Temperature stress tolerance in chickpea rhizobia

Ana Isabel Pereira Alexandre

Thesis presented to obtain the PhD degree

in Biology by the Universidade de Évora

Supervisor: Professor Isabel Solange Oliveira

187 185

Acknowledgements

The present thesis would not have been possible without the important collaboration of several institutions and people to whom I would like to thank:

- to the Universidade de Évora for providing the necessary conditions to carry on my work, namely to the ICAAM (Instituto de Ciências Agrárias e Ambientais Mediterrânicas) for supporting the investigation carried out in the Laboratório de Microbiologia do Solo and also for funding my participation in courses and conferences;

- to the Fundação para a Ciência e a Tecnologia (FCT) for the PhD fellowship (SFRH/BD/18162/2004) and also for the financed investigation projects from Programme POCTI (POCTI/BME/44140/2002) and Programme POCI 2010 (PTDC/BIO/80932/2006), co-financed by EU-FEDER;

- to Professor L. S. Dias for suggestions on the statistical analysis performed on chapter 2.

- to Professor Solange Oliveira for being my supervisor (for the second time), especially for her energy in answering one more question and reading one more manuscript version; and also for the right amount of guidance and freedom, which gave me the opportunity to grow as researcher. Thank you for the discussions and for the friendship.

Among colleagues and friends, many thanks are due:

- to all the people that worked in the Lab, for making it such a nice place to work and helping me in many situations;

- to the Mitra colleagues that I now proudly count as friends for all their care and for making lunch a very refreshing part of the day.

Marta for the lab coaching and for the precious friendly advices; Clarisse for the comic talks in the end of the day; Eng. Alho for the endless good mood and mechanical engineering; Isabel for always having a friendly word and Gertrudes for her daily help in the lab. Rosário for all her warm support and encouraging words; Carla for always keeping the good mood; Joana for the cloning advising (uff!!) and company in the endless days and Zé Manel for all the important silly thinking!

Extra special thanks to all my friends and family for believing in me.

Thank you all for making the difference!

Thesis publications

The present work is based on the following manuscripts:

1. **Alexandre, A.**, Laranjo, M., Young, J.P.W. and Oliveira, S. (2008) "*dnaJ* is a useful phylogenetic marker for alphaproteobacteria" International Journal of Systematic and Evolutionary Microbiology, 58 (12), 2839-2849.

2. **Alexandre, A.**, Brígido, C., Laranjo, M., Rodrigues, S. and Oliveira, S. (2009) "Survey of chickpea rhizobia diversity in Portugal reveals the predominance of species distinct from *Mesorhizobium ciceri* and *Mesorhizobium mediterraneum*" Microbial Ecology, 58, 930-941.

3. **Alexandre, A.** and Oliveira, S. "Most heat tolerant rhizobia show high induction of major chaperone genes upon stress" Accepted for publication in FEMS Microbiology Ecology.

Other publications

1. **Alexandre, A.**, Laranjo, M. and Oliveira, S. (2006) "Natural Populations of Chickpea Rhizobia Evaluated by Antibiotic Resistance Profiles and Molecular Methods" Microbial Ecology, 51, 128-136.

2. Brígido, C., **Alexandre, A.**, Laranjo, M. and Oliveira, S. (2007) "Moderately acidophilic mesorhizobia isolated from chickpea" Letters in Applied Microbiology, 44 (2), 168-174.

3. Laranjo, M., **Alexandre, A.**, Rivas, R., Velázquez, E., Young, J. P. W. and Oliveira, S. (2008) "Chickpea rhizobia symbiosis genes are highly conserved across multiple *Mesorhizobium* species" FEMS Microbiology Ecology, 66 (2), 391-400.

4. Oliveira, S., Laranjo, M., **Alexandre, A.** (2008). "Molecular Tools in Microbial Ecology" In "Ecology Research Trends", Chapter 4, pp. 161-183. Edited by Laia Diaz and Marta Perez, Nova Science Publishers, Inc., New York, USA.

Abbreviations

ADP	adenosine diphosphate
ATP	adenosine triphosphate
BNF	biological nitrogen fixation
bp	base pair
BSA	bovine serum albumin
CSP	cold shock protein
DEPC	diethylpyrocarbonate
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphate
HGT	horizontal gene transfer
HSP	heat shock protein
ILD	Incongruence length difference
kb	kilobase
kDa	kiloDalton
Mbp	Megabase pairs
min	minute
ML	maximum likelihood
MOPS	3-(N-morpholino)propanesulfonic acid
mRNA	messenger ribonucleic acid
NJ	neighbour-joining
PCR	polymerase chain reaction
RAPD	random amplified polymorphic DNA
RFLP	restriction fragment length polymorphism
RNA	ribonucleic acid
rRNA	ribosomal ribonucleic acid

SE	symbiotic effectiveness
sHSP	small heat shock proteins
YEM	yeast extract mannitol

Nucleotide Bases

A	Adenine
C	Cytosine
G	Guanine
T	Thymine
M	Amino (A or C)
N	Any nucleotide (A, C, G or T)
R	Purine (A or G)
S	Strong (G or C)
Y	Pyrimidine (C or T)

Genes

atpD	ATP synthase subunit β
gyrB	DNA gyrase subunit β
recA	DNA recombinase A
groEL	chaperone 60 kDa
groES	co-chaperone 10 kDa
dnaK	chaperone 70 kDa
dnaJ	co-chaperone 40 kDa
gapA	glyceraldehyde-3-phosphate dehydrogenase
rplB	50S ribosomal protein L2
16S rRNA	16S ribosomal RNA

Abstract

The aims of the present thesis were to study the chickpea rhizobia diversity and biogeography using Portugal as case study; to evaluate the temperature stress tolerance of the isolates, and to investigate the molecular basis of stress tolerance. The phylogenetic performance of the co-chaperone *dnaJ* was also addressed, in order to find an alternative marker to 16S rRNA gene.

According to the 16S rRNA gene phylogeny, most isolates were found to be distinct from the typical chickpea rhizobia species, *Mesorhizobium ciceri* and *M. mediterraneum*. Some provinces of origin are associated with particular species groups. *dnaJ* was found to be a useful phylogenetic marker for *Mesorhizobium* and for the *Alphaproteobacteria* class. The evaluation of temperature stress tolerance revealed tolerant and sensitive isolates to both heat and cold. Analysis of the expression of *dnaK* and *groESL* chaperone genes suggested that higher induction of these genes is related to higher tolerance to heat.

Tolerância ao stress térmico de rizóbio de grão-de-bico

Resumo

A presente tese teve como objectivos o estudo da diversidade e biogeografia de rizóbio de grão-de-bico em Portugal, a avaliação da tolerância dos rizóbios ao stress térmico, bem como o estudo das bases moleculares da tolerância ao stress. Estudou-se, ainda, o gene da co-chaperone *dnaJ* do ponto de vista filogenético.

A filogenia baseada no gene 16S rRNA revelou que a maior parte dos rizóbios de grão-de-bico agrupam com outras espécies, que não as típicas desta leguminosa (*Mesorhizobium ciceri* e *M. mediterraneum*). Encontrou-se uma associação entre algumas províncias e determinadas espécies de rizóbio. O gene *dnaJ* revelou-se um bom marcador filogenético para *Mesorhizobium*, bem como para a classe *Alphaproteobacteria*. A avaliação da tolerância à temperatura permitiu diferenciar isolados tolerantes e sensíveis, a altas e baixas temperaturas. A análise da expressão dos genes *dnaK* e *groESL*, sugeriu que uma maior indução destes genes está relacionada com maior tolerância a altas temperaturas.

List of Figures

Chapter 3 – *dnaJ* as a useful phylogenetic marker

Chapter 4 – Tolerance of chickpea rhizobia isolates to temperature stress

List of Tables

Table of Contents

1.

State of the art

The present work focused on an important eukaryotes-bacteria symbiosis: the legume-rhizobia symbiosis. Rhizobia are gram-negative bacteria mostly known by their ability to induce nodule formation in legume plants, including important agricultural species, as chickpea. In the nodules, rhizobia reduce atmospheric nitrogen into compounds that the plant is able to uptake and use as nitrogen source. This biological nitrogen fixation is very important to sustainable agriculture.

As soil bacteria, rhizobia are exposed to biotic and abiotic soil factors and have to compete with other organisms for limited resources. The present study investigated the diversity of native chickpea rhizobia and their temperature stress tolerance, comprising both phenotypic and molecular aspects of such tolerance.

1.1 – Biological Nitrogen Fixation

Nitrogen is the major component of the Earth atmosphere, but is often the limiting nutrient for life, because most organisms can only uptake nitrogen in the form of nitrates or ammonia compounds. Nevertheless, some prokaryotes, grouped under the term diazotrophs, are able to fix atmospheric nitrogen and convert it to molecules that are directly usable by non-fixing organisms.

The biological nitrogen fixation (BNF) is carried out by Bacteria and Archaea that synthesize the nitrogenase enzyme complex. This complex is composed by two components: the dinitrogenase reductase (Fe protein) and the dinitrogenase (Mo-Fe protein). The dinitrogenase is able to reduce the triple bond of the nitrogen atoms of atmospheric N_2, producing two molecules of ammonia and one molecule of hydrogen, using ATP and an electron donor (Fig. 1.1).

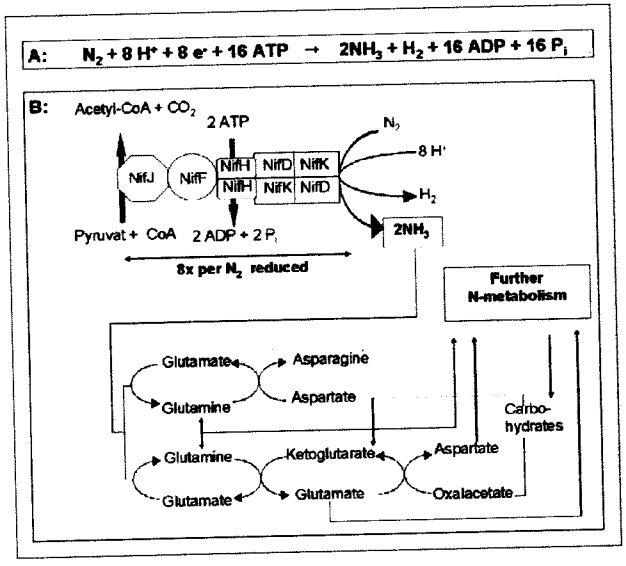


Figure 1.1 – Schematic view of the biological nitrogen fixation process. A – Chemical reaction of atmospheric dinitrogen reduction to ammonia. B – Diagram representation of the nitrogenase enzyme complex mechanism and initial metabolism of ammonia. *nif* genes code for the nitrogenase complex synthesis and catalysis. (from Kneip *et al.*, 2007).

Eukaryotes are not able to fix atmospheric nitrogen, but numerous examples of symbioses between Eukarya and diazotrophic Bacteria can be found (Kneip *et al.*, 2007). Some examples of these symbioses are well studied; others are more unexpected such as symbiosis between nitrogen fixing bacteria and termites or shipworms.

Rhizobia

An important BNF symbiosis that is of crucial importance in sustainable agriculture is the legume-rhizobia symbiosis. This symbiosis leads to the formation of special structures, usually in the plant roots, designated by nodules where the bacteria convert atmospheric nitrogen to ammonia. The ammonia produced in the bacteroids is transported to the plant, where it is assimilated into glutamine or asparagine.

Compared to the N fertilisers that mostly rely on fossil energy to be produced, BNF process is an environmental friendly way to supply N to agro-systems. The nitrogen resulting from this BNF is less prone to leaching than the chemical fertilisers used in agriculture (Jensen & Hauggaard-Nielsen, 2003). In addition, those fertilisers cause soil acidification and eutrophication, so rhizobia play a very important role in sustainable agriculture. Even when legumes are not the main crop, they can be used in rotation with for example cereal crops, in order to increase the N levels available in the soil and consequently improve cereals yields. For instance, chickpea was seen to successfully increase N-availability in cereal systems, as wheat production (Herridge *et al.*, 1995, Aslam *et al.*, 2003).

Other diazotrophs

Bacteria from the genus *Frankia* are gram-positive filamentous bacteria from the class *Actinobacteria*, able to establish nitrogen-fixing symbiosis with non-legume angiosperm families (Benson & Dawson, 2007). This actinorhizal symbiosis is very useful in soil erosion control, as well as in economically important activities as wood production. Rhizobia and *Frankia* are distantly related in terms of phylogeny, with each group belonging to different bacterial phyla. The main difference between *Frankia* and rhizobia is probably the absence of a nodulation genes cluster or symbiotic island in *Frankia* genomes (Normand *et al.*, 2007). In addition, the nodules structure and differentiation are also very distinct.

Cyanobacteria are gram-negative photosynthetic bacteria also able to fix atmospheric nitrogen. Nitrogen fixing cyanobacteria are an important symbiont of coral reefs, (Lesser *et al.*, 2004).

Within the Archaea domain, nitrogen fixation seems to be restricted to the methanogens group. Phylogenetic analysis of *nif* genes coding for the nitrogenase are consistent with an ancient origin of nitrogen fixation, probably before the divergence between Archaea and Bacteria (for review see Leigh, 2005).

1.2 – The legume-rhizobia symbiosis – rules and exceptions

Bacteria were first isolated from a leguminous plant root nodule and confirmed to induce nodule formation in the late XIX century (Beijerinck, 1888, Beijerinck, 1890). Presently, rhizobia can be defined as gram-negative soil bacteria able to establish a nitrogen-fixing symbiotic relationship with legume plants. Nevertheless, rhizobia are also involved in plant interactions other than symbiosis. These bacteria can be endophytic and act as plant growth promoting bacteria (PGPB). Economically important non-legume crops as rice or wheat are the best studied examples of rhizobia acting as PGPB (Biswas *et al.*, 2000, Boivin *et al.*, 1997).

In general, rhizobia are heterotrophic and aerobic non-sporulated rods. Nevertheless, there are *Bradyrhizobium* strains able to grow anaerobically (Polcyn & Lucinski, 2003); photosynthetic rhizobia are also known (So *et al.*, 1994); and furthermore, a methylotrophic strain able to nodulate and fix nitrogen in symbiosis with legumes was also reported (Sy *et al.*, 2001). In the rhizosphere, O_2 concentration is too high to allow proper nitrogenase complex functioning; thereby most rhizobia are efficient diazotrophs only in symbiosis. Even so, *Azorhizobium caulinodans* is an example of a rhizobium able to carry a diazotrophic free-living life (Lee *et al.*, 2008b).

1.2.1 – Rhizobia brief taxonomy

The genus *Rhizobium* was first described in 1889 (Frank, 1889), however only a small part of the bacteria generally designated under the term rhizobia belong to this genus.

Probably due to the agricultural importance of some legumes, the research in legume-rhizobia symbiosis has been an important investigation area. Only recently, wild legumes symbioses received more attention (Zahran, 2001). Some unexpected findings resulted from those investigations, as for example, the fact that bacteria from the *Betaproteobacteria* class are able to nodulate legumes (Moulin *et al.*, 2001). Until then, rhizobia were confined to the *Alphaproteobacteria*, namely in the order *Rhizobiales*, though this order includes many species that are not legume microsymbionts. The diversity of legume nodulating bacteria was clearly underestimated and recent studies report not only new species of rhizobia, but moreover strains from non-typical rhizobia genera that have been added to the rhizobia list. Indeed, rhizobia isolated from wild legumes from tropical regions, where the diversity of leguminous plants is higher, has lead to the description of numerous new species and extended rhizobia to new genera, as is the case of strains from the *Burkholderia* genus (Chen *et al.*, 2003, Chen *et al.*, 2008).

Rhizobia are gram-negative bacteria belonging to the *Proteobacteria* division. The *Alphaproteobacteria* class comprises 11 genera that include bacteria able to induce nodule formation in legume plants: *Agrobacterium*, *Allorhizobium*, *Azorhizobium*, *Bradyrhizobium*, *Devosia*, *Mesorhizobium*, *Methylobacterium*, *Ochrobactrum*, *Phyllobacterium*, *Rhizobium* and *Ensifer* (Euzéby, 1997). *Agrobacterium* genus has been proposed by Young *et al.* (2001) to be emended and reclassified as *Rhizobium*. The members of the genus *Sinorhizobium* were transferred to the genus *Ensifer* by decision of the Judicial Commission of the International Committee on Systematics of Prokaryotes (Tindall, 2008). Three additional genera belonging to the *Betaproteobacteria* class were recently added to the list of genera containing rhizobia species, namely, *Burkholderia* (Moulin *et al.*, 2001), *Cupriavidus* (Chen *et al.*, 2001) and *Herbaspirillum* (Valverde *et al.*, 2003). It is noteworthy to mention that these 14 genera also include a variety of non-symbiotic bacteria.

As the study of non-agricultural legumes is increasing and only a small percentage of leguminous species has been studied so far, the taxonomy of rhizobia will probably evolve, in order to reflect a higher diversity yet to know.

1.2.2 – Rhizobia genomes

The metabolic diversity found within rhizobia might be one explanation for their large and complex genome. In fact, the size of the rhizobial plasmids (often megaplasmids), together with the fact that they sometimes carry essential genes, has raised some questions on the bacterial genome organization and on the distinction between plasmids and chromosomes (Downie & Young, 2001). Until the end of 2009, 16 rhizobial complete genomes have been sequenced, including two rhizobia species from the *Betaproteobacteria* class. The size of those genomes ranges from 5.4 to 9.1 Mb and plasmid number varies between 0 and 7. For example, the *Mesorhizobium loti* MAFF303099 harbors two plasmids, one of 0.35 Mbp (pMla) and other of 0.21 Mbp (pMlb) both with 59% GC content, in addition to the 7.04 Mbp chromosome with 62% GC content. Symbiosis genes are often plasmid encoded, but this strain displays a chromosomal symbiosis island, which means that nitrogen fixation and nodulation genes are clustered in a 610 Kbp DNA fragment in the chromosome. Still, this DNA region is clearly different from the remaining chromosome, since it shows a lower GC content and codon usage more similar to that of the plasmids (Kaneko *et al.*, 2000). Another *Mesorhizobium*, *M. loti* R7a, also shows the symbiosis genes encoded in a

chromosomal symbiotic island, which is able to confer to non-symbiotic mesorhizobia the ability to nodulate *Lotus* species (Sullivan *et al.*, 2002).

Comparative genomics represents an important tool in studying the evolution of the legume-rhizobia symbiosis. Genome integrated comparisons allowed interesting findings, as for example the fact that no gene seems to be common or specific to all rhizobia (Amadou *et al.*, 2008). Thus, genes related to the symbiotic lifestyle were found to be either widely distributed (frequently paralogs of housekeeping genes) or encoded in few rhizobia species. This strongly indicates that the ability to establish nitrogen fixing symbiosis with legume plants evolved multiples times, with distinct strategies (Amadou *et al.*, 2008).

1.2.3 – Nitrogen fixation and nodulation

Nitrogen fixation genes (*nif* genes) are found in all diazotrophs, as they include structural genes for the nitrogenase and are required for both nitrogenase synthesis and catalysis (*nifH, D, K, Y, B, Q, E, N, X, U, S, V, W, Z*) (for review see Dixon & Kahn, 2004). *nif* genes are also involved in regulatory mechanisms, as for example the NifA protein, which is the main transcriptional activator of *nif* genes operons. The *nifH* gene codes for a subunit of the nitrogenase complex and is commonly used as a detection tool for screening the presence of diazotrophs (Burgmann *et al.*, 2004). Rhizobia use the molybdenum-nitrogenase for nitrogen fixation, however the number of *nif* genes found in different species is variable. For example, the *Azorhizobium caulinodans* genome comprises a total 15 different *nif* genes, while *Rhizobium leguminosarum* bv. *viciae* holds the lowest number of *nif* genes found in rhizobia (a total of 8 *nif* genes) (Young *et al.*, 2006). *nifHDK* are structural genes that encode nitrogenase reductase and the α and the β subunits of the dinitrogenase, respectively. The dinitrogenase requires a MoFe cofactor, whose assembly involves the *nifBEN* genes.

Nodules constrain bacteria into a confined space in the host plant and they also allocate the BNF process, providing the anoxic conditions that are essential to the nitrogenase complex functioning. The host-rhizobia signal exchange that leads to nodule formation has been extensively studied. Legume roots exudate flavonoids and isoflavonoids compounds that were traditionally considered as part of the plant defense mechanisms. However, these compounds were later found to be directly involved in the establishment of the legume-rhizobia symbiosis (for review see Dakora & Phillips, 1996). Although flavonoids are mostly known as inducers of the rhizobia nodulation genes (*nod* genes), they also can be *nod* genes inhibitors (Djordjevic *et al.*, 1987). The

interaction of the rhizobial NodD protein with specific legume flavonoids is considered the first level of host recognition. NodD will then activate the *nod* genes required to the synthesis of Nod factors. Nod factors are lipochitooligosaccharides (LCO) considered the second level of host specificity recognition because modifications on its structure, as well as the amounts of Nod factors released, are important in determining the host where nodulation can be initialized (Perret *et al.*, 2000). For instance, *nodC* is involved in the first step of the synthesis of Nod factors and is important in determining the length of the chitin oligosaccharide chain, which is one of the host determinant factors (Kamst *et al.*, 1997). Nod factors are also responsible for induction of cell division in the root cortex and pericycle, which is fundamental to initiate nodule formation.

The nodulation process is very complex, and despite the fact that general mechanisms of signaling between rhizobia and host plants are mostly understood, new strategies and exceptions to these processes arise from recent investigation. For example, the complete genome sequencing of two photosynthetic bradyrhizobia strains revealed the absence of *nod* genes in these rhizobia (Giraud *et al.*, 2007). This finding not only represents an exception to all the rhizobia symbiotically characterized so far, but also raises some questions on the initial step of the nodulation process.

The infection and nodule initiation processes can be very diverse. Still the better understood process is via an intracellular infection-thread, since this is a common pattern for most of the legume genera used in agriculture. Alternatively, rhizobia infection can be intercellular or occur without the formation of a typical infection-thread, as for example via crack entry. In addition, some legumes, like alfalfa, can form bacteria-free nodules (Truchet *et al.*, 1989).

Rhizobia undergo differentiation during nodule development into bacteroids, which are surrounded by a plant-derived membrane called peribacteroid membrane, forming the symbiosomes. Symbiosome membrane mediates nutrient and signal exchanges between the symbiotic partners. Regulation of O_2 concentration within the nodule is particularly important, due to the fact that nitrogenase activity is highly inhibited by oxygen (Hill, 1988). Leghemoglobin is an oxygen carrier produced by the legume, in rhizobia-infected roots, that allows the maintenance of low oxygen concentration in the nitrogenase active nodule and provides O_2 to the respiration sites (Appleby, 1984). Atmospheric N_2 fixation requires carbon and energy sources, which are provided by the plant as sucrose and dicarboxylic acids (carbon source) that are transported to the nodule. Plants have a negative feedback control mechanism that limits the number of root nodules formed and avoids hypernodulation, and consequent extra metabolic cost to the host legume (Oka-Kira & Kawaguchi, 2006).

As an exception to the typical root nodulation, some legumes species are able to establish nitrogen fixing symbioses with formation of stem nodules. Although stem nodulation it restricted to a small number of legume genera, mostly from flooded habitats, the diversity of the corresponding microsymbionts is high. The most studied stem rhizobia, the *Sesbania rostrata* - *Azorhizobium caulinodans* symbiosis, is also able to form root nodules (Boivin *et al.*, 1997).

Host range

Some legume species can establish nitrogen-fixing symbioses with rhizobia from different genera, while other legume species are very restrict for nodulation and only accept as microsymbionts a reduced number of species. Common bean (*Phaseolus vulgaris*) is considered a promiscuous host, as this legume is efficiently nodulated by strains of at least three rhizobia genera (*Bradyrhizobium*, *Rhizobium*, and *Ensifer*), nevertheless the host range depends on the legume cultivar used and conditions tested (Martinez-Romero, 2003).

Specificity can also be addressed from the bacterium perspective. The most classical example would be *Ensifer* sp. NGR234 that is able to nodulate over 120 plant genera, including the non-legume *Parasponia andersonii*. This strain genome (three replicons in a total of 6.9 Mbp) includes a symbiotic plasmid with virtually no essential genes and secretes a family of Nod factors more diverse than all other rhizobia known so far (Schmeisser *et al.*, 2009). In addition, the Nod factors concentration released by the NGR234 is much higher than usual. Finally, the complete genome sequence of this strain revealed that it encodes a large number of different secretion systems. All this features might, at least partially, explain the atypical broad host range of the *Ensifer* sp. NGR234.

The influence that the legume host could exert on the growth of specific rhizobia, as well as the competition among rhizobia in the rhizosphere, have been important issues in rhizobia investigation. Inositol derivatives, generally called rhizopines, are synthesized by the bacteroids of some rhizobia strains and seem to play a role in strains competitiveness. Fry and colleagues (2001) reported that mutant strains of *R. leguminosarum* bv. *viciae* in two genes required for inositol catabolism, were far less competitive than the wild type.

The size and diversity of indigenous rhizobia populations in soil is influenced by the host presence. An increase in the size of common bean-nodulating rhizobia population was associated with the continuous host crop (Andrade *et al.*, 2002). Other studies, in

soybean rhizobia, reported a reduced rhizobia diversity in the presence of the host (Coutinho *et al.*, 1999). In addition, rhizobial species relative abundance, as well as genetic diversity within species, may also shift in response to biotic and abiotic factors. Bradyrhizobia populations associated with soybean crop in Brazil mainly result from the frequent inoculation of this crop with two *Bradyrhizobium* species. However, a study addressing the bradyrhizobia diversity in inoculated fields found a high level of diversity and pointed out soil pH, soil organic matter and clay content as main factors influencing bradyrhizobia diversity (Giongo *et al.*, 2008).

1.2.4 – The host: the Fabaceae family

Legumes belong to the order Fabales and to the family Fabaceae, which is the third largest family of angiosperms. Three main legume subfamilies are distinguished: Caesalpinioideae, Mimosoideae and Papilionoideae. Most of the important agricultural crops, such as soybean (*Glicine max*), common bean (*Phaseolus vulgaris*) and chickpea (*Cicer arietinum*), are members of the Papilionoideae subfamily, while the Mimosoideae includes *Mimosa* spp., which were recently described to be nodulated by rhizobia belonging to the *Betaproteobacteria* (Chen *et al.*, 2005a, Moulin *et al.*, 2001). The Caesalpinioideae subfamily has very few nodulating members.

The exception to the strict legume-rhizobia symbiosis is *Parasponia* (Ulmaceae family), which is so far the only non-legume genus known to be nodulated by rhizobia (Trinick, 1973, Akkermans *et al.*, 1978). *Parasponia*-rhizobia associations usually have lower levels of nitrogen fixation when compared with those observed in legume-rhizobia symbiosis (Becking, 1983). *Parasponia* nodules are more similar to actinorhizal nodules than to the rhizobia induced ones (Patschkowski *et al.*, 1996).

The present study addresses rhizobia nodulating chickpea (*Cicer arietinum*), which is an important pulse crop traditionally grown in many countries, probably due to its adaptability to very different environments. This legume was one of the first grain legumes to be domesticated and its primary centres of origin are south-west Asia and Mediterraneum (Saxena & Singh, 1987). According to recent data available from FAO (Food and Agriculture Organization), India is by far the major chickpea producer with approximately 7.7 millions of tonnes produced in 2008 (http://faostat.fao.org/).

In Portugal, chickpea production decreased significantly in the 80's and it is now reduced to 650 tonnes (FAO data available for 2008), which means that Portugal imports almost all the chickpea consumed (Duarte Maçãs, 2003). Chickpea was traditionally sown in Spring , but the development of new cultivars allowed the transition

of this crop to Autumn, with higher yields that could represent an incentive to reintroduce this grain legume in Portugal (Duarte Maçãs, 2003). This Autumn/Winter sowing was also found to maximize the atmospheric nitrogen fixation carried out by rhizobia.

1.2.5 – The chickpea-rhizobia symbiosis

Despite the fact that chickpea is the second major pulse crop in terms of area cultivated (after common bean) and the third most important in terms of grain yield (Herridge *et al.*, 2008), its symbiosis with rhizobia has not been extensively studied probably because *Cicer arietinum* is not considered a model system, as *Medicago truncatula or Lotus japonicus* and because most of its yield comes from underdeveloped countries.

In the 90's the diversity of rhizobia able to nodulate chickpea begin to be consistently investigated (Nour *et al.*, 1994a, Kuykendall *et al.*, 1993) and Nour and co-workers described the first two chickpea rhizobia species, namely *Rhizobium ciceri* (1994b) and *Rhizobium mediterraneum* (1995). A few years later, these two species were transferred to the genus *Mesorhizobium* (Jarvis *et al.*, 1997), however the diversity of rhizobia able to nodulate chickpea was still scarcely revealed. All the effective rhizobia nodulating chickpea were so far described as belonging to the genus *Mesorhizobium*.

The *Mesorhizobium* genus

The *Mesorhizobium* genus is the more recently described rhizobia genus within the *Rhizobiales* order (Jarvis *et al.*, 1997). Currently, this genus comprises a total of 20 species, but this number is increasing rapidly, as indicated by the fact that half of these species were described in the last three years (Table 1.1). *Mesorhizobium* type strains were isolated from legume root nodules, with the exception of *M. thiogangeticum*, which was obtained from a legume rhizosphere (Ghosh & Roy, 2006). These bacteria with moderately slow growth rate also include non-symbiotic strains, as for example the completely sequenced *Mesorhizobium* sp. BNC1. Due to the high sequence similarity between species within *Mesorhizobium* genus, the use of alternative phylogenetic markers is very important for species definition (Chen *et al.*, 2005b), because some species can not be distinguished using the 16S rRNA sequence alone.

Table 1.1 – Species belonging to the *Mesorhizobium* genus, indicating the respective type strain, country and host of origin, as well as the reference of the species description.

Species	Type strain	Origin	Original host	Reference
M. albiziae	CCBAU 61158^{T}	China	*Albizia kalkora*	(Wang *et al.*, 2007)
M. alhagi	CCNWXJ12-2^{T}	China	*Alhagi sparsifolia*	(Chen *et al.*, *in press*)
M. amorphae	ACCC 19665^{T}	China	*Amorpha fruticosa*	(Wang *et al.*, 1999)
M. australicum	LMG 24608^{T}	Australia	*Biserrula pelecinus*	(Nandasena *et al.*, 2009)
M. caraganae	CCBAU 11299^{T}	China	*Caragana* spp.	(Guan *et al.*, 2008)
M. chacoense	LMG 19008^{T}	Argentina	*Prosopis alba*	(Velázquez *et al.*, 2001)
M. ciceri	UPM-Ca7^{T}	Spain	*Cicer arietinum*	(Nour *et al.*, 1994b)
M. gobiense	CCBAU 83330^{T}	China	*Oxytropis glabra*	(Han *et al.*, 2008)
M. huakuii	CCBAU 2609^{T}	China	*Astragalus sinicus*	(Chen *et al.*, 1991)
M. loti	NZP 2213^{T}	New Zealand	*Lotus corniculatus*	(Jarvis *et al.*, 1982)
M. mediterraneum	UPM-Ca36^{T}	Spain	*Cicer arietinum*	(Nour *et al.*, 1995)
M. metallidurans	STM 2683^{T}	France	*Anthyllis vulneraria*	(Vidal *et al.*, 2009)
M. opportunistum	LMG 24607^{T}	Australia	*Biserrula pelecinus*	(Nandasena *et al.*, 2009)
M. plurifarium	ORS 1032^{T}	Senegal	*Acacia senegal*	(de Lajudie *et al.*, 1998)
M. septentrionale	HAMBI 2582^{T}	China	*Astragalus adsurgens*	(Gao *et al.*, 2004)
M. shangrilense	CCBAU 65327^{T}	China	*Caragana* spp.	(Lu *et al.*, 2009)
M. tarimense	CCBAU 83306^{T}	China	*Lotus frondosus*	(Han *et al.*, 2008)
M. temperatum	HAMBI 2583^{T}	China	*Astragalus adsurgens*	(Gao *et al.*, 2004)
M. thiogangeticum	LMG 22697^{T}	India	*Clitoria ternatea* [(a)]	(Ghosh & Roy, 2006)
M. tianshanense	A-1BS^{T}	China	*Glycyrrhiza pallidiflora*	(Chen *et al.*, 1995)

[(a)] *M. thiogangeticum* was isolated from this species rhizosphere

Chickpea rhizobia diversity

Phenotypic characterization of chickpea rhizobia has been performed using, for example, biochemical and enzymatic tests (L'Taief *et al.*, 2007) and antibiotic resistance patterns (Alexandre *et al.*, 2006). The most common molecular approaches include RAPD fingerprinting, 16S rRNA sequencing or RFLP (Rivas *et al.*, 2006), repetitive extragenic palindromic (REP)-PCR (Ben Romdhane et al., 2008), plasmid profiles (Cadahía *et al.*, 1986) and total protein profile analysis (Alexandre *et al.*, 2006). Preliminary studies of genetic diversity indicated that chickpea rhizobia were too diverse to be comprised in only two *Mesorhizobium* species (Maâtallah *et al.*, 2002, Laranjo *et al.*, 2002). Laranjo and colleagues (2004) showed that isolates obtained from the southern region of Portugal were highly diverse and that some chickpea rhizobia were more related to *M. loti* and *M. tianshanense*. Besides these two species, also isolates close to *M. amorphae* were described as chickpea microsymbionts (Rivas

et al., 2007). Rhizobia populations able to establish symbiosis with chickpea were also studied in Tunisia (L'Taief *et al.*, 2007) and in Morocco (Maâtallah *et al.*, 2002). *Ensifer meliloti* strains isolated from Tunisian soils were found to be able to induce nodule formation in chickpea plants, but these nodules are ineffective (Romdhane *et al.*, 2007).

Phylogenetic analysis of two symbiosis genes (*nifH* and *nodC*) revealed that they are very conserved in chickpea mesorhizobia isolates, since isolates group all together with *M. ciceri* and *M. mediterraneum* type strains, regardless of their species affiliation (Laranjo *et al.*, 2008). Recently, another legume was found to be nodulated by *M. ciceri* strains, namely *Bisserula pelecinus*, however *M. ciceri* type strain (isolated from chickpea) is unable to nodulate *Bisserula* (Nandasena *et al.*, 2007). Furthermore, no cross-inoculation compatibility was found between chickpea *M. ciceri* strains and *Bisserula* mesorhizobia (Nandasena *et al.*, 2004). These findings confirmed chickpea as restricted host for nodulation, which means that chickpea rhizobia despite their diversity, fail to nodulate other legumes as previously reported (Gaur & Sen, 1979).

Inoculation of chickpea crops

Inoculation of different chickpea cultivars with highly efficient strains seems to result in a increased number of nodules and shoot dry weight (Ben Romdhane *et al.*, 2007). In addition, inoculation with multiple strains may further increase the total nitrogen content (Içgen *et al.*, 2002). Even for the biocontrol of plant parasites using plant growth promoting bacteria, rhizobia inoculation represents an advantage for chickpea crops (Siddiqui & Akhtar, 2009). Rhizobia inoculations together with phosphorus solubilizing bacteria, such as *Bacillus subtilis* or *Bacillus megaterium*, could be a good alternative to both nitrogen (N) and phosphorus (P) fertilizers in chickpea crops (Rudresh *et al.*, 2005, Elkoca *et al.*, 2008). On the other hand, rhizobia inoculation seems to activate chickpea genes involved with the production of phenolic compounds/phytoalexins and higher levels of these compounds benefice the plant by restricting disease development (Arfaoui *et al.*, 2007). This rhizobial protection was previously reported as effective on chickpea for the deleterious effects of Fusarium wilt (Arfaoui *et al.*, 2006).

1.3 – The temperature stress response

The most studied stress response is the cellular response to heat shock. Additionally, most of the knowledge on temperature stress response results from studies with the gram-negative bacterial model *Escherichia coli*. The main outcome from those studies was the comprehensive characterization of the most important heat shock proteins (HSP) and how they relate and interact (Fig. 1.2). HSPs include a large set of proteins that are induced by a rapid increase in the environmental temperature and are mainly chaperones and proteases. Interestingly, there are some HSPs conserved from bacteria to humans, which allows some level of extrapolation of the extensively investigated *E. coli* heat shock response to other organisms, in particular to other proteobacteria (Guisbert *et al.*, 2008).

The typical *E. coli* response to a temperature upshift is characterized by a rapid increase in the HSPs synthesis during the induction phase, then the synthesis declines during the adaptation phase and, at last, a new steady state is achieved (Straus *et al.*, 1987). The expression of HSPs genes is mostly controlled by the σ^{32} factor, which is encoded by the *rpoH* gene. σ^{32} functions as a temperature sensor itself, since at high temperature the *rpoH* mRNA losses its secondary structure that was hiding the ribosomal binding site, as well as the translation starting point, and translation becomes possible (Morita *et al.*, 1999). After the induction phase, degradation of σ^{32} mainly results from the activity of an inner membrane protease encoded by the *ftsH* gene (Tomoyasu *et al.*, 1995). Another sigma factor seems to be required to maintain the heat-shock response at higher temperatures, namely σ^{24}, which is encoded by the *rpoE* gene (Raina *et al.*, 1995).

Upon high temperatures, bacteria have to deal with protein denaturation that affects membrane and DNA stabilization, while at low temperatures the major problems are probably the cellular membrane fluidity and the rate of enzymatic reactions, as for example the initiation of translation. It was reported that a cold shock from 37ºC to 10ºC, in *E. coli*, leads to arrest of the synthesis of most proteins. The induction of cold shock proteins (CSP) has been investigated and the induction of small CSP seems to be a common feature between *E. coli* and *B. subtilis* (for review see Graumann & Marahiel, 1996). CSPs have been identified and characterized in *E. coli*, with special interest in CspA that seems to be the only CSP not detectable at normal growth temperature. Other proteins found to be overproduced under cold conditions are related to DNA structure and replication, as *recA* or *gyrA*, others are associated to RNA as the *rbfA* gene (Jones *et al.*, 1987). Contrary to the mostly transcriptional regulation of the heat shock response, the cold shock response seems to be more regulated at

the post-transcriptional level. For example, the stabilization of the *cspA* mRNA was found to be much higher at low temperatures than at 37ºC (Goldenberg *et al.*, 1996).

1.3.1 – The main chaperone systems

Chaperones are known for their role as folding modulators, in sequestering and stabilizing a wide range of polypeptides when the wrong conformational structure is presented. Under standard growth conditions most of these proteins are constitutively expressed because they are essential in the folding of nascent polypeptide chains (Frydman, 2001).

Two chaperone systems are widely studied and play central roles in the heat stress response: DnaK-DnaJ and GroES-GroEL (Fig. 1.2). These proteins have a different designation when referring to eukaryotes, namely HSP70 (DnaK), HSP40 (DnaJ) and HSP60 (GroEL). When stress conditions are too severe, DnaKJ and GroESL systems are unable to avoid protein aggregation and the intervention of another kind of chaperones, termed disaggregating chaperones is required, as is the case of ClpB (Baneyx & Mujacic, 2004). Despite the fact that DnaKJ and GroESL are referred as heat shock proteins, these systems are often activated in response to other cell destabilizations that cause protein misfolding.

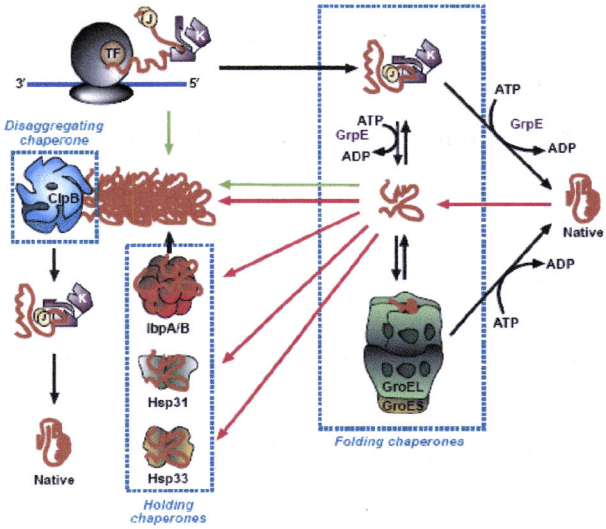


Figure 1.2 – Chaperone-mediated protein folding in *Escherichia coli*, which is probably the most studied bacterium concerning heat shock protein research. The diagram shows the interaction between several types of chaperones, in order to achieve native conformation of both nascent polypeptides and unfolded proteins resulting from stress conditions. (from Baneyx & Mujacic, 2004)

The DnaK-DnaJ system

The importance of the DnaKJ system in bacteria can be demonstrated by the analysis of the diversity and magnitude of phenotypic alterations obtained by mutation studies. Depleted growth rate, alteration of cell division leading to filamentous growth, loss of adaptive resistance to heat and loss of antibiotic resistance phenotype, are some examples of phenotype changes observed in *dnaK* mutants.

The chaperone DnaK usually interacts closely with two other proteins, DnaJ and GrpE, called co-chaperones or co-factors. The target proteins for the DnaKJ system are peptides with exposed hydrophobic residues that under normal conditions would be hidden in the protein core (Rudiger *et al.*, 1997). The activity of the DnaKJ machinery is an ATP-driven cycle, where the binding and hydrolysis of ATP induces conformational changes in the ATPase domain (NH_2-terminal) of DnaK, which are communicated to the peptide-binding domain (COOH-terminal) (for review see Genevaux *et al.*, 2007). This ATP-ADP binding and release is mediated by GrpE. The GrpE activity decreases at high temperatures, which means that the ATPase cycle becomes slower. Furthermore, GrpE acts as a thermosensor since it can undergo a temperature-induced conformational change (Siegenthaler & Christen, 2006).

The highly conserved J-domain (NH_2-terminal) of DnaJ, approximately 70 amino acids long, is found in all DnaJ homologues and stimulates the hydrolysis of DnaK-bound ATP (Wall *et al.*, 1994). Other three distinct regions are usually pointed out in DnaJ: the glycine and phenylalanine (G/F)-rich domain, which is thought to be a flexible linker; a conserved region with the consensus sequence CxxCxGxG that includes two zinc atoms; and a variable length and not conserved C-terminal region. The DnaJ protein with substrate protein bound at the C-terminal domain interacts with DnaK, through its J-domain. DnaJ is able to act as a chaperone on its own (Szabo *et al.*, 1994), but its main role is as co-chaperone interacting with DnaK.

DnaKJ system is also able to participate in the protein secretion machinery of the cell and cell division (Wild *et al.*, 1992, Bukau & Walker, 1989). *In vivo*, a steady-state *E. coli* cell has about 10 times more DnaK than DnaJ (Bardwell *et al.*, 1986), nevertheless viability at 42ºC is affected by even small decreases in the levels of these two proteins (Tomoyasu *et al.*, 1998).

Recently, two *dnaK*-like genes were found in *E. coli*, in addition to the several genes, other than *dnaJ*, that have a J domain. Despite the fact that these *dnaK*-like genes are not heat inducible, some function overlapping was found, what makes this cell machinery more complex than initially expected (Genevaux *et al.*, 2007).

The GroES-GroEL system

GroEL (HSP60) is an ubiquitous protein found in all eubacterial cells and also present in eukaryotes organelles, as for example mitochondrias. GroES-GroEL are considered broad spectrum chaperonines. For instance, in *E. coli* approximately half of the soluble proteins can bind to GroEL when presenting a wrong conformation. Accordingly, null mutations in *groES* or *groEL* are lethal for *E. coli* (Fayet *et al.*, 1989).

Each monomer from the GroEL heptameric ring can be divided in three domains: domain E corresponds to the equatorial section of the double ring and comprises the ATP/ADP/Mg^{2+}-binding pocket; domain A is the apical part, which binds to the co-chaperone GroES and to the substrate; domain I (intermediate) connects the other two domains and provides essential residues for ATP hydrolysis (Karlin & Brocchieri, 2000).

GroES induces conformational changes in GroEL so that the central cavity approximately doubles its size (for review see Chaudhuri *et al.*, 2009). This wider space that acquires hydrophilic properties, allows small polypeptides to fold in a protected environment. The hydrolysis of ATP weakens the interaction between GroES and GroEL and causes the release of the polypeptide. This mechanism is designated by *cis* ring folding and it is not clear whether it requires active folding by GroESL or if the polypeptide folds passively in the chaperone hydrophilic cavity. For larger polypeptides the described *cis* mechanism is not possible, so folding via *trans* mechanism takes place. In this mechanism, individual domains might fold individually or after a single domain becomes correctly folded the rest of the protein will fold for itself. There are also some substrates that bind transiently to GroEL alone.

GroEL seems to act as a temperature sensor given that high temperatures induce a slight conformational change that makes the GroESL machinery to work more like a holdase than a foldase (for review see Muga & Moro, 2008). Moreover, small changes in the nucleotide composition may alter the GroESL functionality and interestingly the consequence of such change might be the extension in the growth temperature range. For example, a six amino acid mutation in the GroESL allowed *E. coli* DH5α to be able to growth at lower temperatures than the wild type (Lee *et al.*, 2008a). The finding that a single chaperone system is able to define the temperature limit for growth was previously demonstrated by Ferrer and colleagues (2003) that transformed *E. coli* with *groESL* genes from a psychrophilic bacterium. The modified strain grew much faster at low temperatures than the parental strain and more interestingly, the lower limit for growth was changed.

1.3.2 – The temperature stress response in rhizobia

Soil environmental conditions are critical to the persistence of rhizobia in the soil, since changes in the rhizosphere environment can affect both growth and saprophytic competence, which will influence competitiveness and persistence (Dowling & Broughton, 1986). In order to improve the yield of legumes in more adverse environments, stress-tolerant cultivars should be combined with stress tolerant rhizobia.

Studies on rhizobia biodiversity are an important approach to find more tolerant strains, even when non-adverse environments are sampled, since populations often contain tolerant strains to non-acting stresses, as resilience to respond to future problems (Giller *et al.*, 1997). The ability of rhizobia to persist in the absence of their host plant is perhaps more dependent of their ability to endure adverse environmental factors than during symbiosis, where the nodule represents a protective environment. Under stress conditions, namely osmotic stress, it has been reported that the rhizobia input is an important factor in explaining the total variance of the nitrogenase activity of chickpea plants (Mhadhbi *et al.*, 2008). Several environmental factors shape the composition and the activity of rhizospheric habitats, namely that of rhizobia populations (Zahran, 1999). Soil pH and temperature are often pointed out as the major factors in determining the bacterial community diversity (Fierer & Jackson, 2006, Staddon *et al.*, 1998).

The optimum temperature for rhizobia growth is 25-30ºC (Zhang *et al.*, 1995), however in both saprophytic and symbiotic life rhizobia are often subject to temperatures out of this range. Most studies in rhizobial temperature stress tolerance focus soybean and common-bean mycrosymbionts. Soybean isolates grow weakly at 40ºC and no isolate was able to grow at 45ºC (Chen *et al.*, 2002). Rhizobia nodulating *P. vulgaris* can survive to 47ºC, but their symbiotic effectiveness is loss at high temperatures; while other isolates tolerant to 40ºC were able to remain infective at that temperature (Karanja & Wood, 1988). Nandal *et al.* (2005) reported that mutants tolerant to high temperature (43ºC), obtained from a thermosensitive *Rhizobium* sp. strain, exhibited a different protein profile from that of the wild type at high temperature, namely the mutant strains showed overexpressed proteins and new proteins. A protein of 63-75 kDa was overproduced in all mutant strains, which probably corresponds to DnaK. In chickpea rhizobia, a 60 kDa protein that could correspond to GroEL was found to be consistently overproduced when isolates were submitted to heat stress (Rodrigues *et al.*, 2006). Strains isolated from a chickpea wild relative (*Cicer anatolicum*), collected from high altitudes, were successful in nodulating chickpea at low temperatures, what

represents an alternative source of chickpea nodulating rhizobia with potential use as inoculants (Ogutcu *et al.*, 2008). Successful increase in chickpea yield in water-limited soils was achieved by using competitive drought and salt tolerant rhizobia as inoculants (Ben Romdhane *et al.*, 2008). Investigations carried out with bean nodulating rhizobia showed a correlation between isolates tolerance to high temperatures and temperatures from the origin site (Karanja & Wood, 1988), however in the case of soybean rhizobia the geographical origin does not correlate with their tolerance to heat stress (Munevar & Wollum, 1981).

Molecular bases of temperature stress tolerance in rhizobia

The molecular bases of temperature stress tolerance remains mostly unstudied in rhizobia, namely the main chaperones systems (DnaKJ and GroESL). Most studies regarding the heat shock response were performed using *Bradyrhizobium japonicum.*

Probably the first study identifying proteins differentially expressed after a heat shock in rhizobia was performed by Michiels and collaborators (1994). This study in bean-nodulating rhizobia, compared a heat tolerant species with another species that was heat sensitive and found common proteins highly expressed after a temperature upshift, particularly proteins of low molecular mass. These small heat shock proteins (sHSP) were further investigated because a high number of heat inducible sHSP seems to be a typical rhizobia feature (Münchbach *et al.*, 1999). sHSP include proteins similar to *E. coli* IbpA and IbpB and other proteins more divergent in sequence and phylogenetic origin. Furthermore, a specific regulation mechanism was found for operons coding these sHSPs (Nocker *et al.*, 2001). This particular type of regulation consists in a *cis*-acting DNA motif designated by ROSE (repression of heat shock gene expression) that confers temperature induction to operons with a typical housekeeping promoter (σ^{70} -type) and was already found in *Bradyrhizobium* species, *Ensifer* sp. NGR234 and *Mesorhizobium loti* MAFF303099 (Narberhaus *et al.*, 1998, Nocker *et al.*, 2001).

B. japonicum shares similar regulation mechanisms with to both *E.coli* (gram-negative) and *B. subtilis* (gram-positive) (Babst *et al.*, 1996). In addition to the large number of sHSP (at least 12) (Münchbach *et al.*, 1999), *B. japonicum* also displays the highest number of *groEL* operons found in rhizobia. From the five *groESL* operons annotated in the *B. japonicum* genome, only $groESL_{1,4,5}$ are heat inducible and they are differently regulated. The $groESL_1$ is σ^{32} dependent and another system controls the $groESL_{4,5}$ regulation, namely the CIRCE element that was initially described in *Bacillus subtilis*

(Babst *et al.*, 1996). These two distinct regulations of the heat shock response showed differences in their induction kinetics.

In *Rhizobium leguminosarum* only one of the three *groEL* homologues is needed for normal growth and corresponds to the highly expressed one (Rodriguez-Quinones *et al.*, 2005). Similarly to what was found in *B. japonicum*, both CIRCE and σ^{32} dependent regulation was found *R. leguminosarum* (Gould *et al.*, 2007). There are few complementation experiments evolving rhizobial chaperones. Ivic and colleagues (1997) complemented an *E. coli groEL* mutant with a *groEL* gene from *R. leguminosarum* bv. *viciae* and observed that this complementation worked at 37ºC but not at 43ºC.

The multiple *groESL* copies found in rhizobia share some degree of functional overlapping, but clearly there is some degree of specialization that is probably more related to different regulation modes than to differences in the protein sequences. Furthermore, it seems that the different regulatory mechanisms of *B. japonicum* are not independent, as the HrcA repressor from the CIRCE element was found to be under the control of σ^{32}-dependent promoter (Minder *et al.*, 2000).

Multiple copies of the *rpoH* gene (gene coding for σ^{32}) were detected for the first time in *B. japonicum* (Narberhaus *et al.*, 1997) and later on, two copies were detected in *S. meliloti* (Ono *et al.*, 2001). In the first case, individual mutants on each of the three homologs maintained the same temperature sensitivity and symbiotic phenotype as the wild type, while in the former species the *rpoH* genes seemed to be involved in both heat shock response and nodulation/fixation phenotype in alfalfa.

In general, and particularly in rhizobia, the DnaKJ system is less studied than the GroESL. The *dnaKJ* operon was characterized in *B. japonicum* and was found to be under the control of σ^{32} factor (Minder *et al.*, 1997). The *dnaK* gene seems to be essential for *B. japonicum* growth, since no *dnaK* mutant could be obtained. The *dnaJ* mutant showed slower growth, particularly at high temperatures but the symbiotic properties were unaffected.

Effects of temperature stress in nodulation and nitrogen fixation

The effects of temperature stress in nodulation and nitrogen fixation have been recognized for a long time, as the first studies addressing this subject can be traced back to the 1960 decade, at least. Even before nodule establishment, root zone temperature influences the rhizobial survival in soil, as well as the exchange of molecular signals between the two symbiotic partners (Sadowsky, 2005). High

temperature was seen to have an inhibitory effect on adherence of bacteria to root hairs, root hair formation and infection thread formation; for example, among tropical grain legumes no nodules were formed on peanut at 40ºC and on soybean at 37ºC (for review see Hungria & Vargas, 2000). Nodule functioning can also be influenced by high temperatures, namely the decreased in the rate of some essential reactions, as the synthesis of leghaemoglobin or the nitrogenase activity, as well as the activity of enzymes involved in the nitrogen assimilation process. On the other hand, low temperatures also affect nodulation, as cold delays nodulation initiation or even completely inhibits the process and seems to influence nodule occupancy (for review see Graham, 1992). In soybean, several studies showed that low root zone temperature influences infection and early nodulation, as well as N_2 fixation, delaying or even shutting down these processes within a temperature shift from 25ºC to 15ºC (Zhang *et al.*, 1995). A more recent study, also with *B. japonicum*, emphasizes the fact that Nod factor production markedly decreases at 17ºC or 15ºC, despite the fact that its biological activity remains unaltered (Duzan *et al.*, 2006). Nevertheless, the temperature threshold for substantial delay in nodulation differs from one legume to another. Bean and soybean have similar threshold, while lentil is more tolerant and nodulation is significantly delayed only at lower temperatures (10ºC) (Lira *et al.*, 2005).

The tolerance of rhizobia to high temperature *in vitro* does not always correlate with the phenotype during symbiosis under the same conditions, at least for bradyrhizobia nodulating soybean and *Rhizobium phaseoli* nodulating common bean (Zahran, 1999). However, other studies found that strain tolerance to high temperature assessed *in vitro* correlated well with the results obtained in symbiotic trials (Kulkarni & Nautiyal, 1999, Kishinevsky *et al.*, 1992). The existence of this correlation is probably due to the fact that a higher symbiotic performance under high temperatures depends more of the ability to survive under heat, than of the ability to fix nitrogen under those conditions, since the nitrogenase activity is drastically repressed after a temperature upshift, regardless of the strain susceptibility to heat (Michiels *et al.*, 1994).

Involvement of chaperone genes in the symbiosis

B. japonicum mutants that individually lack one *groEL* gene have the same symbiotic phenotype as the wild type, (Fischer *et al.*, 1993). Nevertheless, simultaneous mutation of $groEL_3$ and $groEL_4$ affects the symbiotic performance, despite the fact that there is no change in the expression levels of symbiotic genes (Fischer *et al.*, 1999). These two copies are the ones that contribute the most to the GroEL pool in bacteroids

(Fischer *et al.*, 1993) and their mutation seems to affect the levels of NifH and NifDK nitrogenase proteins (Fischer *et al.*, 1999). Under the control of *groEL*$_3$ promoter (σ^{54} and NifA), all other *B. japonicum groEL* copies and even the *E. coli* gene are able to partially complement the symbiotic defected double mutant. Overall, these studies indicate that in *B. japonicum*, GroEL is necessary to the formation of a functional nitrogenase. A few years before these findings, a study with *Klebsiella pneumoniae* already indicated that *groEL* is involved in the rate and level of accumulation of the *nif* system components (Govezensky *et al.*, 1991).

In *Sinorhizobium meliloti* genome, five *groESL* operons were found, and also in this species, one operon was found to be involved in symbiosis (Ogawa & Long, 1995). As in *B. japonicum*, all single mutants are viable but some double mutants are depleted in their symbiotic phenotype (Fischer *et al.*, 1999, Bittner *et al.*, 2007).

Labidi and colleagues (2000) found that the *dnaJ* gene is necessary for effective symbiosis of *R. leguminosarum* bv. *phaseoli* with common bean, contrary to what was seen in *B. japonicum* (Minder *et al.*, 1997).

1.4 – Aims of the study

Despite the studies mentioned before, chickpea rhizobia diversity is still not extensively characterized and little is known about temperature stress tolerance in chickpea rhizobia, especially in what concerns the molecular basis of temperature tolerance. Thus, the main objectives of the present study were:

1 – to evaluate the diversity of chickpea rhizobia populations in Portugal;

2 – to investigate the co-chaperone *dnaJ* gene, as alternative phylogenetic marker to the 16S rRNA gene for the *Mesorhizobium* genus and also for higher taxonomic levels;

3 – to evaluate the tolerance of chickpea rhizobia isolates to temperature stress;

4 – to investigate the molecular bases of thermotolerance in rhizobia, namely by the analysis of the expression of the major chaperone genes: *dnaKJ* and *groESL*.

1.5 – References

Akkermans, A. D. L., S. Abdulkadir & M. J. Trinick, (1978) N2-fixing root nodules in Ulmaceae - *Parasponia* or (and) *Trema* spp.? *Plant Soil* **49**: 711-715.

Alexandre, A., M. Laranjo & S. Oliveira, (2006) Natural populations of chickpea rhizobia evaluated by antibiotic resistance profiles and molecular methods. *Microb. Ecol.* **51**: 128-136.

Amadou, C., G. Pascal, S. Mangenot, M. Glew, C. Bontemps, D. Capela, S. Carrere, S. Cruveiller, C. Dossat, A. Lajus, M. Marchetti, V. Poinsot, Z. Rouy, B. Servin, M. Saad, C. Schenowitz, V. Barbe, J. Batut, C. Medigue & C. Masson-Boivin, (2008) Genome sequence of the beta-rhizobium *Cupriavidus taiwanensis* and comparative genomics of rhizobia. *Genome Res.* **18**: 1472-1483.

Andrade, D. S., P. J. Murphy & K. E. Giller, (2002) The diversity of *Phaseolus*-nodulating rhizobial populations is altered by liming of acid soils planted with *Phaseolus vulgaris* L. in Brazil. *Appl. Environ. Microbiol.* **68**: 4025-4034.

Appleby, C. A., (1984) Leghemoglobin and *Rhizobium* Respiration. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **35**: 443-478.

Arfaoui, A., A. El Hadrami, Y. Mabrouk, B. Sifi, A. Boudabous, I. El Hadrami, F. Daayf & M. Cherif, (2007) Treatment of chickpea with *Rhizobium* isolates enhances the expression of phenylpropanoid defense-related genes in response to infection by *Fusarium oxysporum* f. sp *ciceris*. *Plant Physiol. Biochem.* **45**: 470-479.

Arfaoui, A., B. Sifi, A. Boudabous, I. El Hadrami & M. Cherif, (2006) Identification of Rhizobium isolates possessing antagonistic activity against *Fusarium oxysporum* f.sp *ciceris*, the causal agent of Fusarium wilt of chickpea. *J. Plant Pathol.* **88**: 67-75.

Aslam, M., I. A. Mahmood, M. B. Peoples, G. D. Schwenke & D. F. Herridge, (2003) Contribution of chickpea nitrogen fixation to increased wheat production and soil organic fertility in rain-fed cropping. *Biol. Fertility Soils* **38**: 59-64.

Babst, M., H. Hennecke & H. M. Fischer, (1996) Two different mechanisms are involved in the heat-shock regulation of chaperonin gene expression in *Bradyrhizobium japonicum*. *Mol. Microbiol.* **19**: 827-839.

Baneyx, F. & M. Mujacic, (2004) Recombinant protein folding and misfolding in *Escherichia coli*. *Nat. Biotechnol.* **22**: 1399-1408.

Bardwell, J. C. A., K. Tilly, E. Craig, J. King, M. Zylicz & C. Georgopoulos, (1986) The Nucleotide Sequence of the *Escherichia coli* K12 *dnaJ* Gene - a Gene That Encodes a Heat-Shock Protein. *J. Biol. Chem.* **261**: 1782-1785.

Becking, J. H., (1983) The *Parasponia parviflora* - Rhizobium symbiosis. Host specificity, growth and nitrogen fixation under various conditions. *Plant Soil* **75**: 309-342.

Beijerinck, M. W., (1888) Die Bacterien der Papilionaceenknölchen. *Botanische Zeitung* **46**: 797-804.

Beijerinck, M. W., (1890) Künstliche infektion von Vicia faba mit *Bacillus radicola*. Ernährungsbedingungen dieser Bacterien. *Botanische Zeitung* **48**: 837-843.

Ben Romdhane, S., M. E. Aouani, M. Trabelsi, P. de Lajudie & R. Mhamdi, (2008) Selection of High Nitrogen-Fixing Rhizobia Nodulating Chickpea (*Cicer arietinum*) for Semi-Arid Tunisia. *Journal of Agronomy and Crop Science* **194**: 413-420.

Ben Romdhane, S., F. Tajini, M. Trabelsi, M. Aouani & R. Mhamdi, (2007) Competition for nodule formation between introduced strains of *Mesorhizobium ciceri* and the native populations of rhizobia nodulating chickpea (*Cicer arietinum*) in Tunisia. *World J. Microbiol. Biotechnol.* **23**: 1195-1201.

Benson, D. R. & J. O. Dawson, (2007) Recent advances in the biogeography and genecology of symbiotic Frankia and its host plants. *Physiol. Plant.* **130**: 318-330.

Biswas, J. C., J. K. Ladha, F. B. Dazzo, Y. G. Yanni & B. G. Rolfe, (2000) Rhizobial inoculation influences seedling vigor and yield of rice. *Agron. J.* **92**: 880-886.

Bittner, A. N., A. Foltz & V. Oke, (2007) Only one of five *groEL* genes is required for viability and successful symbiosis in *Sinorhizobium meliloti*. *J. Bacteriol.* **189**: 1884-1889.

Boivin, C., I. Ndoye, F. Molouba, P. deLajudie, N. Dupuy & B. Dreyfus, (1997) Stem nodulation in legumes: Diversity, mechanisms, and unusual characteristics. *Crit. Rev. Plant Sci.* **16**: 1-30.

Bukau, B. & G. C. Walker, (1989) delta-*dnaK52* mutants of *Escherichia coli* have deffects in chromosome segregation and plasmid maintenance at normal growth temperatures. *J. Bacteriol.* **171**: 6030-6038.

Burgmann, H., F. Widmer, W. Von Sigler & J. Zeyer, (2004) New molecular screening tools for analysis of free-living diazotrophs in soil. *Appl. Environ. Microbiol.* **70**: 240-247.

Cadahía, E., A. Leyva & T. Ruiz-Argüeso, (1986) Indigenous plasmids and cultural characteristics of rhizobia nodulating chickpeas (*Cicer arietinum* L.). *Arch. Microbiol.* **146**: 239-244.

Chaudhuri, T. K., V. K. Verma & A. Maheshwari, (2009) GroEL assisted folding of large polypeptide substrates in *Escherichia coli*: Present scenario and assignments for the future. *Prog. Biophys. Mol. Biol.* **99**: 42-50.

Chen, L. S., A. Figueredo, H. Villani, J. Michajluk & M. Hungria, (2002) Diversity and symbiotic effectiveness of rhizobia isolated from field-grown soybean nodules in Paraguay. *Biol. Fertility Soils* **35**: 448-457.

Chen, W.-M., S. M. de Faria, R. Straliotto, R. M. Pitard, J. L. Simoes-Araujo, J.-H. Chou, Y.-J. Chou, E. Barrios, A. R. Prescott, G. N. Elliott, J. I. Sprent, J. P. W. Young & E. K. James, (2005a) Proof that Burkholderia Strains Form Effective Symbioses with Legumes: a Study of Novel Mimosa-Nodulating Strains from South America. *Appl. Environ. Microbiol.* **71**: 7461-7471.

Chen, W.-M., L. Moulin, C. Bontemps, P. Vandamme, G. Béna & C. Boivin-Masson, (2003) Legume Symbiotic Nitrogen Fixation by beta-Proteobacteria Is Widespread in Nature. *J. Bacteriol.* **185**: 7266-7272.

Chen, W. M., S. M. de Faria, J. H. Chou, E. K. James, G. N. Elliott, J. I. Sprent, C. Bontemps, J. P. W. Young & P. Vandamme, (2008) *Burkholderia sabiae* sp nov., isolated from root nodules of *Mimosa caesalpiniifolia*. *Int. J. Syst. Evol. Microbiol.* **58**: 2174-2179.

Chen, W. M., S. Laevens, T. M. Lee, T. Coenye, P. De Vos, M. Mergeay & P. Vandamme, (2001) *Ralstonia taiwanensis* sp nov., isolated from root nodules of *Mimosa* species and sputum of a cystic fibrosis patient. *Int. J. Syst. Evol. Microbiol.* **51**: 1729-1735.

Chen, W. X., G. S. Li, Y. L. Qi, E. T. Wang, H. L. Yuan & J. L. Li, (1991) *Rhizobium huakuii* sp. nov. Isolated from the Root-Nodules of *Astragalus sinicus*. *Int. J. Syst. Bacteriol.* **41**: 275-280.

Chen, W. X., E. Wang, S. Y. Wang, Y. B. Li & X. Q. Chen, (1995) Characteristics of *Rhizobium tianshanense* sp. nov., a moderately and slowly growing root-nodule bacterium isolated from an arid saline environment in Xinjiang, Peoples-Republic-of-China. *Int. J. Syst. Bacteriol.* **45**: 153-159.

Chen, W. X., E. T. Wang & D. Kuykendall, (2005b) Genus VI. *Mesorhizobium*. In: Bergeys Manual of Systematic Bacteriology. New York: Springer, pp. pp. 403-408.

Coutinho, H. L. C., V. M. Oliveira, A. Lovato, A. H. N. Maia & G. P. Manfio, (1999) Evaluation of the diversity of rhizobia in Brazilian agricultural soils cultivated with soybeans. *Applied Soil Ecology* **13**: 159-167.

Dakora, F. D. & D. A. Phillips, (1996) Diverse functions of isoflavonoids in legumes transcend anti-microbial definitions of phytoalexins. *Physiol. Mol. Plant Pathol.* **49**: 1-20.

de Lajudie, P., A. Willems, G. Nick, F. Moreira, F. Molouba, B. Hoste, U. Torck, M. Neyra, M. D. Collins, K. Lindström, B. Dreyfus & M. Gillis, (1998) Characterization of tropical tree rhizobia and description of *Mesorhizobium plurifarium* sp. nov. *Int. J. Syst. Bacteriol.* **48**: 369-382.

Dixon, R. & D. Kahn, (2004) Genetic Regulation of Biological Nitrogen Fixation. *Nature Reviews Microbiology* **2**: 621-631.

Djordjevic, M. A., J. W. Redmond, M. Batley & B. G. Rolfe, (1987) Clovers Secrete Specific Phenolic-Compounds Which Either Stimulate or Repress Nod Gene-Expression in *Rhizobium trifolii*. *EMBO J.* **6**: 1173-1179.

Dowling, D. N. & W. J. Broughton, (1986) Competition for nodulation of legumes. *Annu. Rev. Microbiol.* **40**: 131-157.

Downie, J. A. & J. P. W. Young, (2001) Genome sequencing - The ABC of symbiosis. *Nature* **412**: 597-598.

Duarte Maçãs, I. M. V., (2003) Selecção de linhas de grão de bico (*Cicer arietinum* L.) adaptadas ao ambiente Mediterrânico-critérios morfológicos e fisiológicos. In. Évora: Universidade de Évora, pp. 171.

Duzan, H. M., F. Mabood, A. Souleimanov & D. L. Smith, (2006) Nod Bj-V (C18:1, MeFuc) production by *Bradyrhizobium japonicum* (USDA110, 532C) at suboptimal growth temperatures. *J. Plant Physiol.* **163**: 107-111.

Elkoca, E., F. Kantar & F. Sahin, (2008) Influence of nitrogen fixing and phosphorus solubilizing bacteria on the nodulation, plant growth, and yield of chickpea. *J. Plant Nutr.* **31**: 157-171.

Euzéby, J. P., (1997) List of Bacterial Names with Standing in Nomenclature: a folder available on the Internet. *Int. J. Syst. Evol. Microbiol.*: 590-592. http://www.bacterio.cict.fr.

Fayet, O., T. Ziegelhoffer & C. Georgopoulos, (1989) The *groES* and *groEL* heat shock gene products of *Escherichia coli* are essential for bacterial growth at all temperatures. *J. Bacteriol.* **171**: 1379-1385.

Ferrer, M., T. N. Chernikova, M. M. Yakimov, P. N. Golyshin & K. N. Timmis, (2003) Chaperonins govern growth of *Escherichia coli* at low temperatures. *Nat. Biotechnol.* **21**: 1266-1267.

Fierer, N. & R. B. Jackson, (2006) The diversity and biogeography of soil bacterial communities. *Proc. Natl. Acad. Sci. U. S. A.* **103**: 626-631.

Fischer, H. M., M. Babst, T. Kaspar, G. Acuña, F. Arigoni & H. Hennecke, (1993) One member of a *gro*ESL-like chaperonin multigene family in *Bradyrhizobium japonicum* is co-regulated with symbiotic nitrogen fixation genes. *The Embo Journal* **12**: 2901-2912.

Fischer, H. M., K. Schneider, M. Babst & H. Hennecke, (1999) GroEL chaperonins are required for the formation of a functional nitrogenase in *Bradyrhizobium japonicum*. *Arch. Microbiol.* **171**: 279-289.

Frank, B., (1889) Über die Pilzsymbiose der Leguminosen. *Berichte der Deutschen Botanischen Gesellschaft* **7**: 332-346.

Fry, J., M. Wood & P. S. Poole, (2001) Investigation of myo-inositol catabolism in *Rhizobium leguminosarum* bv. *viciae* and its effect on nodulation competitiveness. *Mol. Plant. Microbe Interact.* **14**: 1016-1025.

Frydman, J., (2001) Folding of newly translated proteins in vivo: the role of molecular chaperones. *Annu. Rev. Biochem.* **70**: 603-647.

Gao, J. L., S. L. Turner, F. L. Kan, E. T. Wang, Z. Y. Tan, Y. H. Qiu, J. Gu, Z. Terefework, J. P. W. Young, K. Lindström & W. X. Chen, (2004) *Mesorhizobium septentrionale* sp. nov. and *Mesorhizobium temperatum* sp. nov., isolated from *Astragalus adsurgens* growing in the northern regions of China. *Int. J. Syst. Evol. Microbiol.* **54**: 2003-2012.

Gaur, Y. D. & A. N. Sen, (1979) Cross inoculation group specificity in *Cicer-Rhizobium* symbiosis. *New Phytol.* **83**: 745-754.

Genevaux, P., C. Georgopoulos & W. L. Kelley, (2007) The Hsp70 chaperone machines of *Escherichia coli*: a paradigm for the repartition of chaperone functions. *Mol. Microbiol.* **66**: 840-857.

Ghosh, W. & P. Roy, (2006) *Mesorhizobium thiogangeticum* sp. nov., a novel sulfur-oxidizing chemolithoautotroph from rhizosphere soil of an Indian tropical leguminous plant. *Int. J. Syst. Evol. Microbiol.* **56**: 91-97.

Giller, K. E., M. H. Beare, P. Lavelle, A. M. N. Izac & M. J. Swift, (1997) Agricultural intensification, soil biodiversity and agroecosystem function. *Applied Soil Ecology* **6**: 3-16.

Giongo, A., A. Ambrosini, L. K. Vargas, J. R. J. Freire, M. H. Bodanese-Zanettini & L. M. P. Passaglia, (2008) Evaluation of genetic diversity of bradyrhizobia strains nodulating soybean [*Glycine max* (L.) Merrill] isolated from South Brazilian fields. *Applied Soil Ecology* **38**: 261-269.

Giraud, E., L. Moulin, D. Vallenet, V. Barbe, E. Cytryn, J.-C. Avarre, M. Jaubert, D. Simon, F. Cartieaux, Y. Prin, G. Bena, L. Hannibal, J. Fardoux, M. Kojadinovic, L. Vuillet, A. Lajus, S. Cruveiller, Z. Rouy, S. Mangenot, B. Segurens, C. Dossat, W. L. Franck, W.-S. Chang, E. Saunders, D. Bruce, P. Richardson, P. Normand, B. Dreyfus, D. Pignol, G. Stacey, D. Emerich, A. Vermeglio, C. Medigue & M. Sadowsky, (2007) Legumes Symbioses: Absence of Nod Genes in Photosynthetic Bradyrhizobia. *Science* **316**: 1307-1312.

Goldenberg, D., I. Azar & A. B. Oppenheim, (1996) Differential mRNA stability of the *cspA* gene in the cold-shock response of *Escherichia coli*. *Mol. Microbiol.* **19**: 241-248.

Gould, P., M. Maguire & P. A. Lund, (2007) Distinct mechanisms regulate expression of the two major *groEL* homologues in *Rhizobium leguminosarum*. *Arch. Microbiol.* **187**: 1-14.

Govezensky, D., T. Greener, G. Segal & A. Zamir, (1991) Involvement of GroEL in *nif* gene regulation and nitrogenase assembly. *J. Bacteriol.* **173**: 6339-6346.

Graham, P. H., (1992) Stress Tolerance in *Rhizobium* and *Bradyrhizobium*, and Nodulation under Adverse Soil-Conditions. *Can. J. Microbiol.* **38**: 475-484.

Graumann, P. & M. A. Marahiel, (1996) Some like it cold: Response of microorganisms to cold shock. *Arch. Microbiol.* **166**: 293-300.

Guan, S. H., W. F. Chen, E. T. Wang, Y. L. Lu, X. R. Yan, X. X. Zhang & W. X. Chen, (2008) *Mesorhizobium caraganae* sp. nov., a novel rhizobial species nodulated with *Caragana* spp. in China. *Int. J. Syst. Evol. Microbiol.* **58**: 2646-2653.

Guisbert, E., T. Yura, V. A. Rhodius & C. A. Gross, (2008) Convergence of molecular, modeling, and systems approaches for an understanding of the *Escherichia coli* heat shock response. *Microbiol. Mol. Biol. Rev.* **72**: 545-+.

Han, T. X., L. L. Han, L. J. Wu, W. F. Chen, X. H. Sui, J. G. Gu, E. T. Wang & W. X. Chen, (2008) *Mesorhizobium gobiense* sp. nov. and *Mesorhizobium tarimense* sp. nov., isolated from wild legumes growing in desert soils of Xinjiang, China. *Int. J. Syst. Evol. Microbiol.* **58**: 2610-2618.

Herridge, D. F., H. Marcellos, W. L. Felton, G. L. Turner & M. B. Peoples, (1995) Chickpea Increases Soil-N Fertility in Cereal Systems through Nitrate Sparing and N-2 Fixation. *Soil Biology & Biochemistry* **27**: 545-551.

Herridge, D. F., M. B. Peoples & R. M. Boddey, (2008) Global inputs of biological nitrogen fixation in agricultural systems. *Plant Soil* **311**: 1-18.

Hill, S., (1988) How Is Nitrogenase Regulated by Oxygen. *Fems Microbiol. Rev.* **54**: 111-129.

Hungria, M. & M. A. T. Vargas, (2000) Environmental factors affecting N_2 fixation in grain legumes in the tropics, with an emphasis on Brazil. *Field Crops Res.* **65**: 151-164.

Içgen, B., G. Ozcengiz & N. G. Alaeddinoglu, (2002) Evaluation of symbiotic effectiveness of various *Rhizobium cicer* strains. *Res. Microbiol.* **153**: 369-372.

Ivic, A., D. Olden, E. J. Wallington & P. A. Lund, (1997) Deletion of *Escherichia coli groEL* is complemented by a *Rhizobium leguminosarum groEL* homologue at 37°C but not at 43°C. *Gene* **194**: 1-8.

Jarvis, B. D. W., C. E. Pankhurst & J. J. Patel, (1982) *Rhizobium loti*, a New Species of Legume Root Nodule Bacteria. *Int. J. Syst. Bacteriol.* **32**: 378-380.

Jarvis, B. D. W., P. van Berkum, W. X. Chen, S. M. Nour, M. P. Fernandez, J. C. Cleyet-Marel & M. Gillis, (1997) Transfer of *Rhizobium loti*, *Rhizobium huakuii*, *Rhizobium ciceri*, *Rhizobium mediterraneum*, and *Rhizobium tianshanense* to *Mesorhizobium* gen. nov. *Int. J. Syst. Bacteriol.* **47**: 895-898.

Jensen, E. S. & H. Hauggaard-Nielsen, (2003) How can increased use of biological N-2 fixation in agriculture benefit the environment? *Plant Soil* **252**: 177-186.

Jones, P. G., R. A. Vanbogelen & F. C. Neidhardt, (1987) Induction of proteins in response to low temperature in *Escherichia coli*. *J. Bacteriol.* **169**: 2092-2095.

Kamst, E., J. Pilling, L. M. Raamsdonk, B. J. J. Lugtenberg & H. P. Spaink, (1997) *Rhizobium* nodulation protein NodC is an important determinant of chitin oligosaccharide chain length in nod factor biosynthesis. *J. Bacteriol.* **179**: 2103-2108.

Kaneko, T., Y. Nakamura, S. Sato, E. Asamizu, T. Kato, S. Sasamoto, A. Watanabe, K. Idesawa, A. Ishikawa, K. Kawashima, T. Kimura, Y. Kishida, C. Kiyokawa, M. Kohara, M. Matsumoto, A. Matsuno, Y. Mochizuki, S. Nakayama, N. Nakazaki, S. Shimpo, M. Sugimoto, C. Takeuchi, M. Yamada & S. Tabata, (2000) Complete genome structure of the nitrogen-fixing symbiotic bacterium *Mesorhizobium loti*. *DNA Res.* **7**: 331-338.

Karanja, N. K. & M. Wood, (1988) Selecting *Rhizobium phaseoli* strains for use with beans (*Phaseolus-Vulgaris* L.) in Kenya: Tolerance of high temperature and antibiotic resistance. *Plant Soil* **112**: 15-22.

Karlin, S. & L. Brocchieri, (2000) Heat shock protein 60 sequence comparisons: Duplications, lateral transfer, and mitochondrial evolution. *Proc. Natl. Acad. Sci. U. S. A.* **97**: 11348-11353.

Kishinevsky, B. D., D. Sen & R. W. Weaver, (1992) Effect of high root temperature on *Bradyrhizobium*-peanut symbiosis *Plant Soil* **143**: 275-282.

Kneip, C., P. Lockhart, C. Voss & U. G. Maier, (2007) Nitrogen fixation in eukaryotes - New models for symbiosis. *BMC Evol. Biol.* **7**.

Kulkarni, S. & C. S. Nautiyal, (1999) Characterization of high temperature-tolerant rhizobia isolated from *Prosopis juliflora* grown in alkaline soil. *Journal of General Applied Microbiology* **45**: 213-220.

Kuykendall, L. D., Y. D. Gaur & S. K. Dutta, (1993) Genetic Diversity among *Rhizobium* Strains from *Cicer arietinum* L. *Lett. Appl. Microbiol.* **17**: 259-263.

L'Taief, B., B. Sifi, M. Gtari, M. Zaman-Allah & M. Lachaal, (2007) Phenotypic and molecular characterization of chickpea rhizobia isolated from different areas of Tunisia. *Can. J. Microbiol.* **53**: 427-434.

Labidi, M., S. Laberge, L. P. Vezina & H. Antoun, (2000) The *dnaJ* (hsp40) locus in *Rhizobium leguminosarum* bv. *phaseoli* is required for the establishment of an effective symbiosis with *Phaseolus vulgaris*. *Mol. Plant. Microbe Interact.* **13**: 1271-1274.

Laranjo, M., A. Alexandre, R. Rivas, E. Velázquez, J. P. W. Young & S. Oliveira, (2008) Chickpea rhizobia symbiosis genes are highly conserved across multiple *Mesorhizobium* species. *FEMS Microbiol. Ecol.* **66**: 391-400.

Laranjo, M., C. Branco, R. Soares, L. Alho, M. Carvalho & S. Oliveira, (2002) Comparison of chickpea rhizobia isolates from diverse Portuguese natural populations based on symbiotic effectiveness and DNA fingerprint. *J. Appl. Microbiol.* **92**: 1043-1050.

Laranjo, M., J. Machado, J. P. W. Young & S. Oliveira, (2004) High diversity of chickpea *Mesorhizobium* species isolated in a Portuguese agricultural region. *FEMS Microbiol. Ecol.* **48**: 101-107.

Lee, J. H., M. A. Heo, J. H. Seo, J. H. Kim, B. G. Kim & S. G. Lee, (2008a) Improving the growth rate of *Escherichia coli* DH5 alpha at low temperature through engineering of GroEL/S chaperone system. *Biotechnol. Bioeng.* **99**: 515-520.

Lee, K. B., P. De Backer, T. Aono, C. T. Liu, S. Suzuki, T. Suzuki, T. Kaneko, M. Yamada, S. Tabata, D. M. Kupfer, F. Z. Najar, G. B. Wiley, B. Roe, T. T. Binnewies, D. W. Ussery, W. D'Haeze, J. Den Herder, D. Gevers, D. Vereecke, M. Holsters & H. Oyaizu, (2008b) The genome of the versatile nitrogen fixer *Azorhizobium caulinodans* ORS571. *BMC Genomics* **9**.

Leigh, J. A., (2005) Genomics of Diazotrophic Archaea. In: Genomes and Genomics of Nitrogen-fixing Organisms. R. Palacios & W. E. Newton (eds). Springer Netherlands, pp. 7-12.

Lesser, M. P., C. H. Mazel, M. Y. Gorbunov & P. G. Falkowski, (2004) Discovery of symbiotic nitrogen-fixing cyanobacteria in corals. *Science* **305**: 997-1000.

Lira, M. D., A. S. T. Lima, J. R. F. Arruda & D. L. Smith, (2005) Effect of root temperature on nodule development of bean, lentil and pea. *Soil Biology & Biochemistry* **37**: 235-239.

Lu, Y. L., W. F. Chen, E. T. Wang, L. L. Han, X. X. Zhang, W. X. Chen & S. Z. Han, (2009) *Mesorhizoblum shangrilense* sp nov., isolated from root nodules of *Caragana* species. *Int. J. Syst. Evol. Microbiol.* **59**: 3012-3018.

Maâtallah, J., E. B. Berraho, S. Muñoz, J. Sanjuan & C. Lluch, (2002) Phenotypic and molecular characterization of chickpea rhizobia isolated from different areas of Morocco. *J. Appl. Microbiol.* **93**: 531-540.

Martinez-Romero, E., (2003) Diversity of *Rhizobium-Phaseolus vulgaris* symbiosis: overview and perspectives. *Plant Soil* **252**: 11-23.

Mhadhbi, H., M. Jebara, A. Zitoun, F. Limam & M. E. Aouani, (2008) Symbiotic effectiveness and response to mannitol-mediated osmotic stress of various chickpea-rhizobia associations. *World Journal of Microbiology & Biotechnology* **24**: 1027-1035.

Michiels, J., C. Verreth & J. Vanderleyden, (1994) Effects of Temperature Stress on Bean-Nodulating *Rhizobium* Strains. *Appl. Environ. Microbiol.* **60**: 1206-1212.

Minder, A. C., H. M. Fischer, H. Hennecke & F. Narberhaus, (2000) Role of HrcA and CIRCE in the heat shock regulatory network of Bradyrhizobium japonicum. *J. Bacteriol.* **182**: 14-22.

Minder, A. C., F. Narberhaus, M. Babst, H. Hennecke & H. M. Fischer, (1997) The *dnaKJ* operon belongs to the sigma(32)-dependent class of heat shock genes in *Bradyrhizobium japonicum*. *Mol. Gen. Genet.* **254**: 195-206.

Morita, M. T., Y. Tanaka, T. S. Kodama, Y. Kyogoku, H. Yanagi & T. Yura, (1999) Translational induction of heat shock transcription factor sigma(32): evidence for a built-in RNA thermosensor. *Genes Dev.* **13**: 655-665.

Moulin, L., A. Munive, B. Dreyfus & C. Boivin-Masson, (2001) Nodulation of legumes by members of the beta-subclass of *Proteobacteria*. *Nature* **411**: 948-950.

Muga, A. & F. Moro, (2008) Thermal Adaptation of Heat Shock Proteins. *Curr. Protein Peptide Sci.* **9**: 552-566.

Münchbach, M., A. Nocker & F. Narberhaus, (1999) Multiple small heat shock proteins in rhizobia. *J. Bacteriol.* **181**: 83-90.

Munevar, F. & A. G. Wollum, (1981) Growth of *Rhizobium japonicum* strains at temperatures above 27°C. *Appl. Environ. Microbiol.* **42**: 272-276.

Nandal, K., A. R. Sehrawat, A. S. Yadav, R. K. Vashishat & K. S. Boora, (2005) High temperature-induced changes in exopolysaccharides, lipopolysaccharides and protein profile of heat-resistant mutants of *Rhizobium* sp (Cajanus). *Microbiol. Res.* **160**: 367-373.

Nandasena, K. G., G. W. O'Hara, R. P. Tiwari, A. Willems & J. G. Howieson, (2009) *Mesorhizobium australicum* sp nov and *Mesorhizobium opportunistum* sp nov., isolated from *Biserrula pelecinus* L. in Australia. *Int. J. Syst. Evol. Microbiol.* **59**: 2140-2147.

Nandasena, K. G., G. W. O'Hara, R. P. Tiwari, A. Willlems & J. G. Howieson, (2007) *Mesorhizobium ciceri* biovar *biserrulae*, a novel biovar nodulating the pasture legume *Biserrula pelecinus* L. *Int. J. Syst. Evol. Microbiol.* **57**: 1041-1045.

Nandasena, K. G., G. W. O'Hara, R. P. Tiwari, R. J. Yates, B. D. Kishinevsky & J. G. Howieson, (2004) Symbiotic relationships and root nodule ultrastructure of the pasture legume *Biserrula pelecinus* L.-a new legume in agriculture. *Soil Biol. Biochem.* **36**: 1309-1317.

Narberhaus, F., R. Kaser, A. Nocker & H. Hennecke, (1998) A novel DNA element that controls bacterial heat shock gene expression. *Mol. Microbiol.* **28**: 315-323.

Narberhaus, F., P. Krummenacher, H. M. Fischer & H. Hennecke, (1997) Three disparately regulated genes for sigma(32)-like transcription factors in *Bradyrhizobium japonicum*. *Mol. Microbiol.* **24**: 93-104.

Nocker, A., N. P. Krstulovic, X. Perret & F. Narberhaus, (2001) ROSE elements occur in disparate rhizobia and are functionally interchangeable between species. *Arch. Microbiol.* **176**: 44-51.

Normand, P., C. Queiroux, L. S. Tisa, D. R. Benson, Z. Rouy, S. Cruveiller & C. Medigue, (2007) Exploring the genomes of Frankia. *Physiol. Plant.* **130**: 331-343.

Nour, S. M., J.-C. Cleyet-Marel, P. Normand & M. P. Fernandez, (1995) Genomic heterogeneity of strains nodulating chickpeas (*Cicer arietinum* L.) and description of *Rhizobium mediterraneum* sp. nov. *Int. J. Syst. Bacteriol.* **45**: 640-648.

Nour, S. M., J. C. Cleyet-Marel, D. Beck, A. Effosse & M. P. Fernandez, (1994a) Genotypic and phenotypic diversity of *Rhizobium* isolated from chickpea (*Cicer arietinum* L.). *Can. J. Microbiol.* **40**: 345-354.

Nour, S. M., M. P. Fernandez, P. Normand & J.-C. Cleyet-Marel, (1994b) *Rhizobium ciceri* sp. nov., consisting of strains that nodulate chickpeas (*Cicer arietinum* L.). *Int. J. Syst. Bacteriol.* **44**: 511-522.

Ogawa, J. & S. R. Long, (1995) The *Rhizobium meliloti groELc* locus is required for regulation of early nod genes by the transcription activator *nodD*. *Genes Dev.* **9**: 714-729.

Ogutcu, H., O. F. Algur, E. Elkoca & F. Kantar, (2008) The determination of symbiotic effectiveness of *Rhizobium* strains isolated from wild chickpeas collected from high altitudes in erzurum. *Turkish Journal of Agriculture and Forestry* **32**: 241-248.

Oka-Kira, E. & M. Kawaguchi, (2006) Long-distance signaling to control root nodule number. *Curr. Opin. Plant Biol.* **9**: 496-502.

Ono, Y., H. Mitsui, T. Sato & K. Minamisawa, (2001) Two RpoH homologs responsible for the expression of heat shock protein genes in *Sinorhizobium meliloti*. *Mol. Gen. Genet.* **264**: 902-912.

Patschkowski, T., A. Schluter & U. B. Priefer, (1996) *Rhizobium leguminosarum* bv *viciae* contains a second *fnr/fixK*-like gene and an unusual *fixL* homologue. *Mol. Microbiol.* **21**: 267-280.

Perret, X., C. Staehelin & W. J. Broughton, (2000) Molecular basis of symbiotic promiscuity. *Microbiol. Mol. Biol. Rev.* **64**: 180-201.

Polcyn, W. & R. Lucinski, (2003) Aerobic and anaerobic nitrate and nitrite reduction in free-living cells of *Bradyrhizobium* sp (Lupinus). *FEMS Microbiol. Lett.* **226**: 331-337.

Raina, S., D. Missiakas & C. Georgopoulos, (1995) The *rpoE* gene encoding the sigma(E) (sigma(24)) heat-shock sigma-factor of *Escherichia coli*. *EMBO J.* **14**: 1043-1055.

Rivas, R., M. Laranjo, P. F. Mateos, S. Oliveira, E. Martinez-Molina & E. Velázquez, (2007) Strains of *Mesorhizobium amorphae* and *Mesorhizobium tianshanense*, carrying symbiotic genes of common chickpea endosymbiotic species, constitute a novel biovar (*ciceri*) capable of nodulating *Cicer arietinum*. *Lett. Appl. Microbiol.* **44**: 412-418.

Rivas, R., A. Peix, P. F. Mateos, M. E. Trujillo, E. Martínez-Molina & E. Velázquez, (2006) Biodiversity of populations of phosphate solubilizing rhizobia that nodulates chickpea in different Spanish soils. *Plant Soil* **287**: 23-33.

Rodrigues, C., M. Laranjo & S. Oliveira, (2006) Effect of heat and pH stress in the growth of chickpea mesorhizobia. *Curr. Microbiol.* **53**: 1-7.

Rodriguez-Quinones, F., M. Maguire, E. J. Wallington, P. S. Gould, V. Yerko, J. A. Downie & P. A. Lund, (2005) Two of the three *groEL* homologues in *Rhizobium leguminosarum* are dispensable for normal growth. *Arch. Microbiol.* **183**: 253-265.

Romdhane, S., M. E. Aouani & R. Mhamdi, (2007) Inefficient nodulation of chickpea (*Cicer arietinum* L.) in the arid and Saharan climates in Tunisia by *Sinorhizobium meliloti* biovar *medicaginis*. *Annals of Microbiology* **57**: 15-19.

Rudiger, S., L. Germeroth, J. SchneiderMergener & B. Bukau, (1997) Substrate specificity of the DnaK chaperone determined by screening cellulose-bound peptide libraries. *EMBO J.* **16**: 1501-1507.

Rudresh, D. L., M. K. Shivaprakash & R. D. Prasad, (2005) Effect of combined application of *Rhizobium*, phosphate solubilizing bacterium and *Trichoderma* spp. on growth, nutrient uptake and yield of chickpea (*Cicer aritenium* L.). *Applied Soil Ecology* **28**: 139-146.

Sadowsky, M., (2005) Soil Stress Factors Influencing Symbiotic Nitrogen Fixation. In: Nitrogen Fixation in Agriculture, Forestry, Ecology, and the Environment. D. Werner & W. E. Newton (eds). Springer Netherlands, pp. 89-112.

Saxena, M. C. & K. B. Singh, (1987) *The Chickpea*. CAB International, Wallingford, UK.

Schmeisser, C., H. Liesegang, D. Krysciak, N. Bakkou, A. Le Quere, A. Wollherr, I. Heinemeyer, B. Morgenstern, A. Pommerening-Roser, M. Flores, R. Palacios, S. Brenner, G. Gottschalk, R. A. Schmitz, W. J. Broughton, X. Perret, A. W. Strittmatter & W. R. Streit, (2009) *Rhizobium* sp. strain NGR234 possesses a remarkable number of secretion systems. *Appl. Environ. Microbiol.* **75**: 4035-4045.

Siddiqui, Z. A. & M. S. Akhtar, (2009) Effect of plant growth promoting rhizobacteria, nematode parasitic fungi and root-nodule bacterium on root-knot nematodes *Meloidogyne javanica* and growth of chickpea. *Biocontrol Sci. Technol.* **19**: 511-521.

Siegenthaler, R. K. & P. Christen, (2006) Tuning of DnaK chaperone action by nonnative protein sensor DnaJ and thermosensor GrpE. *J. Biol. Chem.* **281**: 34448-34456.

So, R. B., J. K. Ladha & J. P. Young, (1994) Photosynthetic symbionts of *Aeschynomene* spp. form a cluster with bradyrhizobia on the basis of fatty acid and rRNA analyses. *Int. J. Syst. Bacteriol.* **44**: 392-403.

Staddon, W. J., J. T. Trevors, L. C. Duchesne & C. A. Colombo, (1998) Soil microbial diversity and community structure across a climatic gradient in western Canada. *Biodivers. Conserv.* **7**: 1081-1092.

Straus, D. B., W. A. Walter & C. A. Gross, (1987) The heat-shock response of *Escherichia coli* regulated by changes in the concentration of sigma-32. *Nature* **329**: 348-351.

Sullivan, J. T., J. R. Trzebiatowski, R. W. Cruickshank, J. Gouzy, S. D. Brown, R. M. Elliot, D. J. Fleetwood, N. G. McCallum, U. Rossbach, G. S. Stuart, J. E. Weaver, R. J. Webby, F. J. de Bruijn & C. W. Ronson, (2002) Comparative sequence analysis of the symbiosis island of *Mesorhizobium loti* strain R7A. *J. Bacteriol.* **184**: 3086-3095.

Sy, A., E. Giraud, P. Jourand, N. Garcia, A. Willems, P. de Lajudie, Y. Prin, M. Neyra, M. Gillis, C. Boivin-Masson & B. Dreyfus, (2001) Methylotrophic *Methylobacterium* bacteria nodulate and fix nitrogen in symbiosis with legumes. *J. Bacteriol.* **183**: 214-220.

Szabo, A., T. Langer, H. Schroder, J. Flanagan, B. Bukau & F. U. Hartl, (1994) The ATP hydrolysis-dependent reaction cycle of the *Escherichia coli* HSP70 system - DnaK, DnaJ, and GrpE. *Proc. Natl. Acad. Sci. U. S. A.* **91**: 10345-10349.

Tindall, B. J., (2008) The genus name *Sinorhizobium* Chen et al. 1988 is a later synonym of *Ensifer* Casida 1982 and is not conserved over the latter genus name, and the species name '*Sinorhizobium adhaerens*' is not validly published. Opinion 84 - Judicial Commission of the International Committee on Systematics of Prokaryotes. *Int. J. Syst. Evol. Microbiol.* **58**: 1973-1973.

Tomoyasu, T., J. Gamer, B. Bukau, M. Kanemori, H. Mori, A. J. Rutman, A. B. Oppenheim, T. Yura, K. Yamanaka, H. Niki, S. Hiraga & T. Ogura, (1995) *Escherichia coli* FtsH is a membrane-bound, ATP-dependent protease which degrades the heat-shock transcription factor sigma(32). *EMBO J.* **14**: 2551-2560.

Tomoyasu, T., T. Ogura, T. Tatsuta & B. Bukau, (1998) Levels of DnaK and DnaJ provide tight control of heat shock gene expression and protein repair in *Escherichia coli*. *Mol. Microbiol.* **30**: 567-581.

Trinick, M. J., (1973) Symbiosis between *Rhizobium* and non-legume, *Trema aspera*. *Nature* **244**: 459-460.

Truchet, G., D. G. Barker, S. Camut, F. Debilly, J. Vasse & T. Huguet, (1989) Alfalfa Nodulation in the Absence of Rhizobium. *Mol. Gen. Genet.* **219**: 65-68.

Valverde, A., E. Velázquez, C. Gutiérrez, E. Cervantes, A. Ventosa & J. M. Igual, (2003) *Herbaspirillum lusitanum* sp. nov., a novel nitrogen-fixing bacterium associated with root nodules of *Phaseolus vulgaris*. *Int. J. Syst. Evol. Microbiol.* **53**: 1979-1983.

Velázquez, E., J. M. Igual, A. Willems, M. P. Fernandez, E. Muñoz, P. F. Mateos, A. Abril, N. Toro, P. Normand, E. Cervantes, M. Gillis & E. Martínez-Molina, (2001) *Mesorhizobium chacoense* sp. nov., a novel species that nodulates *Prosopis alba* in the Chaco Arido region (Argentina). *Int. J. Syst. Evol. Microbiol.* **51**: 1011-1021.

Vidal, C., C. Chantreuil, O. Berge, L. Maure, J. Escarre, G. Bena, B. Brunel & J. C. Cleyet-Marel, (2009) *Mesorhizobium metallidurans* sp nov., a metal-resistant symbiont of *Anthyllis vulneraria* growing on metallicolous soil in Languedoc, France. *Int. J. Syst. Evol. Microbiol.* **59**: 850-855.

Wall, D., M. Zylicz & C. Georgopoulos, (1994) The NH2-Terminal 108 Amino-Acids of the *Escherichia coli* DnaJ Protein Stimulate the ATPase Activity of DnaK and Are Sufficient for Lambda-Replication. *J. Biol. Chem.* **269**: 5446-5451.

Wang, E. T., P. van Berkum, X. H. Sui, D. Beyene, W. X. Chen & E. Martínez-Romero, (1999) Diversity of rhizobia associated with *Amorpha fruticosa* isolated from Chinese soils and description of *Mesorhizobium amorphae* sp. nov. *Int. J. Syst. Bacteriol.* **49**: 51-65.

Wang, F. Q., E. T. Wang, J. Liu, Q. Chen, X. H. Sui, W. F. Chen & W. X. Chen, (2007) *Mesorhizobium albiziae* sp. nov., a novel bacterium that nodulates *Albizia kalkora* in a subtropical region of China. *Int. J. Syst. Evol. Microbiol.* **57**: 1192-1199.

Wild, J., E. Altman, T. Yura & C. A. Gross, (1992) DnaK and DnaJ heat-shock proteins participate in protein export in *Escherichia coli*. *Genes Dev.* **6**: 1165-1172.

Young, J. M., L. D. Kuykendall, E. Martínez-Romero, A. Kerr & H. Sawada, (2001) A revision of *Rhizobium* Frank 1889, with an emended description of the genus, and the inclusion of all species of *Agrobacterium* Conn 1942 and *Allorhizobium undicola* de Lajudie *et al.* 1998 as new combinations: *Rhizobium radiobacter*, *R. rhizogenes*, *R. rubi*, *R. undicola* and *R. vitis*. *Int. J. Syst. Evol. Microbiol.* **51**: 89-103.

Young, J. P., L. Crossman, A. Johnston, N. Thomson, Z. Ghazoui, K. Hull, M. Wexler, A. Curson, J. Todd, P. Poole, T. Mauchline, A. East, M. Quail, C. Churcher, C. Arrowsmith, I. Cherevach, T. Chillingworth, K. Clarke, A. Cronin, P. Davis, A. Fraser, Z. Hance, H. Hauser, K. Jagels, S. Moule, K. Mungall, H. Norbertczak, E. Rabbinowitsch, M. Sanders, M. Simmonds, S. Whitehead & J. Parkhill, (2006) The genome of *Rhizobium leguminosarum* has recognizable core and accessory components. *Genome Biology* **7**: R34.

Zahran, H. H., (1999) *Rhizobium*-legume symbiosis and nitrogen fixation under severe conditions and in an arid climate. *Microbiol. Mol. Biol. Rev.* **63**: 968-989.

Zahran, H. H., (2001) Rhizobia from wild legumes: diversity, taxonomy, ecology, nitrogen fixation and biotechnology. *J. Biotechnol.* **91**: 143-153.

Zhang, F., D. H. Lynch & D. L. Smith, (1995) Impact of low root temperatures in soybean [*Glycine max* (L) Merr] on nodulation and nitrogen fixation. *Environ. Exp. Bot.* **35**: 279-285.

2.

Isolation and characterization of native chickpea rhizobia

This chapter is based on the manuscript:

Alexandre, A., Brígido, C., Laranjo, M., Rodrigues, S. and Oliveira, S. (2009) "Survey of chickpea rhizobia diversity in Portugal reveals the predominance of species distinct from *Mesorhizobium ciceri* and *Mesorhizobium mediterraneum*", *Microbial Ecology*, **58**, 930-941.

2.1 – Summary

Several *Mesorhizobium* species are able to induce effective nodules in chickpea, one of the most important legumes worldwide. Our aims were to examine the diversity and biogeography of chickpea rhizobia, to search for a predominant species and to identify the most efficient microsymbiont, considering Portugal as a case study. One hundred and ten isolates were obtained from continental Portugal and Madeira Island. The 16S rRNA gene phylogeny revealed that isolates are highly diverse, grouping with most *Mesorhizobium* type strains, in four main clusters (A-D). Interestingly, only 33% of the isolates grouped with *M. ciceri* (cluster B) or *M. mediterraneum* (cluster D), the formerly described specific chickpea microsymbionts. Most isolates belong to cluster A, showing higher sequence similarity with *M. huakuii* and *M. amorphae*. The association found between province of origin and species cluster of the isolates suggests biogeography patterns: most isolates from the North, Centre and South belong to clusters B, A and D, respectively. Most of the highly efficient isolates (symbiotic effectiveness >75%) belong to cluster B. A correlation was found between species cluster and origin soil pH of the isolates, suggesting that pH is a key environmental factor, which influences the species geographic distribution. To our knowledge, this is one of the few surveys on chickpea rhizobia and the first systematic assessment of indigenous rhizobia in Portugal.

2.2 – Introduction

Compared to other grain legumes with agricultural importance, few studies addressed the genetic diversity of native chickpea rhizobia (Aouani *et al.*, 2001, Kuykendall *et al.*, 1993, Maâtallah *et al.*, 2002b, Nour *et al.*, 1994a). As far as Portugal is concerned, only a small area in the South of Portugal has been studied (Rodrigues *et al.*, 2006, Laranjo *et al.*, 2001, Alexandre *et al.*, 2006, Laranjo *et al.*, 2002, Laranjo *et al.*, 2004).

Two species were initially identified as specific chickpea microsymbionts: *Mesorhizobium ciceri* (Nour *et al.*, 1994b) and *M. mediterraneum* (Nour *et al.*, 1995). Chickpea has been considered a narrow-host range legume (Broughton & Perret, 1999), mainly because it cannot be nodulated by broad-host range rhizobia, such as *Ensifer* sp. NGR 234 (Perret *et al.*, 2000). Nevertheless, the promiscuity of a given legume is related to the number of Nod factors it can interact with, rather than the diversity of rhizobia, which are able to nodulate such legume (Downie, 1998). Recent studies have been changing the initial suggestion that only two species of *Mesorhizobium* were able to nodulate chickpea, showing that this legume is able to establish symbioses with several other *Mesorhizobium* species, namely *M. amorphae*, *M. loti* and *M. tianshanense* (Laranjo *et al.*, 2004, Rivas *et al.*, 2007). Nevertheless, chickpea is still considered a narrow-host range legume, given that the symbiosis genes (*nodC* and *nifH*) found in isolates belonging to other *Mesorhizobium* species are identical to those carried by *M. ciceri* and *M. mediterraneum*, which probably lead to the production of very similar Nod factors (Rivas *et al.*, 2007, Laranjo *et al.*, 2008). Similar results have been reported for *Mesorhizobium* species that nodulate *Anagyris latifolia* (Donate-Correa *et al.*, 2007) and *Bradyrhizobium* species nodulating *Lupinus* spp. and *Ornithopus* spp. (Jarabo-Lorenzo *et al.*, 2003).

Bacterial phylogeny has relied on the sequence analysis of single core genes, mainly the 16S rRNA gene (Menna *et al.*, 2006) but also on other housekeeping genes, such as *atpD* and *recA* (Young & Park, 2007). However, several studies have shown that single-gene trees may not adequately reflect phylogenetic relationships, so multilocus approaches have been proposed for species identification (Martens *et al.*, 2007, Coenye *et al.*, 2005, Clayton *et al.*, 1995, van Berkum *et al.*, 2006). Still, in the last decade, analysis of the 16S rRNA gene has been, by far, the most widely used approach to define molecular phylogeny and taxonomy of bacteria (Sun *et al.*, 2008, Gevers & Coenye, 2007). The 16S rRNA gene is the only sequence available for most bacterial species, including type strains. Thus, the 16S rRNA gene is a useful tool for placing any new isolate among its closer taxonomic relatives.

Within the *Mesorhizobium* genus the 16S rRNA sequence similarity between species is very high. In the past few years several new species have been added to this genus and still the two most divergent species share over 95% of 16S rRNA gene sequence similarity (*M. albiziae* and *M. thiogangeticum*). Moreover, some *Mesorhizobium* species are 100% identical in terms of the comparable 16S rRNA gene sequence, as the recently described *M. metallidurans* and *M. gobiense*. Nevertheless, the 16S rRNA sequence can be useful for mesorhizobia isolates identification, as long as its limitations are taken into account.

The aims of this study were to examine the biogeography of rhizobia able to nodulate chickpea, to investigate the presence of a predominant chickpea rhizobia species and to identify the most efficient species in the symbiosis, considering Portugal as a case study. A survey on chickpea rhizobia was carried out in continental and insular Portugal. Genetic diversity of native isolates was examined through molecular phylogeny based on 16S rRNA gene sequences and by plasmid profiles analysis. Symbiotic effectiveness of chickpea native isolates was estimated.

2.3 - Methods

Isolates collection

Rhizospheric soil samples were collected from more than 40 sites of the 11 provinces of continental Portugal (Minho, Trás-os-Montes e Alto Douro, Douro Litoral, Beira Litoral, Beira Alta, Beira Baixa, Estremadura, Ribatejo, Alto Alentejo, Baixo Alentejo, Algarve) and from the Madeira and Azores Islands, covering the entire country (total area of 92000km^2). Soil samples were collected from fields not used for chickpea cultivation, with the exception of the Elvas-ENMP site, which is an experimental agricultural field.

Chickpea seeds (winter cultivar Chk 3226) were surface-sterilized with calcium hypochlorite 14%, washed with sterile distilled water and pre-germinated in water-agar for 48 hours. Seeds were sown in sterilised pots containing the soil samples. Plants were maintained in the plant growth chamber under controlled conditions for eight weeks. Nodules were harvested and disinfected, and isolates were obtained as described by Somasegaran and Hoben (Somasegaran & Hoben, 1994). Isolates were re-inoculated, under sterile and controlled conditions, in order to confirm their ability to nodulate chickpea.

Amplification of the 16S rRNA gene

The 16S rRNA gene was amplified for each isolate using primers Y1 (Young *et al.*, 1991) and Y3 (Laranjo *et al.*, 2004), corresponding to positions 20 to 1507 in *Escherichia coli*. Amplification reaction was carried out as previously reported (Laranjo *et al.*, 2004). PCR products were purified using GFX™ PCR DNA and Gel Band Purification kit (GE Healthcare) or ExoSAP-IT®T (usb) following the manufacturer's instructions. Two additional internal primers, namely IntF and IntR (Laranjo *et al.*, 2004), were used to obtain double-stranded and nearly complete sequences.

Phylogeny based on the 16S rRNA gene sequence analysis

Nucleotide sequences were analysed and edited using BioEdit Sequence Alignment Editor (version 7.0.4.1) (Hall, 1999). Alignments were generated using Clustal W (Thompson *et al.*, 1994). The 16S rRNA gene sequences of the isolates were compared to those of the type strains of the following species: *M. albiziae* (DQ100066), *M. amorphae* (AF041442); *M. chacoense* (AJ278249); *M. ciceri* (DQ444456); *M. huakuii* (FJ491264); *M. loti* (X67229); *M. mediterraneum* (AM181745); *M. plurifarium* (Y14158); *M. septentrionale* (AF508207); *M. temperatum* (AF508208); *M. thiogangeticum* (AJ864462), *M. tianshanense* (AF041447); *Rhizobium etli* (U28916); *R. leguminosarum* bv. *viciae* (U29386); *Sinorhizobium medicae* (L39882) and *S. meliloti* (X67222). *Azorhizobium caulinodans* (X67221) and *Bradyrhizobium japonicum* (U69638) were included as outgroups.

MEGA4 (Molecular Evolutionary Genetics Analysis, version 3.1) software (Tamura *et al.*, 2007) was used to infer the molecular phylogeny by the Neighbour-Joining method based on a distance matrix with the distance correction calculated by Kimura's two-parameter model, with 1000 resamplings in the bootstrap analysis.

Plasmid Profiles

Plasmid profiles were analysed by horizontal agarose gel using a two-comb system that allows the in-well lysis method, as described previously (Laranjo *et al.*, 2001).

Symbiotic Effectiveness (SE)

Plant growth chamber trials were performed under controlled conditions in order to evaluate the symbiotic effectiveness of the isolates (Laranjo *et al.*, 2002). Pre-

germinated chickpea seeds, obtained as described above, were sown in sterilized vermiculite and inoculated with a bacterial suspension of each isolate grown in YEM (Vincent, 1970). A nitrogen-free nutrient solution (Broughton & Dilworth, 1971) was used to irrigate the pots three times per week. Uninoculated plants were used as negative control and uninoculated plants supplemented with nitrogen (140 ppm nitrogen as KNO_3, in the nutrient solution) were used as positive control. Three replicates were used for each treatment. Plants were collected after eight weeks and several parameters were measured, such as shoot dry weight, root dry weight, number of nodules and nodules dry weight. Symbiotic effectiveness was estimated according to Gibson (1987), as the ratio between the positive control and each treatment, using shoot dry weight values. The value of the negative control was subtracted from both treatment and positive control values.

Statistical analysis

Statistical analysis was performed using SPSS 15.0 software (SPSS Inc., Chicago, IL, USA). Relationships between categorical variables were determined using the chi-square test of association. Relationships between a continuous variable and an unordered categorical variable were tested using analysis of variance (one-way ANOVA). Results are presented as the test statistic (χ^2), degrees of freedom (df), and probability of equal or greater deviation (P). For samples not satisfying Cochran's criteria (some categories were represented by only one isolate, and more than 20% of the categories were represented by less than five isolates), the exact value of *P*, the critical probability, was computed rather than the asymptotic *P* value, which is an approximation reserved for large samples (Louvrier *et al.*, 1996). Correspondence analysis (CA) was used as an explorative method to study associations and to reveal interdependencies between two variables (Benzécri, 1973). Visualization using CA is based on representing χ^2 distances among variables.

2.4 – Results

Isolation and characterization of native chickpea rhizobia from the following origin sites were performed by C. Brígido, M. Laranjo or S. Rodrigues: Bragança, Lamego, Porto, Guarda, Castelo Branco, Telhado, Setúbal, Sintra, Santarém, Elvas, Évora, Beja, Portimão and Serra d'Água.

Isolates collection

Most of the sampled sites harboured rhizobia able to nodulate chickpea. Nevertheless, no nodules were obtained with the several soil samples collected from Minho province and Azores Islands. A total of 110 chickpea rhizobia isolates, from 23 sites in 10 provinces of continental Portugal and Madeira Island, were used for further studies. Soil characteristics of each site are shown in Table 2.1.

Table 2.1 - Characteristics of soils used to obtain chickpea rhizobia isolates.

Soils	Texture	pH (H_2O)	Phosphorus P_2O_5 (ppm)	Nitrogen NO_3 (ppm)	Potassium K_2O (ppm)
Bragança	medium	6,69	444	681	710
Lamego	medium	6,58	563	53	580
Porto II	medium	6,37	40	35	340
Aveiro	medium	7,08	44	9	66
Aveiro II	medium	6,07	200	68	740
Coimbra	medium	5,66	108	66	224
Leiria	medium	8,19	182	57	356
Guarda	coarse	5,10	476	74	1060
Viseu	medium	5,94	12	11	126
Castelo Branco	medium	6,46	248	132	216
Telhado	medium	7,32	210	86	178
Caldas da Rainha	medium	6,83	82	131	170
Setúbal	medium	8,06	631	18	350
Sintra	medium	7,83	808	352	910
Santarém	medium	7,83	140	101	530
Elvas-CV	medium	6,20	48	64	143
Elvas-ENMP	medium	7,90	529	12	137
Évora	coarse	5,10	56	47	55
Portalegre	medium	5,25	26	87	152
Beja	medium	8,20	108	28	197
Portimão	medium	8,66	744	194	1600
Portimão I	medium	7,20	318	153	180
Serra d'Água	medium	7,63	74	42	560

Soil samples analyses were performed by the Laboratório Químico Agrícola-University of Évora, Portugal.

Phylogeny based on the 16S rRNA gene analysis

GenBank accession numbers for the 16S rRNA gene sequences of all isolates are shown in Table 2.2.

Table 2.2 – Rhizobia isolates used in the present study.

Origin	Isolate	16S rRNA gene accession number	Species cluster	Plasmid Number	SE (%)
Trás-os Montes e Alto Douro	BR-8-Bragança	EU652123	B	2	45
	BR-9-Bragança	EU652124	B	1	43
	BR-15-Bragança	EU652125	B	2	21
	BR-16-Bragança	EU652126	B	2	35
	BR-28-Bragança	EU652127	B	0	48
	LM-1-Lamego	EU652128	A	1	14
	LM-9-Lamego	EU652129	A	1	55
	LM-13-Lamego	EU652130	A	1	11
	LM-18-Lamego	EU652131	B	1	61
	LM-21-Lamego	EU652132	A	1	22
Douro Litoral	PII-1-Porto	EU652133	B	3	58
	PII-2-Porto	EU652134	B	2	71
	PII-3-Porto	EU652106*	B	2	47
	PII-4-Porto	EU652135	B	2	31
Beira Litoral	A-3-Aveiro	EU652136	A	0	36
	A-8b.-Aveiro	EU652107*	A	0	0
	AII-5-Aveiro	EU652137	A	2	26
	AII-7-Aveiro	EU652138	A	2	32
	C-1-Coimbra	EF504313*	A	1 (d)	47
	C-3-Coimbra	EU652108*	A	1	15
	C-7-Coimbra	EU652139	A	1	14
	C-9-Coimbra	EU652140	A	2	20
	C-13-Coimbra	EU652109*	A	1	49
	C-14-Coimbra	EU652110*	A	2	32
	C-15-Coimbra	EU652141	A	1	20
	C-23-Coimbra	EU652142	A	1	23
	C-24-Coimbra	EU652143	A	1	39
	C-25-Coimbra	EU652144	A	2	21
	C-27b.-Coimbra	EF504314*	A	1	62
	L-19-Leiria	EU652111*	A	0	48
Beira Alta	G-1-Guarda	EU652145	B	0	34
	G-4-Guarda	EU652146	B	0	41
	G-10-Guarda	EU652147	B	0	48
	G-24-Guarda	EU652148	B	0	58
	G-55-Guarda	EU652149	B	0	88
	V-5b.-Viseu	EU652112*	A	1	65
	V-15b.-Viseu	EF504315*	A	0	23
	V-18-Viseu	EF504316*	A	0	67
	V-20-Viseu	EF504317*	A	1	67
	V-25b.-Viseu	EU652113*	A	1	69
Beira Baixa	CB-10-Castelo Branco	EU652150	B	0	56
	CB-19-Castelo Branco	EU652151	B	0	30
	CB-23-Castelo Branco	EU652152	B	4	52
	CB-30-Castelo Branco	EU652153	B	4	45
	CB-38-Castelo Branco	EU652154	B	4	61
	CB-75-Castelo Branco	EU652155	B	0	38
	T-1-Telhado	EU652156	A	1	100
	T-3-Telhado	EU652157	A	1	32
	T-4-Telhado	EU652158	A	1	86
	T-5-Telhado	EU652159	A	0	56
	T-7-Telhado	EU652160	A	1	54
	T-8-Telhado	EU652114*	A	1	31
Estremadura	CR-3-Caldas da Rainha	EU652161	C	0	77
	CR-16-Caldas da Rainha	EU652162	C	0	79
	CR-18-Caldas da Rainha	EU652163	C	0	41
	CR-29-Caldas da Rainha	EU652164	C	0	55
	CR-32-Caldas da Rainha	EU652115*	C	0 (d)	57
	ST-2-Setúbal	AY225401*	C	0	4 (c)
	ST-5-Setúbal	EU652165	C	0	21
	ST-8-Setúbal	EU652166	C	0	7
	ST-20-Setúbal	EU652167	C	0	43
	ST-33-Setúbal	EU652168	C	0	44
	S-1-Sintra	EU652169	D	3	53
	S-8-Sintra	EU652116*	B	1	83
	S-15-Sintra	EU652170	B	0	79
	S-26-Sintra	EU652171	B	0	68

Ribatejo	STR-2-Santarém	EU652117*	A	1	40
	STR-4-Santarém	EU652172	A	1	50
	STR-10-Santarém	EU652173	A	1	28
	STR-14-Santarém	EU652118*	C	1	64
	STR-16-Santarém	EU652174	C	2	49
Alto Alentejo	75-Elvas	AY225386*	B	1 (a)	35 (c)
	78-Elvas	AY225387*	A	1 (a)	63 (c)
	79-Elvas	DQ787130	B	1 (a)	47 (a)
	83-Elvas	DQ787131	A	1 (a)	49 (a)
	85-Elvas	AY225388*	A	1 (a)	60 (c)
	CV-1-Elvas	DQ787132	A	0 (b)	28 (b)
	CV-11-Elvas	DQ787133	A	0 (b)	21 (b)
	CV-16-Elvas	AY225389*	B	1 (b)	42 (c)
	CV-18-Elvas	AY225390*	A	1 (b)	72 (c)
	EE-2-ENMP	AY225396*	D	2 (b)	36 (c)
	EE-7-ENMP	AY225397*	B	0 (b)	84 (c)
	EE-12-ENMP	AY225398*	B	1	10 (c)
	EE-14-ENMP	AY225399*	D	4 (b)	32 (c)
	EE-29-ENMP	AY225400*	D	4	21 (c)
	87-Évora	DQ787134	A	1	n.d.
	89a.-Évora	DQ787135	A	4	n.d.
	90-Évora	AY225391*	A	1	49 (c)
	92-Évora	DQ787136	A	2	42 (a)
	93-Évora	AY225392*	C	0	27 (c)
	94-Évora	AY225393*	A	1	33 (c)
	96-Évora	DQ787137	A	2	n.d.
	98-Évora	AY225394*	A	2	72 (c)
	101-Évora	DQ787138	A	2	n.d.
	102-Évora	AY225395*	A	0 (a)	54 (c)
	PT-35-Portalegre	EU652119*	A	1 (d)	56
Baixo Alentejo	6b.-Beja	AY225381*	D	2 (a)	76 (c)
	7a.-Beja	AY225382*	B	2	39 (c)
	27-Beja	AY225383*	B	1 (a)	41 (c)
	29-Beja	AY225384*	D	6 (d)	71 (c)
	64b.-Beja	AY225385*	A	1 (a)	70 (c)
Algarve	PM-1-Portimão	EU652175	D	2	51
	PM-14-Portimão	EU652176	D	2	33
	PM-17-Portimão	EU652120*	D	2	84
	PMI-1-Portimão	EU652177	B	1	80
	PMI-6-Portimão	EU652121*	A	1	81
Madeira	SA-4-Serra d'Água	EU652122*	A	0	63
	SA-9-Serra d'Água	EU652178	A	0	36
	SA-12-Serra d'Água	EU652179	A	0	56
	SA-17-Serra d'Água	EU652180	A	3	16

Province of origin, 16S rRNA gene accession number, species cluster defined from the 16S rRNA gene sequence analysis, plasmid number and symbiotic effectiveness (SE) values are indicated for each isolate.

(a) Laranjo *et al.* (2001); (b) Laranjo *et al.* (2002); (c) Laranjo *et al.* (2008); (d) Brígido *et al.* (2007)

n.d. – not determined

* complete 16S rRNA gene sequence

Analysis of molecular diversity was performed using full length 16S rRNA gene sequence of a set of 43 rhizobia isolates. A dendrogram was generated by the neighbour-joining method (NJ) from a 1357 bp long alignment (257 variable sites). According to the 16S rRNA gene molecular phylogeny (Fig. 2.1), all native isolates were assigned to the genus *Mesorhizobium*. Isolates form a large cluster together with the *Mesorhizobium* type strains, which received 99% bootstrap support. Four main

clusters (A to D) can be identified, each cluster including isolates from at least three different provinces. The largest cluster (A) comprises the type strains of *M. amorphae*, *M. huakuii*, *M. plurifarium* and *M. septentrionale*, as well as 25 isolates from seven provinces (Beira Litoral, Beira Alta, Beira Baixa, Ribatejo, Alto Alentejo, Baixo Alentejo and Algarve) and Madeira Island, and received 76% bootstrap support. Although *M. plurifarium* and *M. septentrionale* type strains are included in cluster A, no isolate was found to group closely to these strains. In terms of 16S rRNA gene sequence, isolate C-14-Coimbra is 100% identical to *M. huakuii* and isolate STR-2-Santarém is 100% identical to *M. amorphae*. Isolate L-19-Leiria shares the same sequence similarity (99.7%) with both *M. amorphae* and *M. huakuii*. Isolates A-8b.-Aveiro and SA-4-Serra d'Água are more similar to *M. amorphae* (99.5% similarity). All other isolates share higher sequence similarity with *M. huakuii* (99.7%) and are divided in subgroups apart from any *Mesorhizobium* type strain; some of them might be different enough to represent new *Mesorhizobium* species. Cluster B includes eight isolates from four provinces (Douro Litoral, Estremadura, Alto Alentejo and Baixo Alentejo), together with *M. loti* and *M. ciceri* type strains. Six of these isolates are closer to *M. ciceri* (99.9-100%) and two isolates are closer to *M. loti* (99.8%). Cluster C includes four isolates from three different provinces (Estremadura, Ribatejo and Alto Alentejo) that share 99.8-99.9% of sequence similarity with *M. tianshanense*, the only type strain included in this cluster. Cluster D comprises the type strains of *M. mediterraneum* and *M. temperatum* and includes six isolates from three provinces (Alto Alentejo, Baixo Alentejo and Algarve), all identical in sequence.

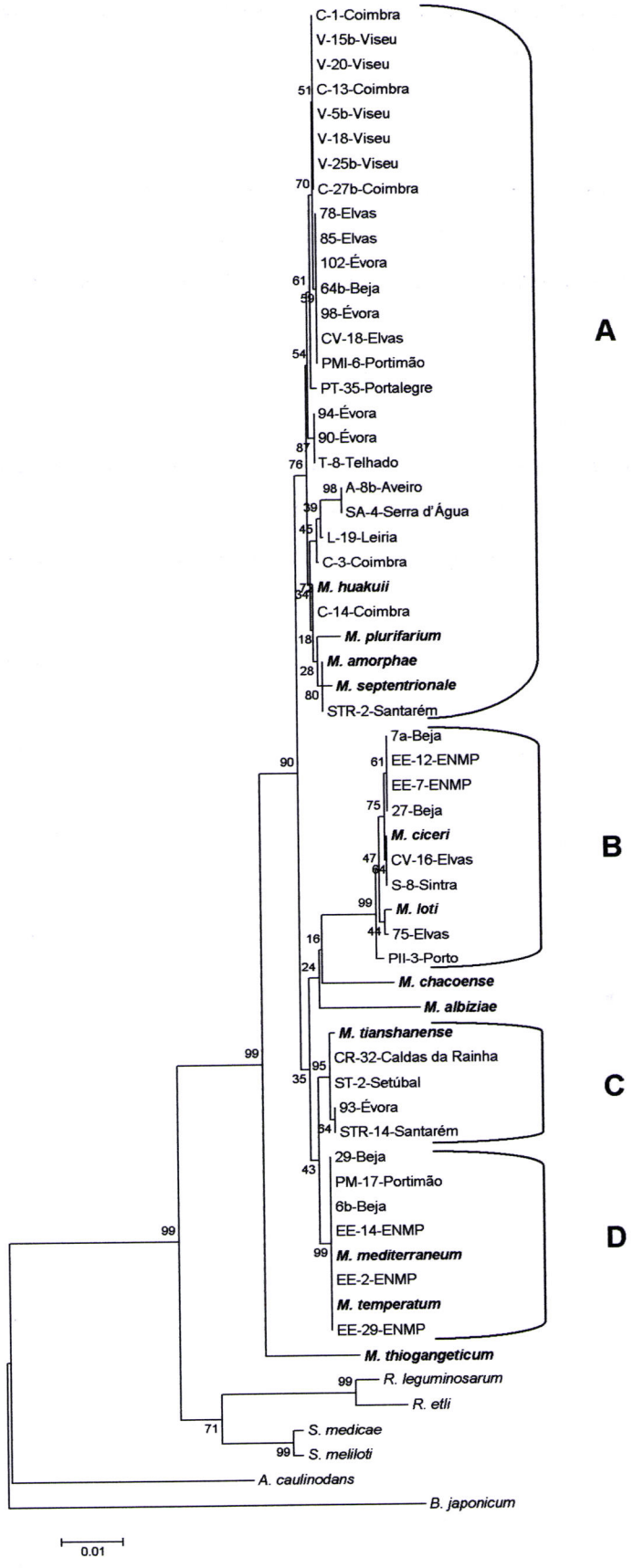
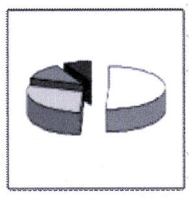


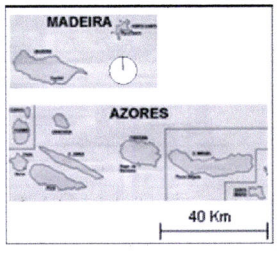
Figure 2.1 - Phylogeny showing the relationships of chickpea rhizobia isolates and rhizobia type strains, based on 16S rRNA gene analysis (alignment length 1357 bp). Neighbour-joining

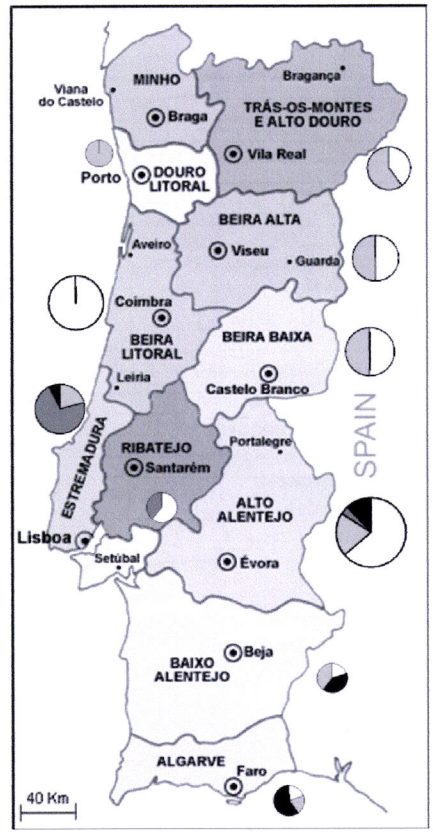
method was used. Bootstrap values are listed at the nodes. The four main clusters generated are marked with letters A to D. The scale bar indicates 1% substitutions per site.

In order to extend the analysis of species diversity, the complete set of 110 Portuguese isolates was used. Taking into account that the complete and the partial 16S rRNA gene sequence analysis, for the set of 43 isolates, generated the same four main clusters (data not shown), the analysis of the total set of 110 isolates was performed using their partial 16S rRNA gene sequence (Table 2.2). A dendrogram was generated from a 578 bp long alignment with 104 variable sites (data not shown). The 110 isolates are comprised in the previously described four main clusters: cluster A (56 isolates); cluster B (32 isolates); cluster C (13 isolates); and cluster D (9 isolates). This analysis revealed that chickpea rhizobia isolates are highly diverse and group with nine *Mesorhizobium* type strains. No isolate groups with *M. chacoense*, *M. albiziae* or *M. thiogangeticum*.



(A)





(B)

□ Cluster A – isolates close to the type strains of *M. huakuii* and *M. amorphae*, *M. plurifarium* and *M. septentrionale.*

□ Cluster B – isolates close to the type strains of *M. ciceri* and *M. loti.*

■ Cluster C – isolates close to the type strain of *M. tianshanense*.

■ Cluster D – isolates close to the type strains of *M. mediterraneum* and *M. temperatum*.

Figure 2.2 – Distribution of the 110 chickpea rhizobia isolates by species clusters, as defined from the 16S rRNA gene-based phylogeny (A). Map of Portugal showing the distribution of isolates in each province, according to their cluster. Pie charts sizes are proportional to the number of isolates in each province (B).

The Fig. 2.2 shows the geographical distribution of isolates by provinces, according to their 16S rRNA gene clusters. In the North of Portugal (Trás-os-Montes e Alto Douro and Douro Litoral), isolates from cluster B prevail, while in the Centre (Beira Alta, Beira Litoral, Beira Baixa, Ribatejo and Alto Alentejo) most isolates are from cluster A and in the South (Baixo Alentejo and Algarve) isolates mainly belong to cluster D. All isolates from Madeira belong to cluster A. Moreover, Estremadura is the only province where isolates from cluster C predominate and is the single Centre province with no isolates from cluster A. Isolates from cluster C are found only in three provinces of the Centre of Portugal (Estremadura, Ribatejo and Alto Alentejo).

The geographic distribution of isolates according to their species cluster is not random given that an association was found between species cluster and origin province of individual isolates (χ^2=126.382, df=30, P<0.001). The CA biplot (data not shown) revealed the existence of three classes of sites, consistent with the distribution of isolates observed in Fig. 2.2. One class, which includes Beira Litoral, Alto Alentejo and Madeira, is associated with cluster A. A second class, Estremadura, is mainly associated with Cluster C. Finally, a class including Trás-os-Montes e Alto Douro, Douro Litoral, Baixo Alentejo and Algarve, is associated with clusters B and D.

Plasmid Profiles

Plasmid profiles were analysed (Fig. 2.3) and for most chickpea native rhizobia at least one plasmid was detected. Plasmid number ranges from zero to six, though no isolate with five plasmids was found (Table 2.2).

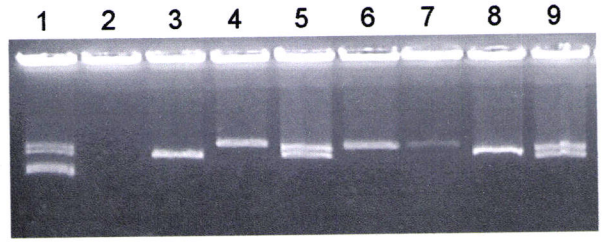


Figure 2.3 – Example of an agarose gel showing plasmid profiles of chickpea rhizobia isolates. Lane 1 - 29-Beja; lane 2 – CR-6-Caldas da Rainha; lane 3 – C-3-Coimbra; lane 4 – C-7-Coimbra; lane 5 – C-9-Coimbra; lane 6 – C-15-Coimbra; lane 7 – C-23-Coimbra; lane 8 – C-24-Coimbra; lane 9 – C-25-Coimbra.

For 39% of the isolates one plasmid was detected. Only in about 9% of the isolates, three or more plasmids were detected. An association was found between plasmid

number and province (χ^2=99.295, df=45, P<0.001). Alto Alentejo is the province with isolates more variable in terms of plasmid number, harbouring zero to four plasmids, while isolates from Douro Litoral, Beira Alta, Ribatejo, Algarve and Madeira show the least variability in plasmid number. There is also an association between plasmid number and species clusters (χ^2=59.740, df=15, P<0.001).

The Fig. 2.4 shows the distribution of isolates in each cluster, according to their plasmid number. Within cluster A, most isolates have one plasmid, while in clusters B and C, predominate isolates with no plasmids. Cluster D only includes isolates with two or more plasmids.

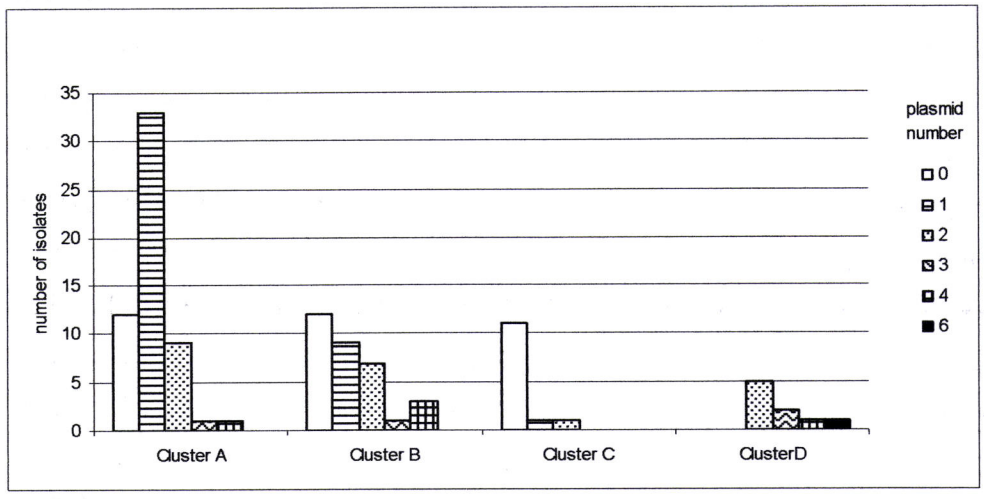


Figure 2.4 – Number of isolates in each cluster according to their plasmid number.

Symbiotic Effectiveness (SE)

Evaluation of SE was performed for all 110 isolates (Table 2.2). SE values range from 0 to 100%. SE trials revealed that 40% of the isolates showed a SE above 50%. Eleven isolates, which represent about 10% of the total number of isolates, were found to be highly effective in fixing N_2 in symbiosis with chickpea (SE values above 75%); most of these isolates belong to cluster B. The most effective isolates were T-1-Telhado from cluster A and G-55-Guarda from cluster B (SE values of 100% and 88%, respectively). Although isolates closer to *M. ciceri* / *M. loti* (cluster B) have the highest mean SE (51%), no correlation was found between SE and species clusters. No correlation was found between SE and plasmid number, contrary to a previous study with a smaller set of isolates from Alentejo provinces (Laranjo *et al.*, 2002).

2.5 - Discussion

The present study is the first survey of chickpea rhizobia native populations covering the Portuguese territory. One hundred and ten isolates were confirmed as chickpea symbionts and identified as *Mesorhizobium* sp., forming a monophyletic cluster with all *Mesorhizobium* type strains in the 16S rRNA gene phylogeny. The analysis of the several 16SrRNA copies for the rhizobia with completely sequenced genomes shows no evidence of divergent copies (data not shown), so this gene is useful in placing any new isolate within the known *Mesorhizobium* species.

The four main clusters of the complete 16S rRNA gene-based phylogeny (Fig. 2.1) show that isolates positioning is scattered within the *Mesorhizobium* genus. Isolates from cluster A, which are more related to *M. huakuii* and *M. amorphae* than to any other type strain, are the most abundant chickpea microsymbionts found in Portuguese soils. This was unexpected since the *M. huakuii* and *M. amorphae* type strains are unable to nodulate chickpea. *M. huakuii* was originally isolated from *Astragalus sinicus* (Chen *et al.*, 1991) that does not exist in Portugal. *M. amorphae* was originally isolated from *Amorpha fruticosa* (Chen *et al.*, 1991), a plant unrelated to *Cicer arietinum*, which is uncommon and considered invasive in Portugal. However, it is probable that most isolates from cluster A belong to a new species. *M. ciceri* and *M. mediterraneum* species groups (clusters B and D, respectively) could be expected to include the majority of native isolates, as these species were described as the specific chickpea microsymbionts (Nour *et al.*, 1995, Nour *et al.*, 1994b). However, only 33% of the isolates grouped with these two type strains. Isolates related to *M. amorphae* (cluster A), *M. loti* (cluster B) and *M. tianshanense* (cluster C) were found, as in previous studies on chickpea rhizobia isolated from Portugal and Spain (Rivas *et al.*, 2007, Laranjo *et al.*, 2004). The present work screened the entire Portuguese territory, confirmed the high diversity of native rhizobia and revealed an unexpected high proportion of isolates unrelated to *M. ciceri* and *M. mediterraneum*.

To our knowledge there is only one study addressing the diversity of chickpea rhizobia covering an entire country, which was performed in Morocco (Maâtallah *et al.*, 2002a). Using PCR-RFLP analysis of the 16S rRNA gene, the authors found a lower diversity of chickpea rhizobia than the one revealed by the present study, since most isolates were described as close to *M. ciceri*, *M. loti* and *M. mediterraneum*. However, four isolates were described as close to *Sinorhizobium* species. More recently, L'Taief and co-workers (2007) isolated chickpea native rhizobia from several regions of Tunisia and found isolates only belonging to either *M. ciceri* or *M. mediterraneum*. Probably, the low diversity found in Tunisia is related to the history of chickpea cultivation on the

sampled sites, as supported by several studies reporting a decrease in rhizobia diversity associated with the presence of the host plant (Coutinho *et al.*, 1999). Accordingly, the high diversity found in Portuguese soils could be explained by the absence of chickpea crop in Portugal (Duarte Maçãs, 2003) and the non-existence of chickpea wild relatives (Talavera *et al.*, 1999); furthermore, there are no records of the use of commercial inoculants that could reduce the natural chickpea rhizobia diversity. Interestingly, isolates from the single site where chickpea has been cultivated (Elvas-ENMP) group with *M. ciceri* or *M. mediterraneum*.

A previous study on chickpea rhizobia diversity in two Portuguese provinces revealed a group of isolates apart from any type strain, still closer to *M. huakuii*, that could represent a new *Mesorhizobium* species (Laranjo *et al.*, 2004). Sequence analysis of other housekeeping genes confirmed the separate position of this group of isolates (Laranjo *et al.*, unpublished results). The present study further supports this putative new species, since 17 new isolates from five different provinces were included into this group. Thus, isolates from cluster A should be further studied in order to investigate their species affiliation.

Considering all isolates, an association was found between province of origin and species clusters of isolates, suggesting that the geographical distribution is not random and that some species are typically found in a certain region (Fig. 2.2). For instance, most isolates found in the Centre of Portugal belong to cluster A. A higher diversity of isolates species was found in the Centre-South and South of the country.

A correlation was found between isolates species cluster and origin soil pH ($P<0,001$), which confirms our previous results obtained with a smaller set of isolates (Brígido *et al.*, 2007). For example, all isolates assigned to the *M. mediterraneum* / *M. temperatum* species cluster D were obtained from the soils with higher pH values. This may indicate that genetic determinants, which allow rhizobia survival in alkaline soil conditions, are species specific.

Considering the correlation found between species cluster and soil pH, it is likely that pH is a key environmental parameter determining the species geographic distribution. This hypothesis is supported by wider studies addressing soil bacterial communities, which indicate soil pH as the variable that best explains the population diversity and overall community composition (Fierer & Jackson, 2006). Several studies in rhizobia showed that the pH affects both survival and competitiveness in soil, as well as the nodulation process (Zahran, 1999). The effect of pH in chickpea rhizobia growth has been addressed in previous studies (Rodrigues *et al.*, 2006, Brígido *et al.*, 2007).

These showed a positive correlation between maximum growth pH and origin soil pH, for isolates belonging to the four species clusters (C-1-Coimbra, PT-35-Portalegre and 64b.-Beja, from cluster A; 75-Elvas from cluster B; CR-32-Caldas da Rainha from cluster C; 29-Beja from cluster D) (Brígido *et al.*, 2007). In addition, using three isolates, a higher symbiotic effectiveness was achieved using watering solution at a pH value closest to the bacterial maximum growth pH (Brígido *et al.*, 2007). Altogether these studies suggest that pH is a key environmental factor for rhizobia population composition, acting on bacteria, both free-living and in symbiosis.

In each 16S rRNA gene-based cluster, isolates with high and low symbiotic effectiveness were found. A large set of isolates with very high SE values (above 75%) are good candidates for field inoculation. Many of these isolates are from the *M. ciceri* cluster (B). Interestingly, the most effective isolate (T-1-Telhado) is closer to *M. huakuii* (cluster A), which is not a chickpea microsymbiont. About 70% of the isolates from cluster B present a SE value above the corresponding type strain *M. ciceri*, which showed a SE of 41%, estimated in a previous study (Laranjo *et al.*, 2008). In the *M. mediterraneum* / *M.temperatum* cluster (D) 56% of the isolates showed a SE above 39%, which is the value described for the type strain of *M. mediterraneum* (Laranjo *et al.*, 2008).

The plasmid number of rhizobia isolates was found to be associated with species cluster, suggesting that this feature might be species constrained. In most isolates from cluster A, one plasmid was detected, similarly to *M. amorphae* (Wang *et al.*, 1999). In cluster D, both isolates and *M. mediterraneum* (Cadahía *et al.*, 1986) showed more than one plasmid. All cluster C isolates, except for two, lack plasmids, similarly to *M. tianshanense* (Chen *et al.*, 1995). Isolates from cluster B seem to be more diverse in plasmid number, including isolates with zero, one and two plasmids. The type strain of *M. ciceri* (cluster B) harbours one plasmid (Cadahía *et al.*, 1986).

Contrary to previous studies, which suggested that most rhizobia nodulating chickpea are *M. ciceri* and *M. mediterraneum*, this wider survey shows a predominance of other species. The obtained isolates collection, highly diverse in terms of species, as well as SE, provides an important source of rhizobia strains to be used, namely as potential inoculants. The present study is the first systematic assessment of *Cicer arietinum* microsymbionts in Portugal and contributes to clarify the biogeography of chickpea rhizobia, providing a global picture of how species are distributed across the country.

2.6 - References

Alexandre, A., M. Laranjo & S. Oliveira, (2006) Natural populations of chickpea rhizobia evaluated by antibiotic resistance profiles and molecular methods. *Microb. Ecol.* **51**: 128-136.

Aouani, M. E., R. Mhamdi, M. Jebara & N. Amarger, (2001) Characterization of rhizobia nodulating chickpea in Tunisia. *Agronomie* **21**: 577-581.

Benzécri, J. P., (1973) *Analyse des données. Tome I: Analyse des correspondances. Tome II: La Classification.* Dunod, Paris.

Brígido, C., A. Alexandre, M. Laranjo & S. Oliveira, (2007) Moderately acidophilic mesorhizobia isolated from chickpea. *Lett. Appl. Microbiol.* **44**: 168-174.

Broughton, W. J. & M. J. Dilworth, (1971) Control of Leghaemoglobin Synthesis in Snake Beans. *Biochem. J.* **125**: 1075-1080.

Broughton, W. J. & X. Perret, (1999) Genealogy of legume-*Rhizobium* symbioses. *Curr. Opin. Plant Biol.* **2**: 305-311.

Cadahía, E., A. Leyva & T. Ruiz-Argüeso, (1986) Indigenous plasmids and cultural characteristics of rhizobia nodulating chickpeas (*Cicer arietinum* L.). *Arch. Microbiol.* **146**: 239-244.

Chen, W. X., G. S. Li, Y. L. Qi, E. T. Wang, H. L. Yuan & J. L. Li, (1991) *Rhizobium huakuii* sp. nov. Isolated from the Root-Nodules of *Astragalus sinicus*. *Int. J. Syst. Bacteriol.* **41**: 275-280.

Chen, W. X., E. Wang, S. Y. Wang, Y. B. Li & X. Q. Chen, (1995) Characteristics of *Rhizobium tianshanense* sp. nov., a moderately and slowly growing root-nodule bacterium isolated from an arid saline environment in Xinjiang, Peoples-Republic-of-China. *Int. J. Syst. Bacteriol.* **45**: 153-159.

Clayton, R., G. Sutton, P. Hinkle, Jr, C. Bult & C. Fields, (1995) Intraspecific variation in small-subunit rRNA sequences in GenBank: why single sequences may not adequately represent prokaryotic taxa. *Int. J. Syst. Bacteriol.* **45**: 595-599.

Coenye, T., D. Gevers, Y. V. Peer, P. Vandamme & J. Swings, (2005) Towards a prokaryotic genomic taxonomy. *Fems Microbiol. Rev.* **29**: 147-167.

Coutinho, H. L. D., H. E. Kay, G. P. Manfio, M. C. P. Neves, J. R. A. Ribeiro, N. G. Rumjanek & J. E. Beringer, (1999) Molecular evidence for shifts in polysaccharide composition associated with adaptation of soybean *Bradyrhizobium* strains to the Brazilian Cerrado soils. *Environ. Microbiol.* **1**: 401-408.

Donate-Correa, J., M. León-Barrios, M. Hernández, R. Pérez-Galdona & M. del Arco-Aguilar, (2007) Different *Mesorhizobium* species sharing the same symbiotic genes nodulate the shrub legume *Anagyris latifolia*. *Syst. Appl. Microbiol.* **30**: 615-623.

Downie, J. A., (1998) Functions of rhizobial nodulation genes. In: The Rhizobiaceae Molecular Biology of Model Plant-Associated Bacteria. H. P. Spaink, A. Kondorosi & P. J. J. Hooykaas (eds). Dordrecht: Kluwer Academic Publishers, pp. 387-402.

Duarte Maçãs, I. M. V., (2003) Selecção de linhas de grão de bico (*Cicer arietinum* L.) adaptadas ao ambiente Mediterrânico-critérios morfológicos e fisiológicos. In. Évora: Universidade de Évora, pp. 171.

Fierer, N. & R. B. Jackson, (2006) The diversity and biogeography of soil bacterial communities. *Proc. Natl. Acad. Sci. U. S. A.* **103**: 626-631.

Gevers, D. & T. Coenye, (2007) Phylogenetic and genomic analysis. In: Manual of Environmental Microbiology. C. J. Hurst, R. L. Crawford, J. L. Garland, D. A. Lipson, A. L. Milles & L. D. Stetzenbach (eds). ASM Press, pp. 1352.

Gibson, A. H., (1987) Evaluation of nitrogen fixation by legumes in the greenhouse and growth chamber. In: Symbiotic Nitrogen Fixation Technology. G. H. Elkan (ed). New York: Marcel Dekker, Inc, pp. 321-363.

Hall, T. A., (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* **41**: 95-98.

Jarabo-Lorenzo, A., R. Perez-Galdona, J. Donate-Correa, R. Rivas, E. Velázquez, M. Hernandez, F. Temprano, E. Martinez-Molina, T. Ruíz-Argüeso & M. Leon-Barrios, (2003) Genetic diversity of bradyrhizobial populations from diverse geographic origins that nodulate *Lupinus* spp. and *Ornithopus* spp. *Syst. Appl. Microbiol.* **26**: 611-623.

Kuykendall, L. D., Y. D. Gaur & S. K. Dutta, (1993) Genetic Diversity among *Rhizobium* Strains from *Cicer arietinum* L. *Lett. Appl. Microbiol.* **17**: 259-263.

L'Taief, B., B. Sifi, M. Gtari, M. Zaman-Allah & M. Lachaal, (2007) Phenotypic and molecular characterization of chickpea rhizobia isolated from different areas of Tunisia. *Can. J. Microbiol.* **53**: 427-434.

Laranjo, M., A. Alexandre, R. Rivas, E. Velázquez, J. P. W. Young & S. Oliveira, (2008) Chickpea rhizobia symbiosis genes are highly conserved across multiple *Mesorhizobium* species. *FEMS Microbiol. Ecol.* **66**: 391-400.

Laranjo, M., C. Branco, R. Soares, L. Alho, M. Carvalho & S. Oliveira, (2002) Comparison of chickpea rhizobia isolates from diverse Portuguese natural populations based on symbiotic effectiveness and DNA fingerprint. *J. Appl. Microbiol.* **92**: 1043-1050.

Laranjo, M., J. Machado, J. P. W. Young & S. Oliveira, (2004) High diversity of chickpea *Mesorhizobium* species isolated in a Portuguese agricultural region. *FEMS Microbiol. Ecol.* **48**: 101-107.

Laranjo, M., R. Rodrigues, L. Alho & S. Oliveira, (2001) Rhizobia of chickpea from southern Portugal: symbiotic efficiency and genetic diversity. *J. Appl. Microbiol.* **90**: 662-667.

Louvrier, P., G. Laguerre & N. Amarger, (1996) Distribution of symbiotic genotypes in *Rhizobium leguminosarum* biovar *viciae* populations isolated directly from soils. *Appl. Environ. Microbiol.* **62**: 4202-4205.

Maâtallah, J., E. Berraho, J. Sanjuan & C. Lluch, (2002a) Phenotypic characterization of rhizobia isolated from chickpea (*Cicer arietinum*) growing in Moroccan soils. *Agronomie* **22**: 321-329.

Maâtallah, J., E. B. Berraho, S. Muñoz, J. Sanjuan & C. Lluch, (2002b) Phenotypic and molecular characterization of chickpea rhizobia isolated from different areas of Morocco. *J. Appl. Microbiol.* **93**: 531-540.

Martens, M., M. Delaere, R. Coopman, P. De Vos, M. Gillis & A. Willems, (2007) Multilocus sequence analysis of *Ensifer* and related taxa. *Int. J. Syst. Evol. Microbiol.* **57**: 489-503.

Menna, P., M. Hungria, F. G. Barcellos, E. V. Bangel, P. N. Hess & E. Martínez-Romero, (2006) Molecular phylogeny based on the 16S rRNA gene of elite rhizobial strains. *Syst. Appl. Microbiol.* **29**: 315-332.

Nour, S. M., J.-C. Cleyet-Marel, P. Normand & M. P. Fernandez, (1995) Genomic heterogeneity of strains nodulating chickpeas (*Cicer arietinum* L.) and description of *Rhizobium mediterraneum* sp. nov. *Int. J. Syst. Bacteriol.* **45**: 640-648.

Nour, S. M., J. C. Cleyet-Marel, D. Beck, A. Effosse & M. P. Fernandez, (1994a) Genotypic and phenotypic diversity of *Rhizobium* isolated from chickpea (*Cicer arietinum* L.). *Can. J. Microbiol.* **40**: 345-354.

Nour, S. M., M. P. Fernandez, P. Normand & J.-C. Cleyet-Marel, (1994b) *Rhizobium ciceri* sp. nov., consisting of strains that nodulate chickpeas (*Cicer arietinum* L.). *Int. J. Syst. Bacteriol.* **44**: 511-522.

Perret, X., C. Staehelin & W. J. Broughton, (2000) Molecular basis of symbiotic promiscuity. *Microbiol. Mol. Biol. Rev.* **64**: 180-201.

Rivas, R., M. Laranjo, P. F. Mateos, S. Oliveira, E. Martinez-Molina & E. Velázquez, (2007) Strains of *Mesorhizobium amorphae* and *Mesorhizobium tianshanense*, carrying symbiotic genes of common chickpea endosymbiotic species, constitute a novel biovar (*ciceri*) capable of nodulating *Cicer arietinum*. *Lett. Appl. Microbiol.* **44**: 412-418.

Rodrigues, C., M. Laranjo & S. Oliveira, (2006) Effect of heat and pH stress in the growth of chickpea mesorhizobia. *Curr. Microbiol.* **53**: 1-7.

Somasegaran, P. & H. J. Hoben, (1994) *Handbook for Rhizobia,* p. 450. Springer-Verlag, New York.

Sun, L., F. B. Qiu, X. X. Zhang, X. Dai, X. Z. Dong & W. Song, (2008) Endophytic bacterial diversity in rice (*Oryza sativa* L.) roots estimated by 16S rDNA sequence analysis. *Microb. Ecol.* **55**: 415-424.

Talavera, S., C. Aedo, S. Castroviejo, C. Romero Zarco, L. Sáez, F. J. Salgueiro & M. Velayos, (1999) *Flora iberica-Plantas vasculares de la Península Ibérica e Islas Baleares*. Real Jardín Botánico, CSIC, Madrid, Spain.

Tamura, K., J. Dudley, M. Nei & S. Kumar, (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) Software Version 4.0. *Mol. Biol. Evol.* **24**: 1596-1599.

Thompson, J. D., D. G. Higgins & T. J. Gibson, (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence. *Nucleic Acids Res.* **22**: 4673-4680.

van Berkum, P., P. Elia & B. D. Eardly, (2006) Multilocus Sequence Typing as an Approach for Population Analysis of *Medicago*-Nodulating Rhizobia. *J. Bacteriol.* **188**: 5570-5577.

Vincent, J. M., (1970) *A Manual for the Practical Study of Root-Nodule Bacteria,* p. 164. Blackwell Scientific Publications, Oxford.

Wang, E. T., P. van Berkum, X. H. Sui, D. Beyene, W. X. Chen & E. Martínez-Romero, (1999) Diversity of rhizobia associated with *Amorpha fruticosa* isolated from Chinese soils and description of *Mesorhizobium amorphae* sp. nov. *Int. J. Syst. Bacteriol.* **49**: 51-65.

Young, J. M. & D. C. Park, (2007) Relationships of plant pathogenic enterobacteria based on partial *atpD*, *carA*, and *recA* as individual and concatenated nucleotide and peptide sequences. *Syst. Appl. Microbiol.* **30**: 343-354.

Young, J. P., H. L. Downer & B. D. Eardly, (1991) Phylogeny of the phototrophic *Rhizobium* strain BTAi1 by polymerase chain reaction-based sequencing of a 16S rRNA gene segment. *J. Bacteriol.* **173**: 2271-2277.

Zahran, H. H., (1999) *Rhizobium*-legume symbiosis and nitrogen fixation under severe conditions and in an arid climate. *Microbiol. Mol. Biol. Rev.* **63**: 968-989.

3.

dnaJ as a useful phylogenetic marker

This chapter is based on the manuscript:

Alexandre, A., Laranjo, M., Young, J.P.W. and Oliveira, S. (2008). *dnaJ* is a useful phylogenetic marker for alphaproteobacteria. International Journal of Systematic and Evolutionary Microbiology. 58: 2839-2849.

3.1 - Summary

In the past, bacterial phylogeny relied almost exclusively on 16S rRNA gene sequence analysis. More recently, multilocus sequences have been used to infer organismal phylogenies. In this study, the *dnaJ* chaperone gene was investigated as a marker for phylogeny studies in alphaproteobacteria. Preliminary analysis of G+C and G+C3s contents showed no clear evidence of horizontal transfer of this gene in proteobacteria. *dnaJ*-based phylogenies were then analysed at three taxonomic levels: *Proteobacteria*, *Alphaproteobacteria* and the genus *Mesorhizobium*. Dendrograms based on DnaJ and 16S rRNA gene sequences revealed the same topology already described for *Proteobacteria*. These results indicate that the DnaJ phylogenetic signal is able to reproduce the accepted relationships among the five *Proteobacteria* classes. At a lower taxonomic level, using 20 alphaproteobacteria species, the 16S rRNA gene-based phylogeny is distinct from the one based on the DnaJ sequence analysis. Although the same clusters are generated, only the topology of the DnaJ tree is consistent with broader phylogenies from recent studies based on concatenated alignments of multiple core genes. For example, the DnaJ tree shows the two *Rhizobiales* clusters, closely related, as expected, while the 16S rRNA gene-based phylogeny shows them distantly related. In order to evaluate the phylogenetic performance of *dnaJ* at the *Mesorhizobium* genus level, a multilocus analysis based on five housekeeping genes (*atpD, gapA, gyrB, recA, rplB*) was performed for ten *Mesorhizobium* type species. In contrast to the 16S rRNA gene, the DnaJ sequence analysis generated a tree similar to the multilocus dendrogram. For identification of chickpea mesorhizobium isolates, a *dnaJ* nucleotide-based tree was used. Despite different topologies, 16S rRNA gene- and *dnaJ*-based trees led to the same species identification. This study suggests that the *dnaJ* gene is a good phylogenetic marker, particularly for the *Alphaproteobacteria* class, since its phylogeny is consistent with phylogenies based on multilocus approaches.

3.2 - Introduction

The phylum *Proteobacteria* is composed of five classes: *Alphaproteobacteria* *Betaproteobacteria*, *Gammaproteobacteria*, *Deltaproteobacteria* and *Epsilonproteobacteria* (Stackebrandt *et al.*, 1988, Garrity & Lilburn, 2005). By the end of 2009, more than 450 complete genomes of proteobacteria were available in public databases (http://www.ncbi.nlm.nih.gov/genomes/lproks.cgi). Alphaproteobacteria exhibit an enormous diversity in their morphological and metabolic characteristics and the Class is presently recognized solely as a clade in the 16S rRNA gene-based phylogeny (Stackebrandt *et al.*, 1988).

Based on 16S rRNA gene trees, *Alphaproteobacteria* have been divided into seven orders: *Caulobacterales*, *Rhizobiales*, *Rhodobacterales*, *Rhodospirillales*, *Rickettsiales*, *Sphingomonadales* and *Parvularculales* (Kersters *et al.*, 2007). The *Alphaproteobacteria* include important bacteria that are widely studied, including, for example, the most important genera of soil bacteria able to live in symbiosis with leguminous plants (order *Rhizobiales*).

Evolutionary relationships among bacteria have been estimated by 16S rRNA gene sequence comparisons, since this was originally considered to reflect, more or less, organismal phylogeny (Olsen & Woese, 1993). More recently, it has been shown that rRNA genes may show sequence heterogeneity and undergo horizontal transfer and genetic recombination (Acinas *et al.*, 2004). Furthermore, due to its very high sequence conservation, the 16S rRNA gene has limited usefulness in resolving closely related species. Another disadvantage of this gene is that most bacteria harbor several copies of 16S rRNA gene, and in some cases different copies tell different stories. For example, *Escherichia coli* 16S rRNA gene copies are divergent (Cilia *et al.*, 1996). For that reason, the search for other genes able to tell an evolutionary story for bacterial species has lead to the use of several housekeeping genes, including *rec*A (Eisen, 1995), *gyrB* (Yamamoto & Harayama, 1995), *rpoD* (Yamamoto *et al.*, 2000), *atpD* (Gaunt *et al.*, 2001) and *glnA* (Turner & Young, 2000), among others, as alternative phylogenetic markers. The 16S rRNA gene phylogeny is not always in agreement with the phylogenies based on housekeeping loci (Xiao *et al.*, 2007). In most cases, 16S rRNA gene and housekeeping genes phylogenies are similar, but the phylogeny generated by protein coding genes frequently shows higher resolution and reliability than the corresponding 16S rRNA gene-based phylogeny (Thompson *et al.*, 2004).

Recent phylogenetic studies based on multiple protein sequences from completely sequenced genomes are believed to reveal the "true" evolutionary history of bacteria.

For example, Fukami-Kobayashi and co-workers (2007) showed a tree of life comprising 167 organisms belonging to *Archaea*, *Bacteria* and *Eukarya*, based on domain organization of all proteins encoded in each species genome; Ciccarelli *et al.* (2006) showed a global phylogeny comprising 191 species (*Archaea*, *Bacteria* and *Eukarya*) based on 31 universal proteins. Gupta *et al.* (2007) studied just proteobacteria, using the amino acid sequences of 10 protein-coding genes. A larger number of proteins (648) was used by Young *et al.* (2006) to generate a phylogeny that shows the relationship of *Rhizobium leguminosarum* and its close relatives with completely sequenced genomes. These studies indicate that a large amount of core gene sequence information produces strongly supported and more resolved phylogenies, frequently in disagreement with the 16S rRNA gene phylogenies. Single gene-based phylogenies may now be validated by comparison with presumed organismal phylogeny obtained from multigenic approaches.

Genes coding for molecular chaperones have been used as alternative phylogenetic markers to 16S rRNA gene. The most commonly used are *groEL* (Baldo *et al.*, 2006) and *dnaK* (Vitorino *et al.*, 2007). *groEL* gene can be useful for gammaproteobacteria phylogenies, however many alphaproteobacteria species harbour multiple copies of this gene, often divergent (Lund, 2009). In most genomes, *dnaK* is a single copy gene often in the same operon as *dnaJ*. This operon can also contain other heat shock genes, such as *grpE* or *hrcA*. For example, in *Bacillus subtilis* (Firmicutes) the *dnaK* operon encodes for *hrcA* and *grpE*, besides *dnaK* and *dnaJ*. In *E. coli* (*Gammaproteobacteria*) the *dnaK* operon is bicistronic comprising the *dnaK* (1917 bp) and *dnaJ* (1131 bp) genes (Saito & Uchida, 1978). The *dnaJ* gene is sometimes found in more than one copy, as in the high G+C gram positive bacteria *Actinobacteria*. It has been reported that these different copies may have different physiological functions and distinct evolutionary paths (Ventura *et al.*, 2005).

A high degree of sequence conservation, functional preservation and universal distribution among bacteria are key features that make the *dnaJ* gene suitable for inferring phylogenetic relationships in bacteria. *dnaJ* has already been successfully used for discrimination of species and subspecies, even within genera in which 16S rRNA gene has not enough resolution, as in pathogenic bacteria, such as *Mycobacterium* (Morita *et al.*, 2004), *Streptococcus* (Itoh *et al.*, 2006) and more recently *Staphylococcus* (Shah *et al.*, 2007) and *Vibrio* (Nhung *et al.*, 2007). To our knowledge, there is only one study using *dnaJ* as a phylogenetic marker within *Proteobacteria*, with the purpose of discriminating *Legionella pneumophila* serogroups (Liu *et al.*, 2003).

Our aim was to find a single core gene with a phylogenetic signal resembling multilocus phylogenies at higher taxonomic levels (*Proteobacteria* and *Alphaproteobacteria*) and with a good phylogenetic signal at a lower taxonomic level, namely for the genus *Mesorhizobium*. We analyzed the phylogenetic signal of *dnaJ* gene from 19 proteobacteria and from 20 alphaproteobacteria and compared it with the 16S rRNA gene phylogeny, as well as with phylogenies from other studies that were based on multiple proteins. We also used the *dnaJ* gene to confirm the previous 16S rRNA-based identification of chickpea rhizobia isolates within the genus *Mesorhizobium* (*Alphaproteobacteria*). The *dnaJ*-based phylogeny was, in this case, compared to that of five concatenated genes (*atpD-gapA-gyrB-recA-rplB*) and 16S rRNA gene phylogenies.

3.3 - Methods

Amplification of the *dnaJ* gene

Bacterial genomes frequently harbor more than one gene annotated as *dnaJ*. To prevent amplification of multiple copies in the same strain and of different copies in different species, a forward primer in the *dnaK* gene was used for PCR amplification (*dnaK* is a single copy gene, with very few exceptions). This guarantees that the target sequence is the *dnaJ* copy found in the *dnaKJ* operon. *dnaJ* amplification by PCR was performed for nine chickpea rhizobia isolates and ten *Mesorhizobium* type strains. Primers *dnaK*-F (5´-CAGATCGAGGTSACCTTCGAC-3') and *dnaJ*-R (5'-CGTCRYCATMGAGATCGGCAC-3') were used for PCR amplification reaction. PCR products were approximately 1600 bp. Total DNA was extracted as described previously (Rivas *et al.*, 2001). All PCR reaction mixtures contained 1 U FideliTaq DNA polymerase (usb), 1X reaction buffer (with 1.5 mM $MgCl_2$), 0.2 mM of each dNTP (Invitrogen), 5% DMSO (Duchefa), 25 pmol of each primer (Invitrogen) and 10 µl DNA. The amplification program consisted of an initial denaturation step of 2 min at 94ºC, followed by 30 cycles of 45 sec at 94ºC, 90 sec at 66-68ºC, 2 min at 68ºC, and a final extension step at 68ºC for 7 min.

Amplification of the 16S rRNA gene

16S rRNA gene was sequenced for four chickpea rhizobia isolates. The 16S rRNA gene was amplified by PCR as mentioned in the previous chapter.

Amplification of the *atpD* gene

The *atpD* gene, which codes for ATP synthase β was amplified using a pair of primers from Gaunt *et al.* (2001), *atpD*-F (5'-ATCGGCGAGCCGGTCGACGA-3') and *atpD*-R. The PCR product was a fragment of approximately 470 bp. All PCR reactions were prepared with 1 U FideliTaq DNA polymerase (usb), 1X reaction buffer (with 1.5 mM $MgCl_2$), 0.2 mM of each dNTP (Invitrogen), 5% DMSO (Duchefa), 25 pmol of each primer (Thermo Electron) and 10 µl DNA. The amplification program consisted of an initial denaturation step of 2 min at 95°C, followed by 30 cycles of 30 sec at 95°C, 30 sec at 60°C, 2 min at 68°C, and a final extension step at 68°C for 5 min.

Amplification of the *gapA* gene

The *gapA* gene, which codes for glyceraldehyde-3-phosphate dehydrogenase was amplified using the primers gap-for (5'-CATGCGGTTGGAGAAGCCCCA-3') and gap-rev (5'-TCGTCGCCGTCAACGATCTCG-3'). The amplified fragment was approximately 880 bp. All PCR reactions were prepared with 1 U FideliTaq DNA polymerase (usb), 1X reaction buffer (with 1.5 mM $MgCl_2$), 0.2 mM of each dNTP (Invitrogen), 5% DMSO (Duchefa), 15 pmol of each primer (Stab Vida) and 7 µl DNA. The amplification program consisted of an initial denaturation step of 2 min at 95°C, followed by 30 cycles of 30 sec at 95°C, 1 min at 63-65°C, 1 min at 68°C, and a final extension step at 68°C for 5 min.

Amplification of the *gyrB* gene

The *gyrB* gene, which codes for DNA gyrase subunit β was amplified using the primers gyrBfor_new (5'-TGCTGCTCACCTTCTTCTTCCG-3') and gyrBrev_new (5'-CCYTTGTAGCGCTGCATGGT-3'). The amplified fragment was approximately 695 bp. All PCR reactions were prepared with 1 U FideliTaq DNA polymerase (usb), 1X reaction buffer (with 1.5 mM $MgCl_2$), 0.2 mM of each dNTP (Invitrogen), 5% DMSO (Duchefa), 15 pmol of each primer (Stab Vida) and 7 µl DNA. The amplification program consisted of an initial denaturation step of 2 min at 95°C, followed by 30 cycles of 30 sec at 95°C, 1 min at 55°C, 1 min at 68°C, and a final extension step at 68°C for 5 min.

Amplification of the *recA* gene

The *recA* gene, which codes for DNA recombinase A was amplified using the primers *recA*-F (5'-ATCGAGCGGTCGTTCGGCAAGGG-3') and *recA*-R (5'-TTGCGCAGCGCCTGGCTCAT-3') (Gaunt *et al.*, 2001). The amplified fragment was approximately 440 bp. All PCR reactions were prepared with 2 U *Taq* DNA polymerase (Fermentas), 1X reaction buffer, 1.5 mM $MgCl_2$, 0.2 mM of each dNTP (Invitrogen), 0.004% BSA (Promega), 25 pmol of each primer (Biomers) and 3 µl DNA. The amplification program consisted of an initial denaturation step of 5 min at 95ºC, followed by 30 cycles of 45 sec at 94ºC, 1 min at 55ºC, and 1 min at 72ºC and a final extension step at 68ºC for 5 min.

Amplification of the *rplB* gene

The *rplB* gene, which codes for 50S ribosomal protein L2 involved in translation, was amplified using the primers L2-for (5'-TTCATCGGCGGTGGTCACAA-3') and L2-rev (5'-CGGAGGTGCGGCCTTCAC-3'). The amplified fragment was approximately 565 bp long. All PCR reactions were prepared with 1 U FideliTaq DNA polymerase (usb), 1X reaction buffer (with 1.5 mM $MgCl_2$), 0.2 mM of each dNTP (Invitrogen), 5% DMSO (Duchefa), 15 pmol of each primer (Stab Vida) and 7 µl DNA. The amplification program consisted of an initial denaturation step of 2 min at 95ºC, followed by 30 cycles of 30 sec at 95ºC, 1 min at 59ºC, 45 sec at 68ºC, and a final extension step at 68ºC for 5 min.

DNA purification and sequencing

PCR products were purified using GFX™ PCR DNA and Gel Band Purification kit (GE Healthcare) or ExoSAP-It (usb) following the manufacturer's instructions. Sequencing reactions were performed by Macrogen, Korea. For 16S rRNA gene sequencing, two extra primers, IntF and IntR, were used as internal primers for double-stranded sequencing (Laranjo *et al.*, 2004). For *dnaJ* sequencing, a forward primer located in the 5'-end of *dnaJ* gene was used (*dnaJ*-F – 5'-GCTGGGCGTGCAAAAGGG-3') together with the *dnaJ*-R primer. For the remaining genes the sequencing primers were the same used for PCR amplification.

Data analysis

Sequences generated in this study were edited using BioEdit 7.0.5.3 (Hall, 1999). The sequences were checked manually for correct alignment.

Gene sequences from completely sequenced genomes were retrieved from NCBI complete microbial genomes database.

In order to detect possible events of horizontal gene transfer (HGT) of the *dnaJ* gene, the G+C content and codon usage frequency were analysed for all proteobacteria species used (Eisen, 2000). The G+C content data for the chromosome was retrieved from the NCBI complete microbial genomes database. G+C content of individual genes was calculated using BioEdit 7.0.5.3 (Hall, 1999). In addition, the G+C content of the synonymous third codon position (G+C3s) was also calculated for the chromosome and the *dnaJ* gene. Both chromosome and gene G+C3s data were obtained using CodonW (Peden, 1999) (http://bioweb.pasteur.fr/seqanal/interfaces/codonw.html). Codon usage frequency analysis of the *dnaJ* gene was performed with Graphical Codon Usage Analyser 2.0 (http://gcua.schoedl.de/), using the codon usage table of each species available at the Codon Usage Database (http://www.kazusa.or.jp/codon/).

Molecular phylogeny was reconstructed using the PHYLIP package (version 3.67) (Felsenstein, 2006) by the maximum likelihood (ML) method with 100 replicates. Neighbour-joining (NJ) (Saitou & Nei, 1987) phylogenies, with 1000 bootstrap resamplings, were generated using MEGA3 (Kumar *et al.*, 2004). Evolutionary distances were calculated by Kimura's two-parameter model for nucleotide alignments (Kimura, 1980), whereas for amino acids alignments the Poisson correction (Nei & Kumar, 2000) was applied.

Dendrograms based on *dnaJ* sequences were always generated using amino acids, with exception for the one concerning the genus *Mesorhizobium*, for which nucleotides were used.

For a multilocus approach to the mesorhizobia phylogeny, five protein-coding genes were chosen: *atpD, gapA, gyrB, recA* and *rplB*. Incongruence length difference (ILD) test (Bull *et al.*, 1993, Cunningham, 1997) was performed to check whether trees for the different genes were sufficiently similar to allow data combination. The ILD test was performed using PAUP*4.0 (Swofford, 2003).

3.4 - Results

In the phylogenetic studies at higher taxonomic levels, sequences from 19 proteobacteria representing all classes (*Alpha-*, *Beta-*, *Gamma-*, *Delta-* and *Epsilonproteobacteria*) were used. For further phylogenetic analysis of alphaproteobacteria, 20 species from all orders, with the exception of *Parvularculales*, were selected. Finally, at a lower taxonomic level, ten type strains and nine isolates from the genus *Mesorhizobium* were used. Bacterial species used in this study are listed in Table 3.1.

Table 3.1 – Strains of the *Proteobacteria* with corresponding accession numbers or gene IDs of sequences used in this study.

			Accession number/Gene ID							
Species	**Strain**	**Genome**	**16S rRNA gene**	***dnaJ***	***dnaK***	***atpD***	***gapA***	***gyrB***	***recA***	***rplB***
Alphaproteobacteria										
Agrobacterium tumefaciens	C58	NC_003062	3244993	1137637	1137638					
Bradyrhizobim japonicum	USDA110	NC_004463	1055154	1049300	1049096	1054445	1055273	1047796	1048547	1051246
Brucella suis	1330[T]	NC_004310	1167560	1167829	1167828					
Caulobacter crescentus	CB15	NC_002696	943247	944076	944075					
Ensifer meliloti	1021	NC_003047	1234653	1231814	1234653	1234718	1234439	1231637	1233464	1233011
Erythrobacter litoralis	HTCC2594	NC_007722	3868981	3870763	3870762					
Gluconobacter oxydans	621H	NC_006677	3248602	3250315	3250314					
Granulibacter bethesdensis	CGDNIH1[T]	NC_008343	4276889	4274426	4274427					
Maricaulis maris	MCS10	NC_008347	4285688	4283983	4283981					
Mesorhizobium amorphae	ACCC19665[T]		AF041442	**EF504296**		AY493453	AM072544	AM076341	AY494816	AM076350
M. chacoense	LMG19008[T]		AJ278249	**EU273806**		AY493460	AM072546	AM076343	AY494825	AM076352
M. ciceri	UPM-Ca7[T]		DQ444456	**EF504297**		AJ294395	AM072545	AM076342	AJ294367	AM076351
M. ciceri	27-Beja			**EF504304**						
M. huakuii	CCBAU 2609[T]		D12797	**EF504298**		AJ294394	AM072547	AM076344	AJ294370	AM076353
M. loti	MAFF303099	NC_002678	3205748	1228214	1228215					
M. loti	LMG6125[T]		X67229	**EU053202**		**EU039868**	**EU273807**	**EU273810**	**EU039875**	**EU273813**
M. loti	75-Elvas			**EF504306**						
M. mediterraneum	UPM-Ca36[T]		L38825	**EF504299**		AM418768	AM072549	AM076346	AJ294369	AM076355
M. mediterraneum	29-Beja			**EF504305**						
M. plurifarium	ORS 1032[T]		Y14158	**EF504300**		AM076366	AM072550	AM076347	AY494824	AM076356
M. septentrionale	HAMBI 2582[T]		AF508207	**EF504301**		**DQ659498**	**EU273808**	**EU273811**	**DQ444304**	**EU273814**
M. temperatum	HAMBI 2583[T]		AF508208	**EF504302**		**DQ659499**	**EU273809**	**EU273812**	**DQ444305**	**EU273815**
M. tianshanense	A-1BS[T]		AF041447	**EF504303**		AM076367	AM072551	AM076348	AJ294368	AM076357
M. tianshanense	93-Évora			**EF504307**						
Mesorhizobium sp.	C1-Coimbra			**EF504308**						
Mesorhizobium sp.	C27b-Coimbra			**EF504309**						
Mesorhizobium sp.	V15b-Viseu			**EF504310**						
Mesorhizobium sp.	V18-Viseu			**EF504311**						
Mesorhizobium sp.	V20-Viseu			**EF504312**						
Nitrobacter winogradskyi	Nb-255[T]	NC_007406	3676957	3676654	3676653					

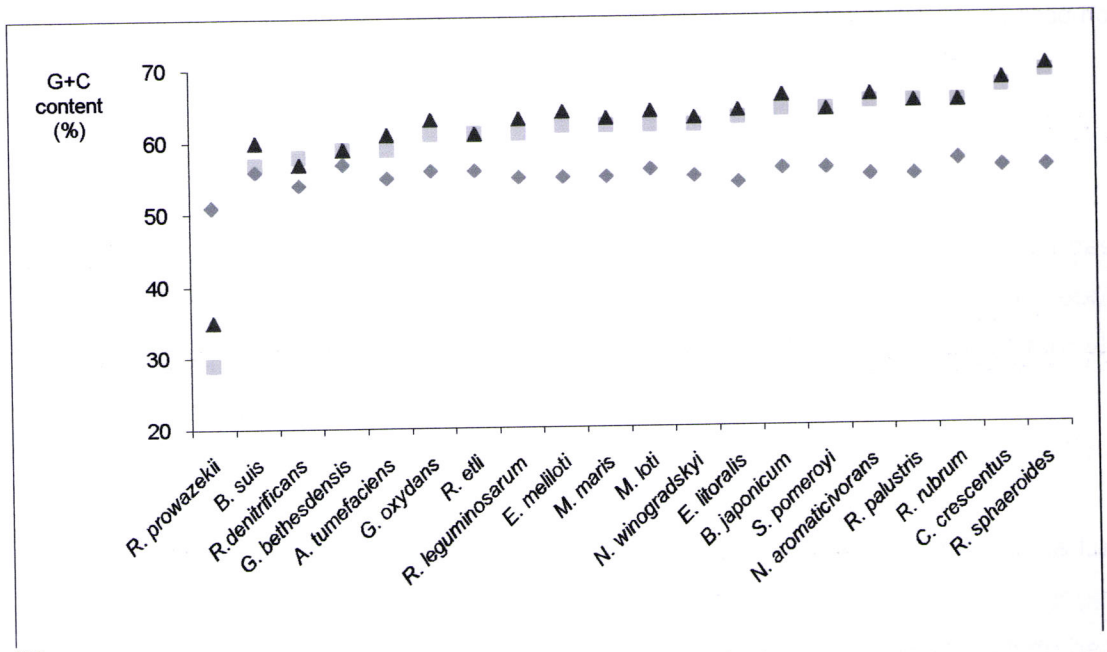
Novosphingobium aromaticivorans	DSM 12444[I]	NC_007794	3917048	3917700	3917701					
Rhizobium etli	CFN42[I]	NC_007761	3893174	3890921	3890922	AJ294404	3891389	3892150	AJ294375	3892587
R. leguminosarum bv. *Viciae*	3841	NC_008380	4403801	4402721	4402722					
Rhodobacter sphaeroides	2.4.1[I]	NC_007493	3718805	3718165	3718166					
Rhodopseudomonas palustris	CGA009	NC_005296	2690886	2691660	2689800					
Rhodospirillum rubrum	ATCC 11170[I]	NC_007643	3833695	3835440	3837010					
Rickettsia prowazekii	Madrid E	NC_000963	883924	883882	883879					
Roseobacter denitrificans	Och 114[I]	NC_008209	4198019	4198416	4198415					
Silicibacter pomeroyi	DSS-3[I]	NC_003911	3196347	3196511	3195579					
Betaproteobacteria										
Azoarcus sp.	EbN1	NC_006513	3180819	3179956						
Bordetella pertussis	Tohama I	NC_002929	3131280	2666530						
Burkholderia ambifaria	AMMD[I]	NC_008390	4308934	4309294						
Nitrosospira multiformis	ATCC25196[I]	NC_007614	3784866	3784939						
Ralstonia solanacearum	GMI1000	NC_003295	1220256	1221481						
Gammaproteobacteria										
Aeromonas hydrophila subsp. *hydrophila*	ATCC7966[T]	NC_008570	4488614	4488918						
Escherichia coli	K12	NC_000913	948332	944753						
Legionella pneumophila	Paris	NC_006368	3116667	3118141						
Pseudomonas aeruginosa	PAO1	NC_002516	3240252	881760						
Vibrio cholerae O1 bv. *eltor*	N16961	NC_002505	2614447	2614523						
Deltaproteobacteria										
Geobacter sulfurreducens	PCA[I]	NC_002939	2685615	2688568						
Pelobacter carbinolicus	DSM 2380[I]	NC_007498	3725044	3722901						
Epsilonproteobacteria										
Campylobacter jejuni	RM1221	NC_003912	3232519	3231902						
Helicobacter pylori	26695	NC_000915	899682	898772						

Accession numbers for sequences resulting from this study are shown in bold.

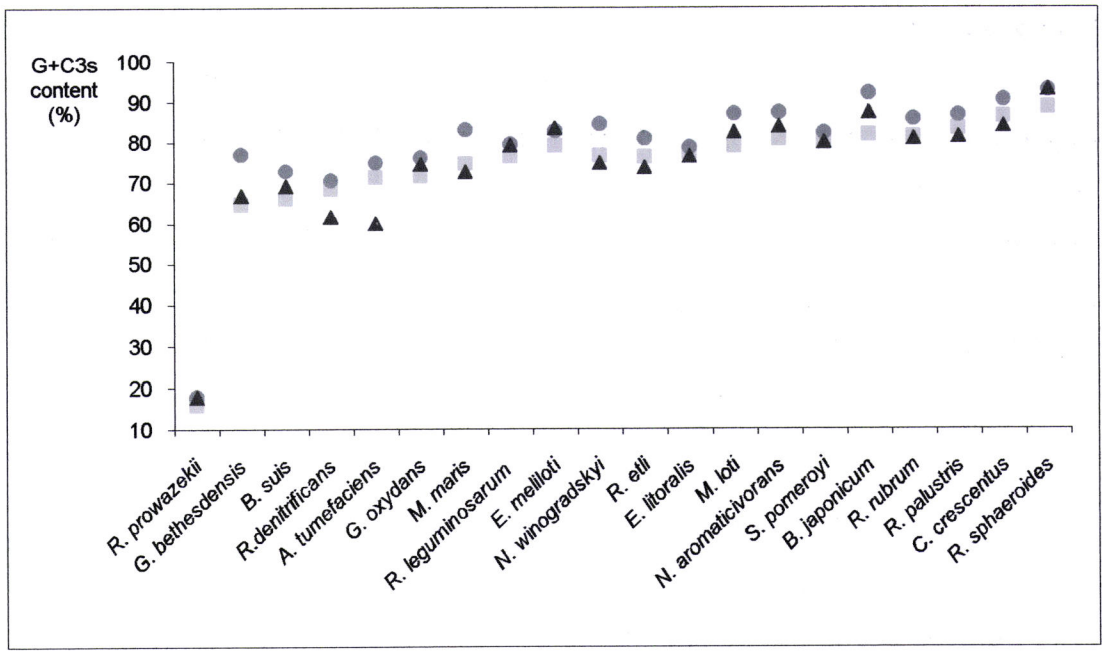
Prior to the use of *dnaJ* as a phylogenetic marker, analysis of G+C content and codon usage frequency was performed, in order to detect possible horizontal gene transfer (HGT) events that would compromise the phylogenetic analysis. The G+C content of the *dnaJ* gene was compared with the G+C content of the entire chromosome, for the alpha- (Fig. 3.1(A)) and other proteobacteria used in this study (data not shown). *dnaJ* G+C content values are similar to those of the chromosome, while the 16S rRNA gene G+C content shows little variation across species (ranging from 50% to 57%), regardless of the higher or lower G+C content of the chromosome (ranging from 29% to 69%). For example, in *Rickettsia prowazekii*, the 16S rRNA gene G+C content (51%) is clearly higher than that of the chromosome (29%), whereas the G+C content of the *dnaJ* gene (36%) is closer to the chromosomal value. Core genes, such as *recA* or *dnaK* present a G+C content similar to that of the whole chromosome.

In order to detect *dnaJ* recently acquired by HGT, the G+C3s content was also analysed (Sharp *et al.*, 2005). The G+C3s of the chromosome shows the average

G+C3s of the chromosomal protein coding genes. Highly expressed genes display higher codon biases than the average, as is the case of genes coding for some chaperones, such as *dnaK* (Karlin *et al.*, 2004). A gene with G+C3s content significantly different from the average of the chromosome and from that of highly expressed genes is probably a recently acquired gene (Karlin *et al.*, 2004). In the alphaproteobacteria (Fig. 3.1(B)), the *dnaJ* gene has, in most cases, a G+C3s content between that of chromosome and that of *dnaK* (a highly expressed gene). The only two species that show a G+C3s for *dnaJ* clearly below the average of the chromosome genes are *Agrobacterium tumefaciens* and *Roseobacter denitrificans*. Therefore, *dnaJ* from these two species could have been acquired by HGT. For the remaining proteobacteria (data not shown), the G+C3s values for *dnaJ* are often very similar to those of the corresponding chromosome and in some cases similar to the value of *dnaK* (for example in *Burkholderia ambifaria*), so no evidence for HGT of *dnaJ* gene was found.



(A)



(B)

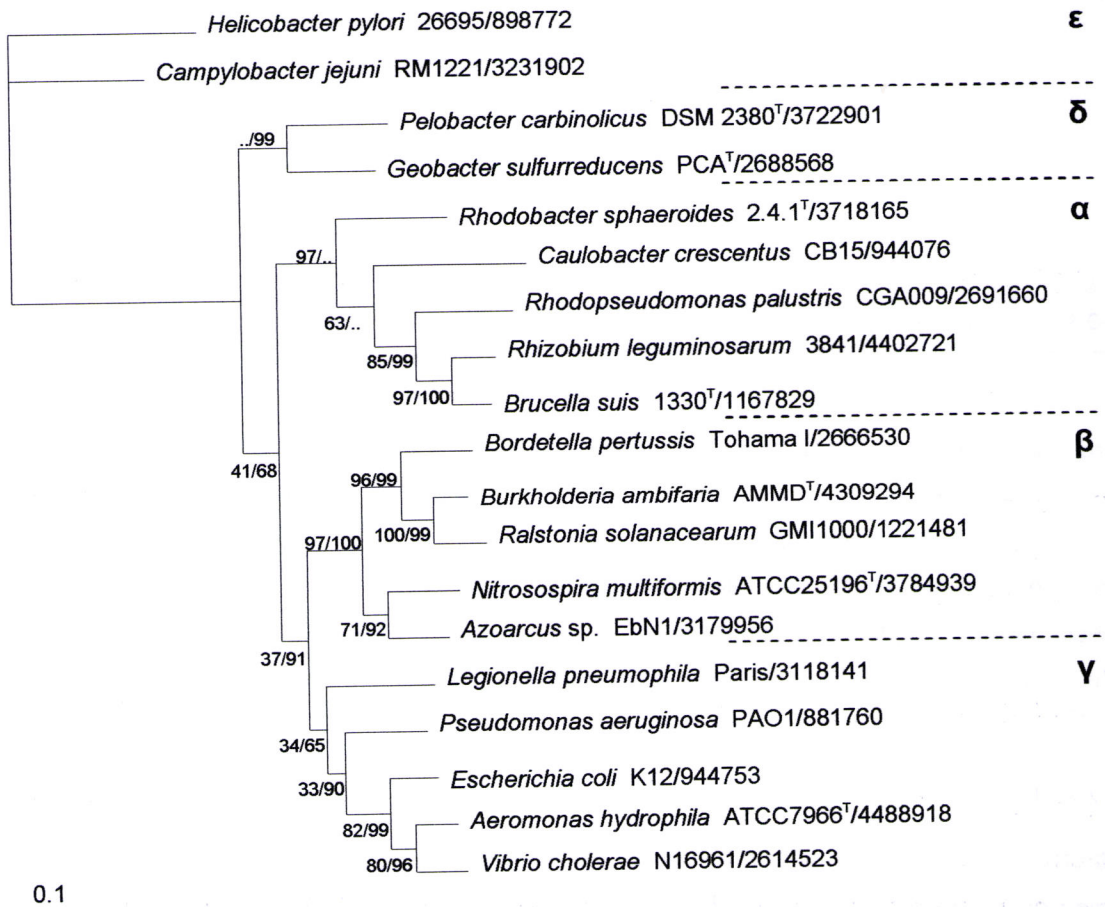
Figure 3.1 - Graphical representation of G+C content (%) for alphaproteobacteria. Total G+C content (chromosome ■ , 16S rRNA gene ◆ and *dnaJ* gene ▲) (A). G+C content of the third codon position (chromosome ■, *dnaK* gene ● and *dnaJ* gene▲) (B).

Another approach to investigate evidence of HGT is to compare the codon usage table of each organism with the codon usage frequency of a given gene. If the *dnaJ* gene includes a high number of codons that are uncommonly used by the bacterium, it is likely that this gene was acquired by HGT. The codon usage analysis did not reveal any discrepancies that could clearly indicate HGT events (data not shown). Nonetheless, the *dnaJ* gene of *A. tumefaciens* shows the highest percentage of uncommonly used codons. This discrepancy on codon usage together with the G+C3s result arise doubts about the origin of the *dnaJ* gene in *A. tumefaciens*, suggesting that it might have been acquired by HGT. Thus, the *A. tumefaciens* position in the *dnaJ*-based phylogeny should be regarded with caution.

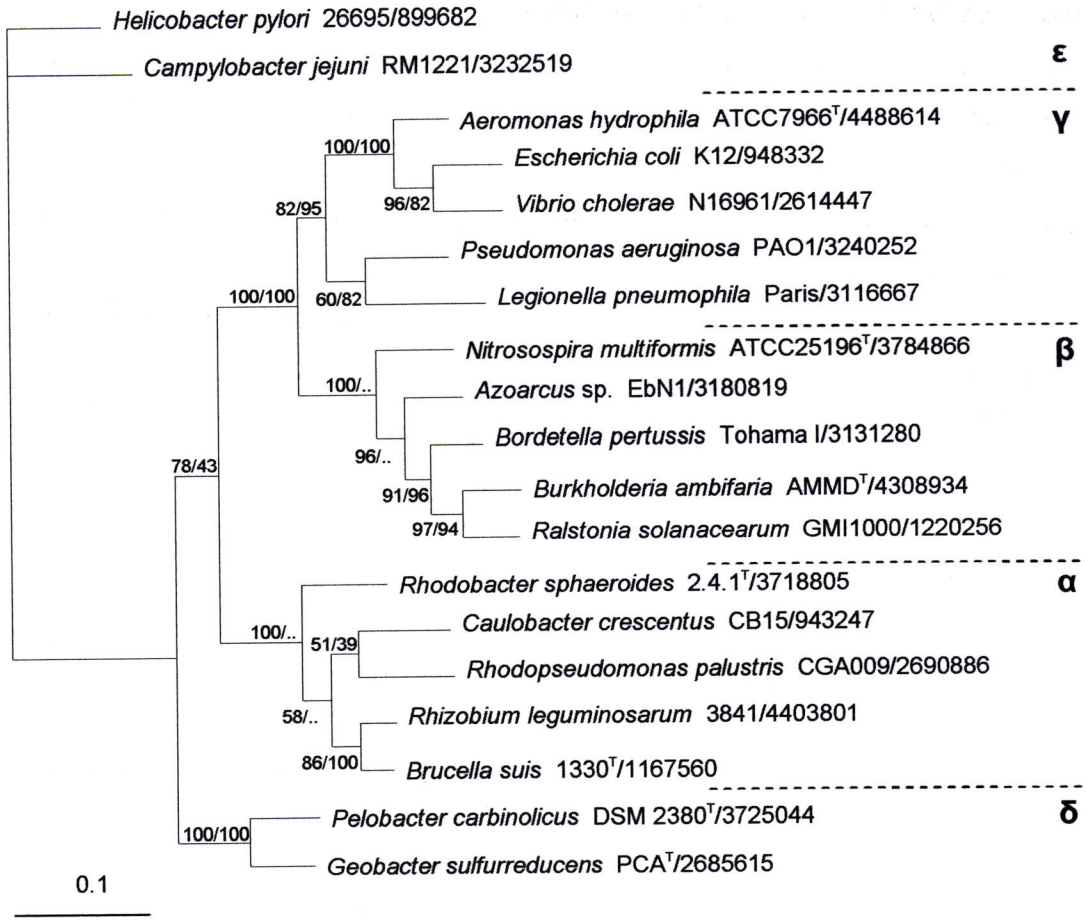
Phylogenetic analyses were performed using both maximum likelihood (ML) and neighbour-joining (NJ) methods. In general, the two methods generated trees with the same topology.

The *Proteobacteria*

In this study, the *Proteobacteria* are represented by 19 species with completely sequenced genomes. The topology of the proteobacterial maximum likelihood tree obtained with DnaJ amino acid sequences (Fig. 3.2(A)) is identical to the one based on 16S rRNA gene sequences (Fig. 3.2(B)). The five classes of *Proteobacteria* are well defined and their branching order is the same in both trees: *Betaproteobacteria* are closer to *Gammaproteobacteria* and these two classes group first with *Alpha-*, then with *Delta-* and finally with *Epsilonproteobacteria*.



(A)



(B)

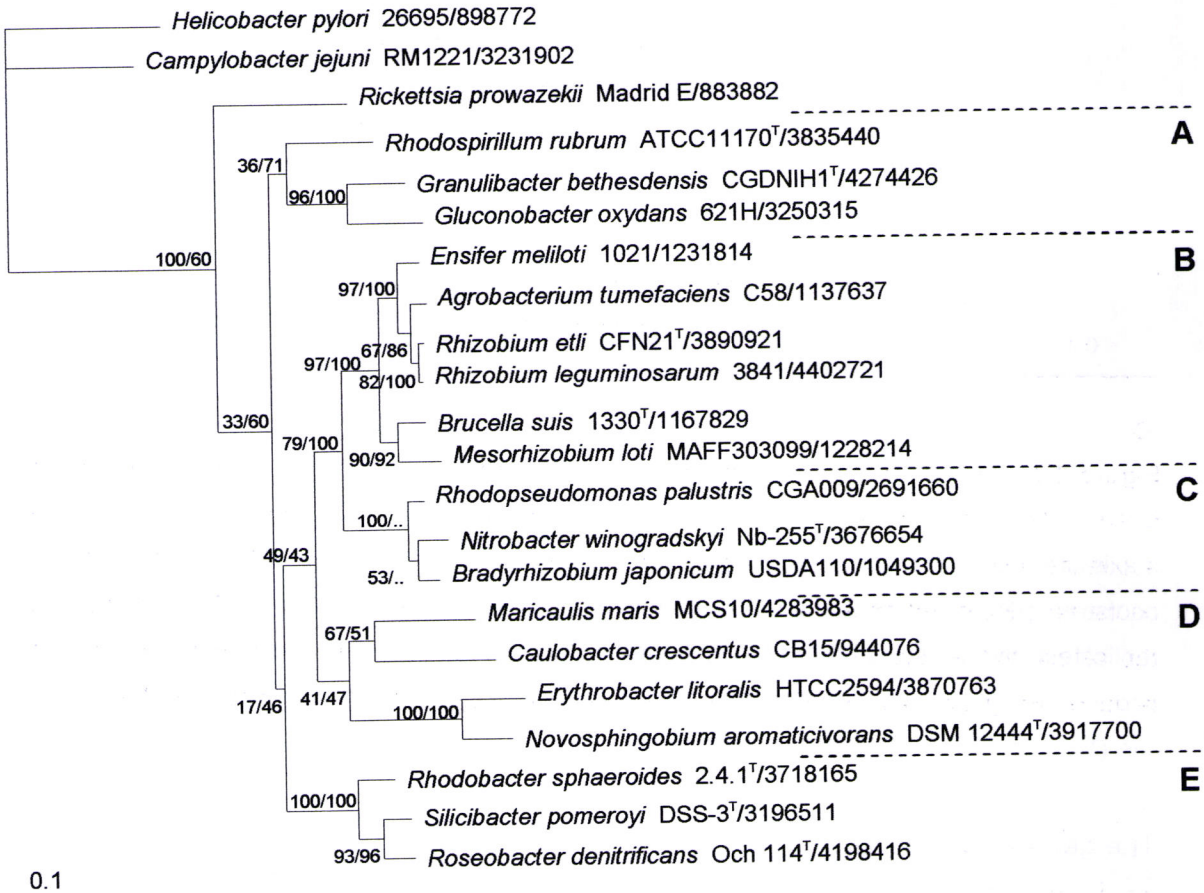
Figure 3.2 - Phylogeny of 19 members of the *Proteobacteria* from all classes (indicated by Greek letters) based on analysis of the DnaJ amino acid sequence (A) and the 16S rRNA gene sequence (B). Trees were generated by maximum likelihood (ML). The first bootstrap percentage indicated on internal branches corresponds to the ML method (100 replicates) and the second to the neighbour-joining method (1000 replicates); dots indicate that nodes were not resolved using that method. Scale bar indicates the number of substitutions per site (ML).

The *Alphaproteobacteria*

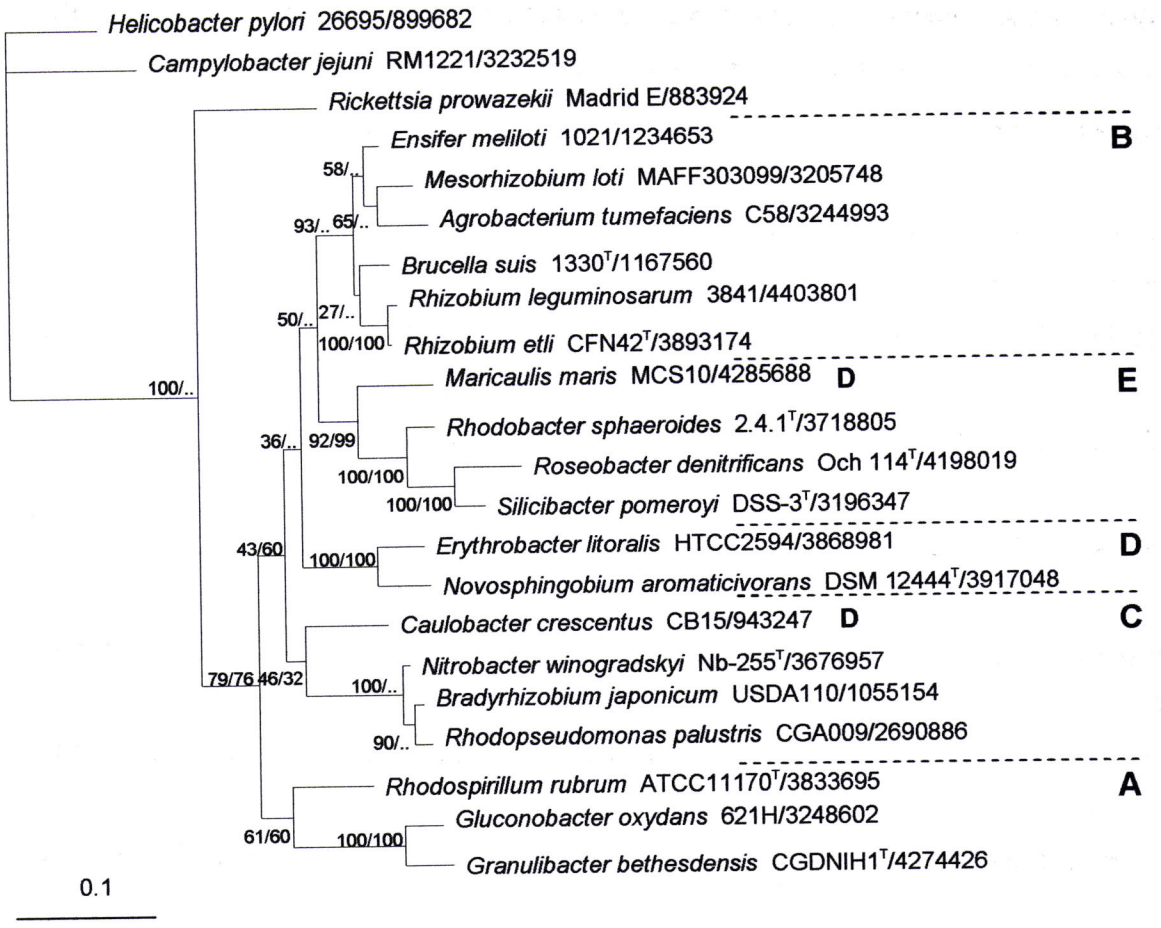
To study the phylogenetic signal of the *dnaJ* gene at a lower taxonomic level, 20 species of *Alphaproteobacteria* with completely sequenced genomes were used.

Both DnaJ and 16S rRNA gene phylogenies (Fig. 3.3) show *Rhodospirillum rubrum*, *Gluconobacter oxydans* and *Granulibacter bethesdensis* (cluster A) and *Rickettsia prowazekii* in a distant position from the other clusters (the letters denote the clusters generated in the DnaJ phylogeny). However, clusters B, C, D and E are differently related in the two phylogenies. In DnaJ phylogeny, cluster B (*Agrobacterium tumefaciens*, *Brucella suis*, *Ensifer meliloti*, *Mesorhizobium loti*, *Rhizobium etli* and *Rhizobium leguminosarum*) is closely related with cluster C (*Bradyrhizobium*

japonicum, *Nitrobacter winogradskyi* and *Rhodopseudomonas palustris*) and then these clusters group first with cluster D (*Caulobacter crescentus*, *Erythrobacter litoralis*, *Maricaulis maris* and *Novosphingobium aromaticivorans*) and finally with cluster E (*Rhodobacter sphaeroides*, *Roseobacter denitrificans* and *Silicibacter pomeroyi*). Contrary to the DnaJ-based phylogeny, the 16S rRNA gene phylogeny shows cluster B closer to cluster E, then these two clusters group with two species of cluster D and finally with cluster C. Furthermore, in the 16S rRNA gene phylogeny, cluster D from the DnaJ phylogeny is dissolved and *M. maris* groups with cluster E while *C. crescentus* groups with cluster C species.



(A)



(B)

Figure 3.3 - Phylogeny of 20 members of the *Alphaproteobacteria*, based on analysis of the DnaJ amino acid sequence (A) and 16S rRNA gene sequence (B). Trees were generated by maximum likelihood (ML). Two *Epsilonproteobacteria* species were used as outgroup. The first bootstrap percentage indicated on internal branches corresponds to the ML method (100 replicates) and the second to the neighbour-joining method (1000 replicates); dots indicate that nodes were not resolved using that method. Scale bar indicates the number of substitutions per site (ML). Letters denote the clusters generated in the DnaJ phylogeny.

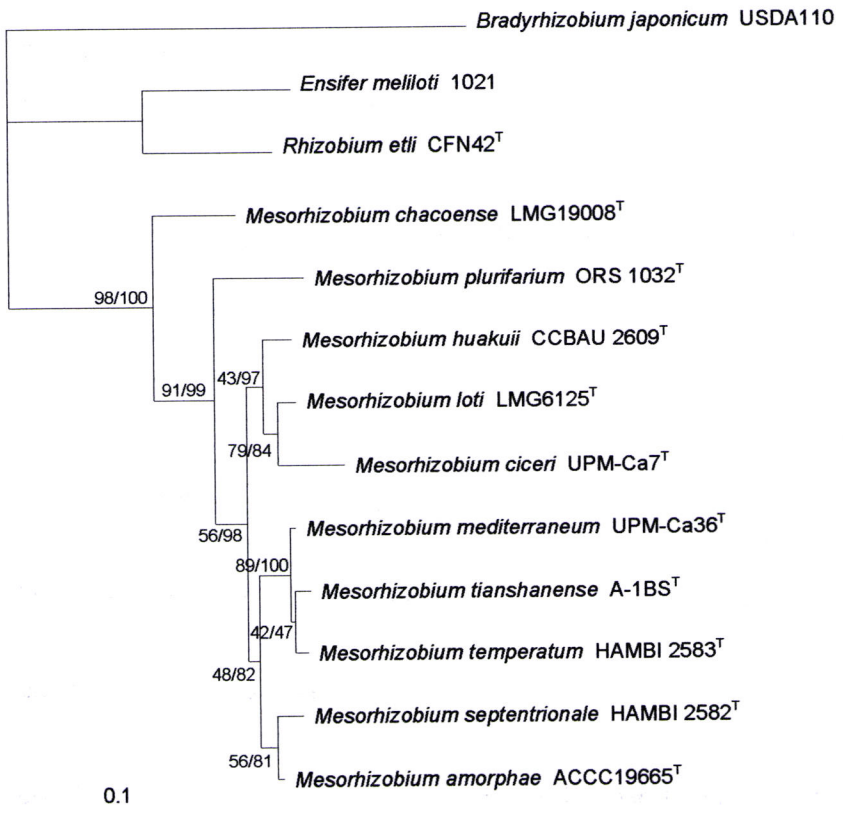
The genus *Mesorhizobium*

In order to obtain a reliable phylogenetic tree for the genus *Mesorhizobium* that could be compared to the *dnaJ*- and 16S rRNA gene-based phylogeny, five core genes (*atpD*, *gapA*, *gyrB*, *recA* and *rplB*) from ten mesorhizobia type strains were partially sequenced. The ILD test was applied to find out which genes could be combined and the results allow the concatenation of the amino acid sequences of all genes. The ML tree based on the concatenated alignment of the five genes (approximately 770 amino acids long), shown in Fig. 3.4(A), reveals the putative organismal phylogeny among *Mesorhizobium* species. Both 16S rRNA gene (Fig. 3.4(C)) tree and DnaJ tree (Fig.

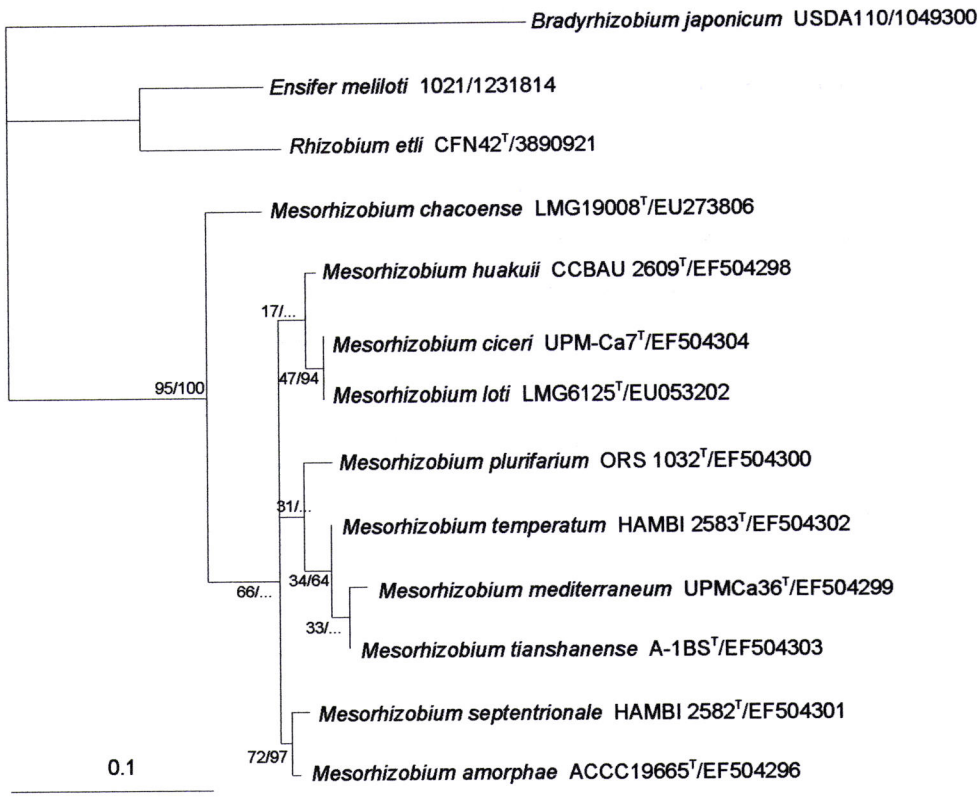
3.4(B)), were compared to the hypothetical organismal phylogeny to evaluate the accuracy of the single gene's phylogenetic signal.

The 16S rRNA gene tree shows a different topology from the multilocus tree. According to the multilocus analysis, *M. huakuii* groups with *M. ciceri* and *M. loti* in the deeper branching of the dendrogram. Contrary to this, the 16S rRNA gene-based phylogeny shows *M. ciceri* and *M. loti* apart from the remaining eight mesorhizobia type species. The concatenated tree shows *M. chacoense* and *M. plurifarium* as the most divergent species. The only two groups generated in both trees are *M. mediterraneum* / *M. temperatum* / *M. tianshanense* and *M. amorphae* / *M. septentrionale*. Among the single gene trees, DnaJ (Fig. 3.4(B)) and GyrB trees (data not shown) show the most consistent topology with the multilocus analysis, although they have some low bootstrap values.

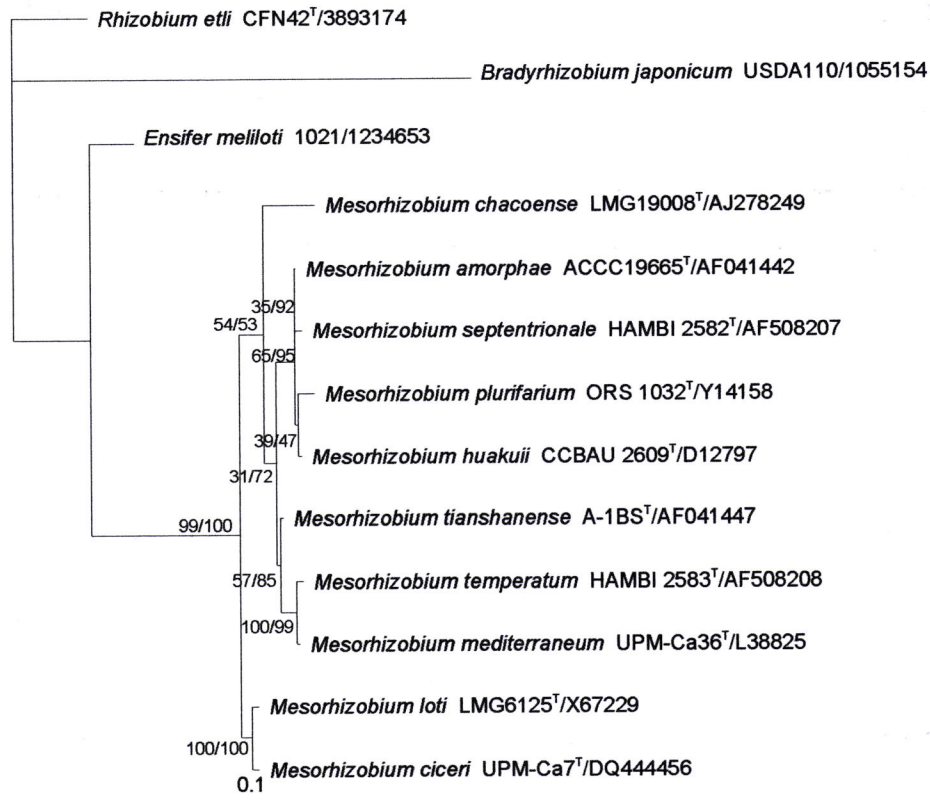
In order to evaluate the suitableness of *dnaJ* sequences in identifying native chickpea rhizobia isolates, a *dnaJ*-based tree including nine isolates was generated (Fig. 3.4(D)). At this taxonomic level, nucleotide sequences were used, since the amino acid based tree showed low resolution and low bootstrap support (data not shown). Isolates grouped closer to the same *Mesorhizobium* type species that in the 16S rRNA-based phylogeny (chapter 2).



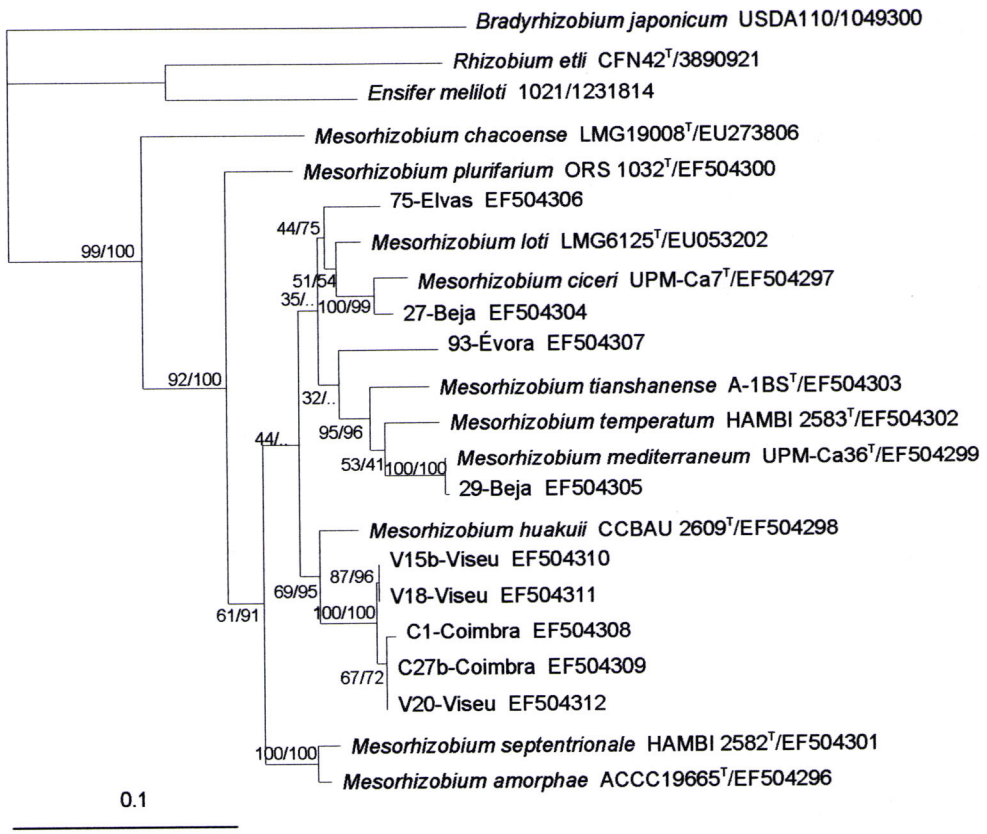
(A)


Bradyrhizobium japonicum USDA110/1049300
Ensifer meliloti 1021/1231814
Rhizobium etli CFN42T/3890921
Mesorhizobium chacoense LMG19008T/EU273806
Mesorhizobium huakuii CCBAU 2609T/EF504298
17/...
Mesorhizobium ciceri UPM-Ca7T/EF504304
47/94
95/100
Mesorhizobium loti LMG6125T/EU053202
Mesorhizobium plurifarium ORS 1032T/EF504300
31/...
Mesorhizobium temperatum HAMBI 2583T/EF504302
34/64
66/...
Mesorhizobium mediterraneum UPMCa36T/EF504299
33/...
Mesorhizobium tianshanense A-1BST/EF504303
Mesorhizobium septentrionale HAMBI 2582T/EF504301
72/97
0.1
Mesorhizobium amorphae ACCC19665T/EF504296

(B)


Rhizobium etli CFN42T/3893174
Bradyrhizobium japonicum USDA110/1055154
Ensifer meliloti 1021/1234653
Mesorhizobium chacoense LMG19008T/AJ278249
Mesorhizobium amorphae ACCC19665T/AF041442
35/92
54/53
Mesorhizobium septentrionale HAMBI 2582T/AF508207
65/95
Mesorhizobium plurifarium ORS 1032T/Y14158
39/47
Mesorhizobium huakuii CCBAU 2609T/D12797
31/72
99/100
Mesorhizobium tianshanense A-1BST/AF041447
57/85
Mesorhizobium temperatum HAMBI 2583T/AF508208
100/99
Mesorhizobium mediterraneum UPM-Ca36T/L38825
Mesorhizobium loti LMG6125T/X67229
100/100
Mesorhizobium ciceri UPM-Ca7T/DQ444456
0.1

(C)



(D)

Figure 3.4 - Phylogeny based on the concatenated amino acid sequence alignment of AtpD-GapA-GyrB-RecA-RplB (A) DnaJ (B) and on 16S rRNA gene sequence (C) for the genus *Mesorhizobium*. Phylogeny based on the nucleotide sequences of *dnaJ* for the genus *Mesorhizobium* and chickpea rhizobial isolates (D). The trees were generated by maximum likelihood (ML). The first bootstrap percentage indicated on internal branches corresponds to the ML method (100 replicates) and the second value corresponds to the neighbour-joining method (1000 replicates). Scale bars indicate the number of substitutions per site (ML).

3.5 - Discussion

In all proteobacteria used in the present study, the *dnaJ* gene was found to be encoded adjacent to *dnaK*, except in epsilonproteobacteria, for which no *dnaKJ* operon was found among the 11 completely sequenced genomes. In the chromosome of each of the proteobacteria used, only one *dnaJ* gene copy was found that had the expected size and included the four characteristic domains of this gene. Other orthologs annotated as *dnaJ*, often dispersed in the chromosome, were found to lack at least one of the characteristic *dnaJ* domains. In two species (*Ensifer meliloti* and *Rhizobium etli*),

a *dnaJ* homolog was found to be encoded on a plasmid, however these homologs were partial *dnaJ* sequences.

No evidence of HGT events was provided by the analysis of G+C content and codon usage frequency for the alphaproteobacteria and the other proteobacteria used. However, the G+C3s content analysis revealed that two species (*Agrobacterium tumefaciens* and *Roseobacter denitrificans*) should be regarded with caution in the *dnaJ* phylogeny. For the remaining species, the possibility of *dnaJ* HGT occurring between organisms with similar G+C content cannot be excluded.

In general, our analysis on G+C and G+C3s content (Fig. 3.1) suggests that *dnaJ* is a core gene and is not commonly subject to HGT between species. These features are important in a good phylogenetic marker, so we used it for phylogenetic analysis of *Proteobacteria*, *Alphaproteobacteria* and *Mesorhizobium*.

The *Proteobacteria*

The branching order of all *Proteobacteria* classes, in both DnaJ and 16S rRNA gene phylogenies (Fig. 3.2), is in agreement with several recent studies that combine data from a large number of protein coding genes, although some different species were used (Fukami-Kobayashi *et al.*, 2007, Ciccarelli *et al.*, 2006, Gupta & Sneath, 2007). It is noteworthy that some of these broader phylogenies showed that *Delta-* and *Epsilonproteobacteria* may form one group, apart from the remaining classes of *Proteobacteria* (Fukami-Kobayashi *et al.*, 2007, Ciccarelli *et al.*, 2006, Gupta, 2000).

Although the same branching order of *Proteobacteria* classes is generated in both phylogenies, the relationships between species within each cluster are different in a few cases. For example, relationships among the gammaproteobacteria *Aeromonas hydrophila, Vibrio cholerae* and *Escherichia coli* in the 16S rRNA gene-based tree are consistent with the phylogeny shown in a previous study using 31 concatenated protein sequences and comprising a large group of gammaproteobacteria (Seshadri *et al.*, 2006), but the DnaJ tree shows a different relationship.

Despite such minor discrepancies, the global congruence found between the 16S rRNA gene and the DnaJ phylogeny and the agreement with other phylogenies based on multilocus data, indicate a good performance of *dnaJ* in reconstructing proteobacteria phylogeny.

The *Alphaproteobacteria*

In contrast to the 16S rRNA gene phylogeny, the phylogeny based on DnaJ sequences (Fig. 3.3) is in agreement with the currently accepted relationships among genera and species within this group of bacteria obtained in other studies based on large amount of sequence data (Gupta, 2005, Ciccarelli *et al.*, 2006, Gupta & Sneath, 2007).

The relationships among different species inferred from DnaJ sequence comparisons are consistent with those inferred in previous studies, namely the phylogeny based on concatenated sequences of 10 proteins (Gupta & Sneath, 2007) and the tree of life generated from the concatenated alignment of 31 universal protein families (Ciccarelli *et al.*, 2006). In the DnaJ-based phylogeny, all *Rhizobiales* cluster together (branch B+C), thus *Bradyrhizobium japonicum* (cluster C) is closer to *Mesorhizobium loti* (cluster B) than to *Caulobacter crescentus* (*Caulobacterales* in cluster D), as reported before (Gupta & Sneath, 2007, Ciccarelli *et al.*, 2006). On the contrary, the 16S rRNA gene-based phylogeny shows *B. japonicum* (cluster C) close to *C. crescentus* (cluster D) and distant from *M. loti* (cluster B). Moreover, within cluster B, the DnaJ-based phylogeny shows *Ensifer meliloti* closer to *Agrobacterium tumefaciens* and *Mesorhizobium loti* closer to *Brucella suis*, which is in full agreement with the phylogeny proposed by Young and co-workers (2006) based on the concatenated sequences of 638 proteins (the present study uses *B. suis* instead of *B. melitensis*) and with the multilocus tree of life presented by Ciccarelli *et al.* (2006), and not concordant with the 16S rRNA gene phylogeny. Therefore, the possible HGT origin of *dnaJ* from *A. tumefaciens* (cluster B) suggested by the G+C content analysis was not confirmed.

The coherence between our results using *dnaJ* and those from broader phylogenies using multiple protein coding genes (Ciccarelli *et al.*, 2006, Gupta & Sneath, 2007), shows that *dnaJ* has a better phylogenetic signal than the 16S rRNA gene in reconstructing the phylogeny of alphaproteobacteria.

The genus *Mesorhizobium*

The high level of sequence conservation of the 16S rRNA gene represents a limitation in the use of this gene for closely related bacteria. In the genus *Mesorhizobium*, other genes have been used for phylogenetic purposes, such as *dnaK* (Stepkowski *et al.*, 2003), *atpD* and *recA* (Vinuesa *et al.*, 2005). The non-coding 16S-23S rRNA intergenic spacer (ITS) has also been used (Rivas *et al.*, 2007). Relationships between mesorhizobia species are different depending on the gene used. Due to this unclear

positioning and to the lack of a multilocus analysis focusing on mesorhizobia, there is no generally accepted phylogeny of this genus.

The present multilocus phylogenetic analysis may contribute to clarify the phylogenetic relationships among *Mesorhizobium* species (Fig. 3.4(A)), namely the proximity of *M. ciceri*, *M. loti* and *M. huakuii*. Interestingly, analysis of phenotypic data from a large set of chickpea mesorhizobia isolates also supported a closer relationship between *M. loti* / *M. ciceri* and *M. huakuii* isolates (Alexandre *et al.*, 2006).

The similarity between the DnaJ based tree and the putative *Mesorhizobium* phylogeny derived from the multilocus analysis performed in this study, suggests that this gene is not commonly subject to HGT between *Mesorhizobium* species.

Although the relative positioning of some type species is different from the ones in the concatenated tree and 16S rRNA gene-based tree, identification of isolates based on *dnaJ* sequences (Fig. 3.4(D)) is consistent with the identification based on 16S rRNA gene (chapter 2). The present phylogenetic analysis of native rhizobia confirmed a high diversity of species able to nodulate chickpea, namely isolates close to *M. ciceri*, *M.huakuii*, *M.loti*, *M. mediterraneum*, *M. temperatum* and *M. tianshanense*, consistent with previous analysis on chickpea mesorhizobia diversity (chapter 2) (Laranjo *et al.*, 2004, Rivas *et al.*, 2007).

For decades, 16S rRNA has been the most ubiquitous gene used for accessing bacterial phylogeny, but this scenario has been changing. Many other core genes are currently used as phylogenetic markers in order to infer species phylogeny and evolution. Bacterial *dnaJ* is a core gene, as confirmed by its chromosomal location and a G+C content similar to that of the chromosome. This gene is present as a single "true" copy usually in the same operon as *dnaK*, which allows its specific amplification by PCR. No evidence that *dnaJ* had been subject to a HGT event was found in the proteobacteria that were examined, and the phylogenetic signal of this gene was consistent with studies based on multiple protein sequences. In contrast to 16S rRNA gene trees, the alphaproteobacterial DnaJ phylogenetic tree showed a global topology resembling the presumed organismal phylogeny inferred from multiple protein sequences from complete genomes. For these reasons, we suggest that a single core gene, such as *dnaJ*, can be used as a phylogenetic marker for proteobacteria at the level of phylum (*Proteobacteria*) and class (*Alphaproteobacteria*). At the genus *Mesorhizobium* level, *dnaJ* gene can be used for identification of isolates, with more resolution than the 16S rRNA gene. This study highlights the usefulness of the *dnaJ* gene as a single alternative phylogenetic marker for alphaproteobacteria.

3.6 - References

Acinas, S. G., L. A. Marcelino, V. Klepac-Ceraj & M. F. Polz, (2004) Divergence and Redundancy of 16S rRNA Sequences in Genomes with Multiple rrn Operons. *J. Bacteriol.* **186**: 2629-2635.

Alexandre, A., M. Laranjo & S. Oliveira, (2006) Natural populations of chickpea rhizobia evaluated by antibiotic resistance profiles and molecular methods. *Microb. Ecol.* **51**: 128-136.

Baldo, L., S. Bordenstein, J. J. Wernegreen & J. H. Werren, (2006) Widespread recombination throughout *Wolbachia* genomes. *Mol. Biol. Evol.* **23**: 437-449.

Bull, J. J., J. P. Huelsenbeck, C. W. Cunningham, D. L. Swofford & P. J. Waddell, (1993) Partitioning and Combining Data in Phylogenetic Analysis. *Syst. Biol.* **42**: 384-397.

Ciccarelli, F. D., T. Doerks, C. von Mering, C. J. Creevey, B. Snel & P. Bork, (2006) Toward automatic reconstruction of a highly resolved tree of life. *Science* **311**: 1283-1287.

Cilia, V., B. Lafay & R. Christen, (1996) Sequence heterogeneities among 16S ribosomal RNA sequences, and their effect on phylogenetic analyses at the species level. *Mol. Biol. Evol.* **13**: 451-461.

Cunningham, C., (1997) Can three incongruence tests predict when data should be combined? *Mol. Biol. Evol.* **14**: 733-740.

Eisen, J. A., (1995) The RecA protein as a model molecule for molecular systematic studies of bacteria: Comparison of trees of RecAs and 16S rRNAs from the same species. *J. Mol. Evol.* **41**: 1105-1123.

Eisen, J. A., (2000) Horizontal gene transfer among microbial genomes: new insights from complete genome analysis. *Curr. Opin. Genet. Dev.* **10**: 606-611.

Felsenstein, J., (2006) PHYLIP (Phylogeny Inference Package). In. Seattle: Distributed by the author, pp.

Fukami-Kobayashi, K., Y. Minezaki, Y. Tateno & K. Nishikawa, (2007) A Tree of Life Based on Protein Domain Organizations. *Mol. Biol. Evol.* **24**: 1181-1189.

Garrity, G. M. & T. Lilburn, (2005) Phylum XIV. Proteobacteria phyl. nov. In: Bergeys Manual of Systematic Bacteriology. New York: Springer

pp. p. 1.

Gaunt, M. W., S. L. Turner, L. Rigottier-Gois, S. A. Lloyd-Macgilp & J. P. W. Young, (2001) Phylogenies of *atpD* and *recA* support the small subunit rRNA-based classification of rhizobia. *Int. J. Syst. Evol. Microbiol.* **51**: 2037-2048.

Gupta, R. S., (2000) The phylogeny of proteobacteria: relationships to other eubacterial phyla and eukaryotes. *Fems Microbiol. Rev.* **24**: 367-402.

Gupta, R. S., (2005) Protein Signatures Distinctive of Alpha Proteobacteria and Its Subgroups and a Model for Alpha-Proteobacterial Evolution. *Crit. Rev. Microbiol.* **31**: 101-135.

Gupta, R. S. & P. H. A. Sneath, (2007) Application of the Character Compatibility Approach to Generalized Molecular Sequence Data: Branching Order of the Proteobacterial Subdivisions. *J. Mol. Evol.* **64**: 90-100.

Hall, T. A., (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* **41**: 95-98.

Itoh, Y., Y. Kawamura, H. Kasai, M. M. Shah, P. H. Nhung, M. Yamada, X. S. Sun, T. Koyana, M. Hayashi, K. Ohkusu & T. Ezaki, (2006) *dnaJ* and *gyrB* gene sequence relationship among species and strains of genus *Streptococcus*. *Syst. Appl. Microbiol.* **29**: 368-374.

Karlin, S., J. Theriot & J. Mrazek, (2004) Comparative analysis of gene expression among low G+C gram-positive genomes. *Proc. Natl. Acad. Sci. U. S. A.* **101**: 6182-6187.

Kersters, K., P. Devos, M. Gillis, P. Vandamme & E. Stackebrandt, (2007) Introduction to the Proteobacteria. In: The Prokaryotes: An Evolving Electronic Resource for the Microbiological Community. M. Dworkin, S. Falkow, E. Rosenberg, K.-H. Schleifer & E. Stackebrandt (eds). New York: Springer-Verlag, pp. 3-37.

Kimura, M., (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* **16**: 111-120.

Kumar, S., K. Tamura & M. Nei, (2004) MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. *Briefings in Bioinformatics* **5**: 150-163.

Laranjo, M., J. Machado, J. P. W. Young & S. Oliveira, (2004) High diversity of chickpea *Mesorhizobium* species isolated in a Portuguese agricultural region. *FEMS Microbiol. Ecol.* **48**: 101-107.

Liu, H., Y. Li, X. Huang, Y. Kawamura & T. Ezaki, (2003) Use of the *dnaJ* gene for the detection and identification of all *Legionella pneumophila* serogroups and description of the primers used to detect 16S rDNA gene sequences of major members of the genus *Legionella*. *Microbiol. Immunol.* **47**: 859-869.

Lund, P. A., (2009) Multiple chaperonins in bacteria - why so many? *Fems Microbiol. Rev.* **33**: 785-800.

Morita, Y., S. Maruyama, H. Kabeya, A. Nagai, K. Kozawa, M. Kato, T. Nakajima, T. Mikami, Y. Katsube & H. Kimura, (2004) Genetic diversity of the *dnaJ* gene in the *Mycobacterium avium* complex. *J. Med. Microbiol.* **53**: 813-817.

Nei, M. & S. Kumar, (2000) *Molecular Evolution and Phylogenetics*. Oxford University Press, New York.

Nhung, P. H., M. M. Shah, K. Ohkusu, M. Noda, H. Hata, X. S. Sun, H. Iihara, K. Goto, T. Masaki, J. Miyasaka & T. Ezaki, (2007) The *dnaJ* gene as a novel phylogenetic marker for identification of *Vibrio* species. *Syst. Appl. Microbiol.* **30**: 309-315.

Olsen, G. J. & C. R. Woese, (1993) Ribosomal RNA: a key to phylogeny. *FASEB J.* **7**: 113-123.

Peden, J. F., (1999) Analysis of codon usage. In. Nottingham: University of Nottingham pp. 215.

Rivas, R., M. Laranjo, P. F. Mateos, S. Oliveira, E. Martinez-Molina & E. Velázquez, (2007) Strains of *Mesorhizobium amorphae* and *Mesorhizobium tianshanense*, carrying symbiotic genes of common chickpea endosymbiotic species, constitute a novel biovar (*ciceri*) capable of nodulating *Cicer arietinum*. *Lett. Appl. Microbiol.* **44**: 412-418.

Rivas, R., E. Velázquez, A. Valverde, P. F. Mateos & E. Martínez-Molina, (2001) A two primers random amplified polymorphic DNA procedure to obtain polymerase chain reaction fingerprints of bacterial species. *Electrophoresis* **22**: 1086-1089.

Saito, H. & H. Uchida, (1978) Organization and Expression of DnaJ and DnaK Genes of *Escherichia coli* K12. *Mol. Gen. Genet.* **164**: 1-8.

Saitou, N. & M. Nei, (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**: 406-425.

Seshadri, R., S. W. Joseph, A. K. Chopra, J. Sha, J. Shaw, J. Graf, D. Haft, M. Wu, Q. H. Ren, M. J. Rosovitz, R. Madupu, L. Tallon, M. Kim, S. H. Jin, H. Vuong, O. C. Stine, A. Ali, A. J. Horneman & J. F. Heidelberg, (2006) Genome sequence of *Aeromonas hydrophila* ATCC 7966(T): Jack of all trades. *J. Bacteriol.* **188**: 8272-8282.

Shah, M. M., H. Iihara, M. Noda, S. X. Song, P. H. Nhung, K. Ohkusu, Y. Kawamura & T. Ezaki, (2007) *dnaJ* gene sequence-based assay for species identification and phylogenetic grouping in the genus *Staphylococcus*. *International Journal of Systematic and Evololutionary Microbiology* **57**: 25-30.

Sharp, P. M., E. Bailes, R. J. Grocock, J. F. Peden & R. E. Sockett, (2005) Variation in the strength of selected codon usage bias among bacteria. *Nucleic Acids Res.* **33**: 1141-1153.

Stackebrandt, E., R. G. E. Murray & H. G. Truper, (1988) Proteobacteria-Classis Nov, a Name for the Phylogenetic Taxon That Includes the Purple Bacteria and Their Relatives. *Int. J. Syst. Bacteriol.* **38**: 321-325.

Stepkowski, T., M. Czaplinska, K. Miedzinska & L. Moulin, (2003) The variable part of the *dnaK* gene as an alternative marker for phylogenetic studies of rhizobia and related alpha Proteobacteria. *Syst. Appl. Microbiol.* **26**: 483-494.

Swofford, D. L., (2003) PAUP*: Phylogenetic analysis using parsimony (* and other methods). In. Sunderland, Massachusetts: Sinauer Associates, pp.

Thompson, C. C., F. L. Thompson, K. Vandemeulebroecke, B. Hoste, P. Dawyndt & J. Swings, (2004) Use of *recA* as an alternative phylogenetic marker in the family Vibrionaceae. *Int. J. Syst. Evol. Microbiol.* **54**: 919-924.

Turner, S. L. & J. P. Young, (2000) The glutamine synthetases of rhizobia: phylogenetics and evolutionary implications. *Mol. Biol. Evol.* **17**: 309-319.

Ventura, M., C. Canchaya, V. Bernini, A. Del Casale, F. Dellaglio, E. Neviani, G. F. Fitzgerald & D. van Sinderen, (2005) Genetic Characterization of the *Bifidobacterium breve* UCC 2003 *hrc*A Locus. *Appl. Environ. Microbiol.* **71**: 8998-9007.

Vinuesa, P., C. Silva, M. J. Lorite, M. L. Izaguirre-Mayoral, E. J. Bedmar & E. Martínez-Romero, (2005) Molecular systematics of rhizobia based on maximum likelihood and Bayesian phylogenies inferred from rrs, *atpD*, *recA* and *nifH* sequences, and their use in the classification of *Sesbania* microsymbionts from Venezuelan wetlands. *Syst. Appl. Microbiol.* **28**: 702-716.

Vitorino, L., I. M. Chelo, F. Bacellar & L. Ze-Ze, (2007) *Rickettsiae* phylogeny: a multigenic approach. *Microbiology-Sgm* **153**: 160-168.

Xiao, X., P. Wang, X. Zeng, D. H. Bartlett & F. P. Wang, (2007) *Shewanella psychrophila* sp nov and *Shewanella piezotolerans* sp nov., isolated from west Pacific deep-sea sediment. *Int. J. Syst. Evol. Microbiol.* **57**: 60-65.

Yamamoto, S. & S. Harayama, (1995) PCR amplification and direct sequencing of *gyrB* genes with universal primers and their application to the detection and taxonomic analysis of *Pseudomonas putida* strains. *Applied Environmental Microbiology* **61**: 1104-1109.

Yamamoto, S., H. Kasai, D. L. Arnold, R. W. Jackson, A. Vivian & S. Harayama, (2000) Phylogeny of the genus *Pseudomonas*: intrageneric structure reconstructed from the nucleotide sequences of *gyrB* and *rpoD* genes. *Microbiology* **146**: 2385-2394.

Young, J. P., L. Crossman, A. Johnston, N. Thomson, Z. Ghazoui, K. Hull, M. Wexler, A. Curson, J. Todd, P. Poole, T. Mauchline, A. East, M. Quail, C. Churcher, C. Arrowsmith, I. Cherevach, T. Chillingworth, K. Clarke, A. Cronin, P. Davis, A. Fraser, Z. Hance, H. Hauser, K. Jagels, S. Moule, K. Mungall, H. Norbertczak, E. Rabbinowitsch, M. Sanders, M. Simmonds, S. Whitehead & J. Parkhill, (2006) The genome of *Rhizobium leguminosarum* has recognizable core and accessory components. *Genome Biology* **7**: R34.

4.

Tolerance of chickpea rhizobia isolates to temperature stress

This chapter is based on the manuscript:

Alexandre, A. and Oliveira, S. "Most heat tolerant rhizobia show high induction of major chaperone genes upon stress" Accepted for publication in FEMS Microbiology Ecology.

4.1 - Summary

The rhizobia-legume symbiosis is affected by environmental conditions such as temperature. Our aims were to evaluate the temperature stress tolerance of chickpea rhizobia and to investigate if tolerance is related to isolates species or origin site. Fifty three chickpea mesorhizobia previously isolated from several provinces of Portugal, and assigned to different species, were used. Temperature tolerance was evaluated by growth in liquid medium under cold (15ºC), heat (37ºC) and heat shock stress (48ºC, 15 minutes). Mesohizobia isolates showed high diversity in their ability to grow under temperature stress. Under 15ºC, the isolates growth ranges from 8% to 68%, while growth at 37ºC ranges from 5% to 46%. After the heat shock, isolates grew between 3% and 89%. Most isolates tolerate heat shock or cold stress better than the heat stress. Isolates from distinct species groups differed significantly in their ability to tolerate temperature stress. Isolates from the specific chickpea microsymbionts species groups (*Mesorhizobium ciceri* and *M. mediterraneum*) showed the highest average growth under all conditions studied. An association was found between some provinces of origin and stress tolerance of the isolates. For instance, isolates from the province Trás-os-Montes e Alto Douro are associated with high tolerance to the tested temperatures.

4.2 - Introduction

The demand for more effective utilization of biologically-fixed N in agricultural systems has prompted studies on rhizobia diversity and tolerance to biotic and abiotic factors. Rhizobia populations are affected by soil environmental factors, such as pH, salinity and temperature. Temperature affects almost all stages of legume-rhizobia symbiosis (Hungria & Vargas, 2000, Zahran, 1999). It is important to search for isolates naturally tolerant to temperature stress since heat-resistant laboratory variants are often affected in other characteristics, as for example in their effectiveness (AbdelGadir & Alexander, 1997, Day *et al.*, 1978). Since the early host-rhizobia molecular signals exchange, until the nodules are already established, temperature variations, especially high temperatures, have been shown to influence symbiosis (for review see Hungria & Vargas, 2000, Zahran, 1999). During symbiosis under high temperatures, the nitrogenase activity decreases and the synthesis of leghaemoglobin also declines. Studies with cowpea have shown that effectiveness of rhizobia strains decreased with periods of temperature increasing from 30ºC to 36-44ºC (Day *et al.*, 1978). The influence of temperature stress is very complex and varies with the host plant and rhizobium strain. Recent studies with bradyrhizobia nodulating soybean have shown that the endurance of strains at high temperatures in culture medium correlates with their symbiotic performance under heat stress (Rahmani *et al.*, 2009), as previously reported in other studies (Munevar & Wollum, 1981).

The optimum temperature range for most rhizobia is described to be between 25ºC and 31ºC (Somasegaran & Hoben, 1994) and the upper limits lie between 32ºC and 47ºC (for review see Hungria & Vargas, 2000). Räsänen and co-workers (2001) studied the heat stress effect in a *Sinorhizobium arboris* strain and reported that high temperature can induce rhizobia to enter a viable but non-culturable state (VBNC).

The present study comprised the evaluation of temperature stress tolerance of chickpea rhizobia from the entire Portuguese territory, in order to identify isolates highly tolerant to both high and low temperatures and to investigate a possible relationship between stress tolerance and the species or the site origin of the isolates.

4.3 - Methods

Bacterial strains and culture medium

A total of 53 native chickpea rhizobia strains, isolated from 18 soil samples, covering 10 Portuguese provinces and Madeira Island, were used (Table 4.1). The two chickpea *Mesorhizobium* type strains were also used (*M. ciceri* UPM-Ca7^T and *M.*

mediterraneum UPM-Ca36^T). All strains were routinely grown in yeast extract mannitol (YEM) broth (Vincent, 1970) and stored in YEM agar slants at 4ºC and kept at -80ºC for long-term storage.

Table 4.1 – Rhizobia isolates used in the present study.

Province of origin	Isolate	16S rRNA species cluster[a]
Trás-os-Montes e Alto Douro	BR-8-Bragança	B
	BR-9-Bragança	B
	BR-15-Bragança	B
	BR-16-Bragança	B
	BR-28-Bragança	B
Douro Litoral	PII-1-Porto	B
	PII-2-Porto	B
	PII-3-Porto	B
	PII-4-Porto	B
Beira Litoral	C-1-Coimbra	A
	C-7-Coimbra	A
	C-14-Coimbra	A
	C-27b-Coimbra	A
	L-19-Leiria	A
Beira Alta	V-5b.-Viseu	A
	V-15b.-Viseu	A
	V-18-Viseu	A
	V-20-Viseu	A
	V-25b.-Viseu	A
Beira Baixa	CB-10-Castelo Branco	B
	CB-19-Castelo Branco	B
	CB-30-Castelo Branco	B
	T-3-Telhado	A
	T-8-Telhado	A
Estremadura	CR-3-Caldas da Rainha	C
	CR-16-Caldas da Rainha	C
	CR-18-Caldas da Rainha	C
	CR-32-Caldas da Rainha	C
	ST-2-Setúbal	C
Ribatejo	STR-2-Santarém	A
	STR-4-Santarém	A
	STR-10-Santarém	A
	STR-14-Santarém	C
	STR-16-Santarém	C
Alto Alentejo	78-Elvas	A
	85-Elvas	A
	EE-14-ENMP	D
	102-Évora	A
	PT-35-Portalegre	A
Baixo Alentejo	6b.-Beja	D
	7a.-Beja	B
	27-Beja	B
	29-Beja	D
	64b.-Beja	A
Algarve	PM-1-Portimão	D
	PM-14-Portimão	D
	PM-17-Portimão	D
	PMI-1-Portimãol	B
	PMI-6-Portimãol	A
Madeira	SA-9-Serra d'Água	A
	SA-12-Serra d'Água	A
	SA-13-Serra d'Água	A
	SA-17-Serra d'Água	A

[a] from chapter 2

A – *M huakuii / M. amorphae* species cluster

B – *M. ciceri / M. loti* species cluster

C – *M. tianshanense* species cluster

D- *M.mediterraneum / M. temperatum* species cluster

Evaluation of temperature stress tolerance

Tolerance of chickpea rhizobia isolates to temperature stress was evaluated in three different conditions: heat stress (37ºC), cold stress (15ºC) and heat shock stress (48ºC for 15 minutes). Stress tolerance was evaluated in liquid YEM broth, by optical density (OD) readings at 540 nm. Initial pre-inocula were standardized to an OD of 0.2. After overnight growth, inocula were standardized to an OD of 0.3 and three replicas were made for each treatment (1:10 dilution of the inocula in a final volume of 4 ml). Final OD readings were performed after 48h at each temperature stress. Continuous growth at 28ºC was used as control condition. In the heat shock stress, cultures were submitted to 48ºC for 15 minutes and then transferred to normal growth temperature (28ºC) for the following 48h.

Data analysis

In order to allow comparisons between isolates, OD readings for each stress were converted into percentage values, considering growth at 28ºC as 100% growth. Average value of the three replicas and standard deviation were calculated.

Statistical analysis was performed using SPSS 15.0 software (SPSS Inc., Chicago, USA). When there was no homogeneity of variances, the Welch test, as well as the Kruskal-Wallis test (non-parametric) were used, instead of the ANOVA test. Different post hoc tests (Tamhane, Dunnett T3 and Games-Howell) were conducted to search for categories that significantly differ from others.

Correspondence analysis was used as an exploratory data analysis technique to detect structure in the relationships between categorical variables (Benzécri, 1973). Isolates were divided into three classes of tolerance to stress (highly tolerant, tolerant and poorly tolerant) and the relationships between these categories and province of origin of the isolates were investigated.

Correlations were examined between temperature of the origin sites and temperature stress tolerance of the isolates. The available temperature data (soil temperature at 10 cm of depth and air temperature) restricted the analysis to a subset of 28 isolates from seven provinces (Trás-os-Montes e Alto Douro, Beira Litoral, Beira Baixa, Alto Alentejo, Baixo Alentejo, Estremadura and Algarve). Temperature data were retrieved from the Portuguese Meteorological Institute and average values (2006-2008) were calculated for summer and winter, for maximum and minimum soil and air temperatures. Spearman's correlation coefficient was used to analyse the data (non parametric statistic).

4.4 - Results

The 53 rhizobia isolates used in the present study (Table 4.1) belong to the collection of chickpea rhizobia from the Soil Microbiology Laboratory (ICAAM-University of Évora, Portugal) and were previously characterized, namely their 16S rRNA gene sequence, plasmid number and symbiotic effectiveness (chapter 2).

In the present study, four to five isolates from each Portuguese province were tested for tolerance to different temperature stresses. This survey revealed that the tolerance of chickpea rhizobia isolates to temperature stress is very diverse (Fig. 4.1). Under 15ºC, the isolates growth ranges from 8% to 68%, while growth at 37ºC ranges from 5% to 46%. After the heat shock, isolates grew between 3% and 89% (detailed percentages shown in Supplementary data). Most isolates tolerate heat shock or cold stress better than the heat stress. The two chickpea rhizobia *Mesorhizobium* type strains (*M. ciceri* and *M. mediterraneum*) are very similar regarding their tolerance to temperature stress and revealed to be highly tolerant to heat shock (>70% of growth) and tolerant to 15ºC. The most tolerant isolate to heat and to cold conditions was BR-9-Bragança, while the most tolerant isolate to continuous heat was BR-15-Bragança. On the other hand, isolates STR-10-Santarém and 78-Elvas are highly sensitive to all the temperature stresses tested (grow <10% in all conditions).

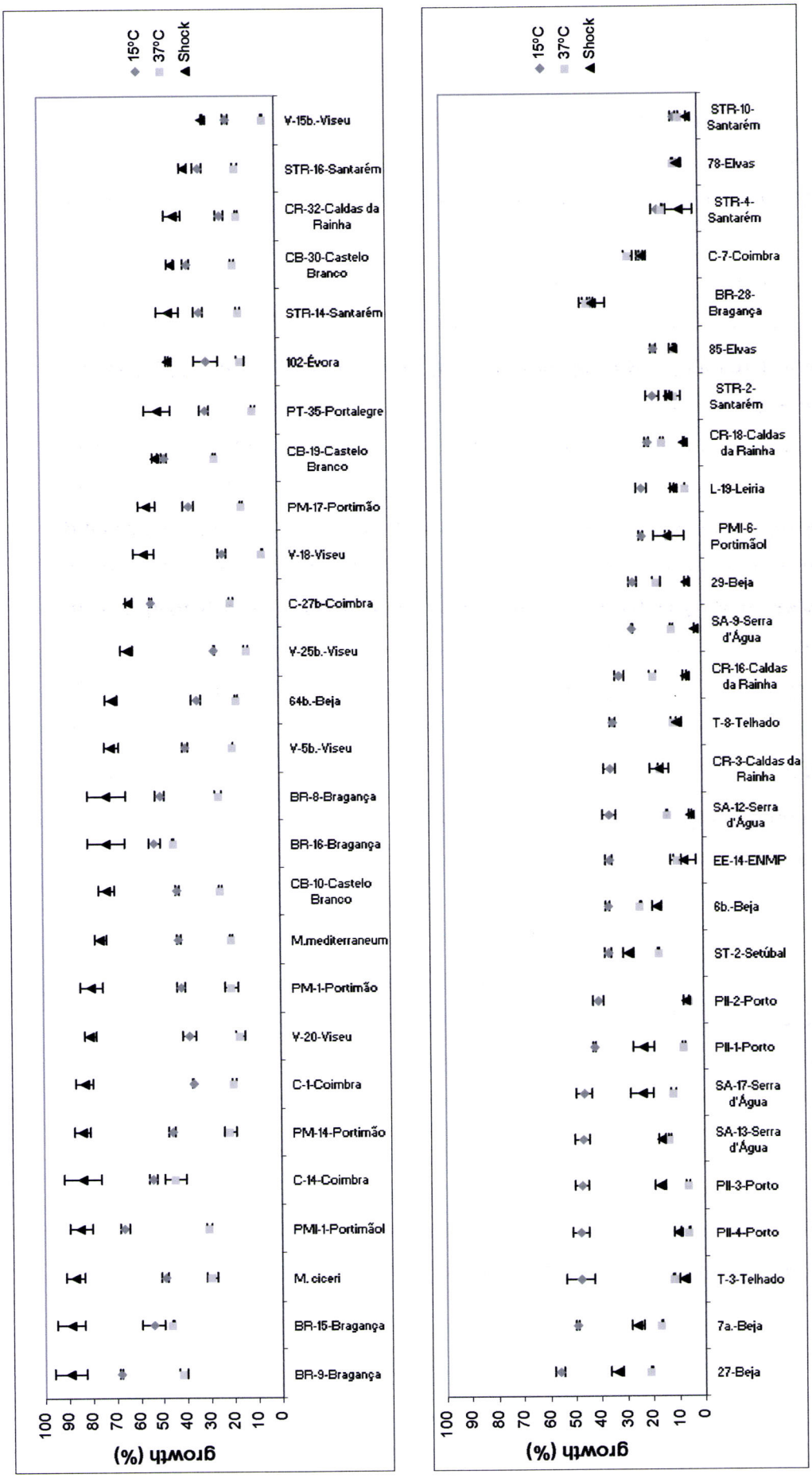


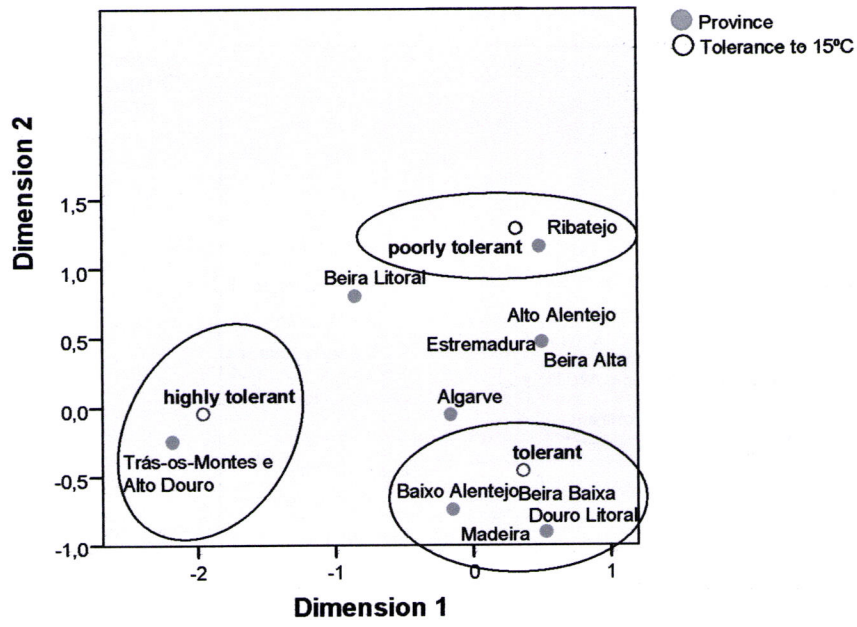
Figure 4.1 – Growth of chickpea rhizobia under different temperature stresses: cold (15ºC), heat (37ºC) and heat shock (48ºC, 15 minutes). Percentages were calculated considering the control condition (28ºC) as 100% growth. Presented values are the average of three replicas (bars show standard deviation).

The isolates used in the present study belong to four main species clusters (Table 4.1), as defined by 16S rRNA sequence analyses in the previous chapter 2. In order to investigate a possible association between the stress tolerance of isolates and their species group, statistical analyses were performed. The Kruskal-Wallis test indicated that there are significant differences between species groups, regarding their growth under each of the three temperature stresses (χ^2=67.345, df=3, P<0.001 for tolerance to cold; χ^2=20.963, df=3, P<0.001 for tolerance to heat and χ^2=11.445, df=3, P<0.05 for tolerance to heat shock). The Welch test also suggests differences between species groups regarding the stress tolerance of the isolates (P< 0.05). Different post hoc tests provided consistent results in identifying the species clusters that were significantly different. Regarding the tolerance to 15ºC, isolates from the species cluster *M. ciceri / M. loti* and isolates from the species cluster *M. mediterraneum / M. temperatum* display the highest growth average and are significantly different from each other and from the remaining species clusters (P< 0.05). In terms of tolerance to heat, significant differences were found between isolates grouping with *M. ciceri / M. loti*, which display the highest growth average and two other species clusters with the lowest growth averages: *M. huakuii / M. amorphae* and *M. tianshanense* (P<0.05). Concerning the tolerance to heat shock, two species clusters differ significantly, namely *M. ciceri / M. loti* species group with the highest growth average and *M. tianshanense* with the lowest (P<0.05).

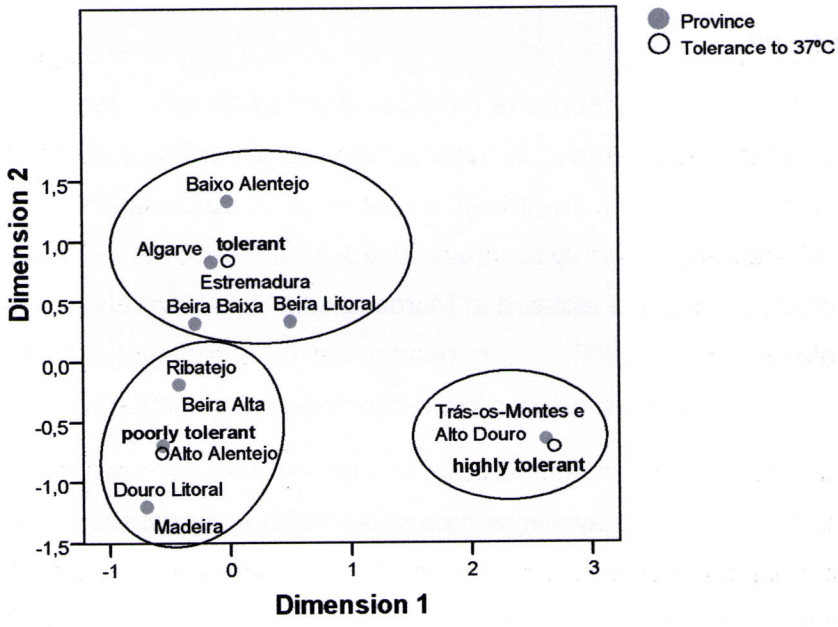
The correspondence analysis (CA) revealed an association between province of origin of the isolates and their growth under each stress (Fig. 4.2). In general, isolates from the province Trás-os-Montes e Alto Douro are associated with high tolerance to cold, heat and heat shock. Additionally, two other provinces are associated with high tolerance to heat shock, namely Algarve and Beira Litoral. On the other hand, isolates from Douro Litoral and Madeira are associated with high sensitivity to heat shock. In terms of cold stress, Ribatejo is the province more associated with low tolerance. The province more associated with low tolerance to heat is Alto Alentejo.

Isolates tolerance data were also tested for correlation with soil and air temperature of the origin sites. The available temperature data allowed us to analyse a subset of 28 isolates from seven provinces and a negative correlation was depicted between the average minimum air temperature in winter and cold tolerance (r=-0.248 P<0.05) (data not shown).

A. Tolerance to cold stress (15ºC)



B. Tolerance to heat stress (37ºC)



C. Tolerance to heat shock (15 minutes at 48ºC)

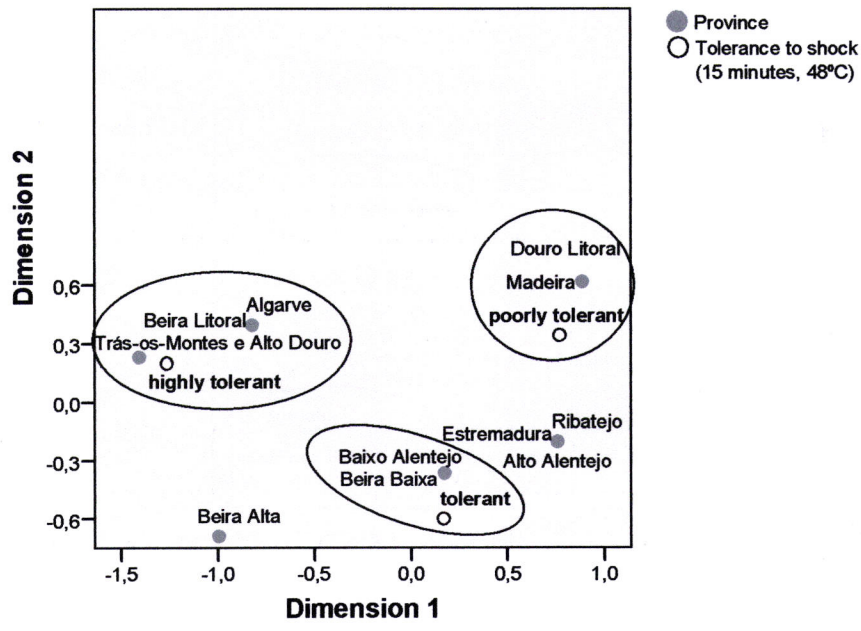


Figure 4.2 - Correspondence analysis biplot of the relationship between tolerance to temperature stresses and province of origin of the isolates.

4.5 - Discussion

The present study shows that native chickpea rhizobia display a large diversity in their tolerance to temperature stress. In general, isolates grew better under low than under high temperature. A maximum growth temperature of 40ºC was reported for both *M. ciceri* and *M. mediterraneum* type strains (Nour *et al.*, 1995, Nour *et al.*, 1994). The growth of chickpea rhizobia isolated in Morocco soils was severely inhibited above 30-37ºC (Maâtallah *et al.*, 2002). Our results are in agreement with these previous findings, since most isolates showed low percentages of growth at 37ºC.

Taking into account that soil temperature in Portugal can go down to negative values (the minimum value in the available data was -1.6ºC in Bragança), one could expect Portuguese rhizobia to be highly tolerant to 15ºC. However only eight isolates showed growth percentages above 50% at 15ºC. Although there are studies indicating that rhizobia survival in soil is not affected at 15ºC (Lindström *et al.*, 1985), rhizobia seem to tolerate the temperature stress in soil better than in culture medium (AbdelGadir & Alexander, 1997, Räsänen *et al.*, 2001).

Despite the fact that only two species of the *Mesorhizobium* genus are reported as specific chickpea microsymbionts, namely *M. ciceri* and *M. mediterraneum* (Nour et al., 1994, Nour et al., 1995), several studies indicated that other *Mesorhizobium* species are able to nodulate this legume, namely isolates from the *M. huakuii* / *M. amorphae* species group and from the *M. tianshanense* species group (chapter 2) (Laranjo *et al.*, 2004, Laranjo *et al.*, 2008, Rivas *et al.*, 2007). In the present study, significant differences were found between isolates from the four species groups, regarding their ability to tolerate temperature stress. Isolates from the species groups that include the specific chickpea microsymbionts, *M. ciceri* / *M. loti* and *M. mediterraneum* / *M. temperatum*, have the highest growth average at 15ºC and are significantly different from the other species. Regarding continuous heat stress and heat shock stress, isolates from the specific chickpea microsymbionts species groups display the two highest tolerance averages, as well. Several other studies in rhizobia report that tolerance seems to be species related, for example pH tolerance (Brígido *et al.*, 2007, Reeve *et al.*, 2006), copper tolerance (Laguerre *et al.*, 2006) and antibiotic resistance (Alexandre *et al.*, 2006).

Contrary to other studies, which indicate that rhizobia tolerant to heat stress were mainly obtained from geographical origins with high temperatures (Rahmani *et al.*, 2009, Zahran, 1999), in the present study no positive correlation was found between rhizobia growth at 37ºC and maximum soil or air temperatures of the origin sites. Previous studies on chickpea rhizobia already reported that temperature tolerance did not seem to be correlated with the temperature in the isolation site (Maâtallah *et al.*, 2002, Rodrigues *et al.*, 2006).

Although isolates tolerance is apparently unrelated to the temperatures of the origin sites, associations were found between some provinces of origin and isolates tolerance to temperature stress. The CA biplots clearly associate isolates from the province Trás-os-Montes e Alto Douro with high tolerance to all tested temperature stresses (Fig. 4.2). Interestingly, this province is in the north of Portugal and presents the lowest air and soil temperatures in the winter, while in the summer shows high temperatures. Thermotolerant isolates from this province, such as BR-15-Bragança and BR-9-Bragança, are good candidates for further studies on the molecular bases of tolerance to both high and low temperatures.

To our knowledge, this is the first survey on chickpea rhizobia temperature tolerance, using diverse isolates and covering a large territory. Thermotolerant chickpea rhizobia isolates with potential agronomical importance were identified.

4.6 - References

AbdelGadir, A. H. & M. Alexander, (1997) Procedures to enhance heat resistance of *Rhizobium*. *Plant Soil* **188**: 93-100.

Alexandre, A., M. Laranjo & S. Oliveira, (2006) Natural populations of chickpea rhizobia evaluated by antibiotic resistance profiles and molecular methods. *Microb. Ecol.* **51**: 128-136.

Benzécri, J. P., (1973) *Analyse des données. Tome I: Analyse des correspondances. Tome II: La Classification.* Dunod, Paris.

Brígido, C., A. Alexandre, M. Laranjo & S. Oliveira, (2007) Moderately acidophilic mesorhizobia isolated from chickpea. *Lett. Appl. Microbiol.* **44**: 168-174.

Day, J. M., R. J. Roughley, A. R. J. Eaglesham, M. Dye & S. P. White, (1978) Effect of high soil temperatures on nodulation of cowpea, *Vigna unguiculata*. *Ann. Appl. Biol.* **88**: 476-481.

Hungria, M. & M. A. T. Vargas, (2000) Environmental factors affecting N_2 fixation in grain legumes in the tropics, with an emphasis on Brazil. *Field Crops Res.* **65**: 151-164.

Laguerre, G., L. Courde, R. Nouaim, I. Lamy, C. Revellin, M. C. Breuil & R. Chaussod, (2006) Response of rhizobial populations to moderate copper stress applied to an agricultural soil. *Microb. Ecol.* **52**: 426-435.

Laranjo, M., A. Alexandre, R. Rivas, E. Velázquez, J. P. W. Young & S. Oliveira, (2008) Chickpea rhizobia symbiosis genes are highly conserved across multiple *Mesorhizobium* species. *FEMS Microbiol. Ecol.* **66**: 391-400.

Laranjo, M., J. Machado, J. P. W. Young & S. Oliveira, (2004) High diversity of chickpea *Mesorhizobium* species isolated in a Portuguese agricultural region. *FEMS Microbiol. Ecol.* **48**: 101-107.

Lindström, K., M. L. Sarsa, J. Polkunen & P. Kansanen, (1985) Symbiotic nitrogen-fixation of Rhizobium (Galega) in acid soils, and its survival in soil under acid and cold stress. *Plant Soil* **87**: 293-302.

Maâtallah, J., E. B. Berraho, S. Muñoz, J. Sanjuan & C. Lluch, (2002) Phenotypic and molecular characterization of chickpea rhizobia isolated from different areas of Morocco. *J. Appl. Microbiol.* **93**: 531-540.

Munevar, F. & A. G. Wollum, (1981) Effect of high root temperature and rhizobium strain on nodulation, nitrogen fixation, and growth of soybeans. *Soil Sci. Soc. Am. J.* **45**: 1113-1120.

Nour, S. M., J.-C. Cleyet-Marel, P. Normand & M. P. Fernandez, (1995) Genomic heterogeneity of strains nodulating chickpeas (*Cicer arietinum* L.) and description of *Rhizobium mediterraneum* sp. nov. *Int. J. Syst. Bacteriol.* **45**: 640-648.

Nour, S. M., M. P. Fernandez, P. Normand & J.-C. Cleyet-Marel, (1994) *Rhizobium ciceri* sp. nov., consisting of strains that nodulate chickpeas (*Cicer arietinum* L.). *Int. J. Syst. Bacteriol.* **44**: 511-522.

Rahmani, H., N. Saleh-Rastin, K. Khavazi, A. Asgharzadeh, D. Fewer, S. Kiani & K. Lindström, (2009) Selection of thermotolerant bradyrhizobial strains for nodulation of soybean (*Glycine max* L.) in semi-arid regions of Iran. *World J. Microbiol. Biotechnol.* **25**: 591-600.

Räsänen, L. A., A. M. Elvang, J. Jansson & K. Lindström, (2001) Effect of heat stress on cell activity and cell morphology of the tropical rhizobium, *Sinorhizobium arboris*. *FEMS Microbiol. Ecol.* **34**: 267-278.

Reeve, W. G., L. Brau, J. Castelli, G. Garau, C. Sohlenkamp, O. Geiger, M. J. Dilworth, A. R. Glenn, J. G. Howieson & R. P. Tiwaril, (2006) The *Sinorhizobium medicae* WSM419 lpiA gene is transcriptionally activated by FsrR and required to enhance survival in lethal acid conditions. *Microbiology-Sgm* **152**: 3049-3059.

Rivas, R., M. Laranjo, P. F. Mateos, S. Oliveira, E. Martinez-Molina & E. Velázquez, (2007) Strains of *Mesorhizobium amorphae* and *Mesorhizobium tianshanense*, carrying symbiotic genes of common chickpea endosymbiotic species, constitute a novel biovar (*ciceri*) capable of nodulating *Cicer arietinum*. *Lett. Appl. Microbiol.* **44**: 412-418.

Rodrigues, C., M. Laranjo & S. Oliveira, (2006) Effect of heat and pH stress in the growth of chickpea mesorhizobia. *Curr. Microbiol.* **53**: 1-7.

Somasegaran, P. & H. J. Hoben, (1994) *Handbook for Rhizobia,* p. 450. Springer-Verlag, New York.

Vincent, J. M., (1970) *A Manual for the Practical Study of Root-Nodule Bacteria,* p. 164. Blackwell Scientific Publications, Oxford.

Zahran, H. H., (1999) *Rhizobium*-legume symbiosis and nitrogen fixation under severe conditions and in an arid climate. *Microbiol. Mol. Biol. Rev.* **63**: 968-989.

5.

Transcriptional analysis of *dnaKJ* and *groESL* genes upon temperature stress

This chapter is based on the manuscript:

Alexandre, A. and Oliveira, S. "Most heat tolerant rhizobia show high induction of major chaperone genes upon stress" Accepted for publication in FEMS Microbiology Ecology.

5.1 - Summary

The ability of rhizobia to endure stress is very important in order to achieve high symbiotic performance in suboptimal conditions. The objective of the present work was to study the molecular bases of temperature stress tolerance in rhizobia, by comparing the expression levels of major chaperone genes *dnaKJ* and *groESL* in thermotolerant and thermosensitive isolates previously characterized. Mesorhizobia isolates were submitted to three temperature stresses: 48ºC for 15 min, 37ºC for 30 min and 15ºC for 30 min. Analysis of the *dnaK* and *groESL* expression by Northern hybridisation, using isolates from three species groups, showed an increase in the transcripts levels with heat but not with cold stress. Interestingly, after a temperature upshift, a higher induction of chaperone genes was detected in tolerant isolates compared to that of sensitive isolates, from the same species. To our knowledge this is the first analysis of chaperone genes expression using tolerant and sensitive strains, from the same species. The present study suggests a relationship between higher transcriptional induction of the major chaperone genes and higher tolerance to heat in rhizobia.

5.2 - Introduction

In free-living state or during symbiosis, rhizobia are affected by environmental factors, such as temperature or pH. Temperature stress response has been widely studied in bacteria, mostly in the gram-negative *Escherichia coli* and in the gram-positive *Bacillus subtilis*. Although molecular chaperones are often mentioned as part of the cell's heat shock response, these proteins are often found to be induced in response to other stresses.

Folding chaperones, as is the case of DnaK and GroEL, recognize exposed hydrophobic domains of target protein that would be hidden in normal protein conformation (for review see Baneyx & Mujacic, 2004). These proteins help denatured proteins to reach its native conformation and are induced by diverse stressful conditions, though constitutively expressed. The DnaK machinery comprises the co-chaperone DnaJ and the nucleotide exchange factor GrpE, while the GroEL machinery includes the co-chaperone GroES (for review see Siegenthaler & Christen, 2006, Chaudhuri *et al.*, 2009).

Many *Alphaproteobacteria* genomes harbor several copies of the *groEL* gene, with root-nodulating bacteria presenting the highest copy number (for review see Lund, 2009). *Bradyrhizobium japonicum*, a soybean symbiont, shows a total of five *groESL* operons (Fischer *et al.*, 1993). These multiple copies have different regulation systems and are differentially induced, namely $groESL_{1,4,5}$ are heat inducible and $groESL_3$ is induced by low oxygen conditions (Fischer *et al.*, 1993, Babst *et al.*, 1996). Accordingly, $groESL_3$ has a distinct regulation mechanism being co-regulated with symbiotic nitrogen fixation genes (Fischer *et al.*, 1993). The $groESL_1$ is σ^{32} dependent and the CIRCE element controls the $groESL_{4,5}$ regulation, which means that regulation mechanisms of heat shock genes described for *E. coli* (positive regulation by the σ^{32} factor) and *B. subtilis* (negative regulation by the CIRCE system) were both found in rhizobia (Babst *et al.*, 1996). In *Rhizobium leguminosarum* only the highly expressed of the three *groEL* homologues is needed for normal growth (Rodriguez-Quinones *et al.*, 2005). In *Sinorhizobium meliloti* genome, five *groESL* operons were found and one operon was found to be involved in the regulation of early *nod* genes (Ogawa & Long, 1995).

In rhizobia, as in most bacteria, *dnaK* is mostly a single copy gene, contrary to its co-chaperone *dnaJ* that in addition to the gene copy found adjacent to *dnaK*, is often found elsewhere in the genome (chapter 3). In *B. japonicum dnaK* seems to have an essential function, given that no *dnaK* knock-out mutant could be obtained, while *dnaJ*

mutants were successfully isolated and showed a reduced growth rate, especially at high temperatures (Minder *et al.*, 1997). *dnaJ* is required for effective symbiosis of *R. leguminosarum* bv. *phaseoli* (Labidi *et al.*, 2000), however in *B. japonicum* the symbiotic performance of *dnaJ* mutants was not altered (Minder *et al.*, 1997).

Expression of *groES* and *groEL* genes from a psychrophilic bacterium was able to increase *E. coli* growth at low temperatures and decrease the lower limit temperature for growth (Ferrer *et al.*, 2003). A recent study in *E. coli* reported that overexpression of native *groESL* system enhanced thermotolerance (Kim *et al.*, 2009). In *B. subtilis*, a mutant in the repressor of the heat shock response regulatory system showed, under normal conditions (37ºC), similar levels of DnaK and GroEL to the wild type under heat stress (48ºC) and was able to recover from heat shocks that were lethal for the wild type strain (Yuan & Wong, 1995). The key roles of DnaKJ and GroESL machineries in determining temperature stress tolerance phenotype prompt us to investigate the chaperone genes expression at the transcriptional level, comparing thermotolerant with thermosensitive rhizobia isolates. A previous study with chickpea rhizobia indicated a consistent overexpression a protein of approximately 60kDa after heat shock that might correspond to GroEL (Rodrigues *et al.*, 2006)

To study the molecular bases of temperature stress response in rhizobia, we analysed the expression of chaperone *dnaKJ* and *groESL* genes upon stress, using a set of temperature tolerant and sensitive chickpea rhizobia isolates, belonging to different *Mesorhizobium* species. Our aim was to compare the expression levels of thermotolerant and thermosensitive isolates, within the same species, to find differences that could be related to the diverse temperature susceptibility phenotypes.

5.3 - Methods

Bacterial strains and growth conditions

A total of 15 chickpea rhizobia isolates were chosen for their different ability to tolerate temperature stress (Table 5.1). The two chickpea *Mesorhizobium* type strains were also used (*M. ciceri* UPMCa7^T and *M. mediterraneum* UPMCa36^T).

To investigate the response of chickpea mesorhizobia to several temperature conditions, cells were grown overnight at 28ºC in YEM (Vincent, 1970) before the different stress conditions were applied. The heat shock stress was performed at 48ºC

for 15 minutes; the heat stress was performed at a lower temperature and for longer time (37ºC for 30 minutes); the cold stress was perform at 15ºC for 30 minutes.

Table 5.1 – Chickpea rhizobia used in this study.

A. Heat shock (48ºC, 15 min)

Species group[a]	Tolerant isolates	Growth (%)[b]	Sensitive isolates	Growth (%)[b]
M. ciceri / M. loti	BR-15-Bragança	92	7a-Beja	26
M. mediterraneum / M. temperatum	PM-1-Portimão	81	29-Beja	5
M. huakuii / M. amorphae	C-1-Coimbra C-14-Coimbra	88 84	78-Elvas L-19-Leiria	8 10

B .Heat (37ºC, 30 min)

Species group[a]	Tolerant isolates	Growth (%)[b]	Sensitive isolates	Growth (%)[b]
M. ciceri / M. loti	BR-15-Bragança BR-28-Bragança	46 43	7a-Beja CB-30-Castelo Branco	17 18
M. huakuii / M. amorphae	C-14-Coimbra	35	L-19-Leiria	6

C. Cold (15ºC, 30 min)

Species group[a]	Tolerant isolates	Growth (%)[b]	Sensitive isolates	Growth (%)[b]
M. ciceri / M. loti	27-Beja BR-9-Bragança	56 68	Not available	
M. huakuii / M. amorphae	C-3-Coimbra	63	78-Elvas 85-Elvas STR-4-Santarém	8 18 10

[a] from chapter 2

[b] from chapter 4

Transcriptional analysis of the *dnaKJ* and *groEL* genes

RNA preparation and electrophoresis

Total RNA was extracted from cells at control temperature (28ºC) and from cells subjected to the temperature stress conditions, according to the protocol for Rapid Isolation of RNA from Gram-negative Bacteria (Ausubel *et al.*, 1997).

Total RNA was separated on 1.5% agarose gel, in 1X MOPS (20 mM MOPS buffer; 5 mM sodium acetate; 2 mM EDTA; pH 7.0) with a denaturing loading buffer (50% deionized formamide; 6.142% formaldehyde; 10% 10X MOPS). A DIG-labeled RNA molecular weight marker (Roche Applied Science) was used in order to confirm the transcripts size. Electrophoresis was carried out at 40 V for 5 hours. Total RNA was transferred overnight by capillary transfer method into positively charged nylon membranes (Roche Applied Science) using 20X SSC (3 M NaCl; 300 mM sodium citrate; pH 7.0). RNA was fixed to the membranes by baking for 30 minutes at 120ºC.

Probes synthesis

The *groEL* gene fragment of approximately 770 bp was amplified using primers groEL-F (5'-GGGCCGCAACGTCGTCATCGACAA-3') and groEL-R (5'-CTTCCAGCATGGCCTTGCGGCGAT-3'). PCR reaction was prepared with 2 U Taq DNA polymerase (Fermentas), 1X reaction buffer, 0.75 mM MgCl2, 0.2 mM of each dNTP (Invitrogen), 15 pmol of each primer (Stabvida) and 1 µl DNA of the *M. mediterraneum* Ca36^T. The amplification program consisted of an initial denaturation step of 3 min at 95ºC, followed by 30 cycles of 2 min at 95ºC, 1 min at 68ºC and 30 sec at 72ºC, and a final extension step at 72ºC for 10 min. The amplified DNA fragment was cloned using the pGEM-T Easy Vector System (Promega) following the manufacturer's instructions. The *dnaKJ* fragment of approximately 1600 bp was amplified as described before (chapter 3) and cloned as mentioned above.

The *groEL* and *dnaKJ* RNA probes were obtained by *in vitro* transcription labelling of the cloned fragment, using the DIG Northern Starter Kit (Roche Applied Science).

The 16S rRNA gene DNA probe was labelled using the DIG High Prime DNA Labelling and Detection Starter Kit II (Roche Applied Science) from purified *M. mediterraneum* Ca36^T 16S rRNA gene PCR product, obtained as previously described (chapter 2).

Northern hybridisation

Membranes were prehibridised in DIG Easy Hyb hybridisation buffer (Roche Applied Science) for 30 minutes. RNA probes were denatured in a boiling bath and added to pre-warmed DIG Easy Hyb hybridisation buffer and all hybridisations were performed overnight at 68ºC. Stringency washes and immunological detection were performed according to the manufacturer's instructions. Membranes were exposed to chemiluminescent detection films (Roche Applied Science) for 30 minutes.

The membranes were re-hybridised with a 16S rRNA gene DNA probe, using the RNA-DNA hybridisation temperature of 50ºC.

Data analysis

Hybridisation signals were analysed using ImageQuant™ TL v7.01 (GE Healthcare). The 16S rRNA hybridisation was used as an internal control of the amount of total RNA loaded in each sample. The ratio between the chaperone gene signal and the 16S

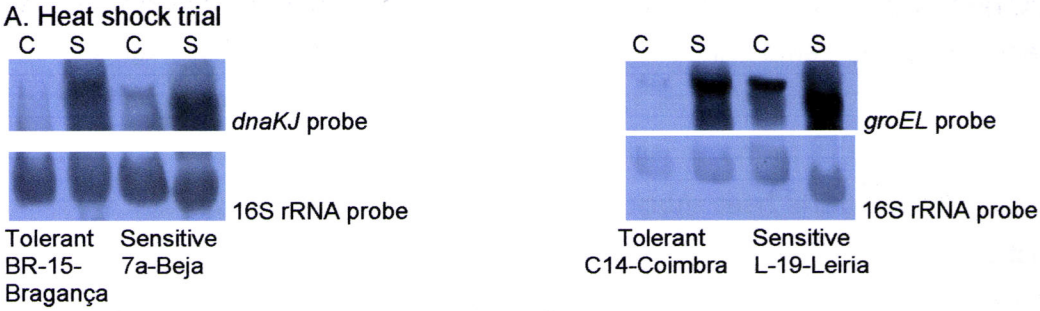
rRNA signal was calculated and the ratio between control and temperature stress conditions yielded the number of folds by which the chaperone mRNA levels were induced under stress.

5.4 - Results

In order to investigate the expression of major chaperone genes in isolates with different susceptibility to temperature stress, the transcriptional levels of the *dnaKJ* and *groESL* genes were analysed by Northern hybridisation. Taking into account previous results from the survey on temperature stress tolerance of chickpea rhizobia isolates (chapter 4), a total of 15 chickpea rhizobia isolates, comprising three species groups, including both tolerant and sensitive isolates for each stress condition, were selected (Table 5.1).

The *dnaK-dnaJ* RNA probe enables the detection of three putative transcripts: the mRNA corresponding to the *dnaK* gene; the mRNA of the *dnaJ* gene; and the bicistronic *dnaKJ* mRNA. Using RNA from rhizobia isolates subjected to the three stress conditions, no detectable bands were found around 3.3 kb, a size that would correspond to that of the *dnaKJ* bicistronic transcript. The most abundant transcript detected corresponded to a RNA with approximately 2 kb (Fig. 5.1), which is consistent with the *dnaK* gene size in *Mesorhizobium loti* MAFF303099 (1917 bp).

The *groEL* RNA probe enables the detection of two possible transcripts: the *groEL* mRNA alone and the bicistronic *groESL* mRNA. Using RNA from rhizobia isolates exposed to the three stress conditions, the most abundant transcript detected with the *groEL* RNA probe corresponded to a size of approximately 2 kb, which is consistent with the *groESL* bicistronic mRNA, since the gene size of *groEL* in *M. loti* MAFF303099 is 1650 bp, whereas the *groES* gene is 300 bp, approximately (Fig. 5.1).



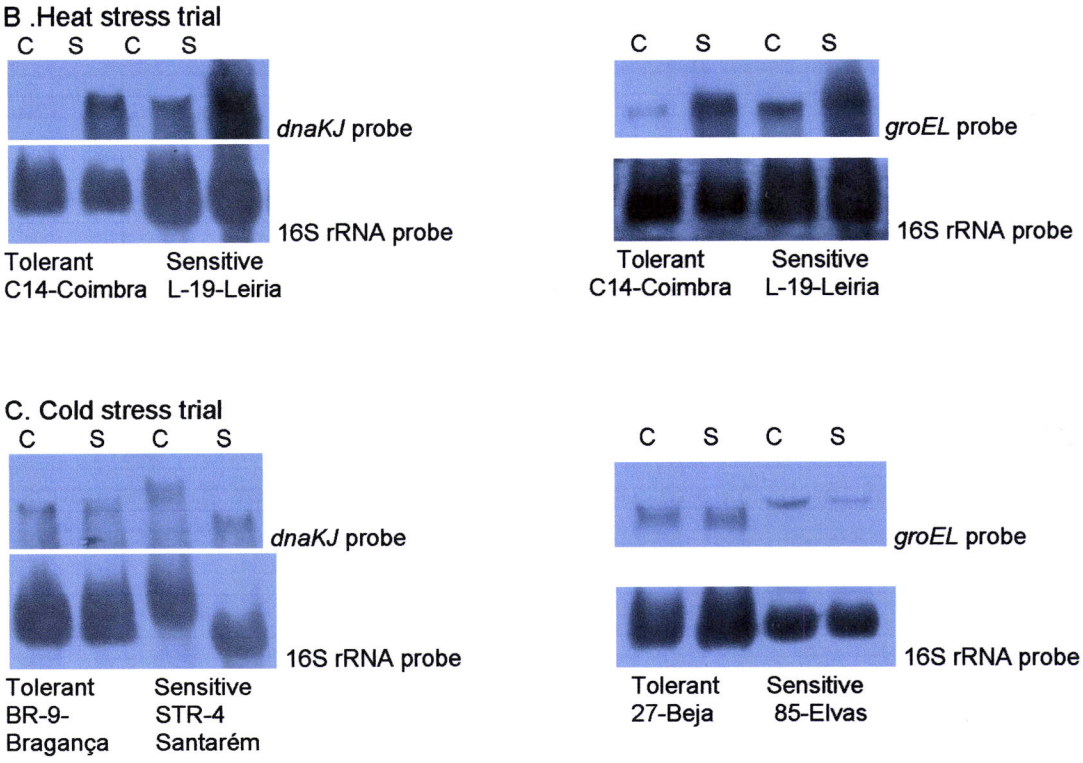


Figure 5.1 – Comparison of the transcription analysis of chaperones genes between a tolerant isolate and a sensitive isolate submitted to temperature stress. For each isolate control conditions (C) were compared to stress conditions (S). Northern blot hybridisations of total RNA with probes specific for *dnaKJ*, *groEL* and 16S rRNA.

As expected, after a temperature upshift to 37ºC or 48ºC, the majority of the isolates showed an increase in the chaperones transcripts levels. Interestingly, after heat shock, the increase of the levels of both *dnaK* and *groESL* transcripts was higher in tolerant isolates, than in sensitive isolates, within the same species cluster (Fig. 5.2). The only exceptions were the tolerant isolate C-14-Coimbra and the sensitive isolate 78-Elvas, with low and high induction levels of *dnaK*, respectively. The highest induction of *dnaK* gene was detected in tolerant isolate BR-15-Bragança, with over 15 fold induction after the heat shock. Sensitive isolate 29-Beja showed repression of the *dnaK* gene, after the temperature upshift (0.5 fold induction). The highest fold induction of the *groESL* gene was detected in the tolerant isolate C-14-Coimbra (approximately 7.5 fold). In the sensitive isolates, almost no induction of the *groESL* was detected following the heat shock, as for example in isolate L-19-Leiria with 1.1 fold induction.

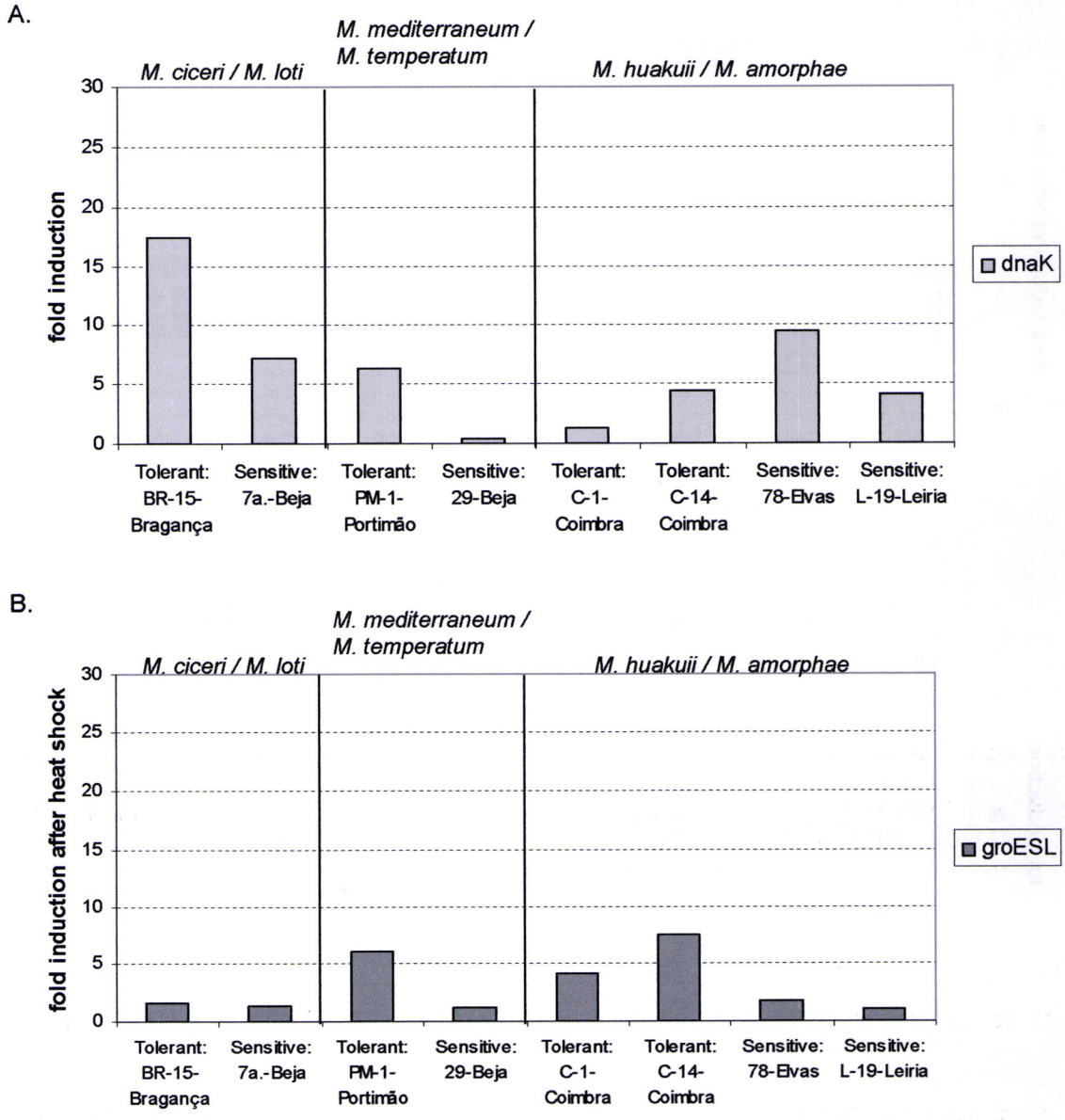
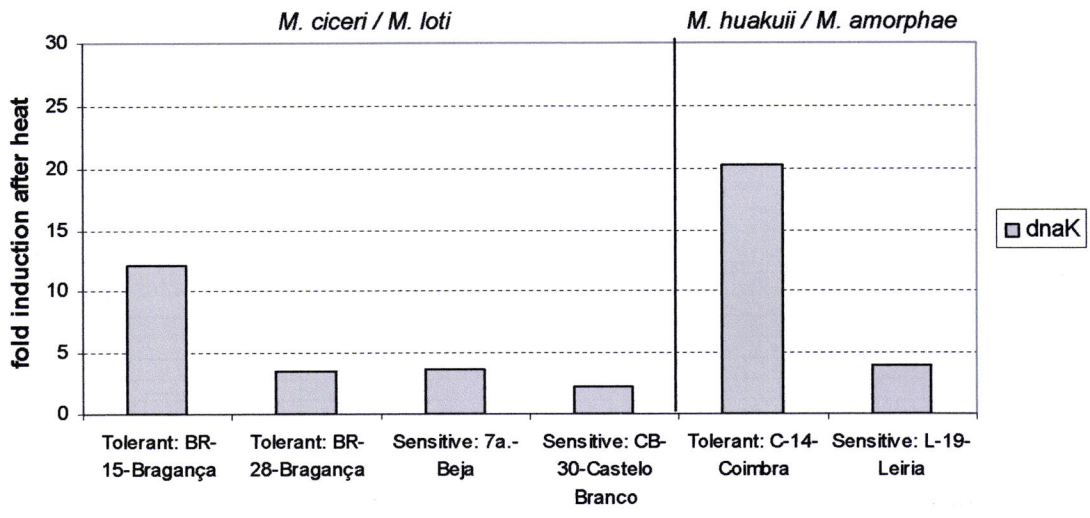


Figure 5.2 – Heat shock induction as determined by Northern blot analysis with *dnaKJ* (A) and *groEL* (B) RNA probes, using total RNA from tolerant and sensitive rhizobia isolates.

Similarly to the results obtained for the heat shock condition, after the heat stress, a higher induction of both *dnaK* and *groESL* was detected in tolerant isolates compared to sensitive isolates within the same species cluster (Fig. 5.3). The only exception was the sensitive isolate 7a.-Beja that showed high induction of the *groESL*, contrary to the other sensitive isolate. The tolerant isolate C-14-Coimbra showed the highest fold induction of *dnaK* after stress (approximately 20 fold), while the highest *groESL* induction was detected in the tolerant isolate BR-15-Bragança (approximately 29 fold).

A.



B.

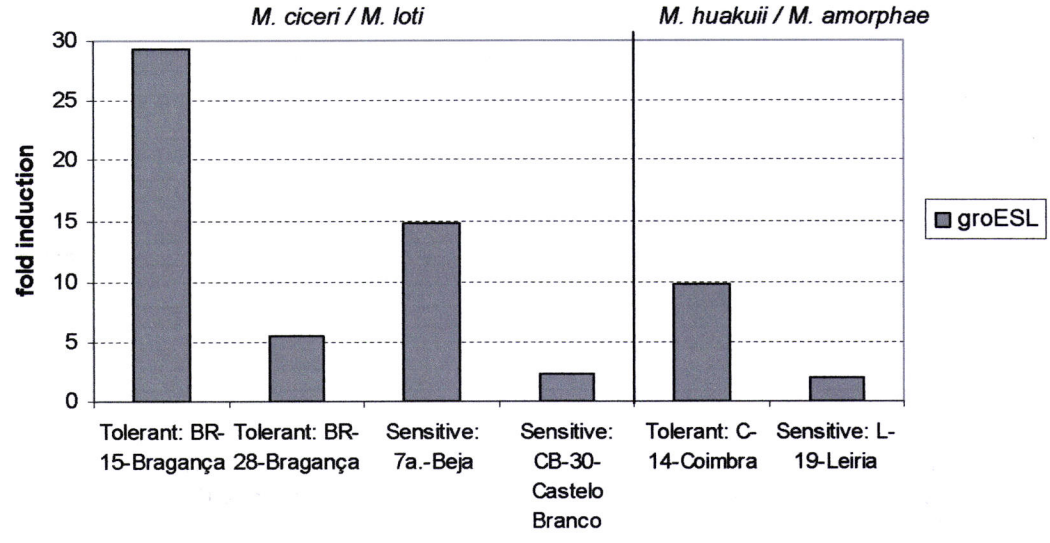
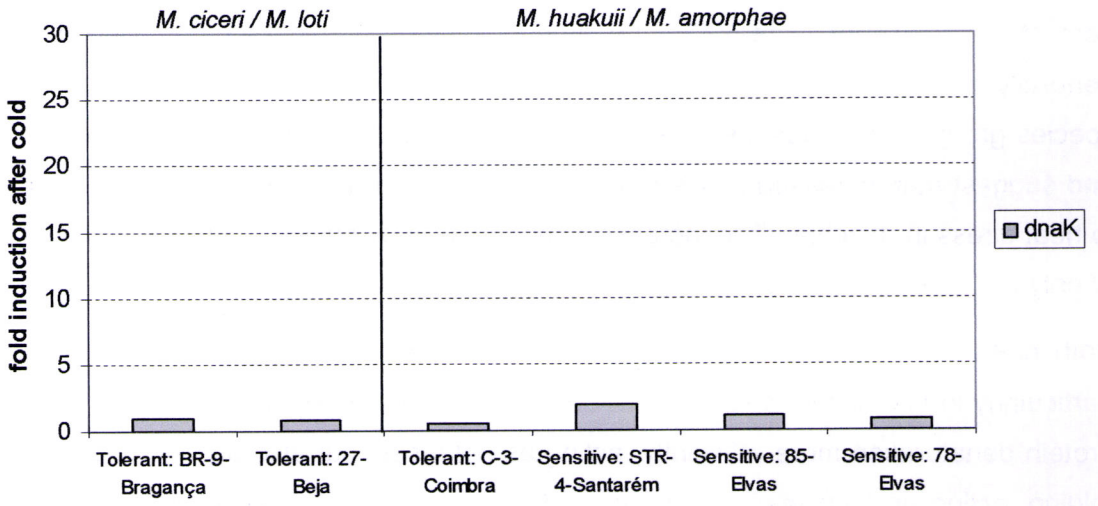


Figure 5.3 – Heat stress induction as determined by Northern blot analysis with *dnaKJ* (A) and *groEL* (B) RNA probes, using total RNA from tolerant and sensitive rhizobia isolates.

In the case of cold stress, no significant variation of *dnaK* or *groESL* mRNA levels was detected in the six tolerant and sensitive isolates analysed (Fig. 5.4). The *dnaK* and *groESL* mRNAs levels detected following the cold stress, were similar to the ones detected under normal growth conditions, regardless of the isolate phenotype or species group. The maximum fold induction detected after cold stress was approximately 2 fold, for the *dnaK* levels in the isolate STR-4-Santarém.

A.



B.

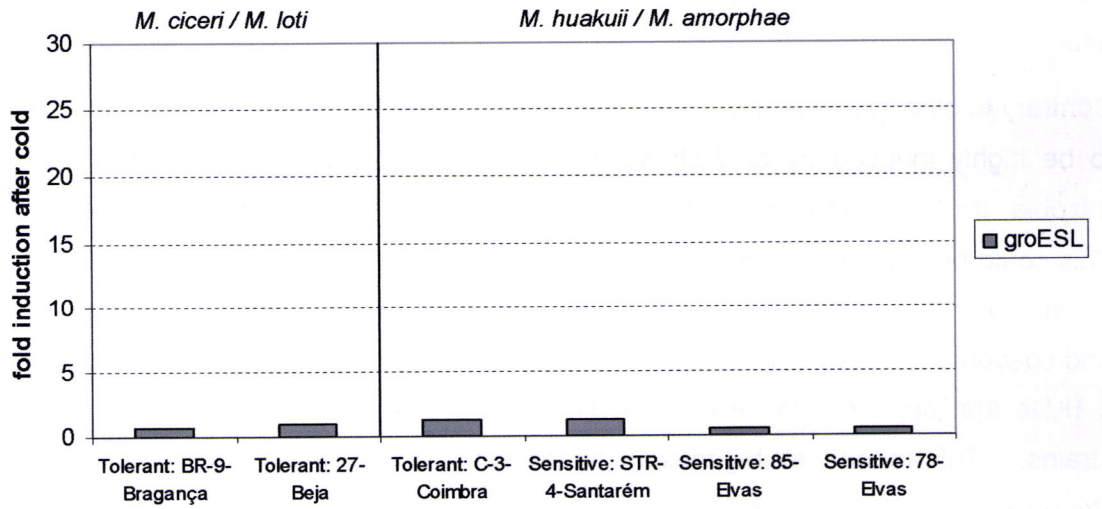


Figure 5.4 – Cold stress induction as determined by Northern blot analysis with *dnaKJ* (A) and *groEL* (B) RNA probes, using total RNA from tolerant and sensitive rhizobia isolates.

5.5 - Discussion

Thermotolerant rhizobia can be useful in more adverse environmental conditions, where temperature is a limiting factor. Therefore, it is important to understand the molecular basis of temperature stress tolerance of these soil bacteria. DnaKJ and GroESL machineries are of major importance when coping with adverse conditions, especially heat stress, and interact with each other in their fundamental task of preventing irreversible misfold of cell proteins.

The transcriptional analysis of 15 chickpea rhizobia isolates points towards a general increase of both *dnaK* and *groESL* mRNA levels after temperature upshift and irrelevant alteration of those levels after cold stress. Interestingly, after heat shock and heat stress, a higher increase in the amount of those chaperones transcripts was generally observed in tolerant isolates, compared to sensitive isolates, within the same species group. These results were obtained using isolates from three species groups and suggest that increased levels of chaperones may contribute to a higher tolerance to heat stress in rhizobia. The increase of heat tolerance caused by higher expression of only the *groESL* genes was already reported in *E. coli* (Kim *et al.*, 2009).

Both *dnaK* and *groEL* genes were shown to be transcriptionally activated by heat, particularly in tolerant isolates. Higher amounts of chaperones could, not only prevent protein denaturation more efficiently, acting as holdases, but also allow proper protein folding, acting as foldases. In addition, it has been seen that these chaperones can increase the mRNAs stability (Yoon *et al.*, 2008). For all these reasons, a higher expression of chaperones genes in rhizobia could contribute to a higher tolerance to heat.

Contrary to other bacteria, such as *Listeria monocytogenes*, for which *groEL* was found to be highly induced by cold stress (Liu *et al.*, 2002), our results suggest that, in rhizobia, *groESL* and also *dnaK* have minor roles in coping with low temperatures. This could be expected since none of the proteins previously associated to cold stress resistance in rhizobia correspond to the genes analysed in the present study. Prévost and co-workers (2003) studied cold-adapted rhizobia and found a cold shock protein of 6.1KDa that appeared to be more synthesized at low temperatures in cold-adapted strains. Tn5 tagged mutagenesis studies with a *Rhizobium* strain isolated from *Sesbania cannabina* revealed two high molecular weight membrane polypeptides (119 and 135 kDa) that are induced under cold shock and possibly have adaptive significance in alleviation of the stress (Sardesai & Babu, 2001).

In mesorhizobia the *dnaK* appears to have a low expression under normal growth conditions compared to the *groESL* operon, since low amount of the *dnaK* transcript was detected in normal growth conditions, contrary to the *groESL* transcript (data not shown). This is consistent to what was reported for *Agrobacterium tumefaciens* (Segal & Ron, 1995). Furthermore, our results show that in most mesorhizobia, *dnaK* induction is higher than *groESL* induction after heat shock, which is the same scenario observed for *A. tumefaciens* (Segal & Ron, 1995).

Analysis of the intergenic region of the *dnaKJ* operon from *M. loti* MAFF303099 shows a putative terminator. A similar terminator could explain the detection of only the *dnaK* transcript in the chickpea mesorhizobia used in the present study. Similarly to our results, Northern hybridisation experiments in *A. tumefaciens* revealed no transcript corresponding to *dnaJ* alone, after the heat shock induction (Segal & Ron, 1995). In *B. japonicum*, no transcription start site was detected in the intergenic region and two putative stem-loop structures located in the intergenic region were suggested to be responsible for the transcription of *dnaK* alone (Minder *et al.*, 1997).

Due to the characteristic high number of *groESL* operons found in rhizobial genomes, it is probable that the *groESL* mRNA levels detected in the present study result from a mRNA pool transcribed from several gene copies. Furthermore, the input of each copy may vary with time and temperature applied. For example, in *B. japonicum* the two most heat-inducible *groESL* operons reach maximal mRNA levels at different temperatures and with different induction kinetics (Babst *et al.*, 1996). This probably outcomes from the fact that these operons are regulated by different mechanisms (Babst *et al.*, 1996). Another study, in *S. meliloti*, identified two *rpoH*-related genes, which codifies for the σ^{32} factor, and showed that they seem to be induced with different timings after the temperature upshift (Ono *et al.*, 2001). Our preliminary data on kinetics of the *groESL* mRNA levels, after temperature upshift, were consistent with these findings, since in some isolates the Northern analysis revealed a decrease followed by an increase in the *groESL* mRNA levels (data not shown). This variation may result from the sum of the input of several operons differentially induced under heat. As in the *Mesorhizobium loti* MAFF303099 genome, five *groESL* operons and two *rpoH*-like genes were found (Kaneko *et al.*, 2000), it is possible that the scenario of diverse regulation mechanism for the heat shock inducible *groESL* operons also exists in chickpea mesorhizobia.

In *B. japonicum*, the three heat inducible *groESL* operons show very different fold induction after a temperature ufshift from 28ºC to 39ºC: 150 fold induction in $groESL_1$, 11 fold induction in $groESL_5$ and 30-50 fold induction in $groESL_4$ (Babst *et al.*, 1996). On the other hand, the fold induction detected for two copies of the *groEL* gene in *R. leguminosarum*, was approximately 3 when cells were heat shocked at 37ºC and 42ºC (Gould *et al.*, 2007). Our results showed that transcriptional induction of *groESL* after temperature upshift could be as high as 29 fold induction and the minimum induction detected was close to 1 (almost no induction).

To our knowledge this is the first study in mesorhizobia focusing on the transcriptional analysis of *dnaKJ* and *groESL* chaperone genes under different temperature stresses.

The present study describes the first analysis of *dnaK* and *groESL* genes expression using both tolerant and sensitive bacteria, from the same species. The results here presented suggest the existence of a relationship between higher levels of transcriptional induction of the *dnaK* and *groESL* chaperone genes and a higher ability of chickpea rhizobia isolates to endure heat stress. Further studies are required to clarify the molecular mechanisms of temperature stress tolerance in rhizobia.

5.6 – References

Ausubel, F. M., R. Brent, R. E. Kingston, D. D. More, J. G. Seidman, J. A. Smith & K. Struhl, (1997) *Current Protocols in Molecular Biology*. Greene Publishing Associates and Wiley-Interscience.

Babst, M., H. Hennecke & H. M. Fischer, (1996) Two different mechanisms are involved in the heat-shock regulation of chaperonin gene expression in *Bradyrhizobium japonicum*. *Mol. Microbiol.* **19**: 827-839.

Baneyx, F. & M. Mujacic, (2004) Recombinant protein folding and misfolding in *Escherichia coli*. *Nat. Biotechnol.* **22**: 1399-1408.

Chaudhuri, T. K., V. K. Verrna & A. Maheshwari, (2009) GroEL assisted folding of large polypeptide substrates in *Escherichia coli*: Present scenario and assignments for the future. *Prog. Biophys. Mol. Biol.* **99**: 42-50.

Ferrer, M., T. N. Chernikova, M. M. Yakimov, P. N. Golyshin & K. N. Timmis, (2003) Chaperonins govern growth of *Escherichia coli* at low temperatures. *Nat. Biotechnol.* **21**: 1266-1267.

Fischer, H. M., M. Babst, T. Kaspar, G. Acuña, F. Arigoni & H. Hennecke, (1993) One member of a *gro*ESL-like chaperonin multigene family in *Bradyrhizobium japonicum* is co-regulated with symbiotic nitrogen fixation genes. *The Embo Journal* **12**: 2901-2912.

Gould, P., M. Maguire & P. A. Lund, (2007) Distinct mechanisms regulate expression of the two major *groEL* homologues in *Rhizobium leguminosarum*. *Arch. Microbiol.* **187**: 1-14.

Kaneko, T., Y. Nakamura, S. Sato, E. Asamizu, T. Kato, S. Sasamoto, A. Watanabe, K. Idesawa, A. Ishikawa, K. Kawashima, T. Kimura, Y. Kishida, C. Kiyokawa, M. Kohara, M. Matsumoto, A. Matsuno, Y. Mochizuki, S. Nakayama, N. Nakazaki, S. Shimpo, M. Sugimoto, C. Takeuchi, M. Yamada & S. Tabata, (2000) Complete genome structure of the nitrogen-fixing symbiotic bacterium *Mesorhizobium loti*. *DNA Res.* **7**: 331-338.

Kim, S. Y., N. Ayyadurai, M. A. Heo, S. Park, Y. J. Jeong & S. G. Lee, (2009) Improving the productivity of recombinant protein in *Escherichia coli* under thermal stress by coexpressing GroELS chaperone system. *J. Microbiol. Biotechnol.* **19**: 72-77.

Labidi, M., S. Laberge, L. P. Vezina & H. Antoun, (2000) The *dnaJ* (hsp40) locus in *Rhizobium leguminosarum* bv. *phaseoli* is required for the establishment of an effective symbiosis with *Phaseolus vulgaris*. *Mol. Plant. Microbe Interact.* **13**: 1271-1274.

Liu, S. Q., J. E. Graham, L. Bigelow, P. D. Morse & B. J. Wilkinson, (2002) Identification of *Listeria monocytogenes* genes expressed in response to growth at low temperature. *Appl. Environ. Microbiol.* **68**: 1697-1705.

Lund, P. A., (2009) Multiple chaperonins in bacteria - why so many? *Fems Microbiol. Rev.* **33**: 785-800.

Minder, A. C., F. Narberhaus, M. Babst, H. Hennecke & H. M. Fischer, (1997) The *dnaKJ* operon belongs to the sigma(32)-dependent class of heat shock genes in *Bradyrhizobium japonicum*. *Mol. Gen. Genet.* **254**: 195-206.

Ogawa, J. & S. R. Long, (1995) The *Rhizobium meliloti groELc* locus is required for regulation of early nod genes by the transcription activator *nodD*. *Genes Dev.* **9**: 714-729.

Ono, Y., H. Mitsui, T. Sato & K. Minamisawa, (2001) Two RpoH homologs responsible for the expression of heat shock protein genes in *Sinorhizobium meliloti*. *Mol. Gen. Genet.* **264**: 902-912.

Prevost, D., P. Drouin, S. Laberge, A. Bertrand, J. Cloutier & G. Levesque, (2003) Cold-adapted rhizobia for nitrogen fixation in temperate regions. *Canadian Journal of Botany-Revue Canadienne De Botanique* **81**: 1153-1161.

Rodrigues, C., M. Laranjo & S. Oliveira, (2006) Effect of heat and pH stress in the growth of chickpea mesorhizobia. *Curr. Microbiol.* **53**: 1-7.

Rodriguez-Quinones, F., M. Maguire, E. J. Wallington, P. S. Gould, V. Yerko, J. A. Downie & P. A. Lund, (2005) Two of the three *groEL* homologues in *Rhizobium leguminosarum* are dispensable for normal growth. *Arch. Microbiol.* **183**: 253-265.

Sardesai, N. & C. R. Babu, (2001) Cold stress induced high molecular weight membrane polypeptides are responsible for cold tolerance in *Rhizobium* DDSS69. *Microbiol. Res.* **156**: 279-284.

Segal, G. & E. Z. Ron, (1995) The *dna*KJ operon of *Agrobacterium tumefaciens*: transcriptional analysis and evidence for a new heat shock promoter. *J. Bacteriol.* **177**: 5952-5958.

Siegenthaler, R. K. & P. Christen, (2006) Tuning of DnaK chaperone action by nonnative protein sensor DnaJ and thermosensor GrpE. *J. Biol. Chem.* **281**: 34448-34456.

Vincent, J. M., (1970) *A Manual for the Practical Study of Root-Nodule Bacteria,* p. 164. Blackwell Scientific Publications, Oxford.

Yoon, H., J. Hong & S. Ryu, (2008) Effects of chaperones on mRNA stability and gene expression in *Escherichia coli*. *J. Microbiol. Biotechnol.* **18**: 228-233.

Yuan, G. & S. L. Wong, (1995) Isolation and characterization of *Bacillus subtilis groE* regulatory mutants: evidence for *orf39* in the *dnaK* operon as a repressor gene in regulating the expression of both *groE* and *dnaK J. Bacteriol.* **177**: 6462-6468.

6.

General discussion

Although chickpea is an important crop worldwide, few studies have addressed the rhizobia nodulating this legume, in comparison with the extensively studied soybean or bean microsymbionts. The present thesis intended to improve the knowledge on chickpea rhizobia by addressing two main subjects: species diversity and tolerance to temperature stress.

The present study includes one of the few surveys on chickpea rhizobia and the first systematic assessment of indigenous rhizobia in Portugal. One hundred and ten chickpea rhizobia isolates obtained from 23 soil samples were identified as *Mesorhizobium* sp., since all isolates are included in a monophyletic cluster with *Mesorhizobium* type strains in the 16S rRNA gene phylogeny.

Native chickpea rhizobia isolates were found to be highly diverse, grouping with several *Mesorhizobium* species in four main clusters, generated by the 16S rRNA gene analysis. Cluster A comprises 56 isolates and the type strains of *M. amorphae*, *M. huakuii*, *M. plurifarium* and *M. septentrionale*. Cluster B includes 32 isolates and *M. loti* and *M. ciceri* type strains. Cluster C includes 13 isolates and only the *M. tianshanense* type strain, while cluster D comprises nine isolates and the type strains of *M. mediterraneum* and *M. temperatum*. Although *M. plurifarium* and *M. septentrionale* type strains are included in cluster A, in the complete 16S rRNA gene phylogeny no isolate was found to group closely or be more similar to these strains. The fact that cluster A was the most abundant species group of chickpea microsymbionts found in Portuguese soils was unexpected, since the *M. huakuii* and *M. amorphae* type strains are unable to nodulate chickpea. Furthermore, the original hosts of these species are unrelated to *Cicer arietinum* and *Astragalus sinicus* (*M. huakuii* host) does not exist in Portugal, while *Amorpha fruticosa* (*M. amorphae* host) is uncommon. We could expect that *M. ciceri* and *M. mediterraneum* species groups (clusters B and D, respectively) included the majority of native isolates, since these species were described as the specific chickpea microsymbionts (Nour *et al.*, 1995, Nour *et al.*, 1994) and were the two species frequently found in other studies on chickpea rhizobia diversity (Maâtallah *et al.*, 2002, L'Taief *et al.*, 2007). To our knowledge only one other study surveyed several regions of a country, in order to study the chickpea rhizobia diversity. The characterization of 48 isolates obtained from 13 sites sampled in Morocco, revealed that more than 60% of the isolates are closely related to either *M. ciceri* or *M. mediterraneum* (Maâtallah *et al.*, 2002).

Several studies have previously addressed the chickpea rhizobia diversity in the southern region of Portugal (provinces of Alto Alentejo and Baixo Alentejo) and found that natural populations are diverse (Laranjo *et al.*, 2002, Laranjo *et al.*, 2001). The 16S rRNA gene sequence analysis supported this high diversity showing that chickpea rhizobia, rather than grouping with just *M. ciceri* and *M. mediterraneum*, also group closer to *M. loti* (one isolate), *M.tianshanense* (two isolates) and *M. huakuii*. (eight isolates) (Laranjo *et al.*, 2004). The present study extended the sampling sites to the entire country and not only confirmed the high diversity, but also showed that in terms

of relative abundance of species, most chickpea rhizobia isolates are more similar to *M. huakuii*, but might be different enough to represent a new species. The existence of at least one new species was recently supported by the phylogenetic analysis of several housekeeping genes (Laranjo *et al.*, unpublished results). The present survey of chickpea rhizobia in Portugal revealed that only 33% of the isolates grouped with *M. ciceri* or *M. mediterraneum*.

The biogeography study of chickpea rhizobia species revealed an association between province of origin and species clusters of the isolates, suggesting that the geographical distribution is not random and that some species are typically found in a certain region. In the North of Portugal isolates from cluster B predominate, while most isolates found in the Centre of Portugal belong to cluster A and in the South isolates mostly belong to cluster D. Studies addressing soil bacterial communities found that pH is a key environmental parameter determining the species geographic distribution (Fierer & Jackson, 2006). Accordingly, in the present study, a correlation was found between isolates species cluster and origin soil pH.

The analysis of the plasmid number of the isolates revealed that plasmid number seems to be associated with species cluster, suggesting that this feature might be species constrained. In each species, isolates with high and low symbiotic effectiveness can be found. A large set of isolates with very high SE values (above 75%) are good candidates for field inoculation, in order to significantly reduce the amount of synthetic N-fertilizers used. For example, isolate T-1-Telhado and isolate G-55-Guarda are very effective with SE values above 85%.

Another aim of the present thesis was to find a core gene that could be a good phylogenetic marker both for *Mesorhizobium* type strains and isolates identification, so the phylogenetic performance of the *dnaJ* gene was analysed. The relationships between the several *Mesorhizobium* type strains vary with the gene used to generate the phylogeny (Vinuesa *et al.*, 2005, Rivas *et al.*, 2007). No multilocus phylogeny was available for the *Mesorhizobium* genus, so several core genes were sequenced in the present study, as an attempt to generate a reliable phylogeny that could be compared to that of *dnaJ* and 16S rRNA gene. The multilocus analysis concatenated the sequence of five core genes, namely *atpD*, *gapA*, *gyrB*, *recA* and *rplB*, representing the first multilocus phylogeny for the *Mesorhizobium* genus that may contribute to clarify the phylogenetic relationships among species. Despite the different relative positioning of some type species, identification of chickpea rhizobia isolates based on *dnaJ* sequences is consistent with the previous identification based on the 16S rRNA gene. Furthermore, the *dnaJ*-based phylogeny allowed the differentiation of isolates that

according to the 16S rRNA phylogeny are identical. This differentiation is particularly relevant because these isolates (three from Viseu and two from Coimbra) belong to cluster A in the 16S rRNA phylogeny (higher sequence similarity with *M. huakuii*), but group apart from the type strains and we suspect that this group represents a new species.

The good performance of the *dnaJ* gene as phylogenetic marker for the genus *Mesorhizobium*, led us to investigate the phylogenetic signal of this gene at higher taxonomic levels, namely *Alphaproteobacteria* and *Proteobacteria*. In all proteobacteria analysed, the *dnaJ* gene was found to be encoded adjacent to *dnaK*, except in epsilonproteobacteria, for which no *dnaKJ* operon was found among the completely sequenced genomes. In general, the analysis on codon usage frequency and G+C and G+C3s content suggests that *dnaJ* is a core gene and not likely to be horizontally transferred between species. These are important features for a gene to be used as phylogenetic marker.

Despite minor discrepancies, a global congruence was found between the 16S rRNA gene and the DnaJ phylogeny for the *Proteobacteria* phylum. The agreement of the *dnaJ*-based phylogeny with other phylogenies based on multilocus data, indicate a good performance of *dnaJ* in reconstructing proteobacteria phylogeny.

Regarding the *Alphaproteobacteria* class, the coherence found between the topology of the dendrogram based on the DnaJ sequences and those from broader phylogenies using multiple protein coding genes (Ciccarelli *et al.*, 2006, Gupta & Sneath, 2007), shows that *dnaJ* has a good phylogenetic signal for the reconstruction of the alphaproteobacteria phylogeny. Contrary to the 16S rRNA gene phylogeny, the DnaJ-based phylogeny is consistent with the currently accepted relationships among genera and species obtained in phylogeny studies that use a large amount of sequence data (Gupta, 2005, Ciccarelli *et al.*, 2006, Gupta & Sneath, 2007).

For these reasons, we suggest that a single core gene, such as *dnaJ*, can be used as a phylogenetic marker for the phylum *Proteobacteria* and class *Alphaproteobacteria*, especially in the case of divergent 16S rRNA copies. At the genus *Mesorhizobium* level, *dnaJ* gene can be used for identification of isolates, in particular when the 16S rRNA gene analysis shows low resolution.

In order to study the tolerance of chickpea rhizobia to temperature stress, isolates representing all provinces of Portugal and Madeira Island were selected. Tolerance was evaluated in 53 isolates by growth in liquid medium under low (15ºC) and high temperature (37ºC) and after heat shock (48ºC, 15 min). In general, isolates grew

better under low than under high temperature, nevertheless tolerant and sensitive isolates to all conditions were found.

The four species groups of isolates were found to differ significantly regarding their ability to tolerate temperature stress. For example, isolates from the species cluster that include the two chickpea type strains display the two highest growth averages at 15ºC and are significantly different from each other and from the remaining species clusters ($P< 0.05$). In terms of tolerance to heat and heat shock these two species clusters (B and D) also display the highest growth averages. Furthermore, associations were found between some provinces of origin and tolerance to temperature stress of the isolates. For instance, the CA biplots associate isolates from the province Trás-os-Montes e Alto Douro with high tolerance to all tested temperature stresses.

The evaluation of chickpea rhizobia temperature tolerance, using diverse isolates and covering a large territory, allowed the identification of thermotolerant isolates with potential agronomical importance. Tolerant rhizobia strains may expand the use of these chickpea-symbionts to more adverse environmental conditions. It has been reported that temperature tolerant soybean rhizobia can improve nitrogen fixation in areas where temperature is the major factor limiting production (Zhang et al., 2003). Isolates tolerant to temperature stress and with a high symbiotic effectiveness were identified in the present work. For example, isolate PMI-1-Portimão has a high symbiotic effectiveness and is among the most tolerant isolates to all the tested conditions, so it represents a good candidate for chickpea field inoculation. In order to improve rhizobia tolerance to stress, it is important to understand the molecular bases of temperature stress tolerance in these soil bacteria.

DnaKJ and GroESL machineries are of major importance when coping with adverse conditions, especially heat stress, and interact with each other in their fundamental task of preventing irreversible misfold of cell proteins (Baneyx & Mujacic, 2004). This led us to investigate the expression levels of these major chaperone genes in several chickpea rhizobia species. The comparison of tolerant and sensitive isolates from the same species may reveal molecular mechanisms underlying the different susceptibility phenotypes to temperature stress. The transcriptional analysis of 15 chickpea rhizobia isolates, comprising three species groups, points towards a general increase of both *dnaK* and *groESL* mRNA levels after temperature upshift and irrelevant alteration of those levels after cold stress. Interestingly, after heat shock and heat stress, a generally higher induction of those chaperones transcripts was observed in tolerant isolates, compared to sensitive isolates, within the same species group. This is the first

study in mesorhizobia focusing on the transcriptional analysis of *dnaKJ* and *groESL* chaperone genes under different temperature stresses. It suggests the existence of a relationship between higher levels of transcriptional induction of major chaperone genes and higher ability of chickpea rhizobia isolates to endure heat stress, i.e. increased levels of chaperones may contribute to a higher tolerance to heat stress in rhizobia.

6.1 - Future Perspectives

Further studies are required to clarify the molecular mechanisms of temperature stress tolerance in rhizobia. In particular, it would be interesting to obtain knock-out mutants in the major chaperone genes, from both tolerant and sensitive strains and to evaluate the alterations in terms of temperature tolerance phenotype. On the other hand, chaperone overexpression studies could also be useful, in order to investigate if fundamental changes in the ability to endure temperature stress or other stresses are observed.

Other stress response genes should be investigated, for example by transcriptome analysis. Despite the fact that there is no completely sequenced genome from any chickpea rhizobia strain, *Mesorhizobium loti* MAFF30399 could be used in alternative, since it is a close phylogenetic relative. The completely sequenced genome of this *M. loti* strain allows the use of genome wide molecular techniques, as microarrays transcriptome analysis, which could indicate putative important genes involved in the tolerance to stress. This findings could be extrapolated to chickpea mesorhizobia and contribute to improve stress tolerance in strains to be used as inoculants.

6.2 – References

Baneyx, F. & M. Mujacic, (2004) Recombinant protein folding and misfolding in *Escherichia coli*. *Nat. Biotechnol.* **22**: 1399-1408.

Ciccarelli, F. D., T. Doerks, C. von Mering, C. J. Creevey, B. Snel & P. Bork, (2006) Toward automatic reconstruction of a highly resolved tree of life. *Science* **311**: 1283-1287.

Fierer, N. & R. B. Jackson, (2006) The diversity and biogeography of soil bacterial communities. *Proc. Natl. Acad. Sci. U. S. A.* **103**: 626-631.

Gupta, R. S., (2005) Protein Signatures Distinctive of Alpha Proteobacteria and Its Subgroups and a Model for Alpha-Proteobacterial Evolution. *Crit. Rev. Microbiol.* **31**: 101-135.

Gupta, R. S. & P. H. A. Sneath, (2007) Application of the Character Compatibility Approach to Generalized Molecular Sequence Data: Branching Order of the Proteobacterial Subdivisions. *J. Mol. Evol.* **64**: 90-100.

L'Taief, B., B. Sifi, M. Gtari, M. Zaman-Allah & M. Lachaal, (2007) Phenotypic and molecular characterization of chickpea rhizobia isolated from different areas of Tunisia. *Can. J. Microbiol.* **53**: 427-434.

Laranjo, M., C. Branco, R. Soares, L. Alho, M. Carvalho & S. Oliveira, (2002) Comparison of chickpea rhizobia isolates from diverse Portuguese natural populations based on symbiotic effectiveness and DNA fingerprint. *J. Appl. Microbiol.* **92**: 1043-1050.

Laranjo, M., J. Machado, J. P. W. Young & S. Oliveira, (2004) High diversity of chickpea *Mesorhizobium* species isolated in a Portuguese agricultural region. *FEMS Microbiol. Ecol.* **48**: 101-107.

Laranjo, M., R. Rodrigues, L. Alho & S. Oliveira, (2001) Rhizobia of chickpea from southern Portugal: symbiotic efficiency and genetic diversity. *J. Appl. Microbiol.* **90**: 662-667.

Maâtallah, J., E. B. Berraho, S. Muñoz, J. Sanjuan & C. Lluch, (2002) Phenotypic and molecular characterization of chickpea rhizobia isolated from different areas of Morocco. *J. Appl. Microbiol.* **93**: 531-540.

Nour, S. M., J.-C. Cleyet-Marel, P. Normand & M. P. Fernandez, (1995) Genomic heterogeneity of strains nodulating chickpeas (*Cicer arietinum* L.) and description of *Rhizobium mediterraneum* sp. nov. *Int. J. Syst. Bacteriol.* **45**: 640-648.

Nour, S. M., M. P. Fernandez, P. Normand & J.-C. Cleyet-Marel, (1994) *Rhizobium ciceri* sp. nov., consisting of strains that nodulate chickpeas (*Cicer arietinum* L.). *Int. J. Syst. Bacteriol.* **44**: 511-522.

Rivas, R., M. Laranjo, P. F. Mateos, S. Oliveira, E. Martinez-Molina & E. Velázquez, (2007) Strains of *Mesorhizobium amorphae* and *Mesorhizobium tianshanense*, carrying symbiotic genes of common chickpea endosymbiotic species, constitute a novel biovar (*ciceri*) capable of nodulating *Cicer arietinum*. *Lett. Appl. Microbiol.* **44**: 412-418.

Vinuesa, P., C. Silva, M. J. Lorite, M. L. Izaguirre-Mayoral, E. J. Bedmar & E. Martínez-Romero, (2005) Molecular systematics of rhizobia based on maximum likelihood and Bayesian phylogenies inferred from rrs, *atpD*, *recA* and *nifH* sequences, and their use in the classification of *Sesbania* microsymbionts from Venezuelan wetlands. *Syst. Appl. Microbiol.* **28**: 702-716.

Zhang, H., B. Prithiviraj, T. C. Charles, B. T. Driscoll & D. L. Smith, (2003) Low temperature tolerant *Bradyrhizobium japonicum* strains allowing improved nodulation and nitrogen fixation of soybean in a short season (cool spring) area. *European Journal of Agronomy* **19**: 205-213.

Supplementary data

Table S1 – Percentage of growth (average of three replicas ± standard deviation) of chickpea rhizobia for the different temperature stresses, considering growth at 28ºC as 100%.

Origin	Isolate	Cold stress (15ºC)	Heat stress (37ºC)	Heat shock (48ºC 15 min)
Trás-os-Montes e Alto Douro	BR-8-Bragança	50.5 ± 2.1	26.0 ± 1.2	73.1 ± 8.1
	BR-9-Bragança	68.3 ± 0.5	42.1 ± 1.7	89.5 ± 6.6
	BR-15-Bragança	54.3 ± 4.7	46.1 ± 0.3	88.9 ± 5.9
	BR-16-Bragança	52.8 ± 2.5	44.9 ± 0.5	73.2 ± 7.6
	BR-28-Bragança	41.8 ± 1.0	43.2 ± 1.4	41.0 ± 4.7
Douro Litoral	PII-1-Porto	42.3 ± 0.5	7.7 ± 0.4	23.0 ± 4.1
	PII-2-Porto	40.7 ± 2.2	5.8 ± 0.1	6.7 ± 0.8
	PII-3-Porto	47.5 ± 2.7	6.3 ± 0.9	17.0 ± 2.0
	PII-4-Porto	47.7 ± 3.1	6.0 ± 0.1	10.5 ± 1.4
Beira Litoral	C-1-Coimbra	37.1 ± 0.3	19.9 ± 0.7	83.0 ± 3.6
	C-7-Coimbra	22.4 ± 1.1	27.1 ± 1.8	21.6 ± 1.3
	C-14-Coimbra	54.3 ± 1.8	44.6 ± 4.3	83.7 ± 7.8
	C-27b-Coimbra	53.7 ± 0.4	19.9 ± 0.7	63.0 ± 1.3
	L-19-Leiria	22.6 ± 2.2	5.7 ± 0.0	10.4 ± 0.8
Beira Alta	V-5b.-Viseu	39.7 ± 1.1	19.7 ± 0.1	70.9 ± 3.0
	V-15b.-Viseu	20.6 ± 1.0	5.2 ± 0.3	30.6 ± 0.3
	V-18-Viseu	23.4 ± 1.7	6.6 ± 0.3	56.5 ± 4.6
	V-20-Viseu	38.3 ± 2.8	17.2 ± 2.0	80.2 ± 2.7
	V-25b.-Viseu	27.1 ± 0.2	13.6 ± 0.9	63.9 ± 2.3
Beira Baixa	CB-10-Castelo Branco	43.4 ± 0.9	25.4 ± 0.6	73.3 ± 3.5
	CB-19-Castelo Branco	47.7 ± 1.0	26.3 ± 0.3	51.5 ± 1.0
	CB-30-Castelo Branco	37.7 ± 1.3	18.0 ± 0.9	44.4 ± 1.4
	T-3-Telhado	48.2 ± 5.2	12.1 ± 0.1	8.1 ± 1.7
	T-8-Telhado	34.6 ± 1.0	10.8 ± 1.0	9.8 ± 2.1
Estremadura	CR-3-Caldas da Rainha	35.7 ± 2.3	16.1 ± 1.1	16.6 ± 3.7
	CR-16-Caldas da Rainha	31.8 ± 1.8	18.7 ± 1.3	6.3 ± 0.7
	CR-18-Caldas da Rainha	20.3 ± 0.9	14.2 ± 0.6	6.1 ± 0.2
	CR-32-Caldas da Rainha	23.5 ± 1.5	16.3 ± 0.4	43.3 ± 3.5
	ST-2-Setúbal	36.8 ± 1.4	17.1 ± 0.4	28.9 ± 1.9
Ribatejo	STR-2-Santarém	18.1 ± 2.5	9.7 ± 2.6	12.0 ± 0.7
	STR-4-Santarém	15.7 ± 2.3	14.1 ± 0.0	7.3 ± 5.2
	STR-10-Santarém	9.2 ± 0.9	7.2 ± 0.2	4.0 ± 0.0
	STR-14-Santarém	32.6 ± 2.1	15.8 ± 0.8	45.5 ± 4.7
	STR-16-Santarém	32.4 ± 2.0	16.7 ± 1.3	38.6 ± 1.6
Alto Alentejo	78-Elvas	8.3 ± 1.5	9.4 ± 0.2	8.1 ± 1.8
	85-Elvas	17.6 ± 0.9	9.4 ± 0.3	10.0 ± 1.5
	EE-14-ENMP	36.2 ± 1.3	9.8 ± 1.3	7.4 ± 4.9
	102-Évora	29.6 ± 4.8	15.1 ± 1.5	45.6 ± 0.6
	PT-35-Portalegre	30.3 ± 2.1	9.9 ± 0.7	50.3 ± 5.8
Baixo Alentejo	6b.-Beja	36.7 ± 0.7	24.2 ± 0.2	17.8 ± 1.9
	7a.-Beja	49.3 ± 0.5	16.8 ± 0.2	26.1 ± 2.2
	27-Beja	56.4 ± 1.6	21.0 ± 0.3	34.2 ± 2.3
	29-Beja	26.2 ± 1.4	17.0 ± 1.4	5.5 ± 0.8
	64b.-Beja	34.7 ± 2.0	18.0 ± 0.5	70.7 ± 2.6

Algarve	PM-1-Portimão	42.1 ± 1.8	20.7 ± 2.6	79.6 ± 4.8
	PM-14-Portimão	46.0 ± 1.3	21.7 ± 2.6	83.7 ± 3.3
	PM-17-Portimão	37.5 ± 2.1	15.0 ± 0.7	55.0 ± 3.8
	PMI-1-Portimãol	66.3 ± 2.0	30.6 ± 1.2	84.7 ± 5.0
	PMI-6-Portimãol	22.9 ± 0.9	12.0 ± 6.0	12.7 ± 0.6
Madeira	SA-9-Serra d'Água	26.6 ± 0.2	11.3 ± 0.5	2.7 ± 0.1
	SA-12-Serra d'Água	36.0 ± 2.6	13.7 ± 0.3	4.5 ± 0.5
	SA-13-Serra dÁgua	47.1 ± 2.7	13.4 ± 0.3	16.3 ± 1.2
	SA-17-Serra d'Água	46.2 ± 3.1	11.7 ± 0.9	23.9 ± 4.4
Type Strains	*M. ciceri* Ca7^T	49.4 ± 1.3	29.4 ± 2.3	87.1 ± 4.0
	M.mediterraneum Ca36^T	42.7 ± 1.0	20.4 ± 0.5	75.8 ± 2.6