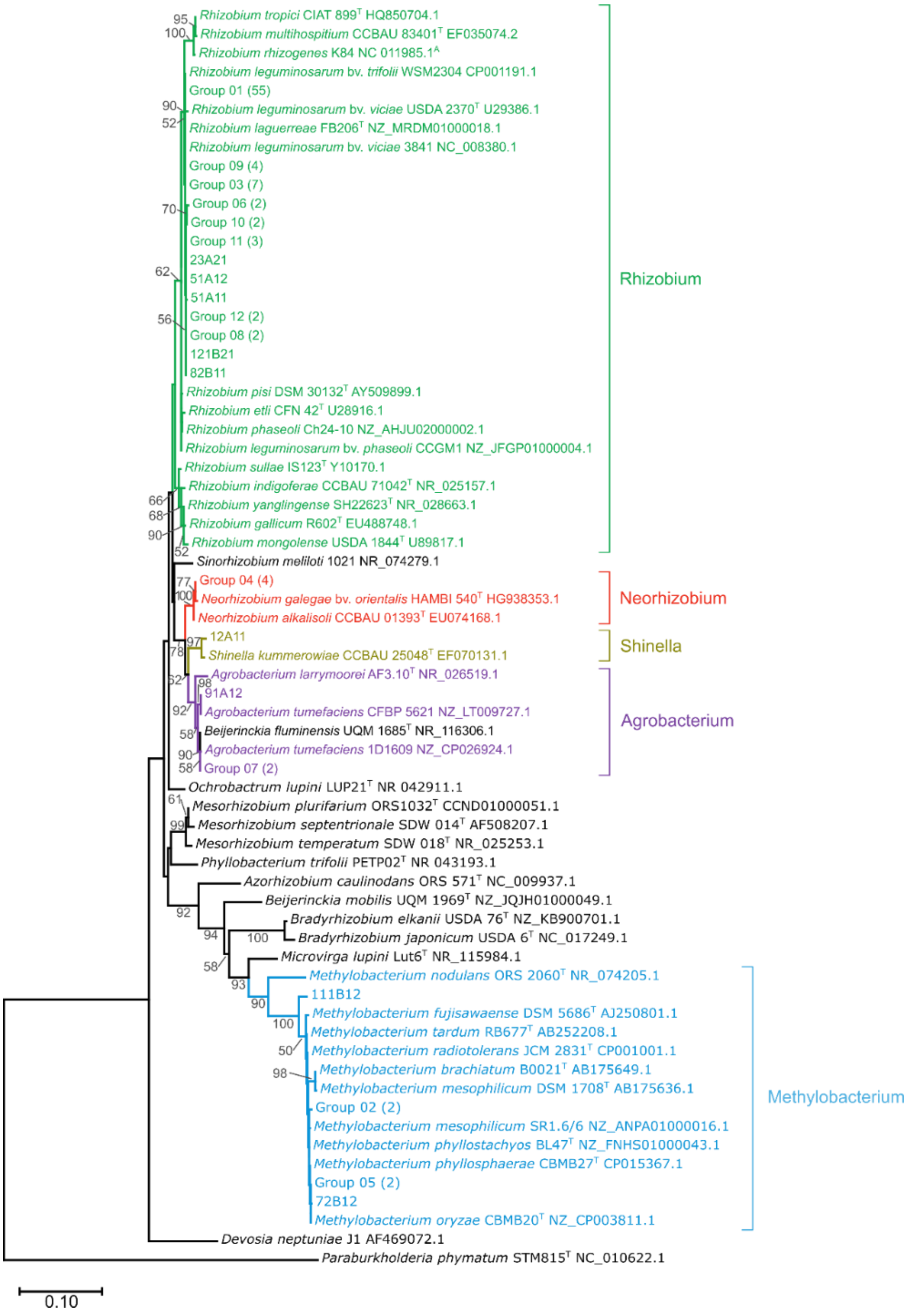


Figure 2.2. Phylogenetic tree of all order Rhizobiales matched isolates for 16S rRNA partial sequences, and only nodes with more than 50 % of support are

shown. The bar represents the number of base substitutions per site estimated by the Maximum Composite Method (Nei and Kumar, 2000). A gamma distribution was used for modelling the rate variation among sites (Shape parameter = 0.15). The number in brackets next to the group number indicates the number of isolates in that group. Type sequences are marked by a superscript ‘T’ after the strain code. The currently accepted name for *A. radiobacter* strain K84 (A) was used as it was mis-labelled on NCBI (Lindström and Young, 2011).

2.3.2 Screening of selected isolates on pea and faba bean

The six selected strains (together with strains with known good symbiotic performance) were screened on pea cvs Corus and Kareni and faba bean cv. Fuego. For the screenings on both pea cultivars, either one, or both of the strains JHI388 and 63A21 showed atypical growth when preparing the TY culture for inoculation. This resulted in a low efficiency for increasing pea biomass when compared with their effectiveness during the preliminary experiments. Thus, the screening was repeated for these strains with a new culture recovered from glycerol on both pea cultivars with some of the test strains (together with the PC and NC), to allow comparison among all isolates and controls. The common treatments between both setups were compared, and apart from the differences between inoculated and non-inoculated control treatments, there were no significant differences between either setup for both pea cultivars (Table 2.4). At the time of harvest of the screening experiments, the root systems of all negative control treatment plants did not have any nodules.

Table 2.4. Two-way ANOVA for the common treatments between the first and second setups made for pea cv. Corus and cv. Kareni. The values for each treatment are the mean biomass dry weight (g) ± standard deviation. Biomass values with different letters in the same column indicate a significant difference between means.

	Treatment	**Pea cv. Corus** **Biomass dry weight (g)**	**Pea cv. Kareni** **Biomass dry weight (g)**
Setup 1	121B21	0.571 ± 0.10^{a}	1.499 ± 0.20^{a}
	51B21	0.692 ± 0.10^{a}	1.478 ± 0.25^{a}
	NC2	0.196 ± 0.05^{b}	0.695 ± 0.05^{b}
	rcr1045	0.496 ± 0.18^{a}	1.527 ± 0.24^{a}
Setup 2	121B21	0.616 ± 0.09^{a}	1.452 ± 0.24^{a}
	51B21	0.599 ± 0.10^{a}	1.653 ± 0.36^{a}
	NC2	0.212 ± 0.08^{b}	0.719 ± 0.13^{b}
	rcr1045	0.542 ± 0.09^{a}	1.418 ± 0.30^{a}
	Variable	**P**	
ANOVA	Treatment	<0.001	<0.001
	Setup	ns	ns
	Treatment x Setup	ns	ns

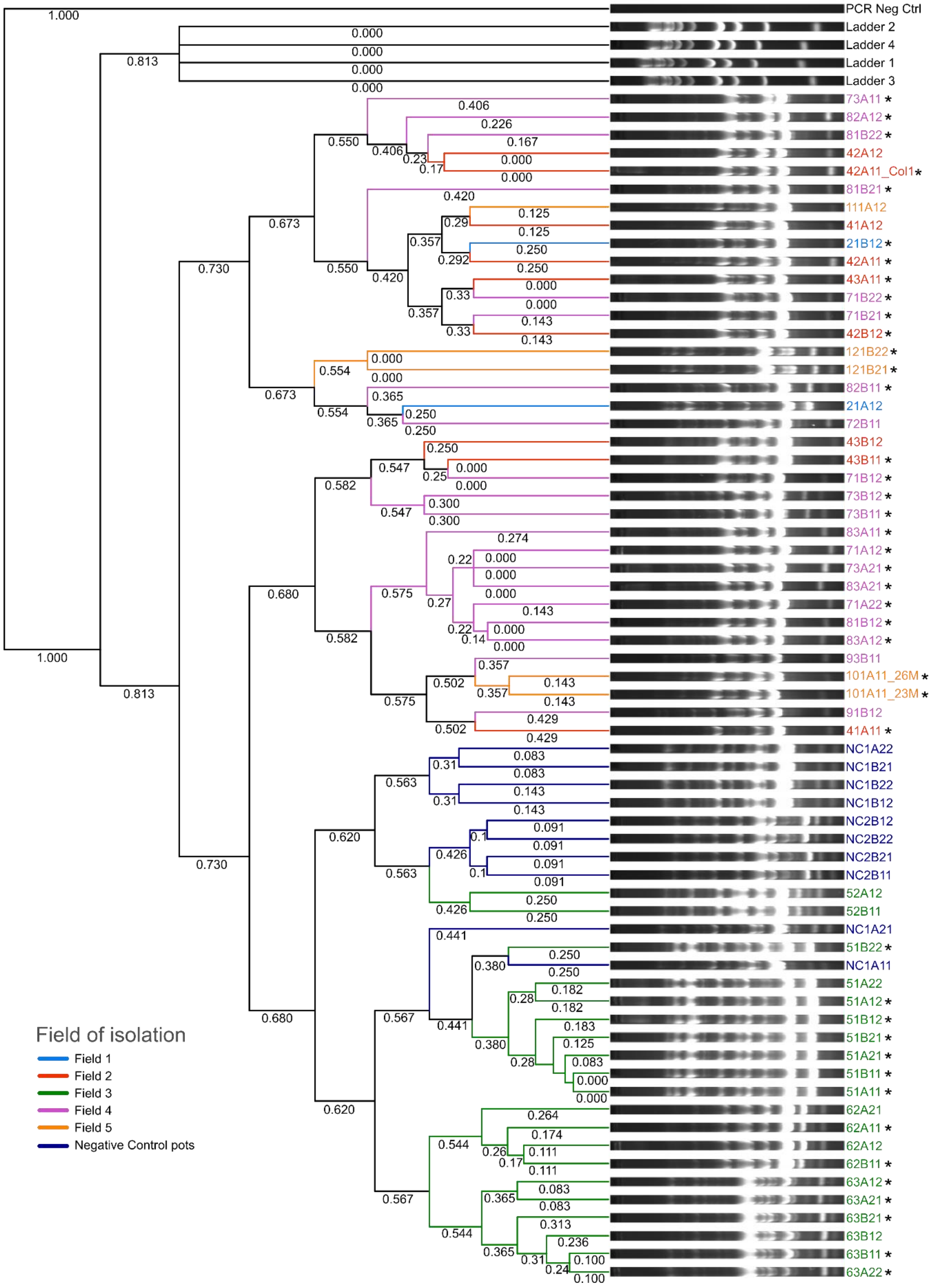


Figure 2.3. Dendrogram estimated by the analysis of the calibrated band lengths produced by BOX PCR with primer BOXA1R. The Jaccard Similarity Coefficient (Jaccard, 1901) was used for estimating the lane formula and UPGMA (Sokal and Michener, 1958) was used for calculating the linkage formula. Strain

code followed by a star indicate the strain was selected for Phase I screenings on pea.

The treatment applied to all crops had a significant effect on the total dry biomass of the plant after 33 days from sowing. On pea cv. Corus, the negative control treatments produced a significantly lower total biomass than most of the test strains and positive controls (Setup 1: df = 8, F = 18.116, P < 0.001; Setup 2: df = 6, F = 19.803, P < 0.001). Only test strain 21B12 and the PC JHI388 with atypical growth on the first setup showed a biomass increase similar to that of the NC treatments (Figure 2.4 A and B). The treatment had a similar result on both pea cv. Kareni setups (Setup 1: df = 9, F = 28.094, P < 0.001; Setup 2: df = 5, F = 10.986, P < 0.001). The biomass increase produced with strain 21B12 and the atypical PC JHI388 was significantly lower (P < 0.05) than that of treatments with other rhizobial strains but was similar to that of both negative controls (Figure 2.4 C). In the second setup of cv. Kareni, JHI388 performed as other rhizobial treatments (Figure 2.4 D). The treatment with strain USDA2364 showed a significantly higher biomass increase on pea cv. Corus than any other treatment on this cultivar (Figure 2.4 B). This was not observed on cv. Kareni where USDA2364 had a biomass increase similar to that of the other strains (Figure 2.4 D). The treatment with the atypical JHI388 produced nodules on both cvs Corus and Kareni but these were smaller and generally white (Figure S2.5 A). Strain 21B12 was also capable of inducing nodulation on both pea cultivars but, in this case, only a few nodules were large and pink with the majority being small and white.

On faba bean cv. Fuego the treatment also had a significant effect on the final biomass of the plant (df = 9, F = 2.3856, P < 0.05). However, the rhizobial treatments on this crop showed less of an increase in final biomass dry weight compared to the negative controls and only plants treated with strain 63A21 had a significantly higher biomass than the pots inoculated with an autoclaved culture (NC2) (Figure 2.4 E). Furthermore, the nodulation capacity of some strains with faba bean were reduced or none. Strain 121B21 was not able to produce any nodules on this crop and all root systems of plants with this treatment were free of nodules. In addition, inoculation with the positive control rcr1045 produced some plants with no nodulation, while others had many very small nodules that were not possible to remove from the root system (Figure S2.5 H).

There were significant differences in the number of nodules formed on each root system depending on the inoculated strain in three of the four pea screenings (Table 2.5). The inoculated strain also influenced the proportion of nodule mass of the total belowground biomass (Nmf) in all crops (Table 2.5).

Similarly, the inoculation with rhizobia also produced a significant decrease in the proportion of belowground mass fraction on both pea cultivars (Table 2.5). Negative control treatments had proportionally larger root systems when compared with inoculated plants regardless of the overall number of nodules. However, this was not observed on faba bean where the belowground mass fraction was similar among plants independent of the treatment.

Table 2.5. Treatment effect on nodule number (Nno), nodule mass fraction (Nmf) and belowground mass fraction (Bgmf). The numbers in each column represent the mean value for each variable. Means with different letters in small case in the superscript within one column indicate significant differences between them. Cells with NA indicate that the treatment was not used in that experiment. Means with a ‘1’ in the Nno column indicate the treatments were not included in the negative binomial generalised linear model as no nodules were generated by the treatment in any of the replicates. Column titles with a capital ‘A’ in the superscript indicate that there was no significant difference between the means of the different treatments at P = 0.05.

Strain	**Pea cv Corus 1**			**Pea cv Corus 2**			**Pea cv Kareni 1**			**Pea cv Kareni 2**			**Faba bean cv Fuego**		
	Nno	**Nmf**	**Bgmf**	**Nno**	**Nmf**	**Bgmf**	**Nno**	**Nmf**	**Bgmf**	**Nno**^{A}	**Nmf**	**Bgmf**	**Nno**^{A}	**Nmf**	**Bgmf**^{A}
121B21	86.4^{c}	0.102^{b}	0.348^{ab}	100.4^{cd}	0.094^{bc}	0.353^{a}	129.6^{ab}	0.109^{b}	0.189^{ab}	156.6	0.113^{c}	0.205^{a}	0^{1}	0^{a}	0.391
21B12	37^{a}	0.127^{b}	0.36^{ab}	NA	NA	NA	134^{ab}	0.163^{d}	0.221^{b}	NA	NA	NA	115.2	0.083^{b}	0.396
51A11	71.8^{bc}	0.104^{b}	0.335^{ab}	NA	NA	NA	124.4^{a}	0.117^{bc}	0.18^{a}	NA	NA	NA	146.4	0.11^{c}	0.373
51B21	76^{c}	0.101^{b}	0.336^{ab}	66.2^{ab}	0.115^{cd}	0.341^{a}	107.6^{a}	0.109^{bc}	0.183^{a}	132.4	0.102^{bc}	0.211^{a}	141.4	0.091^{bc}	0.404
63A21	NA	NA	NA	129.4^{d}	0.125^{d}	0.353^{a}	168.8^{ab}	0.122^{bc}	0.195^{ab}	NA	NA	NA	201	0.152^{d}	0.381
73B11	79^{c}	0.115^{b}	0.329^{a}	NA	NA	NA	143.8^{ab}	0.107^{b}	0.18^{a}	NA	NA	NA	217.6	0.085^{bc}	0.384
JHI388	100.6^{c}	0.107^{b}	0.375^{b}	83.2^{bc}	0.088^{b}	0.337^{a}	200.6^{b}	0.137^{cd}	0.259^{c}	143.8	0.103^{c}	0.186^{a}	94.4	0.073^{b}	0.371
NC1	0^{1}	0^{a}	0.504^{c}	NA	NA	NA	0^{1}	0^{a}	0.31^{d}	NA	NA	NA	0^{1}	0^{a}	0.394
NC2	0^{1}	0^{a}	0.514^{c}	0^{1}	0^{a}	0.508^{b}	0^{1}	0^{a}	0.329^{d}	0^{1}	0a	0.335^{b}	0^{1}	0^{a}	0.421
rcr1045	50^{ab}	0.098^{b}	0.33^{a}	56^{a}	0.095^{bc}	0.347^{a}	117.2^{a}	0.107^{b}	0.175^{a}	119.4	0.107^{c}	0.202^{a}	226.6	0.014^{a}	0.39
USDA2364	NA	NA	NA	108.6^{cd}	0.081^{b}	0.347^{a}	NA	NA	NA	144.6	0.085^{b}	0.202^{a}	NA	NA	NA
Variable								**P**							
Treatment	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	0.370	<0.001	<0.001	ns	<0.001	ns

2.3.3 Nodule image analysis

Most of the nodule shape parameters analysed through image analysis were influenced by the treatment (Table 2.6). On pea, the shape parameters of nodules of all treatments were similar to that of nodules on plants treated with the commercial standard strain rcr1045 (Figure 2.5 A, B, D and E). In the first setup of both cv. Corus and cv. Kareni, inoculation with the JHI388 strain variant that showed a slower growth rate than usual showed significantly different nodules (in size and colour) when compared with PC rcr1045-treated plants. Furthermore, the image analysis confirmed the observation that the nodules formed by this strain were significantly smaller than nodules generated by other treatments. On faba bean, inoculation with PC rcr1045 produced erratic nodulation and the resulting nodules were significantly smaller ($P < 0.05$) than those generated by the majority of other treatments (Figure 2.5 C).

Table 2.6. Likelihood Ratio Test results of the generalised linear mixed effects model on shape descriptors nodule length, perimeter, area, circularity, and solidity with strain treatment as a fixed effect and plant as a random effect.

		Corus 1	**Kareni 1**	**Fuego**	**Corus 2**	**Kareni 2**
	df	6	7	6	5	4
Length	χ^2	26.43157	18.14298	16.5976	4.932678	6.874279
	P	<0.001	<0.05	<0.05	0.424	0.143
Perimeter	χ^2	19.9199	11.62418	18.5975	7.990654	4.4677
	P	<0.001	0.114	<0.01	0.157	0.346
Area	χ^2	34.54819	26.00192	24.62826	16.12297	5.937584
	P	<0.001	<0.001	<0.001	<0.01	0.204
Circularity	χ^2	10.62056	14.12717	1.135487	19.57608	1.333046
	P	0.101	<0.05	0.980	<0.01	0.856
Solidity	χ^2	18.92005	15.89595	2.141031	28.83968	0.592999
	P	<0.01	<0.05	0.906	<0.001	0.964

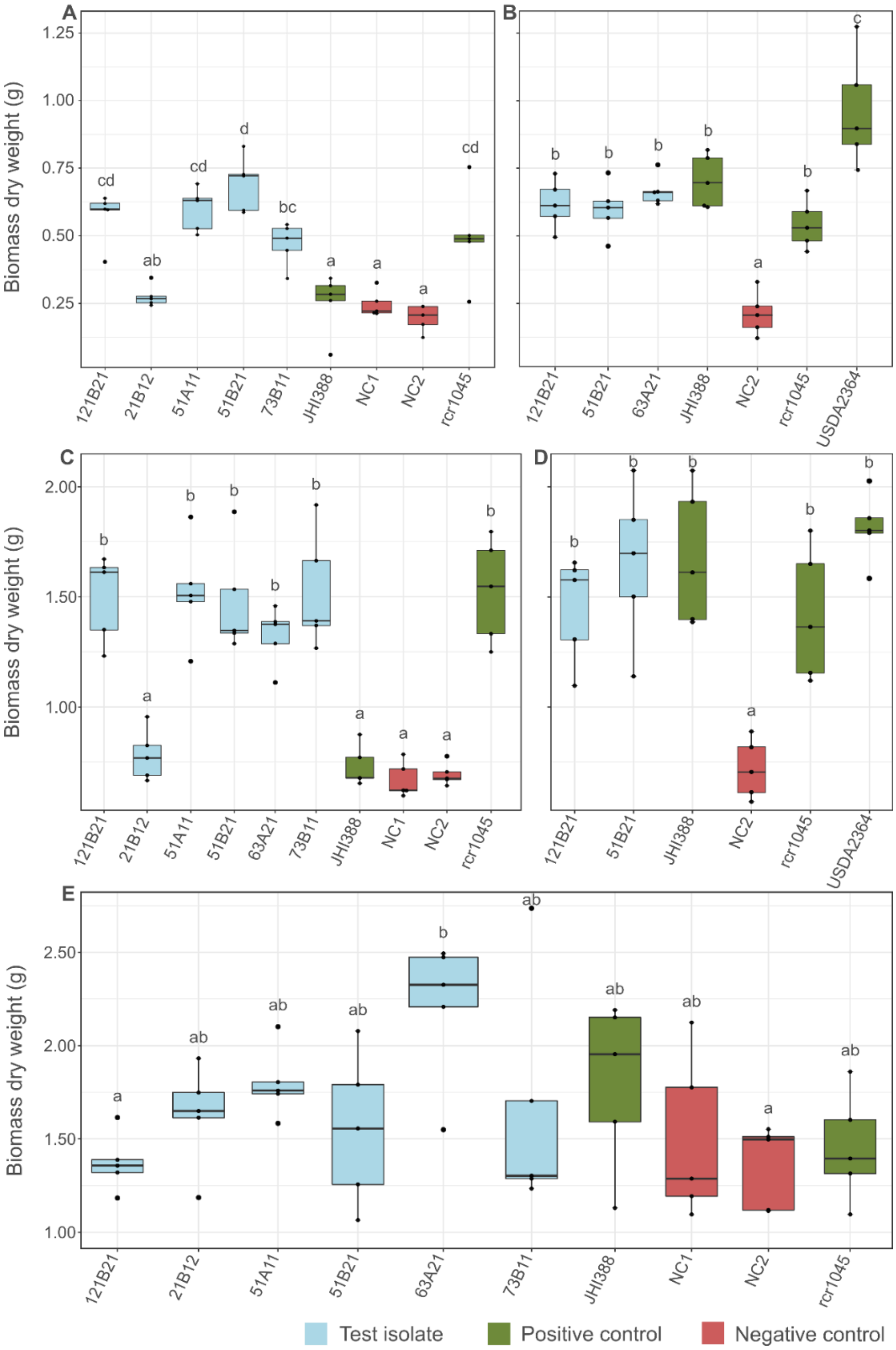


Figure 2.4. Screening experiments on pea cv. Corus (A and B), pea cv. Kareni (C and D) and faba bean cv Fuego (E). Boxes with different letters on top indicate significant differences between treatments estimated by Tukey HSD test at 0.95 confidence interval within the plot. NC1 is a non-inoculated negative control where SDW was used instead of an inoculum and NC2 is an inoculated treatment with an autoclaved grown culture.

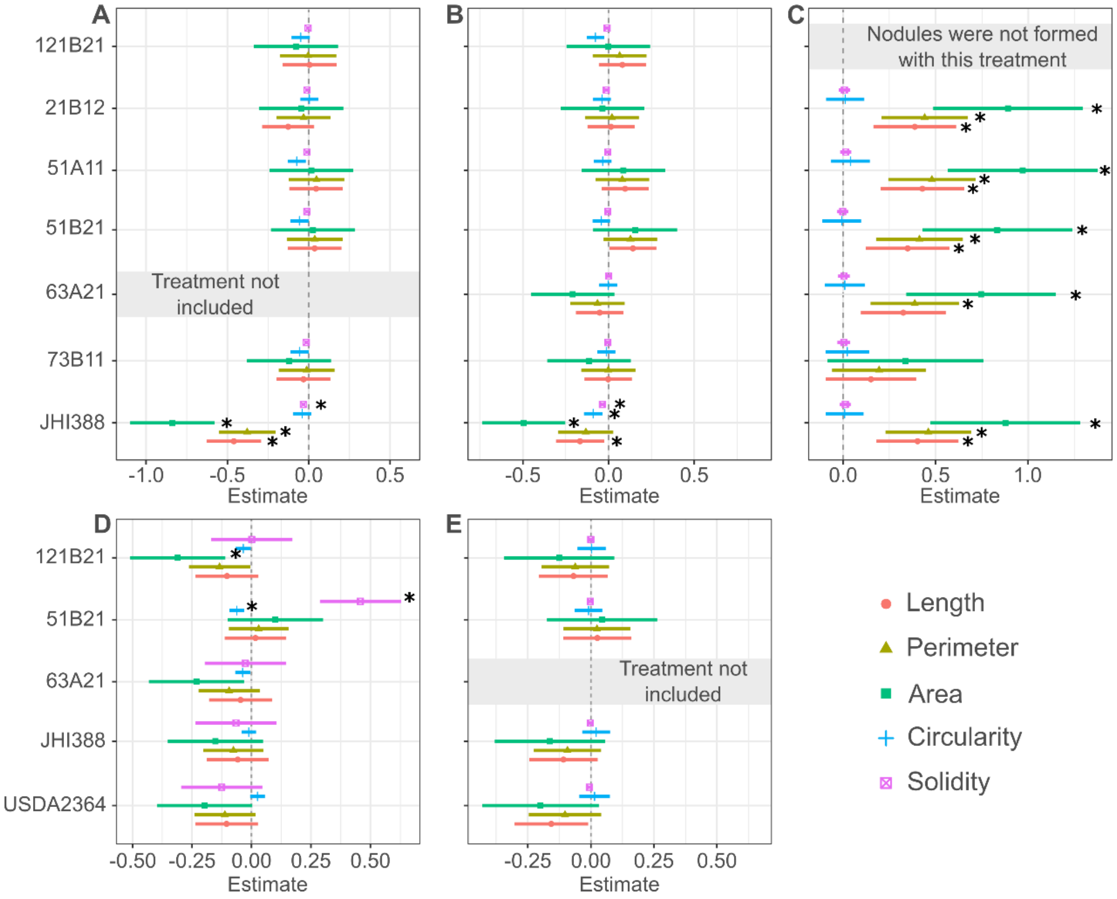


Figure 2.5. Dot-and-whisker plot of nodule image analysis of Phase II screening experiments on pea cv. Corus (A and D), pea cv. Kareni (B and E) and faba bean cv. Fuego (C). Only treatments with nodules were included in the analysis. The dashed line represents the intercept estimate (rcr1045) on the generalised linear mixed effects model, dots represented by the geometric figures are the estimates of each model (intercept estimate - treatment estimate) and whiskers are the confidence interval (CI) of these. CI with a black star at their right indicate a significant difference to treatment rcr1045 on the given parameter at P = 0.05.

2.3.4 Field trials

The final germination percentage (FGP) was affected by both seed treatment (P < 0.001) and pea cultivar (P < 0.001) for both years, although the interaction between treatment and cultivar was only significant in 2019 (P < 0.001). The effect of the treatment on FGP contrasted from the first year to the second (Figure 2.6 A and B). In 2019, the seed treatment only affected FGP of cv. Kareni where seeds inoculated with two of the strains (JHI388 and rcr1045) showed a significant reduction in FGP ($P < 0.05$) compared to the uninoculated seed. In 2020, the effect of inoculation was visible on both cultivars and all treatments showed a significant increase in FGP when compared to the uninoculated control ($P < 0.05$). However, the treatment with different inoculants did not have a significant effect on the time to 50 % germination (T_{50}) and all treatments showed a time similar to that of the uninoculated control (Figure 2.6 C and D). The only significant differences on T_{50} were found between the pea cultivars ($P < 0.001$).

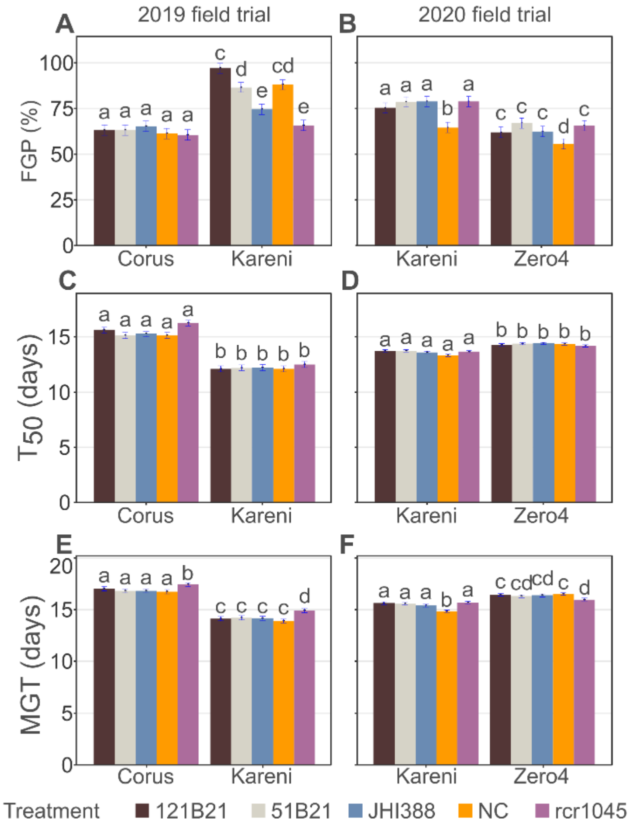


Figure 2.6. Emergence parameters for both years of field trials 2019 (A, C and E) and 2020 (B, D and F). Bars represent the model mean estimates and the error bars represent the standard error for each estimate. Bars with the same letter on top within a plot indicate that there is no statistical significance between the treatments estimates when compared pairwise at $P = 0.05$. FGP - Final germination percentage (A and B), T_{50} - Time to 50% emergence (C and D), MGT - Mean germination time (E and F).

The effect of the treatment on the mean germination time (MGT) was different depending on the pea cultivar and treatment applied to the seed (Figure 2.6 E and F). Kareni showed the fastest MGT for both years, requiring 2.7 days less than cv. Corus ($P < 0.001$) and 0.9 days less than cv. Zero4 ($P < 0.001$). The inoculation only had a significant effect in 2019 ($P < 0.001$) and a pairwise comparison between treatments showed that only treatment with strain rcr1045 produced a significant elongation of the time required for seed to germinate ($P < 0.05$) when compared with any of the other treatments. Moreover, a significant interaction between treatment and cultivar was found in 2020 ($P < 0.001$) where the uninoculated cv. Kareni control took 0.7 days less to emerge than all other inoculated seeds ($P < 0.05$). On cv. Zero4, seeds inoculated with rcr1045 emerged significantly faster than the uninoculated control ($P < 0.05$).

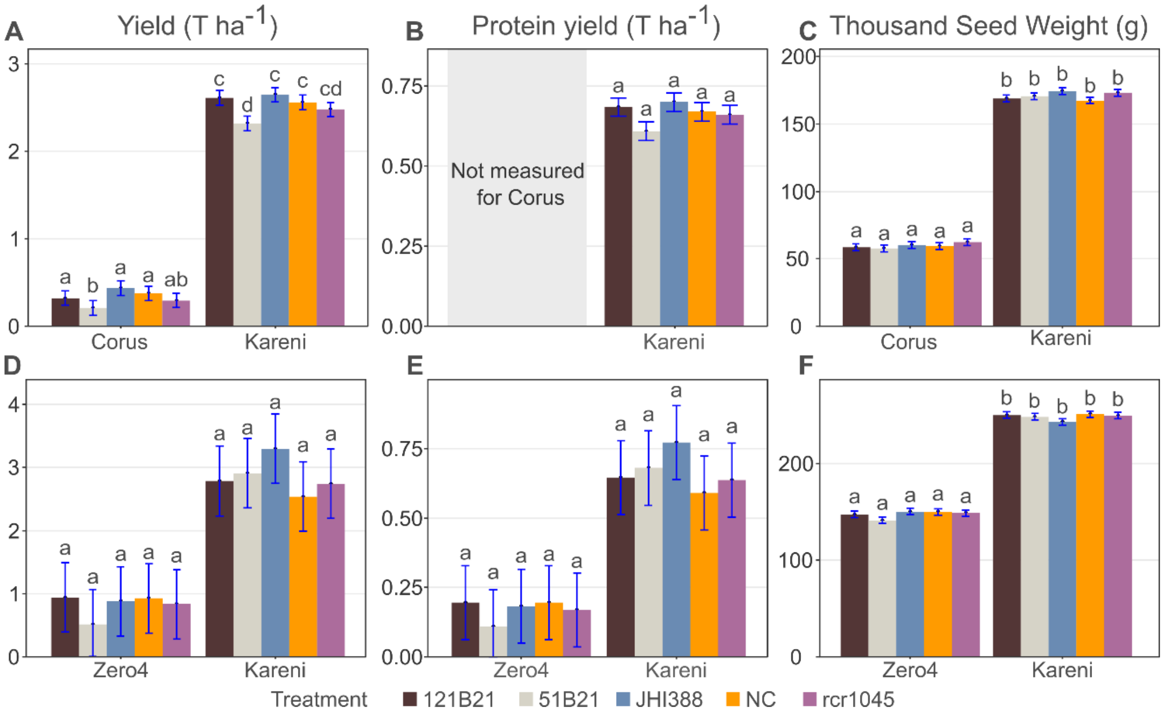


Figure 2.7. Bar plots of harvest data of both years of field trials 2019 (A, B and C) and 2020 (D, E and F). Bars represent the model mean estimates and the error bars represent the standard error for each estimate. Bars with the same letter on top within a plot indicate that there is no statistical significance between the treatment estimates when compared pairwise at $P = 0.05$.

At early- to mid-pod fill the treatments did not influence the shoot dry weight on either of the two cultivars trialled in the second season. Seed treatment influenced the yield of cv. Kareni in 2019 ($P < 0.05$) but this was not repeated the following year (Figure 2.7 A and D). The treatments also affected cv. Corus yield ($P < 0.05$) though most of the yield for this crop was lost due to lodging and not having a suitable harvester available at harvest and therefore the results do not represent the complete yield of this crop. Inoculation did not affect pea protein yield or thousand seed weight (Figure 2.7 B, C, E and F).

2.4 Discussion

As hypothesised, pea grown in pots with soil from Eastern Spain nodulated prolifically with free-living rhizobia from the *Rhizobium leguminosarum* species complex. After successive screenings in a controlled environment using several pea and faba bean cultivars, five strains were identified that had a similar symbiotic performance to current commercial strains. The two best performers were trialled in field conditions where they produced a similar effect to that of the commercial strains tested alongside them. Overall, all five strains identified in this chapter have the potential to become new commercially viable strains to produce commercial inoculants as they perform at least at the same level as currently used commercial strains.

The phylogenetic analysis of 16S rRNA sequences showed that 82 of the isolated strains had a very similar sequence to *Rhizobium* spp. Most of these strains were clustered on a branch with two *R. leguminosarum* biovars and *R. laguerreae*, which indicated that these isolates had a similar 16S rRNA sequence. These two species are known to nodulate both pea and faba bean (Graham, 2008; Saïdi *et al.*, 2014). However, phylogenies based on 16S rRNA alone do not have enough power to separate among those species belonging to the *Rhizobium leguminosarum* species complex or Rlc (Kumar *et al.*, 2015; Young *et al.*, 2021). The Rlc encompasses several Rhizobium species capable of inducing nodulation on pea and faba bean in addition to other *R. leguminosarum* hosts in the legume tribe Viciae. All of these species are subdivided into eighteen genospecies, and a phylogenetic analysis based on the concatenated sequences of the genes *atpD*, *gyrB* and *recA*, was carried out to allow discrimination among them (Young *et al.*, 2021). Thus, with the phylogenetic analysis run in this study it is only possible to state that these strains are likely to be representatives of the Rlc.

The band pattern analysis of the BOX PCR products provided the information for selecting those strains for further screenings. The ten strains isolated from negative control pots during the trapping experiment were clustered into two main groups along with field F3 strains. This suggests that the contamination of the trapping experiment originated from the same source and due to the similarity between the F3 and NC isolates, it is likely the contamination occurred during the experimental setup, e.g. via splashing of slurry from field F3.

The strains tested in this study, including the positive controls, showed a differential response depending on the host, which indicates a specificity or preference between host and symbiont as previously described (Laguerre *et al.*, 2003; Mutch and Young, 2004; Boivin *et al.*, 2020). Thus, while some strains produced very little biomass on peas and average performance on faba bean (i.e. 21B12), others that performed well on peas did not interact well with faba bean with some strains not being able to induce nodulation on this host (i.e. 121B21, rcr1045). Strains formulated in commercial inoculants must not only be capable of nodulating and fixing N effectively, but ideally they also need to be able to do it with as wide a range of hosts as possible (Keyser and Li, 1992; Brockwell *et al.*, 1995; Howieson and Dilworth, 2016).

This has proven to be a challenge as pea and faba bean have different preferences for symbiont genotypes (Boivin *et al.*, 2020), with faba bean being the most selective of the two crops (Laguerre *et al.*, 2003). Therefore, as the results here show, despite both hosts nodulating with similar strains and, in some cases, with the same isolate, finding a rhizobium that performs outstandingly with both crops is challenging and it may be necessary to produce inoculum for each host separately rather than having a universal fit-for-all product. Manufacturing a product for each crop will significantly increase the cost of production of inoculants, thus, an alternative to this is the co-formulation of two or more strains to extend the range of hosts the inoculant can successfully nodulate. This is a strategy that is frequently used, e.g. nodulating strains are combined with other plant growth promoting rhizobacteria (PGPR) or an arbuscular mycorrhiza fungi (AMF) to achieve a synergistic effect (Dileep Kumar *et al.*, 2001; Figueiredo *et al.*, 2008; Gao *et al.*, 2012; Hungria *et al.*, 2013; Ju *et al.*, 2019). However, co-formulation of two nodulating strains may also have detrimental effects and reduce the effectiveness of the inoculant if the strains are not selected carefully (Mendoza-Suárez *et al.*, 2020). The combination of two or more strains that nodulate with pea and faba bean but with a differential degree of effectiveness on both crops may reduce the overall effectiveness of the inoculant when compared with a single-strain formulation because the two strains will compete against each other for nodule occupancy and thus limit host access to the more effective strain (Mendoza-Suárez *et al.*, 2020). This can be aggravated if the Nod group has not been taken into account which could facilitate the nodulation of less effective strain which is better at nodulating a particular crop genotype (Boivin *et al.*, 2021). Nevertheless, recent research in Rleg consortia inoculants, where more than two strains are formulated together, has shown that the presence of a high performing strain, either in terms of plant biomass or nodule number or weight, is enough to improve the performance of an inoculant (Fields *et al.*, 2021).

The effect of inoculation was less visible on faba bean and only one treatment produced a significant biomass increase, but only when compared with the autoclaved culture negative control (NC2). These strains were initially isolated from pea (which might have biased the trapping of strains towards pea-preferred genotypes), and therefore had a lower effect on faba bean, which is a more selective crop (Laguerre *et al.*, 2003). In addition, the seed of faba bean is larger than pea and it is possible that the greater amount of nutrients stored in faba bean seeds may have helped boost biomass production from N stored in the seed and thus reduced the demand for N supplied by BNF. The duration of the screening process could have been extended to last longer than the 33 days from sowing to exhaust the seed reserves and increase reliance on BNF in this N-free medium. However, within this time, the plants did reach physiological maturity and most of the NC and 121B21-inoculated plants were chlorotic indicating that the reserves of the seeds had already been exhausted (Figure S2.5 E).

The total number of nodules generated on pea was different depending on the strain used as inoculant. The number of root nodules is controlled by the legume host and is dependent on the capacity of nodules to sustain plant growth (Smit and Bisseling, 2008). Thus, this measurement may also be used as a proxy of effectiveness on N-fixation as nodule number is positively correlated to root and shoot weight (Sinclair *et al.*, 1991). However, the screening experiments indicated that nodule number was not as reliable as the increase in total plant biomass as an indirect estimation of N-fixation, e.g. plants inoculated with the atypical JHI388 variant (which had the largest number of nodules but low total biomass), or plants inoculated with 21B12 (which had the lowest nodule number without substantial biomass increase when compared with the negative controls).

In pea, the autoregulation of nodulation is triggered by the start of the nodulation process in the root and the strength of inhibition is correlated with the nodule developmental stage and the size of the nodule where large and active nodules produce a stronger inhibition on other nodules to form (Li *et al.*, 2009). The nodules of plants inoculated with the atypical JHI388 variant strain were generally small and white, indicating low N-fixation activity (Vikman and Vessey, 1993). Therefore, if the nodules formed by this strain were incapable of fixing nitrogen, the plant did not divert sugars to them and allow the nodule to grow and fix N into ammonia which inhibits nodulation, thus each of these nodules produced a low inhibition of nodulation which ultimately increased the overall number of nodules.

The nodule image analysis revealed that nodule size was reduced compared to the commercial standard strain only when the strain was incompatible with the host (e.g., rcr1045 on faba bean), or had an abnormal growth (i.e., JHI388 on the first setup with cvs Corus or Kareni). Despite this, the morphological parameters circularity and solidity were generally similar between all treatments indicating that regardless of the efficiency of nodulation and N-fixation nodule morphology was not altered. Nodule size and colour have been previously correlated to nitrogenase activity and nodule respiration (Vikman and Vessey, 1993); however, the data presented in this chapter suggest that nodule morphology is not dependent on the functionality of the rhizobia *per se* but rather the result of the interaction between the host plant and the rhizobia, at least at early stages of nodule ontogenesis.

The belowground mass fraction was significantly higher in all negative control treatments on peas, which is likely to be due to the lack of N in the growth media. Pea plants in the screening experiments were limited to the N reserves stored in their seed unless they nodulated with a strain of rhizobia. Thus, the host plant needs to invest more resources to search for N in the soil by reducing N allocation to leaves and photosynthesis. This re-allocation is evidenced by the increased degree of chlorosis of pea seedlings following emergence. The increased Bgmf was not observed on faba bean, which may be driven by the larger quantities of N reserves in its larger seed.

The inoculation of peas before sowing with the two isolated strains significantly improved emergence of the crop in the field when compared to uninoculated control in a dry year, and in a wetter year, they enhanced emergence when compared with the positive control treatments. However, none of the seed treatments influenced the yield of the crop. The first weeks from sowing up to final emergence in the first year were wetter than the same period in the second year. This seasonal variation affected the emergence of pea seeds which in the second year required more time to emerge. The drier weather during crop establishment also influenced the final emergence, which was reduced in the second year. Inoculation with rhizobia assisted seedling emergence and boosted it to similar percentages as the year before.

Previous research has found that inoculation with strains of *R. leguminosarum* can improve seed germination on faba bean and common bean in field and glasshouse experiments (Kumar et al., 2016; Senberga *et al.*, 2018); although the effect on final emergence varied depending on the inoculated strain. In the drier year, all strains improved emergence to a similar degree, but in the wetter year the two strains had better emergence compared to the positive controls. This may seem contradictory given that the two positive control strains were isolated in the UK (Dye, 1978; Maluk *et al.*, 2022) and would be expected to have performed better in a wet year compared to strains isolated from a semi-arid environment, but the increase in emergence was not translated into a significant increase in shoot biomass nor yield. Therefore, the two strains tested in field conditions have similar performance to that of the positive controls and hence are strong candidates to become commercially used strains.

In conclusion, the soils in Eastern Spain can support many pea-compatible rhizobia strains, some of which have shown symbiotic efficiency comparable to that of current commercial strains. The combination of multiple strains with different symbiotic capacities may provide a solution to single elite strains with a narrow host range but the specific strains to be combined must be examined in detail to avoid efficiency losses due to competition among strains (Mendoza-Suárez *et al.*, 2020). The environment from which these strains were isolated may have selected traits facilitating a level of desiccation tolerance which is greater than provided by some current commercial inoculants, specifically, those which originated from wetter environments such as the British Isles. Thus future research in this area should investigate whether this set of strains show enhanced tolerance to this environmental stress.

Chapter 3 | Assessment of desiccation tolerance in two populations of *Rhizobium leguminosarum*

Abstract

The use of rhizobial inoculants for cultivating legumes has become an established practice in many countries. Inoculants are the combination of one or more effective strains, and a liquid or solid ‘carrying medium’ or ‘carrier’. However, these products are characterised by a short shelf-life due to the rapid die-off of cells caused by desiccation stress. Rhizobia isolated from areas with high water deficit have been shown to possess superior tolerance to desiccation, and using such strains is likely to extend the shelf-life of commercial inoculants. Therefore, the aim of this study was to compare the desiccation tolerance of rhizobia strains isolated from a temperate environment with relatively high levels of precipitation (the UK), with those from a semi-arid (low humidity) and hot environment (Spain). A total of 108 strains of *Rhizobium leguminosarum* were screened for desiccation tolerance using a high-throughput method from which a “desiccation factor” was derived for each strain by dividing the area under the growth curve (AUC) of the stressed culture by that of the non-stressed culture. All strains survived the desiccation stress, and strains isolated from Spain had a higher tolerance to desiccation, reaching a desiccation factor of 51 %. Although desiccation affected the growth rate of strains from each location, those isolated from the semi-arid environment showed improved desiccation tolerance. This indicates that agricultural soils of semi-arid environments may serve as a suitable source of rhizobia strains with increased propensity to tolerate desiccation, and if utilised in commercial inoculants may maintain viable population densities, and so extend the shelf-life and/or effectiveness of the product.

Keywords

Rhizobia, desiccation tolerance, anhydrobiosis, inoculant, shelf-life

3.1 Introduction

The symbiotic relationship between legume crops, such as peas (*Pisum sativum* L.) and faba beans (*Vicia faba* L.), with root nodule bacteria (rhizobia) can fix between 200 and 300 Kg ha^{-1} of atmospheric N, and once the crop has been harvested up to 90 Kg of this fixed N is left in the soil for the next crop (Sessitsch *et al.*, 2002; Maluk *et al.*, 2022). Since the discovery of this interaction in the 19th century the use of compatible rhizobia for growing legume crops in the form of inoculants has become an established practice in many countries (Catroux *et al.*, 2001; Herridge, 2008).

Legume inoculants combine a carrier with one or more so-called ‘elite’ strains of rhizobia that have a known effective symbiosis with the target legume crop (Bashan *et al.*, 2014). The carrier is a liquid or solid medium comprising a mixture of bacterial nutrients, protectants and other substances with a variety of aims, e.g., facilitating adhesion to the seed surface or enhancing survival of rhizobia during storage (Bashan *et al.*, 2014). The most commonly used carrier is peat which has the capacity to support the growth of rhizobia and regulate the pH of the medium, while peat liquid extract is also capable of protecting rhizobia from some abiotic stresses (Deaker *et al.*, 2011; Casteriano *et al.*, 2013; Mahmood *et al.*, 2016). At the formulation stage, a liquid culture of rhizobia is mixed with the carrier to manufacture the inoculant and it is at this stage where most of the candidate, symbiotically-efficient, rhizobial strains die (Bashan *et al.*, 2014).

The quality of a commercial inoculant is measured by the extent to which the rhizobia population density (number of viable cells per gram or litre of carrier) is maintained between the time of packaging (or formulation), and after storage time, the time of field- or seed-application (Howieson and Dilworth, 2016). On-seed shelf-life may also be considered as an important quality attribute. Following the formulation stage, inoculants start losing cell viability (Date, 2001) mainly due to desiccation stress (Deaker *et al.*, 2012; Casteriano *et al.*, 2013), which ultimately affects the shelf-life of the product. Inoculant manufacturers desire high viable rhizobia population densities, and which may be maintained for as long as possible to guard against losses during distribution, storage, and application (Bashan *et al.*, 2014). Furthermore, the application of the inoculant to seed or soil before or at sowing will impose further (desiccation) stress on the formulated rhizobia, reducing even further the number of viable bacteria. Thus, desiccation tolerance is a desirable trait of candidate strains for the creation of new commercial inoculants with improved shelf-live (Bashan *et al.*, 2014; Molina-Romero *et al.*, 2017).

Rhizobium leguminosarum (Rleg) is a Gram-negative symbiotic rhizobia of peas and faba beans (Howieson and Dilworth, 2016) and like other Gram-negative bacteria it has an external lipid membrane above the thin peptidoglycan membrane that makes it incapable of generating spores, and hence desiccation stress is often lethal for these rhizobia (Berninger *et al.*, 2018). During desiccation, cellular damage is caused by oxidative stress, phase transition after rehydration and browning reactions which are melanoid formations under the depletion of water between carbonyl groups of saccharides and the amino groups of proteins and nucleic acids (García, 2011). Rhizobia can protect themselves from such damage by the accumulation of trehalose, which reduces the amount of water loss from the cytoplasm, and/or the synthesis of proteins that protect and repair the cell envelope, and the production of exopolysaccharides (Cytryn *et al.*, 2007; McIntyre *et al.*, 2007; Donati *et al.*, 2011). This indicates that different strains may also have different adaptive capacities to desiccation. The screening of strains for improved shelf-life potential is, therefore, necessary alongside that of symbiotic performance. A study carried out in New Zealand showed that Rleg strains isolated from areas with higher annual soil moisture deficit were naturally more tolerant to desiccation than strains isolated from areas with less soil moisture deficit (van Ham *et al.*, 2016). Hence, if strains isolated from different locations of the same country have shown a differential response to desiccation stress, it is likely that isolating rhizobia from dry and/or hot environments may yield more desiccation-tolerant strains than more temperate and/or humid environments.

The main aim of this study was to assess the desiccation tolerance of rhizobia strains isolated from contrasting climatic and biogeographical regions, and to compare their relative desiccation tolerance. For this, strains isolated from the UK, a country characterised by a temperate climate of cool and wet winters, plus warm and wet summers were compared with strains isolated from eastern Spain, characterised by a semi-arid Mediterranean climate with mild winters and hot, dry summers. The objective of the study was to test the hypothesis that strains isolated from Spain are naturally better at tolerating desiccation when exposed to 24 h of this stress than strains from the UK.

3.2 Materials and methods

3.2.1 Rhizobial strains used and culturing

The rhizobial strains used in this study were either isolated from eastern Spain (described in Chapter 2) or were part of the collection from the James Hutton Institute, which included strains from a range of geographic regions; all strains were stored in glycerol at -80 °C. Each strain was streaked onto Petri dishes with YMA-CR medium and incubated at 28 °C for 48 h. A single colony from each Petri dish was picked off and inoculated into a 30 mL tube with 5 mL of TY (pH 6.8) medium and incubated at 30 °C for 48 h at 150 rpm. After incubation, culture concentrations were adjusted to 3 x 10^9 CFU mL^{-1} by regression to a standard curve generated using serial dilutions of a culture of the reference strain *R. leguminosarum* bv. *viciae* 3841 (Young *et al.*, 2006) of known concentration and absorbance at 600 nm (Multiskan GO; Thermo Fisher Scientific, USA).

3.2.2 Assessment of short- and long-term desiccation tolerance

Prior to running assays, an assessment was made of the evaporation that each well of a transparent 96-well plate experiences during incubation. For this, all wells of a 96-well plate were loaded with 200 µL of sterile distilled water, covered with a lid and incubated in a Multiskan GO at 30 °C for 47 h with 3 minutes of shaking at 20 Hz every 15 minutes. An absorbance reading at 975 nm was made every hour. The volume of each well was calculated at every timepoint by regression to a standard curve of known volumes of sterile distilled water and their absorbances at 975 nm. All the outermost wells (A1-12, H1-12, B1 and 12, C1 and 12, D1 and 12, E1 and 12, F1 and 12 and G1 and 12) were considered to be unusable as they lost an average of 14.6 % of their volume in contrast to just 0.19 % of volume loss for all the inner wells (Figure S3.1).

The assessment of desiccation tolerance was performed with a modification of the high-throughput methods described by Slininger and Schisler (2013) and van Ham *et al.* (2016). Six strains were assessed in each desiccation tolerance assay.

All 96-well plates and lids used in the desiccation tolerance assays were sterilised for 20 minutes using UV irradiation.

In the initial screening, strains were exposed to desiccation stress for a short period. On the first day, 10 µL of each 3 x 10^9 CFU mL^{-1} culture were added in five replicates to the wells in the 96-well plate. The plate (without its lid) was then placed in a sterile glass desiccation chamber with pre-dried silica gel at the bottom and incubated at 20 °C for 24 h. The remaining cultures were kept in closed tubes and incubated under the same conditions but outside the desiccation chamber. After 24 h, 10 µL of the un-stressed cultures were aliquoted into the remaining wells in the 96-well plate. The wells with dried rhizobia were filled with 200 µL of sterile TY medium while 190 µL of sterile TY was used for wells with un-stressed cultures. The outer wells were filled with 200 µL of sterile distilled water, apart from four wells which contained 200 µL of sterile TY to act as blanks. The plate was covered with its lid and then shaken at 20 Hz at room temperature for 15 min in order to re-suspend dry rhizobia and to mix thoroughly the contents in each well. Following this, the plate was inserted into a Multiskan GO and incubated for 47 h at 30 °C with three minutes shaking at 20 Hz every 15 min; during this an absorbance reading at 600 nm was taken every h. After each assay, a droplet (1-2 µL) of each well was inoculated onto a Petri dish with YMA-CR medium and incubated for 48 h at 28 °C to visually assess possible contamination and the presence of living rhizobia cells in those wells where there was no significant increment of absorbance from the background level (i.e. the culture did not reach log phase). The growth rate was assessed for all the non-stressed cultures, and two different groups were delimited depending on whether the growth rate was higher (fast growing) or lower (slow growing) than the median growth rate of the dataset.

The desiccation tolerance was assessed by quantifying and comparing the desiccation factor (DF) of stressed and non-stress strains as calculated using Equation 3.1:

$$DF = \frac{Area\ under\ the\ curve\ stressed\ culture}{Area\ under\ the\ curve\ unstressed\ culture} \qquad \text{Eq. 3.1}$$

Based on their symbiotic interaction with host legume plants and on their short-term desiccation tolerance, eight strains were selected for assessment of their tolerance to a longer period of desiccation stress. The eight strains were randomly aliquoted into seven different plates. Six of the plates were stored in a custom sterile drying chamber (Figure S3.2) made from a hermetically sealed plastic storage box. Inside the box the plates were placed on a grid above the silica gel in the base, mimicking the setup of a conventional drying chamber. The plates were stored without their lids and in the same plane to ensure uniform drying of all wells of each plate. The box was then stored in an incubator at 20 °C. After 24 h, all the wells were completely dry, and a lid was placed onto each plate so as to avoid contamination when opening the chamber at each timepoint. After this, the plates with their respective lids were stacked, one on top of the lid of the other, and stored in a sterile glass desiccation chamber with freshly dried silica gel at 20 °C. At every timepoint the plates were stacked in a different order and the silica gel was replaced with freshly dried silica gel. All wells with cultures in the remaining plate were filled with 240 µL of sterile TY making a final volume of 250 µL. The plate was then shaken at room temperature at 20 Hz for 15 minutes. After shaking, 50 µL of each of the five wells with the same strain were pooled in a sterile 1.5 mL plastic capped tube and these pooled cultures were used to determine the concentration of rhizobia. The plate was then inserted into the Multiskan GO for the same time and the same settings as the plates used in the short-term assay, with the addition of a photometric step at 15 min. after the start of the incubation.

The concentration of rhizobia was assessed using a 1:10 serial dilution method by plating 50 µL of them onto Petri dishes with YMA-CR and incubating them at 28 °C until visible colonies appeared. A single colony from each plate was then used to inoculate 5 mL TY medium in a 30 mL screw-cap tube. Cultures were grown at 30 °C and 2.5 Hz for 48 h. One mL of each of the cultures in the log phase was mixed with 1 mL of 50 % glycerol, frozen in liquid nitrogen and stored at -80 °C. The remaining culture was used for DNA extractions using the method described in Section 2.2.3.2.

The plates in the drying chamber were stored for up to 133 d under the drying conditions. At 2, 7, 14, 21, 28 and 133 d after initiation of the assay one of the plates was used to determine the desiccation tolerance using the method described above with each well with dry rhizobia being loaded with 250 µL of sterile TY before shaking at room temperature.

3.2.3 Bacterial genomic DNA isolation

For DNA extraction, 4 mL of each log phase cell culture was pelleted by centrifugation at 1900 RCF for 15 min. The pellet was resuspended and lysed in a lysis buffer (10 mM Tris-HCl, 1 mM EDTA, 0.5 % SDS and 19.05 units mL^{-1} Proteinase K (Merck Millipore, Germany)) for 1 h at 37 °C followed by the addition of an equal volume of phenol:chloroform:isoamilalcohol and mixed well by vortexing. The top layer of the resulting solution containing DNA was recovered after centrifugation at 13,200 RCF for 12 min. The DNA was precipitated using 0.078 M sodium acetate in 100 % propanol and incubated overnight at -20 °C. Next, the DNA was pelleted by centrifugation at 13,200 RCF for 12 min. and washed in 70 % ethanol. Ethanol was removed after centrifugation at 13,200 RCF for 2 min. The DNA pellet was dried at 28 °C for 30 min. and then resuspended in 50 µL sterile distilled water. The quality and quantity of the DNA was analysed using a NanoDrop ND-1000 (NanoDrop Technologies, Inc., Wilmington, USA).

3.2.4 BOX PCR analysis

Polymerase chain reaction using the BOX motif (BOX PCR) was used to confirm the identity of the strains grown at each timepoint of the long-term desiccation assay. The Go Taq ® G2 DNA Polymerase (Promega, USA) kit was used according to the manufacturers recommendations for preparing the master mix with 0.6 µL at 10 mM of each dNTP (Invitrogen, USA), 2.4 µL of primer BOXA1R (Versalovic *et al.*, 1994) at 10 µM and with 25 ng of the extracted DNA in a final reaction volume of 28 µL. The PCR product was then loaded onto a 2 % [w/v] agarose gel (8 x 8 cm) in 1x TBE with SYBR Safe (Invitrogen) following the manufacturers recommendations and run in 1x TBE buffer at 50 V and 400 A for 1 h with a 1 kb ladder (Promega). The gels were analysed and photographed under UV light in a UVP BioDoc-It^2 Imager from Analytikjena. The banding patterns for each strain were compared to confirm the identity of each strain.

3.2.5 Data processing and statistical analyses

R software v 4.0.3 (R Core Team, 2019) was used for all statistical analyses. Package dplyr v 1.0.2 (Wickham *et al.*, 2019) was used for processing and organising data for analysis. Raw absorbance data were analysed with package grofit v 1.1.1-1 (Kahm *et al.*, 2010) for calculating growth curve parameters by spline regression and 0.55 smoothing factor. Growth parameters λ (lag time), µ (growth rate), A (maximum cell growth) and area under the curve (AUC) were extracted from each replicate (Figure S3.3). Generalised linear models with mixed effects (GLMM) and linear mixed effect models (LMM) were run with package lme4 v 1.1-26 (Bates *et al.*, 2015) and generalised linear models (GLM) with R base v 4.0.3 (R Core Team, 2019). Principal component analyses were carried out with R base, and visualised using packages factoextra v 1.0.7 (Kassambara and Mundt, 2020).

Normal distribution was assessed by visual inspection of the histogram and QQ-plot of the dataset. Data normalisation was carried out by arcsin of the square root for desiccation factor and when normalisation was not possible, generalised models were used. GLMMs and LMMs models had the area of the plate where the replicate was located as a random effect to account for any variability depending on the position in the 96-well plate. For the statistical analysis of λ of the different cultures, rounded λ values to the closest minute were run in a GLMM with Poisson family and log link formula. The software package 'grofit' v1.1.1-1 was used to calculate the half maximal exposure to desiccation from a 1000 bootstraps run using the raw data from the long-term desiccation assay. Half maximal exposure was calculated for each repetition and a generalised linear model with Poisson family, with a log link formula to assess the strain fixed effect, and a post-hoc Tukey HSD test to compare pairwise between strains.

3.3 Results

A total of 108 *Rhizobium leguminosarum* strains isolated from Spanish (51) and UK (57) soils were screened for desiccation tolerance (Figure S3.4). The main result of this study supports the initial hypothesis that strains from soils with a greater intrinsic water deficit will tolerate desiccation better. Thus, strains isolated from Spanish soils showed a greater tolerance to desiccation than those isolated from UK soils. However, all strains survived exposure to a desiccation stress of 24 h, with 99.34 % reaching log phase within 47 h of incubation in TY growth medium (only one UK strain, JHI761, was unable to reach log phase in all replicates).

Strains isolated from Spain generally grew slower than those from the UK ($P < 0.001$) and required an additional 22.8 minutes to reach log phase regardless of the speed at which they grew (Figure 3.1A). Moreover, the rate of growth was not a unique characteristic for populations from either country as both had strains with each type of growth behaviour (Figure 3.1B). However, the populations from the UK had a higher number of fast-growing Rlv, whilst those from Spain had more slow-growing strains. Both slow- and fast-growing strains of rhizobia showed a shorter lag phase when isolated from the UK when compared with the same types isolated from Spain. Furthermore, there was a significant interaction between the country of origin and the growth rate type of the strains ($P < 0.001$). Comparing the lag phases of these two growth types between the two countries showed that differences in lag phases were greater between faster- than slower-growing Rlv types.

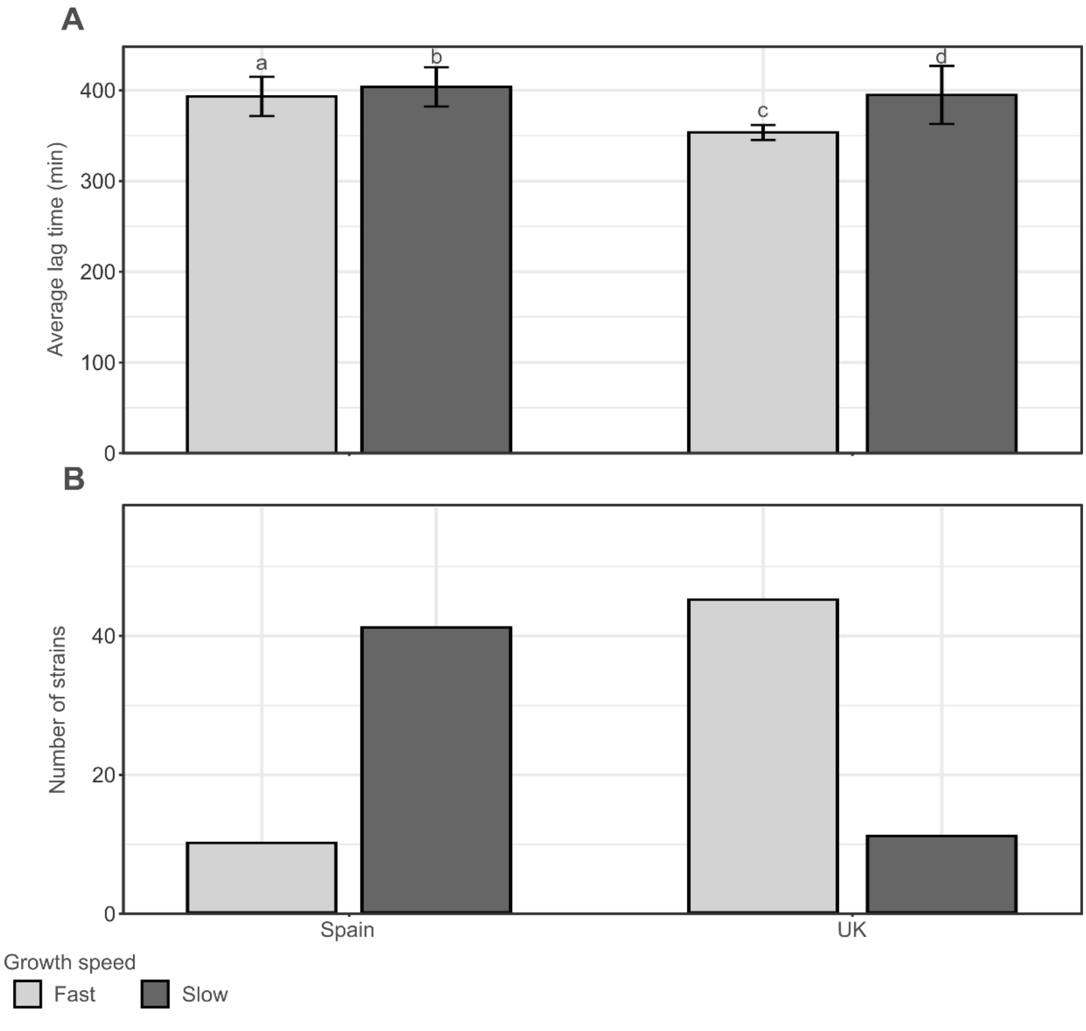


Figure 3.1. Lag time (A) and the strain growth speed composition from either Spain or the UK (B). The error bars in A represent the 95 % confidence interval and bars with different letters indicate significant differences between pairwise comparison of means by Tukey HSD ($P < 0.05$).

A principal component analysis of the growth parameters showed that the first two components explained 97.9 % of the observed variation and that the populations from both countries were very similar as indicated by the large area of overlap between both ellipses delimiting the variation within populations (Figure 3.2A). However, the speed of growth better characterised the population as seen by the smaller overlap between the slow and fast-growing ellipses (Figure 3.2B). Moreover, the PCA revealed a high positive correlation among growth rate, maximum cell growth and Area Under the Curve (AUC), as all three vectors were positioned in the same direction and with a similar elongation.

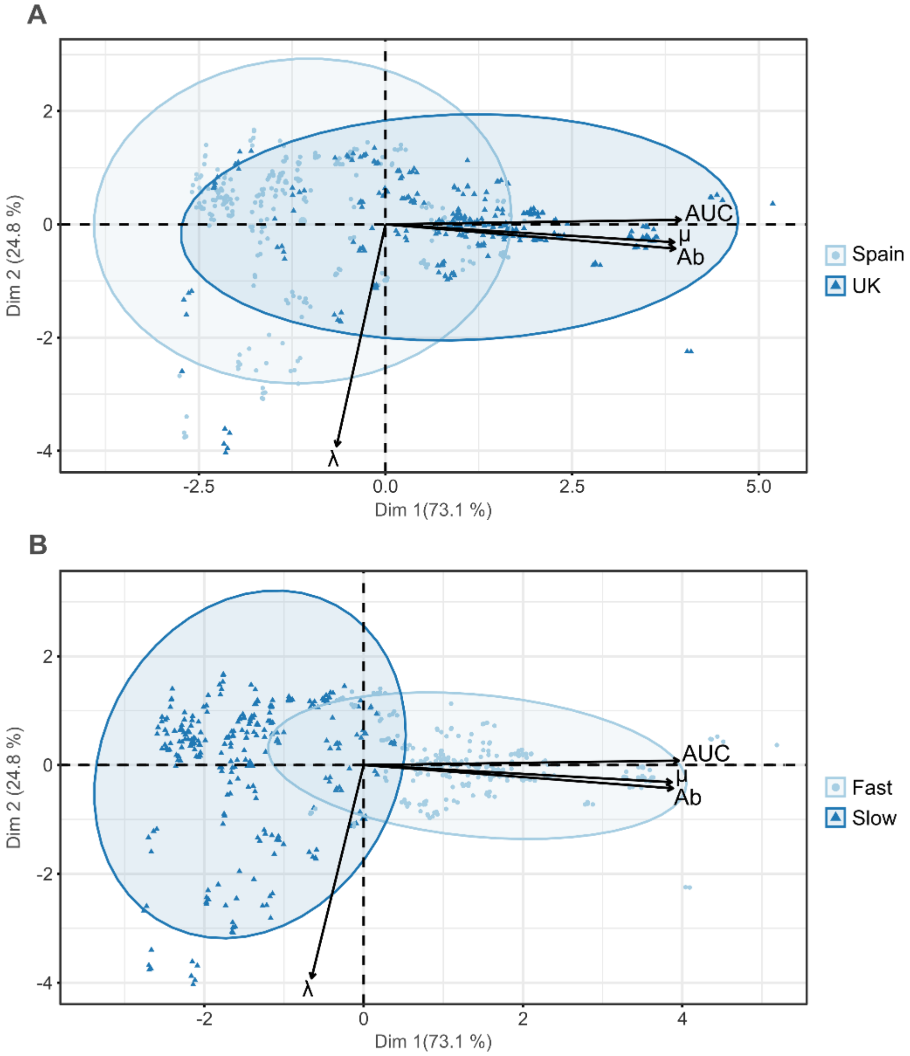


Figure 3.2. PCA plots of non-desiccation-stressed cultures categorised according to country (A), rate of growth (B), Lag time (λ), growth rate (μ), maximum cell growth (Ab), and area under the curve (AUC). Ellipses represent the area containing 95 % of all samples of each group.

Strains isolated from Spain showed a higher desiccation factor, indicating their greater tolerance to desiccation than strains from the UK (P < 0.05) (Figure 3.3). There were also significant differences (P < 0.001) between fast- and slow-growing rhizobia after 24 h of complete drying where slow-growing Rlv strains showed a lower desiccation factor. In pairwise comparisons of both fast- and slow-growing Rlv between both countries, strains from Spain showed a superior tolerance to desiccation (P < 0.05) in both cases. Fast-growing strains isolated from Spain achieved a 51.1 % desiccation factor which was a 1.55-fold higher tolerance than slow-growing UK-isolated strains which had the lowest desiccation factor.

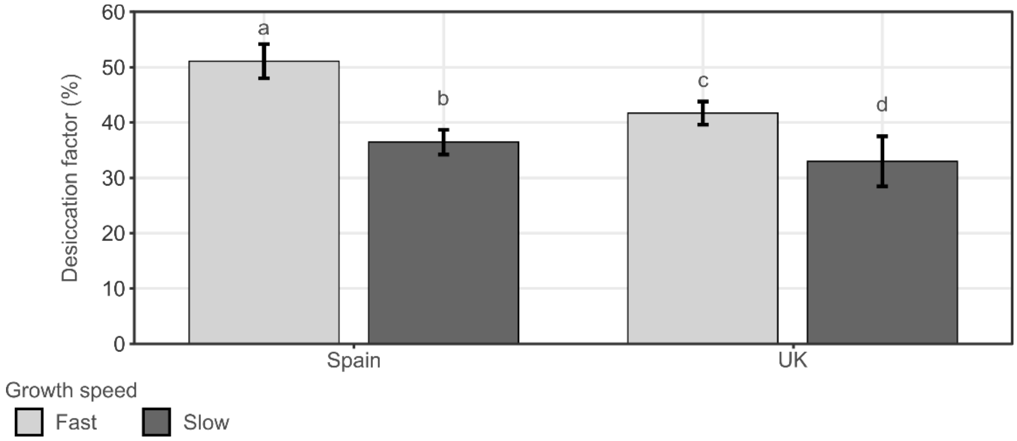


Figure 3.3. Average desiccation factors of fast- and slow-growing rhizobia strains isolated from either Spain or the UK. The error bars represent the 95 % coefficient interval, and different letters above them indicates a significant difference between groups when compared pairwise (P < 0.05).

The conditions under which desiccation tolerance was assessed in this study involved the complete evaporation of water from the media, which may have affected the biological functions of the surviving rhizobia. Therefore, to assess whether the strains were affected, the growth rate of the cultures before and after the desiccation event was compared (Figure 3.4). It was found that water deprivation (i.e. complete drying) did, indeed, affect growth rates (P < 0.001). Moreover, independent of the country of origin and the speed of growth of the bacteria, all strains were affected in equal measure by complete desiccation such that no significant differences were found among groups.

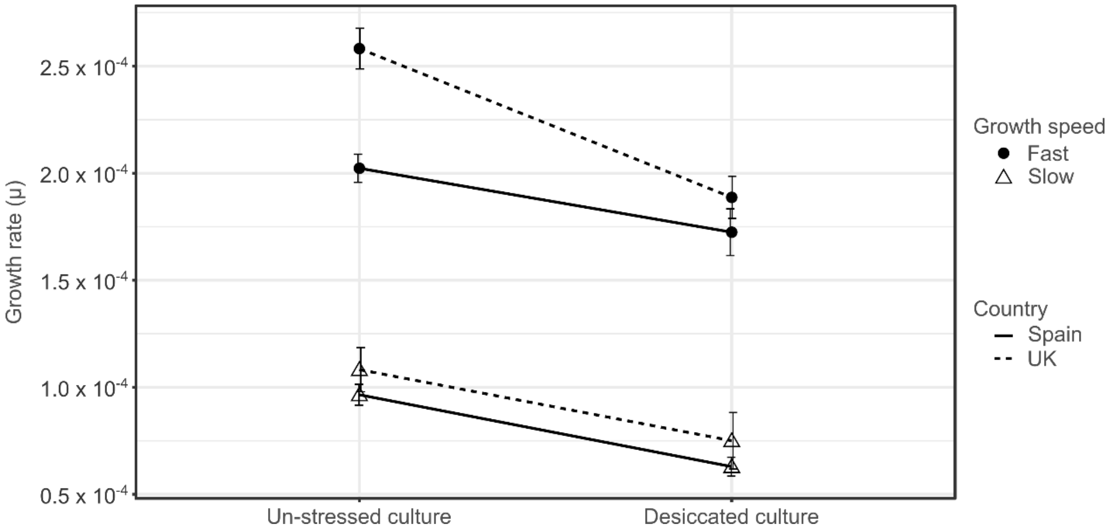


Figure 3.4. Effect of drying for 24 h on fast- and slow-growing rhizobia strains isolated from Spain or the UK. The error bars on each datapoint represent the 95 % coefficient interval.

Long-term exposure to desiccation revealed that the strains most tolerant to the 24 h exposure were also the most successful at tolerating desiccation for extended periods (Figure S3.5). The desiccation factor was strongly correlated with the number of surviving rhizobia (rho = 0.91, $P < 0.001$). All strains survived the longest exposure to desiccation (133 days) though in some cases they did not survive in sufficient numbers to reach log phase during the 47 h incubation period (strains JHI388 and 63A21). Or, they only achieved log phase in some of the replicates (strain rcr1045). The best performing strain, JHI1118, reached the end of the experiment with 0.33 % of the initial bacteria surviving. However, the strain with the highest percentage remaining after 133 days of desiccation exposure was 21B12 with 1.72 % of the initial rhizobia still viable. The time needed to decrease the desiccation factor below 50 % was strongly determined by the strain ($P < 0.001$); strain JHI1118 was able to maintain the desiccation factor above 50 % for 71.9 days, which was 60 days longer than the second-best performing strain (121B21) (Figure 3.5).

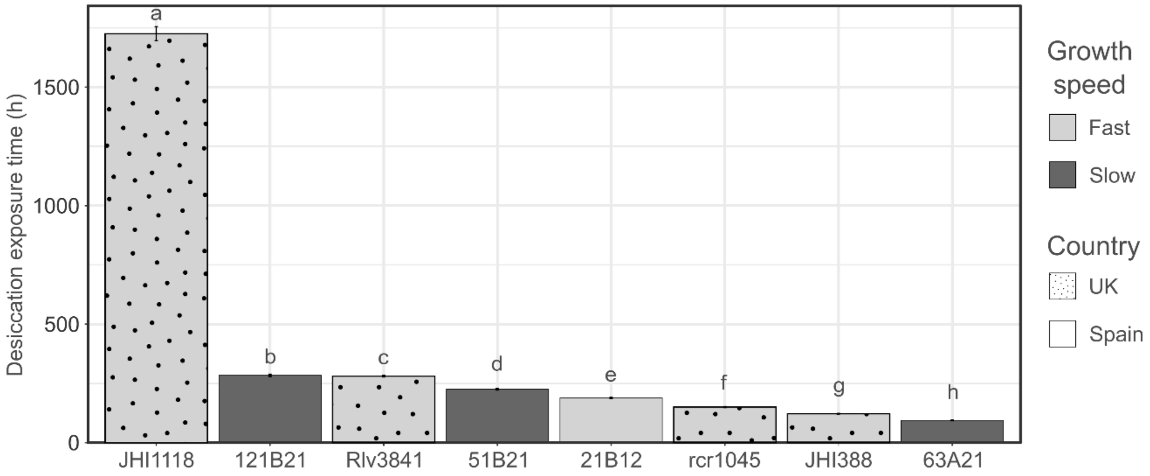


Figure 3.5. Half maximal exposure time to desiccation. The error bars represent the 95 % coefficient interval and different letters indicate a significant difference between strains when compared pairwise (P < 0.05).

3.4 Discussion

This study supports the hypothesis that rhizobial strains isolated from a semi-arid environment, such as the one found in eastern Spain, are generally better at withstanding desiccation stress than strains isolated from wetter environments like those isolated from UK soils that experience fewer periods of water scarcity. *Rhizobium* is classed within the rhizobia paraphyletic group as a fast-growing genus relative to other rhizobia such as *Bradyrhizobium* (Howieson and Dilworth, 2016). However, within a group of fast-growing rhizobia belonging to the same clade, it is possible to find sub-groups of strains that grow at very different rates. Furthermore, this study has shown that the growth rate is dependent on both the strain and the environment, which may present selective conditions for different growth rates given the differential proportion of these two types of strains in both populations.

In soil, drought events are known to affect total bacterial mass (Alster *et al.*, 2013) and bacterial community composition by increasing the presence of Gram-positive phyla (Barnard *et al.*, 2013) as a consequence of resource limitations (Naylor and Coleman-Derr, 2018). The environmental conditions in the UK may have selected for faster growing Rleg as the soils are rarely deficient in moisture. In contrast, in the areas of Spain where the strains used in this study were isolated, the availability of water is determined by the season. In summer there may be very little rain for several weeks with hot temperatures that favour evaporation. Whereas in autumn high volumes of precipitation can overwhelm soil percolation and extensive surface run-off or flooding can become an issue. Thus, these two contrasting environments determine the life cycle of the rhizobia in their respective regions. While in the UK there is a constant availability of water throughout the year which facilitates the presence and availability of resources in the soil, and hence faster growing strains thrive. This increased growth rate benefits strains that utilise resources quickly, whilst the extreme seasonality of the Spanish precipitation may encourage slower life cycles with a slower metabolism that may facilitate survival during long periods of water deprivation (reduced resource availability). Furthermore, in an environment with plenty of water it is likely that other microorganisms will also be growing faster and competing for soil nutrients, hence a faster growth cycle would enable Rleg to better compete for soil nutrients in these environments.

Other factors, such as pH, may also substantially contribute to the life cycle of these bacteria as it is one of the main environmental factors affecting rhizobial growth (Hirsch, 1996; Graham, 2008). While UK soils have a pH towards the acidic side of the scale, the Spanish soils used in this study had considerably higher pH close to pH 9. The growth assays in this study used TY broth at pH 6.8 and the increased number of slow-growing rhizobia of Spanish origin may have been influenced by this change in pH, as they are adapted to a higher pH which may have led them to grow more slowly. If this was the case, it should have affected all strains from Spain equally rather than just a sub-group of them, and strains of rhizobia demonstrating both growth phenotypes were indeed found in soils from both countries. However, although the UK populations used in this study comprised a more-or-less nation-wide strain representation, the strains from Spain were only sampled from five different fields in eastern Spain. Therefore, it could be that these fields were rich in the slow-growing Rleg and a different population composition may have been obtained from other parts of Spain.

Desiccation can cause severe damage to rhizobia, and the reduction in the growth rate observed for all strains might be a consequence of the rehydration process, independent of the country of origin or growth-rate type (García, 2011). Rhizobia surviving desiccation can undergo a reduction in the volume of the cytoplasm which generates mechanical stresses on the cellular membrane and negatively impacts metabolism. Such physical and metabolic effects can be fatal for non-adapted Rleg types (Berninger *et al.*, 2018). Upon rehydration, rhizobia need to re-balance their cellular activity and a reduction in growth rate is therefore observed due to the prioritisation of resources to repair damage suffered during anhydrobiosis. This reduction in growth rate may also indicate that desiccation stress selects for slower-growing rhizobia which would explain why there were more slower-growing types in the population from Spain compared to that from the UK.

The cultures used in the assays originated from a single colony to reduce genetic variability within the sample, and it was thus considered that all single cells in the culture had the same genome. Despite this, although a single colony is formed by thousands of bacteria generated from a single ancestor by binary fission, during replication mutations may be generated by different means such as tautomeric shift of bases, oxidative damage or deamination and depurination (Najafi and Pezeshki, 2014). When rhizobia were grown in TY before the start of the desiccation assay, they were also undergoing binary fission; without a selective pressure in this environment, all new mutants (e.g., with a differential response to desiccation) could have survived and started generating copies of themselves. Thus, when the droplet of the adjusted culture was desiccated at the bottom of the plate well, a selective pressure was applied to this small community of rhizobia some of which may contain variants that enabled them to better withstand desiccation. If growing at a slower rate increases the chances of surviving a desiccation event, this may be causing the resulting liquid culture to grow slightly slower.

Strains were classified as fast or slow growers based on an empirical observation of growth curves of the full dataset and the distribution of the data growth rate. The growth rate data of un-stressed cultures showed a bi-modal distribution which often indicates the presence of two different subgroups in the dataset. Initially, the country of origin was thought to be the factor that caused these two different growth speeds, strains isolated from one environment showing slower growth rate than others. Nevertheless, after a PCA analysis, there was no clear separation among populations, thus other grouping factor was thought to be involved here. The threshold for separating both growth speeds was chosen as the median growth rate of the dataset as it was thought to be the least subjective of the methods for separating among strains at time of analysis. However, this threshold is inherent to the dataset and a different median value would have been reached in a different set of strains. Moreover, the dataset does not contain the whole population and it is possible that the bimodal distribution seen is the result of fortuity on the sampling method which differentially selected for the extremes of a normally distributed dataset causing its shape to become apparently bimodal.

When comparing strains exhibiting slow or fast growth phenotypes, the fast-growing types appeared to withstand desiccation better than slower-growing types. This may be an intrinsic bias of the method for measuring desiccation tolerance by assessing the growth curve. It is possible that when a fast-growing rhizobium that does not tolerate desiccation is compared with a slow-growing strain that has a better tolerance to desiccation, this difference in growth between the two strains may overcome the differential tolerance to desiccation unless a standardisation method that accounts for this is used. In addition, the strains that demonstrated high tolerance to desiccation may be expressing genes that assist with water retention and avoid its loss to evaporation such as genes involved in the biosynthesis of trehalose or hydroxyectoine (Manzanera *et al.*, 2002; Streeter, 2003), or genes that are involved in the synthesis of membrane repair proteins (Cytryn *et al.*, 2007; McIntyre *et al.*, 2007).

Rhizobium leguminosarum is a bacterium that secretes many exopolysaccharides (EPS) and these molecules can assist in tolerating desiccation stress through the creation of biofilms (Costerton *et al.*, 1995; Donati *et al.*, 2011). Thus, it is possible that those strains that were better at tolerating desiccation had an increased production of EPS which also contributed to surviving the period of anhydrobiosis. A liquid medium with a higher concentration of EPS will take longer to fully evaporate, and even when dry may still contain higher quantities of water molecules than a culture containing less EPS. In this study, cultures were adjusted to the same concentrations of CFU mL^{-1} in order to standardise the number of bacteria. However, the concentration of EPS in the solution was not estimated and it is possible that this may have facilitated those strains secreting higher concentrations of EPS to better tolerate desiccation stress and to survive for longer. Although pelleting and re-suspending the rhizobia in fresh sterile medium would have eliminated this variability from the data, centrifugation may also have been a source of additional stress and damage to cells (Peterson *et al.*, 2012). Therefore, given the possible biases to assess desiccation tolerance with these types of growth assays, assessing just the growth curve of the stressed culture is unlikely to determine whether a strain is desiccation tolerant. Thus, using the growth curve of the non-stressed culture to calculate a ‘desiccation factor’ standardises the measurement of desiccation tolerance and, despite reducing the number of strains that can be tested simultaneously, it does remove the variability that EPS secretion and other strain-specific growth characteristics may add to the final measurement. Importantly, the presence and quantity of EPS is a trait that is intrinsic to any given rhizobia strain and will significantly add to their overall tolerance to desiccation in a real-life scenario during commercial inoculant formulation and application.

Rhizobia isolated from the Spanish semi-arid eastern coast have a better tolerance to desiccation than those strains isolated from the UK, which confirms previous reports where strains isolated from regions with higher soil moisture deficit were better at tolerating desiccation, e.g., in New Zealand (van Ham *et al.*, 2016). In the present study, the soil used to isolate rhizobia from both Spain and the UK was not treated in any way to improve the sampling of increased desiccation tolerant strains: such as via initial air drying of soil to increase the chances of finding these phenotypes (van Ham *et al.*, 2016). Nevertheless, as trapping and isolating rhizobia from soil takes several days this period may increase the chances of finding more desiccation tolerant strains and especially for soils from semi-arid environments. Furthermore, treating the nodules before isolating new strains may provide a further step to select strains with improved desiccation tolerance. Strain JHI1118 showed the best desiccation tolerance of the UK strains. This strain was isolated from a nodule from a dried herbarium specimen of *V. faba*. It is, therefore, possible that by selecting for isolates based on climate and extent of nodule drying prior to isolation may increase the likelihood of isolating more desiccation tolerant strains. It should be emphasised, however, that additional attributes of the strains should then be assessed via screening with the host for assessing their abilities to promote or enhance plant nodulation and symbiotic N fixation. Desiccation tolerant strains may not necessarily perform well when tested *in planta* as has been shown for example with JHI1118 (Maluk *et al.*, 2022).

The inoculant industry desires that the strains in the formulation tolerate desiccation, and so allow for a longer shelf-life whether stored/packaged, or after application to soil or seed. Desiccation is considered the main cause of reduced efficacy in inoculants (Deaker *et al.*, 2012), and a good formulation (i.e. containing the right range of additives) may maintain the population density of viable rhizobia. However, including a strain that naturally tolerates desiccation should increase the shelf-life even further. The results from strains that are already being used in commercial inoculants (i.e., rcr1045 and WSM455) show they have a medium to low desiccation tolerance compared to the other strains tested in this study, indicating that there is room for improvement. This is probably due to the fact that the inoculant industry has generally focused more on finding good performers from the symbiotic perspective, and have concentrated their efforts on keeping these strains alive as long as they can with formulation additives such as trehalose, peat, or peat-extract as a carrier (McIntyre *et al.*, 2007; Howieson and Dilworth, 2016; Atieno *et al.*, 2018).

This focus on the formulation is partially due to the high costs of isolating and characterising strains for the dual traits of desiccation-tolerance and symbiotic efficiency compared to improving the formulation for an already known efficient strain. In addition, formulation is easier and less expensive for manufacturers than bioprospecting for isolates of rhizobia from semi-arid environments. Furthermore, the formulation is the part of the inoculant which is bound by intellectual property, and thus provides the manufacturer with the unique selling point that may confer protected commercial advantage against competitors. Notwithstanding this, the combination of formulating a strain that tolerates desiccation with a carrier that improves this capacity may present a synergistic effect that could considerably increase the final product shelf-life, and effectiveness *in situ*.

In conclusion, this study has shown that rhizobia isolated from a semi-arid environment are better at withstanding desiccation. Therefore, future research in this area should investigate the environmental selection pressures and genetics underpinning this capacity. Given that the over 100 strains described here have been functionally characterised for desiccation tolerance, their whole-genome sequencing may be used, by association mapping, to identify genomic regions responsible for this capacity. If success is achieved in this respect, such knowledge may be applied to allow the molecular-assisted identification of strains with enhanced desiccation tolerance.

Chapter 4 | Genomic analysis of rhizobia from pea and faba bean and identification of genes involved with desiccation tolerance

Abstract

Legume inoculants are affected by the die-off of cells which reduces their shelf-life. This affects both the distribution time and application period of the inoculant limiting the global reach and use of these products. The main cause of this loss in viable rhizobia is desiccation. Pea and faba bean nodulating rhizobia do not sporulate, leaving them defenceless during desiccation. With the advent of high throughput sequencing techniques, the study of complex phenotypes, such as desiccation tolerance, at a genome wide scale has been facilitated. The aim of this study was to identify genes involved in desiccation tolerance of pea and faba bean symbionts from the *Rhizobium leguminosarum* species complex (Rlc). For this, the genomes of 71 strains of rhizobia isolated from Spain and the UK which showed differential response to desiccation, were sequenced. After classifying the strains in one of the 18 Rlc genospecies by pairwise average nucleotide identity (ANI) calculations, a variant call was made with strains belonging to gsC which generated over 300,000 single nucleotide polymorphisms (SNP). A genome-wide association study (GWAS) conducted on the bi-allelic SNPs of the gsC strains revealed a strategy in common with other rhizobia under desiccation events i.e. increasing the concentration of the cytoplasm by the *de novo* synthesis of osmolytes like trehalose, or uptake of osmolytes from the medium, and curating and protecting the genetic material by using proteins involved in the replication and transcription processes. However, the data showed that desiccation tolerance might be linked with particular genospecies with some tolerating this stress better than others. Thus, future research should focus on increasing the number of genomes of each genospecies and running a GWAS independently on each genospecies to assess in full the set of tools to tolerate desiccation of this group of bacteria.

Keywords

Desiccation, anhydrobiosis, GWAS, *Rhizobium leguminosarum*

4.1 Introduction

Rhizobium leguminosarum species complex (Rlc) species are native to UK soils and spontaneously nodulate peas (*Pisum sativum* L.) and faba beans *Vicia faba* L.), although in many cases the capacity of these soil dwelling bacteria to nodulate and fix nitrogen may be far from optimal (Mutch and Young, 2004; Macdonald *et al.*, 2011; Maluk *et al.*, 2022). Thus, selecting specific strains (i.e., so-called 'elite' strains) from the Rlc that can improve root colonisation, N-fixation and plant growth promotion and applying them to the seed in the form of seed inoculants can improve the outcomes of this symbiosis. In the seed inoculant industry, the final product must be capable of sustaining the viability of the formulated organisms. Therefore, desiccation - the main cause of cell die-off and shelf-life reduction in commercial inoculants, is a considerable problem (Deaker *et al.*, 2012; Bashan *et al.*, 2014; Berninger *et al.*, 2018). Desiccation tolerance is a complex trait that is mediated by many changes in cell function, and although the roles of some prominent genes or molecules have been identified, there are likely many other physiological changes involved with the survival of rhizobia during desiccation (Casteriano *et al.*, 2013).

Under anhydrobiosis, the rhizobial cell suffers changes in the permeability of the membrane, combined with protein and membrane damage caused by reactive oxygen species (ROS), browning (Maillard) reactions, and phase transition upon rehydration, all of which can be lethal (Potts, 1994, 2001; García, 2011). Gram-negative bacteria have many strategies to withstand desiccation, such as increasing the expression of genes involved in DNA mismatch repair proteins (*mutS*) or genes coding for outer membrane proteins (*oprH*) (Pazos-Rojas *et al.*, 2019). Furthermore, silencing genes involved in the repair of the 3-dimensional structure of the DNA, such as *uvrABC*, has been shown to make *Ensifer (Sinorhizobium) meliloti* strains more sensitive to desiccation events than wild-type strains (Humann *et al.*, 2009). The *de novo* synthesis or accumulation of osmoprotectants, e.g., trehalose, is a common strategy used by many rhizobia exposed to desiccation stress to protect themselves from cellular damage (McIntyre *et al.*, 2007; Reina-Bueno *et al.*, 2012). There are four main pathways for trehalose biosynthesis in bacteria mediated by the genes *otsAB*, *treS*, *treYZ* and *treT* with that mediated by *otsAB* as the most widespread route (McIntyre *et al.*, 2007; Sugawara *et al.*, 2010; Ruhal *et al.*, 2013).

A transcriptome analysis of *Bradyrhizobium japonicum* under anhydrobiosis found the upregulation of over 200 genes, including a greater than two-fold increase in expression of *otsAB* and *treS* (Cytryn *et al.*, 2007). In addition, the uptake and accumulation of the osmoprotectant betaine or its precursor choline-O-sulfate which is mediated by *betS* and *betR* respectively, has been shown to play an important role during early osmotic adjustment and anhydrobiosis in *S. meliloti* and in *Klebsiella variicola* (Boscari *et al.*, 2002; Rodríguez-Andrade *et al.*, 2019). Similarly, as protection from oxidative damage, *B. japonicum* will upregulate genes involved in cellular motility, and in the synthesis of exopolysaccharide and chaperone proteins (Donati *et al.*, 2011).

Genes involved in desiccation tolerance in the Rlc have only been studied via gene mutagenesis. This approach has been successful for identifying genes in the biosynthesis of trehalose (McIntyre *et al.*, 2007), ATP-binding proteins (Vanderlinde *et al.*, 2010), proteases (Gilbert *et al.*, 2007), lipopolysaccharides (Vanderlinde *et al.*, 2009) and membrane repair and cell envelope proteins (Neudorf and Yost, 2017; Atieno *et al.*, 2018) that are involved in desiccation tolerance of Rlc species. Furthermore, chaperone synthesis together with membrane repair protein and ribosomal protein synthesis have also been linked to enhanced desiccation tolerance in rhizobia grown in an aqueous peat extract, a medium that can stimulate physiological changes capable of preparing cells for desiccation stress (Casteriano *et al.*, 2013).

Recent advances in whole genome sequencing and computing have allowed the examination of entire bacterial genomes and have facilitated genome-wide analyses of complex traits. In rhizobia, genome-wide studies have been used for the identification of alleles related to symbiotic traits or the use of different carbon sources (Epstein *et al.*, 2018), as well as for the study of introgression events in Rlc strains (Cavassim *et al.*, 2020). Thus, the aim of this chapter is to utilise a whole genome approach to increase our understanding of the genes involved in desiccation tolerance of the Rlc, and to identify genetic markers that can be used for the rapid selection of desiccation-tolerant strains for their potential use as seed inoculants.

4.2 Materials and methods

4.2.1 Genomes analysed

The genomes of 70 strains with a wide range of desiccation tolerance (assessed in Chapter 3) were sequenced. From these, 27 were isolated from Spain, 38 from the UK, and 5 of commercial interest that originated in other countries (Table 4.1). Furthermore, the genome of the reference strain Rlv3841 (Young *et al.*, 2006) was also included in the dataset.

Table 4.1. Strains used in this study. Column DT indicates the desiccation tolerance of the strain.

Strain	Country	Location	Host	DT	Accession No
111A12	Spain	Ontinyent	*P. sativum*	Low	-
121B21	Spain	Ontinyent	*P. sativum*	High	-
21A12	Spain	Valencia	*P. sativum*	High	-
21B12	Spain	Valencia	*P. sativum*	High	-
41A11	Spain	Valencia	*P. sativum*	High	-
42B12	Spain	Valencia	*P. sativum*	Low	-
43A11	Spain	Valencia	*P. sativum*	Low	-
43B11	Spain	Valencia	*P. sativum*	High	-
43B12	Spain	Valencia	*P. sativum*	Low	-
51A11	Spain	Valencia	*P. sativum*	High	-
51B21	Spain	Valencia	*P. sativum*	High	-
63A21	Spain	Valencia	*P. sativum*	Low	-
71A12	Spain	Ontinyent	*P. sativum*	High	-
73A11	Spain	Ontinyent	*P. sativum*	High	-
73B11	Spain	Ontinyent	*P. sativum*	Low	-
73B12	Spain	Ontinyent	*P. sativum*	High	-
81B22	Spain	Ontinyent	*P. sativum*	High	-
83A12	Spain	Ontinyent	*P. sativum*	High	-
93B11	Spain	Ontinyent	*P. sativum*	High	-
JHI10	UK	Angus	*P. sativum*	High	GCF_010668925.1
JHI1084	USA	Yelm	*Lathyrus sativus*	High	GCF_010668425.1
JHI1093	UK	Angus	*Lathyrus linifolius*	Low	GCF_010668385.1
JHI1096	UK	Angus	*L. linifolius*	High	GCF_010668355.1
JHI1118	UK	Warwickshire	*V. faba*	High	GCF_010668395.1
JHI1236	UK	Yorkshire	*V. faba*	Low	GCF_010668085.1
JHI1238	UK	Yorkshire	*V. faba*	Low	GCF_010668055.1
JHI1253	UK	Orkney	*P. sativum*	Low	GCF_010668345.1
JHI1259	UK	Orkney	*P. sativum*	Low	GCF_010668315.1
JHI1266	UK	Orkney	*P. sativum*	Low	GCF_010668285.1
JHI13	UK	Angus	*P. sativum*	Low	GCF_010668945.1
JHI1415	UK	Wiltshire	*Lens culinaris*	High	GCF_010668265.1
JHI1422	UK	Wiltshire	*L. culinaris*	High	GCF_010668245.1

Table 4.1. (Continuation).

Strain	Country	Location	Host	DT	Accession No
JHI1438	UK	Angus	*P. sativum*	High	-
JHI1587	UK	Cambridge	*P. sativum*	Low	GCF_010668175.1
JHI1592	UK	Skye	*P. sativum*	High	GCF_010668145.1
JHI1600	UK	Wiltshire	*P. sativum*	High	GCF_010668195.1
JHI24	UK	Angus	*V. tetrasperma*	Low	GCF_010668905.1
JHI2442	USA	Virginia	*P. sativum*	Low	-
JHI2449	UK	Norfolk	*P. sativum*	High	GCF_010668165.1
JHI2450	UK	Norfolk	*P. sativum*	High	GCF_010668125.1
JHI2451	UK	Norfolk	*P. sativum*	High	GCF_010668065.1
JHI370	UK	Angus	*V. faba*	Low	GCF_010668785.1
JHI387	UK	Angus	*V. faba*	Low	GCF_010668765.1
JHI388	UK	Angus	*V. faba*	Low	GCF_010668735.1
JHI42	UK	Angus	*V. faba*	Low	GCF_010668865.1
JHI535	UK	Wiltshire	*V. faba*	High	GCF_010668705.1
JHI536	Unknown	Inoculant	*Inoculant*	High	-
JHI54	UK	Angus	*V. sativa*	Low	GCF_010668875.1
JHI585	UK	Wiltshire	*V. faba*	High	GCF_010668715.1
JHI782	Greece	Mykonos	*V. faba*	Low	GCF_010668685.1
JHI783	UK	Hertfordshire	*P. sativum*	Low	GCF_010668665.1
JHI787	Ethiopia	Enemay	*V. faba*	Low	GCF_010668635.1
JHI788	UK	Roxburghshire	*V. faba*	Low	GCF_010668585.1
JHI925	UK	Angus	*V. faba*	High	GCF_010668625.1
JHI944	UK	Angus	*V. faba*	High	GCF_010668605.1
JHI953	UK	Angus	*V. faba*	High	GCF_010668555.1
JHI960	UK	Oxfordshire	*V. faba*	High	GCA_010668545.1
JHI963	UK	Oxfordshire	*V. faba*	Low	GCF_010668515.1
JHI973	UK	Angus	*P. sativum*	Low	GCF_010668505.1
JHI974	UK	Angus	*P. sativum*	Low	GCF_010668475.1
JHI979	UK	Angus	*P. sativum*	Low	GCF_010668465.1
JHI985	UK	Angus	*P. sativum*	Low	GCF_010668445.1
Rlv3841	UK	Norfolk	*P. sativum*	High	GCF_000009265.1
VFCR2A2	Spain	Valencia	*V. faba*	High	-
VFF1R1A2	Spain	Valencia	*V. faba*	High	-
VFF1R2A1	Spain	Valencia	*V. faba*	High	-
VFF1R2B1	Spain	Valencia	*V. faba*	High	-
VFF2R2A1	Spain	Valencia	*V. faba*	High	-
VFHR1A2	Spain	Ontinyent	*V. faba*	High	-
VFSR2A2	Spain	Ontinyent	*V. faba*	High	-
VFSR2B1	Spain	Ontinyent	*V. faba*	High	-

4.2.2 Extraction and sequencing of genomic DNA

A culture at log phase (OD_{600} 0.2 – 0.8) of each strain grown in tryptone yeast broth (5 g tryptone, 3 g yeast extract (Fermtech, Merck), 1 L SDW and pH 6.8) for

48 h at 30 °C and 2 Hz was centrifuged at 1900 RCF for 15 minutes. The pelleted cells were then lysed in 10 mM Tris-HCl, 1 mM EDTA, 0.5% SDS and 19.05 units mL^{-1} Proteinase K for 1 h at 37 °C. Following this, an equal volume of phenol:chloroform:isoamilalcohol was mixed into the lysed cells and vortexed thoroughly and centrifuged at 13200 RCF for 12 minutes. Between 170 and 180 µL of DNA were pipetted out from the resulting top layer and placed in a clean tube to which a 0.078 M sodium acetate in 100 % propanol was added prior to vertexing well. The mixture was then incubated at -20 °C overnight. Following this, a centrifugation at 13200 RCF for 12 minutes pelleted the DNA which was subsequently washed and resuspended in 70 % ethanol before a second centrifugation in the same conditions. The final supernatant was discarded, and the remaining ethanol allowed to evaporate completely in an incubator at 28 °C. The DNA pellet was then resuspended in sterile distilled water and a NanoDrop ND-1000 (NanoDrop Technologies Inc., USA) was used to assess nucleic acid concentration before storage at -20 °C. DNA was sent to MicrobesNG (Birmingham, UK) for whole-genome shotgun sequencing by Illumina (Illumina Inc., USA) following internal protocol v20210419 (MicrobesNG, 2021). In summary, for each DNA sample a genomic DNA library was prepared using the Nextera XT Library Prep Kit (Illumina, USA) following the manufacturer’s protocol but increasing the template DNA 2-fold and the PCR elongation step to 45 s. The library was prepared and the DNA quantified in a Hamilton Microlab STAR automatic handling system (Hamilton Bonaduz AG, Switzerland). Following this, the libraries are quantified with a Kapa Biosystems Library Quantification Kit for Illumina and sequenced using Illumina sequencers HiSeq and NovaSeq with a 250 bp paired end protocol. Then the reads are treimmed using Trimmomatic 0.30 (Bolger *et al.*, 2014) with a quality cutoff of Q15 followed by a de novo assembly using SPAdes v3.7 (Bankevich *et al.*, 2012) and annotation with Prokka 1.11 (Seemann, 2014). A contig and trimmed read files were subsequently provided for each strain.

4.2.3 Genome assembly and annotation

The contig files provided by MicrobesNG and those downloaded from NCBI were cleaned of all sequences with less than 200 nucleotides and the assembly was done following Cavassim *et al.* (2020) assembly method. For this, scaffolds were built using the Python script Jigome (Cavassim *et al.*, 2020) which orientates, aligns and concatenates contigs into scaffolds by matching the end of a contig with the beginning of the following one until it no longer can be extended. In short, a set of 3215 core genes and *repA* alleles (plasmid replication initiation gene) were used to putatively assign each scaffold to a chromosome or plasmid, respectively (Cavassim *et al.*, 2020). Then, a set of 47 genomes (Cavassim *et al.*, 2020) was used to align each contig and continue elongating it when possible or to place an arbitrary spacer of 20 ‘N’ symbols when contigs did not overlap. For labelling, the *dnaA* (DNA start of replication) gene was searched for, and the first scaffold contig started at the ATG region of this gene with the other chromosomic scaffolds numbered subsequently. As the Rlc chromosome is circular, the chromosomic scaffold numbered 00 is situated immediately upstream of the *dnaA* ATG region. The plasmid scaffolds were labelled according to the *repA* allele they carry for which a library of 20 *repA* alleles is used. Those scaffolds that were impossible to assign to the chromosome or plasmid were labelled as fragments in order of decreasing size. The quality of the assembly was assessed using QUAST v5.0.2 (Mikheenko *et al.*, 2018) and the annotation was performed using Prokka v1.14.6 (Seemann, 2014), both of which were used with default options.

4.2.4 Identification of orthologous genes for preliminary population analysis

Orthologous genes were identified among genomes with proteinortho v6.0.23 (Lechner *et al.*, 2011) with the synteny option activated. All gene groups with only one gene were removed and the remaining groups were analysed with Syntenizer3000 (Cavassim *et al.*, 2020) for the identification of syntenic genes by comparing the 40-gene neighbourhood of each gene group among genomes. Syntenic genes were removed from the group and moved into a new group, and only orthologous genes with the same neighbourhood remained in each gene group.

The disambiguated gene groups were aligned using Clustalo v1.2.4 (Sievers and Higgins, 2018) and *codon_aware_clustal.py* (Cavassim *et al.*, 2020) which translates the gene sequences into proteins, aligns them and translates them back to DNA by inserting three gap symbols (‘-‘) per amino acid gap found. Single nucleotide polymorphisms (SNP) were extracted from the aligned genes using *rhizob_ld.py* (Cavassim *et al.*, 2020) which searches SNPs among the codon-aware alignments. In this step, only gene groups found in at least 35 strains were used

(Cavassim *et al.*, 2020), multi-allelic SNPs were removed leaving only bi-allelic SNPs; a matrix with major alleles encoded with a ‘1’ and minor alleles with a ‘0’ was then generated. This script subsequently used the generated matrix to produce a Principal Component Analysis plot.

4.2.5 Variant calling and GWAS on gsC strains

Prior to running GWAS the population genetic homogeneity was assessed on a PCA of SNPs (Section 4.2.4) and given the marked separation among genospecies GWAS was only performed on the genomes from the strains belonging to the most frequent genospecies (gsC). The variant call was performed with bowtie2 v2.3.5.1 (Langmead and Salzberg, 2012) using the trimmed paired-end fastq files of each strain and the reference genome designated by Young *et al*. (2021) (*R. ruizarguesonsis* UPM1133^T) as a representative for the genospecies as a template. The contigs of the reference strain were annotated as belonging to the chromosome, plasmid or fragment as described above using Jigome (Cavassim *et al.*, 2020). The output generated by bowtie2 was transformed into BAM format, sorted and indexed with SAMtools v1.9 (Danecek *et al.*, 2021). Then, a variant call was made using FreeBayes v0.9.21.7 (Garrison and Marth, 2012) with a flag for marking samples as haploid.

Variant data quality control was performed with vcftools v0.1.16 (Danecek *et al.*, 2011) and vcfR v1.12.0 (Knaus and Grünwald, 2017) in three QC steps. For this, all indels and multiallelic variants were filtered out in a first step with vcftools. On the second step, all variants genotyped in less than 50% of individuals and had a minimum quality score of 30 were removed with vcftools. On the third step, all variants below 10 % and above 90 % of the dataset read depth, samples with over 55 % missing data and all variants with missing data were removed with vcfR.

Prior to running the genome-wide association study (GWAS), a Neighbour Joining phylogenetic tree was generated with ape v5.5 (Paradis and Schliep, 2019) with 10,000 bootstraps using the average nucleotide identity (ANI) values as a distance matrix (Sánchez-Cañizares *et al.*, 2018) among strains and strain JHI536 as an outgroup. The tree was rooted to JHI536 and its tip removed afterwards. The desiccation factor (DF) was binarized depending on whether they had a higher (1) or lower (0) DF than the average DF. The average DF was calculated separately for each type of growth rate (estimated in Chapter 3), thus, for fast-growing rhizobia high desiccation tolerance was assigned when they had a DF > 0.44, whilst for slow-growing rhizobia the threshold was set at DF > 0.36. The GWAS analysis was performed with hogwash v1.2.5 (Saund and Snitkin, 2020) with 50,000 permutations and 0.005 false discovery rate. Both phyC (Farhat *et al.*, 2013) and Synchronous (Saund and Snitkin, 2020) tests were run with the same hogwash parameters. Significant SNPs were mapped on the reference genome and annotation information was extracted using the general feature format (gff) file of the reference strain.

4.2.6 Genospecies assignation

The strains belonging to the Rlc have recently been classified into 18 genospecies(Kumar *et al.*, 2015; Young *et al.*, 2021) therefore a phylogenetic analysis of the concatenated *atpD, gyrB* and *recA* genes of each strain in addition to those of reference strains available on NCBI (Table S4.1) was carried for a preliminary genospecies assignation. Python scripts *find_genes.py* and *concat_seqs.py* (Young *et al.*, 2021) were used for the location of the genes in each genome and their concatenation respectively. For this, the corresponding protein sequence of gsA strain SM130B (GCA_004304475.1) (Young *et al.*, 2021) was blasted on each genome and extracted using blast+ v2.12.0 (Camacho *et al.*, 2009). Then the sequences of each gene were aligned with clustalO v1.2.4 (Sievers and Higgins, 2018), an arbitrary spacer of 3 ‘N’ symbols was added to each end of the sequence and the genes were concatenated in the order *atpD-gyrB-recA* .

Following this, pairwise distances among strains was calculated with Mega X v10.0.5 (Kumar *et al.*, 2018) using the ‘number of differences’ method. When identical sequences were found, only one sequence of each repeated group was left as a representative of the group. A preliminary phylogeny was then estimated using FastTree v2.1.10 (Price *et al.*, 2010) followed by an optimal model estimation with ModelTest-NG v0.1.6 (Darriba *et al.*, 2019) and the final phylogeny was calculated by maximum likelihood with 1000 bootstraps using the best fit model and the preliminary phylogeny as a starting topology with RaxML-NG v1.0.3 (Kozlov *et al.*, 2019).

A genospecies (gs) was assigned to each strain depending on the reference strains of known gs in each branch. Furthermore, the pairwise average nucleotide identity (ANI) was calculated using the scaffolds built with Jigome (Section 4.2.3) and fastANI v1.1 (Jain *et al.*, 2018) and an ANI heatmap was built using *ANI_heatmap.py* (Young *et al.*, 2021) for corroborating the phylogenetic genospecies assignation.

4.2.7 *nodD* type assignation

Following on the characterisation of the strains, the *nodD* type of each strain was determined (Boivin et al 2020 2021). For this, the *nodD* sequence of reference strain Rlv3841 (Young *et al.*, 2006) was used to extract the sequences for this gene from each genome using *find_genes.py* script (Young *et al.*, 2021). Reference sequences for each type of *nodD* (Boivin *et al.*, 2020) were downloaded from NCBI (Table S4.2). A codon-aware alignment was performed by Muscle (Edgar, 2004) on MegaX v10.0.5 (Kumar *et al.*, 2018) with default settings. The alignment gamma distribution was calculated before running a Neighbour-Joining (Saitoh, 1987) phylogenetic analysis using the maximum composite likelihood model and the calculated gamma distribution and 1000 bootstraps to test the phylogeny.

4.2.8 Analysis of desiccation genes

A total of 20 genes known to be involved in desiccation tolerance of rhizobia or free-living N-fixing bacteria were extracted from each genome (Table S4.3). Each gene locus was searched in the annotation of genome of Rlv3841 (Young *et al.*, 2006). When the gene was not annotated, the nucleotide sequence of the species studied in the literature was downloaded and used to find the homologous sequence using the online blastn tool with default parameters apart from the organism option which was defined to Rlv3841 (taxid: 216596) (*https://blast.ncbi.nlm.nih.gov/Blast.cgi*). Then, the Rlv3841 sequence for each gene was used to search the gene locus in each genome as previously done with *nodD* (Section 4.2.7). Genes were concatenated as described in Section 4.2.6 in alphabetic order and phylogenetic analysis was carried out following the same procedure described in Section 4.2.7.

4.3 Results

4.3.1 Genome assembly and characterisation

The Rlc assembled genomes had an average genome comprised of 7.5 Mb, which was distributed between the chromosome (5.2 Mb), several plasmids (1.8 Mb), and several fragments (0.6 Mb), i.e. scaffolds that were not possible to label as either belonging to chromosome nor plasmid (Table S4.4). The genomes had an average of 60.82 % GC content and a median coverage of 29. The average N50 and L50 were 903615 and 4, respectively.

Only ten out of the 18 genospecies defined by Young *et al*. (2021) were found in the dataset. The *atpD-gyrB-recA* phylogenetic analysis showed that two of the strains did not belong to the Rlc, one being clustered with *R. anhuiense* (JHI536) and the second one grouped with *R. tropici* (JHI1118) (Figure S4.1). The ANI analysis of the assembled genomes confirmed most of the genospecies assignations by the concatenated genes phylogenetic analysis apart from two of the strains previously assign to gsC which showed ANI values < 96 % with other gsC strains and similar ANI values to strains from gsJ (Figure 4.1A and Table S4.5 for ANI values). The distribution of the Rlc strains into certain genospecies depended on the country of isolation (Figure 4.1B). While gsC and gsB were the most common genospecies in the UK, gsR was the most common one in Spain. Importantly, strains from Spain and the UK belonged to completely different genospecies.

The nodulation gene *nodD* was found in all Rlc strains and in the *R. anhuiense* strain (JHI536). The only strain that did not have *nodD*, or a homologous protein, was strain JHI1118, which clearly did not belong to the Rlc. The type of *nodD* was not bound to a genospecies with many of them sharing *nodD* types, for example, type A1 was present on gsC, K and L or type B1 was present in gsB, C, E, J, N, Q and R (Figure 4.2A and Figure S4.2). However, some genospecies (J, L, O, Q and R) only possessed one type of *nodD*. *nodD* type B1 was the most frequent type and accounted for 62 % of all sequences found. In contrast, strains isolated from Spain only had B types of *nodD* which, apart from type B2, were also present in strains isolated from the UK (Figure 4.2B). The UK-isolated group harboured all types of *nodD*.

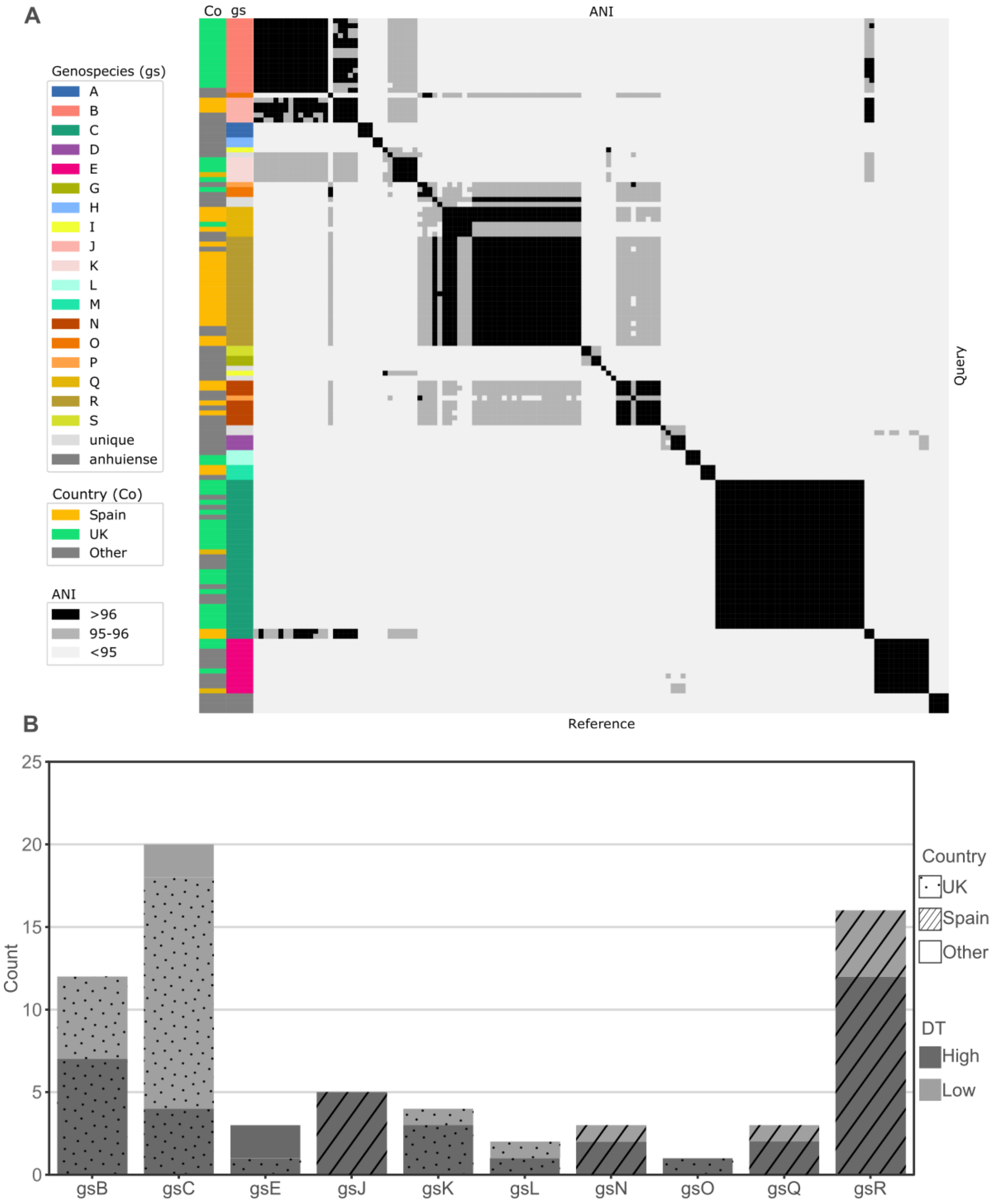


Figure 4.1. Genospecies assignation results from ANI calculations (A). In the heatmap rows are the query genomes and in the columns reference genome used as comparison. ANI values > 96 % are coloured black. The two bars on the left indicate the country (Co) and genospecies (gs) assigned on the *atpd-gyrB-recA* tree of each query genome. The second plot (B) shows the country and desiccation tolerance strains in each genospecies.

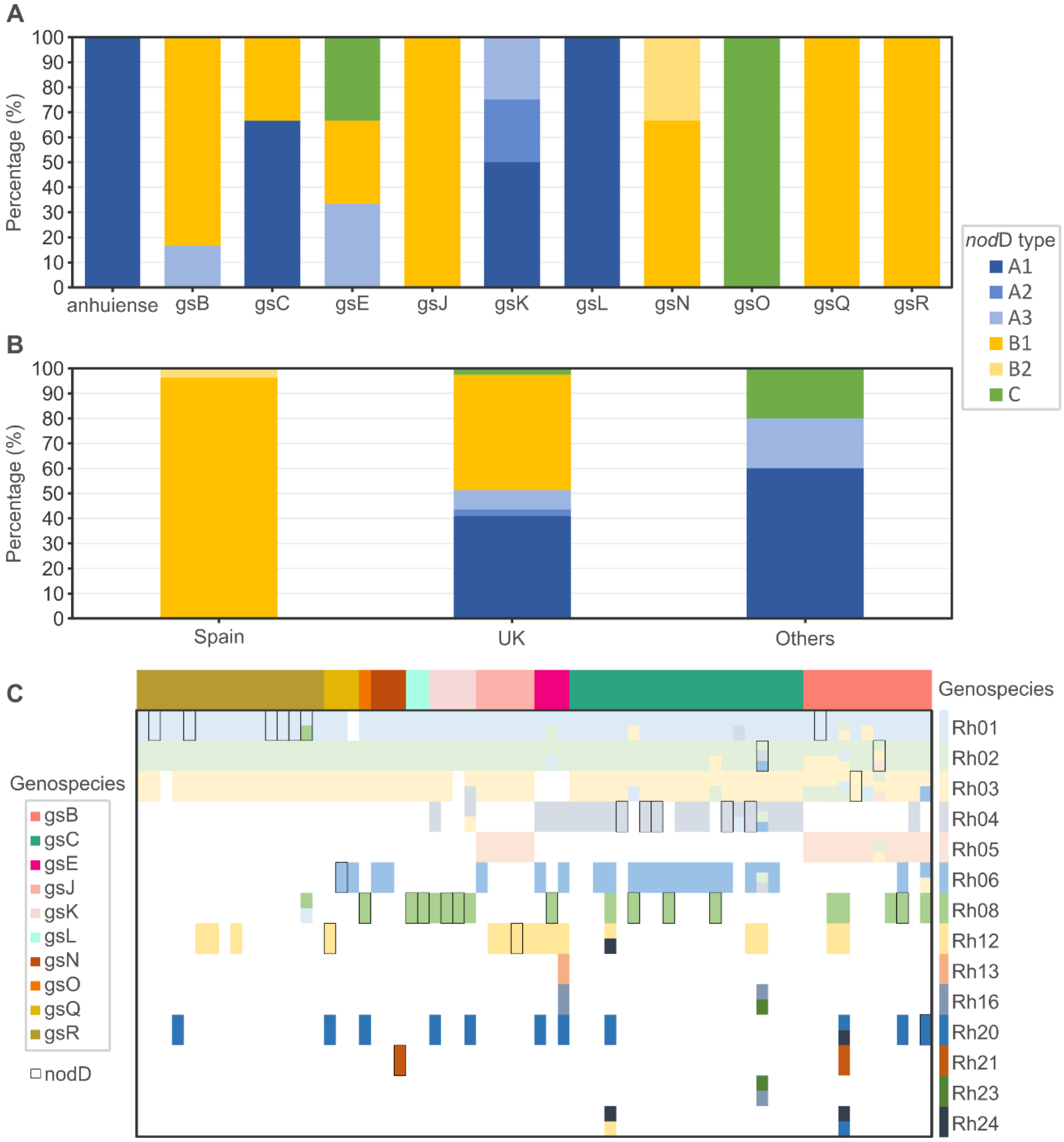


Figure 4.2. Type of nodD found in the dataset per genospecies (A) and per country (B). Plasmid groups found in each genospecies (C). Each column in panel C represents a strain, and the coloured rectangles represent the presence of the plasmid (detailed in the rows), and those with a black outline indicate the presence of *nodD* in that plasmid. Multi-coloured rectangles indicate plasmids with two *repA* types.

The gene *repA* was present in all Rlc strains and 14 different *repA* groups were found in total, each representing a different type of plasmid (Figure 4.2C). Plasmids Rh1-3 were the most frequently encountered, being present in almost every genome in the dataset. Plasmid types were not bound to a particular genospecies and many of them were shared among different genospecies.

Moreover, the presence of *nodD* was not limited to a single plasmid type and it changed location even within the same genospecies.

Almost all strains showed the presence of genes which have been indicated as being involved in desiccation tolerance (Figure 4.3 A). Most of these genes (17) were located in the chromosome and only three of them, *otsB*, *treS* and *treY*, were mainly located in plasmids. The phylogenetic analysis of the concatenated sequences of these genes clustered the strains in the different genospecies with strong branch support for each group (Figure S4.3). Apart from *otsA* of the gsE strains, all other genes had a high similarity with the reference strain with the gsB strains showing the highest resemblance to the reference sequence of Rlv3841 (Figure 4.3B).

A total of 510,300 genes were predicted and grouped in 23,854 orthologous groups from which 6,403 groups were orphan genes (genes which were only found in one of the genomes analysed) (Table S4.4). Nine strains had more than 200 orphan genes which accounted for 42.14 % of the total number of orphan genes. The strains had between 20 and 1,441 paralogous genes in the genome with the top eleven strains contributing 62.97 % of the total of paralogous genes found. Paralogous genes were found in 1,639 orthologous groups. Variants were called on 4,788 aligned gene groups and 296,035 SNPs were discovered. The principal component analysis on the identified bi-allelic SNPs showed that the first two components explained 54.42 % of the variability found and revealed that the dataset is widely spread with clusters mainly comprising strains belonging to the same genospecies (Figure 4.4).

4.3.2 Genospecies variant call and GWAS

The variant call assessment on the 20 gsC strains found 346,333 SNPs. After the third QC step 20,947 SNPs present in all gsC strains remained. A total of 164 SNPs were found to have a significant association with the desiccation tolerance phenotype with the synchronous test while only 29 were found significant with the phyC test (Figure 4.5 A and B). All phyC test significant SNPs were also found significant in the synchronous test. Significant SNPs found by the synchronous test were mostly found in coding regions (71.3 %) although some were found in pseudogenes (2.4 %) and non-coding regions (26.2 %) of the genome. (Table S4.6).

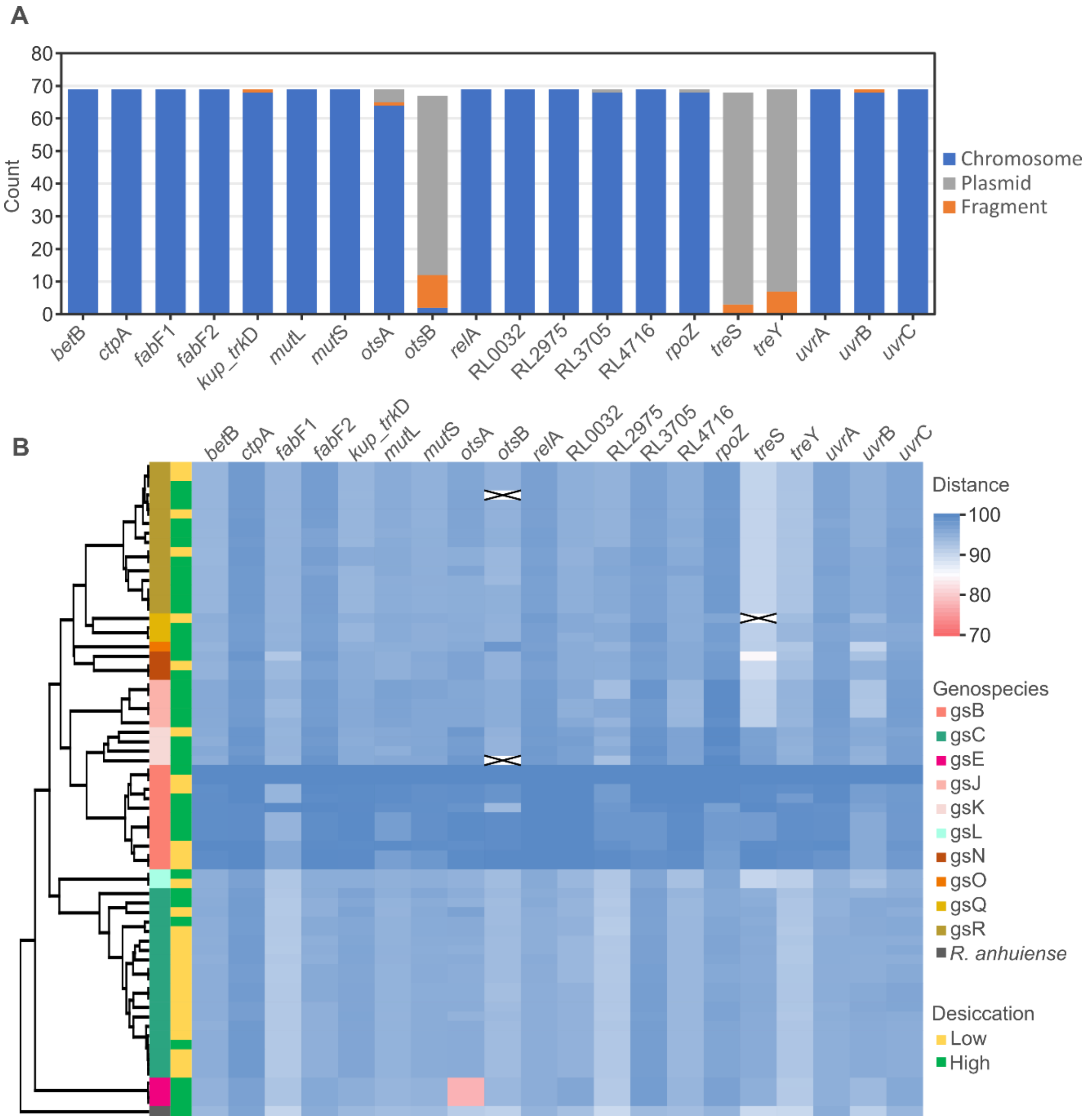


Figure 4.3. Location of desiccation genes in the genome of the Rlc strains (A). Heatmap of similarity of each strain gene with the reference gene from Rlv 3841 (B). The clustering method used was UPGMA with the concatenated sequence of all genes in the order shown. The black ‘X’ indicates genes that are missing.

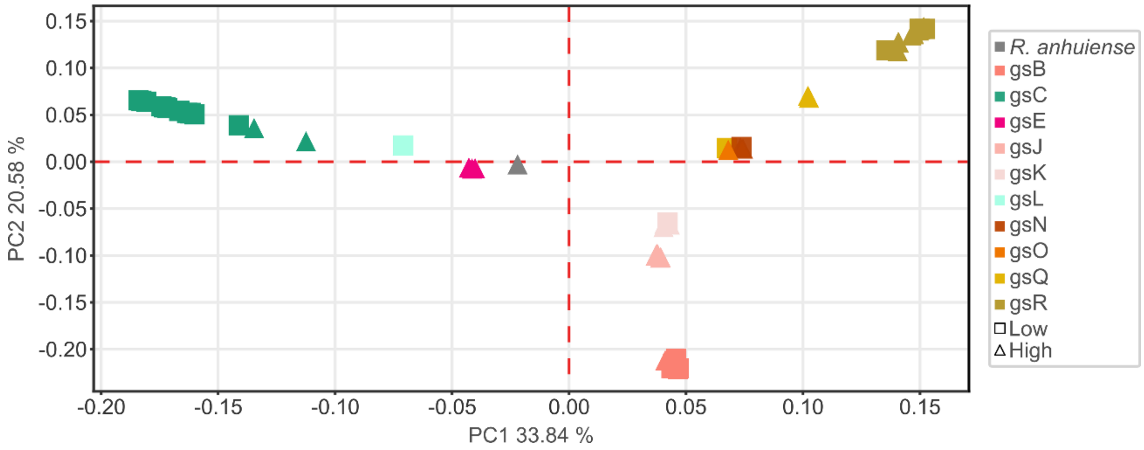


Figure 4.4. Principal component analysis of SNPs. Desiccation tolerance and genospecies are marked by symbol and colour respectively.

No SNPs with a significant association with the desiccation tolerance phenotype were found in chromosomal contigs. Significant SNPs located in genes or coding regions were found to belong to 120 unique genomic features. From them, only 19 were annotated genes whilst the rest were features inferred by gene prediction (Table 4.2). The predicted or known products of these genes were found to be involved in a range of processes (Table S4.6). These processes included membrane transport of substrates such as ABC transporters for sugar or carbohydrates (WP_018072987.1, WP_018480520.1, WP_024321462.1), taurine (*tauA*), nitrate (WP_131614440.1) or potassium (*kdpC*); DNA replication or transcription like plasmid replication genes (*repA*) and transcriptional regulators (WP_027687940.1, WP_018496327.1, WP_018069891.1); osmoprotectant biosynthesis like trehalose (*treS*); and, motility and chemotaxis genes (Figure 4.5 C).

Table 4.2. Annotated gene hits on the reference gsC strain UPM1133 (*Rhizobium ruizarguesonis*).

Gene	**Function**	**Reference**
tauA	Involved in sulphur membrane transport under lack of this element in the medium.	Javaux *et al.* (2007), Qu *et al.* (2019)
pcaC	Involved in the aerobic pyrene degradation pathway	Eulberg *et al.* (1998), Elyamine *et al.* (2021)
pcaG	Involved in the aerobic pyrene degradation pathway	Eulberg *et al.* (1998), Elyamine *et al.* (2021)
accC	Involved in byosynthesis of biotin and fatty acids	Abdel-Hamid and Cronan (2007)
alr	Involved in biosynthesis of bacterial peptidoglycan	Tauch *et al.* (2002), Oikawa *et al.* (2006)
cobG	Involved in the cobalamin synthesis	Martens *et al.* (2002), Ngabonziza *et al*. (2020)
doeA	Involved in ectoine degradation	Schwibbert *et al.* (2011)
ggt	Involved in microbial adaptation to hostile conditions such as drought, production of PGA	Bajaj and Singhal (2011), Scoffone *et al.* (2013), Najar and Das (2015)
glnT	Glutamine synthesis and Nitrogen assimilation	Chiurazzi *et al.* (1992), Forchhammer (2007)
iolE	Involved in myo-inositol catabolism	Yoshida *et al.* (2006), Kohler *et al.* (2010),
kdpB	High affinity K transporter under severe K limitation or osmotic upshift	Wood (1999), Ballal *et al.* (2007), Kannaiah *et al.* (2019)
kdpC	High affinity K transporter under severe K limitation or osmotic upshift	Wood (1999), Ballal *et al.* (2007), Kannaiah *et al.* (2019)
kduI	Conversion of galacturonate and glucuronate under osmotic stress	Rothe *et al.* (2013), Vorobjeva *et al.* (2020)
repA	Initiation of DNA replication in plasmids	Wetzel *et al.* (2015)
treS	Involved in trehalose biosynthesis	Sugawara *et al.* (2010)
tsdA	Involved in the thiosulfate oxidation pathway	Denkmann *et al.* (2012), Brito *et al.* (2015), Koch and Dahl (2018)
ugpC	Involved in the uptake of trehalose	Rivera-Araya *et al.* (2020)
zwf	involved in resistance to paraquat (oxidative stress reagents)	Ma *et al.* (1998), Kim *et al.* (2008), Kawai *et al.* (2015)

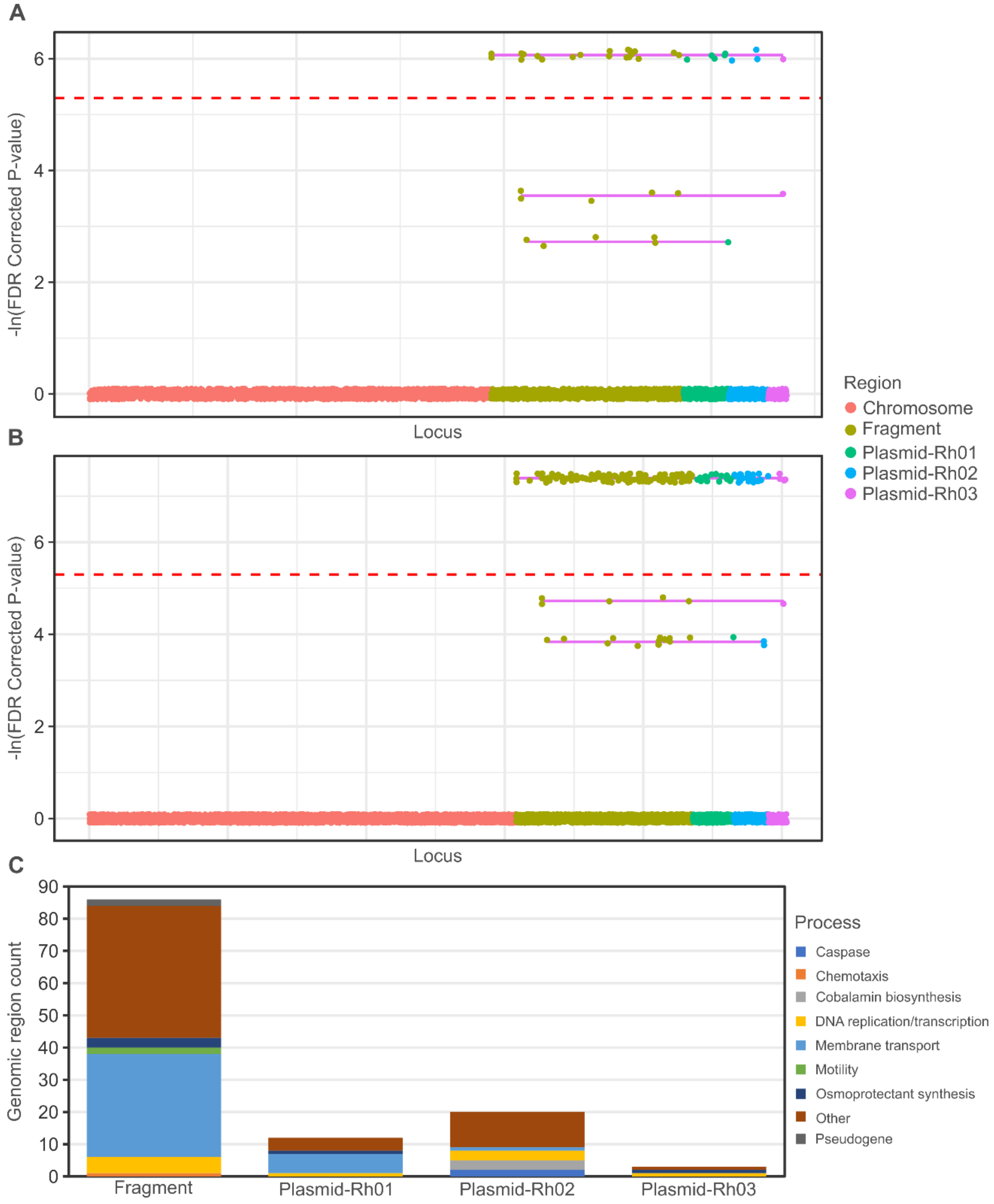


Figure 4.5. Manhattan plots of gsC SNPs resulting from the phyC and synchronous tests (A and B respectively) and distribution of cellular processes for identified SNPs (C). The dashed lines on A and B mark the significance threshold and the purple lines the P-values of the SNPs which have been jittered for ease of visualisation of very close datapoints.

4.4 Discussion

Genes involved in membrane transport, DNA replication and transcription, and osmoprotectant synthesis can have significant roles during desiccation tolerance in gsC of the Rlc. Some of these genes have known functions when bacteria are under severe stress conditions, yet the function of many of them remains unclear. These results show that many of the processes involved in desiccation tolerance are yet to be described, and that different organisms can have a different set of tools for withstanding desiccation. Nevertheless, the common theme of desiccation tolerance is to increase the osmolarity of the cytoplasm, and to protect their DNA.

Increasing cytoplasm osmolarity and protecting DNA are strategies shared with other bacterial species when under desiccation stress, although the mechanisms to achieve this may be different, since desiccation tolerance is a complex trait that affects the whole individual simultaneously impacting all biological functions and structures. This is because desiccation causes a severe imbalance of cellular homeostasis, membranes can lose their permeability, and metabolism is slowed down or even stopped which may lead to a cellular death (Potts, 2001; García, 2011). Desiccation can also lead to bacteria entering a capped state (i.e., viable but non-culturable), which removes their capacity for duplication (Vriezen *et al.*, 2012; Bravo *et al.*, 2016; Cholley *et al.*, 2020). Most of the genes known to have a role in desiccation tolerance of rhizobia or N-fixing bacteria shown in Table S4.3 are found in the chromosome of the strains studied here. In contrast, none of the significant SNPs identified on the gsC have been found in the chromosome. However, most of the SNPs were identified in fragments. The reference genome for gsC is the type strain *Rhizobium ruizarguesonsis* UPM1133 (Jorrin *et al.*, 2020; Young *et al.*, 2021) whose genome assembly is fragmented into 154 contigs. Most of these contigs were possible to assign to either the chromosome or a plasmid using a set of Rlc core genes and *repA* types (Cavassim *et al.*, 2020), but many remained un-assigned and were hence labelled as fragments. Therefore, if a more gsC-focused assembly was performed with a set of gsC-only genomes, rather than using the set of Rlc genomes on Jigome during genome assembly, it may be that many of these fragments can be found as part of the chromosome.

Of the known genes that play a role in desiccation tolerance, the GWAS analysis on gsC strains only found a significant SNP on the plasmid-bound *treS* gene. The product of this gene catalyses the transglucosylation of maltose into trehalose, a disaccharide frequently synthesised under desiccation, as it generates hypertonicity of the cytoplasm, thus reducing the water loss caused by an anhydrobiotic environment (Streeter, 2003; Reina-Bueno *et al.*, 2012; Ruhal *et al.*, 2013). Trehalose has been shown to be a widespread osmolyte under desiccation stress or in high salinity environments in many bacteria (McIntyre *et al.*, 2007; Sugawara *et al.*, 2010; Reina-Bueno *et al.*, 2012; Rivera-Araya *et al.*, 2020). All strains studied here are equipped with genes of the trehalose biosynthetic pathways mediated by *otsAB, treYZ* and *treS* illustrating the importance of this disaccharide for rhizobia. Moreover, the GWAS results also showed a significant hit on *ugpC*, a gene involved in trehalose uptake (Rivera-Araya *et al.*, 2020). The reference strain for gsC, like all the other gsC strains studied here, has two copies of *ugpC* in its genome, and both copies showed a significant hit on the GWAS analysis further highlighting the relevance of trehalose under desiccation stress.

Control over cytoplasm osmolarity seems to be the most recurrent adaptation for rhizobia to withstand desiccation. In addition to genes involved with the synthesis and uptake of trehalose, membrane transport of other osmolytes such as potassium (*kdpB* and *kdpC*), and for betaine/L-proline (WP_130663140.1) also seemed to have a significant role in desiccation tolerance of gsC strains. Both potassium and betaine/L-proline are known osmolytes involved in the control of cytoplasm osmolarity under desiccation stress (Wood, 1999; Boscari *et al.*, 2002; Ballal *et al.*, 2007). Furthermore, many other putative membrane transporters appeared to play a significant role in gsC desiccation tolerance, which could indicate that other osmolytes may be incorporated into the cytoplasm to counter the adverse environment.

In a water deficient environment, membranes and DNA can become damaged due to low availability of water and oxidative stress (García, 2011). Genes involved in biosynthetic pathways of DNA repair, replication and transcription were also shown to play a significant role in desiccation tolerance, e.g., *repA* which is involved in the initiation of plasmid replication (Cavassim *et al.*, 2020). Furthermore, *accC* is involved in the regulation of biotin (vitamin H) biosynthesis by repressing the *bio* operon which synthesises biotin, a vitamin necessary for the synthesis of fatty acids and the metabolism of amino acids, cholesterol and urea (Streit and Entcheva, 2003; Abdel-Hamid and Cronan, 2007). Similarly, *cobG* plus two significant loci inferred by gene prediction, are involved in the biosynthesis of cobalamin (vitamin B_{12}) which is necessary for the synthesis of Acetyl-CoA and ribonucleotide reductase, the latter necessary for DNA synthesis (Martens *et al.*, 2002). Vitamins H and B_{12} are both known to stimulate rhizobial growth even when present in very low concentrations in media (Watson *et al.*, 2001); they also appear to be important for the desiccation-tolerant phenotype. Finally, a gene involved in resistance to paraquat (*zwf*) was found to be significant: this gene is expressed when the bacterial cell is exposed to substances (e.g. paraquat) that generate ROS (Ma *et al.*, 1998; Kawai *et al.*, 2015).

This study used GWAS analysis to better understand the desiccation tolerance of gsC as it was the most common genospecies of the Rlc in this dataset. Importantly however, for GWAS analyses to have a meaningful population-wide interpretation the use of hundreds or thousands of genomes is required (Epstein *et al.*, 2018; Farhat *et al.*, 2019; Boivin *et al.*, 2020; Saber and Shapiro, 2020). Moreover, due to the elevated clonality in bacterial populations, linkage disequilibrium and population stratification are strong populational effects that can influence the result of a GWAS analysis and give a positive correlation with a phenotype to a gene that is not involved with it (Chen and Shapiro, 2015). Several methods have been developed for bacteria in an attempt to overcome these population effects such as cluster-based techniques (Chen and Shapiro, 2015), dimensionality reduction methods (Salipante *et al.*, 2015), multi-locus elastic nets (Saber and Shapiro, 2020), and phylogenetic tree-based approaches (Collins and Didelot, 2018; Saund and Snitkin, 2020). To overcome these populational effects, the two methods used in this study are based on phylogenetic convergence, a phylogenetic tree approach, which assesses whether a mutation appears more often in different edges of the tree when the phenotype of interest is present than what is expected by chance (Saund and Snitkin, 2020). All analyses carried out on the Rlc genomes (SNPs PCA, housekeeping gene phylogeny and ANI) have shown that genospecies are well defined blocks that are substantially different from their phylogenetic neighbours, as previously described by Young et al (2021). Thus, using the whole set of genomes would not have been appropriate for this GWAS analysis because it may have led to the identification of significant SNPs that were indicators of genospecies rather than the desiccation tolerance phenotype. Moreover, it seems that the extent of desiccation tolerance within a genospecies varies among genospecies (e.g., gsC and gsR) and the composition of genospecies also varies between countries (e.g., Spain and the UK). This might indicate that the difference in desiccation tolerance is deeply bound to environmental adaptation. Thus, using different genospecies with such genetic and populational differences may be inappropriate for a correct and meaningful interpretation of the results of a GWAS analysis.

The findings in this study support the idea that the Rlc is in fact a group formed by several, genetically differentiated, species. Young *et al.* (2021) hypothesised that many of the Rlc genospecies, if not all of them, would eventually become a separate species within the Rlc due to the clear genetic variation among them. Some of these genospecies are visibly different from one another; however, conventional phylogenetic assessments with one or a few concatenated housekeeping genes do not give sufficient resolution to properly differentiate among all genospecies as seen here in the *atpD-gyrB-recA* concatenated gene phylogenetic analysis. Similar to the results found by Young *et al*. (2021), the phylogenetic analysis of these housekeeping genes struggled to correctly assign a few of the strains used, and failed to give enough support to gsI, O and P, with some of their representatives appearing in distant branches of the tree. On the other hand, genome-wide approaches, like ANI, have been consistent in clearly identifying genospecies (Kumar *et al.*, 2015; Boivin *et al.*, 2020, 2021; Cavassim *et al.*, 2020; Flores-Félix *et al.*, 2020; Young *et al.*, 2021).

In contrast to the separation of strains based on genospecies, the *nodD* types are shared across genospecies. The nodulation genes are located on the Sym-plasmid of Rlc strains (Young *et al.*, 2006), and are frequently exchanged between individuals (Cavassim *et al.*, 2020); these genes confer on the bacteria the ability to nodulate with the legume host which has a preference for a specific nod type (Boivin *et al.*, 2020, 2021) . However, the *nodD* variability was higher in strains isolated from the UK, which showed all types of *nodD*, whilst strains isolated from Spain possessed mainly type B1, apart from one strain with type B2. Type B1 is the most frequent *nodD* type found in strains isolated from the nodules of faba bean, whilst type A1 is frequently found in strains isolated from pea (Boivin *et al.*, 2020). This is surprising given that most of the Spain-isolated genomes were equipped with type B1, despite being trapped using pea. The different range of *nodD* types found between Spain and the UK may be due to either local agronomic practices or sampling effort. Whilst the set of UK strains was comprised of isolates from several different parts of Great Britain (Maluk *et al.*, 2022), the isolates from Spain were isolated from five different fields on the east coast of Spain. Moreover, the cropping history of the Spanish fields from whence these strains were isolated had a recent history of faba bean cultivation (Chapter 2), apart from one of them that was cropping peas at the time of sampling, which might have increased the population of type B1 in these soils. Thus, while the spread of *nodD* types in the set of UK strains may be representative of the entire UK population, it is necessary to consider that the set of strains from Spain is only representative of the east coast of Spain rather than the whole country.

In conclusion, the Rlc is a complex formed by genetically well differentiated genospecies, yet it contains mobile elements that move beyond genospecies boundaries. The strategies used by gsC for withstanding desiccation follow similar pathways to those previously reported in the literature for other rhizobia, where the hyper-concentration of the cytoplasm is used to avert water loss and the likely death of these organisms. However, although the number of genomes used in this study was low, the data suggest that tolerance to desiccation might be bound to genospecies, with some genospecies showing a higher occurrence of tolerant types. Future work in this area should focus on running a GWAS on desiccation tolerance on a larger dataset for each genospecies to discern the intricacies of desiccation tolerance of these Gram-negative bacteria and on confirming the significant genes relation on desiccation tolerance with targeted gene silencing or transcriptomic analyses. Finally, future research should also look at the effect of environmental drivers of population change, e.g., does the history of leguminous cropping significantly modify the populations or the genetic pool of soilborne rhizobia?

Chapter 5 | Quantifying rhizobia concentrations in a six-year crop rotation system: a case study at the JHI Centre for Sustainable Cropping

Abstract

Intensive agriculture is the most common form of arable farmland management in Europe. This management method aims to optimise crop productivity, although often comes with a cost to the environment. With the total human population expected to reach 10 billion by 2055, the demand for food will increase and with it the environmental impact of this agricultural practice. Therefore, a shift towards a more sustainable production system is necessary. The Centre for Sustainable Cropping (CSC) was established at the James Hutton Institute as a whole-system experimental platform aiming to improve environmental outputs whilst maintaining the economic outputs of an arable system. A six-year crop rotation was established at the beginning of the experiment in 2009, with faba bean being used as a leguminous crop within the rotation to provide benefit from its nitrogen fixation capability. Molecular methods were used to quantify the population density and diversity of soil rhizobia, specifically *Rhizobium leguminosarum* (the symbiotic nitrogen fixing organisms associated with faba bean), in the CSC fields over the first six years of the rotation. The main aim of this study was to assess the impact of the two different CSC management regimes, which are ‘conventional’ or ‘integrated’, on this group of soil bacteria. Implementation of the integrated management regime increased the concentration of faba bean-compatible rhizobia in soil by 15 %. Furthermore, the presence of faba bean in the crop rotation produced a change in the rhizobia population diversity, where the proportion of faba bean-nodulating rhizobia (i.e., *R. leguminosarum* sv. viciae) increased to more than 30 % of the total population four years after the legume cropping. In conclusion, the implementation of the integrated management regime increased the soil population density of faba bean-nodulating rhizobia.

Key words

Rhizobia, legumes, faba bean, sustainable agriculture, *Rhizobium leguminosarum* sv. viciae

5.1 Introduction

The current conventional management of agricultural land in Europe is based on intensive agriculture, and it covers almost half of the European Union land area (Henle *et al.*, 2008). The intensification of arable agriculture has allowed an increase in yields yet at a cost to the environment, as this type of agricultural land management often has a negative impact on biodiversity and ecosystem services (Hawes *et al.*, 2016). Therefore, in a world where the human population is increasing year-on-year with a predicted population of almost 10 billion people by 2055 (United Nations, 2019), the demand for farmers to produce more food is going to increase, and with it, a negative impact on the environment. Thus, it is necessary to shift towards more sustainable agricultural approaches to reduce the environmental impact whilst simultaneously optimising agronomical outputs.

Legume-associated nodule-forming rhizobia fix nitrogen (N) into ammonia which is converted by the host plant into proteins and other useful compounds (Howieson and Dilworth, 2016). Some of this fixed N will remain in the soil when the legume senesces and mineralises after its grains are harvested, and will be available for other plants to use (Iannetta *et al.*, 2016; Maluk *et al.*, 2022). Moreover, grain legumes have high nutritional values such as high-fibre, and -protein (Hall *et al.*, 2017; Ferreira *et al.*, 2021) and can represent an important source of income for farmers. Thus, legumes may also play an important role in faciliting more sustainable agriculture by reducing chemical fertiliser input and providing improved economic resilience for farmers.

Studying management modifications in terms of environmental output may result in benefits for some ecosystem services but be detrimental to others (Kleijn and Sutherland, 2003; Kleijn *et al.*, 2006). This highlights the necessity of reporting all positive and negative outputs of a management change, and the use of a multidisciplinary approach for assessing the impact of this change (Carey *et al.*, 2003). In view of this, the Centre for Sustainable Cropping (CSC) was established at the James Hutton Institute (Angus, Scotland) in 2009 as a whole-system experimental platform (Hawes *et al.*, 2016, 2019). The CSC implemented an integrated management strategy, which aimed to improve environmental outputs whilst maintaining economic outputs compared to the conventional intensive agriculture practices in the local area (on a six-year rotation). The CSC incorporates a suite of ecological, environmental, and economic indicators for monitoring the management change outputs (Hawes *et al.*, 2016, 2019). The results from the first rotation of the platform (2011-2016) have already been assessed in different areas such as crop production and quality (Freitag *et al.*, 2018; Hawes *et al.*, 2019), and the different management regimes impact on the economic, ecological, and environmental factors (Hawes *et al.*, 2019) including, weed presence (Hawes *et al.*, 2018), erosional soil organic carbon (SOC) presence on soil microbial biomass (Dungait *et al.*, 2013), and nitrogen fixation (Maluk *et al.*, 2022). Apart from the genetic and symbiotic characterisation of isolated root nodule rhizobia carried out by Maluk et al (2022), and the estimation of SOC on total soil microbial mass (Dungait *et al.*, 2013), the effect of integrated management in the CSC on soil microorganisms has not yet been determined.

Peas (*Pisum sativum* L.) and faba beans (*Vicia sativa* L.) are the two main grain legumes grown in the UK. They only form nodules with rhizobia belonging to the *Rhizobium leguminosarum* species complex, which are equipped with a symbiotic plasmid that has compatible nodulation genes (Kumar *et al.*, 2015; Cavassim *et al.*, 2020; Young *et al.*, 2021). The occurrence and effectiveness of nodulation on peas and faba beans depends on several factors. Firstly, the presence of pea and faba bean rhizobia in the soil is essential; the presence of these rhizobia is ubiquitous in UK soils with population densities between 10^2 – 10^5 per g of soil (Hirsch, 1996; Macdonald *et al.*, 2011; Maluk *et al.*, 2022), which is value-range similar to that found in other countries (Drew *et al.*, 2012; Mothapo *et al.*, 2013). However, soil densities of host-specific pea or faba bean rhizobia do not necessarily correlate with functional performance in terms of N fixation and competition with other strains, as wild strains of rhizobia can compete for nodule occupancy, and their efficiency at fixing N can be variable (Graham, 2008; Boivin *et al.*, 2020; Mendoza-Suárez *et al.*, 2020; Maluk *et al.*, 2022). Thus, the use of inoculants, containing a formulation of a single or a combination of multiple rhizobia selected for enhanced nodulation and N-fixation with a solid or liquid carrier (Howieson and Dilworth, 2016), are an efficient way to supplement the crop with effective compatible rhizobia.

Soil rhizobia population density has traditionally been quantified by the most probable number (MPN) method (Howieson and Dilworth, 2016). This involves growing the host plant in serially diluted soil and comparing the number of subsequent nodules with that of a plant inoculated with known concentrations of a compatible rhizobia. The MPN method has low-technology-capacity requirement , although its accurancy is reliant on the ability of the rhizobia and host plant interacting with each other. However, this interaction may be constrained by the presence of nodulation inhibitors in the soil, such as NO_3 (Macdonald *et al.*, 2011; Reid *et al.*, 2011; Howieson and Dilworth, 2016). Therefore, the use of molecular methods for quantifying total rhizobia population density in soil using DNA may provide a more accurate estimation whilst reducing estimation error by inhibition.

Previous studies have used the MPN method to quantify changes in rhizobial populations in soil due to changing agricultural practices (i.e., fields on continuous fallow or continuous intensive wheat crop) and have mainly focused on studying a particular practice change rather than from a whole-system integrated approaches which the CSC offers (Nutman and Ross, 1970; Nutman and Hearne, 1979; Thies *et al.*, 1995; Hirsch, 1996; Caballero-Mellado and Martinez-Romero, 1999; Gibbs *et al.*, 2006; Chaudri *et al.*, 2008). The fields in the CSC complex have not cultivated peas or faba beans for over 50 years and thus, represent an ideal scenario for studying the effects of management change, together with the incorporation of legumes into the rotation on the soil population structure of rhizobia.

Therefore, the concentration of *Rhizobium leguminosarum* (Rleg) and its symbiovar *R. leguminosarum* bv. viciae (Rlv) in soil was measured at the CSC field complex during the first crop rotation (2011-2016) with the overarching aim of assessing the impact of the management change on these bacteria. For this, five study objectives were designed to assess: (i) the correlation between Rleg and Rlv concentration; (ii) the effect of the crops used in the rotation on these rhizobia; (iii) how the management change affected pea- and faba bean- compatible rhizobia; and (iv) whether the effect of having faba beans inserted in the rotation benefitted the presence of Rleg and Rlv over the period since faba beans were first cropped.

5.2 Materials and methods

5.2.1 The CSC experimental site and treatments

The samples analysed in this chapter originate from the long-term experimental platform Centre for Sustainable Cropping (CSC) at the James Hutton

Institute. A full description of the platform environment is described in Hawes *et al*. (2016, 2018, 2019). In brief, the CSC is a 42 ha whole-system experimental platform designed to assess the impacts of management changes that consider environmental and economic factors and ecological processes. The platform aims to implement an integrated management on an arable system that, while maintaining yield and production quality, improves biodiversity and soil health and reduces environmental pollution.

The CSC is situated at Balruddery Farm, Dundee, Scotland (56°29'03.4"N, 3°07'53.9"W) (Figure 5.1A), with an average annual precipitation of 800 mm and annual minimum and maximum average temperatures between 5 and 12 °C respectively (data from a 30-year average 1971-2000). The soil has an average pH of 5.7 and is classified in the Balrownie series with textures ranging from sandy loam to sandy silt loam (Hawes *et al.*, 2018).

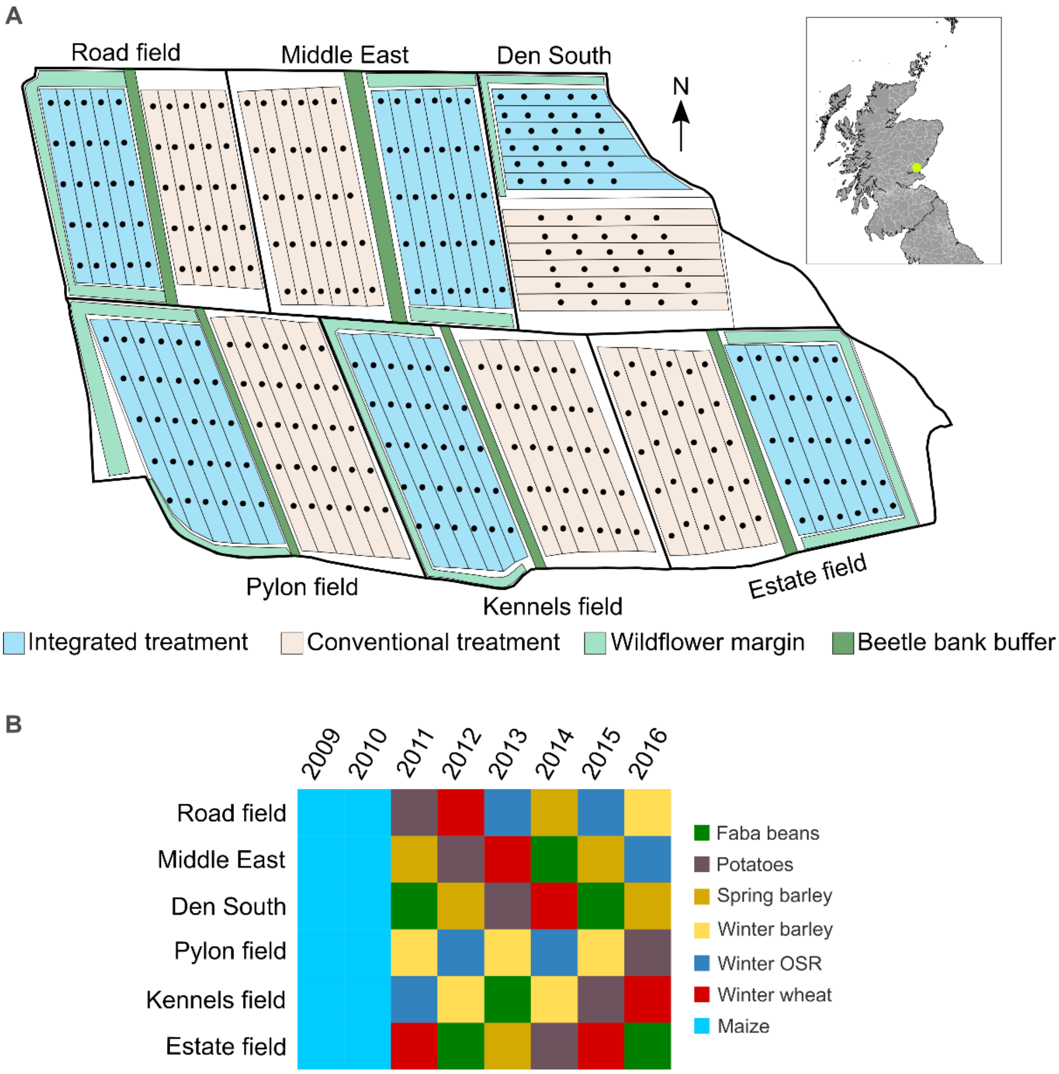


Figure 5.1. Centre for Sustainable Cropping (CSC) at Balruddery farm layout (A) and crop rotation in each field during the first rotation (B). The black dots on A represent fixed GPS locations where soil was sampled annually.

The rotation started in 2011 after two baseline years in 2009 and 2010 where all fields were sown with maize (*Zea mays* L.). The site is composed of six contiguous fields organised in two rows and three columns (Figure 5.1A) with an established six-year rotation of crops normally grown in the wider area: spring faba beans (*Vicia faba* L.), spring and winter barley (*Hordeum vulgare* L.), winter wheat (*Triticum sp.*), winter oil seed rape (*Brassica napus* L.) and potatoes (*Solanum tuberosum* L.) (Figure 5.1B). Each field is divided in half, where either an integrated or conventional management (or treatment) is permanently used in each. The conventional treatment involves a cropping system that mirrors agronomical practices in the area (i.e. inversion soil tillage and use of manufacturer recommended levels of fertilisers, herbicides and crop protectants). In contrast, the integrated treatment encompasses a range of measures such as improving soil structure using non-inversion tillage and reducing erosion by the use of cover crops or incorporation of straw, reducing N-fertiliser use by introducing N-fixing legume crops, cover crops and the application of compost (depending on crop and soil N concentration each year), or increasing the number of pollinators and natural enemies in the system by sowing flower margins with the aim of improving biodiversity and reducing non-renewable inputs whilst conserving yields.

Each half of each field was divided into five or six 18 m wide strips where a different cultivar of the selected crop was sown (Figure 5.1A) and where one of the cultivars was the ‘industry standard’ at the time with desirable traits (i.e., mostly yield and quality). The same selection of cultivars was sown in each field half. Due to harvest and sowing time incompatibilities between spring and winter crops, the rotation had to be altered to optimise sowing time (Hawes *et al.*, 2018) and this resulted in the rotation of faba beans being limited to only four fields (Middle East, Den South, Kennela and Estate) during the first rotation (Figure 5.1B).

5.2.2 Soil sampling

Along each replicate strip, five permanent GPS locations were sampled in March of each year from 2011 before the spring crop was sown (Figure 5.1A). At each location, the soil from an area approx. 20 x 20 cm was mixed to a depth of 20 cm. About 2 L of this soil was passed through a 10 mm sieve, then at least 300 g were passed again through a 2 mm sieve. Some of this was used to fill two 2 mL tubes that were then stored at –80 °C. All remaining soil was kept at 4 °C until used for further analyses (see Appendix 5.2 | Correlation analysis of Rleg and Rlv concentration with soil properties and chemical analyses).

5.2.3 Quantification of rhizobia by qPCR

5.2.3.1 Soil DNA isolation

For the four fields that included faba beans in the rotation, six random soil samples per half of each field for each year were selected from the soil samples stored at -80°C by a random number generator with at least one sample from each strip and from one of the 5 GPS locations within a strip; soil from Pylon field, that had not housed faba bean as a crop, was used as reference (Figure 5.1A). From each soil sample, 0.25 g was added to a bead tube for DNA isolation with the

DNeasy PowerSoil DNA extraction kit (QIAGEN, USA) following the manufacturers recommendations but with the following modifications: an aliquot (10 µL) of a solution with an artificial DNA fragment (spike) of a known concentration was added as an internal standard for assessing the efficiency of DNA extraction (Daniell *et al.*, 2012); and, all centrifugation steps were done at 9000 RCF. After isolation, DNA was stored at -20°C until used.

5.2.3.2 qPCR standard preparation

Gene standards for 16S rRNA and *nodD* were generated from soil DNA extracts. A PCR was run with the soil DNA for amplifying both gene regions using GoTaq® DNA (Promega Cat M7845) as described in Section 2.2.3.3, in a G-Storm GS1 thermal cycler (GRI Ltd, Braintree, UK). The PCR product was subsequently purified in a 2.5 % agarose gel and cloned into *Escherichia coli* DH5 α competent cells (Invitrogen, USA) with pGEM®-T Easy Vector System (Promega, USA). Successfully transformed *E. coli* was grown in Luria-Bertani broth (LB) medium (Bertani, 1951) with 10 mg mL^{-1} of ampicillin; the plasmid was isolated using QIAprep Spin Miniprep Kit (Qiagen, USA) and quantified with the Quant-iT Pico Green dsDNA assay kit (Invitrogen, USA) and diluted accordingly to the appropriate concentration ranging from 10^1 to 10^8 copies µL^{-1}. In order to generate the spike standards, spike-plasmids were extracted from transformed *E. coli* cultures received from (Daniell *et al.*, 2012); and quantified and diluted following the same procedure as above.

5.2.3.3 qPCR reaction preparation and run

For the quantification of the soil rhizobial load, a qPCR method was used. The PCR reaction master mix was prepared following the manufacturer recommendation with the LightCycler® 480 SYBR Green I Master kit (Roche, Switzerland), with the addition of 0.5 µL of 20 mg mL^{-1} Bovine Serum Albumin (Roche) and 1 µL of 10 µM of each forward and reverse primer per reaction (Table 5.1). Following the loading of the master mix into a white LightCycler® 480 Multiwell Plate 96 (Roche), 1 µL of the DNA template was added to each well. The qPCR was run in a LightCycler® 480 II thermal cycler (Roche) following the program detailed in Table 5.2.

Table 5.1. Primers used for the qPCR reactions. Y = T or C, R = A or G, K = T or G, M = A or C, W = A or T, N = any base.

Primer	**Sequence**	**Reference**
MUT-F	5' - CCT ACG GGA GGC ACG TC - 3'	Daniell *et al.* (2012)
MUT-R	5' - ATT ACC GCG GCT GGA CC - 3'	
Rleg 16SqPCR - F979	5' - CCC GGC TAC YTG CAG AGA TG - 3'	Macdonald *et al.* (2011)
Rleg 16SqPCR - R1264	5' - TAG CTC ACA CTC CGC TGC TC - 3'	
nodD viciae qPCR - F88	5' - TGC AGA GAC GGG AGC TAR TTC - 3'	
nodD viciae qPCR - R443	5' - GAC GCA CAC CAG TCT CTC TTC G - 3'	

5.2.3.4 LightCycler raw data processing, from crossing points (CP) to number of copies

Crossing points (CP) and melting curve data were calculated with the LightCycler 480 Software (v1.5) from raw data generated during the qPCR run. Individual reaction efficiencies were calculated with LinRegPCR v2020.0 (Ruijter *et al.*, 2009) using the qPCR fluorescence raw data.

For each plate, a regression line was calculated for spike, 16S rRNA and *nodD* standards, and the copy numbers per well were interpolated using the calculated CP and the regression line intercept and slope; and efficiency for each regression line was calculated using Equation 5.1

$$Efficiency\ (E) = 10^{(-1/slope)} \qquad \text{(Eq. 5.1)}$$

For the quantification of the spike DNA, the copy number was corrected for the individual reaction efficiency and the spike correction factor (S_{cf}) was calculated by dividing the estimated number of copies in the well by the expected number of copies in 1 µL of DNA (2×10^7 copies μL^{-1}).

Table 5.2. qPCR program used for all amplified regions.

Process	Step	Temperature	Duration	Cycles
Denaturation	Denaturation	95ºC	15 minutes	1
PCR	Denaturation Annealing Elongation Quantification	94ºC 58ºC 72ºC 81ºC	20 seconds 30 seconds 30 seconds 5 seconds	42
Melting curve	Denaturation Annealing Denaturation	95ºC 55ºC 95ºC	10 seconds 15 seconds ∞	1

For the quantification of 16S rRNA and *nodD*, after the individual reaction efficiency correction had been calculated, the concentration was divided by the S_{cf} to account for number of copies lost during DNA isolation. Following this, gene copy number in the *Rhizobium leguminosarum* genome were accounted for by transforming the copy number to number of bacteria. Accordingly, 16S rRNA copies were divided by 3, whereas for *nodD* they remained the same as only one copy of this gene is usually found in Rlv (Macdonald *et al.*, 2011). Finally, the number of bacteria was corrected for the total DNA extraction volume (50 µL) and divided by the moisture corrected soil weight to obtain the final number of bacteria per gram of dry soil, which was used for all subsequent statistical analyses.

5.2.4 Data analysis

For data analysis RStudio v 1.2.50001 (Boston, USA) was used implementing R software v 3.6.1 (R Core Team, 2019) and package dplyr v 0.8.3 (Wickham *et al.*, 2019) for data handling and processing and ggplot2 v 3.3.0 (Wickham, 2016) for data visualisation. For statistical analysis, linear mixed effects models followed by Least Significant Difference (LSD) and adjusting p-values by BH method (Benjamini and Hochberg, 1995) to reduce false positives were used with packages lme4 v1.1.21 (Bates *et al.*, 2015), car v 3.0.6 (Fox and Weisberg, 2019) and predictmeans v 1.0.4 (Luo *et al.*, 2020).

For the CSC linear mixed effects models, the concentrations of Rleg and Rlv and the proportion of Rlv:Rleg (ratio) were log-transformed, and the two main models were tested on each dependent variable: in the first model, concentration was assessed as explained by the field management and the previous crop in the rotation and any interaction between them; in the second model, concentration was assessed as explained by the time since the last legume crop. Both models had the variability between year, field, field half and technical replicate accounted for as a nested random effect on each model (i.e., Quantification~Management*Previous.Crop + (1|Year/Field/Half/Sample)).

5.3 Results

Integrated management had a beneficial effect on Rlv concentration after the 6-year period but it did not influence the concentration of Rleg. The concentration of Rlv in soil was increased significantly by an average of 14.75 % compared to the conventionally managed halves ($P < 0.01$) (Figure 5.2 B). In contrast, the concentration of Rleg remained similar in both field-halves (Figure 5.2A). Despite the increased concentration of Rlv in the integrated halves, the relative proportion of Rlv to Rleg remained the same under both management treatments at an average of c. 15 % ($P = 0.496$) (Figure 5.2C).

Each field had a different sequence of crops during the first rotation (Figure 5.1B). Thus, the concentrations of Rleg and Rlv were also compared between fields (Figure 5.2 D-E). Although one of the fields did not have any legumes cropped in it, the other four did have legumes sown at least once in the rotation. These fields showed a variation in the Rleg density with significant differences between the Rleg density of the legume naïve field (Pylon) and some of the other fields ($P < 0.01$) (Figure 5.2D). The highest concentration of Rleg was found in the legume-naïve field which had 1.39×10^6 Rleg g^{-1} soil dw. The Rleg density in this field was similar to that of Kennels field, which had legumes in the third year, but significantly higher than the concentration in the remaining three fields (Figure 5.2D). The concentration of Rlv differed significantly between fields ($P < 0.01$). In this case, the only field that had two previous harvests of faba beans also had the lowest densities of Rlv at 1.53×10^5 Rlv g^{-1} soil dw ($P < 0.05$). This concentration was 20.68 % lower than the Rlv concentration found in Estate field which had one of the highest concentrations of Rlv despite faba beans only being grown once during the second year of the rotation (Figure 5.2E).

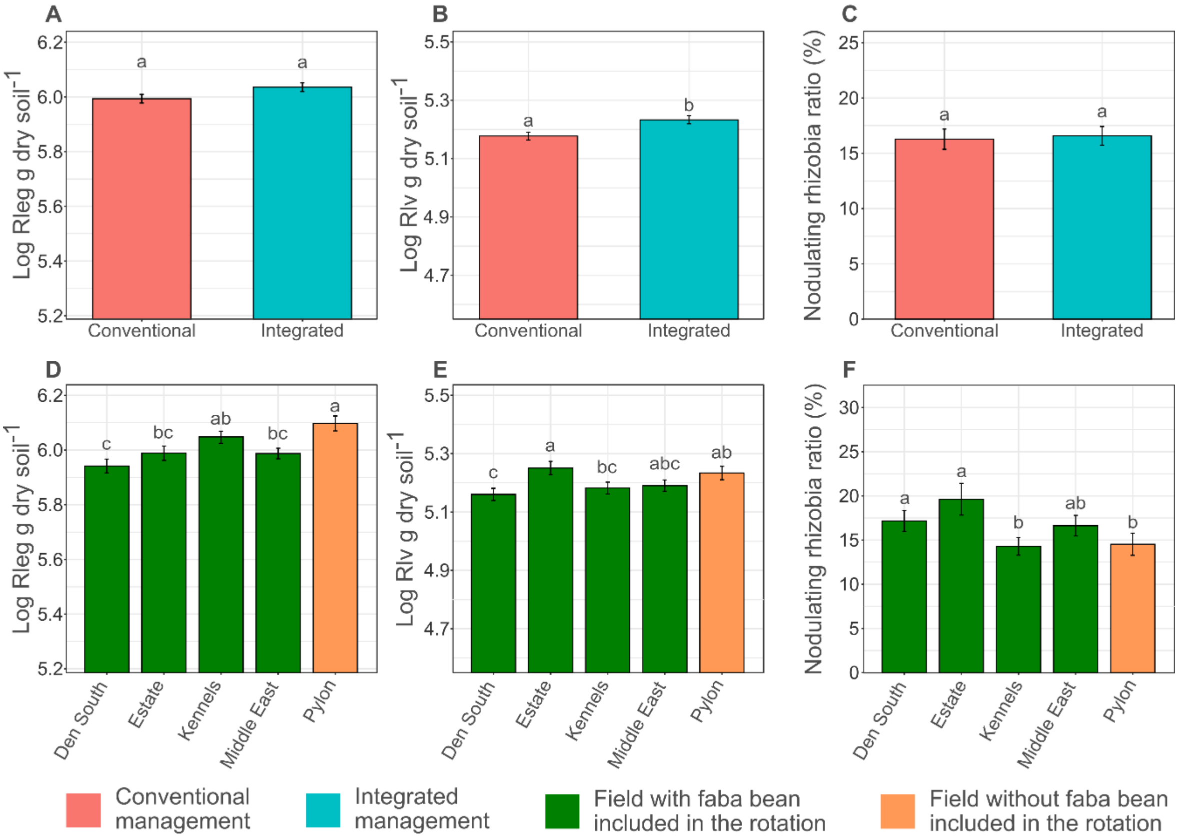


Figure 5.2. Bar plots with standard error bars of the effect of the insertion of legumes within the crop rotation at the management (A-C) and field (D-E) level. Same letters on top of each bar within a plot indicate that no statistical difference is found when bars were compared pairwise.

The ratio between Rlv and Rleg concentrations (Figure 5.2 F) was influenced by the addition of legumes into the rotation ($P < 0.01$). The pairwise comparison between fields showed two main groups, the first group (Den South and Estate field), showed the highest proportion of faba bean-nodulating rhizobia within the total *R. leguminosarum* population, with a proportion of Sym-plasmid-equipped Rleg of 16.15 and 17.82 % respectively. The second group (Pylon and Kennels fields) showed the lowest proportion of nodulating Rleg with 13.36 and 13.54 % of Sym-plasmid-equipped Rleg, respectively. The remaining field did not show a significant difference between either of the previous groups with 15.98 % of the Rleg population having presence of the Sym-plasmid.

Overall, the concentrations of Rleg and Rlv showed a strong positive correlation on both conventional (Pearson's r = 0.599, $P < 0.001$) and integrated

(Pearson's r = 0.701, $P < 0.001$) field halves (Figure 5.3). The average concentration of Rleg and Rlv was 1.14 x 10^6 and 1.72 x 10^5 individuals g^{-1} soil dw, respectively.

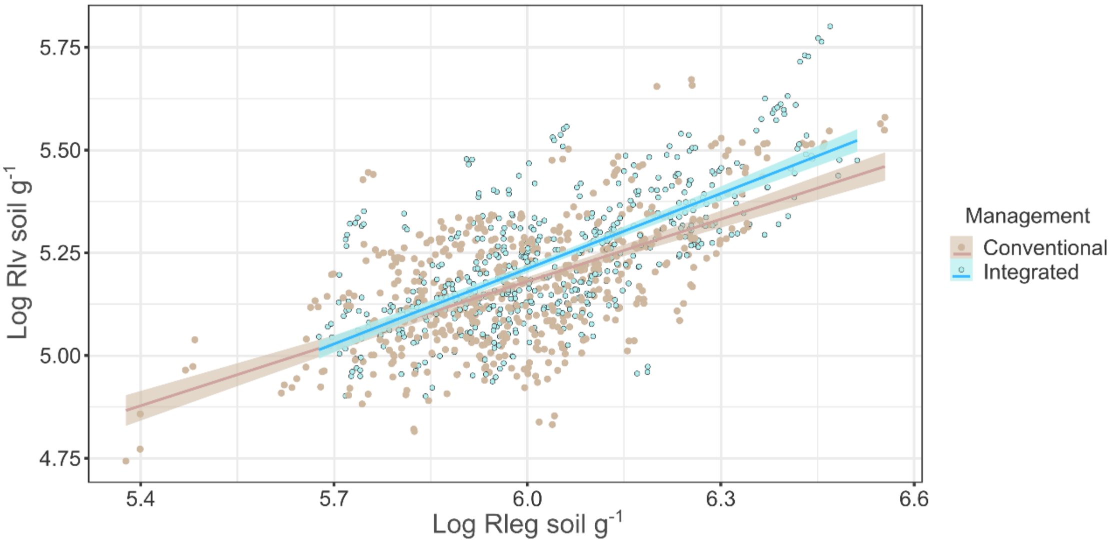


Figure 5.3. Pearson's correlation analysis for Rleg (16S rRNA) and Rlv (*nodD*) per gram of dry soil in each management. The blue and brown lines represent the linear correlation between Rleg and Rlv on both conventional and integrated managements. The shaded area of the same colour shows the confidence interval for each correlation.

The addition of legumes into the rotation produced a change in the composition of the faba bean-nodulating rhizobia population, shifting it towards an increased proportion of Sym-plasmid-equipped Rleg (Figure 5.4). The addition of the legume crop into the rotation also correlated with a decline in the concentration of Rleg over time ($P < 0.05$). This decline was observed for both conventional and integrated management, with a reduction in Rleg of 25.82 and 11.25 %, respectively, since the faba beans had been cropped. Despite this, there was insufficient statistical evidence to support a different rate of reduction of Rleg concentration over time between both management practices (Figure 5.4 A and B). The time elapsed since the legume crop was grown, however, did not have a significant effect on the concentration of Rlv in the soil, and only the effect of management on the concentration of Rlv was evident for this population ($P < 0.05$) (Figures 5.5 C and D). Consequently, the ratio between Rlv and Rleg concentrations showed a strong positive trend over time ($P < 0.001$) (Figures 5.5 E and F) where the soil concentration of *R. leguminosarum* individuals with a Sym-plasmid increased by 35.38 % and 32.35 % in the conventional and integrated halves, respectively.

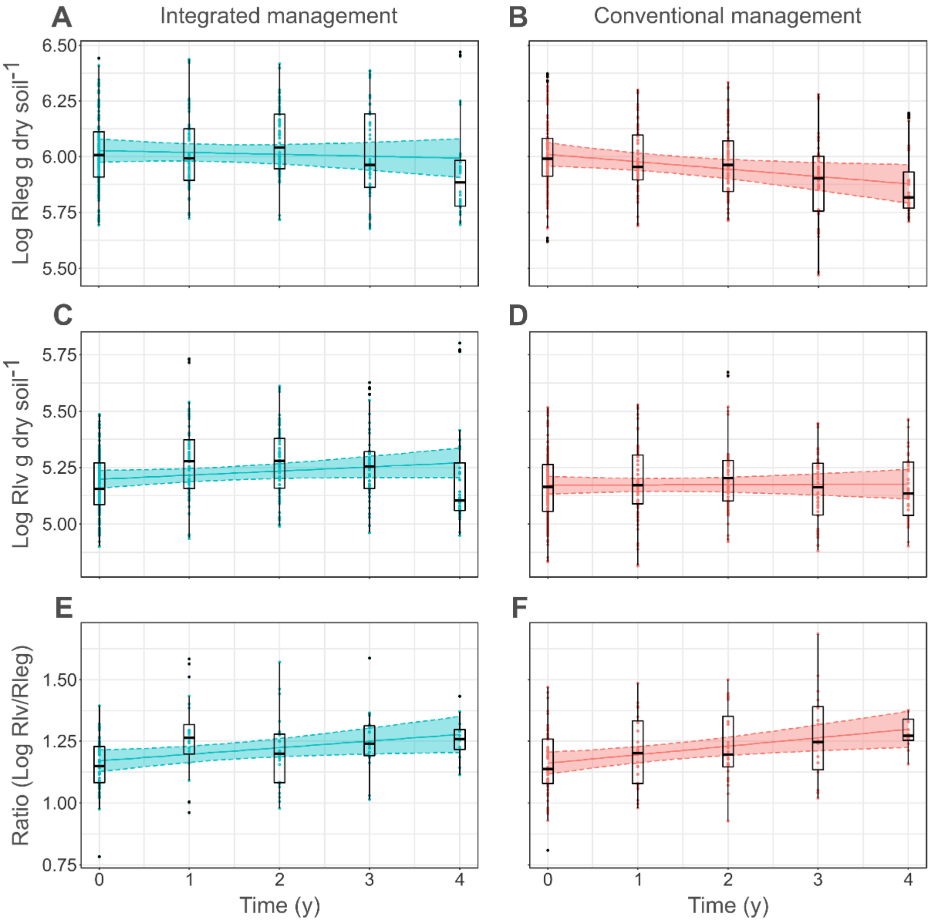


Figure 5.4. Scatter plot illustrating Rleg (A and B), Rlv (C and D) and ratio (E and F) quantification dynamics over time since the insertion of legumes in the rotation. The solid line represents the model estimated direction of the quantification over time and the shaded area on both sides of the line represents a 95% confidence interval. The boxplots at each timepoint show the distribution of quantifications at each timepoint.

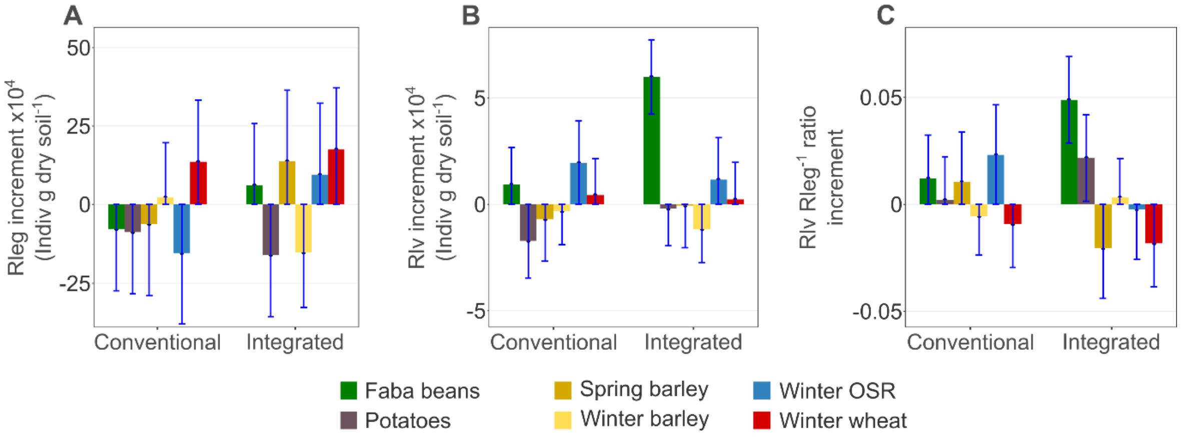


Figure 5.5. Bar plot with standard error bars of Rleg (A), Rlv (B) and ratio (C) increments before and after plot. There were not statistical differences between the means of each crop, managements or their intersection.

These long-term results contrasted with those obtained when the concentrations of Rleg and Rlv were compared before and after each year. There were no significant differences in Rleg or Rlv concentrations in the soil after the crop had been harvested compared to before the crop was sown, and the ratio between the two measurements remained the same.

5.4 Discussion

This study has shown that an integrated arable management system can have a positive effect on populations of Rleg equipped with the sv. viciae nodulation plasmid (i.e., Rlv), compared to more conventional management systems. However, the type of management system does not have a similar effect on the concentrations of Rleg populations. The addition of faba beans into crop rotations can lead to a decline in the total Rleg population but not Rlv populations, which results in an increase in the proportion of the Rleg population capable of nodulating faba beans.

The high correlation between the densities of Rleg and Rlv indicates that *nodD* is mainly found in *Rhizobium* species (and other currently recognised legume-nodulating rhizobial types; Peix *et al*., 2015), although there have been recent reports of horizontal gene transfer (HGT) between *Rhizobium* and other bacterial genera (i.e., *Agrobacterium*) in the wild that are not normally associated with nodulation (Delamuta *et al.*, 2020; Youseif *et al.*, 2021). The *Rhizobium leguminosarum* clade (Rlc) is comprised of eighteen genetically distinct genospecies (Kumar *et al.*, 2015; Young *et al.*, 2021). The distinction between genospecies is based on the phylogenetic analysis of housekeeping genes and the sym-plasmids are not bound to a given genospecies giving support to the likelihood of HGT among genospecies (Kumar *et al.*, 2015; Cavassim *et al.*, 2020; Youseif *et al.*, 2021). HGT is a relatively common adaptation mechanism in rhizobia whereby some individuals incorporate genetic material from another individual (Andrews *et al.*, 2018). Notwithstanding the widespread presence of sym-plasmids across the eighteen genospecies, it seems to be less common to find HGT events between different bacterial genera. In addition, Maluk *et al*. (2022) isolated and characterised root nodule bacteria harvested from faba bean crops sown at the CSC experimental platform and their study demonstrated that, based on a phylogenetic analysis of the 16S rRNA, *recA* and *atpD* genes, all isolated strains were *Rhizobium leguminosarum*. Thus, we can probably disregard the possibility that *nodD* was frequently present in many different genera at the CSC soil, as there would have been little or no correlation between both genes identifying Rleg (16S rRNA) and Rlv (*nodD*).

The Rleg population density in the CSC soil was ubiquitous across all fields but was consistently higher than Rlv (Macdonald *et al.*, 2011; Mauchline *et al.*, 2018) (Table S5.4). There are several factors that can influence this, for example, *R. leguminosarum* has three main symbiovars (sv), sv *viciae* (Rlv), sv *trifolii* (Rlt) and sv *phaseoli* (Rlp), each one with a characteristic sym-plasmid that enables them to nodulate with plants in the genera *Pisum*, *Vicia*, *Lathyrus* and *Lens (Rlv)*, or, *Trifolium* (Rlt) or *Phaseolus* (Rlp) (Dilworth *et al.*, 2008). However, it is likely that Rleg primers do not distinguish between the different symbiovars, which will give an overall population estimate regardless of the sym-plasmid in the bacterial genome (Macdonald *et al.*, 2011). The 16S rRNA primers can also amplify other *Rhizobium* species (Macdonald *et al.*, 2011) (Appendix 5.1 | On the qPCR efficiency and melting curve assessment); and as the boundaries between *Rhizobium* species, in particular those belonging to the Rlc, are currently under scrutiny due to the recent increase of available full genomes from around the world many of the *Rhizobium* genospecies might be elevated to species status once more genomes become available (Kumar *et al.*, 2015; Young *et al.*, 2021). Nevertheless, higher Rleg populations densities were also found when either metagenomic and metatranscriptomic approaches were used for the quantification of *R. leguminosarum* and its sv *trifolii* (Mauchline *et al.*, 2018). *Rhizobium leguminosarum* is a successful soilborne bacteria which can thrive saprophytically in soil; this is demonstrated by its complicated genome which is comprised of a chromosome and many accessory genes organised in plasmids that allow them to metabolise many different types of substrate (Young *et al.*, 2006), and the capability of strains without sym-plasmids to survive and establish in a recently colonised soil (Clark *et al.*, 2002). Therefore, there could be potentially as many different Rleg individuals as there are different microhabitats found in soil, each one exploiting their characteristic micro-niche for which they might or might not need nodulation and nitrogen fixation genes.

The Rlv population densities also gave an estimation of the total Rleg density which are capable of triggering the nodulation process, yet this is not an absolute nodulation estimation *per se* but rather an indicator of 'nodulation potential' . This is because nodulation is a complex process with many genes playing their role in both the bacterium and the host plant (Sessitsch *et al.*, 2002; Ferguson *et al.*, 2010, 2019; Ryu *et al.*, 2012; Howieson and Dilworth, 2016) together with some soil characteristics e.g., pH and N levels, which can also play a key role in rhizobial growth and nodulation inhibition (Hirsch, 1996; Graham, 2008). A correlation for this was found between soil inorganic N concentration, %N in soil, and pH with the Rlv Rleg^{-1} ratio (Appendix 5.2 **|** Correlation analysis of Rleg and Rlv concentration with soil properties and chemical analyses). As with Rleg, Rlv was abundant at the CSC and its concentration is similar to that measured in other areas of the UK and around the world (Nutman and Ross, 1970; Catroux and Amarger, 1992; Hirsch, 1996; Mothapo *et al.*, 2013). Consequently, the ratio calculated between Rlv and Rleg is a measurement that provides information on the proportion of the total Rleg population that is equipped with the sv. viciae nodulation genes.

At the CSC, the Rlv population represents about one sixth of the total Rleg population which is similar to that found for sv *trifolii* (Jarvis *et al.*, 1989) and almost eight times higher than that of sv *phaseoli* (Segovia *et al.*, 1991) using the MPN method. MPN and molecular methods such as the one used in this study are positively correlated, but Rlt is often the most common Rleg biovar found in soils, and can represent more than 70% of the overall population (Macdonald *et al.*, 2011; Mauchline *et al.*, 2018). Despite this, the ratio of Rlv found at the CSC was several-fold higher compared to that in the control soil analysed by Macdonald et al (2011) which had a percentage of Rlv in the Rleg population of about 2.12 %. The absence of the legume host in the field is known to reduce the numbers of nodulating rhizobia over time (Nutman and Ross, 1970; Nutman and Hearne, 1979), and the soil analysed by Macdonald *et al.* (2011) had been for 17 years with a permanent grass cover (Gibbs *et al.*, 2006). This might in part, explain the low Rlv Rleg^{-1} ratio Macdonald *et al.* (2011) found. However, this explanation is contrary to ratios found at the CSC fields, because they have a known cropping history without Rlv host legumes sown since the 1960s, yet higher population densities of Rleg and Rlv were found. However, it is possible that the pedoclimatic conditions at the CSC favour saprophytic Rlv forms to thrive, as some soils are capable of holding significant rhizobial population densities regardless of the presence of any particular legume crop (Hirsch, 1996). Furthermore, the CSC fields are located in a prolific farming area in Easter Scotland where leguminous crops are often grown in nearby fields. The dust arisingfrom agricultural practices, such as during drilling or combining, in adjacent or nearby fields may contain significant numbers of Rlv (Parker *et al.*, 1977) which may recharge the bacterial pools every season with a constant influx of rhizobia. In addition, wild Rlv legume hosts such as *Lathyrus* and *Vicia* species are present in the field margins at the CSC and may have acted as a primary inoculum of this rhizobia and helped to maintain the population (Maluk *et al.*, 2022). However, the wild legume types occur only rarely withn arable fields across the UK, and within the CSC fields only a few *Trifolium* and *Vicia* volunteers had been recorded (Hawes *et al.*, 2018).

The annual effect of the different crops sown in the CSC rotation did not show any significant results, though differences were observed among crops. The CSC is a long-term experimental platform, and its second full rotation is scheduled to finish in 2022. Consequently, the addition of data from further rotation cycles to this dataset will add additional replicates which may modify this result for the effect of different crops on the concentration of Rleg and Rlv. As part of the integrated management, soil-disturbing arable practices such as inversion tillage were not implemented, and ploughing was reduced to one application every six years for the potato crop. Additionally, a cover crop was grown over-winter (Hawes *et al.*, 2018). Thus, the increases in Rleg often found in the integrated field halves may be a result of reducing these arable practices which are known to negatively affect rhizobia population after continuous fallow or intensive cereal farming (Nutman and Hearne, 1979). Furthermore, the integrated management showed a significant increase in Rlv concentration over the six-year period compared with the conventionally managed field halves. Apart from low soil disturbance measures, the integrated management also implemented compost amendments and straw incorporation which have shown to increase soil pH and organic matter content (OM) content (Hawes *et al.*, 2018). These soil parameters have previously been linked to the stability of microbial communities in soils (Rousk *et al.*, 2010; Griffiths and Philippot, 2013) and show a positive correlation between Rlv population densities, and the Rlv Rleg^{-1} ratio (Table S5.3). Furthermore, the fact that significant differences between integrated and conventional management were found for Rlv over the six-year rotation but not after each individual crop, reflects the disruption of the previous systemic processes that the implementation of a different management system has had on the field. In other words, it is likely that the system is reaching a new equilibrium wherein Rlv population density is higher in the soils due to differing arable practices. This disruption to the *status quo* of the arable system and subsequent re-normalising to a new equilibrium has also been shown to affect plant productivity and composition (Freitag *et al.*, 2018). Conversely, the use of fertilisers and phytosanitary products can reduce the genetic diversity of the *Rhizobium* population too (Nutman and Hearne, 1979; Caballero-Mellado and Martinez-Romero, 1999; Ahemad and Khan, 2013). So, the increases in Rlv in the field halves with integrated management, regardless of the lack of different concentrations of Rleg, may be evidence of this change in genetic diversity even though it is not shown in the Rlv Rleg^{-1} ratio measurements. Moreover, significant differences were found amongst fields for all three parameters, but rather than observing a low number of both Rleg and Rlv under a crop rotation without legumes as previously reported for intensive cereal farming (Nutman and Hearne, 1979), or an increase in Rlv after the legume crop (Kucey and Hynes, 1989; Hirsch and Spokes, 1994), the legume-naïve field showed one of the largest concentrations for both genes, and it was, in fact, Den South field that had the lowest numbers despite it being sown with faba beans twice during the first rotation. In contrast, the Rlv Rleg^{-1} ratio amongst fields showed a population change between those which had faba beans and the field that had not, and it is in this case where Den South and Estate fields show larger proportions of Rlv per total population of Rleg.

This evidence indicates that the insertion of legumes in the crop rotation may have had a higher impact on the equilibrium between Rleg and its nodulating counterpart, Rlv, than other changes in arable practices. In fact, there were no differences between the two types of management since faba beans were included the rotation, apart from the consistently higher population densities of Rlv in the integrated field halves. Consequently, the insertion of legumes into the rotation did have an impact on the total Rleg population which declined over time, and this, coupled with a constant Rlv population, resulted in an overall increase of the proportion of nodulating rhizobia. This findings concurs with Herold *et al.* (2018) who reported a decrease in total 16S rRNA concentration in soil after the field had been recently rotated to ley grass, and it is also in accordance with Hirsch (1996) who consistently found a constant 3-fold increase in nodulating bacteria over a five year period since peas were grown. Therefore, the results indicate that there is a population composition change from the moment that legumes are introduced into the rotation, and this change is maintained over (at least) a 4-year period. However, rather than this being due to an increase in the number of Rlv in soil, it is the result of a decrease in those Rleg that are not symbiotically active. This increases the proportion of nodulating rhizobia, augmenting the probability that the next legume crop finds compatible rhizobia that will nodulate and fix nitrogen. It is plausible then that this population shift may be driven by an ’infection and release’ effect (Provorov and Vorobyov, 2000) after the crop has been harvested. This could be produced by nodules from dying roots becoming incorporated into the soil, and the release of strains of rhizobia equipped with a sym-plasmid which have multiplied asexually inside the nodule from the original legume-nodulating rhizobia variant that first entered the host root (Thies *et al.*, 1995; Provorov and Vorobyov, 2000). This, coupled with a competitive advantage of these nodule inhabitants which enables them to use new carbon sources present in the soil, which has been produced and released by the legume roots as exudates, for example homoserine, which only Sym-plasmid-equipped Rlv can catabolise (van Egeraat, 1975; Hirsch, 1996; Graham, 2008). Thus, it seems that after the legume crop has been harvested, the soil is flooded with a fresh batch of legume-compatible rhizobia that are conferred with various advantages which select against the persistence of Sym-plasmid-less variants, sweeping away previous genetic variation as previously hypothesised by Kumar *et al.* (2015). This is a known ecological process called ‘periodic selection’, and which has been reported for other bacterial species too (Cohan, 2002; Wiedenbeck and Cohan, 2011; Kopac *et al.*, 2014). Numerous strains of rhizobia have been isolated from root nodules at the CSC, which have shown a wide range of symbiotic efficiency (N-fixation and plant growth promotion capabilities) when compared to uninoculated control plants in N-free glasshouse experiments (Maluk *et al.*, 2022). This suggests that efficiency is not necessarily linked with the ability to nodulate effectively with the host plant. Thus, despite having a natural and able population of compatible rhizobia in soil, inoculation of a legume crop with an "elite" strain (or strains) selected for successful in-soil persistence, competitiveness to nodulate and that exhibits high levels of nodule occupancy, N-fixation, and plant growth promotion might be a good strategy to increase the proportion of effective rhizobia to significantly enhance crop yield and biomass.

In conclusion, implementing an integrated management approach like the one demonstrated at the CSC has proven to benefit the number of beneficial nodulating rhizobia in soil. The incorporation of faba beans in the rotation produced a rhizobia population composition shift in the years following the legume crop, increasing the proportion of Sym-plasmid-equipped *R. leguminosarum*. The efficiency of the population is uncertain in terms of N-fixation and plant growth promotion and a wide range of efficiencies are found in natural populations. It is therefore suggested that future research in this area should examine the effect of inoculation as a driver of this population change - towards a rhizobial population compositions with a higher proportion of "elite" strains, and the long-term impact of this on improving crop yield.

Chapter 6 | General discussion

6.1 Overall outcomes of the project

The main aim of this project was to isolate and characterise novel rhizobia strains tolerant to desiccation for their use as inoculants with improved shelf-life for peas and faba beans, and with a view to optimising crop, growth, fitness, nodulation, and biological nitrogen fixation (BNF). For this, rhizobia were isolated from a semi-arid environment and tested for their ability to affect plant performance and compared to commercial and high-performing standard strains. Experiments were performed in glasshouse and field environments, and *in vitro* to assess their natural capacity to withstand desiccation, the main cause of a reduction in inoculant shelf-life. This research has successfully isolated strains that have a similar effect on plant growth promotion to standard commercial strains in both field and glasshouse trials. The novel strains isolated and reported here offer the additional advantage of having an improved tolerance to desiccation. These novel strains therefore have the potential to extend inoculant shelf-life and provide both manufacturers and farmers with a product that retains its efficacy for longer periods, thus allowing longer distribution distances and storage times without reducing in-field effectiveness. Finally, this project quantified natural rhizobial populations within different crops and cropping systems over several consecutive seasons. This revealed that pea-nodulating rhizobia are successful saprophytes and can thrive in soil in the absence of the host, although following the addition of a compatible legume crop, concentrations of host-specific rhizobia (i.e. with the nodulating capacity) were elevated.

6.2 Pea and faba bean rhizobial genospecies

In the last decade, high throughput sequencing technologies like Illumina (Illumina Inc, USA) and Oxford Nanopore (Oxford Nanopore Technologies, UK) have reduced the cost of sequencing whilst increasing the quantity, quality, and length of the sequences. This has revolutionised molecular systematics allowing the sequencing and use of whole genomes, instead of individual housekeeping genes, and has led to an increased understanding of the relationship between pea, faba bean, and their symbiotic rhizobia.

Molecular analyses of several housekeeping genes and DNA-DNA hybridisation experiments allowed the characterisation and differentiation of genetically similar species within rhizobia (Saïdi *et al.*, 2014; Jiao *et al.*, 2015). With the advent of high-throughput sequencing technologies and the use of whole genomes to assess similarity among rhizobial species, further differentiation of very similar species has been made possible. This has been the case for *R. leguminosarum* which once was considered a single species with different symbiovars, but is now considered to be a complex of genetically well-defined organisms or genospecies with shared parts of the genome in the form of plasmids, now known as the *Rhizobium leguminosarum* species complex (Rlc) (Mutch and Young, 2004; Boivin *et al.*, 2020, 2021; Jorrin *et al.*, 2020; Young *et al.*, 2021).

During the different genotypic and functional characterisations of the rhizobia used in the project (Chapter 2 - 4) there was evidence of strain differentiation (i.e., growth speeds, BOX patterns, and/or tolerance to desiccation). The findings in Chapters 2 and 3 indicated that the group of strains studied was not homogeneous, and following whole genome sequencing of many of these strains (Chapter 4), further genotypic evidence supporting this heterogeneity was revealed. In accordance with the current genospecies divisions proposed by Kumar *et al.* (2015) and recently extended by Young *et al.* (2021), the sequenced genomes fell within 10 of the 18 Rlc genospecies with clear boundaries set at the 96 % ANI threshold (Figure 4.1A); these were clearly delimited by the principal components analysis of SNPs which were identified on orthologous genes among these strains (Figure 4.4).

Collating all these data, it seems that genospecies comprise distinct desiccation tolerant and growth speed types (Figure 6.1). Significantly, the UK and Spain strains did not have any genospecies in common (Figure 4.1B), and so it was expected that functional and genetic differences between genospecies would be found from each locality. For instance, strains from genospecies C (a UK genospecies) are mainly fast growers with low tolerance to desiccation, while strains from genospecies R (a Spanish genospecies) are generally slow growers with high tolerance to desiccation. These two genospecies have among their members the type strains for two recently Rlc-excised species: *R. laguerreae* FB206 (gsR; Saïdi *et al.*, 2014) and *R. ruizarguesonis* UPM1133 (gsC; Jorrin *et al.*, 2020) isolated from faba bean and pea, respectively. Thus, it is likely that speciation has driven the differences between Spain and UK strains of rhizobia, in addition to them being genetically distant organisms. However, as discussed in chapter 3, the classification between fast and slow growers is inherent to the studied dataset and the threshold separating both groups is likely to change in a different dataset. Therefore, additional sampling and growth curve parameters assessment would be necessary, particularly within genospecies, to further understand phenotypical differences among these genospecies.

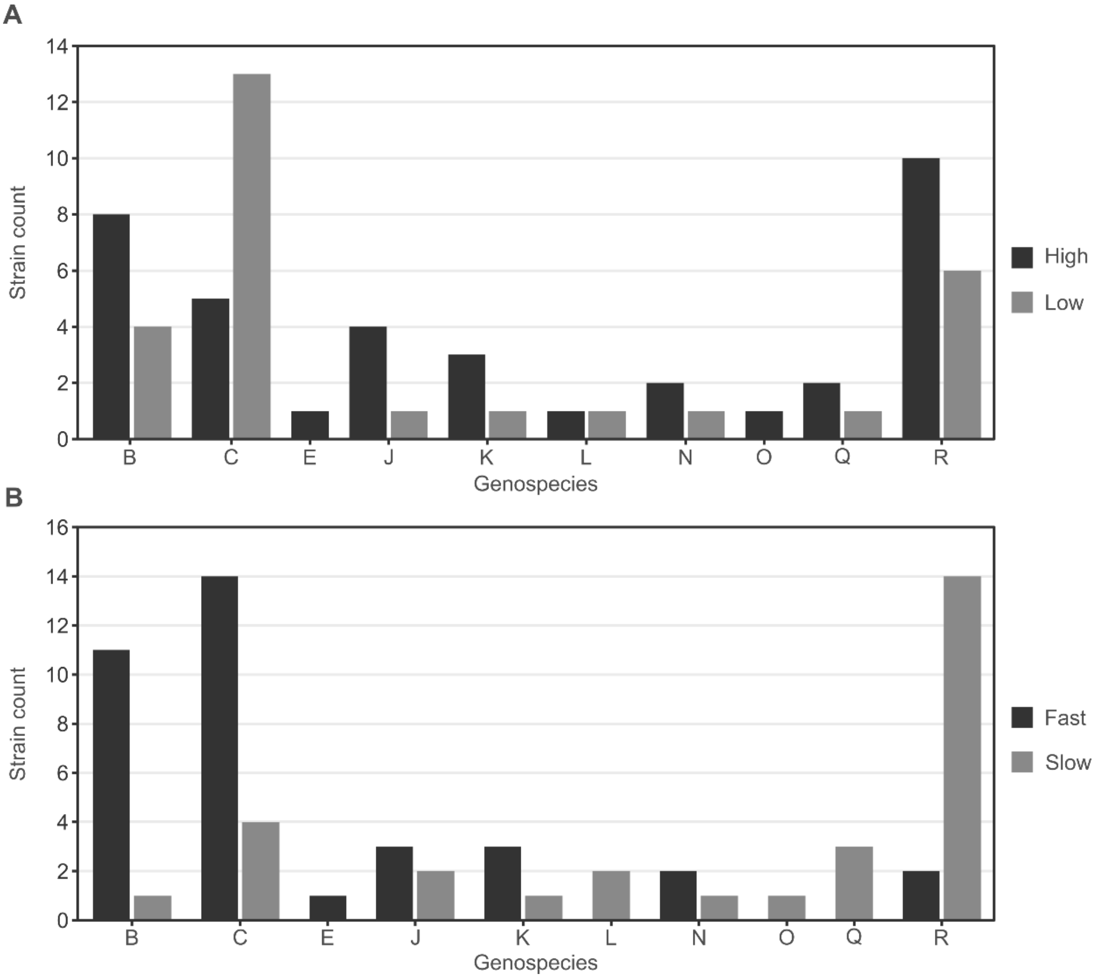


Figure 6.1. Desiccation tolerance (A) and growth speed (B) composition of Rlc genospecies.

The differences of genospecies composition between Spain and the UK could correlate with a north-south division in Europe where some genospecies might be more common in certain regions to which they have become adapted. A similar suggestion was made by Cavassim *et al.* (2020) regarding genospecies A-E, whereby the authors considered them as, "*likely* [to] *represent a large part of northern European* R. leguminosarum *diversity*". Most of the strains in genospecies A-E identified in this project were isolated from the UK only, i.e. none came from

Spain. Although two gsC strains were isolated from Greece and Ethiopia and two gsE strains were isolated from the USA.

Recent publications have assessed the genospecies of hundreds of Rlc genomes (Boivin *et al.*, 2020, 2021; Cavassim *et al.*, 2020; Young *et al.*, 2021) and these are available on public databases. Genospecies C is the most common genospecies where 37.47 % of the total number of genomes assessed belonged to this group (Table 6.1). Europe is the continent with the most genomes available, with Denmark, the UK, France, and Spain representing 71.84 % of the total number of genomes assessed.

Peas and faba beans were domesticated in the Middle East from where they were dispersed into Europe, northern Africa, Asia, and finally globally (Zohary *et al.*, 2012; Kosterin, 2014). Wild relatives like vetches (*Vicia* spp.), vetchlings (*Lathyrus* spp.), and wild peas (*Pisum* spp.) are common European flora. Thus, it can be expected that compatible rhizobia were already present in the soil, likely easing the ancestral dispersion of pea and faba bean in this continent. However, these plants are not native to America or Oceania wherein peas and faba beans were introduced by European explorers, and hence their cultivation there is only a few hundred years old. Consequently, inoculation in such locations is often recommended as the soil does not harbour native populations of compatible rhizobia (Matthews and Marcellos, 2003; Bing, 2015; Giller *et al.*, 2016).

The global genospecies distribution (Figure 6.2) shows that the dominant genospecies varies depending on the country. While in France, Denmark, and Italy the most common genospecies encountered is genospecies C, the most frequent in the UK and Spain are genospecies B and R, respectively. Furthermore, genospecies A is only found in European, Oceanic and American countries providing evidence of the dispersion of the legume host from Europe to the two latter continents, either as a passenger on pea and faba bean seeds, or via the application of ‘inoculants’ and subsequent naturalisation. A similar explanation may be given for the presence of genospecies H and M; they are naturally present only in Greece and Spain, but also in Australia. It is likely that their similar environmental conditions (derived from their Mediterranean-type climates), are likely to have facilitated the importation and naturalisation of Greek and Spanish strains for their use as inoculants in Australian pea and faba bean fields. This is the case of genospecies

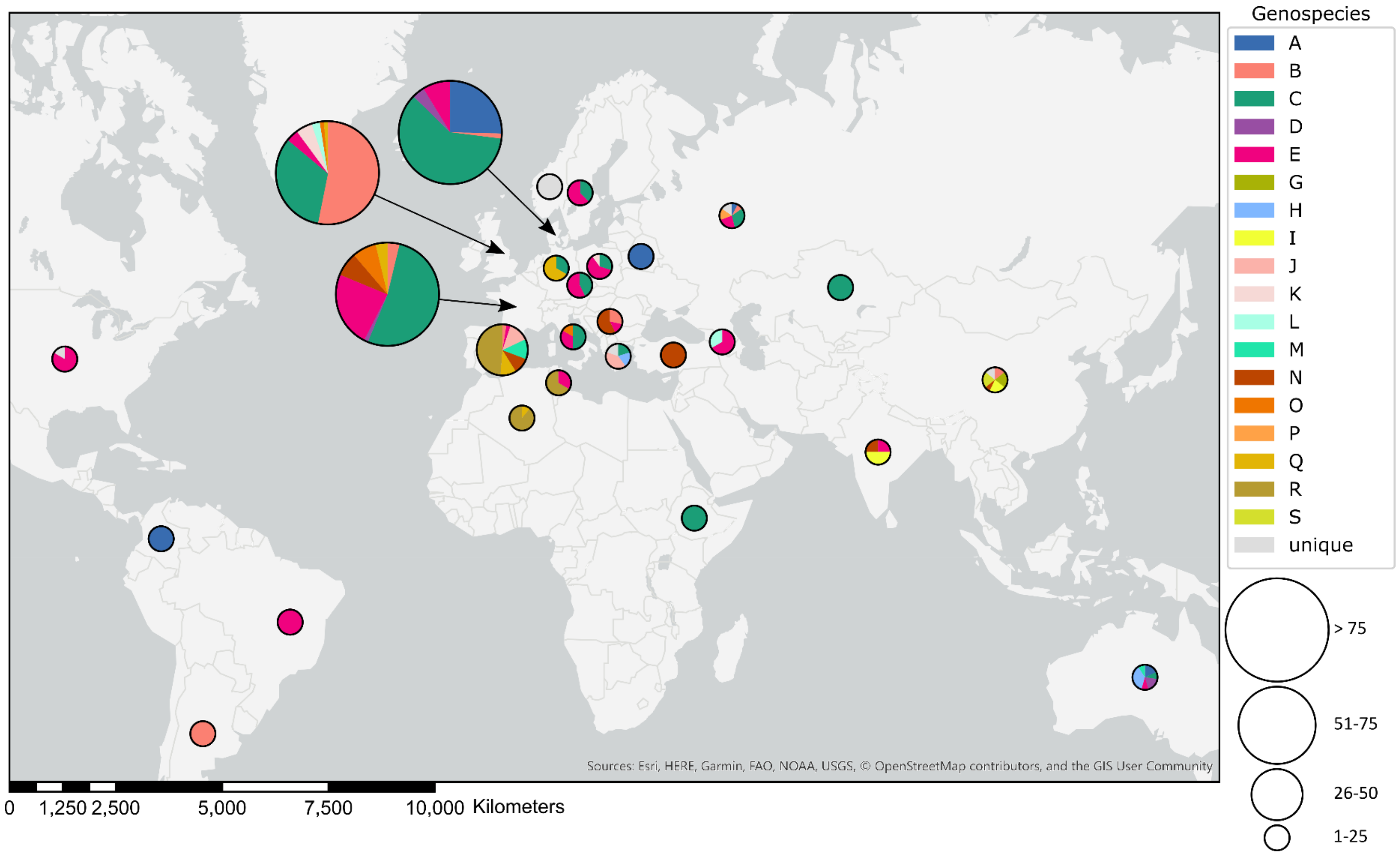
H strain WSM1325, isolated from the Greek island of Serifos and used in Australia as a clover (*Trifolium* spp.) inoculant, a common forage crop in Australia (Reeve *et al.*, 2010), or WSM1455 (genospecies J) isolated from the Greek island Mykonos and often used in Australia for faba bean, peas and lentils (Herridge *et al.*, 2008).

Middle Eastern rhizobia are under-represented in this dataset with only 4 available genomes belonging to genospecies E, L and N. Similarly, African (more specifically, North African) rhizobia populations are very infrequent in the available data with genomes mainly belonging to genospecies R. However, there seems to be a difference in the dominant genospecies depending on the region. While in northern European countries genospecies A-C and E are the most frequent, in the Mediterranean Basin genospecies R is by far the most frequently found genospecies. Furthermore, genospecies A, K and L are not found in Mediterranean Basin countries but are encountered in northern European countries, whereas genospecies J, M, N and R show the opposite pattern.

The environment is a selective pressure that affects all life forms. For edaphic dwellers like rhizobia, the soil chemical composition, pH, and texture together with climatic factors like precipitation, temperature, and solar irradiation impose a selective pressure over these organisms and after many centuries the bacteria evolve into functionally distinct types, or species. The genotypic and phenotypic evidence found in this project gives support to the genospecies separation proposed by Kumar *et al*. (2015) and Young *et al.* (2021), and their likely separation into formally described novel species with different adaptations to environmental stresses. Understanding the taxonomy of this closely related group of genospecies will provide the opportunity to start identifying those that are better at tolerating certain stresses like desiccation tolerance. For example, genospecies R seems better at tolerating desiccation than genospecies C, a tolerance that is most likely due to its adaptation to drier habitats (Chapter 3 and 4).

Table 6.1. Genospecies distribution per continent. Data from Cavassim *et al.* (2020), Boivin *et al.* (2020,2021), Young *et al.* (2021) and this work.

Continent	**A**	**B**	**C**	**D**	**E**	**G**	**H**	**I**	**J**	**K**	**L**	**M**	**N**	**O**	**P**	**Q**	**R**	**S**	**unique**	**Total**
Africa	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	1	10	0	0	13
America	1	1	0	0	8	0	0	0	0	0	0	0	0	0	0	0	0	0	1	11
Asia	0	2	2	0	3	3	0	5	0	0	1	0	2	0	0	0	0	3	2	23
Europe	34	51	165	6	59	0	1	0	7	5	2	5	15	8	2	10	19	0	4	393
Oceania	2	0	1	2	1	0	4	0	0	0	0	1	0	0	0	0	0	0	0	11
Total	37	54	169	8	72	3	5	5	7	5	3	6	17	8	2	11	29	3	7	451



Sources: Esri, HERE, Garmin, FAO, NOAA, USGS, © OpenStreetMap contributors, and the GIS User Community

Figure 6.2. Genospecies population composition of Rlc per country. Data collated from Boivin *et al.* (2020, 2021), Cavassim *et al.* (2020), and Young *et al.* (2021), and the present study. The dataset contains 451 genomes. The number of genomes in each country is marked by the pie chart diameter as indicated in the legend.

6.3 Are any of the novel strains isolated in this project better potential inoculant candidates than current standard strains?

The strains used in this study have been screened for symbiotic performance (Chapter 2) and desiccation tolerance (Chapter 3) along with current strains used commercially in inoculants, and alongside other potential candidates. Three of the strains isolated during this project exhibited a level of desiccation tolerance which was superior to the most tolerant of the existing commercial strains, USDA2364 (Figure 6.3A) - while still giving a similar increase in plant biomass on both cultivars of pea (Figure 6.3B and C). These strains belonged to genospecies J, Q and N, while the positive controls belonged to genospecies C and E. Two of these strains (121B21 and 51B21) were tested in field experiments over two consecutive seasons (Chapter 2) although they did not show yield increases over the positive control (existing commercial strains), or the un-inoculated control. Also, in the drier year they showed improved seedling emergence compared to the un-inoculated control, and in the wetter year they also improved emergence compared to the positive control strains.

These results indicate that while the plant interaction may be at an optimal state in terms of plant growth promotion using current commercialised strains, there is room for improvement in terms of tolerance to desiccation, and that sourcing strains from drier areas yields rhizobia that are better adapted to withstand this stress without an apparent reduction in plant performance.

6.4 Should there be a genospecies control when comparing strains for any phenotype?

It is possible that symbiotic performance is a cross-genospecies phenotype determined by different Sym-plasmids being shared between genospecies (Boivin *et al.*, 2020, 2021). That is, and in terms of symbiotic performance, a genospecies-specific control would be unnecessary. However, the data in this project showed that certain genospecies are better adapted to desiccation tolerance than others, and the same could apply to other phenotypes, particularly if the phenotype is controlled by chromosomic genes.

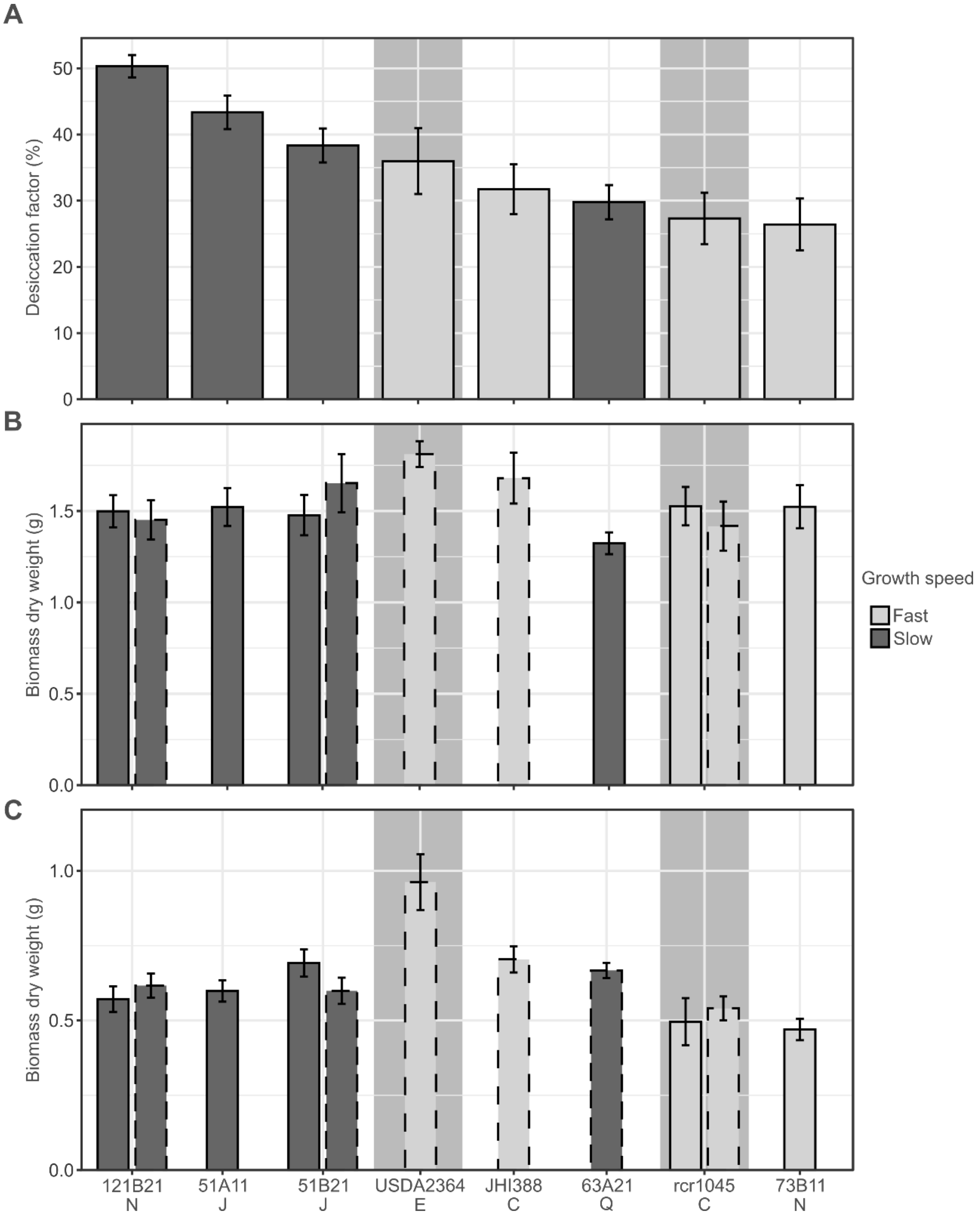


Figure 6.3. Selection of strains screened for desiccation tolerance (A) and biomass increase on pea cvs Kareni (B) and Corus (C), and their growth speed. Data extracted from Chapters 2 and 3 for the best performing isolates in biomass production and their respective desiccation tolerance factor. Bars with a shaded area in the background are commercially used strains. The two biomass screening experiments are represented by bars with dashed or solid lines for the first and

second experiments, respectively. The letters under the strain code indicate their genospecies.

Comparing the three most frequent genospecies of the strains analysed in the present study (B, C and R) it is possible to observe that some have a higher average desiccation factor than others (Figure 6.4A), but these differences are less explicit when comparing the average above-ground biomass (Figure 6.4B). Therefore, before comparing these strains their functional difference needs to be standardised. For example, this can be done by dividing the desiccation factor of each strain by the desiccation factor of a genospecies-specific reference strain.

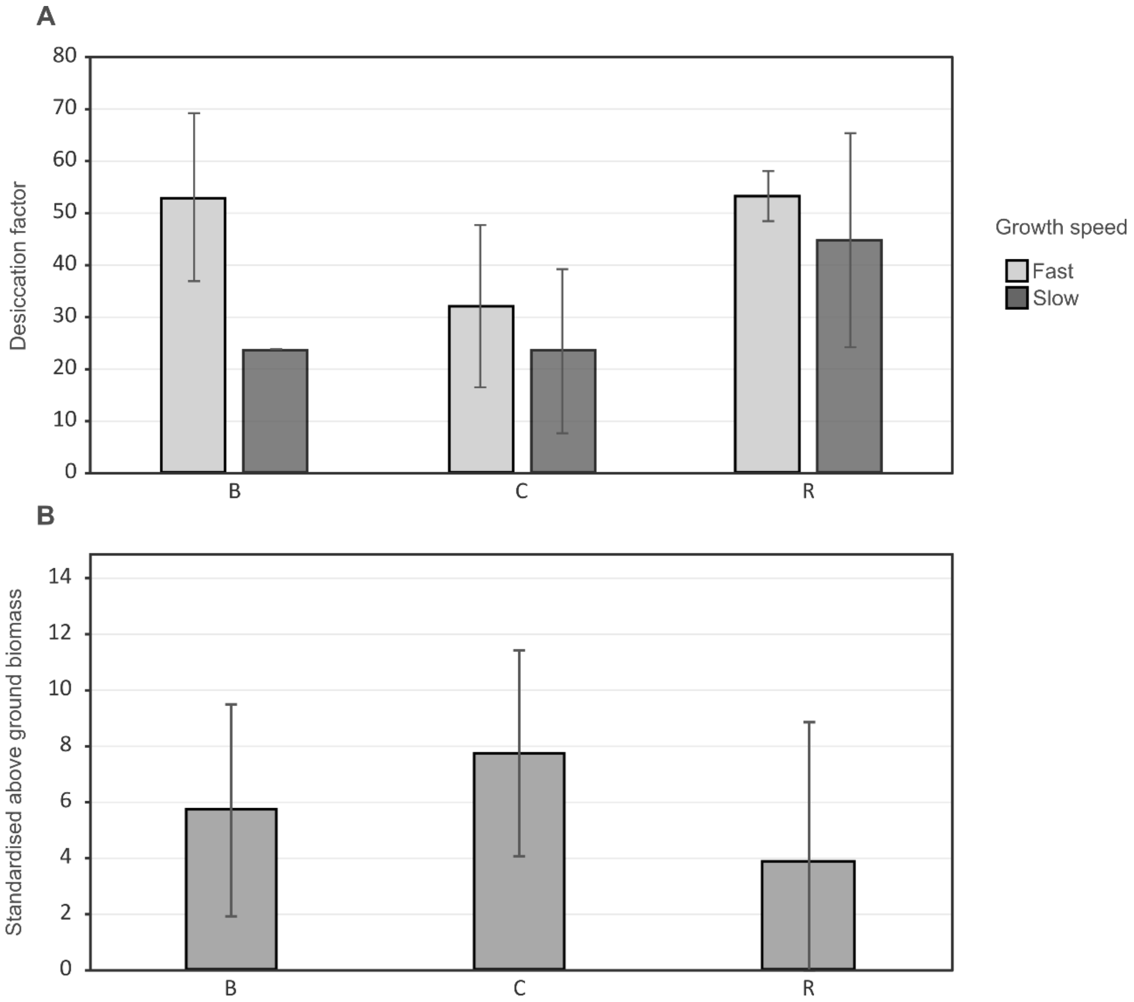


Figure 6.4. Desiccation factors of fast and slow growing strains from the three most frequent genospecies (A). Standardised above ground biomass on pea cv Corus of the three most frequent genospecies (B). The bars represent the average and the lines the standard deviation. Data extracted from Chapters 2 and 3 and Maluk *et al.* (2022).

6.5 What are the next steps for these novel strains?

Starting from the isolation of wild symbiotic rhizobia, screening for a strain that meets all the requirements for a commercially viable inoculant requires many years of research and development. Therefore, those candidate strains with potential for commercialisation that have been isolated during this project still require further characterisation and field trialling before they are ready for the market.

Rhizobium leguminosarum is native to the UK and its soils support hundreds of thousands of pea-compatible strains (Chapter 6, Mutch and Young, 2004; Macdonald *et al.*, 2011; Maluk *et al.*, 2022). Thus, when a strain is added to the soil or sown into it as part of a seed coating, the strain must be able to compete with these natural populations of rhizobia. High competitiveness for nodule occupancy is a desirable characteristic for a candidate strain, otherwise native soil-borne rhizobia, and potentially with a lower capacity for BNF, may nodulate the target legume and reduce or negate any benefit of the inoculant (Sánchez-Cañizares and Palacios, 2013; Mendoza-Suárez *et al.*, 2021; Westhoek *et al.*, 2021).

In contrast to previous reports, arable soils can sustain large quantities of Rlv even after many decades of absence of the legume host (Chapter 6). The fields at Balruddery Farm used for running the trials in Chapter 2 were in the vicinity of the Centre for Sustainable Cropping (CSC) platform studied in Chapter 6. Therefore, it is likely that they have a similar (quite high) concentration of compatible rhizobia to the CSC platform fields. These native rhizobia are likely to have competed for nodule occupancy with the inoculated strains reducing the overall effectiveness of the inoculation. Thus, assessing the competitiveness for nodule occupancy of these strains is necessary to assess their suitability for being the biological component of a new inoculant.

Another important step in the development of an inoculant, and the most intellectual property-sensitive component, is the formulation of the carrier. This is because the carrier is the substrate wherein the inoculant strain will live until its use, and formulating an optimal medium for it can considerably extend the inoculant shelf-life (Streeter, 2003), and/or improve the inoculant efficacy (Kozar *et al.*, 2019).

Therefore, two areas of interest for progressing these candidate strains into a commercial product are the assessment of their competition for nodule occupancy, and the design of an optimum carrier for their storage and dissemination.

6.6 Are optimal symbiotic and desiccation tolerance possible in the same strain?

Nodulation and nitrogen fixation genes are plasmid-bound in Rlc species (Young *et al.*, 2006). However, most of the known genes involved in desiccation tolerance are on the chromosome (Chapter 4). It can be hypothesised, therefore, that finding a strain with optimal N-fixation and plant interaction abilities with optimal desiccation tolerance is possible, as it should be only a matter of finding the right chromosome together with the right plasmid. This has been reported recently for the complex traits of competitivity for nodule occupancy and BNF though the genes involved in both traits are located in the Sym-plasmid (Mendoza-Suárez *et al.*, 2020).

Conversely, there might be a trade-off between optimal nodulation, optimal BNF, and optimal desiccation tolerance (Figure 6.5). A comparative correlation analysis was therefore conducted of the desiccation tolerance factors for all strains and above-ground biomass data of inoculated pea cv Corus from Maluk *et al.* (2022) and tests of strains isolated from Spain (reported here). Despite the screening process being the same between our two studies, to reduce the error inherent in variation among each experimental setup, the aboveground biomass was standardised by dividing the test strain biomass production by that of the un-inoculated control from the same experiment. Additionally, when the same strain was present in several experiments, an average of all the standardised biomass production was calculated. The results of the Pearson’s correlation showed that there was a significant negative correlation ($P < 0.05$) between both variables. This indicates that there might be a trade-off between desiccation tolerance and plant growth promoting (BNF) potential, and that a strain which enable excellent nodulation and BNF, may not necessarily have a high tolerance to desiccation.

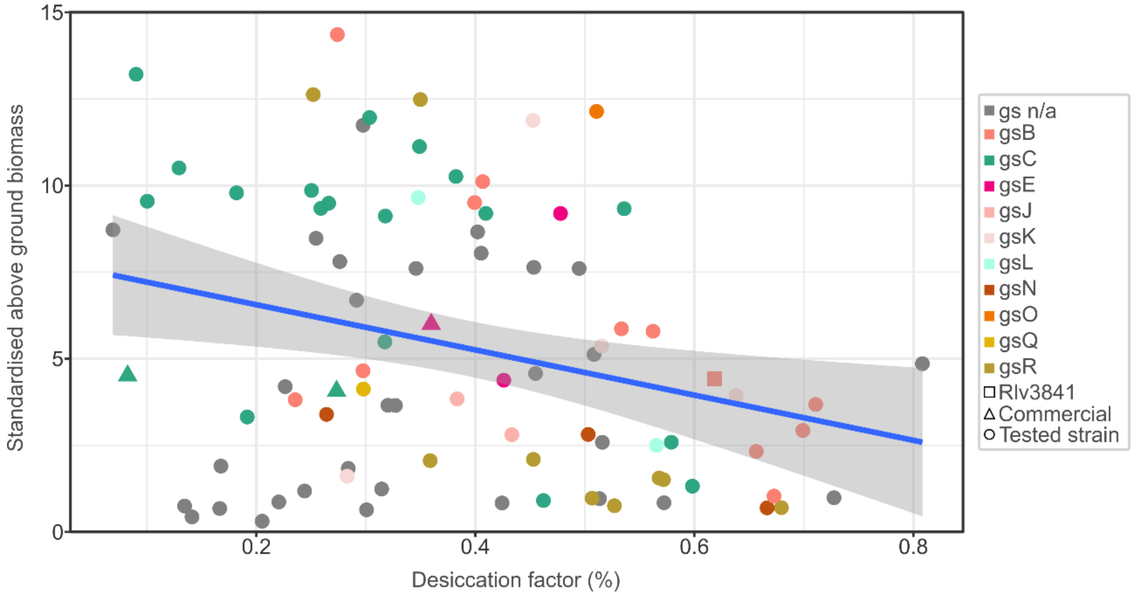


Figure 6.5. Correlation plot of desiccation factor and the standardised above-ground biomass production on pea cv Corus. The blue line represents the correlation between both variables and the shaded area is the standard error of this correlation. Legend acronyms: gs - genospecies, n/a - not assessed.

The two phenotypes studied here are complex traits whose function is the product of many interacting genes and proteins, and likely influenced by environmental factors too. Symbiotic efficiency is mediated by several gene families such as *nod, nif* and *fix* which are responsible for nodulation and nitrogen fixation and are located on the Sym plasmid (Young *et al.*, 2006). Conversely, desiccation tolerance is also mediated by a range of genes, but the linkage between genotype and phenotype is sometimes indirect, as many different mechanisms can interact to achieve the same phenotype. There are therefore many strategies a rhizobia cell can evoke to tolerating desiccation, e.g. the accumulation of trehalose, and/or other osmolytes seems to be a common strategy (Streeter, 2003; Cytryn *et al.*, 2007; Reina-Bueno *et al.*, 2012), and there are many other mechanisms such as DNA, membrane, or protein protection (Humann *et al.*, 2009; Humann and Kahn, 2015). Thus, for both traits to be co-optimised many appropriate genes need to be present in the same strain at the same time, and this might be difficult to achieve.

However, the data synthesis presented here does highlight that what may be easier to identify, is a strain that has a good balance (though not maximal) of phenotypes for both symbiotic interaction (i.e. BNF), and desiccation tolerance.

6.7 Concluding remarks

This project has successfully isolated rhizobial strains with comparable symbiotic interactions to standard commercial strains in terms of plant growth promotion mediated *via* BNF. It has demonstrated that the symbiotic efficiency of commercial strains has been maximised for peas, with the highest performing strains yielding a similar biomass production on pea in a N-free environment in growth room experiments. In contrast, this research has shown that the desiccation tolerance of such commercial standards is below optimal, whereas the candidate strains isolated in this study have a better tolerance to desiccation without compromising their symbiotic performance. This desiccation tolerance makes them potentially superior for use in commercial inoculants as they are capable of remaining effective for longer periods of storage.

Desiccation tolerance is a complex trait and the findings from this study support known strategies for withstanding this environmental stress in rhizobia. Furthermore, this study found evidence indicating that strains isolated from countries with drier environmental conditions are likely to tolerate better anhydrobiosis than those isolated from locations where water is more abundant. This suggests opportunities to search for optimal desiccation tolerant strains in hot and dry areas of the world for their use as inoculants.

Finally, this project has shown that fields with long absences of legume hosts are still able to sustain considerable populations of rhizobia. This highlights the capacity of rhizobia to live saprophytically in soil without the necessity of interacting with their legume hosts. However, this study has also demonstrated that the reintroduction of a compatible legume host after long absences produces an at least 4-year lasting effect that increases the proportion of nodulating rhizobia in soil year-on-year, resulting in a population shift in favour of host-compatible rhizobia.

References

Abdel-Hamid, A. M. and Cronan, J. E. (2007) ‘Coordinate expression of the acetyl coenzyme A carboxylase genes, *accB* and *accC*, is necessary for normal regulation of biotin synthesis in *Escherichia coli*’, *Journal of Bacteriology*, 189(2), pp. 369–376. doi: 10.1128/JB.01373-06.

Ahemad, M. and Khan, M. S. (2013) ‘Pesticides as antagonists of rhizobia and the legume-rhizobium symbiosis: a paradigmatic and mechanistic outlook’, *Biochem Mole Biol*, 1, pp. 63–75.

Albareda, M., Rodríguez-Navarro, D. N., Camacho, M. and Temprano, F. J. (2008) ‘Alternatives to peat as a carrier for rhizobia inoculants: Solid and liquid formulations’, *Soil Biology and Biochemistry*, 40(11), pp. 2771–2779. doi: 10.1016/j.soilbio.2008.07.021.

Alster, C. J., German, D. P., Lu, Y. and Allison, S. D. (2013) ‘Microbial enzymatic responses to drought and to nitrogen addition in a southern California grassland’, *Soil Biology and Biochemistry*, 64, pp. 68–79. doi: https://doi.org/10.1016/j.soilbio.2013.03.034.

Altamia, M. A., Shipway, J. R., Stein, D., Betcher, M. A., Fung, J. M., Jospin, G., Eisen, J., Haygood, M. G. and Distel, D. L. (2020) ‘*Teredinibacter waterburyi* sp. nov., a marine, cellulolytic endosymbiotic bacterium isolated from the gills of the wood-boring mollusc *Bankia setacea* (Bivalvia: Teredinidae) and emended description of the genus *Teredinibacter*’, *International journal of systematic and evolutionary microbiology*. 2020/02/20. Microbiology Society, 70(4), pp. 2388–2394. doi: 10.1099/ijsem.0.004049.

Alves, B. J. R., Boddey, R. M. and Urquiaga, S. (2003) ‘The success of BNF in soybean in Brazil’, *Plant and Soil*, 252(1), pp. 1–9. doi: 10.1023/A:1024191913296.

Andrews, M. and Andrews, M. E. (2017) ‘Specificity in legume-rhizobia symbioses’, *International Journal of Molecular Sciences*, 18(4), p. 705. doi: 10.3390/ijms18040705.

Andrews, M., De Meyer, S., James, E. K., St\kepkowski, T., Hodge, S., Simon, M. F. and Young, J. P. W. (2018) ‘Horizontal transfer of symbiosis genes within and between rhizobial genera: occurrence and importance’, *Genes*. Multidisciplinary Digital Publishing Institute, 9(7), p. 321.

Angamuthu, M., Shankar, V. K. and Murthy, S. N. (2018) ‘Water activity and its significance in topical dosage forms’, *Journal of Pharmaceutical Sciences*. Elsevier, 107(6), pp. 1656–1666. doi: 10.1016/j.xphs.2018.02.013.

Atieno, M., Wilson, N., Casteriano, A., Crossett, B., Lesueur, D. and Deaker, R. (2018) ‘Aqueous peat extract exposes rhizobia to sub-lethal stress which may prime cells for improved desiccation tolerance’, *Applied Microbiology and Biotechnology*, 102(17), pp. 7521–7539. doi: 10.1007/s00253-018-9086-2.

Bajaj, I. and Singhal, R. (2011) ‘Poly (glutamic acid) – An emerging biopolymer of commercial interest’, *Bioresource Technology*, 102(10), pp. 5551–5561. doi: https://doi.org/10.1016/j.biortech.2011.02.047.

Ballal, A., Basu, B. and Apte, S. K. (2007) ‘The Kdp-ATPase system and its regulation’, *Journal of Biosciences*, 32(3), pp. 559–568. doi: 10.1007/s12038-007-0055-7.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., Lesin, V. M., Nikolenko, S. I., Pham, S., Prjibelski, A. D., Pyshkin, A. V, Sirotkin, A. V, Vyahhi, N., Tesler, G., Alekseyev, M. A. and Pevzner, P. A. (2012) ‘SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing’, *Journal of Computational Biology*, 19(5), pp. 455–477. doi: 10.1089/cmb.2012.0021.

Barbedo, J. G. A. (2012) ‘Method for automatic counting root nodules using digital images’, in *2012 12th International Conference on Computational Science and Its Applications*, pp. 159–161. doi: 10.1109/ICCSA.2012.39.

Barłóg, P., Grzebisz, W. and Łukowiak, R. (2018) ‘Faba bean yield and growth dynamics in response to soil potassium availability and sulfur application’, *Field Crops Research*, 219, pp. 87–97. doi: https://doi.org/10.1016/j.fcr.2018.01.027.

Barnard, R. L., Osborne, C. A. and Firestone, M. K. (2013) ‘Responses of soil bacterial and fungal communities to extreme desiccation and rewetting’, *The ISME Journal*, 7(11), pp. 2229–2241. doi: 10.1038/ismej.2013.104.

Bashan, Y. (1986) ‘Alginate beads as synthetic inoculant carriers for slow release of bacteria that affect plant growth’, *Applied and Environmental Microbiology*, 51(5), pp. 1089–1098.

Bashan, Y. (1998) ‘Inoculants of plant growth-promoting bacteria for use in agriculture’, *Biotechnology Advances*, 16(4), pp. 729–770. doi: 10.1016/S0734-9750(98)00003-2.

Bashan, Y., De-Bashan, L. E., Prabhu, S. R. and Hernandez, J.-P. (2014) ‘Advances in plant growth-promoting bacterial inoculant technology: formulations and practical perspectives (1998–2013)’, *Plant and Soil*, 378(1–2), pp. 1–33. doi: 10.1007/s11104-013-1956-x.

Bashan, Y., Hernandez, J.-P., Leyva, L. and Bacilio, M. (2002) ‘Alginate microbeads as inoculant carriers for plant growth-promoting bacteria’, *Biology and Fertility of Soils*, 35(5), pp. 359–368. doi: 10.1007/s00374-002-0481-5.

Bates, D., Mächler, M., Bolker, B. and Walker, S. (2015) ‘Fitting linear mixed-effects models using {lme4}’, *Journal of Statistical Software*, 67(1), pp. 1–48. doi: 10.18637/jss.v067.i01.

Bell, C. D., Soltis, D. E. and Soltis, P. S. (2010) ‘The age and diversification of the angiosperms re-revisited’, *American Journal of Botany*, 97(8), pp. 1296–1303. doi: https://doi.org/10.3732/ajb.0900346.

Bell, L. and Labuza, T. (2000) *Moisture sorption: practical aspects of isotherm measurement and use*. 2nd edn. American Association of Cereal Chemists.

Benidire, L., Lahrouni, M., Daoui, K., el Abidine Fatemi, Z., Carmona, R. G., Göttfert, M. and Oufdou, K. (2018) ‘Phenotypic and genetic diversity of Moroccan rhizobia isolated from *Vicia faba* and study of genes that are likely to be involved in their osmotolerance’, *Systematic and Applied Microbiology*, 41(1), pp. 51–61. doi: https://doi.org/10.1016/j.syapm.2017.09.003.

Benjamini, Y. and Hochberg, Y. (1995) ‘Controlling the false discovery rate: A practical and powerful approach to multiple testing’, *Journal of the Royal Statistical Society. Series B (Methodological)*. [Royal Statistical Society, Wiley], 57(1), pp. 289–300. Available at: http://www.jstor.org/stable/2346101.

van Berkum, P., Beyene, D., Vera, F. T. and Keyser, H. H. (1995) ‘Variability among *Rhizobium* strains originating from nodules of *Vicia faba*.’, *Applied and Environmental Microbiology*, 61(7), pp. 2649 LP – 2653. Available at: http://aem.asm.org/content/61/7/2649.abstract.

Bernhard, A. (2010) ‘The nitrogen cycle: processes, players, and human impact’, *Nature Education Knowledge*, 3(10), p. 25. Available at: https://www.nature.com/scitable/knowledge/library/the-nitrogen-cycle-processes-players-and-human-15644632.

Berninger, T., Lopez, O. G., Bejarano, A., Preininger, C. and Sessitsch, A. (2018) ‘Maintenance and assessment of cell viability in formulation of non-sporulating bacterial inoculants’, *Microbial Biotechnology*. WILEY, 11(2), pp. 277–301. doi: 10.1111/1751-7915.12880.

Berninger, T., Mitter, B. and Preininger, C. (2016) ‘The smaller, the better? The size effect of alginate beads carrying plant growth-promoting bacteria for seed coating’, *Journal of Microencapsulation*, 33(2), pp. 127–136. doi: 10.3109/02652048.2015.1134690.

Bertani, G. (1951) ‘Studies on lysogenesis. I. The mode of phage liberation by lysogenic *Escherichia coli*’, *Journal of bacteriology*, 62(3), pp. 293–300. Available at: https://pubmed.ncbi.nlm.nih.gov/14888646.

Bing, D. (2015) *Pea*, *The Canadian Encyclopedia. Historica Canada*. Available at: www.thecanadianencyclopedia.ca/en/article/pea (Accessed: 12 December 2021).

Bitanyi, H. F. (1983) *Competition studies in* Rhizobium leguminosarum. Imperial College of Science and Technology, London.

Boivin, C., Ndoye, I., Molouba, F., de Lajudie, P., Dupuy, N., Dreyfus, B. and de Bruijn, D. F. J. (1997) ‘Stem nodulation in legumes: Diversity, mechanisms, and unusual characteristics’, *Critical Reviews in Plant Sciences*. Taylor & Francis, 16(1), pp. 1–30. doi: 10.1080/07352689709701944.

Boivin, S. *et al.* (2020) ‘Host-specific competitiveness to form nodules in *Rhizobium leguminosarum* symbiovar *viciae*’, *New Phytologist*, 226(2), pp. 555–568. doi: 10.1111/nph.16392.

Boivin, S., Mahé, F., Debellé, F., Pervent, M., Tancelin, M., Tauzin, M., Wielbo, J., Mazurier, S., Young, P. and Lepetit, M. (2021) ‘Genetic variation in host-specific competitiveness of the symbiont *Rhizobium leguminosarum* symbiovar *viciae*’, *Frontiers in Plant Science*, 12, p. 1790. doi: 10.3389/fpls.2021.719987.

Bolger, A. M., Lohse, M. and Usadel, B. (2014) ‘Trimmomatic: a flexible trimmer for Illumina sequence data’, *Bioinformatics*, 30(15), pp. 2114–2120. doi: 10.1093/bioinformatics/btu170.

Boscari, A., Mandon, K., Dupont, L., Poggi, M.-C. and Rudulier, D. Le (2002) ‘*BetS* is a major glycine betaine/proline betaine transporter required for early osmotic adjustment in *Sinorhizobium meliloti*’, *Journal of Bacteriology*, 184(10), pp. 2654–2663. doi: 10.1128/JB.184.10.2654-2663.2002.

Bravo, Z., Orruño, M., Parada, C., Kaberdin, V. R., Barcina, I. and Arana, I. (2016) ‘The long-term survival of *Acinetobacter baumannii* ATCC 19606T under nutrient-deprived conditions does not require the entry into the viable but non-culturable state’, *Archives of Microbiology*, 198(5), pp. 399–407. doi: 10.1007/s00203-016-1200-1.

Brear, E., Day, D. and Smith, P. (2013) ‘Iron: an essential micronutrient for the legume-rhizobium symbiosis’, *Frontiers in Plant Science*, 4, p. 359. doi: 10.3389/fpls.2013.00359.

Brito, J. A., Denkmann, K., Pereira, I. A. C., Archer, M. and Dahl, C. (2015) ‘Thiosulfate dehydrogenase (TsdA) from *Allochromatium vinosum*: Structural and functional insights into thiosulfate oxidation’, *Journal of Biological Chemistry*. Elsevier, 290(14), pp. 9222–9238. doi: 10.1074/jbc.M114.623397.

Brockwell, J. and Bottomley, P. J. (1995) ‘Recent advances in inoculant technology and prospects for the future’, *Soil Biology and Biochemistry*. Oxford: Pergamon-Elsevier Science Ltd, 27(4–5), pp. 683–697. doi: 10.1016/0038-0717(95)98649-9.

Brockwell, J., Bottomley, P. J. and Thies, J. E. (1995) ‘Manipulation of rhizobia microflora for improving legume productivity and soil fertility: A critical assessment’, *Plant and Soil*, 174(1–2), pp. 143–180. doi: 10.1007/BF00032245.

de Bruijn, F. J. (2015) ‘Introduction’, in *Biological Nitrogen Fixation*. John Wiley & Sons, Ltd, pp. 1–4. doi: https://doi.org/10.1002/9781119053095.ch1.

Bullard, G. K., Roughley, R. J. and Pulsford, D. J. (2005) ‘The legume inoculant industry and inoculant quality control in Australia: 1953-2003’, *Australian Journal of Experimental Agriculture*, 45(3), pp. 127–140. Available at: https://doi.org/10.1071/EA03159.

Burchill, W., James, E. K., Li, D., Lanigan, G. J., Williams, M., Iannetta, P. P. M. and Humphreys, J. (2014) ‘Comparisons of biological nitrogen fixation in association with white clover (*Trifolium repens L.*) under four fertiliser nitrogen inputs as measured using two 15N techniques’, *Plant and Soil*, 385(1), pp. 287–302. doi: 10.1007/s11104-014-2199-1.

Burris, R. H. (1980) ‘The global nitrogen budget: science or seance?’ University Park Press.

Burris, R. H. (1994) ‘Historical developments in biological nitrogen fixation’, in Frey, K. J. (ed.) *Historical perspectives in plant science*. Iowa State University Press, pp. 23–41.

Caballero-Mellado, J. and Martinez-Romero, E. (1999) ‘Soil fertilization limits the genetic diversity of *Rhizobium* in bean nodules’, *Symbiosis*. Balaban Publishers.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K. and Madden, T. L. (2009) ‘BLAST+: architecture and applications’, *BMC Bioinformatics*, 10(1), p. 421. doi: 10.1186/1471-2105-10-421.

Carey, P. D., Short, C., Morris, C., Hunt, J., Priscott, A., Davis, M., Finch, C., Curry, N., Little, W., Winter, M., Parkin, A. and Firbank, L. G. (2003) ‘The multi-disciplinary evaluation of a national agri-environment scheme’, *Journal of Environmental Management*, 69(1), pp. 71–91. doi: https://doi.org/10.1016/S0301-4797(03)00120-8.

Carpenter, S. R. (2005) ‘Eutrophication of aquatic ecosystems: Bistability and soil phosphorus’, *Proceedings of the National Academy of Sciences*. National Academy of Sciences, 102(29), pp. 10002–10005. doi: 10.1073/pnas.0503959102.

Casteriano, A. (2014) *Physiological mechanisms of desiccation tolerance in rhizobia*. University of Sydney.

Casteriano, A., Wilkes, M. A. and Deaker, R. (2013) ‘Physiological changes in rhizobia after growth in peat extract may be related to improved desiccation tolerance’, *Applied and Environmental Microbiology*, 79(13), pp. 3998–4007. doi: 10.1128/AEM.00082-13.

Catroux, G. and Amarger, N. (1992) ‘Rhizobia as soil inoculants in agriculture’, in Fry, J. C. and Day, M. J. (eds) *Release of genetically engineered and other micro-organisms*. Cambridge University Press, pp. 1–13.

Catroux, G., Hartmann, A. and Revellin, C. (2001) ‘Trends in rhizobial inoculant production and use’, *Plant and Soil*, 230(1), pp. 21–30. doi: 10.1023/A:1004777115628.

Cavassim, M. I. A., Moeskjær, S., Moslemi, C., Fields, B., Bachmann, A., Vilhjálmsson, B. J., Schierup, M. H., W. Young, J. P. and Andersen, S. U. (2020) ‘Symbiosis genes show a unique pattern of introgression and selection within a *Rhizobium leguminosarum* species complex’, *Microbial Genomics*. Microbiology Society, 6(4). doi: https://doi.org/10.1099/mgen.0.000351.

Cerri, M. R., Frances, L., Laloum, T., Auriac, M.-C., Niebel, A., Oldroyd, G. E. D., Barker, D. G., Fournier, J. and de Carvalho-Niebel, F. (2012) ‘*Medicago truncatula* ERN transcription factors: Regulatory interplay with NSP1/NSP2

GRAS factors and expression dynamics throughout rhizobial infection', *Plant Physiology*, 160(4), pp. 2155–2172. doi: 10.1104/pp.112.203190.

Chao, W. L. and Alexander, M. (1984) 'Mineral soils as carriers for rhizobium inoculants', *Applied and Environmental Microbiology*, 47(1), pp. 94–97.

Chaudri, A., McGrath, S., Gibbs, P., Chambers, B., Carlton-Smith, C., Bacon, J., Campbell, C. and Aitken, M. (2008) 'Population size of indigenous *Rhizobium leguminosarum* biovar *trifolii* in long-term field experiments with sewage sludge cake, metal-amended liquid sludge or metal salts: Effects of zinc, copper and cadmium', *Soil Biology and Biochemistry*, 40(7), pp. 1670–1680. doi: https://doi.org/10.1016/j.soilbio.2008.01.026.

Chen, P. E. and Shapiro, B. J. (2015) 'The advent of genome-wide association studies for bacteria', *Current Opinion in Microbiology*, 25, pp. 17–24. doi: https://doi.org/10.1016/j.mib.2015.03.002.

Chibeba, A. M., Kyei-Boahen, S., de Fátima Guimarães, M., Nogueira, M. A. and Hungria, M. (2018) 'Feasibility of transference of inoculation-related technologies: A case study of evaluation of soybean rhizobial strains under the agro-climatic conditions of Brazil and Mozambique', *Agriculture, Ecosystems & Environment*, 261, pp. 230–240. doi: https://doi.org/10.1016/j.agee.2017.06.037.

Chirife, J. and Fontana, A. J. (2007) 'Introduction: historical highlights of water activity research', in Barbosa-Cánovas, G. V., Fontana, A. J., Schmidt, S. J., and Labuza, T. P. (eds) *Water activity in foods. Fundamentals and applications.* Blackwell Publishing Ltd, pp. 3–13.

Chiurazzi, M., Meza, R., Lara, M., Lahm, A., Defez, R., Iaccarino, M. and Espín, G. (1992) 'The *Rhizobium leguminosarum* biovar *phaseoli glnT* gene, encoding glutamine synthetase III', *Gene*, 119(1), pp. 1–8. doi: https://doi.org/10.1016/0378-1119(92)90060-3.

Cholley, A. C., Traoré, O., Hennequin, C. and Aumeran, C. (2020) '*Klebsiella pneumoniae* survival and regrowth in endoscope channel biofilm exposed to glutaraldehyde and desiccation', *European Journal of Clinical Microbiology &*

Infectious Diseases, 39(6), pp. 1129–1136. doi: 10.1007/s10096-020-03818-7.

Clark, I. M., Mendum, T. A. and Hirsch, P. R. (2002) ‘The influence of the symbiotic plasmid pRL1JI on the distribution of GM rhizobia in soil and crop rhizospheres, and implications for gene flow’, *Antonie van Leeuwenhoek*, 81(1), pp. 607–616. doi: 10.1023/A:1020574009445.

Cohan, F. M. (2002) ‘What are bacterial species?’, *Annual Review of Microbiology*. Annual Reviews, 56(1), pp. 457–487. doi: 10.1146/annurev.micro.56.012302.160634.

Collins, C. and Didelot, X. (2018) ‘A phylogenetic method to perform genome-wide association studies in microbes that accounts for population structure and recombination’, *PLOS Computational Biology*. Public Library of Science, 14(2), pp. 1–21. doi: 10.1371/journal.pcbi.1005958.

Costerton, J. W., Lewandowski, Z., Caldwell, D. E., Korber, D. R. and Lappin-Scott, H. M. (1995) ‘Microbial biofilms’, *Annual Review of Microbiology*, 49(1), pp. 711–745. doi: 10.1146/annurev.mi.49.100195.003431.

Cousin, R. (1997) ‘Peas (*Pisum sativum* L.)’, *Field Crops Research*, 53(1), pp. 111–130. doi: https://doi.org/10.1016/S0378-4290(97)00026-9.

Crews, T. E. and Peoples, M. B. (2004) ‘Legume versus fertilizer sources of nitrogen: ecological tradeoffs and human needs’, *Agriculture, Ecosystems & Environment*, 102(3), pp. 279–297. doi: https://doi.org/10.1016/j.agee.2003.09.018.

Cytryn, E. J., Sangurdekar, D. P., Streeter, J. G., Franck, W. L., Chang, W.-S., Stacey, G., Emerich, D. W., Joshi, T., Xu, D. and Sadowsky, M. J. (2007) ‘Transcriptional and physiological responses of *Bradyrhizobium japonicum* to desiccation-induced stress’, *JOURNAL OF BACTERIOLOGY*, 189(19), pp. 6751–6762. doi: 10.1128/JB.00533-07.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., Durbin, R. and Group, 1000 Genomes Project Analysis (2011) ‘The variant call format and VCFtools', *Bioinformatics*, 27(15), pp. 2156–2158. doi: 10.1093/bioinformatics/btr330.

Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M. and Li, H. (2021) 'Twelve years of SAMtools and BCFtools', *GigaScience*, 10(2). doi: 10.1093/gigascience/giab008.

Daniell, T. J., Davidson, J., Alexander, C. J., Caul, S. and Roberts, D. M. (2012) 'Improved real-time PCR estimation of gene copy number in soil extracts using an artificial reference', *Journal of Microbiological Methods*, 91(1), pp. 38–44. doi: https://doi.org/10.1016/j.mimet.2012.07.010.

Darriba, D., Posada, D., Kozlov, A. M., Stamatakis, A., Morel, B. and Flouri, T. (2019) 'ModelTest-NG: A new and scalable tool for the selection of DNA and protein evolutionary models', *Molecular Biology and Evolution*, 37(1), pp. 291–294. doi: 10.1093/molbev/msz189.

Date, R. A. (2001) 'Advances in inoculant technology: A brief review', *Australian Journal of Experimental Agriculture*, 41(3), p. 321. doi: 10.1071/EA00006.

Deaker, R. (2004) 'Legume seed inoculation technology - a review', *Soil Biology and Biochemistry*, 36(8), pp. 1275–1288. doi: 10.1016/j.soilbio.2004.04.009.

Deaker, R., Hartley, E. and Gemell, G. (2012) 'Conditions affecting shelf-life of inoculated legume seed', *Agriculture*, 2(4), pp. 38–51. doi: 10.3390/agriculture2010038.

Deaker, R., Hartley, E., Gemell, G., Herridge, D. F. and Karanja, N. (2016) 'Inoculant production and quality control', in Howieson, J. G. and Dilworth, M. J. (eds) *Working with rhizobia*. Australian Centre for International Agricultural Research (ACIAR), pp. 167–186.

Deaker, R., Kecskés, M. L., Rose, M. T., Amprayn, K., Ganisan, K., Tran, T. K. C., Vu, T. N., Phan, T. C., Nguyen, T. H., Kennedy, I. R. and others (2011) *Practical methods for the quality control of inoculant biofertilisers.* Australian Centre for International Agricultural Research (ACIAR).

Deaker, R., Roughley, R. J. and Kennedy, I. R. (2007) ‘Desiccation tolerance of rhizobia when protected by synthetic polymers’, *Soil Biology & Biochemistry*, 39(2), pp. 573–580. doi: 10.1016/j.soilbio.2006.09.005.

Delamuta, J. R. M., Scherer, A. J., Ribeiro, R. A. and Hungria, M. (2020) ‘Genetic diversity of Agrobacterium species isolated from nodules of common bean and soybean in Brazil, Mexico, Ecuador and Mozambique, and description of the new species *Agrobacterium fabacearum* sp. nov.’, *International Journal of Systematic and Evolutionary Microbiology*. Microbiology Society, 70(7), pp. 4233–4244. doi: https://doi.org/10.1099/ijsem.0.004278.

Denkmann, K., Grein, F., Zigann, R., Siemen, A., Bergmann, J., van Helmont, S., Nicolai, A., Pereira, I. A. C. and Dahl, C. (2012) ‘Thiosulfate dehydrogenase: a widespread unusual acidophilic c-type cytochrome’, *Environmental Microbiology*, 14(10), pp. 2673–2688. doi: https://doi.org/10.1111/j.1462-2920.2012.02820.x.

Denton, M. D., Pearce, D. J. and Peoples, M. B. (2013) ‘Nitrogen contributions from faba bean (*Vicia faba* L.) reliant on soil rhizobia or inoculation’, *Plant and Soil*, 365(1), pp. 363–374. doi: 10.1007/s11104-012-1393-2.

Diez-Mendez, A., Menéndez, E., García-Fraile, P., Celador-Lera, L., Rivas, R. and Mateos, P. F. (2015) ‘Rhizobium cellulosilyticum as a co-inoculant enhances Phaseolus vulgaris grain yield under greenhouse conditions’, *Symbiosis*, 67(1), pp. 135–141. doi: 10.1007/s13199-015-0372-9.

Dileep Kumar, B. S., Berggren, I. and Mårtensson, A. M. (2001) ‘Potential for improving pea production by co-inoculation with fluorescent *Pseudomonas* and *Rhizobium*’, *Plant and Soil*, 229(1), pp. 25–34. doi: 10.1023/A:1004896118286.

Dilworth, M. J., James, E. K., Sprent, J. I. and Newton, W. E. (2008) *Nitrogen-fixing leguminous symbioses*. Edited by M. J. Dilworth, E. K. James, J. I. Sprent, and W. E. Newton. Springer Netherlands. doi: 10.1007/978-1-4020-3548-7.

Donati, A. J., Jeon, J.-M., Sangurdekar, D., So, J.-S. and Chang, W.-S. (2011) ‘Genome-wide transcriptional and physiological responses of *Bradyrhizobium japonicum* to paraquat-mediated oxidative stress’, *Applied* *and Environmental Microbiology*, 77(11), pp. 3633–3643. doi: 10.1128/AEM.00047-11.

Downie, J. A. (2014) ‘Legume nodulation’, *Current Biology*, 24(5), pp. R184–R190. doi: https://doi.org/10.1016/j.cub.2014.01.028.

Doyle, J. J. (2011) ‘Phylogenetic perspectives on the origins of nodulation’, *Molecular Plant-Microbe Interactions®*, 24(11), pp. 1289–1295. doi: 10.1094/MPMI-05-11-0114.

Doyle, J. J. (2016) ‘Chasing unicorns: Nodulation origins and the paradox of novelty’, *American Journal of Botany*, 103(11), pp. 1865–1868. doi: https://doi.org/10.3732/ajb.1600260.

Drew, E. A., Denton, M. D., Sadras, V. O. and Ballard, R. A. (2012) ‘Agronomic and environmental drivers of population size and symbiotic performance of *Rhizobium leguminosarum* bv. *viciae* in Mediterranean-type environments’, *Crop and Pasture Science*, 63(5), pp. 467–477. Available at: https://doi.org/10.1071/CP12032.

Dungait, J. A. J., Ghee, C., Rowan, J. S., McKenzie, B. M., Hawes, C., Dixon, E. R., Paterson, E. and Hopkins, D. W. (2013) ‘Microbial responses to the erosional redistribution of soil organic carbon in arable fields’, *Soil Biology and Biochemistry*, 60, pp. 195–201. doi: https://doi.org/10.1016/j.soilbio.2013.01.027.

Dye, M. (1978) ‘The Rothamsted rhizobium culture collection and inoculant use in the UK’, in *Rothamsted Experimental Station Report for 1978 Part 2*.

Edgar, R. C. (2004) ‘MUSCLE: multiple sequence alignment with high accuracy and high throughput’, *Nucleic Acids Research*, 32(5), pp. 1792–1797. doi: 10.1093/nar/gkh340.

van Egeraat, A. W. S. M. (1975) ‘The possible role of homoserine in the development of *Rhizobium leguminosarum* in the rhizosphere of pea seedlings’, *Plant and Soil*. Springer, 42(2), pp. 381–386. Available at: http://www.jstor.org/stable/42946903.

Ellis, R. H. and Roberts, E. H. (1980) ‘Improved equations for the prediction of seed longevity’, *Annals of Botany*, 45(1), pp. 13–30. doi: 10.1093/oxfordjournals.aob.a085797.

Elyamine, A. M., Kan, J., Meng, S., Tao, P., Wang, H. and Hu, Z. (2021) ‘Aerobic and anaerobic bacterial and fungal degradation of pyrene: Mechanism pathway including biochemical reaction and catabolic genes’, *International Journal of Molecular Sciences*, 22(15). doi: 10.3390/ijms22158202.

Epstein, B., Abou-Shanab, R. A. I., Shamseldin, A., Taylor, M. R., Guhlin, J., Burghardt, L. T., Nelson, M., Sadowsky, M. J., Tiffin, P. and Oh, J. (2018) ‘Genome-wide association analyses in the model rhizobium *Ensifer meliloti*’, *mSphere*, 3(5), pp. e00386-18. doi: 10.1128/mSphere.00386-18.

Eulberg, D., Lakner, S., Golovleva, L. A. and Schlömann, M. (1998) ‘Characterization of a protocatechuate catabolic gene cluster from *Rhodococcus opacus* 1CP: Evidence for a merged enzyme with 4-carboxymuconolactone-decarboxylating and 3-oxoadipate enol-lactone-hydrolyzing activity’, *Journal of Bacteriology*, 180(5), pp. 1072–1081. doi: 10.1128/JB.180.5.1072-1081.1998.

FAOSTAT (2021) *FAOSTAT statistical database*, *Food and Agriculture Organization of the United Nations Statistical database*. [Rome] : FAO, c1997-. Available at: http://www.fao.org/faostat/en/#home.

Farhat, M. R. *et al.* (2013) ‘Genomic analysis identifies targets of convergent positive selection in drug-resistant *Mycobacterium tuberculosis*’, *Nature Genetics*, 45(10), pp. 1183–1189. doi: 10.1038/ng.2747.

Farhat, M. R., Freschi, L., Calderon, R., Ioerger, T., Snyder, M., Meehan, C. J., de Jong, B., Rigouts, L., Sloutsky, A., Kaur, D., Sunyaev, S., van Soolingen, D., Shendure, J., Sacchettini, J. and Murray, M. (2019) ‘GWAS for quantitative resistance phenotypes in *Mycobacterium tuberculosis* reveals resistance genes and regulatory regions’, *Nature Communications*, 10(1), p. 2128. doi: 10.1038/s41467-019-10110-6.

Farooq, M., Basra, S. M. A., Ahmad, N. and Hafeez, K. (2005) ‘Thermal hardening: a new seed vigor enhancement tool in rice’, *Journal of Integrative Plant Biology*, 47(2), pp. 187–193. doi: 10.1111/j.1744-7909.2005.00031.x.

Felsenstein, J. (1985) ‘Confidence limits on phylogenies: An approach using the bootstrap’, *Evolution*. John Wiley & Sons, Ltd, 39(4), pp. 783–791. doi: 10.1111/j.1558-5646.1985.tb00420.x.

Ferguson, B. J., Indrasumunar, A., Hayashi, S., Lin, M., Lin, Y., Reid, D. E. and Gresshoff, P. M. (2010) ‘Molecular analysis of legume nodule development and autoregulation’, *Journal of Integrative Plant Biology*, 52(1), pp. 61–76. doi: 10.1111/j.1744-7909.2010.00899.x.

Ferguson, B. J., Mens, C., Hastwell, A. H., Zhang, M., Su, H., Jones, C. H., Chu, X. and Gresshoff, P. M. (2019) ‘Legume nodulation: The host controls the party’, *Plant, Cell & Environment*, 42(1), pp. 41–51. doi: https://doi.org/10.1111/pce.13348.

Ferreira, H., Vasconcelos, M., Gil, A. M. and Pinto, E. (2021) ‘Benefits of pulse consumption on metabolism and health: A systematic review of randomized controlled trials’, *Critical Reviews in Food Science and Nutrition*. Taylor & Francis, 61(1), pp. 85–96. doi: 10.1080/10408398.2020.1716680.

Fields, B., Moffat, E. K., Friman, V.-P. and Harrison, E. (2021) ‘The impact of intra-specific diversity in the rhizobia-legume symbiosis’, *Microbiology*. Microbiology Society, 167(4). doi: https://doi.org/10.1099/mic.0.001051.

Figueiredo, M. V. B., Burity, H. A., Martínez, C. R. and Chanway, C. P. (2008) ‘Alleviation of drought stress in the common bean (*Phaseolus vulgaris* L.) by co-inoculation with *Paenibacillus polymyxa* and *Rhizobium tropici*’, *Applied Soil Ecology*, 40(1), pp. 182–188. doi: https://doi.org/10.1016/j.apsoil.2008.04.005.

Flores-Félix, J. D., Carro, L., Cerda-Castillo, E., Squartini, A., Rivas, R. and Velázquez, E. (2020) ‘Analysis of the interaction between *Pisum sativum* L. and *Rhizobium laguerreae* strains nodulating this legume in northwest Spain’, *Plants*, 9(12). doi: 10.3390/plants9121755.

Fonouni-Farde, C., Kisiala, A., Brault, M., Emery, R. J. N., Diet, A. and Frugier, F. (2017) 'DELLA1-Mediated gibberellin signaling regulates cytokinin-dependent symbiotic nodulation', *Plant Physiology*, 175(4), pp. 1795–1806. doi: 10.1104/pp.17.00919.

Food and Drug Administration (FDA) (2015) *Water activity (aw) in foods*. Available at: https://www-fda-gov.ezproxy.stir.ac.uk/iceci/inspections/inspectionguides/inspectiontechnicalguides/ucm072916.htm (Accessed: 20 April 2018).

Forchhammer, K. (2007) 'Glutamine signalling in bacteria', *Front Biosci*, 12(2069), pp. 10–2741.

Fox, J. and Weisberg, S. (2019) *An {R} Companion to Applied Regression*. Third. Thousand Oaks {CA}: Sage. Available at: https://socialsciences.mcmaster.ca/jfox/Books/Companion/.

Freitag, S., Verrall, S. R., Pont, S. D. A., McRae, D., Sungurtas, J. A., Palau, R., Hawes, C., Alexander, C. J., Allwood, J. W., Foito, A., Stewart, D. and Shepherd, L. V. T. (2018) 'Impact of conventional and integrated management systems on the water-soluble vitamin content in potatoes, field beans, and cereals', *Journal of Agricultural and Food Chemistry*, 66(4), pp. 831–841. doi: 10.1021/acs.jafc.7b03509.

Friesen, T., Hill, G., Pugsley, T. and Holloway, G. (2004) 'Optimization of the convective air drying of *Penicillium bilaii* for improved efficiency', *Drying Technology*, 22(5), pp. 1153–1172. doi: 10.1081/drt-120038585.

Friesen, T., Hill, G., Pugsley, T., Holloway, G. and Zimmerman, D. (2005) 'Experimental determination of viability loss of *Penicillium bilaiae* conidia during convective air-drying', *Applied Microbiology and Biotechnology*, 68(3), pp. 397–404. doi: 10.1007/s00253-004-1866-1.

Friesen, T. J., Holloway, G., Hill, G. A. and Pugsley, T. S. (2006) 'Effect of conditions and protectants on the survival of *Penicillium bilaiae* during storage', *Biocontrol Science and Technology*, 16(1), pp. 89–98. doi: 10.1080/09583150500258263.

Galloway, J. N. and Cowling, E. B. (2002) 'Reactive nitrogen and the World: 200 years of change', *AMBIO: A Journal of the Human Environment*. Royal Swedish Academy of Sciences, 31(2), pp. 64–71. doi: 10.1579/0044-7447-31.2.64.

Galloway, J. N., Townsend, A. R., Willem, E. J., Mateete, B., Zucong, C., 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*. American Association for the Advancement of Science, 320(5878), pp. 889–892. doi: 10.1126/science.1136674.

Gao, X., Lu, X., Wu, M., Zhang, H., Pan, R., Tian, J., Li, S. and Liao, H. (2012) 'Co-inoculation with rhizobia and AMF inhibited soybean red crown rot: From field study to plant defense-related gene expression analysis', *PLOS ONE*. Public Library of Science, 7(3), pp. 1–10. doi: 10.1371/journal.pone.0033977.

García, A. H. (2011) 'Anhydrobiosis in bacteria: From physiology to applications', *Journal of Biosciences*, 36(5), pp. 939–950. doi: 10.1007/s12038-011-9107-0.

Garrison, E. and Marth, G. (2012) 'Haplotype-based variant detection from short-read sequencing'.

Gibbs, P. A., Chambers, B. J., Chaudri, A. M., McGrath, S. P., Carlton-Smith, C. H., Bacon, J. R., Campbell, C. D. and Aitken, M. N. (2006) 'Initial results from a long-term, multi-site field study of the effects on soil fertility and microbial activity of sludge cakes containing heavy metals', *Soil Use and Management*, 22(1), pp. 11–21. doi: https://doi.org/10.1111/j.1475-2743.2006.00003.x.

Gilbert, K. B., Vanderlinde, E. M. and Yost, C. K. (2007) 'Mutagenesis of the carboxy terminal protease *CtpA* decreases desiccation tolerance in *Rhizobium leguminosarum*', *FEMS Microbiology Letters*, 272(1), pp. 65–74. doi: 10.1111/j.1574-6968.2007.00735.x.

Giller, K. E., Herridge, D. F. and Sprent, J. I. (2016) 'The legume-rhizobia symbiosis and assessing the need ot inoculate', in Howieson, J. G. and Dilworth, M. J. (eds) *Working with rhizobia*. Australian Centre for International Agricultural Research (ACIAR).

Glenn, A. R., Poole, P. S. and Hudman, J. F. (1980) ‘Succinate uptake by free-living and bacteroid forms of *Rhizobium leguminosarum*’, *Microbiology*. Microbiology Society, 119(1), pp. 267–271. doi: https://doi.org/10.1099/00221287-119-1-267.

Goss, G. R., Baldwin, H. M. and Riepl, R. G. (2003) ‘Clays as biological carriers’, in Downer, RA and Mueninghoff, JC and Volgas, G. (ed.) *Pesticide formulations and delivery systems: meeting the challenges of the current crop protection industry*. American Society for Testing and Materials, pp. 24–34. doi: 10.1520/STP11110S.

Gourion, B., Berrabah, F., Ratet, P. and Stacey, G. (2015) ‘Rhizobium–legume symbioses: the crucial role of plant immunity’, *Trends in Plant Science*. Elsevier Current Trends, 20(3), pp. 186–194. doi: 10.1016/J.TPLANTS.2014.11.008.

Graham, P. H. (2008) ‘Ecology of the root-nodule bacteria of legumes’, in Dilworth, M. J., James, E. K., Sprent, J. I., and Newton, W. E. (eds) *Nitrogen-fixing Leguminous Symbioses*. Dordrecht: Springer Netherlands, pp. 23–58. doi: 10.1007/978-1-4020-3548-7_2.

Graham, P. H. and Vance, C. P. (2003) ‘Legumes: Importance and constraints to greater use’, *Plant Physiology*. American Society of Plant Biologists, 131(3), pp. 872–877. doi: 10.1104/pp.017004.

Griesmann, M. *et al.* (2018) ‘Phylogenomics reveals multiple losses of nitrogen-fixing root nodule symbiosis’, *Science*, 361(6398), p. eaat1743. doi: 10.1126/science.aat1743.

Griffiths, B. S. and Philippot, L. (2013) ‘Insights into the resistance and resilience of the soil microbial community’, *FEMS Microbiology Reviews*, 37(2), pp. 112–129. doi: 10.1111/j.1574-6976.2012.00343.x.

Gruber, N. and Galloway, J. N. (2008) ‘An Earth-system perspective of the global nitrogen cycle’, *Nature*. Nature Publishing Group, 451, p. 293. Available at: http://dx.doi.org/10.1038/nature06592.

Hall, C., Hillen, C. and Garden Robinson, J. (2017) ‘Composition, nutritional value, and health benefits of pulses’, *Cereal Chemistry*, 94(1), pp. 11–31. doi: https://doi.org/10.1094/CCHEM-03-16-0069-FI.

Hall, T. A. (1999) ‘BioEdit: a user-friendly biologival sequence alignemtn editor and analysis program for Windows 95/98/NT’, in. Nucleic Acids Symposium Series No. 41, p. 4. Available at: http://jwbrown.mbio.ncsu.edu/JWB/papers/1999Hall1.pdf.

van Ham, R., O’Callaghan, M., Geurts, R., Ridgway, H. J., Ballard, R., Noble, A., Macara, G. and Wakelin, S. A. (2016) ‘Soil moisture deficit selects for desiccation tolerant *Rhizobium leguminosarum* bv. *trifolii*’, *Applied Soil Ecology*, 108, pp. 371–380. doi: 10.1016/j.apsoil.2016.09.016.

Hawes, C., Alexander, C. J., Begg, G. S., Iannetta, P. P. M., Karley, A. J., Squire, G. R. and Young, M. (2018) ‘Plant responses to an integrated cropping system designed to maintain yield whilst enhancing soil properties and biodiversity’, *Agronomy*. Multidisciplinary Digital Publishing Institute, 8(10), p. 229.

Hawes, C., Begg, G. S., Iannetta, P. P. M., Karley, A. J. and Squire, G. R. (2016) ‘A whole-systems approach for assessing measures to improve arable ecosystem sustainability’, *Ecosystem Health and Sustainability*. Taylor & Francis, 2(12), p. e01252. doi: 10.1002/ehs2.1252.

Hawes, C., Young, M. W., Banks, G., Begg, G. S., Christie, A., Iannetta, P. P. M., Karley, A. J. and Squire, G. R. (2019) ‘Whole-systems analysis of environmental and economic sustainability in arable cropping systems: A case study’, *Agronomy*. Multidisciplinary Digital Publishing Institute, 9(8), p. 438.

Henle, K., Alard, D., Clitherow, J., Cobb, P., Firbank, L., Kull, T., McCracken, D., Moritz, R. F. A., Niemelä, J., Rebane, M., Wascher, D., Watt, A. and Young, J. (2008) ‘Identifying and managing the conflicts between agriculture and biodiversity conservation in Europe–A review’, *Agriculture, Ecosystems & Environment*, 124(1), pp. 60–71. doi: https://doi.org/10.1016/j.agee.2007.09.005.

Herold, M. B., Giles, M. E., Alexander, C. J., Baggs, E. M. and Daniell, T. J. (2018) ‘Variable response of *nirK* and *nirS* containing denitrifier communities to long-term pH manipulation and cultivation’, *FEMS Microbiology Letters*, 365(7). doi: 10.1093/femsle/fny035.

Herridge, D. F. (2008) ‘Inoculation technology for legumes’, in Dilworth, M. J., James, E. K., Sprent, J. I., and Newton, W. E. (eds) *Nitrogen-fixing Leguminous Symbioses*. Dordrecht: Springer Netherlands, pp. 77–115. doi: 10.1007/978-1-4020-3548-7_4.

Herridge, D. F., Peoples, M. B. and Boddey, R. M. (2008) ‘Global inputs of biological nitrogen fixation in agricultural systems’, *Plant and Soil*, 311(1), pp. 1–18. doi: 10.1007/s11104-008-9668-3.

Hirsch, P. R. (1996) ‘Population dynamics of indigenous and genetically modified rhizobia in the field’, *New Phytologist*, 133(1), pp. 159–171. doi: 10.1111/j.1469-8137.1996.tb04351.x.

Hirsch, P. R. and Spokes, J. D. (1994) ‘Survival and dispersion of genetically modified rhizobia in the field and genetic interactions with native strains’, *FEMS Microbiology Ecology*, 15(1–2), pp. 147–159. doi: 10.1111/j.1574-6941.1994.tb00239.x.

Hothorn, T., Bretz, F. and Westfall, P. (2008) ‘Simultaneous inference in general parametric models’, *Biometrical Journal*, 50(3), pp. 346–363.

Howieson, J. G. and Dilworth, M. J. (2016) *Working with rhizobia*. Canberra : Australian Centre for International Agricultural Research. Available at: http://aciar.gov.au/publication/mn173.

Howieson, J. G., Malden, J., Yates, R. J. and O’Hara, G. W. (2000) ‘Techniques for the selection and development of elite inoculant strains of *Rhizobium leguminosarum* in Southern Australia’, *Symbiosis*. Balaban Publishers.

Huang, L.-K. and Wang, M.-J. J. (1995) ‘Image thresholding by minimizing the measures of fuzziness’, *Pattern Recognition*. Pergamon, 28(1), pp. 41–51. doi: 10.1016/0031-3203(94)E0043-K.

Humann, J. L. and Kahn, M. L. (2015) ‘Genes involved in desiccation resistance of rhizobia and other bacteria’, in *Biological Nitrogen Fixation*. John Wiley & Sons, Ltd, pp. 397–404. doi: https://doi.org/10.1002/9781119053095.ch39.

Humann, J. L., Ziemkiewicz, H. T., Yurgel, S. N. and Kahn, M. L. (2009) ‘Regulatory and DNA repair genes contribute to the desiccation resistance of *Sinorhizobium meliloti* Rm1021’, *Applied and Environmental Microbiology*, 75(2), pp. 446–453. doi: 10.1128/AEM.02207-08.

Hungria, M., Nogueira, M. A. and Araujo, R. S. (2013) ‘Co-inoculation of soybeans and common beans with rhizobia and azospirilla: strategies to improve sustainability’, *Biology and Fertility of Soils*, 49(7), pp. 791–801. doi: 10.1007/s00374-012-0771-5.

Hungria, M., O’Hara, G. W., Zilli, J. E., Araujo, R. S., Deaker, R. and Howieson, J. G. (2016) ‘Isolation and growth of rhizobia’, in Howieson, J. G. and Dilworth, M. J. (eds) *Working with rhizobia*. Australian Centre for International Agricultural Research (ACIAR).

Iannetta, P. P. M., Hawes, C., Begg, G. S., Maaß, H., Ntatsi, G., Savvas, D., Vasconcelos, M., Hamann, K., Williams, M., Styles, D., Toma, L., Shrestha, S., Balázs, B., Kelemen, E., Debeljak, M., Trajanov, A., Vickers, R. and Rees, R. M. (2021) ‘A multifunctional solution for wicked problems: Value-chain wide facilitation of legumes cultivated at bioregional scales is necessary to address the climate-biodiversity-nutrition nexus’, *Frontiers in Sustainable Food Systems*, 5, p. 239. doi: 10.3389/fsufs.2021.692137.

Iannetta, P. P. M., Young, M., Bachinger, J., Bergkvist, G., Doltra, J., Lopez-Bellido, R. J., Monti, M., Pappa, V. A., Reckling, M., Topp, C. F. E., Walker, R. L., Rees, R. M., Watson, C. A., James, E. K., Squire, G. R. and Begg, G. S. (2016) ‘A comparative nitrogen balance and productivity analysis of legume and non-legume supported cropping systems: The potential role of biological nitrogen fixation’, *Frontiers in Plant Science*, 7, p. 1700. doi: 10.3389/fpls.2016.01700.

Jaccard, P. (1901) ‘Distribution de la flore alpine dans le bassin des Dranses et dans quelques régions voisines’, *Bulletin de la Société Vaudoise des Sciences Naturelles*, 37, pp. 241–272.

Jain, C., Rodriguez-R, L. M., Phillippy, A. M., Konstantinidis, K. T. and Aluru, S. (2018) ‘High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries’, *Nature Communications*, 9(1), p. 5114. doi: 10.1038/s41467-018-07641-9.

James, E. K. (2017) ‘Nitrogen fixation’, in Thomas, B., Murray, B. G., and Murphy, D. J. (eds) *Encyclopedia of Applied Plant Sciences (Second Edition)*. Second Edi. Oxford: Academic Press, pp. 271–277. doi: https://doi.org/10.1016/B978-0-12-394807-6.00124-6.

Jarvis, B. D. W., Ward, L. J. H. and Slade, E. A. (1989) ‘Expression by soil bacteria of nodulation genes from Rhizobium leguminosarum biovar trifolii’, *Applied and Environmental Microbiology*. American Society for Microbiology Journals, 55(6), pp. 1426–1434. Available at: https://aem.asm.org/content/55/6/1426.

Javaux, C., Joris, B. and De Witte, P. (2007) ‘Functional characteristics of *TauA* binding protein from *TauABC Escherichia coli* system’, *The Protein Journal*, 26(4), pp. 231–238. doi: 10.1007/s10930-006-9064-x.

Jensen, E. S., Peoples, M. B., Boddey, R. M., Gresshoff, P. M., Hauggaard-Nielsen, H., J.R. Alves, B. and Morrison, M. J. (2012) ‘Legumes for mitigation of climate change and the provision of feedstock for biofuels and biorefineries. A review’, *Agronomy for Sustainable Development*, 32(2), pp. 329–364. doi: 10.1007/s13593-011-0056-7.

Jensen, E. S., Peoples, M. B. and Hauggaard-Nielsen, H. (2010) ‘Faba bean in cropping systems’, *Field Crops Research*, 115(3), pp. 203–216. doi: https://doi.org/10.1016/j.fcr.2009.10.008.

Jha, C. K. and Saraf, M. (2012) ‘Evaluation of multispecies plant-growth-promoting consortia for the growth promotion of *Jatropha curcas* L.’, *Journal of Plant Growth Regulation*, 31(4), pp. 588–598. doi: 10.1007/s00344-012-9269-5.

Jiao, Y. S., Yan, H., Ji, Z. J., Liu, Y. H., Sui, X. H., Wang, E. T., Guo, B. L., Chen, W. X. and Chen, W. F. (2015) ‘*Rhizobium sophorae* sp. nov. and *Rhizobium sophoriradicis* sp. nov., nitrogen-fixing rhizobial symbionts of the medicinal legume *Sophora flavescens*’, *International Journal of Systematic and*

Evolutionary Microbiology. Microbiology Society, 65(Pt_2), pp. 497–503. doi: https://doi.org/10.1099/ijs.0.068916-0.

John, R. P., Tyagi, R. D., Brar, S. K., Surampalli, R. Y. and Prévost, D. (2011) ‘Bio-encapsulation of microbial cells for targeted agricultural delivery’, *Critical Reviews in Biotechnology*, 31(3), pp. 211–226. doi: 10.3109/07388551.2010.513327.

Johnston, A. W. B. and Beringer, J. E. (1975) ‘Identification of the Rhizobium strains in pea root nodules using genetic markers’, *Microbiology*. Microbiology Society, 87(2), pp. 343–350. doi: https://doi.org/10.1099/00221287-87-2-343.

Johnston, A. W. B., Todd, J. D., Curson, A. R., Lei, S., Nikolaidou-Katsaridou, N., Gelfand, M. S. and Rodionov, D. A. (2007) ‘Living without Fur: the subtlety and complexity of iron-responsive gene regulation in the symbiotic bacterium *Rhizobium* and other α-proteobacteria’, *BioMetals*, 20(3), pp. 501–511. doi: 10.1007/s10534-007-9085-8.

Jorrin, B., Palacios, J. M., Peix, Á. and Imperial, J. (2020) ‘*Rhizobium ruizarguesonis* sp. nov., isolated from nodules of *Pisum sativum* L’, *Systematic and Applied Microbiology*, 43(4), p. 126090. doi: https://doi.org/10.1016/j.syapm.2020.126090.

Ju, W., Liu, L., Fang, L., Cui, Y., Duan, C. and Wu, H. (2019) ‘Impact of co-inoculation with plant-growth-promoting rhizobacteria and rhizobium on the biochemical responses of alfalfa-soil system in copper contaminated soil’, *Ecotoxicology and Environmental Safety*, 167, pp. 218–226. doi: https://doi.org/10.1016/j.ecoenv.2018.10.016.

Kahm, M., Hasenbrink, G., Lichtenberg-Frat’e, H., Ludwig, J. and Kschischo, M. (2010) ‘{grofit}: Fitting biological growth curves with {R}’, *Journal of Statistical Software*, 33(7), pp. 1–21. Available at: http://www.jstatsoft.org/v33/i07/.

Kannaiah, S., Livny, J. and Amster-Choder, O. (2019) ‘Spatiotemporal organization of the *E. coli* transcriptome: Translation independence and engagement in regulation’, *Molecular Cell*, 76(4), pp. 574-589.e7. doi: https://doi.org/10.1016/j.molcel.2019.08.013.

Kassambara, A. and Mundt, F. (2020) ‘factoextra: Extract and Visualize the Results of Multivariate Data Analyses’. Available at: https://cran.r-project.org/package=factoextra.

Kawai, Y., Mercier, R., Wu, L. J., Domínguez-Cuevas, P., Oshima, T. and Errington, J. (2015) ‘Cell growth of wall-free L-form bacteria is limited by oxidative damage’, *Current Biology*, 25(12), pp. 1613–1618. doi: https://doi.org/10.1016/j.cub.2015.04.031.

Keyser, H. H. and Li, F. (1992) ‘Potential for increasing biological nitrogen fixation in soybean’, *Plant and Soil*, 141(1–2), pp. 119–135. doi: 10.1007/BF00011313.

Khalid, A., Arshad, M. and Zahir, Z. A. (2004) ‘Screening plant growth-promoting rhizobacteria for improving growth and yield of wheat’, *Journal of Applied Microbiology*, 96(3), pp. 473–480. doi: 10.1046/j.1365-2672.2003.02161.x.

Kiers, E. T., West, S. K. and Denison, R. F. (2008) ‘Maintaining cooperation in the legume-rhizobia symbiosis: Identifying selection pressures and mechanisms’, in Dilworth, M. J., James, E. K., Sprent, J. I., and Newton, W. E. (eds) *Nitrogen-fixing Leguminous Symbioses*, pp. 59–76. doi: 10.1007/978-1-4020-3548-7_3.

Kim, J., Jeon, C. O. and Park, W. (2008) ‘Dual regulation of *zwf*-1 by both 2-keto-3-deoxy-6-phosphogluconate and oxidative stress in *Pseudomonas putida*’, *Microbiology*. Microbiology Society, 154(12), pp. 3905–3916. doi: https://doi.org/10.1099/mic.0.2008/020362-0.

Kleijn, D., Baquero, R. A., Clough, Y., Díaz, M., De Esteban, J., Fernández, F., Gabriel, D., Herzog, F., Holzschuh, A., Jöhl, R., Knop, E., Kruess, A., Marshall, E. J. P., Steffan-Dewenter, I., Tscharntke, T., Verhulst, J., West, T. M. and Yela, J. L. (2006) ‘Mixed biodiversity benefits of agri-environment schemes in five European countries’, *Ecology Letters*, 9(3), pp. 243–254. doi: https://doi.org/10.1111/j.1461-0248.2005.00869.x.

Kleijn, D. and Sutherland, W. J. (2003) ‘How effective are European agri-environment schemes in conserving and promoting biodiversity?’, *Journal of*

Applied Ecology, 40(6), pp. 947–969. doi: https://doi.org/10.1111/j.1365-2664.2003.00868.x.

Knaus, B. J. and Grünwald, N. J. (2017) ‘vcfr: a package to manipulate and visualize variant call format data in R’, *Molecular Ecology Resources*, 17(1), pp. 44–53. doi: https://doi.org/10.1111/1755-0998.12549.

Knott, C. M. (1987) ‘A key for stages of development of the pea (*Pisum sativum*)’, *Annals of Applied Biology*, 111(1), pp. 233–245. doi: 10.1111/j.1744-7348.1987.tb01450.x.

Kobayashi, H. and Broughton, W. J. (2008) ‘Fine-tuning of symbiotic genes in rhizobia: Flavonoid signal transduction cascade’, in Dilworth, M. J., James, E. K., Sprent, J. I., and Newton, W. E. (eds) *Nitrogen-fixing Leguminous Symbioses*. Dordrecht: Springer Netherlands, pp. 117–152. doi: 10.1007/978-1-4020-3548-7_5.

Koch, T. and Dahl, C. (2018) ‘A novel bacterial sulfur oxidation pathway provides a new link between the cycles of organic and inorganic sulfur compounds’, *The ISME Journal*, 12(10), pp. 2479–2491. doi: 10.1038/s41396-018-0209-7.

Koenen, E. J. M., Ojeda, D. I., Steeves, R., Migliore, J., Bakker, F. T., Wieringa, J. J., Kidner, C., Hardy, O. J., Pennington, R. T., Bruneau, A. and Hughes, C. E. (2020) ‘Large-scale genomic sequence data resolve the deepest divergences in the legume phylogeny and support a near-simultaneous evolutionary origin of all six subfamilies’, *New Phytologist*, 225(3), pp. 1355–1369. doi: https://doi.org/10.1111/nph.16290.

Kohler, P. R. A., Zheng, J. Y., Schoffers, E. and Rossbach, S. (2010) ‘Inositol catabolism, a key pathway in Sinorhizobium meliloti for competitive host nodulation’, *Applied and Environmental Microbiology*, 76(24), pp. 7972–7980. doi: 10.1128/AEM.01972-10.

Kopac, S., Wang, Z., Wiedenbeck, J., Sherry, J., Wu, M. and Cohan, F. M. (2014) ‘Genomic heterogeneity and ecological speciation within one subspecies of *Bacillus subtilis*’, *Applied and environmental microbiology*. Am Soc Microbiol, 80(16), pp. 4842–4853.

Köpke, U. and Nemecek, T. (2010) ‘Ecological services of faba bean’, *Field Crops Research*, 115(3), pp. 217–233. doi: https://doi.org/10.1016/j.fcr.2009.10.012.

Kosanke, J. W., Osburn, R. M., Shuppe, G. I. and Smith, R. S. (1992) ‘Slow rehydration improves the recovery of dried bacterial-populations’, *Canadian Journal of Microbiology*. NATL RESEARCH COUNCIL CANADA, 38(6), pp. 520–525. doi: 10.1139/m92-086.

Kosterin, O. E. (2014) ‘The lost ancestor of the broad bean (*Vicia faba* L.) and the origin of plant cultivation in the Near East’, *Vavilov Journal of Genetics and Breeding*, 18(4/1), pp. 831–840. doi: https://doi.org/10.18699/VJ15.118.

Kozar, S. F., Symonenko, E. P., Volkohon, V. V and Volkogon, M. V (2019) ‘Nanocarboxylates of molybdenum and of iron enhance the functional activity of *Rhizobium radiobacter* 204’, *Applied Nanoscience*, 9(5), pp. 795–800. doi: 10.1007/s13204-018-00939-6.

Kozlov, A. M., Darriba, D., Flouri, T., Morel, B. and Stamatakis, A. (2019) ‘RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference’, *Bioinformatics*, 35(21), pp. 4453–4455. doi: 10.1093/bioinformatics/btz305.

Kucey, R. M. N. and Hynes, M. F. (1989) ‘Populations of *Rhizobium leguminosarum* biovars *phaseoli* and *viceae* in fields after bean or pea in rotation with nonlegumes’, *Canadian Journal of Microbiology*, 35(6), pp. 661–667. doi: 10.1139/m89-107.

Kumar, N., Lad, G., Giuntini, E., Kaye, M. E., Udomwong, P., Shamsani, N. J., Young, J. P. W. and Bailly, X. (2015) ‘Bacterial genospecies that are not ecologically coherent: population genomics of *Rhizobium leguminosarum*’, *Open Biology*, 5(1), p. 140133. doi: 10.1098/rsob.140133.

Kumar, P., Pandey, P., Dubey, R. C. and Maheshwari, D. K. (2016) ‘Bacteria consortium optimization improves nutrient uptake, nodulation, disease suppression and growth of the common bean (*Phaseolus vulgaris*) in both pot and field studies’, *Rhizosphere*, 2, pp. 13–23. doi: https://doi.org/10.1016/j.rhisph.2016.09.002.

Kumar, S. and Filipski, A. (2007) ‘Multiple sequence alignment: in pursuit of homologous DNA positions’, *Genome research*. Cold Spring Harbor Lab, 17(2), pp. 127–135.

Kumar, S., Stecher, G., Li, M., Knyaz, C. and Tamura, K. (2018) ‘MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms’, *Molecular Biology and Evolution*, (35), pp. 1547–1549.

Kumar, Stecher, G. and Tamura, K. (2016) ‘MEGA7: Molecular evolutionary genetics analysis Version 7.0 for bigger datasets’, *Molecular Biology and Evolution*, 33(7), pp. 1870–1874.

Laguerre, G., Louvrier, P., Allard, M. R. and Amarger, N. (2003) ‘Compatibility of rhizobial genotypes within natural populations of *Rhizobium leguminosarum* biovar *viciae* for nodulation of host legumes’, *Applied and Environmental Microbiology*, 69(4), pp. 2276–2283. doi: 10.1128/AEM.69.4.2276-2283.2003.

Laguerre, G., Nour, S. M., Macheret, V., Sanjuan, J., Drouin, P. and Amarger, N. (2001) ‘Classification of rhizobia based on *nodC* and *nifH* gene analysis reveals a close phylogenetic relationship among *Phaseolus vulgaris* symbionts’, *Microbiology*. Microbiology Society, 147(4), pp. 981–993. doi: https://doi.org/10.1099/00221287-147-4-981.

de Lajudie, P. M. *et al.* (2019) ‘Minimal standards for the description of new genera and species of rhizobia and agrobacteria’, *International Journal of Systematic and Evolutionary Microbiology*. Microbiology Society, 69(7), pp. 1852–1863. doi: https://doi.org/10.1099/ijsem.0.003426.

de Lajudie, P. M. and Young, J. P. W. (2017) ‘International committee on systematics of Prokaryotes subcommittee for the taxonomy of *Rhizobium* and *Agrobacterium* Minutes of the meeting, Budapest, 25 August 2016’, *International Journal of Systematic and Evolutionary Microbiology*. Microbiology Society, 67(7), pp. 2485–2494. doi: https://doi.org/10.1099/ijsem.0.002144.

Langmead, B. and Salzberg, S. L. (2012) ‘Fast gapped-read alignment with Bowtie 2’, *Nature Methods*, 9(4), pp. 357–359. doi: 10.1038/nmeth.1923.

Laranjo, M., Alexandre, A. and Oliveira, S. (2017) ‘Global transcriptional response to salt shock of the plant microsymbiont *Mesorhizobium loti* MAFF303099’, *Research in Microbiology*, 168(1), pp. 55–63. doi: 10.1016/j.resmic.2016.07.006.

Lechner, M., Findeiß, S., Steiner, L., Marz, M., Stadler, P. F. and Prohaska, S. J. (2011) ‘Proteinortho: Detection of (co-)orthologs in large-scale analysis’, *BMC Bioinformatics*, 12(1), p. 124. doi: 10.1186/1471-2105-12-124.

Li, D., Kinkema, M. and Gresshoff, P. M. (2009) ‘Autoregulation of nodulation (AON) in *Pisum sativum* (pea) involves signalling events associated with both nodule primordia development and nitrogen fixation’, *Journal of Plant Physiology*, 166(9), pp. 955–967. doi: https://doi.org/10.1016/j.jplph.2009.03.004.

Li, H.-L., Wang, W., Mortimer, P. E., Li, R.-Q., Li, D.-Z., Hyde, K. D., Xu, J.-C., Soltis, D. E. and Chen, Z.-D. (2015) ‘Large-scale phylogenetic analyses reveal multiple gains of actinorhizal nitrogen-fixing symbioses in angiosperms associated with climate change’, *Scientific Reports*, 5(1), p. 14023. doi: 10.1038/srep14023.

Lindström, K. and Young, J. P. W. (2011) ‘International committee on systematics of Prokaryotes subcommittee on the taxonomy of *Agrobacterium* and *Rhizobium*: minutes of the meeting, 7 September 2010, Geneva, Switzerland’, *International journal of systematic and evolutionary microbiology*. Microbiology Society, 61(12), pp. 3089–3093.

Lira, M. d. A. and Smith, D. L. (2000) ‘Use of a standard TWAIN scanner and software for nodule number determination on different legume species’, *Soil Biology and Biochemistry*, 32(10), pp. 1463–1467. doi: https://doi.org/10.1016/S0038-0717(00)00052-3.

Liu, S., Rojas, R. V, Gray, P., Zhu, M.-J. and Tang, J. (2018) ‘*Enterococcus faecium* as a *Salmonella* surrogate in the thermal processing of wheat flour: Influence of water activity at high temperatures’, *Food Microbiology*, 74, pp. 92–99. doi: https://doi.org/10.1016/j.fm.2018.03.001.

López-García, S. L., Vázquez, T. E. E., Favelukes, G. and Lodeiro, A. R. (2002) ‘Rhizobial position as a main determinant in the problem of competition for nodulation in soybean’, *Environmental Microbiology*, 4(4), pp. 216–224. doi: 10.1046/j.1462-2920.2002.00287.x.

LPWG (2013) ‘Legume phylogeny and classification in the 21st century: progress, prospects and lessons for other species-rich clades’, *Taxon*, 62(2), pp. 217–248. doi: doi.org/10.5167/uzh-78167.

LPWG (2017) ‘A new subfamily classification of the Leguminosae based on a taxonomically comprehensive phylogeny – The Legume Phylogeny Working Group (LPWG)’, *Taxon*, 66(1), pp. 44–77. doi: 10.12705/661.3.

Luo, D., Ganesh, S. and Koolaard, J. (2020) ‘predictmeans: Calculate predicted means for linear models’. Available at: https://cran.r-project.org/package=predictmeans.

Ma, J.-F., Hager, P. W., Howell, M. L., Phibbs, P. V and Hassett, D. J. (1998) ‘Cloning and characterization of the *Pseudomonas aeruginosa zwf* gene encoding Glucose-6-Phosphate Dehydrogenase, an enzyme important in resistance to Methyl Viologen (Paraquat)’, *Journal of Bacteriology*, 180(7), pp. 1741–1749. doi: 10.1128/JB.180.7.1741-1749.1998.

Macdonald, C. A., Clark, I. M., Hirsch, P. R., Zhao, F.-J. and McGrath, S. P. (2011) ‘Development of a real-time PCR assay for detection and quantification of *Rhizobium leguminosarum* bacteria and discrimination between different biovars in Zinc-contaminated soil’, *Applied and Environmental Microbiology*. American Society for Microbiology Journals, 77(13), pp. 4626–4633. doi: 10.1128/AEM.02232-10.

Magallon, S. and Sanderson, M. J. (2001) ‘Absolute diversification rates in angiosperm clades’, *Evolution*, 55(9), pp. 1762–1780. doi: 10.1111/j.0014-3820.2001.tb00826.x.

Mahmood, A., Turgay, O. C., Farooq, M. and Hayat, R. (2016) ‘Seed biopriming with plant growth promoting rhizobacteria: a review’, *FEMS Microbiology Ecology*. Edited by G. Muyzer, 92(8), p. fiw112. doi: 10.1093/femsec/fiw112.

Maluk, M., Ferrando-Molina, F., Lopez del Egido, L., Langarica-Fuentes, A., Yohannes, G. G., Young, M. W., Martin, P., Gantlett, R., Kenicer, G., Hawes, C., Begg, G. S., Quilliam, R. S., Squire, G. R., Young, J. P. W., Iannetta, P. P. M. and James, E. K. (2022) 'Fields with no recent legume cultivation have sufficient nitrogen-fixing rhizobia for crops of faba bean (*Vicia faba* L.)', *Plant and Soil*. doi: https://doi.org/10.1007/s11104-021-05246-8.

Manikandan, R., Saravanakumar, D., Rajendran, L., Raguchander, T. and Samiyappan, R. (2010) 'Standardization of liquid formulation of *Pseudomonas fluorescens* Pf1 for its efficacy against Fusarium wilt of tomato', *Biological Control*. Academic Press, 54(2), pp. 83–89. doi: 10.1016/J.BIOCONTROL.2010.04.004.

Manzanera, M., García de Castro, A., Tøndervik, A., Rayner-Brandes, M., Strøm, A. R. and Tunnacliffe, A. (2002) 'Hydroxyectoine is superior to trehalose for anhydrobiotic engineering of *Pseudomonas putida* KT2440.', *Applied and environmental microbiology*. American Society for Microbiology, 68(9), pp. 4328–33. doi: 10.1128/AEM.68.9.4328-4333.2002.

Martens, J.-H., Barg, H., Warren, M. and Jahn, D. (2002) 'Microbial production of vitamin B12', *Applied Microbiology and Biotechnology*, 58(3), pp. 275–285. doi: 10.1007/s00253-001-0902-7.

Matthews, P. and Marcellos, H. (2003) 'Faba bean', in *Agfact P4.2.7*. Second. Division of Plant Industries, New South Wales Agriculture, pp. 1–12. Available at: http://www.dpi.nsw.gov.au/__data/assets/pdf_file/0004/157729/faba-bean-pt1.pdf.

Mauchline, T. H., Hayat, R., Clark, I. M. and Hirsch, P. R. (2018) 'Old meets new: most probable number validation of metagenomic and metatranscriptomic datasets in soil', *Letters in Applied Microbiology*, 66(1), pp. 14–18. doi: 10.1111/lam.12821.

Maunoury, N., Kondorosi, A., Kondorosi, E. and Mergaert, P. (2008) 'Cell biology of nodule infection and development', in Dilworth, M. J., James, E. K., Sprent, J. I., and Newton, W. E. (eds) *Nitrogen-fixing Leguminous Symbioses*.

Dordrecht: Springer Netherlands, pp. 153–189. doi: 10.1007/978-1-4020-3548-7_6.

McIntyre, H. J., Davies, H., Hore, T. A., Miller, S. H., Dufour, J.-P. and Ronson, C. W. (2007) ‘Trehalose biosynthesis in *Rhizobium leguminosarum* bv. *trifolii* and its role in desiccation tolerance’, *Applied and Environmental Microbiology*, 73(12), pp. 3984–3992. doi: 10.1128/AEM.00412-07.

Mendoza-Suárez, M. A., Geddes, B. A., Sánchez-Cañizares, C., Ramírez-González, R. H., Kirchhelle, C., Jorrin, B. and Poole, P. S. (2020) ‘Optimizing rhizobium-legume symbioses by simultaneous measurement of rhizobial competitiveness and N2 fixation in nodules’, *Proceedings of the National Academy of Sciences*, 117(18), pp. 9822 LP – 9831. doi: 10.1073/pnas.1921225117.

Mendoza-Suárez, M., Andersen, S. U., Poole, P. S. and Sánchez-Cañizares, C. (2021) ‘Competition, nodule occupancy, and persistence of inoculant strains: key factors in the rhizobium-legume symbioses’, *Frontiers in Plant Science*, 12, p. 1684. doi: 10.3389/fpls.2021.690567.

MicrobesNG (2021) *MicrobesNG - Genome sequencing service methods v20210419*, *MicrobesNG*. Available at: https://microbesng.com/documents/24/MicrobesNG_Sequencing_Service_Methods_v20210419.pdf (Accessed: 7 December 2021).

Mikheenko, A., Prjibelski, A., Saveliev, V., Antipov, D. and Gurevich, A. (2018) ‘Versatile genome assembly evaluation with QUAST-LG’, *Bioinformatics*, 34(13), pp. i142–i150. doi: 10.1093/bioinformatics/bty266.

Molina-Romero, D., Baez, A., Quintero-Hernández, V., Castañeda-Lucio, M., Fuentes-Ramírez, L. E., Bustillos-Cristales, M. del R., Rodríguez-Andrade, O., Morales-García, Y. E., Munive, A. and Muñoz-Rojas, J. (2017) ‘Compatible bacterial mixture, tolerant to desiccation, improves maize plant growth’, *Plos One*. Public Library of Science, 12(11), pp. 1–21. doi: 10.1371/journal.pone.0187913.

Montes-Grajales, D., Jiménez, B., Rogel, M. A., Alagón, A., Esturau-Escofet, N., Esquivel, B., Mart\'\inez-Romero, J. and Mart\'\inez-Romero, E. (2019)

‘Nitrogen-fixing *Klebsiella variicola* in feces from herbivorous tortoises’, *bioRxiv*. Cold Spring Harbor Laboratory. doi: 10.1101/666818.

Mothapo, N. V, Grossman, J. M., Sooksa-nguan, T., Maul, J., Bräuer, S. L. and Shi, W. (2013) ‘Cropping history affects nodulation and symbiotic efficiency of distinct hairy vetch (*Vicia villosa* Roth.) genotypes with resident soil rhizobia’, *Biology and Fertility of Soils*, 49(7), pp. 871–879. doi: 10.1007/s00374-013-0781-y.

Mugnier, J. and Jung, G. (1985) ‘Survival of bacteria and fungi in relation to water activity and the solvent properties of water in biopolymer gels.’, *Applied and environmental microbiology*, 50(1), pp. 108–14. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16346829 (Accessed: 13 June 2018).

Mutch, L. A. and Young, J. P. W. (2004) ‘Diversity and specificity of *Rhizobium leguminosarum* biovar *viciae* on wild and cultivated legumes’, *Molecular Ecology*. Oxford: Blackwell Publishing LTD, 13(8), pp. 2435–2444. doi: 10.1111/j.1365-294X.2004.02259.x.

Najafi, M. B. H. and Pezeshki, P. (2014) ‘Bacterial mutation; Types, mechanisms and mutant detection methods: a review’, *European Scientific Journal, ESJ*, 9(10). doi: 10.19044/esj.2013.v9n10p%p.

Najar, I. N. and Das, S. (2015) ‘Poly-glutamic acid (PGA)-Structure, synthesis, genomic organization and its application: A Review’, *International Journal of Pharmaceutical Sciences and Research*. International Journal of Pharmaceutical Sciences and Research, 6(6), p. 2258.

Naylor, D. and Coleman-Derr, D. (2018) ‘Drought stress and root-associated bacterial communities’, *Frontiers in Plant Science*, 8, p. 2223. doi: 10.3389/fpls.2017.02223.

Nei, M. and Kumar, S. (2000) *Molecular evolution and phylogenetics*. Oxford university press.

Neudorf, K. D. and Yost, C. K. (2017) ‘An uncharacterized gene coding a conserved lytic transglycosylase domain (RL4716) is required for proper cell envelope function in *Rhizobium leguminosarum*’, *FEMS Microbiology Letters*. Oxford: Oxford University Press, 364(5). doi: 10.1093/femsle/fnx035.

Ngabonziza, J. C. S. *et al.* (2020) ‘A sister lineage of the *Mycobacterium tuberculosis* complex discovered in the African Great Lakes region’, *Nature Communications*, 11(1), p. 2917. doi: 10.1038/s41467-020-16626-6.

Nutman, P. S. and Hearne, R. (1979) ‘Persistence of nodule bacteria in soil under long-term cereal cultivation’, *Rothamsted Experimental Station Report for 1979 Part 2*, pp. 77–90.

Nutman, P. S. and Ross, G. J. (1970) ‘Rhizobium in the soils of the Rothamsted and Woburn farms.’, *Report of the Rothamsted Experimental Station, 1969*, pp. 148–167.

O’Hara, G. W., Zilli, J. E., Poole, P. . S. and Hungria, M. (2016) ‘Taxonomy and physiology of rhizobia’, in Howieson, J. G. and Dilworth, M. J. (eds) *Working with rhizobia*. Australian Centre for International Agricultural Research (ACIAR), pp. 125–144.

Oikawa, T., Tauch, A., Schaffer, S. and Fujioka, T. (2006) ‘Expression of *alr* gene from *Corynebacterium glutamicum* ATCC 13032 in *Escherichia coli* and molecular characterization of the recombinant alanine racemase’, *Journal of Biotechnology*, 125(4), pp. 503–512. doi: https://doi.org/10.1016/j.jbiotec.2006.04.002.

Oldroyd, G. E. D. (2013) ‘Speak, friend, and enter: signalling systems that promote beneficial symbiotic associations in plants’, *Nature Reviews Microbiology*, 11(4), pp. 252–263. doi: 10.1038/nrmicro2990.

Oldroyd, G. E. D., Murray, J. D., Poole, P. S. and Downie, J. A. (2011) ‘The rules of engagement in the legume-rhizobial symbiosis’, *Annual Review of Genetics*, 45(1), pp. 119–144. doi: 10.1146/annurev-genet-110410-132549.

Oliveira, M., Castro, C., Coutinho, J. and Trindade, H. (2021) ‘Grain legume-based cropping systems can mitigate greenhouse gas emissions from cereal under Mediterranean conditions’, *Agriculture, Ecosystems & Environment*, 313, p. 107406. doi: https://doi.org/10.1016/j.agee.2021.107406.

Paau, A. S. (1989) ‘Improvement of rhizobium inoculants’, *Applied and Environmental Microbiology*, 55(4), pp. 862–865.

Paradis, E. and Schliep, K. (2019) ‘ape 5.0: an environment for modern phylogenetics and evolutionary analyses in {R}’, *Bioinformatics*, 35, pp. 526–528.

Parker, C., Trinick, M. and Chatel, D. (1977) ‘Rhizobia as soil and rhizosphere inhabitants’, in Hardy, R. and Gibson, A. (eds) *A treatise on dinitrogen fixation, section IV: agronomy and ecology*. John Wiley & Sons: New York, pp. 311–352.

Pastor-Bueis, R., Sánchez-Cañizares, C., James, E. K. and González-Andrés, F. (2019) ‘Formulation of a highly effective inoculant for common bean based on an autochthonous elite strain of *Rhizobium leguminosarum* bv. *phaseoli*, and genomic-based insights into its agronomic performance’, *Frontiers in Microbiology*, 10, p. 2724. doi: 10.3389/fmicb.2019.02724.

Patel, M. A., Luthra, S., Shamblin, S. L., Arora, K. K., Krzyzaniak, J. F. and Taylor, L. S. (2018) ‘Effect of excipient properties, water activity, and water content on the disproportionation of a pharmaceutical salt’, *International Journal of Pharmaceutics*, 546(1), pp. 226–234. doi: https://doi.org/10.1016/j.ijpharm.2018.05.035.

Paul, E., Fages, J., Blanc, P., Goma, G. and Pareilleux, A. (1993) ‘Survival of alginate-entrapped cells of *Azospirillum lipoferum* during dehydration and storage in relation to water properties’, *Applied Microbiology and Biotechnology*, 40(1), pp. 34–39.

Pazos-Rojas, L. A., Muñoz-Arenas, L. C., Rodriguez-Andrade, O., López-Cruz, L. E., López-Ortega, O., Lopes-Olivares, F., Luna-Suarez, S., Baez, A., Morales-Garcia, Y. E., Quintero-Hernández, V., Villalobos-López, M. A., De la Torre, J. and Muñoz-Rojas, J. (2019) ‘Desiccation-induced viable but nonculturable state in *Pseudomonas putida* KT2440, a survival strategy’, *PloS one*. Public Library of Science San Francisco, CA USA, 14(7), p. e0219554.

Peel, M. C., Finlayson, B. L. and Mcmahon, T. A. (2007) ‘Updated world map of the Köppen-Geiger climate classification’, *Hydrology and Earth System Sciences Discussions*. European Geosciences Union, 4(2), pp. 439–473. Available at: https://hal.archives-ouvertes.fr/hal-00298818.

Peoples, M. B., Giller, K. E., Jensen, E. S. and Herridge, D. F. (2021) ‘Quantifying country-to-global scale nitrogen fixation for grain legumes: I. Reliance on nitrogen fixation of soybean, groundnut and pulses’, *Plant and Soil*, 469(1), pp. 1–14. doi: 10.1007/s11104-021-05167-6.

Peoples, M. B., Herridge, D. F. and Ladha, J. K. (1995) ‘Biological nitrogen-fixation - an efficient source of nitrogen for sustainable agricultural production’, *Plant and Soil*, 174(1–2), pp. 3–28. doi: 10.1007/BF00032239.

Pérez-Cuevas, A. J. (1994) *Atlas climático de la Comunidad Valenciana, 1961-1990*. Valencia: Generalitat Valenciana.

Peterson, B. W., Sharma, P. K., van der Mei, H. C. and Busscher, H. J. (2012) ‘Bacterial cell surface damage due to centrifugal compaction’, *Applied and Environmental Microbiology*, 78(1), pp. 120–125. doi: 10.1128/AEM.06780-11.

PGRO (2013) ‘PGRO Pulse Agronomy guide’. Processors and Growers Research Organisation. Available at: https://www.pgro.org/downloads/Pulse_Agronomy_Guide_2013.pdf.

PGRO (2017) ‘PGRO Pulse Agronomy guide’. Processors and Growers Research Organisation. Available at: https://www.pgro.org/downloads/PGRO-AGRONOMY-GUIDE-2017.pdf.

Potts, M. (1994) ‘Desiccation tolerance of prokaryotes.’, *Microbiological reviews*. American Society for Microbiology, 58(4), pp. 755–805. Available at: http://www.ncbi.nlm.nih.gov/pubmed/7854254 (Accessed: 5 June 2018).

Potts, M. (2001) ‘Desiccation tolerance: a simple process?’, *Trends in Microbiology*, 9(11), pp. 553–559. doi: https://doi.org/10.1016/S0966-842X(01)02231-4.

Preissel, S., Reckling, M., Schläfke, N. and Zander, P. (2015) ‘Magnitude and farm-economic value of grain legume pre-crop benefits in Europe: A review’, *Field Crops Research*, 175, pp. 64–79. doi: https://doi.org/10.1016/j.fcr.2015.01.012.

Price, M. N., Dehal, P. S. and Arkin, A. P. (2010) ‘FastTree 2 – Approximately maximum-likelihood trees for large alignments’, *PLOS ONE*. Public Library of Science, 5(3), pp. 1–10. doi: 10.1371/journal.pone.0009490.

Provorov, N. A. and Vorobyov, N. I. (2000) ‘Population genetics of rhizobia: Construction and analysis of an “infection and release” model’, *Journal of Theoretical Biology*, 205(1), pp. 105–119. doi: https://doi.org/10.1006/jtbi.2000.2051.

Qu, F., ElOmari, K., Wagner, A., De Simone, A. and Beis, K. (2019) ‘Desolvation of the substrate-binding protein *TauA* dictates ligand specificity for the alkanesulfonate ABC importer *TauABC*’, *Biochemical Journal*, 476(23), pp. 3649–3660. doi: 10.1042/BCJ20190779.

R Core Team (2019) ‘R: A language and environment for statistical computing’. Vienna, Austria. Available at: https://www.r-project.org/.

Raja, W., Rathaur, P., John, S. A. and Ramteke, P. W. (2012) ‘*Azolla-Anabaena* association and its significance in supportable agriculture ’, *Hacettepe Journal of Biology and Chemistry*. Ankara: Hacettepe University, pp. 1–6.

Ben Rebah, F. Ben, Tyagi, R. D. and Prévost, D. (2002) ‘Wastewater sludge as a substrate for growth and carrier for rhizobia: the effect of storage conditions on survival of *Sinorhizobium meliloti*’, *Bioresource Technology*, 83(2), pp. 145–151. doi: https://doi.org/10.1016/S0960-8524(01)00202-4.

Ben Rebah, F., Prevost, D., Yezza, A. and Tagi, R. (2007) ‘Agro-industrial waste materials and wastewater sludge for rhizobial inoculant production: A review’, *Bioresource Technology*, 98(18), pp. 3535–3546. doi: 10.1016/j.biortech.2006.11.066.

Rees, D. C. and Howard, J. B. (2000) ‘Nitrogenase: standing at the crossroads’, *Current Opinion in Chemical Biology*, 4(5), pp. 559–566. doi: https://doi.org/10.1016/S1367-5931(00)00132-0.

Rees, R. M. *et al.* (2013) ‘Nitrous oxide emissions from European agriculture - an analysis of variability and drivers of emissions from field experiments’, *Biogeosciences*, 10(4), pp. 2671–2682. doi: 10.5194/bg-10-2671-2013.

Reeve, W. *et al.* (2010) 'Complete genome sequence of *Rhizobium leguminosarum* bv. *trifolii* strain WSM1325, an effective microsymbiont of annual Mediterranean clovers', *Standards in genomic sciences*. Michigan State University, 2(3), pp. 347–356. doi: 10.4056/sigs.852027.

Reid, D. E., Ferguson, B. J., Hayashi, S., Lin, Y.-H. and Gresshoff, P. M. (2011) 'Molecular mechanisms controlling legume autoregulation of nodulation', *Annals of Botany*, 108(5), pp. 789–795. Available at: http://dx.doi.org/10.1093/aob/mcr205.

Reid, D., Liu, H., Kelly, S., Kawaharada, Y., Mun, T., Andersen, S. U., Desbrosses, G. and Stougaard, J. (2018) 'Dynamics of ethylene production in response to compatible Nod factor', *Plant Physiology*, 176(2), pp. 1764–1772. doi: 10.1104/pp.17.01371.

Reina-Bueno, M., Argandoña, M., Nieto, J. J., Hidalgo-García, A., Iglesias-Guerra, F., Delgado, M. J. and Vargas, C. (2012) 'Role of trehalose in heat and desiccation tolerance in the soil bacterium *Rhizobium etli*', *BMC Microbiology*, 12(1), p. 207. doi: 10.1186/1471-2180-12-207.

Requena, N., Jimenez, I., Toro, M. and Barea, J. M. (1997) 'Interactions between plant-growth-promoting rhizobacteria (PGPR), arbuscular mycorrhizal fungi and *Rhizobium* spp. in the rhizosphere of *Anthyllis cytisoides*, a model legume for revegetation in mediterranean semi-arid ecosystems', *New Phytologist*, 136(4), pp. 667–677. doi: 10.1046/j.1469-8137.1997.00786.x.

Rivera-Araya, J., Huynh, N. D., Kaszuba, M., Chávez, R., Schlömann, M. and Levicán, G. (2020) 'Mechanisms of NaCl-tolerance in acidophilic iron-oxidizing bacteria and archaea: Comparative genomic predictions and insights', *Hydrometallurgy*, 194, p. 105334. doi: https://doi.org/10.1016/j.hydromet.2020.105334.

Rockström, J. *et al.* (2009) 'Planetary boundaries: Exploring the safe operating space for humanity', *Ecology and Society*, 14(2).

Rockström, J., Williams, J., Daily, G., Noble, A., Matthews, N., Gordon, L., Wetterstrand, H., DeClerck, F., Shah, M., Steduto, P., de Fraiture, C., Hatibu, N., Unver, O., Bird, J., Sibanda, L. and Smith, J. (2017) 'Sustainable intensification of agriculture for human prosperity and global sustainability', *Ambio*, 46(1), pp. 4–17. doi: 10.1007/s13280-016-0793-6.

Rodríguez-Andrade, O., Corral Lugo, A., Morales-García, Y., Quintero-Hernandez, V., Rivera-Urbalejo, A., Molina-Romero, D., Contreras, R., Bernal, P. and Muñoz-Rojas, J. (2019) 'Identification of *Klebsiella variicola* T29A genes involved in tolerance to desiccation', *The Open Microbiology Journal*, 13, pp. 256–267. doi: 10.2174/1874285801913010256.

Rothe, M., Alpert, C., Loh, G. and Blaut, M. (2013) 'Novel insights into *E. coli*'s hexuronate metabolism: *KduI* facilitates the conversion of galacturonate and glucuronate under osmotic stress conditions', *PLOS ONE*. Public Library of Science, 8(2), pp. 1–14. doi: 10.1371/journal.pone.0056906.

Rousk, J., Bååth, E., Brookes, P. C., Lauber, C. L., Lozupone, C., Caporaso, J. G., Knight, R. and Fierer, N. (2010) 'Soil bacterial and fungal communities across a pH gradient in an arable soil', *The ISME Journal*, 4(10), pp. 1340–1351. doi: 10.1038/ismej.2010.58.

RStudio Team (2020) 'RStudio: Integrated development environment for R'. Boston, MA. Available at: http://www.rstudio.com/.

Rubiales, D., Rojas-Molina, M. M. and Sillero, J. C. (2016) 'Characterization of resistance mechanisms in faba bean (*Vicia faba*) against broomrape species (*Orobanche* and *Phelipanche* spp.)', *Frontiers in Plant Science*, 7, p. 1747. doi: 10.3389/fpls.2016.01747.

Rueden, C. T., Schindelin, J., Hiner, M. C., DeZonia, B. E., Walter, A. E., Arena, E. T. and Eliceiri, K. W. (2017) 'ImageJ2: ImageJ for the next generation of scientific image data', *BMC Bioinformatics*, 18(1), p. 529. doi: 10.1186/s12859-017-1934-z.

Ruhal, R., Kataria, R. and Choudhury, B. (2013) 'Trends in bacterial trehalose metabolism and significant nodes of metabolic pathway in the direction of trehalose accumulation', *Microbial Biotechnology*, 6(5), pp. 493–502. doi: https://doi.org/10.1111/1751-7915.12029.

Ruijter, J. M., Ramakers, C., Hoogaars, W. M. H., Karlen, Y., Bakker, O., van den Hoff, M. J. B. and Moorman, A. F. M. (2009) 'Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data', *Nucleic Acids Research*, 37(6), pp. e45–e45. doi: 10.1093/nar/gkp045.

Ryu, H., Cho, H., Choi, D. and Hwang, I. (2012) 'Plant hormonal regulation of nitrogen-fixing nodule organogenesis', *Molecules and Cells*, 34(2), pp. 117–126. doi: 10.1007/s10059-012-0131-1.

Saber, M. M. and Shapiro, B. J. (2020) 'Benchmarking bacterial genome-wide association study methods using simulated genomes and phenotypes', *Microbial genomics*. Microbiology Society, 6(3), p. e000337. doi: 10.1099/mgen.0.000337.

Saïdi, S., Ramírez-Bahena, M.-H., Santillana, N., Zúñiga, D., Álvarez-Martínez, E., Peix, A., Mhamdi, R. and Velázquez, E. (2014) '*Rhizobium laguerreae* sp. nov. nodulates *Vicia faba* on several continents', *International Journal of Systematic and Evolutionary Microbiology*. Microbiology Society, 64(Pt_1), pp. 242–247. doi: 10.1099/ijs.0.052191-0.

Saitoh, N. (1987) 'The neighbour-joining method: a new method for reconstructing phylogenetic trees', *Mol Biol Evol*, 10, pp. 471–483.

Salipante, S. J., Roach, D. J., Kitzman, J. O., Snyder, M. W., Stackhouse, B., Butler-Wu, S. M., Lee, C., Cookson, B. T. and Shendure, J. (2015) 'Large-scale genomic sequencing of extraintestinal pathogenic *Escherichia coli* strains', *Genome research*. Cold Spring Harbor Lab, 25(1), pp. 119–128.

Sánchez-Cañizares, C., Jorrín, B., Durán, D., Nadendla, S., Albareda, M., Rubio-Sanz, L., Lanza, M., González-Guerrero, M., Prieto, R. I., Brito, B., Giglio, M. G., Rey, L., Ruiz-Argüeso, T., Palacios, J. M. and Imperial, J. (2018) 'Genomic diversity in the endosymbiotic bacterium *Rhizobium leguminosarum*', *Genes*, 9(2). doi: 10.3390/genes9020060.

Sánchez-Cañizares, C. and Palacios, J. (2013) 'Construction of a marker system for the evaluation of competitiveness for legume nodulation in *Rhizobium* strains', *Journal of Microbiological Methods*, 92(3), pp. 246–249. doi: https://doi.org/10.1016/j.mimet.2012.12.022.

Saund, K. and Snitkin, E. S. (2020) ‘Hogwash: three methods for genome-wide association studies in bacteria’, *Microbial Genomics*. Microbiology Society, 6(11). doi: https://doi.org/10.1099/mgen.0.000469.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P. and Cardona, A. (2012) ‘Fiji: an open-source platform for biological-image analysis’, *Nature Methods*. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved., 9, p. 676. Available at: https://doi.org/10.1038/nmeth.2019.

Schneider, C. A., Rasband, W. S. and Eliceiri, K. W. (2012) ‘NIH Image to ImageJ: 25 years of image analysis’, *Nature Methods*. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved., 9, p. 671. Available at: https://doi.org/10.1038/nmeth.2089.

Schoebitz, M., López, M. D. and Roldán, A. (2013) ‘Bioencapsulation of microbial inoculants for better soil–plant fertilization. A review’, *Agronomy for Sustainable Development*, 33(4), pp. 751–765. doi: 10.1007/s13593-013-0142-0.

Schwibbert, K., Marin-Sanguino, A., Bagyan, I., Heidrich, G., Lentzen, G., Seitz, H., Rampp, M., Schuster, S. C., Klenk, H.-P., Pfeiffer, F., Oesterhelt, D. and Kunte, H. J. (2011) ‘A blueprint of ectoine metabolism from the genome of the industrial producer *Halomonas elongata* DSM 2581T’, *Environmental Microbiology*, 13(8), pp. 1973–1994. doi: https://doi.org/10.1111/j.1462-2920.2010.02336.x.

Scoffone, V., Dondi, D., Biino, G., Borghese, G., Pasini, D., Galizzi, A. and Calvio, C. (2013) ‘Knockout of *pgdS* and *ggt* genes improves γ-PGA yield in *B. subtilis*’, *Biotechnology and Bioengineering*, 110(7), pp. 2006–2012. doi: https://doi.org/10.1002/bit.24846.

Seemann, T. (2014) ‘Prokka: rapid prokaryotic genome annotation’, *Bioinformatics*, 14(30), pp. 2068–9. doi: 10.1093/bioinformatics/btu153.

Segovia, L., Piñero, D., Palacios, R. and Mart\'\inez-Romero, E. (1991) ‘Genetic structure of a soil population of nonsymbiotic *Rhizobium leguminosarum*.’, *Applied and Environmental Microbiology*. American Society for Microbiology Journals, 57(2), pp. 426–433. Available at: https://aem.asm.org/content/57/2/426.

Senberga, A., Dubova, L. and Alsina, I. (2018) ‘Germination and growth of primary roots of inoculated bean (*Vicia faba*) seeds under different temperatures’, *Agronomy Research*, 16(1), pp. 243–253. doi: https://doi.org/10.15159/AR.18.024.

Sessitsch, A., Howieson, J. G., Perret, X., Antoun, H. and Martinez-Romero, E. (2002) ‘Advances in rhizobium research’, *Critical Reviews in Plant Sciences*, 21(4), pp. 323–378. doi: 10.1080/0735-260291044278.

Shi, G. H., Dong, X., Lytle, M., Kemp, C. A. J., Behme, R. J., Hinds, J. and Xiao, Z. (2018) ‘Two contrasting failure modes of enteric coated beads’, *AAPS Pharmscitech*, 19(4), pp. 1827–1836. doi: 10.1208/s12249-018-1000-9.

Sievers, F. and Higgins, D. G. (2018) ‘Clustal Omega for making accurate alignments of many protein sequences’, *Protein Science*, 27(1), pp. 135–145. doi: https://doi.org/10.1002/pro.3290.

Sillero, J. C., Villegas-Fernández, A. M., Thomas, J., Rojas-Molina, M. M., Emeran, A. A., Fernández-Aparicio, M. and Rubiales, D. (2010) ‘Faba bean breeding for disease resistance’, *Field Crops Research*, 115(3), pp. 297–307. doi: https://doi.org/10.1016/j.fcr.2009.09.012.

Sinclair, T. R., Soffes, A. R., Hinson, K., Albrecht, S. L. and Pfahler, P. L. (1991) ‘Genotypic variation in soybean nodule number and weight’, *Crop Science*, 31(2), p. cropsci1991.0011183X003100020014x. doi: https://doi.org/10.2135/cropsci1991.0011183X003100020014x.

Slininger, P. J. and Schisler, D. A. (2013) ‘High-throughput assay for optimising microbial biological control agent production and delivery’, *Biocontrol Science and Technology*, 23(8), pp. 920–943. doi: 10.1080/09583157.2013.808739.

Smil, V. (2000) *Enriching the Earth*. Cambridge, Massachusetts: MIT Press.

Smit, P. and Bisseling, T. (2008) ‘Genetics: A way to unravel molecular mechanisms controlling the rhizobial-legume symbiosis’, in Dilworth, M. J., James, E. K., Sprent, J. I., and Newton, W. E. (eds) *Nitrogen-fixing Leguminous Symbioses*. Dordrecht: Springer Netherlands, pp. 191–210. doi: 10.1007/978-1-4020-3548-7_7.

Smith, P. *et al.* (2008) ‘Greenhouse gas mitigation in agriculture’, *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1492), pp. 789–813. doi: 10.1098/rstb.2007.2184.

Smith, R. S. (1992) ‘Legume inoculant formulation and application’, *Canadian Journal of Microbiology*, 38(6), pp. 485–492. doi: doi.org/10.1139/m92-080.

Sokal, R. R. and Michener, C. D. (1958) ‘A statistical method for evaluating systematic relationships’, *The University of Kansas Science Bulletin*, 38(22), pp. 1409–1438.

Solt, F. and Hu, Y. (2018) ‘dotwhisker: Dot-and-Whisker plots of regression results’. Available at: https://cran.r-project.org/package=dotwhisker.

Soltis, D. E., Soltis, P. S., Chase, M. W., Mort, M. E., Albach, D. C., Zanis, M., Savolainen, V., Hahn, W. H., Hoot, S. B., Fay, M. F., Axtell, M., Swensen, S. M., Prince, L. M., Kress, W. J., Nixon, K. C. and Farris, J. S. (2000) ‘Angiosperm phylogeny inferred from 18S rDNA, rbcL, and atpB sequences’, *Botanical Journal of the Linnean Society*, 133(4), pp. 381–461. doi: 10.1111/j.1095-8339.2000.tb01588.x.

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.’, *Proceedings of the National Academy of Sciences*. National Academy of Sciences, 92(7), pp. 2647–2651. doi: 10.1073/pnas.92.7.2647.

Sprent, J. I. (2009) ‘Bateria nodulating legumes’, in *Legume nodulation: A global perspective*. Wiley-Blackwell.

Sprent, J. I., Ardley, J. and James, E. K. (2017) ‘Biogeography of nodulated legumes and their nitrogen-fixing symbionts’, *New Phytologist*, 215(1), pp. 40–56. doi: 10.1111/nph.14474.

Sprent, J. I., Ardley, J. K. and James, E. K. (2013) ‘From North to South: A latitudinal look at legume nodulation processes’, *South African Journal of Botany*. Elsevier, 89, pp. 31–41. doi: 10.1016/j.sajb.2013.06.011.

Stapelfeldt, H., Nielsen, B. R. and Skibsted, L. H. (1997) ‘Effect of heat treatment, water activity and storage temperature on the oxidative stability of whole milk powder’, *International Dairy Journal*, 7(5), pp. 331–339. doi: https://doi.org/10.1016/S0958-6946(97)00016-2.

Statham, P. J. (2012) ‘Nutrients in estuaries — An overview and the potential impacts of climate change’, *Science of The Total Environment*, 434, pp. 213–227. doi: https://doi.org/10.1016/j.scitotenv.2011.09.088.

Steffen, W., Richardson, K., Rockstrom, J., Cornell, S. E., Fetzer, I., Bennett, E. M., Biggs, R., Carpenter, S. R., de Vries, W., de Wit, C. A., Folke, C., Gerten, D., Heinke, J., Mace, G. M., Persson, L. M., Ramanathan, V., Reyers, B. and Sorlin, S. (2015) ‘Planetary boundaries: Guiding human development on a changing planet’, *Science*, 347(6223). doi: 10.1126/science.1259855.

Stephens, J. and Rask, H. (2000) ‘Inoculant production and formulation’, *Field Crops Research*. Elsevier, 65(2–3), pp. 249–258. doi: 10.1016/S0378-4290(99)00090-8.

Streeter, J. G. (2003) ‘Effect of trehalose on survival of *Bradyrhizobium japonicum* during desiccation’, *Journal of Applied Microbiology*, 95(3), pp. 484–491. doi: 10.1046/j.1365-2672.2003.02017.x.

Streit, W. R. and Entcheva, P. (2003) ‘Biotin in microbes, the genes involved in its biosynthesis, its biochemical role and perspectives for biotechnological production’, *Applied Microbiology and Biotechnology*, 61(1), pp. 21–31. doi: 10.1007/s00253-002-1186-2.

Sugawara, M., Cytryn, E. J. and Sadowsky, M. J. (2010) ‘Functional role of *Bradyrhizobium japonicum* trehalose biosynthesis and metabolism genes during physiological stress and nodulation’, *Applied and Environmental Microbiology*, 76(4), pp. 1071–1081. doi: 10.1128/AEM.02483-09.

Tadapaneni, R. K., Xu, J., Yang, R. and Tang, J. (2018) ‘Improving design of thermal water activity cell to study thermal resistance of *Salmonella* in low-moisture foods’, *LWT*, 92, pp. 371–379. doi: https://doi.org/10.1016/j.lwt.2018.02.046.

Tamura, K., Nei, M. and Kumar, S. (2004) ‘Prospects for inferring very large phylogenies by using the neighbor-joining method’, *Proceedings of the National Academy of Sciences*. National Acad Sciences, 101(30), pp. 11030–11035.

Tauch, A., Götker, S., Pühler, A., Kalinowski, J. and Thierbach, G. (2002) ‘The alanine racemase gene alr is an alternative to antibiotic resistance genes in cloning systems for industrial *Corynebacterium glutamicum* strains’, *Journal of Biotechnology*, 99(1), pp. 79–91. doi: https://doi.org/10.1016/S0168-1656(02)00159-1.

Taurian, T., Anzuay, M. S., Angelini, J. G., Tonelli, M. L., Ludueña, L., Pena, D., Ibáñez, F. and Fabra, A. (2010) ‘Phosphate-solubilizing peanut associated bacteria: screening for plant growth-promoting activities’, *Plant and Soil*, 329(1), pp. 421–431. doi: 10.1007/s11104-009-0168-x.

Thies, J. E., Woomer, P. L. and Singleton, P. W. (1995) ‘Enrichment of *Bradyrhizobium* spp populations in soil due to cropping of the homologous host legume’, *Soil Biology and Biochemistry*, 27(4), pp. 633–636. doi: https://doi.org/10.1016/0038-0717(95)98643-3.

Todd, J. D., Wexler, M., Sawers, G., Yeoman, K. H., Poole, P. S. and Johnston, A. W. B. (2002) ‘RirA, an iron-responsive regulator in the symbiotic bacterium *Rhizobium leguminosarum* The GenBank accession number for the RirA sequence is CAC35510.’, *Microbiology*. Microbiology Society, 148(12), pp. 4059–4071. doi: https://doi.org/10.1099/00221287-148-12-4059.

Tong, W., Li, X., Huo, Y., Zhang, L., Cao, Y., Wang, E., Chen, W., Tao, S. and Wei, G. (2018) ‘Genomic insight into the taxonomy of *Rhizobium* genospecies that nodulate *Phaseolus vulgaris*’, *Systematic and Applied Microbiology*, 41(4), pp. 300–310. doi: https://doi.org/10.1016/j.syapm.2018.03.001.

Udvardi, M., Below, F. E., Castellano, M. J., Eagle, A. J., Giller, K. E., Ladha, J. K., Liu, X., Maaz, T. M., Nova-Franco, B., Raghuram, N., Robertson, G. P., Roy, S., Saha, M., Schmidt, S., Tegeder, M., York, L. M. and Peters, J. W. (2021) ‘A research road map for responsible use of agricultural nitrogen’, *Frontiers in Sustainable Food Systems*, 5, p. 165. doi: 10.3389/fsufs.2021.660155.

Udvardi, M. and Poole, P. S. (2013) ‘Transport and metabolism in legume-rhizobia symbioses’, *Annual Review of Plant Biology*, 64(1), pp. 781–805. doi: 10.1146/annurev-arplant-050312-120235.

United Nations (2019) *Probabilistic population projections Rev. 1 based on the World population prospects 2019*, *Department of Economic and Social Affairs, Population Division*. Available at: http://population.un.org/wpp/ (Accessed: 5 January 2021).

VandenBosch, K. A. and Stacey, G. (2003) ‘Summaries of legume genomics projects from around the globe. Community resources for crops and models’, *Plant Physiology*, 131(3), pp. 840–865. doi: 10.1104/pp.103.020388.

Vanderlinde, E. M., Harrison, J. J., Muszyński, A., Carlson, R. W., Turner, R. J. and Yost, C. K. (2010) ‘Identification of a novel ABC transporter required for desiccation tolerance, and biofilm formation in *Rhizobium leguminosarum* bv. *viciae* 3841’, *FEMS Microbiology Ecology*, 71(3), pp. 327–340. doi: 10.1111/j.1574-6941.2009.00824.x.

Vanderlinde, E. M., Muszyński, A., Harrison, J. J., Koval, S. F., Foreman, D. L., Ceri, H., Kannenberg, E. L., Carlson, R. W. and Yost, C. K. (2009) ‘*Rhizobium leguminosarum* biovar *viciae* 3841, deficient in 27-hydroxyoctacosanoate-modified lipopolysaccharide, is impaired in desiccation tolerance, biofilm formation and motility’, *Microbiology (Reading, England)*. 2009/05/21. Microbiology Society, 155(Pt 9), pp. 3055–3069. doi: 10.1099/mic.0.025031-0.

van Velzen, R., Doyle, J. J. and Geurts, R. (2019) ‘A Resurrected scenario: Single gain and massive loss of Nitrogen-fixing nodulation’, *Trends in Plant Science*, 24(1), pp. 49–57. doi: https://doi.org/10.1016/j.tplants.2018.10.005.

Versalovic, J., Schneider, M., De Bruijn, F. and Lupski, J. R. (1994) ‘Genomic fingerprinting of bacteria using repetitive sequence-based polymerase chain reaction’, *Methods in molecular and cellular biology*, 5(1), pp. 25–40.

Vikman, P. and Vessey, J. K. (1993) ‘Ontogenetic changes in root nodule subpopulations of common bean (*Phaseolus vulgaris* L.): III. Nodule formation, growth and degradation’, *Journal of Experimental Botany*, 44(3), pp. 579–586. doi: 10.1093/jxb/44.3.579.

Vincent, J. M. (1962) ‘Influence of calcium and magnesium on the growth of *Rhizobium*’, *Microbiology*. Microbiology Society, 28(4), pp. 653–663. doi: https://doi.org/10.1099/00221287-28-4-653.

Vincent, J. M. (1970) *A manual for the practical study of the root-nodule bacteria.* IBP Handbk 15 Oxford and Edinburgh: Blackwell Scientific Publications.

Vitousek, P. M., Cassman, K., Cleveland, C., Crews, T., Field, C. B., Grimm, N. B., Howarth, R. W., Marino, R., Martinelli, L., Rastetter, E. B. and Sprent, J. I. (2002) ‘Towards an ecological understanding of biological nitrogen fixation’, in Boyer, E. W. and Howarth, R. W. (eds) *The Nitrogen Cycle at Regional to Global Scales*. Dordrecht: Springer Netherlands, pp. 1–45. doi: 10.1007/978-94-017-3405-9_1.

Vorobjeva, N. N., Kurilova, S. A., Petukhova, A. F., Nazarova, T. I., Kolomijtseva, G. Y., Baykov, A. A. and Rodina, E. V (2020) ‘A novel, cupin-type phosphoglucose isomerase in *Escherichia coli*’, *Biochimica et Biophysica Acta (BBA) - General Subjects*, 1864(7), p. 129601. doi: https://doi.org/10.1016/j.bbagen.2020.129601.

Vriezen, J. A. C., de Bruijn, F. J. and Nusslein, K. (2007) ‘Responses of rhizobia to desiccation in relation to osmotic stress, oxygen, and temperature’, *Applied and Environmental Microbiology*, 73(11), pp. 3451–3459. doi: 10.1128/AEM.02991-06.

Vriezen, J. A. C., de Bruijn, F. J. and Nüsslein, K. R. (2012) ‘Desiccation induces viable but non-culturable cells in *Sinorhizobium meliloti* 1021’, *AMB Express*, 2(1), p. 6. doi: 10.1186/2191-0855-2-6.

Wang, H., Moore, M. J., Soltis, P. S., Bell, C. D., Brockington, S. F., Alexandre, R., Davis, C. C., Latvis, M., Manchester, S. R. and Soltis, D. E. (2009) ‘Rosid radiation and the rapid rise of angiosperm-dominated forests’, *Proceedings* *of the National Academy of Sciences*. National Academy of Sciences, 106(10), pp. 3853–3858. doi: 10.1073/pnas.0813376106.

Wardlaw, J., Davis, C., Monie, C. and Reay, G. (2019) *Pesticide usage in Scotland, outdoor vegetable crops 2019*. Edinburgh. Available at: www.sasa.gov.uk/pesticides.

Watson, R. J., Heys, R., Martin, T. and Savard, M. (2001) '*Sinorhizobium meliloti* cells require biotin and either cobalt or methionine for growth', *Applied and Environmental Microbiology*, 67(8), pp. 3767–3770. doi: 10.1128/AEM.67.8.3767-3770.2001.

Weisburg, W. G., Barns, S. M., Pelletier, D. A. and Lane, D. J. (1991) '16S ribosomal DNA amplification for phylogenetic study.', *Journal of Bacteriology*, 173(2), pp. 697–703. doi: 10.1128/JB.173.2.697-703.1991.

Werner, G. D. A., Cornwell, W. K., Sprent, J. I., Kattge, J. and Kiers, E. T. (2014) 'A single evolutionary innovation drives the deep evolution of symbiotic N2-fixation in angiosperms', *Nature Communications*, 5(1), p. 4087. doi: 10.1038/ncomms5087.

Westhoek, A., Clark, L. J., Culbert, M., Dalchau, N., Griffiths, M., Jorrin, B., Karunakaran, R., Ledermann, R., Tkacz, A., Webb, I., James, E. K., Poole, P. S. and Turnbull, L. A. (2021) 'Conditional sanctioning in a legume-*Rhizobium* mutualism', *Proceedings of the National Academy of Sciences*. National Academy of Sciences, 118(19). doi: 10.1073/pnas.2025760118.

Wetzel, M. E., Olsen, G. J., Chakravartty, V. and Farrand, S. K. (2015) 'The *repABC* plasmids with quorum-regulated transfer systems in members of the Rhizobiales divide into two structurally and separately evolving groups', *Genome Biology and Evolution*, 7(12), pp. 3337–3357. doi: 10.1093/gbe/evv227.

White, P. J. and Brown, P. H. (2010) 'Plant nutrition for sustainable development and global health', *Annals of Botany*, 105(7), pp. 1073–1080. doi: 10.1093/aob/mcq085.

Wickham, H. (2016) *ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York. Available at: https://ggplot2.tidyverse.org.

Wickham, H., François, R., Henry, L. and Müller, K. (2019) 'dplyr: A grammar of data manipulation'. Available at: https://cran.r-project.org/package=dplyr.

Wiedenbeck, J. and Cohan, F. M. (2011) 'Origins of bacterial diversity through horizontal genetic transfer and adaptation to new ecological niches', *FEMS Microbiology Reviews*, 35(5), pp. 957–976. doi: 10.1111/j.1574-6976.2011.00292.x.

Wood, J. M. (1999) 'Osmosensing by bacteria: Signals and membrane-based sensors', *Microbiology and Molecular Biology Reviews*, 63(1), pp. 230–262. doi: 10.1128/MMBR.63.1.230-262.1999.

Yamada, A., Inoue, T., Wiwatwitaya, D., Ohkuma, M., Kudo, T. and Sugimoto, A. (2006) 'Nitrogen fixation by termites in tropical forests, Thailand', *Ecosystems*, 9(1), pp. 75–83. doi: 10.1007/s10021-005-0024-7.

Yates, R. J., Howieson, J. G., Hungria, M., Bala, A., O'Hara, G. W. and Terpolilli, J. (2016) 'Authentication of rhizobia and assessment of the legume symbiosis in controlled plant growth systems', in Howieson, J. and Dilworth, M. (eds) *Working with rhizobia*. Australian Centre for International Agricultural Research (ACIAR).

Yoshida, K., Kim, W.-S., Kinehara, M., Mukai, R., Ashida, H., Ikeda, H., Fujita, Y. and Krishnan, H. B. (2006) 'Identification of a functional 2-keto-myo-inositol dehydratase gene of *Sinorhizobium fredii* USDA191 required for myo-inositol utilization', *Bioscience, Biotechnology, and Biochemistry*, 70(12), pp. 2957–2964. doi: 10.1271/bbb.60362.

Young, J. P. W. (2000) 'Molecular evolution in diazotrophs: Do the genes agree?', in Pedrosa, F. O., Hungria, M., Yates, G., and Newton, W. E. (eds) *Nitrogen fixation: From molecules to crop productivity*. Dordrecht: Springer Netherlands, pp. 161–164. doi: 10.1007/0-306-47615-0_82.

Young, J. P. W. *et al.* (2006) 'The genome of *Rhizobium leguminosarum* has recognizable core and accessory components', *Genome Biology*, 7(4), p. R34. doi: 10.1186/gb-2006-7-4-r34.

Young, J. P. W. *et al.* (2021) ‘Defining the *Rhizobium leguminosarum* species complex’, *Genes 2021, Vol. 12, Page 111*. Multidisciplinary Digital Publishing Institute, 12(1), p. 111. doi: 10.3390/GENES12010111.

Youseif, S. H., Abd El-Megeed, F. H., Mohamed, A. H., Ageez, A., Veliz, E. and Martínez-Romero, E. (2021) ‘Diverse *Rhizobium* strains isolated from root nodules of *Trifolium alexandrinum* in Egypt and symbiovars’, *Systematic and Applied Microbiology*, 44(1), p. 126156. doi: https://doi.org/10.1016/j.syapm.2020.126156.

Zhang, Q. and Yan, T. (2012) ‘Correlation of intracellular trehalose concentration with desiccation resistance of soil *Escherichia coli* populations’, *Applied and Environmental Microbiology*, 78(20), pp. 7407–7413. doi: 10.1128/AEM.01904-12.

Zhu, J., Jiang, X., Guan, D., Kang, Y., Li, L., Cao, F., Zhao, B., Ma, M., Zhao, J. and Li, J. (2021) ‘Effects of rehydration on physiological and transcriptional responses of a water-stressed rhizobium’, *Journal of Microbiology*. doi: 10.1007/s12275-022-1325-7.

Zohary, D., Hopf, M. and Weiss, E. (2012) ‘Pulses’, in *Domestication of plants in the Old World: The origin and spread of domesticated plants in Southwest Asia, Europe, and the Mediterranean Basin*. Oxford university press, pp. 75–99. doi: 10.1093/acprof:osobl/9780199549061.001.0001.

Appendices

Appendix 2.1 | Screening and selection of rhizobia

From the 82 strains of Rhizobium spp. isolated from the soil, an initial broad screening was carried out with over half of these strains, which were selected based upon their field of origin and BOX PCR band pattern (Figure 2.3). The aim was to reduce the number of test strains to six based on their symbiotic interaction with the trapping host plant.

The screening experimental setup was carried out following the same method described in Section 2.2.4. A total of six replica experiments were set up where pure cultures of the test strains were inoculated on pea cv. Corus seedlings at sowing. These experiments were run with four positive controls (PC) (Table S2.1) and two negative controls (NC) (uninoculated, NC1, and an autoclaved grown PC culture, NC2) to allow comparison between the different experiments. Each treatment was replicated three times.

After set up the pots were transferred to a glasshouse and checked every day until seedlings started to emerge at which point they were thinned to one seedling per pot. Plants were checked every other day when the maximum and minimum temperatures were recorded (Figure S2.1) and pots watered when necessary with N-free solution (Burchill *et al.*, 2014).

Plants were harvested when 60 % of them reached growth stage 203 (first flower open) (Knott, 1987), usually between 7-8 weeks after sowing. The harvest was carried out following the same procedure as described in Section 2.2.4 but nodules were counted manually. The total biomass dry weight was standardised against the mean biomass dry weight of the PC treatment with least variance (rcr1045) by dividing each replicate biomass dry weight of each treatment by the PC mean biomass dry weight.

In total, 43 representative Rhizobium strains from all fields and all clades of the BOX PCR dendrogram were screened (Figure 2.5). All plants in all experiments produced nodules, including the uninoculated negative controls. However, the treatment had an effect in all experiments and several strains produced significantly more total dry biomass than negative controls (Figure S2.2).

Of the positive control treatments, commercial strain rcr1045 was the treatment with the least variance of biomass dry weight between experiments (total biomass mean = 1.375 g, $\sigma^2 = 0.105$) and thus it was used for standardising all treatment replicates in each individual experiment (Figure S2.3). Fourteen of the test strains showed an average standardised biomass higher than plants inoculated with strain rcr1045. From these, the four strains with the highest mean standardised biomass and two that had a similar performance to Rlv3841 were selected for further screening.

This initial set of screening experiments were affected by widespread contamination, which meant that both negative control treatments produced nodules at harvest. However, the NC plants were generally smaller than PC and almost 70 % of the isolates produced bigger plants than NC treatments. The source of this contamination is unknown, although it was later discovered that some precipitation had leaked into the glasshouse and this is the probable cause of this contamination. Nevertheless, quality control measures and a controlled growth environment were implemented for further screening to avoid leaks or any other contamination that may compromise the sterility of the experiments

Appendix 2.2 | Optimisation of a macro for automated counting and measuring of nodules using digital images

In the pea and faba bean growth room experiments, nodules were harvested, arranged on a flat surface and a picture was captured for image analysis. The images were analysed on ImageJ using a macro developed along this project.

Images were captured during two different setups; in the first a white background and two tabletop lamps on each side were used, and in the second a light blue background and a ring flashlight attached to the lens were used (Figure S2.4). Images were taken as described in Section 2.2.4.

For each set of photographs, a macro was implemented built on the "Batch Measure" script (https://imagej.nih.gov/ij/macros/) with modifications for the segmentation and measurement of the nodules in these experimental conditions. Images were analysed on FIJI ImageJ v1.52n (Schindelin *et al.*, 2012; Schneider *et al.*, 2012; Rueden *et al.*, 2017).

The macro nodule counts were validated against manual nodule counts using Spearman's correlation. Nodule length was estimated as the highest value between the height and the width of the bounding rectangle. Circularity was calculated using formula (Eq. S2.1), which is a measure of object roundness, where the value varies from 0 to 1, with 1 being a perfect circle.

$$Circularity = \frac{4 \times \pi \times area}{Perimeter^2} \qquad \text{(Eq. S2.1)}$$

Solidity is a measure of how smooth an object is. It varies from 0 to 1 where 0 indicates a very irregular shape, and it was calculated using formula (Eq. S2.2).

$$Solidity = \frac{Area}{Convex\ area} \qquad \text{(Eq. S2.2)}$$

where convex area is the area of the minimum polygon that encloses the nodule shape. Shape parameters are represented in Figure S2.6.

The first macro (Macro I) was designed to separate the nodules from a white background and was optimised for each type of nodule. All three segmentation scripts perform the same image process with an optimised set of thresholds for the different nodule types. On the first step, the original image is split into the three colour channels, red (R), blue (B) and green (G), and generates an image resulting from the arithmetic calculation of these three channels called “First selection”; ((R-G)+(R-B)) + 3x((G-B)+(R-B)). Following this, a threshold optimised for each type of nodule is applied to “First selection” which transforms the image into binary (only white and black pixels) where the nodules (black) are now separated from the background (white). Afterwards, all holes are filled, and the image is cleaned of noise (small black pixel particles in between nodules) by eroding three times followed by three dilations. The cleaned “First selection” is then added to the original image to crop the nodules out of the original picture.

This new image with the cropped nodules is then split into its HSB channels (Hue, Saturation and Brightness) for a finer segmentation of the nodules. Each resulting image has an optimised threshold for the type of nodule applied and then the resulting binary images have the holes filled followed by two erosions and two dilations for reducing the background noise. Then a calculation with the three images is run, (H AND S) AND B, to produce a binary image called “Second selection”. Then “Second selection” is added to the original image in order to crop the nodules out of the original image.

Finally, the macro proceeds to analyse all particles in the binary “Second selection” and creates a CSV file with all the results.

Macro I generates four main outputs: a .CSV file with the results of the image analysis, an RGB of the cropped nodules, a binary image of the cropped nodules (“Second selection”), and a mask image that highlights and identifies each analysed particle which permits the identification of which measurement in the results file corresponds to each nodule on the image.

The second macro (Macro II) was designed and optimised for analysing nodule images taken with blue background and illuminated with the Macro Ring Flash. This macro has only one image processing step where it splits the image into its colour channels (R, G and B), then applies a “Huang” threshold (Huang and Wang, 1995) to the blue channel and converts the resulting image into binary followed by two erosion and two dilation steps. Macro II then generates the same four outputs as Macro I; .CSV file with the results of the image analysis, cropped nodules in RGB and binary, and a mask for nodule identification.

The first batch of images (pea cvs Corus Setup 1 and Kareni Setup 1 and faba bean cv. Fuego) were taken with white background and lateral flashlights and were analysed with Macro I. The second batch of images (cvs Corus setup 2 and Kareni setup 2) were taken with blue background and a ring flashlight situated directly on top of the object. The quality of image segmentation of the macros was validated against manual nodule counts by Spearman's correlation. Both macros showed a high correlation with manual counts (Macro I: rho = 0.984, $P < 0.001$; Macro II: rho = 0.988, $P < 0.001$).

The linear correlation is similar to that achieved by previous studies which have used image analysis for counting nodules (Vikman and Vessey, 1993; Lira and Smith, 2000; Barbedo, 2012).

Further, as previously suggested by Barbedo (2012) , attention must be paid to the illumination of the subject before taking the images, otherwise nodule segmentation is complicated as shown with the complexity of macro I where multiple algorithms had to be implemented before nodules were segmented. On the contrary, when the lighting was homogeneous, the separation of nodules from the background was simpler. Moreover, Barbedo (2012) also suggests using a lighter background for the images and that when the image is taken in RGB format, the blue channel gives the best contrast. In agreement with this, we found that a light blue background and blue channel performed the best. Therefore, using a light blue background and uniform illumination simplified in great measure the algorithm used for nodule segmentation, thus reducing the computation time which accelerates the process of image analysis.

Appendix 5.1 | On the qPCR efficiency and melting curve assessment

The average spike recovery was 45.84% indicating that during the DNA isolation procedure over 50% of the copies in soil were lost which was 3.36-fold higher than that reported previously (Daniell *et al.*, 2012). The improvement in the recovery of DNA during isolation may be due to the different extraction methods used. The efficiency measured for each individual reaction was, respectively, for the spike, 16S rRNA and *nodD*: 1.68, 1.71 and 1.76 on average. Simultaneously, the standard regression line efficiency for each of the target DNA fragments was 1.92 on average for all three targets.

Moreover, the melting curve analysis run at the end of each PCR revealed that some of the DNA samples showed one or more melting temperatures around the expected melting temperature indicated by the melting curve of the standard (Figure S5.1). Whilst 16S rRNA had a prominent peak at a similar temperature where the standard peak is situated, the *nodD* melting curves show more variability. The forward primers designed by Macdonald et al (2011) are degenerate, which means that one of the primers in each set (the forward primer in both cases, see Table 5.2) has a bi-variant position which will join to two different alleles of the gene, thus altering slightly the melting temperature of the resulting amplicon. Notwithstanding this, the degenerate forward primers might not fully explain the variability observed in *nodD* products. Additionally, Macdonald et al (2011) reported that the 16S rRNA primers might amplify other *Rhizobium* species which is in accordance with current understanding that this traditionally-used gene for prokaryote identification is better suited for generic rather than species-specific identification, and a multi-locus or whole genome approach is currently used for species identification (Kumar *et al.*, 2015; Tong *et al.*, 2018; Cavassim *et al.*, 2020; Young *et al.*, 2021). Thus, a BLASTn search in the NCBI database for both 16S rRNA primers was performed, and this showed that all sequences that had a match for both primers (Table S5.1) belonged to the genus *Rhizobium* with *R. leguminosarum* the most common species (Figure S5.2 A). Similarly, the same analysis carried with both *nodD* primers only produced *Rhizobium* species with, again, *R. leguminosarum* the most frequent (Figure S5.2 C).

An alignment of all the database 16S rRNA and *nodD* amplicon sequences revealed that whilst the 16S rRNA amplicon sequences (285 bp) showed a 96.1% of nucleotide conservation, the *nodD* amplicon (357 bp) only showed 82.6% of conservation thus explaining the higher variability in melting temperatures for *nodD*. Furthermore, both searches found sequences that only showed a match for one of the primers (Figure S5.2 B and D) and despite this not having an effect on the quantification or melting curve steps of the qPCR (the dye used, SybrGreen, only binds to dsDNA) it might affect the reaction efficiency, essentially acting as inhibitors, which in part explains the slightly lower efficiencies observed in individual samples when compared with the standard regression curve efficiency.

In conclusion, despite some minor inhibition that may have been caused by DNA which only had an annealing locus for one of the primers, the results obtained in the melting curve analyses are very likely explained by the results obtained from the NCBI database search.

Appendix 5.2 | Correlation analysis of Rleg and Rlv concentration with soil properties and chemical analyses

Each year, the soil of the CSC farm complex was sampled in March following the procedure described in chapter 6. A subsample of the soil sampled from each permanent GPS location was sent for chemical analysis either to Yara (Lincolnshire, UK) or analysed at the James Hutton Institute (Dundee, UK) depending on the analysis and its availability each year (Table S5.1). The results of these soil properties and chemical analyses for each of the GPS locations processed during the measurement of Rleg and Rlv concentration was extracted from the CSC database, and their correlation was assessed with Rleg and Rlv concentrations. The correlation analysis was carried by the Kendal's correlation analysis using the R base package v 3.6.1 (R Core Team, 2019).

The correlation analysis between both soil properties and chemical analyses and rhizobia concentrations showed significant correlations between 15 of these soil analyses and Rleg, Rlv and the Rlv Rleg^{-1} ratio, yet in most of the cases the correlations were only found for one of the management types or when looking at the overall concentration for the field (Table S5.4). The strongest positive correlation was found with the concentrations of Rlv and Rleg, and the concentration of Molybdenum (Mo) in the integrated field halves, which had a Kendall's tau of 0.282 and 0.277 respectively. The strongest negative correlation was found when contrasting both rhizobia concentrations with the Iron (Fe) concentration which had a Kendall's tau of around -0.28 in both cases.

The Rlv Rleg^{-1} ratio was negatively correlated with the concentrations of soil nitrate (NO_3) in the conventional field halves and Magnesium (Mg) and Phosphorus (P) in the integrated field halves. The Calcium (Ca) concentration showed a consistent negative correlation with the Rlv Rleg^{-1} ratio in both the conventional and integrated field halves and when the field concentrations were considered as a whole. The organic matter content (OM) and the soil pH had a positive correlation with the Rlv Rleg^{-1} ratio in the integrated field halves.

The concentration of Rleg was the least influenced by any of the analysed soil parameter concentrations, it only being significantly correlated with Ca, Mo, Fe, and percent N and Carbon (C).

The soil data were only available for a limited number of years across the six years of the rotation, in some cases, only being available for one of the years (Table S5.1). Notwithstanding this, significant correlations between these analyses and the rhizobia concentrations were found for several of the analysed parameters. Elements such as Fe, Mg, Mo and S are essential for the nitrogen fixation process as they participate in the binding of dinitrogen, the electron chain, and are a fundamental part of the nitrogenase enzyme (Rees and Howard, 2000; Brear *et al.*, 2013). The concentration of these elements has shown the highest correlations with Rleg and Rlv concentrations. Furthermore, Fe also plays an important regulatory function in the expression of genes involved in the growth and nitrogen fixation of rhizobia (Todd *et al.*, 2002; Johnston *et al.*, 2007; Brear *et al.*, 2013). It is suggested therefore that the high concentrations found in the soils at the CSC might be inhibiting rhizobial growth and ultimately the concentration of these bacteria in the soil. Other negative correlations found were with Ca which is a limiting nutrient for rhizobia growth (Vincent, 1962), and ammonia and nitrate which interact in the autoregulation of the nodulation process in the legume root (Reid *et al.*, 2011). This contrasts with positive correlations found for calcium carbonate and total N percentage.

Supplementary figures

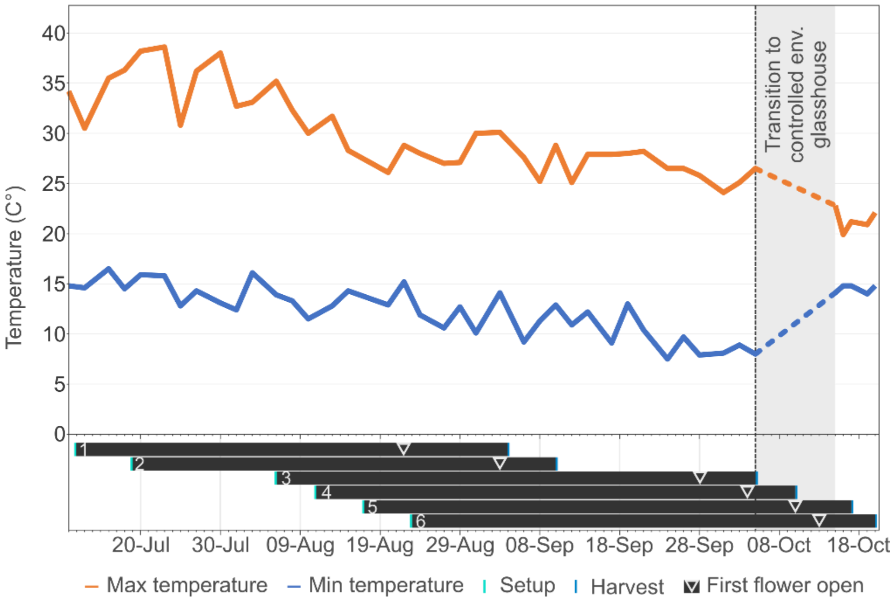


Figure S2.1. Maximum and minimum temperatures recorded during the screening experiments run in 2018. In the bottom panel, the duration of each experiment has been marked with a bar and the triangle indicates the time when the first flower opened in each experiment.

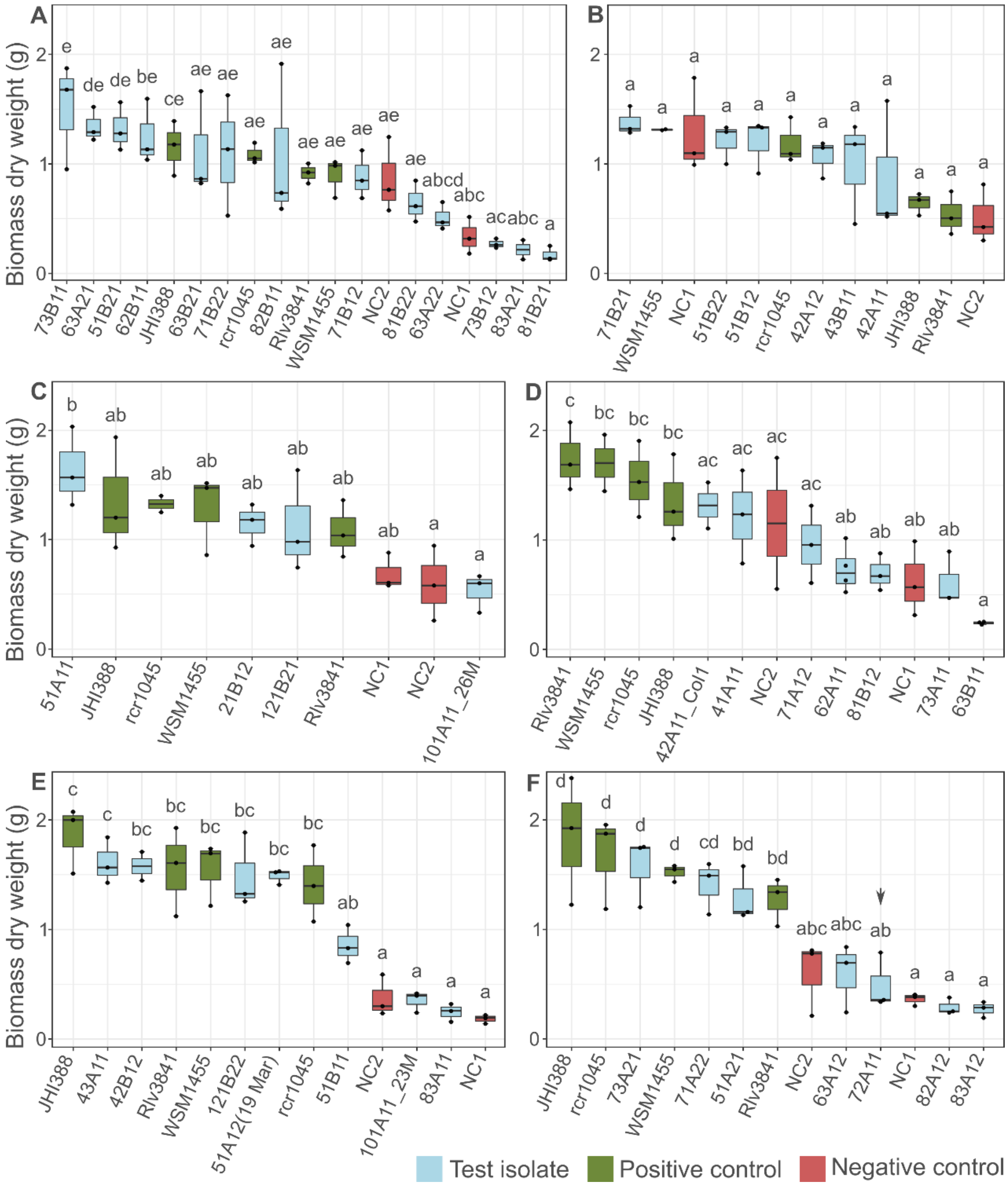


Figure S2.2. Box plot for biomass dry weight of all phase I screening experiments. Within each plot, boxes with different letters indicate significant differences between both treatments estimated by Tukey HSD test at 0.95 confidence interval. NC1 is a uninoculated negative control where SDW was used instead of an inoculum and NC2 is an inoculated treatment with an autoclaved grown culture. The arrow indicates a *Neorhizobium* strain tested.

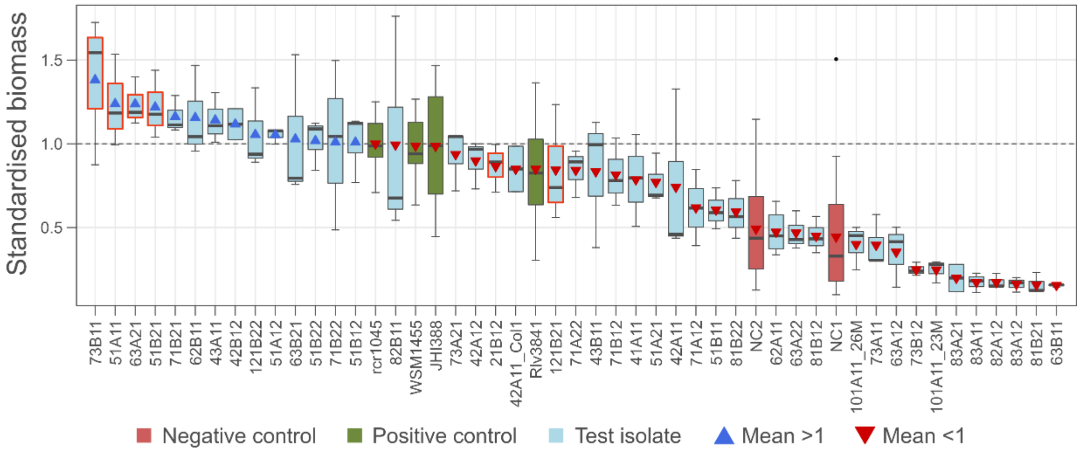


Figure S2.3. Boxplot of the standardised biomass dry weight. Standardisation was calculated by dividing each plant biomass dry weight by the mean biomass dry weight of the positive control rcr1045 within each individual experiment. Red bounded boxes are the strains that were selected for phase II screenings. The red and blue arrows within each box indicate whether the mean biomass of the treatment is greater (blue) or lower (red) than 1.

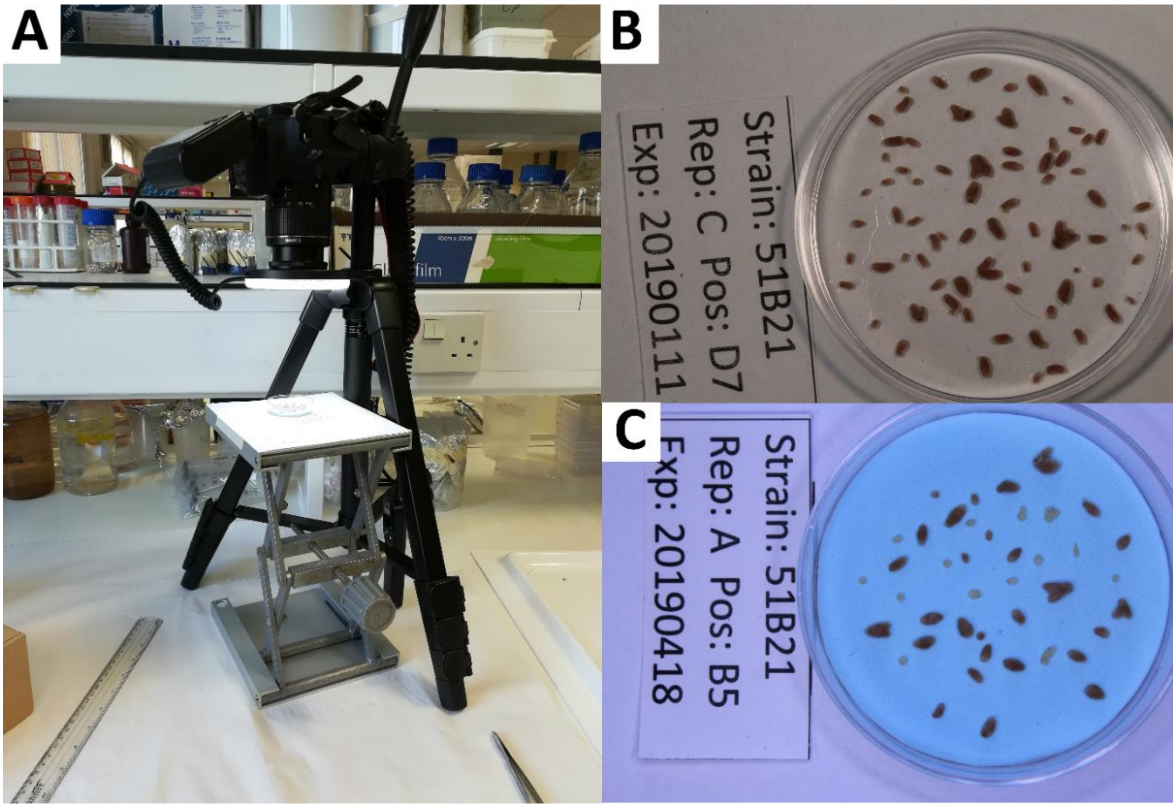


Figure S2.4. Nodule image capture display with Ring Flash RF-600D (A) and images captured without (B) and with ring and blue background (C).

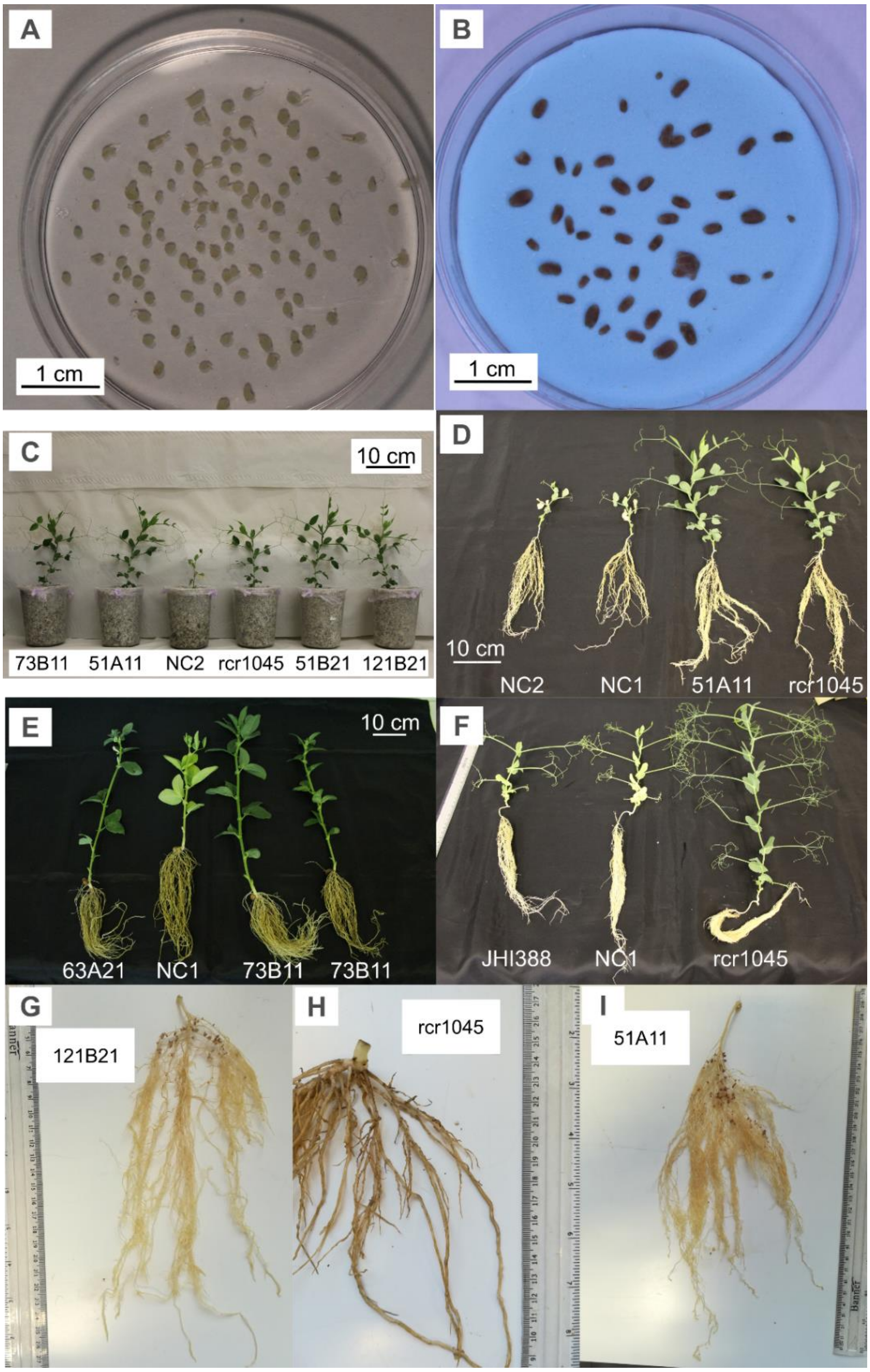
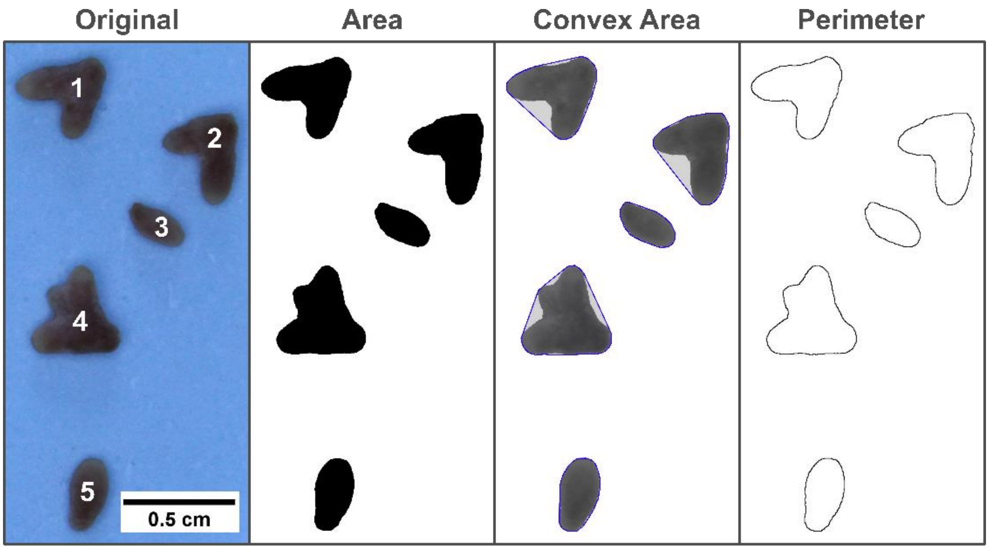


Figure S2.5. Screening experiments images. Nodules formed by strain JHI388 with atypical growth (A) and normal growth (B). Plants at harvest of pea cv. Corus in pots (C), uprooted (D), faba bean cv Fuego (G) and pea cv Kareni (F). Nodulated roots of pea cv Corus (G), faba bean cv Fuego inoculated with rcr1045 with small and un-harvestable nodules (H) and pea cv Kareni (I).



Nodule #	Area	Convex area	Perimeter	Length	Circularity	Solidity
1	0.082	0.097	1.325	0.349	0.587	0.855
2	0.081	0.096	1.312	0.381	0.592	0.855
3	0.031	0.032	0.729	0.280	0.730	0.971
4	0.099	0.112	1.396	0.378	0.636	0.884
5	0.045	0.047	0.870	0.336	0.745	0.974

Figure S2.6. Visual representation of the main shape parameters measurement and their measure. Area units are expressed in cm^2 and lengths in cm.

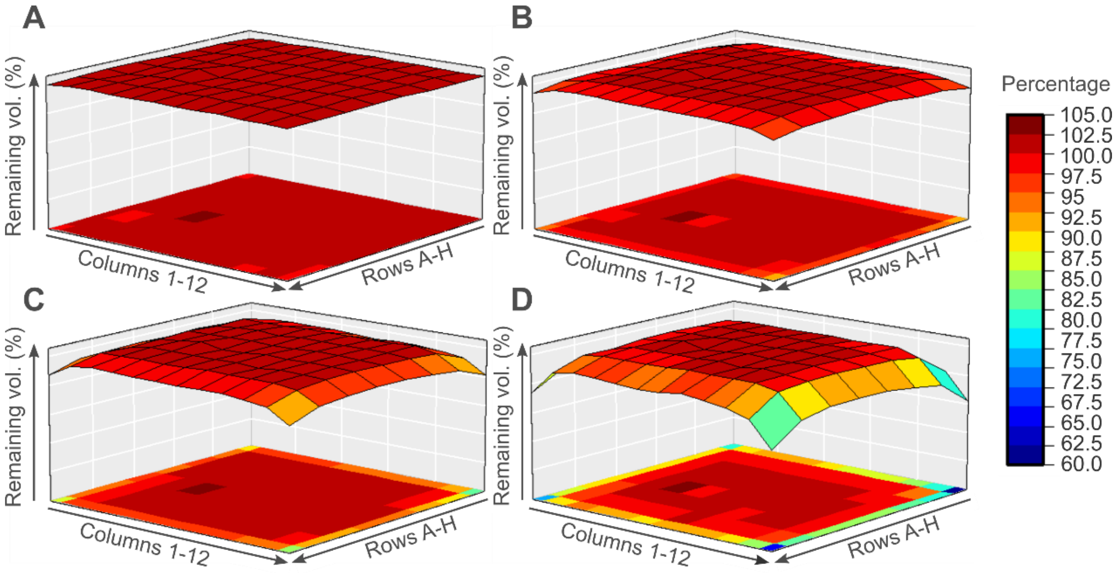


Figure S3.1. Evaporation pressures on each of the 96 wells of a plate during the 47 h incubation at Rleg growth conditions. Each plot represents the remaining percentage of the initial volume after 2 h (A), 12 h (B), 24 h (C) and 48 h (D).

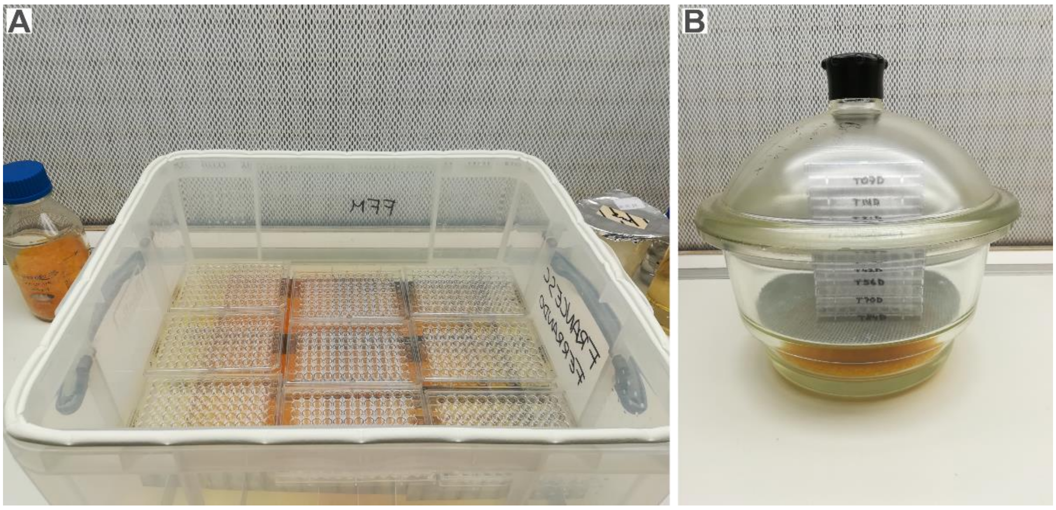


Figure S3.2. Custom made drying chamber that allocated all the necessary plates for the long-term desiccation assay (A). The plates were displayed flat on a grid with silica gel at the bottom mimicking a conventional drying chamber (B) to allow for uniform desiccation of all wells.

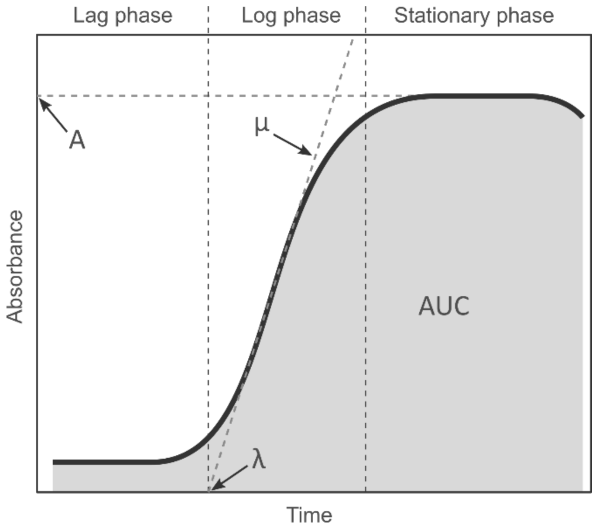


Figure S3.3. Growth curve parameters calculated with grofit. A - maximum cell growth, μ - growth rate, λ - lag time and AUC (shaded area) - area under the curve.

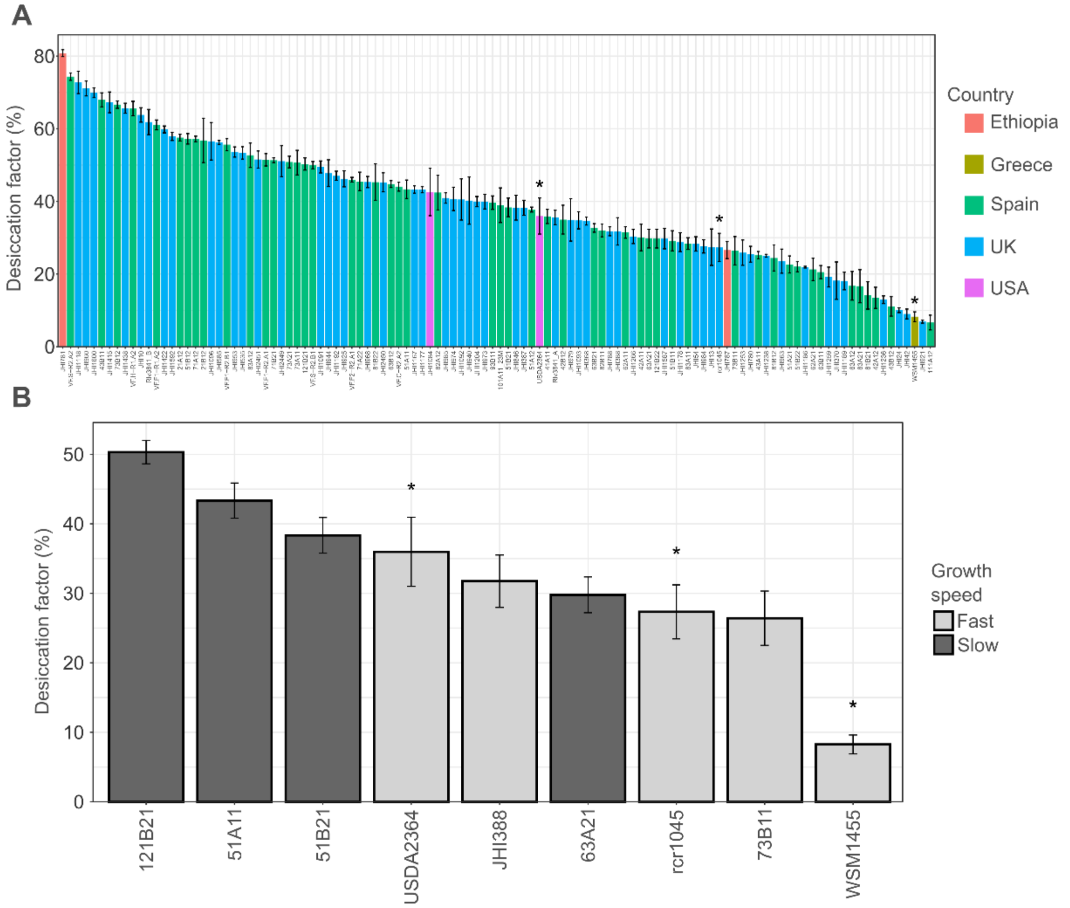


Figure S3.4. Desiccation factor of all assessed strains (A) and strains of known good symbiotic performance (B). The error bars on top of each bar represent the 95 % confidence interval. Bars marked with a star (*) are strains currently being used in commercial inoculants.

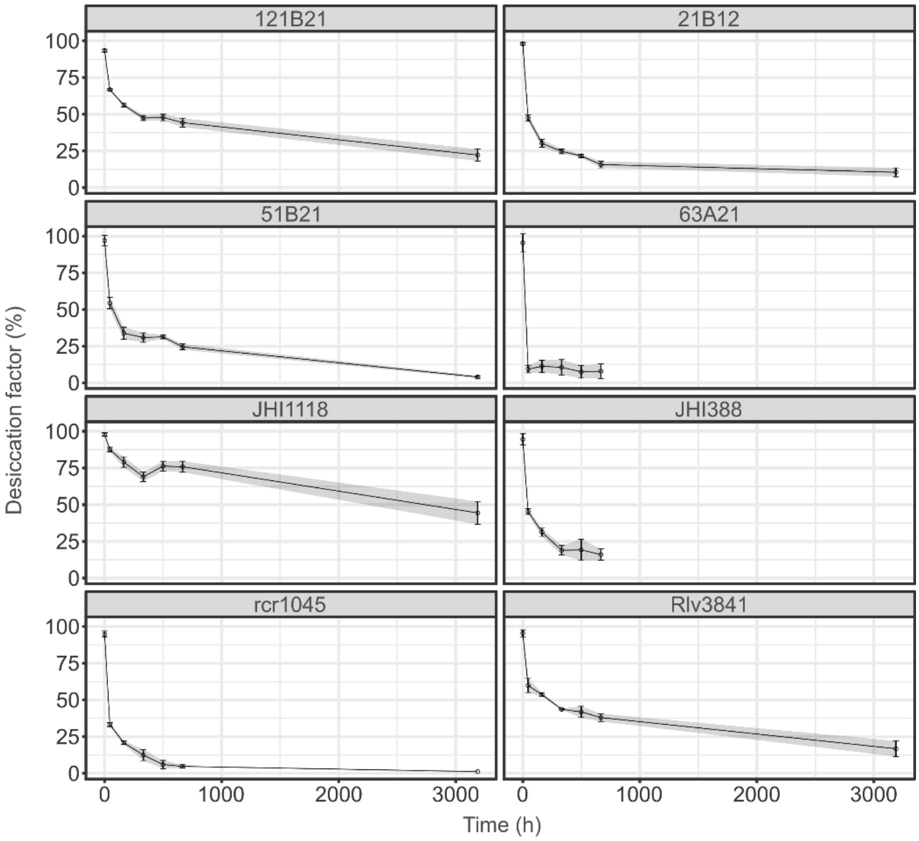


Figure S3.5. Long term exposure of Rlv strains to desiccation (2 - 133 days). The error bars at each datapoint and the shadowed area on both sides of the line represent the 95 % confidence interval.

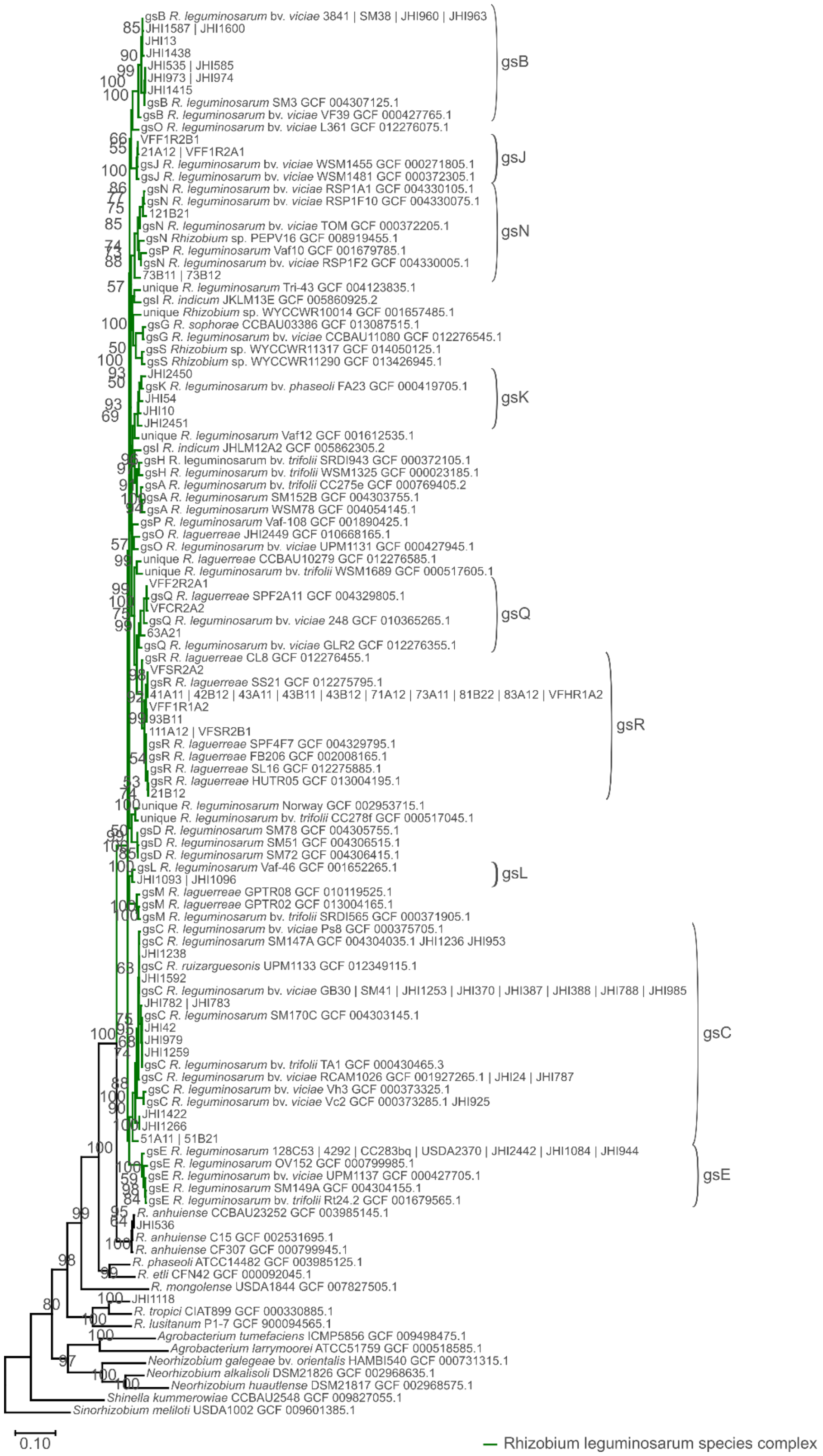


Figure S4.1. Phylogenetic tree of the concatenated sequence of genes atpD-gyrB-recA inferred by Maximum Likelihood. The values next to the nodes indicate the bootstrap value. The analysis involved 114 sequences and 4956 positions. The tree is drawn to scale with the bar indicating the number of base substitutions per site.

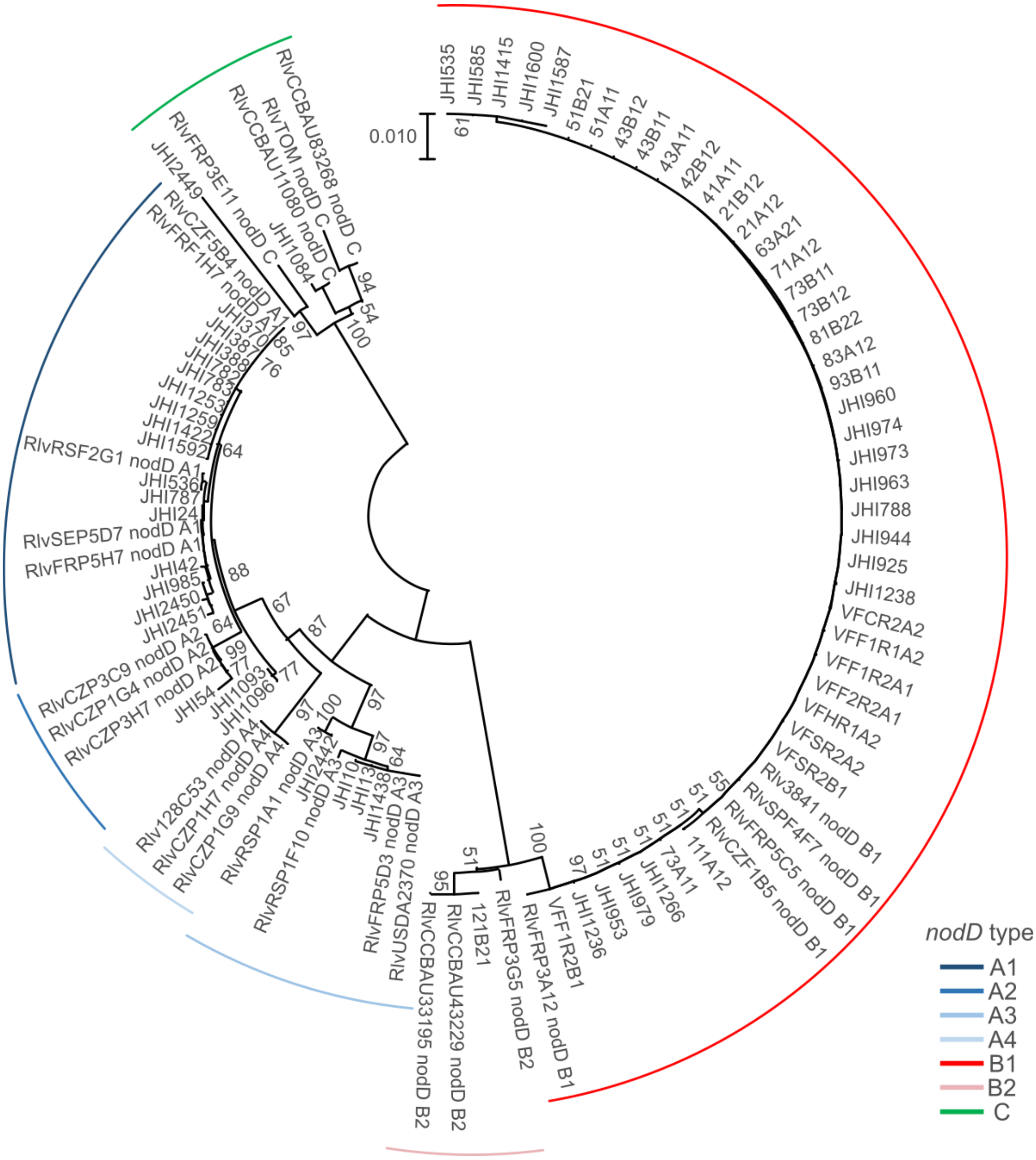


Figure S4.2. Phylogenetic tree of nodD sequences inferred by Neighbour-Joining. The values next to the nodes indicate the bootstrap value. The analysis involved 96 sequences and 933 positions. The tree is drawn to scale with the bar indicating the number of base substitutions per site.

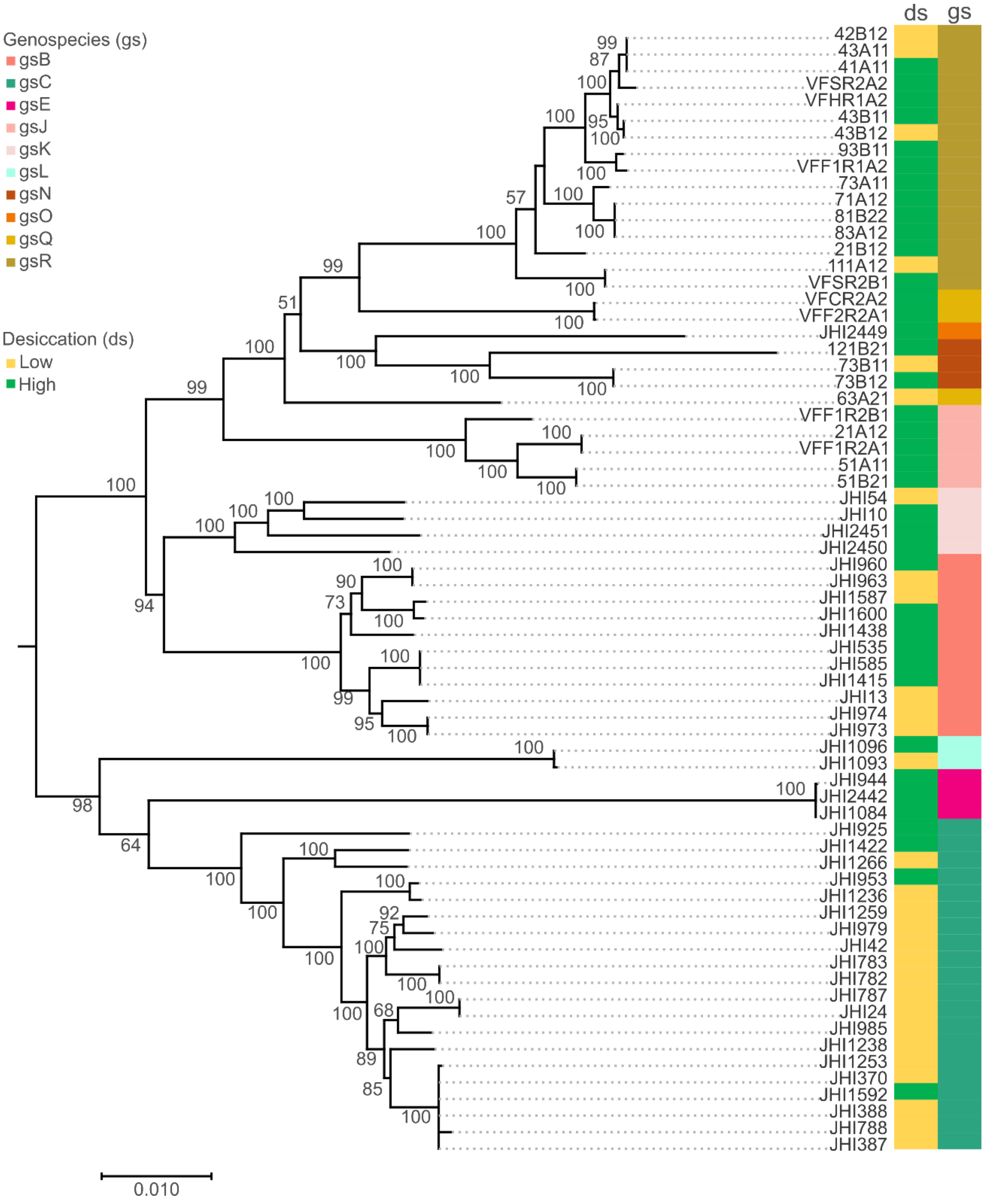


Figure S4.3. Phylogenetic tree of the concatenated sequences of literature desiccation genes sequences inferred by Neighbour-Joining. The values next to the nodes indicate the bootstrap value. The analysis involved 69 sequences and 33,612 positions. The tree is drawn to scale with the bar indicating the number of base substitutions per site.

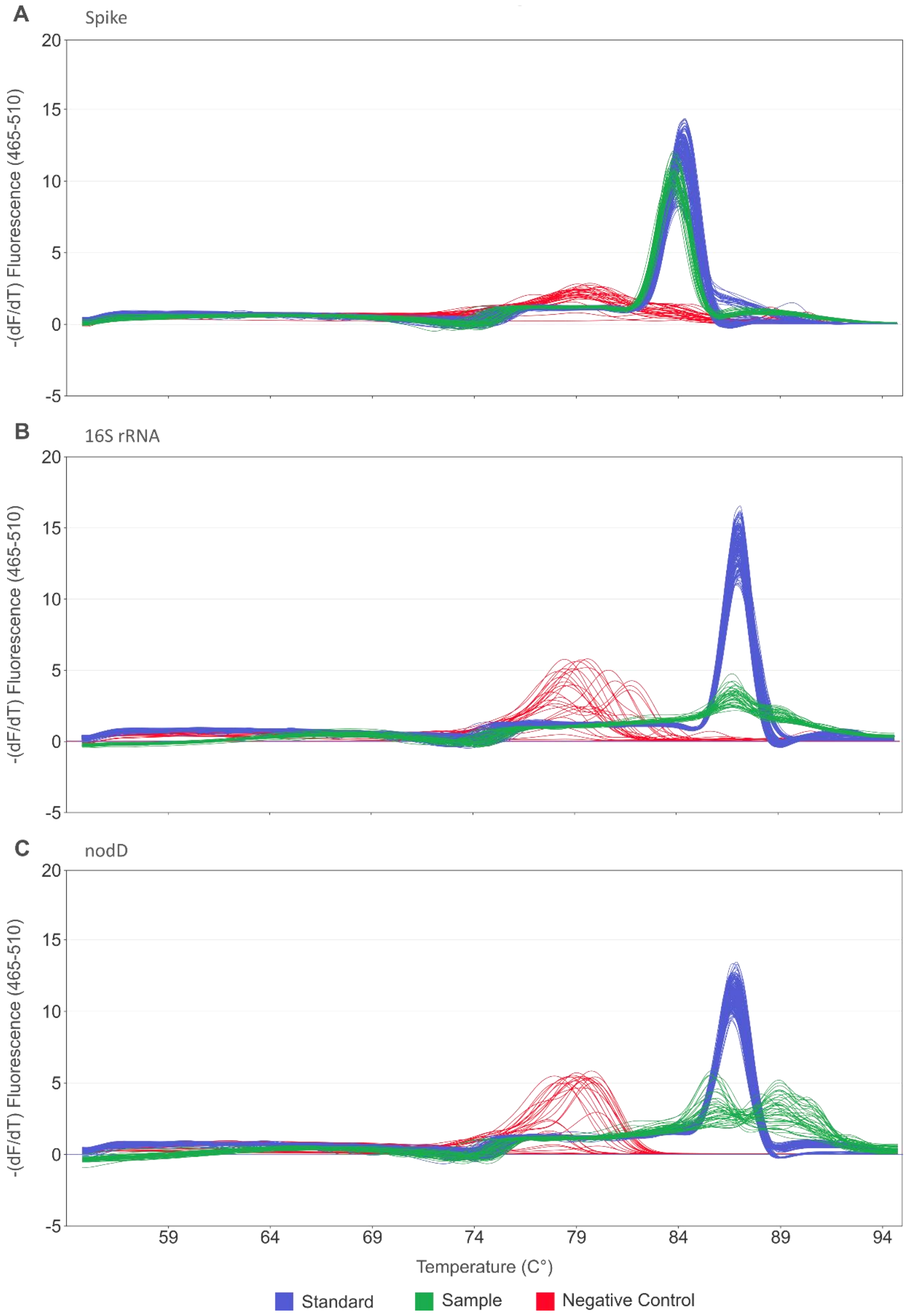


Figure S5.1. General overview of melting curves obtained after the PCR reaction for Spike (A), 16S rRNA (B) and *nodD* (C). While negative controls (red) and standard s(blue) contain all wells for all plates run, the sample curves (green) show a random selection of 40 wells to aid visualisation of the different curves.

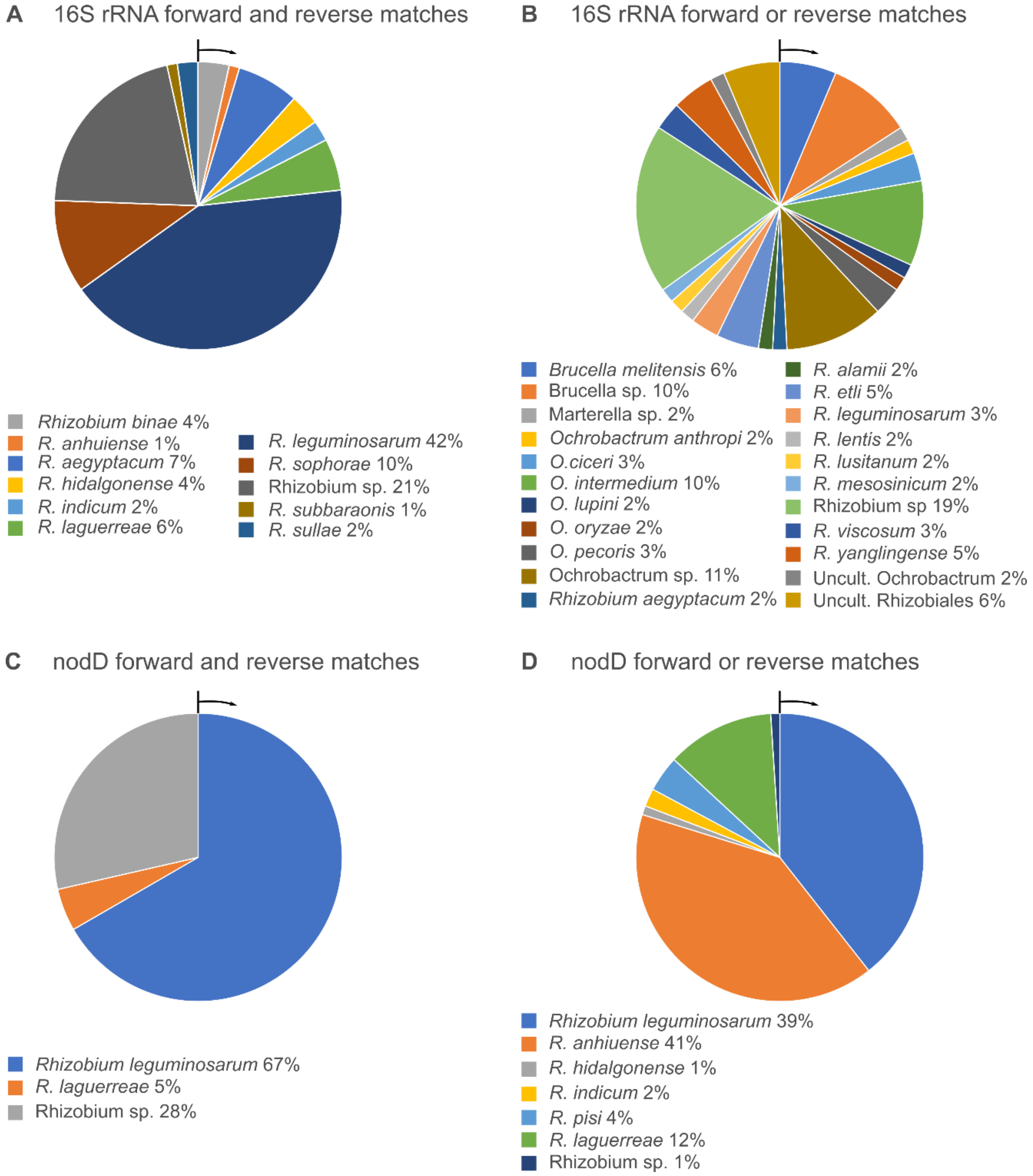


Figure S5.2. Results of the primer-matching sequences on NCBI for 16S rRNA (A and B) and *nodD* (C and D). Figures A and C show species which showed an exact match for both forward and reverse primers. Figures B and D show species which showed an exact match only for one of the primers, either forward or reverse. The legend under each pie chart show the species name followed by the percentage it represents over the total of sequences for each gene. The items in the legends are shown in clockwise order of appearance on the pie chart starting from the black bar and arrow.

Supplementary tables

Table S2.1. Positive control strains used in screening experiments.

Strain	Isolated from	Country of origin	References
Rlv 3841	Soil	UK	Johnston and Beringer (1975), Glenn *et al.* (1980)
rcr1045	*Pisum sativum* L.	Ireland	Dye (1978), Bitanyi (1983)
WSM1455	*Vicia faba* L.	Greece	Howieson *et al*. (2000), Bullard *et al.* (2005), Herridge (2008)
JHI388	*Pisum sativum* L.	Scotland	Maluk *et al.* (2022)
USDA2364	*Pisum sativum* L.	USA	van Berkum *et al.* (1995)

Table S2.2. Results of the NCBI Blast of the 16S rRNA gene sequence for each strain.

Isolate ID	**Location**	**Field**	**Species**	**Max Score**	**Total Score**	**Query Cover**	**E value**	**Per Ident**	**Accession**
11B11	Valencia	1	*Rhizobium* sp.	2503	2503	100%	0	99.85%	MN006388.1
12A11	Valencia	1	*Shinella* sp.	2505	2505	99%	0	99.85%	KJ510217.1
12B11	Valencia	1	*Rhizobium* sp.	2242	2242	100%	0	99.92%	MN498075.1
21A12	Valencia	1	*Rhizobium leguminosarum* bv. viciae	2512	7537	99%	0	99%	CP022564.1
21B12	Valencia	1	*Rhizobium leguminosarum* bv. trifolii	2538	2538	99%	0	99%	HQ836161.1
23A11	Valencia	1	*Rhizobium* sp.	2495	2495	99%	0	99.85%	EF549399.1
23A21	Valencia	1	*Rhizobium leguminosarum* bv. viciae	2468	7404	100%	0	99.70%	CP022564.1
23B11	Valencia	1	*Rhizobium leguminosarum* bv. trifolii	2527	2527	99%	0	99.93%	JF810501.1
32A11	Valencia	1	*Rhizobium* sp.	2495	2495	100%	0	99.85%	MN006388.1
32B11	Valencia	1	*Rhizobium* sp.	2483	2483	100%	0	99.93%	MN006388.1
41A11	Valencia	2	*Rhizobium leguminosarum*	2527	2527	99%	0	99%	JN105994.1
41A12	Valencia	2	*Rhizobium leguminosarum* bv. trifolii	2525	2525	99%	0	99%	JF810501.1
41A13	Valencia	2	Uncultured bacterium	2481	2481	100%	0	99%	MG744662.1
42A11	Valencia	2	*Rhizobium leguminosarum* bv. trifolii	2532	2532	99%	0	99%	JF810501.1
42A11_Col_1	Valencia	2	*Rhizobium leguminosarum* bv. trifolii	2531	2531	99%	0	99%	JF810501.1
42A12	Valencia	2	*Rhizobium leguminosarum* bv. trifolii	2532	2532	99%	0	99%	JF810501.1
42B12	Valencia	2	*Rhizobium* sp.	2523	2523	99%	0	100%	KM999134.1
43A11	Valencia	2	*Rhizobium leguminosarum*	2503	2503	100%	0	99%	KY784928.1
43B11	Valencia	2	*Rhizobium leguminosarum* bv. trifolii	2531	2531	99%	0	99%	JF810501.1
43B12	Valencia	2	*Rhizobium* sp.	2446	2446	100%	0	99%	KM999134.1
51A11	Valencia	3	*Rhizobium leguminosarum* bv. viciae	2486	7460	99%	0	99%	CP022564.1
51A12	Valencia	3	*Rhizobium* sp.	2497	2497	99%	0	99%	MF624031.1

Table S2.2. (Continuation)

Isolate ID	Location	Field	Species	Max Score	Total Score	Query Cover	E value	Per Ident	Accession
51A21	Valencia	3	*Rhizobium leguminosarum* bv. viciae	2477	7432	99%	0	99%	CP022564.1
51A22	Valencia	3	*Rhizobium leguminosarum*	2372	2372	100%	0	99%	EU256420.1
51B11	Valencia	3	*Rhizobium leguminosarum*	2486	2486	99%	0	99%	KF662884.2
51B12	Valencia	3	*Rhizobium leguminosarum*	2470	2470	100%	0	99%	EU256420.1
51B21	Valencia	3	*Rhizobium leguminosarum* bv. viciae	2484	7454	98%	0	99%	CP022564.1
51B22	Valencia	3	*Rhizobium leguminosarum*	2390	2390	100%	0	99%	EU256420.1
52A12	Valencia	3	*Rhizobium leguminosarum* bv. trifolii	2529	2529	99%	0	99%	JF810501.1
52B11	Valencia	3	*Rhizobium leguminosarum* bv. trifolii	2527	2527	99%	0	99%	HQ836161.1
53B11	Valencia	3	*Rhizobium leguminosarum*	2473	2473	100%	0	99.70%	EU256420.1
53B12	Valencia	3	*Rhizobium leguminosarum*	2497	2497	100%	0	100.00%	EU256420.1
53B21	Valencia	3	*Rhizobium* sp.	2329	2329	100%	0	99.76%	MN498075.1
53B22	Valencia	3	*Rhizobium leguminosarum*	2492	2492	100%	0	99.93%	EU256420.1
61B12_06Jun	Valencia	3	*Rhizobium leguminosarum* bv. trifolii	2523	2523	100%	0	99.93%	JF810501.1
61B12_15May	Valencia	3	*Rhizobium* sp.	2490	2490	100%	0	99.93%	KM999134.1
61B21	Valencia	3	*Rhizobium* sp.	2486	2486	100%	0	99.93%	KM999134.1
62A11	Valencia	3	*Rhizobium laguerreae*	2494	2494	100%	0	100%	FJ595999.3
62A12	Valencia	3	*Rhizobium leguminosarum* bv. trifolii	2519	2519	99%	0	99%	JF810501.1
62A21	Valencia	3	*Rhizobium* sp.	2521	2521	99%	0	99%	KM999134.1
62B11	Valencia	3	*Rhizobium leguminosarum*	2501	2501	100%	0	99%	KY784928.1
63A11	Valencia	3	*Sphingomonas* sp.	2473	2473	100%	0	99%	HM484354.2
63A12	Valencia	3	*Rhizobium leguminosarum*	2510	2510	100%	0	99%	KY784928.1
63A21	Valencia	3	*Rhizobium laguerreae*	2405	2405	100%	0	100%	FJ595999.3
63A22	Valencia	3	*Rhizobium leguminosarum*	2510	2510	100%	0	99%	KY784928.1

Table S2.2. (Continuation)

Isolate ID	Location	Field	Species	Max Score	Total Score	Query Cover	E value	Per Ident	Accession
63B11	Valencia	3	*Rhizobium leguminosarum*	2508	2508	100%	0	99%	KY784928.1
63B12	Valencia	3	*Rhizobium leguminosarum*	2484	2484	100%	0	99%	KY784928.1
63B21	Valencia	3	*Rhizobium laguerreae*	2473	2473	100%	0	99%	FJ595999.3
71A11	Ontinyent	4	*Paenibacillus* sp.	2494	2494	100%	0	99%	KR051041.1
71A12	Ontinyent	4	*Rhizobium leguminosarum*	2494	2494	100%	0	99%	KY784928.1
71A22	Ontinyent	4	*Rhizobium leguminosarum*	2508	2508	100%	0	99%	KY784928.1
71B12	Ontinyent	4	*Rhizobium* sp.	2494	2494	100%	0	99%	MF624038.1
71B21	Ontinyent	4	*Rhizobium leguminosarum*	2499	2499	100%	0	99%	KY784928.1
71B22	Ontinyent	4	*Rhizobium laguerreae*	2479	2479	100%	0	99%	FJ595999.3
72A11	Ontinyent	4	*Neorhizobium* sp.	2518	2518	100%	0	99%	MH064335.1
72A12	Ontinyent	4	*Neorhizobium* sp.	2497	2497	99%	0	100%	MH064335.1
72A21	Ontinyent	4	*Neorhizobium* sp.	2512	2512	99%	0	99%	MH064335.1
72A22	Ontinyent	4	*Neorhizobium* sp.	2497	2497	100%	0	99%	MH064335.1
72B11	Ontinyent	4	*Rhizobium* sp.	2494	2494	100%	0	99%	MF624038.1
72B12	Ontinyent	4	*Methylobacterium* sp.	2405	2405	100%	0	99%	MG807376.1
73A11	Ontinyent	4	*Rhizobium laguerreae*	2466	2466	100%	0	99%	FJ595999.3
73A12	Ontinyent	4	*Methylobacterium* sp.	2394	2394	100%	0	100%	MG807376.1
73A21	Ontinyent	4	*Rhizobium* sp.	2501	2501	100%	0	99%	MF624038.1
73B11	Ontinyent	4	*Rhizobium leguminosarum*	2098	2098	100%	0	98%	GU552880.1
73B12	Ontinyent	4	*Rhizobium leguminosarum*	2272	2272	100%	0	99%	KY587906.1
81B11	Ontinyent	4	*Brevundimonas vesicularis*	2342	4685	100%	0	99%	CP022048.2
81B12	Ontinyent	4	*Rhizobium laguerreae*	2497	2497	99%	0	100%	FJ595999.3
81B21	Ontinyent	4	*Rhizobium laguerreae*	2468	2468	100%	0	100%	FJ595999.3
81B22	Ontinyent	4	*Rhizobium laguerreae*	2436	2436	99%	0	100%	FJ595999.3

Table S2.2. (Continuation)

Isolate ID	**Location**	**Field**	**Species**	**Max Score**	**Total Score**	**Query Cover**	**E value**	**Per Ident**	**Accession**
82A12	Ontinyent	4	*Rhizobium leguminosarum*	2490	2490	100%	0	99%	KY784928.1
82B11	Ontinyent	4	*Rhizobium leguminosarum*	2350	7050	100%	0	99%	CP025012.1
83A11	Ontinyent	4	*Rhizobium laguerreae*	2484	2484	99%	0	99%	FJ595999.3
83A12	Ontinyent	4	*Rhizobium leguminosarum*	2510	2510	100%	0	99%	KY784928.1
83A21	Ontinyent	4	*Rhizobium laguerreae*	2475	2475	99%	0	100%	FJ595999.3
83B12	Ontinyent	4	*Bacillus pumilus*	2619	2619	100%	0	99%	KC692196.1
91A12	Ontinyent	4	*Agrobacterium tumefaciens*	2510	2510	99%	0	99.93%	KP762564.1
91B12	Ontinyent	4	*Rhizobium leguminosarum*	2508	2508	100%	0	99%	KY784928.1
92A12	Ontinyent	4	*Agrobacterium tumefaciens*	2521	2521	100%	0	99%	MH236271.1
92A12_ER	Ontinyent	4	*Agrobacterium* sp.	2453	2453	100%	0	99%	LC385681.1
93A12	Ontinyent	4	*Rhizobium* sp.	2494	2494	100%	0	99.93%	MN006388.1
93B11	Ontinyent	4	*Rhizobium laguerreae*	2451	2451	100%	0	99%	FJ595999.3
101A11_19M	Ontinyent	5	*Rhizobium laguerreae*	2473	2473	100%	0	99%	FJ595999.3
101A11_23M	Ontinyent	5	*Rhizobium laguerreae*	2451	2451	100%	0	99%	FJ595999.3
101B11	Ontinyent	5	*Cupriavidus gilardii*	2556	2556	100%	0	100%	AY860231.1
101B21	Ontinyent	5	*Methylobacterium* sp.	2451	2451	100%	0	99%	MG798746.1
111A12	Ontinyent	5	*Rhizobium laguerreae*	2470	2470	100%	0	100%	FJ595999.3
111B11	Ontinyent	5	*Paenibacillus* sp.	2558	2558	100%	0	99%	KC236524.1
111B12	Ontinyent	5	*Agrobacterium tumefaciens*	1600	1675	97%	0	88%	LT630451.1
121A12	Ontinyent	5	*Methylobacterium* sp.	2451	2451	100%	0	99%	MG798746.1
121B21	Ontinyent	5	*Rhizobium leguminosarum*	2475	2475	100%	0	99%	KY587906.1
121B22	Ontinyent	5	*Rhizobium leguminosarum*	2466	2466	100%	0	99%	KY587906.1
NC1A11	Neg ctrl	-	Uncultured *Rhizobium*	2379	2379	100%	0	99%	MH236575.1
NC1A21	Neg ctrl	-	Uncultured *Rhizobium*	2464	2464	100%	0	100%	MH236575.1

Table S2.2. (Continuation)

Isolate ID	**Location**	**Field**	**Species**	**Max Score**	**Total Score**	**Query Cover**	**E value**	**Per Ident**	**Accession**
NC1A22	Neg ctrl	-	*Rhizobium leguminosarum*	2481	2481	100%	0	99%	MF624030.1
NC1B11	Neg ctrl	-	*Cupriavidus gilardii*	2579	2579	100%	0	100%	AY860231.1
NC1B12	Neg ctrl	-	Uncultured *Rhizobium*	2479	2479	99%	0	99%	MH236575.1
NC1B21	Neg ctrl	-	Uncultured *Rhizobium*	2507	2507	99%	0	99%	MH236575.1
NC1B22	Neg ctrl	-	Uncultured *Rhizobium*	2499	2499	99%	0	99%	MH236575.1
NC2A12	Neg ctrl	-	*Methylobacterium* sp.	2427	2427	100%	0	100%	MG807376.1
NC2B11	Neg ctrl	-	*Rhizobium* sp.	2507	2507	99%	0	99%	EF437252.1
NC2B12	Neg ctrl	-	*Rhizobium leguminosarum*	2481	2481	100%	0	100%	KY587906.1
NC2B21	Neg ctrl	-	*Rhizobium leguminosarum*	2512	2512	99%	0	99%	KY587906.1
NC2B22	Neg ctrl	-	*Rhizobium leguminosarum*	2497	2497	100%	0	99%	KY587906.1

Table S4.1. Reference genomes used for the concatenated *atpD-gyrB-recA* phylogenetic and ANI analyses. Gs - Genospecies.

Species	**Strain**	**Accession number**	**Gs**
Rhizobium etli	CFN42	GCF_000092045.1	-
Rhizobium tropici	CIAT899	GCF_000330885.1	-
Agrobacterium larrymoorei	ATCC51759	GCF_000518585.1	-
Neorhizobium galegeae bv. *orientalis*	HAMBI540	GCF_000731315.1	-
Neorhizobium huautlense	DSM21817	GCF_002968575.1	-
Neorhizobium alkalisoli	DSM21826	GCF_002968635.1	-
Rhizobium phaseoli	ATCC14482	GCF_003985125.1	-
Rhizobium mongolense	USDA1844	GCF_007827505.1	-
Agrobacterium tumefaciens	ICMP5856	GCF_009498475.1	-
Sinorhizobium meliloti	USDA1002	GCF_009601385.1	-
Shinella kummerowiae	CCBAU2548	GCF_009827055.1	-
Rhizobium lusitanum	P1-7	GCF_900094565.1	-
Rhizobium leguminosarum bv. *trifolii*	CC275e	GCF_000769405.2	A
Rhizobium leguminosarum	WSM78	GCF_004054145.1	A
Rhizobium leguminosarum	SM152B	GCF_004303755.1	A
Rhizobium leguminosarum bv. *viciae*	3841	GCF_000009265.1	B
Rhizobium leguminosarum bv. *viciae*	VF39	GCF_000427765.1	B
Rhizobium leguminosarum	SM38	GCF_004306065.1	B
Rhizobium leguminosarum	SM3	GCF_004307125.1	B
Rhizobium leguminosarum bv. *viciae*	Vc2	GCF_000373285.1	C
Rhizobium leguminosarum bv. *viciae*	Vh3	GCF_000373325.1	C
Rhizobium leguminosarum bv. *viciae*	Ps8	GCF_000375705.1	C
Rhizobium leguminosarum bv. *viciae*	GB30	GCF_000419745.1	C
Rhizobium leguminosarum bv. *trifolii*	TA1	GCF_000430465.3	C
Rhizobium leguminosarum bv. *viciae*	RCAM1026	GCF_001927265.1	C
Rhizobium leguminosarum	SM170C	GCF_004303145.1	C
Rhizobium leguminosarum	SM147A	GCF_004304035.1	C
Rhizobium leguminosarum	SM41	GCF_004305845.1	C
Rhizobium ruizarguesonis	UPM1133	GCF_012349115.1	C
Rhizobium leguminosarum	SM78	GCF_004305755.1	D
Rhizobium leguminosarum	SM72	GCF_004306415.1	D
Rhizobium leguminosarum	SM51	GCF_004306515.1	D
Rhizobium leguminosarum bv. *viciae*	128C53	GCF_000373425.1	E
Rhizobium leguminosarum bv. *phaseoli*	4292	GCF_000379005.1	E
Rhizobium leguminosarum bv. *viciae*	UPM1137	GCF_000427705.1	E
Rhizobium leguminosarum bv. *trifolii*	CC283bq	GCF_000515375.1	E
Rhizobium leguminosarum	OV152	GCF_000799985.1	E
Rhizobium leguminosarum bv. *trifolii*	Rt24.2	GCF_001679565.1	E
Rhizobium leguminosarum bv. *viciae*	USDA2370	GCF_003058385.1	E
Rhizobium leguminosarum	SM149A	GCF_004304155.1	E
Rhizobium leguminosarum bv. *viciae*	CCBAU11080	GCF_012276545.1	G
Rhizobium sophorae	CCBAU03386	GCF_013087515.1	G
Rhizobium leguminosarum bv. *trifolii*	WSM1325	GCF_000023185.1	H
Rhizobium leguminosarum bv. *trifolii*	SRDI943	GCF_000372105.1	H

Table S4.1. (Continuation)

Species	**Strain**	**Accession number**	**Gs**
Rhizobium indicum	JKLM13E	GCF_005860925.2	I
Rhizobium indicum	JHLM12A2	GCF_005862305.2	I
Rhizobium leguminosarum bv. *viciae*	WSM1455	GCF_000271805.1	J
Rhizobium leguminosarum bv. *viciae*	WSM1481	GCF_000372305.1	J
Rhizobium leguminosarum bv. *phaseoli*	FA23	GCF_000419705.1	K
Rhizobium leguminosarum	Vaf-46	GCF_001652265.1	L
Rhizobium leguminosarum bv. *trifolii*	SRDI565	GCF_000371905.1	M
Rhizobium laguerreae	GPTR08	GCF_010119525.1	M
Rhizobium laguerreae	GPTR02	GCF_013004165.1	M
Rhizobium leguminosarum bv. *viciae*	TOM	GCF_000372205.1	N
Rhizobium leguminosarum bv. *viciae*	RSP1F2	GCF_004330005.1	N
Rhizobium leguminosarum bv. *viciae*	RSP1F10	GCF_004330075.1	N
Rhizobium leguminosarum bv. *viciae*	RSP1A1	GCF_004330105.1	N
Rhizobium sp.	PEPV16	GCF_008919455.1	N
Rhizobium leguminosarum bv. *viciae*	UPM1131	GCF_000427945.1	O
Rhizobium laguerreae	JHI2449	GCF_010668165.1	O
Rhizobium leguminosarum bv. *viciae*	L361	GCF_012276075.1	O
Rhizobium leguminosarum	Vaf10	GCF_001679785.1	P
Rhizobium leguminosarum	Vaf-108	GCF_001890425.1	P
Rhizobium laguerreae	SPF2A11	GCF_004329805.1	Q
Rhizobium leguminosarum bv. *viciae*	248	GCF_010365265.1	Q
Rhizobium leguminosarum bv. *viciae*	GLR2	GCF_012276355.1	Q
Rhizobium laguerreae	FB206	GCF_002008165.1	R
Rhizobium laguerreae	SPF4F7	GCF_004329795.1	R
Rhizobium laguerreae	SS21	GCF_012275795.1	R
Rhizobium laguerreae	SL16	GCF_012275885.1	R
Rhizobium laguerreae	CL8	GCF_012276455.1	R
Rhizobium laguerreae	HUTR05	GCF_013004195.1	R
Rhizobium sp.	WYCCWR11290	GCF_013426945.1	S
Rhizobium sp.	WYCCWR11317	GCF_014050125.1	S
Rhizobium leguminosarum	CF307	GCF_000799945.1	-
Rhizobium anhuiense	C15	GCF_002531695.1	-
Rhizobium anhuiense	CCBAU23252	GCF_003985145.1	-
Rhizobium leguminosarum bv. *trifolii*	CC278f	GCF_000517045.1	unique
Rhizobium leguminosarum bv. *trifolii*	WSM1689	GCF_000517605.1	unique
Rhizobium leguminosarum	Vaf12	GCF_001612535.1	unique
Rhizobium sp.	WYCCWR10014	GCF_001657485.1	unique
Rhizobium leguminosarum	Norway	GCF_002953715.1	unique
Rhizobium leguminosarum	Tri-43	GCF_004123835.1	unique
Rhizobium laguerreae	CCBAU10279	GCF_012276585.1	unique

Table S4.2. Boivin et al (2020) *nodD* types reference sequences used for the phylogenetic analysis for *nodD* type assignation.

Species	Strain	Accession	Type
Rhizobium leguminosarum bv. *viciae*	FRF1H7	SJMX01000082.1	A1
Rhizobium leguminosarum bv. *viciae*	FRP5H7	SJML01000007.1	A1
Rhizobium leguminosarum bv. *viciae*	SEP5D7	SJLW01000027.1	A1
Rhizobium leguminosarum bv. *viciae*	CZF5B4	SJNL01000008.1	A1
Rhizobium leguminosarum bv. *viciae*	RSF2G1	SJMJ01000015.1	A1
Rhizobium leguminosarum bv. *viciae*	CZP1G4	SJNI01000014.1	A2
Rhizobium leguminosarum bv. *viciae*	CZP3C9	SJND01000016.1	A2
Rhizobium leguminosarum bv. *viciae*	CZP3H7	SJNA01000045.1	A2
Rhizobium leguminosarum bv. *viciae*	USDA2370	MRDL01000023.1	A3
Rhizobium leguminosarum bv. *viciae*	RSP1A1	SJMI01000030.1	A3
Rhizobium leguminosarum bv. *viciae*	FRP5D3	SJMM01000012.1	A3
Rhizobium leguminosarum bv. *viciae*	RSP1F10	SJMF01000035.1	A3
Rhizobium leguminosarum bv. *viciae*	128C53	ARDW01000025.1	A4
Rhizobium leguminosarum bv. *viciae*	CZP1G9	SJNH01000008.1	A4
Rhizobium leguminosarum bv. *viciae*	CZP1H7	SJNG01000007.1	A4
Rhizobium leguminosarum bv. *viciae*	CZF1B5	SJNN01000062.1	B1
Rhizobium leguminosarum bv. *viciae*	FRP3A12	SJMU01000027.1	B1
Rhizobium leguminosarum bv. *viciae*	FRP5C5	SJMN01000002.1	B1
Rhizobium leguminosarum bv. *viciae*	Rlv3841	AM236084.1	B1
Rhizobium laguerreae	SPF4F7	SJNO01000036.1	B1
Rhizobium leguminosarum bv. *viciae*	CCBAU33195	WIEM01000039.1	B2
Rhizobium anhuiense	CCBAU43229	WIFM01000024.1	B2
Rhizobium leguminosarum bv. *viciae*	FRP3G5	SJMR01000006.1	B2
Rhizobium leguminosarum bv. *viciae*	CCBAU11080	WIEN01000030.1	C
Rhizobium leguminosarum bv. *viciae*	CCBAU83268	WIFD01000010.1	C
Rhizobium leguminosarum bv. *viciae*	FRP3E11	SJMT01000030.1	C
Rhizobium leguminosarum bv. *viciae*	TOM	AQUC01000005.1	C

Table S4.3. Genes involved in desiccation stress response found in the literature on rhizobia or N-fixing organisms.

Gene	**Homologous on Rlv3841**	**Processes**	**Organism studied on**	**Reference**
betR	no homologue	Osmoprotectant production	*Klebsiella variicola*	Rodriguez-Andrade *et al.* (2019)
mutL	-	DNA repair	*Pseudomonas putida*	Pazos-Rojas *et al.* (2019)
mutS	-	DNA repair	*Pseudomonas putida*	Pazos-Rojas *et al.* (2019)
oprH	no homologue	DNA repair	*Pseudomonas putida*	Pazos-Rojas *et al.* (2019)
otsA	-	Osmoprotectant production	*Rhizobium leguminosarum/ R. laguerreae/ Ensifer meliloti*	Benidire *et al.* (2018)
otsB	-	Osmoprotectant production	*Rhizobium leguminosarum/ R. laguerreae/ Ensifer meliloti*	Benidire *et al.* (2018)
kup/trkD	-	Stress responses	*Rhizobium leguminosarum/ R. laguerreae/ Ensifer meliloti*	Benidire *et al.* (2018)
betB	-	Osmoprotectant production	*Sinorhizobium meliloti*	Boscari *et al.* (2002)
betS/betP	no homologue	Osmoprotectant production	*Sinorhizobium meliloti*	Boscari *et al.* (2002)
rpoE2	-	Stress responses	*Sinorhizobium meliloti*	Humann *et al.* (2009)
treS	-	Osmoprotectant production	*Bradyrhizobium japonicum*	Sugawara *et al.* (2010)
RL4716	-	LPS/EPS production	*Rhizobium leguminosarum*	Neudorf *et al.* (2017)
asnO	no homologue	Stress responses	*Sinorhizobium meliloti*	Vriezen *et al.*(2012)
ngg	no homologue	Stress responses	*Sinorhizobium meliloti*	Vriezen *et al.*(2012)
uvrA	-	DNA repair	*Sinorhizobium meliloti*	Humann *et al.* (2009)
uvrB	-	DNA repair	*Sinorhizobium meliloti*	Humann *et al.* (2009)
uvrC	-	DNA repair	*Sinorhizobium meliloti*	Humann *et al.* (2009)
ecfG	*rpoZ*	Stress responses	*Bradyrhizobium japonicum*	Gourion *et al.* (2015)
hpr	RL0032	Stress responses	*Sinorhizobium meliloti*	Humann *et al.* (2009)
phyR	RL3705	Stress responses	*Bradyrhizobium japonicum*	Gourion *et al.* (2015)
phyR	RL3705	Stress responses	*Methylobacterium extorquens*	Gourion *et al.* (2015)
relA	-	Stress responses	*Sinorhizobium meliloti*	Humann *et al.* (2009)
rpoE2	*rpoZ*	Stress responses	*Sinorhizobium meliloti*	Humann *et al.* (2009)

Table S4.3. (Continuation)

Gene	Homologous on Rlv3841	Processes	Organism studied on	Reference
rsiB1	*phyR*	Stress responses	*Sinorhizobium meliloti*	Humann and Kahn (2015)
ctpA	-	Protein structure	*Rhizobium leguminosarum*	Gilbert *et al.* (2007)
fabF1	-	LPS/EPS production	*Rhizobium leguminosarum*	Vanderlinde *et al.* (2010)
fabF2	-	LPS/EPS production	*Rhizobium leguminosarum*	Vanderlinde *et al.* (2010)
RL2975	-	LPS/EPS production	*Rhizobium leguminosarum*	Vanderlinde *et al.* (2010)
otsA	-	Osmoprotectant production	*Rhizobium leguminosarum*	McIntyre *et al.* (2007)
treY	-	Osmoprotectant production	*Rhizobium leguminosarum*	McIntyre *et al.* (2007)

Table S4.4. Genome assembly results of Rlc strains after assembly with Jigome, analysis of quality with Quast and annotation with Prokka

		Jigome					Quast					Prokka				
Strain	Genospecies	Median coverage	Chromosomal length (bp)	Plasmid length (bp)	Fragment length (bp)	Total length (bp)	Contigs	GC (%)	N50	L50	N's per 100 kbp	CDS	Gene	rRNA	tRNA	tmRNA
111A12	R	18.85	5577104	1115525	305687	6998316	39	60.91	1237194	2	0.57	6734	6793	9	49	1
121B21	N	25.98	4987946	1446620	939887	7374453	49	60.78	491596	6	0.54	7075	7128	3	49	1
21A12	J	18.04	5001444	2189208	152849	7343501	26	60.97	781277	4	1.36	7014	7084	12	57	1
21B12	R	31.18	4800570	1476001	842995	7119566	21	60.92	1213867	2	1.97	6812	6877	13	51	1
41A11	R	9.90	5220887	1302422	777208	7300517	32	60.9	595573	3	1.92	7044	7106	9	52	1
42B12	R	18.93	6373321	1345053	291154	8009528	19	60.96	948744	3	1.25	7760	7834	9	64	1
43A11	R	51.45	5121469	1483587	289571	6894627	22	60.94	948690	2	0.58	6649	6712	9	53	1
43B11	R	13.52	5214944	1054206	1554476	7823626	35	60.93	507475	4	0.78	7409	7477	12	55	1
43B12	R	25.06	5661339	1028487	286644	6976470	18	60.91	788179	2	0.57	6677	6745	12	55	1
51A11	J	13.20	5109458	1907184	464118	7480760	34	60.96	330366	7	2.41	7124	7185	9	51	1
51B21	J	43.01	5104438	2120438	249269	7474145	29	60.97	447850	5	1.61	7132	7189	6	50	1
63A21	Q	63.11	5014513	889681	486121	6390315	30	61.04	756306	4	0.63	6133	6197	9	54	1
71A12	R	24.11	5575676	1319423	244541	7139640	45	60.89	1085670	3	1.12	6821	6884	9	53	1
73A11	R	9.54	5442738	901817	847493	7192048	54	60.86	597409	4	1.39	6922	6984	9	52	1
73B11	N	43.23	4841288	1224375	977225	7042888	31	60.97	488689	4	1.14	6729	6786	5	51	1
73B12	N	10.08	5780128	1217072	974008	7971208	51	60.96	441934	7	2.01	7660	7736	5	70	1
81B22	R	17.23	4754384	1613829	737039	7105252	48	60.91	701539	4	1.13	6784	6846	9	52	1
83A12	R	16.83	5008294	1347951	758680	7114925	45	60.89	728845	2	1.13	6809	6872	9	53	1
93B11	R	76.82	4891302	1662311	424557	6978170	42	60.9	963099	2	1.15	6723	6782	6	52	1
JHI10	K	25.66	4871811	2330406	704436	7906653	54	60.83	4156248	1	1.01	7532	7596	9	54	1
JHI1084	E	36.97	5179994	2441110	719495	8340599	134	60.39	1173206	3	3.12	8091	8158	8	58	1
JHI1093	L	42.77	5112396	1505388	732854	7350638	34	60.73	531759	4	3.54	7094	7163	12	56	1
JHI1096	L	28.71	5116044	1515012	719445	7350501	35	60.73	531884	5	2.72	7094	7163	12	56	1
JHI1236	C	11.78	4973142	1706791	951138	7631071	29	60.77	855290	4	0.79	7268	7333	9	55	1
JHI1238	C	13.97	5101750	1364116	1084653	7550519	38	60.8	652897	4	1.59	7141	7207	9	56	1
JHI1253	C	45.67	5067405	2282423	743885	8093713	16	60.65	1742435	2	2.47	7705	7769	9	54	1
JHI1259	C	40.66	5159559	2757774	303515	8220848	54	60.56	1983245	2	1.95	7897	7966	9	59	1
JHI1266	C	30.31	5297521	1218121	1338936	7854578	66	60.71	654412	5	3.31	7465	7537	12	59	1
JHI13	B	32.72	4969512	2410484	538525	7918521	57	60.8	720061	2	1.52	7530	7595	9	55	1
JHI1415	B	37.18	5231856	2531624	289035	8052515	16	60.88	980925	2	0.25	7562	7626	9	54	1
JHI1422	C	19.69	5402298	1591994	677835	7672127	30	60.75	655785	4	2.61	7291	7356	9	55	1
JHI1438	B	26.99	5180470	2029478	872071	8082019	84	60.8	477366	6	1.98	7635	7704	12	56	1
JHI1587	B	14.87	5390622	2176075	209532	7776229	29	60.83	423387	7	0	7445	7510	9	55	1
JHI1592	C	21.05	5132897	2569533	157570	7860000	18	60.66	1782404	2	1.27	7485	7550	9	55	1
JHI1600	B	49.80	4900377	2344482	239187	7484046	12	60.93	1513495	2	0.53	7097	7161	9	54	1
JHI24	C	57.15	5164943	2634349	10239	7809531	12	60.69	2321390	2	1.02	7445	7508	9	53	1

Table S4.4. (Continuation)

		Jigome					Quast					Prokka				
Strain	**Genospecies**	**Median coverage**	**Chromosomal length (bp)**	**Plasmid length (bp)**	**Fragment length (bp)**	**Total length (bp)**	**Contigs**	**GC (%)**	**N50**	**L50**	**N's per 100 kbp**	**CDS**	**Gene**	**rRNA**	**tRNA**	**tmRNA**
JHI2442	E	46.53	4786659	2363776	504897	7655332	29	60.6	962391	3	1.83	7231	7294	6	56	1
JHI2449	O	8.33	4952563	1601255	731983	7285801	100	60.81	365547	6	1.65	6964	7025	5	55	1
JHI2450	K	14.43	4869193	1503785	632264	7005242	30	60.93	934214	3	0.57	6671	6729	6	51	1
JHI2451	K	9.35	4480818	2414357	494211	7389386	38	60.85	650685	5	0.54	7010	7067	6	50	1
JHI370	C	21.46	5163256	2517448	32388	7713092	19	60.76	1067459	3	0.78	7331	7395	9	54	1
JHI387	C	40.42	5173307	2299691	256187	7729185	14	60.76	1783577	2	1.04	7348	7412	9	54	1
JHI388	C	21.41	5101565	2266703	293903	7662171	22	60.8	1782547	2	1.31	7283	7347	9	54	1
JHI42	C	18.11	5053470	1451100	1454073	7958643	49	60.61	488844	5	3.02	7576	7640	9	54	1
JHI535	B	32.34	5197764	2175429	48258	7421451	14	60.96	1512289	2	0.27	7008	7072	9	54	1
JHI54	K	17.63	5924925	1222193	1182720	8329838	48	60.62	703922	4	2.16	7868	7929	8	52	1
JHI585	B	19.89	5044283	2279184	313657	7637124	18	60.97	653410	4	0.26	7181	7245	9	54	1
JHI782	C	41.74	4900278	2193935	856598	7950811	21	60.66	2444018	2	1.26	7565	7628	9	53	1
JHI783	C	42.00	4900100	1813685	1240916	7954701	19	60.66	1637483	2	2.51	7563	7626	9	53	1
JHI787	C	23.80	5164721	2106816	542689	7814226	21	60.69	1784913	2	0.51	7456	7519	9	53	1
JHI788	C	37.73	5148615	1763034	1229856	8141505	20	60.74	764038	3	1.47	7739	7807	9	58	1
JHI925	C	18.72	5395528	1474435	1784618	8654581	187	60.51	380490	6	8.1	8277	8350	12	60	1
JHI944	E	42.58	5355529	2129002	277868	7762399	66	60.53	1008232	2	1.55	7463	7529	8	57	1
JHI953	C	19.78	5014054	2061733	638155	7713942	35	60.73	727827	4	0.78	7432	7497	9	55	1
JHI960	B	34.02	5484702	2164971	267337	7917010	76	60.83	885529	3	1.02	7529	7595	9	56	1
JHI963	B	23.64	5433173	1706340	651570	7791083	72	60.88	771138	4	1.28	7406	7471	9	55	1
JHI973	B	18.10	5627923	1725861	135856	7489640	10	60.91	978130	2	1.34	7049	7112	9	53	1
JHI974	B	17.66	5772418	1288540	420020	7480978	17	60.91	661953	4	1.34	7050	7113	9	53	1
JHI979	C	18.20	5141588	2155144	815690	8112422	30	60.73	807509	3	1.48	7733	7800	12	54	1
JHI985	C	27.75	5166652	2173511	506022	7846185	47	60.74	598168	4	1.27	7527	7590	8	54	1
VFCR2A2	Q	27.59	5391629	1627999	818925	7838553	74	60.78	282455	8	1.28	7518	7603	14	70	1
VFF1R1A2	R	16.15	5166465	1096893	1253370	7516728	79	60.73	261996	10	1.33	7260	7325	12	52	1
VFF1R2A1	J	53.33	4980807	2093921	261718	7336446	33	60.97	413341	5	1.09	7006	7076	12	57	1
VFF1R2B1	J	68.23	4994571	1045263	1201901	7241735	108	61.04	169806	14	0	6895	6950	3	51	1
VFF2R2A1	Q	23.14	4983995	1667056	613059	7264110	136	60.8	200446	12	0.55	7005	7068	9	53	1
VFHR1A2	R	40.19	5092094	992584	610256	6694934	58	60.9	243701	9	0.6	6443	6500	5	51	1
VFSR2A2	R	31.15	4815620	1912220	291034	7018874	32	60.93	523693	5	0.29	6771	6833	9	52	1
VFSR2B1	R	58.72	5127325	921783	1041287	7090395	86	60.87	285275	8	0.85	6783	6848	11	53	1

Table S4.5 Pairwise ANI values among genomes.

Query	3841	SM38	JHI960	JHI963	JHI1587	JHI1600	JHI13	JHI1438	JHI535	JHI585
3841	100	99.2934	99.8792	99.7566	98.8883	98.8582	98.7379	98.6236	98.7971	98.7793
SM38	99.3062	100	99.2607	99.3197	98.7846	98.7821	98.6608	98.7693	98.7866	98.7788
JHI960	99.8883	99.2259	100	99.8263	98.9256	98.8403	98.6417	98.5627	98.8064	98.7915
JHI963	99.8115	99.276	99.8418	100	98.9825	98.8326	98.5947	98.5415	98.8061	98.8136
JHI1587	98.868	98.8169	98.9889	99.0176	100	99.4411	98.6915	98.5158	98.8576	98.8519
JHI1600	98.8258	98.7693	98.8507	98.8465	99.4296	100	98.6455	98.4921	98.8572	98.8447
JHI13	98.6828	98.6018	98.5856	98.5645	98.6678	98.7001	100	98.8388	98.6049	98.6049
JHI1438	98.5872	98.732	98.5341	98.5465	98.531	98.4741	98.8826	100	98.802	98.7822
JHI535	98.8404	98.7857	98.811	98.8375	98.8458	98.8384	98.618	98.7862	100	99.9973
JHI585	98.8084	98.7905	98.8302	98.848	98.8532	98.8309	98.6374	98.8031	99.998	100
JHI973	99.0458	98.8775	99.021	99.0225	98.9329	98.8922	98.7795	98.8206	98.9412	98.9275
JHI974	99.0275	98.8653	99.0121	99.0192	98.9474	98.8934	98.8001	98.8006	98.9324	98.9166
JHI1415	98.8115	98.7973	98.8409	98.8507	98.8059	98.7642	98.5937	98.7861	99.9421	99.9504
SM3	98.8559	99.0437	98.8606	98.771	98.6496	98.6094	98.56	98.5212	98.5773	98.5837
VF39	98.7069	98.5642	98.7389	98.7862	98.9566	99.0619	98.668	98.449	98.7325	98.7332
L361	94.4697	94.4845	94.4936	94.5381	94.4795	94.5091	94.4878	94.4349	94.5275	94.5204
VFF1R2B1	95.9368	95.9678	95.9687	95.9841	96.0084	95.999	96.0409	95.9041	96.0703	96.069
21A12	96.0502	96.0947	96.0726	96.0584	96.0139	96.0875	96.0054	95.9377	96.1382	96.1616
VFF1R2A1	96.0426	96.122	96.0735	96.0791	96.0185	96.1107	96.0125	95.9658	96.1315	96.1394
WSM1455	95.9622	95.9975	96.0067	96.0094	96.0562	96.1255	95.97	95.9139	96.021	95.9999
WSM1481	96.0003	96.0207	95.9502	95.9799	96.0359	96.0998	95.9621	95.9114	96.0625	96.0823
WSM78	93.6954	93.8265	93.7025	93.7124	93.6429	93.6723	93.7821	93.7437	93.8128	93.7941
SM152B	93.8083	93.9273	93.7984	93.8595	93.8554	93.7906	93.9322	93.8547	93.8511	93.8629
CC275e	93.8487	93.9103	93.8308	93.8024	93.8057	93.823	93.8631	93.8737	93.8858	93.9099
WSM1325	94.0075	94.0972	93.9678	94.008	93.9461	94.0177	94.0472	93.9328	94.1423	94.1238
SRDI943	94.1756	94.1537	94.1261	94.1633	94.0335	94.1465	94.1718	94.0853	94.2654	94.1865
JHLM12A2	94.6315	94.6736	94.5891	94.6109	94.6836	94.5855	94.6266	94.5916	94.6708	94.651
Vaf12	95.0401	95.079	95.0861	95.0679	95.093	95.1435	95.1718	95.1326	95.1745	95.2005
JHI2451	95.435	95.554	95.4722	95.48	95.5069	95.575	95.622	95.6137	95.6406	95.6737
JHI10	95.627	95.6113	95.6571	95.6566	95.62	95.7085	95.9467	95.8644	95.773	95.7839
JHI54	95.3261	95.4975	95.391	95.3741	95.4844	95.5406	95.4549	95.5238	95.5894	95.5737
FA23	95.6056	95.5773	95.5174	95.53	95.5842	95.5682	95.634	95.6357	95.6709	95.6223
JHI2450	95.5473	95.612	95.5589	95.5635	95.599	95.5903	95.7215	95.6697	95.744	95.7603
Vaf-108	94.1659	94.2824	94.096	94.1304	94.1807	94.2476	94.1718	94.2076	94.3117	94.3218
JHI2449	94.4374	94.4062	94.5047	94.5148	94.5227	94.5134	94.5157	94.4902	94.4656	94.4468
UPM1131	94.3985	94.3817	94.3691	94.4309	94.3393	94.3954	94.4882	94.4771	94.5091	94.4629
CCBAU10279	94.2296	94.3021	94.2638	94.2969	94.215	94.2323	94.2868	94.2482	94.3236	94.2921
WSM1689	94.3771	94.3634	94.3583	94.3561	94.2937	94.2162	94.3133	94.2826	94.297	94.3623
VFF2R2A1	94.2452	94.1683	94.2481	94.2948	94.3051	94.3064	94.2204	94.1959	94.3252	94.3231
SPF2A11	94.3056	94.1563	94.315	94.3239	94.3272	94.3417	94.2775	94.1821	94.2715	94.2735
VFCR2A2	94.1898	94.1532	94.2241	94.2858	94.2891	94.3287	94.2457	94.1596	94.2875	94.2838
248	94.4595	94.3238	94.5218	94.5863	94.4905	94.4847	94.3947	94.3969	94.5181	94.5106
63A21	94.4424	94.3155	94.4835	94.499	94.4501	94.4622	94.3601	94.3383	94.5399	94.5188
GLR2	94.3359	94.2679	94.3095	94.282	94.2807	94.3261	94.2796	94.1844	94.3475	94.3462
CL8	94.207	94.2025	94.2496	94.2762	94.3814	94.3022	94.2247	94.1812	94.2675	94.2553
VFSR2A2	94.4099	94.1556	94.3926	94.4475	94.3655	94.4536	94.2025	94.1274	94.4287	94.4338
SS21	94.1246	94.0878	94.1656	94.1192	94.0916	94.1707	94.1003	94.0652	94.2239	94.2407
41A11	94.3509	94.249	94.4321	94.458	94.3831	94.4303	94.347	94.2623	94.4123	94.36
42B12	94.4412	94.3501	94.4382	94.466	94.4095	94.4904	94.3865	94.3402	94.4416	94.4651
43A11	94.3963	94.2935	94.4049	94.4078	94.3566	94.4232	94.3305	94.2585	94.4034	94.3634
43B11	94.4412	94.1588	94.4663	94.4426	94.4499	94.5285	94.2318	94.1821	94.4857	94.5159
43B12	94.4484	94.2217	94.4889	94.5085	94.4643	94.5242	94.2773	94.1583	94.5213	94.5106
71A12	94.3365	94.3023	94.322	94.3847	94.3927	94.356	94.3209	94.2364	94.3805	94.3316
73A11	94.3656	94.192	94.3201	94.3118	94.3321	94.3895	94.2656	94.1854	94.3244	94.3768
81B22	94.3475	94.2994	94.2937	94.3782	94.3224	94.3509	94.293	94.2074	94.3239	94.3512
83A12	94.3253	94.2896	94.3583	94.3555	94.3464	94.3144	94.2804	94.2093	94.3175	94.2985
VFHR1A2	94.2676	94.187	94.2937	94.3745	94.2756	94.3015	94.2639	94.1948	94.3445	94.3514
VFF1R1A2	94.2371	94.1808	94.2584	94.2624	94.2651	94.4016	94.1571	94.0844	94.3993	94.4215
93B11	94.3251	94.2873	94.3454	94.3289	94.296	94.3095	94.2287	94.1904	94.3791	94.34
111A12	94.4136	94.2787	94.4033	94.4097	94.4031	94.3881	94.296	94.2756	94.4253	94.4212
VFSR2B1	94.3831	94.2984	94.3942	94.383	94.3483	94.383	94.2693	94.2679	94.4135	94.4048
SPF4F7	94.426	94.1977	94.3906	94.418	94.3484	94.4111	94.2355	94.1365	94.406	94.3975
FB206	94.2887	94.1684	94.3138	94.3799	94.3229	94.2848	94.1135	94.0608	94.3929	94.3902
SL16	94.2035	94.1677	94.199	94.1854	94.254	94.3233	94.171	94.0702	94.3321	94.3147
HUTR05	94.2258	94.2174	94.1772	94.2334	94.193	94.222	94.2184	94.1744	94.2342	94.2353
21B12	94.3283	94.3106	94.3911	94.3949	94.3205	94.3691	94.2504	94.2073	94.3686	94.3813
WYCCWR11290	93.8372	93.7589	93.7906	93.8102	93.7649	93.6991	93.8286	93.8239	93.8544	93.8362
WYCCWR11317	93.7831	93.9116	93.8202	93.8525	93.8332	93.8706	93.8172	93.8742	93.9709	93.9281
CCBAU11080	93.7498	93.6946	93.7764	93.8031	93.7702	93.7686	93.7691	93.7919	93.7858	93.7951
CCBAU03386	93.8306	93.8592	93.8125	93.7681	93.773	93.7669	93.7875	93.7325	93.845	93.8359

Table S4.5. (Continuation)

Query	3841	SM38	JHI960	JHI963	JHI1587	JHI1600	JHI13	JHI1438	JHI535	JHI585
WYCCWR10014	94.0716	94.0858	94.0173	94.0764	94.032	94.0364	94.1093	94.0817	94.1605	94.1153
JKLM13E	94.5559	94.7112	94.5936	94.5648	94.6051	94.5998	94.6311	94.6528	94.6821	94.6993
Tri-43	94.1864	94.2388	94.2048	94.2377	94.2418	94.2063	94.1767	94.1824	94.2291	94.2398
73B11	94.5021	94.4286	94.5221	94.497	94.4833	94.4658	94.4948	94.4106	94.556	94.5149
73B12	94.5674	94.4963	94.57	94.583	94.509	94.5184	94.5261	94.4633	94.5717	94.5752
RSP1F2	94.3728	94.328	94.3778	94.3794	94.3935	94.4337	94.5536	94.4365	94.4664	94.4165
Vaf10	94.1601	94.1348	94.0965	94.2189	94.2189	94.2486	94.1844	94.224	94.3474	94.3283
PEPV16	94.291	94.2775	94.3235	94.3785	94.4026	94.3765	94.3749	94.3631	94.4339	94.428
TOM	94.4416	94.4671	94.3986	94.4268	94.4343	94.3986	94.349	94.3645	94.5041	94.4936
121B21	94.3845	94.3804	94.4136	94.4067	94.4094	94.4276	94.3797	94.3136	94.5289	94.5455
RSP1F10	94.4288	94.3361	94.3786	94.4286	94.402	94.4264	94.6818	94.6348	94.4838	94.4989
RSP1A1	94.4289	94.3487	94.349	94.4278	94.419	94.4566	94.739	94.6667	94.5481	94.5179
Norway	93.3019	93.3565	93.2522	93.2895	93.2637	93.2894	93.3504	93.3241	93.3977	93.3719
CC278f	93.0666	93.1551	93.0246	93.1019	93.1038	93.1113	93.0394	93.0556	93.1926	93.1863
SM78	93.2799	93.3669	93.2786	93.3127	93.245	93.2771	93.2282	93.2507	93.3427	93.3683
SM51	93.1525	93.3071	93.1978	93.16	93.1723	93.1815	93.2011	93.0928	93.2763	93.2504
SM72	93.2731	93.4514	93.3342	93.3272	93.293	93.2893	93.2956	93.2996	93.3413	93.3467
Vaf-46	93.0496	93.0601	92.9669	93.0769	93.093	93.0989	93.1691	93.1531	93.1954	93.1763
JHI1093	93.1696	93.1554	93.2048	93.2314	93.1934	93.2196	93.3415	93.3234	93.2457	93.2498
JHI1096	93.2098	93.1885	93.2004	93.2383	93.1776	93.2005	93.3252	93.3652	93.2233	93.2642
GPTR08	92.653	92.6467	92.5504	92.5245	92.5469	92.5914	92.5793	92.6326	92.6498	92.6404
GPTR02	92.6758	92.8216	92.7468	92.7476	92.6559	92.732	92.694	92.7045	92.7633	92.7396
SRDI565	92.6267	92.7385	92.649	92.717	92.691	92.6747	92.6982	92.7422	92.7032	92.7347
Ps8	93.7044	93.6331	93.6852	93.6733	93.6592	93.7336	93.6867	93.7558	93.7604	93.7492
JHI1236	93.7817	93.7049	93.7362	93.7534	93.694	93.7555	93.7397	93.7548	93.8079	93.8072
JHI953	93.5871	93.6403	93.6377	93.6283	93.599	93.686	93.6008	93.7	93.805	93.7948
SM147A	93.5749	93.7293	93.6027	93.6797	93.6176	93.6489	93.5222	93.606	93.7322	93.6709
JHI1238	93.6917	93.6476	93.6393	93.6601	93.6328	93.6973	93.6305	93.6474	93.7408	93.7306
UPM1133	93.55	93.5489	93.5219	93.537	93.624	93.616	93.623	93.7132	93.6375	93.6595
JHI1592	93.6522	93.603	93.6377	93.6093	93.6301	93.6421	93.6417	93.6134	93.7664	93.7915
SM41	93.6283	93.6714	93.6392	93.6006	93.5968	93.6137	93.5779	93.5681	93.7018	93.653
JHI1253	93.7443	93.5585	93.6434	93.6122	93.6806	93.6125	93.7544	93.7851	93.6978	93.7475
JHI370	93.6337	93.6301	93.6096	93.5509	93.5923	93.6211	93.5842	93.6489	93.7727	93.7111
JHI387	93.6508	93.6696	93.5959	93.6138	93.6361	93.6497	93.6327	93.6732	93.7675	93.7575
JHI388	93.6735	93.6236	93.6105	93.6217	93.6685	93.6431	93.6345	93.7058	93.758	93.7705
JHI788	93.7287	93.6832	93.7571	93.6867	93.6405	93.7092	93.7242	93.7603	93.8186	93.8238
JHI985	93.685	93.6818	93.6098	93.6205	93.6098	93.6682	93.6015	93.6729	93.6999	93.7027
GB30	93.5828	93.616	93.667	93.6946	93.6322	93.7092	93.7031	93.7357	93.7607	93.7127
JHI782	93.6462	93.6412	93.5907	93.6453	93.6232	93.5975	93.7015	93.7648	93.7758	93.7568
JHI783	93.6545	93.6557	93.6467	93.6722	93.668	93.5699	93.7463	93.7526	93.7238	93.6911
SM170C	93.6046	93.6819	93.5241	93.5791	93.6348	93.6046	93.5425	93.5546	93.697	93.6599
JHI42	93.6621	93.5974	93.6582	93.6697	93.6635	93.6961	93.6938	93.7382	93.8153	93.7867
JHI979	93.6632	93.6183	93.6379	93.5786	93.6579	93.6369	93.6222	93.6543	93.792	93.77
JHI1259	93.6936	93.641	93.6711	93.6895	93.6643	93.6373	93.7098	93.775	93.7403	93.7447
TA1	93.5688	93.6853	93.6125	93.6257	93.6383	93.6424	93.5999	93.6052	93.6512	93.6532
JHI24	93.622	93.6689	93.6366	93.6625	93.6689	93.6805	93.661	93.7164	93.7527	93.7031
JHI787	93.603	93.5997	93.6161	93.6212	93.6837	93.6803	93.6572	93.6902	93.771	93.7248
RCAM1026	93.7112	93.7212	93.7173	93.7401	93.6955	93.7193	93.7212	93.8062	93.7614	93.7492
Vh3	93.7624	93.7727	93.712	93.7437	93.7222	93.758	93.7859	93.8632	93.8319	93.816
JHI925	93.9713	93.6096	94.0355	94.0327	94.0284	93.9209	93.6722	93.71	93.9834	94.0067
Vc2	93.6699	93.6935	93.6236	93.6638	93.6951	93.6663	93.8136	93.7593	93.8496	93.8027
JHI1422	93.6863	93.6982	93.6505	93.6666	93.6899	93.7382	93.6735	93.774	93.7815	93.8222
JHI1266	93.7379	93.5609	93.6796	93.6491	93.6813	93.6618	93.779	93.8415	93.7694	93.7687
51A11	95.987	96.0306	95.9921	95.9885	95.9499	96.019	95.9038	95.877	96.0478	96.0709
51B21	95.9554	96.0294	95.9783	95.9634	95.9532	96.0038	95.9326	95.8566	96.0494	96.0569
128C53	92.7534	92.6904	92.7142	92.8072	92.7822	92.833	92.9649	92.9604	92.8437	92.8298
4292	92.7608	92.8131	92.7891	92.8283	92.8007	92.8097	92.8127	92.7694	92.8982	92.8462
CC283bq	92.7574	92.8016	92.6318	92.6321	92.6688	92.7586	92.6844	92.6798	92.8738	92.8526
USDA2370	92.8175	92.6139	92.8481	92.7855	92.8009	92.7337	93.027	93.0455	92.8076	92.806
JHI2442	92.84	92.7682	92.7717	92.8565	92.8234	92.807	92.9978	93.0219	92.9067	92.8796
JHI1084	92.8647	92.7563	92.8259	92.8779	92.8709	92.8011	92.8452	92.8695	92.8946	92.8455
JHI944	93.1068	92.7254	93.2175	93.1325	93.0317	92.9766	92.8717	92.8433	93.2439	93.2037
OV152	92.8768	92.8745	92.8729	92.8895	92.8711	92.7953	92.8125	92.7923	92.9283	92.9483
UPM1137	92.7261	92.8513	92.8118	92.8345	92.8442	92.8789	92.8827	92.8641	92.9457	92.906
SM149A	92.9712	92.8752	92.7318	92.857	92.832	92.8471	92.8393	92.7217	92.9008	92.8484
Rt24.2	92.7689	92.8876	92.7656	92.8164	92.8876	92.8263	92.7785	92.7792	92.8838	92.8722
CCBAU23252	91.0425	91.0307	91.0684	91.1378	91.145	91.1073	91.0881	91.0205	91.1582	91.1616
JHI536	91.0841	91.1326	91.1754	91.172	91.1923	91.2076	91.3165	91.2911	91.1835	91.248
C15	91.092	91.179	91.1272	91.0635	91.1448	91.1625	91.0864	91.1408	91.215	91.2436
CF307	91.1316	91.2006	91.125	91.1616	91.1914	91.2102	91.1684	91.141	91.2454	91.2277

Table S4.5. (Continuation)

Query	JHI973	JHI974	JHI1415	SM3	VF39	L361	VFF1R2B1	21A12	VFF1R2A1	WSM1455
3841	98.9642	99.007	98.7811	98.8402	98.64	94.4534	96.0101	96.0706	96.0485	95.9464
SM38	98.8355	98.8184	98.7327	99.0708	98.5541	94.4665	95.9554	96.0434	96.0587	96.0106
JHI960	98.9901	98.9826	98.7832	98.8729	98.7199	94.4918	95.9692	96.0316	96.0396	95.9538
JHI963	98.9846	98.998	98.7738	98.8865	98.8229	94.492	96.0096	95.9997	96.01	95.9745
JHI1587	98.9093	98.883	98.7867	98.6706	98.958	94.4822	96.0112	96.032	96.0473	96.0195
JHI1600	98.888	98.8791	98.793	98.5979	99.086	94.4188	96.0096	96.0469	96.0542	96.0671
JHI13	98.7706	98.7604	98.5657	98.5415	98.601	94.4157	95.988	95.9743	95.98	95.9162
JHI1438	98.7683	98.8113	98.7495	98.5683	98.3675	94.3667	95.9416	95.9145	95.9088	95.8528
JHI535	98.9389	98.954	99.9451	98.6121	98.7089	94.5135	96.0564	96.1241	96.115	96.0488
JHI585	98.9366	98.9491	99.9378	98.6022	98.7572	94.4907	96.0507	96.1116	96.141	96.0702
JHI973	100	99.9959	98.9803	98.7871	98.8731	94.4807	96.0874	96.0368	96.0457	96.021
JHI974	99.9975	100	98.9416	98.8245	98.8642	94.5061	96.0921	96.0408	96.0447	95.9993
JHI1415	98.968	98.9566	100	98.5831	98.6695	94.4868	96.0613	96.1142	96.1213	96.0291
SM3	98.7828	98.8054	98.5293	100	98.554	94.4226	95.9126	95.9436	95.9312	95.9088
VF39	98.8621	98.8692	98.6449	98.5726	100	94.3762	95.9639	95.9893	95.9482	96.101
L361	94.5579	94.5439	94.4779	94.4469	94.4953	100	94.5422	94.5604	94.5537	94.5467
VFF1R2B1	96.0574	96.0507	96.0369	95.9532	95.9886	94.5449	100	98.424	98.4285	98.4434
21A12	96.0961	96.1007	96.0823	96.0649	96.0335	94.5581	98.4184	100	99.9945	98.6979
VFF1R2A1	96.0851	96.1029	96.0778	96.0713	96.0406	94.5783	98.4379	99.9976	100	98.6899
WSM1455	96.0265	96.0297	95.9899	95.9859	96.0814	94.5124	98.4031	98.6685	98.6742	100
WSM1481	95.9987	95.9802	96.0139	95.918	96.0671	94.447	98.3565	98.5091	98.5035	98.6876
WSM78	93.79	93.7196	93.714	93.7761	93.7074	93.1641	93.6876	93.6712	93.6574	93.588
SM152B	93.8906	93.9175	93.8486	93.9203	93.8577	93.1825	93.7592	93.7905	93.8078	93.6873
CC275e	93.903	93.9058	93.8815	93.9154	93.8682	93.3332	93.809	93.7401	93.7143	93.7244
WSM1325	94.0394	94.0203	94.091	94.109	94.0021	93.4908	94.1209	94.0727	94.0795	94.0266
SRDI943	94.1139	94.1607	94.1835	94.1584	94.0623	93.6201	94.1739	94.2687	94.2448	94.0772
JHLM12A2	94.6078	94.6271	94.5897	94.5324	94.5548	94.2184	94.6886	94.7313	94.7293	94.7117
Vaf12	95.1434	95.1723	95.1111	95.0209	95.0421	94.8562	95.1458	95.229	95.2073	95.1272
JHI2451	95.5779	95.63	95.6789	95.4797	95.5068	94.8236	95.5584	95.5606	95.5808	95.5036
JHI10	95.7307	95.7149	95.7241	95.5907	95.5222	94.7581	95.574	95.6706	95.6766	95.5794
JHI54	95.5442	95.5177	95.5344	95.4406	95.3958	94.7031	95.5214	95.5017	95.4832	95.4498
FA23	95.6001	95.6335	95.6488	95.5911	95.4777	94.8509	95.5265	95.5856	95.5547	95.6063
JHI2450	95.6895	95.7322	95.6857	95.5807	95.5979	94.8141	95.6955	95.6841	95.6758	95.6132
Vaf-108	94.3585	94.3473	94.289	94.1822	94.0525	95.9637	94.3386	94.3508	94.3542	94.2217
JHI2449	94.426	94.4416	94.3335	94.3832	94.3103	96.6847	94.4762	94.5218	94.5282	94.4452
UPM1131	94.4483	94.438	94.3786	94.3561	94.3755	96.2511	94.3957	94.505	94.5065	94.421
CCBAU10279	94.3425	94.331	94.3146	94.323	94.2454	95.7039	94.3743	94.3321	94.3304	94.2724
WSM1689	94.3767	94.3215	94.3309	94.393	94.297	94.8534	94.463	94.4152	94.3936	94.3455
VFF2R2A1	94.3114	94.2899	94.1908	94.1909	94.192	95.2345	94.2803	94.2703	94.2729	94.1952
SPF2A11	94.3325	94.2768	94.2333	94.2145	94.3155	95.2465	94.3012	94.3803	94.3388	94.1937
VFCR2A2	94.3922	94.3662	94.2698	94.1169	94.2425	95.331	94.3211	94.3245	94.299	94.2202
248	94.4482	94.4348	94.5053	94.2752	94.3904	95.3326	94.4468	94.5337	94.538	94.4633
63A21	94.4674	94.4466	94.4884	94.3696	94.48	94.9956	94.5396	94.4594	94.4614	94.5424
GLR2	94.2983	94.3133	94.3105	94.1738	94.3126	95.3689	94.4525	94.369	94.3443	94.3745
CL8	94.3144	94.2739	94.2479	94.2412	94.235	95.4411	94.3062	94.3112	94.3308	94.2781
VFSR2A2	94.445	94.4324	94.4449	94.1947	94.3111	95.4583	94.2791	94.369	94.3446	94.2294
SS21	94.1843	94.1679	94.1458	94.0882	94.125	95.416	94.3201	94.3372	94.3226	94.1936
41A11	94.436	94.3925	94.3394	94.2274	94.3701	95.4784	94.3219	94.3496	94.3537	94.2747
42B12	94.4389	94.4073	94.444	94.311	94.3457	95.4975	94.4079	94.4152	94.4244	94.3471
43A11	94.3841	94.3241	94.3606	94.2234	94.3271	95.415	94.3409	94.3649	94.3175	94.3159
43B11	94.555	94.4944	94.4456	94.1422	94.4047	95.3951	94.3449	94.3623	94.3532	94.2134
43B12	94.4899	94.4733	94.4667	94.1868	94.4243	95.4341	94.3272	94.3595	94.3227	94.2012
71A12	94.3752	94.3496	94.2668	94.2851	94.3256	95.4661	94.3188	94.3492	94.3352	94.3096
73A11	94.4204	94.3811	94.2912	94.1729	94.24	95.3986	94.3082	94.3611	94.3405	94.2854
81B22	94.3384	94.3389	94.2903	94.2975	94.3131	95.486	94.3224	94.3303	94.3015	94.2652
83A12	94.3426	94.3405	94.2596	94.211	94.2427	95.4205	94.2995	94.3151	94.3005	94.2868
VFHR1A2	94.3132	94.3262	94.2842	94.1796	94.2478	95.4373	94.3313	94.3779	94.3937	94.26
VFF1R1A2	94.3901	94.4127	94.3206	94.1082	94.2822	95.3302	94.2723	94.3515	94.3387	94.2554
93B11	94.3273	94.3307	94.2687	94.2035	94.2516	95.3482	94.2229	94.3627	94.3562	94.2598
111A12	94.3761	94.358	94.4022	94.2639	94.3066	95.4004	94.3389	94.3963	94.3993	94.3323
VFSR2B1	94.3948	94.3773	94.3284	94.2257	94.3258	95.429	94.3218	94.3809	94.3644	94.2814
SPF4F7	94.3904	94.3882	94.3379	94.1448	94.3206	95.3872	94.312	94.3263	94.3399	94.202
FB206	94.3389	94.3513	94.3288	94.1444	94.3233	95.3639	94.2405	94.3007	94.3223	94.1824
SL16	94.2678	94.254	94.2602	94.1339	94.2242	95.4103	94.3013	94.3036	94.3058	94.2475
HUTR05	94.2078	94.2373	94.1295	94.1958	94.1734	95.3779	94.2802	94.313	94.3113	94.2025
21B12	94.4442	94.4267	94.3324	94.2422	94.3178	95.372	94.3024	94.3915	94.3522	94.231
WYCCWR11290	93.7893	93.7952	93.7828	93.7555	93.7809	93.7346	93.769	93.7132	93.7058	93.7023
WYCCWR11317	93.8707	93.8412	93.8608	93.9434	93.8515	93.8429	93.8614	93.8857	93.8283	93.8516
CCBAU11080	93.784	93.779	93.785	93.7451	93.7361	93.558	93.7845	93.8171	93.7988	93.7737
CCBAU03386	93.8438	93.847	93.7722	93.8978	93.7465	93.7001	93.8283	93.8243	93.81	93.7766

Table S4.5. (Continuation)

Query	JHI973	JHI974	JHI1415	SM3	VF39	L361	VFF1R2B1	21A12	VFF1R2A1	WSM1455
WYCCWR10014	94.0906	94.1262	94.1138	94.0551	94.061	93.7453	94.1207	94.0818	94.0649	94.0792
JKLM13E	94.6691	94.6789	94.6617	94.587	94.5489	94.2087	94.6642	94.7434	94.7691	94.7045
Tri-43	94.281	94.2186	94.2407	94.1505	94.1562	94.2076	94.2248	94.296	94.2407	94.2034
73B11	94.4844	94.4577	94.4446	94.4062	94.4356	95.7482	94.5221	94.5334	94.5686	94.4334
73B12	94.5302	94.5005	94.5758	94.5226	94.5376	95.8177	94.5279	94.5688	94.5878	94.4996
RSP1F2	94.4149	94.3813	94.3856	94.2781	94.3382	95.6519	94.4185	94.4821	94.4997	94.407
Vaf10	94.2649	94.277	94.2825	94.1774	94.2422	95.3908	94.2364	94.2941	94.2949	94.2197
PEPV16	94.3851	94.3333	94.352	94.2839	94.3348	95.6764	94.3689	94.344	94.3494	94.307
TOM	94.4497	94.4388	94.4858	94.4169	94.3805	95.6273	94.4711	94.4421	94.4553	94.4022
121B21	94.5025	94.48	94.4414	94.3126	94.3461	95.6871	94.467	94.5556	94.5262	94.4363
RSP1F10	94.4705	94.4303	94.4712	94.368	94.3314	95.7245	94.4772	94.4413	94.4381	94.3818
RSP1A1	94.5382	94.5191	94.4502	94.3375	94.3959	95.7101	94.5092	94.5328	94.5141	94.4174
Norway	93.3827	93.3509	93.3463	93.3194	93.2574	93.0291	93.2942	93.2985	93.2945	93.2338
CC278f	93.1876	93.1089	93.1556	93.0833	93.0627	92.7972	93.0967	93.2192	93.1676	93.0452
SM78	93.2765	93.3516	93.2762	93.2765	93.2712	92.8251	93.2383	93.2164	93.2174	93.1595
SM51	93.2256	93.2384	93.1516	93.2445	93.182	92.8567	93.1421	93.1878	93.1835	93.0365
SM72	93.3132	93.3327	93.2673	93.3334	93.2795	92.8799	93.2584	93.2033	93.2469	93.1533
Vaf-46	93.2055	93.2267	93.0656	92.9773	93.0986	93.1045	93.1277	93.2278	93.1995	92.9858
JHI1093	93.2627	93.2205	93.1759	93.1344	93.1983	93.114	93.1863	93.2097	93.2121	93.1161
JHI1096	93.2654	93.2021	93.1631	93.1299	93.1461	93.1197	93.183	93.2007	93.1884	93.1625
GPTR08	92.6578	92.6139	92.6291	92.6628	92.6552	92.653	92.5759	92.4928	92.4922	92.4307
GPTR02	92.7619	92.7481	92.6708	92.7865	92.6846	92.7714	92.6574	92.7261	92.7486	92.5727
SRDI565	92.7597	92.7753	92.6885	92.7194	92.7307	92.7179	92.6661	92.6087	92.6057	92.5592
Ps8	93.7298	93.7384	93.7271	93.5347	93.6503	93.1707	93.5497	93.5602	93.5695	93.5061
JHI1236	93.7724	93.7931	93.7795	93.621	93.7194	93.0839	93.6236	93.6441	93.6581	93.5277
JHI953	93.6911	93.6982	93.7816	93.5449	93.5799	93.0168	93.5082	93.5995	93.602	93.4992
SM147A	93.684	93.6868	93.6653	93.6553	93.6278	93.0552	93.5194	93.5121	93.5183	93.5162
JHI1238	93.7469	93.7497	93.7518	93.5599	93.6487	93.0652	93.6056	93.5615	93.5937	93.4712
UPM1133	93.6991	93.637	93.6949	93.4609	93.5573	93.1152	93.5807	93.6187	93.6016	93.401
JHI1592	93.7492	93.7166	93.7446	93.616	93.5478	93.044	93.5345	93.611	93.5789	93.51
SM41	93.7299	93.7251	93.6411	93.646	93.5998	93.0425	93.5144	93.6109	93.6493	93.5194
JHI1253	93.7174	93.75	93.7204	93.5736	93.601	93.1137	93.5153	93.5807	93.5397	93.5892
JHI370	93.7323	93.685	93.7453	93.5564	93.5308	93.0604	93.4496	93.5926	93.5518	93.5537
JHI387	93.7196	93.7354	93.7672	93.5902	93.4925	93.0793	93.4353	93.5666	93.5406	93.4846
JHI388	93.7564	93.7143	93.7387	93.6048	93.5833	93.0613	93.4773	93.5736	93.5769	93.5063
JHI788	93.7468	93.732	93.7906	93.6832	93.6862	93.1923	93.6353	93.5993	93.6457	93.5051
JHI985	93.6843	93.6785	93.682	93.5661	93.5598	93.0828	93.4874	93.5759	93.5898	93.5563
GB30	93.7113	93.7557	93.717	93.6219	93.6423	93.1613	93.5941	93.6146	93.6261	93.5548
JHI782	93.7268	93.7657	93.7954	93.6118	93.6253	93.05	93.4867	93.4557	93.4863	93.4645
JHI783	93.7586	93.6987	93.7944	93.5933	93.6374	93.0715	93.4594	93.4875	93.5105	93.4617
SM170C	93.6604	93.6571	93.6115	93.5631	93.5471	93.0124	93.4774	93.4997	93.524	93.4157
JHI42	93.7501	93.7403	93.7061	93.583	93.5923	93.0069	93.521	93.5997	93.5859	93.5148
JHI979	93.7105	93.7246	93.7858	93.4598	93.6474	93.1234	93.5561	93.5646	93.5837	93.4965
JHI1259	93.6875	93.7118	93.7978	93.5085	93.5374	93.0228	93.517	93.5065	93.5197	93.4882
TA1	93.6134	93.6172	93.6331	93.6261	93.5493	93.0849	93.5949	93.5664	93.5806	93.4569
JHI24	93.7038	93.6845	93.6746	93.5335	93.5653	93.0582	93.4673	93.5293	93.5265	93.4572
JHI787	93.6714	93.6667	93.6814	93.5574	93.5875	93.0663	93.4346	93.5486	93.5504	93.5034
RCAM1026	93.7732	93.7361	93.6969	93.6124	93.7676	93.1442	93.6364	93.6633	93.6551	93.5947
Vh3	93.8347	93.835	93.8689	93.6746	93.6912	93.069	93.52	93.5789	93.5955	93.5327
JHI925	94.0061	94.0216	94.0093	93.5819	93.833	93.0638	93.4989	93.5705	93.6065	93.4551
Vc2	93.8496	93.7939	93.864	93.6562	93.6265	92.9974	93.5662	93.5561	93.5343	93.4422
JHI1422	93.9024	93.8652	93.8469	93.6958	93.6783	93.0726	93.5558	93.6173	93.6509	93.524
JHI1266	93.717	93.7361	93.751	93.4813	93.5875	93.0923	93.5858	93.5518	93.5261	93.5601
51A11	96.0335	96.0437	96.0174	95.9777	95.941	94.5137	98.2003	98.354	98.3777	98.2109
51B21	96.0291	96.0738	95.9942	95.9727	95.9223	94.5197	98.2076	98.37	98.3769	98.202
128C53	92.8248	92.8044	92.7352	92.7064	92.7007	92.4693	92.6695	92.7912	92.7783	92.6446
4292	92.8569	92.8649	92.9031	92.7675	92.7508	92.4359	92.7482	92.7839	92.8301	92.6591
CC283bq	92.7927	92.8212	92.7887	92.7579	92.7051	92.3928	92.6515	92.7705	92.7523	92.6748
USDA2370	92.7485	92.7407	92.7417	92.6441	92.7624	92.3356	92.7315	92.7374	92.7842	92.6924
JHI2442	92.8307	92.8153	92.7877	92.7583	92.7543	92.5136	92.761	92.7727	92.8079	92.6889
JHI1084	92.7852	92.7955	92.7916	92.6777	92.7551	92.443	92.7261	92.7114	92.7272	92.6668
JHI944	93.0941	93.1008	93.1281	92.6501	93.1046	92.5043	92.7195	92.8136	92.8388	92.7044
OV152	92.8725	92.8703	92.9094	92.8572	92.799	92.459	92.7457	92.8415	92.8668	92.6916
UPM1137	92.8566	92.8894	92.8502	92.683	92.8347	92.402	92.8134	92.8354	92.8337	92.7413
SM149A	92.7654	92.8278	92.7701	92.8025	92.7596	92.3621	92.7552	92.7962	92.8099	92.6267
Rt24.2	92.8254	92.8574	92.7886	92.8703	92.8961	92.4116	92.7528	92.857	92.8583	92.7242
CCBAU23252	91.148	91.0971	91.0636	91.0551	91.1393	90.8992	91.0983	91.1085	91.0869	91.0246
JHI536	91.2224	91.1659	91.1451	91.1362	91.1569	90.9443	91.0951	91.1345	91.1107	91.1325
C15	91.2309	91.213	91.1752	91.1056	91.1431	90.9219	91.1421	91.1395	91.1593	91.1226
CF307	91.1856	91.2019	91.1522	91.1645	91.2381	90.9711	91.1219	91.1883	91.1486	91.181

Table S4.5. (Continuation)

Query	WSM1481	WSM78	SM152B	CC275e	WSM1325	SRDI943	JHLM12A2	Vaf12	JHI2451
3841	95.9861	93.6931	93.7827	93.8056	94.0457	94.2132	94.5569	95.0923	95.4788
SM38	96.0292	93.7302	93.844	93.9489	94.1185	94.1258	94.6911	95.0715	95.5855
JHI960	95.9261	93.6916	93.7922	93.7681	93.9834	94.1524	94.5649	95.0581	95.4763
JHI963	95.9438	93.663	93.8126	93.8205	94.0117	94.1724	94.5682	95.0379	95.4981
JHI1587	96.1049	93.5718	93.8191	93.8105	94.0233	94.08	94.6787	95.0242	95.5418
JHI1600	96.1333	93.6512	93.7312	93.7778	94.1202	94.1875	94.6875	95.104	95.6098
JHI13	95.8879	93.6899	93.9073	93.8472	94.0179	94.1153	94.6509	95.1102	95.6177
JHI1438	95.9647	93.6726	93.7644	93.8701	94.0195	94.0323	94.6283	95.0821	95.6027
JHI535	96.0314	93.7645	93.8601	93.961	94.0838	94.1643	94.6796	95.1558	95.6359
JHI585	96.0637	93.7477	93.8282	93.9305	94.14	94.2087	94.7058	95.1855	95.6638
JHI973	95.9831	93.6527	93.7991	93.8423	94.0811	94.1144	94.609	95.1196	95.621
JHI974	96.0101	93.6745	93.7744	93.8776	94.0678	94.1698	94.5853	95.0999	95.6594
JHI1415	96.0216	93.7167	93.8274	93.8808	94.1014	94.1703	94.7041	95.098	95.6789
SM3	95.9105	93.6332	93.8537	93.8646	94.0397	94.0768	94.5441	94.9857	95.4877
VF39	96.1185	93.7102	93.8091	93.876	94.044	94.0791	94.6334	95.0144	95.5067
L361	94.5366	93.1516	93.2175	93.3002	93.5588	93.7004	94.2754	94.8108	94.8981
VFF1R2B1	98.3827	93.6993	93.7561	93.8048	94.1613	94.2157	94.6866	95.2007	95.5863
21A12	98.5447	93.6695	93.7541	93.7433	94.1371	94.2874	94.7124	95.2193	95.6885
VFF1R2A1	98.552	93.6241	93.7943	93.7579	94.0462	94.2538	94.7136	95.1951	95.6774
WSM1455	98.711	93.5866	93.6949	93.707	94.1013	94.1598	94.6623	95.1721	95.5475
WSM1481	100	93.5556	93.6945	93.6619	94.0351	94.0991	94.683	95.0363	95.5217
WSM78	93.5614	100	96.2057	96.4587	94.3771	94.3893	93.7109	93.8363	93.9009
SM152B	93.7184	96.2637	100	96.9813	94.4779	94.5379	93.8137	93.836	93.863
CC275e	93.7097	96.4915	97.0141	100	94.7583	94.7666	93.9057	94.0008	94.0834
WSM1325	94.033	94.3996	94.4819	94.6275	100	98.47	94.0331	94.3583	94.4889
SRDI943	94.1057	94.3589	94.425	94.7163	98.5007	100	94.1091	94.4074	94.5261
JHLM12A2	94.6711	93.6861	93.7818	93.8919	94.0127	94.1097	100	95.1373	95.1046
Vaf12	95.0983	93.8661	93.8418	93.9353	94.3509	94.3651	95.1389	100	95.7499
JHI2451	95.5043	93.8839	93.8728	94.0475	94.4182	94.5327	95.0603	95.7565	100
JHI10	95.5653	93.7983	93.8057	94.0341	94.4105	94.4631	94.9853	95.6449	97.4941
JHI54	95.4347	93.7431	93.8347	93.9915	94.3121	94.3495	94.8372	95.6079	97.3388
FA23	95.5756	93.8126	93.8995	94.1775	94.4608	94.5396	95.0272	95.7281	98.4395
JHI2450	95.5539	93.8342	93.8962	94.0608	94.5221	94.6149	95.139	95.7813	97.7623
Vaf-108	94.158	93.0013	93.0734	93.1505	93.4335	93.5313	94.1013	95.1453	94.6279
JHI2449	94.5025	93.1396	93.1987	93.2627	93.5743	93.6247	94.3077	94.7342	94.7188
UPM1131	94.4223	93.1292	93.2021	93.2683	93.5065	93.6018	94.2624	95.03	94.8845
CCBAU10279	94.2751	92.957	93.0532	93.1411	93.3037	93.392	93.9975	94.6974	94.6128
WSM1689	94.2937	92.9758	93.1528	93.2338	93.3934	93.5028	94.0459	94.5371	94.7155
VFF2R2A1	94.2705	92.8691	92.9078	92.9931	93.2502	93.3492	93.9144	94.4627	94.46
SPF2A11	94.2977	92.9631	92.9703	93.0997	93.1888	93.3403	93.847	94.4825	94.4829
VFCR2A2	94.2769	92.8515	92.9372	92.9009	93.2561	93.3307	93.8864	94.4746	94.4376
248	94.5234	93.1094	93.1213	93.2008	93.4718	93.5105	94.0758	94.4694	94.3971
63A21	94.5376	93.0729	93.1355	93.1737	93.4813	93.5978	94.0825	94.5484	94.6184
GLR2	94.37	92.9394	92.9297	93.0765	93.4038	93.4113	93.8971	94.4594	94.4066
CL8	94.2992	92.9949	92.9385	93.1332	93.3257	93.3233	93.9164	94.5558	94.4829
VFSR2A2	94.1893	92.8408	92.9132	93.0217	93.2771	93.2819	93.9339	94.452	94.5573
SS21	94.1598	92.8906	92.9222	93.0056	93.2552	93.3037	93.8666	94.4181	94.5158
41A11	94.2671	93.0289	92.9763	93.101	93.3612	93.4455	93.9337	94.5419	94.53
42B12	94.3636	93.0618	93.0566	93.2304	93.4271	93.4876	94.044	94.5968	94.5715
43A11	94.2888	92.9732	92.9628	93.0496	93.2974	93.3683	93.986	94.5019	94.5488
43B11	94.2329	92.8709	92.9284	93.0596	93.3066	93.4402	93.9082	94.4372	94.5313
43B12	94.2375	92.9487	92.9709	93.0698	93.3082	93.4259	93.9961	94.452	94.5715
71A12	94.3477	92.9578	93.0265	93.1659	93.2909	93.4791	93.9851	94.5656	94.532
73A11	94.2939	92.9566	92.9788	93.0894	93.2605	93.3665	94.0431	94.4956	94.5207
81B22	94.2654	92.9653	93.0258	93.1291	93.3178	93.4542	94.0151	94.5339	94.4948
83A12	94.2588	92.9923	92.9921	93.0805	93.3643	93.4845	93.9803	94.5059	94.5667
VFHR1A2	94.2748	92.9901	92.952	93.0451	93.2838	93.3713	93.8986	94.4594	94.5423
VFF1R1A2	94.1824	92.9369	92.9557	93.0202	93.2933	93.3293	93.934	94.4566	94.4887
93B11	94.2556	92.9148	92.966	93.1565	93.2821	93.3719	94.0103	94.5778	94.5626
111A12	94.2288	92.8852	93.0294	93.1236	93.3542	93.4359	94.077	94.5239	94.5606
VFSR2B1	94.1884	92.8905	93.0206	93.0821	93.2937	93.3908	94.0116	94.4728	94.5108
SPF4F7	94.2342	92.8934	92.9295	93.0738	93.2852	93.3654	93.9425	94.4145	94.5203
FB206	94.1226	92.7797	92.8539	92.9838	93.2961	93.289	93.932	94.4164	94.533
SL16	94.2102	92.9299	92.9222	93.08	93.2144	93.316	93.9416	94.5377	94.4558
HUTR05	94.183	92.9197	92.9739	93.0392	93.2281	93.3123	93.8976	94.5339	94.5087
21B12	94.231	92.9567	92.9722	93.0504	93.3372	93.3738	93.9751	94.4938	94.6157
WYCCWR11290	93.7911	92.8857	92.9582	93.0135	93.1066	93.042	94.1853	94.0108	94.0336
WYCCWR11317	93.8285	93.0769	93.1295	93.0642	93.2566	93.2514	94.2033	94.0124	94.0449
CCBAU11080	93.737	92.9689	92.8694	92.8862	93.0728	93.1338	94.3706	93.8082	93.9593
CCBAU03386	93.8071	92.9427	93.0357	92.9368	93.1064	93.1847	94.0343	93.8316	93.8873

Table S4.5. (Continuation)

Query	WSM1481	WSM78	SM152B	CC275e	WSM1325	SRDI943	JHLM12A2	Vaf12	JHI2451
WYCCWR10014	94.1182	94.2054	94.0627	93.9388	93.6469	93.7547	94.297	94.1999	94.3795
JKLM13E	94.7293	93.6706	93.6442	93.8393	94.0178	94.0697	97.9959	95.0596	95.1406
Tri-43	94.2334	93.4201	93.4745	93.4034	93.6129	93.6401	94.7597	94.4263	94.4113
73B11	94.5258	93.15	93.1978	93.2884	93.4925	93.6093	94.2428	94.7958	94.7761
73B12	94.5608	93.2389	93.2947	93.3188	93.5545	93.6193	94.2542	94.8163	94.8368
RSP1F2	94.3748	93.1402	93.2575	93.199	93.488	93.528	94.2825	94.7895	94.89
Vaf10	94.1698	93.0476	93.0369	93.1051	93.3605	93.4644	94.0956	95.0378	94.6245
PEPV16	94.3255	93.0059	93.0658	93.2088	93.3583	93.4789	94.2084	94.5697	94.7911
TOM	94.4446	93.0756	93.1994	93.2613	93.4905	93.5374	94.1986	94.7095	94.7919
121B21	94.372	93.0637	93.139	93.1935	93.4015	93.5204	94.1263	94.639	94.8471
RSP1F10	94.4633	93.1574	93.2438	93.3537	93.5507	93.606	94.2454	94.7685	94.8713
RSP1A1	94.5099	93.1505	93.2276	93.2807	93.5163	93.564	94.3443	94.7957	94.8268
Norway	93.1854	94.0896	94.1059	94.1466	93.4719	93.5175	93.3625	93.5451	93.5707
CC278f	93.0406	94.1744	94.0727	94.0595	93.1183	93.2171	93.2022	93.1838	93.258
SM78	93.1657	93.9024	93.9349	94.0201	93.3049	93.3722	93.199	93.2069	93.2672
SM51	93.0473	93.8794	93.967	93.9579	93.3104	93.3372	93.1876	93.1423	93.1962
SM72	93.2145	93.9796	94.0031	94.0578	93.3632	93.3464	93.34	93.2534	93.3931
Vaf-46	93.0445	92.9278	93.0961	93.0661	92.8752	93.0093	92.8699	93.3721	93.2647
JHI1093	93.1008	92.9461	93.1121	93.0791	93.0095	93.0185	92.9002	93.3073	93.3662
JHI1096	93.1265	92.9822	93.0503	93.0792	92.9344	93.0191	92.8839	93.246	93.2863
GPTR08	92.4687	92.4304	92.705	92.6428	92.4075	92.5337	92.3375	92.6301	92.6812
GPTR02	92.5262	92.5456	92.6428	92.5956	92.5041	92.5163	92.3973	92.6455	92.6626
SRDI565	92.5975	92.5364	92.713	92.7494	92.586	92.5938	92.4671	92.758	92.732
Ps8	93.482	93.523	93.6009	93.6521	93.3723	93.4375	93.2082	93.4291	93.5647
JHI1236	93.5615	93.4674	93.6687	93.6721	93.4726	93.3171	93.2222	93.4564	93.5523
JHI953	93.5309	93.4803	93.5749	93.6071	93.4349	93.4023	93.1043	93.4114	93.5517
SM147A	93.5178	93.4956	93.5973	93.6676	93.4492	93.4997	93.1684	93.2911	93.5287
JHI1238	93.5085	93.4681	93.6254	93.7084	93.3808	93.4392	93.189	93.5266	93.5846
UPM1133	93.497	93.4816	93.588	93.6348	93.3261	93.3795	93.1681	93.4621	93.6438
JHI1592	93.5598	93.4383	93.6214	93.5994	93.421	93.5117	93.183	93.3618	93.5894
SM41	93.4953	93.4991	93.6901	93.7019	93.3747	93.474	93.1845	93.317	93.4862
JHI1253	93.4947	93.455	93.6386	93.6676	93.4023	93.4381	93.2604	93.4044	93.5508
JHI370	93.5359	93.4623	93.601	93.5501	93.4395	93.477	93.147	93.4443	93.6042
JHI387	93.4869	93.4636	93.5599	93.587	93.3861	93.4593	93.1058	93.4036	93.5605
JHI388	93.5376	93.4532	93.5816	93.5962	93.4056	93.4588	93.1722	93.399	93.5515
JHI788	93.5703	93.5134	93.689	93.67	93.5052	93.5194	93.2674	93.4625	93.6259
JHI985	93.4742	93.4289	93.5963	93.6741	93.3883	93.4056	93.0999	93.3472	93.544
GB30	93.5413	93.4718	93.6388	93.6362	93.3804	93.4	93.2261	93.4038	93.6144
JHI782	93.4958	93.5002	93.5631	93.621	93.276	93.3456	93.1878	93.4458	93.5517
JHI783	93.4763	93.5401	93.5677	93.6221	93.3159	93.4715	93.1844	93.4135	93.527
SM170C	93.4935	93.4127	93.7012	93.6873	93.4089	93.3975	93.1319	93.3102	93.5026
JHI42	93.5134	93.4753	93.5261	93.4771	93.368	93.4421	93.0791	93.5297	93.6109
JHI979	93.5346	93.4843	93.6266	93.6261	93.407	93.4524	93.1602	93.4046	93.5183
JHI1259	93.4776	93.4958	93.6631	93.6642	93.3552	93.4602	93.1558	93.3783	93.4657
TA1	93.5155	93.4504	93.6949	93.7523	93.5185	93.5276	93.0856	93.3674	93.5125
JHI24	93.5357	93.3921	93.524	93.5796	93.3284	93.3906	93.2124	93.5284	93.8523
JHI787	93.4728	93.4261	93.5099	93.5382	93.3093	93.3806	93.2414	93.5306	93.9007
RCAM1026	93.5446	93.5895	93.6134	93.7264	93.3664	93.5079	93.3035	93.5133	93.634
Vh3	93.5851	93.4451	93.574	93.6514	93.3174	93.3564	93.2842	93.4091	93.5789
JHI925	93.4662	93.3767	93.5017	93.4833	93.2528	93.3175	93.1062	93.2984	93.4233
Vc2	93.3967	93.433	93.5564	93.4722	93.1917	93.3611	93.1944	93.4792	93.6921
JHI1422	93.5159	93.4834	93.5761	93.5916	93.3471	93.4071	93.1853	93.3664	93.5801
JHI1266	93.5452	93.4677	93.5382	93.5689	93.3713	93.4218	93.217	93.3266	93.5347
51A11	98.1588	93.6317	93.6791	93.6536	94.0734	94.1702	94.5833	95.2216	95.61
51B21	98.1421	93.6259	93.6523	93.6573	94.1406	94.1939	94.6202	95.1859	95.6501
128C53	92.6275	93.4847	93.3265	93.2449	92.7161	92.7894	92.8703	92.868	92.9727
4292	92.7254	93.4631	93.3272	93.3057	92.7773	92.8615	92.7847	92.7972	92.8376
CC283bq	92.7286	93.5333	93.3362	93.357	92.752	92.8196	92.7479	92.7862	92.8352
USDA2370	92.7472	93.3713	93.2248	93.2608	92.746	92.8404	92.7995	92.8543	93.0254
JHI2442	92.6991	93.521	93.314	93.3229	92.7797	92.8186	92.9173	92.8058	93.0195
JHI1084	92.697	93.4899	93.4116	93.3347	92.8582	92.8558	93.2094	92.8121	92.9404
JHI944	92.7053	93.4918	93.3138	93.3294	92.779	92.8565	92.8431	92.7839	92.8659
OV152	92.7269	93.3473	93.4345	93.3901	92.8028	92.8289	92.9578	92.9302	92.9134
UPM1137	92.7424	93.4383	93.2952	93.3539	92.8363	92.8373	92.9768	92.9425	93.0046
SM149A	92.6783	93.5005	93.4305	93.3173	92.8291	92.8348	92.8803	92.7382	92.8382
Rt24.2	92.8052	93.4864	93.4217	93.3585	92.8793	92.9296	92.8612	92.8167	92.9127
CCBAU23252	91.0851	91.1411	91.1077	91.0509	90.9436	91.0277	91.1909	91.1749	91.1869
JHI536	91.0818	91.183	91.0937	91.0836	90.9994	91.0216	91.2458	91.3847	91.4039
C15	91.1393	91.181	91.1839	91.1093	91.0299	91.0041	91.2225	91.1684	91.1686
CF307	91.175	91.2667	91.1926	91.2732	91.1188	91.1722	91.1548	91.1393	91.2018

Table S4.5. (Continuation)

Query	JHI10	JHI54	FA23	JHI2450	Vaf-108	JHI2449	UPM1131	CCBAU10279	WSM1689
3841	95.5849	95.4932	95.5624	95.5723	94.193	94.4203	94.31	94.286	94.2597
SM38	95.5322	95.5713	95.5627	95.6235	94.2716	94.3622	94.3725	94.3166	94.3025
JHI960	95.6282	95.4064	95.543	95.522	94.1591	94.4774	94.4329	94.2953	94.2843
JHI963	95.5798	95.4153	95.5581	95.5496	94.176	94.4742	94.419	94.3164	94.33
JHI1587	95.559	95.5113	95.5753	95.5712	94.2004	94.4458	94.3048	94.2584	94.2399
JHI1600	95.6571	95.5876	95.5715	95.6481	94.2977	94.4633	94.3555	94.2468	94.2305
JHI13	95.8921	95.543	95.6505	95.7464	94.1488	94.4542	94.4549	94.2752	94.3203
JHI1438	95.8585	95.5189	95.6422	95.6254	94.2251	94.4484	94.4279	94.2635	94.2748
JHI535	95.6864	95.6312	95.6492	95.7284	94.3812	94.4174	94.3952	94.2981	94.2469
JHI585	95.7299	95.6368	95.6264	95.6771	94.4063	94.456	94.4597	94.3199	94.2995
JHI973	95.6877	95.6574	95.6176	95.7296	94.3519	94.4182	94.4186	94.3018	94.2597
JHI974	95.7292	95.6442	95.6264	95.7207	94.3523	94.4363	94.437	94.3298	94.2859
JHI1415	95.6697	95.6211	95.6121	95.73	94.32	94.362	94.3699	94.3271	94.3579
SM3	95.5655	95.5552	95.5011	95.5849	94.1314	94.3454	94.2618	94.3077	94.2506
VF39	95.5546	95.4717	95.5102	95.5622	94.1816	94.3069	94.3488	94.2308	94.276
L361	94.8616	94.7748	94.833	94.8241	95.971	96.7539	96.2238	95.6396	94.8695
VFF1R2B1	95.6069	95.5253	95.5404	95.705	94.3056	94.503	94.4809	94.3114	94.4139
21A12	95.6682	95.6146	95.611	95.7066	94.389	94.5322	94.4388	94.2858	94.358
VFF1R2A1	95.6421	95.602	95.634	95.7396	94.382	94.5586	94.4846	94.3053	94.4142
WSM1455	95.6572	95.5808	95.6054	95.6423	94.2997	94.4691	94.3785	94.213	94.2641
WSM1481	95.5533	95.546	95.5726	95.5614	94.2738	94.4077	94.3307	94.2553	94.2505
WSM78	93.8164	93.8138	93.8989	93.8539	93.0138	93.1853	93.1235	93.0076	92.9037
SM152B	93.8597	93.9005	93.9446	93.9932	93.1029	93.2701	93.2084	93.0351	93.1556
CC275e	94.0177	94.037	94.1729	94.0987	93.128	93.3536	93.2668	93.1589	93.197
WSM1325	94.4574	94.3716	94.4851	94.5127	93.5315	93.5312	93.5045	93.2939	93.4035
SRDI943	94.5762	94.4884	94.5502	94.579	93.5546	93.578	93.5123	93.3602	93.4396
JHLM12A2	95.0022	94.9853	94.9986	95.1153	94.0998	94.3584	94.252	94.0263	93.943
Vaf12	95.696	95.603	95.6577	95.7576	95.1428	94.7295	94.9447	94.597	94.5096
JHI2451	97.4678	97.3748	98.3276	97.7549	94.6451	94.66	94.7734	94.52	94.6261
JHI10	100	97.8149	97.519	97.5147	94.3664	94.7084	94.6783	94.4563	94.5755
JHI54	97.8586	100	97.3882	97.4244	94.2285	94.6002	94.5837	94.3815	94.5426
FA23	97.524	97.3651	100	97.7372	94.3983	94.7697	94.7606	94.5587	94.6828
JHI2450	97.5526	97.4408	97.7012	100	94.679	94.7154	94.7951	94.5956	94.6468
Vaf-108	94.4315	94.2861	94.4639	94.6672	100	96.0412	95.9581	95.508	94.6686
JHI2449	94.8079	94.6482	94.8386	94.7229	95.9926	100	96.2962	95.6991	94.8288
UPM1131	94.7633	94.6891	94.7688	94.8098	96.0087	96.357	100	95.6012	94.7414
CCBAU10279	94.5607	94.5429	94.6673	94.6252	95.5065	95.7566	95.6022	100	95.4629
WSM1689	94.5627	94.6057	94.7299	94.6451	94.6922	94.8277	94.7963	95.5009	100
VFF2R2A1	94.4258	94.3018	94.4597	94.4725	94.8726	95.3182	95.1952	95.8217	95.3734
SPF2A11	94.3967	94.4647	94.4423	94.481	95.0903	95.302	95.1438	95.9068	95.3315
VFCR2A2	94.25	94.446	94.4741	94.4699	94.8496	95.2579	95.1229	95.8845	95.3075
248	94.389	94.4379	94.4919	94.4207	95.0161	95.3815	95.156	95.3254	94.7604
63A21	94.517	94.5242	94.4851	94.5966	94.6756	95.0381	94.8416	94.9971	94.5748
GLR2	94.4045	94.3489	94.3725	94.4718	94.7433	95.135	94.982	95.0831	94.7121
CL8	94.4379	94.4151	94.4391	94.5632	95.2296	95.5216	95.3814	96.4813	95.9012
VFSR2A2	94.4972	94.4989	94.4761	94.5031	95.2002	95.4526	95.3165	96.5198	95.885
SS21	94.4452	94.4078	94.4907	94.4677	95.1646	95.4373	95.2865	96.4386	95.7562
41A11	94.5152	94.4999	94.5152	94.5302	95.2839	95.4938	95.3312	96.5712	95.8557
42B12	94.6077	94.5575	94.6861	94.5768	95.3352	95.5933	95.4745	96.5967	95.9146
43A11	94.52	94.5227	94.5201	94.4982	95.2785	95.5276	95.3927	96.5715	95.8754
43B11	94.5204	94.4864	94.5205	94.4981	95.2439	95.4596	95.3045	96.5838	95.9367
43B12	94.5033	94.5188	94.4999	94.5344	95.1969	95.5318	95.3646	96.6161	95.8978
71A12	94.5467	94.455	94.522	94.5729	95.1416	95.5405	95.3981	96.5843	95.9682
73A11	94.4963	94.5206	94.5274	94.5007	95.2829	95.4914	95.4513	96.6051	95.9193
81B22	94.5202	94.448	94.5234	94.5465	95.208	95.5056	95.3998	96.5221	95.9687
83A12	94.5273	94.4725	94.4747	94.562	95.1695	95.4806	95.4067	96.5664	96.0395
VFHR1A2	94.5287	94.5119	94.4885	94.5801	95.1808	95.4759	95.372	96.5289	95.9074
VFF1R1A2	94.3799	94.3084	94.4009	94.4834	95.0825	95.434	95.2827	96.4807	95.8518
93B11	94.521	94.3927	94.4401	94.5857	95.2699	95.5165	95.4273	96.5551	95.8565
111A12	94.4434	94.4052	94.4568	94.5363	95.2249	95.5499	95.372	96.5088	95.9169
VFSR2B1	94.4356	94.4615	94.4352	94.5622	95.2354	95.5248	95.4141	96.5317	95.8769
SPF4F7	94.4459	94.4423	94.4613	94.5289	95.2039	95.4525	95.3649	96.5199	95.9182
FB206	94.4234	94.4238	94.4471	94.5047	95.0735	95.4107	95.3226	96.5456	95.9076
SL16	94.4102	94.3857	94.4748	94.4958	95.1884	95.429	95.3195	96.5181	95.8095
HUTR05	94.438	94.4115	94.5062	94.5037	95.3038	95.5293	95.3957	96.6034	95.9176
21B12	94.5217	94.3557	94.45	94.493	95.2041	95.5266	95.3344	96.5613	95.8923
WYCCWR11290	93.967	93.8766	93.992	94.0618	93.5772	93.954	93.7801	93.5396	93.534
WYCCWR11317	94.0428	94.012	94.0543	94.1354	93.7256	93.8752	93.8307	93.592	93.65
CCBAU11080	93.8594	93.847	93.979	94.0068	93.4086	93.6816	93.67	93.4614	93.4057
CCBAU03386	93.9326	93.8938	93.8591	94.0164	93.5005	93.6101	93.6497	93.5191	93.5235

Table S4.5. (Continuation)

Query	JHI10	JHI54	FA23	JHI2450	Vaf-108	JHI2449	UPM1131	CCBAU10279	WSM1689
WYCCWR10014	94.3318	94.3481	94.4458	94.3927	93.5616	93.6879	93.7436	93.5121	93.5307
JKLM13E	95.0797	95.0271	95.0852	95.1775	94.1662	94.3183	94.2951	93.9597	93.9817
Tri-43	94.4307	94.4248	94.4239	94.4562	94.056	94.2172	94.1861	93.9463	93.8987
73B11	94.7415	94.6892	94.7654	94.8447	95.6153	95.8433	95.8778	95.6158	94.9894
73B12	94.7087	94.7464	94.764	94.8906	95.6192	95.8852	95.9397	95.6481	94.9898
RSP1F2	94.7919	94.786	94.7881	94.8864	95.6509	95.7542	95.7375	95.5874	94.7866
Vaf10	94.441	94.3529	94.46	94.6779	96.2416	95.3389	95.5942	95.2503	94.6341
PEPV16	94.6963	94.6419	94.7095	94.7675	95.413	95.7232	95.69	95.6337	94.7641
TOM	94.6854	94.6519	94.6525	94.8258	95.4519	95.7274	95.6845	95.5129	94.7196
121B21	94.671	94.5596	94.7012	94.8246	95.4806	95.7273	95.754	95.5985	94.8879
RSP1F10	94.9243	94.7625	94.908	94.9893	95.4895	95.7449	95.8226	95.5621	94.8706
RSP1A1	94.9417	94.7454	94.8595	94.8258	95.4557	95.6247	95.7186	95.6355	94.8369
Norway	93.4897	93.5054	93.509	93.5867	92.9471	92.8596	93.0556	92.6656	92.7337
CC278f	93.1439	93.0608	93.2523	93.3693	92.4512	92.6526	92.7191	92.4451	92.5723
SM78	93.2661	93.2923	93.2996	93.3389	92.6889	92.7872	92.7563	92.5904	92.6475
SM51	93.2693	93.2103	93.349	93.3055	92.7063	92.7471	92.79	92.5756	92.6067
SM72	93.3121	93.1993	93.3367	93.4304	92.7164	92.8107	92.8566	92.6243	92.6975
Vaf-46	93.1004	93.1114	93.1219	93.2583	93.2687	93.0361	93.1126	92.8406	92.6221
JHI1093	93.2755	93.2262	93.1208	93.2839	93.1011	93.0366	93.0553	92.8182	92.6339
JHI1096	93.2328	93.1849	93.0832	93.2812	93.086	93.0361	93.0365	92.7874	92.6701
GPTR08	92.4772	92.4799	92.686	92.6265	92.4707	92.5856	92.5343	92.5142	92.5144
GPTR02	92.5506	92.5855	92.6915	92.6173	92.5526	92.6244	92.5753	92.5025	92.466
SRDI565	92.644	92.7269	92.7726	92.7117	92.514	92.5683	92.536	92.5483	92.5065
Ps8	93.5333	93.5398	93.4694	93.6213	92.9724	93.0688	92.9871	92.7654	92.7849
JHI1236	93.5996	93.4376	93.5446	93.5954	92.9361	92.952	92.9005	92.7754	92.7416
JHI953	93.5637	93.409	93.5065	93.5956	92.8851	92.9856	92.9656	92.6716	92.7506
SM147A	93.4455	93.4603	93.4859	93.4554	92.8882	92.9554	92.8687	92.7091	92.8102
JHI1238	93.5792	93.4903	93.5113	93.6072	92.9724	92.9377	92.9087	92.747	92.7595
UPM1133	93.5604	93.5651	93.5723	93.6679	92.96	92.9431	93.0027	92.7757	92.6483
JHI1592	93.5779	93.4722	93.4559	93.6045	92.9089	92.9704	92.9772	92.7668	92.7403
SM41	93.427	93.4352	93.4943	93.494	92.9696	92.8823	92.9094	92.7683	92.8137
JHI1253	93.4855	93.4571	93.5629	93.5299	92.8814	92.9771	92.9421	92.7877	92.7402
JHI370	93.576	93.5193	93.464	93.5681	92.8759	92.9711	92.9318	92.7408	92.7022
JHI387	93.5134	93.5206	93.4266	93.555	92.9111	92.9931	92.9813	92.7046	92.6808
JHI388	93.5659	93.5223	93.5146	93.566	92.8761	92.9527	92.9496	92.7236	92.7572
JHI788	93.6079	93.5164	93.532	93.5495	93.0248	93.0418	93.0677	92.8254	92.7871
JHI985	93.4917	93.4394	93.4915	93.4764	92.8927	92.9642	92.9274	92.6873	92.7054
GB30	93.5687	93.5194	93.5673	93.5353	92.9545	92.9853	92.9207	92.7959	92.756
JHI782	93.5629	93.4395	93.5117	93.5526	92.8461	92.9874	92.9383	92.8373	92.8173
JHI783	93.5652	93.4079	93.5445	93.5064	92.798	93.0023	92.9548	92.8264	92.8638
SM170C	93.4119	93.458	93.468	93.5216	92.9162	92.9199	92.8719	92.7182	92.8293
JHI42	93.7106	93.4906	93.3841	93.6542	92.9643	93.0157	93.0094	92.7387	92.7064
JHI979	93.5959	93.5303	93.4834	93.5277	92.8706	92.9969	92.9855	92.7907	92.8017
JHI1259	93.5405	93.4138	93.5573	93.5769	92.8827	93.0203	92.966	92.7917	92.8026
TA1	93.4295	93.4214	93.5808	93.537	92.9584	92.9809	92.9736	92.7141	92.8296
JHI24	93.6473	93.521	93.5186	93.8583	92.9751	93.036	93.0904	92.7378	92.8043
JHI787	93.6722	93.5598	93.4319	93.859	92.9487	93.0322	93.1167	92.764	92.7014
RCAM1026	93.616	93.6186	93.6219	93.6562	93.0688	93.0218	93.0199	92.7608	92.843
Vh3	93.5647	93.4739	93.6179	93.5686	92.854	92.951	93.0296	92.8358	92.8286
JHI925	93.4857	93.3539	93.4212	93.5049	92.7856	93.0678	92.8696	92.6851	92.7797
Vc2	93.5931	93.5281	93.5726	93.8318	92.943	92.9447	93.036	92.7812	92.6789
JHI1422	93.5501	93.5035	93.5202	93.5569	92.8463	92.9586	92.9658	92.7905	92.7959
JHI1266	93.5111	93.3304	93.6433	93.4923	92.8322	92.9921	92.9661	92.8369	92.8458
51A11	95.5404	95.5296	95.5371	95.6086	94.2628	94.4829	94.4199	94.1916	94.2671
51B21	95.6033	95.5433	95.5356	95.6174	94.2825	94.5217	94.4145	94.1626	94.2734
128C53	92.9278	92.8518	93.0399	93.0165	92.408	92.6131	92.5009	92.2345	92.3125
4292	92.8683	92.7858	93.2643	92.883	92.2973	92.3814	92.4026	92.1381	92.2775
CC283bq	92.7453	92.6177	92.7366	92.8824	92.3783	92.3668	92.501	92.1531	92.2874
USDA2370	93.1569	92.8562	92.9356	93.053	92.3274	92.4439	92.5941	92.2247	92.2798
JHI2442	92.9954	92.8378	92.9187	93.1229	92.3455	92.5805	92.5127	92.1864	92.3222
JHI1084	92.8419	92.7674	92.8352	92.9719	92.2141	92.586	92.3939	92.2351	92.3093
JHI944	92.8725	92.7332	92.8279	92.8537	92.2367	92.5404	92.4629	92.2608	92.3204
OV152	92.9239	92.8509	92.9261	92.9991	92.2978	92.4632	92.4845	92.1555	92.1735
UPM1137	92.9725	92.9578	92.933	93.054	92.4006	92.4656	92.5837	92.2071	92.2372
SM149A	92.7257	92.7706	92.8439	92.9304	92.2501	92.4323	92.3773	92.117	92.3413
Rt24.2	92.8279	92.9036	92.9274	92.8707	92.2878	92.4092	92.4626	92.2207	92.2976
CCBAU23252	91.2318	91.2045	91.0874	91.2514	90.83	90.9252	90.9261	90.6899	90.7145
JHI536	91.3759	91.337	91.2087	91.3856	90.933	91.003	91.0661	90.8016	90.7613
C15	91.2032	91.147	91.2211	91.2167	90.743	90.8942	90.9336	90.7919	90.8241
CF307	91.2519	91.2624	91.2376	91.2638	90.8526	90.8701	90.962	90.822	90.8963

Table S4.5. (Continuation)

Query	VFF2R2A1	SPF2A11	VFCR2A2	248	63A21	GLR2	CL8	VFSR2A2	SS21	41A11
3841	94.1991	94.2713	94.2209	94.411	94.4177	94.3469	94.1726	94.4487	94.0732	94.3208
SM38	94.1439	94.1935	94.1329	94.2955	94.2834	94.3125	94.2006	94.2165	94.1428	94.2394
JHI960	94.2397	94.189	94.1957	94.4798	94.472	94.3612	94.2057	94.4419	94.1146	94.2948
JHI963	94.2831	94.2923	94.294	94.4697	94.5473	94.3571	94.1977	94.4311	94.1218	94.3545
JHI1587	94.2788	94.3025	94.2794	94.4715	94.4544	94.3492	94.2833	94.4005	94.1259	94.3143
JHI1600	94.3034	94.2994	94.309	94.4618	94.4514	94.3399	94.202	94.4756	94.1686	94.3633
JHI13	94.1697	94.2428	94.1399	94.3188	94.3794	94.2238	94.1528	94.2316	94.1147	94.253
JHI1438	94.178	94.1702	94.0756	94.262	94.23	94.1739	94.1038	94.1779	94.0887	94.1851
JHI535	94.2484	94.2967	94.2534	94.4203	94.4408	94.3702	94.1528	94.4042	94.1714	94.2945
JHI585	94.2263	94.2943	94.2635	94.4555	94.5128	94.3906	94.1994	94.4041	94.1791	94.3132
JHI973	94.3133	94.2756	94.3315	94.3887	94.4615	94.4034	94.232	94.4291	94.1864	94.3157
JHI974	94.3134	94.2933	94.3013	94.4119	94.4531	94.406	94.2017	94.4389	94.194	94.292
JHI1415	94.2255	94.2943	94.2162	94.4736	94.4978	94.3672	94.1842	94.4252	94.1123	94.3572
SM3	94.1264	94.0896	94.1031	94.2914	94.2815	94.2339	94.1059	94.1971	94.128	94.1518
VF39	94.1992	94.2089	94.1241	94.3457	94.3932	94.318	94.2256	94.3737	94.0711	94.2892
L361	95.2343	95.1927	95.2236	95.3372	94.932	95.3712	95.427	95.4225	95.4029	95.4358
VFF1R2B1	94.253	94.3492	94.2143	94.4608	94.5056	94.4152	94.2355	94.2479	94.2894	94.3082
21A12	94.2748	94.362	94.2285	94.4221	94.4208	94.3383	94.3201	94.3608	94.2825	94.3362
VFF1R2A1	94.2233	94.3763	94.2522	94.4164	94.4498	94.3859	94.2733	94.3555	94.2481	94.2709
WSM1455	94.1901	94.2023	94.2094	94.3962	94.4857	94.3901	94.1556	94.2257	94.1555	94.1907
WSM1481	94.2271	94.1663	94.1997	94.4115	94.4835	94.3554	94.145	94.1585	94.0831	94.1492
WSM78	92.9151	92.9595	92.9132	93.1206	93.1401	92.9669	92.9185	92.8793	92.8902	92.9648
SM152B	92.9268	92.9531	92.921	93.1346	93.0693	93.0387	92.9017	92.9266	92.9383	93.0354
CC275e	92.8942	93.0319	92.944	93.1253	93.0883	93.0357	93.0383	93.0532	92.9754	93.0865
WSM1325	93.2789	93.2298	93.268	93.4057	93.3741	93.3647	93.3214	93.2421	93.1441	93.2908
SRDI943	93.3183	93.2854	93.2933	93.4992	93.5246	93.4398	93.2592	93.3239	93.2315	93.3437
JHLM12A2	93.9527	93.9024	93.9107	94.0646	94.0056	93.9532	93.8668	93.9188	93.8988	94.0077
Vaf12	94.5099	94.4848	94.4149	94.5141	94.4319	94.4326	94.4205	94.4668	94.381	94.5099
JHI2451	94.4212	94.4436	94.4616	94.454	94.4809	94.4263	94.4476	94.4996	94.4496	94.5079
JHI10	94.4145	94.329	94.2859	94.3788	94.4562	94.3873	94.3603	94.4996	94.3862	94.4534
JHI54	94.3238	94.4145	94.355	94.3603	94.3866	94.347	94.3597	94.4624	94.3159	94.4137
FA23	94.4374	94.3919	94.4488	94.475	94.4082	94.4163	94.4131	94.3686	94.4288	94.4162
JHI2450	94.5128	94.5086	94.4801	94.419	94.616	94.5212	94.4855	94.532	94.4238	94.4716
Vaf-108	94.9201	95.0833	94.8987	95.0162	94.6494	94.7917	95.1705	95.231	95.1027	95.1918
JHI2449	95.3225	95.3096	95.1887	95.3877	94.9928	95.126	95.4198	95.4	95.3802	95.464
UPM1131	95.1294	95.1459	95.069	95.1534	94.8433	94.9586	95.3467	95.2895	95.2601	95.326
CCBAU10279	95.8393	95.8567	95.9291	95.3254	95.0656	95.2012	96.478	96.5252	96.4949	96.509
WSM1689	95.3519	95.366	95.3507	94.7216	94.6224	94.7556	95.8714	95.8755	95.7507	95.8321
VFF2R2A1	100	97.9037	99.488	96.0663	96.2985	96.2866	96.8049	96.8502	96.6917	96.779
SPF2A11	97.889	100	97.9597	96.097	96.157	96.2792	96.7935	96.8273	96.8026	96.8736
VFCR2A2	99.4935	97.9441	100	96.1114	96.3775	96.2139	96.8931	96.9046	96.74	96.8337
248	96.112	96.0858	96.1206	100	96.4591	96.5985	95.5189	95.6243	95.499	95.6376
63A21	96.2815	96.1157	96.376	96.5389	100	96.9816	95.3833	95.411	95.2659	95.4
GLR2	96.2856	96.2989	96.2659	96.5707	96.9279	100	95.4365	95.4974	95.4216	95.5172
CL8	96.854	96.8357	96.8376	95.5992	95.4449	95.5237	100	98.136	98.1253	98.1304
VFSR2A2	96.8486	96.8844	96.8814	95.5973	95.3705	95.5283	98.0691	100	99.49	99.3645
SS21	96.7038	96.7725	96.7102	95.5244	95.3176	95.4461	98.1243	99.4844	100	99.4555
41A11	96.8431	96.9244	96.8708	95.6725	95.3746	95.5768	98.1493	99.4389	99.4507	100
42B12	96.9254	96.9946	96.9268	95.7408	95.4721	95.6424	98.1997	99.4546	99.4898	99.9935
43A11	96.8841	96.8932	96.8604	95.6463	95.3859	95.5604	98.1299	99.448	99.4898	99.9967
43B11	96.8988	96.9335	96.9047	95.6178	95.4665	95.5102	98.1554	99.5499	99.4126	99.3662
43B12	96.8856	96.9266	96.8867	95.6333	95.4834	95.5001	98.1742	99.5426	99.3889	99.3672
71A12	96.8153	96.9346	96.8237	95.6412	95.3562	95.5489	98.2987	98.7692	98.6751	98.7667
73A11	96.7873	96.8827	96.8475	95.635	95.3503	95.4501	98.2681	98.7227	98.6239	98.6604
81B22	96.7694	96.9279	96.783	95.5937	95.3363	95.5739	98.2835	98.7493	98.6741	98.7924
83A12	96.8238	96.9142	96.8279	95.6201	95.396	95.5232	98.3047	98.7595	98.6796	98.7938
VFHR1A2	97.0182	96.8291	97.0009	95.5273	95.6016	95.5477	98.1705	99.3614	99.4129	99.3939
VFF1R1A2	96.6599	96.8058	96.6964	95.5369	95.284	95.3619	98.1422	98.5831	98.3759	98.5844
93B11	96.7871	96.8261	96.7726	95.5557	95.3412	95.4399	98.1478	98.5581	98.5033	98.5977
111A12	96.8139	96.848	96.8785	95.5903	95.3403	95.5742	98.2989	98.3066	98.1605	98.2415
VFSR2B1	96.8498	96.8646	96.855	95.631	95.3934	95.5439	98.2707	98.2941	98.1381	98.2459
SPF4F7	96.8354	96.8582	96.8392	95.5496	95.347	95.4341	98.1578	98.2756	98.184	98.32
FB206	96.7965	96.7583	96.7818	95.4916	95.3446	95.3844	98.2346	98.4036	98.1536	98.2344
SL16	96.7383	96.7664	96.7633	95.4975	95.363	95.5167	98.2778	98.5912	98.7136	98.547
HUTR05	96.7807	96.7722	96.7845	95.4628	95.2691	95.4912	98.2991	98.2865	98.1604	98.2042
21B12	96.7283	96.8024	96.7059	95.5834	95.3389	95.479	98.1446	98.3579	98.2013	98.3147
WYCCWR11290	93.6813	93.6413	93.6462	93.6425	93.6259	93.5635	93.615	93.5738	93.5254	93.639
WYCCWR11317	93.6246	93.5816	93.6248	93.6732	93.6441	93.716	93.5587	93.6132	93.5504	93.5762
CCBAU11080	93.439	93.3837	93.3237	93.4666	93.442	93.41	93.4243	93.4927	93.3491	93.4651
CCBAU03386	93.5071	93.5555	93.5388	93.3939	93.5462	93.3773	93.5449	93.5411	93.4477	93.509

Table S4.5. (Continuation)

Query	VFF2R2A1	SPF2A11	VFCR2A2	248	63A21	GLR2	CL8	VFSR2A2	SS21	41A11
WYCCWR10014	93.3822	93.4394	93.3974	93.6106	93.5981	93.5083	93.4039	93.3863	93.3781	93.5131
JKLM13E	93.935	93.9276	93.946	94.0088	94.077	94.0207	93.91	93.8995	93.9159	93.9828
Tri-43	93.9128	93.8488	93.9198	93.924	93.978	93.9319	93.8925	93.8813	93.8902	93.8583
73B11	95.1821	95.1799	95.2237	94.9561	94.7956	94.7684	95.5464	95.576	95.5151	95.5817
73B12	95.2368	95.2507	95.3078	95.0144	94.8184	94.8028	95.571	95.5863	95.525	95.5735
RSP1F2	95.1432	95.094	95.076	94.8454	94.6385	94.7711	95.5373	95.5639	95.4575	95.5465
Vaf10	94.7663	94.8524	94.7089	94.6847	94.4785	94.5702	94.9763	95.0391	94.8889	95.0023
PEPV16	95.0465	95.1067	95.0309	94.7977	94.5944	94.6227	95.4734	95.5277	95.4694	95.4718
TOM	95.0224	95.0059	95.069	94.8225	94.6395	94.7124	95.3879	95.4096	95.4018	95.3962
121B21	95.0386	95.1254	94.9713	94.8393	94.5486	94.6678	95.4632	95.5131	95.6384	95.4958
RSP1F10	95.082	95.1238	95.0564	94.9364	94.6326	94.6481	95.5093	95.4241	95.4001	95.4435
RSP1A1	95.1628	95.1166	94.9452	94.9064	94.5579	94.7389	95.4217	95.3776	95.3923	95.4459
Norway	92.6996	92.6999	92.6785	92.7979	92.7328	92.7639	92.6711	92.7058	92.6089	92.6837
CC278f	92.429	92.5518	92.451	92.6081	92.5952	92.6331	92.485	92.5623	92.4143	92.5018
SM78	92.5655	92.6143	92.5208	92.6938	92.6619	92.6332	92.5764	92.5312	92.4564	92.5256
SM51	92.5118	92.5497	92.4728	92.6565	92.6237	92.5518	92.4457	92.4947	92.3845	92.4373
SM72	92.6058	92.6695	92.5781	92.7076	92.7208	92.6286	92.6108	92.6056	92.4526	92.5767
Vaf-46	92.7341	92.7572	92.6807	93.1578	92.9745	92.9345	92.7303	92.6989	92.6716	92.6882
JHI1093	92.8781	92.8682	92.813	93.1272	93.0901	92.9632	92.7327	92.7598	92.6273	92.7536
JHI1096	92.8692	92.8923	92.8009	93.1078	93.067	92.9985	92.7276	92.7765	92.5961	92.7147
GPTR08	92.4543	92.4789	92.4761	92.668	92.6779	92.6461	92.4249	92.2842	92.3508	92.4059
GPTR02	92.5457	92.571	92.5014	92.8089	92.8183	92.7115	92.4502	92.4562	92.3939	92.5023
SRDI565	92.4616	92.5381	92.5034	92.7832	92.7556	92.7321	92.4411	92.4171	92.4308	92.4816
Ps8	92.8411	92.8375	92.7741	93.0243	92.9616	92.9995	92.715	92.7487	92.6885	92.7801
JHI1236	92.8024	92.8088	92.7964	93.0718	92.9986	92.8947	92.7245	92.7273	92.725	92.8296
JHI953	92.7963	92.8274	92.7717	92.9835	93.0114	92.7734	92.7522	92.7842	92.5846	92.7884
SM147A	92.7571	92.8816	92.7177	92.9804	92.9445	92.9544	92.6971	92.7004	92.6171	92.725
JHI1238	92.8162	92.8212	92.7674	92.9891	92.9854	92.9091	92.7207	92.8047	92.634	92.803
UPM1133	92.6807	92.7333	92.6082	92.9251	92.9897	92.9559	92.647	92.6372	92.58	92.718
JHI1592	92.7252	92.8416	92.7042	93.0399	92.93	92.838	92.7037	92.7919	92.6395	92.7359
SM41	92.7233	92.8003	92.7075	92.9296	92.9106	92.8509	92.6598	92.6818	92.5833	92.7554
JHI1253	92.7816	92.7978	92.7359	93.0558	92.9581	92.821	92.7109	92.7547	92.6594	92.7443
JHI370	92.7402	92.7979	92.6916	93.001	92.916	92.7718	92.6963	92.7587	92.652	92.6989
JHI387	92.6814	92.8259	92.6649	93.0046	92.8722	92.8605	92.6504	92.7219	92.6551	92.7016
JHI388	92.6983	92.8495	92.6607	93.0939	92.9229	92.8902	92.6831	92.7949	92.6603	92.7095
JHI788	92.868	92.8738	92.8386	93.0987	93.0595	92.9473	92.7343	92.8042	92.7452	92.7994
JHI985	92.7495	92.7708	92.7708	92.9551	92.93	92.9065	92.7291	92.7693	92.6304	92.7567
GB30	92.8153	92.8288	92.7852	93.0428	92.9779	93.0236	92.7427	92.7771	92.6622	92.8018
JHI782	92.7502	92.7918	92.756	92.9002	92.8827	92.8985	92.6335	92.787	92.6189	92.7663
JHI783	92.7728	92.7512	92.7716	92.9771	92.9067	92.9028	92.71	92.7798	92.6648	92.7983
SM170C	92.6695	92.7354	92.6689	92.9089	92.8932	92.8191	92.7022	92.6644	92.574	92.7497
JHI42	92.6744	92.7882	92.6649	93.0016	92.9232	92.8623	92.7068	92.768	92.6646	92.7659
JHI979	92.8057	92.8447	92.7531	93.052	93.0132	92.8978	92.7615	92.742	92.6157	92.837
JHI1259	92.8101	92.8524	92.8068	93.0586	92.9627	92.8758	92.7176	92.7437	92.6349	92.7725
TA1	92.792	92.8618	92.761	93.0034	92.9199	92.8929	92.7492	92.653	92.6506	92.7894
JHI24	92.7348	92.7768	92.664	92.9945	92.9504	92.843	92.7913	92.766	92.6738	92.7487
JHI787	92.7402	92.8441	92.6808	93.0302	92.9388	92.9011	92.7237	92.783	92.6853	92.797
RCAM1026	92.81	92.8949	92.7978	93.0398	92.9582	93.0045	92.7982	92.7979	92.7671	92.8498
Vh3	92.6602	92.8162	92.7754	92.9787	92.978	92.9748	92.7358	92.7175	92.6703	92.7221
JHI925	92.8509	92.8865	92.8094	93.114	92.9817	92.9427	92.7085	93.0365	92.6558	92.7684
Vc2	92.8711	92.8445	92.7786	92.9454	92.9858	92.935	92.6712	92.7657	92.6219	92.6876
JHI1422	92.835	92.7836	92.7833	92.9965	92.97	92.9995	92.747	92.7605	92.6826	92.7429
JHI1266	92.9331	92.8567	92.8324	93.0145	93.0439	92.9189	92.7	92.8288	92.634	92.8326
51A11	94.281	94.2824	94.2069	94.4432	94.3778	94.3575	94.1947	94.3033	94.196	94.2164
51B21	94.2638	94.269	94.222	94.4231	94.4039	94.4084	94.2846	94.3101	94.2458	94.2543
128C53	92.2129	92.2722	92.113	92.3622	92.4722	92.2376	92.2005	92.2013	92.1522	92.2285
4292	92.1684	92.2628	92.1838	92.3812	92.4324	92.3012	92.2666	92.2734	92.144	92.2677
CC283bq	92.206	92.2697	92.1227	92.3384	92.3784	92.2423	92.2155	92.185	92.0866	92.1704
USDA2370	92.2275	92.2201	92.15	92.3868	92.4775	92.3215	92.262	92.2478	92.124	92.2264
JHI2442	92.2953	92.325	92.2991	92.4939	92.4812	92.3969	92.3163	92.2885	92.208	92.3172
JHI1084	92.2693	92.2609	92.2408	92.4018	92.4437	92.3649	92.2192	92.2297	92.1167	92.2593
JHI944	92.377	92.3203	92.4185	92.4826	92.6025	92.3497	92.2677	92.5443	92.235	92.3199
OV152	92.1441	92.2376	92.141	92.3743	92.3834	92.2621	92.2691	92.1944	92.1804	92.2187
UPM1137	92.2041	92.2551	92.2208	92.3795	92.5179	92.3339	92.2476	92.2741	92.1983	92.2786
SM149A	92.2295	92.2197	92.1305	92.3895	92.3937	92.2179	92.1809	92.244	92.1356	92.262
Rt24.2	92.2616	92.2302	92.2232	92.3895	92.4703	92.2393	92.2184	92.2553	92.1718	92.2491
CCBAU23252	90.8194	90.7782	90.7663	90.8845	91.0386	90.8295	90.7256	90.7932	90.886	90.7556
JHI536	90.8792	90.8438	90.8261	90.9035	91.0393	90.9024	90.843	90.8952	90.8039	90.877
C15	90.7763	90.8185	90.779	90.8708	91.0495	90.7778	90.7705	90.8168	90.7152	90.8102
CF307	90.824	90.7873	90.7772	90.8958	90.9259	90.7666	90.7735	90.872	90.8034	90.7763

Table S4.5. (Continuation)

Query	42B12	43A11	43B11	43B12	71A12	73A11	81B22	83A12	VFHR1A2	VFF1R1A2
3841	94.2507	94.3163	94.3598	94.3642	94.2594	94.2828	94.2337	94.2768	94.3028	94.2126
SM38	94.2132	94.2918	94.1831	94.1913	94.1976	94.1683	94.2746	94.2133	94.1957	94.1389
JHI960	94.3642	94.381	94.4367	94.411	94.3622	94.2687	94.3692	94.3542	94.2925	94.2484
JHI963	94.3374	94.3656	94.4965	94.4531	94.2778	94.2728	94.3263	94.3354	94.3412	94.2672
JHI1587	94.2502	94.3437	94.4025	94.4312	94.3326	94.2848	94.2755	94.3253	94.3032	94.2973
JHI1600	94.3677	94.4434	94.4166	94.4264	94.2567	94.3433	94.2814	94.2978	94.3383	94.3318
JHI13	94.1772	94.2554	94.1885	94.1982	94.2407	94.2136	94.2487	94.2317	94.1908	94.0517
JHI1438	94.0994	94.1812	94.1031	94.1161	94.1335	94.1429	94.1532	94.1267	94.1441	93.9651
JHI535	94.2787	94.3316	94.3933	94.3804	94.2752	94.2796	94.3212	94.2803	94.3111	94.3071
JHI585	94.2409	94.3215	94.4065	94.3907	94.3078	94.3369	94.3223	94.3193	94.3028	94.3136
JHI973	94.2675	94.3388	94.4323	94.4425	94.302	94.3229	94.2744	94.2582	94.2666	94.3638
JHI974	94.2458	94.3387	94.4553	94.4497	94.3059	94.2879	94.2908	94.2901	94.306	94.3712
JHI1415	94.2432	94.3405	94.4193	94.4227	94.2638	94.2594	94.2618	94.2715	94.2824	94.3277
SM3	94.0497	94.1709	94.1445	94.1587	94.1295	94.1148	94.2018	94.1683	94.1645	94.0436
VF39	94.205	94.2678	94.3192	94.3605	94.2456	94.2088	94.279	94.2102	94.2359	94.2282
L361	95.4016	95.4188	95.359	95.3868	95.3791	95.3701	95.3821	95.3522	95.3673	95.388
VFF1R2B1	94.2566	94.3226	94.2708	94.2955	94.3058	94.2625	94.3733	94.3056	94.3018	94.3107
21A12	94.2311	94.2902	94.2238	94.263	94.2608	94.2547	94.2528	94.249	94.3205	94.2414
VFF1R2A1	94.2001	94.2565	94.2222	94.224	94.2464	94.3004	94.2448	94.2521	94.3017	94.197
WSM1455	94.125	94.2407	94.1306	94.1487	94.2409	94.2262	94.2498	94.2365	94.219	94.1986
WSM1481	94.1098	94.1717	94.1728	94.1732	94.1703	94.2112	94.1939	94.1661	94.2168	94.1395
WSM78	92.922	92.9531	92.8998	92.9131	92.9558	92.9351	92.9531	92.9354	92.9108	92.9766
SM152B	93.012	93.0425	92.9246	92.9156	93.0037	93.0342	93.0186	92.9454	92.9786	92.9366
CC275e	93.0199	93.0532	92.9629	93.0505	93.0623	93.0091	93.0541	93.0182	93.0842	93.013
WSM1325	93.2515	93.2988	93.2171	93.2274	93.2947	93.253	93.2883	93.2985	93.2435	93.2441
SRDI943	93.3116	93.3417	93.2894	93.3368	93.3667	93.3367	93.3726	93.3616	93.3189	93.3039
JHLM12A2	93.9793	94.0054	93.9495	93.9819	93.9065	93.9006	93.9154	93.9654	93.9195	93.8753
Vaf12	94.4387	94.4863	94.4332	94.4201	94.5095	94.5052	94.5279	94.5052	94.4149	94.4767
JHI2451	94.4635	94.5176	94.5392	94.538	94.4664	94.496	94.4775	94.4456	94.5155	94.4614
JHI10	94.4995	94.5408	94.4451	94.4497	94.4653	94.4225	94.4498	94.4332	94.4557	94.366
JHI54	94.3228	94.4283	94.4076	94.4174	94.4239	94.3588	94.4538	94.3845	94.3491	94.2747
FA23	94.3703	94.4286	94.457	94.4272	94.4324	94.5029	94.4705	94.4396	94.4066	94.3392
JHI2450	94.4223	94.4994	94.5003	94.531	94.474	94.4661	94.5218	94.5387	94.5421	94.4767
Vaf-108	95.1269	95.1941	95.1478	95.1047	95.1492	95.1822	95.1195	95.1326	95.1394	95.0177
JHI2449	95.3927	95.427	95.4094	95.4009	95.434	95.4632	95.4812	95.4243	95.4174	95.3714
UPM1131	95.2772	95.3405	95.2874	95.2899	95.3463	95.4152	95.3803	95.352	95.359	95.3019
CCBAU10279	96.4693	96.4955	96.5123	96.5263	96.5195	96.5436	96.5292	96.5045	96.5097	96.423
WSM1689	95.8012	95.8281	95.8603	95.9024	95.9276	95.8683	95.9653	95.9541	95.913	95.8655
VFF2R2A1	96.7666	96.8119	96.8541	96.8192	96.7917	96.7985	96.8349	96.814	96.9319	96.6404
SPF2A11	96.8709	96.8946	96.8722	96.8645	96.8799	96.856	96.8837	96.8862	96.8217	96.7857
VFCR2A2	96.8256	96.848	96.9096	96.9075	96.8405	96.8647	96.8511	96.852	96.9986	96.722
248	95.6343	95.6569	95.5809	95.5953	95.5849	95.5901	95.6007	95.5957	95.5179	95.4521
63A21	95.3613	95.4018	95.4454	95.4572	95.327	95.3685	95.3767	95.3608	95.5685	95.2901
GLR2	95.4249	95.4925	95.4222	95.4171	95.4745	95.409	95.5258	95.4912	95.4829	95.4025
CL8	98.1381	98.1528	98.2002	98.1873	98.3019	98.2954	98.3059	98.2908	98.1893	98.1489
VFSR2A2	99.386	99.4062	99.5556	99.5523	98.7972	98.7275	98.798	98.7875	99.4063	98.559
SS21	99.4281	99.4702	99.4121	99.4124	98.67	98.5867	98.6916	98.6675	99.3816	98.3835
41A11	99.9969	99.9966	99.3791	99.3768	98.7723	98.6586	98.8015	98.7931	99.4213	98.5959
42B12	100	99.995	99.3823	99.3848	98.8258	98.7297	98.8426	98.8282	99.4065	98.6762
43A11	99.9968	100	99.3895	99.3853	98.7817	98.6742	98.7875	98.8038	99.3969	98.6266
43B11	99.3805	99.3829	100	99.9978	98.8208	98.7847	98.8284	98.8267	99.4473	98.5879
43B12	99.383	99.3777	99.9963	100	98.8201	98.7925	98.8082	98.8382	99.4439	98.6173
71A12	98.746	98.7729	98.7707	98.7834	100	99.4774	99.9768	99.9866	98.846	98.6037
73A11	98.6844	98.692	98.7886	98.7489	99.4732	100	99.4721	99.4746	98.878	98.5128
81B22	98.773	98.7805	98.7781	98.7731	99.9772	99.493	100	99.9889	98.8842	98.5952
83A12	98.7814	98.7962	98.7948	98.7952	99.9836	99.4634	99.9826	100	98.899	98.5828
VFHR1A2	99.4232	99.4244	99.4542	99.4493	98.8844	98.9028	98.906	98.899	100	98.6416
VFF1R1A2	98.5822	98.6177	98.6106	98.6068	98.5771	98.5077	98.5773	98.5535	98.6317	100
93B11	98.6067	98.6226	98.6031	98.6285	98.6035	98.5448	98.5904	98.5948	98.6183	99.7584
111A12	98.2414	98.2466	98.333	98.3445	98.4462	98.5628	98.4553	98.4605	98.3285	98.329
VFSR2B1	98.2185	98.2371	98.341	98.3148	98.4752	98.5574	98.4803	98.4728	98.2959	98.3542
SPF4F7	98.3159	98.3275	98.2953	98.3212	98.3332	98.3126	98.3402	98.3594	98.2486	98.2917
FB206	98.194	98.2581	98.3541	98.3951	98.4903	98.4701	98.4471	98.4838	98.3018	98.279
SL16	98.5802	98.5754	98.7082	98.6867	98.7991	98.8043	98.7992	98.7869	98.7193	98.2926
HUTR05	98.1829	98.2085	98.3911	98.3849	98.3142	98.3291	98.3501	98.3442	98.2791	98.29
21B12	98.2863	98.3107	98.4124	98.3973	98.5293	98.5599	98.5096	98.5403	98.3389	98.6022
WYCCWR11290	93.613	93.6704	93.5644	93.5536	93.5606	93.5687	93.5931	93.5689	93.6006	93.5911
WYCCWR11317	93.6044	93.65	93.6393	93.6283	93.5871	93.5522	93.6215	93.5974	93.6355	93.6107
CCBAU11080	93.4671	93.5045	93.539	93.5113	93.5424	93.5066	93.5497	93.5552	93.5058	93.4234
CCBAU03386	93.4828	93.5371	93.4938	93.516	93.5406	93.5817	93.5462	93.5066	93.5822	93.4221

Table S4.5. (Continuation)

Query	42B12	43A11	43B11	43B12	71A12	73A11	81B22	83A12	VFHR1A2	VFF1R1A2
WYCCWR10014	93.4937	93.5348	93.4328	93.4608	93.4416	93.4511	93.5011	93.4379	93.4354	93.3665
JKLM13E	93.9078	93.9375	93.9106	93.9494	93.9706	93.9025	93.9863	93.9484	93.9406	93.9402
Tri-43	93.8024	93.877	93.867	93.9091	93.9144	93.8899	93.8934	93.9372	93.8337	93.8484
73B11	95.5356	95.5959	95.5486	95.5444	95.5757	95.606	95.585	95.5811	95.5943	95.5444
73B12	95.5784	95.6163	95.5573	95.5718	95.6243	95.6558	95.687	95.6208	95.6124	95.6301
RSP1F2	95.5467	95.5958	95.493	95.5059	95.4854	95.5125	95.5167	95.4977	95.5117	95.501
Vaf10	95.0203	95.0595	94.9742	95.0264	94.9791	95.0705	94.9934	94.9866	94.9936	94.8713
PEPV16	95.4674	95.5146	95.4926	95.5145	95.4977	95.4996	95.4993	95.5069	95.5045	95.4723
TOM	95.4366	95.4582	95.3991	95.3941	95.4025	95.4559	95.3823	95.3736	95.4233	95.4099
121B21	95.4808	95.5379	95.4599	95.4834	95.4554	95.3997	95.498	95.4479	95.471	95.4053
RSP1F10	95.4762	95.5106	95.448	95.4359	95.4737	95.4973	95.4811	95.4495	95.5229	95.4758
RSP1A1	95.4202	95.4656	95.3622	95.4197	95.5001	95.4617	95.4764	95.4611	95.439	95.4548
Norway	92.6279	92.626	92.6537	92.6919	92.7087	92.6819	92.6595	92.73	92.6297	92.5873
CC278f	92.4953	92.5841	92.5245	92.5461	92.5652	92.5624	92.6282	92.5114	92.5328	92.4903
SM78	92.4709	92.5319	92.5189	92.5388	92.5275	92.5545	92.5513	92.5184	92.5595	92.549
SM51	92.4363	92.4925	92.4935	92.4637	92.5467	92.5587	92.5639	92.5273	92.5402	92.4383
SM72	92.6251	92.6582	92.6186	92.642	92.6438	92.6217	92.6545	92.6431	92.6776	92.52
Vaf-46	92.6374	92.6811	92.6985	92.6792	92.6977	92.7658	92.7461	92.6933	92.6786	92.612
JHI1093	92.7014	92.7583	92.6987	92.6983	92.7863	92.7678	92.7843	92.8092	92.7691	92.7029
JHI1096	92.7084	92.7371	92.6966	92.7134	92.7959	92.7465	92.803	92.8252	92.7885	92.6866
GPTR08	92.3702	92.4513	92.3731	92.4129	92.4434	92.3753	92.4835	92.4673	92.3512	92.3603
GPTR02	92.431	92.4953	92.4464	92.4687	92.5759	92.4975	92.5686	92.5807	92.4827	92.4454
SRDI565	92.4021	92.4791	92.4652	92.433	92.3972	92.3739	92.4854	92.425	92.4213	92.3651
Ps8	92.7387	92.7562	92.751	92.7894	92.7979	92.7519	92.7736	92.7842	92.7075	92.7239
JHI1236	92.7136	92.7879	92.6923	92.7187	92.7972	92.7714	92.766	92.8259	92.7057	92.7278
JHI953	92.7122	92.807	92.6935	92.7146	92.7687	92.7476	92.7746	92.7456	92.7335	92.7449
SM147A	92.6169	92.7519	92.6752	92.7072	92.7043	92.7439	92.7434	92.7244	92.7069	92.6784
JHI1238	92.6746	92.777	92.6946	92.7954	92.7021	92.7498	92.784	92.7407	92.7012	92.7084
UPM1133	92.62	92.7126	92.6925	92.6929	92.6225	92.6674	92.6848	92.6707	92.6906	92.6508
JHI1592	92.6616	92.7421	92.7253	92.7312	92.6863	92.6828	92.6814	92.6433	92.7116	92.6612
SM41	92.6202	92.6857	92.6542	92.6974	92.7185	92.7166	92.7747	92.7204	92.6293	92.5968
JHI1253	92.6957	92.7818	92.7253	92.7201	92.7761	92.7588	92.8312	92.7625	92.7002	92.651
JHI370	92.6673	92.7613	92.7277	92.7453	92.6861	92.672	92.7528	92.6393	92.7119	92.6283
JHI387	92.6849	92.7521	92.6621	92.7148	92.7153	92.6718	92.6977	92.6725	92.6682	92.6591
JHI388	92.6658	92.7682	92.7432	92.746	92.7495	92.673	92.7204	92.708	92.7413	92.7308
JHI788	92.7097	92.7667	92.7937	92.8183	92.791	92.8296	92.7977	92.7896	92.79	92.8121
JHI985	92.6738	92.8017	92.6829	92.7094	92.7109	92.7695	92.7616	92.705	92.735	92.6956
GB30	92.7364	92.7948	92.7387	92.8049	92.7675	92.7508	92.7876	92.8074	92.7085	92.77
JHI782	92.7138	92.7657	92.6978	92.7212	92.7484	92.7107	92.7415	92.7206	92.6765	92.6349
JHI783	92.7186	92.8385	92.7453	92.7405	92.721	92.7228	92.7505	92.7637	92.7627	92.6733
SM170C	92.6524	92.7293	92.6471	92.6224	92.6133	92.6432	92.6821	92.6438	92.6608	92.6093
JHI42	92.6698	92.7196	92.7136	92.7308	92.7231	92.7006	92.6928	92.6796	92.741	92.6945
JHI979	92.7301	92.7648	92.7635	92.7773	92.8115	92.7726	92.8152	92.8042	92.7564	92.72
JHI1259	92.689	92.738	92.6753	92.678	92.7324	92.7358	92.746	92.6895	92.682	92.67
TA1	92.7194	92.7998	92.654	92.6976	92.743	92.6855	92.7604	92.7502	92.7203	92.7447
JHI24	92.6746	92.7903	92.7629	92.7665	92.6722	92.7594	92.7112	92.666	92.7399	92.6785
JHI787	92.734	92.7903	92.765	92.7946	92.6651	92.6919	92.7004	92.6657	92.725	92.7003
RCAM1026	92.7551	92.7991	92.7681	92.816	92.8224	92.7352	92.8545	92.8676	92.8307	92.8171
Vh3	92.6864	92.7335	92.7431	92.7762	92.7174	92.7215	92.8051	92.7511	92.7663	92.7563
JHI925	92.7158	92.7819	93.0067	92.9881	92.814	92.7454	92.7576	92.7521	92.7284	92.8288
Vc2	92.707	92.7622	92.7545	92.7457	92.712	92.7406	92.7687	92.7875	92.6874	92.7097
JHI1422	92.6798	92.8019	92.6644	92.7074	92.7392	92.7683	92.7931	92.7551	92.7361	92.8079
JHI1266	92.7339	92.8395	92.7732	92.7572	92.8454	92.7794	92.862	92.8201	92.7579	92.7482
51A11	94.1684	94.2268	94.2467	94.2499	94.2357	94.2555	94.2405	94.2572	94.3226	94.2332
51B21	94.2279	94.2342	94.2163	94.2875	94.3143	94.3133	94.309	94.2884	94.3178	94.2728
128C53	92.1957	92.2268	92.2341	92.225	92.2424	92.1972	92.2288	92.2572	92.1955	92.128
4292	92.1212	92.2152	92.2602	92.2944	92.194	92.1926	92.2098	92.1566	92.1981	92.2483
CC283bq	92.174	92.2226	92.124	92.139	92.2276	92.1394	92.2231	92.1819	92.139	92.1305
USDA2370	92.1966	92.2617	92.2294	92.2335	92.1637	92.1602	92.262	92.2286	92.2333	92.2183
JHI2442	92.3251	92.3412	92.3285	92.2969	92.3213	92.228	92.3554	92.2787	92.3173	92.2644
JHI1084	92.2251	92.3126	92.2261	92.2677	92.1979	92.2182	92.2444	92.205	92.2863	92.107
JHI944	92.246	92.3103	92.4847	92.5152	92.3179	92.2964	92.3545	92.3595	92.349	92.473
OV152	92.1332	92.2056	92.1999	92.2146	92.1462	92.198	92.162	92.156	92.2094	92.1881
UPM1137	92.1908	92.3096	92.3557	92.2982	92.2628	92.2977	92.388	92.268	92.2582	92.2349
SM149A	92.1759	92.2958	92.2455	92.24	92.1856	92.1655	92.1717	92.1689	92.1871	92.157
Rt24.2	92.1945	92.2898	92.2056	92.2816	92.2325	92.2276	92.2198	92.2405	92.2321	92.2712
CCBAU23252	90.7156	90.797	90.7567	90.8082	90.7742	90.7437	90.7623	90.7797	90.7109	90.7421
JHI536	90.8453	90.8694	90.8317	90.8334	90.8206	90.8245	90.8081	90.7865	90.789	90.7192
C15	90.7953	90.8527	90.8259	90.8329	90.7648	90.8003	90.762	90.7755	90.8195	90.6916
CF307	90.7911	90.8536	90.7788	90.8267	90.7711	90.8053	90.7929	90.776	90.6971	90.8103

Table S4.5. (Continuation)

Query	93B11	111A12	VFSR2B1	SPF4F7	FB206	SL16	HUTR05	21B12	WYCCWR11290
3841	94.2952	94.4129	94.3553	94.4186	94.3139	94.1871	94.1848	94.2788	93.75
SM38	94.2732	94.2868	94.2418	94.2099	94.1677	94.1303	94.2603	94.2224	93.773
JHI960	94.3008	94.3617	94.3255	94.3424	94.3142	94.1495	94.212	94.2951	93.6892
JHI963	94.2889	94.3582	94.3111	94.3946	94.3011	94.2117	94.1479	94.3279	93.7784
JHI1587	94.2962	94.3987	94.3476	94.2997	94.2447	94.2021	94.1259	94.2905	93.7183
JHI1600	94.2951	94.4121	94.4395	94.3975	94.2987	94.2152	94.2449	94.284	93.7455
JHI13	94.2452	94.3036	94.2309	94.2325	94.1148	94.173	94.195	94.177	93.7879
JHI1438	94.1434	94.2213	94.2354	94.0929	94.0733	94.0999	94.162	94.0907	93.8043
JHI535	94.2566	94.3184	94.2929	94.3317	94.3763	94.1596	94.1673	94.3176	93.836
JHI585	94.2926	94.3461	94.3031	94.3143	94.3604	94.2297	94.2164	94.3276	93.8252
JHI973	94.2504	94.2694	94.2911	94.3265	94.2896	94.1374	94.2455	94.3022	93.7836
JHI974	94.2572	94.3289	94.2986	94.3321	94.3523	94.1887	94.1926	94.3027	93.7348
JHI1415	94.263	94.3247	94.3129	94.3725	94.3451	94.1887	94.1824	94.2914	93.797
SM3	94.1245	94.1886	94.1483	94.1957	94.0827	94.0866	94.177	94.1412	93.7024
VF39	94.208	94.2613	94.2708	94.2076	94.2616	94.0888	94.1493	94.2185	93.711
L361	95.3285	95.4304	95.4151	95.4002	95.3622	95.3532	95.3687	95.2739	93.7625
VFF1R2B1	94.2492	94.3077	94.3217	94.2836	94.2114	94.2512	94.2922	94.3011	93.7845
21A12	94.3258	94.3122	94.3433	94.2307	94.2495	94.2822	94.2335	94.3598	93.7285
VFF1R2A1	94.2997	94.3171	94.3583	94.2747	94.2264	94.259	94.2284	94.3503	93.7443
WSM1455	94.1749	94.3219	94.25	94.1774	94.1553	94.1655	94.1745	94.224	93.6999
WSM1481	94.1876	94.2319	94.2132	94.159	94.132	94.094	94.1424	94.1747	93.7397
WSM78	92.8953	92.891	92.8579	92.8839	92.8188	92.94	92.9728	92.9031	92.9143
SM152B	93.0093	93.0318	93.0401	92.9237	92.8485	92.9632	93.0248	92.9754	92.9568
CC275e	93.0934	93.0246	93.0428	93.004	92.9698	92.9786	93.0206	93.0234	93.0175
WSM1325	93.2728	93.2682	93.231	93.2313	93.2054	93.1674	93.2698	93.2814	93.1254
SRDI943	93.3997	93.3204	93.3228	93.3926	93.2281	93.2563	93.3211	93.3376	93.0336
JHLM12A2	93.9641	94.0203	94.0242	93.9395	93.9031	93.8985	93.9266	93.9131	94.2131
Vaf12	94.528	94.4764	94.4949	94.4216	94.4459	94.419	94.5448	94.4635	93.9253
JHI2451	94.5613	94.4868	94.4555	94.4077	94.4048	94.3804	94.4761	94.5234	93.9835
JHI10	94.4522	94.4412	94.4371	94.3976	94.3216	94.4061	94.4131	94.4442	93.9007
JHI54	94.3769	94.4061	94.3583	94.4072	94.3016	94.3531	94.3963	94.335	93.8216
FA23	94.4014	94.4383	94.386	94.4265	94.3635	94.3921	94.4433	94.4144	93.9806
JHI2450	94.5569	94.5295	94.5417	94.4887	94.492	94.4895	94.4986	94.522	94.0114
Vaf-108	95.1954	95.1798	95.1324	95.1442	95.0768	95.108	95.1641	95.1242	93.6062
JHI2449	95.4522	95.4929	95.4563	95.4052	95.3825	95.346	95.4716	95.5295	93.9437
UPM1131	95.3895	95.4235	95.3586	95.3359	95.2606	95.2484	95.3977	95.3735	93.7691
CCBAU10279	96.5133	96.5135	96.5339	96.5169	96.5293	96.4756	96.575	96.4869	93.6727
WSM1689	95.8689	95.9169	95.9474	95.9389	95.8978	95.8155	95.9278	95.9196	93.5538
VFF2R2A1	96.7021	96.7966	96.7907	96.7756	96.7884	96.6596	96.7164	96.682	93.638
SPF2A11	96.757	96.8328	96.8662	96.8173	96.7242	96.7574	96.7355	96.8223	93.6021
VFCR2A2	96.7329	96.8323	96.8299	96.889	96.808	96.6844	96.7826	96.6728	93.6147
248	95.5347	95.6005	95.564	95.5549	95.4975	95.4506	95.471	95.5478	93.5628
63A21	95.3275	95.3353	95.3365	95.3545	95.3408	95.3125	95.2829	95.3015	93.6554
GLR2	95.4465	95.5421	95.5045	95.3911	95.4534	95.4584	95.4887	95.3786	93.5307
CL8	98.1759	98.3107	98.2638	98.1571	98.2237	98.3026	98.2784	98.156	93.6203
VFSR2A2	98.5574	98.3057	98.3132	98.2847	98.3883	98.5788	98.3143	98.3563	93.5718
SS21	98.4779	98.2092	98.1553	98.1573	98.1521	98.7412	98.2044	98.2011	93.534
41A11	98.6122	98.2369	98.2398	98.3514	98.2974	98.6248	98.2048	98.3458	93.6035
42B12	98.6804	98.2926	98.2995	98.4021	98.3318	98.6742	98.2931	98.3587	93.687
43A11	98.6342	98.2506	98.2188	98.3408	98.2621	98.6135	98.2497	98.3304	93.6694
43B11	98.615	98.3588	98.344	98.3624	98.3779	98.7017	98.3453	98.3671	93.58
43B12	98.6313	98.3538	98.355	98.3591	98.3801	98.7124	98.3717	98.4212	93.5882
71A12	98.6238	98.4853	98.4772	98.3691	98.445	98.7785	98.3757	98.515	93.5957
73A11	98.5514	98.5672	98.5891	98.3312	98.4952	98.7981	98.3555	98.5481	93.6581
81B22	98.5983	98.4688	98.5018	98.3542	98.4381	98.8276	98.3555	98.5039	93.5281
83A12	98.5927	98.478	98.4786	98.2979	98.4797	98.8097	98.3738	98.5189	93.6328
VFHR1A2	98.64	98.3094	98.292	98.2696	98.33	98.7038	98.3174	98.3475	93.6005
VFF1R1A2	99.7543	98.3173	98.3601	98.2863	98.3472	98.345	98.3437	98.5815	93.5808
93B11	100	98.3544	98.3636	98.2841	98.3683	98.4412	98.3535	98.6335	93.6322
111A12	98.3953	100	99.9802	98.4495	98.4938	98.4107	98.3342	98.4917	93.585
VFSR2B1	98.3469	99.9847	100	98.4885	98.4738	98.4011	98.2846	98.4978	93.6206
SPF4F7	98.2847	98.4354	98.4224	100	98.4533	98.2842	98.382	98.384	93.6207
FB206	98.3362	98.4965	98.4934	98.438	100	98.4884	98.4086	98.449	93.5369
SL16	98.3927	98.3722	98.3779	98.2971	98.4824	100	98.3313	98.2946	93.537
HUTR05	98.3011	98.3366	98.3021	98.3518	98.3928	98.3141	100	98.635	93.5798
21B12	98.6524	98.5088	98.5002	98.3964	98.4777	98.3244	98.6641	100	93.5881
WYCCWR11290	93.6283	93.6173	93.5824	93.5797	93.5228	93.5139	93.5674	93.5794	100
WYCCWR11317	93.6202	93.6146	93.6266	93.5985	93.5872	93.5575	93.5693	93.5795	98.0588
CCBAU11080	93.4476	93.525	93.5699	93.5353	93.4077	93.4616	93.5013	93.4843	95.6116
CCBAU03386	93.5102	93.5529	93.5436	93.4545	93.4769	93.4828	93.4431	93.5263	95.5162

Table S4.5. (Continuation)

Query	93B11	111A12	VFSR2B1	SPF4F7	FB206	SL16	HUTR05	21B12	WYCCWR11290
WYCCWR10014	93.429	93.4367	93.4173	93.3943	93.3741	93.4078	93.4996	93.4629	93.5888
JKLM13E	93.9272	93.9616	93.9928	93.8876	93.8599	93.9276	93.8918	93.9676	94.193
Tri-43	93.9043	93.9101	93.914	93.8169	93.8895	93.8734	93.8592	93.8296	94.6463
73B11	95.6071	95.5854	95.573	95.5902	95.5816	95.5055	95.5528	95.5951	93.7703
73B12	95.6474	95.6375	95.598	95.5954	95.6245	95.5675	95.5889	95.6067	93.8116
RSP1F2	95.5662	95.5538	95.5649	95.4763	95.4939	95.4514	95.5002	95.4939	93.9485
Vaf10	95.0624	95.0924	95.0916	95.0491	95.0233	94.9356	95.0487	94.9689	93.6384
PEPV16	95.5178	95.5248	95.535	95.5066	95.4826	95.4514	95.4579	95.5422	93.9332
TOM	95.4798	95.4806	95.4584	95.3858	95.3869	95.4075	95.3957	95.4638	93.8621
121B21	95.5109	95.4607	95.4453	95.4932	95.4964	95.5173	95.4491	95.487	93.7893
RSP1F10	95.4909	95.5297	95.5286	95.3859	95.3976	95.4376	95.4837	95.4812	93.9033
RSP1A1	95.5229	95.449	95.4873	95.4379	95.3993	95.3667	95.4831	95.46	93.9048
Norway	92.6802	92.6772	92.6971	92.6889	92.6318	92.6115	92.6278	92.6272	92.8706
CC278f	92.5454	92.5174	92.4874	92.5438	92.4606	92.3342	92.5451	92.4822	92.6108
SM78	92.5934	92.5705	92.5771	92.5528	92.449	92.4777	92.6436	92.4992	92.7425
SM51	92.546	92.5547	92.5289	92.5875	92.5025	92.477	92.5418	92.4956	92.7429
SM72	92.6178	92.6303	92.6482	92.5994	92.5119	92.5522	92.6175	92.5641	92.8304
Vaf-46	92.7527	92.7273	92.7583	92.7823	92.6816	92.6711	92.7578	92.6794	92.4892
JHI1093	92.7617	92.83	92.8317	92.7552	92.6871	92.6818	92.8043	92.6867	92.5066
JHI1096	92.75	92.7744	92.8132	92.7294	92.7357	92.6648	92.7669	92.724	92.4692
GPTR08	92.4405	92.4014	92.3666	92.3299	92.3385	92.3619	92.4125	92.3193	91.9608
GPTR02	92.4791	92.4223	92.4267	92.4645	92.3982	92.4298	92.532	92.4304	91.9737
SRDI565	92.4304	92.4422	92.3627	92.3991	92.3136	92.3939	92.505	92.4422	92.0926
Ps8	92.8023	92.7749	92.8349	92.7642	92.7754	92.6613	92.7268	92.7543	92.6377
JHI1236	92.8045	92.7818	92.7879	92.7807	92.702	92.6653	92.6883	92.7479	92.6395
JHI953	92.8204	92.7749	92.8304	92.8042	92.6166	92.6515	92.6833	92.6858	92.555
SM147A	92.7595	92.7872	92.7555	92.7215	92.6647	92.6722	92.738	92.6973	92.6557
JHI1238	92.7253	92.7767	92.7859	92.743	92.6417	92.728	92.67	92.6898	92.6281
UPM1133	92.6563	92.6387	92.6426	92.6813	92.6004	92.5985	92.6652	92.6612	92.6688
JHI1592	92.6932	92.7063	92.7215	92.6687	92.5778	92.689	92.6359	92.6866	92.5728
SM41	92.7115	92.7254	92.775	92.6985	92.6012	92.6769	92.7324	92.6594	92.6329
JHI1253	92.698	92.7331	92.7393	92.6976	92.6745	92.5805	92.7195	92.7036	92.6707
JHI370	92.6931	92.749	92.761	92.667	92.5837	92.7052	92.6463	92.6636	92.5391
JHI387	92.7401	92.6742	92.6643	92.6953	92.533	92.6589	92.6287	92.6692	92.5785
JHI388	92.7677	92.7481	92.7373	92.7293	92.5555	92.6545	92.681	92.7332	92.6395
JHI788	92.8291	92.7992	92.7992	92.8461	92.8039	92.7451	92.7782	92.7762	92.6867
JHI985	92.7637	92.7766	92.7501	92.7509	92.6341	92.6696	92.725	92.7051	92.5964
GB30	92.7771	92.6707	92.7154	92.715	92.6761	92.7133	92.7413	92.7448	92.6695
JHI782	92.6406	92.718	92.6759	92.6738	92.6503	92.6651	92.6564	92.6459	92.5924
JHI783	92.7338	92.7873	92.7811	92.7616	92.6771	92.7071	92.7168	92.6822	92.5834
SM170C	92.7133	92.713	92.6639	92.6703	92.5159	92.597	92.6687	92.6971	92.6044
JHI42	92.8298	92.7485	92.7829	92.7855	92.5773	92.6574	92.7226	92.7132	92.6575
JHI979	92.8305	92.7402	92.7206	92.7776	92.6764	92.6419	92.6552	92.7397	92.6088
JHI1259	92.7195	92.7101	92.7408	92.7118	92.7105	92.6395	92.7345	92.6997	92.6367
TA1	92.7715	92.8124	92.8035	92.7914	92.6538	92.6708	92.7619	92.7489	92.6729
JHI24	92.7803	92.7372	92.7735	92.7287	92.5703	92.7111	92.7181	92.7042	92.6415
JHI787	92.8115	92.7161	92.771	92.7787	92.6574	92.739	92.7736	92.7513	92.6546
RCAM1026	92.7798	92.7996	92.8223	92.8625	92.7994	92.7883	92.805	92.7809	92.7209
Vh3	92.8011	92.7839	92.7246	92.8054	92.7473	92.7503	92.7592	92.7606	92.6316
JHI925	92.7985	92.7375	92.7244	92.7968	92.8113	92.6473	92.7081	92.8708	92.5331
Vc2	92.7563	92.807	92.762	92.7521	92.6601	92.6779	92.764	92.7369	92.6376
JHI1422	92.8049	92.76	92.7674	92.7175	92.7691	92.7076	92.6791	92.7498	92.6006
JHI1266	92.7864	92.8084	92.796	92.8094	92.7626	92.6688	92.6946	92.818	92.6292
51A11	94.2592	94.3196	94.363	94.2572	94.2502	94.2513	94.2675	94.3269	93.6769
51B21	94.2982	94.325	94.2952	94.2386	94.2207	94.2602	94.2804	94.306	93.7146
128C53	92.3071	92.2569	92.2774	92.2854	92.1946	92.1307	92.2023	92.204	92.5284
4292	92.2201	92.2083	92.2463	92.2647	92.1137	92.1484	92.2217	92.2119	92.2848
CC283bq	92.1601	92.1759	92.2006	92.1602	92.1384	92.0425	92.1664	92.107	92.3335
USDA2370	92.2024	92.221	92.2352	92.2434	92.2136	92.1391	92.1962	92.2641	92.4141
JHI2442	92.2928	92.3049	92.298	92.297	92.2293	92.1823	92.2154	92.2467	92.4328
JHI1084	92.2798	92.2283	92.2372	92.215	92.2526	92.145	92.2172	92.2373	92.4168
JHI944	92.3914	92.3278	92.2549	92.4496	92.4248	92.189	92.2506	92.379	92.3193
OV152	92.2662	92.2198	92.2179	92.2563	92.0896	92.2123	92.2629	92.2523	92.4108
UPM1137	92.2474	92.3013	92.2958	92.2636	92.2	92.2705	92.3044	92.2995	92.4198
SM149A	92.175	92.2369	92.196	92.2022	92.2159	92.0768	92.2456	92.1975	92.355
Rt24.2	92.2872	92.2623	92.2317	92.2539	92.1817	92.1746	92.2727	92.2644	92.325
CCBAU23252	90.8335	90.7723	90.7839	90.8062	90.6938	90.861	90.7133	90.7656	91.2081
JHI536	90.8773	90.8015	90.8307	90.8273	90.8134	90.7949	90.75	90.7845	91.4109
C15	90.7895	90.8357	90.8292	90.7971	90.7442	90.7176	90.7253	90.7394	91.2165
CF307	90.8812	90.7507	90.7655	90.7971	90.7768	90.8198	90.7672	90.7461	91.2645

Table S4.5. (Continuation)

Query	WYCCWR11317	CCBAU11080	CCBAU03386	WYCCWR10014	JKLM13E	Tri-43	73B11	73B12
3841	93.7893	93.7777	93.7076	94.1327	94.6158	94.2004	94.5017	94.4997
SM38	93.8536	93.6899	93.8117	94.1667	94.6993	94.2605	94.4541	94.4359
JHI960	93.7718	93.8274	93.7168	94.066	94.5458	94.2398	94.4807	94.3815
JHI963	93.8672	93.8077	93.7611	94.1278	94.5995	94.2117	94.493	94.4744
JHI1587	93.8098	93.6751	93.6869	94.0575	94.6579	94.1969	94.4273	94.363
JHI1600	93.8138	93.7739	93.8022	94.0493	94.6316	94.2059	94.4659	94.4076
JHI13	93.822	93.7082	93.7575	94.0865	94.6362	94.1591	94.4547	94.4484
JHI1438	93.8377	93.7612	93.7482	94.0947	94.6695	94.166	94.4102	94.3843
JHI535	93.9008	93.794	93.8194	94.2044	94.6681	94.2205	94.5125	94.5039
JHI585	93.9131	93.8131	93.8712	94.1876	94.7313	94.2432	94.535	94.4758
JHI973	93.802	93.7657	93.8344	94.1449	94.6684	94.2504	94.4662	94.4556
JHI974	93.7867	93.7445	93.8661	94.1337	94.6738	94.1733	94.4755	94.4594
JHI1415	93.9113	93.7828	93.8825	94.1497	94.6943	94.2249	94.5172	94.4652
SM3	93.8799	93.7195	93.8125	94.096	94.5947	94.1645	94.402	94.3244
VF39	93.7958	93.7058	93.7725	94.0835	94.6742	94.1931	94.435	94.4233
L361	93.8187	93.574	93.7083	93.7536	94.2	94.1364	95.7287	95.7229
VFF1R2B1	93.8601	93.7212	93.7798	94.1438	94.7162	94.1692	94.4797	94.4329
21A12	93.8538	93.8135	93.7652	94.1444	94.6747	94.2849	94.4698	94.441
VFF1R2A1	93.8275	93.7815	93.766	94.1154	94.6727	94.2356	94.4968	94.4356
WSM1455	93.782	93.7328	93.6773	94.0676	94.6632	94.1642	94.4634	94.4435
WSM1481	93.7797	93.7193	93.6779	94.038	94.7794	94.1762	94.4233	94.4024
WSM78	93.0028	92.9411	92.942	94.2727	93.6378	93.427	93.1296	93.1281
SM152B	93.0393	92.9509	92.9673	94.1167	93.6458	93.488	93.2567	93.2194
CC275e	93.0538	92.9322	92.9586	94.0158	93.8734	93.4418	93.3134	93.281
WSM1325	93.2078	93.0023	93.0691	93.7055	94.0522	93.6629	93.5031	93.4484
SRDI943	93.2261	93.0421	93.0482	93.7727	94.1085	93.6158	93.5616	93.5383
JHLM12A2	94.1982	94.3727	94.0302	94.2849	98.0014	94.7316	94.1841	94.1772
Vaf12	94.0174	93.794	93.8553	94.2514	95.0725	94.4423	94.7299	94.7155
JHI2451	93.9891	93.9227	93.8388	94.3738	95.0267	94.3631	94.7439	94.692
JHI10	94.0106	93.8966	93.8068	94.3051	95.0249	94.4456	94.6603	94.6061
JHI54	93.9242	93.7088	93.7685	94.2426	94.9494	94.3906	94.6155	94.5484
FA23	93.9946	93.9659	93.8662	94.4257	95.0646	94.444	94.8031	94.7348
JHI2450	94.071	94.0116	93.9462	94.3846	95.1895	94.4842	94.7355	94.7228
Vaf-108	93.7227	93.3704	93.5694	93.5068	94.0784	94.0594	95.472	95.4834
JHI2449	93.8535	93.7063	93.5539	93.6622	94.274	94.1441	95.7872	95.7164
UPM1131	93.8335	93.6391	93.6368	93.7822	94.24	94.1847	95.8224	95.7832
CCBAU10279	93.6497	93.5077	93.5227	93.5163	93.9898	93.8991	95.6009	95.5791
WSM1689	93.6543	93.4784	93.5106	93.6272	94.046	93.9397	94.9574	94.9339
VFF2R2A1	93.6124	93.4216	93.5192	93.4674	93.9983	93.9637	95.1585	95.1458
SPF2A11	93.5508	93.4185	93.4795	93.4563	93.9357	93.8545	95.1921	95.1489
VFCR2A2	93.5937	93.3991	93.496	93.4184	94.0068	93.9282	95.1872	95.1617
248	93.6321	93.5122	93.4724	93.6135	94.1039	93.9066	94.9557	94.9238
63A21	93.6578	93.5334	93.5752	93.622	94.0743	94.0288	94.7796	94.7551
GLR2	93.5931	93.3031	93.3955	93.4595	94.0072	93.8873	94.7452	94.7789
CL8	93.5993	93.4967	93.5351	93.5637	93.9929	93.9633	95.5107	95.5358
VFSR2A2	93.6152	93.526	93.5606	93.4425	93.968	93.9315	95.5316	95.4854
SS21	93.5725	93.456	93.4663	93.4329	93.9092	93.9471	95.5033	95.5366
41A11	93.6933	93.5478	93.5305	93.452	94.0322	93.9837	95.628	95.6
42B12	93.7275	93.6138	93.6344	93.6037	94.1168	94.0854	95.6876	95.6734
43A11	93.6876	93.5255	93.5518	93.4961	94.0252	93.9502	95.589	95.5968
43B11	93.6307	93.5544	93.5892	93.5635	93.9954	93.9303	95.4893	95.4304
43B12	93.6419	93.5768	93.5719	93.5486	94.0196	93.9708	95.5203	95.4816
71A12	93.5416	93.5338	93.5714	93.4824	94.0275	93.9565	95.5906	95.6376
73A11	93.5715	93.5288	93.5317	93.4917	94.0125	93.8884	95.5892	95.6173
81B22	93.5635	93.613	93.5415	93.5525	94.0325	93.9889	95.5947	95.5966
83A12	93.6065	93.599	93.5755	93.4476	94.0301	93.9444	95.6074	95.6141
VFHR1A2	93.6209	93.4853	93.6592	93.5107	93.9976	93.92	95.5994	95.5676
VFF1R1A2	93.5977	93.3708	93.5141	93.4812	93.9825	93.9163	95.5793	95.5451
93B11	93.6336	93.4913	93.5426	93.5466	93.959	93.9516	95.591	95.5984
111A12	93.6018	93.5657	93.5911	93.4571	94.0909	94.0191	95.5403	95.5453
VFSR2B1	93.5811	93.5419	93.5544	93.4336	94.0211	94.0104	95.5723	95.5484
SPF4F7	93.6012	93.5749	93.5151	93.4515	93.8639	93.925	95.5627	95.537
FB206	93.5623	93.399	93.4327	93.4659	93.9188	93.9272	95.5304	95.5628
SL16	93.5963	93.4668	93.4711	93.5268	93.968	93.9737	95.5254	95.5159
HUTR05	93.5753	93.4925	93.5709	93.5504	93.9188	94.0094	95.5174	95.5153
21B12	93.5689	93.5062	93.4749	93.5353	94.0418	93.9794	95.5382	95.548
WYCCWR11290	98.0525	95.6157	95.5096	93.5771	94.1922	94.6135	93.7814	93.7831
WYCCWR11317	100	95.5141	95.4854	93.7098	94.1865	94.755	93.7819	93.7874
CCBAU11080	95.3909	100	98.6635	93.5768	94.3413	94.3628	93.7881	93.6945
CCBAU03386	95.4271	98.694	100	93.6323	94.0716	94.353	93.6946	93.7076

Table S4.5. (Continuation)

Query	WYCCWR11317	CCBAU11080	CCBAU03386	WYCCWR10014	JKLM13E	Tri-43	73B11	73B12
WYCCWR10014	93.6001	93.4677	93.5433	100	94.2288	94.1625	93.6851	93.7326
JKLM13E	94.1384	94.3621	94.1116	94.221	100	94.7253	94.2214	94.1553
Tri-43	94.7203	94.3924	94.3816	94.1652	94.751	100	94.155	94.1606
73B11	93.7609	93.7553	93.7355	93.7702	94.2905	94.1946	100	99.9931
73B12	93.8026	93.8379	93.8079	93.8252	94.2717	94.2127	99.996	100
RSP1F2	93.8811	93.8014	93.7796	93.7966	94.3498	94.2404	97.3172	97.2614
Vaf10	93.7306	93.4811	93.6327	93.5235	94.182	94.1406	95.5055	95.5364
PEPV16	93.7227	93.7473	93.6384	93.6038	94.2388	94.1811	97.4551	97.4473
TOM	93.8124	93.7641	93.67	93.6619	94.2064	94.1801	97.5999	97.554
121B21	93.8082	93.5929	93.6717	93.7295	94.1196	94.1315	97.9164	97.9363
RSP1F10	93.886	93.8285	93.7322	93.8585	94.2824	94.222	97.4545	97.4693
RSP1A1	93.8719	93.8836	93.766	93.8193	94.2896	94.2298	97.3452	97.3359
Norway	92.8634	92.7577	92.8265	93.6134	93.343	93.2307	92.9216	92.9013
CC278f	92.7675	92.5812	92.7045	93.8111	93.2964	93.1241	92.7808	92.7009
SM78	92.8634	92.5878	92.7032	93.559	93.1271	93.1972	92.7867	92.7078
SM51	92.8565	92.6195	92.6876	93.5453	93.1711	93.0627	92.7407	92.6887
SM72	92.8894	92.7337	92.7832	93.6104	93.2406	93.2431	92.8367	92.7369
Vaf-46	92.4824	92.258	92.3257	92.6235	92.8563	92.7672	92.9038	92.8268
JHI1093	92.4365	92.3721	92.3654	92.7828	92.8991	92.8322	92.8913	92.8434
JHI1096	92.391	92.3926	92.3844	92.7714	92.8524	92.7869	92.912	92.8871
GPTR08	91.9726	91.9235	91.922	92.2949	92.2762	92.348	92.4839	92.4312
GPTR02	92.0516	91.8814	91.9361	92.244	92.2728	92.4083	92.5642	92.5231
SRDI565	92.1028	91.9644	92.057	92.3346	92.3997	92.5261	92.4939	92.4075
Ps8	92.6339	92.5581	92.489	93.1342	93.1192	92.9955	92.967	92.9128
JHI1236	92.6702	92.5601	92.5503	93.1298	93.2435	92.9707	92.9038	92.874
JHI953	92.6852	92.5237	92.4802	93.0398	93.0993	92.9933	92.9012	92.882
SM147A	92.6555	92.5413	92.5176	93.0873	93.164	93.0186	92.8956	92.8295
JHI1238	92.6801	92.5351	92.5476	93.1079	93.2067	92.934	92.9654	92.9555
UPM1133	92.6374	92.5103	92.4433	93.0361	93.1429	92.9255	92.85	92.8413
JHI1592	92.6738	92.5306	92.6062	93.0676	93.1352	93.0166	92.8714	92.8335
SM41	92.661	92.5305	92.6357	93.0996	93.1793	92.972	92.9018	92.8229
JHI1253	92.6998	92.5525	92.4888	93.0934	93.1712	92.9453	92.9669	92.8945
JHI370	92.6204	92.4909	92.5511	93.0172	93.1929	92.9943	92.9128	92.8326
JHI387	92.6563	92.4813	92.5924	93.0689	93.1432	92.9493	92.8652	92.8636
JHI388	92.6531	92.5001	92.5492	93.0121	93.1192	92.9396	92.8598	92.8387
JHI788	92.6973	92.6679	92.6176	93.1391	93.2207	93.0491	92.9907	92.9757
JHI985	92.6388	92.5105	92.4444	93.0734	93.1497	92.9287	92.9671	92.8957
GB30	92.643	92.4728	92.4748	93.1152	93.1546	92.9807	93.0271	92.9755
JHI782	92.6125	92.5176	92.527	93.0309	93.1443	92.9485	92.8304	92.8192
JHI783	92.6262	92.5745	92.5029	93.0964	93.1501	92.9464	92.8788	92.8437
SM170C	92.6612	92.4179	92.5423	93.0923	93.0101	92.9867	92.8607	92.7994
JHI42	92.6831	92.4072	92.4728	92.9809	93.177	92.9703	92.8022	92.8061
JHI979	92.6874	92.5533	92.5078	93.0711	93.0982	92.8846	92.9018	92.9345
JHI1259	92.6483	92.6061	92.4923	93.0358	93.1897	92.9256	92.8343	92.8488
TA1	92.736	92.4687	92.5218	93.1214	93.1891	92.9787	92.9372	92.8838
JHI24	92.6543	92.5161	92.4991	93.002	93.1449	92.9612	92.8744	92.8524
JHI787	92.6307	92.5594	92.4821	93.0669	93.2135	92.9629	92.9083	92.845
RCAM1026	92.7819	92.6333	92.5882	93.1535	93.3034	93.0016	92.9924	92.987
Vh3	92.6346	92.5382	92.4985	93.0217	93.1896	93.0028	92.9019	92.8823
JHI925	92.5774	92.5313	92.4719	92.9366	93.0449	92.8428	92.8697	92.8215
Vc2	92.6641	92.5553	92.5424	92.9393	93.1729	92.9299	92.8653	92.8335
JHI1422	92.6467	92.5298	92.5646	93.0134	93.1814	92.886	92.9378	92.9126
JHI1266	92.692	92.6876	92.5726	93.1007	93.2381	92.9158	92.9562	92.9723
51A11	93.7719	93.6846	93.7701	94.1004	94.611	94.2004	94.4117	94.3458
51B21	93.7847	93.7677	93.7737	94.0907	94.6448	94.225	94.4231	94.4018
128C53	92.4208	92.4397	92.4425	93.4209	92.8705	92.755	92.3645	92.3479
4292	92.4074	92.3407	92.3774	93.4804	92.7588	92.7051	92.4091	92.3964
CC283bq	92.378	92.3301	92.2928	93.4637	92.834	92.7201	92.3039	92.2376
USDA2370	92.376	92.4318	92.279	93.3691	92.9277	92.7127	92.4099	92.3994
JHI2442	92.3945	92.5251	92.3484	93.4209	92.9375	92.7753	92.408	92.395
JHI1084	92.3652	92.8644	92.4152	93.4478	93.2607	92.6797	92.3726	92.3823
JHI944	92.3582	92.3603	92.3939	93.451	92.8329	92.7518	92.4257	92.4675
OV152	92.5437	92.4338	92.5346	93.5202	92.863	92.8704	92.3333	92.3536
UPM1137	92.4114	92.4561	92.3788	93.3506	92.9684	92.7443	92.4214	92.4427
SM149A	92.3862	92.4208	92.3119	93.4256	92.8534	92.7205	92.298	92.3491
Rt24.2	92.416	92.4234	92.4435	93.517	92.8669	92.7507	92.3147	92.2812
CCBAU23252	91.1845	90.9759	90.9413	91.0336	91.1911	91.3091	90.9165	90.8165
JHI536	91.2934	91.147	91.1107	91.1033	91.2557	91.331	90.935	90.8727
C15	91.2105	91.0698	91.0203	91.1234	91.1576	91.3502	90.8871	90.7805
CF307	91.2257	91.09	91.1108	91.3207	91.2183	91.2972	90.9528	90.9083

Table S4.5. (Continuation)

Query	RSP1F2	Vaf10	PEPV16	TOM	121B21	RSP1F10	RSP1A1	Norway	CC278f	SM78
3841	94.3303	94.2216	94.2802	94.3396	94.427	94.4734	94.4312	93.2722	93.0122	93.2792
SM38	94.3715	94.2227	94.2621	94.4642	94.4303	94.3671	94.4052	93.3641	93.0795	93.4117
JHI960	94.3377	94.1393	94.2885	94.3208	94.3816	94.4074	94.3446	93.3017	92.9805	93.2559
JHI963	94.3149	94.2405	94.3651	94.3862	94.4366	94.4102	94.3886	93.3154	92.9518	93.2141
JHI1587	94.3333	94.2317	94.3585	94.3568	94.3738	94.4271	94.3089	93.2969	93.0558	93.2397
JHI1600	94.3759	94.2657	94.3227	94.2971	94.4456	94.4605	94.4584	93.2204	93.0329	93.2755
JHI13	94.4494	94.2266	94.3252	94.2742	94.426	94.6224	94.7158	93.4188	92.9715	93.2611
JHI1438	94.4299	94.1748	94.2693	94.2643	94.3423	94.6297	94.6654	93.3684	93.0175	93.187
JHI535	94.3534	94.369	94.3862	94.3845	94.519	94.4738	94.5531	93.3505	93.0872	93.3059
JHI585	94.4558	94.3754	94.3622	94.419	94.4923	94.5095	94.5675	93.3681	93.0747	93.276
JHI973	94.4195	94.377	94.359	94.4037	94.4592	94.4691	94.5946	93.3362	93.0369	93.2783
JHI974	94.3895	94.3115	94.379	94.3669	94.4531	94.4032	94.5629	93.302	93.0534	93.3101
JHI1415	94.4084	94.2872	94.3625	94.4008	94.456	94.4603	94.5668	93.3562	93.0734	93.2428
SM3	94.3498	94.1203	94.2269	94.2944	94.2997	94.2639	94.3872	93.319	92.9753	93.3514
VF39	94.3233	94.1917	94.3983	94.2879	94.3368	94.4252	94.3525	93.2357	92.9527	93.213
L361	95.7041	95.4576	95.6539	95.5984	95.7166	95.7145	95.7073	93.0444	92.7065	92.9223
VFF1R2B1	94.385	94.2893	94.4222	94.4053	94.5086	94.5253	94.5332	93.3166	93.0251	93.2249
21A12	94.4962	94.2987	94.3389	94.3767	94.5397	94.5037	94.4838	93.3538	93.1176	93.233
VFF1R2A1	94.5007	94.2954	94.3298	94.3586	94.5219	94.4819	94.4704	93.3596	93.1056	93.2051
WSM1455	94.3488	94.2623	94.2795	94.345	94.4155	94.4364	94.4435	93.328	92.9982	93.1304
WSM1481	94.3344	94.2177	94.3125	94.3156	94.3948	94.4147	94.4423	93.2203	92.9585	93.0672
WSM78	93.19	93.0988	93.0889	93.0719	93.1462	93.238	93.2024	94.0448	94.1342	93.9233
SM152B	93.2279	93.0623	93.1108	93.2108	93.1894	93.281	93.2554	94.0944	93.9715	93.8936
CC275e	93.2842	93.1294	93.1516	93.2545	93.2475	93.3288	93.2108	94.2219	93.8997	94.0136
WSM1325	93.4353	93.3932	93.3586	93.4538	93.4686	93.5557	93.4456	93.5292	93.1213	93.3623
SRDI943	93.4792	93.4702	93.4182	93.4814	93.5926	93.6227	93.5175	93.5689	93.1024	93.3143
JHLM12A2	94.2634	94.1949	94.1381	94.1181	94.1694	94.2702	94.3198	93.348	93.0922	93.2317
Vaf12	94.768	95.0468	94.5962	94.6246	94.6577	94.8025	94.7983	93.4578	93.1242	93.1848
JHI2451	94.786	94.6444	94.6041	94.7322	94.7635	94.847	94.8451	93.5801	93.1867	93.2582
JHI10	94.7345	94.463	94.658	94.5927	94.6378	94.903	94.8862	93.473	93.1143	93.21
JHI54	94.7008	94.306	94.5613	94.5881	94.4862	94.695	94.6342	93.4081	93.0301	93.193
FA23	94.7778	94.5351	94.6393	94.5679	94.7327	94.8649	94.8566	93.4924	93.1695	93.4154
JHI2450	94.9073	94.7573	94.7986	94.7716	94.7794	94.9145	94.8998	93.6597	93.1684	93.3916
Vaf-108	95.5845	96.2467	95.4162	95.365	95.4424	95.4509	95.4182	92.9579	92.3084	92.6252
JHI2449	95.7178	95.3932	95.6932	95.6535	95.6836	95.683	95.6555	92.8743	92.606	92.7757
UPM1131	95.7732	95.5214	95.6621	95.7319	95.7548	95.8114	95.6883	92.9718	92.5761	92.7487
CCBAU10279	95.6381	95.3597	95.6447	95.5256	95.5772	95.5661	95.5592	92.6993	92.4501	92.5097
WSM1689	94.7765	94.6365	94.8121	94.726	94.8449	94.8791	94.852	92.7673	92.5912	92.6567
VFF2R2A1	95.21	94.7733	95.0613	95.0544	95.0242	95.1157	95.117	92.6432	92.3878	92.5898
SPF2A11	95.0477	94.8327	95.0752	95.0032	95.1011	95.0873	95.1643	92.687	92.4366	92.6268
VFCR2A2	95.0004	94.714	95.0393	95.0799	94.9243	95.0649	95.0246	92.6451	92.3596	92.5926
248	94.8251	94.7402	94.7645	94.8564	94.9206	94.8562	94.8897	92.8052	92.6058	92.7551
63A21	94.6398	94.5671	94.5705	94.6224	94.621	94.6816	94.695	92.817	92.7112	92.7184
GLR2	94.7456	94.5511	94.6606	94.6674	94.6995	94.6333	94.687	92.7337	92.4981	92.6729
CL8	95.512	95.0259	95.4362	95.4019	95.4486	95.5165	95.4457	92.7101	92.4348	92.612
VFSR2A2	95.5113	95.0666	95.5023	95.4439	95.4586	95.5035	95.4667	92.6606	92.4416	92.55
SS21	95.4741	94.9935	95.377	95.3155	95.5827	95.3897	95.4259	92.6462	92.3711	92.4272
41A11	95.5582	95.126	95.5515	95.4585	95.54	95.5006	95.5144	92.678	92.4879	92.5839
42B12	95.6318	95.1178	95.5511	95.5216	95.6009	95.5898	95.5717	92.7329	92.5208	92.6334
43A11	95.5592	95.0809	95.4594	95.4407	95.5143	95.5458	95.4782	92.7177	92.4552	92.5655
43B11	95.4232	95.0066	95.5056	95.4044	95.481	95.4574	95.4527	92.666	92.4404	92.5729
43B12	95.4386	95.0354	95.4467	95.414	95.4658	95.4998	95.4126	92.6712	92.5201	92.582
71A12	95.5633	95.0854	95.5168	95.404	95.5248	95.5286	95.5155	92.7038	92.524	92.608
73A11	95.5548	95.1078	95.4802	95.4047	95.5177	95.5563	95.491	92.7423	92.5608	92.612
81B22	95.4982	95.0389	95.5062	95.4434	95.5168	95.4872	95.4555	92.7325	92.5264	92.5999
83A12	95.5325	95.0428	95.4698	95.393	95.5435	95.5125	95.5003	92.7331	92.586	92.5667
VFHR1A2	95.5413	94.9991	95.4964	95.4171	95.5446	95.5394	95.492	92.6494	92.4433	92.5948
VFF1R1A2	95.4761	94.9614	95.4498	95.3869	95.5063	95.4405	95.424	92.5797	92.4457	92.5033
93B11	95.6172	95.0804	95.4547	95.5128	95.5816	95.5547	95.4918	92.7184	92.3868	92.5015
111A12	95.6031	95.0907	95.4807	95.4889	95.5398	95.569	95.5448	92.6254	92.4707	92.58
VFSR2B1	95.5698	95.0861	95.4153	95.4447	95.4918	95.5558	95.499	92.6852	92.4429	92.5132
SPF4F7	95.478	94.9765	95.4581	95.3906	95.5152	95.4759	95.4889	92.6197	92.4074	92.511
FB206	95.4725	95.0047	95.4563	95.3142	95.4706	95.4815	95.4319	92.6279	92.3795	92.4723
SL16	95.491	94.9815	95.4303	95.3829	95.5633	95.4774	95.4371	92.6576	92.3161	92.4838
HUTR05	95.542	95.1	95.4399	95.3917	95.4699	95.4717	95.4443	92.6821	92.4751	92.623
21B12	95.4932	95.0274	95.5141	95.4107	95.4574	95.5014	95.4655	92.6721	92.4416	92.6423
WYCCWR11290	93.9204	93.6936	93.783	93.8339	93.7303	93.9766	93.9221	92.8115	92.5567	92.7887
WYCCWR11317	93.9677	93.8317	93.8437	93.7445	93.8102	93.9111	93.9373	92.9727	92.7827	92.8829
CCBAU11080	93.7953	93.4818	93.659	93.7063	93.5852	93.848	93.8105	92.7573	92.5505	92.6506
CCBAU03386	93.8215	93.6349	93.5964	93.6555	93.601	93.7923	93.7864	92.7378	92.6407	92.7491

Table S4.5. (Continuation)

Query	RSP1F2	Vaf10	PEPV16	TOM	121B21	RSP1F10	RSP1A1	Norway	CC278f	SM78
WYCCWR10014	93.756	93.549	93.6261	93.6352	93.6546	93.8414	93.7534	93.5428	93.704	93.5433
JKLM13E	94.3088	94.2159	94.1776	94.1502	94.1329	94.2831	94.3333	93.3495	93.1497	93.1399
Tri-43	94.2806	94.1389	94.1729	94.1158	94.1946	94.2684	94.2886	93.2774	93.114	93.1602
73B11	97.2902	95.5663	97.4593	97.5776	97.9582	97.4849	97.3512	92.9178	92.6171	92.822
73B12	97.33	95.6003	97.4847	97.6528	97.9534	97.5145	97.3652	92.9741	92.7322	92.9175
RSP1F2	100	95.6243	97.1031	97.4674	97.0638	97.5674	97.7104	93.0843	92.6626	92.8607
Vaf10	95.5488	100	95.3931	95.4663	95.4091	95.4906	95.489	92.9354	92.3682	92.723
PEPV16	97.2433	95.4489	100	97.5037	97.4523	97.381	97.1727	92.9192	92.5179	92.7098
TOM	97.4109	95.472	97.4464	100	97.6651	97.6847	97.3707	92.8303	92.5538	92.7713
121B21	97.1096	95.4502	97.4165	97.6462	100	97.3989	97.2693	92.8425	92.5592	92.6795
RSP1F10	97.6023	95.5221	97.3484	97.7375	97.3733	100	97.6962	92.9693	92.6304	92.7811
RSP1A1	97.7378	95.4786	97.1567	97.3997	97.2108	97.7677	100	93.016	92.5674	92.7898
Norway	93.0984	92.8963	92.8637	92.8794	92.9053	92.9053	92.9834	100	95.7673	95.69
CC278f	92.7247	92.5478	92.6543	92.674	92.6107	92.6143	92.7054	95.7247	100	95.0768
SM78	92.8425	92.6652	92.6734	92.7391	92.7244	92.832	92.7748	95.6129	95.0018	100
SM51	92.8278	92.581	92.7129	92.7597	92.7127	92.7944	92.8067	95.5983	94.9394	99.0621
SM72	92.8807	92.7002	92.7945	92.8223	92.7757	92.962	92.8923	95.6318	95.1077	99.0187
Vaf-46	92.8824	93.1446	92.7409	92.8203	92.8484	92.8879	92.8813	93.2992	92.7692	93.3906
JHI1093	93.0348	92.8841	92.7768	92.8881	92.7931	92.9671	93.0315	93.533	92.8867	93.4256
JHI1096	93.0032	92.824	92.8301	92.8791	92.7975	92.9796	93.0219	93.4971	92.8667	93.3997
GPTR08	92.4241	92.369	92.3428	92.3169	92.3491	92.3916	92.3317	92.5645	92.3395	92.7352
GPTR02	92.4314	92.3158	92.4225	92.4205	92.4347	92.4909	92.4834	92.5865	92.256	92.8085
SRDI565	92.5127	92.3803	92.4569	92.4078	92.4058	92.435	92.4742	92.6904	92.4414	92.8534
Ps8	92.8943	92.8418	92.8616	92.8746	92.8138	93.0012	92.9761	94.0039	93.4366	94.0756
JHI1236	92.8911	92.7673	92.799	92.784	92.8624	93.0092	92.9815	93.9804	93.5221	94.0644
JHI953	92.9159	92.7234	92.7291	92.74	92.9043	92.923	92.8961	93.8981	93.4732	94.0512
SM147A	92.8273	92.7915	92.7649	92.8428	92.7922	92.9298	92.8751	93.8426	93.4162	94.0129
JHI1238	92.9075	92.8495	92.8142	92.8099	92.9135	92.9944	92.9764	94.0104	93.4756	94.1103
UPM1133	93.0353	92.8568	92.7168	92.7661	92.8147	92.977	92.8876	94.0473	93.4454	94.0273
JHI1592	92.9592	92.7958	92.7197	92.7233	92.8162	92.8973	92.8858	93.9668	93.4837	94.0828
SM41	92.8792	92.792	92.7359	92.839	92.8791	92.9164	92.7424	93.9411	93.366	94.1295
JHI1253	92.9593	92.7927	92.8148	92.8283	92.8035	93.0842	93.0699	93.9852	93.4199	94.0272
JHI370	92.905	92.7689	92.778	92.7826	92.8822	92.9365	92.8908	93.9683	93.4825	94.0793
JHI387	92.9384	92.7634	92.6965	92.7634	92.8792	92.8905	92.8495	93.9676	93.4827	94.1264
JHI388	92.9837	92.808	92.7427	92.7573	92.8383	92.9295	92.8536	94.0147	93.4901	94.1003
JHI788	92.9331	92.9395	92.8184	92.8238	92.9072	93.0229	93.0701	94.0879	93.5379	94.1787
JHI985	92.881	92.8549	92.8264	92.8492	92.8481	92.9007	92.7463	94.0133	93.4209	94.0375
GB30	92.9804	92.8416	92.8816	92.8378	92.9189	93.0145	93.0129	94.0083	93.4566	94.1202
JHI782	92.8873	92.7098	92.7891	92.8165	92.7656	92.9967	92.9101	93.9392	93.3902	94.0352
JHI783	92.8464	92.7051	92.7599	92.8145	92.8064	93.0374	92.9268	93.9025	93.3913	94.0465
SM170C	92.8178	92.7148	92.7804	92.768	92.813	92.9232	92.8158	93.8523	93.3478	94.1725
JHI42	93.0382	92.8049	92.7612	92.8167	92.827	93.0024	92.9107	93.9765	93.4671	94.0701
JHI979	92.8707	92.8251	92.7701	92.8242	92.8665	92.9002	92.9371	93.9874	93.4437	94.0459
JHI1259	92.9222	92.7356	92.7826	92.7716	92.7898	93.1825	93.0687	93.8761	93.4319	93.9834
TA1	92.837	92.7441	92.7836	92.8294	92.8402	92.9121	92.822	93.9231	93.3417	94.0721
JHI24	93.0993	92.8506	92.7079	92.757	92.8493	93.0093	93.0245	94.0509	93.4436	94.092
JHI787	93.1021	92.8204	92.7946	92.8117	92.8835	93.0596	93.0785	94.0795	93.4382	94.1345
RCAM1026	93.0091	92.9098	92.9368	92.9173	92.9617	93.0806	93.0879	94.0297	93.6169	94.1436
Vh3	92.8834	92.8119	92.8801	92.8658	92.84	93.003	92.9125	93.8121	93.3226	93.9238
JHI925	92.8315	92.644	92.7239	92.8009	92.6932	92.8676	92.7596	93.5372	93.0619	93.7227
Vc2	92.9849	92.9028	92.808	92.7719	92.7412	93.0015	92.9782	93.7558	93.1873	93.8241
JHI1422	92.9394	92.77	92.814	92.8709	92.9252	92.9705	93.1438	93.8468	93.4193	93.929
JHI1266	92.8657	92.7338	92.8352	92.8244	92.8626	93.204	93.1164	93.7742	93.4164	93.9991
51A11	94.3668	94.242	94.2581	94.2822	94.4583	94.4171	94.4314	93.2936	93.0911	93.1513
51B21	94.3633	94.2538	94.3389	94.2779	94.4153	94.4222	94.436	93.253	93.0915	93.2203
128C53	92.6559	92.357	92.3955	92.3521	92.3623	92.7191	92.6591	94.484	94.9015	94.8196
4292	92.4716	92.1165	92.26	92.3716	92.3289	92.4145	92.3772	94.5847	94.9713	94.843
CC283bq	92.4008	92.3689	92.308	92.3213	92.3445	92.3004	92.448	94.4886	94.7682	94.8531
USDA2370	92.7038	92.2839	92.4039	92.3484	92.3913	92.8176	92.8293	94.5259	94.8902	94.7774
JHI2442	92.5751	92.3816	92.3543	92.3258	92.3923	92.5977	92.6611	94.4561	94.9504	94.8491
JHI1084	92.4548	92.1753	92.3946	92.4226	92.2796	92.5807	92.5535	94.4207	94.8526	94.8585
JHI944	92.3885	92.2475	92.3859	92.4018	92.4224	92.4176	92.5248	94.4384	94.8774	94.8418
OV152	92.5361	92.0828	92.2958	92.3016	92.3587	92.3863	92.4085	94.6412	95.0143	94.9352
UPM1137	92.5908	92.3326	92.2983	92.3618	92.3873	92.59	92.5179	94.6217	94.9633	94.838
SM149A	92.332	92.2045	92.2329	92.3044	92.2422	92.4733	92.4984	94.4652	94.8772	95.1285
Rt24.2	92.4234	92.2533	92.3568	92.353	92.3749	92.3955	92.4799	94.507	94.9341	95.2047
CCBAU23252	90.9096	90.8011	90.7411	90.7754	91.1214	90.9121	90.8347	91.4618	91.3261	91.5941
JHI536	91.3805	90.8422	90.8481	90.8634	90.916	91.107	91.1326	91.7178	91.3908	91.6161
C15	90.8894	90.7609	90.7861	90.7462	90.7546	90.9191	90.8296	91.4835	91.4169	91.6472
CF307	90.9699	90.8511	90.7885	90.7836	90.8728	90.9042	90.9521	91.5313	91.4514	91.6887

Table S4.5. (Continuation)

Query	SM51	SM72	Vaf-46	JHI1093	JHI1096	GPTR08	GPTR02	SRDI565	Ps8	JHI1236
3841	93.1667	93.2701	93.0914	93.1448	93.1532	92.6062	92.6248	92.6327	93.66	93.741
SM38	93.3373	93.4713	93.0689	93.0939	93.1822	92.669	92.7785	92.7229	93.7075	93.6911
JHI960	93.1735	93.2674	93.1042	93.102	93.1077	92.5434	92.6232	92.6494	93.6177	93.7155
JHI963	93.1524	93.2374	93.0799	93.1493	93.1175	92.5975	92.6146	92.5569	93.6664	93.7047
JHI1587	93.1883	93.2165	93.1028	93.1648	93.1632	92.5736	92.6328	92.633	93.6929	93.6872
JHI1600	93.2058	93.2833	93.0912	93.114	93.1125	92.5519	92.639	92.5818	93.7394	93.7611
JHI13	93.1525	93.2917	93.0891	93.2554	93.2587	92.6068	92.5699	92.655	93.6209	93.6251
JHI1438	93.1153	93.2036	93.1953	93.1816	93.1912	92.6784	92.6552	92.7485	93.7838	93.7922
JHI535	93.2227	93.3147	93.229	93.2077	93.1804	92.6899	92.7229	92.633	93.8137	93.8652
JHI585	93.2742	93.2851	93.2595	93.183	93.1701	92.6823	92.6644	92.6302	93.8555	93.855
JHI973	93.217	93.2943	93.244	93.1892	93.186	92.6552	92.6797	92.6662	93.7301	93.7479
JHI974	93.223	93.2999	93.1767	93.2792	93.2419	92.6828	92.6642	92.6784	93.8009	93.8026
JHI1415	93.1859	93.2464	93.2764	93.2134	93.2067	92.6431	92.6515	92.617	93.8117	93.8381
SM3	93.3226	93.4105	93.0292	93.0762	93.0406	92.6635	92.6766	92.6248	93.6099	93.5949
VF39	93.1284	93.2331	93.1975	93.1662	93.1612	92.6256	92.6155	92.6399	93.6751	93.6556
L361	92.8411	92.9833	93.1696	93.1192	93.0983	92.6076	92.7105	92.6982	93.1974	93.1026
VFF1R2B1	93.1867	93.2243	93.1172	93.0942	93.0795	92.6336	92.6483	92.6277	93.5646	93.5647
21A12	93.2064	93.2526	93.1897	93.1253	93.1385	92.5757	92.6275	92.5878	93.5622	93.648
VFF1R2A1	93.2058	93.2506	93.1751	93.1903	93.1974	92.5679	92.6717	92.6071	93.5946	93.6582
WSM1455	93.1212	93.1712	93.0505	93.0773	93.0668	92.4859	92.5077	92.5875	93.5009	93.4969
WSM1481	93.0344	93.1337	93.0483	93.0851	93.0719	92.4716	92.4619	92.5589	93.5324	93.4906
WSM78	93.8521	93.9	92.9885	92.9721	92.9774	92.4079	92.4202	92.5503	93.5549	93.5431
SM152B	93.9706	94.0255	93.0941	93.078	93.0619	92.57	92.5985	92.6468	93.6598	93.6485
CC275e	93.9157	94.0692	93.0742	93.1282	93.1073	92.6079	92.6497	92.7451	93.6879	93.6887
WSM1325	93.3073	93.383	92.9607	93.01	92.9945	92.4086	92.494	92.5237	93.4437	93.392
SRDI943	93.3225	93.3104	92.9698	93.032	93.0262	92.481	92.48	92.5752	93.5017	93.4215
JHLM12A2	93.1549	93.3565	92.9448	92.9341	92.9445	92.3053	92.3439	92.4845	93.1952	93.2834
Vaf12	93.1465	93.1878	93.4013	93.1829	93.2415	92.591	92.6135	92.672	93.5	93.4808
JHI2451	93.191	93.2794	93.2715	93.2889	93.2567	92.6017	92.6488	92.7214	93.5768	93.5357
JHI10	93.2113	93.3121	93.0264	93.2129	93.2228	92.4744	92.5504	92.6364	93.592	93.5693
JHI54	93.1859	93.2926	93.0564	93.1334	93.1122	92.5681	92.5559	92.5691	93.5195	93.3914
FA23	93.3158	93.3573	93.0826	93.072	93.1038	92.6328	92.6756	92.7407	93.5667	93.5111
JHI2450	93.3022	93.3651	93.2968	93.2917	93.2505	92.5827	92.5713	92.683	93.6091	93.5429
Vaf-108	92.6092	92.6713	93.3025	93.0703	93.0734	92.4174	92.3403	92.4958	92.9324	92.9758
JHI2449	92.7473	92.8436	93.0371	93.1035	93.0799	92.4905	92.6115	92.5325	93.1081	93.0618
UPM1131	92.7246	92.8261	93.1349	93.0178	93.0644	92.4528	92.4852	92.5751	93.0893	92.9766
CCBAU10279	92.5266	92.6583	92.9241	92.8736	92.8537	92.4674	92.4507	92.5329	92.8055	92.807
WSM1689	92.6222	92.718	92.6765	92.6983	92.67	92.5034	92.5915	92.5335	92.907	92.8354
VFF2R2A1	92.5475	92.6254	92.8399	92.8581	92.8374	92.4607	92.4883	92.4695	92.8371	92.8495
SPF2A11	92.6202	92.6814	92.8239	92.8569	92.8137	92.4913	92.566	92.5159	92.9376	92.8838
VFCR2A2	92.5113	92.6185	92.6934	92.7791	92.7826	92.4668	92.4541	92.5707	92.8623	92.857
248	92.7066	92.8348	93.166	93.1233	93.1255	92.7523	92.7104	92.8251	93.0544	93.0383
63A21	92.7131	92.7785	93.0947	93.1978	93.1527	92.7806	92.7859	92.8206	93.0943	93.088
GLR2	92.6132	92.7101	92.858	92.9813	92.9928	92.5636	92.6038	92.6995	92.9927	92.923
CL8	92.5647	92.6997	92.8946	92.7724	92.7436	92.4815	92.5274	92.5066	92.8693	92.8539
VFSR2A2	92.5038	92.6126	92.781	92.8311	92.8115	92.3233	92.4598	92.3987	92.8755	92.8284
SS21	92.4631	92.5138	92.7157	92.7218	92.6793	92.3265	92.4232	92.4111	92.7957	92.7488
41A11	92.5821	92.6818	92.7962	92.823	92.8153	92.4638	92.5413	92.4909	92.9003	92.8141
42B12	92.6292	92.7048	92.9015	92.9321	92.9079	92.5217	92.5763	92.56	93.0035	92.9565
43A11	92.5774	92.637	92.7968	92.849	92.815	92.4697	92.4712	92.5043	92.8871	92.9035
43B11	92.4723	92.6636	92.7467	92.7742	92.7626	92.3686	92.4426	92.4502	92.8296	92.8382
43B12	92.5064	92.6857	92.743	92.8045	92.7883	92.3787	92.484	92.4704	92.8093	92.8713
71A12	92.6037	92.6915	92.7607	92.7976	92.7546	92.4909	92.5205	92.5188	92.8985	92.8694
73A11	92.5529	92.6835	92.8225	92.8602	92.831	92.4219	92.4378	92.4214	92.8749	92.8367
81B22	92.5924	92.654	92.8022	92.746	92.7604	92.4695	92.5353	92.5088	92.8959	92.9206
83A12	92.5691	92.6725	92.8233	92.7641	92.7026	92.4608	92.5505	92.4765	92.8956	92.8766
VFHR1A2	92.5521	92.749	92.7684	92.7448	92.7482	92.3818	92.4954	92.4561	92.8888	92.8432
VFF1R1A2	92.4875	92.611	92.7486	92.6842	92.6926	92.2732	92.3916	92.3961	92.7942	92.7781
93B11	92.5266	92.6345	92.8413	92.7957	92.7869	92.4564	92.4937	92.4142	92.8559	92.8792
111A12	92.5953	92.6871	92.8191	92.8509	92.7897	92.3885	92.4527	92.4184	92.845	92.8784
VFSR2B1	92.5752	92.5906	92.7803	92.7905	92.8064	92.4081	92.3932	92.3221	92.8719	92.8447
SPF4F7	92.5619	92.6684	92.772	92.7982	92.7755	92.4694	92.4123	92.4206	92.8624	92.8478
FB206	92.4807	92.5766	92.7373	92.7501	92.7033	92.3019	92.3652	92.3938	92.7941	92.7039
SL16	92.5202	92.5626	92.7191	92.7606	92.7354	92.3816	92.4642	92.4647	92.8243	92.8148
HUTR05	92.6013	92.6796	92.8488	92.7823	92.81	92.44	92.5008	92.5259	92.8329	92.7664
21B12	92.5476	92.6092	92.8412	92.7612	92.771	92.3635	92.427	92.4631	92.8582	92.7868
WYCCWR11290	92.7277	92.8311	92.516	92.4797	92.4893	91.9152	91.9387	91.956	92.6815	92.665
WYCCWR11317	92.8651	92.9046	92.4971	92.3739	92.4294	91.9421	92.0263	92.0589	92.6617	92.6497
CCBAU11080	92.6348	92.7343	92.2748	92.3833	92.3645	91.9395	91.9171	91.953	92.5659	92.5474
CCBAU03386	92.7309	92.8534	92.4032	92.351	92.359	91.8989	91.9242	91.9488	92.4417	92.5032

Table S4.5. (Continuation)

Query	SM51	SM72	Vaf-46	JHI1093	JHI1096	GPTR08	GPTR02	SRDI565	Ps8	JHI1236
WYCCWR10014	93.4919	93.6036	92.6497	92.7076	92.6771	92.199	92.205	92.3388	93.1419	93.0733
JKLM13E	93.0996	93.3078	92.8536	92.8802	92.8774	92.3995	92.2977	92.4606	93.1764	93.1708
Tri-43	93.1432	93.2302	92.846	92.77	92.7379	92.3365	92.3428	92.503	92.9879	93.0267
73B11	92.7803	92.8648	92.9134	92.9458	92.9282	92.5081	92.4806	92.5234	93.0423	92.9367
73B12	92.8793	92.9363	93.0109	92.9927	92.9904	92.6103	92.5811	92.5903	93.0283	93.0092
RSP1F2	92.8485	92.8991	92.9997	93.1067	93.0699	92.4164	92.4366	92.4824	92.9665	92.9062
Vaf10	92.6325	92.7301	93.1697	92.8566	92.899	92.3154	92.3277	92.3942	92.7757	92.8257
PEPV16	92.6484	92.7488	92.7422	92.815	92.7517	92.3564	92.3417	92.3263	92.9483	92.8915
TOM	92.7115	92.8226	92.8414	92.9045	92.9095	92.4208	92.4106	92.3556	92.938	92.8774
121B21	92.7342	92.793	92.875	92.8773	92.7989	92.4163	92.4139	92.3724	92.8826	92.8905
RSP1F10	92.7439	92.8887	92.9994	93.0299	93.0229	92.4672	92.4441	92.4047	93.0017	92.9421
RSP1A1	92.8415	92.8153	92.9214	93.0286	92.9707	92.38	92.4334	92.4765	93.0032	93.0067
Norway	95.6191	95.6287	93.3385	93.4935	93.4823	92.5375	92.6373	92.6833	94.0461	94.0039
CC278f	95.0031	95.1385	92.8879	92.9759	92.946	92.4359	92.3285	92.4706	93.6895	93.6017
SM78	99.0866	99.0405	93.2972	93.4438	93.3832	92.7199	92.6897	92.8639	94.0589	94.0634
SM51	100	98.9879	93.3241	93.3408	93.3067	92.6724	92.7098	92.7052	94.0483	94.0321
SM72	98.9703	100	93.319	93.3788	93.3703	92.7541	92.6727	92.812	94.0051	93.985
Vaf-46	93.3556	93.3305	100	97.6829	97.6648	94.4618	94.4802	94.5203	94.7879	94.7461
JHI1093	93.4716	93.402	97.7071	100	99.9961	94.5836	94.5263	94.6235	94.9045	94.7917
JHI1096	93.3849	93.3199	97.6829	99.9952	100	94.5289	94.5034	94.6049	94.8624	94.8187
GPTR08	92.6935	92.7959	94.4968	94.6022	94.6169	100	97.738	97.9183	93.9347	93.8962
GPTR02	92.7702	92.7322	94.5473	94.5715	94.5669	97.7414	100	97.7463	93.8973	93.8802
SRDI565	92.781	92.8363	94.5868	94.6355	94.6779	97.9445	97.6696	100	93.9764	93.989
Ps8	93.9855	93.9809	94.8402	94.8047	94.8215	93.8357	93.8305	93.8683	100	97.988
JHI1236	93.9225	94.0046	94.765	94.7669	94.7592	93.8636	93.88	93.8897	97.9218	100
JHI953	93.9057	93.9897	94.6686	94.7891	94.8113	93.7316	93.7763	93.8296	97.9721	99.4676
SM147A	94.0262	93.9805	94.5973	94.7809	94.7775	93.815	93.91	93.9196	97.8482	99.5112
JHI1238	94.0948	94.0783	94.7856	94.8493	94.8463	93.8413	93.7969	93.9109	98.0241	98.1879
UPM1133	93.9355	93.9329	94.8054	94.8349	94.8525	93.7246	93.6676	93.8593	98.0126	97.8512
JHI1592	94.0408	94.0476	94.7776	94.8239	94.8547	93.7643	93.8419	93.7814	98.0463	98.1612
SM41	93.9826	94.0052	94.6862	94.7439	94.7864	93.8129	93.8374	93.8407	97.922	98.0318
JHI1253	93.9564	93.9714	94.8301	94.7667	94.7691	93.7495	93.7852	93.8565	98.0704	98.1308
JHI370	93.9821	94.0114	94.8029	94.8256	94.8288	93.7448	93.8579	93.7884	98.09	98.1646
JHI387	94.0764	94.0354	94.7925	94.8038	94.8074	93.7391	93.8344	93.8111	98.0442	98.1352
JHI388	94.0518	94.082	94.8315	94.8221	94.822	93.8123	93.8008	93.8571	98.0921	98.1663
JHI788	94.0174	94.0765	94.8364	94.8955	94.9201	93.9229	93.9627	93.9328	98.0202	98.1619
JHI985	93.936	93.9871	94.7147	94.8157	94.7558	93.7271	93.7887	93.8138	97.9159	98.0877
GB30	94.0339	93.991	94.8274	94.9062	94.8827	93.8531	93.8418	93.9293	98.0683	98.1588
JHI782	94.034	93.9889	94.7655	94.84	94.8574	93.8281	93.8431	93.9006	98.2871	97.9897
JHI783	94.0476	93.9892	94.7697	94.8174	94.8324	93.8765	93.8193	93.902	98.2586	97.9969
SM170C	94.0114	94.0134	94.6586	94.6589	94.6927	93.8533	93.7991	93.8614	98.1397	97.998
JHI42	93.8999	93.9218	94.8718	94.8514	94.8415	93.7167	93.8182	93.7935	98.5029	98.0704
JHI979	93.9623	93.9936	94.7451	94.808	94.8316	93.8201	93.729	93.9566	98.9326	98.2496
JHI1259	93.9456	93.9194	94.7589	94.7508	94.7284	93.8631	93.8206	93.9155	99.0408	98.1135
TA1	93.9914	93.953	94.7104	94.7278	94.7162	93.8753	93.8464	93.8836	98.9429	97.9859
JHI24	93.9584	94.0029	94.8514	94.8751	94.8871	93.7824	93.8467	93.8654	98.0117	98.0582
JHI787	94.0001	94.0225	94.8008	94.8299	94.8752	93.7473	93.7401	93.8524	97.997	98.0527
RCAM1026	94.076	94.1055	94.8669	94.8658	94.878	93.9733	93.8999	93.9028	98.1402	98.1451
Vh3	93.779	93.9143	94.7325	94.8299	94.8087	93.8395	93.8968	93.8446	96.565	96.5803
JHI925	93.6946	93.7317	94.5192	94.6476	94.6846	93.7382	93.7604	93.7406	96.4047	96.464
Vc2	93.7415	93.7733	94.8255	94.8866	94.8931	93.8626	93.8149	93.873	96.5206	96.5785
JHI1422	93.8214	93.8594	94.7095	94.7287	94.7122	93.7985	93.7779	93.8197	97.1576	97.2281
JHI1266	93.8823	93.8896	94.6735	94.6031	94.6391	93.7651	93.7753	93.9006	97.3672	97.5243
51A11	93.05	93.248	93.1251	93.1811	93.1479	92.4646	92.5949	92.6391	93.5611	93.5839
51B21	93.0926	93.2356	93.1291	93.1463	93.1254	92.546	92.6774	92.6401	93.5631	93.6213
128C53	94.7217	94.8282	92.9229	93.1218	93.1383	92.1798	92.0882	92.2889	93.2357	93.3057
4292	94.7347	94.9551	92.7853	92.9006	92.897	92.2611	92.163	92.2729	93.3672	93.3246
CC283bq	94.7416	94.9116	92.923	92.8527	92.8737	92.2797	92.1641	92.2364	93.2373	93.2553
USDA2370	94.7395	94.9186	92.838	92.9306	93.0147	92.1494	92.0977	92.3257	93.3142	93.2959
JHI2442	94.7425	94.8794	92.9172	93.0069	92.999	92.1807	92.1651	92.3505	93.3533	93.3427
JHI1084	94.7496	94.8714	92.8052	92.8624	92.8796	92.1166	92.234	92.2736	93.2286	93.3004
JHI944	94.762	94.9347	92.8475	92.8461	92.8936	92.1678	92.1715	92.2484	93.2722	93.2748
OV152	94.8405	95.0211	92.7877	92.8705	92.8948	92.1439	92.0982	92.2569	93.3509	93.4144
UPM1137	94.7512	94.8796	92.969	93.0353	93.0125	92.1847	92.1214	92.2981	93.3992	93.3302
SM149A	95.0438	95.2087	92.6841	92.8313	92.8001	92.3747	92.1655	92.3909	93.1955	93.2375
Rt24.2	95.0592	95.2249	92.791	92.7752	92.8475	92.3005	92.2823	92.4107	93.2549	93.3057
CCBAU23252	91.5166	91.6574	91.0116	91.0003	90.9377	90.5012	90.6199	90.6221	91.2827	91.2788
JHI536	91.5279	91.6441	91.0824	91.1827	91.1862	90.5219	90.6331	90.6322	91.3121	91.3438
C15	91.5921	91.6558	90.9025	90.9718	90.962	90.5657	90.5403	90.7595	91.2886	91.2682
CF307	91.6137	91.7849	91.0036	90.9955	91.0228	90.5725	90.6851	90.7342	91.3535	91.3113

Table S4.5. (Continuation)

Query	JHI953	SM147A	JHI1238	UPM1133	JHI1592	SM41	JHI1253	JHI370	JHI387	JHI388
3841	93.6444	93.6593	93.6504	93.6312	93.6179	93.5786	93.7246	93.6605	93.6749	93.671
SM38	93.6399	93.7817	93.6613	93.6497	93.6882	93.7696	93.6146	93.738	93.6717	93.7102
JHI960	93.5627	93.6234	93.6584	93.5808	93.5874	93.6233	93.6265	93.5966	93.5804	93.6017
JHI963	93.5714	93.6373	93.65	93.5726	93.6155	93.6287	93.6683	93.5967	93.6027	93.6079
JHI1587	93.6591	93.6379	93.6373	93.6254	93.6467	93.6915	93.6283	93.6636	93.6373	93.6327
JHI1600	93.6936	93.6604	93.7349	93.6279	93.6903	93.7046	93.6694	93.7503	93.6342	93.6945
JHI13	93.4902	93.5308	93.6053	93.6137	93.5398	93.6412	93.7163	93.5894	93.5189	93.5648
JHI1438	93.6144	93.6656	93.7379	93.7409	93.6393	93.6594	93.8083	93.697	93.6189	93.6709
JHI535	93.7613	93.7493	93.7678	93.7492	93.7802	93.7666	93.7762	93.8258	93.7888	93.7858
JHI585	93.761	93.749	93.7689	93.6875	93.7734	93.7296	93.7902	93.7923	93.7838	93.796
JHI973	93.7336	93.7086	93.7237	93.6883	93.7748	93.7505	93.7387	93.757	93.7413	93.7517
JHI974	93.7495	93.7165	93.7829	93.7138	93.7845	93.7446	93.8636	93.8351	93.7872	93.8243
JHI1415	93.7644	93.7118	93.7592	93.6918	93.8219	93.7517	93.7614	93.8066	93.7944	93.8213
SM3	93.4973	93.6128	93.521	93.4762	93.6319	93.6372	93.592	93.5843	93.5402	93.5544
VF39	93.5594	93.5832	93.6371	93.5936	93.5837	93.5835	93.6171	93.6064	93.5929	93.6182
L361	93.055	93.1225	93.0955	93.2167	93.1251	93.0744	93.1307	93.126	93.1342	93.0832
VFF1R2B1	93.4737	93.4785	93.5848	93.5927	93.4252	93.5811	93.5054	93.4654	93.4471	93.4613
21A12	93.5406	93.57	93.6408	93.6235	93.6176	93.6617	93.5698	93.571	93.6301	93.6036
VFF1R2A1	93.5943	93.5859	93.6158	93.6746	93.6129	93.6375	93.5645	93.612	93.6409	93.6091
WSM1455	93.463	93.4487	93.4899	93.4292	93.5727	93.5376	93.5363	93.5257	93.5156	93.5542
WSM1481	93.4031	93.4792	93.5361	93.4886	93.5271	93.5149	93.4791	93.5232	93.5114	93.501
WSM78	93.5055	93.5797	93.5373	93.5252	93.4935	93.5214	93.5686	93.5427	93.49	93.4789
SM152B	93.4935	93.6127	93.5987	93.5927	93.5721	93.6792	93.6995	93.5555	93.5929	93.5928
CC275e	93.5482	93.6817	93.7364	93.6157	93.5716	93.7542	93.6157	93.5443	93.572	93.5976
WSM1325	93.3373	93.407	93.362	93.3569	93.3769	93.4444	93.3696	93.3837	93.3361	93.3089
SRDI943	93.3605	93.5296	93.4463	93.3571	93.5419	93.5308	93.479	93.4481	93.4614	93.4622
JHLM12A2	93.136	93.1918	93.2759	93.1917	93.1458	93.2022	93.2365	93.1262	93.1541	93.1773
Vaf12	93.4104	93.3995	93.4694	93.5257	93.396	93.389	93.3977	93.4398	93.4543	93.4381
JHI2451	93.5357	93.4932	93.5136	93.6589	93.609	93.5538	93.5583	93.6294	93.5773	93.602
JHI10	93.5396	93.4244	93.5581	93.512	93.5686	93.5156	93.5002	93.5836	93.5661	93.5937
JHI54	93.3675	93.4398	93.4981	93.4673	93.462	93.4456	93.4231	93.4691	93.472	93.438
FA23	93.4133	93.483	93.4666	93.5525	93.4999	93.5178	93.665	93.4827	93.4975	93.5116
JHI2450	93.5754	93.4722	93.5253	93.6669	93.5803	93.439	93.5709	93.519	93.6028	93.5384
Vaf-108	92.9155	92.888	92.9664	92.9074	92.932	92.9604	92.7451	92.8727	92.8796	92.8994
JHI2449	92.9174	92.9492	92.9141	93.0494	92.976	92.9515	92.9859	92.9671	92.9741	92.9703
UPM1131	93.0077	92.8461	92.9267	93.0342	93.0098	92.8765	92.9348	92.9611	93.0106	92.9697
CCBAU10279	92.7184	92.7967	92.7642	92.7378	92.7324	92.8408	92.8002	92.8031	92.722	92.7697
WSM1689	92.7555	92.8281	92.8241	92.7985	92.7963	92.874	92.8248	92.8261	92.837	92.8438
VFF2R2A1	92.8222	92.8061	92.8136	92.7733	92.7556	92.7735	92.8491	92.7907	92.7967	92.7837
SPF2A11	92.8266	92.8427	92.8402	92.8282	92.8296	92.8282	92.8427	92.8302	92.8115	92.7876
VFCR2A2	92.7697	92.7639	92.8084	92.6807	92.7552	92.7729	92.7858	92.7977	92.7404	92.7514
248	93.0472	93.012	93.084	92.9856	93.0473	93.0149	93.0217	93.0188	93.0464	93.0304
63A21	92.9836	93.0041	93.0613	93.0721	92.9875	93.0038	93.0232	93.0312	92.9844	93.0094
GLR2	92.8772	92.9299	92.8618	92.932	92.909	92.9472	92.8901	92.9814	92.9151	92.9463
CL8	92.8577	92.8297	92.8014	92.7858	92.8408	92.789	92.8038	92.8408	92.8393	92.8361
VFSR2A2	92.7699	92.7861	92.8798	92.7593	92.8062	92.7791	92.8237	92.8696	92.8697	92.8796
SS21	92.7343	92.7355	92.76	92.7583	92.7419	92.7543	92.7634	92.7422	92.7166	92.7228
41A11	92.815	92.743	92.8318	92.7917	92.8011	92.7826	92.8087	92.8023	92.8077	92.8106
42B12	92.9159	92.9039	92.8919	92.8504	92.9194	92.8949	92.9542	92.9333	92.902	92.929
43A11	92.8646	92.8212	92.84	92.7879	92.8081	92.8069	92.8514	92.8242	92.8157	92.8028
43B11	92.8324	92.8095	92.8554	92.7296	92.8209	92.7831	92.731	92.7893	92.8203	92.8045
43B12	92.8636	92.7944	92.85	92.7788	92.7603	92.7635	92.7925	92.8109	92.8011	92.7968
71A12	92.7831	92.8496	92.8994	92.8255	92.802	92.8255	92.8696	92.8245	92.8428	92.865
73A11	92.8385	92.8214	92.7776	92.8122	92.81	92.7947	92.76	92.8063	92.8519	92.8497
81B22	92.8129	92.7933	92.8626	92.798	92.8214	92.8679	92.8675	92.8169	92.8693	92.8025
83A12	92.827	92.8152	92.8712	92.7945	92.7928	92.8092	92.8335	92.817	92.8128	92.7917
VFHR1A2	92.8918	92.8052	92.8047	92.738	92.8044	92.7799	92.7139	92.8181	92.8175	92.8321
VFF1R1A2	92.7827	92.7433	92.7888	92.6359	92.7541	92.7283	92.6678	92.8121	92.8102	92.8206
93B11	92.854	92.8222	92.8287	92.8073	92.8534	92.875	92.8043	92.8326	92.9048	92.8564
111A12	92.8549	92.8329	92.8207	92.8318	92.8446	92.8606	92.7859	92.8116	92.8356	92.8176
VFSR2B1	92.7943	92.797	92.8023	92.7883	92.8494	92.8416	92.824	92.8106	92.8596	92.878
SPF4F7	92.7981	92.8442	92.8151	92.7437	92.7909	92.7768	92.7736	92.7721	92.8165	92.8357
FB206	92.6815	92.6934	92.7068	92.7093	92.6303	92.6797	92.7506	92.6909	92.6424	92.6309
SL16	92.7995	92.7767	92.84	92.715	92.8011	92.7748	92.722	92.7988	92.8336	92.8088
HUTR05	92.7273	92.8304	92.6667	92.7973	92.7399	92.7582	92.7958	92.7633	92.7279	92.782
21B12	92.7811	92.7998	92.8082	92.7145	92.7927	92.7772	92.8013	92.7782	92.7405	92.7951
WYCCWR11290	92.5972	92.6762	92.7168	92.7461	92.6441	92.652	92.7277	92.6534	92.6293	92.6343
WYCCWR11317	92.6301	92.6904	92.752	92.6442	92.6603	92.678	92.709	92.6633	92.6642	92.7258
CCBAU11080	92.4966	92.4029	92.5755	92.5022	92.4958	92.5142	92.5917	92.5254	92.5014	92.5179
CCBAU03386	92.51	92.5087	92.5415	92.4942	92.6133	92.5851	92.5183	92.5478	92.5095	92.5309

Table S4.5. (Continuation)

Query	JHI953	SM147A	JHI1238	UPM1133	JHI1592	SM41	JHI1253	JHI370	JHI387	JHI388
WYCCWR10014	93.0757	93.1376	93.11	93.0538	93.0128	93.1227	93.126	93.019	93.0394	92.996
JKLM13E	93.0717	93.1374	93.1847	93.1486	93.0837	93.1358	93.159	93.0701	93.0965	93.0793
Tri-43	92.9866	93.0483	92.9653	93.0579	93.022	92.9996	92.9722	92.9751	93.0463	92.9998
73B11	92.9145	92.8833	92.9852	92.8816	92.8962	92.9865	93.0069	92.9142	92.9483	92.9357
73B12	92.9784	93.0118	93.062	92.9468	92.9933	93.0821	93.049	92.9901	92.9926	92.9723
RSP1F2	92.9158	92.9555	92.9234	93.1577	93.0397	92.933	92.9821	92.9947	93.0407	93.0486
Vaf10	92.6867	92.7526	92.8065	92.93	92.7697	92.7601	92.7742	92.7339	92.7991	92.7827
PEPV16	92.7503	92.8063	92.8338	92.8223	92.7499	92.8793	92.8196	92.7609	92.7614	92.773
TOM	92.7134	92.9045	92.8705	92.8562	92.8143	92.8455	92.8838	92.836	92.8273	92.8413
121B21	92.7815	92.8631	92.8829	92.8278	92.8805	92.8837	92.8578	92.8403	92.8826	92.8294
RSP1F10	92.8554	92.8626	92.9925	92.9664	92.9668	92.9347	93.1011	92.8847	92.9711	92.9028
RSP1A1	92.9246	92.9134	92.9857	92.9724	92.9625	92.8565	93.1496	92.9447	92.9678	92.9529
Norway	93.9414	93.9563	94.034	94.0593	93.9969	93.9553	93.9086	93.9859	93.9838	93.9794
CC278f	93.4947	93.4259	93.5924	93.5384	93.5916	93.4125	93.5494	93.6026	93.5781	93.5535
SM78	94.0583	94.133	94.1341	94.0766	94.1291	94.1826	94.1124	94.1109	94.1213	94.1302
SM51	93.8835	94.0363	94.0622	93.9772	94.0298	94.0302	93.9733	94.0043	94.0429	94.0311
SM72	93.9527	94.0211	94.0662	93.8969	94.0416	94.0532	94.0447	94.0348	94.0287	94.0461
Vaf-46	94.7023	94.7304	94.7817	94.7934	94.8623	94.7281	94.7821	94.8333	94.854	94.8455
JHI1093	94.8206	94.7929	94.8386	94.853	94.8408	94.8191	94.7733	94.841	94.8036	94.8221
JHI1096	94.7918	94.7896	94.8745	94.8947	94.8756	94.816	94.8139	94.8848	94.8598	94.868
GPTR08	93.7536	93.9322	93.9404	93.841	93.8516	93.8805	93.8542	93.8439	93.8581	93.8168
GPTR02	93.8742	93.8708	93.8897	93.8204	93.9136	93.8531	93.8495	93.8899	93.9059	93.9089
SRDI565	93.9268	93.9737	93.9943	93.9028	93.8174	93.9205	93.9536	93.8947	93.8583	93.8569
Ps8	98.0361	97.7938	97.987	97.9909	98.0272	98.0054	98.0637	98.0682	98.0501	98.0609
JHI1236	99.4798	99.4905	98.1823	97.861	98.1314	98.0805	98.1426	98.1397	98.1482	98.1409
JHI953	100	99.3059	98.0654	97.9321	98.1554	97.9665	98.1447	98.1829	98.1918	98.1717
SM147A	99.2809	100	98.1061	97.9461	98.1753	98.0828	98.1706	98.1847	98.1472	98.137
JHI1238	98.1058	98.1383	100	98.4293	98.813	98.9286	98.8458	98.82	98.8065	98.7997
UPM1133	97.883	97.8973	98.3662	100	98.2699	98.3364	98.3142	98.2871	98.2743	98.2763
JHI1592	98.1693	98.1892	98.7791	98.3412	100	98.7722	99.7613	99.9821	99.9776	99.9613
SM41	97.9133	98.0662	98.8474	98.3103	98.7198	100	98.7916	98.8081	98.8139	98.7891
JHI1253	98.106	98.2295	98.834	98.3774	99.7681	98.8146	100	99.7955	99.7953	99.7929
JHI370	98.2131	98.1478	98.8028	98.3594	99.9743	98.8747	99.7756	100	99.9956	99.9563
JHI387	98.2251	98.1401	98.7802	98.3654	99.9791	98.8966	99.8052	99.9935	100	99.9507
JHI388	98.1993	98.1481	98.7705	98.3473	99.9572	98.8616	99.7995	99.9591	99.9563	100
JHI788	98.1617	98.0513	98.9492	98.4619	99.329	98.8926	99.3156	99.3684	99.3616	99.3541
JHI985	97.9664	97.9921	98.6998	98.2128	98.7563	98.877	98.7607	98.7923	98.7369	98.7564
GB30	98.0742	98.023	98.851	98.4538	99.3301	98.9601	99.3572	99.3175	99.3175	99.3309
JHI782	98.0734	98.0827	98.1462	98.194	98.1948	98.1334	98.2343	98.2364	98.1932	98.1759
JHI783	98.0727	98.0999	98.1639	98.1683	98.2201	98.165	98.2432	98.1984	98.2189	98.2008
SM170C	98.0246	98.0691	98.0888	97.991	98.0182	98.121	98.0571	98.0486	98.0233	98.0007
JHI42	98.1843	97.9993	98.0213	98.0442	98.0796	97.9533	98.0783	98.1352	98.0877	98.1173
JHI979	98.0763	98.0377	98.3068	98.1455	98.1478	98.1384	98.1731	98.1632	98.1347	98.1307
JHI1259	98.0866	97.8562	98.1924	97.9716	97.9931	97.9137	98.1037	98.0141	98.0382	98.011
TA1	97.976	98.0953	98.0757	97.9891	98.0283	98.0919	97.9998	98.0114	97.9662	97.9546
JHI24	98.1181	98.0963	98.4592	98.2683	98.6204	98.5064	98.5461	98.63	98.6193	98.6072
JHI787	98.1018	98.0705	98.5084	98.3155	98.6187	98.4762	98.5782	98.6385	98.6103	98.6236
RCAM1026	98.1331	98.1242	98.6558	98.3704	98.7107	98.7569	98.707	98.7146	98.6857	98.6691
Vh3	96.5441	96.5239	96.5042	96.472	96.5331	96.4738	96.5217	96.5373	96.5594	96.6029
JHI925	96.3425	96.3496	96.4166	96.2407	96.3131	96.3135	96.2725	96.3461	96.3403	96.3397
Vc2	96.4823	96.437	96.5111	96.5484	96.569	96.4731	96.4832	96.5367	96.545	96.5969
JHI1422	97.1796	97.1674	97.0838	96.966	97.1259	97.0744	97.1534	97.1463	97.1231	97.1167
JHI1266	97.5227	97.3092	97.4564	97.1819	97.3895	97.2445	97.4516	97.4	97.4058	97.3693
51A11	93.535	93.5493	93.537	93.5775	93.5431	93.6288	93.5038	93.5438	93.5741	93.5813
51B21	93.5361	93.5599	93.5844	93.5892	93.5441	93.5782	93.4427	93.4773	93.581	93.5819
128C53	93.2671	93.2222	93.2608	93.4586	93.2516	93.2987	93.3587	93.2957	93.2786	93.3075
4292	93.3366	93.3063	93.3494	93.3164	93.3658	93.3095	93.3383	93.4397	93.3909	93.3979
CC283bq	93.2653	93.2215	93.3323	93.3461	93.2062	93.2834	93.1831	93.2551	93.2573	93.2277
USDA2370	93.1814	93.1565	93.2836	93.2809	93.3271	93.2457	93.3015	93.2632	93.2639	93.2444
JHI2442	93.2786	93.2483	93.3397	93.3152	93.2982	93.2794	93.3856	93.2932	93.3304	93.3045
JHI1084	93.1441	93.2133	93.2761	93.3031	93.2138	93.2216	93.3487	93.2645	93.2158	93.2335
JHI944	93.1567	93.1603	93.3374	93.2401	93.2611	93.2603	93.2871	93.254	93.2159	93.2188
OV152	93.3479	93.3458	93.3472	93.2653	93.3598	93.3729	93.3002	93.3119	93.348	93.2846
UPM1137	93.2908	93.295	93.2863	93.4015	93.3524	93.2927	93.369	93.405	93.3674	93.3865
SM149A	93.1441	93.3095	93.2553	93.1762	93.2532	93.3445	93.2511	93.2224	93.2292	93.257
Rt24.2	93.2332	93.2936	93.2945	93.2144	93.3084	93.3512	93.3064	93.3	93.3226	93.3009
CCBAU23252	91.2547	91.2485	91.274	91.2942	91.24	91.285	91.2042	91.3024	91.2288	91.2684
JHI536	91.2355	91.2106	91.2678	91.4581	91.2894	91.2118	91.2477	91.2754	91.2904	91.2525
C15	91.21	91.2677	91.283	91.205	91.2627	91.3094	91.3281	91.2805	91.3179	91.2986
CF307	91.3255	91.2998	91.3088	91.3415	91.3326	91.2879	91.2847	91.3449	91.3433	91.3164

Table S4.5. (Continuation)

Query	JHI788	JHI985	GB30	JHI782	JHI783	SM170C	JHI42	JHI979	JHI1259	TA1
3841	93.6372	93.7397	93.6838	93.6543	93.6985	93.6017	93.6265	93.6979	93.6717	93.6838
SM38	93.6325	93.7382	93.7036	93.6732	93.6698	93.6738	93.6699	93.6312	93.6387	93.7271
JHI960	93.6534	93.713	93.6354	93.6652	93.6258	93.5816	93.5286	93.622	93.6208	93.597
JHI963	93.6582	93.7762	93.6497	93.6369	93.6708	93.6022	93.572	93.6298	93.6717	93.6741
JHI1587	93.6204	93.6979	93.6373	93.6783	93.6497	93.7018	93.6617	93.6863	93.6643	93.7392
JHI1600	93.726	93.7257	93.7322	93.6898	93.6458	93.6587	93.6583	93.6876	93.6901	93.7546
JHI13	93.6205	93.5576	93.6305	93.6797	93.6605	93.5676	93.5795	93.567	93.7152	93.6358
JHI1438	93.7577	93.7265	93.7526	93.7802	93.7587	93.6904	93.6579	93.6934	93.8345	93.7231
JHI535	93.7851	93.7839	93.7704	93.7503	93.8055	93.7055	93.7742	93.7969	93.834	93.7781
JHI585	93.7258	93.7489	93.7483	93.7721	93.748	93.6586	93.774	93.7322	93.7333	93.7721
JHI973	93.7204	93.7414	93.7539	93.7547	93.7464	93.7738	93.7446	93.7516	93.7205	93.6938
JHI974	93.8221	93.7675	93.8026	93.8131	93.7761	93.7615	93.7924	93.7957	93.7745	93.7215
JHI1415	93.7906	93.7342	93.802	93.8452	93.8552	93.7319	93.721	93.7813	93.8712	93.753
SM3	93.6174	93.6139	93.5971	93.617	93.554	93.6119	93.5573	93.491	93.5168	93.6673
VF39	93.6889	93.6382	93.6778	93.5735	93.4983	93.5876	93.5112	93.6063	93.5849	93.5823
L361	93.1656	93.1356	93.1842	93.0802	93.0636	93.0445	93.0667	93.1213	93.0379	93.083
VFF1R2B1	93.5912	93.5358	93.5358	93.5024	93.5402	93.4964	93.4828	93.5312	93.547	93.5981
21A12	93.6227	93.6315	93.6835	93.5203	93.5673	93.602	93.5457	93.6563	93.6302	93.6434
VFF1R2A1	93.6196	93.6453	93.6111	93.5408	93.5767	93.6136	93.5651	93.6301	93.615	93.633
WSM1455	93.5501	93.5588	93.5586	93.5269	93.4901	93.4625	93.4536	93.5439	93.5016	93.5046
WSM1481	93.5115	93.53	93.5527	93.4617	93.451	93.4913	93.4711	93.4793	93.5141	93.5391
WSM78	93.4911	93.4921	93.5866	93.5443	93.5546	93.4647	93.4252	93.4961	93.5946	93.5885
SM152B	93.6232	93.6337	93.6393	93.5737	93.5959	93.6822	93.5591	93.6053	93.6352	93.6486
CC275e	93.69	93.6941	93.6604	93.5944	93.6253	93.6479	93.5318	93.6195	93.6678	93.7372
WSM1325	93.2899	93.3628	93.3638	93.2693	93.2955	93.4467	93.3553	93.3622	93.3534	93.4213
SRDI943	93.4318	93.4711	93.421	93.434	93.4629	93.5162	93.4078	93.4405	93.3919	93.5062
JHLM12A2	93.1894	93.2313	93.2044	93.1806	93.1434	93.1778	93.1406	93.1553	93.1237	93.1604
Vaf12	93.4552	93.4562	93.5048	93.4306	93.3858	93.3554	93.4681	93.3702	93.3903	93.3829
JHI2451	93.5977	93.6185	93.5883	93.5617	93.4934	93.531	93.6424	93.5004	93.4709	93.5715
JHI10	93.5212	93.5853	93.6032	93.4675	93.4896	93.466	93.5329	93.5099	93.5086	93.4513
JHI54	93.3992	93.4481	93.5369	93.4371	93.4191	93.389	93.4959	93.4431	93.3256	93.4721
FA23	93.5384	93.469	93.5421	93.5117	93.4812	93.5005	93.4176	93.5195	93.5628	93.5449
JHI2450	93.5111	93.5731	93.5528	93.5132	93.5084	93.4756	93.6644	93.5545	93.541	93.575
Vaf-108	92.9204	92.8939	92.9206	92.7959	92.7819	92.8438	92.862	92.8327	92.8885	92.9471
JHI2449	92.9696	92.9916	93.0449	93.001	92.9954	92.9681	92.9879	92.9984	92.9947	93.0103
UPM1131	92.9568	92.993	92.9645	93.0333	92.9913	92.922	93.1065	93.0036	93.0109	92.992
CCBAU10279	92.8166	92.8311	92.8383	92.8071	92.7965	92.7743	92.7593	92.7399	92.8404	92.8193
WSM1689	92.8236	92.7702	92.8509	92.9153	92.952	92.931	92.7105	92.86	92.92	92.994
VFF2R2A1	92.9078	92.8069	92.8832	92.8054	92.7725	92.7483	92.6193	92.7736	92.8584	92.824
SPF2A11	92.8354	92.8885	92.8801	92.7184	92.7748	92.792	92.7976	92.9068	92.8534	92.8899
VFCR2A2	92.87	92.7779	92.8598	92.8302	92.7948	92.7661	92.7106	92.8899	92.841	92.8069
248	93.0866	92.9944	93.068	93.0037	92.9993	92.9967	93.0063	92.9948	93.0298	93.0349
63A21	93.1114	92.9977	93.1066	93.0254	92.991	92.9895	92.9593	93.0593	93.0307	93.0086
GLR2	92.9417	92.9806	92.9733	92.9119	92.9216	92.8634	92.8217	92.9205	92.8579	92.9389
CL8	92.7845	92.8106	92.8002	92.7935	92.7547	92.8065	92.785	92.8696	92.8583	92.8297
VFSR2A2	92.8349	92.7909	92.8054	92.8425	92.8645	92.7154	92.8068	92.8887	92.7501	92.7365
SS21	92.7582	92.7753	92.7696	92.7433	92.738	92.7119	92.6799	92.7357	92.7179	92.7653
41A11	92.8381	92.8221	92.8714	92.7938	92.8171	92.8214	92.7506	92.8292	92.9044	92.8402
42B12	92.9155	92.9578	92.9232	92.9123	92.9284	92.8982	92.8884	92.9485	92.9411	92.9504
43A11	92.8155	92.8784	92.8455	92.8164	92.8541	92.8206	92.7848	92.8372	92.8717	92.8958
43B11	92.8519	92.8504	92.878	92.8009	92.7928	92.7615	92.8427	92.8165	92.8516	92.8333
43B12	92.8231	92.8474	92.8091	92.7601	92.8019	92.7037	92.8154	92.8231	92.8107	92.7667
71A12	92.8688	92.8176	92.8551	92.8134	92.8141	92.794	92.7572	92.8918	92.8723	92.8171
73A11	92.8254	92.8554	92.8465	92.76	92.7481	92.7681	92.7814	92.8216	92.8283	92.8471
81B22	92.843	92.8617	92.8449	92.8155	92.7932	92.8253	92.7228	92.8898	92.8804	92.8803
83A12	92.8301	92.8355	92.8945	92.8547	92.8324	92.8067	92.7295	92.8684	92.8993	92.8564
VFHR1A2	92.873	92.7893	92.8255	92.7635	92.7763	92.7173	92.7904	92.8621	92.813	92.8361
VFF1R1A2	92.8122	92.7905	92.7942	92.7678	92.7108	92.6755	92.7128	92.7887	92.7848	92.7781
93B11	92.8563	92.8484	92.8176	92.8094	92.7588	92.7488	92.7591	92.8444	92.7835	92.8629
111A12	92.8267	92.8228	92.7993	92.8363	92.8468	92.7961	92.8163	92.8752	92.8792	92.8596
VFSR2B1	92.7656	92.8307	92.8249	92.8193	92.8121	92.7835	92.7451	92.835	92.8453	92.8552
SPF4F7	92.8224	92.8313	92.8364	92.7788	92.791	92.7536	92.7776	92.7897	92.7722	92.8217
FB206	92.724	92.7287	92.7607	92.6832	92.6967	92.655	92.5305	92.7502	92.6787	92.6923
SL16	92.8137	92.786	92.8443	92.8311	92.7724	92.7461	92.723	92.7531	92.7684	92.7774
HUTR05	92.7243	92.808	92.778	92.7495	92.7547	92.7321	92.673	92.6629	92.7483	92.8357
21B12	92.7841	92.7378	92.7838	92.8163	92.8227	92.8279	92.7304	92.8687	92.8306	92.827
WYCCWR11290	92.6717	92.7213	92.7062	92.7196	92.6941	92.6454	92.6927	92.6507	92.7185	92.663
WYCCWR11317	92.6581	92.6799	92.6814	92.7197	92.6599	92.7336	92.6692	92.663	92.6404	92.7331
CCBAU11080	92.6331	92.558	92.5379	92.5505	92.5404	92.5012	92.4017	92.4893	92.5826	92.4121
CCBAU03386	92.6401	92.5645	92.5402	92.5123	92.4936	92.5396	92.4568	92.4428	92.4925	92.5178

Table S4.5. (Continuation)

Query	JHI788	JHI985	GB30	JHI782	JHI783	SM170C	JHI42	JHI979	JHI1259	TA1
WYCCWR10014	93.0788	93.0942	93.1408	93.1009	93.0595	93.0939	93.0093	93.0113	93.1324	93.1062
JKLM13E	93.1926	93.2007	93.1877	93.1504	93.1077	93.1201	93.0965	93.0981	93.2055	93.1453
Tri-43	92.9907	92.9541	93.0431	92.9355	92.9742	93.0301	92.9727	92.8835	92.9284	93.0128
73B11	92.9812	93.0376	92.9897	92.9297	92.9565	92.9541	92.861	92.9241	92.9226	93.0122
73B12	92.9931	93.0681	93.0041	92.9619	92.9548	92.9638	92.9149	93.0167	93.0273	93.0109
RSP1F2	92.9188	92.9813	92.9728	92.9464	92.9472	92.9326	93.0486	92.8736	92.9315	92.936
Vaf10	92.8438	92.8301	92.833	92.6973	92.6793	92.7936	92.7845	92.7502	92.7512	92.8456
PEPV16	92.8426	92.8771	92.8932	92.9128	92.9175	92.8384	92.7175	92.8312	92.927	92.9121
TOM	92.9012	92.9203	92.8975	92.8971	92.896	92.859	92.8378	92.882	92.8269	92.8851
121B21	92.9015	92.8564	92.9123	92.8536	92.8234	92.8767	92.8192	92.8196	92.7626	92.8645
RSP1F10	92.9333	92.9389	92.9941	93.0201	93.0532	92.9458	92.9166	92.8873	93.1578	92.8886
RSP1A1	92.9892	93.0002	93.0392	92.9818	92.9818	92.9211	92.9287	92.9549	93.1996	92.9582
Norway	94.002	94.0724	94.0177	93.9401	93.9426	93.9532	94.0318	93.9038	93.8768	93.9949
CC278f	93.5236	93.4832	93.6045	93.5012	93.4759	93.4928	93.5461	93.5854	93.4524	93.5184
SM78	94.0792	94.0685	94.1172	94.0801	94.108	94.219	94.0803	94.0883	94.0734	94.1566
SM51	93.9779	94.0134	94.0524	94.0272	93.9897	94.0147	93.9237	94.0297	93.9357	94.0709
SM72	94.0567	94.0004	94.0377	94.0379	93.9831	93.9683	93.8804	93.9756	93.9466	94.0473
Vaf-46	94.8127	94.7376	94.8291	94.7741	94.794	94.6853	94.7384	94.7469	94.7278	94.6838
JHI1093	94.8616	94.858	94.8191	94.7585	94.8044	94.7784	94.826	94.7519	94.7064	94.7776
JHI1096	94.8557	94.8427	94.81	94.8318	94.8205	94.7179	94.8658	94.7794	94.7319	94.7347
GPTR08	93.9315	93.8519	93.9851	93.9237	93.9229	93.8952	93.8209	93.864	93.9099	93.8804
GPTR02	93.9235	93.9467	93.9095	93.8989	93.8709	93.7912	93.8239	93.8014	93.8629	93.8636
SRDI565	93.9264	93.9176	93.9531	93.9408	93.9526	93.8989	93.8111	93.9215	93.9066	93.9489
Ps8	97.9755	97.9391	98.0555	98.3208	98.3071	98.1766	98.4476	98.9505	99.0421	98.9317
JHI1236	98.11	98.087	98.1403	98.0043	98.0044	98.0365	98.0027	98.2234	98.1231	97.9406
JHI953	98.0768	97.9161	98.0699	98.0499	98.0564	98.1396	98.1551	98.0654	98.0689	97.9741
SM147A	98.0191	98.0334	98.0784	98.0516	98.046	98.1	97.9734	98.0281	97.8659	98.0892
JHI1238	98.9249	98.7152	98.8575	98.1571	98.1553	98.1687	98.0115	98.3219	98.2052	98.0884
UPM1133	98.3828	98.2033	98.4657	98.1776	98.1635	98.0217	98.0333	98.0878	97.9701	97.9875
JHI1592	99.2644	98.7718	99.3158	98.1913	98.1978	98.1055	98.0721	98.1229	97.9957	98.0271
SM41	98.776	98.8593	98.9478	98.1099	98.0581	98.0965	97.9392	98.064	97.8887	98.1054
JHI1253	99.2667	98.7874	99.3119	98.2105	98.2212	98.1456	98.0827	98.1413	98.1445	98.0326
JHI370	99.2976	98.777	99.3064	98.2194	98.2054	98.1001	98.1152	98.1109	97.9814	97.9614
JHI387	99.3077	98.7525	99.3204	98.16	98.1841	98.1104	98.071	98.1111	97.9961	98.0048
JHI388	99.2975	98.8192	99.3159	98.2049	98.1835	98.1021	98.0892	98.1452	98.052	98.0363
JHI788	100	98.7757	99.4889	98.1854	98.182	98.1238	98.0225	98.2149	98.0954	97.9655
JHI985	98.6874	100	98.7327	97.987	97.9863	97.9491	97.9381	97.9343	97.9516	97.9282
GB30	99.4963	98.7127	100	98.2612	98.253	98.1098	98.0461	98.1883	98.101	97.9635
JHI782	98.2238	97.9503	98.2438	100	99.9982	98.1392	98.1205	98.4698	98.4297	98.3348
JHI783	98.1916	97.9702	98.2521	99.9984	100	98.1412	98.1271	98.463	98.463	98.3758
SM170C	98.043	97.9108	98.0949	98.1389	98.1255	100	98.4403	98.2296	98.2216	98.4384
JHI42	97.9401	97.9294	98.0389	98.1944	98.1698	98.5188	100	98.384	98.3646	98.3932
JHI979	98.2082	97.9521	98.1766	98.4476	98.4582	98.269	98.3687	100	98.7794	98.7699
JHI1259	98.0613	97.9655	98.1323	98.4225	98.4127	98.1989	98.3847	98.7693	100	99.1358
TA1	97.9688	97.9288	98.0428	98.2939	98.3153	98.4532	98.3431	98.7745	99.1822	100
JHI24	98.51	98.3649	98.5445	98.0891	98.1237	98.1943	98.1707	98.0979	97.9455	98.0471
JHI787	98.531	98.3789	98.5675	98.0924	98.1444	98.2339	98.2112	98.1132	97.9339	98.0688
RCAM1026	98.8227	98.5223	98.8619	98.276	98.2794	98.2283	98.1619	98.3059	98.2028	98.124
Vh3	96.5015	96.6982	96.5706	96.6118	96.6015	96.5433	96.4824	96.5571	96.5548	96.5605
JHI925	96.3904	96.419	96.3926	96.3153	96.3164	96.3406	96.2959	96.4113	96.3906	96.3329
Vc2	96.4883	96.5109	96.4983	96.599	96.5397	96.4265	96.5458	96.5372	96.5809	96.4282
JHI1422	97.071	97.0594	97.0882	97.0755	97.0802	97.0763	97.0511	97.1055	97.0561	96.9738
JHI1266	97.3868	97.1909	97.4291	97.354	97.3307	97.3308	97.1597	97.4136	97.5877	97.2419
51A11	93.5783	93.5864	93.5895	93.5383	93.5679	93.5138	93.5297	93.5247	93.5291	93.5727
51B21	93.5517	93.5948	93.6078	93.5322	93.5439	93.5733	93.5288	93.5654	93.5013	93.5741
128C53	93.3077	93.2591	93.3342	93.2304	93.304	93.2427	93.3518	93.1844	93.4057	93.2251
4292	93.3183	93.3331	93.4014	93.2186	93.2361	93.2999	93.2338	93.2998	93.2888	93.3328
CC283bq	93.2774	93.2194	93.3027	93.1977	93.1943	93.1734	93.1421	93.1182	93.1029	93.2376
USDA2370	93.3773	93.3339	93.3064	93.1639	93.1797	93.1537	93.1879	93.1736	93.3732	93.1117
JHI2442	93.4192	93.3179	93.2969	93.278	93.3099	93.2271	93.1892	93.2309	93.3529	93.2406
JHI1084	93.2802	93.2595	93.3018	93.2393	93.2327	93.1328	93.1158	93.1405	93.2832	93.2188
JHI944	93.2523	93.2763	93.3114	93.2195	93.2282	93.1956	93.1324	93.226	93.2299	93.1991
OV152	93.3335	93.308	93.3499	93.2801	93.2688	93.297	93.2875	93.2912	93.196	93.3149
UPM1137	93.3011	93.3249	93.3887	93.3134	93.3327	93.293	93.3319	93.2843	93.2204	93.2551
SM149A	93.176	93.205	93.2822	93.234	93.2379	93.2874	93.1479	93.2042	93.1934	93.2637
Rt24.2	93.3069	93.2241	93.2712	93.2472	93.1898	93.3865	93.2541	93.2127	93.207	93.3288
CCBAU23252	91.2833	91.2635	91.2387	91.2294	91.2365	91.2066	91.2518	91.2072	91.167	91.2508
JHI536	91.2973	91.3173	91.2918	91.3076	91.2686	91.1771	91.4361	91.2839	91.2788	91.309
C15	91.2968	91.2672	91.2903	91.2648	91.2203	91.2936	91.1915	91.2141	91.2147	91.3148
CF307	91.2971	91.3707	91.2802	91.273	91.3015	91.3733	91.2835	91.2777	91.2995	91.3675

Table S4.5. (Continuation)

Query	JHI24	JHI787	RCAM1026	Vh3	JHI925	Vc2	JHI1422	JHI1266	51A11	51B21
3841	93.6071	93.584	93.6823	93.8078	94.0167	93.7293	93.7494	93.7827	95.9712	95.9694
SM38	93.6259	93.6223	93.6987	93.818	93.7028	93.7442	93.7445	93.6367	95.9919	96.0113
JHI960	93.6273	93.6186	93.697	93.7608	94.0657	93.7068	93.7157	93.6975	95.9837	95.9804
JHI963	93.6479	93.6231	93.7435	93.8398	94.0199	93.6892	93.7281	93.6861	95.9672	95.9761
JHI1587	93.6363	93.6716	93.7289	93.8243	94.0028	93.7183	93.7394	93.7429	95.9327	95.9355
JHI1600	93.697	93.7073	93.7443	93.8391	94.0033	93.7662	93.8603	93.7347	95.9618	95.9605
JHI13	93.6576	93.6634	93.7076	93.7178	93.7283	93.7887	93.6928	93.7461	95.8812	95.896
JHI1438	93.7023	93.7119	93.7495	93.8654	93.692	93.8159	93.7872	93.8461	95.8247	95.8239
JHI535	93.7943	93.8111	93.8623	93.7925	94.0924	93.8624	93.8677	93.8019	96.0354	96.0481
JHI585	93.7393	93.7742	93.8388	93.8346	94.0347	93.9049	93.8728	93.7716	96.0478	96.0641
JHI973	93.694	93.6934	93.754	93.848	94.0232	93.8157	93.9543	93.7514	96.0436	96.0333
JHI974	93.6859	93.7138	93.824	93.836	93.9867	93.7945	93.9479	93.7685	96.0363	96.0053
JHI1415	93.7131	93.7439	93.8243	93.9404	94.1124	93.8642	93.9237	93.7832	95.9933	96.0374
SM3	93.4998	93.4774	93.5958	93.713	93.6794	93.6358	93.735	93.5265	95.9074	95.8952
VF39	93.5414	93.5527	93.7081	93.6742	93.8694	93.6105	93.6918	93.5925	95.9284	95.9725
L361	93.1735	93.1817	93.1705	93.1267	93.1014	93.1102	93.1199	93.0972	94.5534	94.54
VFF1R2B1	93.4085	93.4177	93.5625	93.581	93.5657	93.6292	93.5982	93.6312	98.2085	98.2222
21A12	93.5199	93.5259	93.6019	93.5972	93.6143	93.5706	93.7415	93.5705	98.3665	98.3699
VFF1R2A1	93.6143	93.5935	93.6269	93.6231	93.6096	93.5252	93.7301	93.5965	98.3818	98.3738
WSM1455	93.4147	93.463	93.5361	93.5731	93.4763	93.4745	93.5633	93.5657	98.1496	98.1875
WSM1481	93.4669	93.5018	93.5252	93.5204	93.4868	93.4547	93.5706	93.5489	98.1267	98.1094
WSM78	93.4393	93.4305	93.5729	93.47	93.4377	93.4337	93.5245	93.5469	93.6135	93.6014
SM152B	93.5253	93.499	93.6443	93.5944	93.5242	93.5798	93.617	93.6237	93.6685	93.6862
CC275e	93.593	93.5795	93.6901	93.7241	93.5933	93.567	93.6693	93.5786	93.6906	93.6818
WSM1325	93.3816	93.3542	93.3749	93.3354	93.2779	93.2361	93.4126	93.4117	94.0621	94.0863
SRDI943	93.4576	93.4626	93.4424	93.3636	93.3548	93.3335	93.4598	93.4543	94.2483	94.2191
JHLM12A2	93.1885	93.238	93.3194	93.2187	93.1718	93.1985	93.2355	93.2052	94.6593	94.6471
Vaf12	93.5828	93.5771	93.5288	93.423	93.3429	93.5463	93.4618	93.4031	95.2482	95.2433
JHI2451	93.7735	93.7891	93.6032	93.6045	93.4787	93.7126	93.5574	93.5036	95.582	95.5851
JHI10	93.5993	93.5847	93.5514	93.6283	93.4673	93.5689	93.5403	93.5797	95.5843	95.5932
JHI54	93.4629	93.4734	93.5298	93.4942	93.3514	93.5292	93.458	93.3544	95.4519	95.4583
FA23	93.4491	93.461	93.5812	93.5588	93.4159	93.5331	93.5646	93.6616	95.5729	95.5899
JHI2450	93.8267	93.8204	93.5546	93.6061	93.5203	93.8776	93.5838	93.556	95.6411	95.6648
Vaf-108	92.9869	92.9661	92.9457	92.8829	92.7881	92.8847	92.9013	92.902	94.3005	94.3108
JHI2449	93.0244	93.017	93.0462	92.9945	93.0631	93.0051	92.9766	93.0182	94.4633	94.4924
UPM1131	93.1108	93.1099	93.0221	92.9861	92.8962	93.0578	93.0048	93.0162	94.4359	94.4359
CCBAU10279	92.8329	92.8043	92.8322	92.8289	92.7826	92.8101	92.8495	92.8722	94.2846	94.3313
WSM1689	92.8194	92.8302	92.8546	92.8889	92.8092	92.7762	92.8707	92.9367	94.4185	94.3916
VFF2R2A1	92.6943	92.6881	92.8468	92.7648	92.8623	92.8654	92.8562	92.9406	94.2938	94.2872
SPF2A11	92.7715	92.803	92.9034	92.8646	92.8867	92.812	92.8855	92.8821	94.2985	94.3491
VFCR2A2	92.6814	92.7069	92.8194	92.8071	92.8306	92.7562	92.8266	92.8702	94.2471	94.2962
248	93.0638	93.066	93.0653	93.1012	93.1134	92.9805	93.0946	93.1425	94.4743	94.4922
63A21	93.028	93.0376	93.0916	93.0502	93.0942	93.0376	93.0654	93.0595	94.4167	94.4547
GLR2	92.8684	92.8391	92.9537	93.0142	92.9579	92.9292	92.9841	92.9873	94.3781	94.4062
CL8	92.889	92.9058	92.8036	92.825	92.8215	92.8629	92.8245	92.8785	94.3492	94.3797
VFSR2A2	92.823	92.8231	92.8377	92.8053	93.0709	92.7941	92.8101	92.8618	94.3966	94.4095
SS21	92.7522	92.7755	92.7638	92.8231	92.6804	92.6938	92.8255	92.7287	94.2643	94.3182
41A11	92.7802	92.7748	92.8642	92.7842	92.9115	92.8057	92.8888	92.9593	94.3167	94.3479
42B12	92.8521	92.885	92.9542	92.84	92.9526	92.8938	92.9541	92.9607	94.3713	94.3948
43A11	92.744	92.8006	92.8518	92.8322	92.8825	92.8105	92.8397	92.8912	94.3508	94.3421
43B11	92.8836	92.905	92.8736	92.842	93.0882	92.8499	92.7838	92.8594	94.3145	94.3574
43B12	92.8495	92.8305	92.8342	92.7603	93.0393	92.9041	92.8111	92.8681	94.32	94.3327
71A12	92.792	92.8055	92.8777	92.8641	92.8722	92.8031	92.8561	92.877	94.3295	94.4083
73A11	92.7905	92.797	92.8304	92.7581	92.8379	92.7605	92.8206	92.8563	94.3405	94.3635
81B22	92.7965	92.7862	92.8323	92.8694	92.8676	92.7887	92.8382	92.8571	94.2784	94.3293
83A12	92.7985	92.7946	92.8661	92.7936	92.854	92.8039	92.8534	92.8851	94.2671	94.3343
VFHR1A2	92.8578	92.8214	92.865	92.8189	92.7846	92.8215	92.8206	92.9411	94.4052	94.4207
VFF1R1A2	92.7392	92.7405	92.7872	92.7625	92.8806	92.7191	92.8018	92.7705	94.3281	94.3534
93B11	92.8544	92.8633	92.8322	92.8301	92.878	92.7392	92.8182	92.8235	94.3546	94.364
111A12	92.7851	92.7946	92.8408	92.8273	92.8777	92.8041	92.8466	92.9342	94.3799	94.4416
VFSR2B1	92.8224	92.8527	92.8395	92.8369	92.8246	92.7883	92.7829	92.8654	94.3631	94.3784
SPF4F7	92.7675	92.7681	92.8742	92.8756	92.9117	92.7502	92.7617	92.8834	94.3587	94.3746
FB206	92.6127	92.6333	92.7843	92.7561	92.8118	92.73	92.7956	92.7331	94.3244	94.3453
SL16	92.8001	92.806	92.8207	92.7355	92.7177	92.7706	92.7778	92.8019	94.3719	94.4169
HUTR05	92.7725	92.792	92.7923	92.7593	92.7259	92.7336	92.8182	92.788	94.2784	94.2986
21B12	92.8119	92.7985	92.7895	92.8145	92.9687	92.8197	92.8114	92.8467	94.4019	94.4046
WYCCWR11290	92.68	92.6944	92.7094	92.5964	92.6413	92.7716	92.6196	92.6751	93.7026	93.7654
WYCCWR11317	92.6976	92.749	92.6787	92.6481	92.6437	92.6864	92.7166	92.6918	93.7956	93.8561
CCBAU11080	92.5826	92.5453	92.5479	92.564	92.4564	92.5798	92.5812	92.6713	93.7129	93.7368
CCBAU03386	92.5208	92.4954	92.5859	92.5133	92.5009	92.5548	92.5575	92.5091	93.8064	93.8164

Table S4.5. (Continuation)

Query	JHI24	JHI787	RCAM1026	Vh3	JHI925	Vc2	JHI1422	JHI1266	51A11	51B21
WYCCWR10014	93.0066	92.9889	93.0881	92.9482	93.0052	92.9833	93.0706	93.0875	94.0545	94.0446
JKLM13E	93.0799	93.115	93.2435	93.1939	93.0865	93.1521	93.1707	93.1548	94.7003	94.6419
Tri-43	93.0252	93.0052	93.0807	93.008	92.8935	93.002	92.9531	92.9795	94.2247	94.2616
73B11	92.9354	92.9484	93.0497	92.964	92.9436	92.8962	92.9796	92.9712	94.4534	94.4944
73B12	92.9995	92.9893	93.0262	92.9714	93.013	92.9915	93.0378	93.0413	94.5386	94.5718
RSP1F2	93.1778	93.1129	93.0491	92.9585	92.8224	93.0432	93.0203	92.9742	94.4726	94.4463
Vaf10	92.844	92.849	92.902	92.8726	92.7243	92.8115	92.84	92.6919	94.2687	94.3097
PEPV16	92.8412	92.8365	92.8767	92.919	92.7303	92.8574	92.8287	92.8124	94.2857	94.3217
TOM	92.8135	92.8241	92.9128	92.9593	92.8636	92.8755	92.9021	92.8497	94.3101	94.4008
121B21	92.86	92.8512	92.8444	92.8837	92.7919	92.8209	92.9556	92.8828	94.4767	94.4476
RSP1F10	93.0313	93.0489	93.024	92.9707	92.8461	93.0484	92.9896	93.2702	94.3595	94.3927
RSP1A1	93.1052	93.1064	93.0476	92.9859	92.7985	93.0254	93.1606	93.1591	94.4498	94.4756
Norway	94.1177	94.1337	94.042	93.8431	93.6113	93.7652	93.8882	93.8008	93.2531	93.2852
CC278f	93.5477	93.6069	93.6917	93.4078	93.2524	93.3883	93.4782	93.4812	93.1639	93.1548
SM78	94.1315	94.1296	94.1645	93.8884	93.804	93.8267	94.0021	94.0204	93.2236	93.2316
SM51	93.9612	93.9715	94.0845	93.8031	93.6759	93.7156	93.8568	93.9502	93.1023	93.1499
SM72	94.0264	94.0532	94.0931	93.8552	93.7587	93.7638	93.9388	93.9302	93.2198	93.2644
Vaf-46	94.7515	94.7467	94.8057	94.7145	94.5234	94.7818	94.7003	94.6347	93.1748	93.1901
JHI1093	94.8315	94.9028	94.8916	94.762	94.659	94.8644	94.737	94.7155	93.1564	93.1407
JHI1096	94.8351	94.8915	94.9212	94.7619	94.6491	94.8696	94.7186	94.7276	93.1808	93.151
GPTR08	93.8102	93.8177	93.9671	93.8693	93.7744	93.91	93.857	93.8373	92.4681	92.5028
GPTR02	93.8036	93.82	93.9754	93.933	93.8805	93.8393	93.872	93.8455	92.6411	92.6986
SRDI565	93.9771	93.9515	93.9845	93.9787	93.8643	93.9165	93.9175	93.9284	92.5609	92.6255
Ps8	97.966	97.9791	98.1036	96.5699	96.4763	96.571	97.1985	97.4607	93.5196	93.547
JHI1236	98.0653	98.1051	98.1523	96.6288	96.5166	96.565	97.2906	97.5903	93.628	93.653
JHI953	98.1498	98.1417	98.1256	96.5724	96.3762	96.4825	97.1908	97.5449	93.6263	93.649
SM147A	98.1304	98.1392	98.0826	96.4848	96.4405	96.458	97.1386	97.3305	93.5417	93.5188
JHI1238	98.5025	98.5243	98.6569	96.5415	96.4652	96.5873	97.1276	97.4669	93.5801	93.6027
UPM1133	98.296	98.2814	98.3667	96.4919	96.2895	96.514	96.8601	97.1485	93.5258	93.5609
JHI1592	98.6458	98.6303	98.6592	96.5464	96.3575	96.5458	97.0887	97.367	93.5772	93.615
SM41	98.4465	98.4549	98.7236	96.4628	96.3042	96.4458	97.0236	97.2044	93.5899	93.6019
JHI1253	98.5581	98.5499	98.6823	96.532	96.2813	96.4812	97.1637	97.4172	93.5481	93.5683
JHI370	98.6119	98.6077	98.6727	96.5364	96.421	96.5205	97.1498	97.4008	93.5637	93.6044
JHI387	98.6021	98.6111	98.6684	96.5715	96.401	96.5088	97.1475	97.3993	93.5502	93.5822
JHI388	98.6206	98.6345	98.6323	96.5563	96.429	96.5463	97.0894	97.3912	93.5644	93.5875
JHI788	98.6301	98.5981	98.8329	96.5732	96.4314	96.5426	97.1104	97.5089	93.5985	93.6198
JHI985	98.425	98.4261	98.4573	96.6616	96.3705	96.4412	97.0784	97.1912	93.5671	93.6156
GB30	98.5494	98.5476	98.819	96.5633	96.4128	96.5375	97.1357	97.4144	93.578	93.5961
JHI782	98.1356	98.1595	98.2943	96.5832	96.4377	96.579	97.1043	97.3564	93.4888	93.4774
JHI783	98.1114	98.1483	98.2777	96.5736	96.4339	96.5838	97.077	97.354	93.4846	93.4995
SM170C	98.1503	98.1341	98.1205	96.5071	96.3493	96.4268	97.0656	97.354	93.4721	93.4781
JHI42	98.146	98.1521	98.1399	96.5668	96.364	96.555	97.0606	97.1892	93.5948	93.5993
JHI979	98.115	98.127	98.3003	96.5966	96.4641	96.5939	97.096	97.4548	93.5507	93.563
JHI1259	97.9336	97.9576	98.1608	96.5843	96.4565	96.6016	97.0532	97.5828	93.5902	93.5997
TA1	98.0586	98.0178	98.0893	96.5472	96.3957	96.4706	97.0055	97.2569	93.5476	93.5852
JHI24	100	99.9975	99.1347	96.493	96.3868	96.6716	97.0053	97.2132	93.5502	93.5676
JHI787	99.9982	100	99.1379	96.4925	96.357	96.7085	97.0079	97.2185	93.5637	93.5188
RCAM1026	99.2329	99.2176	100	96.6059	96.5123	96.633	97.1385	97.409	93.6218	93.6393
Vh3	96.4977	96.4652	96.5662	100	97.9797	98.0618	96.6983	96.5722	93.5976	93.6104
JHI925	96.3312	96.3042	96.4634	97.9325	100	99.1397	96.5601	96.313	93.5395	93.5398
Vc2	96.671	96.6607	96.5312	98.0409	99.1611	100	96.5744	96.5798	93.567	93.582
JHI1422	97.0409	97.0467	97.1315	96.6583	96.5637	96.6266	100	98.4693	93.6539	93.6867
JHI1266	97.198	97.1795	97.3873	96.5664	96.4551	96.5269	98.4929	100	93.5478	93.5745
51A11	93.5118	93.5316	93.5997	93.5755	93.5721	93.5677	93.6512	93.5736	100	99.9942
51B21	93.5341	93.5281	93.584	93.5906	93.5965	93.5945	93.6708	93.5774	99.9946	100
128C53	93.372	93.3071	93.3129	93.171	93.0004	93.3038	93.27	93.3103	92.689	92.7021
4292	93.3603	93.3685	93.3961	93.2588	93.1902	93.3008	93.3876	93.3464	92.7793	92.8563
CC283bq	93.2175	93.1798	93.329	93.1762	93.0588	93.1537	93.2973	93.1017	92.7653	92.7487
USDA2370	93.3868	93.3549	93.3443	93.2343	93.2087	93.3174	93.2415	93.2963	92.6677	92.7141
JHI2442	93.4367	93.4327	93.3531	93.2172	93.1539	93.369	93.3648	93.3657	92.8088	92.8089
JHI1084	93.2495	93.2667	93.3868	93.1125	93.1906	93.2242	93.3158	93.3835	92.7492	92.7255
JHI944	93.2618	93.2206	93.3642	93.2235	93.6716	93.1291	93.3882	93.264	92.7406	92.7969
OV152	93.3978	93.3744	93.377	93.2379	93.1645	93.2336	93.2815	93.2452	92.7529	92.7927
UPM1137	93.4608	93.4475	93.4432	93.224	93.2121	93.2927	93.2885	93.2934	92.8012	92.8095
SM149A	93.2421	93.2447	93.2995	93.1431	93.1283	93.1308	93.2833	93.2598	92.7514	92.8087
Rt24.2	93.3143	93.3289	93.3037	93.1433	93.1828	93.15	93.2886	93.2312	92.8287	92.8427
CCBAU23252	91.2851	91.2736	91.3383	91.2522	91.2663	91.3184	91.3136	91.233	91.1183	91.1682
JHI536	91.5172	91.5742	91.2867	91.3032	91.2554	91.583	91.3944	91.3211	91.2298	91.2477
C15	91.2702	91.2541	91.3458	91.2989	91.3335	91.2841	91.3395	91.2936	91.1275	91.1505
CF307	91.3653	91.3578	91.3125	91.3111	91.3555	91.3713	91.4339	91.3474	91.2029	91.2408

Table S4.5. (Continuation)

Query	128C53	4292	CC283bq	USDA2370	JHI2442	JHI1084	JHI944	OV152	UPM1137	SM149A
3841	92.8378	92.8446	92.7754	92.8124	92.9087	92.8526	93.2252	92.9331	92.8565	92.9871
SM38	92.7346	92.7937	92.8311	92.6823	92.7802	92.724	92.7605	92.9354	92.7667	92.8751
JHI960	92.7171	92.7483	92.5664	92.7916	92.8338	92.7792	93.1765	92.8863	92.7542	92.7168
JHI963	92.8045	92.8016	92.67	92.7169	92.9023	92.826	93.1086	92.8985	92.834	92.8208
JHI1587	92.8239	92.8393	92.7383	92.7618	92.8798	92.818	93.0832	92.8799	92.8028	92.7676
JHI1600	92.8663	92.8932	92.7947	92.7668	92.8527	92.828	93.0653	92.8602	92.9142	92.8083
JHI13	93.0031	92.7823	92.6388	92.9773	93.0153	92.8312	92.7711	92.838	92.886	92.8867
JHI1438	92.9044	92.7793	92.7197	93.0439	92.9445	92.8298	92.839	92.8434	92.9257	92.8215
JHI535	92.8186	92.8229	92.7882	92.7669	92.9045	92.7697	93.1585	92.9128	92.9159	92.8132
JHI585	92.8393	92.8956	92.8584	92.7692	92.8915	92.8253	93.1684	92.9623	92.9161	92.89
JHI973	92.8122	92.8043	92.8223	92.8217	92.8686	92.8563	93.1537	92.9909	92.9093	92.824
JHI974	92.7874	92.8313	92.8141	92.7584	92.8571	92.7877	93.1125	92.9301	92.884	92.7101
JHI1415	92.8719	92.8995	92.8186	92.7883	92.8674	92.8046	93.1787	92.9503	92.8932	92.7641
SM3	92.6531	92.7162	92.7565	92.5344	92.6511	92.6164	92.6756	92.927	92.6924	92.8117
VF39	92.7427	92.7691	92.8352	92.8049	92.8177	92.7072	93.1329	92.8599	92.8851	92.7553
L361	92.428	92.3677	92.3936	92.4192	92.4711	92.4281	92.451	92.4479	92.4207	92.4411
VFF1R2B1	92.7524	92.7447	92.7191	92.8012	92.8344	92.7905	92.8566	92.7785	92.8561	92.7399
21A12	92.843	92.8605	92.8339	92.7538	92.8684	92.8205	92.871	92.808	92.8863	92.812
VFF1R2A1	92.854	92.8675	92.8759	92.8028	92.8943	92.8162	92.8815	92.8039	92.8883	92.815
WSM1455	92.7246	92.7303	92.7157	92.7676	92.7974	92.7169	92.7948	92.6765	92.7068	92.6858
WSM1481	92.6469	92.6376	92.7093	92.7282	92.7381	92.6818	92.7695	92.7363	92.6995	92.6118
WSM78	93.481	93.4656	93.4715	93.4024	93.4646	93.4065	93.423	93.4056	93.4604	93.5192
SM152B	93.3835	93.3736	93.415	93.3293	93.3394	93.3193	93.334	93.4871	93.3506	93.4572
CC275e	93.3772	93.4641	93.3791	93.3868	93.4379	93.332	93.4062	93.429	93.3848	93.3218
WSM1325	92.7777	92.8406	92.7641	92.7109	92.8656	92.7765	92.796	92.7712	92.784	92.7988
SRDI943	92.7705	92.8418	92.7861	92.7653	92.8545	92.7828	92.8055	92.8464	92.8144	92.8478
JHLM12A2	92.9105	92.8314	92.7829	92.8263	92.9554	93.2303	92.8723	92.932	92.9132	92.8045
Vaf12	92.9248	92.8296	92.718	92.9109	92.9321	92.7407	92.8653	92.9318	93.045	92.7721
JHI2451	92.9593	92.863	92.7875	93.0035	93.0254	92.8691	92.8092	92.9727	93.0067	92.8144
JHI10	92.9329	92.8664	92.7141	93.206	92.9661	92.8012	92.8931	92.8781	93.0111	92.8442
JHI54	92.8536	92.7582	92.6209	92.8022	92.8787	92.623	92.7538	92.869	92.9458	92.7246
FA23	93.0218	93.3349	92.8717	92.9035	92.9781	92.9054	92.8585	92.9739	92.9195	92.9812
JHI2450	93.079	92.9198	92.8598	93.0038	93.1125	92.93	92.8774	92.9996	93.007	92.9371
Vaf-108	92.3362	92.283	92.3674	92.2754	92.3174	92.1419	92.3123	92.2987	92.3784	92.319
JHI2449	92.5768	92.4635	92.3221	92.4316	92.5423	92.6182	92.5238	92.4605	92.4815	92.4549
UPM1131	92.5705	92.4827	92.4544	92.5651	92.5546	92.4511	92.4336	92.5023	92.5918	92.4027
CCBAU10279	92.2154	92.2339	92.1226	92.201	92.2444	92.1844	92.2014	92.261	92.2797	92.1945
WSM1689	92.26	92.3336	92.3262	92.2373	92.39	92.3322	92.4204	92.2512	92.3053	92.4225
VFF2R2A1	92.2862	92.1514	92.2122	92.2994	92.387	92.3013	92.3723	92.1662	92.2601	92.235
SPF2A11	92.2923	92.2033	92.2206	92.2582	92.3305	92.2655	92.3877	92.2742	92.2874	92.2619
VFCR2A2	92.1887	92.2154	92.1324	92.1911	92.2732	92.1554	92.3789	92.1337	92.232	92.1536
248	92.4684	92.4061	92.36	92.3352	92.4373	92.4577	92.5065	92.4562	92.4482	92.3756
63A21	92.5918	92.5201	92.4305	92.5545	92.5691	92.5547	92.6611	92.4154	92.5293	92.457
GLR2	92.334	92.3187	92.2776	92.3796	92.3301	92.3364	92.3941	92.3442	92.3654	92.2861
CL8	92.288	92.28	92.2129	92.2778	92.3385	92.2438	92.3446	92.3245	92.2999	92.2243
VFSR2A2	92.2753	92.2648	92.2198	92.269	92.3232	92.2616	92.5739	92.1717	92.2864	92.2759
SS21	92.1971	92.1755	92.0828	92.237	92.2418	92.17	92.2324	92.2388	92.2857	92.1344
41A11	92.3939	92.311	92.2597	92.2987	92.3955	92.3785	92.4466	92.3059	92.2848	92.3354
42B12	92.4447	92.3436	92.3179	92.3879	92.4326	92.4281	92.4835	92.3554	92.4003	92.3846
43A11	92.3223	92.2581	92.2419	92.2675	92.3589	92.2889	92.4275	92.26	92.3286	92.2904
43B11	92.3094	92.2813	92.0971	92.244	92.3242	92.2649	92.5345	92.2014	92.2704	92.1971
43B12	92.3075	92.2908	92.166	92.2738	92.3141	92.2666	92.609	92.1811	92.2946	92.1703
71A12	92.2939	92.3078	92.2676	92.2837	92.3756	92.3024	92.3966	92.2615	92.3078	92.2682
73A11	92.3178	92.2992	92.1971	92.2388	92.3338	92.232	92.4084	92.3005	92.3102	92.2191
81B22	92.277	92.3035	92.2552	92.2605	92.4286	92.2658	92.3932	92.2238	92.3094	92.265
83A12	92.3095	92.2807	92.3496	92.2935	92.3736	92.2752	92.3854	92.1993	92.3146	92.2299
VFHR1A2	92.285	92.3179	92.2213	92.3092	92.3674	92.3299	92.3808	92.3318	92.2987	92.2479
VFF1R1A2	92.22	92.2415	92.1462	92.1795	92.2979	92.1529	92.4349	92.3006	92.258	92.1944
93B11	92.3069	92.2904	92.1684	92.2506	92.2904	92.2199	92.3745	92.2673	92.2656	92.2804
111A12	92.3475	92.2492	92.1984	92.2674	92.347	92.2729	92.3987	92.2255	92.2704	92.2715
VFSR2B1	92.2919	92.2257	92.1845	92.2332	92.3703	92.2692	92.3376	92.2282	92.2703	92.2516
SPF4F7	92.2876	92.2569	92.2136	92.211	92.35	92.2199	92.4709	92.27	92.2899	92.2153
FB206	92.271	92.176	92.1473	92.1652	92.2722	92.1456	92.403	92.0705	92.2687	92.1534
SL16	92.1943	92.22	92.12	92.1893	92.2341	92.1746	92.2104	92.2665	92.2225	92.1466
HUTR05	92.2625	92.31	92.1795	92.1896	92.2343	92.1949	92.2113	92.2586	92.2695	92.254
21B12	92.3062	92.2878	92.2133	92.2595	92.2895	92.2278	92.3302	92.2408	92.2972	92.2111
WYCCWR11290	92.5328	92.4396	92.3021	92.4397	92.4975	92.5288	92.4045	92.3926	92.4531	92.3808
WYCCWR11317	92.4772	92.4524	92.3616	92.4289	92.5094	92.3843	92.3827	92.5029	92.4142	92.4727
CCBAU11080	92.3535	92.3674	92.3028	92.4783	92.4905	92.7597	92.4277	92.3725	92.4734	92.3036
CCBAU03386	92.4124	92.3722	92.294	92.3595	92.3768	92.2825	92.3745	92.4944	92.4822	92.3146

Table S4.5. (Continuation)

Query	128C53	4292	CC283bq	USDA2370	JHI2442	JHI1084	JHI944	OV152	UPM1137	SM149A
WYCCWR10014	93.4422	93.5168	93.4283	93.4181	93.4866	93.3699	93.3964	93.4797	93.4469	93.4953
JKLM13E	92.883	92.8482	92.849	92.9034	92.974	93.2453	92.8452	92.9211	92.9652	92.8278
Tri-43	92.7495	92.7337	92.6913	92.7849	92.7677	92.7386	92.7111	92.8764	92.7144	92.7814
73B11	92.4695	92.4318	92.4227	92.4599	92.5823	92.4707	92.5733	92.3982	92.5025	92.345
73B12	92.564	92.5018	92.4438	92.4952	92.5842	92.5057	92.5886	92.4841	92.5009	92.4658
RSP1F2	92.7419	92.5271	92.4153	92.7247	92.6759	92.4612	92.5289	92.5402	92.6381	92.4276
Vaf10	92.3563	92.2161	92.4073	92.3315	92.3156	92.1471	92.3313	92.1797	92.3568	92.2763
PEPV16	92.4046	92.3262	92.2635	92.3567	92.3498	92.3847	92.3129	92.3778	92.3916	92.2509
TOM	92.4439	92.407	92.3181	92.456	92.3903	92.4448	92.4062	92.3379	92.4306	92.3272
121B21	92.4807	92.386	92.3329	92.3909	92.4214	92.3541	92.4547	92.3889	92.3992	92.3055
RSP1F10	92.64	92.3399	92.2434	92.7872	92.6411	92.576	92.427	92.3772	92.508	92.4593
RSP1A1	92.6407	92.3873	92.3852	92.814	92.7148	92.5363	92.4962	92.4225	92.512	92.4969
Norway	94.5347	94.5737	94.5535	94.5429	94.539	94.4451	94.4518	94.6572	94.5724	94.4462
CC278f	95.0145	95.138	94.9162	95.0027	95.0705	94.8715	94.9348	95.1128	95.0326	94.9388
SM78	94.8241	94.902	94.8242	94.8335	94.8038	94.7876	94.8535	94.9799	94.8592	95.1643
SM51	94.7559	94.7834	94.7501	94.6574	94.7765	94.6985	94.7879	94.8126	94.7535	95.0441
SM72	94.873	94.9196	94.9234	94.8188	94.914	94.834	94.8954	94.9813	94.8938	95.1955
Vaf-46	92.8687	92.801	92.8493	92.7124	92.9337	92.7298	92.7524	92.8945	92.9289	92.7225
JHI1093	93.1404	92.922	92.8799	92.9178	93.0357	92.8417	92.8197	92.8721	93.0594	92.821
JHI1096	93.1435	92.9196	92.8133	92.9218	93.0588	92.8369	92.839	92.8261	93.0514	92.7735
GPTR08	92.2197	92.3101	92.2936	92.2103	92.2808	92.147	92.2652	92.1059	92.2529	92.3789
GPTR02	92.2329	92.2245	92.2262	92.0957	92.2268	92.1455	92.2134	92.1457	92.1641	92.2974
SRDI565	92.3032	92.296	92.2985	92.3073	92.3903	92.2811	92.3579	92.2292	92.2923	92.4054
Ps8	93.2996	93.2903	93.1696	93.2088	93.2876	93.2007	93.2759	93.2583	93.2558	93.1156
JHI1236	93.253	93.3963	93.1982	93.2993	93.3404	93.2082	93.2936	93.3839	93.3048	93.2741
JHI953	93.2973	93.3505	93.2393	93.2488	93.309	93.1918	93.2707	93.3473	93.3216	93.2265
SM147A	93.2102	93.2509	93.2437	93.1787	93.2071	93.1381	93.2177	93.319	93.2693	93.2317
JHI1238	93.3303	93.3617	93.3224	93.3459	93.4186	93.2503	93.2733	93.346	93.3123	93.2612
UPM1133	93.468	93.3135	93.2618	93.2848	93.3645	93.2251	93.2397	93.2672	93.3271	93.1192
JHI1592	93.3117	93.3462	93.2801	93.2759	93.297	93.1566	93.2354	93.3822	93.3567	93.2288
SM41	93.2574	93.2747	93.2305	93.2077	93.2444	93.1694	93.235	93.2913	93.27	93.2885
JHI1253	93.3599	93.356	93.2412	93.3442	93.4567	93.3258	93.2572	93.3167	93.335	93.2758
JHI370	93.2673	93.3608	93.2725	93.2265	93.3441	93.1077	93.2049	93.3415	93.3818	93.2086
JHI387	93.3123	93.3201	93.2318	93.2185	93.2978	93.1037	93.2152	93.2987	93.4135	93.2312
JHI388	93.3113	93.3195	93.2521	93.2498	93.3196	93.1456	93.2519	93.3512	93.4022	93.2347
JHI788	93.369	93.3605	93.3144	93.4121	93.4327	93.3874	93.3973	93.4103	93.4128	93.3111
JHI985	93.1955	93.2296	93.1992	93.1493	93.2355	93.2114	93.1986	93.2748	93.242	93.1535
GB30	93.3317	93.3468	93.2267	93.311	93.3549	93.276	93.2894	93.3195	93.345	93.2468
JHI782	93.1669	93.1698	93.1189	93.1658	93.2161	93.1178	93.184	93.2233	93.1924	93.1316
JHI783	93.1665	93.1993	93.1392	93.1868	93.1767	93.1434	93.2559	93.2606	93.2791	93.2332
SM170C	93.2053	93.2561	93.2134	93.1336	93.2403	93.0472	93.1842	93.263	93.1845	93.1896
JHI42	93.4215	93.2934	93.2141	93.3011	93.2665	93.1337	93.2451	93.3003	93.336	93.1418
JHI979	93.1986	93.2822	93.1463	93.1647	93.2637	93.0965	93.2238	93.2415	93.3155	93.1805
JHI1259	93.3968	93.2376	93.0855	93.2802	93.3765	93.2994	93.2055	93.1667	93.2299	93.2007
TA1	93.2221	93.2557	93.2302	93.1273	93.2628	93.1134	93.1837	93.3196	93.1956	93.2627
JHI24	93.3885	93.2901	93.2279	93.2909	93.3603	93.2064	93.2004	93.3304	93.4008	93.167
JHI787	93.451	93.3395	93.2359	93.3457	93.4227	93.2389	93.2473	93.2806	93.4379	93.1885
RCAM1026	93.428	93.3985	93.3467	93.3895	93.4391	93.3912	93.3445	93.3601	93.4135	93.3607
Vh3	93.1935	93.2023	93.1489	93.1847	93.2554	93.1737	93.1214	93.2461	93.2812	93.1197
JHI925	93.1259	93.1845	92.9503	93.1806	93.0813	93.0309	93.6359	93.1149	93.1011	93.0072
Vc2	93.2857	93.2318	93.1761	93.3171	93.3619	93.1618	93.1194	93.2302	93.2668	93.0908
JHI1422	93.2745	93.2909	93.2143	93.2832	93.3478	93.2375	93.3192	93.2889	93.2868	93.2019
JHI1266	93.3484	93.2443	93.1348	93.3982	93.3423	93.3468	93.2999	93.2382	93.2895	93.159
51A11	92.7764	92.8222	92.8473	92.7611	92.877	92.7117	92.8304	92.788	92.8608	92.7295
51B21	92.7471	92.7725	92.8385	92.7728	92.8339	92.7315	92.8085	92.7097	92.9043	92.7637
128C53	100	99.7804	99.5509	99.6434	99.703	99.5378	99.5881	98.0238	98.1664	98.1291
4292	99.7732	100	99.7564	99.8717	99.8173	99.7879	99.815	98.0923	98.3705	98.0369
CC283bq	99.6022	99.8044	100	99.5153	99.7007	99.4712	99.5884	98.0056	98.1565	97.9054
USDA2370	99.6377	99.8412	99.5238	100	99.7751	99.6031	99.6007	97.9189	98.2424	97.8958
JHI2442	99.6962	99.8654	99.7116	99.7614	100	99.696	99.7358	97.9566	98.3132	98.0318
JHI1084	99.5687	99.7713	99.471	99.5602	99.6759	100	99.5783	97.9494	98.2478	97.8953
JHI944	99.687	99.8078	99.5313	99.6463	99.7035	99.5279	100	98.0097	98.2886	98.0007
OV152	98.0172	98.0763	97.9808	97.8907	97.9934	97.909	97.9746	100	98.2279	98.0592
UPM1137	98.1734	98.3584	98.1867	98.2785	98.2629	98.2171	98.252	98.1893	100	97.993
SM149A	98.1576	98.0849	97.9634	97.8767	98.0602	97.8783	97.9658	98.0137	98.0298	100
Rt24.2	98.1452	98.311	98.2001	98.2175	98.2539	98.1996	98.2689	98.169	99.0157	98.4803
CCBAU23252	91.5073	91.4663	91.4788	91.4786	91.492	91.5213	91.4916	91.5484	91.5535	91.4144
JHI536	91.805	91.5529	91.5208	91.7913	91.7798	91.6087	91.5522	91.6377	91.7924	91.4781
C15	91.444	91.5862	91.4992	91.5169	91.5248	91.4585	91.5155	91.5852	91.581	91.4036
CF307	91.5899	91.5857	91.6054	91.536	91.5726	91.5373	91.5758	91.5935	91.5947	91.5131

Table S4.5. (Continuation)

Query	Rt24.2	CCBAU23252	JHI536	C15	CF307
3841	92.8325	91.1	91.1363	91.0172	91.1616
SM38	93.0039	91.0253	91.0807	91.1331	91.2119
JHI960	92.7511	91.1089	91.1269	91.0767	91.1751
JHI963	92.8237	91.1056	91.23	91.0668	91.1571
JHI1587	92.8017	91.1461	91.1773	91.0483	91.1152
JHI1600	92.8579	91.1463	91.1849	91.1161	91.203
JHI13	92.8141	91.1493	91.2573	91.0753	91.1799
JHI1438	92.8114	91.1304	91.3116	91.0337	91.132
JHI535	92.8608	91.2151	91.2066	91.1321	91.1935
JHI585	92.8742	91.2197	91.2018	91.2028	91.2226
JHI973	92.8666	91.1286	91.1589	91.1253	91.2084
JHI974	92.8048	91.1604	91.1947	91.1091	91.1525
JHI1415	92.9067	91.1708	91.1484	91.0849	91.2318
SM3	92.8503	91.0327	91.0503	91.042	91.1853
VF39	92.8247	91.0723	91.1708	91.1293	91.1586
L361	92.3992	90.9778	91.0252	90.8669	90.9191
VFF1R2B1	92.8187	91.1036	91.028	91.1103	91.1901
21A12	92.8929	91.0885	91.1408	91.0922	91.2141
VFF1R2A1	92.8956	91.1207	91.1217	91.1029	91.1676
WSM1455	92.7152	91.0808	91.071	91.0454	91.177
WSM1481	92.705	91.0724	91.0329	91.0107	91.1391
WSM78	93.4656	91.1594	91.1843	91.1823	91.2336
SM152B	93.392	91.0923	91.1272	91.1403	91.1615
CC275e	93.3808	91.1796	91.1681	91.1697	91.3028
WSM1325	92.8421	91.0077	90.9973	90.9757	91.0719
SRDI943	92.8547	91.0199	91.0843	90.9852	91.1338
JHLM12A2	92.7969	91.18	91.2494	91.2236	91.206
Vaf12	92.8109	91.2038	91.321	91.1301	91.1854
JHI2451	92.804	91.164	91.3787	91.146	91.2105
JHI10	92.8484	91.2011	91.3198	91.0793	91.2639
JHI54	92.7774	91.1535	91.2452	91.0691	91.2576
FA23	92.8494	91.1469	91.2476	91.1686	91.2442
JHI2450	92.9169	91.2194	91.4272	91.1511	91.3073
Vaf-108	92.2906	90.8174	90.8965	90.6714	90.862
JHI2449	92.3991	90.9586	91.0079	90.836	90.9502
UPM1131	92.4716	90.9292	91.0777	90.9285	90.9116
CCBAU10279	92.1286	90.728	90.8499	90.8548	90.7912
WSM1689	92.379	90.7675	90.8303	90.785	90.9444
VFF2R2A1	92.2565	90.8394	90.8968	90.8302	90.8302
SPF2A11	92.1881	90.8202	90.7375	90.6873	90.7955
VFCR2A2	92.1894	90.8112	90.8248	90.8992	90.8429
248	92.321	90.9399	90.902	90.8457	90.8978
63A21	92.447	91.1038	91.0166	91.0517	91.0113
GLR2	92.2823	90.8443	90.9019	90.751	90.8334
CL8	92.2475	90.7895	90.7582	90.716	90.8073
VFSR2A2	92.1875	90.8458	90.838	90.8654	90.8476
SS21	92.1958	90.8925	90.7627	90.635	90.8022
41A11	92.231	90.8468	90.8361	90.834	90.7968
42B12	92.2945	90.8764	90.9391	90.8712	90.9146
43A11	92.2024	90.8024	90.8653	90.7628	90.8671
43B11	92.1815	90.8535	90.8407	90.8562	90.8617
43B12	92.2405	90.798	90.7958	90.7874	90.8764
71A12	92.2309	90.7789	90.8086	90.8275	90.7862
73A11	92.2596	90.7801	90.8495	90.7688	90.8581
81B22	92.2353	90.7819	90.8063	90.8063	90.85
83A12	92.2202	90.7962	90.7904	90.7948	90.8314
VFHR1A2	92.2982	90.8359	90.8823	90.835	90.8565
VFF1R1A2	92.2185	90.6898	90.6806	90.6575	90.8082
93B11	92.1932	90.735	90.8343	90.7714	90.8337
111A12	92.2548	90.8248	90.7813	90.731	90.8225
VFSR2B1	92.223	90.8339	90.762	90.7189	90.8836
SPF4F7	92.2126	90.8495	90.8116	90.7889	90.8405
FB206	92.157	90.7523	90.7812	90.7176	90.8268
SL16	92.1752	90.9143	90.7955	90.7229	90.8827
HUTR05	92.2978	90.6412	90.6807	90.6959	90.758
21B12	92.2904	90.8265	90.7903	90.6924	90.811
WYCCWR11290	92.3658	91.2348	91.3525	91.2218	91.2383
WYCCWR11317	92.4061	91.1164	91.2633	91.2394	91.3217
CCBAU11080	92.3555	90.9963	91.1187	91.0823	91.098
CCBAU03386	92.44	91.0209	91.0933	91.0563	91.1997

Table S4.5. (Continuation)

Query	Rt24.2	CCBAU23252	JHI536	C15	CF307
WYCCWR10014	93.4768	91.0882	91.1781	91.1692	91.2637
JKLM13E	92.823	91.1793	91.1796	91.1848	91.2082
Tri-43	92.7433	91.2712	91.2949	91.2712	91.2936
73B11	92.3936	90.8484	90.8481	90.822	90.9181
73B12	92.4336	90.8961	90.9442	90.8983	90.9946
RSP1F2	92.3869	90.8966	91.3556	90.8264	90.914
Vaf10	92.2852	90.816	90.8758	90.7064	90.7776
PEPV16	92.2497	90.7245	90.8564	90.7314	90.7842
TOM	92.2775	90.8143	90.8952	90.8072	90.8323
121B21	92.3177	91.1479	90.8594	90.7346	90.859
RSP1F10	92.3219	90.9297	91.0651	90.8963	90.9196
RSP1A1	92.394	90.7966	91.1275	90.8044	90.84
Norway	94.5233	91.5572	91.745	91.5015	91.5956
CC278f	94.9497	91.3997	91.4675	91.4301	91.5902
SM78	95.116	91.5932	91.5857	91.553	91.6588
SM51	95.0536	91.5509	91.5167	91.5408	91.6119
SM72	95.2033	91.5818	91.5904	91.5648	91.7056
Vaf-46	92.8129	90.9825	91.0275	90.8819	91.0052
JHI1093	92.8893	91.0241	91.2503	90.9487	91.0119
JHI1096	92.8553	91.0092	91.196	90.9026	90.9992
GPTR08	92.2498	90.4934	90.5427	90.5864	90.6613
GPTR02	92.2852	90.5367	90.527	90.5415	90.5995
SRDI565	92.3448	90.6075	90.5761	90.6298	90.6419
Ps8	93.1736	91.2747	91.2746	91.2188	91.2914
JHI1236	93.279	91.2623	91.2257	91.1784	91.3055
JHI953	93.2985	91.2813	91.2069	91.1982	91.2637
SM147A	93.2852	91.2487	91.233	91.2018	91.2503
JHI1238	93.3271	91.2612	91.2433	91.2267	91.2765
UPM1133	93.2178	91.2737	91.4769	91.1652	91.2644
JHI1592	93.2823	91.1935	91.2613	91.1938	91.268
SM41	93.3257	91.1563	91.1805	91.2214	91.22
JHI1253	93.2456	91.1819	91.2203	91.204	91.2491
JHI370	93.263	91.1866	91.2082	91.1695	91.2461
JHI387	93.2272	91.2515	91.2898	91.1884	91.3041
JHI388	93.2523	91.2409	91.2234	91.2477	91.3066
JHI788	93.3081	91.3627	91.3279	91.2657	91.3468
JHI985	93.2446	91.2028	91.2491	91.2327	91.3362
GB30	93.2709	91.2215	91.2532	91.2306	91.2576
JHI782	93.1741	91.2401	91.2419	91.1488	91.2178
JHI783	93.2153	91.2409	91.2064	91.1473	91.2262
SM170C	93.3058	91.1696	91.1666	91.1329	91.3339
JHI42	93.2258	91.2968	91.3904	91.2612	91.3199
JHI979	93.2282	91.2317	91.3171	91.1802	91.3017
JHI1259	93.2177	91.2277	91.2624	91.1892	91.1915
TA1	93.3173	91.2403	91.1973	91.1618	91.3284
JHI24	93.2618	91.2603	91.4786	91.2023	91.3057
JHI787	93.2823	91.2692	91.4579	91.2048	91.2957
RCAM1026	93.3996	91.264	91.3052	91.2188	91.3265
Vh3	93.2232	91.2873	91.2903	91.263	91.2919
JHI925	93.0887	91.3338	91.3128	91.201	91.3109
Vc2	93.1387	91.2531	91.555	91.2542	91.319
JHI1422	93.253	91.3082	91.3451	91.1943	91.3587
JHI1266	93.193	91.2616	91.1785	91.199	91.2552
51A11	92.8022	91.0837	91.1558	91.0786	91.2267
51B21	92.8721	91.0987	91.2028	91.0822	91.2564
128C53	98.1404	91.4618	91.717	91.4694	91.5364
4292	98.2897	91.4649	91.5467	91.5286	91.538
CC283bq	98.1918	91.3564	91.4749	91.4457	91.5234
USDA2370	98.2417	91.4465	91.746	91.4408	91.5033
JHI2442	98.307	91.5237	91.6569	91.4225	91.5496
JHI1084	98.2045	91.4853	91.5496	91.4464	91.5108
JHI944	98.3054	91.4969	91.5068	91.4516	91.5289
OV152	98.1585	91.5815	91.6215	91.584	91.5955
UPM1137	98.9872	91.5281	91.6814	91.5077	91.5762
SM149A	98.4684	91.3447	91.4577	91.3092	91.5036
Rt24.2	100	91.4539	91.4874	91.4736	91.5266
CCBAU23252	91.4352	100	98.6027	98.7931	98.463
JHI536	91.5256	98.5744	100	98.5981	98.5688
C15	91.5351	98.8109	98.6392	100	98.4108
CF307	91.6098	98.3846	98.4983	98.4162	100

Table S4.6. Complete hit result table of genes with significant SNPs. Gene names in brackets preceeded by a 'p' indicate pseudogenes, genes inferred by protein homology which are incomplete or with a stop codon in the middle of the sequence.

Protein/gene ID	Gene name	UniProt	Product	Process	Genome location
C3Y91_RS00465	p(WP_018483041.1)	UPI000381F47F	efflux RND transporter permease subunit	Pseudogene	fragment
C3Y91_RS04085	p(cobF)	UPI00027D4FD7	precorrin-6A synthase (deacetylating)	Pseudogene	plasmid-Rh02
C3Y91_RS10365	p(WP_012490139.1)	B3Q3I6	Putative dehydrogenase protein	Pseudogene	fragment
C3Y91_RS23865	p(WP_019858718.1)	UPI0003804A95	GcvT family protein	Pseudogene	fragment
WP_018068543.1	protein_coding	A0A4Q8YSB1	autoinducer 2 ABC transporter substrate-binding protein	Membrane transport	fragment
WP_018069891.1	protein_coding	UPI0003776405	AraC family transcriptional regulator	DNA replication/transcription	fragment
WP_018071070.1	protein_coding	A0A7G6PZE5	endonuclease/exonuclease/phosphatase family protein	Other	fragment
WP_018071162.1	protein_coding	A0A6B3JFH6	ABC transporter permease	Membrane transport	fragment
WP_018071233.1	protein_coding	A0A6B3J6S8	ABC transporter substrate-binding protein	Membrane transport	fragment
WP_018071502.1	protein_coding	A0A4V2IQN2	MBL fold metallo-hydrolase	Other	plasmid-Rh02
WP_018071582.1	protein_coding	A0A6B3J4R3	Bax inhibitor-1/YccA family protein	Other	plasmid-Rh02
WP_018071653.1	protein_coding	A0A6B3J4X4	Tad domain-containing protein	Other	fragment
WP_018071665.1	glnT	A0A4Q8YTN2	type III glutamate--ammonia ligase	Other	fragment
WP_018072931.1	protein_coding	A0A4Q1THM3	haloacid dehalogenase type II	Other	plasmid-Rh01
WP_018072987.1	protein_coding	UPI000369ACEA	sugar ABC transporter permease	Membrane transport	plasmid-Rh01
WP_018480207.1	zwf	A0A4Q8Y7N7	glucose-6-phosphate dehydrogenase	Other	fragment
WP_018480358.1	protein_coding	A0A7K3V0I2	SMP-30/gluconolactonase/LRE family protein	Other	fragment
WP_018480520.1	protein_coding	A0A4Q8YTS1	sugar ABC transporter ATP-binding protein	Membrane transport	fragment
WP_018483352.1	protein_coding	A0A4Q8ZUV2	ABC transporter substrate-binding protein	Membrane transport	fragment
WP_018483394.1	kdpB	A0A6P0BL74	potassium-transporting ATPase subunit KdpB	Membrane transport	fragment
WP_018493609.1	protein_coding	UPI00035EB03C	aromatic acid/H+ symport family MFS transporter	Membrane transport	fragment
WP_018493610.1	protein_coding	A0A6P0BLY9	MarR family transcriptional regulator	Other	fragment
WP_018494241.1	protein_coding	A0A4V2IK13	GntR family transcriptional regulator	DNA replication/transcription	fragment
WP_018494407.1	protein_coding	A0A4Q8XUP3	FCD domain-containing protein	Other	plasmid-Rh02
WP_018496327.1	protein_coding	A0A6B3JIZ3	IclR family transcriptional regulator	DNA replication/transcription	plasmid-Rh03
WP_020397477.1	protein_coding	A0A4Q8YGH1	ABC transporter ATP-binding protein	Membrane transport	fragment
WP_020397511.1	protein_coding	A0A7G6PZZ8	hypothetical protein	Other	plasmid-Rh02
WP_024319643.1	protein_coding	A0A7G6Q267	sugar ABC transporter permease	Membrane transport	plasmid-Rh01
WP_024319743.1	protein_coding	A0A4Q8ZX09	aldo/keto reductase	Membrane transport	plasmid-Rh01
WP_024319744.1	protein_coding	A0A7G6Q359	6-chlorohydroxyquinol-1,2-dioxygenase	Other	plasmid-Rh01
WP_024319810.1	protein_coding	A0A6G7MR82	sugar ABC transporter ATP-binding protein	Membrane transport	fragment
WP_024319811.1	protein_coding	A0A6N9ZSN0	ABC transporter permease	Membrane transport	fragment
WP_024319843.1	protein_coding	UPI00040A2D32	D-amino acid dehydrogenase	Other	fragment

Table S4.6. (Continuation)

Protein/gene ID	Gene name	UniProt	Product	Process	Genome location
WP_024319844.1	alr	UPI000462BCB0	alanine racemase	Other	fragment
WP_024320102.1	kduI	A0A6P0ADP8	5-dehydro-4-deoxy-D-glucuronate isomerase	Other	fragment
WP_024320175.1	protein_coding	A0A4Q8ZU96	ABC transporter permease	Membrane transport	fragment
WP_024321462.1	protein_coding	UPI000462D9AD	carbohydrate ABC transporter permease	Membrane transport	fragment
WP_024321466.1	protein_coding	A0A6P0AGE8	DeoR/GlpR family DNA-binding transcription regulator	DNA replication/transcription	fragment
WP_024321535.1	protein_coding	A0A4Q8YPL5	methyl-accepting chemotaxis protein	Chemotaxis	fragment
WP_024321569.1	protein_coding	UPI00040CC489	hydrolase	Other	fragment
WP_024321570.1	ggt	A0A7K3V220	gamma-glutamyltransferase	Osmoprotectant synthesis	fragment
WP_024321596.1	protein_coding	UPI0004056E09	ROK family transcriptional regulator	Other	fragment
WP_024321617.1	protein_coding	A0A4Q8ZXD8	YihY/virulence factor BrkB family protein	Other	fragment
WP_024321672.1	protein_coding	A0A444HNC3	Adenylate/guanylate cyclase domain-containing protein	DNA replication/transcription	plasmid-Rh02
WP_024321713.1	repA	A0A4Q8Z0Z5	plasmid partitioning protein RepA	DNA replication/transcription	plasmid-Rh02
WP_024321745.1	protein_coding	UPI0003F5CD49	DUF4384 domain-containing protein	Other	plasmid-Rh02
WP_024321758.1	protein_coding	A0A4Q8ZVI5	Pilus assembly protein	Motility	fragment
WP_024322052.1	iolE	A0A6P0ASW1	myo-inosose-2 dehydratase	Other	fragment
WP_027687940.1	protein_coding	A0A4Q8ZY58	LacI family DNA-binding transcriptional regulator	DNA replication/transcription	fragment
WP_027688003.1	protein_coding	UPI000404FB23	5-dehydro-4-deoxyglucarate dehydratase	Other	fragment
WP_027688017.1	protein_coding	UPI000426BA67	SIS domain-containing protein	Other	fragment
WP_027688281.1	protein_coding	UPI000427A9F9	carbohydrate ABC transporter permease	Membrane transport	plasmid-Rh01
WP_027688879.1	protein_coding	A0A4Q8ZTC2	ATP-grasp domain-containing protein	Other	fragment
WP_027689924.1	protein_coding	UPI0004830B5C	ABC transporter substrate-binding protein	Membrane transport	fragment
WP_027690125.1	protein_coding	UPI0004162509	FMN-binding glutamate synthase family protein	Other	fragment
WP_029767995.1	protein_coding	UPI0003FD9F86	HD-GYP domain-containing protein	Other	fragment
WP_033183136.1	protein_coding	A0A6M5ZST6	sugar ABC transporter permease	Membrane transport	fragment
WP_130655402.1	protein_coding	UPI0010301605	ABC transporter ATP-binding protein	Membrane transport	fragment
WP_130657758.1	protein_coding	UPI00103164D3	glutathione S-transferase family protein	Other	plasmid-Rh03
WP_130663140.1	protein_coding	UPI00102F63B0	glycine betaine/L-proline ABC transporter ATP-binding protein	Membrane transport	fragment
WP_130663142.1	protein_coding	UPI00102F4BD3	aldehyde dehydrogenase	Other	fragment
WP_130663186.1	protein_coding	A0A6G7MRY6	redoxin domain-containing protein	Other	fragment
WP_130663304.1	protein_coding	UPI00103130A8	replication initiation protein RepC	DNA replication/transcription	plasmid-Rh01
WP_130663448.1	protein_coding	UPI001031DB26	EAL domain-containing protein	Other	fragment
WP_130663843.1	protein_coding	UPI001031928A	aspartate aminotransferase family protein	Other	plasmid-Rh01
WP_130665636.1	ugpC	UPI001031E164	sn-glycerol-3-phosphate ABC transporter ATP-binding protein UgpC	Membrane transport	fragment
WP_130665639.1	protein_coding	UPI0010305B5C	Ldh family oxidoreductase	Other	fragment

Table S4.6. (Continuation)

Protein/gene ID	Gene name	UniProt	Product	Process	Genome location
WP_130665642.1	protein_coding	UPI0010301993	mandelate racemase/muconate lactonizing enzyme family protein	Other	fragment
WP_130665649.1	protein_coding	A0A6G7MS07	NUDIX domain-containing protein	Other	fragment
WP_130665708.1	protein_coding	UPI001030E3D2	dihydrodipicolinate synthase family protein	Other	fragment
WP_130665734.1	doeA	UPI0010312317	ectoine hydrolase DoeA	Other	plasmid-Rh01
WP_130665966.1	accC	UPI00103123A4	acetyl-CoA carboxylase biotin carboxylase subunit	Osmoprotectant synthesis	plasmid-Rh03
WP_130671142.1	protein_coding	UPI0010305067	TetR/AcrR family transcriptional regulator	DNA replication/transcription	fragment
WP_130676365.1	protein_coding	UPI00103015DC	sugar ABC transporter substrate-binding protein	Membrane transport	fragment
WP_130676390.1	protein_coding	UPI00144517B6	MFS transporter	Membrane transport	fragment
WP_130680579.1	protein_coding	UPI001031C253	aromatic ring-hydroxylating dioxygenase subunit alpha	Other	fragment
WP_130687842.1	protein_coding	UPI001030EC95	alpha-glucosidase/alpha-galactosidase	Other	fragment
WP_130697438.1	protein_coding	UPI001030502A	ABC transporter substrate-binding protein	Membrane transport	fragment
WP_130698974.1	protein_coding	UPI00102F76C2	precorrin-3B C(17)-methyltransferase	Other	plasmid-Rh02
WP_130764027.1	protein_coding	UPI00102FBFFD	acetolactate synthase large subunit	Other	fragment
WP_130764183.1	protein_coding	UPI001030060C	beta-N-acetylhexosaminidase	Other	fragment
WP_130765699.1	protein_coding	UPI0010314FD0	carbohydrate ABC transporter permease	Membrane transport	fragment
WP_130800988.1	protein_coding	UPI0010313D04	bifunctional rhamnulose-1-phosphate aldolase/short-chain dehydrogenase	Other	fragment
WP_130828059.1	tauA	UPI00102F698A	taurine ABC transporter substrate-binding protein	Membrane transport	fragment
WP_131614440.1	protein_coding	UPI00103F617B	nitrate ABC transporter substrate-binding protein	Membrane transport	fragment
WP_163857423.1	protein_coding	UPI0013D1E26F	ABC transporter substrate-binding protein	Membrane transport	plasmid-Rh01
WP_164567644.1	protein_coding	UPI0013C1D49E	cytochrome-c peroxidase	Other	fragment
WP_168307136.1	protein_coding	UPI0014429517	aldehyde dehydrogenase family protein	Other	fragment
WP_168575007.1	pcaC	UPI001446A972	4-carboxymuconolactone decarboxylase	Other	plasmid-Rh02
WP_168575009.1	pcaG	UPI001448438F	protocatechuate 3%2C4-dioxygenase subunit alpha	Other	plasmid-Rh02
WP_168575019.1	protein_coding	UPI001444DDBD	adenylate/guanylate cyclase domain-containing protein	DNA replication/transcription	plasmid-Rh02
WP_168575034.1	protein_coding	UPI0014471A60	PAS domain-containing hybrid sensor histidine kinase/response regulator	Other	plasmid-Rh02
WP_168575043.1	protein_coding	UPI00144824A8	alpha-2-macroglobulin family protein	Other	plasmid-Rh02
WP_168575051.1	cobG	UPI001446CB04	precorrin-3B synthase	Cobalamin biosynthesis	plasmid-Rh02
WP_168575052.1	protein_coding	UPI001444DA6C	Cobalamin biosynthesis precorrin-8X methylmutase CobH/CbiC	Cobalamin biosynthesis	plasmid-Rh02
WP_168575059.1	protein_coding	UPI001445773B	L-aspartate oxidase	Other	plasmid-Rh02
WP_168575063.1	protein_coding	UPI001444B575	caspase family protein	Caspase	plasmid-Rh02
WP_168575065.1	protein_coding	UPI0014454E65	caspase family protein	Caspase	plasmid-Rh02
WP_168575071.1	protein_coding	UPI001444C517	ABC transporter ATP-binding protein/permease	Membrane transport	plasmid-Rh02
WP_168575125.1	protein_coding	UPI001446EB0D	fumarylacetoacetate hydrolase family protein	Other	fragment
WP_168575137.1	kdpC	UPI001446BA69	potassium-transporting ATPase subunit KdpC	Membrane transport	fragment

Table S4.6. (Continuation)

Protein/gene ID	Gene name	UniProt	Product	Process	Genome location
WP_168575138.1	protein_coding	UPI0014488B7B	sensor histidine kinase KdpD	Membrane transport	fragment
WP_168575141.1	protein_coding	UPI0014455FA6	sulfate ABC transporter substrate-binding protein	Membrane transport	fragment
WP_168575146.1	protein_coding	UPI00144517B6	MFS transporter	Membrane transport	fragment
WP_168575232.1	protein_coding	UPI001444ADA7	pilus assembly protein CpaB	Motility	fragment
WP_168575234.1	protein_coding	UPI001447A8D6	sarcosine oxidase subunit alpha family protein	Osmoprotectant synthesis	fragment
WP_168575241.1	protein_coding	UPI0014451CDA	cellobiose phosphorylase	Osmoprotectant synthesis	fragment
WP_168575251.1	protein_coding	UPI001446CFF5	MOSC domain-containing protein	Other	fragment
WP_168575268.1	protein_coding	UPI00144590D4	EAL domain-containing protein	Other	fragment
WP_168575278.1	protein_coding	UPI0014458283	AMP-binding protein	Other	fragment
WP_168575610.1	treS	UPI001446B1F0	maltose alpha-D-glucosyltransferase	Osmoprotectant synthesis	plasmid-Rh01
WP_168575692.1	protein_coding	UPI001444E23D	NAD(P)-dependent alcohol dehydrogenase	Other	fragment
WP_168575862.1	ugpC	UPI0014478B3F	sn-glycerol-3-phosphate ABC transporter ATP-binding protein UgpC	Membrane transport	fragment
WP_168575955.1	protein_coding	UPI00144672E9	imidazolonepropionase	Other	fragment
WP_168576260.1	protein_coding	UPI001447F5BF	sensor histidine kinase	Membrane transport	fragment
WP_168576265.1	protein_coding	UPI001446B6B3	HlyD family secretion protein	Membrane transport	fragment
WP_172643707.1	protein_coding	UPI00041C43E5	ABC transporter ATP-binding protein	Membrane transport	fragment
WP_174826993.1	protein_coding	UPI001582E37B	sugar phosphate isomerase/epimerase	Other	fragment
WP_174827004.1	protein_coding	UPI0015833C91	Para-hydroxybenzoic acid efflux pump subunit AaeB/fusaric acid resistance protein	Membrane transport	fragment
WP_174827059.1	protein_coding	UPI001583244C	efflux RND transporter permease subunit	Membrane transport	plasmid-Rh01
WP_174827074.1	tsdA	UPI001581EF02	gamma-resorcylate decarboxylase	Other	fragment

Table S5.1. List of unique sequences with at least one primer match for 16S rRNA and *nodD* from the first 100 matches for each primer used in the qPCR reactions. The NCBI primer match column values indicate whether the sequence has a matching sequence for both primers (Full match) or only for one of them (Partial match).

Gene	Order	Genera	Species	Accession No	NCBI primer match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium aegyptiacum	MT846025.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium aegyptiacum	MT846023.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium aegyptiacum	MT846019.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium aegyptiacum	MT846018.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium aegyptiacum	MT846017.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium aegyptiacum	MT534142.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium anhuiense	MT476932.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium binae	MT846020.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium binae	MT846022.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium binae	MT846026.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium hidalgonense	CP054027.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium hidalgonense	MT370019.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium hidalgonense	MT476933.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium indicum	CP054021.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium indicum	CP054031.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium laguerreae	MT370020.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium laguerreae	MT370021.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium laguerreae	MT370022.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium laguerreae	MT370023.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium laguerreae	MT370024.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium leguminosarum	MT605966.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium leguminosarum	MT605968.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium leguminosarum	MT605969.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium leguminosarum	MT605967.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium leguminosarum	CP050514.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium leguminosarum	CP050549.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium leguminosarum	CP050555.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium leguminosarum	CP050562.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium leguminosarum	MT775520.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium leguminosarum	MT775521.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium leguminosarum	MT775522.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium leguminosarum	MT775523.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium leguminosarum	MT775515.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium leguminosarum	MT775516.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium leguminosarum	MT775517.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium leguminosarum	MT775518.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium leguminosarum	MT775519.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium leguminosarum	MT825127.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium leguminosarum	MT825128.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium leguminosarum	MT825134.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium leguminosarum	MT825116.1	Full match

Table S5.1. (Continuation)

Gene	Order	Genera	Species	Accession No	NCBI primer match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium leguminosarum	MT825120.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium leguminosarum	MT825121.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium leguminosarum	MT645954.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium leguminosarum	MT645962.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium leguminosarum	MT645960.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium leguminosarum	MT775524.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium leguminosarum	CP053205.2	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium leguminosarum	CP053439.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium leguminosarum	MT572952.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium leguminosarum	MT605962.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium leguminosarum	MT605963.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium leguminosarum	MT605964.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium leguminosarum	MT605965.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium leguminosarum	MT900580.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium leguminosarum	MT917183.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium sophorae	MT645953.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium sophorae	MT645955.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium sophorae	MT645961.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium sophorae	MT645963.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium sophorae	MT645959.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium sophorae	MT645958.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium sophorae	MT645957.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium sophorae	MT645956.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium sophorae	MT645952.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium sp.	MT572931.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium sp.	MT860405.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium sp.	MT860406.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium sp.	MT860408.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium sp.	MT860409.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium sp.	MT860410.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium sp.	MT860412.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium sp.	MT860413.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium sp.	MT860414.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium sp.	MT860415.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium sp.	MT651617.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium sp.	MT793077.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium sp.	MT476928.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium sp.	MT476929.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium sp.	MT476930.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium sp.	MT370016.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium sp.	MT370017.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium sp.	MT370018.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium subbaraonis	MT544595.1	Full match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium sullae	MT776720.1	Full match

Table S5.1. (Continuation)

Gene	Order	Genera	Species	Accession No	NCBI primer match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium sullae	MT776721.1	Full match
16S rRNA	Rhizobiales	Brucella	Brucella melitensis	MT611102.1	Partial match
16S rRNA	Rhizobiales	Brucella	Brucella melitensis	MT611103.1	Partial match
16S rRNA	Rhizobiales	Brucella	Brucella melitensis	MT611104.1	Partial match
16S rRNA	Rhizobiales	Brucella	Brucella melitensis	MT611105.1	Partial match
16S rRNA	Rhizobiales	Brucella	Brucella sp.	CP061088.1	Partial match
16S rRNA	Rhizobiales	Brucella	Brucella sp.	CP061089.1	Partial match
16S rRNA	Rhizobiales	Brucella	Brucella sp.	MT991987.1	Partial match
16S rRNA	Rhizobiales	Brucella	Brucella sp.	CP047232.1	Partial match
16S rRNA	Rhizobiales	Brucella	Brucella sp.	CP047233.1	Partial match
16S rRNA	Rhizobiales	Brucella	Brucella sp.	MT991988.1	Partial match
16S rRNA	Rhizobiales	Martelella	Martelella sp.	MT830287.1	Partial match
16S rRNA	Rhizobiales	Ochrobactrum	Ochrobactrum anthropi	LC557006.1	Partial match
16S rRNA	Rhizobiales	Ochrobactrum	Ochrobactrum ciceri	MT984446.1	Partial match
16S rRNA	Rhizobiales	Ochrobactrum	Ochrobactrum ciceri	MT984445.1	Partial match
16S rRNA	Rhizobiales	Ochrobactrum	Ochrobactrum intermedium	LC557008.1	Partial match
16S rRNA	Rhizobiales	Ochrobactrum	Ochrobactrum intermedium	MT649859.1	Partial match
16S rRNA	Rhizobiales	Ochrobactrum	Ochrobactrum intermedium	MW007813.1	Partial match
16S rRNA	Rhizobiales	Ochrobactrum	Ochrobactrum intermedium	MT605439.1	Partial match
16S rRNA	Rhizobiales	Ochrobactrum	Ochrobactrum intermedium	CP061039.1	Partial match
16S rRNA	Rhizobiales	Ochrobactrum	Ochrobactrum intermedium	CP061040.1	Partial match
16S rRNA	Rhizobiales	Ochrobactrum	Ochrobactrum lupini	MT765157.1	Partial match
16S rRNA	Rhizobiales	Ochrobactrum	Ochrobactrum oryzae	LC557005.1	Partial match
16S rRNA	Rhizobiales	Ochrobactrum	Ochrobactrum pecoris	MT758011.1	Partial match
16S rRNA	Rhizobiales	Ochrobactrum	Ochrobactrum pecoris	MT760048.1	Partial match
16S rRNA	Rhizobiales	Ochrobactrum	Ochrobactrum sp.	MT850128.1	Partial match
16S rRNA	Rhizobiales	Ochrobactrum	Ochrobactrum sp.	MT742989.1	Partial match
16S rRNA	Rhizobiales	Ochrobactrum	Ochrobactrum sp.	MT673841.1	Partial match
16S rRNA	Rhizobiales	Ochrobactrum	Ochrobactrum sp.	MT994341.1	Partial match
16S rRNA	Rhizobiales	Ochrobactrum	Ochrobactrum sp.	MT994342.1	Partial match
16S rRNA	Rhizobiales	Ochrobactrum	Ochrobactrum sp.	MT994335.1	Partial match
16S rRNA	Rhizobiales	Ochrobactrum	Ochrobactrum sp.	MT754793.1	Partial match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium aegyptiacum	MT980913.1	Partial match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium alamii	MT775434.1	Partial match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium etli	MT825122.1	Partial match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium etli	MT825123.1	Partial match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium etli	MT825136.1	Partial match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium leguminosarum	MT572951.1	Partial match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium leguminosarum	MT573155.1	Partial match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium lentis	MT846024.1	Partial match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium lusitanum	CP050308.1	Partial match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium mesosinicum	MT775436.1	Partial match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium sp.	MT416001.1	Partial match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium sp.	MT762867.1	Partial match

Table S5.1. (Continuation)

Gene	Order	Genera	Species	Accession No	NCBI primer match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium sp.	MT759825.1	Partial match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium sp.	MT830879.1	Partial match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium sp.	MT830881.1	Partial match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium sp.	MT860407.1	Partial match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium sp.	MT657358.1	Partial match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium sp.	MT793076.1	Partial match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium sp.	MT793080.1	Partial match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium sp.	MT793101.1	Partial match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium sp.	MT649296.1	Partial match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium sp.	MT707042.1	Partial match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium viscosum	MT534118.1	Partial match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium viscosum	MT534119.1	Partial match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium yanglingense	MT974169.1	Partial match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium yanglingense	MT974170.1	Partial match
16S rRNA	Rhizobiales	Rhizobium	Rhizobium yanglingense	MT974171.1	Partial match
16S rRNA	Rhizobiales	Uncultured	Uncultured Ochrobactrum	MT858322.1	Partial match
16S rRNA	Rhizobiales	Uncultured	Uncultured Rhizobiales	MT858053.1	Partial match
16S rRNA	Rhizobiales	Uncultured	Uncultured Rhizobiales	MT858289.1	Partial match
16S rRNA	Rhizobiales	Uncultured	Uncultured Rhizobiales	MT858088.1	Partial match
16S rRNA	Rhizobiales	Uncultured	Uncultured Rhizobiales	MT858268.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium laguerreae	MF572077.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium laguerreae	MF572080.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium laguerreae	MF572083.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	CP025015.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	CP025505.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	CP050554.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	KC679657.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	KC679658.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	KC679659.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	KC679660.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	KC679662.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	KC679663.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	KC679664.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	KC679665.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	KC679667.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	KC679668.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	KC679669.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	KC679670.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	KC679671.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	KC679672.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	KC679673.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	KC679674.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	KF264444.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	KF264445.1	Full match

Table S5.1. (Continuation)

Gene	Order	Genera	Species	Accession No	NCBI primer match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	KF264447.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	KF264448.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	KF264449.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	KF264450.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	MF572078.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	MK514429.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	MK514430.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	MK514431.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	MK514432.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	MK514433.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	MK514434.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	MK514435.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	MK514436.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	MK514437.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	MK514438.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	MK514439.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	MK514440.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	MK514441.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium sp.	MG546103.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium sp.	MG546104.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium sp.	MG546105.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium sp.	MG546106.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium sp.	MG546107.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium sp.	MG546108.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium sp.	MG546109.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium sp.	MG546110.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium sp.	MG546111.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium sp.	MG546112.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium sp.	MG546113.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium sp.	MN219408.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium sp.	MF572081.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium sp.	MF572082.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium sp.	MF572084.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium sp.	MF572085.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium sp.	MF572086.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium sp.	MF572087.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	CP022669.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	CP048285.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	CP050564.1	Full match
nodD	Rhizobiales	Rhizobium	Rhizobium anhuiense	KY425293.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium anhuiense	KY425295.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium anhuiense	KY425296.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium anhuiense	KY425297.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium anhuiense	KY425298.1	Partial match

Table S5.1. (Continuation)

Gene	Order	Genera	Species	Accession No	NCBI primer match
nodD	Rhizobiales	Rhizobium	Rhizobium anhuiense	KY425299.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium anhuiense	KY425300.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium anhuiense	KY425301.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium anhuiense	KY425302.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium anhuiense	KY425303.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium anhuiense	KY425304.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium anhuiense	KY425305.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium anhuiense	KY425306.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium anhuiense	KY425307.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium anhuiense	KY425308.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium anhuiense	KY425309.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium anhuiense	KY425310.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium anhuiense	KY425312.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium anhuiense	KY425313.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium anhuiense	KY425315.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium anhuiense	KY425317.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium anhuiense	KY425318.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium anhuiense	KY425319.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium anhuiense	KY425322.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium anhuiense	KY425323.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium anhuiense	KY425324.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium anhuiense	KY425326.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium anhuiense	KY425327.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium anhuiense	KY425328.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium anhuiense	KY425329.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium anhuiense	KY425330.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium anhuiense	KY425336.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium anhuiense	KY425338.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium anhuiense	KY425340.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium anhuiense	KY425341.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium anhuiense	KY425342.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium anhuiense	KY425347.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium anhuiense	KY425348.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium anhuiense	KY425349.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium anhuiense	KY425350.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium hidalgonense	CP054030.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium indicum	CP054024.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium indicum	CP054035.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium laguerreae	KY425288.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium laguerreae	KY425311.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium laguerreae	KY425331.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium laguerreae	KY425332.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium laguerreae	KY425333.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium laguerreae	KY425334.1	Partial match

Table S5.1. (Continuation)

Gene	Order	Genera	Species	Accession No	NCBI primer match
nodD	Rhizobiales	Rhizobium	Rhizobium laguerreae	KY425335.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium laguerreae	KY425337.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium laguerreae	KY425343.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium laguerreae	KY425344.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium laguerreae	KY425345.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium laguerreae	KY425346.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	CP016290.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	CP018235.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	CP022669.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	CP048285.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	CP049734.2	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	CP050564.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	KC679661.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	KC679675.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	KC679680.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	KF264446.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	KT944070.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	KY930298.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	KY930299.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	KY930300.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	KY930301.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	KY930302.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	KY930303.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	KY930304.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	KY930305.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	KY930306.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	KY930307.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	KY930308.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	KY930309.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	KY930310.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	KY930311.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	KY930312.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	KY930313.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	KY930314.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	KY930315.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	KY930316.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	KY930317.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	KY930318.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	KY930319.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	KY930320.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	KY930321.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	KY930322.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	KY425290.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	KY425292.1	Partial match

Table S5.1. (Continuation)

Gene	Order	Genera	Species	Accession No	NCBI primer match
nodD	Rhizobiales	Rhizobium	Rhizobium leguminosarum	KY440257.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium pisi	KY425316.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium pisi	KY425321.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium pisi	KY425325.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium pisi	KY440259.1	Partial match
nodD	Rhizobiales	Rhizobium	Rhizobium sp.	MF572079.1	Partial match

Table S5.2. Chemical analyses performed each year. Organic matter and $CaCO_3$ analyses correspond to loss on ignition percentages at 450°C and 900°C. CEC stands for Cation Exchange Capacity which indicates the capacity of the soil to retain cations. %N and %C are measurements of the total N and C percentage in the soil sample. All remaining elements or inorganic compounds were measured in mg kg^{-1}.

Analysis	2011	2012	2013	2014	2015	2016	No years
NO_3	-	✓	✓	✓	-	✓	4
NH_4	-	✓	✓	✓	-	✓	4
Ca	-	✓	✓	-	-	✓	3
K	-	✓	✓	-	-	✓	3
Mg	-	✓	✓	-	-	✓	3
P	-	✓	✓	-	-	✓	3
pH	-	✓	✓	✓	-	✓	4
Organic matter	-	✓	✓	✓	-	-	3
$CaCO_3$	-	✓	✓	✓	-	-	3
%N	-	-	✓	✓	✓	✓	4
%C	-	-	✓	✓	✓	✓	4
S	-	-	-	-	-	✓	1
Mn	-	-	-	-	-	✓	1
Cu	-	-	-	-	-	✓	1
B	-	-	-	-	-	✓	1
Zn	-	-	-	-	-	✓	1
Mo	-	-	-	-	-	✓	1
Fe	-	-	-	-	-	✓	1
Na	-	-	-	-	-	✓	1
CEC	-	-	-	-	-	✓	1
Nitrogen	-	-	-	-	-	✓	1

Table S5.3. Results of the Kendall’s correlation of chemical analyses that had at least one significant correlation with Rleg, Rlv or their ratio. Elemental analyses marked with an ‘*’ indicate only one year of data available. The numbers represent Kendall’s tau. The darker the colour shade the stronger positive (blue) or negative (red) correlation. Cells with “ns” indicate non-significant correlations.

Analysis	**Rlv Rleg^{-1} % Ratio**			**Rleg**			**Rlv**		
	General	**Conventional**	**Integrated**	**General**	**Conventional**	**Integrated**	**General**	**Conventional**	**Integrated**
NO_3	ns	-0.155	ns	ns	ns	ns	ns	ns	ns
NH_4	-0.108	ns	ns	ns	ns	ns	ns	ns	ns
Ca	-0.141	-0.180	-0.162	ns	0.148	ns	ns	ns	ns
K	ns	ns	ns	ns	ns	ns	0.109	ns	ns
Mg	ns	ns	-0.219	ns	ns	ns	-0.102	ns	-0.195
P	ns	ns	-0.157	ns	ns	ns	ns	ns	-0.152
pH	0.110	ns	0.161	ns	ns	ns	ns	ns	ns
Organic matter	ns	ns	0.151	ns	ns	ns	0.130	ns	ns
$CaCO_3$	ns	ns	ns	ns	ns	ns	0.122	ns	ns
%N	ns	ns	ns	0.122	ns	ns	0.140	ns	ns
%C	ns	ns	ns	0.119	ns	ns	0.145	ns	ns
*S	ns	ns	ns	ns	ns	ns	0.207	ns	ns
*Cu	0.179	ns	ns	ns	ns	ns	ns	ns	ns
*Mo	ns	ns	ns	ns	ns	0.277	ns	ns	0.282
*Fe	ns	ns	ns	-0.188	ns	-0.279	ns	ns	-0.284

Table S5.4. In soil concentration of Rleg and Rlv in soils of different origins. Rleg and Rlv columns represent the concentration in individuals g^{-1} of dry soil of Rleg and Rlv respectively. PF samples were provided by a collaboration with the PeaYEN project.

Soil	**Country**	**Location**	**Soil use**	**Rleg**	**Rlv**	**Rlv $Rleg^{-1}$ % Ratio**
E1	UK	Yatesbury, Wilts	Arable organic	$1.90 X 10^6$	$1.65 X 10^5$	8.663688
E2	UK	Cornwall	Wild moorland	$5.12 X 10^4$	$5.41 X 10^3$	10.54927
E3	UK	Cornwall	Wild moorland	$1.80 X 10^6$	$2.01 X 10^5$	11.17383
E4	UK	Yatesbury, Wilts	Arable organic	$2.35 X 10^6$	$2.56 X 10^5$	10.88029
E5	UK	Hockwold, Norfolk	Arable	$8.10 X 10^5$	$9.93 X 10^4$	12.25987
E6	UK	White Horse Hill, Oxon	Native chalkland	$1.72 X 10^6$	$7.73 X 10^4$	4.48396
E7	UK	Yatesbury, Wilts	Arable organic	$1.97 X 10^6$	$2.17 X 10^5$	11.0044
PF01	UK	March, Cambridgeshire	Arable	$2.91 X 10^6$	$1.87 X 10^5$	6.431324
PF02	UK	Wimblington, Cambridgeshire	Arable	$1.53 X 10^6$	$3.18 X 10^5$	20.81692
PF03	UK	Ramsey, Cambridgeshire	Arable	$2.18 X 10^6$	$1.41 X 10^5$	6.468574
PF04	UK	Thorney, Peterborough	Arable	$1.80 X 10^6$	$1.53 X 10^5$	8.474807
PF05	UK	Barley, Royston	Arable	$1.88 X 10^6$	$1.59 X 10^5$	8.436683
PF06	UK	Salisbury	Arable	$2.49 X 10^6$	$2.00 X 10^5$	8.031641
PF07	UK	Sutton bridge, Spalding	Arable	$1.79 X 10^6$	$1.87 X 10^5$	10.4541
PF08	UK	Upwell, Wisbech	Arable	$1.81 X 10^6$	$2.15 X 10^5$	11.87361
PF09	UK	Stiffkey, Wells-next-the-Sea	Arable	$2.65 X 10^6$	$1.86 X 10^5$	7.037545
PF10	UK	Nuffield, Henley-on-Thames	Arable	$1.88 X 10^6$	$1.94 X 10^5$	10.32825
PF11	UK	Tillingham, Southminster	Arable	$1.51 X 10^6$	$1.24 X 10^5$	8.175861
PF12	UK	Pettistree, Woodbridge	Arable	$1.80 X 10^6$	$1.58 X 10^5$	8.741072
PF13	UK	Wallington, Baldock	Arable	$2.40 X 10^6$	$1.61 X 10^5$	6.723151
PF14	UK	South Cockerington, Louth	Arable	$1.69 X 10^6$	$1.79 X 10^5$	10.55621
PF15	UK	Great Wilbraham, Cambridge	Arable	$1.78 X 10^6$	$1.92 X 10^5$	10.79306
PF16	UK	Shearman's Wath, Horncastle	Arable	$2.04 X 10^6$	$1.98 X 10^5$	9.699129
PF17	UK	Roughton, Norwich	Arable	$1.39 X 10^6$	$1.30 X 10^5$	9.377057
PF18	UK	Acle, Norwich	Arable	$1.80 X 10^6$	$2.10 X 10^5$	11.65542
PF19	UK	Flitcham, King's Lynn	Arable	$1.51 X 10^6$	$3.45 X 10^5$	22.77627
F1	Spain	Valencia	Arable	$2.20 X 10^7$	$2.53 X 10^6$	11.48477
F2	Spain	Valencia	Arable	$1.38 X 10^7$	$1.33 X 10^6$	9.641025
F3	Spain	Valencia	Arable	$3.40 X 10^7$	$1.67 X 10^6$	4.920017
F4	Spain	Ontinyent	Arable	$2.29 X 10^7$	$1.08 X 10^6$	4.722107

F5	Spain	Ontinyent	Arable	2.70 X 10^7	1.52 X 10^6	5.650904