

THESE

En vue de l'obtention du : **DOCTORAT**

Centre de Recherche : Biotechnologies Végétale et Microbienne, Biodiversité et Environnement

Structure de Recherche : Laboratoire de Microbiologie et de Biologie Moléculaire

Discipline : Biologie

Spécialité : Microbiologie et Biotechnologie Environnementale

Présentée et soutenue le 24/09/2022 par :

Meryem BELFQUIH

Génomique fonctionnelle pour l'étude des mécanismes d'adaptation aux stress salin et hydrique chez *Ensifer aridi*, une nouvelle espèce de rhizobium isolée du désert au Maroc (Merzouga)

JURY

Mohammed EL HASSOUNI	PES, Université Sidi Mohamed Ben Abdellah, Faculté des Sciences - Dhar EL Mahraz-Fès	Président/Rapporteur
Ilham BOUHMOUCH	PES, Université Mohammed V, Faculté des Sciences - Rabat	Rapporteuse/ Examinatrice
Imane WAHBY	PES, Université Mohammed V, Faculté des Sciences - Rabat	Rapporteuse/Examinatrice
Naima RHALLABI	PES, Université Hassan II, Faculté des Sciences et Techniques - Mohammedia	Examinatrice
Antoine LE QUERE	Expert, Dr, Université Montpellier, Institut de Recherche pour le Développement - France	Co-Directeur
Abdelkarim FILALI MALTOUF	PES, Expert, Académie Hassan II des Sciences et Techniques	Co-Directeur
Bouchra BELKADI	PES, Université Mohammed V, Faculté des Sciences - Rabat	Directrice de Thèse

Année Universitaire : 2021 /2022

Dédicaces

A mes chers parents

Mon défunt cher père Abdellah Belfquih et ma chère mère Mina Farayi

Aucun hommage ne saurait transmettre à sa juste valeur ma reconnaissance et mon profond amour. Je mets entre vos mains, le fruit de longues années d'études. Chaque ligne de cette thèse, chaque mot et chaque lettre vous exprime ma profonde gratitude pour votre grand soutien, dévouements exemplaires et les énormes sacrifices que vous m'avez consentis tout au long de ma vie.

A mon cher frère Hatim et ma chère sœur Fatima Zahra

Merci de m'avoir donné tant d'affection et de m'avoir soutenue tout le long de mes études. Qu'il soit pour vous le témoignage de ma grande reconnaissance pour vos encouragements, aide et soutien permanent.

A mon cher beau-frère Zine-Dine Miguel-Angel

Et ma très chère nièce Nora

Merci de m'avoir donné une nouvelle source de soutien et d'amour,

A mes chers amis et collègues au Maroc et en France

En souvenir des agréables moments passés ensemble, je vous dédie ce travail qui témoigne ma profonde reconnaissance. Vous avez été une bonne compagnie tout au long de la préparation de ma thèse. Merci d'avoir partagé avec moi les moments les plus difficiles de ce travail.

A mes Encadrants

et à Tous mes enseignants tout au long de mes études à la faculté des sciences Rabat

Toute ma famille

Tous ceux qui de près ou de loin, ont permis à cette thèse d'aboutir

Respect, gratitude, reconnaissance sont les moindres sentiments que je peux vous témoigner

Remerciements

Ce travail a été réalisé grâce au concours de nombreuses personnes auxquelles je tiens à rendre hommage :

*Un grand merci aux membres, permanents et stagiaires, mes professeurs chercheurs au Laboratoire de Microbiologie et biologie moléculaire (LMBM) pour leur accueil, leur gentillesse, leur sympathie et leur soutien infaillible, j'adresse ici aussi mes vifs remerciements à Mme Fatou DIOUF pour sa participation si joyeuse à mon travail pratique durant mes premiers mois de stage au laboratoire LMBM, son soutien et sa gentillesse ont été un appui et un réconfort précieux pour moi. Je tiens à remercier vivement Monsieur le Directeur du laboratoire LMBM et le Directeur du centre BioBio Monsieur **Jamal AURAG**.*

*Concernant mes séjours de thèse en France, je tiens dans un premier temps à remercier vivement Monsieur le Directeur de Recherche chez IRD Monsieur **Robin DUPONNOIS**, qui m'a invité à rejoindre son laboratoire des symbioses tropicales et méditerranéennes (LSTM). Qu'il trouve ici l'expression de mes sincères remerciements pour l'intérêt qu'il a bien voulu porter à notre collaboration. Sa grande expérience et ses précieux conseils durant mes séjours à Montpellier ont permis le bon déroulement et l'aboutissement de ce travail. Je lui suis également reconnaissante pour sa grande disponibilité et son dévouement incomparable.*

*Je tiens à remercier ma directrice de thèse, Madame, **Bouchra BELKADI** Professeur à la Faculté des Sciences de Rabat pour sa bienveillance et son aide à l'élaboration et le suivi de ce projet malgré ses nombreuses occupations. Vos critiques et vos encouragements m'ont été toujours précieux. Aucun remerciement ne saurait exprimer ma profonde reconnaissance et ma grande estime.*

*Je remercie profondément Monsieur **Abdelkarim FILALI-MALTOUF** co-directeur de thèse, Professeur à la Faculté des Sciences Rabat, Expert et Membre de l'Académie Hassan II pour tous les précieux conseils, les encouragements et le soutien. Pour le partage de ses connaissances et ses expériences dans la recherche scientifique, tout en m'accordant sa confiance. Veuillez cher professeur bien trouver et accepter mes sincères remerciements pour votre temps et pour votre bienveillance et gentillesse*

*J'adresse mes sincères remerciements à mon co-directeur de thèse, Monsieur **Antoine LE QUERE**, Docteur chercheur à l'institut de Recherche pour le développement FRANCE, pour m'avoir permis de réaliser cette thèse dans les meilleures conditions possibles. Je vous remercie également pour l'intérêt et le suivi assidu dont vous m'avez fait preuve. Je suis très reconnaissante de votre disponibilité, vos conseils constructifs, et surtout vos encouragements constants à mon égard et la confiance que vous m'avez accordée. Veuillez trouver ici l'expression de ma respectueuse gratitude.*

*Mes plus vifs remerciements vont à Monsieur **Mohammed EL HASSOUNI**, Professeur à la Faculté des Sciences Dhar El Mahraz Fes, Université Sidi Mohamed Ben Abdellah, pour m'avoir fait l'honneur de présider le jury de cette thèse. Qu'elle veuille bien trouver ici l'expression de ma sincère reconnaissance.*

*J'exprime toute ma considération et mes remerciements sincères à Madame **Naima RHALLABI**, Professeur à la Faculté des Sciences Mohammedia, d'avoir aimablement accepté d'être examinatrice de ma thèse. Ses remarques et suggestions m'ont permis d'apporter des améliorations à la qualité de mon rapport.*

*Merci à Madame **Ilham BOUHMOUCH**, Professeur à la Faculté des Sciences Rabat a bien voulu accepter d'être le rapporteur de mon travail en dépit de ses tâches multiples. Qu'elle veuille trouver ici l'expression de mes remerciements les plus sincères et de ma profonde gratitude.*

*Je remercie profondément Madame **Imane WAHBY**, Professeur à la Faculté des Sciences Rabat, d'accepter d'être rapporteur de mon travail de thèse. Veuillez cher professeur bien trouver et accepter mes sincères remerciements pour votre temps et pour votre bienveillance et gentillesse.*

*Ensuite, j'exprime toute ma gratitude et tiens à remercier chaleureusement Monsieur le professeur **Marc LEPETIT** et les membres de mon équipe de recherche en France de m'avoir bien accueilli dans l'équipe Mécanismes d'adaptation et la symbiose aux contraintes de l'environnement et de me soutenir pendant mon séjour passé au laboratoire des symbioses tropicales et méditerranéennes. Merci pour vous et pour toutes les autres équipes du laboratoire français pour votre accueil chaleureux au sein de votre laboratoire à Montpellier.*

Je vous remercie aussi pour m'avoir fait partager vos nombreuses connaissances, votre détermination, votre disponibilité, vos soutiens sans failles, la confiance que vous m'avez donnée, et votre amitié.

*Je dois aussi une grande reconnaissance à tous mes chers amis et collègues qui m'ont apporté un soutien humain et/ou scientifique : **Ayoub Abbadi, Meriem Harrouch, Nassima Ait Lahmidi, Fatima zahra Maghnia, Nadia Boukraa, Basma Boukour, Sakina Bouifden, Myriam Benkirane, Sarra Machri, Meryem Makkaoui, Malika Oubohssaine, Chaimaa Yatrib, Kaoutar Taha, Imane Elattar, Imane chamkhi, Mohammed ElRoubi, Mohamed Ait Hamza, Mohamed Hnini, karim Rabeh, Houcine Ouakrim, Mehdi Belouad,***

*Je remercie chaleureusement les responsables du projet **Allocations de recherche pour une thèse au Sud (ARTS)** pour m'avoir octroyé des bourses qui m'ont permis de réaliser les différents stages en France.*

Enfin, bien qu'il me soit impossible de citer tous ceux qui m'ont apporté leur soutien, de près ou de loin, je voudrais cependant qu'ils soient assurés de ma reconnaissance

Résumé

La bactérie fixatrice d'azote *Ensifer aridi*, isolée du désert Marocain, est dotée de capacités de survie multiples. Elle s'adapte à des conditions environnementales particulièrement difficiles. Cette nouvelle espèce représente un modèle d'étude intéressant pour étudier les mécanismes d'adaptation aux environnements difficiles (Sakrouhi et al., 2016). L'étude transcriptomique RNAseq a généré des données sur des gènes différentiellement exprimés dans les situations de stress salin et hydrique. Dans cette étude des mutants ont été générés sur les gènes sélectionnés, avec une analyse de leurs transcriptions et une caractérisation comparative de leurs réponses phénotypiques. La mutation des gènes, *nesR* et *sahR*, a affecté drastiquement la symbiose, la tolérance au sel, la motilité et la tolérance aux stress oxydatifs, aux acides et aux détergents des mutants. En revanche la double mutation des deux gènes régulateurs a généré une tolérance accrue aux stress détergents, oxydatifs et acides, couplée avec une diminution de la capacité de tolérance au sel et la symbiose. La mutation des gènes impliqués dans la biosynthèse et le transport du tréhalose, *otsB* et *thuR*, a amélioré la symbiose et la croissance des deux mutants dans un milieu hyper-osmotique. La mutation de *rpoE2* n'a pas affecté la tolérance au stress salin et la capacité symbiotique. Suggérant donc que la bactérie *E. aridi* implique des mécanismes alternatifs que le système Général Stress Réponse pour s'y adapter au stress. La mutation du gène *iolR*, répresseur des gènes impliqués dans le catabolisme d'inositol, n'a pas généré une différence phénotypique. Finalement, l'analyse de l'expression des gènes par l'utilisation des fusions *gfp*-promoteur suggère des circuits d'interactions régulatrices très complexes.

Mots-clefs : *Ensifer aridi*, hyper-osmoadaptation, ARN seq, métabolismes, génomique fonctionnelle.

Abstract

The nitrogen-fixing bacterium *Ensifer aridi*, isolated from the Moroccan desert, has multiple survival capacities. It adapts to particularly difficult environmental conditions. This new species represents an interesting study model for studying the mechanisms of adaptation to harsh environments (Sakrouhi et al., 2016). The RNAseq transcriptomic study generated data on differentially expressed genes in saline and water stress situations. In this study mutants were generated on the selected genes, with an analysis of their transcriptions and a comparative characterization of their phenotypic responses. The mutation of the genes, *nesR* and *sahR*, has drastically affected symbiosis, salt tolerance, motility and tolerance to oxidative stresses, to acids and to detergents of the mutants. On the other hand, the double mutation of the two regulatory genes generated an increased tolerance to detergent, oxidative and acid stresses, coupled with a decrease in salt tolerance capacity and symbiosis. The mutation of genes involved in the biosynthesis and transport of trehalose, *otsB* and *thuR*, improved the symbiosis and growth of the two mutants in a hyper-osmotic medium. The *rpoE2* mutation did not affect saline stress tolerance and symbiotic ability. Suggesting, therefore that the bacterium *E. aridi* involves alternative mechanisms such as the General Stress System Response to adapt to stress. The mutation of the *iolR* gene, a repressor of genes involved in inositol catabolism, did not generate a phenotypic difference. Finally, the analysis of gene expression using *gfp*-promoter fusions suggests very complex circuits of regulatory interactions.

key-words: *Ensifer aridi*, hyper-osmoadaptation, ARN seq, metabolisms, genomic function.

Résumé détaillé

La bactérie fixatrice d'azote *Ensifer aridi* LMR001, isolée du désert de Merzouga au Maroc, est dotée de capacités de survie multiples. Elle s'adapte à des conditions environnementales particulièrement difficiles : salinité élevée, fluctuations de la température et la dessiccation. D'où ce fait, cette nouvelle espèce représente un modèle d'étude intéressant pour étudier les mécanismes d'adaptation à ces types d'environnement (Sakrouhi et al., 2016). L'étude transcriptomique faite sur cette bactérie, par l'utilisation d'une approche globale sans a priori de RNAseq, a généré des données sur des gènes différentiellement exprimés dans les situations de stress salin et hydrique et impliqués dans plusieurs métabolismes et fonctions cellulaires. Par exemple, l'accumulation du tréhalose ; la production des exopolysaccharides ; la mobilité ; la chimiotaxie ; la régulation du cycle de la méthionine ; le catabolisme d'inositol et d'autre source de carbone ou la régulation de la réponse générale au stress GSR ont ainsi pu être identifiés. Afin d'explorer les rôles de ces métabolismes et fonctions dans l'adaptation au stress, dans cette étude des mutants ont été générés sur les gènes sélectionnés, avec une analyse de leurs transcriptions et une caractérisation comparative de leurs réponses phénotypiques. Nous avons constaté que la mutation des gènes régulateurs, *nesR* et *sahR*, du cycle de la méthionine a affecté drastiquement la symbiose, la tolérance au sel, la motilité en présence de NaCl et la tolérance aux stress oxydatifs, acides et détergents. La double mutation des deux gènes régulateurs de la méthionine cycle a généré une tolérance accrue aux stress détergents, oxydatifs et acides, couplée avec une diminution plus sévère de tolérance au sel et un impact plus important sur le développement de la symbiose comparée aux simples mutants. Nous avons également relevé que la mutation des gènes impliqués dans le transport et l'utilisation du tréhalose *thuR* et dans la biosynthèse du tréhalose *otsB*, a amélioré la symbiose et la croissance des deux mutants dans un milieu hyper-osmotique. Ces résultats suggèrent le rôle important du tréhalose dans l'amélioration des capacités symbiotiques des microbes et de tolérance aux stress sévères. En revanche la mutation de *rpoE2*, le gène codant pour le sigma facteur régulateur du système de la réponse générale au stress, n'a pas affecté la tolérance du mutant au stress salin et sa capacité à développer une symbiose. Ces données suggèrent donc que la bactérie *E. aridi* implique des mécanismes alternatifs que le système Général Stress Réponse GSR (RsiB1, RsiA1, RpoE2) pour s'adapter aux changements environnementaux. La mutation du gène *iolR*, répresseur des gènes impliqués dans le catabolisme d'inositol n'a pas généré une différence phénotypique par rapport à la souche sauvage. Finalement, l'analyse de l'expression des gènes régulateurs et des gènes régulés par l'utilisation des fusions transcriptionnelles de *gfp* aux promoteurs suggère des

circuits d'interactions régulatrices très complexes entre les voies métaboliques de l'inositol, méthionine, tréhalose et GSR.

Mots-clefs : *Ensifer aridi*, hyper-osmoadaptation, ARN seq, métabolismes, génomique fonctionnelle

Abstract detailed

Ensifer arid LMR001, a new species nitrogen-fixing bacterium isolated from desert of Merzouga in Morocco, has multiple survival capacities. It adapts to particularly difficult environmental conditions: high salinity, temperature fluctuations and desiccation. Hence, this new species represents an interesting model to study mechanisms of adaptation related to these types of environment (Sakrouhi et al., 2016). A transcriptomic study carried out on the *Ensifer aridi* LMR001, by using the RNAseq technique revealed genes involved in several metabolisms and functions were differentially expressed in response to salt and water stress. Among others, genes responsible for the accumulation of trehalose; the production of exopolysaccharides; mobility; chemotaxis; regulation of the methionine cycle; inositol and other carbon source catabolism and regulation of the general GSR stress response were found. In order to explore the roles of these metabolisms and functions in adaptation to stress, mutants generated on selected genes were generated, their transcriptions were analysed and a comparative characterization of their phenotypic responses was done. We found that the single mutation of the regulatory genes of the methionine cycle, *nesR* and *sahR*, has drastically affected symbiosis capacity, salt tolerance, motility in the presence of NaCl, and tolerance to oxidative, acid, and detergent stresses. While the double mutation of the two same regulatory genes generated an increased tolerance to detergent, oxidative and acidic stresses, coupled with a more severe decrease of tolerance to salt and a more severe capacity for symbiosis development compared to the single mutants. We also noted that the mutations of the trehalose transport and utilization regulator *thuR* or the trehalose biosynthesis *otsB* enhanced the symbiosis and growth of the two mutants under hyper-osmotic stress. On the other hand, the mutation of *rpoE2*, the gene coding for the sigma factor regulator of the general stress response system, did not affect the mutant's tolerance to salt stress and its ability to develop symbiosis. This suggests that the *E. aridi* bacterium involves other alternatives than the General Stress Response GSR system (RsiB1, RsiA1 and RpoE2) to adapt to environmental changes. The mutation of the *iolR* gene, a repressor of the catabolic inositol pathway, didn't generate a difference in phenotypic responses compared to the wild strain. Finally, the analysis of the expression of regulatory genes and genes regulated by the use of transcriptional fusions of the promoters to *gfp* provided very complex regulatory interaction circuits between the metabolic pathways of inositol, methionine, trehalose and the regulator GSR transcriptional factors.

key-words: *Ensifer aridi*, hyper-osmoadaptation, ARN seq, metabolisms, genomic function

Productions scientifiques

Articles publiés :

Meryem Belfquih, Abdelkarim Filali-Maltouf, et Antoine Le Quéré. 2022. « Analysis of *Ensifer aridi* Mutants Affecting Regulation of Methionine, Trehalose, and Inositol Metabolisms Suggests a Role in Stress Adaptation and Symbiosis Development» *Microorganisms* 10 (2): 298. <https://doi.org/10.3390/microorganisms10020298>.

Meryem Belfquih, Ilham Sakrouhi, Hassan Ait-Benhassou, Emeric Dubois, Dany Severac, Abdelkarim Filali-Maltouf, et Antoine Le Quere. 2021. « *Ensifer aridi* LMR001T Symbiosis and Tolerance to Stress Do Not Require the Alternative Sigma Factor RpoE2 ». *Agronomy* 11 (9): 1787. <https://doi.org/10.3390/agronomy11091787>.

Sakrouhi, Ilham, **Meryem Belfquih**, Laïla Sbabou, Patricia Moulin, Gilles Bena, Abdelkarim Filali-Maltouf, et Antoine Le Quéré. 2016. « Recovery of Symbiotic Nitrogen Fixing *Acacia Rhizobia* from Merzouga Desert Sand Dunes in South East Morocco – Identification of a Probable New Species of *Ensifer* Adapted to Stressed Environments». *Systematic and Applied Microbiology* 39 (2): 122-31. <https://doi.org/10.1016/j.syapm.2016.01.001>.

Communications:

Meryem Belfquih, Ilham Sakrouhi, Jawahir Zemrani, Antoine Le Quere, Abdelkarim Filali-Maltouf, **2020**. Regulation of Flagellar Biosynthesis, Motility And Chemotaxis In Response To Cell Membrane Of *Ensifer aridi* Under Stress Effect. Congress of the University Sidi Mohamed Ben Abdelah (faculty of sciences Fes) with Biotech Association. (Biotech), Fes (Morocco), 27-29 February 2020.

Meryem Belfquih, Ilham Sakrouhi, Hassan Ait-Benhassou, Emeric Dubois, Dany Severac, Abdelkarim Filali-Maltouf, et Antoine Le Quere, **2019**. Genomic and phenotypic characterization of sigma factor RpoE2 general stress response role in *Ensifer aridi* bacteria isolated from morocco. Congress of the Moroccan association of Microbiology (AMM) with the university Al Akhawayn, Microbial biotechnology and socio-economic development and opportunities, (BIOMIODS), Ifrane (Morocco), 16-18 October 2019

Meryem Belfquih, Abdelkarim Filali-Maltouf, et Antoine Le Quere, **2018**. Climate risk study: (Drought, flood, marine submersion and Storm) and development of sustainable development objectives. Second Summer School of Sustainable Development Goals (SDGs), organized by the research institute for development (IRD), in partnership with Aix-Marseille University (France). Marseille (FRANCE), 2-7 July 2018

Meryem Belfquih, Ilham Sakrouhi, Abdelkarim Filali-Maltouf, et Antoine Le Quere, **2017**. Study of the mechanisms involved in the response of the symbiotic desert bacterium *Ensifer aridi* LMR001 to hyperosmotic stress. Congress of the Center of the international cooperation in agronomic research for development, Montferrier-Baillarguet. Le Printemps Baillarguet 9^{ème} édition (CIRAD). Montpellier (France), 14 April 2017

Liste des figures

Figure 1: Schéma représentant les étapes de l'infection des légumineuses par <i>Rhizobium leguminosarum</i> . a présente les premiers stades de l'infection des poils absorbants par les rhizobiums et la croissance des nodules sur les racines de la plante.	8
Figure 2: Présentation du processus de la nodulation chez les plantes qui forment des nodules indéterminés (type de nodule formé par les bactéries du genre <i>Ensifer</i>)..	10
Figure 3: L'alignement génomique des 6 génomes d' <i>Ensifer</i> .sp du désert.	16
Figure 4: Neighbor Joining arbre montrant des relations phylogénétiques entre 31 souches rhizobiennes d' <i>Ensifer</i> .sp..	19
Figure 5: Présente la distribution géographique mondiale de l'espèce <i>Vachellia</i> . Les deux espèces <i>V.farnesiana</i> , <i>V.karro</i> et <i>V.nilotica</i> ne sont pas montré sur la carte géographique de l' <i>Australie</i> , pour bien présenter que la distribution des espèces de <i>Vachellia</i> native du pays. Les données sont publié sur le site www.gbif.org (Comben et al., 2020).	22
Figure 6: Représentation du mécanisme majeur de la régulation, PhyP-PhyR-NepR-σE, de la transription des ARN liés au stress.	119
Figure 7: Représentation schématique de l'organisation des régions chromosomiques de la région des gènes <i>Sigma-EcfG</i> (<i>rpoE2</i>), <i>hyR</i> (<i>rsiB</i>), <i>nepR</i> (<i>rsiA</i>) et <i>phyK</i> ou <i>phyP</i> (sensor kinase) dans la souche <i>E.aridi</i> LMR001 par rapport à <i>E.meliloti</i> 1021.	121

Liste des abréviations

°C : degré Celsius

ADN : Acide désoxyribonucléique

AdoHcy: adénosine homocystéine

AdoMet: adénosine méthionine

ANI : average nucliotide identity

ANOVA : Analyse de la variance

ARN : acide ribonucléique

ARNm : acide ribonucléique messenger

ARNs: les acides ribonucléiques

ATP: adénosine triphosphate

BLAST: Basic local alignment search tool

C : carbone

CDS : les séquences codantes

CLC Genomics: cross-platform desktop software compatible with computer

CO: monoxide de carbone

dDDH: digital AND-AND hybridation

EDTA : acide éthylène diamine tétra-acétique

EPS I: exopolysaccharide type 1

EPS II: exopolysaccharide type 2

EPS: exopolysaccharide

GPS: Global Positioning System

HCD: high cell density

Hiseq 2000: Illumina high sequencing 2000 system

IRD : institut de recherche pour le développement

Kpb: Kilo Paire de base

LCD: low cell density

LMW: low Molecular weight

LPS: lipopolysaccharide

MAUVE system: Alignment software for comparing two or more draft or finished *genomes*.

Mb : millions de paires de bases

MCP : Methyl Accepting Chemotaxis Proteins

MLSA : multilocus sequence analysis

Mpb: millions de paires de bases

NaCl: chlorure de sodium

NADP : nicotinamide adenine dinucleotide phosphate

Nap : nitrate réductase périplasmique

NCBI: national center for biotechnology information
NGS: next generation sequencing
Nir: nitrite réductase
NJ: Neighbor Joining
NO: monoxyde d'azote
Nod: protéine Nod
O2: dioxygène
OH: Hydroxyle
pb: Paire de base
PEG : polyethylene glycol
pH: potential of hydrogen
Phase G1: phase of growth 1
Pi : phosphate inorganique
pSymA: plasmide symbiotique A
pSymB: plasmide symbiotique B
QS : quorum sensing
RAST: rapid annotations using subsystems technology
RNA- seq: acide ribonucléique séquençage
SAM: S-adenosyl-methionine
T6P: tréhalose 6 phosphate
Tre6P: tréhalose 6 phophate
TY: Yeast Tryptone
UI: unite internationale
UV: ultraviolet
WGS: Whole Genome Shotgun
YM: Yeast Mannitol

SOMMAIRE

DEDICACES	I
REMERCIEMENTS	II
RESUME.....	V
ABSTRACT	VI
RESUME DETAILLE.....	VII
ABSTRACT DETAILED	IX
PRODUCTIONS SCIENTIFIQUES.....	X
LISTE DES FIGURES	XII
LISTE DES ABREVIATIONS	XIII
INTRODUCTION GÉNÉRALE	1
CHAPITRE I.....	5
REVUE BIBLIOGRAPHIQUE.....	5
1. Diversité des organismes fixateurs d'azote : les rhizobiums	6
1.1 Symbiose fixatrice d'azote : <i>Rhizobium-légumineuse</i>	7
1.2 Symbiose fixatrice d'azote : <i>Ensifer-légumineuse</i>	8
2. Caractérisation génomique, phylogénique du genre <i>Ensifer</i>	12
2.1 Description du genre <i>Ensifer</i>	12
2.2 Génome de <i>Ensifer</i>	12
3. Caractérisation génomique, phylogénique et origine de <i>Ensifer aridi</i> sp. Nov.....	14
3.1 Description de la nouvelle espèce <i>E.aridi</i>	14

3.2 Effet du Stress sur la souche <i>Ensifer aridi</i> LMR001	19
4. La symbiose dans les environnements sous contraintes abiotiques	21
4.1 <i>Vachellia</i> , des espèces d'arbres d'intérêt pour le Maroc (<i>V tortillis</i> et <i>V gummifera</i>)	21
4.2 Bactéries fixatrices d'azote symbiotiques <i>Ensifers</i> tolérance et exigence	23
5. Mécanismes d'osmo-adaptation chez les bactéries	24
5.1 Effet des facteurs abiotiques sur l'activité bactérienne	24
5.2 Réponse générale au stress	26
5.3 La stratégie « salt in cytoplasm », une stratégie spécifique aux micro-organismes halotolérants (extrêmes)	27
5.3.1 Les mécanismes d'accumulation intracellulaire de solutés compatibles solubles osmotiquement actifs	27
5.3.1.1 Rôle des ions potassium (la réponse primaire)	27
5.3.1.2 Les solutés organiques compatibles (les réponses secondaires)	28
CHAPITRE II	30
LES BACTERIES FIXATRICES D'AZOTE ISOLEES DU DESERT AU MAROC ET SYMBIOTIQUE DE VACHELLIA POURRAIENT APPARTENIR A UNE NOUVELLE ESPECE DE SINORHIZOBIUM (ENSIFER)	30
Article 1, Sakrouhi et al., 2016: Recovery of symbiotic nitrogen fixing acacia rhizobia from Merzouga Desert sand dunes in Southeast Morocco -Identification of a probable new species of Ensifer adapted to stressed environments.	31
1. Introduction	69
2. Résultats et discussion	69
3. Conclusion	70
CHAPITRE III	71
ÉTUDE TRANSCRIPTOMIQUE ET ANALYSE FONCTIONNELLE DU GENE <i>RPOE2</i> REGULATEUR IMPLIQUE DANS LA REPOSE GENERALE AU STRESS CHEZ <i>ENSIFER ARIDI</i> EN REPOSE AU STRESS SALIN	71

Article 2, Belfquih et al., 2021: <i>Ensifer aridi</i> LMR001T Symbiosis and Tolerance to Stress Do Not Require the Alternative Sigma Factor RpoE2	72
1. Introduction.....	117
2. Résultats et discussion	117
3. Conclusion.....	123
CHAPITRE IV	125
APPROCHE DE MUTAGENESE DIRIGEE DES GENES CIBLES DANS LE METABOLISME DE LA METHIONINE, DU TREHALOSE ET D'INOSITOL PAR L'ETUDE TRANSCRIPTOMIQUE ET ANALYSE FONCTIONNELLE DES GENES DE REGULATION	125
Article 3, Belfquih et al., 2022: Analysis of <i>Ensifer aridi</i> Mutants Affecting Regulation of Methionine, Trehalose, and Inositol Metabolisms Suggests a Role in Stress Adaptation and Symbiosis Development	126
1. Introduction.....	161
2. Résultats et discussion	161
3. Conclusion.....	167
CHAPITRE V.....	169
DISCUSSION, CONCLUSION GENERALE ET PERSPECTIVES	169
1. Discussion.....	170
2. Conclusion générale et perspectives	181
RÉFÉRENCES BIBLIOGRAPHIQUES.....	183
ANNEXES : MATERIEL ET METHODES UTILISEES DANS CE TRAVAIL DE THESE	195

1.	Culture bactérienne et test phénotypique : Tolérance aux stress abiotiques	195
1.1	Stress salin en milieu solide/ en milieu liquide	195
1.2	Test du stress acide (HCL), détergent (SDS) et oxydatif (H₂O₂)	195
1.3	Test de motilité	195
1.4	Test de nodulation.....	196
2.	Étude transcriptomique.....	196
3.	Mutagenèse dirigée et le clonage des promoteurs de gènes ciblés	198
	A. PRINCIPE : mutagenèse dirigée par délétion et construction des promoteurs.....	198
	B. METHODOLOGIE : Mutagenèse dirigée par PCR et construction des promoteurs.....	199
	B.1. Dessin d’amorce et architecture moléculaire pour l’amplification et l’obtention des parties 5’ et 3’ du gène d’intérêt à muter par PCR successives.....	199
	B.2. Création d’un antibiogramme d’<i>Ensifer aridi</i>	200
	B.3. Extraction et purification des acides nucléiques	200
	B.4. Caractérisation moléculaire	202
	B.5. Création de l’amplicon de la séquence H1-H2 obtenus par PCR	204
	B.6. Manipulation des fragments amplifiés	206
	B.7. Transformation bactérienne.....	209
	B.8. Mutagenèse par double recombinaison homologue	211
4.	Étude de l’expression de gènes cibles en utilisant le gène rapporteur <i>gfp</i>	214

INTRODUCTION GÉNÉRALE

Le Maroc, dans son aire géographique, bénéficie de deux types de bioclimats: méditerranéen au nord et aride au sud (Grouzis et al., 2003). Les zones arides et semi arides sont généralement caractérisées par des conditions de vie difficiles : une carence en éléments nutritifs, une dominance des facteurs abiotiques très stressants pour les formes vivantes (une salinité élevée, un manque d'eau, une température et une dessiccation extrêmes), une instabilité physique des sols causées par des vents forts et une exposition permanente aux rayons UV. Tous ces facteurs agissent directement sur la structure, la qualité et la fertilité du sol et donnent une production agricole drastiquement fragilisée (Ayangbenro et al., 2021). La problématique actuelle est que l'aridité du sud marocain s'accroît et gagne de plus en plus en surface en se propageant vers le nord à cause de l'accumulation des effets de serre (Grouzis et al., 2003). Actuellement, les zones arides abritent environ 2 milliards de personnes et occupent une superficie estimée à environ 66,7 millions km² de la planète. Ces chiffres devraient aussi augmenter encore en taille dans le futur, en raison des scénarios climatiques actuels (Botumile et al., 2020). L'évolution des effets climatiques avec la persistance des activités humaines non responsables vont de plus faire accélérer les changements climatiques et stimuler encore la dégradation des écosystèmes avec une détérioration de la qualité des sols, une diminution de la couverture végétale et une augmentation des taux d'érosion des sols. Ceci va générer une diminution rapide des teneurs en éléments nutritifs des sols et des pertes énormes dans la production agricole avec une fragilisation de la biodiversité des habitats sauvages (Grouzis et al., 2003, Ayangbenro et al., 2021). Les causes sous-jacentes de la dégradation de ces zones arides fragiles sont cependant révélées étroitement liées aux facteurs qui se renforcent mutuellement, à savoir la pauvreté, la croissance démographique, la faible croissance économique (Botumile et al., 2020). Ceci impose la nécessité d'introduire en urgence un programme de protection de ces zones définies essentielles à l'équilibre de l'écosystème, par la mise en place d'approches écologiques opérationnelles et particulièrement adaptées à chaque problématique (Ayangbenro et al., 2021).

Au cours de ces 50 dernières années, avec le réchauffement de la planète, les températures de surface de la biosphère ont augmenté de 0,6 °C et les prévisions indiquent que la température va continuer à augmenter de façon exponentielle dans les prochaines années. Avec l'augmentation de la température ; beaucoup d'autres variables changeront, telles que la pression atmosphérique, l'humidité du sol, la fonte des neiges, la montée de niveau des océans et la dégradation des grandes forêts avec une abondance des zones arides et bien plus encore (Ayangbenro et al., 2021).

Maintenant, les retours à l'équilibre initial des écosystèmes à grande échelle sont déclarés impossibles et l'augmentation de la température devrait engendrer une foule d'autres variabilités. Cependant, il est encore possible de stabiliser voir de limiter ces changements (Balser et al., 2006). Jusqu'à présent, les terres arides n'ont pas bénéficié de toute l'attention nécessaire. Toutes les stratégies employées se heurtent trop souvent sur des connaissances insuffisantes, concernant la biodiversité et le rôle écologique de chaque espèce vivant dans ces écosystèmes (Grouzis et al., 2003). Le manque de données des dégâts sur ces zones est un obstacle, qui limite la progression de la mise en œuvre des stratégies du développement durable (Rocha et al., 2020). Les diverses recherches faites sur les zones arides ont été concrétisées, au Sahel, par la mise en œuvre des programmes d'aménagement sylvopastoraux et agroforestiers. Des plantations expérimentales voire industrielles des plantes de *Retama sp.*, *Vachellia sp.*, *Lotus sp.*, *Lupinus sp.*, *Medicago sp.*, etc, ont été installées au Nord-Ouest de l'Afrique, pour fertiliser des sols et stabiliser des dunes et lutter contre la désertification (Rejili et al., 2020, Botumile et al., 2020). Ce projet est considéré comme le « miracle du Sahel » en raison de ses capacités à stabiliser et améliorer la qualité des sols et à produire d'importante quantité de fourrage à partir des arbres comme *Vachellia.sp.* Au Maroc, deux espèces d'arbres légumineuses originaires des zones arides et subhumides, de l'Australie occidentale et l'Afrique du Sud, appelées respectivement *V. cyanophylla* et *V. horrida* sont utilisées avec succès pour la stabilisation des dunes et le reboisement des zones arides (Zerhari et al., 2000).

La flore sauvage des légumineuses du nord-ouest de l'Afrique représente une diversité spécifique et intra-spécifique très riche avec une abondance large en surface d'occupation (Rejili et al., 2020). Cette présence importante peut être utilisée pour augmenter la disponibilité de l'eau et aussi fertiliser le sol via la formation d'interaction symbiotique avec les microorganismes bénéfiques dans les zones drastiquement dégradées (Grouzis et al., 2003). De ce fait, les légumineuses sélectionnées et installées dans les sols pauvres vont agir avec un effet direct sur le développement de la strate herbacée, sur le plan de la diversité floristique et sur le plan de la phyto-masse, et enrichir la terre en azote et en éléments nutritifs. Actuellement, la biodiversité microbienne des rhizobiums dans ces zones non-stables est devenue très rare et considérée en voie de disparition (Rejili et al., 2020). Ces mêmes conditions climatiques sont aussi un avantage sélectif très important. La pression sélective agit en permanence sur la conservation des ressources en micro-organismes spécifiquement adaptées aux stress extrêmes (Zerhari et al., 2000). Ainsi, le chevauchement entre les effets de ces conditions abiotiques croissantes et la pression anthropique génère une pression sélective sur la biodiversité locale et donne par conséquence une évolution adaptative du patrimoine génétique de ces micro-organismes dans ces environnements (Grouzis et

al., 2003). Ces écosystèmes complexes abritent donc des divers types de micro-organismes rares et à intérêt, avec des propriétés génétiques, métaboliques et physiologiques riches. Ceci suggère que les micro-organismes fertilisants isolés de ces sols sont porteurs des divers vertus métaboliques et peuvent être utilisés pour améliorer les propriétés physico-chimiques des sols des zones arides et participer au maintien et à l'équilibre durable de ces écosystèmes (Ayangbenro et al., 2021). D'où l'importance d'introduire un programme scientifique pour isoler, sélectionner et caractériser des partenaires rhizobiums adaptées aux zones arides et étudier les possibilités des interactions bénéfiques avec les plantes locales (Rocha et al., 2020).

Les embranchements bactériens, que l'on trouve généralement dominants dans ces zones arides et semi-arides, comprennent les *Actinobacteria*, les *Bacteroidetes*, les *Firmicutes*, les *Proteobacteria* et les *Cyanobacteria*. D'autres non dominants comprennent les phylums, comme les *Chloroflexi*, les *Deinococcus*, les *Verrucomicrobiae* et les *archaea phyla* (qui sont *Euryarchaeota*, *Crenarchaeota* et *Thaumarchaeota*) (Ayangbenro et al., 2021). Ces espèces bactériennes indigènes de ces régions adoptent des mécanismes d'adaptation biochimiques, physiologiques et moléculaires complexes, pour atténuer les différents stress biotiques et abiotiques et prospérer dans des conditions aussi sévères. Par exemple, chaque microorganisme réagit différemment aux changements d'humidité. Les bactéries Gram-négatives sont considérées sensibles aux changements d'humidité du sol, tandis que les champignons sont considérés comme plus tolérants à une faible disponibilité en eau. L'efficacité des microorganismes est aussi réduite par les températures élevées du sol, le stress salin et osmotique, l'acidité du sol, l'alcalinité et les métaux lourds (Burghardt et al., 2020, Ayangbenro et al., 2021). De ce fait, une application réussie des micro-organismes sélectionnés contribuera indirectement à l'amélioration de la rétention d'eau du sol, le stockage du carbone par le biais de la fertilisation des plantes et contribuera directement à la stabilité des niveaux de la disponibilité des éléments nutritifs, à l'élimination des polluants, au maintien de la santé des sols et la conservation de la biodiversité floristique locale. Ces organismes offrent également des opportunités dans le contrôle des maladies transmises par le sol et la dégradation de la matière organique récalcitrante (Ayangbenro et al., 2021). Pour cela, la connaissance de la diversité des populations bactériennes dans ces types de sols et la sélection des souches à caractères spécifiques (compétitivité, adaptation facile à certains environnements, efficacité de l'assimilation d'azote et des éléments nutritifs) est devenue une nécessité urgente pour augmenter les bénéfices de l'inoculation et promouvoir la durabilité de la biodiversité dans les zones arides (Elbouthiri et al., 2010, Burghardt et al., 2020). Ce projet a été établi pour approfondir un certain nombre de questions clés liées aux caractéristiques adaptatives des espèces de rhizobiums spécifiquement adaptées aux conditions arides, et à intérêts écologiques voir

agronomiques. Des souches à intérêt ont été isolées, évaluées et étudiées afin d'être testées et utilisées comme inoculums dans les grandes zones arides du Maroc et ailleurs dans d'autres zones dégradées. De plus, des avancées méthodologiques ont ouvert par la suite de nouvelles pistes de recherche. Ainsi, les études phylogénétiques et phénotypiques ont suggéré que les bactéries isolées de sable de dunes du désert de Merzouga nodulant deux espèces de *Vachellia* appartenaient à une nouvelle espèce appartenant au genre *Ensifer* (*Sinorhizobium*) (Sakrouhi et al., 2016). Ce programme a aussi fait l'objet de recherches des déterminants génétiques liés à des capacités adaptatives notamment liées aux stress hyperosmotique de ces nouveaux rhizobiums à l'état libre dans le sol et à la symbiose. Une détermination des traits physiologiques et phénotypiques des bactéries a été faite, avec une étude génomique et métabolique inter-intra espèce. Ces études ont permis de relever des nouvelles connaissances sur des gènes et des traits fonctionnels pertinents, et indique que l'environnement (semi-) désertique influe sur la répartition géographique des souches appartenant à cette nouvelle dont le nom *Ensifer aridi* a été proposé (Le Quéré et al., 2017). En effet, le séquençage à haut débit du génome de 6 souches appartenant à *Ensifer aridi*, et l'étude transcriptomique effectuée sur la souche *E. aridi* LMR001 nous ont permis d'identifier des métabolismes et des mécanismes actifs sous l'effet du stress hydrique et du stress salin, de cibler les gènes/fonctions révélés différentiellement exprimés en réponse adaptative au stress. Ainsi, nous avons sélectionné des gènes que nous avons muté afin de réaliser des analyses fonctionnelles afin d'étudier l'implication de métabolismes ou fonctions impliquées en réponse à divers stress.

CHAPITRE I

REVUE BIBLIOGRAPHIQUE

I. La fixation biologique de l'azote

L'interaction entre les rhizobiums et leurs plantes hôtes légumineuses aboutit à la formation d'organes racinaires spécialisés appelés nodules dans lesquels des bactéries endo-symbiotiques se différencient en bactéroïdes et fixent l'azote atmosphérique au profit de la plante. En effet les rhizobiums possèdent la capacité de réduire l'azote atmosphérique en ammonium au bénéfice de la plante via l'activité d'une enzyme appelée nitrogénase. La voie de dénitrification est souvent décrite comme le principal moteur de la production de NO avec le nitrate comme substrat (Ruiz et al., 2019). Cette voie comprend le nitrate réductase périplasmique (Nap) qui réduit le nitrate en nitrite, et le nitrite réductase (Nir) qui réduit le nitrite en NO. Fait intéressant, la présence de la molécule monoxyde d'azote (NO : une molécule gazeuse diffusible) a été détectée à chaque étape de l'interaction symbiotique. Le monoxyde d'azote est important pour de nombreuses voies de signalisations dans divers processus chez les plantes. Chez les plantes, il contrôle la germination des graines, la croissance des racines, la floraison et la fermeture des stomates. Le NO est aussi impliqué de l'hypersensibilité de la plante lors de la réponse de défense contre les attaques des pathogènes et impliqué dans les différentes étapes de la symbiose (plante : rhizobiums). L'interaction plante- rhizobiums, en raison de ses intérêts qu'ils portent pour le développement durable du système agronomique et plus généralement pour l'environnement, est étudiée depuis de nombreuses années. Les études ont pour objectif de déterminer les systèmes moléculaires actifs: avant et durant la préinfection, au cours de l'infection, pendant le processus de différenciation en bactéroïde au sein de structures végétales spécialisées appelées nodosités, et finalement durant la fixation de l'azote (Ruiz et al., 2019).

1. Diversité des organismes fixateurs d'azote : les rhizobiums

Les bactéries du sol, appelées communément rhizobiums, ont une capacité de former des symbioses fixatrices d'azote avec les légumineuses. Elles sont dispersées généralement dans sept familles d' alphaprotéobactérie et une famille de bêtaprotéobactérie (Fagorzi et al., 2020). Le groupe des rhizobiums est constitué par des bactéries appartenant aux genres : *Rhizobium*, *Bradyrhizobium*, *Azorhizobium*, *Mesorhizobium*, *Allorhizobium*, *Methylobacterium* et *Ensifer* (*Sinorhizobium*) dans la classe *Alphaproteobacteria* ainsi que dans deux branches de la Classe des *Betaroteobacteria* (*Burkholderia*, *Cupriavidus*) et quelques souches non classées au sein des *gamma-protéobactéries*. À l'heure actuelle, de plus en plus des espèces sont continuellement identifiées, 15 genres et plus de 120 espèces de rhizobium, avec un taux de découverte de 10 espèces par an (Abdullahi et al., 2020).

1.1 Symbiose fixatrice d'azote : *Rhizobium-légumineuse*

Le processus de l'interaction symbiotique plante-rhizobium commence lors de la pré-infection, par une phase de reconnaissance entre les deux partenaires (Patriarca *et al.*, 2004). La préinfection est une phase où les rhizobia sont attirées vers les poils absorbants de la racine, par chimiotactisme. Les rhizobia se dirigent ainsi vers de nombreux composés exsudés par la plante dans leur rhizosphère. Ces composés sont des substances de type flavonoïdes et iso flavonoïdes. Une production plus importante de ces composés est observée en condition de carence azotée. Les flavonoïdes présents dans les exsudats racinaires induisent ainsi l'expression des gènes *nod* bactériens qui gouvernent la production des facteurs Nod, des lipochitooligosaccharides (Figure 1).

Les étapes de l'interaction symbiotique nécessitent un dialogue moléculaire spécifique entre les deux symbiotes, qui contrôle et gère les processus de développement du nodule (Calatrava-Morales *et al.*, 2018). Durant la première étape de la pré-infection, les légumineuses sécrètent une variété de composés dont des flavonoïdes qui attirent les bactéries vers leurs racines et activent la production d'une molécule signal lipochitooligosaccharide connue sous le nom de facteur Nod via le régulateur transcriptionnel NodD1 qui active les gènes *nodABC* impliqués dans la biosynthèse de ces facteurs Nod. Ensuite, la présence du facteur Nod provoque l'activation d'une cascade des voies de signalisation qui provoquent un enroulement des poils racinaires (appelés « crosse de Berger ») et la division des cellules corticales de la racine. La déformation des poils absorbants piège les bactéries qui induisent l'invasion des racines à travers une structure en forme de tube appelée cordon d'infection. Ce processus nécessite la production d'exopolysaccharide (EPS) : succinoglycane (EPS I) ou galactoglucane (EPS II). Les cordons d'infection se développent vers les cellules corticales en division et se ramifient, pour guider les bactéries vers les cellules végétales en division dans le primordium nodulaire en division. Ce processus endocytotique est maintenu génère des symbiosomes qui sont constitués de membrane plasmique d'origine végétale qui renferment les rhizobiums. Le développement des cordons d'infection dans les cellules du cortex externe puis interne se fait à l'aide des cellules voisines, constituées de ponts cytoplasmiques alignés de façon radiale. Ces structures guident la croissance des cordons d'infection en direction du primordium nodulaire en formation (Rasanen, 2002). Après ces étapes d'infection de la racine, les bactéries colonisent donc ces organes appelés nodules dans lesquels le micro-symbionte stoppe la division et se différencie en bactéroïdes pour se spécialiser dans la réduction du diazote atmosphérique en composés mono-azotés (Ammonium) assimilable pour la plante (Calatrava-Morales *et al.*, 2018).

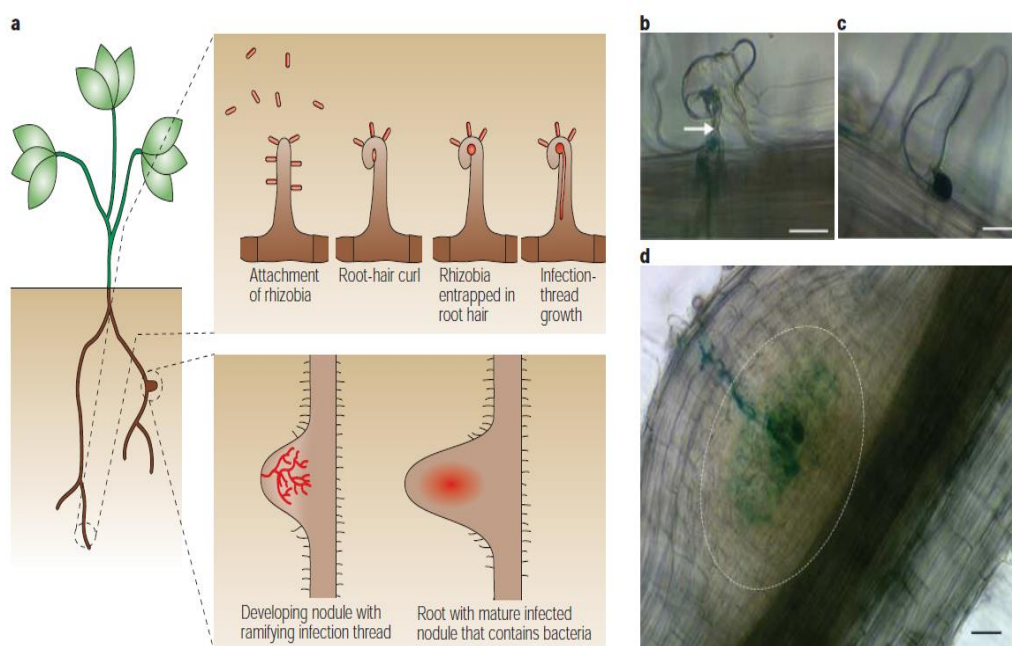


Figure 1: Schéma représentant les étapes de l'infection des légumineuses par *Rhizobium leguminosarum*. a | présente les premiers stades de l'infection des poils absorbants par les rhizobiums et la croissance des nodules sur les racines de la plante. b, c et d | sont des images qui montrent l'expression du gène *lacZ* (exprimé de manière constitutive) dans les cellules des bactéries *Rhizobium leguminosarum biovar viciae*, durant l'infection du *Pisum sativum* ou *Vicia hirsuta*. La racine a été colorée histochimiquement par l'activité de l'enzyme β -galactosidase, révélant les bactéries infectantes en bleu. b | photo microscopique qui montre un fil d'infection commençant sur la racine de *P. sativum* qui se replie et piège les bactéries pour former un foyer d'infection. c | photo microscopique montrant un fil d'infection non courbé. d | Les fils d'infection se développent à travers les cellules racinaires jusqu'au primordium du nodule en croissance (entouré sur l'image par un cercle). L'image montre un jeune nodule sur *V. hirsuta* qui a été infecté comme expliqué ci-dessus. Les barres des parties a et b représentent 20 μ m et dans la partie c représente 100 μ m (Oldroyd et Downie 2004).

1.2 Symbiose fixatrice d'azote : *Ensifer-légumineuse*

Ensifer meliloti est considérée comme une bactérie modèle du genre *Ensifer*. C'est l'espèce la plus étudiée parmi les autres espèces qui appartiennent au genre des *Ensifers*. Elle peut établir une symbiose fixatrice d'azote avec des légumineuses appartenant aux genres *Medicago*, *Melilotus* et *Trigonella*. Elle a été largement étudiée en raison de son importance agricole et écologique. L'interaction symbiotique *E. meliloti* - *Medicago truncatula*, le système modèle, se traduit par la

formation de nodules de type indéterminé sur la racine des plantes (figure 2) (Rodriguez-Llorente et al., 2009). L'étape de la pré-infection qui précède la symbiose, commence par le groupement des bactéries autour de la rhizosphère. À la suite de la colonisation bactérienne des racines, une augmentation de la densité de population bactérienne déclenche le quorum sensing (QS). Le QS est un mécanisme de signalisation de cellule à cellule qui permet aux bactéries de modifier collectivement leur comportement. Les fonctions régulées par le QS chez les rhizobium comprennent le transfert du plasmide, la sécrétion des polysaccharides de surface, la répression de la motilité, l'adaptation à la phase stationnaire de croissance et aux conditions biotique et abiotiques stressantes auxquelles sont confrontés les rhizobiums lors de l'interaction symbiotiques (Calatrava-Morales et al., 2018).

L'infection racinaire par les *Ensifiers* induit la dédifférenciation et la division des cellules du cortex qui vont former le nodule de type indéterminé. La division des cellules du cortex interne de la racine forme spécialement des nodules de type indéterminé alors que les nodules de type déterminé sont formés à partir du cortex externe. La persistance du méristème chez les espèces à nodules de type déterminé est très éphémère, et on remarque que la croissance en longueur du nodule est limitée et que la croissance en épaisseur est marquée par une hypertrophie des cellules corticales et des divisions de cellules contenant déjà des rhizobia. Ce processus de formation se traduit par une forme sphérique et un état de différenciation identique pour toutes les cellules. En ce qui concerne le nodule de type indéterminé, les cellules nodulaires ne sont pas toutes au même stade développemental et plusieurs zones se distinguent (Figure 2). La zone méristématique reste persistante ce qui se traduit par une forme allongée des structures nodulaires. Dans les deux cas, les cellules du cortex se divisent de manière anticline puis péricline (Timmers *et al.*, 2000).

A l'entrée des bactéries par endocytose à l'intérieur de la cellule végétale (cytoplasme), chaque bactérie se retrouve dans une vésicule composée de membrane d'origine végétale appelée symbiosome (Figure 2). A l'intérieur de cette enveloppe les bactéries subissent une série de modifications morphologiques et cellulaires et se différencient en bactéroïdes, qui permet de fixer l'azote. Dans le cas de la formation de nodules indéterminés comme c'est le cas chez la symbiose *Esnifer-Medicago*, cette différenciation correspond à des modifications de la perméabilité de la membrane bactérienne et du génome qui subit une série d'endo-duplication, contribuant ainsi à une augmentation importante des tailles du génome (nombre de copies) et des cellules. Dans ce cas, les bactéries ont également perdu leur capacité de reproduction et sont donc des bactéries définitivement différenciées. A la maturation, les bactéroïdes mûrs continuent à accumuler de grandes quantités de PHB ou polyhydroxybutyrate (Juergen Prell *et al.*, 2006).

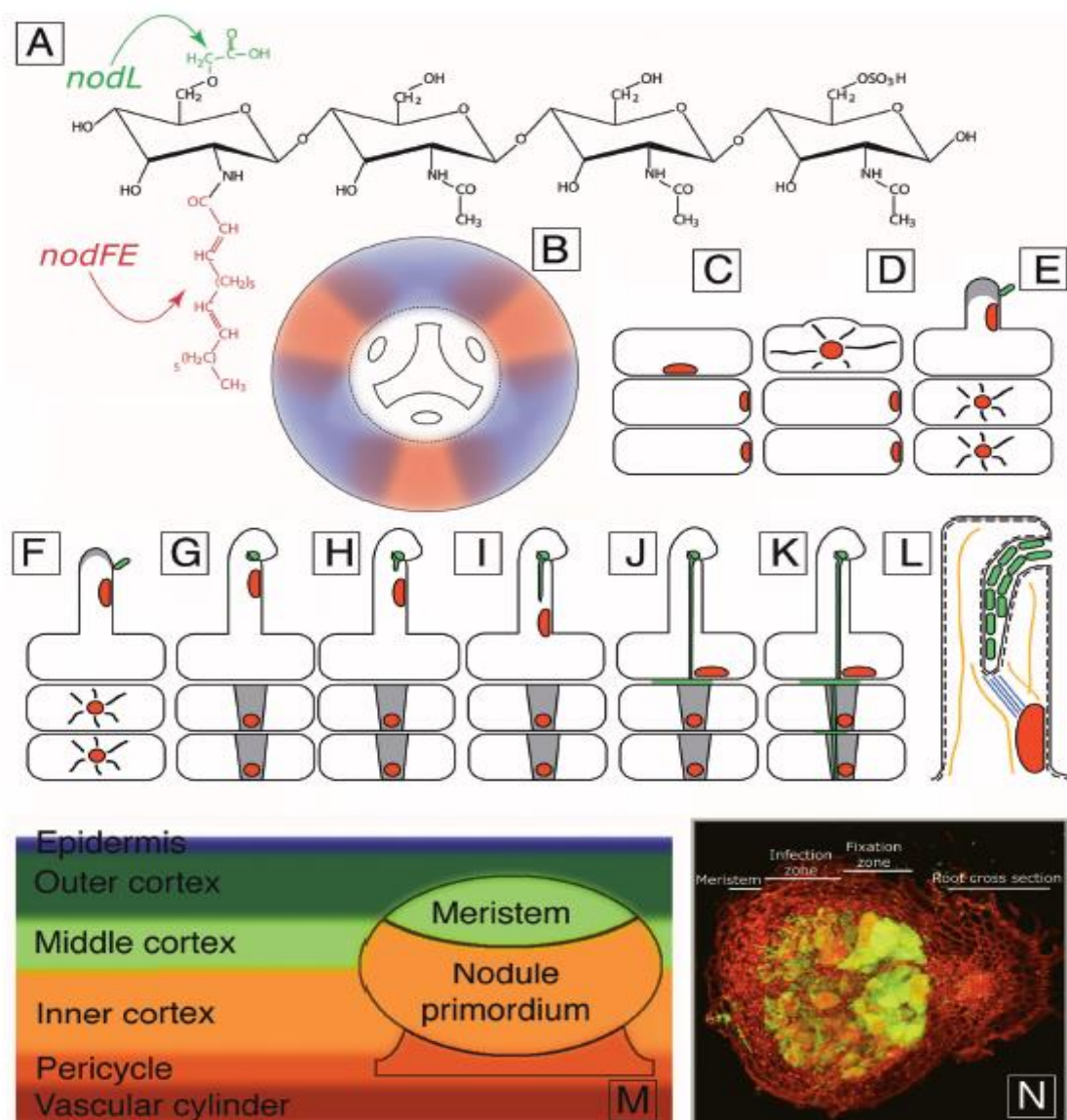


Figure 2: Présentation du processus de la nodulation chez les plantes qui forment des nodules indéterminés (type de nodule formé par les bactéries du genre *Ensifer*). (A) Exemple de la composition moléculaire du facteur Nod synthétisé par *E. meliloti*. La flèche supérieure indique le groupe acétyle ajouté par NodL, et la flèche inférieure indique la molécule lipidique dont la longueur et le degré de saturation, sont modulés par NodF et NodE. (B) Coupe transversale schématique montrant la racine, qui présente les gradients d'activation d'un facteur aux pôles du protoxylème (coloration bleu) et un inhibiteur aux pôles du protophloème (coloration rouge). Ces gradients déterminent les cellules racinaires en activité en réponse à l'infection racinaire par les rhizobiums. Les nodules se forment généralement à côté des pôles du protoxylème, qui se trouvent aux extrémités de la structure (en forme Y) localisées au centre, au-dessus des pôles du

protophloème (représentés comme des ovales). (C) schéma d'une cellule épidermique sur deux cellules corticales externes sous-jacentes ; la cellule épidermique a un noyau positionné juste en face de l'endroit où un nouveau poil racinaire se formera. (D) schéma représentant l'initiation des poils absorbants dans la cellule épidermique. (E) En réponse au facteur Nod, une liaison d'une cellule rhizobienne à un poil radiculaire de type I se produit et une activation de l'activité des cellules corticales sous-jacentes commence. (F) par la suite, la croissance des poils absorbants continue, stade II. (G) Les poils absorbants se courbent en stade II sous l'influence du facteur Nod et des microcolonies rhizobiennes s'installent dans la boucle formée au niveau de la courbure du poil racinaire. Les cellules corticales sous-jacentes se sont polarisées et des ponts cytoplasmiques (PIT) se sont formés et sont représentés en gris. (H) Infection et prolongation du fil d'infection. (I) Le développement du fil d'infection le long des poils absorbants, et le noyau descend le long des poils absorbants devant le fil. (J) Fusion du fil d'infection avec la paroi cellulaire épidermique et croissance des rhizobiums dans l'espace intracellulaire entre la cellule épidermique et la cellule corticale sous-jacente. (K) La croissance du fil d'infection à travers l'espace inter-cytoplasmique dans les cellules corticales externes. (L) présentation agrandie des poils absorbants ; la boucle a été déroulée pour montrer que topologiquement la localisation des bactéries dans le fil d'infection et à l'extérieur des poils. La paroi cellulaire végétale et la membrane cellulaire végétale sont représentées par des lignes noires et en pointillés, respectivement. Les microtubules sont situés entre le noyau et le fil infectieux pointe (bleu). Les câbles d'actine sont représentés par des brins orange. (M) Schéma montrant les tissus racinaires et un jeune nodule non mature. La dérivation des tissus nodulaires à partir des tissus radiculaires est indiquée. (N) Une section longitudinale d'un nodule de luzerne de 10 jours. Le nodule a été infecté par *E. meliloti* exprimant la GFP, et le réseau de fils d'infection peut être vu derrière la région du méristème du nodule ; La zone I correspondant au méristème, la zone II cortex nodulaire : la zone d'infection, la zone II et III du cortex correspond à l'interzone, et la zone IV correspond à la zone de sénescence. Les sites d'infection initiale des bactéries infectieuses dans le nodule peuvent être vus sur la périphérie du nodule à gauche. L'iodure de propidium (rouge) a été utilisé pour contre-colorer le tissu végétal (Gage , 2004).

2. Caractérisation génomique, phylogénique du genre *Ensifer*

2.1 Description du genre *Ensifer*

Le genre *Ensifer* (anciennement appelé *Sinorhizobium*) est récemment reconnu comme formant un seul clade phylogénétique avec les *Sinorhizobiums*, conformément à la règle 38 du Code bactériologique. Il comprend actuellement 17 espèces qui sont largement réparties dans les régions arides (Rejili et al., 2020). La bactérie est capable de générer une interaction par la présence de gènes symbiotiques, codant pour la synthèse de signaux spécifiques (facteurs nod). Ces derniers agissent en déclenchant une série de réponses chez la plante conduisant à la formation des nodules au sein desquels la bactérie fixe le diazote atmosphérique en une forme assimilable par la plante (Rejili et al., 2020). Les sujets d'études de ces réseaux fonctionnels ont fait l'intérêt des microbiologistes durant ces dernières années, et ont donné des progrès significatifs sur les connaissances des échanges très complexe entre la plante et les rhizobia durant l'infection la formation des nodules et le fonctionnement symbiotique.

Le genre *Ensifer* regroupe de nombreuses bactéries dont la majorité peut fixer l'azote en symbiose avec diverses légumineuses. Les souches symbiotiques sont presque toutes situées dans un même clade monophylétique et portent les gènes de nodulation de la fixation d'azote (*nodABC* et *nifHDK*). Les bactéries du genre *Ensifer* peuvent servir à titre d'exemple pour étudier l'évolution de l'association mutualiste entre une cellule procaryotes et une cellule eucaryotes, pour étudier aussi l'évolution de la phylogénie génomique entre des bactéries symbiotiques comme *Ensifer meliloti* et *Ensifer fredii* et des bactéries non symbiontes comme *Ensifer morelense* et *Ensifer adhaerens* (Fagorzi et al., 2020). Une approche taxonomique plus récente basée sur la phylogénie du génome a suggéré que le genre *Ensifer* devrait être séparé du genre *Sinorhizobium*. En effet, la proposition originale de *Sinorhizobium* (Chen et al., 1988b), a été remise en question, au motif qu'il n'y avait pas grand-preuve pour discriminer ce genre de *Sinorhizobium*. Actuellement, les chercheurs sont divisés sur la validité taxonomique de *Sinorhizobium* et d'*Ensifer* qui vient d'être remis en question (Kuzmanovic et al., 2022).

2.2 Génome de *Ensifer*

Le génome de la souche modèle appartenant au genre *Ensifer Sinorhizobium meliloti* (Rm1021), comprend trois réplicons, un chromosome, et deux mégaplasmides *pSymA* et *pSymB* mais l'analyse comparative génomique de plusieurs souches appartenant à cette espèce montre qu'il peut également contenir des plasmides et que l'espèce possède un large set de gènes accessoires

incluant des régulons potentiellement impliqués dans l'interaction suggérant une certaine variabilité de performances symbiotiques entre les souches (Galardini et al., 2011). Il comprend divers régulateurs transcriptionnels impliqués dans le développement et le fonctionnement symbiotique. Comme par exemples des régulateurs de la perception des exsudats racinaires et la nodulation précoce (NodD, NodR), l'adaptation microaéroophile (FixK), la synthèse de la nitrogénase (FixJ, NifA), l'absorption du fer (Fur), la biosynthèse d'EPS (ChvI) et la compétition avec l'invasion des plantes (ResR, NesR). Le système de la régulation des gènes et des métabolismes dans les situations de stress chez l'espèce *Ensifer* est très complexe. Par exemple, les gènes régulés par la protéine FixK ne s'expriment pas tous dans les conditions de microaérophilie et sont aussi contrôlés par d'autres régulateurs. Les gènes de régulation sont divisés en plusieurs groupes fonctionnels, dont le transport d'électrons, la symbiose, le métabolisme de l'azote. L'analyse des régulons suggère qu'en général NodD1 régule *nodL*, *nodF*, *nodA*, *nodM* et *syrM*, la FixK régule *fixNOQP 1*, *fix- NOQP 2*, *fixGIS*, *arcABC*, *napEFDABC*, *norBCE*, *cycB2* et *degP4*, NifA régule les opérons *nifHDKEX*, *fixABCX* et *nifB*, NodR régule *nodA*, *nodD1*, *nodD2* et *nodM*, et Fur régule le gène *sitA*, NesR régule l'opéron *metHK* et le gène *ahcY* et ChvI le gène *ropB*. Une co-régulation par différents facteurs a été observée sur plusieurs gènes. L'étude comparative génomique a permis d'identifier chez la *E. meliloti*, plusieurs gènes régulés par plus d'un régulateur incluant *nodD1*, *nodD2*, *nodA*, *nodM*, la copie de *nodM* (AK83) et deux autres gènes non caractérisés. Une co-régulation a également été trouvée (NodR sur NodD1 et FixJ sur FixK) indiquant que certains des régulateurs symbiotiques sont liés entre eux dans un réseau pour coordonner l'expression de multiples gènes nécessaires à l'infection et la fixation de l'azote. FixK contrôle le plus grand nombre de gènes impliqués dans le processus de la symbiose, ainsi que ceux impliqués dans le métabolisme de l'azote et le transport d'électrons nécessaires dans l'environnement micro-aéroophile du bactéroïde (Galardini et al., 2011).

Parmi les gènes accessoires, certains sont impliqués dans le processus de la symbiose. Parmi 290 groupes d'orthologues, 61 d'entre eux se trouvent au niveau du génome accessoire. Ces 61 groupes d'orthologues accessoires sont divisés en 36 entités, avec 32 gènes groupés en 7 opérons. Au final, trois classes principales ont été identifiées : Le premier comprenant des gènes impliqués dans la croissance microaéroophile des bactéroïdes et contrôlés par le régulateur transcriptionnel FixK. Le second comprenant d'autres gènes qui sont soit directement ou indirectement liés à la symbiose : par exemple impliqué dans la définition du spectre d'hôtes, la compétitivité, l'effectivité et l'efficacité de fixation biologique de l'azote, etc. Le dernier comprenant des protéines avec des informations limitées sur leur fonction, bien que probablement impliqués également dans le

processus symbiotique. Des gènes accessoires impliqués dans la micro-aérophilie sont retrouvés dans deux souches symbiotiques (Rm1021 et BL225C) mais absents dans la souche AK83 (fix-) et pourraient être important pour le fonctionnement symbiotique. Ils sont principalement regroupés dans trois parties du réplicon *pSymA* dont deux connues pour être induites en conditions de microaérophilie. Ils comprennent un gène de type *fixK* qui code pour un régulateur transcriptionnel, une troisième copie de l'opéron *fixNOQP* qui code pour une chaîne de transport d'électrons à haute affinité pour l'oxygène et plusieurs opérons liés au métabolisme de l'azote, avec les gènes *nor* (nitrique réduction des oxydes), *nir* (réduction du nitrite), *nos* (nitreux réduction des oxydes), *nnr* (régulation des opérons *nir* et *nor*), *nrt* (transport des nitrates) et deux gènes également liés à l'environnement microaérophile *cycB2* et *hemN*. Fait intéressant, les chercheurs ont également identifié différents nombres de copies des gènes *fixK* situés dans les trois génomes. Avec un groupe orthologue (N.280) présent dans le core-génome, FixK2 (groupe orthologue N.7042) et une troisième copie de type FixK (groupe orthologue N.5488). D'autres gènes sont présents uniquement dans le génome accessoire comme le gène qui code pour une déshydrogénase à chaîne courte (SDR) dont la mutation conduit à la formation de nodules blancs et allongés ou finalement, le groupe de gènes qui codent pour la biosynthèse de la rhizobactine qui n'affectent pas directement la fixation de l'azote, mais affecte à long terme la croissance des plantes (Galardini et al., 2011). Quatre gènes accessoires sont exclusivement présents chez la souche AK83, dont un gène qui code pour une nickel perméase/ hydrogénase impliquée dans le recyclage de l'hydrogène lors de la fixation d'azote. Un gène Cadherin-like impliqué dans le système de sécrétion de type I, nécessaire pour l'agrégation cellulaire et l'attachement à la racine. Un gène qui code pour la phosphocholine cytidyllyltransférase impliquée dans le métabolisme de la phosphatidylcholine, localisée dans la membrane bactérienne et important pour l'adhésion aux cellules eucaryotes. Et finalement le gène *nodM*, qui joue un rôle essentiel dans les premières étapes de l'invasion rhizobienne de la plante hôte font également partie de ces gènes accessoires spécifiques de la souche AK83 (Galardini et al., 2011).

3. Caractérisation génomique, phylogénique et origine de *Ensifer aridi* sp. Nov

3.1 Description de la nouvelle espèce *E.aridi*

La découverte de nouvelles souches d'*Ensifer* groupant avec les souches Marocaines et Indiennes mais cette fois isolées des sols mexicains dans la région de Baja California à partir de nodules de *Phaseolus filiformis* poussant sous un climat semi-aride a confirmé l'adaptation de cette nouvelle espèce de *Ensifer* à l'environnement (semi-) désertique et montre qu'elle est présente au moins sur

trois continents (Rocha et al., 2020). Afin de tester par des analyses de génomique comparatives si toutes ces souches appartiennent bien à une même et nouvelle espèce de *Ensifer* et rechercher les bases génétiques pouvant expliquer sa répartition géographique, les génomes de deux souches Mexicaines (LEM451, LEM457), deux souches Marocaines (LMR001, LMR013) et une souche Indienne (TP6) ont été séquencés et comparés en intégrant le génome d'une première souche indienne groupant avec nos souches et qui était déjà séquencée (TW10) (Le Quéré et al., 2017). Les souches indiennes furent isolées des nodules de *Tephrosia purpurea* (TP6) et *T. wallichii* (TW10), les deux souches mexicaines (LEM451 et LEM457) isolées des nodules de haricot sauvage (*Phaseolus filiformis*) les 2 souches Marocaines (LMR001 et LMR013) isolées respectivement des nodules *Vachellia gummifera* et *V. raddiana*.

Les génomes des 5 souches LMR001, LMR013, LEM451, LEM457 et TP6 ont été obtenus en utilisant la technologie Illumina (Hiseq2000) consistant pour chaque souche en 2 séries séquentielles appariées de 100 cycles de fragments génomiques. Le nombre des séquences obtenues après l'analyse varie de 34 à 71 millions selon les souches. Les séquences filtrées ont été assemblées avec le logiciel CLC Genomics. Les tailles des génomes des souches étudiées varient de 6,4 à 6,8 millions de paires de bases (Mbp). Les résultats des projets Whole Genome Shotgun (WGS) pour les souches LMR001, LMR013, LEM451, LEM457 et TP6 ont été déposés dans les bases de données DDBJ/ENA/GenBank sous les accessions respectives LUAV, LUFU, LUFV, LUFW and LUFX. Les versions correspondantes décrites dans ce document sont les versions LUAV01000000, LUFU01000000-LUFX01000000 (Le Quéré et al., 2017). La technique d'analyse de génomique comparative a permis de délimiter les souches isolées et de prédire des gènes et leurs fonctions associées. Les 6 génomes ont été comparés avec d'autres génomes déjà disponibles dans les bases de données génomiques des bactéries appartenant aux espèces *E. meliloti*, *E. medicae* et *E. fredii*. (Le Quéré et al., 2017).. Les analyses phénotypiques ont montré qu'elles étaient capables d'utiliser le glycérol, le l- arabinose, le d- fructose, le l- rhamnose, l'inositol, le d- mannitol, le d- maltose, le d- saccharose, le d- tréhalose, le d- raffinose, le d- turanose ou le d- arabitol comme sources de carbone (Sakrouhi et al., 2016 ; Le Quéré et al., 2017). Les souches appartenant à cette nouvelle espèce présentent des génomes accessoires variables, ce qui peut expliquer les gammes variables des plantes hôtes ainsi que des sources de C qu'elles peuvent utiliser (Le Quéré et al., 2017). La teneur en G + C de la souche LMR001 est de 61,7 %. Les comparaisons avec les espèces de *Ensifer* définies et séquencées des valeurs moyennes d'identité nucléotidique (ANI) et d'hybridation ADN-ADN digitale (dDDH) ont été effectuées et démontrent clairement que les souches d'*Ensifer* isolées des déserts sur trois continents appartiennent bien à une nouvelle espèce du genre *Ensifer*. Les valeurs les plus élevées avec les

espèces définies ont été trouvées avec *E. saheli* LMG 783-T, avec 85,4 % et 29,8 % de valeurs ANI et dDDH, ce qui

Ces rhizobiums isolés des déserts chauds d'Asie, d'Afrique et d'Amérique, de sols à pH alcalin, sableux soumis à des températures élevées et développant des interactions symbiotiques effectives avec diverses plantes hôtes, appartiennent à une nouvelle espèce qui a été proposée d'être nommée *Ensifer aridi* en raison de sa capacité d'adaptation aux conditions arides avec LMR001 comme souche type (Rocha et al., 2020).

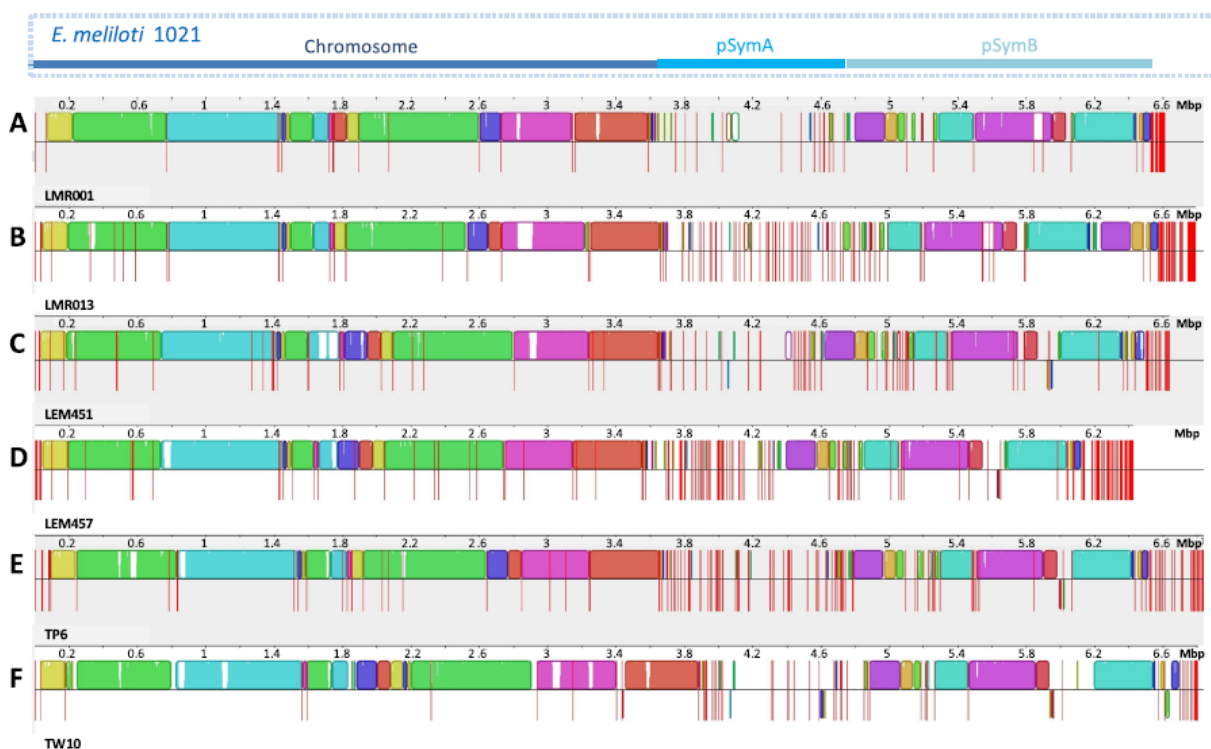


Figure 3: L'alignement des contigs obtenus pour les 6 génomes d'*Ensifer*.sp du désert. Le génome fini d'*Ensifer meliloti* 1021 constitué de réplicons concaténés dans l'ordre : chromosome, pSymA et pSymB (indiqué en haut de la figure), utilisé comme référence pour aligner les 79 Scaffolds de la souche *Ensifer aridi* LMR001 (a) à l'aide de l'aligneur progressif MAUVE. Le génome de la LMR001 pseudo-assemblé et réordonné a été ensuite utilisé comme référence pour réorganiser les scaffolds de la LMR013 (b), la LEM451 (c), la LEM457 (d), la TP6 (e) et la TW10 (f). Les échelles au-dessus de chaque carte font référence à la taille du génome en millions de paires de bases (Mbp). Les scaffolds sont indiqués par des barres rouges verticales et les blocs localement colinéaires sont affichés en utilisant des couleurs similaires à travers les génomes. Les espaces incolores indiquent les régions génomiques accessoires.

L'alignement des régions conservées du génome de la souche *E. aridi* LMR001 avec les 5 autres génomes a montré la présence de grandes régions conservées. Ces régions correspondent globalement avec le chromosome et le chromide *pSymB* de la souche *E. meliloti* 1021 et montre une différence en particulier au niveau de la région centrale correspondante au plasmide symbiotique de *E. meliloti*. (*pSymA*) qui contient les gènes *nod* et *nif* impliqués dans le développement et le fonctionnement de la symbiose. L'identification de gènes *rep*, *tra* et *trb* impliqués dans la répllication et le transfert des plasmides suggèrent que la nouvelle espèce de *Ensifer* contient également de larges plasmides. Les résultats de l'analyse sur gel d'Eckhardt ont clairement montré la présence de réplicons extra-chromosomiques (méga plasmides) (allant approximativement de 0,1 à 1,4 Mbp) dans tous les isolats étudiés. Les souches Marocaines ont présenté 2 réplicons en plus du chromosome comme pour la souche modèle d'*Ensifer meliloti*. Un grand réplicon d'environ 1,3 Mpb de taille et un plus petit d'environ 200 Kpb pour la souche (LMR001) et de 400 Kpb pour la souche (LMR013). Le profil de l'une des souches mexicaines LEM451 indiquait également la présence de 2 réplicons supplémentaires (d'environ 800 Kpb et 250 Kpb). Tandis que le profil obtenu à partir de la seconde souche (LEM457) montrait la présence d'un seul plasmide d'environ 100 Kpb. Enfin, comme pour les souches marocaines, le profil de la souche indienne TP6 indiquait la présence de 2 réplicons supplémentaires d'environ 1,3 et 0,5 Mpb (Le Quéré et al., 2017). L'analyse de la taille du génome core et l'analyse pan-génomique ont indiqué une proportion plus faible de protéines conservées par toutes les souches de *E. aridi*. Ce niveau de conservation était plus élevé lorsque des souches de même origine géographique étaient comparées. Une présence d'une plus grande diversité de protéines était observée dans les génomes accessoires de cette espèce avec une grande conservation du core génome. Ces données soutiennent l'évidence de l'acquisition locale des grands plasmides et explique la diversité des hôtes avec laquelle ils s'associent (Le Quéré et al., 2017).

La taille et le contenu génique des génomes des souches de la nouvelle espèce d'*Ensifer aridi* ont été comparées à celles de trois espèces d'*Ensifer*. (*E. meliloti*, *E. medicae* et *E. fredii*) pour lesquelles au moins 6 génomes étaient accessibles au moment de l'étude après annotation automatique en utilisant la pipeline « Rapid Annotation using Subsystem Technology » (RAST). À l'aide de la méthode développée par Ozer et ses collègues, les tailles des génomes conservés au niveau de chaque espèce (core-génome) et accessoires ainsi que le % GC de chaque fraction a été estimé et les gènes attribués (Predicted Encoding Genes, PEGs) à chaque fraction (prédits par l'annotation RAST) ont été récupéré (tableau 2 et suppléments sur l'article). La taille du core-génome varie d'une espèce à l'autre. Le plus petit est celui d'*E. aridi* avec 5,12 Mbp alors que *E. fredii* et *E. meliloti* ont des tailles respectives de 5,55 et 5,76 Mbp et le plus grand core-génome a

été trouvé pour *E. medicae* avec 5,95 Mbp. Les core-génomes représentaient 75 à 87% des génomes et le GC % des génomes accessoires étaient plus faibles (58,8-60,1%) que celles correspondant aux core-génomes (61,4 - 62,8%) pour l'ensemble des 4 espèces étudiées (Le Quéré et al., 2017). Les séquences fusionnées de 103 gènes conservés chez les souches étudiées ont été analysées phylogénétiquement par Neighbor Joining (NJ). Encore une fois et comme attendu par l'analyse ANI, toutes les 6 souches sont regroupées et ont montré très peu de divergence de séquence comme indiqué par la faible longueur de branche dans le clade comparé à celles trouvées entre des souches appartenant à des espèces différentes, avec des valeurs de bootstrap de 100 %. Comme déjà identifié par l'analyse MLSA par Sakrouhi et al., (2016), le plus proche parent de l'espèce actuellement décrite est bien *E. saheli*, une espèce récupérée d'un nodule de *Sesbania* dans un environnement aride et chaud (Le Quéré et al., 2017). Les distances évolutives intra-espèces obtenues, en comparant les 103 gènes de ménage situés en majorité sur les chromosomes, ont indiqué une proximité étroite entre les deux souches marocaines (LMR001 et LMR013) et les deux souches mexicaines (LEM451 et LEM457). Ces 4 souches qui se sont regroupées ont montré entre 108 et 156 de sites variables sur les 118 746 nucléotides alignés et étaient clairement séparées des isolats indiens (TP6 et TW10) par 377 à 426 sites variables sur 118 746 nucléotides alignés alors que les souches indiennes ne possèdent que 7 sites variables sur les 118 746 nucléotides alignés correspondant à ces gènes conservés. (Le Quéré et al., 2017). L'arbre généré en utilisant des gènes situés dans la région montrant des similitudes avec *pSymB* de *E. meliloti* 1021 a indiqué une plus grande divergence entre la souche mexicaine LEM451 et les souches LEM457, LMR001 et LMR013 ; ainsi qu'une nouvelle fois, une forte conservation des séquences entre les deux souches indiennes TP6 et TW10. Les phylogénies des ensembles de gènes concaténés situés dans des régions présentant des homologies avec le chromosome d'*E. meliloti* n'ont pas regroupé les souches marocaine et mexicaine selon leur origine géographique indiquant une relation évolutive plus étroite (Le Quéré et al., 2017).

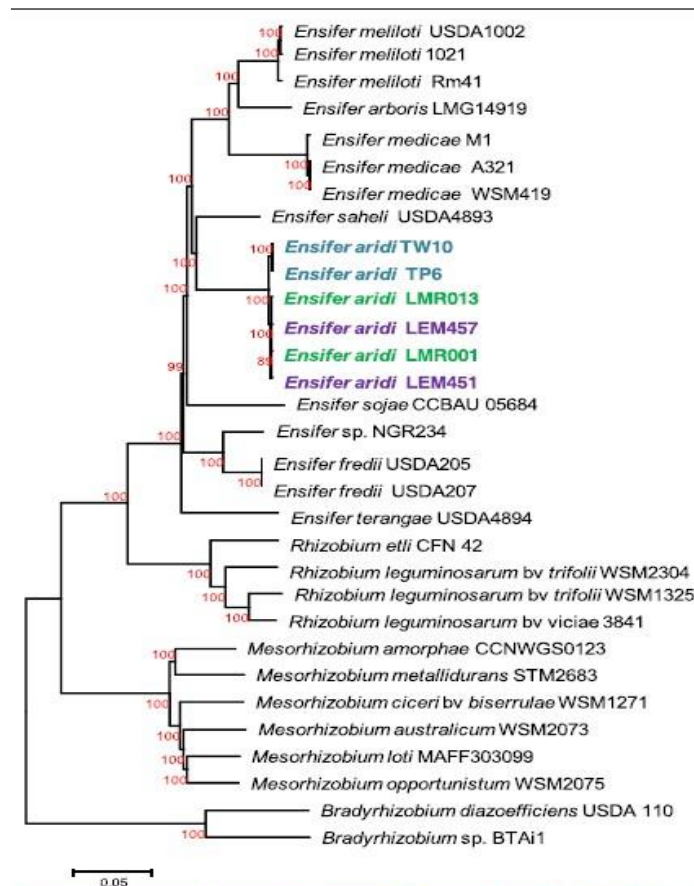


Figure 4 : Arbre construit par Neighbor Joining montrant des relations phylogénétiques entre 31 souches rhizobiennes d'*Ensifers.sp* . L'arbre a été calculé à l'aide de 103 séquences de gènes conservées qui ont été concaténées et alignées. L'arbre est dessiné à l'échelle, avec des longueurs de branches dans les mêmes unités que celles des distances évolutives utilisées. Les distances évolutives ont été calculées à l'aide de la méthode K2P et sont exprimées en unités du nombre de substitutions de bases par site. Le pourcentage d'arbres répliqués dans lesquels les taxons associés se sont regroupés dans le test de bootstrap (1000 répliques) est indiqué à côté des branches (rouge). Les souches marocaines, mexicaines et indiennes sont respectivement surlignées en vert, violet et bleu.

3.2 Effet du Stress sur la souche *Ensifer aridi* LMR001

Les rhizobiums fixateurs de N₂ ont des cycles de vie très complexes (Poole et al., 2018). Dans la rhizosphère, les rhizobiums doivent rivaliser avec d'autres membres de la communauté microbienne et résister aux conditions de stress abiotiques (Li et al., 2016). Durant les différentes phases du développement symbiotique, les rhizobiums doivent tolérer les conditions d'acidité et d'oxydation, contrôler les échanges de signaux métaboliques et résister à la réponse immunitaire

de la plante (Dicenzo et al., 2019). Même après le développement complet du nodule, les souches sont toujours sous stress. Dans quelques cas, des rhizobiums trichent et cohabitent avec d'autres souches plus compétentes, soit dans les mêmes nodules ou dans des nodules différents de la même plante. Cependant, les plantes limitent le développement des nodules non fixateurs en les sanctionnant directement pour sélectionner les meilleures, afin de préserver son potentiel énergétique tout en optimisant l'apport d'azote fixé. Les rhizobiums interagissent en permanence avec une diversité microbienne très large dans le sol, par antagonisme ou mutualisme. En cas de rivalité, les bactéries piègent les nutriments ou produisent des bactériocines (activité antibactérienne) contre des taxons non-apparentés pour augmenter la compétitivité et se propager sur plus de surfaces. Les rhizobiums rencontrent également de nombreux bactériophages dans la nature, qui peuvent réduire la densité d'une population ou/et permettre aux souches résistantes de persister (Dicenzo et al., 2019). Dans ces environnement stressants, les bactéries doivent aussi résister à la dessiccation/ la haute température/ salinité et variation de pH (Santos et da Costa 2002). Sous l'effet du stress hydrique, l'activité métabolique de la cellule bactérienne évolue vers une conservation de l'énergie par l'utilisation des sources de carbone, telles que les acides gras/ glucose, qui sont utilisées pour produire des rendements en ATP plus élevés (Liu et al., 2005; Finn et al., 2013). Inversement, les molécules d'eau perdues provoquent une accumulation de solutés et de sels compatibles qui remplacent l'eau et stabilisent les macromolécules et les membranes de la cellule et empêchent également la formation de radicaux hydroxyles en diminuant les vitesses de diffusion intracellulaire (Halverson et Firestone 2000). Au niveau de la membrane, des modifications phospholipidiques se manifestent et aident la membrane cellulaire à mieux retenir l'eau intracellulaire. La bactérie sécrète ainsi plus des exopolysaccharides (EPS) et de l'alginate pour se protéger des stress abiotiques et améliorer la teneur en humidité dans la colonie. Ces sécrétions facilitent le maintien d'un environnement hydraté autour du biofilm et à créer un microenvironnement humide autour de la cellule (Chang et al., 2007 ; Finn et al., 2013 ; Nowak, Greń, et Mrozik 2016; Lebre, De Maayer, et Cowan 2017). L'activation de ce système adaptatif est encore peu détaillée et varie en fonction des souches. Une étude transcriptomique a été réalisée pour explorer les mécanismes liés aux réponses adaptatives de la bactérie *Ensifer aridi* LMR001 en réponse à deux conditions de stress environnementales (stress hyperosmotique) en utilisant du PEG8000 (12% W/V) simulant un stress hydrique et une concentration de 250mM de NaCl générant un stress salin (Belfquih et al., 2021). Les résultats indiquent des changements de niveau de régulation de gènes liés à l'activité du cycle de la méthionine (gérant du S-adenosyl Methioinine, le donneur universel de groupement méthyls, impliqués dans la régulation de nombreuses fonctions cellulaires), le transport et l'accumulation du tréhalose (un composé

compatible soluble osmoprotecteur), le transport et le catabolisme d'inositol (source de carbone), la transduction des signaux liés au stress membranaire et le contrôle de la mobilité.

4. La symbiose dans les environnements sous contraintes abiotiques

4.1 *Vachellia*, des espèces d'arbres d'intérêt pour le Maroc (*V tortillis* et *V gummifera*)

Les légumineuses (Fabaceae) sont la troisième plus grande famille de plantes à fleurs répertoriée dans le monde, avec environ 730 genres et plus 19 400 espèces, à importance économique et écologique mondiale. Les légumineuses ont une teneur particulièrement élevée en azote et en protéines qui représente une alternative écologique intéressante pour les pays touchés par l'aridité. De plus, elles ont une capacité à développer une interaction mutualiste naturelle avec des bactéries symbiotiques fixatrices d'azote (appelées rhizobiums). Les légumineuses pérennes à grains telles que l'arbre *vachellia*, (*Caesalpinioideae*, *Mimosoideae*, *Fabaceae* ou *Papilionaceae*), ont une importance économique et écologique mondiale. La symbiose nécessite ainsi la présence de partenaires microbiens compatibles, compétitifs et efficaces (Rocha et al., 2020). La fixation biologique du diazote (N₂) est le deuxième processus biologique le plus important sur terre, après la photosynthèse. Près de 80 % des gaz atmosphériques est de l'azote sous forme de diazote. De plus, environ 78 000 tonnes de gaz inerte sont présentes au niveau de la mer. Malgré une présence très importante en azote dans l'atmosphère et en mer, cet azote est non disponible pour les plantes (Abdullahi et al., 2020). En général, l'azote fixé par la symbiose légumineuse- rhizobium, varie entre 16 et 145 kg N ha⁻¹ an⁻¹ dans les sols fertiles et de 15 et 123 kg N ha⁻¹ an⁻¹ dans les sols pauvres, alors que l'utilisation annuelle moyenne d'engrais chimiques a été rapportée à 89,7 kg par ha an⁻¹ (Abdullahi et al., 2020). Les *vachellias* africains bénéficient d'une attention toute particulière des chercheurs et décideurs en raison de leur rôle capitale sur le plan écologique (aménagement et lutte contre la désertification grâce en partie à leur capacité à fixer l'azote dans des terres sèches et chaudes) et économique (ressources agro-forestières, production de substances naturelles) (Grouzis et al., 2003). L'espèce de plante *vachellia tortilis* subsp. *Raddiana*, est une légumineuse qui n'a pas fait jusqu'ici l'objet d'un grand intérêt. *Vachellia tortilis* subsp. *Raddiana*, comme quelques autres ligneux *Balanites aegyptiaca*, *vachellia ehrenbergiana*, *Vachellia gummifera*...), est capable de croître et de survivre dans des zones à pluviométrie moyenne annuelle très faible (inférieure à 100 mm) et capable d'assimiler efficacement l'azote dans ce type d'environnement (Grouzis et al., 2003). Les légumineuses *Vachellia.sp* ne sont pas seulement utilisées pour enrichir le sol des zones arides mais peuvent aussi fournir de nombreux types de

produits (fourrage, fumier, bois de chauffage, produits pharmaceutiques, miel, brise-vent, remise des terres dégradée, et produits comestibles) (Davies, 2005).

Au Maroc, deux espèces d'arbres originaires des zones arides et subhumides, de l'Australie occidentale et l'Afrique du Sud, appelées respectivement, *V. cyanophylla* et *V. horrida*, sont utilisées avec succès pour la stabilisation des dunes et le reboisement (Zerhari et al., 2000). Le Maroc possède aussi trois espèces endémiques de *Vachellia*, dont *Vachellia gummifera* (gommier du Maroc) une espèce non saharienne, et deux espèces *V. raddiana*, *Vachellia flava* qui sont sahariens (Blanco et al., 2016). L'espèce *Vachellia tortilis* (Forssk) Hayne subsp. *raddiana* (Savi) Brenan (communément nommé *A. raddiana*) constitue l'espèce la plus fréquente. Il s'agit d'une légumineuse pérenne à taxon autochtone, originaire du continent Africain (plus spécifiquement des zones sahéliennes et sahariennes). Elle joue un rôle important en Afrique pour la restauration des sols affectés, par le biais de la symbiose avec les Rhizobiums décrites par une croissance rapide dans les conditions de sécheresse (Ba et al., 2002).

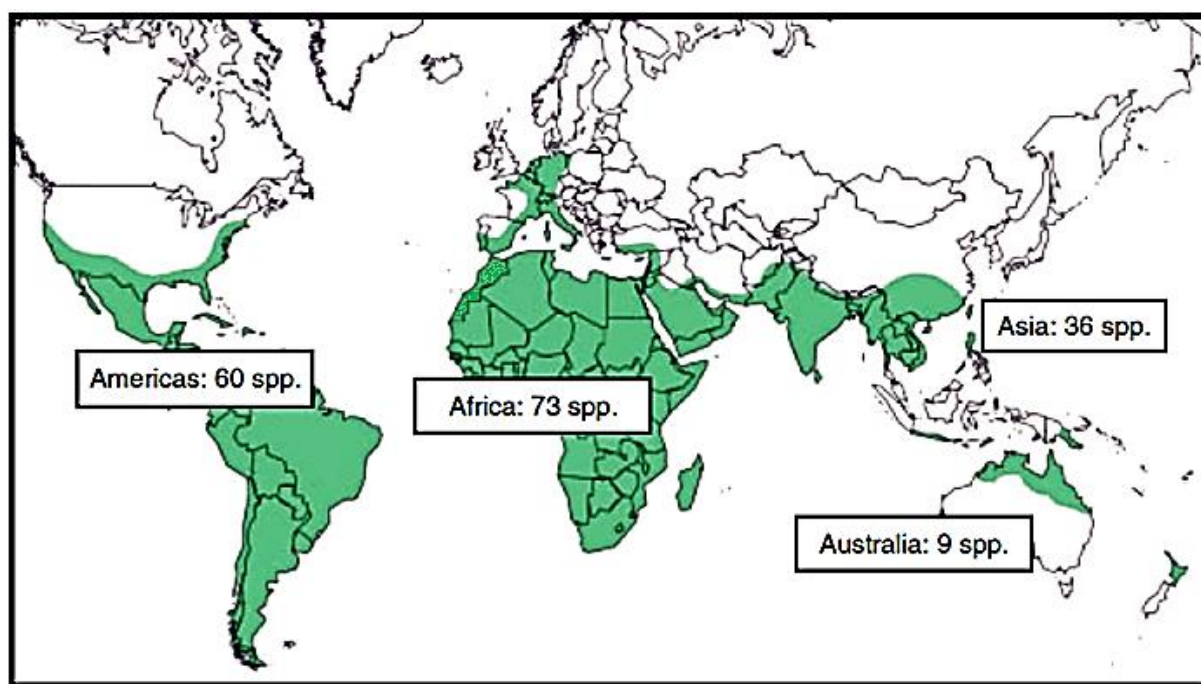


Figure 5 : Présente la distribution géographique mondiale de l'espèce *Vachellia*. Les deux espèces *V.farnesiana*, *V.karro* et *V.nilotica* ne sont pas montré sur la carte géographique de l'Australie, pour bien présenter que la distribution des espèces de *Vachellia* native du pays. Les données sont publié sur le site www.gbif.org (Comben et al., 2020).

Actuellement, les lignés les plus proches avec le *Vachellia* Australien sont originaires de la bio-région du Malé, ce qui généré l'hypothèse d'une dispersion de *Vachellia* sur une longue distance (Comben et al., 2020). Les graines de certaines espèces mimosoïdes ont une longue durée de vie

et peuvent rester viables même après 1 décennie d'immersion dans l'eau salée. Ceci démontre le potentiel de la dispersion de l'espèce à longue distance dans les courants océaniques. Des études phylogénétiques antérieures de *Vachellia* ont suggéré que le genre a divergé de son clade sœur (contenant *Sénégalia* Raf, *Mariosousa* Seigler & Ebinger et *Acaciella* Britton & Rose) il y a 40 millions d'années. Cette scission postérieure à la fragmentation du supercontinent Gondwana, suggère que la dispersion transocéanique à longue distance a joué un rôle important dans la distribution pantropicale contemporaine de *Vachellia*. La récente répartition biogéographique des plantes légumineuses *Vachellia* Wight & Arn, pantropicale, a commencé depuis l'Asie à l'Australie, en passant par l'Afrique jusqu'au continent Américain. Avec 73 espèces ayant été décrites d'Afrique, 60 des Amériques, 36 d'Asie du Sud et 9 espèces du nord tropical de l'Australie. En effet, le rayonnement ultérieur au sein du genre a commencé il y a 18 millions d'années. La divergence entre le clade afro-asiatique et le clade américain s'est produite il y a 17 à 13 millions d'années (Abdullahi A. A, Aliyu I. A., And Gabasawa A. I. 2020, Comben et al., 2020, Rocha et al., 2020).

4.2 Bactéries fixatrices d'azote symbiotiques *Ensifers* tolérance et exigence

Les sols des régions du sud Marocain, arides et sahariennes, contiennent des rhizobiums qui nodulent les plantes *Vachellia*, espèces locales et introduites. La majorité de ces souches rhizobiums appartenaient au genre des *Ensifers*, avec un taux de croissance qui varie entre très rapide (54min) à très lent (plus que 16 heures). Les taux de croissance ne corrélaient pas avec les conditions climatiques de la région, où les souches ont été isolées. Particulièrement, les souches identifiées et étudiées ont montré des capacités à tolérer des conditions de stress extrêmes, des concentrations en NaCl très élevées, des variations de pH et à tolérer des fluctuations de température (Zerhari et al., 2000, Sankhla et al., 2017).

Les bactéries du genre *Ensifer*, sont des bactéries, Gram négatives, non sporulantes, leur taille peut varier de 0.5–1.0 à 1.2–3.0 µm. Avec une morphologie pléomorphe qui apparait dans des conditions de croissance défavorables, elles contiennent généralement des granules de polyhydroxybutyrate qui sont réfractées par microscopie à contraste de phase. Aérobie, possédant un métabolisme de type respiratoire avec l'oxygène comme accepteur terminal d'électrons, elles sont mobiles grâce à des flagelles polaires ou subpolaires et peuvent posséder d'un à six flagelles péritriches, provoquant des déplacements aléatoires. Les espèces *E. fredii*, *E. saheli*, *E. terangaie* et *E. xinjiangense* sont monotriches, dotées d'un seul flagelle avec une position d'insertion polaire ou subpolaire. Les *Ensifers* tolèrent, sur un milieu riche à base d'extraits de

levure et de Mannitol (YM) une concentration de sel (NaCl) allant de 1,0% à 4,5% de (Kuykendall, Hashem, et Wang 2005). La température optimale de croissance varie de 10 à 44 °C. La plupart des souches peuvent croître à 35 °C; et certaines souches peuvent croître à 10 °C et d'autres comme *E. saheli* et *E. teranga* peuvent croître à des températures allant jusqu'à 42-44 °C (Kuykendall, Hashem, et Wang 2005). Le pH optimal pour la croissance est presque neutre, mais certaines souches peuvent se développer à un pH de 10,5 et d'autres tolèrent un pH de 5,0. Les colonies sont circulaires, convexes, généralement opaques, parfois translucides, surélevées et mucilagineuses. Les colonies ont généralement un diamètre de 2 à 4 mm en 3 à 5 jours de croissance sur YM agar. La turbidité est prononcée après 2 à 3 jours dans un bouillon aéré ou agité. Les temps de génération sont de 3 à 6 h. Chimio-organo-trophe, elles utilisent une large gamme d'hydrates de carbone et de sels d'acides organiques comme sources de carbone en produisant de l'acide sans formation de gaz. À l'exception de la cellulose et l'amidon qui ne sont pas utilisés comme sources de carbone. Ces bactéries peuvent produire une réaction acide dans un milieu contenant du mannitol et des sels minéraux. La croissance sur milieu glucidique s'accompagne souvent d'une abondante production de mucus polysaccharidique extracellulaire. Les sels d'ammonium, les nitrates, les nitrites et la plupart des acides aminés peuvent servir de sources d'azote. Toutes les souches d'*Ensifer* nécessitent du pantothénate et de l'acide nicotinique. La peptone est peu utilisée. La caséine et la gélose ne sont pas hydrolysées. Les souches présentent divers degrés de spécificité d'hôte, déterminés par le type de facteurs Nod produits. Dans les nodules racinaires, les bactéries de l'espèce *Ensifer* se présentent sous forme d'endophyte présentant des formes pléomorphes, (bactéroïdes) qui réduisent ou fixent l'azote atmosphérique gazeux en une forme combinée utilisable par la plante hôte (Kuykendall, Hashem, et Wang 2005).

5. Mécanismes d'osmo-adaptation chez les bactéries

5.1 Effet des facteurs abiotiques sur l'activité bactérienne

Les rhizobiums à l'état libre sont plus tolérants aux stress abiotiques que les légumineuses (Zahran, 2003). Des facteurs abiotiques tels que le pH du sol, la salinité, la sécheresse ou la température peuvent interférer et affecter directement le processus de symbiose ou de nodulation, voire affecter l'activité de fixation de l'azote des rhizobiums (Kinkema et al., 2006). La majorité des légumineuses nécessitent un pH neutre ou légèrement acide pour établir une symbiose efficace dans le sol. Un sol à forte acidité affecte la solubilité des éléments minéraux et perturbe la nutrition minérale, affectant d'une part la croissance de la plante hôte, et d'autre part, l'efficacité des rhizobiums en réduisant ainsi les nodules. Le pH d'un sol alcalin a un effet négatif sur la

disponibilité de certains minéraux comme le fer et le manganèse tant pour le rhizobium que pour les plantes hôtes (Bordeleau et Prévost, 1994). Le stress salin affecte la synthèse de certains composants cellulaires de la bactérie, protéines et lipopolysaccharides. Les processus d'infection, l'initiation, le développement et l'activité des nodules et la fixation symbiotique de l'azote sont également très affecté par la salinité. Il a été démontré que le SNF (Symbiotic Nitrogen Fixation) est le processus le plus affecté par le sel (Rao et al., 2002). Ces données suggèrent l'intérêt de la sélection des légumineuses et des rhizobiums dans les zones à problèmes de salinité (Payakapong et al., 2006). La plage de tolérance au sel des rhizobiums peut varier en fonction de la souche. Certains sont inhibés par des concentrations de 100 mM de NaCl, tandis que ceux qui sont les plus tolérants peuvent survivre à des concentrations extrêmes de 1700 mM (Zahran, 2003). Les bactéries du genre *Ensifera* sont considérées comme très tolérantes et peuvent se développer dans des milieux contenant 300 à 700 mM de NaCl (Muller et Pereira 1995). Le stress hydrique affecte aussi la fixation symbiotique de l'azote à différents degrés : (i) Croissance et formation granulaires ; (ii) métabolisme du carbone et de l'azote ; (iii) activité nitrogénase ; (iv) perméabilité nodale à l'oxygène (Sadowsky 2005).

La sécheresse inhibe également les processus de la nodulation et la fixation d'azote. Les chercheurs ont montré qu'il existait une relation linéaire entre la présence de l'eau dans le sol et l'activité de fixation de l'acétylène utilisée pour estimer la fixation du diazote. Ceci relève la forte corrélation entre l'alimentation en eau et la fixation symbiotique du N₂. L'irrigation stimule donc la fixation d'azote et aussi stimule l'attachement symbiotique, en particulier chez les plantes non fertilisées. Si la fixation de l'azote semble être très sensible à toute limitation de l'approvisionnement en eau. Selon (Ben Rebah et al., 2007), l'irrigation d'appoint améliore de plus le rendement fourrager de 125 à 170% selon l'origine de la souche. Dans des conditions de pluviométrie limitée, l'inoculation des plantes par des souches adaptées a tout de même permis des améliorations significatives de la production fourragère (entre 87 et 91%) montrant ainsi l'importance des symbiotes. La température élevée provoque la déshydratation et la dégradation des enzymes dans les principales voies métaboliques des bactéries et limite la conductivité des sols. Cependant, les basses températures entraînent une gélification de l'eau intracellulaire et parfois une inactivation irréversible des enzymes. Il a été rapporté que des températures extrêmement basses inhibent l'expression des gènes nodulaires et donc infectent et forment des nodules (Zhang et al., 1996). Le stress thermique (pas extrême) induit souvent l'expression de protéine chaperonne et de protéine de choc thermique (HSP), assurant la protection d'enzymes importantes pour la physiologie microbienne (Cloutier et al., 1992). En revanche, une température élevée affecte la différenciation des rhizobiums ainsi que la fonction des nodules (Zahran, 1999).

5.2 Réponse générale au stress

La réponse au stress pour la plupart des bactéries étudiées regroupe multiples mécanismes de tolérance et d'adaptation pour survivre à des conditions extrêmes. Ces réactions entraînent des modifications de la morphologie bactérienne, de l'enveloppe cellulaire et des composants cytoplasmiques. L'adaptation au stress osmotique chez les bactéries se résume par les nombreux processus par lesquels les bactéries tendent à contrôler les mouvements (flux et écoulement) de l'eau et contrôler la biosynthèse/transport des solutés pour réguler la pression osmotique. Pour maintenir une perméabilité interne de la cellule bactérienne à une valeur supérieure à celle du milieu de culture (Stand, 2002). La disponibilité en eau de la cellule peut affecter sa survie et/ou sa croissance. Chez les micro-organismes, deux types de stratégies de régulation osmotique ont été décrites pour faire face au choc hyperosmotique : l'accumulation de sels dans le cytoplasme ("sels cytoplasmiques") et l'accumulation de solutés de solubilité cytoplasmique compatible permettant de maintenir la pression osmotique intra-cellulaire et empêcher la perturbation de l'ensemble des mécanismes et métabolismes bactériens (OByrne et Booth, 2002). Ainsi, les principaux changements morphologiques de la bactérie sous stress osmotique sont : formation de formes résistantes telles que les spores, altération des structures extracellulaires liées à l'activité cellulaire ou donnant aux bactéries la capacité d'adhérer aux surfaces ou même de former une couche biologique. Ces changements peuvent se produire dans de nombreuses bactéries. Chez *Bacillus subtilis*, par exemple, les spores se forment en réponse à des carences en nutriments (Errington, 1993). Chez *Escherichia coli*, la plupart des signaux environnementaux, en particulier les concentrations élevées de métaux, induisent des réponses au stress et conduisent à la formation de biofilms (Landini, 2009). Chez *Campylobacter jejuni*, la présence de flagellés rend la bactérie mobile et capable d'échapper à une acidité extrême (Reid et al., 2008). Dans le cas des rhizobiums, les bactéries ne sont pas capables de former des spores mais la présence de l'enveloppe joue un rôle essentiel dans la physiologie cellulaire et la réponse au stress. En effet, l'enveloppe bactérienne des rhizobiums est responsable de plusieurs processus tels : le transport, la production d'énergie, échange d'informations entre les cellules et la transduction de signaux. Sa localisation forme une interface entre le milieu et le milieu intracellulaire. Elle est donc la plus exposée aux stress dont elle constitue la première ligne de défense. Le stress osmotique affecte aussi le processus d'assemblage ou de dénaturation des protéines de la membrane externe (EMP) et induit une accumulation de protéines repliées dans la membrane externe et peut être toxique pour les bactéries. Ainsi, la réponse au stress nécessite une régulation physiologique de l'enveloppe cellulaire et en particulier la régulation de la production de nouvelles protéines par la membrane

externe (Kiliç et al., 2009). A l'intérieur de la cellule, les stress physico-chimiques altèrent rudement les protéines, l'ADN et l'ARN, créant des dommages importants ou par accumulation de produits toxiques (Fridovich, 1998). Deux mécanismes de réponse au stress peuvent être distingués : d'une part la protection et la réparation des composants macromoléculaires, et d'autre part, la détoxification du cytoplasme et le maintien de l'homéostasie. Dans la section suivante, un focus sur les réponses moléculaires impliquées dans le stress osmotique qui assure la défense et l'homéostasie est développé.

5.3 La stratégie « salt in cytoplasm », une stratégie spécifique aux micro-organismes halotolérants (extrêmes)

La stratégie bactérienne, appelée « salt in cytoplasm », est spécifique aux bactéries Halo et conduit à l'équilibre osmotique en maintenant une concentration en sel dans le cytoplasme équivalent à celle du milieu extérieur. La forte concentration de sels dans le cytoplasme nécessite des mécanismes pour stabiliser les structures protéiques en réduisant la répulsion entre les charges négatives et en augmentant leur hydratation. De plus, des concentrations élevées d'ions dans le cytoplasme ont le potentiel de perturber la physiologie cellulaire en provoquant l'agrégation par des interactions hydrophobes des macromolécules en réduisant la disponibilité en eau (Sleator et Hill, 2002). Cela nécessite des mécanismes d'adaptation spécifiques que l'on retrouve chez les bactéries spécialisées dans ce type de milieu, comme l'utilisation d'une plus grande proportion d'excès d'acides dans les protéines pour limiter l'hydrophobicité et également l'accumulation d'ions K^+ . Ainsi, les propriétés de l'hémophilie dépendent de la capacité des bactéries à accumuler des molécules protectrices afin de rétablir l'équilibre osmotique.

5.3.1 Les mécanismes d'accumulation intracellulaire de solutés compatibles solubles osmotiquement actifs

La bactérie rétablit une pression de turgescence positive en réponse au stress hyper-osmotique en accumulant du potassium dans le cytoplasme (réponse primaire) et/ou des solutés compatibles (réponse secondaire). Cette accumulation peut avoir lieu par biosynthèse endogène ou par transport depuis le milieu extérieur (Sleator et Hill, 2002).

5.3.1.1 Rôle des ions potassium (la réponse primaire)

Les ions potassium (K^+) jouent un rôle essentiel pour Les microorganismes mésophiles, dans le maintien de la pression de turgescence. La concentration en potassium augmente ainsi proportionnellement à l'augmentation de l'osmolalité du milieu extérieur. Il existe généralement

une corrélation positive entre la teneur en ions potassium dans le cytoplasme et la tolérance bactérienne à une osmolalité élevée. Il a été reconnu que le potassium sert également de « messenger secondaire » essentiel pour l'établissement efficace des systèmes cellulaires face aux effets de stress osmotique (Sleator et Hill, 2002).

5.3.1.2 Les solutés organiques compatibles (les réponses secondaires)

Les solutés compatibles sont moins perturbateurs pour la cellule bactérienne que le potassium (Bremee et Kramer, 2000). Ce sont des molécules organiques très solubles dans l'eau qui ont une charge nulle à pH physiologique et n'interagissent pas avec les protéines. Ils n'interfèrent pas avec les fonctions cellulaires importantes même à des concentrations élevées (>1M) (Sleator et Hill, 2002) et stabilisent indirectement les enzymes contre les effets des agents dénaturants tels que les sels, la température. La nature des solutés compatibles accumulés par les bactéries dépend de la présence de ces composés ou de leurs précurseurs dans le milieu extérieur, de l'existence de systèmes de transport et de synthèse. Il s'agit de sucres (comme le saccharose, le tréhalose ou le glucose) ou de polyols (glycérol, arabinol, mannitol, etc.), d'acides aminés (proline, glutamate, glycine) et de certains dérivés d'acides aminés (taurine, glycine bêtaïne, triméthylamine, -N-oxyde, ectoïne...). La glycine bêtaïne (GB) est mentionnée parmi les protecteurs osmotiques potentiels, qui peuvent s'accumuler dans le cytosol par absorption directe ou par synthèse de choline, de choline-O-sulfate ou de la carnitine (Sleator et Hill, 2002). En revanche, un grand nombre de bactéries transportent la bêtaïne dans les cellules pour être utilisée comme agent osmotique. Seules quelques bactéries (par exemple, *Actinopolyspora Halophila* et *Halomonas elongata*) et un méthanogène (*Methanohalophilus portucalensis FDF1*) sont capables de synthétiser la bêtaïne en oxydant la choline ou le méthylate de glycine (Sleator et Hill, 2002). Deux systèmes de transport des bêtaïnes ont été entièrement caractérisés : le système Hut impliqué dans le transport de la glycine-bêtaïne et le système BetS dans le transport de bêtaïne, choline et carnitine (Boscari *et al.*, 2002). Le transport de ces solutés est présent au sein des espèces de *Rhizobiaceae*. En général, il est rare que la glycine-bêtaïne soit le seul composé soluble compatible présent chez une halobactérie soumise à un stress salin. Elle est souvent accompagnée d'accumulation des sucres (saccharose, tréhalose), de polyalcools (sorbitol, glycérol) ou d'autres acides aminés et de leurs dérivés (glutamate, ectoïne). L'ectoïne est synthétisée par de nombreuses bactéries de deux genres halotolérantes et halophiles isolées de milieux dits hypersalins et son accumulation dépend du stade de croissance de la cellule bactérienne. Chez *Chromohalobacter israelensis*, l'ectoïne s'est accumulée uniquement lorsque les cellules étaient cultivées dans un milieu avec une concentration en sel supérieure à 0,6 M, et que les cellules se développaient de

façon exponentielle. Par conséquent, de nombreuses espèces de *Rhizobium* accumulent du glutamate dans leurs cellules (Zahran, 2001). Les acides aminés glutamate et proline sont des acides aminés anioniques et leur accumulation implique l'accumulation intracellulaire d'ions potassium. Chez certaines bactéries Gram-négatives se développant sous un stress osmotique, le glutamate possède plus de 90 acides aminés libres. Les disaccharides, saccharose et tréhalose, sont utilisés comme osmo-protéctants chez *E. meliloti* (Zahran, 1999). Le tréhalose est utilisé par les organismes pour lutter contre les effets de la dessiccation, mais il sert également comme un osmolyte. Chez *Chromohalobacter israelensis*, le tréhalose n'est important que lorsque les cellules sont cultivées à des concentrations inférieures à 0,6 M dans du NaCl. Chez la bactérie sulfato-réductrice *Desulfovibrio halophilus*, le tréhalose est le principal osmolyte utilisé lorsque les bactéries sont cultivées dans un milieu additionné avec (2,5 M) de NaCl et en absence de sources de bétaine. Le disaccharide non modifié, en particulier le saccharose, est un facteur important pour la survie des cellules en phase stationnaire sous stress salin et il est synthétisé dans les cyanobactéries et les protéobactéries. Il restaure également les voies métaboliques actives sous stress nutritionnel pendant la phase stationnaire.

CHAPITRE II

Les Bactéries fixatrices d'azote isolées du désert au Maroc et symbiotique de *Vachellia* pourraient appartenir à une nouvelle espèce de *Sinorhizobium* (*Ensifer*)

Article 1, Sakrouhi et al., 2016: Recovery of symbiotic nitrogen fixing acacia rhizobia from Merzouga Desert sand dunes in Southeast Morocco -Identification of a probable new species of *Ensifer* adapted to stressed environments.

Auteur: Ilham Sakrouhi^a, Meryem Belfquih^{a,b}, Laïla Sbabou^a, Patricia Moulin^{a,1}, Gilles Bena^{a,2}, Abdelkarim Filali-Maltouf^a & Antoine Le Quéré^{a,b,*}

^a Laboratoire de Microbiologie et de Biologie Moléculaire, Faculté des Sciences, Université Mohammed V - Agdal, Rabat, Morocco.

^b Institut de Recherche pour le Développement, Laboratoire des Symbioses Tropicales et Méditerranéennes, UMR113 IRD-CIRAD-SupAgro-UM-INRA, Campus International de Baillarguet, 34 398 Montpellier Cedex 5, France.

* Corresponding author. Tel. : +33 4 67593824 ; Fax: +33 4 67 593802. E-mail address: Antoine.le-quere@ird.fr (A. Le Quéré).

¹ Present address: Institut de Recherche pour le Développement (IRD), UMR210 Eco&Sols, Campus SupAgro, Place Viala, bât. 12, 34060 Montpellier Cedex 2, France.

² Present address: Institut de Recherche pour le Développement (IRD), UMR186 IPME, Centre IRD France-Sud, 911, avenue Agropolis, BP 64501, 34394 Montpellier cedex 5, France.

Running title: Recovery of rhizobia from desert sand dunes

Note : Nucleotide sequence data reported are available in the genbank database under the accession numbers : for the partial 16S rDNA (KP993276 - KP993295); for the partial *recA* gene (KP993231 - KP993239; KT991374 – KT991387); for the partial *dnaK* gene (KP993240 - KP993245); for the partial *glnA* gene (KP993246 - KP993254); for the partial *gltA* gene (KP993255 - KP993263); for the partial *nifH* gene (KP993264 - KP993268) and for the partial *nodC* gene (KP993269 - KP993275).

Abstract

Bacteria capable of nodulating *Acacia tortilis* and *A. gummifera* could be recovered from sand dunes collected in the Moroccan Merzouga desert. The trapping approach enabled the recovery of 17 desert rhizobia that all clustered within the *Ensifer* (*Sinorhizobium*) genus. Four isolates of the dominant genotype comprising 15 strains as well as 2 divergent strains were further characterized by MLSA. Phylogenetic analyzes indicated that the dominant genetic type was belonging to a new and yet undefined species within the *Ensifer* genus. Interestingly, housekeeping gene phylogenies showed that this possibly new species is also present in another desert but in India. Phylogenetic

analyses of *nifH* and *nodC* sequences showed high sequence conservation among the Moroccan strains belonging to the dominant genotype but high divergence with sequences from Indian isolates suggesting acquisition of symbiotic genes through Horizontal Gene Transfer. These desert rhizobia were capable of growing in media containing high salt concentrations, under high pH and most of the strains showed growth at 45°C. Only recovered from desert type of Biome, yet, this new taxon appears particularly adapted to such harsh environment.

Keywords: Hot Desert, Africa, *Acacia* rhizobia, *Ensifer*, Symbiosis, Phylogeny

Introduction

One of the most severe and widespread environmental problem that the agricultural industry now faces is the degradation of soil quality. This problem is partially due to intensification of cultures and human activities which, in vulnerable regions, are already impacted by climate changes that synergistically accelerates soil degradation and desertification [48]. If desertification affects many regions in the world and has been perceived as a major environmental concern for decades, its emphasis has led to over estimations in the past [59]. Nevertheless, a recent study reports that 10-20% of drylands, covering about 40% of land surfaces, are subjected to desertification which may directly impact over a quarter of million people in developing countries [50]. Photographs taken in 1868 and in 2008 in the semi arid northern region of Ethiopia show that restoration of a plant cover in degraded environments may be reached [45] and demonstrates the importance of land use management and the need of policies and actions that should integrate biophysical and socio economical issues for a sustainable development in arid and semi arid areas [50], [55].

Aiming at promoting plant cover restoration in arid and semi arid zones, acacias appear particularly interesting as several species are adapted to drought and salinity [7], [69]. Furthermore, they are able to stabilize and fertilize soils via nitrogen fixing and mycorrhizal symbioses [16], [20] and constitute sources of wood, fodder, medicine and gum [40]. The study of *Acacia* rhizobial diversity and the selection of efficient N-fixers may help in identifying the best plant-microsymbiont couples to be used in reforestation programs which require inoculation and preliminary growth in nurseries prior planting to be optimal [7]. Studies on the diversity of nitrogen fixing rhizobia that develop the symbiosis with acacias naturally growing in Africa have shown that this genus is associated with a broad range of species belonging to genera *Mesorhizobium*, *Ensifer*, *Rhizobium* and *Bradyrhizobium* [2], [6], [12], [13], [14], [15], [17], [28], [30], [35], [39], [44], [46], [52], [72]. Moreover, previous work on Australian acacias has showed that harsh

environmental conditions such as limited organic matter, heat, desiccation or salinity can affect acacia rhizobia diversity [3], [60].

With the objectives to promote acacia growth in arid regions in reforestation programs, we need to identify adapted symbiotic microorganisms that could, upon inoculation facilitate their installation and maintenance. To do so, a possibility consists in the identification of such bacteria directly from degraded plant free areas such as deserts that should lead to sustainment of the symbiotic system. Interestingly, hot deserts have been shown to hide a great diversity of microorganisms as exemplified by previous studies in America [19], [43], Africa [4], [8] or in Asia [1]. Furthermore, the comprehensive study on soil microbial communities based on comparative metagenomic data from various types of biomes including hot desert [19] clearly suggests that taxonomic and functional gene richness in hot desert samples are comparable to those found in forest Biomes. Notably, the 4 genera found to nodulate with *Acacia* (*Mesorhizobium*, *Bradyrhizobium*, *Rhizobium* and *Ensifer*) were identified among the 50 most abundant genera in all biomes studied with the exception of *Mesorhizobium* in cold desert samples [19].

Despite the great microbial diversity of soils, rhizobia can be quickly isolated by using a nodule trapping approach. It enables to rapidly select strains capable to develop a beneficial interaction with symbiotic leguminous plants of interest. Previous studies have showed that some bacteria capable of nodulating acacia species of interest in Morocco (one endemic species *A. gummifera* and the widely distributed *A. tortilis*) could be isolated from desert sand dunes of the Merzouga desert [18], [71]. Given the broad diversity of rhizobia capable to nodulate acacias, metagenomic studies indicating the presence of rhizobia in desert sand including that of Merzouga [23], and previous studies reporting isolation of symbiotic bacteria from *Acacia* nodules, it is expected that sand dunes of Merzouga contains rhizobia.

In this report, the presence and the diversity of bacteria originating from desert sand dunes that could develop the symbiosis with *Acacia* was studied. A new collection of desert rhizobia was generated and the taxonomical positioning among rhizobia was estimated using 16S rDNA and Multi Locus Sequence Analysis (MLSA) with 4 housekeeping marker genes *glnA*, *recA*, *gltA* and *dnaK*. The phylogenetic analysis of symbiotic genes *nodC* and *nifH* from selected isolates is presented. Finally, the fact that viable rhizobial strains could be recovered from such environment and the presence of isolates sharing the same taxonomical position in a distant desert region are discussed.

Materials and methods

Sampling and physico-chemical characterization of sand dune samples

Around 20 liters of sand were collected from four different sites in the dunes of the Merzouga desert at a depth of 15-25 cm and corresponding GPS coordinates were recorded. The four sites were completely free of vegetation. For the characterization of the total C and N content from collected samples, an aliquot of sand was air dried, passed through a 2mm sieve, grinded to 0.1mm and analyzed using a thermo-Fisher Flash EA 1112. Available phosphate content in soil sampled was measured using dried and 2mm sieved samples by colorimetric flux AA3 following a procedure described previously [11]. The pH characterization of sand samples was calculated as follow; 50 ml of distilled water was added to 20g of sand. After one hour of agitation, the solution was decanted for 30 minutes and the pH recorded.

Selection of rhizobia from desert sand dunes

The rhizobial isolates were obtained by trapping using two species of *Acacia* (*Acacia gummifera* and *Acacia tortilis* subsp. *raddiana*) on sand. Seeds of *Acacia* were surface sterilized using a solution of sulfuric acid at 95% (V/V) for 2 and half hours and washed 6 times with sterile water. The absence of microbes in water from the final wash was verified by plating onto fresh Yeast Extract – Mannitol (YEM) agar medium [63]. Seeds were then germinated on water agar (1%, W/V) medium at 28 °C in the dark for 72 h.

Seedlings were then placed aseptically either directly in sand samples from the Merzouga desert in closed magenta jars or in sterilized sand. In the latter case, 25 g of sand samples from the Merzouga desert were first mixed with 50 ml of physiological water for 90 minutes to resuspend bacterial cells, then 25 ml were used as inocula to trap nodule forming rhizobia while sterile physiological water was used to inoculate control plants. *Acacia* seedlings were watered using a sterile nitrogen-free Jensen solution [63] and a layer of sterile sand was added on top of the inoculated sand so as to avoid contaminations.

After 8-10 weeks in a growth chamber under a light/dark photoperiod of 16h/8h at a temperature of 22-23°C, the root systems were examined. When observed, nodules were excised and surface sterilized using a solution of 1% (W/V) CaCl_2O_2 for 3-15 min depending on the size of the nodule followed by six washings with sterile distilled water. Control plant roots never showed any nodule.

To isolate rhizobia, nodules were cut with a sterile razor blade, picked in the middle and directly streaked onto solid YEM medium. Small nodules were crushed in a drop of sterile distilled water and transferred to solid YEM. Pure colonies were obtained by repeated streaking onto YEM and/or TY media [5] and stored in 20% (V/V) glycerol at – 80°C. Once purified, isolates were further tested for their capacity to re-nodulate their original hosts.

Tolerance of isolated strains to salt, temperature and pH

To test the salt tolerance of isolated bacteria, strains were pre-grown in liquid cultures of YEM medium up to stationary phase, and the optical density at 630 nm was adjusted to 0.7. Ten μ l were then used to inoculate 200 μ l of YEM medium supplemented with NaCl to reach final mM salt concentrations ranging from 100 to 700 mM in steps of 100 in micro-well plates. The tolerance to salinity was estimated after 10 days growth in rotator shaker at 33°C by optical density measurements at 630nm using the BIOTEK Elx808 Absorbance microplate reader. To test growth abilities of the strains under acidic and alkaline pH, YEM medium was supplemented with 25Mm Homopipes to buffer the pH to 4 or 25mM TAPS (N-Tris(hydroxymethyl)methyl-3-aminopropanesulfonic acid) to buffer the pH to 9. The pH of media was checked after autoclaving and did not change by more than ± 0.1 unit. The ability to grow at high temperatures was tested on the strains following the same pre-cultivation procedure as described above. However YEM agar plates were used. Five μ l of the pre-cultures were spotted onto the plates. After drying of the inoculated cell suspensions under sterile flow hood, plates were incubated at 33, 37, 40, 45, 47.5 or 50 °C for 10 days and growth at the highest temperature was recorded. Two biological repetitions with three technical replicates were done for each strain.

Molecular Methods

Genomic DNA of bacteria was extracted following a previously described method [9]. Partial 16S rDNA were amplified and sequenced using the primers 16S-27f (5'- AGA GTT TGA TCM TGG CTC AG -3') [67] and 16S-1080r (5'- GGG ACT TAA CCC AAC ATC T -3') [56]. Secondly, for selected isolates, the partial gene sequences of the house keeping genes *recA*, *dnaK*, *glnA* and *gltA* were amplified and sequenced using the primer pairs (*recA*6F/555R, *dnaK*1466F/1777R, *glnA*572F/1142R and *gltA*428F/1111R) as described previously [37].

Sequence divergence of symbiotic genes from selected isolates was also studied using *nifH* and *nodC* marker gene amplification and sequencing. The *nodC* gene fragments were amplified and sequenced following previously published method [32]. The *nifH* sequences were obtained following a 2 step Nested PCR approach. First, the primer pair CDHPnifH53f [54] / nifHI [32] was used and cycling conditions were as follow: 95°C 5 min and 5 cycles : 94°C for 30 sec, 50°C for 30 sec, 72°C for 1 min followed by 30 cycles : 94°C for 30 sec, 56°C for 30 sec, 72°C for 1 min and a terminal incubation at 72°C for 7 min. One microliter of the resulting PCR product was then used as template in a second PCR using primers CDHPnifH53 / CDHPnifH723R [54] and the

following cycling conditions: 95°C 5 min followed by 35 cycles: 94°C for 30 sec, 55°C for 45 sec, 72°C for 1min and a terminal incubation at 72°C for 7 min.

The PCR products were purified and sequenced by Genoscreen Inc (Lille, France). The sequences produced in this study are deposited in the GenBank database under the following accession numbers: 16S rRNA (KP993276 - KP993295); dnaK (KP993240 - KP993245); glnA (KP993246 - KP993254); gltA (KP993255 - KP993263); recA (KP993231 - KP993239 and KT991374 – KT991387); nifH (KP993264 - KP993268); nodC (KP993269 - KP993275) (See Supplementary tables S1 and S2 for details).

Genomic fingerprinting was performed by REP-PCR using REP1R-I and REP2-I primers [62] as described previously [41].

Phylogenetic analyses

Sequences were manually corrected using BioEdit v9.0 software [27], aligned using Muscle as implemented in MEGA v5 [58]. Phylogenies were constructed using MEGA v5.2 by UPGMA using the Kimura 2 Parameter method or by Maximum Likelihood (ML) method using the best fitted ML model and parameters as indicated by the lowest Bayesian Information Criterion score. Tree topologies were tested using 1000 bootstrap replicates. We included in the phylogenetic reconstructions sequences retrieved from GeneBank and corresponding to defined *Ensifer* species or to other rhizobial species. Furthermore, to estimate the within-species distances among *Ensifer* species, we also included several sequences originating from isolates belonging to the same species in the MLSA. The list of strains and accessions used in the present study are reported in Supplementary tables S1 and S2.

Results

Estimation of total C and N content, available phosphate and pH of sand dune sampled

Sand samples from the Merzouga desert were collected at 4 locations (see GPS information from Supplementary table S3). They contained total carbon levels ranging from 0.10 to 0.65 % and nitrogen levels below the detection threshold (< 0.02 %). Regarding the available phosphorus, values ranged between 3 to 8 mg/kg. The pH of the 4 Merzouga desert sand samples was alkaline and ranged from 8.7 to 9 (Table S3).

Isolation of *Acacia* nodule bacteria from desert sand dune and tolerance to salt heat and alkaline pH

A total of 21 nodules could be recovered out of more than 300 seedlings. Only one could be found when using sand sample N° 4 (with *A. gummifera*) while the remaining 20 nodules were collected from *A. tortilis* root systems using sand samples from the site N° 2. Pure isolates were obtained from 20 nodules, cultivated and stored in glycerol 20% at -80°C. These 20 isolates were further tested for their ability to re-nodulate *Acacia* as well as for their tolerance to stresses induced by salt, high temperatures or pH (Table 1). All the strains showed some growth in YEM medium complemented with NaCl concentrations up to 400mM and one strain was capable of growing in YEM medium complemented with NaCl at 700mM. All isolates but 3 were capable of growing at 45°C and 10 isolates showed visible growth at 50°C. Finally, all isolates could grow in a medium buffered to an alkaline pH of 9 while no growth was observed at an acidic pH of 4. Finally, 17 strains were found to re nodulate *Acacia* seedlings (Table 1).

Phylogenetic analyses of 16S rDNA and housekeeping genes

To get insights into the diversity and the phylogenetic relationships among strains isolated from desert sand dunes, the 16S ribosomal subunit was first partially sequenced (720 bp) and analyzed using MEGA5. The phylogenetic positions of the 20 isolated strains as estimated by UPGMA showed that they all belong to the alpha-proteobacteria class within the family *Rhizobiaceae* (Supplementary figure S1). Sequences from 17 isolates out of 20 studied presented an identical partial 16S rDNA sequence with strains isolated from wild legumes in the Thar Desert (strains *Ensifer* sp. JNVU Ra9, TP6 and TW10) as well as to *E. kostiensis*. These 17 partial 16S-rRNA gene sequences were 99.8 and 99.4 % similar to Indian strain TP13 and *E.saheli* respectively. Additionally, 2 strains were found within the *Ensifer* clade. The strain LMR002 presented an identical partial 16S rRNA sequence to *E. garamanticus* (ORS 1400T) isolated from nodules of the herbaceous legume *Argyrolobium uniflorum* in the infra-arid zone of Tunisia [70] and *Ensifer* genospecies IV strains represented by *Ensifer* sp. AC25e which was isolated in Southern Ethiopia from root nodules of *Acacia tortillis*[14]. The second strain (LMR021) showed an identical partial 16S rDNA sequence to those of *E. fredii*, *E. xinjiangensis*, and *E. americanum* type strains as well as with *Ensifer* genospecies V (represented by *Ensifer* sp. AC38d2) comprising a large group of 22 Ethiopian isolates nodulating *Acacia tortilis* and *A. seyal*[14]. Finally, one strain (LMR014) grouped within the *Agrobacterium* clade and showed 98.8% and 97.3% of sequence similarities with *A. radiobacter* and *A. rubi* type strains respectively (Supplementary figure S1).

Given the high conservation of partial 16S rRNA gene sequences used here for assessment of the genera of newly identified isolates and to further estimate the diversity and phylogenetic

position, the widely used housekeeping *recA* gene, which is available for all type strains within the *Ensifer* genus, was partially sequenced for all 20 desert isolates and a ML phylogenetic tree was produced (Figure 1). Again, all 17 desert strains sharing the same partial 16S rRNA sequences possessed an identical partial *recA* sequence which was also found in the Indian strain *Ensifer* sp. JNVU Ra9. These 18 *recA* sequences were highly similar to the 2 Indian strains (*Ensifer* sp. JNVU TP6 and TW10) showing 97.3% of conserved residues while the strain TP13 shared 93.8% identical nucleotides (Figure 1B). The closest described species to this large group of desert rhizobia were *Ensifer saheli* (92.5% similarity) and *E. kostiensis* that showed, in contrast to the 16S rRNA, a greater *recA* sequence divergence with 91.4% of similarity. Given the high conservation of 16S rRNA and *recA* gene sequences between the 17 strains isolated from the Moroccan desert strains *Ensifer* sp. JNVU TP6 and Ra9, their genomic fingerprints were compared by REP-PCR. Using the REP 1R-1 and REP 2-1 primers, we found that all these strains shared a majority of bands with however 7 different profiles among the Moroccan isolates and 2 additional ones for the Indian strains (Supplementary figure S2) which confirms the close evolutionary relationships between these desert rhizobia that present however variable repetitive extragenic palindromic elements. The *recA* phylogeny confirmed the close proximity between the strain LMR002, the Ethiopian strain *Ensifer* sp. AC25e (99.7% similarity) and *E. garamanticus* (99.4% similarity). In contrast to the partial 16S rRNA based phylogeny that grouped the strain LMR021 together with *E. fredii*, *E. xinjuangensis* and *E. americanum*, the *recA* phylogeny indicated a divergent evolutionary relationship with these *Ensifer* species and suggested that it may belong to another genomospecies. Finally, the *recA* phylogeny showed that strain LMR014 was related to *Agrobacterium radiobacter* (sharing 93% of sequence similarity with a bootstrap value of 95%). Incapable to renodulate acacia seedlings when inoculated as pure culture, LMR014 was discarded from further analyses.

To gain resolution in the evolutionary relationships among isolated desert rhizobia, 4 strains representing the dominant genotype (LMR001, LMR006, LMR008 and LMR013), and the 2 divergent genotypes (LMR002 and LMR021) were selected and further studied using partial sequencing and phylogenetic analyzes of 4 housekeeping genes (*dnaK*, *glnA*, *gltA* and *recA*). The maximum likelihood (ML) phylogenetic tree obtained using concatenated sequences of the 4 markers is shown in figure 2, and trees based on single housekeeping genes *recA*, *dnaK*, *glnA* and *gltA* are shown in figure 1, supplementary figure S3-5 respectively. The ML trees confirmed the close proximity between strains representing the dominant genotype from the Merzouga desert (LMR001, LMR006, LMR008 and LMR013) together with strains isolated from the Thar desert in India (Ra9, TP6, TP13 and TW10), TP13 being slightly more divergent. We found sequence

similarities ranging from 99 to 100 % between concatenated housekeeping gene sequences from these 4 Moroccan strains and Indian isolates Ra9, TP6 and TW10. These 7 strains showed 96 % concatenated sequence similarities to TP13 concatenated sequence and the closest described species was *E. saheli* that showed 92.1 % similarity (Table S4). Pair wise evolutionary distances estimated on concatenated sequences within the clade grouping the 7 closest desert strains was extremely low and comparable to other intraspecies distances (Figure 2, Supplementary table S4) suggesting that these isolates may belong to the same new and yet undefined *Ensifer* species while the strain TP13 may represent a sister species.

Analyses of housekeeping genes *recA*, *glnA* and *gltA* from LMR002 confirmed the close proximity with *Ensifer garamanticus* ORS 1400T isolated from *Argyrolobium uniflorum* in Tunisia with sequence similarities above 99 % for *recA* and *glnA* and above 98 % for *gltA* (figure 1, supplementary figures S4 and S5 respectively). Regarding the strain LMR021, based on core gene phylogenies, we found that it may represent another and yet undefined *Ensifer* species with as closest neighbor *E. americanum* with which it presented 94 % and 94.3 % sequence similarities when *glnA* and *gltA* were compared while *recA* was more divergent. Globally, trees obtained using single core genes presented similar topologies except for species *Ensifer numiduicus*, *E. kostiensis* and *E. americanum* which showed variable evolutionary relationships depending on the marker used.

Phylogenetic analyses of symbiotic genes

To get insights into the diversity of the symbiotic genes among rhizobia isolated from the Merzouga desert, *nifH* and *nodC* genes were partially sequenced (605 bp and 817 bp respectively) and used to construct phylogenetic trees. Using these symbiotic markers, the tree topologies differed from those obtained using core genes (Figure 3). Symbiotic markers from Indian strains TP6 and TW10 isolated from wild *Tephrosia* plants in the Thar desert grouped with those of *E. sojiae* CCBAU 05699 and were highly divergent from *nifH* and *nodC* from the Moroccan strains. Both symbiotic markers from the dominant genotype representative strains recovered from the Merzouga desert (as shown for LMR001, LMR006 and LMR013) grouped together and diverged from those of LMR002 and LMR021. In the case of the *nifH* phylogeny (Figure 3A), sequences from LMR001, LMR006 and LMR013 (sharing more than 99.4% sequence similarities) were grouping (Bootstrap value of 80%) with *Ensifer* sp. HAMBI 1499 (> 99.1% similarity), *Ensifer kostiensis* strains HAMBI 1476 and HAMBI 1489 and *E. arboris* strain HAMBI 1700 (> 98% sequence similarities) which were all isolated in Africa from *Acacia* or *Prosopis* nodules. While

nifH sequence from LMR002 was highly divergent to other *nifH* sequences, the phylogenetic tree placed LMR021 sequence in a cluster that included all reported *Ensifer* *nifH* sequences found in Ethiopian isolates recovered from *Acacia* nodules [14] which was supported by a high bootstrap value of 97%, with as closest reference strain the *nifH* sequence of *Ensifer teranga* USDA4894; and more precisely into a subcluster (99% bootstrap value) including Ethiopian strains AC14d, AC27d1, AC01e, AC01c2 and AC47a that showed more than 99% similarities (Figure 3A). Beside LMR002 *nifH* sequence which was not related to any other sequences under study, *nifH* from other desert strains were included in a large cluster that was shown to represent the African continent in a previous study [14]. However, Indian strains recovered from native *Mimosa* (such as *Ensifer* sp. JNVU MHM12) were also included in this cluster.

Regarding the phylogenetic tree inferred from partial *nodC* sequences, more divergence was obtained (Figure 3B). Strains LMR-001 / -006 / -008 and -013 showed *nodC* sequences ($\geq 99.8\%$ similarity) that were closely related ($> 97.4\%$ similarity) to the Senegalese strain *E. arboris* LMG 14919 and from the Ethiopian *Ensifer* sp. AC01e which was representative of 13 strains [14]. Additionally, closely related *nodC* sequences from Indian native desert *Mimosa* symbionts (*Ensifer* sp. JNVU MHM-12, -3 and -4) were found ($\geq 96.6\%$ similarity) and formed a larger clade that was separated (bootstrap of 79%) from another group which included *nodC* sequences from LMR021 and 6 Ethiopian *Ensifer* strains (AC27a1, AC52C, AC47a, AC27d1, AC14d and AC38d2) that represent a set of 52 strains [14] with which it was 95.5 – 97.4% similar. As for *nifH* phylogeny, LMR002 *nodC* sequence was highly divergent as were those from the Indian TP6 and TW10 isolates that displayed proximity with those of *E. sojae* CCBAU 05699 and *Ensifer* sp. NGR234 (showing 94.3 and 94.6 % sequence similarity respectively).

Discussion

Sand dunes from the Merzouga desert contains viable symbiotic N₂ fixing nodulating bacteria

A microbial community study comparing several biomes showed that hot deserts present high diversity [19] and a recent study by Giongo et al., [22] has reported the successful recovery of rhizobia including *Ensifer* sp. from desert sand when using *Acacia albidia*, *A. spinosa* and *Vigna* as legume baits. One of the main objectives in the present study were to test whether rhizobial symbionts of the widespread *Acacia tortilis* subsp. *raddiana* or the endemic *A. gummifera* were present in desert sand dunes from the Merzouga desert in Morocco and examine their taxonomy. The use of leguminous trees such as *Acacia* particularly adapted to desert environments [7], [69] present particular interest for local populations in arid and semi arid zones such as in Morocco.

The two species *A. tortilis* and *A. gummifera* were chosen as hosts so as to constitute a collection of symbiotic isolates which may further be studied for their potential use as bio-fertilizer to facilitate their installation in such eco-climatic zones. Only 21 nodules were recovered out of more than 300 plants. The very low nodulation frequency was somehow expected as the sites sampled were located in arid zones receiving extremely low rain-falls and totally free of vegetation thus inhibiting the production of organic matter. The absence of a plant cover may explain the very low N and available P content found in the sand samples (under the detection threshold 0.02% and ranging from 3 to 8 mg/kg, respectively, table S3) which limit growth of organisms. The total carbon content was relatively high (ranging from 0.09 to 0.66%) and may be due to the mineral form (probably calcium carbonate) which is favored under alkaline pH. The number of nodules recovered varied greatly with sites sampled indicating a spatial heterogeneity of rhizobial abundance in the Merzouga desert. A previous study using direct count from multiple sites and in several desert regions has observed some variability in microbial abundance among samples collected within the same desert region [24]. Furthermore, desert sand dunes are moving and winds can occasionally bring bacteria which may explain rhizobial frequency variations across sites. The reason why we found a single nodule on *A. gummifera* is not known, and increasing the length of the nodulation assays with the number of plants may help in recovering higher nodule number when using this *Acacia* species. Nevertheless, the strain LMR001 was also capable to nodulate *A. tortilis*.

The 20 isolates were all fast growing and highly mucous when grown on Yeast – Mannitol Agar medium. Their tolerance to salt and high temperatures was studied and showed some variability. However, they were all growing in media containing high salt (400 mM) and they were all capable of growing at 37°C with a majority showing positive growth at 45°C on solid media which indicates an overall tolerance to high temperatures. This is not surprising as these isolates were collected from sand dunes where temperature often reaches 40°C even at 15 cm of depth. Finally, their tolerance to an alkaline pH of 9 was also expected as this is in the range of pH found from sand dune sampled. Our results show an overall tolerance to salt induced hyper-osmotic stress, temperature and alkaline pH that are major constraints found in arid soils, and demonstrate the interest to study rhizobia originating from such less explored harsh environments.

The mechanisms whereby rhizobial cells are able to sustain in deserts, which impose high UV radiations, extreme temperatures and low access to essential components such as water or nutrients, are not known. Microorganisms that are particularly adapted to deserts include Gram-positive bacteria or spore-forming species that allow the organisms to sustain under unfavorable conditions for long periods of time and until appropriate conditions are restored in their proximate

environment. However, rhizobia are gram negative, non-spore forming and the mechanisms whereby they can survive remains to be addressed even though some clues on their metabolic and phenotypic responses to desiccation have been studied [10], [65] and showed that the physiological state preceding the desiccation impacted their survival [65]. Notably, it can be expected that environmental conditions found in deserts and which precede desiccation including a progressive water loss concomitant to an osmotic stress may contribute to their desiccation tolerance and survival. Additionally, one study reported the formation of cyst like resting forms in a *S. meliloti* strain [34] which may further be investigated as this has only been observed once to date.

The majority of *Acacia* rhizobial symbionts from the Merzouga desert belong to a yet undefined *Ensifer* species that was also recovered from the Thar desert in India

Based on the analysis of partial 16S rDNA sequences, all the isolates capable of forming nodules on *Acacia* upon re-nodulation assays were found within the *Ensifer* genus among which 15 presented identical sequences. Using *recA* as taxonomical marker, this dominant genotype did not show close proximity to any defined species while the 2 remaining strains LMR002 and LMR021 were found closely related to *Ensifer garamanticus* and *E. fredii* respectively. All strains that were capable to renodulate acacia here belong to the fast growing *Ensifer* genus. Several *Ensifer* species nodulating acacias have already been described including *E. teranga*[13], *E. mexicanus*[33], *E. americanum*[61], [66] and *E. kostiensis*[44] and studies on acacia rhizobia diversity suggest that several new species may be defined in the future [6], [14], (this study). Interestingly, acacia rhizobia belonging to this genus were found preferentially associated to Australian acacias in soils subjected to heat, water stress or salinity [3], [60] while *Bradyrhizobium* and *Rhizobium* strains were dominant in more temperate regions [31] which demonstrates the interest to investigate less explored stressed environments where adapted rhizobia can be recovered.

Surprisingly, 16S rDNA and *recA* partial sequences obtained from the most frequent genotype was identical or highly similar to that of isolates recently identified from another desert in India which nodulated diverse wild leguminous species, suggesting that this new clade is also present in other parts of the world. Four marker genes shown to fit evolutionary relationships within the *Ensifer* genus [37] confirmed the close proximity between Indian and Moroccan desert *Ensifer* strains. The low sequence divergence detected between these strains was also confirmed and, together with the genetic distance with the other *Ensifer* species, suggested that these isolates may belong to a common and yet undefined *Ensifer* species.

Being recovered from two distant deserts, this putative new species seems particularly adapted to such environment. One possible explanation regarding their presence in such distant geographical locations is dispersal through sand storms, which can transport high number and diverse microorganisms present in desert dust storm regions [24] through long distances and allow trans-Mediterranean and trans oceanic trips [25], [26], [29], [53], [68]. Such transportation vector has already been hypothesized for rhizobia following a population genetics study which demonstrated a high level of migration of *Bradyrhizobium canariense* between Morocco and the Canarian archipelago [64].

The divergent strain LMR002 was, on the basis of 3 core gene phylogenies (*recA*, *glnA* and *gltA*, Figures 1, S2 and S4), related to *Ensifer garamanticus* ORS 1400T isolated from herbaceous legumes in another arid zone in Southern Tunisia by Zakhia and colleagues [70] and was, based on 16S rDNA and *recA* phylogenies, also closely related to other isolates identified from *Acacia* nodules in southern Ethiopia [14]. Interestingly, in the later study, many isolates (representing several genospecies) showed identical partial 16S rDNA sequences with our last isolate LMR021 but based on the *recA* phylogeny, this strain appeared to belong to another genospecies, not yet described. Despite the lack of data on several studied housekeeping gene sequences from Ethiopian or Tunisian isolates, our results indicate that rhizobia found in the Merzouga desert possess close relatives that are all originating from arid zones in Africa or Asia which suggests environmental driven adaptation and speciation.

The reason why strains recovered from the Merzouga and Thar deserts that from a new clade within the *Ensifer* genus have not been yet observed in other types of environments remains to be investigated. Such scarcity is puzzling as one would expect a wide spread of this new group, especially when desert sand and dust is considered as a major source of materials transported over long distances which should favor spreading of this group worldwide and more specially in regions on the path of major atmospheric fluxes [25]. One possible explanation is a poor competitiveness of isolates native of deserts as compared with those present in areas subjected to lower abiotic stresses, higher nutrient availability and whose presence with their hosts is ancient. Noteworthy, one of the hypotheses resulting from the comparative metagenomic work performed by Fierer and colleagues [19] was that abiotic stresses such as those found in hot deserts are more important in shaping the microbial community than competitiveness as shown by a systematically reduced proportion of antibiotic resistance genes found in such ecosystems. Such extrapolation to sub-population levels may partly explain why new rhizobial strains identified from the Merzouga desert are not yet recovered from other less perturbed areas as these may be less competitive as compared to native rhizobia. Finally, regarding the very low level of nodulation rate in our study,

and the fact that two among four samples did not resulted in any nodules, suggest that this new nodulating species might be recovered from other hot deserts only with an important and multisite trapping scheme.

Analyses of symbiotic marker genes suggest that the new desert *Ensifer* group identified has acquired host-determinant genes through horizontal transfer

Isolates belonging to the new *Ensifer* clade that grouped desert Indian strains (TP6, TW10) with Moroccan ones (LMR001, LMR006, LMR013) harbored divergent *nodC* and *nifH* genes. This diversity was correlated to the geographical origin of isolates. Such result was not surprising as symbiotic bacteria were recovered from different hosts (*Tephrosia* vs *Acacia*) and additional studies aiming at defining the host ranges of these desert rhizobia may reveal their classification into different symbiovars as shown in other species of rhizobia [51]. It is well known that symbiotic genes are carried by mobile genetic elements, like genomic islands or symbiotic plasmids, that may allow bacteria to promote nodulation upon acquisition through HGT [21], [38], [42], [47], [49]. The symbiotic genes from the dominant group identified here were closely related to those of *Ensifer* sp. HAMBI 1499, *E. kostiensis*, and *E. arboris* strains that were all recovered from *Acacia* or *Prosopis* (sharing the same inoculation group) nodules in Africa suggesting acquisition through HGT.

The results we presented here raise several questions and notably on the rhizobial mode of dispersion that may rely on sand storm events which, combined to atmospheric transport could maintain a flux of microorganisms throughout the world. However, the fact that these bacteria are recovered uniquely from such stressed environments suggests an adaptation of these strains and/or a poor competitiveness in environments with lower abiotic stress pressure. A possible explanation could be that for those organisms that are able to sustain in such hot deserts, abiotic stress adaptation including high soil pH and low moisture may prevail over competitiveness whose determinants may have been lost during evolution. Conversely, analyses of symbiotic markers suggest that this possible new species is capable to acquire heterologous genetic material through HGT. However, much work remains to be done to test these hypotheses.

Interestingly, in the frame of a large sequencing project aiming at building a Genomic Encyclopedia of Archae and Bacteria (GEBA) that will be available through the Integrated Microbial Genome system [36], the genomes of 100 root nodulating bacteria coordinated by Professor Wayne Reeves are being sequenced. Interestingly, one of the strains selected was the Indian isolate *Ensifer* sp. TW10 that cluster with the dominant genotype identified in the present study. The TW10 genome has been released [57] and may offer a good basis to study the genomic

background relative to the adaptive traits explaining why this new group has so far only been isolated from hot deserts. Finally, this possibly new species may present interest in biotechnology. Tolerant to salinity, high temperatures and pH, the analysis of symbiotic genes shows that this new desert rhizobial species are capable of acquiring variable symbiotic genes which opens up the range of legumes to which it may be associated with in arid and semi-arid areas.

Acknowledgments

This study was financially supported by the Institute of research for the Development (IRD) through the LMI LBMV and Programme d'Urgence de l'Académie Hassan II from Morocco. We would like to thank L'OREAL – UNESCO for the promotion of Women in Science that awarded Ilham Sakrouhi a prize for her PhD project; as well as the IRD for the ARTS PhD grant awarded to Meryem Belfquih. Finally, the authors would like to thank Pr. Jamal Aurag and Pr. El Bekkay Berraho for their scientific support as well as Prof Hukam Singh Gehlot for providing the strains of *Ensifer* sp. JNVU Ra9, TP6 and TP13 isolated from the Thar desert in India.

References

- [1] An, S., Couteau, C., Luo, F., Neveu, J., Dubow, M.S. (2013) Bacterial Diversity of Surface Sand Samples from the Gobi and Taklamaken Deserts. *Microb. Ecol.* 66, 850-860.
- [2] Ba, S., Willems, A., De Lajudie, P., Roche, P., Jeder, H., Quatrini, P., Neyra, M., Ferro, M., Prome, J.C., Gillis, M., Boivin-Masson, C., Lorquin, J. (2002) Symbiotic and taxonomic diversity of rhizobia isolated from *Acacia tortilis* subsp. *raddiana* in Africa. *Syst. Appl. Microbiol.* 25, 130-145.
- [3] Barnett, Y.M., Catt, P.C. (1991) Distribution and Characteristics of Root-Nodule Bacteria Isolated from Australian *Acacia*-Spp. *Plant and Soil* 135, 109-120.
- [4] Benzerara, K., Chapon V, Moreira D, Lopez-Garcia P, Guyot F & Heulin T (2006) Microbial diversity on the Tatahouine meteorite. *Meteor. Planet. Sci.* 41, 1249-1265.
- [5] Beringer, J.E. (1974) R factor transfer in *Rhizobium leguminosarum*. *J. Gen. Microbiol.* 84, 188-198.
- [6] Boukhatem, Z.F., Domergue, O., Bekki, A., Merabet, C., Sekkour, S., Bouazza, F., Duponnois, R., De Lajudie, P., Galiana, A. (2012) Symbiotic characterization and diversity of rhizobia associated with native and introduced acacias in arid and semi-arid regions in Algeria. *FEMS Microbiol. Ecol.* 80, 534-547.
- [7] Brockwell, J., Searle, S.D., Jeavons, A.C. And Waayers, M., Nitrogen fixation in acacias: an untapped resource for sustainable plantations, farm forestry and land reclamation. in: ACIAR Monograph, 2005, pp. 132.

- [8] Chanal, A., Chapon, V., Benzerara, K., Barakat, M., Christen, R., Achouak, W., Barras, F., Heulin, T. (2006) The desert of Tataouine: an extreme environment that hosts a wide diversity of microorganisms and radiotolerant bacteria. *Environ. Microbiol.* 8, 514-525.
- [9] Chen, W.P., Kuo, T.T. (1993) A simple and rapid method for the preparation of gram-negative bacterial genomic DNA. *Nucleic Acids Res.* 21, 2260.
- [10] Cytryn, E.J., Sangurdekar, D.P., Streeter, J.G., Franck, W.L., Chang, W.S., Stacey, G., Emerich, D.W., Joshi, T., Xu, D., Sadowsky, M.J. (2007) Transcriptional and physiological responses of *Bradyrhizobium japonicum* to desiccation-induced stress. *J. Bacteriol.* 189, 6751-6762.
- [11] Dabin, P. (1980) Phosphorus deficiency in tropical soils as a constraint on agricultural output, Priorities for alleviating soil-related constraints to food production in the tropics. IRRI, Los Banos (Philippines), 17.
- [12] De Lajudie, P., Willems, A., Nick, G., Moreira, F., Molouba, F., Hoste, B., Torck, U., Neyra, M., Collins, M.D., Lindstrom, K., Dreyfus, B., Gillis, M. (1998) Characterization of tropical tree rhizobia and description of *Mesorhizobium plurifarum* sp. nov. *Int. J. Syst. Evol. Microbiol.* 48, 369-382.
- [13] De Lajudie, P., Willems, A., Pot, B., Dewettinck, D., Maestrojuan, G., Neyra, M., Collins, M.D., Dreyfus, B., Kersters, K., Gillis, M. (1994) Polyphasic taxonomy of rhizobia; emendation of the genus *Sinorhizobium* and description of *Sinorhizobium meliloti* com. nov., *Sinorhizobium saheli* sp. nov., and *Sinorhizobium teranga* sp. nov. *Int. Syst. Bacteriol.* 44, 715-733.
- [14] Degefu, T., Wolde-Meskel, E., Frostegard, A. (2012) Phylogenetic multilocus sequence analysis identifies seven novel *Ensifer* genospecies isolated from a less-well-explored biogeographical region in East Africa. *Int. J. Syst. Evol. Microbiol.* 62, 2286-2295.
- [15] Diouf, D., Fall, D., Chaintreuil, C., Ba, A.T., Dreyfus, B., Neyra, M., Ndoye, I., Moulin, L. (2010) Phylogenetic analyses of symbiotic genes and characterization of functional traits of *Mesorhizobium* spp. strains associated with the promiscuous species *Acacia seyal* Del. *J. Appl. Microbiol.* 108, 818-830.
- [16] Duponnois, R., Plenchette, C., Prin, Y., Ducousso, M., Kisa, M., Ba, A.M., Galiana, A. (2007) Use of mycorrhizal inoculation to improve reafforestation process with Australian *Acacia* in Sahelian ecozones. *Ecol. Eng.* 29, 105-112.
- [17] Dupuy, N., Willems, A., Pot, B., Dewettinck, D., Vandenbruaene, I., Maestrojuan, G., Dreyfus, B., Kersters, K., Collins, M.D., Gillis, M. (1994) Phenotypic and genotypic characterization of bradyrhizobia nodulating the leguminous tree *Acacia albida*. *Int. J. Syst. Bacteriol.* 44, 461-473.

- [18] Essendoubi, M., Brhada, F., Eljamali, J.E., Filali-Maltouf, A., Bonnassie, S., Georgeault, S., Blanco, C., Jebbar, M. (2007) Osmoadaptative responses in the rhizobia nodulating *Acacia* isolated from south-eastern Moroccan Sahara. *Environ. Microbiol.* 9, 603-611.
- [19] Fierer, N., Leff, J.W., Adams, B.J., Nielsen, U.N., Bates, S.T., Lauber, C.L., Owens, S., Gilbert, J.A., Wall, D.H., Caporaso, J.G. (2012) Cross-biome metagenomic analyses of soil microbial communities and their functional attributes. *Proc. Natl. Acad. Sci. U.S.A.* 109, 21390-21395.
- [20] Galiana, A., Bouillet, J.P., Ganry, F. (2004) The importance of biological nitrogen fixation by trees in agroforestry. *Symbiotic Nitrogen Fixation*. Enfield : Science Publishers, pp. 185-199.
- [21] Gibson, K.E., Kobayashi, H., Walker, G.C. (2008) Molecular Determinants of a Symbiotic Chronic Infection. *Annu. Rev. Genet.* 42, 413-441.
- [22] Giongo, A., Favet, J., Lapanje, A., Gano, K.A., Kennedy, S., Davis-Richardson, A.G., Brown, C., Beck, A., Farmerie, W.G., Cattaneo, A., Crabb, D.B., Aung, Y.Y., Kort, R., Brumsack, H.J., Schnetger, B., Broughton, W.J., Gorbushina, A.A., Triplett, E.W. (2013) Microbial hitchhikers on intercontinental dust: high-throughput sequencing to catalogue microbes in small sand samples. *Aerobiol.* 29, 71-84.
- [23] Gommeaux, M., Barakat, M., Montagnac, G., Christen, R., Guyot, F., Heulin, T. (2010) Mineral and Bacterial Diversities of Desert Sand Grains from South-East Morocco. *Geomicrobiol. J.* 27, 76-92.
- [24] Gonzalez-Martin, C., Teigell-Perez, N., Lyles, M., Valladares, B., Griffin, D.W. (2013) Epifluorescent direct counts of bacteria and viruses from topsoil of various desert dust storm regions. *Res. Microbiol.* 164, 17-21.
- [25] Griffin, D.W. (2007) Atmospheric movement of microorganisms in clouds of desert dust and implications for human health. *Clin. Microbiol. Rev.* 20, 459-477.
- [26] Griffin, D.W., Garrison, V. H., Herman, J. R., Shinn, E. A. (2001) African desert dust in the Caribbean atmosphere: Microbiology and public health. *Aerobiol.* 17, 11.
- [27] Hall, T.A. (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Acids. Symp. Ser.* 41, 4.
- [28] Haukka, K., Lindstrom, K., Young, J.P.W. (1998) Three phylogenetic groups of *nodA* and *nifH* genes in *Sinorhizobium* and *Mesorhizobium* isolates from leguminous trees growing in Africa and Latin America. *Appl. Environ. Microbiol.* 64, 419-426.
- [29] Katra, I., Arotsker, L., Krasnov, H., Zaritsky, A., Kushmaro, A., Ben-Dov, E. (2014) Richness and Diversity in Dust Stormborne Biomes at the Southeast Mediterranean. *Sci. Rep.* 4: 5265.

- [30] Khbaya, B., Neyra, M., Normand, P., Zerhari, K., Filali-Maltouf, A. (1998) Genetic diversity and phylogeny of rhizobia that nodulate acacia spp. in morocco assessed by analysis of rRNA genes. *Appl. Environ. Microbiol.* 64, 4912-4917.
- [31] Lafay, B., Burdon, J.J. (2001) Small-subunit rRNA genotyping of rhizobia nodulating Australian *Acacia* spp. *Appl. Environ. Microbiol.* 67, 396-402.
- [32] Laguerre, G., Nour, S.M., Macheret, V., Sanjuan, J., Drouin, P., Amarger, N. (2001) Classification of rhizobia based on *nodC* and *nifH* gene analysis reveals a close phylogenetic relationship among *Phaseolus vulgaris* symbionts. *Microbiology*. 147, 981-993.
- [33] Lloret, L., Ormeno-Orrillo, E., Rincon, R., Martinez-Romero, J., Rogel-Hernandez, M.A., Martinez-Romero, E. (2007) *Ensifer mexicanus* sp. nov. a new species nodulating *Acacia angustissima* (Mill.) Kuntze in Mexico. *Syst. Appl. Microbiol.* 30, 280-290.
- [34] Loiko, N.G., Kryazhevskikh, N.A., Suzina, N.E., Demkina, E.V., Muratova, A.Y., Turkovskaya, O.V., Kozlova, A.N., Galchenko, V.F., El'-Registan, G.I. (2011) Resting forms of *Sinorhizobium meliloti*. *Microbiology*. 80, 472-482.
- [35] Lortet, G., Mear, N., Lorquin, J., Dreyfus, B., Delajudie, P., Rosenberg, C., Boivin, C. (1996) Nod factor thin-layer chromatography profiling as a tool to characterize symbiotic specificity of rhizobial strains: Application to *Sinorhizobium saheli*, *S. teranga*, and *Rhizobium* sp strains isolated from *Acacia* and *Sesbania*. . *Mol. Plant-Microbe Interact.* 9, 736-747.
- [36] Markowitz, V.M., Szeto, E., Palaniappan, K., Grechkin, Y., Chu, K., Chen, I.M., Dubchak, I., Anderson, I., Lykidis, A., Mavromatis, K., Ivanova, N.N., Kyrpides, N.C. (2008) The integrated microbial genomes (IMG) system in 2007: data content and analysis tool extensions. *Nucleic Acids Res.* 36, D528-533.
- [37] Martens, M., Delaere, M., Coopman, R., De Vos, P., Gillis, M., Willems, A. (2007) Multilocus sequence analysis of *Ensifer* and related taxa. *Int. J. Syst. Evol. Microbiol.* 57, 489-503.
- [38] Masson-Boivin, C., Giraud, E., Perret, X., Batut, J. (2009) Establishing nitrogen-fixing symbiosis with legumes: how many rhizobium recipes? *Trends Microbiol.* 17, 458-466.
- [39] Mcinroy, S.G., Campbell, C.D., Haukka, K.E., Odee, D.W., Sprent, J.I., Wang, W.J., Young, J.P., Sutherland, J.M. (1999) Characterisation of rhizobia from African acacias and other tropical woody legumes using Biolog and partial 16S rRNA sequencing. *FEMS Microbiol. Lett.* 170, 111-117.
- [40] Midgley, J.J., Bond, W.J. (2001) A synthesis of the demography of African acacias. *J. Trop. Ecol.* 17, 871-886.

- [41] Mohapatra, B.R., Broersma, K., Mazumder, A. (2007) Comparison of five rep-PCR genomic fingerprinting methods for differentiation of fecal *Escherichia coli* from humans, poultry and wild birds. FEMS Microbiol. Lett. 277, 98-106.
- [42] Moran, N.A., Degnan, P.H., Santos, S.R., Dunbar, H.E., Ochman, H. (2005) The players in a mutualistic symbiosis: insects, bacteria, viruses, and virulence genes. Proc. Natl. Acad. Sci. U.S.A. 102, 16919-16926.
- [43] Neilson, J.W., Quade, J., Ortiz, M., Nelson, W.M., Legatzki, A., Tian, F., Lacombe, M., Betancourt, J.L., Wing, R.A., Soderlund, C.A., Maier, R.M. (2012) Life at the hyperarid margin: novel bacterial diversity in arid soils of the Atacama Desert, Chile. Extremophiles. 16, 553-566.
- [44] Nick, G., De Lajudie, P., Eardly, B.D., Suomalainen, S., Paulin, L., Zhang, X., Gillis, M., Lindstrom, K. (1999) *Sinorhizobium arboris* sp. nov. and *Sinorhizobium kostiense* sp. nov., isolated from leguminous trees in Sudan and Kenya. Int. J. Syst. Bacteriol. 49, 1359-1368.
- [45] Nyssen, J., Haile, M., Naudts, J., Munro, N., Poesen, J., Moeyersons, J., Frankl, A., Deckers, J., Pankhurst, R. (2009) Desertification? Northern Ethiopia re-photographed after 140 years. Sci Total Environ. 407, 2749-2755.
- [46] Odee, D.W., Haukka, K., Mcinroy, S.G., Sprent, J.I., Sutherland, J.M., Young, J.P.W. (2002) Genetic and symbiotic characterization of rhizobia isolated from tree and herbaceous legumes grown in soils from ecologically diverse sites in Kenya. Soil Biol. Biochem. 34, 801-811.
- [47] Oldroyd, G.E., Murray, J.D., Poole, P.S., Downie, J.A. (2011) The rules of engagement in the legume-rhizobial symbiosis. Annu. Rev. Genet. 45, 119-144.
- [48] Puigdefabregas, J. (1998) Ecological impacts of global change on drylands and their implications for desertification. Land Degrad. & Dev. 9, 393-406.
- [49] Remigi, P., Capela, D., Clerissi, C., Tasse, L., Torchet, R., Bouchez, O., Batut, J., Cruveiller, S., Rocha, E.P., Masson-Boivin, C. (2014) Transient hypermutagenesis accelerates the evolution of legume endosymbionts following horizontal gene transfer. PLoS Biol. 12, e1001942.
- [50] Reynolds, J.F., Smith, D.M., Lambin, E.F., Turner, B.L., 2nd, Mortimore, M., Batterbury, S.P., Downing, T.E., Dowlatabadi, H., Fernandez, R.J., Herrick, J.E., Huber-Sannwald, E., Jiang, H., Leemans, R., Lynam, T., Maestre, F.T., Ayarza, M., Walker, B. (2007) Global desertification: building a science for dryland development. Science. 316, 847-851.
- [51] Rogel, M.A., Ormeno-Orrillo, E., Martinez Romero, E. (2011) Symbiovars in rhizobia reflect bacterial adaptation to legumes. Syst. Appl. Microbiol. 34, 96-104.

- [52] Romdhane, S.B., Nasr, H., Samba-Mbaye, R., Neyra, M., Ghorbal, M.H., De Lajudie, P. (2006) Genetic diversity of *Acacia tortilis* ssp. *raddiana* rhizobia in Tunisia assessed by 16S and 16S-23S rDNA genes analysis. J. Appl. Microbiol. 100, 436-445.
- [53] Smith, D.J., Timonen, H.J., Jaffe, D.A., Griffin, D.W., Birmele, M.N., Perry, K.D., Ward, P.D., Roberts, M.S. (2013) Intercontinental dispersal of bacteria and archaea by transpacific winds. Appl. Environ. Microbiol. 79, 1134-1139.
- [54] Steward, G.F., Jenkins, B.D., Ward, B.B., Zehr, J.P. (2004) Development and testing of a DNA macroarray to assess nitrogenase (nifH) gene diversity. Appl. Environ. Microbiol. 70, 1455-1465.
- [55] Stringer, L.C., Dyer, J.C., Reed, M.S., Dougill, A.J., Twyman, C., Mkwambisi, D. (2009) Adaptations to climate change, drought and desertification: local insights to enhance policy in southern Africa. Environ. Sci. Policy. 12, 748-765.
- [56] Sy, A., Giraud, E., Samba, R., De Lajudie, P., Gillis, M., Dreyfus, B. (2001) Nodulation of certain legumes of the genus *Crotalaria* by the new species *Methylobacterium*.. Can. J. Microbiol. 47, 503-508.
- [57] Tak, N., Gehlot, H.S., Kaushik, M., Choudhary, S., Tiwari, R., Tian, R., Hill, Y., Brau, L., Goodwin, L., Han, J., Liolios, K., Huntemann, M., Palaniappan, K., Pati, A., Mavromatis, K., Ivanova, N., Markowitz, V., Woyke, T., Kyrpides, N., Reeve, W. (2013) Genome sequence of *Ensifer* sp. TW10; a *Tephrosia wallichii* (Biyani) microsymbiont native to the Indian Thar Desert. Stand. Genomic Sci. 9, 304-314.
- [58] Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S. (2011) MEGA5: Molecular Evolutionary Genetics Analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28, 2731-2739.
- [59] Thomas, D.S.G. (1993) Sandstorm in a teacup - Understanding desertification. Geogr. J. 159, 318-331.
- [60] Thrall, P.H., Broadhurst, L.M., Hoque, M.S., Bagnall, D.J. (2009) Diversity and salt tolerance of native *Acacia* rhizobia isolated from saline and non-saline soils. Austral. Ecol. 34, 950-963.
- [61] Toledo, I., Lloret, L., Martinez-Romero, E. (2003) *Sinorhizobium americanus* sp. nov., a new *Sinorhizobium* species nodulating native *Acacia* spp. in Mexico. Syst. Appl. Microbiol. 26, 54-64.
- [62] Versalovic, J., Koeuth, T., Lupski, J.R. (1991) Distribution of repetitive DNA sequences in eubacteria and application to fingerprinting of bacterial genomes. Nucleic Acids Res. 19, 6823-6831.

- [63] Vincent, J.M., A Manual for the Practical Study of the Root-Nodule Bacteria. Blackwell scientific pub., Oxford, 1970.
- [64] Vinuesa, P., Silva, C., Werner, D., Martinez-Romero, E. (2005) Population genetics and phylogenetic inference in bacterial molecular systematics: the roles of migration and recombination in *Bradyrhizobium* species cohesion and delineation. Mol. Phylogenet. Evol. 34, 29-54.
- [65] Vriezen, J.a.C., De Bruijn, F.J., Nusslein, K. (2007) Responses of rhizobia to desiccation in relation to osmotic stress, oxygen, and temperature. Appl. Environ. Microbiol. 73, 3451-3459.
- [66] Wang, Y.C., Wang, F., Hou, B.C., Wang, E.T., Chen, W.F., Sui, X.H., Chen, W.X., Li, Y., Zhang, Y.B. (2013) Proposal of *Ensifer psoraleae* sp. nov., *Ensifer sesbaniae* sp. nov., *Ensifer morelense* comb. nov. and *Ensifer americanum* comb. nov. Syst. Appl. Microbiol. 36, 467-473.
- [67] Weisburg, W.G., Barns, S.M., Pelletier, D.A., Lane, D.J. (1991) 16S ribosomal DNA amplification for phylogenetic study. J. Bacteriol. 173, 697-703.
- [68] Yamaguchi, N., Ichijo, T., Sakotani, A., Baba, T., Nasu, M. (2012) Global dispersion of bacterial cells on Asian dust. Sci. Rep. 2: 525.
- [69] 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.
- [70] Zakhia, F., Jeder, H., Domergue, O., Willems, A., Cleyet-Marel, J.C., Gillis, M., Dreyfus, B., De Lajudie, P. (2004) Characterisation of wild Legume Nodulating Bacteria (LNB) in the infra-arid zone of Tunisia. Syst. Appl. Microbiol. 27, 380-395.
- [71] Zerhari, K., Aurag, J., Khbaya, B., Kharchaf, D., Filali-Maltouf, A. (2000) Phenotypic characteristics of rhizobia isolates nodulating acacia species in the arid and Saharan regions of Morocco. Lett. Appl. Microbiol. 30, 351-357.
- [72] Zhang, X.P., Harper, R., Karsisto, M., Lindstrom, K. (1991) Diversity of rhizobium bacteria isolated from the root nodules of leguminous trees. Int. J. Syst. Bacteriol. 41, 104-113.

Phylogenetic tree showing relationships between various *Ensifer* and *Rhizobium* strains. The tree is rooted with *Mesorhizobium mediterraneum* LMG 17148T. The strains are color-coded: green for *Ensifer* isolates, yellow for *Ensifer* *americanum*, orange for *Ensifer* *saheli*, red for *Ensifer* *numidicus*, pink for *Ensifer* *kostiensis*, light blue for *Ensifer* *medicae*, dark blue for *Ensifer* *arboris*, light green for *Ensifer* *kummerowiae*, dark green for *Ensifer* *melloti*, and grey for *Ensifer* *americanum* and *Ensifer* *psoraleae*. The tree is supported by bootstrap values (0.1 to 1.0) and posterior probabilities (0.808 to 1.0). The color scale at the bottom indicates the percentage of identity from 0.808 to 1.0.

Strains and their support values (bootstrap/posterior probability):

- LMR009: 1.0/0.89
- LMR017: 1.0/0.89
- Ensifer* sp. JNVU RA9: 1.0/0.89
- LMR020: 1.0/0.89
- LMR019: 1.0/0.89
- LMR018: 1.0/0.89
- LMR016: 1.0/0.89
- LMR015: 1.0/0.89
- LMR013: 1.0/0.89
- LMR012: 1.0/0.89
- LMR010: 1.0/0.89
- LMR008: 1.0/0.89
- LMR007: 1.0/0.89
- LMR006: 1.0/0.89
- LMR004: 1.0/0.89
- LMR003: 1.0/0.89
- LMR001: 1.0/0.89
- LMR005: 1.0/0.89
- Ensifer* sp. JNVU TP6: 0.973/0.896
- Ensifer* sp. JNVU TW10: 0.973/0.896
- E. americanum* CFNEI 156T: 0.906/0.888
- Ensifer* sp. JNVU TP13: 0.938/0.893
- Ensifer* sp. AC01e: 0.925/0.928
- E. sahelii* LMG 7837T: 0.925/0.92
- E. sahelii* LMG 8310: 0.925/0.92
- E. sahelii* LMG 11864: 0.922/0.917
- E. numidicus* ORS 1407T: 0.906/0.898
- E. numidicus* ORS 1444: 0.906/0.898
- E. numidicus* ORS 1410: 0.906/0.898
- E. kostiensis* LMG 19227T: 0.914/0.896
- E. kostiensis* LMG 19225: 0.914/0.896
- E. kostiensis* LMG 1491T: 0.914/0.896
- E. medicae* LMG 19920T: 0.872/0.861
- E. medicae* LMG 19921: 0.872/0.861
- E. medicae* LMG 18864: 0.869/0.864
- E. arboris* LMG 14919T: 0.877/0.882
- E. arboris* LMG 19223: 0.874/0.88
- E. arboris* LMG 14917: 0.877/0.882
- E. kummerowiae* CCBAU 71714T: 0.888/0.877
- E. melloti* LMG 6133T: 0.89/0.872
- E. melloti* LMG 4289: 0.89/0.872
- E. melloti* LMG 6130: 0.89/0.872
- Ensifer* sp. AC14d: 0.906/0.917
- LMR021: 0.89/ID
- Ensifer* sp. AC38d2: 0.866/0.893
- E. xinjiangensis* R16438: 0.877/0.914
- E. xinjiangensis* LMG 17930T: 0.872/0.901
- E. fredii* LMG 6217T: 0.872/0.904
- E. fredii* LMG 8317: 0.872/0.904
- E. fredii* LMG 6218: 0.872/0.904
- E. fredii* LMG 6216: 0.872/0.904
- E. sojae* CCBAU 05684T: 0.904/0.914
- Ensifer* sp. AC52c: 0.898/0.92
- E. psoraleae* CCBAU 65732T: 0.885/0.906
- E. garamanticus* ORS 1400T: 0.885/0.888
- E. garamanticus* STM 354: 0.885/0.888
- E. garamanticus* ORS 1401: 0.885/0.888
- LMR002: 0.89/ID
- Ensifer* sp. AC25e: 0.893/0.896
- E. mexicanus* HAMB1 2910T: 0.89/0.888
- E. terangae* LMG 7834T: 0.882/0.896
- E. terangae* LMG 11859: 0.882/0.896
- E. terangae* LMG 6464: 0.882/0.896
- E. morelense* Lc04T: 0.85/0.896
- Ensifer* sp. LMG 20571: 0.869/0.896
- E. adhaerens* gv. LMG 20216T: 0.877/0.893
- E. adhaerens* gv. LMG 20582: 0.877/0.893
- E. adhaerens* gv. BR 7457: 0.893/0.893
- E. adhaerens* gv. A. LMG 10007: 0.888/0.901
- E. adhaerens* gv. A. R9451: 0.888/0.901
- E. adhaerens* gv. A. LMG 9954: 0.888/0.901
- E. sesbaniae* CCBAU 65729T: 0.88/0.904
- Ensifer* sp. AC47a: 0.885/0.909
- Rhizobium huastense* LMG 18254T: 0.845/0.841
- Rhizobium gallicum* R4387T: 0.864/0.861
- Rhizobium leguminosarum* LMG 14904T: 0.858/0.842
- Rhizobium rhizogenes* LMG 150T: 0.864/0.842
- Rhizobium galegae* LMG 6214T: 0.864/0.856
- Rhizobium giardinii* R4385T: 0.842/0.834
- Agrobacterium rubi* LMG 17935T: 0.813/0.81
- Agrobacterium radiobacter* LMG 140T: 0.845/0.821
- LMR014: 0.872/0.84
- Bradyrhizobium elkanii* LMG 6134T: 0.826/0.821
- Mesorhizobium mediterraneum* LMG 17148T: 0.874/0.848
- Bradyrhizobium japonicum* LMG 6138T: 0.813/0.813

52

replicates) are shown in red next to nodes. The scale bar indicates the estimated number of substitutions per site. (B) The % similarity obtained using 375 bp aligned partial *recA* gene sequences between the Moroccan strains (shown on the top of the matrix as LMR001 (17) for the sequence corresponding to the 17 isolates with identical partial 16S rRNA and *recA* gene sequences, LMR021, LMR002 and LMR014) and the strains used to infer the ML tree following the same order and color codes are shown. The % similarity between compared sequences are indicated in the matrix and highlighted with a color code as indicated at the bottom of the matrix and ID refers to self comparisons. Strains from described *Ensifer* species are also indicated on the right of the matrix accordingly. Accession numbers used are shown in table S1.

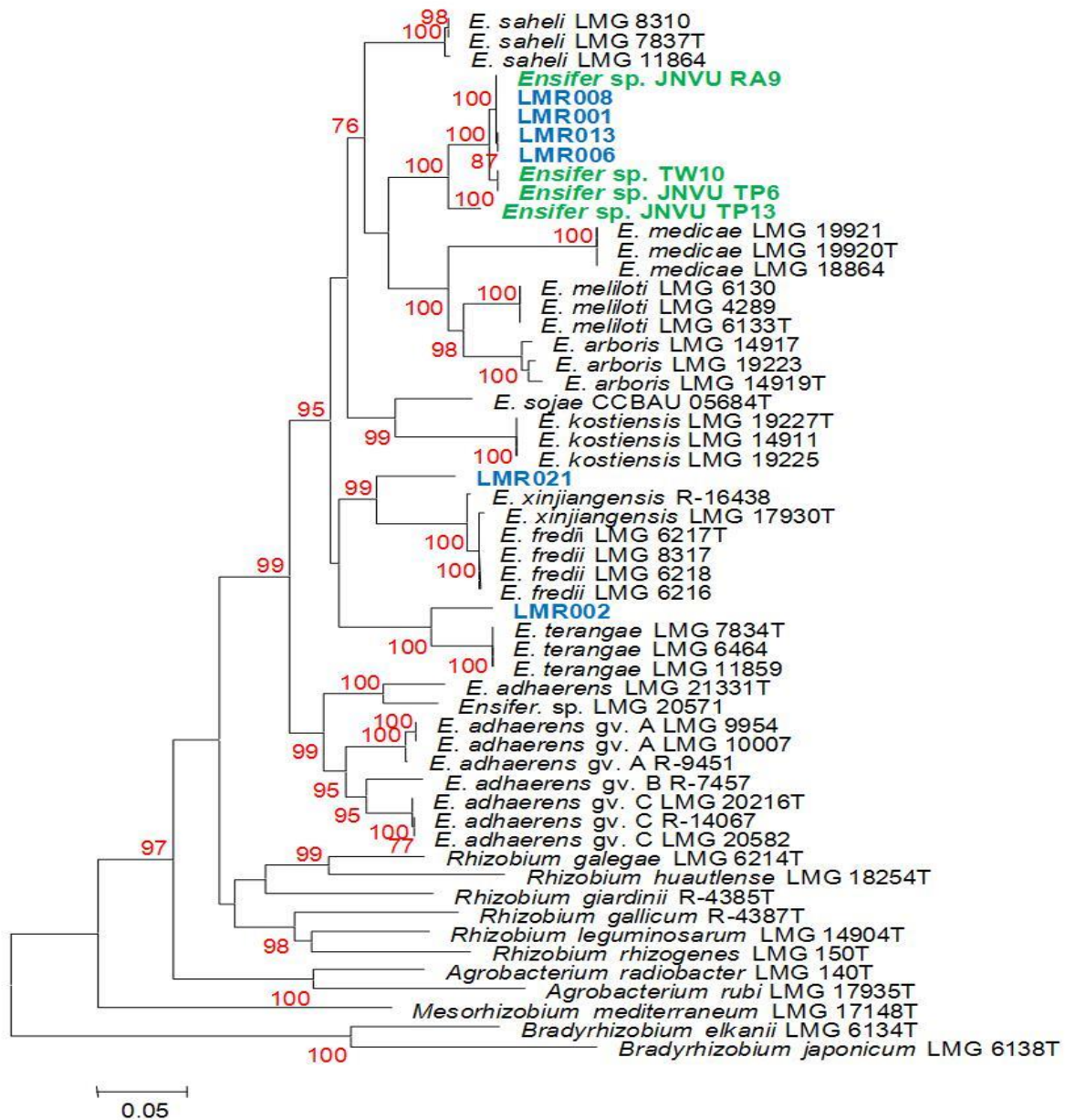


Figure 2. Maximum-Likelihood Phylogenetic tree based on the concatenated *dnaK-glnA-gltA-recA* gene sequences showing evolutionary relationships among selected desert rhizobial strains from Merzouga (Blue) or from Thar (Green) and strains from recognized species for which these 4 housekeeping genes are available (Black). The tree was constructed using the GTR + G + I model. Bootstrap values above 70% (from 1000 replicates) are shown in red next to nodes. The scale bar indicates the estimated number of substitutions per site. Accession numbers used are shown in table S1.

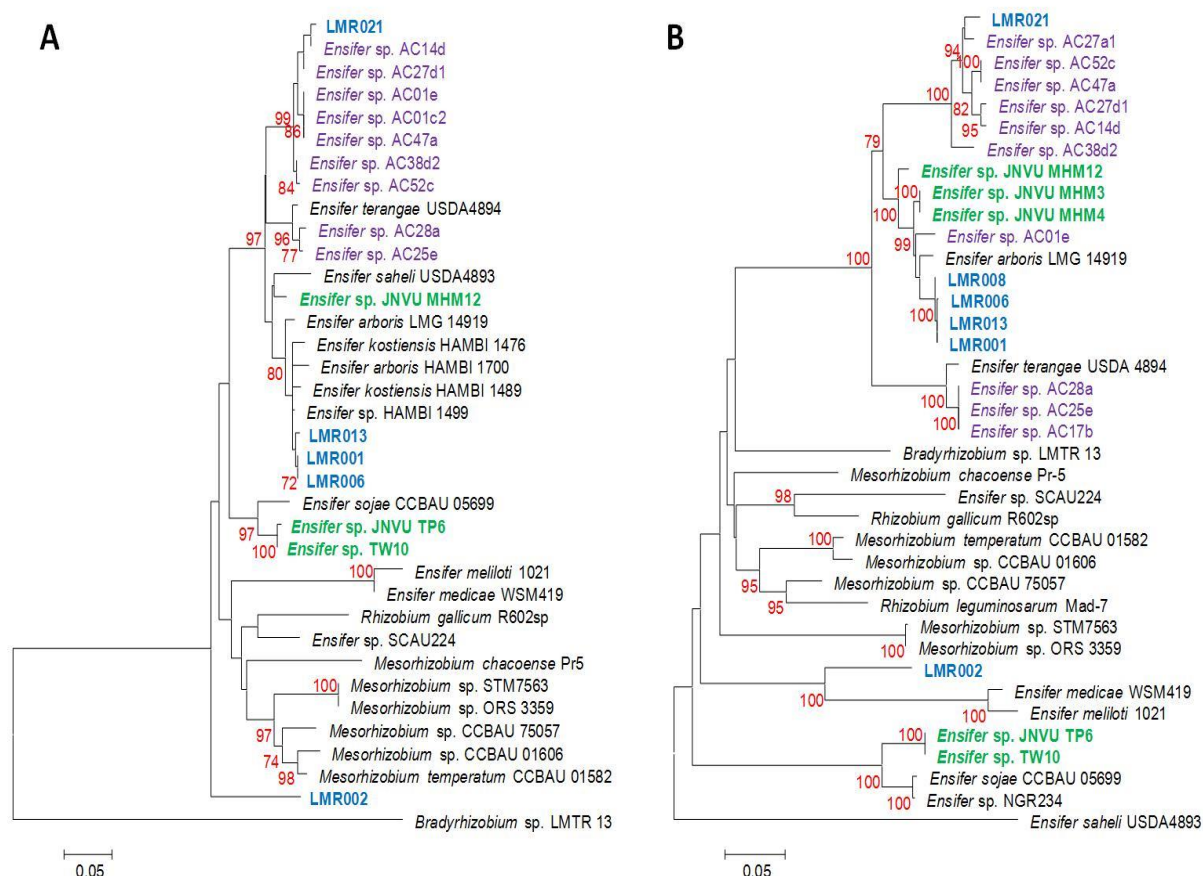


Figure 3. Maximum-Likelihood Phylogenetic trees based on *nifH* (A) and *nodC* (B) partial gene sequences showing evolutionary relationships among selected rhizobial strains recovered from the Merzouga desert (Blue), the Thar desert (Green), South Ethiopia (Purple) and selected strains (Black). The tree was constructed using the models GTR + G + I (A) or GTR + G (B). Bootstrap values above 70% (from 1000 replicates) are shown in red next to nodes. The scale bars indicate the estimated number of substitutions per site. Accession numbers used are shown in table S2.

Table 1. Phenotypic characterization of desert bacteria recovered from Acacia nodules

Strain	Temperature (°C) *	Salt (NaCl mM) **								pH **			Re- nodulation test
		0	100	200	300	400	50 0	60 0	70 0	4	7	9	
LMR00 1	33/37/45/47.5	++ +	++ +	++ +	++	++	+/-	-	-	-	++ +	++ +	+
LMR00 2	33/37/45/47.5/50	++ +	++	++	++	+	+	-	-	-	++ +	++ +	+
LMR00 3	33/37/45	++ +	++	++	++	++	+	+	+/-	-	++ +	++ +	+
LMR00 4	33/37/45/47.5/50	++ +	++	++	++	+	+/-	-	-	-	++ +	++ +	+
LMR00 5	33/37/45/47.5/50	++ +	++ +	++ +	++	+	+/-	-	-	-	++ +	++ +	+
LMR00 6	33/37/45/47.5/50	++ +	++	+	+	+	-	-	-	-	++ +	++ +	+
LMR00 7	33/37/45/47.5/50	++ +	++	++	++	+	+	-	-	-	++ +	++ +	+
LMR00 8	33/37/45/47.5/50	++ +	++ +	++ +	++	++	++	-	-	-	++ +	++ +	+
LMR00 9	33/37/45/47.5/50	++ +	++	+	+	+	+	-	-	-	++ +	++ +	+
LMR01 0	33/37/45/47.5/50	++ +	++ +	++ +	++	+/-	-	-	-	-	++ +	++ +	+
LMR01 2	33/37/45/47.5	++ +	++	++	++	+	+	-	-	-	++ +	++ +	+
LMR01 3	33/37/45/47.5	++ +	++ +	++ +	++	++	++	-	-	-	++ +	++ +	+
LMR01 4	33/37	++ +	++	+	+	+	+	-	-	-	++ +	++ +	-
LMR01 5	33/37/45/47.5/50	++ +	++ +	++ +	++	+ /-	+/-	-	-	-	++ +	++ +	-
LMR01 6	33/37	++ +	++	+	+	+	-	-	-	-	++ +	++ +	+
LMR01 7	33/37/45/47.5	++ +	++	++	++	+	+	-	-	-	++ +	++ +	+
LMR01 8	33/37/45	++ +	++ +	++ +	++	++	++	-	-	-	++ +	++ +	+
LMR01 9	33/37/45	++ +	++	+	+	+	+	-	-	-	++ +	++ +	-
LMR02 0	33/37/45/47.5/50	++ +	++ +	++ +	++	+	-	-	-	-	++ +	++ +	+
LMR02 1	33/37	++ +	++	+	+	+	-	-	-	-	++ +	++ +	+

* Growth was tested at 33 / 37 / 45 / 47.5 and 50°C

** The relative growth of bacteria in YEM medium complemented with salt at concentrations ranging from 100-700 mM is indicated by the number of "+" ; "+/-" indicates low growth and no-growth is indicated by "-".

Supplementary data

Title: Recovery of symbiotic nitrogen fixing acacia rhizobia from Merzouga Desert sand dunes in South East Morocco - Identification of a probable new species of *Ensifer* adapted to stressed environments

Authors: Ilham Sakrouhi, Meryem Belfquih, Laïla Sbabou, Patricia Moulin, Gilles Bena, Abdelkarim Filali-Maltouf & Antoine Le Quéré

Supplementary figures

Figure S1. Estimation of desert rhizobia diversity based on partial 16S rRNA sequence analysis

Figure S2. REP-PCR profiles of 17 Moroccan and 2 Indian desert rhizobia

Figure S3. ML tree based on *dnaK* partial sequences

Figure S4. ML tree based on *glnA* partial sequences

Figure S5. ML tree based on *gltA* partial sequences

Supplementary tables

Table S1. Strains and accessions used for core gene phylogenies

Table S2. Strains and accessions used for symbiotic gene phylogenies

Table S3. Description of Merzouga desert sites sampled

Table S4. Matrix showing estimates of evolutionary divergence between concatenated *recA-gltA-dnaK-glnA* sequences

Supplementary Figure S1

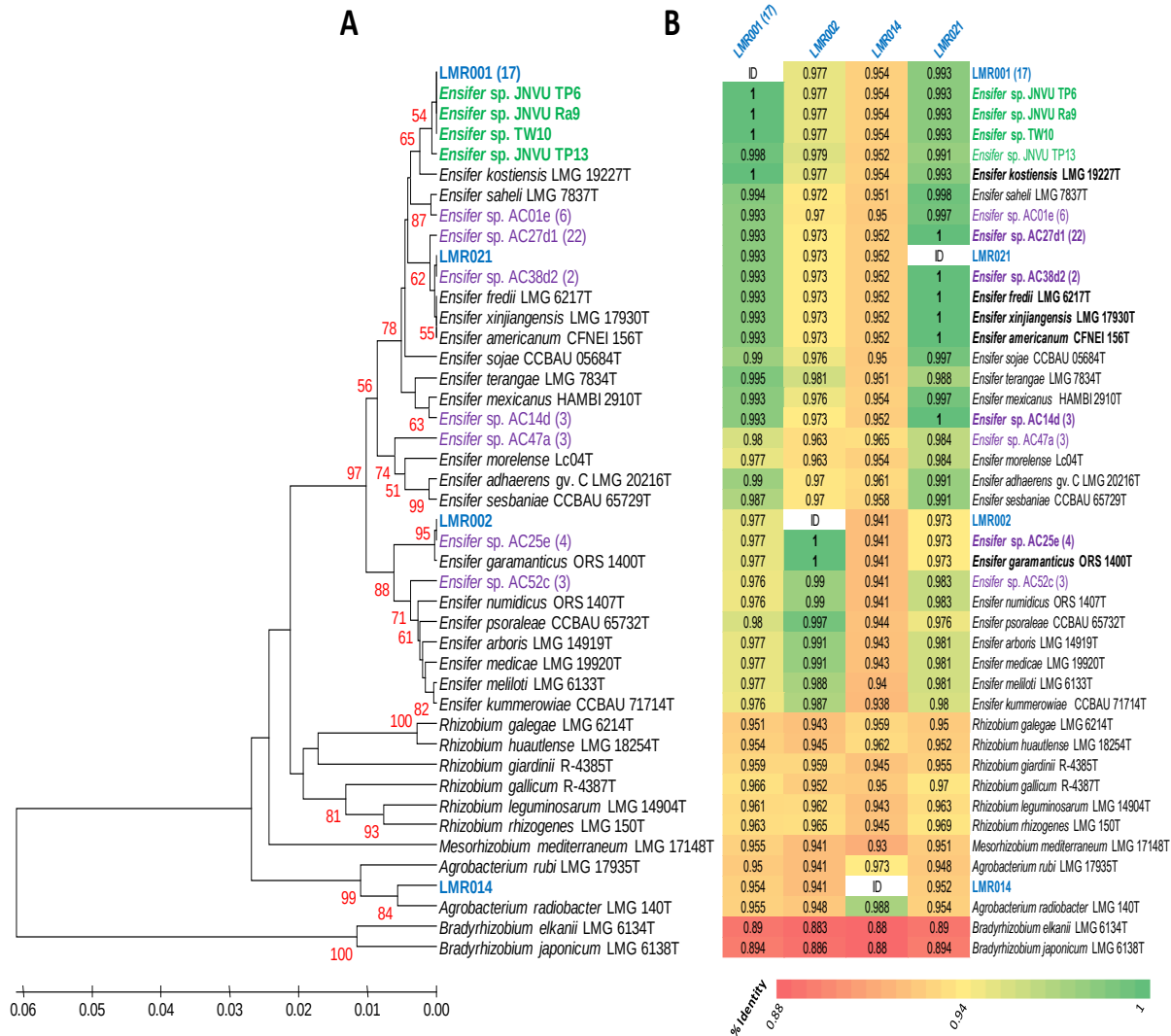


Figure S1. Estimation of desert rhizobia diversity based on partial 16S rRNA sequence analysis. **(A)** Phylogenetic tree based on 16S rDNA partial gene sequences of strains isolated from the Merzouga desert (blue), the Thar Desert (green) and south Ethiopia (purple) and selected reference strains (Black). The evolutionary history was inferred using the UPGMA method. Evolutionary distances were computed using the kimura 2-parameter method and are shown by branch length as the number of base substitutions per site indicated by the scale (bottom of the tree). Numbers in brackets refer to number of isolates showing identical partial 16S rDNA sequences. Results from bootstrap test (with 1000 replicates) are shown as % in red next to nodes when above 50%. **(B)** The % similarity obtained using 720 bp aligned partial 16S rDNA sequences between the Moroccan strains (shown on the top of the matrix) and the strains used to infer the UPGMA tree (indicated on the right of the matrix and ordered as in the tree) using the same color code are shown. The % of identical residues between compared sequences are indicated in the matrix and highlighted with a color code as indicated on the bottom of the matrix and ID refers to self comparisons.

Supplementary Figure S2

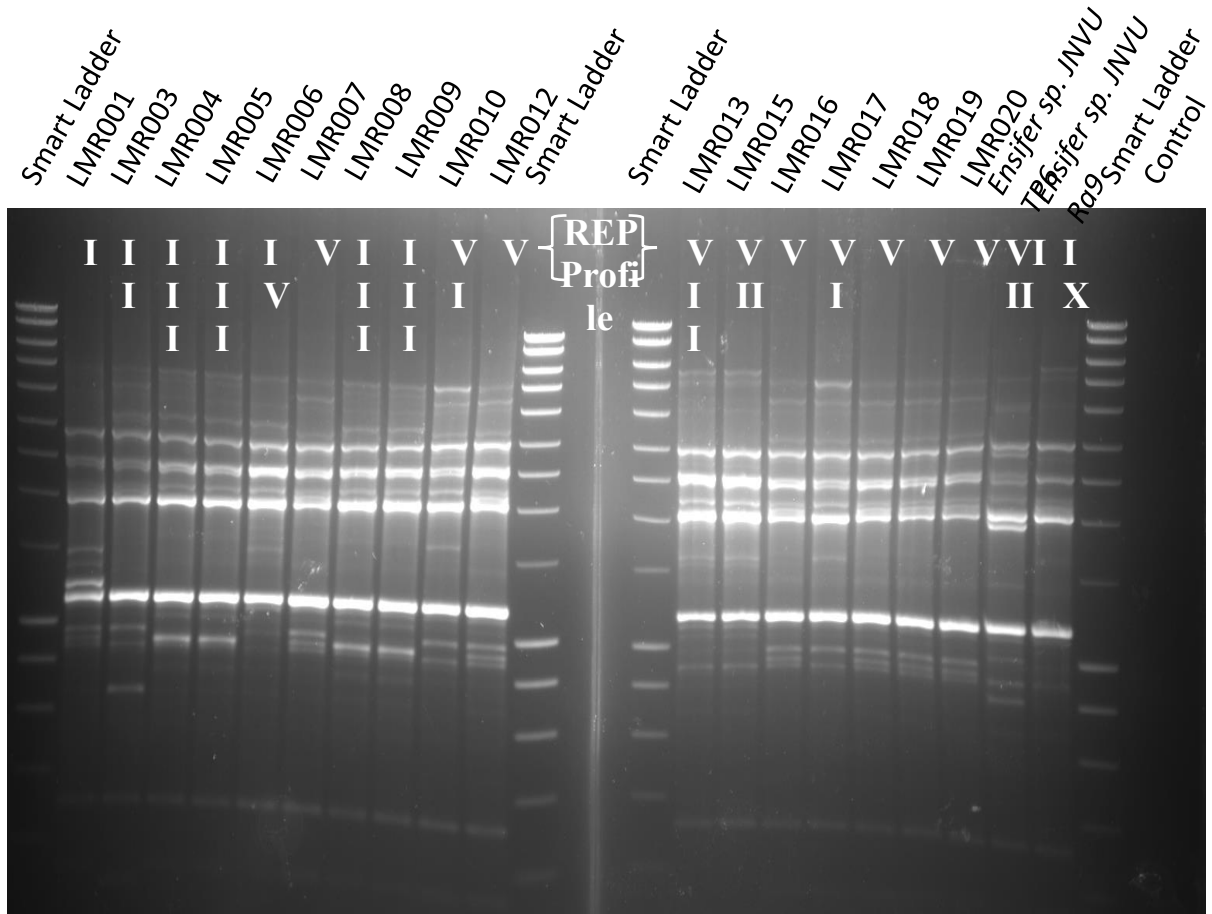


Figure S2. REP-PCR profiles of 17 Moroccan and 2 Indian desert rhizobia showing identical partial 16s rDNA sequences and more than 97% similarity on partial recA sequences. The names of the strains are indicated above the photos and their respective REP profiles obtained using REP 1R-1 / REP 2-1 primers are indicated on the photos as white Latin numbers. Smart Ladder (Eurogentec) was used and is indicated on top of the gel on the corresponding lanes. Finally the control (last lane on the right of the gels) resulted in no amplification.

Supplementary Figure S3

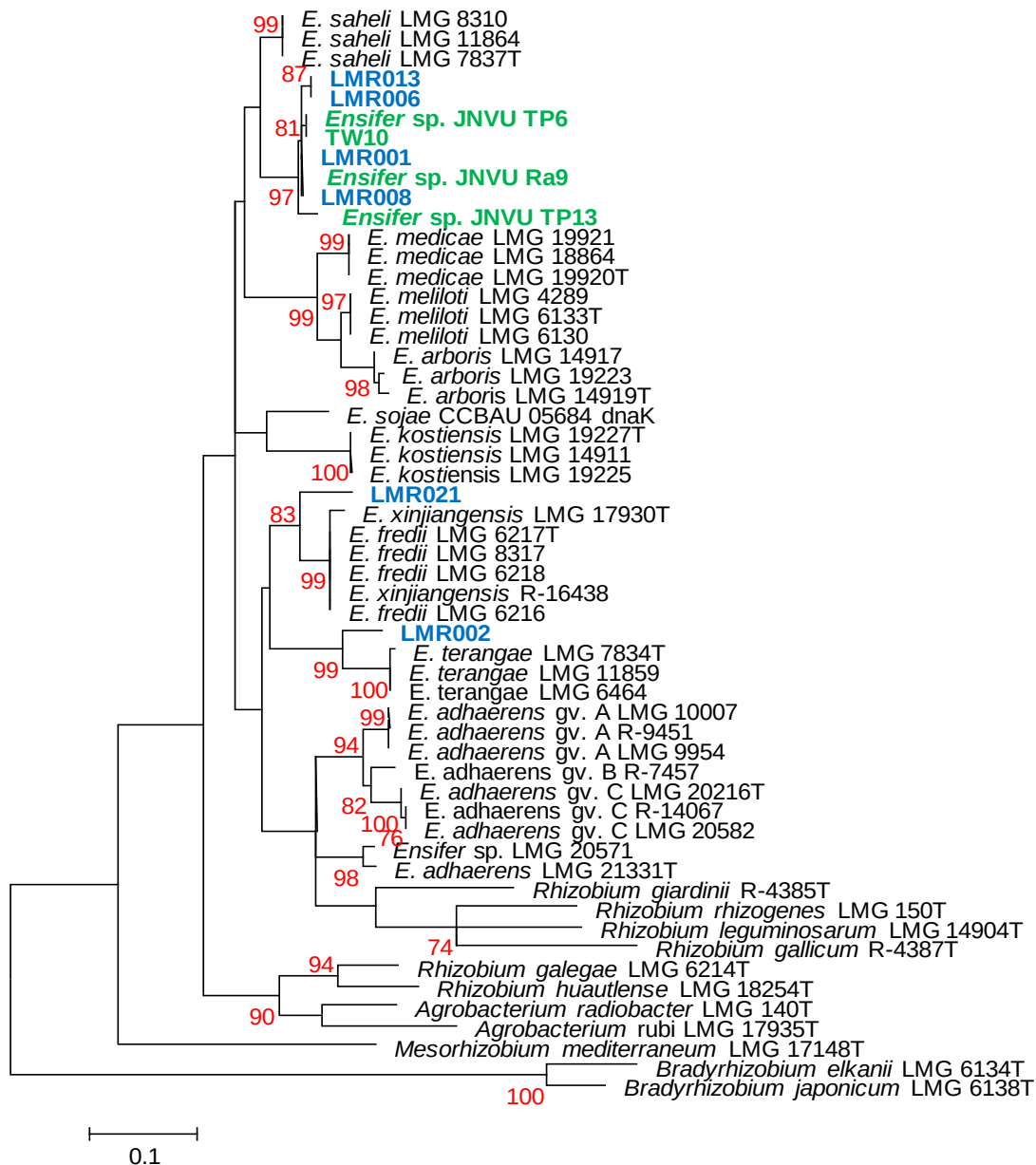


Figure S3. ML tree based on *dnaK* partial sequences. Maximum-Likelihood Phylogenetic tree based on *dnaK* partial gene sequences showing evolutionary relationships among selected desert rhizobial strains from Merzouga (Blue) or from Thar (Green) and strains from recognized species (Black). The tree was constructed using the GTR + G model. Bootstrap values above 70% (from 1000 replicates) are shown in red next to nodes. The scale bar indicates the estimated number of substitutions per site.

Supplementary Figure S4

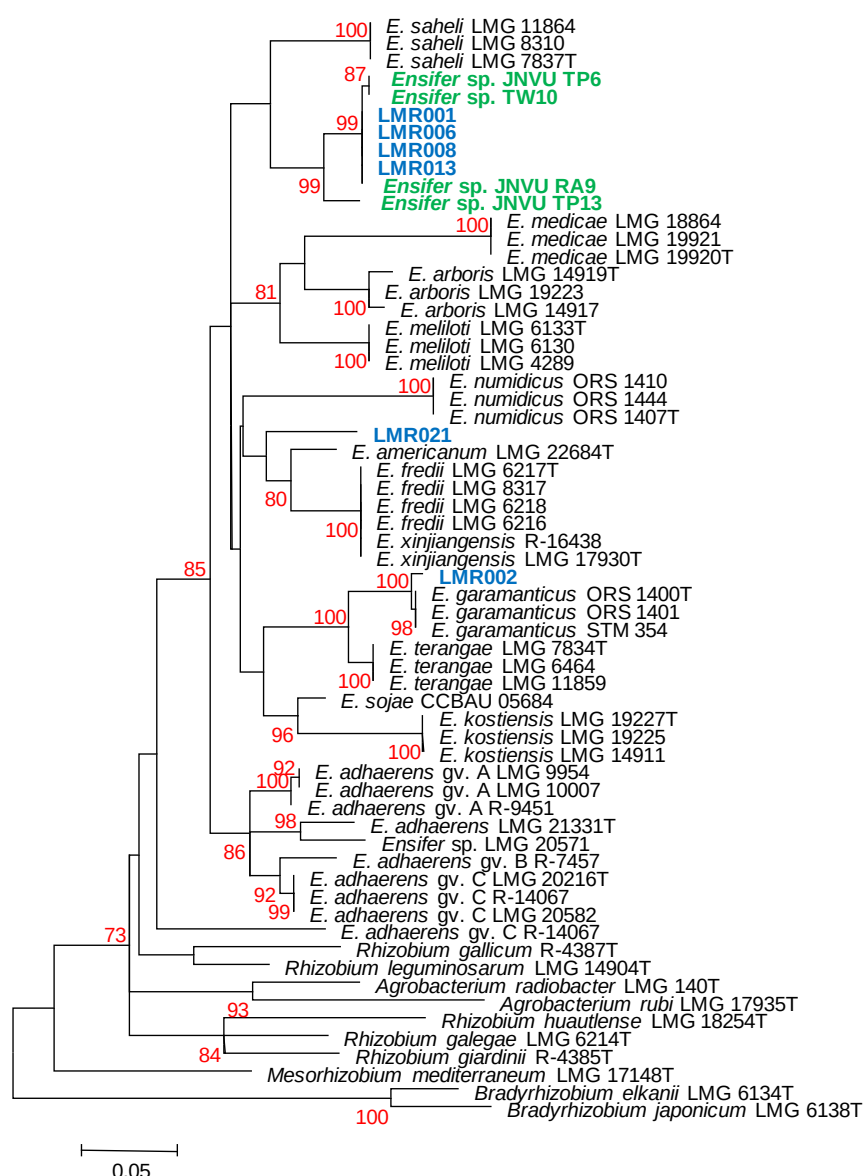


Figure S4. ML tree based on *glnA* partial sequences. Maximum-Likelihood Phylogenetic tree based on *glnA* partial gene sequences showing evolutionary relationships among selected desert rhizobial strains from Merzouga (Blue) or from Thar (Green) and strains from recognized species (Black). The tree was constructed using the GTR + G + I model. Bootstrap values above 70% (from 1000 replicates) are shown in red next to nodes. The scale bar indicates the estimated number of substitutions per site.

Supplementary Figure S5

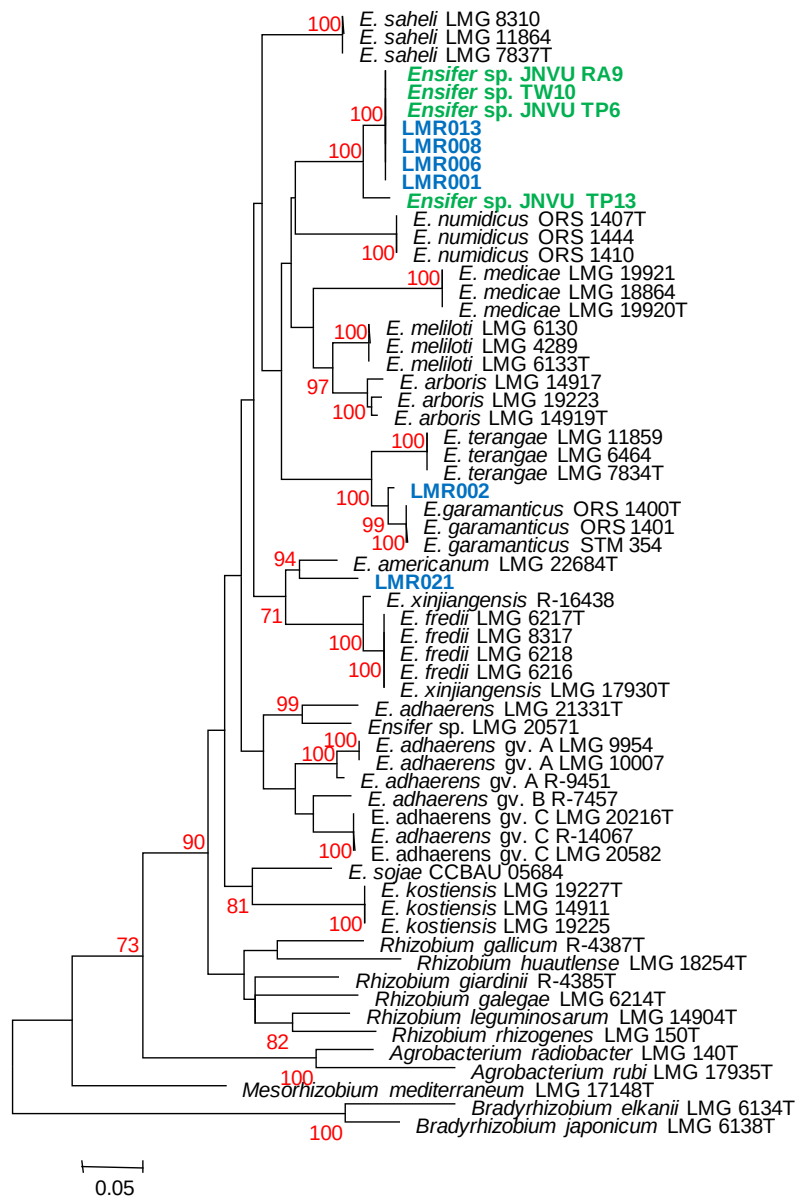


Figure S5. ML tree based on *gltA* partial sequences. Maximum-Likelihood Phylogenetic tree based on *gltA* partial gene sequences showing evolutionary relationships among selected desert rhizobial strains from Merzouga (Blue) or from Thar (Green) and strains from recognized species (Black). The tree was constructed using the GTR + G + I model. Bootstrap values above 70% (from 1000 replicates) are shown in red next to nodes. The scale bar indicates the estimated number of substitutions per site.

Supplementary table S1. Strains and accessions used for core gene phylogenies

Reference strains	Accession numbers*				Origin	References
	rRNA	recA	glnA	dna K	gltA	
<i>Ensifer adhaerens</i> gv. A						
UMG 9954 (=BR819)	AJ505599	AM182970	AM182112	AM181647	<i>Leucaena leucocephala</i> , Brazil	Marteens et al., 2007
UMG 10007 (=BR606)	AJ505600	AM182969	AM182111	AM181646	<i>Phtheobolium dulce</i> , Brazil	Marteens et al., 2007
R-9451 (=HAMB1 1631)	AJ505598	AM182968	AM182110	AM181645	<i>Sesbania grandiflora</i> , Sri Lanka	Marteens et al., 2007
<i>Ensifer adhaerens</i> gv. B R-7457 (=SD19)	AJ505597	AM182967	AM182109	AM181644	<i>Medicago sativa</i> , Spain	Marteens et al., 2007
<i>Ensifer adhaerens</i> gv. C						
UMG 20216 [†] (=ATCC 33212 [†])	AJ505595	AM182966	AM182108	AM181643	Soil, central PA, USA	Marteens et al., 2007
R-14067 (=ATCC 33499)	AM182150	AM182964	AM182106	AM181641	Soil, central PA, USA	Marteens et al., 2007
UMG 20582 (=R-6387)	AJ505596	AM182965	AM182107	AM181642	Agricultural soil, Pittem, Belgium	Marteens et al., 2007
<i>Ensifer morelense</i> L04 [†] (=UMG 21331 [†])	AM181737	AM182971	AM182113	AM181648	<i>Leucaena leucocephala</i> , Mexico	Marteens et al., 2007
<i>Ensifer</i> sp. UMG 20571 (=R-4955)	AJ20776	AM182151	AM182114	AM181649	Agricultural soil, Pittem, Belgium	Marteens et al., 2007
<i>Ensifer carbonis</i>						
UMG 14619 [†] (=HAMB1 1552 [†])	AM181744	AM182130	AM182944	AM181621	<i>Prosopis chilensis</i> , Kosti, Sudan	Marteens et al., 2007
UMG 14617 (=HAMB1 1396)	AM182131	AM182945	AM182087	AM181622	<i>Prosopis chilensis</i> , Kenya	Marteens et al., 2007
UMG 19223 (=HAMB1 1704)	AM182129	AM182943	AM182085	AM181620	<i>Acacia senegal</i> , Khartoum, Sudan	Marteens et al., 2007
<i>Ensifer fredii</i>						
UMG 6217 [†] (=USDA 205 [†])	X6731	AM182145	AM182959	AM182101	<i>Glycine max</i> , Hovnan, China	Marteens et al., 2007
UMG 8317 (=USDA 191)		AM182144	AM182958	AM182100	Soil, Shanghai, China	Marteens et al., 2007
UMG 6218 (=USDA 206)		AM182146	AM182960	AM182102	<i>Glycine max</i> , Hovnan, China	Marteens et al., 2007
UMG 6216 (=UMG 8317)		AM182147	AM182961	AM182103	Soil, Shanghai, China	Marteens et al., 2007
<i>Ensifer kostiensis</i>						
UMG 19227 [†] (=HAMB1 1485 [†])	AM181748	AM182142	AM182956	AM182098	<i>Acacia senegal</i> , Kosti, Sudan	Marteens et al., 2007
UMG 14911 (=HAMB1 1502)		AM182143	AM182957	AM181634	<i>Acacia senegal</i> , Tendeti, Sudan	Marteens et al., 2007
UMG 19225 (=HAMB1 1484)		AM182141	AM182955	AM182097	<i>Prosopis chilensis</i> , Kosti, Sudan	Marteens et al., 2007
<i>Ensifer mediceae</i>						
UMG 19920 [†] (=EA 321 [†])	L39882	AM182135	AM182949	AM182091	<i>Medicago truncatula</i> , Auile, France	Marteens et al., 2007
UMG 19921 (=R-481)		AM182136	AM182950	AM182092	<i>Medicago truncatula</i>	Marteens et al., 2007
UMG 18864 (=HAMB1 1809)		AM182137	AM182951	AM182093	<i>Medicago truncatula</i> , Syria	Marteens et al., 2007
<i>Ensifer mellolii</i>						
UMG 6133 [†] (=NP 4027 [†])	X6722	AM182133	AM182947	AM182089	<i>Medicago sativa</i> , VA, USA	Marteens et al., 2007
UMG 4289 (=INRA 2011)		AM182134	AM182948	AM182090	<i>Medicago sativa</i>	Marteens et al., 2007
UMG 6130 (=NZP 4009)		AM182132	AM182946	AM182088	<i>Medicago sativa</i> , Australia	Marteens et al., 2007
<i>Ensifer saheli</i>						
UMG 7837 (=ORS 606 [†])	X68390	AM182138	AM182952	AM182094	<i>Sesbania cannabina</i> , Dakar, Senegal	Marteens et al., 2007
UMG 11864 (=ORS 600)		AM182140	AM182954	AM182096	<i>Sesbania pachycarpa</i> , Senegal	Marteens et al., 2007
UMG 8310 (=ORS 611)		AM182139	AM182953	AM182095	<i>Sesbania grandiflora</i> , Dakar, Senegal	Marteens et al., 2007
<i>Ensifer terrae</i>						
UMG 7834 [†] (=ORS 1009 [†])	X68388	AM182153	AM182974	AM182116	<i>Acacia baobab</i> , Dakar, Senegal	Marteens et al., 2007
UMG 11859 (=ORS 52)		AM182154	AM182975	AM182117	<i>Sesbania rostrata</i> , Senegal	Marteens et al., 2007
UMG 6464 (=ORS 51)		AM182152	AM182973	AM182115	<i>Sesbania rostrata</i> , Dakar, Senegal	Marteens et al., 2007
<i>Ensifer xingjiangensis</i>						
UMG 17930 [†] (=CCBAU 110 [†])	AM181732	AM182148	AM182962	AM182104	<i>Glycine max</i> , Xinjiang, China	Marteens et al., 2007
R-16438 (=CCBAU 83834)		AM182149	AM182963	AM182105	—	Marteens et al., 2007

...

Supplementary table S1. Strains and accessions used for core gene phylogenies (continued)

Reference strains	Accession numbers*				Origin	References
	rRNA	rec A	gln A	dhx K	glt A	
<i>Ensifer numidicus</i> ORS 1407 [†] (=LMG 24690 [†]) ORS 1444 (=LMG 24691) ORS 1410 (=LMG 24736)	AY500254	AM946576 AM946578 AM946577	AM946560 AM946562 AM946561		AM946594 AM946596 AM946595	Zakaria et al., 2004 Zakaria et al., 2004 Zakaria et al., 2004
<i>Ensifer garmaniticus</i> ORS 1400 [†] (=LMG 24692 [†]) ORS 1401 (=LMG 24693) STM 354 (=LMG 24694)	AY500255	AM946573 AM946574 AM946575	AM946557 AM946558 AM946559		AM946591 AM946592 AM946593	Zakaria et al., 2004 Zakaria et al., 2004 Zakaria et al., 2004
<i>Ensifer mexicanus</i> HANBI 2910 [†]	NR_115768	DQ411943				Lloret et al., 2007
<i>Ensifer americanus</i> CFN E1156 [†] (=LMG 22684 [†])	AF506513					Toledo et al., 2003
<i>Ensifer sspae</i> CGBAU 05684 [†]	NR_117540	DQ411951	AM946556	AQOT01000000	AM946590	Li et al., 2011
<i>Ensifer kummerowiae</i> CGBAU 71714 [†]	NR_114614	AQOT01000000	AQOT01000000	AQOT01000000	AQOT01000000	Wei et al., 2002
<i>Ensifer psoraleae</i> CGBAU 65732 [†] (=LMG 26835 [†])	EU618039	DQ411945				Wang et al., 2013
<i>Ensifer sesbaniae</i> CGBAU 65729 [†] (=LMG 26833 [†])	JF834143	EU622106 JX028823				Wang et al., 2013
<i>Mesorhizobium loti</i> LMG 6125 [†] (=NZP 2213 [†])	X67229	AM182156		AM182977	AM181654	Martens et al., 2007
<i>Mesorhizobium mediterraneum</i> LMG 17148 [†] (=UPM-Co36 [†])	AM181745	AM182157		AM182978	AM181655	Martens et al., 2007
<i>Bradyrhizobium elkanii</i> LMG 6134 [†] (=USDA 76 [†])	AF829442	AM182155		AM182976	AM181653	Martens et al., 2007
<i>Bradyrhizobium japonicum</i> LMG 6138 [†] (=NZP 5549 [†])	X66024	AM182158		AM182979	AM181656	Martens et al., 2007
<i>Rhizobium galegae</i> LMG 6214 [†] (=HAMBI 540 [†])	X67226	AM182127		AM182941	AM181618	Martens et al., 2007
<i>Rhizobium galegae</i> R-4387 [†] (=6023p [†])	AM181756	AM182124		AM182938	AM181615	Martens et al., 2007
<i>Rhizobium galegae</i> R-4385 [†] (=H152 [†])	AM181755	AM182123		AM182937	AM181614	Martens et al., 2007
<i>Rhizobium huatdense</i> LMG 18254 [†] (=502 [†])	AF058552	AM182128		AM182942	AM181619	Martens et al., 2007
<i>Rhizobium leguminosarum</i> LMG 14904 [†] (=ATCC 10004 [†])	AM181757	AM182125		AM182939	AM181616	Martens et al., 2007
<i>Rhizobium rhizogenes</i> LMG 150 [†] (=ATCC 11325 [†])	AY945955	AM182126		AM182940	AM181617	Martens et al., 2007
<i>Agrobacterium radiobacter</i> LMG 140 [†] (=ATCC 19398 [†])	AM181758	AM182121		AM182935	AM181612	Martens et al., 2007
<i>Agrobacterium rubi</i> LMG 17935 [†] (=ATCC 13335 [†])	AM181759	AM182122		AM182936	AM181613	Martens et al., 2007
<i>Argemone uniflorum</i> , Tunisia						
<i>Argemone uniflorum</i> , Tunisia						
<i>Lotus creticus</i> , Tunisia						
<i>Argemone uniflorum</i> , Tunisia						
<i>Argemone uniflorum</i> , Tunisia						
<i>Medicago sativa</i> , Tunisia						
<i>Acacia angustissima</i> , Mexico						
<i>Acacia sp.</i> , Mexico						
<i>Glycine max</i> , Hebei, China						
<i>Kummerowia stipulacea</i> , Loess Plateau, China						
<i>Psoralea carthagenica</i> , Dehong, China						
<i>Sesbania arnabiana</i> , Dehong, China						
<i>Lotus corniculatus</i> var. <i>terrestris</i> , New Zealand						
<i>Cicer arietinum</i> L, Spain						
<i>Glycine max</i> , USA						
<i>Glycine max</i> , Japan						
<i>Galega orientalis</i> , Finland						
<i>Phaseolus vulgaris</i> , France						
<i>Phaseolus vulgaris</i> , France						
<i>Sesbania herbacea</i> , Sierra de Huautla, Mexico						
<i>Pisum sativum</i> , IL, USA						
<i>Rubus ursinus</i> var. <i>loganobaccus</i> , USA						

Supplementary table S1. Strains and accessions used for core gene phylogenies (continued)

Other strains	Accession numbers*					Origin	References
	rRNA	recA	gln A	dna K	glt A		
Strains isolated from desert sand dunes in Morocco	KP993276	KP993231	KP993246	KP993240	KP993255	<i>Acacia gummifera</i> , Merzouga Desert, Morocco	This study
	KP993277	KP993232	KP993247	KP993241	KP993256	<i>Acacia tortilis</i> ssp. <i>Raddiana</i> , Merzouga Desert, Morocco	This study
	KP993278	KP993233				<i>Acacia tortilis</i> ssp. <i>Raddiana</i> , Merzouga Desert, Morocco	This study
	KP993279	KP993234				<i>Acacia tortilis</i> ssp. <i>Raddiana</i> , Merzouga Desert, Morocco	This study
	KP993280	KP993235				<i>Acacia tortilis</i> ssp. <i>Raddiana</i> , Merzouga Desert, Morocco	This study
	KP993281	KP993236				<i>Acacia tortilis</i> ssp. <i>Raddiana</i> , Merzouga Desert, Morocco	This study
	KP993282	KP993237				<i>Acacia tortilis</i> ssp. <i>Raddiana</i> , Merzouga Desert, Morocco	This study
	KP993283	KP993238				<i>Acacia tortilis</i> ssp. <i>Raddiana</i> , Merzouga Desert, Morocco	This study
	KP993284	KP993239				<i>Acacia tortilis</i> ssp. <i>Raddiana</i> , Merzouga Desert, Morocco	This study
	KP993285	KP993240				<i>Acacia tortilis</i> ssp. <i>Raddiana</i> , Merzouga Desert, Morocco	This study
	KP993286	KP993241				<i>Acacia tortilis</i> ssp. <i>Raddiana</i> , Merzouga Desert, Morocco	This study
	KP993287	KP993242				<i>Acacia tortilis</i> ssp. <i>Raddiana</i> , Merzouga Desert, Morocco	This study
	KP993288	KP993243				<i>Acacia tortilis</i> ssp. <i>Raddiana</i> , Merzouga Desert, Morocco	This study
	KP993289	KP993244				<i>Acacia tortilis</i> ssp. <i>Raddiana</i> , Merzouga Desert, Morocco	This study
	KP993290	KP993245				<i>Acacia tortilis</i> ssp. <i>Raddiana</i> , Merzouga Desert, Morocco	This study
	KP993291	KP993246				<i>Acacia tortilis</i> ssp. <i>Raddiana</i> , Merzouga Desert, Morocco	This study
	KP993292	KP993247				<i>Acacia tortilis</i> ssp. <i>Raddiana</i> , Merzouga Desert, Morocco	This study
	KP993293	KP993248				<i>Acacia tortilis</i> ssp. <i>Raddiana</i> , Merzouga Desert, Morocco	This study
	KP993294	KP993249				<i>Acacia tortilis</i> ssp. <i>Raddiana</i> , Merzouga Desert, Morocco	This study
	KP993295	KP993250				<i>Acacia tortilis</i> ssp. <i>Raddiana</i> , Merzouga Desert, Morocco	This study
	KP993296	KP993251				<i>Acacia tortilis</i> ssp. <i>Raddiana</i> , Merzouga Desert, Morocco	This study
	KP993297	KP993252				<i>Acacia tortilis</i> ssp. <i>Raddiana</i> , Merzouga Desert, Morocco	This study
	KP993298	KP993253				<i>Acacia tortilis</i> ssp. <i>Raddiana</i> , Merzouga Desert, Morocco	This study
	KP993299	KP993254				<i>Acacia tortilis</i> ssp. <i>Raddiana</i> , Merzouga Desert, Morocco	This study
	KP993300	KP993255				<i>Acacia tortilis</i> ssp. <i>Raddiana</i> , Merzouga Desert, Morocco	This study
	KP993301	KP993256				<i>Acacia tortilis</i> ssp. <i>Raddiana</i> , Merzouga Desert, Morocco	This study
	KP993302	KP993257				<i>Acacia tortilis</i> ssp. <i>Raddiana</i> , Merzouga Desert, Morocco	This study
	KP993303	KP993258				<i>Acacia tortilis</i> ssp. <i>Raddiana</i> , Merzouga Desert, Morocco	This study
	KP993304	KP993259				<i>Acacia tortilis</i> ssp. <i>Raddiana</i> , Merzouga Desert, Morocco	This study
	KP993305	KP993260				<i>Acacia tortilis</i> ssp. <i>Raddiana</i> , Merzouga Desert, Morocco	This study
Strains isolated from wild legumes in the Indian Thar Desert	KP993306	KP993261				<i>Acacia tortilis</i> ssp. <i>Raddiana</i> , Merzouga Desert, Morocco	This study
	KP993307	KP993262				<i>Acacia tortilis</i> ssp. <i>Raddiana</i> , Merzouga Desert, Morocco	This study
	KP993308	KP993263				<i>Acacia tortilis</i> ssp. <i>Raddiana</i> , Merzouga Desert, Morocco	This study
	KP993309	KP993264				<i>Acacia tortilis</i> ssp. <i>Raddiana</i> , Merzouga Desert, Morocco	This study
	KP993310	KP993265				<i>Acacia tortilis</i> ssp. <i>Raddiana</i> , Merzouga Desert, Morocco	This study
	KP993311	KP993266				<i>Acacia tortilis</i> ssp. <i>Raddiana</i> , Merzouga Desert, Morocco	This study
	KP993312	KP993267				<i>Acacia tortilis</i> ssp. <i>Raddiana</i> , Merzouga Desert, Morocco	This study
	KP993313	KP993268				<i>Acacia tortilis</i> ssp. <i>Raddiana</i> , Merzouga Desert, Morocco	This study
	KP993314	KP993269				<i>Acacia tortilis</i> ssp. <i>Raddiana</i> , Merzouga Desert, Morocco	This study
	KP993315	KP993270				<i>Acacia tortilis</i> ssp. <i>Raddiana</i> , Merzouga Desert, Morocco	This study
Strains isolated from <i>Acacia</i> in Southern Ethiopia	KP993316	KP993271				<i>Acacia tortilis</i> ssp. <i>Raddiana</i> , Merzouga Desert, Morocco	This study
	KP993317	KP993272				<i>Acacia tortilis</i> ssp. <i>Raddiana</i> , Merzouga Desert, Morocco	This study
	KP993318	KP993273				<i>Acacia tortilis</i> ssp. <i>Raddiana</i> , Merzouga Desert, Morocco	This study
	KP993319	KP993274				<i>Acacia tortilis</i> ssp. <i>Raddiana</i> , Merzouga Desert, Morocco	This study
	KP993320	KP993275				<i>Acacia tortilis</i> ssp. <i>Raddiana</i> , Merzouga Desert, Morocco	This study
	KP993321	KP993276				<i>Acacia tortilis</i> ssp. <i>Raddiana</i> , Merzouga Desert, Morocco	This study
	KP993322	KP993277				<i>Acacia tortilis</i> ssp. <i>Raddiana</i> , Merzouga Desert, Morocco	This study
	KP993323	KP993278				<i>Acacia tortilis</i> ssp. <i>Raddiana</i> , Merzouga Desert, Morocco	This study
	KP993324	KP993279				<i>Acacia tortilis</i> ssp. <i>Raddiana</i> , Merzouga Desert, Morocco	This study
	KP993325	KP993280				<i>Acacia tortilis</i> ssp. <i>Raddiana</i> , Merzouga Desert, Morocco	This study

*Sequences in bold correspond to new sequence data
§Partial 16S sequencing was first used for genotyping

References cited in Supplementary table S1

Degefu et al., , (2012) Int. J. Syst. Evol. Microbiol. 62, 2286-2295.
Gehlot et al., , (2012) Plant Soil. 357, 227-243.
Li et al., , (2011) Int. J. Syst. Evol. Microbiol. 61, 1981-1988.
Lloret et al., , (2007) Syst. Appl. Microbiol. 30, 280-90.
Martens et al., , (2007) Int. J. Syst. Evol. Microbiol. 57, 489-503.
Toledo et al., , (2003) Syst. Appl. Microbiol. 26, 54-64.
Wei et al., , (2002) Int. J. Syst. Evol. Microbiol. 52, 2231-2239.
Zakhia et al., , (2004) Syst. Appl. Microbiol. 27, 380-395.

Supplementary table S2. Strains and accessions used for symbiotic gene phylogenies

Reference strains	nifH	nodC	Origin	References
<i>Ensifer arboris</i> LMG 14919	NZ_ATY801000000*	NZ_ATY801000000*	<i>Prosopis chilensis</i> , Sudan	Reeve et al., 2013
<i>Bradyrhizobium</i> sp. LMTR 13	KF896161	KF896159	<i>Phaseolus lunatus</i> , Peru	Duran et al., 2014
<i>Ensifer medicae</i> WSM419	CP000740*	CP000740*	<i>Medicago truncatula</i>	Reeve et al., 2010
<i>Ensifer meliloti</i> 1021	AE006469*	AE006469*	<i>Medicago sativa</i>	Barnett et al., 2001
<i>Ensifer saheli</i> USDA4893	chromosome USDA4893v1_U4893v1*	chromosome USDA4893v1_U4893v1*	<i>Sesbania cannabina</i> , Senegal	Sugawara et al., (2013)
<i>Ensifer sojae</i> CCB4U 05699	AQ1000000000*	AQ1000000000*	<i>Glycine max</i> , China	Tian et al., (2012)
<i>Ensifer</i> sp. NGR234		NC_000914.2*	<i>Lablab purpureus</i> , New Guinea	Schmeisser et al., (2009)
<i>Ensifer terangaie</i> USDA4894		chromosome USDA4894v1_U4894v1*	<i>Acacia laeta</i> , Senegal	Sugawara et al., (2013)
<i>Mesorhizobium chacoense</i> Pr-5	DQ450927	EF209424	<i>Prosopis alba</i> , Argentina	Laranjo et al., (2008)
<i>Mesorhizobium</i> sp. ORS 3359	CCNC01000179.1*	CCNC01000179.1*	<i>Acacia seyal</i> , Senegal	Diouf et al., (2015)
<i>Mesorhizobium</i> sp. ORS 3324	CCMW01000077.1*	CCMW01000077.1*	<i>Acacia seyal</i> , Senegal	Diouf et al., (2015)
<i>Mesorhizobium</i> sp. CCB4U 01606	JN991113	JN991131	<i>Caragana intermedia</i> , China	Li et al., (2012)
<i>Mesorhizobium</i> sp. CCB4U 75057	JN991121	JN991139	<i>Caragana intermedia</i> , China	Li et al., (2012)
<i>Mesorhizobium</i> sp. CCB4U 01582	JN991118	JN991136	<i>Caragana intermedia</i> , China	Li et al., (2012)
<i>R. gallium</i> R602sp	NZ_AHRD000000000*	NZ_AHRD000000000*	<i>Phaseolus vulgaris</i> , France	Kypides et al., (unpublished)
<i>Ensifer</i> sp. JNUU MHM12	IQ951773	IQ951775	<i>Mimosa himalayana</i> , India	Gehl et al., (2013)
<i>Ensifer</i> sp. JNUU MHM3		IQ951769	<i>Mimosa himalayana</i> , India	Gehl et al., (2013)
<i>Ensifer</i> sp. JNUU MHM4		IQ951771	<i>Mimosa himalayana</i> , India	Gehl et al., (2013)
<i>Ensifer</i> sp. SCAU224		JX073925	<i>Mimosa himalayana</i> , India	Gehl et al., (2013)
<i>R. leguminosarum</i> Mad-7		EF373056	<i>Leucaena leucocephala</i> , China	Xu et al., (2013)
<i>Ensifer kostiensis</i> HAMB1 1476	Z95217		<i>Ononis tridentata</i>	Arenal et al., (unpublished)
<i>Ensifer kostiensis</i> HAMB1 1489T	Z95216		<i>Prosopis chilensis</i> , Sudan	Haukka et al., (1998)
<i>Ensifer arboris</i> HAMB1 1700	Z95215		<i>Acacia senegal</i> , Sudan	Haukka et al., (1998)
Other strains	nifH	nodC	Origin	References
Strains isolated from desert sand dunes in Morocco				
<i>Ensifer</i> sp. LMR001	KP993264	KP993269	<i>Acacia gummifera</i> , Merzouga Desert, Morocco	This study
<i>Ensifer</i> sp. LMR002	KP993265	KP993270	<i>Acacia tortilis</i> ssp. <i>Raddiana</i> , Merzouga Desert, Morocco	This study
<i>Ensifer</i> sp. LMR006	KP993266	KP993271	<i>Acacia tortilis</i> ssp. <i>Raddiana</i> , Merzouga Desert, Morocco	This study
<i>Ensifer</i> sp. LMR008		KP993272	<i>Acacia tortilis</i> ssp. <i>Raddiana</i> , Merzouga Desert, Morocco	This study
<i>Ensifer</i> sp. LMR013	KP993267	KP993273	<i>Acacia tortilis</i> ssp. <i>Raddiana</i> , Merzouga Desert, Morocco	This study
<i>Ensifer</i> sp. LMR021	KP993268	KP993274	<i>Acacia tortilis</i> ssp. <i>Raddiana</i> , Merzouga Desert, Morocco	This study
Strains isolated from wild legumes in the Indian Thar Desert				
<i>Ensifer</i> sp. JNUU TP6 (TP6)	JN571912	KP993275	<i>Tephrosia purpurea</i> , Thar Desert, India	Gehl et al., 2012
<i>Ensifer</i> sp. TW10 (TW10)	AZNX01000000 *	AZNX01000000 *	<i>Tephrosia wallichii</i> , Thar Desert, India	Gehl et al., 2012
Strains isolated from Acacia in Southern Ethiopia				
<i>Ensifer</i> sp. AC01e	JF450346	JF450417	<i>Acacia seyal</i> , southern Ethiopia	Degefu et al., 2012
<i>Ensifer</i> sp. AC01c2	JF450343		<i>Acacia seyal</i> , southern Ethiopia	Degefu et al., 2012
<i>Ensifer</i> sp. AC14d	JF450360	JF450431	<i>Acacia seyal</i> , southern Ethiopia	Degefu et al., 2012
<i>Ensifer</i> sp. AC17b	JF450366	JF450437	<i>Acacia seyal</i> , southern Ethiopia	Degefu et al., 2012
<i>Ensifer</i> sp. AC25e	JF450385	JF450456	<i>Acacia tortilis</i> , southern Ethiopia	Degefu et al., 2012
<i>Ensifer</i> sp. AC27a1	JF450390	JF450457	<i>Acacia tortilis</i> , southern Ethiopia	Degefu et al., 2012
<i>Ensifer</i> sp. AC27d1	JF450393	JF450461	<i>Acacia tortilis</i> , southern Ethiopia	Degefu et al., 2012
<i>Ensifer</i> sp. AC28a	JF450398	JF450464	<i>Acacia tortilis</i> , southern Ethiopia	Degefu et al., 2012
<i>Ensifer</i> sp. AC38d2	JF450400	JF450469	<i>Acacia abyssinica</i> , southern Ethiopia	Degefu et al., 2012
<i>Ensifer</i> sp. AC47a	JF450409	JF450471	<i>Sesbania sesban</i> , southern Ethiopia	Degefu et al., 2012
<i>Ensifer</i> sp. AC52c	JF450409	JF450480	<i>Sesbania sesban</i> , southern Ethiopia	Degefu et al., 2012

Sequences in bold correspond to new sequence data
 * Corresponding sequences retrieved from whole genome sequences

References cited in Supplementary table S2

Barnett et al., , (2001) Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888.
 Degefu et al., , (2012) Int. J. Syst. Evol. Microbiol. 62, 2286-2295.
 Diouf et al., , (2015) PloS One. 6;10(2):e0117667
 Duran et al., , (2014) Int. J. Syst. Evol. Microbiol. 64 , 2072-2078.
 Gehlot et al., , (2013) Ann. Bot. 112, 179-196.
 Haukka et al., , (1998) Appl. Environ. Microbiol. 64, 419-426.
 Laranjo et al., , (2008) FEMS Microbiol Ecol. 66, 391-400.
 Li et al., , (2012) Syst. Appl. Microbiol. 35, 239-245.
 Schmeisser et al., , (2009) Appl. Environ. Microbiol. 75, 4035-4045.
 Sugawara et al., , (2013) Genome Biol. 14:R17.
 Tian et al., , (2012) Proc. Natl. Acad. Sci. U.S.A. 109, 8629-8634.
 Reeve et al., , (2013) Stand Genomic Sci. 31, 473-83.

Reeve et al., , (2010) Stand Genomic Sci. 2, 77-86.

Xu et al., , (2013) Appl. Microbiol. Biotechnol. 97, 783-793.

Supplementary table S3: description of Merzouga desert sites sampled

Site	GPS coordinates	Type	pH	Total N (%)	Total C (%)	Assimilable P (mg/kg)
1	N 31°13'48,1"/W 3°58'15,7"	Sand dune without vegetation	9,2	< 0.02	0,65	5
2	N 31°03'06,9"/W 3°59'31,2"	Sand dune without vegetation	8,9	< 0.02	0,15	8
3	N 31°03'36,7"/W 3°59'07,7"	Sand dune without vegetation	8,7	< 0.02	0,12	5
4	N 31°5'42,7"/W 3°57'56,1"	Sand dune without vegetation	8,7	< 0.02	0,10	3

[illegible]

68

1. Introduction

Nous avons utilisé deux espèces de *Vachellia* (*V. gummifera* et *V. ornilis ssp. Radianna*) pour isoler des rhizobiums présents dans les sables de dunes de Merzouga au Maroc. Dans le cadre de cette étude, une caractérisation phénotypique et phylogénétique a été réalisée. Un type génétique dominant a été identifié et indique qu'elles appartiennent à une nouvelle espèce (Sakrouhi et al., 2016). Cette nouvelle espèce est particulièrement adaptée aux environnements arides tel que retrouvé dans le désert marocain chaud.

2. Résultats et discussion

La souche *Ensifer sp.* LMR001, a été isolée avec d'autres souches de rhizobiums à partir des dunes du désert de Merzouga par une technique de piégeage en utilisant deux espèces de légumineuses adaptées aux conditions (semi-) désertiques : *Vachellia gummifera* et *Vachellia tortilis ssp. raddiana*. Les échantillons de sable ont été collectés dans le désert de Merzouga et ont été analysés. L'estimation de la teneur totale en carbone et en azote, du phosphate disponible et du pH de la dune collectés à 4 endroits différents est faite (voir les données GPS dans l'article, tableau supplémentaire). L'analyse des sables montre des niveaux de carbone total très faibles allant de 0,10 à 0,65 % et des niveaux d'azote inférieurs au seuil de détection (< 0,02 %). Les données de la disponibilité du phosphore montraient des valeurs qui varient entre 3 et 8 mg/kg avec un pH alcalin de 8,7 à 9. Ces résultats confirment que les sols de ces zones sans couvert végétal sont très pauvres et non-fertiles.

Au final, 17 rhizobiums ont ainsi pu être isolés. L'analyse phylogénétique utilisant les gènes de ménages (*dnaK-glnA-gltA-recA*) ont montré que tous appartiennent au genre *Ensifer* (*Sinorhizobium*). Les données phylogénétiques montrent que 15 souches partagent le même groupe taxonomique et 2 souches sont divergentes. L'étude phylogénique montre également que les 15 souches partageant le même clade groupaient avec d'autres souches isolées d'un autre désert mais en Inde (le désert de Thar). Cette étude suggère que ces souches sont présentes sur plusieurs continents mais spécifiquement dans des zones arides ou/et sablonneuses. L'analyse phylogénétique basée sur les gènes symbiotiques *nodC* et *nifH* a révélé une haute conservation des séquences entre les souches isolées de Merzouga appartenant au même génotype dominant et qui diffèrent cependant des isolats indiens qui possèdent des gènes de ménage quasiment identiques suggérant l'acquisition de ces gènes symbiotiques par un transfert horizontal de gènes entre les bactéries localement (Sakrouhi et al., 2016). Ces bactéries appartenant probablement à une nouvelle espèce de *Ensifer* isolée au Maroc a montré

une grande capacité d'adaptation aux conditions extrêmes du désert (Sakrouhi et al., 2016). La bactérie est Gram négatif, aérobies, mobiles et ne produit pas de spores. Les colonies sont circulaires, légèrement à très muqueuses, blanches opaques, convexes et de 1 à 3 mm de diamètre après 48 h de croissance sur milieu YMA à 30 °C. Elle peut pousser à des températures particulièrement élevées (pouvant atteindre jusqu'à 47.5 °C), et tolère des pH allant de 6 à 9,5.

3. Conclusion

Des rhizobiums natifs de sable de désert ont pu être isolés. Une majorité des souches avec d'autres souches qui a été isolée d'un autre désert mais en Asie et appartiennent à un nouveau taxon. Les résultats obtenus à partir des phylogénies basées sur les gènes symbiotiques et sur les séquences des gènes conservés suggèrent donc que les gènes de la symbiose ont été acquis par un échange local via des transferts latéraux de gènes. Les souches isolées sont fortement tolérantes au stress thermique et salin et au pH alcalin et représentent un nouveau modèle de rhizobium adapté au désert.

CHAPITRE III

Étude transcriptomique et analyse fonctionnelle du gène *rpoE2* régulateur impliqué dans la réponse générale au stress chez *Ensifer aridi* en réponse au stress salin

Article 2, Belfquih et al., 2021: *Ensifer aridi* LMR001T Symbiosis and Tolerance to Stress Do Not Require the Alternative Sigma Factor RpoE2

Auteur: Meryem Belfquih 1,2, Ilham Sakrouhi 1, Hassan Ait-Benhassou 1,3, Emeric Dubois 4,5, Dany Severac 4,5, Abdelkarim Filali-Maltouf 1,* and Antoine Le Quere 1,2

¹ Laboratory of Microbiology and Molecular Biology (LMBM), Department of Biology, Faculty of Sciences, University Mohammed V in Rabat, Rabat, 10100, Morocco; belfquih.meryem@gmail.com (M.B.); ilham.sakrouhi@hotmail.fr (I.S.); h.aitbenhassou@mascir.com (H.A.-B.); filalimaltouf@gmail.com (A.F.-M.); antoine.le-quere@ird.fr (A.L.Q.)

² Laboratoire des Symbioses Tropicales et Méditerranéennes (LSTM), University of Montpellier, IRD, CIRAD, INRAE, Institut Agro, Montpellier, 34398, France

³ Medical Biotechnology Center, Moroccan Foundation for Advanced Science, Innovation and Research (MAScIR), Rabat Design Center, Rabat 10100, Morocco

⁴ Institut de Génomique Fonctionnelle (IGF), University of Montpellier, CNRS, INSERM, Montpellier, 34094, France; emeric.dubois@mgx.cnrs.fr (E.D.); dany.severac@mgx.cnrs.fr (D.S.)

⁵ Montpellier GenomiX, France Génomique Montpellier, 34094, France

* Correspondence: (A.F.-M.)

Citation: Belfquih, M.; Sakrouhi, I.; Ait-Benhassou, H.; Dubois, E.; Severac, D.; Filali-Maltouf, A.; Le Quere, A. *Ensifer aridi* LMR001^T Symbiosis and Tolerance to Stress do not Require the Alternative Sigma Factor RpoE2. *Agronomy* **2021**, *11*, x. <https://doi.org/10.3390/xxxxx>

Academic Editor(s): Anna Tsyganova and Viktor E. Tsyganov

Received: 20 July 2021

Accepted: 2 September 2021

Published: date

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Abstract: The recently proposed species *Ensifer aridi* represents an interesting model to

study adaptive mechanisms explaining its maintenance under stressful pedo-climatic conditions. To get insights into functions associated with hyperosmotic stress adaptation in *E. aridi*, we first performed RNAseq profiling of cells grown under sub-lethal stresses applied by permeating (NaCl) and non-permeating (PEG8000) solutes that were compared to a transcriptome from unstressed bacteria. Then an *a priori* approach, consisting of targeted mutagenesis of the gene encoding alternative sigma factor (*rpoE2*), involved in the General Stress Response combined with phenotyping and promoter *gfp* fusion-based reporter assays of selected genes was carried out to examine the involvement of *rpoE2* in symbiosis and stress response. The majority of motility and chemotaxis genes were repressed by both stresses. Results also suggest accumulation of compatible solute trehalose under stress and other metabolisms such as inositol catabolism or the methionine cycling-generating S-adenosyl methionine appears strongly induced notably under salt stress. Interestingly, many functions regulated by salt were shown to favor competitiveness for nodulation in other rhizobia, supporting a role of stress genes for proper symbiosis' development and functioning. However, despite activation of the general stress response and identification of several genes possibly under its control, our data suggest that *rpoE2* was not essential for stress tolerance and symbiosis' development, indicating that *E. aridi* possesses alternative regulatory mechanisms to adapt and respond to stressful environments.

Keywords: *Ensifer aridi*; hyperosmotic stress response; alternative Sigma factor; RNAseq

1. Introduction

The rhizobium strain LMR001^T (=LMG 31426^T; =HAMBI 3707^T) was isolated from desert sand dune in the Moroccan Merzouga desert [1]. This strain belongs to a new species within the *Ensifer* genus (formerly *Sinorhizobium*) that was proposed to be named *E. aridi* as isolates belonging to this new taxon were primarily isolated from arid or semi-arid environments in Africa, Asia, and America [1–3]. These bacteria are gram-negative, rod-shaped, motile, fast growing, mucous and develop a symbiotic interaction with

diverse legume species able to survive under environmental constraints such as those found in deserts. More recently, *E. aridi* strains were recovered in Pakistan from mung bean grown in soils receiving as little as 12 mm of mean rainfall during the crop season [4], and strains have also been recovered from *Retama monosperma* growing in abandoned lead mine soils in Morocco [5]. The ability to develop the symbiosis with locally growing legume species was proposed to be due to accessory genome variations among strains, which most probably result from large plasmid exchanges through horizontal gene transfer [6].

Studying nitrogen-fixing rhizobia diversity and mechanisms underlying their adaptive responses to environmental constraints such as drought or saline stresses to improve the fertility of soils and arid land resilience was raised decades ago [7]. In a context of climate change, soil degradation, and demographic growth, there is now an emergency to act and propose solutions, as warned by the latest IPCC report on Climate change and land, published in August 2019. Capable to utilize a wide range of carbon sources and to develop the symbiosis with plants adapted to environmental constraints, *Ensifer aridi* represents an interesting model to study mechanisms involved in the adaptation to aridity.

Microorganisms such as rhizobia are adapted to contrasted growth conditions (soil, rhizosphere, *in planta*) and may be recovered from contrasted pedo-eco-climatic environments. They may face oxidative, pH, acid, hyperosmotic, and/or thermal stresses during their life cycle and have developed numerous mechanisms to thrive. Among functions associated to environmental stress adaptation, bacteria use various mechanisms including cell membrane structure and composition modifications [8], tuning its permeability and transporter activity to ensure proper nutrition and ionic exchanges. For example, phospholipids and the production of surface polysaccharides may be induced to prevent water loss and maintain a hydrated microenvironment around the cell [9–11]. Bacteria may respond by a modification in the transport and/or the metabolism of several molecules such as potassium ions, sugars, or amino acid derivatives that may act as compatible solutes to counteract osmotic pressure and/or

protect bacteria against intracellular damages notably in rhizobia [12]. Among them, trehalose, a non-reducing disaccharide, appears promising as it may improve plant stress tolerance through symbiosis-mediated accumulation [13]. Bacteria may also trigger repair mechanisms to protect their macromolecules against damages by reactive oxygen species [14]. Finally, several transcription factors are also involved to modulate gene expression and enable the bacteria to rapidly respond to environmental changes through specific regulatory pathways.

Indeed, when subjected to environmental stressors, alphaproteobacteria, the bacterial class to which rhizobia belong, induce a conserved General Stress Response (GSR) that involves a regulatory system composed of an Extra Cytoplasmic Function (ECF) sigma factor (σ^{ECFG}), an anti-sigma factor (NepR), and an anti-anti sigma factor (PhyR), whose activity depends on its phosphorylation by the Histidine kinase PhyK [15]. More recently, Lori and colleagues showed the involvement of a single domain response regulator MrrA as an intermediate between histidine kinases and PhyK, suggesting that PhyK may act as a phosphotransferase rather than a phosphate kinase [16]. Inducing molecules interacting with the primary sensors and some regulatory proteins involved in the GSR regulation remain to be determined. However, the phosphorylation of the multi-domain alternative two component system protein PhyR leads to its binding to the anti-sigma factor NepR, which is under normal growth conditions bound to σ^{ECFG} , thus releasing its interaction with RNA polymerase, which triggers GSR [15].

Here, to explore the regulatory response of the new species *E. aridi* to salt- or water-induced hyperosmotic stresses, a genome-wide transcriptional study was performed. A mutant in the *Ensifer aridi* LMR001 σ^{ECFG} encoding gene (*rpoE2*) was generated and studied to evaluate its involvement in the GSR-mediated regulation of salt-induced osmotic upshift of selected metabolisms' transport and regulation or to assess the importance of the GSR to cope with such stress and to develop the symbiosis with compatible hosts.

2. Materials and Methods

2.1. Bacterial Growth

The bacterial strains and plasmids used in this study are listed in Table 1. *Ensifer* strains and derivatives were cultured in tryptone yeast extract (TY) medium [17] at 28 °C. *Ensifer* cultures were routinely started from glycerol stocks upon a primary growth on TY agar plates. *E. coli* strains were grown at 37 °C in Luria-Bertani medium [18]. Solid media contained 2% agar. Gentamycin (Gm, 10 µg mL⁻¹), kanamycin (Km, 50 µg mL⁻¹), ampicillin (Ap, 50 µg mL⁻¹), rifampicin (Rif, 100 µg mL⁻¹), and 5-bromo-4-chloro-3-indolyl-beta-D-galactosidase (X-gal, 80 µg mL⁻¹) were added to the media for selection as required.

2.2. Phenotypic Characterization of the Δ rhoE2 Strain

2.2.1. Plant Nodulation Assay

Seeds of *Vachellia tortilis* subsp. *raddiana* and of the endemic Moroccan *Vachellia gummifera*[19] were scarified and surface sterilized in concentrated sulfuric acid (95%) with agitation for 1 h for *V. tortilis* and 30 min for *V. gummifera* and rinsed thoroughly with sterile water. Seeds were then germinated on water agar plates at 25 °C for 5 days in the dark. After germination, seedlings were planted in pots containing sand that was previously sterilized by autoclaving for 1 h at 120 °C. *Ensifer aridi* LMR001 WT and Δ rhoE2 were pre-grown in TY medium up to an optical density (OD) 600 nm of 0.7 and then washed and re-suspended in sterilized water prior to inoculation using 1 mL of washed cells directly after planting. Six individual plants were inoculated with each rhizobial strain or sterile water for the negative and positive controls. Plants were incubated in a growth chamber at 26 °C under a photoperiod of 16 h day and 8 h night and watered using N-free plant growth BD medium for watering [20] that was supplemented with KNO₃ (3 mM) as a source of nitrogen for the positive controls. After 6 months of growth, the plants were harvested, and nodule number, plant shoot, and root fresh weights were recorded.

2.2.2. Motility Assay

LMR001 and $\Delta rpoE2$ were grown at 28 °C in TY broth with shaking at 180 RPM to an OD (600 nm) of 0.7. After incubation, cells were pelleted, washed twice with sterile water, and resuspended in 10 volumes of sterile water. Swimming test assays were performed using 20 mL of TY media that were complemented either with 0, 5.85, or 11.7 g.L⁻¹ NaCl (generating, respectively, $\approx$ 0, 0.1, and 0.25 M final NaCl concentrations) and 0.2% agar. Each medium was poured into Petri dishes, upon cooling to 50 °C in a water bath to avoid water loss by evaporation, and allowed to solidify in a horizontal laminar flow for 15 min. Finally, aliquots (5 μ L) of washed cells were gently dispensed into the center of the plates and incubated at 28 °C for 5 days and the diameters of motile bacteria were measured after 2, 3, and 5 days of growth. The average diameter and the standard deviations from three biological replicates were calculated.

2.2.3. Salinity Tolerance

Ensifer aridi LMR001 and $\Delta rpoE2$ derivatives were first grown in liquid TY cultures up to an OD (600 nm) of 0.7, and 1 mL was used to inoculate 50 mL of liquid TY medium as well as TY complemented with 14.6, 17.5, or 23.4 g L⁻¹ of NaCl (generating, respectively, 0.25 M, 0.3 M, and 0.4 M final salt concentrations). Cultures were incubated in a rotary shaker (180 RPM) at 28 °C and the OD (600 nm) was recorded after 4, 8, 24, 48, 120, and 144 h incubation. The mean ODs (600 nm) of three independent cultures were calculated for both strains and at each time point with their respective standard deviations.

2.2.4. Sensitivity to Detergent, Acid, and Oxidative Stresses

Sensitivities to oxidative, acid, and detergent stresses induced, respectively, by H₂O₂, HCl, and SDS were compared between the wild-type LMR001 and the $\Delta rpoE2$ strains. Briefly, 4 mL of bacterial cultures of *E. aridi* or the $\Delta rpoE2$ strain (pre-grown to OD 600 nm $\approx$ 0.7 in TY liquid medium) were added to 16 mL of soft TY agar containing 0.7% of agar precooled to 45 °C and poured onto TY agar plates. Once solidified, a paper disc (6-mm diameter) was placed on the plate and 5 μ L of H₂O₂ (2 M), HCl (5.5 M), or SDS (10% w/v) was spotted onto it. Plates were incubated at 28 °C for three days prior to

recording the diameters of the inhibition zones, and the average of three replicates were averaged.

2.3. RNAseq Analysis and Characterization of Differentially Regulated Genes in LMR001 Upon NaCl- or PEG8000-Induced Hyperosmotic Stresses

For the RNAseq analysis, the method was adapted from previously published work [21]. Briefly, six biological replicates were first grown in TY medium up to an OD (600 nm) of 0.9 (+/- 0.05) in 250 mL Erlenmeyer flasks at 30 °C with 140 rpm. For each replicate, 15 mL were added to an equal volume of pre-warmed TY containing either nothing (Controls), NaCl at 500 mM (Salt treatment, 250 mM final concentration), or PEG8000 at 24% (W:V) (PEG8000 treatment, 12% final concentration), and were further incubated at 30 °C and 140 rpm. After 2 ½ hours of incubation, which is less than a generation time, 1/10th volume of stop buffer (ethanol containing phenol at a final concentration of 5%) was added to each culture (6 replicates × 3 treatments) to inhibit RNAses and centrifuged 4 min at 8000 g and 4 °C. The pellets were snap-frozen in liquid nitrogen and kept in the freezer at -80 °C until use.

Total RNAs were purified from each replicate and treated with the RiboPure™ kit (AMBION), which included a DNase I treatment, according to manufacturer's recommendations. RNA extracts from the six replicates were then pooled and subjected to two successive rounds of rRNA subtractions with Microbe Express™ kit (AMBION), following manufacturer's recommendations. Purified mRNAs were sent to Montpellier Genomix for sequencing. Three libraries corresponding to the three treatments (Control, Salt, and PEG8000) were prepared, following illumina kit instructions. Briefly, for each library, 100 ng of mRNA were first fragmented chemically, and reverse transcribed using random primers and SuperScript II™ to generate first-strand cDNAs. The second strand was synthesized using DNA polymerase I and RNase H. The cDNA fragments were then end-repaired, a single 'A' base was added, and adapters were ligated. These products were enriched by PCR to select for cDNA fragments containing adaptors and to create the final cDNA library. The efficacy of the library construction was checked in a quality control step that involved measuring the adapter-cDNA size and concentration on an

Agilent DNA 1000 chip. Sequencing libraries were denatured with sodium hydroxide and diluted to 6 pM in hybridization buffer for loading onto a single lane of an Illumina HiSeq 2000 flowcell V1.5. Cluster formation, primer hybridization, and single-read, 100 sequencing cycles were performed on cBot and HiSeq2000 (Illumina, San Diego, CA, USA), respectively.

After sequencing, reads that passed the sequencing platform quality filter were imported into CLC Genomics workbench (V5.5.1) and mapped to the RAST-predicted genomic objects [6] including putative protein encoding genes (PEGs) (listed in Table S1) and RNAs from *Ensifer aridi* LMR001 using 95% of sequence similarity over 70% of read length. Reads that mapped several PEGs were discarded. Finally, the read table obtained upon mapping the three libraries onto PEGs were normalized using a LOESS transformation that appeared to be the most adapted method, as the PEG8000 treatment resulted in an enhanced recovery of the least abundant transcripts as compared to the control or salt treatments (see Results section). Genes showing an average coverage of 50X and more than 2-fold differences between pairwise comparisons were considered as differentially regulated.

To explore the major changes associated to the transcriptomes, the predicted proteome of the LMR001 strain was functionally annotated using COG and KEGG classifications [22,23]. Differentially regulated genes were then used to estimate major classes and functions associated to salt- and PEG8000-induced hyperosmotic stresses.

2.4. Real-Time PCR on Selected Genes

The list of genes targeted, as well as primers used for real-time PCR experiments, are listed in Table 1. LMR001 was cultivated either in TY or in TY complemented with NaCl at 250 mM, as described for the RNAseq experiment. Three biological replicates for each growth condition were extracted using SV Total RNA Isolation kit (Promega), following manufacturer's recommendations. RNA samples were treated with Turbo DNase (Applied Biosystems) to remove any DNA contaminant from RNA samples. For each sample and replicate, 1 µg of total RNA was reverse transcribed with 400 U of SuperScript II (Invitrogen) and random hexamer primers. The qPCRs were performed on a

Stratagene MXP3005P system using Power SYBER green master mix (Applied Biosystems). PCR cycling conditions consisted of a primary 10-min incubation at 95 °C, followed by 40 cycles consisting of 15 s at 95 °C and 30 s at 60 °C. Primer specificity and efficiencies and the formation of primer dimers were estimated by dissociation curves. The expression levels of a selection of 12 genes were standardized by using the gene *hrcA* as reference. PCR efficiency (E) for each amplicon was calculated using the linear regression method on the log (fluorescence)-per-cycle-number data using Stratagene MXpro software. All qPCRs were performed in two technical replicates using total RNA isolated from three biological samples so as to assess biological variability. Data for each sample were expressed relative to the expression levels of *hrcA* by using the mathematical model described previously, which determines the relative quantification of a target gene in comparison to a reference (ref) gene between treatment and control samples [24]. The relative expression ratio (R) of a target gene is calculated based on efficiency (E) and the threshold cycle (Ct) of an unknown sample versus a control and expression in comparison to a reference gene. The average relative quantity for each gene under study was calculated and log₂-transformed. Means and standard deviations of these final log₂ ratios were calculated using data from three biological replicates.

2.5. LMR001 Δ rpoE2 Mutant Construction

The regions bordering the *rpoE2* gene (PEG2539) were first amplified by PCR with custom-designed primers (see Table 1). The primer pair rpoE2-P2539-A-XI/rpoE2-P2539-B-HIII was used to amplify a 478 bp sequence located upstream of the *rpoE2* gene including the first 37 bases of the coding sequence. Furthermore, the primer rpoE2-P2539-B-HIII contained in 5' a HindIII site as well as 13 bases that were complementary to the 5' end of the primer rpoE2-P2539-C-HIII that was used together with rpoE2-P2539-D-XI to amplify a 473 sequence of the 3' region of the *rpoE2* gene including the last 77 nucleotides. The PCR was performed using, as templates, genomic DNA extracts from LMR001 obtained, following the method reported previously [25], using 10 mL of bacterial liquid cultures at early stationary phase. First, PCR reactions were performed to amplify both the 5' and the 3' flanking regions using the proof-reading enzyme Phusion

(Thermo-Fischer) and recommended chemistry, except for primers (1 μ M final concentrations for forward and reverse primers). The cycling conditions included a primary denaturation step of 30 s at 98 °C followed by 35 cycles (98 °C, 10 s/56 °C, 30 s/72 °C, 45 s) and a final elongation step of 7 min at 72 °C. PCR products were purified using Illustra™ GFX PCR DNA and Gel Band purification kits (GE Healthcare), following manufacturer's recommendations. Then a second PCR reaction was set up, using as primers and matrix, purified PCR products obtained that contain complementary sequences, thus enabling priming of the amplification. The 25- μ L reactions contained 5 μ L of 5X GoTaq buffer (Promega), 1.25 μ L of MgCl₂ (25 mM), 2.5 μ L of dNTP mix (2.5 mM each), 0.125 μ L GoTaq (5 U/ μ L), 2 μ L of each purified PCR products obtained, and nuclease free water. A primary PCR cycling consisting of a denaturation step of 95 °C for 5 min followed by 5 cycles (95 °C, 30 s/50 °C, 45 s/72 °C, 50 s) and 5 min at 72 °C was performed so as to generate chimeric DNA molecules consisting of merged 5' and 3' regions. Then, 15 μ L of the mix (3 μ L of 5X GoTaq buffer, 0.75 μ L of MgCl₂ (25 mM), 1.5 μ L of dNTP mix, 0.125 μ L of GoTaq (5 U/ μ L), 4 μ L of external primers rpoE2-P2539-A-XI and rpoE2-P2539-D-XI, and 1.625 μ L of nuclease free water) was added to the PCR reaction and subjected to the following cycling conditions: a denaturation step of 95 °C for 5 min followed by 25 cycles (95 °C, 30 s/55 °C, 45 s/72 °C, 50 s) and 5 min at 72 °C. The PCR product was purified from 1% agar gel using promega Wizard® SV Gel and PCR Clean-Up System, following recommended protocol. As a result of the two amplifications' process, the PCR-amplified and -purified product was cloned as a fragment into a pGEM-T Easy Vector (Promega kit) and transferred into XL2-Blue Ultra-competent Cells. The resulting positive clones were selected and the plasmid was extracted by using the Wizard® Plus SV Minipreps DNA Purification System, digested by XbaI enzyme and ligated into suicide vector pJQ200SK [26]. Enzymes were purchased from Promega. The latter plasmid was used to transform *Ensifer aridi* LMR001 cells by bipartite conjugation method using S17-1 *E. coli* strains as a donor [27]. Transformants were selected on TY plates containing Rif and 5% sucrose to allow selection of knockout clones that had undergone a double cross-over. Double cross-over gene replacement in

mutants was verified by screening for antibiotic resistant/ sensibility phenotypes (Rif and Gm, respectively) and also PCR to verify deletion by using specific external and internal primers.

2.6. Use of Promoter-Gfp Fusions for Transcriptional Analyses in the Wild-Type and the Δ rhoE2 Strains

The promoter region of genes involved in the regulation of the le cycle de la méthionine(*nesR*, PEG5735; *sahR*, PEG2239), in the regulation of the GSR (*rsiA1*, PEG2540; *rsiB1*, PEG2541), in the trehalose transport (*thuE*,PEG6268) metabolism (*otsB*, PEG4868; *treZ*, PEG5323), a glycosyl transferase (PEG4866) located upstream of the *otsBA* operon in reverse orientation or the regulation of trehalose transport and endogenous utilization (*thuR*, PEG6269), and in the inositol catabolism (*iolC*, PEG373) and regulation (*iolR*, PEG374) was amplified using primers listed in Table 1. PCR products were cleaned up using illustra™ GFX™ PCR DNA and Gel Band Purification Kits (GE Healthcare). The purified PCR products were ligated into TOPO® TA Cloning® vector (INVITROGEN), following the manufacturer's recommendations, and were transferred into *E. coli* XL2 Blue Ultra-competent Cells. Plasmid DNA was extracted using the Wizard Plus SV Minipreps DNA Purification System kit (Promega), following the manufacturer's recommendations. Plasmids containing the promoters were sequenced by Genoscreen (Lille, France) to verify sequence integrity. Inserts were excised using the restriction enzyme EcoRI and ligated into a pPROBE NT vector, which allows for transcriptional fusions with the green fluorescent protein (GFP) encoding gene *gfp*[28], previously digested with EcoRI and dephosphorylated. The ligation mix was transferred into electro-competent *E. coli* strain S17-1. The selection of clones containing the promoters in the proper orientation was done by using the forward primers of the promoters with the primer *gfp*-Rev50 located in the reverse orientation of the 5' end of the *gfp* gene, which results in positive PCR amplification when correct. The pPROBE NT constructs were transferred into LMR001 Rif^R or Δ rhoE2 strains by biparental mating. Briefly, Donor S17-1 cells containing the pPROBE NT derivatives were grown on LB agar containing Kanamycine (50 µg mL⁻¹) and the recipient cells were grown onto TY Agar containing

Rifampicine ($100\text{ }\mu\text{g mL}^{-1}$). Half a loop of donor cells and 2–3 loops of recipient cells were resuspended in separate Eppendorf tubes containing 1 mL of TY medium each to wash cells from antibiotics. The tubes were mixed by mild vortexing and centrifuged at 7000 g for 7 min and resuspended in 0.5 mL of fresh TY medium. Then, 100 μL of each of the donor and the recipient cells were mixed onto fresh TY plates and grown for 2 days at 28 °C. Half loops of cells were then streaked onto a TY agar plate containing the two antibiotics (Kanamycine and Rifampicine) for selection of transconjugants. Individual colonies were finally isolated and verified by PCR prior conservation in 20% glycerol at $-80\text{ }^{\circ}\text{C}$. When needed, strains carrying the promoter pPROBE constructs were grown on TY agar plates supplemented with appropriate antibiotics at 28 °C. Using sterile toothpicks, fresh colonies were transferred to 10 mL liquid TY containing Kanamycine and grown overnight with shaking at 28 °C. Cells were washed in fresh TY medium and resuspended in an equal volume of fresh TY medium prior to use in transcriptional assays. For each transcriptional assay, the same bacterial inocula (70 μL) were used to inoculate 140 μL TY medium supplemented or not with NaCl (250 mM final concentration) in 96-well plates. The plates were transferred in a spectrofluorometer microplate reader (TECAN Infinite M200) set at 28 °C and subjected to the following cycling conditions. After a preliminary shaking of 900 s using orbital shaking mode with 1-mm amplitude, the mean of three reads of optical density (600 nm) and three reads of fluorescence levels (excitation filter at 485 nm and emission filter at 535 nm) were recorded every 30 min for all wells and at least 9 h. The experiment was performed three times. For the analysis, each plate was treated separately and relative fluorescence values were recorded as follow. First, the fluorescence data were transformed for each time point as a percentage of the maximal fluorescence levels observed in the experiment to enable the comparison between each replicate. The relative fluorescence values were finally obtained by dividing the normalized fluorescence obtained by the corrected OD600 nm measures at each time point. Data from the bacterium containing empty vector were finally subtracted and the means from the three independent repetitions were finally calculated.

Table 1. Strains, plasmids, and primers used in the study.

Strains Plasmids Primers	Description/Primer Sequence	Source/Reference
Strains		
<i>Escherichia coli</i> strains		
XL2 Blue Ultra-competent Cells S17-1	<i>endA1supE44 thi-1 hsdR17 recA1 gyrA96 relA1 lac</i> [F' proAB <i>lacIq</i> ΔM15 Tn10 (Tetr) Amy Camr] <i>recA</i> [SmR], thi, pro, RP4-2-Tc:Mu: <i>aph::Tn7</i> λpir.	Stratagene [27]
<i>Ensifer aridi</i> strains		
LMR001 ^T	<i>Ensifer aridi</i> LMR001 ^T (=LMG 31426 ^T ; =HAMBI 3707 ^T)	[1,2]
LMR001 Δ <i>rpoE2</i>	<i>rpoE2</i> deletion mutant of <i>Ensifer aridi</i> LMR001 ^T	This work
Plasmids		
pGEM-T Easy	Amp ^R , pUC origin, Multi Cloning Sites), <i>lacZ</i> gene fusion, β-lactamase coding region, <i>lac</i> operon sequences.	Promega
TOPO vector	Km ^R , Amp ^R , pUC origin, <i>lacZ</i> α- <i>ccdB</i> gene fusion. poisomerase enzyme	Invitrogen
pPROBE NT (pNT)	Km ^R , pBBR1 replicon, Multi Cloning Sites upstream of <i>gfp</i> reporter gene	[28]
pJQ200SK+	Gm ^R , SacB; Ori origin replication, mob region, Multi Cloning Sites	[26]
pNT- <i>sahRp</i>	pPROBE NT containing <i>sahR</i> (PEG2239) promoter fused to <i>gfp</i> gene, KmR	This work
pNT- <i>nesRp</i>	pPROBE NT containing <i>nesR</i> (PEG5735) promoter fused to <i>gfp</i> gene, KmR	This work
pNT- <i>rsiA1p</i>	pPROBE NT containing <i>rsiA1</i> (PEG2540) promoter fused to <i>gfp</i> gene, KmR	This work
pNT- <i>rsiB1p</i>	pPROBE NT containing <i>rsiB1</i> (PEG2541) promoter fused to <i>gfp</i> gene, KmR	This work
pNT- <i>thuEp</i>	pPROBE NT containing <i>thuE</i> (PEG6268) promoter fused to <i>gfp</i> gene, KmR	This work
pNT- <i>otsBp</i>	pPROBE NT containing <i>otsB</i> (PEG4868) promoter fused to <i>gfp</i> gene, KmR	This work
pNT- <i>treZp</i>	pPROBE NT containing <i>treZ</i> (PEG5323) promoter fused to <i>gfp</i> gene, KmR	This work
pNT-PEG4866p	pPROBE NT containing a glycosyl transferase encoding gene (PEG4866) promoter fused to <i>gfp</i> gene, KmR	This work
pNT- <i>thuRp</i>	pPROBE NT containing <i>thuR</i> (PEG6269) promoter fused to <i>gfp</i> gene, KmR	This work
pNT- <i>iolCp</i>	pPROBE NT containing <i>iolC</i> (PEG373) promoter fused to <i>gfp</i> gene, KmR	This work
pNT- <i>iolRp</i>	pPROBE NT containing <i>iolR</i> (PEG374) promoter fused to <i>gfp</i> gene, KmR	This work
Primers		
Targeted gene primers for qPCR		
<i>iolB</i> (PEG370) Forward	5'-CGGGCACACGTCCTCCTATC-3'	This work

<i>iolB</i> (PEG370) Reverse	5'-CCCTTGGGTACCAGCGTGAC-3'	This work
<i>flgB</i> (PEG666) Forward	5'-CCCAGGTGACCGAGGTGAGT-3'	This work
<i>flgB</i> (PEG666) Reverse	5'-CCCCGCATTGAGCTCGTAGT-3'	This work
<i>hrcA</i> (PEG397) Forward	5'-ACCAAGGGGATCGCATCGAC-3'	This work
<i>hrcA</i> (PEG397) Reverse	5'-GGCTCGCCGCTTTCCAGATA-3'	This work
<i>metK</i> (PEG438) Forward	5'-CCCAGCCGCTGTCGATCTAT-3'	This work
<i>metK</i> (PEG438) Reverse	5'-GGCGAGAGGTCCATCGTCTT-3'	This work
<i>cheA</i> (PEG646) Forward	5'-TCCGCTCCGTTTTTCGAGTTC-3'	This work
<i>cheA</i> (PEG646) Reverse	5'-GCAACGGTCCGATCCTCTTC-3'	This work
<i>gltA</i> (PEG1564) Forward	5'-ATTGCGCTGACGGACGAGTAC-3'	This work
<i>gltA</i> (PEG1564) Reverse	5'-GCGTGATGCCGGAATAGAAG-3'	This work
<i>ahcY</i> (PEG3556) Forward	5'-AAGGGCAACCGCATCATTC-3'	This work
<i>ahcY</i> (PEG3556) Reverse	5'-GCTTCGGCAGCACGTAGAC-3'	This work
<i>metH</i> (PEG3206) Forward	5'-TGGGAGCTGAAGGGCGTCTA-3'	This work
<i>metH</i> (PEG3206) Reverse	5'-TCGGTGAAAAGGCGGATGTC-3'	This work
<i>bhMT</i> (PEG1969) Forward	5'-GCGGGTCCGAGATCCACTATT-3'	This work
<i>bhMT</i> (PEG1969) Reverse	5'-CGCAGCAGCCTCCGATGAT-3'	This work
<i>sahR</i> (PEG2239) Forward	5'-ATTTCCGCCTGCGTCAAGAG-3'	This work
<i>sahR</i> (PEG2239) Reverse	5'-CGGCGTTGCGACTGAAATAG-3'	This work
L11MT (PEG2189) Forward	5'-ACAAGGTGAAGGCGGGTGAG-3'	This work
L11MT (PEG2189) Reverse	5'-GAGGTCGAGGACGTTGCAGAA-3'	This work
<i>xylF</i> (PEG2505) Forward	5'-GCTCGGCACCCAGACTGTTT-3'	This work
<i>xylF</i> (PEG2505) Reverse	5'-GGTCTGCACGCCTTCGATCT-3'	This work
<i>thuE</i> (PEG6268) Forward	5'-GCTCCGGCGCTCTACTATCG-3'	This work
<i>thuE</i> (PEG6268) Reverse	5'-GGCGTTTGCCTGGAAGACA-3'	This work

**Primers for *rpoE2*
deletion mutant
construction**

RPOE2-P2539-A-XI	5'-CCTCTAGACATCGCCTGAGGTCTGAGAT-3'	This work
	5'-	
RPOE2-P2539-B-HIII	CGCCCTTCACTTGAAGCTTGCATCTCACGCTTGA ACTCT-3'	This work
RPOE2-P2539-C-HIII	5'-AAGCTTCAAGTGAAGGGCGAGAAC-3'	This work
RPOE2-P2539-D-XI	5'-CCTCTAGATTGACCCATTGCCATTCACG-3'	This work

**Targeted gene
promoter/Primers
for promoter::*gfp*
fusion cloning**

<i>iolRp</i> Forward/ <i>iolR</i> -1	5'-GACTTCCGCGAGGGCTAC-3'	This work
<i>iolRp</i> Reverse/ <i>iolR</i> -B	5'-TCGGCCTCGACGGAGGGACATCGGTGGTTCG-3'	This work
<i>iolCp</i> Forward/ <i>iolC</i> -1	5'-CTCGCTTTCGGCAAGC-3'	This work
<i>iolCp</i> Reverse/ <i>iolC</i> -B	5'-GGCGACGGCTTGGCCGATCGTGATGATGTC-3'	This work
<i>sahRp</i> Forward/ <i>metReg</i> -1	5'-CAATCCGAAGCGGGTG-3'	This work
<i>sahRp</i> Reverse/ <i>metReg</i> -B	5'-GTTTTCTTCCTGGCGTTCCTCTGTCACCC-3'	This work
<i>nesRp</i> Forward/ <i>nesR</i> -15'	5'-GGTTTCCGGGCTGCAC-3'	This work
<i>nesRp</i> Reverse/ <i>nesR</i> -B5'	5'-TTAAGCCGCGGCGGTGAACTGGTTCCTGATC-3'	This work
<i>rsiA1p</i> Forward/ <i>rsiA1</i> -1	5'-GAAGGTAAGGAAGGAATG-3'	This work
<i>rsiA1p</i> Reverse/ <i>rsiA1</i> -B	5'-GAATCGATTGTCCCGCAATCTGTACATTTCG-3'	This work
<i>risB1p</i> Forward/ <i>rsiA1</i> -B	5'-GAATCGATTGTCCCGCAATCTGTACATTTCG-3'	This work
<i>risB1p</i> Reverse/ <i>rsiA1</i> -1	5'-GAAGGTAAGGAAGGAATG-3'	This work
<i>otsBp</i> Forward/ <i>peg4866</i> -B	5'-AGCTTATCGAGAGCAATTTGCGCGATCTTC-3'	This work
<i>otsBp</i> Reverse/ <i>peg4866</i> -1	5'-CTCGATCATGGAGAGG-3'	This work
PEG4866p Forward/ <i>peg4866</i> -1	5'-CTCGATCATGGAGAGG-3'	This work
PEG4866p Reverse/ <i>peg4866</i> -B	5'-AGCTTATCGAGAGCAATTTGCGCGATCTTC-3'	This work
<i>treZp</i> Forward/ <i>treZ</i> -1	5'-GAGCTCGGGTTTGATG-3'	This work
<i>treZp</i> Reverse/ <i>treZ</i> -B	5'-CGAAAGGTTGACCGTCGCTGTCTCTCGCAT-3'	This work
<i>thuRp</i> Forward/ <i>thuR</i> -1	5'-CTCGAAGGCTCCTCAGC-3'	This work
<i>thuRp</i> Reverse/ <i>thuR</i> -B	5'-AGCAATTGCAGCGCGAACTCCTTGAGCTTC-3'	This work
<i>thuEp</i> Forward/ <i>thuR</i> -B	5'-AGCAATTGCAGCGCGAACTCCTTGAGCTTC-3'	This work
<i>thuEp</i> Reverse/ <i>thuR</i> -1	5'-CTCGAAGGCTCCTCAGC-3'	This work
<i>gfp</i> Reverse/ <i>gfp</i> -Rev50	5'-ACATCACCATCTAATTCAAC-3'	This work

3. Results and Discussion

To explore the possible mechanisms that *Ensifer aridi* possibly uses to face hyperosmotic stresses, a transcriptomic approach was realized using NaCl as a permeating solute and PEG8000 as a non-permeating solute that, respectively, resulted in salt- and water-induced osmotic upshifts. To focus on functions associated with these stresses and disregard mechanisms linked to cell death, mild stresses were triggered using sublethal concentrations for less than a generation time. First, growth kinetics of LMR001 were compared in TY medium containing various concentrations of Salt (NaCl) and PEG8000. Concentration of 250 mM of salt and 12% (*w/v*) of PEG8000 were chosen, as these concentrations reduced the growth of the bacterium without inhibiting its growth. Indeed, six independent cultures were pre-grown in TY Medium up to an OD (600 nm) of 0.9. Each pre-culture was further used to inoculate an equal volume of either TY or TY complemented with Salt 0.5 M (0.25 M final concentration) or with PEG8000 24%*w/v* (12% final concentration). After 2 ½ hours of incubation, the mean OD (600 nm) of the control rose from 0.52 to 0.94 ($\Delta OD = 0.42$), which was significantly higher than the ΔOD obtained in media containing salt and PEG8000 (-10% in PEG8000 growth medium and -40% in salt containing medium; *p*-values < 0.05, Student Test), confirming that bacteria were subjected to a mild stress that did not inhibit totally their growth.

All cultures were subjected to a total RNA extraction and purified RNAs were pooled according to the treatments prior to rRNA subtractions. Finally, enriched mRNA samples from controls (TY), salt-stressed bacteria (250 mM NaCl), or Water-stressed bacteria (12% PEG8000) were sent for sequencing at the Montpellier MGX sequencing platform. Sequencing consisted of 100 cycles on Hiseq2000 that generated, respectively, 109.95-, 95.17-, and 105.89-million sequence reads of 100 bp (>31 Gb). For each library, the reads were mapped to the LMR001 RAST Predicted Encoding Genes (PEGs) and rRNAs as well as the spike phiX using CLC genomics Workbench (see Materials and Methods for parameters used). Results of the sequencing and mapping are shown in Figure S1. Even though the number of reads that mapped to rRNAs were high (55.39, 35.1, and 39.23 million reads for control, salt, and PEG libraries, respectively), the procedure enabled us to recover significant numbers of reads that mapped to LMR001-predicted encoding gene

sequences (respectively 34.91, 35.12, and 42.79 million reads). If we consider that 95% of the RNAs in cells are corresponding to ribosomal RNAs, while only 5% are mRNAs, the rRNA subtraction procedure enabled the removal of 91.6%, 94.7%, and 95.2% of rRNAs (respectively, for the libraries TY, TY + 250 mM salt, and TY + 12% PEG8000). These rates are comparable to those obtained in *Mesorhizobium*[21] and indicate that the procedure is also well suited to *Ensifer aridi*.

Scatter plots showing distributions of reads obtained upon mapping on LMR001 PEGs in control (X Axis) as a function of salt or PEG8000 treatments (Y Axes) (Figure S2A,B, respectively) or M-A plots (showing the logarithmic values of mean read numbers for control and salt or control and PEG8000, as a function of the log₂ ratios (Figure S2C and S2D, respectively)) indicated that while salt and control showed rather centered distributions, the sample PEG8000 returned a strong bias for lowly expressed transcripts that were over-represented in this library as compared with control and salt libraries. This bias may result from the presence of PEG8000 remaining in pelleted cells as these molecules favor precipitation of nucleic acids [29,30]. To estimate whether the presence of PEG8000 in the samples also influenced the recovery of particular size fractions of the RNAs, we compared the relative abundances of read counts in the three libraries upon normalization to the total read numbers mapped to PEGs for all predicted genes and ordering based on PEG size, their mean sequencing coverage (obtained from the three libraries), and, finally, by predicted gene order, which is independent of size or coverage. Figure S3 shows the relative abundances of reads obtained in the three libraries (Y axis) for all genes upon ordering, as described above (X axes), as a piled 100% Surface area graphics. The comparison of profiles confirmed that only a bias toward predicted genes lowly transcribed was found. These results clearly show that the PEG8000 facilitated the recovery of rare transcripts using the method used here, and that this enrichment was higher as the coverage was lowering (Figure S3B).

Furthermore, box plots showing the distributions of raw read counts (as log values) in the three libraries also indicated differences in sequencing medians and quartile values (Figure S2E). To normalize for differences in sequencing depth between the three libraries

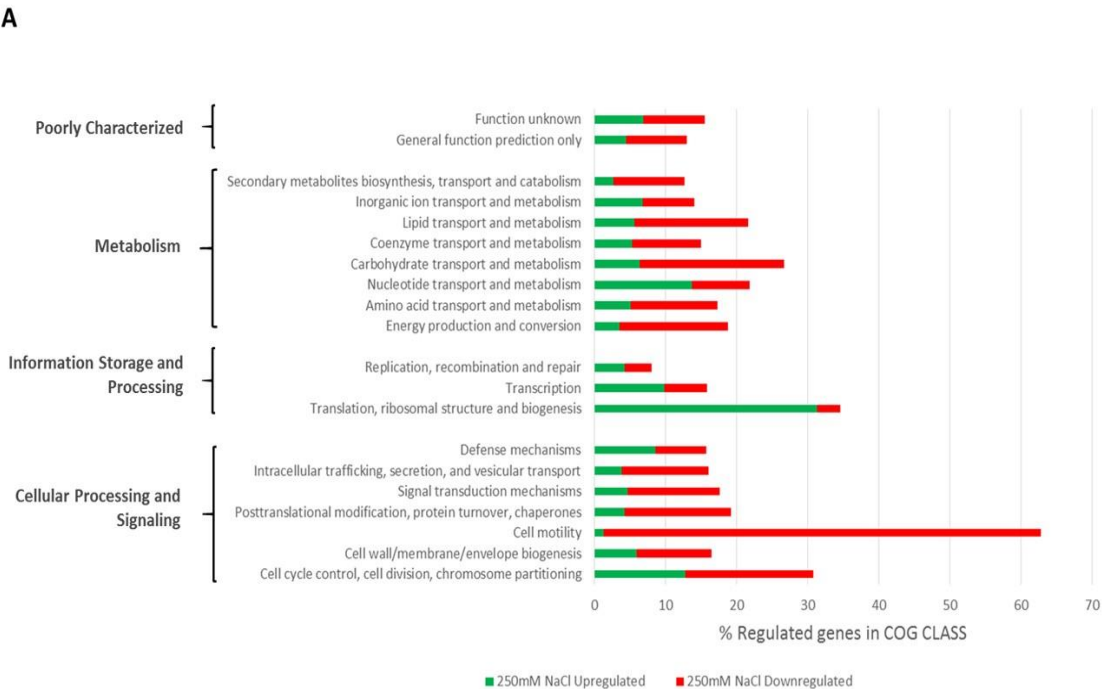
and obtain statistics on differentially regulated genes without multiple replicates, DESeq can be used [31]. However, given the strong bias obtained from the PEG8000 library that generates incomparable expression levels for the genes that are the least transcribed, we could not use this package, and lack statistics from the RNAseq data. In order to normalize for the bias related to lowly expressed genes' enrichment in the PEG8000 library and for the sequencing depth differences in the three libraries, we performed locally estimated scatter plot smoothing (LOESS) using a window of 0.1 (Log read counts) for both comparisons, Control vs. Salt or PEG8000. This normalization step enabled us to retrieve expected read counts from all three libraries and to obtain the relative abundances of each gene in the three libraries. As expected, distributions of normalized data presented more centered distributions (Figure S2F–J, Figure S3D–F) even though a slight bias for the least transcribed genes remained visible for genes whose coverage was below 30X mean coverage (Figure S3E). To support the RNAseq data, a selection of genes was analyzed by RT-qPCR using fresh bacterial cultures.

The relative expression of 12 genes showing contrasting expressions levels upon salt stress was monitored using *hcrA* as normalizer by qPCR assays on samples obtained from three independent samples. This analysis indicated a good correlation with RNAseq data ($R^2 = 0.917$) (Figure S4), which confirms that RNAseq data can be used to explore the major changes associated with stress under study, even though qPCR data indicated lower differential expressions as compared to RNAseq, the latter representing a direct measure of transcript levels.

To explore functions associated to salt- and water-induced hyperosmotic stresses, predicted genes showing 50X of mean coverage between two compared treatments (Salt vs. Control or PEG8000 vs. Control) and presenting fold changes of 2 or greater were first selected. Analyses of COGs and KEGG orthologous assignation of differentially regulated genes enabled us to estimate major metabolic and cellular functional changes associated to stresses. Table S1 shows the entire list of LMR001-predicted encoding genes (PEGs), their predicted annotations (RAST-SEED based, COG classification, and KEGG orthology), and RNAseq data.

Globally, we found that salt stress induced a stronger transcriptional response as compared to PEG8000. Addition of salt in the medium triggered an induction of 473 PEGs and a repression of 708 PEGs, of which 331 and 549, respectively, were classified into at least one COG class. In the PEG8000 treatment, only 75 PEGs presented more than 50 x mean coverage and a 2-fold induction, and 237 were repressed, of which, 67 and 122, respectively, were assigned to COG classes. This difference may result from a stronger stress induced by the short-term exposure to salt, as shown by the higher impact salt had on delta growth after 2 ½ h of incubation. It could also possibly be due to distinct global responses induced by the two solutes used to induce osmotic upshifts already shown to be partially conserved in *Sphingomonas*[32].

The distribution of COGs recovered from regulated genes in relation to the total number of COGs in each class is shown in Figure 1. Besides an upregulation of approximately 30% of genes involved in translation ribosomal structure and biogenesis encoding, for example, ribosomal proteins, we found that a majority of genes involved in chemotaxis and motility were repressed by both stresses. Indeed, nearly 60% of genes with COG assigned to that class were repressed upon stress.



B

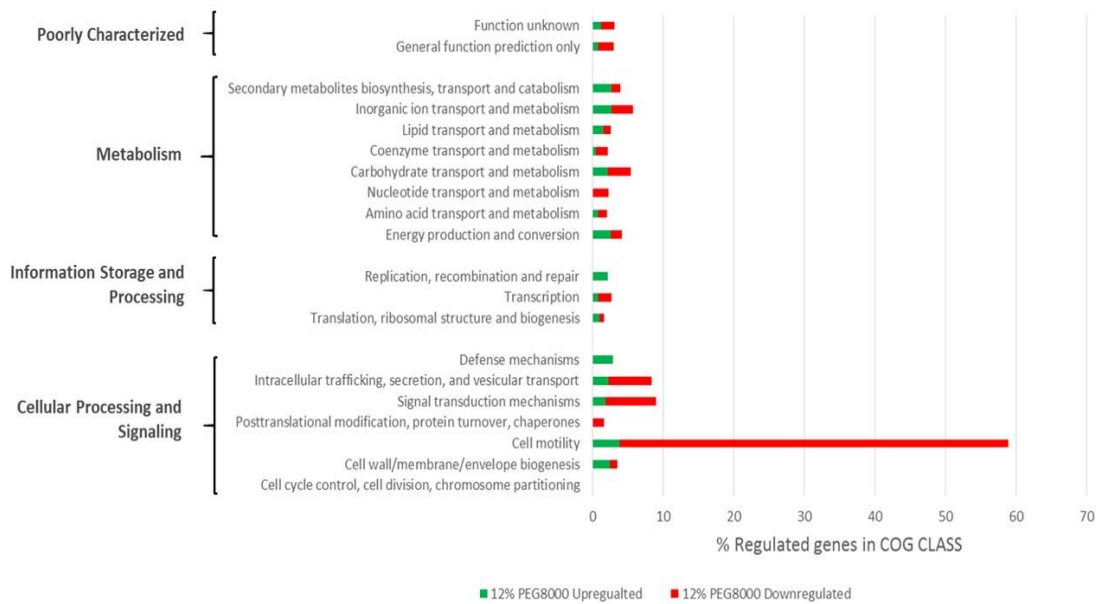


Figure 1. COG-based functional classification of genes regulated (>50X mean coverage and a minimum of 2-fold change) upon salt (A) or PEG8000 (B) treatments. The horizontal bars represent the % of genes regulated in the corresponding COG Class and, for each bar, the relative proportion of genes that are up- or down-regulated is shown in green and red, respectively. The COG class description and processes to which they belong are shown on the left of graphs.

Hyperosmotic stresses have already been shown to repress motility gene transcription in many bacteria [32–36] including rhizobia [37]. Interestingly the repression of motility is often associated to an induction of genes involved in the synthesis of biofilm components such as surface polysaccharides or proteins [32,35,38,39].

The regulation of motility genes and surface components have been studied in *Ensifer meliloti* and several transcriptional regulators have been characterized. Regulators such as CbrA, VisN/VisR, and Rem alter the regulation of many genes involved in motility or surface components [40,41], which were shown to be modulated by the ExpR/SinI quorum-sensing system [42], emphasizing the complexity of the regulatory circuits involved. Fine-tuning regulation of motility and cell surface components is of particular importance for bacteria, notably, host-associated ones such as rhizobia that have to adapt to contrasted environments (soil-rhizosphere-nodule) [43]. Among genes

that were repressed by the osmotic upshift, VisR/VisN- and Rem-predicted encoding genes (respectively, PEG655/PEG654 and PEG688) were identified. This may explain the downstream down-regulation of motility genes identified here; even though the repression was less marked under PEG80000 condition; with notably *visR* and *visN* that were less than 2-fold repressed as compared to the control (see Table S1). Even though the design did not enable us to obtain statistical significance on the RNAseq dataset, we also found several genes involved in exopolysaccharides to be induced upon salt stress (*exoQ*, *exoF*, *exoY*, *exoH*, *exoO*, and *exoN*, respectively, PEG5904-5906, PEG5914, PEG5919, and PEG5920), supporting a possible developmental switch, as shown in *S. meliloti*[40].

Among the metabolic pathways known to respond to hyperosmotic stresses, we found that two trehalose denovo biosynthesis pathways were induced. The genes encoding for Alpha,alpha-trehalose-phosphate synthase (*otsA*, PEG4869) and the Trehalose-6-phosphate phosphatase (*otsB*, PEG4868) that converts UDP glucose and glucose 6P into trehalose by two successive reactions were induced more strongly under salt stress (fold changes of 3.7 and 5.2, respectively) than upon PEG8000-induced hyperosmotic stresses (fold changes of 2 and 1.6, respectively). The second pathway that converts maltodextrin to trehalose by a Malto-oligosyltrehalose synthase (*treY*, PEG5229) and a Malto-oligosyltrehalose trehalohydrolase (*treZ*, PEG5323) appeared, in contrast, more induced under PEG treatment with respective foldchanges of 4.9 and 4.3 when these were induced 1.8 and 3.4 times by salt. We also found a strong repression of the trehalose ABC-type import system *thuEFGK* (PEG6268, 6266, 6265, and 6264) and the trehalose utilization genes *thuAB* (PEG6263 and 6262), which showed from 9- to 20-fold reduction in expression levels upon salt exposure. These genes were also repressed upon PEG8000 exposure but to a lesser extent (fold changes ranging from 1.5 to 2.7). The induction of trehalose de novo synthesis upon osmotic upshift has been documented in many bacteria, including rhizobia [12,44–46]. *Ensifer aridi* most probably also uses this common non-reducing disaccharide upon osmotic upshift. However, in *Ensifer aridi* LMR001, we found that the ABC-type transport system encoded by *thuEFGK* as well as trehalose catabolic genes *thuAB*[47,48]was strongly repressed, while *thuE*, *thuG*, and *thuA*

genes were found induced upon osmotic upshift in *Ensifer meliloti*[37]. This difference may result from different growth media used and, notably, the presence of trehalose in the medium that appears to induce both the *thuEFGK* and *thuAB* genes [47,48]. Interestingly, a repression of the trehalose utilization genes combined with an upregulation of trehalose denovo synthesis should result in increased intracellular trehalose levels that may favor rapid osmo-adaptation to stressful conditions. Interestingly, *S. meliloti* mutants in the *thuAB* genes improved the efficiency of nodulation and, more precisely, the early root hair IT-infection competitiveness [48], which may, however, depend on the host genotype [49]. Symbiotic bacteria that are prone to trehalose biosynthesis such as *E. aridi* may provide advantageous adaptation to plants under abiotic stresses [13].

Among metabolisms that were altered by the salt stress, we found a large locus (PEG2271—PEG2284) shown to be involved in erythritol, adonitol, and L-arabitol transport and catabolism [50]. Predicted genes *eryA* (PEG2278) and *eryB* (PEG2275) showed reduced transcription (respectively, 6.7- and 3.7-fold repression) upon salt exposure, while *eryC* (PEG2271) presented a 2.2-fold repression but a 30X mean coverage. The operon predicted to encode erythritol ABC-type transport system (PEG2283-2279) located just upstream of the catabolic genes was even more strongly repressed by salt (9.5–24-fold repression) when the predicted erythritol transcriptional repressor gene *eryD* (PEG2284) presented a 10-fold repression, which was surprising, as we would expect an induction. Erythritol transport and catabolism was found to be important for nodule occupancy in *Rhizobium leguminosarum*[51] but not in *Sinorhizobium meliloti*, as shown by competitiveness the Triose Phosphate isomerase *tpiB* mutant that cannot utilize erythritol [52]. Given that the erythritol catabolic genes are induced by the presence of erythritol, the repression observed could be due to an increased and rapid utilization of this polyol under salt stress, leading to reduced external erythritol concentration that may also explain the strong repression observed for *thuEFGK* and *thuAB* genes. Other possible explanations may be the involvement of additional transcriptional regulators involved, such as a putative EryR homolog (PEG2270) that was, however, not regulated in the

present study or, finally, a possible alternative utilization of an intermediary product of the erythritol catabolism suggested to be involved in its regulation [53].

In contrast to erythritol, inositol transport and metabolism were strongly induced by salt. First a myo-inositol ABC transport system (PEG5305-5307) showed 4.8- to 8.1-fold induction upon salt exposure, as compared to the control treatment. Furthermore, genes encoding 2 myo inositol-2 dehydrogenases *idhA* (PEG2503) and *iolY* (PEG375) that covert, respectively, D-chiro inositol and scyllo-inositol into 2-keto myo-inositol presented salt inductions (respectively, 6.7- and 1.8-fold changes). The genes *iolE* (PEG371), *iolD* (PEG372), *iolB* (PEG370), and *iolC* (PEG373) that are involved in the successive conversion of 2-keto-myo-inositol into 3-D-trihydroxycyclohexane-1,2-dione, 5-deoxy-glucuronic acid, 2-deoxy-5-keto-D-gluconic acid, and 2-deoxy-5-keto-D-gluconic acid 6-phosphate (DKGP) showed 3.6- to 5.6-fold inductions in the salt treatment, as compared to control. These results support active inositol import and catabolism in salt-treated cells that are under the control of the predicted transcriptional repressor IolR (PEG374). Again, *iolR* was surprisingly also up-regulated by salt with a 2.2-fold induction. It was shown that 2 deoxy 5 keto D gluconic acid 6 P (DKGP), a key product of inositol catabolism, participates in the regulation of many genes involved in this metabolism by interacting with the repressor IolR through competitive binding [54,55], suggesting accumulation of this intermediate product. Interestingly, mutations in genes involved the catabolism of myo-inositol, such as *idhA*, were shown to negatively impact symbiosis development and, particularly, competitiveness for nodule occupancy with their host plants, as shown in *S. meliloti* [56], *S. fredii* [57], and *Rhizobium leguminosarum* bv. *viciae* [58], which suggest involvement of inositol derivatives for proper symbiosis establishment and functioning. The induction of inositol catabolism, together with potential trehalose accumulation, as evidenced by trehalose metabolic and transport genes' expression levels found here upon salt stress together with their respective involvement in symbiosis establishment, further supports a role of such stress-responsive genes in improving symbiosis establishment and functioning [59,60].

We also found an induction by salt stress exposure of genes involved in the méthionine cycle. It included the adenosyl homocysteinase *ahcY* (PEG3356, 11.4-fold induction) that converts S-adenosyl-L-homocysteine (SAH) into homocysteine that is, in turn, converted to methionine, either from glycine betaine by a betaine--homocysteine S-methyltransferase BHMT (PEG1969, 18.5-fold induction) or from a sequential conversion of 5, 10-methylene tetrahydrofolate by a reductase encoded by *metF* (PEG2238, 4.6-fold induction) to methyl tetrahydrofolate that is converted with homocysteine as substrate into méthionine and tetrahydrofolate with a 5-methyltetrahydrofolate-homocysteine methyltransferase by *metH* (PEG3206, 6.7-fold induction). Finally, we found a strong induction (20.5-fold induction by salt) of the S-adenosylméthionine (SAM) synthetase-encoding gene *metK* (PEG438) that converts méthionine into SAM, a universal methyl donor used as cofactor in methyl transferase reactions involved in methylation of nucleic acids and protein or in secondary metabolite biosynthesis such as polyamines. Furthermore, two transcriptional factors previously shown to regulate the le cycle de la méthionine were also found differentially regulated. The LuxR-type solo *nesR* gene (PEG5735), shown to affect the le cycle de la méthionine by Patankar and Gonzalez [61], presented a 2-fold induction upon salt stress, while the *sahR* gene encoding an ArsR-type transcriptional regulator was found 9 times more expressed in the salt condition. SahR was shown to activate several genes of the méthionine cycle, including itself upon binding to a SAH, which releases its binding from their promoter, thus enabling their transcriptional activation [62]. An active le cycle de la méthionine was shown to be important for osmoprotection in several rhizobia including *S. fredii* [63] or in *S. meliloti* [19,61]. Interestingly, a mutation in *metH*, part of the main pathway for méthionine synthesis in *S. meliloti* strain 102F34, or *nesR* in Rm2011 reduced competitiveness for nodule occupancy when the mutants were challenged to the wild type [19,61]. Such observation was also observed in a *metZ* mutant of *R. etli* impaired for méthionine biosynthesis when inoculated to *P. vulgaris* [64]. Furthermore, a *cobO* mutant, incapable to biosynthesize the vitamin B12 necessary for MetH-mediated méthionine biosynthesis in *S. fredii* HH103, was severely affected in nodulation.

Altogether, these data support a role of the methionine nodulation competitiveness in the rhizobium legume symbiosis and, again, further suggest a possible improvement of symbiosis performance by rhizobial stress-response genes [60].

Salt stress also induced the transcription of the alternative extracytoplasmic function sigma factor *rpoE2* (PEG2539), as well as the anti-sigma factor gene *rsiA1*, a *nepR* homolog (PEG2540), and *rsiB1*, a *phyR* homolog (PEG2541), that were at least induced by 3 folds. This result indicates that the General Stress Response (GSR) was activated in *Ensifer aridi*, as already shown in other alphaproteobacteria [65]. The effects of mutations in these genes have been studied in various organisms including rhizobial genera, such as *Bradyrhizobium* [66,67], *Sinorhizobium* [68,69], or *Rhizobium* [70], and have shown their involvement as master regulators of the stress response, which may also be found in *Ensifer aridi*. Other regulatory systems have been evidenced, such as the *lovK-LovR* Two-Component System in *Caulobacter* [71] or in *S. meliloti*, in which homologs of the *rsiA1* and *rsiB1* (anti-sigma and anti-anti-sigma factors, respectively) have recently been shown to modulate *rpoE2* regulation [68]. This shows that the stress response is complex and finely tuned, which is not surprising, as bacteria have to cope with different but sometimes combined stresses, which may utilize common molecular mechanisms depending on the stress. It is interesting to note that we also found homologs of *rsiA1* and *rsiB1*, named, respectively, *rsiA2* (PEG753) and *rsiB2* (PEG752), in our strain, which may act similarly as shown for *S. meliloti*. However, our transcriptomic data showed that the stresses applied here did not affect their transcriptional status, which remained unchanged. Additional studies are required to decipher their regulatory functions. Surprisingly, while bradyrhizobia mutated in genes responsible for the alternative extracytoplasmic function, sigma factor functioning showed strong symbiosis development deficiencies [66,67], symbiosis development and functioning was not impaired by such mutations in *Sinorhizobium meliloti* [69,72] or *Rhizobium etli* [70]. Interestingly, Lang and colleagues [72] demonstrated that the absence of such symbiosis deficiency in *S. meliloti* was not due to extracytoplasmic function sigma factors (*ecfG*) redundancy. They obtained similar symbiosis phenotypes when all *ecfG* were mutated and compared to the wild-type strain.

Regarding the responses of *ecfG* mutants to stresses, studies on *S. meliloti* showed contrasting results. In a primary study, Sauviac and colleagues [69] showed that *rpoE2* mutants did not present phenotypic differences with the wild type upon various stresses (heat, salt, acidic pH, or H₂O₂) in both exponential and stationary growth phases. Later, Flechard and colleagues [45,73] showed that a *rpoE2* mutant in a stationary growth phase was sensitive to hydrogen peroxide and hyperosmotic stresses. More recently, Lang and colleagues confirmed Sauviac's results as no or subtle differences were found in the *rpoE2* mutant or a mutant lacking all ECF-like genes, which is surprising, as many stress-responsive genes were identified in the RpoE2 regulon [69,72]. These differences may result from distinct mutation procedures and/or media used in these studies, which included, for example, a deletion of the anti-sigma factors in the study by Lang and colleagues [72]. In *Bradyrhizobium*, such mutations increased sensitivity of *B. japonicum* to heat shock and desiccation [66]. Furthermore, recent work on *B. diazoefficiens* showed that *ecfG* mutation resulted in increased sensitivity to various stresses including hyperosmotic and ionic stresses [74]. More interestingly, in both *Bradyrhizobium diazoefficiens* and *Sinorhizobium meliloti*, a *rpoE2* mutation was shown to repress trehalose 6 Phosphate pathway, which was involved in hyperosmotic stresses' tolerance [45] and was sufficient to impair early symbiotic development in the more sensitive *B. diazoefficiens* species [74], further supporting that trehalose is a key compatible solute for *Bradyrhizobium* stress adaptation and nodulation, as shown in *B. japonicum* [46]. The absence of a symbiotic phenotype observed in the *rpoE2* mutant of *S. meliloti* may result from a high tolerance to hyperosmotic stress in this species as compared to the *Bradyrhizobium* species. Such stress may be encountered, notably, in infection threads, which is supported by the restoration of symbiosis by addition of exogenous trehalose in *B. diazoefficiens* Δ *otsA* strain still capable to import it or in a strain capable to endogenously synthesize other compatible osmoprotectants [74]. To explore the function of RpoE2 and its possible role in the regulation of several functions shown to be transcriptionally regulated by salt stress as well as in the tolerance of the LMR001 bacterium to various

stresses and symbiosis development, a deletion mutant of the *rpoE2* gene ($\Delta rpoE2$) was generated and phenotypically characterized.

First, the growth of $\Delta rpoE2$ was compared to that of LMR001 in TY medium complemented with salt at 250, 300, and 400 mM final concentrations (Figure 2A). We found that, even though the WT strain reached the stationary phase earlier in TY medium containing 250 mM of NaCl as compared to $\Delta rpoE2$, both strains reached comparable optical densities after 120 h of incubation. Secondly, when we looked at the motility of the two strains, while we found a significant reduction of motility for the two strains upon addition of salt at both concentrations tested (0.1 and 0.25 M), which corroborates transcriptional data observed by RNAseq, no clear differences were observed between the two strains (Figure 2B). Sensitivities of the two strains to detergent, oxidative, and acid stresses were estimated by the inhibitory growth diameter induced by the diffusion in the agar medium of 5 μ L concentrated SDS, H₂O₂, or HCl that were spotted on small paper discs placed in the center of soft agar plates previously inoculated with the strains (see Materials and Methods). The concentration of these molecules decreased as they diffused away from the discs imbibed with the reagents. A measure of the growth inhibition diameter, thus, enabled us to compare tolerance of the strains. We did not find any difference between the wild-type and the $\Delta rpoE2$ strains (Figure 2C), suggesting that the *rpoE2* mutation did not alter cell membrane stability under such stresses. Finally, both the LMR001 and $\Delta rpoE2$ strains developed the symbiosis with the tree species *V. gummifera* and *V. tortilis* and there were no significant differences associated to plant weights or nodulation (Table 2). These results show that the ECF sigma factor RpoE2 is not essential for motility, symbiosis development, and tolerance to the stresses tested, suggesting that other functions and/or regulators are involved in these processes, as suggested in *Sinorhizobium meliloti* [69,72].

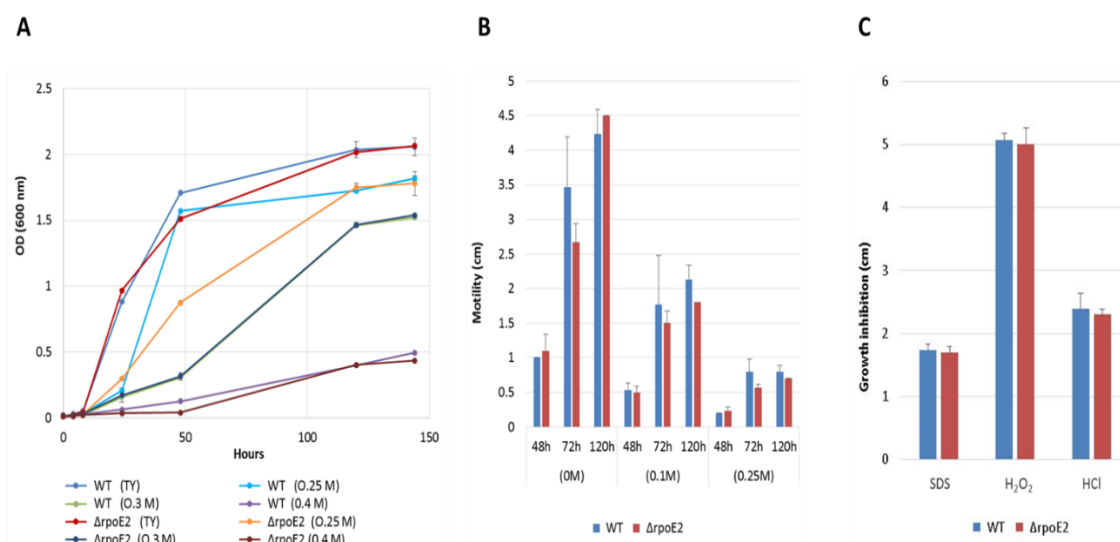


Figure 2. Effect of the *rpoE2* mutation on stress tolerance and motility of LMR001. *Ensifer aridi* LMR001 (WT) and the LMR001 *rpoE2* deletion mutant ($\Delta rpoE2$) growths were compared in liquid TY medium complemented with 0, 0.1, 0.25, 0.3, or 0.4 M of NaCl (A). Their motility was assessed in semi-solid TY complemented with salt at 0, 0.1, or 0.25 M by measurement of the diameter of the swimming colonies after 2, 3, or 5 days of incubation (B). Finally, the diameters of the growth inhibition zones were compared for the two strains inoculated in top agar plates, to the center of which patches imbued with 5 μ L of SDS (10%w/v), H₂O₂ (5.5 M), or HCl (2 M) were placed (C).

Table 2. Effect of the *rpoE2* mutation on symbiosis development with *Vachellia*.

Plant/Treatment	Shoot Weight (g)		Root Weight (g)		Nodule Number	
	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
<i>Vachellia gummifera</i>						
LMR001 WT	2.27	0.21	0.90	0.09	33	3.22
LMR001 $\Delta rpoE2$	2.33	0.37	1.03	0.19	35.67	5.24
Positive Control	3.13	0.10	2.83	0.26	0	
Negative Control	1.37	0.14	1.07	0.10	0	
<i>Vachellia tortilis</i>						
LMR001 WT	2.72	0.09	0.83	0.05	39	1.79
LMR001 $\Delta rpoE2$	2.16	0.21	1.08	0.15	38	3.94
Positive Control	3.73	0.58	2.33	0.26	0	
Negative Control	1.83	0.14	1.10	0.09	0	

We further explored the involvement of the extra cytoplasmic function sigma factor in the transcriptional regulation of genes involved in the inositol catabolism (*iolC* and *iolR*), le cycle de la méthionineregulation (*nesR* and *sahR*), trehalose transport, and metabolism (*thuR*, *thuE*, *otsB*, and *treZ*) including a glycosyl transferase (PEG4866) located upstream of the *otsBA* operon in reverse orientation or in the regulation of the

GSR (*rsiA1*, *rsiB1*). Their promoter regions were fused in front of the reporter gene encoding GFP in a mobilizable vector, and resulting constructions were further transferred into LMR001 and $\Delta rpoE2$. Both strains containing the pPROBE derivatives were then grown either in TY medium or in TY complemented with NaCl at a final concentration of 250 mM in 96-microtiter plates and fluorescence with the optical density (600 nm) and were recorded every half hour.

Figure 3 shows the relative expression levels of the GFP under the control of selected promoters recorded every 30 min for 9 h. The relative expression of the GFP varied with the promoter and, for some of them, with the strain. The approach did not allow us to detect differences between salt and control growth conditions for a strain that would require less stable GFP and a more sensitive approach to detect early response to hyperosmotic stress. However, the methodology enabled us to identify transcriptional changes resulting from the mutation. We found that the mutation resulted in an increased transcription of four genes, while it reduced the expression of three genes. An increased transcription of *otsB* was obtained in the $\Delta rpoE2$ strain, indicating that RpoE2-mediated GSR in *E. aridi* involves the T6P pathway, as shown in *S. meliloti*[45] or in *B. diazoefficiens*[74]. We also found that *forrsiA1*, the anti-sigma factor that is located upstream of and overlaps the 5' end of the *rpoE2* gene forming an operon was less expressed in the mutant, which suggests, again, that, as shown in other alphaproteobacteria [65], *rpoE2* is autoregulated in *E. aridi*. Given the high similarities, both in terms of syntheny and sequence between *S. meliloti* and *E. aridi* for many genes discussed here (for example, 97.8% similarity for RpoE2), we tested whether the same promoter motif GGAACN₁₆₋₁₇CGTT that likely represents the -35 and -10 elements of the RpoE2 responsive genes [69,72] could be found in the target promoter sequences of *E. aridi* genes studied here. This motif was found upstream of *rsiA1* and *rsiB1*, supporting a conserved RpoE2-mediated regulatory mechanism. It was also found twice in the promoter region of the *treZ* gene, whose transcription was, however, not found significantly different when comparing the WT and $\Delta rpoE2$ strains, and in the range of background values. Regarding the *otsB* promoter region, we found two motifs that were

close to the consensus sequence: (GGAACN₁₇CGAT and AGAACN₁₇CGAT). Finally, one motif with three nucleotide differences as compared to the consensus sequence was found in front of the *thuR* gene (GGAAAN₁₇AGAT). However, the involvement of these regions requires additional work to assess whether these are directly regulated by RpoE2 and to exclude the possibility that a hidden mutation resulting from the mutagenesis impacted our data. Surprisingly, the most severe transcriptional repression resulting from the mutation was found for *thuR*, encoding a HTH DNA-Binding LacI-type repressor showing homologies to *aglR* that may be involved in transport and utilization of trehalose through *thuEFGK* and *thuAB* expression [47,48]. It is interesting to note the presence of *thuR* consensus sequence motifs found among rhizobiales upstream of these operons in LMR001 as well as in front of *aglR* and *aglEFGAK* that share partially their motifs [75], which may, however, explain cross-regulation of these transport systems by *aglR*[47]. Given (1) ThuR homology to AglR, (2) the trehalose transport specificity observed of ThuE, ThuF, ThuG, and ThuK, shown by Jensen and colleagues, and (3) evolutionary and comparative analyses within rhizobiales of LacI-family transcription factors that predicts ThuR as a regulator of two operons (*thuEFGKAB* involved in trehalose transport and utilization as well as *thuR*) [47,75], our data further support that ThuR is also autoregulated and involved in the regulation of transport and utilization of trehalose in *E. aridi*. Nevertheless, additional work is required to characterize the regulatory pathways involved and, notably, whether *thuR* is directly or indirectly regulated by RpoE2

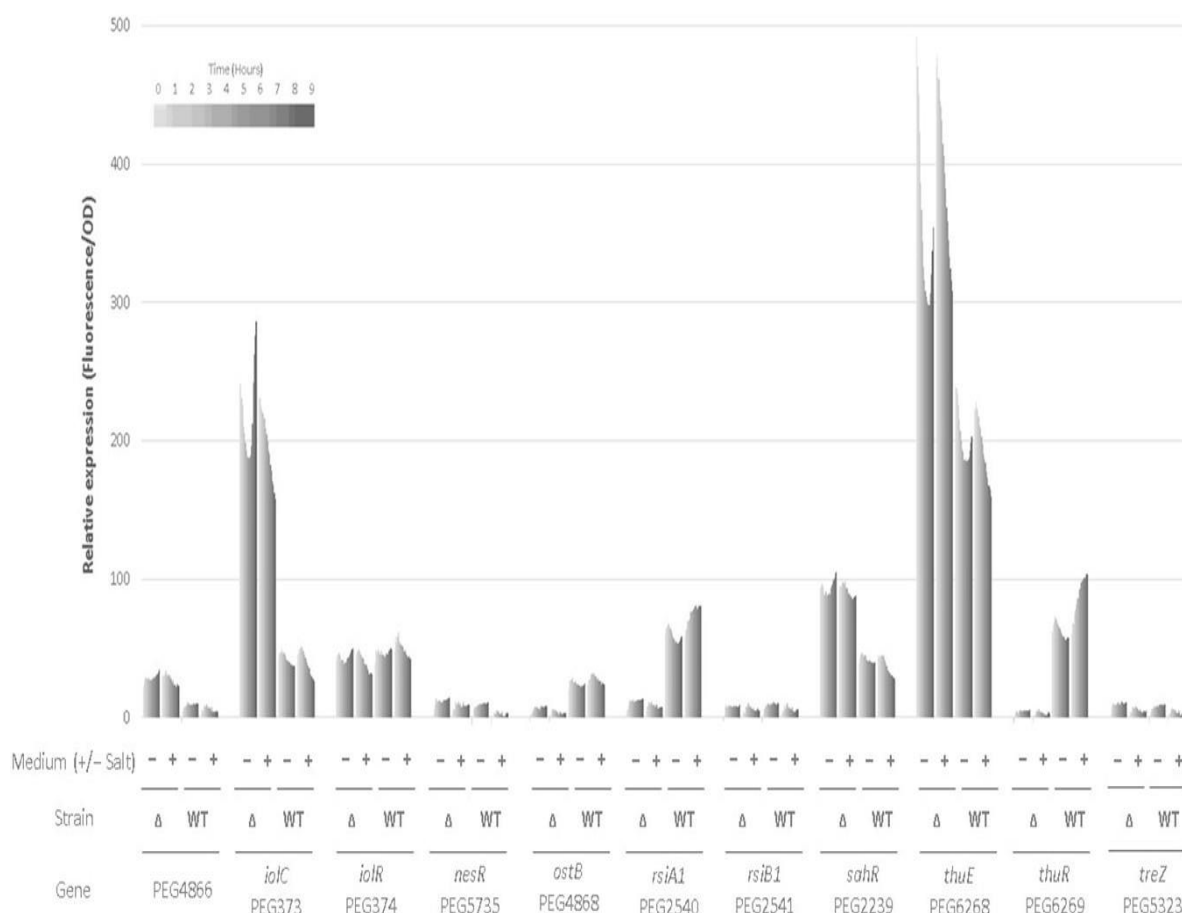


Figure 3. Effect of the *rpoE2* mutation on transcriptional activity of selected genes. The histograms show the relative expression (see Materials and Methods) of *gfp* under the control of promoters from selected genes mobilized in wild-type *Ensifer aridi* LMR001 (WT) or the $\Delta rpoE2$ derivative (Δ) strains grown in TY medium containing salt (NaCl at 250 mM) or not (+/-), as indicated at the bottom of the figure. Relative expressions were recorded every 30 min for 9 h and are shown by gray bars of increasing darkness with time (colors indicated on the top left of the figure).

Among genes whose transcriptional activities were increased by the mutation, we found *thuE*, which may result from the *thuR* repression in the $\Delta rpoE2$ strain. Interestingly, the glycosyl transferase (PEG4866) located upstream of the *otsBA* operon in reverse orientation also showed an increased transcriptional activity in $\Delta rpoE2$, which may suggest RpoE2-mediated repression in the wild-type strain. Surprisingly, *iolC* and *sahR* genes involved, respectively, in the catabolism of myo-inositol and the regulation of the le cycle de la méthionine were also found more expressed in $\Delta rpoE2$ background while they were found induced upon salt stress in the wild-type LMR001 strain. Whether these

inductions result from a direct RpoE2-mediated repression remains to be addressed, but our data suggest a complex stress response in *E. aridi*, which involves RpoE2 that can, nevertheless, be complemented.

4. Conclusions

To explore the mechanisms by which *Ensifer aridi* LMR001 copes with hyperosmotic stresses, 31 Gb of sequence data were generated by three RNAseq libraries constructed from cells grown either in control medium or upon addition of the permeating solute NaCl at 250 mM or the non-permeating solute PEG8000 at 12%*w/v*. As shown in other bacteria, hyperosmotic stresses resulted in the repression of motility genes, which was confirmed by phenotypic tests. These effects may be considered for improving the rhizobium–legume symbiosis in soils containing salt, notably, through inoculation of adapted bacteria to ensure contact between non-motile rhizobia and roots. We also found that the extra-cytoplasmic function sigma factor-encoding gene *rpoE2* was induced and probably involved, as shown in other rhizobia, in its autoregulation and the induction of trehalose denovo biosynthesis. A concomitant repression of trehalose utilization genes *thuAB* suggests intracellular accumulation of trehalose in stressed cells. Salt also induced inositol catabolism and the méthionine cycle, suggesting a possible role of inositol derivatives and possibly methylation in the early response to such stress. In addition, a combination of directed mutagenesis of *rpoE2* and the use of promoter-*gfp* fusion reporter system suggested that RpoE2 regulates positively *thuR*, a LacI-type regulator probably involved in trehalose import and utilization. In contrast, the mutation resulted in increased expression of genes involved in inositol catabolism or the regulation of the cycle de la méthionine that, given their induction in the wild-type bacterium upon salt stress, supports the involvement of alternative regulatory circuits. However, additional experiments are required to ascertain that $\Delta rpoE2$ does not carry hidden mutations, which could have impacted our data. Intriguingly, previous studies showed that a mutation of trehalose utilization genes in *S. meliloti* resulted in improved competitiveness for nodulation, which is also favored with active inositol catabolism and

méthioninecycle, suggesting that hyperosmotic stress occurs in the early infection of leguminous hosts. The *rpoE2* deletion did not result in increased stress sensitivity or symbiosis functioning, suggesting the involvement of alternative regulatory circuits in *Ensifer aridi* for stress adaptation.

Supplementary Materials: The following are available online at www.mdpi.com/xxx/s1. Figure S1: Sequencing and mapping of the three RNAseq libraries. Figure S2: Distributions of RNAseq data. Figure S3: Effect of PEG order, mean coverage, or size on the relative abundance of reads obtained in the three libraries. Figure S4: Comparison of the relative expression levels of selected genes by RNAseq and qPCR. Table S1: Predicted LMR001-encoding gene annotation and mapping.

Author Contributions: Conceptualization, A.L.Q. and A.F.-M.; methodology, A.L.Q., M.B., E.D., D.S., and H.A.-B.; formal analysis, M.B. and A.L.Q.; investigation, M.B., A.L.Q., and I.S.; resources, E.D. and D.S.; data curation, E.D., D.S., and A.L.Q.; writing—original draft preparation, M.B. and A.L.Q.; writing—review and editing, all authors.; visualization, A.L.Q. and M.B.; supervision, A.L.Q. and A.F.-M.; project administration, A.L.Q. and A.F.-M.; funding acquisition, A.L.Q. and A.F.-M. All authors have read and agreed to the published version of the manuscript

Funding: This research was funded by the French Institute of Research for the Development (IRD) through the LMI LBMV and an ARTS PhD grant for M.B.

Institutional Review Board Statement: Not applicable

Informed Consent Statement: Not applicable

Data Availability Statement: *Ensifer aridi* LMR001 genome browser, Fasta nucleotide and protein sequences and annotations can be accessed through the RAST server <https://rast.nmpdr.org/> (accessed November 11, 2020) using “guest” as login and password under job nb “188440” with ID “6666666.88372”. RNAseq data have been deposited in SRA in NCBI and can be accessed through the Bioproject PRJNA312299.

Acknowledgments: The authors would like to acknowledge Samir El Qaidi for methodological support for the construction of the mutant and Abderrahim Ferradous from the Centre Régional de Recherches Forestière in Marrakech for *Vachellia* seeds. M.G.X. acknowledges financial support from France Génomique National infrastructure, funded as part of “Investissement d’Avenir” program managed by Agence Nationale pour la Recherche (contract ANR-10-INBS-09)

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Sakrouhi, I.; Belfquih, M.; Sbabou, L.; Moulin, P.; Bena, G.; Filali-Maltouf, A.; Le Quéré, A. Recovery of Symbiotic Nitrogen Fixing Acacia Rhizobia from Merzouga Desert Sand Dunes in South East Morocco—Identification of a Probable New Species of *Ensifer* Adapted to Stressed Environments. *Syst. Appl. Microbiol.* **2016**, *39*, 122–131, doi:10.1016/j.syapm.2016.01.001.
2. Rocha, G.; Le Queré, A.; Medina, A.; Cuéllar, A.; Contreras, J.-L.; Carreño, R.; Bustillos, R.; Muñoz-Rojas, J.; Villegas, M. del C.; Chaintreuil, C.; et al., Diversity and Phenotypic Analyses of Salt- and Heat-Tolerant Wild Bean *Phaseolus filiformis* Rhizobia Native of a Sand Beach in Baja California and Description of *Ensifer aridi* Sp. Nov. *Arch. Microbiol.* **2020**, *202*, 309–322, doi:10.1007/s00203-019-01744-7.
3. Tak, N.; Awasthi, E.; Bissa, G.; Meghwal, R.R.; James, E.K.; Sprent, J.S.; Gehlot, H.S. Multi Locus Sequence Analysis and Symbiotic Characterization of Novel *Ensifer* Strains Nodulating *Tephrosia* Spp. in the Indian Thar Desert. *Syst. Appl. Microbiol.* **2016**, *39*, 534–545, doi:10.1016/j.syapm.2016.08.002.
4. Hakim, S.; Imran, A.; Mirza, M.S. Phylogenetic Diversity Analysis Reveals *Bradyrhizobium yuanmingense* and *Ensifer aridi* as Major Symbionts of Mung Bean (*Vigna radiata* L.) in Pakistan. *Braz. J. Microbiol.* **2021**, *52*, 311–324, doi:10.1007/s42770-020-00397-9.
5. Lamin, H.; Alami, S.; Bouhnik, O.; ElFaik, S.; Abdelmoumen, H.; Bedmar, E.J.; Missbah-El Idrissi, M. Nodulation of *Retama monosperma* by *Ensifer aridi* in an Abandoned Lead Mine Soils in Eastern Morocco. *Front. Microbiol.* **2019**, *10*, 1456, doi:10.3389/fmicb.2019.01456.

6. Le Quéré,, A.; Tak, N.; Gehlot, H.S.; Lavire, C.; Meyer, T.; Chapulliot, D.; Rath, S.; Sakrouhi, I.; Rocha, G.; Rohmer, M.; et al., Genomic Characterization of *Ensifer aridi*, a Proposed New Species of Nitrogen-Fixing Rhizobium Recovered from Asian, African and American Deserts. *BMC Genom.* **2017**, *18*, 85, doi:10.1186/s12864-016-3447-y.
7. Zahran, H.H. *Rhizobium* -Legume Symbiosis and Nitrogen Fixation under Severe Conditions and in an Arid Climate. *Microbiol. Mol. Biol. Rev.***1999**, *63*, 968–989, doi:10.1128/MMBR.63.4.968-989.1999.
8. Denich, T.J.; Beaudette, L.A.; Lee, H.; Trevors, J.T. Effect of Selected Environmental and Physico-Chemical Factors on Bacterial Cytoplasmic Membranes. *J. Microbiol. Methods***2003**, *52*, 149–182, doi:10.1016/S0167-7012(02)00155-0.
9. Chang, W.-S.; van de Mortel, M.; Nielsen, L.; Nino de Guzman, G.; Li, X.; Halverson, L.J. Alginate Production by *Pseudomonas putida* Creates a Hydrated Microenvironment and Contributes to Biofilm Architecture and Stress Tolerance under Water-Limiting Conditions. *J. Bacteriol.***2007**, *189*, 8290–8299, doi:10.1128/JB.00727-07.
10. Finn, S.; Condell, O.; McClure, P.; Amézquita, A.; Fanning, S. Mechanisms of Survival, Responses and Sources of *Salmonella* in Low-Moisture Environments. *Front. Microbiol.***2013**, *4*, doi:10.3389/fmicb.2013.00331.
11. Lebre, P.H.; De Maayer, P.; Cowan, D.A. Xerotolerant Bacteria: Surviving through a Dry Spell. *Nat. Rev. Microbiol.***2017**, *15*, 285–296, doi:10.1038/nrmicro.2017.16.
12. Vriezen, J.A.C.; de Bruijn, F.J.; Nüsslein, K. Responses of Rhizobia to Desiccation in Relation to Osmotic Stress, Oxygen, and Temperature. *Appl. Environ. Microbiol.***2007**, *73*, 3451–3459, doi:10.1128/AEM.02991-06.
13. Sharma, M.P.; Grover, M.; Chourasiya, D.; Bharti, A.; Agnihotri, R.; Maheshwari, H.S.; Pareek, A.; Buyer, J.S.; Sharma, S.K.; Schütz, L.; et al., Deciphering the Role of Trehalose in Tripartite Symbiosis Among Rhizobia, Arbuscular Mycorrhizal Fungi, and Legumes for Enhancing Abiotic Stress Tolerance in Crop Plants. *Front. Microbiol.***2020**, *11*, 509919, doi:10.3389/fmicb.2020.509919.
14. Bonilla, C.Y. Generally Stressed Out Bacteria: Environmental Stress Response Mechanisms in Gram-Positive Bacteria. *Integr. Comp. Biol.***2020**, *60*, 126–133, doi:10.1093/icb/icaa002.
15. Fiebig, A.; Herrou, J.; Willett, J.; Crosson, S. General Stress Signaling in the *Alphaproteobacteria*. *Annu. Rev. Genet.***2015**, *49*, 603–625, doi:10.1146/annurev-genet-112414-054813.

16. Lori, C.; Kaczmarczyk, A.; de Jong, I.; Jenal, U. A Single-Domain Response Regulator Functions as an Integrating Hub To Coordinate General Stress Response and Development in *Alphaproteobacteria*. *mBio***2018**, *9*, doi:10.1128/mBio.00809-18.
17. Beringer, J.E. R Factor Transfer in *Rhizobium leguminosarum*. *Microbiology***1974**, *84*, 188–198, doi:10.1099/00221287-84-1-188.
18. Sambrook, J.; Fritsch, E.; Maniatis, T. *Molecular Cloning: A Laboratory Manual: Volume. 2*, 2nd ed.; Cold Spring Harbor: New York, USA. 1989; ISBN 978-0-87969-309-1.
19. Barra, L.; Fontenelle, C.; Ermel, G.; Trautwetter, A.; Walker, G.C.; Blanco, C. Interrelations between Glycine Betaine Catabolism and Méthionine Biosynthesis in *Sinorhizobium meliloti* Strain 102F34. *J. Bacteriol.***2006**, *188*, 7195–7204, doi:10.1128/JB.00208-06.
20. Broughton, W.J.; Dilworth, M.J. Control of Leghaemoglobin Synthesis in Snake Beans. *Biochem. J.***1971**, *125*, 1075–1080, doi:10.1042/bj1251075.
21. Maynaud, G.; Brunel, B.; Mornico, D.; Durot, M.; Severac, D.; Dubois, E.; Navarro, E.; Cleyet-Marel, J.-C.; Le Quéré, A. Genome-Wide Transcriptional Responses of Two Metal-Tolerant Symbiotic *Mesorhizobium* Isolates to Zinc and Cadmium Exposure. *BMC Genom.* **2013**, *14*, 292, doi:10.1186/1471-2164-14-292.
22. Kanehisa, M.; Goto, S.; Sato, Y.; Kawashima, M.; Furumichi, M.; Tanabe, M. Data, Information, Knowledge and Principle: Back to Metabolism in KEGG. *Nucl. Acids Res.***2014**, *42*, D199–D205, doi:10.1093/nar/gkt1076.
23. Tatusov, R.L. The COG Database: New Developments in Phylogenetic Classification of Proteins from Complete Genomes. *Nucl. Acids Res.***2001**, *29*, 22–28, doi:10.1093/nar/29.1.22.
24. Pfaffl, M.W. A New Mathematical Model for Relative Quantification in Real-Time RT-PCR. *Nucl. Acids Res.***2001**, *29*, e45, doi:10.1093/nar/29.9.e45.
25. Chen, W.; Kuo, T. A Simple and Rapid Method for the Preparation of Gram-Negative Bacterial Genomic DNA. *Nucl. Acids Res.***1993**, *21*, 2260, doi:10.1093/nar/21.9.2260.
26. Quandt, J.; Hynes, M.F. Versatile Suicide Vectors Which Allow Direct Selection for Gene Replacement in Gram-Negative Bacteria. *Gene***1993**, *127*, 15–21, doi:10.1016/0378-1119(93)90611-6.
27. Simon, R.; Priefer, U.; Pühler, A. A Broad Host Range Mobilization System for In Vivo Genetic Engineering: Transposon Mutagenesis in Gram Negative Bacteria. *Nat. Biotechnol.***1983**, *1*, 784–791, doi:10.1038/nbt1183-784.

28. Miller, W.G.; Leveau, J.H.J.; Lindow, S.E. Improved *Gfp* and *InaZ* Broad-Host-Range Promoter-Probe Vectors. *Mol. Plant Microbe Interact.***2000**, *13*, 1243–1250, doi:10.1094/MPMI.2000.13.11.1243.
29. Zewei An Co-Extraction of High-Quality RNA and DNA from Rubber Tree (*Hevea brasiliensis*). *Afr. J. Biotechnol.***2012**, *11*, doi:10.5897/AJB12.136.
30. Roose-Amsaleg, C.L.; Garnier-Sillam, E.; Harry, M. Extraction and Purification of Microbial DNA from Soil and Sediment Samples. *Appl. Soil Ecol.***2001**, *18*, 47–60, doi:10.1016/S0929-1393(01)00149-4.
31. Anders, S.; Huber, W. Differential Expression Analysis for Sequence Count Data. *Genome Biol.***2010**, *11*, R106, doi:10.1186/gb-2010-11-10-r106.
32. Alavi, P.; Starcher, M.R.; Zachow, C.; Müller, H.; Berg, G. Root-Microbe Systems: The Effect and Mode of Interaction of Stress Protecting Agent (SPA) *Stenotrophomonas rhizophila* DSM14405T. *Front. Plant Sci.***2013**, *4*, doi:10.3389/fpls.2013.00141.
33. Dressaire, C.; Moreira, R.N.; Barahona, S.; Alves de Matos, A.P.; Arraiano, C.M. BolA Is a Transcriptional Switch That Turns Off Motility and Turns On Biofilm Development. *mBio***2015**, *6*, doi:10.1128/mBio.02352-14.
34. Li, S.; Liang, H.; Wei, Z.; Bai, H.; Li, M.; Li, Q.; Qu, M.; Shen, X.; Wang, Y.; Zhang, L. An Osmoregulatory Mechanism Operating through OmpR and LrhA Controls the Motile-Sessile Switch in the Plant Growth-Promoting Bacterium *Pantoea alhagi*. *Appl Environ. Microbiol.***2019**, *85*, doi:10.1128/AEM.00077-19.
35. Steil, L.; Hoffmann, T.; Budde, I.; Völker, U.; Bremer, E. Genome-Wide Transcriptional Profiling Analysis of Adaptation of *Bacillus subtilis* to High Salinity. *J. Bacteriol.***2003**, *185*, 6358–6370, doi:10.1128/JB.185.21.6358-6370.2003.
36. Zhou, A.; Baidoo, E.; He, Z.; Mukhopadhyay, A.; Baumohl, J.K.; Benke, P.; Joachimiak, M.P.; Xie, M.; Song, R.; Arkin, A.P.; et al., Characterization of NaCl Tolerance in *Desulfovibrio vulgaris* Hildenborough through Experimental Evolution. *ISME J.***2013**, *7*, 1790–1802, doi:10.1038/ismej.2013.60.
37. Domínguez-Ferreras, A.; Pérez-Arnedo, R.; Becker, A.; Olivares, J.; Soto, M.J.; Sanjuán, J. Transcriptome Profiling Reveals the Importance of Plasmid PSymB for Osmoadaptation of *Sinorhizobium meliloti*. *J. Bacteriol.***2006**, *188*, 7617–7625, doi:10.1128/JB.00719-06.

38. Bahlawane, C.; McIntosh, M.; Krol, E.; Becker, A. *Sinorhizobium meliloti* Regulator MucR Couples Exopolysaccharide Synthesis and Motility. *Mol. Plant Microbe Interact.***2008**, *21*, 1498–1509, doi:10.1094/MPMI-21-11-1498.
39. Liu, Y.; Gao, W.; Wang, Y.; Wu, L.; Liu, X.; Yan, T.; Alm, E.; Arkin, A.; Thompson, D.K.; Fields, M.W.; et al., Transcriptome Analysis of *Shewanella oneidensis* MR-1 in Response to Elevated Salt Conditions. *J. Bacteriol.***2005**, *187*, 2501–2507, doi:10.1128/JB.187.7.2501-2507.2005.
40. Gibson, K.E.; Barnett, M.J.; Toman, C.J.; Long, S.R.; Walker, G.C. The Symbiosis Regulator CbrA Modulates a Complex Regulatory Network Affecting the Flagellar Apparatus and Cell Envelope Proteins. *J. Bacteriol.***2007**, *189*, 3591–3602, doi:10.1128/JB.01834-06.
41. Sourjik, V.; Muschler, P.; Scharf, B.; Schmitt, R. VisN and VisR Are Global Regulators of Chemotaxis, Flagellar, and Motility Genes in *Sinorhizobium (Rhizobium) meliloti*. *J. Bacteriol.***2000**, *182*, 782–788, doi:10.1128/JB.182.3.782-788.2000.
42. Hoang, H.H.; Gurich, N.; González, J.E. Regulation of Motility by the ExpR/Sin Quorum-Sensing System in *Sinorhizobium meliloti*. *J. Bacteriol.***2008**, *190*, 861–871, doi:10.1128/JB.01310-07.
43. Janczarek, M. Environmental Signals and Regulatory Pathways That Influence Exopolysaccharide Production in Rhizobia. *Int. J. Mol. Sci.***2011**, *12*, 7898–7933, doi:10.3390/ijms12117898.
44. Domínguez-Ferreras, A.; Soto, M.J.; Pérez-Arnedo, R.; Olivares, J.; Sanjuán, J. Importance of Trehalose Biosynthesis for *Sinorhizobium meliloti* Osmotolerance and Nodulation of Alfalfa Roots. *J. Bacteriol.***2009**, *191*, 7490–7499, doi:10.1128/JB.00725-09.
45. Flechard, M.; Fontenelle, C.; Blanco, C.; Goude, R.; Ermel, G.; Trautwetter, A. RpoE2 of *Sinorhizobium meliloti* Is Necessary for Trehalose Synthesis and Growth in Hyperosmotic Media. *Microbiology***2010**, *156*, 1708–1718, doi:10.1099/mic.0.034850-0.
46. Sugawara, M.; Cytryn, E.J.; Sadowsky, M.J. Functional Role of *Bradyrhizobium Japonicum* Trehalose Biosynthesis and Metabolism Genes during Physiological Stress and Nodulation. *Appl. Environ. Microbiol.***2010**, *76*, 1071–1081, doi:10.1128/AEM.02483-09.
47. Jensen, J.B.; Peters, N.K.; Bhuvaneshwari, T.V. Redundancy in Periplasmic Binding Protein-Dependent Transport Systems for Trehalose, Sucrose, and Maltose in *Sinorhizobium meliloti*. *J. Bacteriol.***2002**, *184*, 2978–2986, doi:10.1128/JB.184.11.2978-2986.2002.

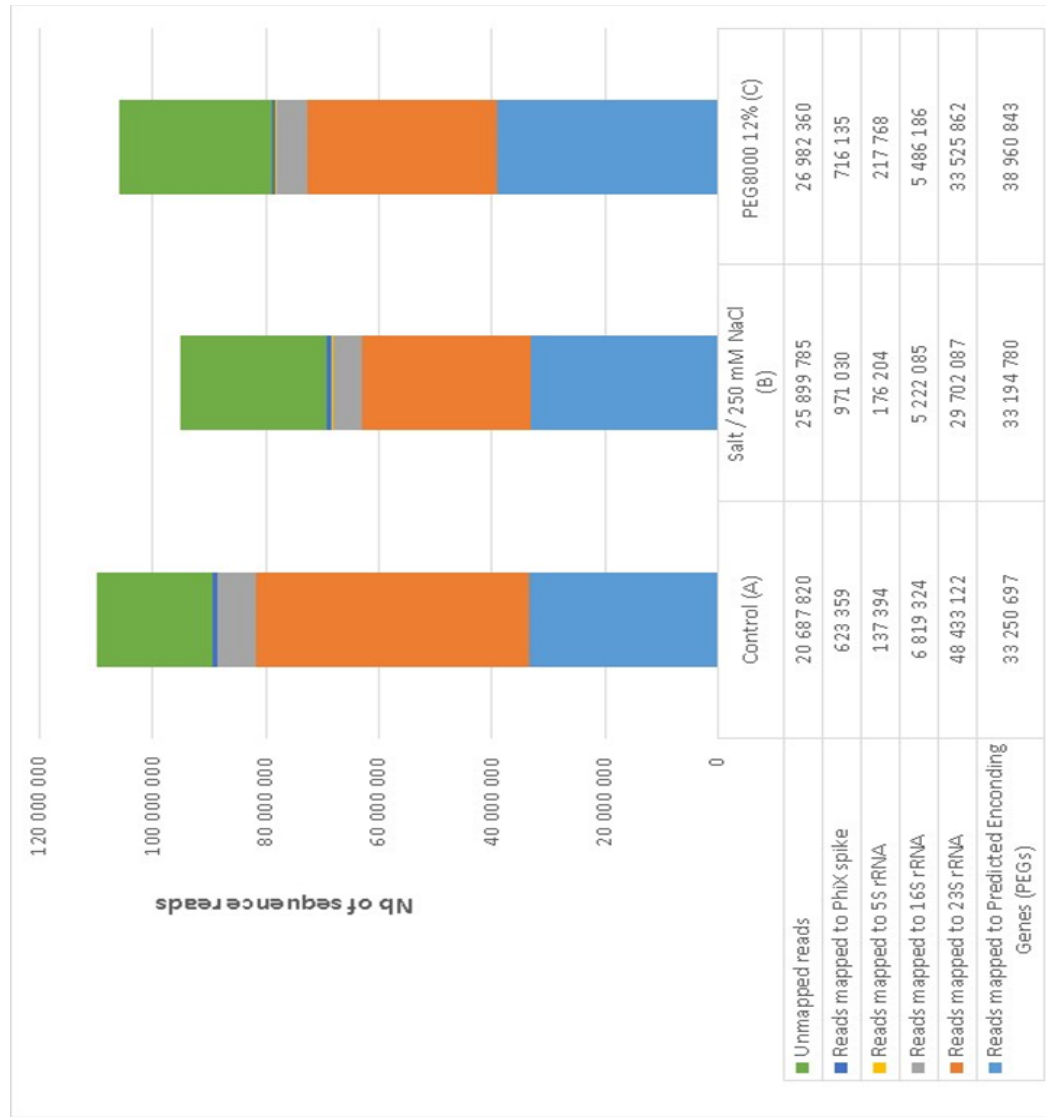
48. Jensen, J.B.; Ampomah, O.Y.; Darrah, R.; Peters, N.K.; Bhuvaneswari, T.V. Role of Trehalose Transport and Utilization in *Sinorhizobium meliloti*--Alfalfa Interactions. *Mol. Plant Microbe Interact.***2005**, *18*, 694–702, doi:10.1094/MPMI-18-0694.
49. Ampomah, O.Y.; Jensen, J.B.; Bhuvaneswari, T.V. Lack of Trehalose Catabolism in *Sinorhizobium* Species Increases Their Nodulation Competitiveness on Certain Host Genotypes. *New Phytol.***2008**, *179*, 495–504, doi:10.1111/j.1469-8137.2008.02460.x.
50. Geddes, B.A.; Oresnik, I.J. Genetic Characterization of a Complex Locus Necessary for the Transport and Catabolism of Erythritol, Adonitol and l-Arabitol in *Sinorhizobium meliloti*. *Microbiology***2012**, *158*, 2180–2191, doi:10.1099/mic.0.057877-0.
51. Yost, C.K.; Rath, A.M.; Noel, T.C.; Hynes, M.F. Characterization of Genes Involved in Erythritol Catabolism in *Rhizobium leguminosarum* Bv. *viciae*. *Microbiology* **2006**, *152*, 2061–2074, doi:10.1099/mic.0.28938-0.
52. Poysti, N.J.; Oresnik, I.J. Characterization of *Sinorhizobium meliloti* Triose Phosphate Isomerase Genes. *J. Bacteriol.***2007**, *189*, 3445–3451, doi:10.1128/JB.01707-06.
53. Geddes, B.A.; Pickering, B.S.; Poysti, N.J.; Collins, H.; Yudistira, H.; Oresnik, I.J. A Locus Necessary for the Transport and Catabolism of Erythritol in *Sinorhizobium meliloti*. *Microbiology***2010**, *156*, 2970–2981, doi:10.1099/mic.0.041905-0.
54. Kohler, P.R.A.; Choong, E.-L.; Rossbach, S. The RpiR-Like Repressor IolR Regulates Inositol Catabolism in *Sinorhizobium meliloti*. *J. Bacteriol.***2011**, *193*, 5155–5163, doi:10.1128/JB.05371-11.
55. Herrou, J.; Crosson, S. Myo -Inositol and D -Ribose Ligand Discrimination in an ABC Periplasmic Binding Protein. *J. Bacteriol.***2013**, *195*, 2379–2388, doi:10.1128/JB.00116-13.
56. Kohler, P.R.A.; Zheng, J.Y.; Schoffers, E.; Rossbach, S. Inositol Catabolism, a Key Pathway in *Sinorhizobium meliloti* for Competitive Host Nodulation. *Appl. Environ. Microbiol.***2010**, *76*, 7972–7980, doi:10.1128/AEM.01972-10.
57. Jiang, G.; Krishnan, A.H.; Kim, Y.-W.; Wacek, T.J.; Krishnan, H.B. A Functional Myo -Inositol Dehydrogenase Gene Is Required for Efficient Nitrogen Fixation and Competitiveness of *Sinorhizobium fredii* USDA191 To Nodulate Soybean (*Glycine max* [L.] Merr.). *J. Bacteriol.***2001**, *183*, 2595–2604, doi:10.1128/JB.183.8.2595-2604.2001.
58. Fry, J.; Wood, M.; Poole, P.S. Investigation of Myo -Inositol Catabolism in *Rhizobium leguminosarum* Bv. *viciae* and Its Effect on Nodulation Competitiveness. *Mol. Plant Microbe Interact.***2001**, *14*, 1016–1025, doi:10.1094/MPMI.2001.14.8.1016.

59. Gourion, B.; Berrabah, F.; Ratet, P.; Stacey, G. Rhizobium–Legume Symbioses: The Crucial Role of Plant Immunity. *Trends Plant Sci.***2015**, *20*, 186–194, doi:10.1016/j.tplants.2014.11.008.
60. Rodrigo da-Silva, J.; Alexandre, A.; Brígido, C.; Oliveira, S. Can Stress Response Genes Be Used to Improve the Symbiotic Performance of Rhizobia? *AIMS Microbiol.* **2017**, *3*, 365–382, doi:10.3934/microbiol.2017.3.365.
61. Patankar, A.V.; González, J.E. An Orphan LuxR Homolog of *Sinorhizobium meliloti* Affects Stress Adaptation and Competition for Nodulation. *Appl. Environ. Microbiol.***2009**, *75*, 946–955, doi:10.1128/AEM.01692-08.
62. Novichkov, P.S.; Li, X.; Kuehl, J.V.; Deutschbauer, A.M.; Arkin, A.P.; Price, M.N.; Rodionov, D.A. Control of MéthionineMetabolism by the SahR Transcriptional Regulator in Proteobacteria: Control of MéthionineMetabolism in Proteobacteria. *Environ. Microbiol.***2014**, *16*, 1–8, doi:10.1111/1462-2920.12273.
63. Jiang, J.Q.; Wei, W.; Du, B.H.; Li, X.H.; Wang, L.; Yang, S.S. Salt-Tolerance Genes Involved in Cation Efflux and Osmoregulation of *Sinorhizobium fredii* RT19 Detected by Isolation and Characterization of Tn5 Mutants. *FEMS Microbiol. Lett.* **2004**, *239*, 139–146, doi:10.1016/j.femsle.2004.08.029.
64. Taté,, R.; Riccio, A.; Caputo, E.; Iaccarino, M.; Patriarca, E.J. The *Rhizobium etli metZ* Gene Is Essential for MéthionineBiosynthesis and Nodulation of *Phaseolus vulgaris*. *Mol. Plant Microbe Interact.***1999**, *12*, 24–34, doi:10.1094/MPMI.1999.12.1.24.
65. Francez-Charlot, A.; Kaczmarczyk, A.; Fischer, H.-M.; Vorholt, J.A. The General Stress Response in Alphaproteobacteria. *Trends Microbiol.***2015**, *23*, 164–171, doi:10.1016/j.tim.2014.12.006.
66. Gourion, B.; Sulser, S.; Frunzke, J.; Francez-Charlot, A.; Stiefel, P.; Pessi, G.; Vorholt, J.A.; Fischer, H.-M. The PhyR- σ^{EcFG} Signalling Cascade Is Involved in Stress Response and Symbiotic Efficiency in *Bradyrhizobium japonicum*. *Mol. Microbiol.***2009**, *73*, 291–305, doi:10.1111/j.1365-2958.2009.06769.x.
67. Ledermann, R.; Bartsch, I.; Müller, B.; Wülser, J.; Fischer, H.-M. A Functional General Stress Response of *Bradyrhizobium diazoefficiens* Is Required for Early Stages of Host Plant Infection. *Mol. Plant Microbe Interact.***2018**, *31*, 537–547, doi:10.1094/MPMI-11-17-0284-R.

68. Bastiat, B.; Sauviac, L.; Bruand, C. Dual Control of *Sinorhizobium meliloti* RpoE2 Sigma Factor Activity by Two PhyR-Type Two-Component Response Regulators. *J. Bacteriol.***2010**, *192*, 2255–2265, doi:10.1128/JB.01666-09.
69. Sauviac, L.; Philippe, H.; Phok, K.; Bruand, C. An Extracytoplasmic Function Sigma Factor Acts as a General Stress Response Regulator in *Sinorhizobium meliloti*. *J. Bacteriol.***2007**, *189*, 4204–4216, doi:10.1128/JB.00175-07.
70. Jans, A.; Vercruysse, M.; Gao, S.; Engelen, K.; Lambrichts, I.; Fauvart, M.; Michiels, J. Canonical and Non-Canonical EcfG Sigma Factors Control the General Stress Response in *Rhizobium etli*. *MicrobiologyOpen***2013**, *2*, 976–987, doi:10.1002/mbo3.137.
71. Foreman, R.; Fiebig, A.; Crosson, S. The LovK-LovR Two-Component System Is a Regulator of the General Stress Pathway in *Caulobacter crescentus*. *J. Bacteriol.***2012**, *194*, 3038–3049, doi:10.1128/JB.00182-12.
72. Lang, C.; Barnett, M.J.; Fisher, R.F.; Smith, L.S.; Diodati, M.E.; Long, S.R. Most *Sinorhizobium meliloti* Extracytoplasmic Function Sigma Factors Control Accessory Functions. *mSphere***2018**, *3*, doi:10.1128/mSphereDirect.00454-18.
73. Flechard, M.; Fontenelle, C.; Trautwetter, A.; Ermel, G.; Blanco, C. *Sinorhizobium meliloti* *rpoE2* Is Necessary for H₂ O₂ Stress Resistance during the Stationary Growth Phase. *FEMS Microbiol. Lett.* **2009**, *290*, 25–31, doi:10.1111/j.1574-6968.2008.01401.x.
74. Ledermann, R.; Emmenegger, B.; Couzigou, J.-M.; Zamboni, N.; Kiefer, P.; Vorholt, J.A.; Fischer, H.-M. *Bradyrhizobium diazoefficiens* Requires Chemical Chaperones To Cope with Osmotic Stress during Soybean Infection. *mBio***2021**, *12*, doi:10.1128/mBio.00390-21.
75. Ravcheev, D.A.; Khoroshkin, M.S.; Laikova, O.N.; Tsoy, O.V.; Sernova, N.V.; Petrova, S.A.; Rakhmaninova, A.B.; Novichkov, P.S.; Gelfand, M.S.; Rodionov, D.A. Comparative Genomics and Evolution of Regulons of the LacI-Family Transcription Factors. *Front. Microbiol.***2014**, *5*, doi:10.3389/fmicb.2014.00294.

Supplementary figures

Figure S1. Sequencing and Mapping of the 3 RNAseq libraries. Histograms show for the number of sequenced reads obtained by HiSeq for the 3 libraries (indicated at the bottom of each bar. For each bar, the reads that mapped to PEGs, rRNAs, PhiX spike sequence or those were not mapped using 95% sequence identity over 70% of sequence lengths (other reads) are shown using the color code indicated in the table at the bottom of the histogram together with their respective numbers in the 3 libraries.



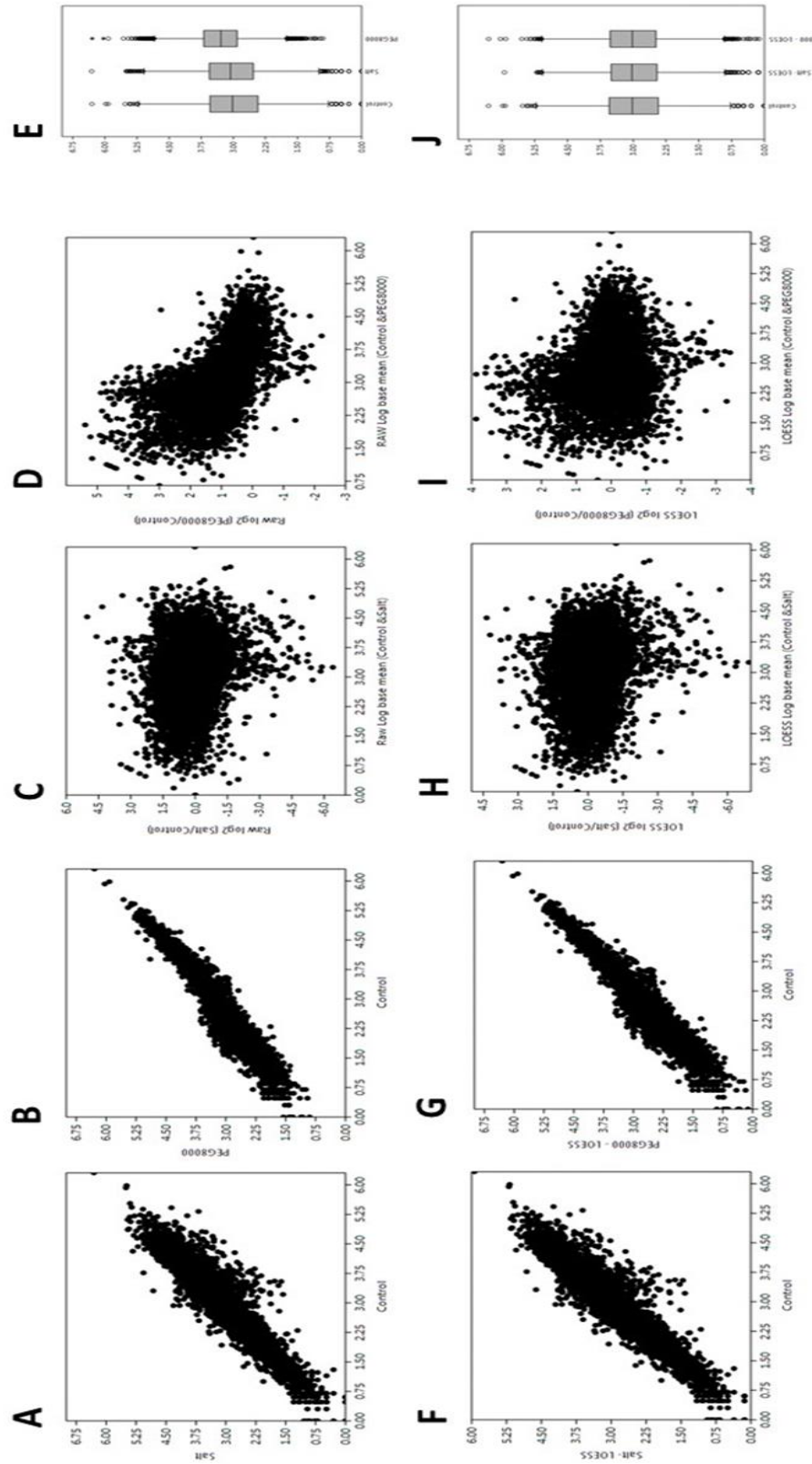


Figure S2. Distributions of RNAseq data. Raw data (A-E) and LOESS normalized data (F-J) are represented using Log scales. Scatter plots show the read counts obtained in control (X axes) and salt stress (A, F) or water stress (B, G) treatments for all predicted PEGs. MA plots for salt stress (C, H) or water stress (D, I) and box plots (E, J) represent the distributions of PEG read counts in each library (Control, salt and PEG8000).

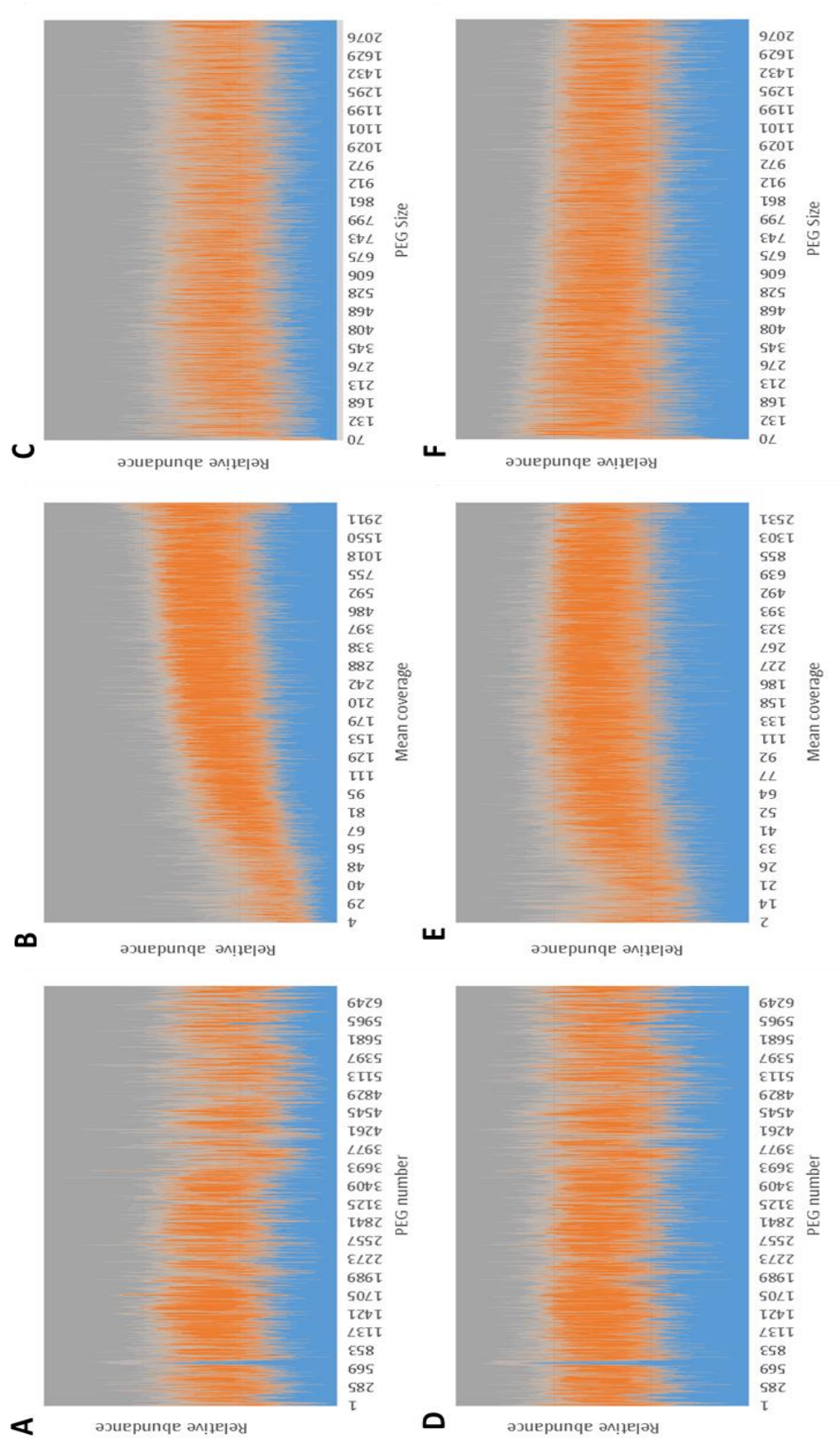


Figure S3. Effect of PEG order, mean coverage or size on the Relative abundance of reads obtained in the 3 libraries Control (Blue), Salt (Orange) PEG8000(Grey). Relative abundances obtained using raw data upon mapping onto PEGs (A-C) and upon LOESS normalization (D-F) are represented using Surface areas 100% piled graphics. Relative abundances obtained upon ordering PEGs according to their predicted position on the genome (A, D) their mean coverage obtained from the three libraires (B, E) or their sizes (C, F) are shown.

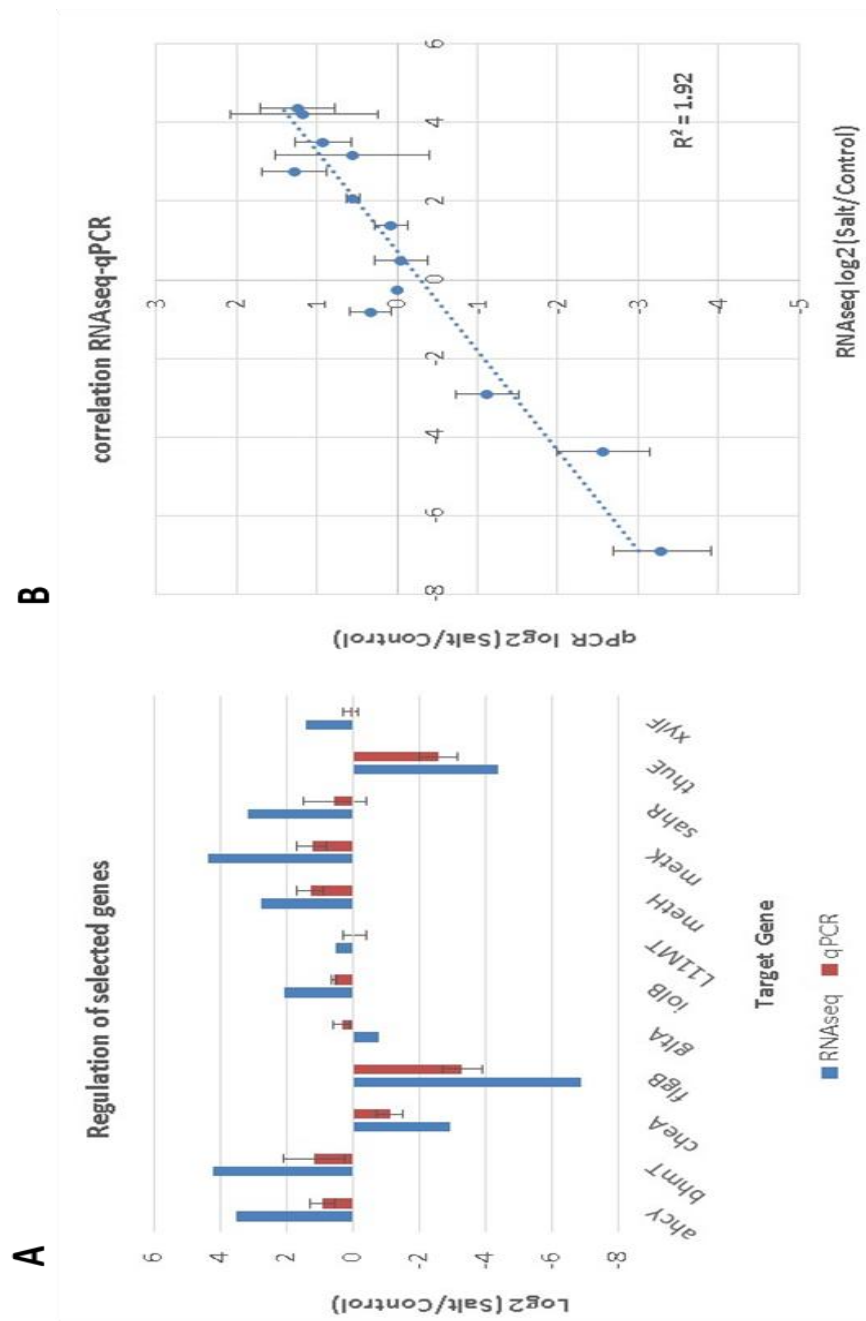


Figure S4: Comparison of the relative expression levels of selected genes by RNAseq and qPCR. The log2 fold changes of expression in salt treated cells as compared to the Control TY cells for the selected genes (Shown below bars in histogram) were calculated using the reference genes *hrcA*. Individual relative expressions for the 12 genes are shown in A; the RNAseq differential expression values are shown as blue bars and red bars show the mean relative log2 fold changes (Salt / Control) as analyzed by qPCR using 3 RNA samples, and error bars correspond to standard deviations. The scatter plot (B) shows the average log2 fold changes (Salt / Control) from the RNAseq data (X Axis) against the mean qPCR values from three replicates on RNA samples from 3 biological replicates (Y Axis) which was used to calculate the correlation factors (R^2) indicated on the graph.

1. Introduction

Divers microorganismes, comme les rhizobiums, ont la capacité de s'adapter à des conditions de croissance contrastées et difficiles dans les sols/rhizosphère (à l'état libre), ou à l'intérieur de la racine de la plante. Face à ces conditions de vies changeantes et fluctuantes, ces rhizobiums ont développé des réponses adaptatives spécifiques, des compétences et des mécanismes leur permettant de résister et de se maintenir dans l'environnement. La souche *Ensifer aridi* LMR001 T (=LMG 31426 T ;= HAMBI 3707 T) a été isolée des dunes de sable désertique dans le désert Marocain de Merzouga, où les conditions environnementales sont défavorables pour la survie et la croissance de la plupart des organismes (Sakrouhi et al. 2016). La bactérie *E. aridi* LMR001 est capable d'utiliser un large éventail de sources de carbone, de développer la symbiose avec des plantes adaptées aux contraintes environnementales et d'agir avec des mécanismes spécifiquement adaptés à ces environnements arides et semi-arides. De ce fait, *Ensifer aridi* représente un modèle intéressant pour étudier les réponses adaptatives impliquées dans la tolérance à l'aridité. Ainsi, pour définir et étudier les fonctions associées à l'adaptation de la bactérie *Ensifer aridi* (modèle d'étude). Un séquençage a été effectué des séquences d'ARNs exprimées dans des conditions de stress hydrique et de stress salin. L'étude transcriptomique « RNAseq » nous a permis de sélectionner les métabolismes différentiellement exprimés et de cibler les gènes afin d'évaluer leur rôle fonctionnel dans la réponse adaptative de cette bactérie aux stress étudiés. Parmi ces gènes, on trouve le gène *rpoE2* régulateur du système de la réponse générale au stress. L'article présente l'analyse transcriptomique par RNAseq ainsi que des résultats sur le rôle du gène *rpoE2* dans la réponse au stress salin et l'adaptation de la bactérie *Ensifer aridi* aux changements dans son milieu de vie.

2. Résultats et discussion

Les données du séquençage HiSeq2000 des ARNs de la bactérie *E. aridi* LMR001 de 100 millions de séquences ont été traitées. Elles ont été mappées sur le génome de LMR001 annotées par le pipeline RAST pour identifier 6294 séquences codantes (CDS) (Le Quéré et al., 2017). L'analyse des résultats a permis d'identifier les CDS induites ou réprimées par chaque type de stress. Le stress salin a induit 59 CDS et a réprimé 189 CDS ; alors que le stress hydrique a induit 25 CDS et réprimé 38 CDS. On remarque que le stress salin a affecté l'expression de plus de gènes que le stress hydrique et que les gènes réprimés sont plus abondants que les gènes induits. L'annotation par RAST a montré aussi qu'un certain nombre de processus métaboliques a été affecté par les stress chez la souche *Ensifer aridi* LMR001. En effet, nous avons pu observer que plusieurs gènes

impliqués dans les métabolismes et/ou le transport du tréhalose, de plusieurs sources en carbone comme le myo inositol,, des acides aminés dont la méthionine, la modification de la surface cellulaire, la mobilité ou la réponse au stress oxydatif étaient différentiellement exprimés en réponse au stress hyperosmotique (Belfquih et al., 2021).

La deuxième approche expérimentale a consisté en la mutagenèse ciblée du gène codant pour le facteur sigma alternatif (*rpoE2*) impliqué dans la régulation de la réponse générale au stress (GSR) dont les régulateurs sont apparus différentiellement exprimés à la suite du stress dans l'étude transcriptomique. Elle a permis d'estimer le rôle de ce facteur sigma dans la réponse au stress hyperosmotique et dans le développement de la symbiose avec des hôtes compatibles. Généralement, les *alphaprotéobactéries* (la classe à laquelle appartient *Ensifer aridi*), induisent un système de régulation composé d'un sigma facteur de fonction extra cytoplasmique (ECF) appelé σ ECFG, codé par le gène *rpoE2*. L'activité du sigma facteur *RpoE2* implique aussi l'activité d'anti-sigma NepR, codé par le gène *rsiA1* et l'activité d'un autre facteur anti-anti sigma PhyR, codé par le gène *rsiB1* dont l'activité dépend de la phosphorylation par l'histidine kinase PhyK (Bastiat et al., 2010).

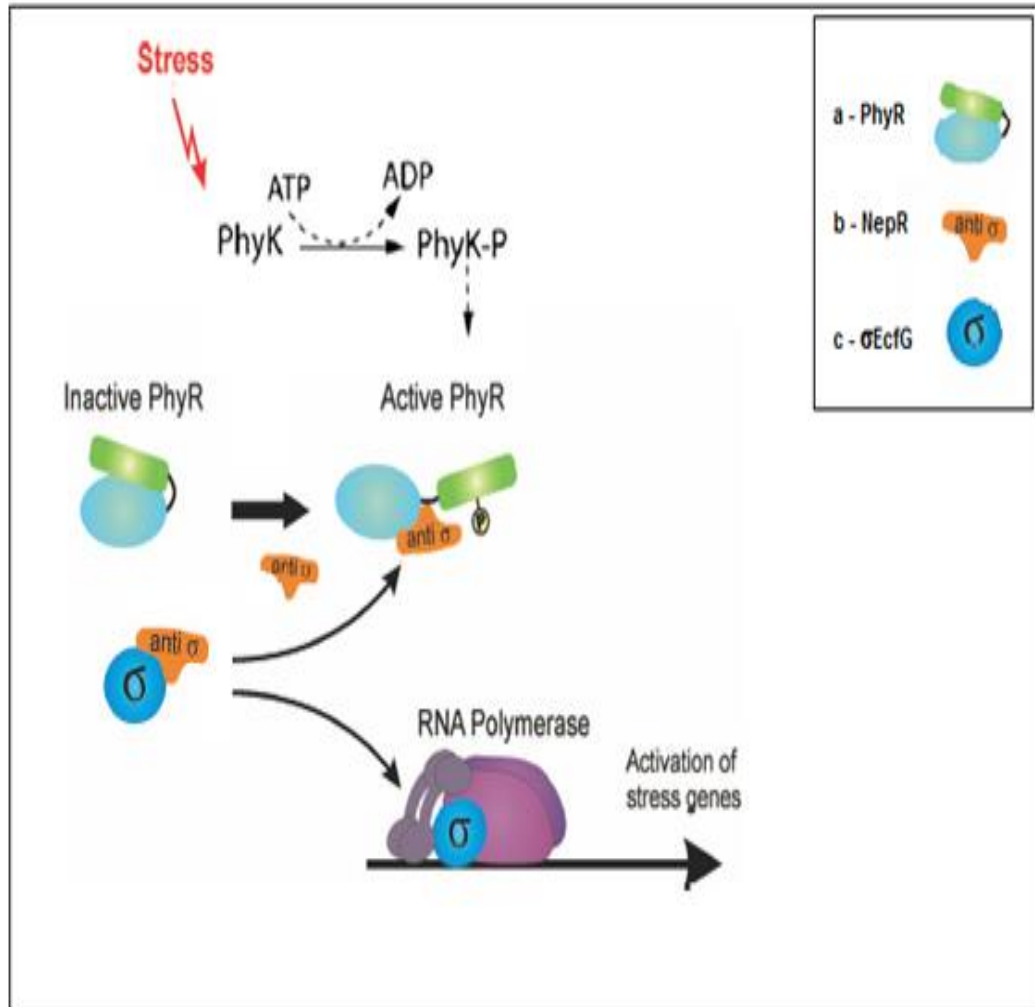


Figure 6: Représentation du mécanisme majeur de la régulation, PhyP-PhyR-NepR-σE, de la transcription des ARN liés au stress. Ce mécanisme contrôle la réponse générale au stress chez *M. extorquens*. En cas de situation de non-stress, la protéine PhyR est inactive et le facteur sigma ECF est séquestré par son facteur anti-sigma, NepR. En réponse à un stress, PhyR est phosphorylée et interagit avec NepR, libérant ainsi EcfG1 lui permettant de s'associer à l'ARN polymérase et de transcrire les gènes de stress (Francez-Charlot et al., 2009).

Sous effet du stress, la phosphorylation de la protéine PhyR du système à deux composants conduit à sa liaison au facteur anti-sigma NepR. Ce dernier dans des conditions normales de croissance de non-stress est lié au σECFG. Cependant, la fixation PhyR-NepR libère le σECFG qui peut ainsi par conséquent mobiliser l'ARN polymérase l'expression de nombreux gènes de la réponse générale au stress GSR (Francez-Charlot et al., 2009).

L'examen des bases de données de séquences génomiques indique la présence de séquences qui codent pour des orthologues de NepR et PhyR chez *E. meliloti* mais aussi chez *E. medicae* et chez *E. fredii* NGR234 (respectivement *rsiA1* et *rsiB1*). L'originalité chez *E. meliloti* réside sur le fait

qu'elle possède deux paires de facteurs *rsiB/rsiA* suggérant une complexité du mécanisme de régulation de la voie de signalisation RpoE2-dépendante. Les annotations et les noms des gènes figurent sur le site Web de *E. meliloti* (<http://sequence.toulouse.inra.fr/E.meliloti>) (Bastiat et al., 2010). Les protéines PhyR (*rsiB*), NepR (*rsiA*) et σ^{EcfG} ou σ^{24} (*rpoE2*) sont conservées dans pratiquement toutes les *alphaprotéobactéries* et leurs gènes se trouvent généralement situés sur un même locus (Gourion, Francez-Charlot, et Vorholt 2008; Staroń et al., 2009 ; Kaczmarczyk et al., 2011). Comme pour *E. meliloti*, deux copies des gènes codant pour PhyR et NepR ont été identifiées par une recherche BLAST sur le protéome de la bactérie *Ensifer aridi* LMR001. L'organisation des loci chez *Ensifer aridi* LMR001 est illustrée sur la figure 7. Le gène PhyR (*rsiB*) est situé en anti-sens des deux gènes *nepR* (*rsiA*) et *rpoE2*, et code pour la protéine PhyR qui porte un domaine σ^{like} en N-terminal et en C-terminal une partie du domaine de réception Rec. Cette protéine est aussi appelée par l'anti-anti-sigma factor. Le gène *nepR* (*rsiA1*) localisé en amont du gène *rpoE2* est un gène codant pour l'anti-sigma factor et est co-transcrit. Ces gènes sont conservés parmi plusieurs espèces d'*Ensifer*. Bien que la signification de la présence des deux paires de régulation PhyR (*rsiB1*, *rsiB2*), NepR (*rsiA1*, *rsiA2*) chez *E. aridi* soit jusqu'à présent inconnue, une hypothèse pourrait être que ces deux facteurs anti-anti-sigma soient phosphorylés par des histidine kinases différentes permettant ainsi éventuellement de détecter des stimuli différents en réponse à des environnements changeants (Bastiat et al., 2010).

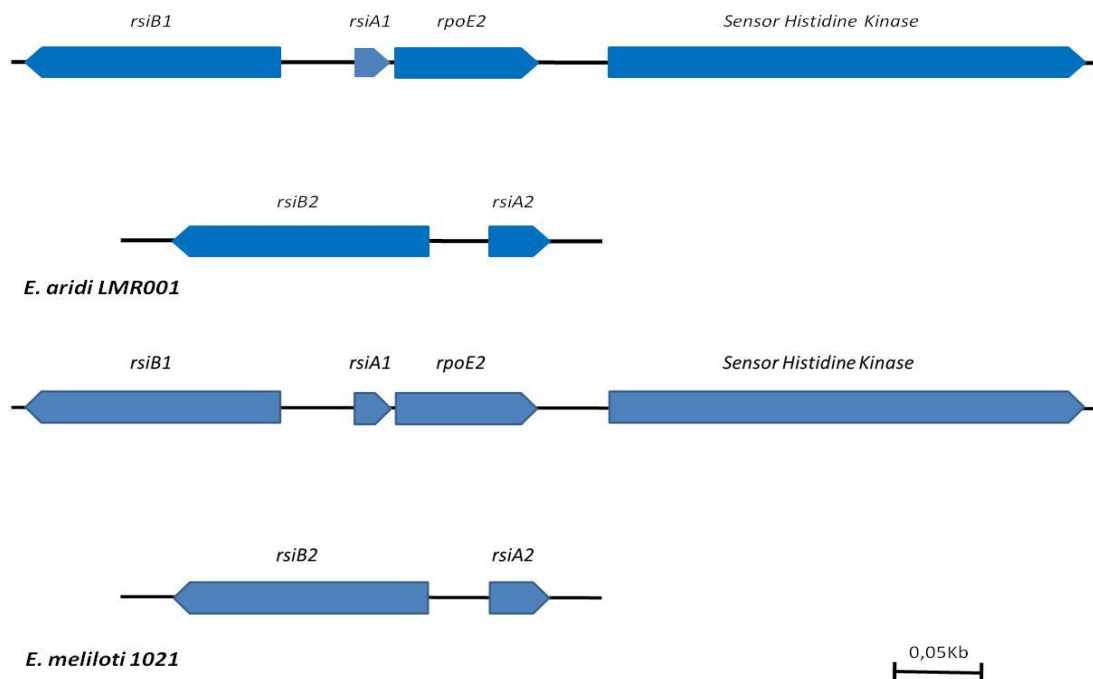


Figure 7: Représentation schématique de l'organisation des régions chromosomiques de la région des gènes *Sigma-EcfG (rpoE2)*, *PhyR (rsiB)*, *nepR (rsiA)* et *phyK* ou *phyP* (sensor kinase) dans la souche *E.aridi* LMR001 par rapport à *E.meliloti* 1021. Les flèches représentent l'origine des cadres de lecture ORF et leurs directions représentent la direction de la transcription. La position de tous les gènes de la souche *E.meliloti* est présentée dans l'article (Bastiat et al., 2010).

Les résultats de l'analyse transcriptomique ont révélé une induction du système PhyP-PhyR-NepR- σ E sous effet du stress salin plus forte que le stress hydrique. Cette différence d'expression peut être due à une contrainte plus forte induite par l'exposition à court terme au sel. Une augmentation de la régulation d'environ 30 % des gènes impliqués dans la traduction de la structure ribosomique et de la biogenèse a été révélée et une répression de 60% des gènes impliqués dans la chimiotaxie et motilité. Chez les bactéries, la répression des gènes de la mobilité est souvent associée à une induction de gènes impliqués dans la synthèse des composants du biofilm tels que les polysaccharides de surface ou les protéines (Wadhams et al., 2004, Verstraeten et al., 2008). Parmi les gènes qui ont été réprimés par l'effet du stress osmotique, on trouve chez *E. aridi* les gènes qui codent pour, VisR/VisN et Rem (codées respectivement par PEG655/PEG654 et PEG688) et d'autres gènes induits impliqués dans la production des exopolysaccharides ; (*exoQ*, *exoF*, *exoY* *exoH*, *exoO* et *exoN*, respectivement PEG5904-5906, PEG5914, PEG5919 et PEG5920) (Janczarek 2011) .

D'autres voies métaboliques ont été aussi montrées différenciellement exprimées en réponse aux stress hyperosmotiques. Les deux voies de biosynthèse de novo du tréhalose : une voie générée par l'expression des gènes codant pour alpha, alpha - tréhalose -phosphate synthase (*otsA*, PEG4869) et le tréhalose-6-phosphatase (*otsB*, PEG4868). Ces deux enzymes convertissent le glucose UDP et le glucose 6P en tréhalose par deux réactions successives. La deuxième voie est produit par la conversion de la maltodextrine en tréhalose par une Malto-oligosyltréhalose synthase (*treY*, PEG5229) et une Malto-oligosyltréhalose tréhalohydrolase (*treZ*, PEG5323) (Avonce et al., 2006, Avonce et al., 2010, Gopal et al., 2010). Nous avons également constaté une forte répression des gènes de transport du tréhalose *thuEFGK* de type ABC (PEG6268, 6266, 6265 et 6264) et les gènes d'utilisation du tréhalose *thuAB* (PEG6263 et 6262) (Ampomah et al., 2013). On a aussi trouvé parmi les métabolismes altérés par le stress salin, une répression d'un grand locus (PEG2271—PEG2284) impliqué dans le transport et le catabolisme de l'érythritol. L'opéron qui code pour le système de transport de type ABC de l'érythritol (PEG2283-2279) est situé juste en amont des gènes cataboliques. Il a été aussi montré une répression de l'ordre de 10 fois du gène répresseur appelé *eryD* (PEG2284). En revanche, le transport et le catabolisme du myo-inositol

ont montré une forte induction par le sel avec notamment une forte activation du système de transport myo-inositol ABC (PEG5305-5307), et les gènes *idhA* (PEG2503) et *iolY* (PEG375) qui codent pour 2 myo inositol-2 déshydrogénases. Les autres gènes *iolE* (PEG371), *iolD* (PEG372), *iolB* (PEG370) et *iolC* (PEG373) qui interviennent dans la conversion successive de 2- céto-myo-inositol en 3-D-trihydroxycyclohexane-1,2-dione, acide 5-désoxy-glucuronique, 5-céto-D-gluconique et acide 2-désoxy-5-céto-D-gluconique 6-phosphate (DKGP) étaient aussi induits (Kohler et al., 2010, Kohler et al., 2013). Les résultats du séquençage des ARNs ont aussi révélé une induction des gènes impliqués dans le cycle de la méthionine. Il comprenait l'adénosyle homocystéinase *ahcY* (PE3356), une bétaine-homocystéine S – méthyltransférase *BHMT* (PEG1969), une réductase codée par *metF* (PEG2238), et une 5-méthyltétrahydrofolate-homocystéine méthyltransférase codée par *metH* (PEG3206). Enfin, nous avons trouvé une forte induction du gène *metK* codant pour la S- adénosylméthionine (SAM) synthétase (PEG438) (Novichkov et al., 2014). de plus, deux facteurs transcriptionnels régulateurs du cycle de la méthionine, appelées NesR et SahR, ont été également trouvés différemment régulés. Un doublement de la transcription du gène *nesR* par le stress salin par rapport au témoin tandis que *sahR* a été trouvé 9 fois plus exprimé dans le sel.

Des études ont ensuite été réalisées pour évaluer le rôle du gène *rpoE2* impliqué dans la GSR par une approche de mutagénèse dirigée. La comparaison des réponses phénotypiques de mutant par rapport à la souche sauvage a ensuite été effectué. Premièrement l'étude de la croissance bactérienne en réponse au stress salin a montré que la *LMR001* sauvage a atteint plus rapidement la phase stationnaire de croissance en milieu TY complémenté avec du sel par rapport à la souche mutée *rpoE2*. Après 120 heures d'incubation, les mutants ont montré une croissance comparable à la sauvage dans le milieu stressé. Le test de la motilité des deux souches a révélé une réduction significative de la motilité pour les deux souches lors de l'ajout de sel aux deux concentrations testées (0,1 et 0,25 M). La sensibilité des deux souches aux stress détergents, oxydatifs et acides n'a montré aucune différence entre les sauvages et les mutants Suggérant donc que la mutation du gène *rpoE2* n'a pas altéré la croissance, la mobilité et la tolérance aux stress acides et oxydatifs. La réponse de la symbiose des deux souches avec l'espèce d'arbre *V. gummifera* et *V. tortilis* n'a montré aucune différence significative entre les données des poids des plantes ou à la nodulation. Ces données indiquent que facteur sigma ECF RpoE2 n'est pas essentiel pour le développement de la symbiose et pour la tolérance au stress pour *E. aridi LMR001*.

Un autre test transcriptionnel a été aussi réalisé, comparant l'expression des gènes sélectionnés entre la bactérie sauvage et le mutant en réponse au stress salin. Le test a révélé une variabilité de l'expression du gène rapporteur *gfp* lié aux promoteurs des gènes ciblés par l'étude. Nous avons

constaté que la mutation entraînait une transcription élevée de quatre gènes, et une réduction de l'expression de trois gènes. Une induction du gène *otsB* dans le mutant *rpoE2*, indiquant que la GSR médiée par RpoE2 dans *E. aridi* implique la voie T6P. Nous avons aussi trouvé une répression du gène *rsiA1*, expliqué par la présence du site de fixation GGAACN 16-17 CGTT en amont des gènes *rsiA1*, *rsiB1*. Ces résultats suggèrent que le gène *rpoE2* est autorégulé chez *E. aridi*. Nous avons également constaté que le motif promoteur de transcription de la RpoE2 a été représenté deux fois dans la région promotrice du gène *treZ* gène. Bien que la transcription du gène *treZ* est faible et proche des valeurs basales de fluorescence, sa transcription n'apparaît pas significativement différente après mutation du gène *rpoE2*. Deux motifs ont été également trouvés dans la région promotrice du gène *otsB* et un motif avec trois différences nucléotidiques en amont du gène *thuR*. Cependant, des études supplémentaires sont encore nécessaires pour vérifier et évaluer la régulation directe de ces gènes par RpoE2. La répression transcriptionnelle du gène *thuR* s'est montrée la plus sévère chez les mutants. Ce gène code pour un répresseur de type HTH DNA-Binding LacI, homologue avec l'*aglR* et impliqué dans le transport et l'utilisation du tréhalose par la régulation de la transcription de l'opéron *thuEFGK* et des deux gènes *thuAB*. La présence d'un partage partiel du même type de séquence de motif sur les régions promotrices du gène *thuR* et du gène *aglR* (régulateur de l'opéron *aglEFGAK*) peut montrer une régulation croisée de ces deux systèmes de transport. Le régulateur ThuR est homologue au régulateur AglR. Nos données suggèrent que ThuR est également autorégulé et impliqué dans la régulation du transport et de l'utilisation du tréhalose chez *E. aridi*. De plus, la réalisation de d'autres études supplémentaires reste nécessaire pour décrypter les voies de régulations impliquées dans la régulation du gène *thuR* et évaluer l'implication directe ou indirecte du régulateur RpoE2 dans le transport du tréhalose.

3. Conclusion

L'étude transcriptomique réalisée sur *E. aridi* LMR001, sous stress hydrique et salin, a révélé une augmentation de l'expression d'un ensemble des gènes liés à la GSR et notamment le système complexe impliquant PhyP-PhyR-NepR-σE. Sur ces résultats, nous avons effectué une étude fonctionnelle sur le gène *rpoE2*, qui est un gène de la famille des facteurs de la fonction extracytoplasmique ECF (de classe 2). Ce gène code pour le sigma facteur E (σ^{24}). Nous avons relevé au niveau de la réponse transcriptionnelle, une activité plus forte du gène *rpoE2* en réponse au stress salin par rapport au stress hydrique. Aussi l'expression des gènes de la mobilité et la chemotaxie s'est montrée réprimée et cette répression est corrélée avec une induction des gènes de la synthèse des composants du biofilm. Les autres voies métaboliques ; synthèse et transport du tréhalose, catabolisme d'inositol et méthionine cycle ; sont révélés aussi différenciellement

exprimées sous effet des deux stress. Les résultats phénotypiques et physiologiques ont montré que la mutation du gène *rpoE2* n'a pas affecté la tolérance et l'activité symbiotique de la souche *E. aridi* mutée. Ces données suggèrent que le gène *rpoE2* n'est pas essentiel à la bactérie *E. aridi* en réponse au stress appliqué ainsi que pour la colonisation et l'infection des plantes hôtes. De plus, nous avons constaté la présence de motifs conservés de liaison de RpoE2 en amont des gènes *rsiA1-rpoE2*, des gènes *rsiB1*, *treZ*, *otsB* et *thuR* qui suggèrent malgré tout un rôle direct ou indirect dans la régulation de la réponse au stress via l'activité du *rpoE2* chez *E. aridi* en réponse au stress salin qui est cependant complémentée par d'autres mécanismes.

CHAPITRE IV

Approche de mutagenèse dirigée des gènes ciblés dans le métabolisme de la méthionine, du tréhalose et d'inositol par l'étude transcriptomique et analyse fonctionnelle des gènes de régulation

Article 3, Belfquih et al., 2022: Analysis of *Ensifer aridi* Mutants Affecting Regulation of Methionine, Trehalose, and Inositol Metabolisms Suggests a Role in Stress Adaptation and Symbiosis Development

Auteur : Meryem Belfquih 1,2, Abdelkarim Filali-Maltouf 1,* and Antoine Le Quéré 1,2,*

¹ Laboratory of Microbiology and Molecular Biology (LMBM), Department of Biology, Faculty of Sciences, University Mohammed V in Rabat, Rabat, 10100, Morocco; belfquih.meryem@gmail.com (M.B.); filalimaltouf@gmail.com (A.F.-M.); antoine.le-quere@ird.fr (A.L.Q.)

² Laboratoire des Symbioses Tropicales et Méditerranéennes (LSTM), University of Montpellier, IRD, CIRAD, INRAE, Institut Agro, Montpellier, 34398, France

* Correspondence: A.F-M & A.L.Q.

Abstract: Isolated from desert, the nitrogen fixing bacterium *Ensifer aridi* LMR001 is capable to survive under particularly harsh environmental conditions. To get insights in molecular mechanisms involved in stress adaptation, a recent study using RNAseq revealed that the RpoE2 mediated general stress response was activated under mild saline stress but appeared non-essential for the bacterium to thrive under stress and develop the symbiosis. Functions associated with the stress response included the metabolisms of trehalose, methionine and inositol. To explore the roles of these metabolisms in stress adaptation and symbiosis development, and the possible regulatory mechanisms involved, mutants were generated notably in regulators and their transcriptions were studied in various mutant backgrounds. We found that mutations in regulatory genes *nesR* and *sahR* of the le cycle de la methionine generating S-Adenosyl Methionine negatively impacted symbiosis, tolerance to salt and motility in the presence of NaCl. When both regulators were mutated, an increased tolerance to detergent, oxidative and acid stresses was found suggesting a modification of the cell wall components which may explain these phenotypes and support a major role of the fine-tuning methylation for symbiosis and stress adaptation of the bacterium. In contrast, we also found that mutations in the predicted trehalose transport and utilization regulator *ThuR* and the trehalose phosphate phosphatase *OtsB* encoding genes improved

symbiosis and growth in liquid medium containing 0.4 M of NaCl of LMR001 Δ *otsB* suggesting that trehalose metabolism control and possibly trehalose-6 phosphate cellular status may be biotechnologically engineered for improved symbiosis under stress. Finally, transcriptional fusions of *gfp* to promoters of selected genes and expression studies in the various mutant backgrounds suggest complex regulatory interplay between inositol, methionine and trehalose metabolic pathways.

Keywords: *Ensifer aridi*; nitrogen fixing symbiosis; trehalose; S-Adenosyl Methionine cycling; inositol catabolism)

1. Introduction

Primarily found in deserts and capable to develop the symbiosis with plants native and adapted to stress, *Ensifer aridi* represents an interesting species to study genes and functions that may explain its particular adaptation, broad host range and phylogeographic distribution [1–5]. Next generation sequencing technologies (NGS) have become accessible to study any organism at the genome-wide level. Comparative genomic analyzes of 6 *Ensifer aridi* strains isolated in Asia, Africa or America from diverse legumes showed that accessory genes were most probably located on large plasmids acquired through horizontal gene transfers that probably enabled the bacterium to develop the symbiosis with local hosts [6].

Among the most rapid methods to identify genes and functions associated to a particular stress, transcriptomics represents a method of choice. Genome-wide transcriptional responses resulting from abiotic stresses have been studied using microarray technology or NGS-based RNAseq in several rhizobial genera including for example *Bradyrhizobium* [7], *Rhizobium* [8, 9], *Mesorhizobium* [10–14] or *Sinorhizobium* (*Ensifer*) [15–20]. The bacterial response to an osmotic upshift is complex and varies according to the strain and species. Nevertheless, to counteract osmotic pressure and the resulting oxidative stress, bacteria commonly accumulate osmoprotectants, compatible

solutes, chaperones and enzymes involved in ROS scavenging and they modify their cellular membranes and envelopes [21].

According to a transcriptome analysis performed in *E. aridi* to explore the response of this bacterium to hyperosmotic stress, a number of genes were found differentially regulated including the alternative extra-cytoplasmic function sigma factor RpoE2 which suggested activation of the General Stress Response (GSR) [15]. However, the mutation of the *rpoE2* gene suggested that the response was complex and that alternative regulatory mechanisms may be recruited by the bacterium to cope with various stresses. Among genes that were found differentially regulated, a strong repression of motility genes could be identified and validated phenotypically by swimming assays [15]. We also found that salt stress altered regulation for a number of genes involved in the le cycle de la methionine generating S-Adenosyl Methionine (SAM), the universal methyl donor, in inositol transport and catabolism or in trehalose *de novo* biosynthesis and transport [15]. To further explore the possible roles these metabolisms, have in stress adaptation and symbiosis of *E. aridi*, deletion mutants of regulatory genes involved in the le cycle de la methionine (*nesR*, *sahR*), the trehalose transport and utilization (*thuR*) and the inositol catabolism (*iolR*) or in the trehalose denovo biosynthesis (*otsB*) were generated and phenotyped. Finally, to explore the interdependency between regulated pathways, the transcriptional responses of studied genes were assessed in the various mutant backgrounds upon mobilization of vectors that harbored selected promoter fusions to Green Fluorescence Protein (GFP) encoding gene.

2. Materials and Methods

2.1. Bacterial strains and plasmids

The bacterial strains and plasmids used in this study are listed in Table 1. *Ensifer* strains and derivatives were cultured in Tryptone Yeast extract (TY) medium [22] at 28 °C. *E. coli* strains were grown at 37 °C in Luria-Bertani (LB) medium [23]. Agar at 2% (W/V) was added when needed. For selection of bacterial transformants and transconjugants, sucrose (5%, W/V) gentamycin (Gm, 10 µg mL⁻¹), kanamycin (Km, 50

$\mu\text{g mL}^{-1}$), ampicillin (Ap, $50 \mu\text{g mL}^{-1}$), rifampicin (Rif, $100 \mu\text{g mL}^{-1}$) and 5-bromo-4-chloro-3-indolyl-beta-D-galactosidase (X-gal, $80 \mu\text{g mL}^{-1}$) were added to the media when required.

Table 1. List of bacteria and plasmid used in the study.

Strain/plasmid name	Description	Source
Bacterial strains		
LMR001 ^T	<i>Ensifer aridi</i> LMR001 ^T isolated from <i>Vachellia gummifera</i> grown in Merzouga desert sand dune	[4]
LMR001 Δ <i>rpoE2</i>	<i>rpoE2</i> deletion mutant of LMR001	[15]
LMR001 Δ <i>nesR</i>	<i>nesR</i> deletion mutant of LMR001	This work
LMR001 Δ <i>sahR</i>	<i>sahR</i> deletion mutant of LMR001	This work
LMR001 Δ <i>nesR</i> Δ <i>sahR</i>	<i>nesR</i> deletion mutant of LMR001 Δ <i>sahR</i>	This work
LMR001 Δ <i>iolR</i>	<i>iolR</i> deletion mutant of LMR001	This work
LMR001 Δ <i>thuR</i>	<i>thuR</i> deletion mutant of LMR001	This work
LMR001 Δ <i>otsB</i>	<i>otsB</i> deletion mutant of LMR001	This work
<i>E. coli</i> DH5 α	<i>hsdR17 endA1 thi-1 gyrA96 relA1 recA1 supE44</i> DlacU169 (f80 <i>lacZ</i> DM15)	[24]
<i>E. coli</i> S17-1	<i>recA</i> [SmR], <i>thi</i> , pro, RP4-2-Tc:Mu: <i>aph::Tn7</i> λ pir.	[25]
<i>E. coli</i> XL2 Blue Ultra-competent	<i>endA1 supE44 thi-1 hsdR17 recA1 gyrA96 relA1 lac</i> [F' <i>proAB</i> <i>lacIqZ</i> Δ M15 Tn10 (TetR) Amy CamR]	Stratagene
Plasmids		
pJQ200SK	pACYC184-derived (p15A) suicide vector (GmR)	[26]
pRK2013	Helper plasmid containing the ColE1 replicon with RK2 <i>tra</i> genes (KmR)	[27]
pGEM-T Easy	pUC origin, Multi Cloning Sites (MCS), <i>lacZ</i> gene fusion, β -lactamase coding region, <i>lac</i> operon sequences (ApR)	Promega
pBluescript KS+	ColEI, F1 Origin, MCS, <i>lacZ</i> gene fusion, β -lactamase coding region, <i>lac</i> operon sequences (ApR)	Stratagene
TOPO vecteur	pUC origin, région <i>Plac</i> promoteur, <i>lacZ</i> α - <i>ccdB</i> gene fusion. Topoisomerase enzyme (KmR, ApR)	Invitroge n

pPROBE-NT	pVS1-derived (p15a) vector, <i>gfp</i> (GmR)	[28]
pNT- <i>sahRp</i>	pPROBE NT containing <i>sahR</i> (PEG2239) promoter fused to <i>gfp</i> (KmR)	[15]
pNT- <i>nesRp</i>	pPROBE NT containing <i>nesR</i> (PEG5735) promoter fused to <i>gfp</i> (KmR)	[15]
pNT- <i>rsiA1p</i>	pPROBE NT containing <i>rsiA1</i> (PEG2540) promoter fused to <i>gfp</i> (KmR)	[15]
pNT- <i>rsiB1p</i>	pPROBE NT containing <i>rsiB1</i> (PEG2541) promoter fused to <i>gfp</i> (KmR)	[15]
pNT- <i>thuEp</i>	pPROBE NT containing <i>thuE</i> (PEG6268) promoter fused to <i>gfp</i> (KmR)	[15]
pNT- <i>otsBp</i>	pPROBE NT containing <i>otsB</i> (PEG4868) promoter fused to <i>gfp</i> (KmR)	[15]
pNT- <i>thuRp</i>	pPROBE NT containing <i>thuR</i> (PEG6269) promoter fused to <i>gfp</i> (KmR)	[15]
pNT- <i>iolCp</i>	pPROBE NT containing <i>iolC</i> (PEG373) promoter fused to <i>gfp</i> (KmR)	[15]
pNT- <i>iolRp</i>	pPROBE NT containing <i>iolR</i> (PEG374) promoter fused to <i>gfp</i> (KmR)	[15]

2.2. Mutants' constructions

The cloning of *iolR* (PEG374), *thuR* (PEG6269), *nesR* (PEG5735) and *otsB* (PEG4868) 5' and 3' regions were performed using a two-step PCR using primers listed in Table 2. The PCR were done using, as templates, genomic DNA extracts from LMR001 [29], using 10 mL of bacterial liquid cultures at early stationary phase. First, PCR reactions were performed to amplify separately both the 5' and the 3' flanking regions using the proof-reading enzyme Phusion™ (Thermo-Fischer) and recommended chemistry, except for primers (1 µM final concentrations for forward and reverse primers). The cycling conditions included a primary denaturation step of 30 s at 98 °C followed by 35 cycles (98 °C, 10 s/56 °C, 30 s/72 °C, 45 s) and a final elongation step of 7 min at 72 °C. PCR products

were purified using Illustra™ GFX PCR DNA and Gel Band purification kits (GE Healthcare), following manufacturer's recommendations. Then for each gene to be mutated, a second PCR reaction was set up, using as primers and matrix, purified PCR products from the 5' and 3' regions obtained that contain complementary sequences included in the 5' end of internal primers (see table 2), thus enabling priming of the amplification. The 25-μL reactions contained 5 μL of 5× GoTaq buffer (Promega), 1.25 μL of MgCl₂ (25 mM), 2.5 μL of dNTP mix (2.5 mM each), 0.125 μL GoTaq (5 U/μL), 2 μL of each purified PCR products obtained, and nuclease free water. A primary PCR cycling consisting of a denaturation step of 95 °C for 5 min followed by 5 cycles (95 °C, 30 s/50 °C, 45 s/72 °C, 50 s) and 5 min at 72 °C was performed so as to generate chimeric DNA molecules consisting of merged 5' and 3' regions. Then, 15 μL of the mix (3 μL of 5× GoTaq buffer, 0.75 μL of MgCl₂ (25 mM), 1.5 μL of dNTP mix, 0.125 μL of GoTaq (5 U/μL), 4 μL of external primers and 1.625 μL of nuclease free water) was added to the PCR reaction and subjected to the following cycling conditions: a denaturation step of 95 °C for 5 min followed by 25 cycles (95 °C, 30 s/55 °C, 45 s/72 °C, 50 s) and 5 min at 72 °C. The PCR products were purified from 1% agar gel using Promega Wizard® SV Gel and PCR Clean-Up System, following recommended protocol. As a result of the two amplifications' process, the PCR-amplified and -purified products were cloned as chimeric fragments into pGEM-T Easy Vector (Promega kit) and transferred into XL2-Blue Ultra-competent cells. The resulting positive clones were selected and the plasmids were extracted by using the Wizard® Plus SV Minipreps DNA Purification System, digested by XbaI enzyme and ligated into suicide vector pJQ200SK [26]. Enzymes were purchased from Promega. The latter plasmid was used to transform *Ensifer aridi* LMR001 cells by bipartite conjugation method by using *E. coli* S17-1 as donor (Simon et al., 1983) or by tri-parental mating procedure with *E. coli* DH5α [24] as donor and the helper plasmid pRK2013 [27]. Transconjugants were selected on TY plates containing Rif and 5% sucrose to allow selection of knockout clones that had undergone a double cross-over. Double cross-over gene replacement in mutants was verified by screening for antibiotic resistant/ sensibility

phenotypes (Rif and Gm, respectively) and also PCR to verify deletion by using specific external and internal primers.

Table 2. Primers used in this study.

Target gene (bp)	Primer*	Sequence (5' - 3')*	5' region size (bp) / Distance to start (nt)	3' region size (bp) / Distance to stop (nt)	Deletion (bp)
<i>iolR</i> (861)	IOLR-P374-A-XI	CCTCTAGACGAGATTAC GGCTCCAAGG TCAAAGACACCACAAGCT	457 / -10	658 / -183	689
	IOLR-P374-B-HIII	TCTCCAGCTTGCTCGTGTTC			
	IOLR-P374-C-HIII	AAGCTTGTGGTGTCTTT GACCGACTC			
	IOLR-P374-D-XI	CCTCTAGAGCAATTGTCG CGATAGAAGA			
<i>otsB</i> (750)	OSTB-P4868-A-XI	GGTCTAGATGCGCGATCT TCATGAACAA TTCATCCGTCAGAAAGCT	454 / +75	500 / -172	503
	OSTB-P4868-B-HIII	TGTCGGAGCAATGTGCGAG AAG			
	OSTB-P4868-C-HIII	AAGCTTTCTGACGGATG AAGGAATGTT			
	OSTB-P4868-D-XI	GGTCTAGAATAGCCGGAC ATCTCATGCC			
<i>thuR</i> (1020)	THUR-P6269-A-XI	GGTCTAGAGGTCCTCAGA AGCATTGTCA GTTCGTGAATCTCAAGCT	473 / +87	453 / -69	865
	THUR-P6269-B-HIII	TCTCGCTGACCTCGGGAT AG			
	THUR-P6269-C-HIII	AAGCTTGAGATTACGA ACTCTGGCC			
	THUR-P6269-D-XI	GGTCTAGATAGGTGCGAT GAACATGACG			
<i>nesR</i> (783)	NESR-P5735-A-XI	CGCTCTAGAGGCTCGATA TCACGCCAC TTAAGCCGCGGCGGTGAA	186 / +39	160 / +6	749
	NESR-P5735-B	CTGGTTCCTGATC			
	NESR-P5735-C	GCCGCGGCTTAAACACC			
	NESR-P5735-D-XI	GCGTCTAGAACGACAAGC GGCAGCTTGC			
<i>sahR</i> (1023)	SAHR-P2239-A-AI	GGGCCCCACAACATCTC GAAGG CTGCAGAGCATCCAAAC	1062 / +36	1132 / -14	973
	SAHR-P2239-B-PI	CAAGCG			
	SAHR-P2239-C-PI	CTGCAGGGGGAGTTTGA GAATGA			
	SAHR-P2239-D-XI	TCTAGACCGAATTAGGG ACTATAATTCCG			

* Bold font indicate restriction sites (XI: XbaI, HIII: HindIII, AI: ApaI, PI: PstI). Red font shows complementary sequences used for the fusion of 5' and 3' regions.

For the construction of the *sahR* (PEG2239) mutant, the 5' and 3' regions were first amplified using respectively the primer pairs SAHR-P2239-A-AI / SAHR-P2239-B-PI and

SAHR-P2239-C-PI / SAHR-P2239-D-XI (Table 2). PCR products were then cloned into pGEM-T Easy Vector following manufacturer's recommendations. Ligations were used to transform competent DH5 α cells. The plasmids were extracted and clones containing the inserts oriented from ApaI toward NsiI in the MCS were selected using plasmid and insert specific primers. DNA inserts in purified plasmids were sequenced and further digested with PstI. The clone that contained the 5' region generated a linearized plasmid that was further dephosphorylated and the plasmid that contained the 3' region generated a band corresponding to the linearized vector and a second band that corresponded to the 3' region cloned that was further extracted from an agarose gel. The latter insert was further ligated into the vector containing the 5' region PstI digested. Ligation mix was used to transform DH5 α cells. A clone containing the 5' and the 3' regions in the correct orientation was selected by positive screening using specific primers. Finally, the insert was excised using external restriction sites ApaI and XbaI and cloned into pJQ200SK prior conjugation into LMR001 to generate the $\Delta sahR$ strain as described above. Finally, the double mutant $\Delta sahR \Delta nesR$ was produced by mutating the $\Delta sahR$ strain upon double recombination of the pJQ200SK clone containing the *nesR* 5' and 3' regions as described above.

2.3. Phenotypic Characterization of the mutants

2.3.1. Plant Nodulation Assay

Seeds of *Vachellia tortilis* subsp. *raddiana* and endemic Moroccan *Vachellia gummifera* were surface sterilized, germinated and cultivated as described previously [15]. Briefly, after germination (5 days in the dark), young seedlings were transferred in pots containing sterilized sand and *Ensifer aridi* strains were inoculated using 1mL of washed cells at an optical density (600 nm) $\approx$ 0.7 corresponding to exponential growth phase. Six individual plants were inoculated with each rhizobial strain and sterile water was added for the negative and positive controls. Plants were grown at 26 °C with 16 hours day and 8 hours night and watered using BD medium for watering [30] supplemented with 3 mM

of KNO₃ for positive controls. Plants were harvested after 6 months of growth and the nodule number; plant shoot and root fresh weights were recorded.

2.3.2. Motility, tolerance to salt, detergent, acid and oxidative stresses

The mutants were phenotyped as described previously [15]. Briefly, bacterial motility was calculated by measuring the growth diameter of motile bacteria after 2, 3 or 5 days of incubation at 28 °C in soft TY agar (0.2 % W/V) complemented or not with NaCl at 0.1 or 0.25 M. Salt tolerance was estimated by growing bacteria in TY liquid medium and TY complemented with 0.25, 0.3 or 0.4 M NaCl and the OD was recorded after 0, 4, 8, 24, 56, 120, 144, 168 and 216 hours of growth at 28 °C with shaking. Five µL of 2 weeks old cultures were spotted onto TY agar plates to estimate long term survival of bacteria in these different media. Sensitivities of the mutants to oxidative, acid, and detergent stresses were compared to those of the wild type LMR001 strain. Briefly, bacterial cultures (pre-grown to OD 600 nm $\approx$ 0.7 in TY liquid medium) were added to 16 mL of soft TY agar containing 0.7 % of agar precooled to 45 °C and poured onto TY agar plates. Once solidified, a paper disc (6-mm diameter) soaked with 5 µL of H₂O₂ (2 M), HCl (5.5 M), or SDS (10 % w/v) was placed onto the surface of the plates that were incubated at 28 °C for three days prior to recording the diameters of the inhibition zones. Means of three replicates were calculated.

2.4. Transcriptional fusion assays

The promoter of selected genes were cloned previously in the broad host range mobilisable vector pROBE-NT and studied in *E. aridi* LMR001 and $\Delta rpoE2$ strains [15]. These included the promoter regions of le cycle de la methionine regulatory genes *nesR* (PEG5735), *sahR* (PEG2239), GSR regulatory genes *rsiA1* (PEG2540), *rsiB1* (PEG2541), genes involved in the trehalose import system and utilization *thuEFGKAB* (PEG6268-PEG6262) and its regulator *thuR* (PEG6269), or its endogenous biosynthesis *otsBA* (PEG4868, PEG4869), *treZ* (PEG5323) and finally genes involved in inositol catabolism *iolC* (PEG373) and its regulation *iolR* (PEG374). Here, to explore the mutations' effects on

the transcriptional regulation of these genes, pPROBE derivatives were transferred by biparental conjugation into the new mutants using S17-1 *E. coli* as donor. To estimate the relative expression of targeted genes' transcriptions, the median relative expression of the 11 primary measures obtained every half hour was calculated for each mutant upon correction to controls and normalization by OD (600 nm) as previously described [15] and the data obtained from two kinetics were averaged.

3. Results and Discussion

To know whether the mutated genes were important for symbiosis development, the mutants were inoculated on two compatible hosts *V. gummifera* and *V. tortilis* and their growth and nodulation were compared to those obtained with the LMR001 wild type strain (Figure 1).

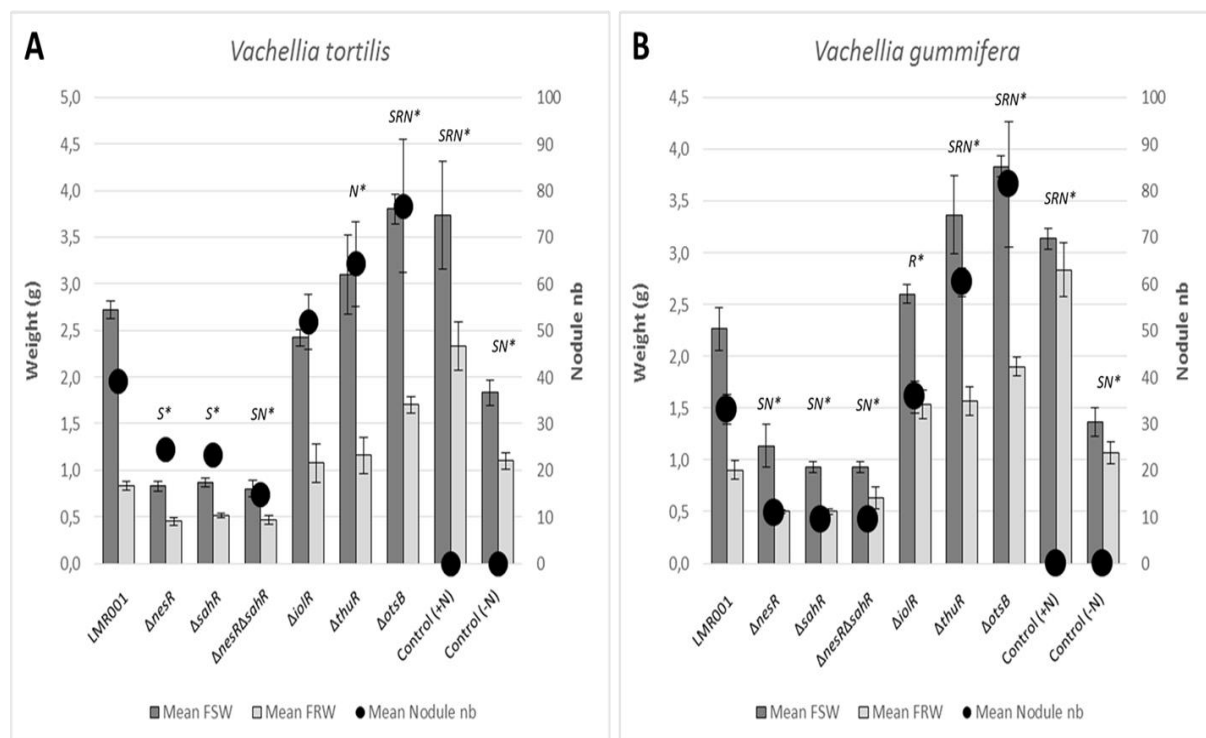


Figure 1. Effects of the mutations on symbiosis development with *Vachellia*. The Mean Fresh Shoot and root Weights (respectively shown as FSW, dark grey and FRW, light grey; left axis) and nodule numbers (filled black circles; right axis) in *V. tortilis* (A) and *V. gummifera* (B) are shown for the mutants and non-inoculated control plants either watered using nitrogen complemented nutrient solution (+N) or without nitrogen (-N) indicated at the bottom of the bar graphs. Error bars correspond to standard deviations and significant differences with LMR001

data (p -value $< 0,05$, Tukey's Test) are indicated by "*" "preceded by the letter "S", "R", and/or "N" for mean Shoot, Root weights and Nodule numbers respectively.

We found that deletions of genes involved in the regulation of the le cycle de la méthionine(*nesR* and *sahR*) reduced significantly the plant shoot weights for both *Vachellia* species as compared to the wild type of strain. Furthermore, the nodulation was also significantly lower to that of the LMR001 strain when the 2 regulatory genes were mutated in *V. tortilis*. In *V. gummifera*, all three mutants ($\Delta nesR$, $\Delta sahR$ and the double mutant $\Delta nesR\Delta sahR$) Significantly reduced the nodule number suggesting that methionine cycling regulation generating the universal methyl donor SAM is important for proper symbiosis development and functioning. The regulation of the le cycle de la methionine is complex and involves several transcription factors such as NesR and SahR or RNA regulatory elements (SAM riboswitch) that remain to be investigated in *Ensifer aridi*. Nevertheless, given strong homologies with the LuxR solo NesR of *S. meliloti* [31] and the ArsR family repressor SahR found in diverse proteobacteria including rhizobiales [32] for which a conserved DNA-binding domain could be identified upstream LMR001 *sahR*, *ahcY*, *metH*, *bhmT* and *metK* (Figure 2); our data suggest that a fine tuning of the SAM generating le cycle de la methionine is essential for symbiosis establishment and functioning. Interestingly, it was recently shown in *S. meliloti* that the SAM synthetase MetK was also binding stress inducible RNAs previously shown to interact with the RNA chaperone Hfq and more surprisingly other mRNAs [33] which suggests that MetK mediated post-transcriptional regulation may occur not only through epigenetics via the production of the universal methyl donor SAM but also by interacting with various RNAs and mRNAs.

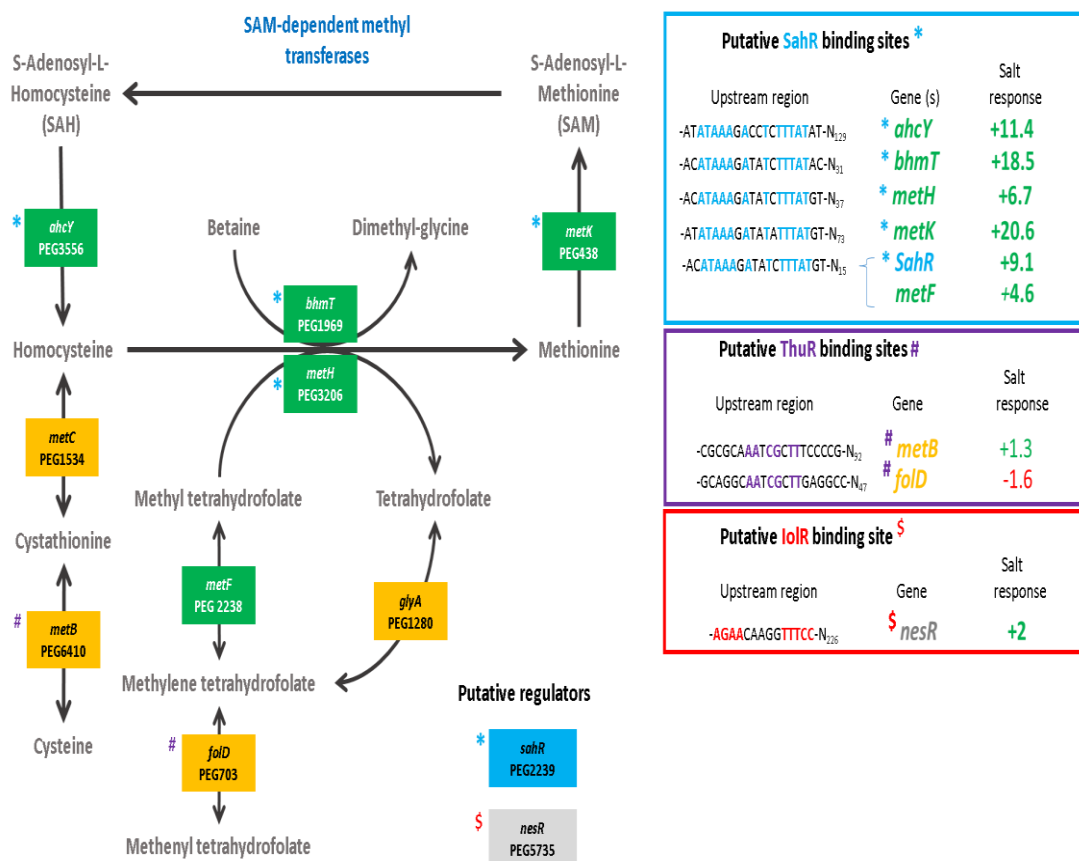


Figure 2. Metabolism of Methionine. *Ensifer aridi* predicted encoding genes (PEGs) involved in the metabolism of methionine are shown in the metabolic pathway. The presence of putative SahR, ThuR and IolR binding sites upstream of genes (shown in the right boxes) are indicated respectively by “ * ”, “ # ” and “ \$ ” on the left of genes when identified and the position of conserved residues are bolded and colored accordingly in boxes. The salt response refers to fold change previously found by RNAseq [15]. Positive fold changes indicating upregulation (green) and negative ones showing repression (red) in salt containing medium are bolded when above 2.

The plant test also showed that a mutation in *thuR* or *otsB*, involved in the trehalose transport and/or metabolism, stimulated nodulation with both plants tested while roots and shoots were also significantly increased in *V. gummifera*. Because ThuR is a LacI-GalR family regulator predicted to repress transport and catabolism genes (*thuEFGKAB*) through binding to upstream conserved motifs [34] also identified in *E. aridi* (Figure 3), the *thuR* deletion is expected to increase trehalose uptake capacity and utilization which may improve intracellular fitness and thus symbiotic efficiency, however, it remains to be functionally demonstrated. Strikingly, the *otsB* mutation significantly improved both

plant weight and nodule number with the two plants. The trehalose Phosphate Phosphatase OtsB is involved in the second step of the trehalose biosynthesis pathway which utilize Glucose-6P and UDP- glucose as primary substrates (Figure 3). A deletion of this gene should therefore result in an accumulation of trehalose-6P in the cell. Interestingly, an increased trehalose-6P synthase production by rhizobial strains genetically manipulated was reported to increase nodulation and plant biomasses in *Phaseolus vulgaris* [35] or chickpea upon salt stress [36]. Even though much remains to be done to characterize factors involved in Trehalose-6P mediated plant growth and stress tolerance improvements, this phosphosugar is gaining increasing interest. It was recently shown to act not only as a compatible solute but also as signaling molecule in plant notably by modulating SnRK1 phosphorylation status, a central protein kinase involved in the coordination of the plant development that can lead to improved growth notably under stress conditions [37, 38]. Additional work is required to determine whether trehalose-6P accumulates in the $\Delta otsB$ strain notably during the symbiosis development and functioning and whether the effects are linked to improved fitness and/or signaling.

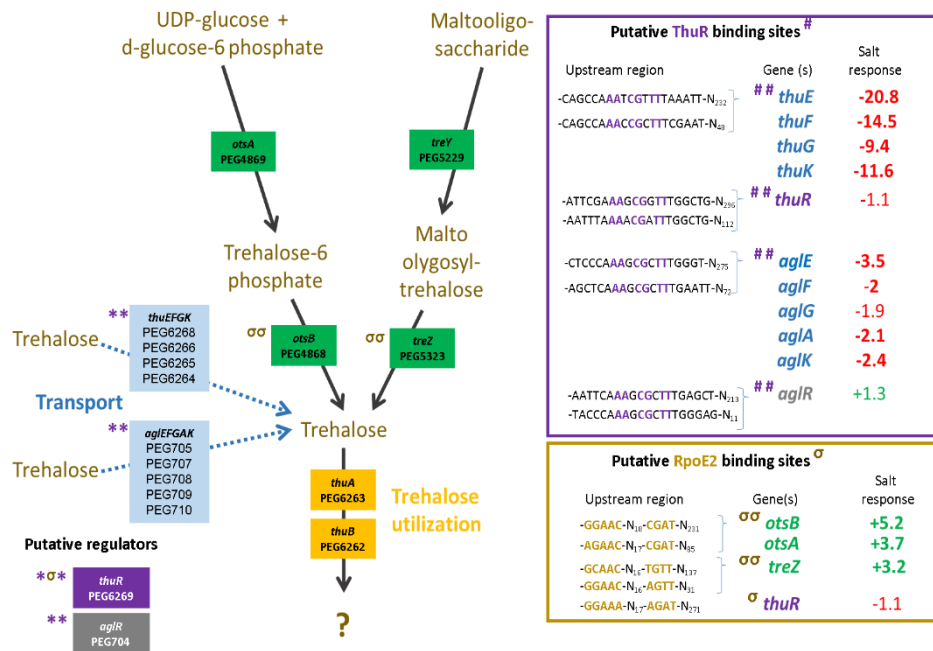


Figure 3. Metabolism of trehalose. *Ensifer aridi* predicted encoding genes (PEGs) involved in the metabolism and transport of trehalose are shown in the metabolic pathway. The presence of putative ThmR and RpoE2 binding sites upstream of

genes (shown in the right boxes) are indicated respectively by “ # ” and “ σ ” on the left of genes when identified and the position of conserved residues are bolded and colored accordingly in boxes. The salt response refers to fold change previously found by RNAseq [15]. Positive fold changes indicating upregulation (green) and negative ones showing repression (red) in salt containing medium are bolded when above 2.

The tolerance of the mutants to salt-induced osmotic upshift was compared to that of the LMR001 strain (Figure 4). The growth kinetics of tested strains in TY medium or when it was supplemented with 0.25 or 0.3 M NaCl were very similar. However, in the medium containing 0.4 M of salt, the growth of regulator mutants $\Delta nesR$, $\Delta sahR$ and the double mutant $\Delta nesR\Delta sahR$ were strongly impaired suggesting that a fine-tuning regulation of the le cycle de la methionine is important for salt tolerance in *Ensifer aridi*. The deletion of the trehalose uptake and utilization regulator *thuR* resulted in a continuous growth in the medium containing the highest salt concentration. Interestingly, while the *otsB* mutation resulted in reduced growth as compared to LMR001 in TY containing 0.4 M of salt after 2 days, the mutation enabled the bacterium to reach a higher OD at stationary phase which almost reached the density observed in other media suggesting that the strain impaired in trehalose phosphate phosphatase activity improves salt tolerance in LMR001. To estimate the viability of bacteria in tested media, 5 microliters of 2-weeks old cultures were spotted onto TY agar medium. While strains were all able to survive in media containing 0.3M NaCl or less, different growth phenotypes were found when bacteria were grown in the medium containing the highest salt concentration.

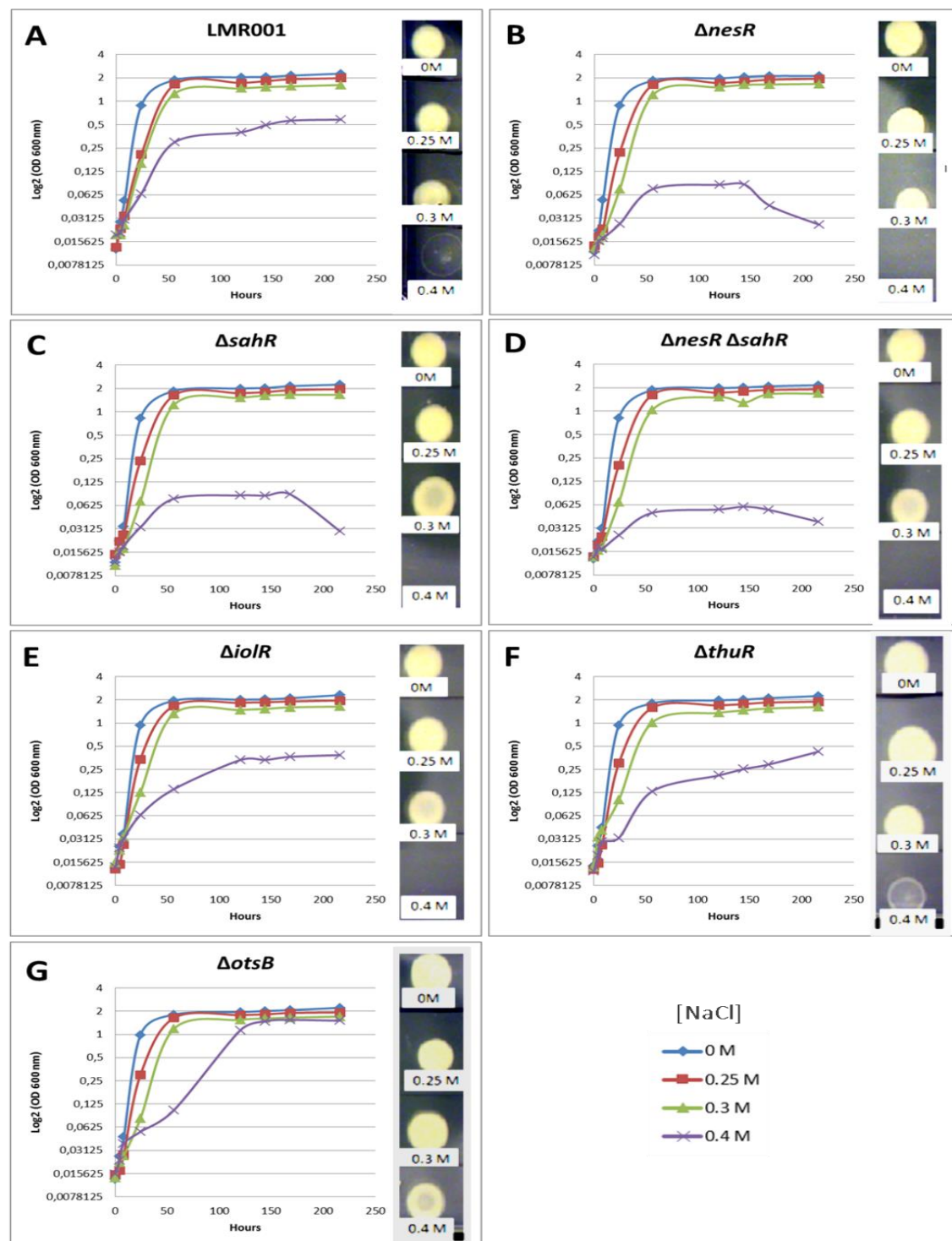


Figure 4. Effects of the mutations on salt tolerance. The growth kinetics are shown for the LMR001 (A), the strains $\Delta nesR$ (B), $\Delta sahR$ (C), $\Delta nesR \Delta sahR$ (D), $\Delta iolR$ (E), $\Delta thuR$ (F) and $\Delta otsB$ (G) in TY complemented with 0, 0.25, 0.3 or 0.4 M of NaCl (color code indicated at the bottom right of the figure). The OD (600 nm) corresponding to the mean of three replicates are shown using log 2 scale up to 9 days of growth (line graphs) and 5 μ L of 2-weeks old cultures were spotted onto

TY to estimate long term survival (pictures shown on the right of each line graph).

We found that after 2 weeks of growth in the medium supplemented with NaCl at 0.4 M, the LMR001 strain was still alive even though it showed a weak growth as compared to cells grown in other media. All 3 mutants involved in the cycle de la methionine regulation did not survive showing that at higher osmotic strength, a proper regulation of the methionine cycling generating the universal methyl donor SAM is essential. The $\Delta iolR$ strain that reached a lower OD at stationary phase as compared to LMR001 strain did not survive prolonged incubation in TY supplemented with 0.4 M NaCl. Interestingly, a deletion of *thuR* and more importantly *otsB* resulted in an improved long-term survival in the high salt concentrated medium as compared to the LMR001 strain. These results are globally in line with plant tests and further support the important role of stress responsive genes in symbiosis functioning [39] .

The motility of *Ensifer aridi* was previously shown to decrease in the presence of salt in the medium [15]. To test whether the mutated genes modified the swimming capacity of LMR001, their motility was compared to that of the LMR001 strain in soft agar plates that contained either TY medium or TY complemented with 0.1 or 0.25 M of NaCl. The Figure 5 shows the results of the swimming assay for all strains. As shown for the LMR001 strain, the presence of salt in the medium reduced the motility in a dose dependent manner. However, we found that after 5 days of growth in the TY medium, the $\Delta otsB$ showed continuous motility that goes past that of the LMR001 strain. Whether this phenotype results from an increased trehalose-6P remains to be addressed. A similar phenotype was found for the double mutant $\Delta nesR \Delta sahR$ when, in contrast, single mutants of the cycle de la methionine regulators $\Delta nesR$ and $\Delta sahR$ presented a lower motility. These results are surprising as we would expect that the double mutation should result in a stronger effect. When salt was added in the medium, the motility was more strongly reduced for the double mutant as compared to the LMR001 strain and the single mutants and even totally impaired in the medium containing 0.25 M of salt showing the importance for the bacterium to finely tune the cycle de la methionine for proper

functioning. Given the involvement of methyl-accepting chemotaxis proteins in the regulation of diverse cellular activities including flagellum biosynthesis, it is not surprising that an alteration of the SAM generating methionine salvage pathway regulation generates motility phenotypes. Nevertheless, additional work is required to characterize the effects of the mutations on SAM and whether these are responsible for the observed motility effects. Finally, after 5 days of growth, the $\Delta iolR$ and the $\Delta thuR$ mutants showed significant lower motility in the medium containing 0.1 M of NaCl as compared to the LMR001 strain suggesting that both trehalose and inositol affect swimming, however, mechanisms underlying these effects remain to be determined.

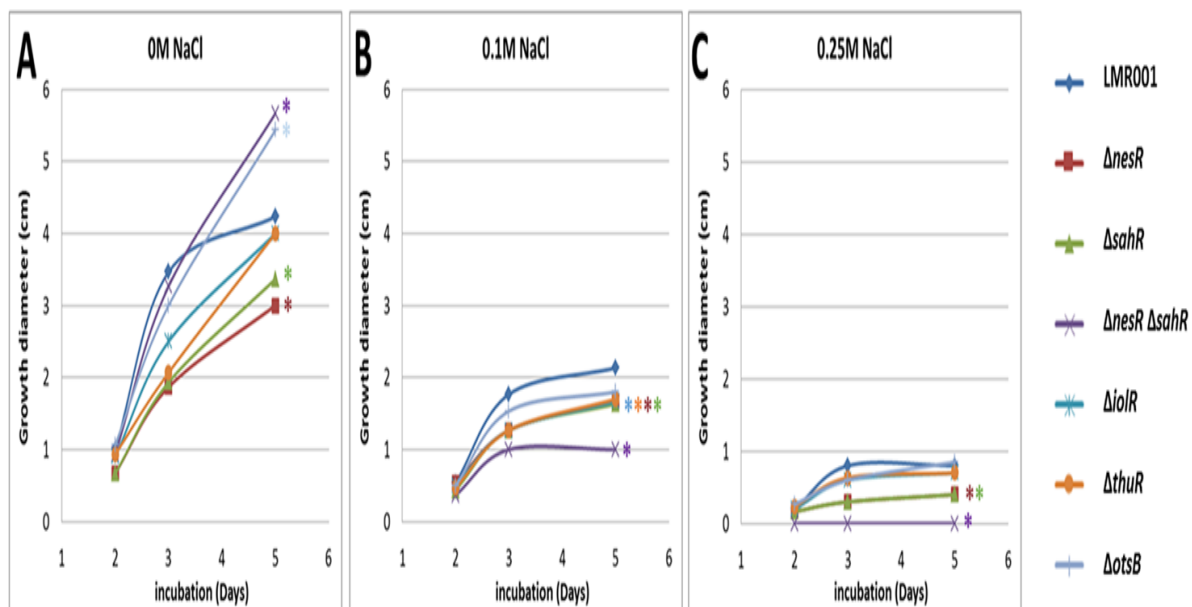


Figure 5. Effects of the mutations and osmotic strength on bacterial swimming. The motility was estimated by measuring the growth diameter from triplicate TY soft agar plates supplemented with salt 0 M (A), 0.1 M (B) or 0.25 M (C). The mean motility after 2, 3 and 5 days are shown for the LMR001 and the 6 mutants (indicated on the right of the figure). Significant differences with LMR001 motility after 5 days (p-value < 0.05, Tukey's Test) are indicated by "*" on the right of lines using the same color code.

The tolerance to detergent (SDS), oxidative (H_2O_2) and acid (HCl) stresses of the studied strains was evaluated by comparing the growth inhibition zone resulting from their diffusion from a disc soaked with these chemicals (See materials and methods section). We found that only a double mutation of the two le cycle de la methionine

regulators resulted in a significant difference when compared to the LMR001 strain (Figure 6). The reduced inhibition zone found in the double mutant suggests that its membrane structure is altered and result in improved resistance to the three stresses which require additional work to be characterized.

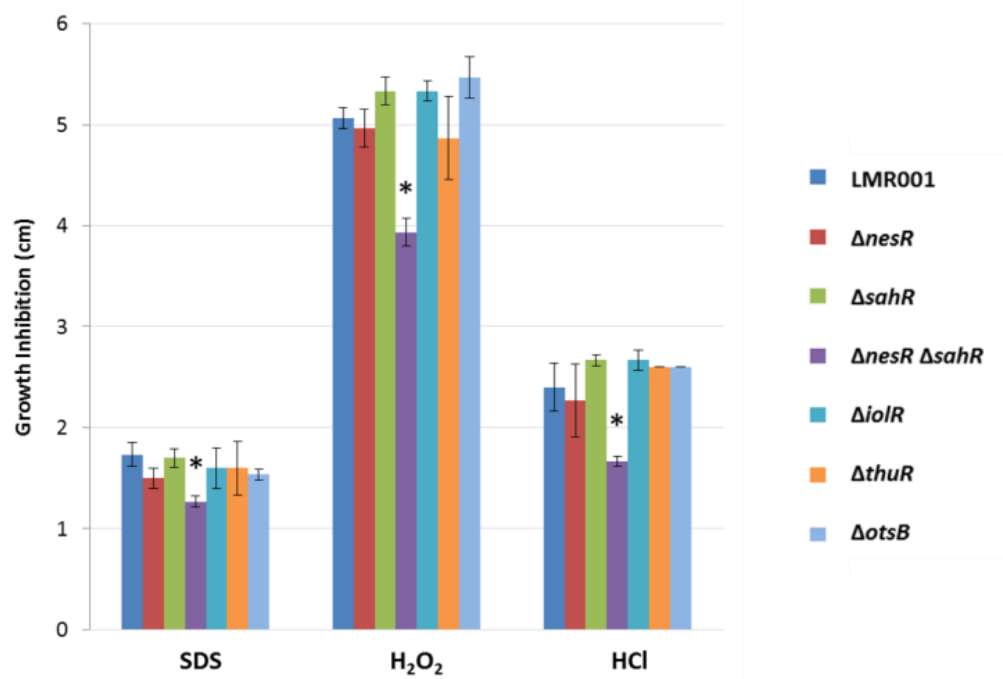


Figure 6. Effects of the mutations on detergent, oxidative and acids stresses. The mean diameters of the growth inhibition zone are shown for the 6 mutants (indicated on the right of the figure). The size of the inhibition zone is inversely correlated to the tolerance as the chemicals (indicted at the bottom of the graph) diffuse away from the disc into the agar (See materials and methods). Error bars correspond to standard deviations and significant differences with LMR001 data (p-value < 0,05, Tukey's Test) are indicated by "*" above the corresponding bar.

To further describe the effects mutations had on the transcriptional control of the inositol, the methionine and the trehalose metabolisms, the relative expression levels of selected genes were assessed using the pPROBE derivatives previously screened in the $\Delta rpoE2$ strain [15]. The figure 7 shows the mean relative expression obtained for all mutants. We found that in the $\Delta iolR$ and $\Delta thuR$ mutant backgrounds, *iolR* promoter activity was higher than in the wild type strain while in other mutants, the relative expression remained low and close to the background level (Figure 7A). The regulator *IolR* is a RpiR-like repressor shown to be involved in the regulation of the inositol

catabolism in *S. meliloti* [40, 41]. A regulatory model was proposed which involves IolR binding to promoter regions of target genes that are released when the intermediary product of the inositol catabolism 2-deoxy-5-keto-D-gluconic acid 6-phosphate (DKGP) is produced [41]. The promoters of genes involved in studied pathways of *E. aridi* were therefore screened for putative regulator binding sites identified by Kohler and colleagues [41]. Interestingly, we found putative binding sites upstream of *idhA*, *iolY*, *iolR* and *iolC* in LMR001 (Figure 8) suggesting a similar regulation in *Ensifer aridi* as all these genes were previously found to be induced upon salt induced hyperosmotic stress [15] (Figure 8). In line with this model, we found that the mutation of *iolR* resulted in an induction of its own transcription supporting a feedback regulatory loop (Figure 7A). However, we found that the *iolC* transcription was not induced in the *iolR* mutant background (Figure 7B) indicating that IolR is necessary for proper *iolC* transcription thus suggesting a complex regulation of the inositol catabolism that may involve additional factors for proper regulation. Nevertheless, it is also possible that the *iolR* deletion resulted in an unexpected modification of the downstream genes involved in inositol catabolism that could impair the production of the product effector KDGP suggested to be required for antagonizing IolR repression. It should be noted that Kohler and colleagues [41] showed that even though IolR binding domains were found upstream the *iolY*, *iolR* and *iolC* genes, all the *iolYRCDEB* genes were transcribed as a single mRNA forming an operon. Recently, a *iolR* mutant in *Aeromonas* was shown to act both as a repressor and as an activator for a high number of genes [42]. Interestingly, a sequence close to the consensus IolR binding site was found in the intergenic region of the *rsiA1-rsiB1* genes involved in the GSR (Figure S1). This putative IolR binding site overlaps with the predicted RpoE2 binding motif [20, 43] located upstream of *rsiB1* (Figure S1) which may suggest a possible implication of the IolR repressor in the modulation of the GSR. It should be noted that 2 motifs close to the consensus sequence of the *ecfG* sigma factor RpoE2 binding site were also predicted upstream of the *iolR* gene, the IolR putative binding site overlapping the primary putative RpoE2 binding site (Figure 8) which would support a possible cross regulation. By screening the genome of *E. aridi* for the presence

of additional putative IolR binding sites, we found several consensus sequences. Among them, we identified the sequence motif upstream of a predicted ABC type inositol transport system (PEG5307-5306-5305) that was also found induced by salt stress [15] (Figure 8).

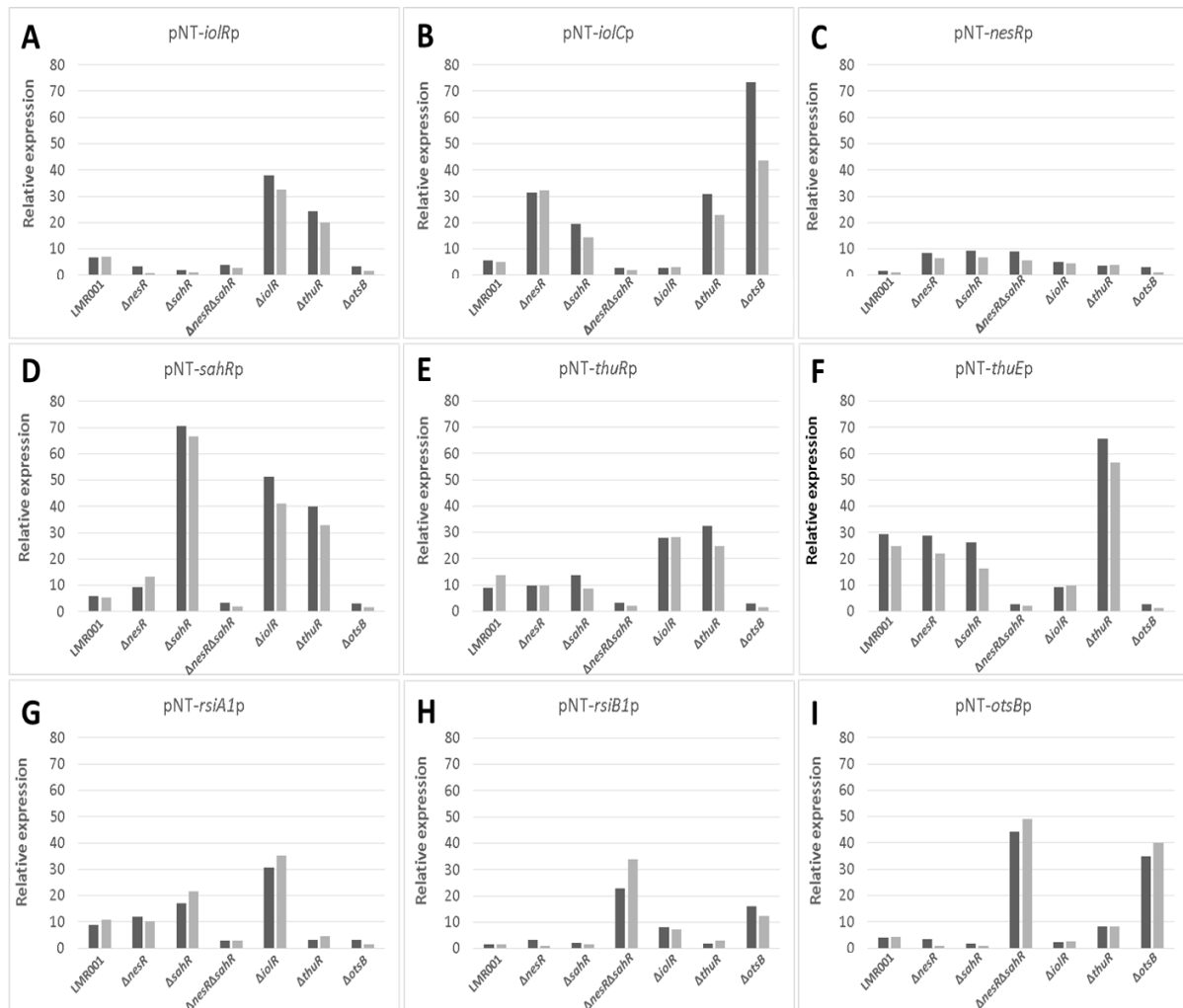


Figure 7. Effects of the mutations on the relative expression level of selected genes using promoter-gfp fusion reporter system. The relative expression of *iolR* (A), *iolC* (B), *nesR* (C), *sahR* (D), *thuR* (E), *thuE* (F), *rsiA1* (G), *rsiB1* (H) and *otsB* (I) in the studied strains (indicated at the bottom of each bar graph) in TY or TY complemented with salt (250 mM NaCl) are shown as dark and light grey bars respectively. The mean relative expressions were calculated using the median fluorescence of 11 measures recorded every half hour that were further corrected to controls and normalized to the OD (600 nm).

Furthermore, another putative IolR binding site was also found upstream of PEG771, a LacI type transcriptional repressor predicted to control inositol transport and

utilization in *R. etli* [34]. Interestingly, a regulatory motif predicted as a the binding motif of *R. etli* IolR (RHE_CH02362) by Regprecise was found twice in the intergenic region of PEG771-772 suggesting that inositol metabolism and transport regulation involves additional genes in *E. aridi* that may be modulated by IolR. Among putative IolR binding sites, we also found a putative IolR binding site upstream of the *nesR* start codon (Figure 2) which could suggest IolR-dependent le cycle de la methionine regulation. However, this motif was located within the 5' end of the *pip* gene and *nesR* transcription using the GFP reporter system was close to background levels which make it difficult to assess. We also found that the *iolC* transcription was higher in the strains $\Delta nesR$, $\Delta sahR$, $\Delta thuR$ and $\Delta otsB$ while its transcription was close to background levels in the double mutant $\Delta nesR\Delta sahR$ (Figure 7B). Whether these increased *iolC* transcriptions are directly linked to the methionine and trehalose cycles remains to be determined but these data strongly support a complex regulation of these metabolisms that appear interdependent and additional work is required to test the targets of IolR that may not be restricted to the regulation of inositol catabolism.

The relative expression of *nesR*, encoding a LuxR solo showed to positively regulate the méthioninesalvage pathway generating SAM in *S. meliloti* [31] appeared more expressed in the $\Delta nesR$, the $\Delta sahR$ strains and the double mutant $\Delta nesR\Delta sahR$ as compared to the other strains (Figure 7C). However, the expression remained low which makes it difficult to interpret. In contrast, the alphaproteobacterial ARS-type regulator SahR predicted to act as a repressor [32] showed contrasting relative expression levels in the various strains (Figure 7D). Novichkov and colleagues showed in *Desulfovibrio* that SahR binds to operators of genes involved in SAM cycling depending on cellular level of the effector S-Adenosyl Homocysteine (SAH) [32]. Here, the highest *sahR* relative expression was found in the $\Delta sahR$ strain supporting autoregulation. Accordingly, the promoter region of *sahR* in LMR001 possessed the predicted binding motif conserved among rhizobia that could also be identified upstream of the known target genes *ahcY*, *metH*, *metK* and *bhmT* [32] (Figure 2). Given the fact that all these genes were found upregulated by salt previously [15], it is tempting to speculate on the role of SahR as a

repressor of the SAM cycling in *E. aridi*. This raises the question regarding the regulatory mechanisms of this important pathway that involves, once again several regulators, which is nevertheless not surprising given the importance of SAM as a universal methyl donor and the downstream activities it is involved in for cell functioning. Screening of the LMR001 genome for the presence of other putative SahR binding sites did not return any other hit suggesting that SahR is primarily involved in SAM cycling regulation. However, the *sahR* promoter activity appeared also increased in the $\Delta iolR$ and the $\Delta thuR$ strains as previously shown in $\Delta rpoE2$ [15] which further shows the pleiotropic effects alterations of the methionine cycling have on the other metabolisms.

Regarding the relative expression of *thuR* (Figure 7E) and *thuE* (Figure 7F), we previously found that *thuR* transcription was reduced in the $\Delta rpoE2$ strain in contrast to that of *thuE* which supported a function of ThuR as a repressor of its own transcription and that of the *thuEFGKAB* operon that possess 2 predicted ThuR binding motifs in their promoter regions (Figure 3) [15, 34]. Here, we found that the *thuR* mutation resulted in inductions of both *thuR* and *thuE* while the induction of *thuR* in the $\Delta iolR$ strain showed a concomitant repression of *thuE* when compared to LMR001. However, a strong repression of *thuR* in the $\Delta otsB$ and the double mutant $\Delta nesR\Delta sahR$ were found which did not result in *thuE* inductions. Nevertheless, as shown in other LacI type regulators such as the close regulator AglR, the DNA binding of such type of transcription factors depends on the presence of specific sugars as effectors whose concentrations and availability may be altered in the later strains. Indeed, given the inductibility of *thuEFGK* by trehalose [43], it appears probable that the trehalose is involved in ThuR binding affinity to target promoters and a mutation of the trehalose-6P Phosphatase may alter the availability of the effector to the repressor. Furthermore, as shown above, the double mutant $\Delta nesR\Delta sahR$ presented a stronger tolerance to detergent, oxidative and acid stresses suggesting a modification of the cell wall in this mutant that may modify cellular exchanges that could alter the availability of effectors to the regulator. However, additional work is required to further characterize the regulation of *thu* genes. It is interesting to note the presence of a motif (GGAAA-N17-AGAT) that resemble RpoE2

sigma factor binding domains between these two ThuR recognition motifs upstream the *thuR* start codon which would suggest that *thuR* transcription is RpoE2 mediated, in accordance with *thuR* transcriptional repression reported in the $\Delta rpoE2$ strain [15]. The genome of the LMR001 strain was also screened for additional putative ThuR binding sites [34]. Among the conserved residues that matched LMR001 intergenic regions, putative binding motifs were found upstream of *folD* and *metC* involved in the methionine metabolism (Figure 2). Interestingly, in $\Delta thuR$, *sahR* transcription was induced which was shown to be activated through interaction of the SAH, the intermediary product of the le cycle de la methionine acting as SahR effector [32], supporting implication of ThuR in the methionine cycling. Nevertheless, whether ThuR is directly involved in the regulation of methionine precursors' homeostasis remains to be determined experimentally. In addition, we also found the putative ThuR binding motif once upstream of the *iolR* gene and twice before the *iolY* start codon (Figure 8) that supports the involvement of ThuR in the regulation of the inositol catabolism. Again, *iolR* transcription was also found activated in $\Delta thuR$ further supporting interplay between regulators of the methionine and trehalose metabolisms.

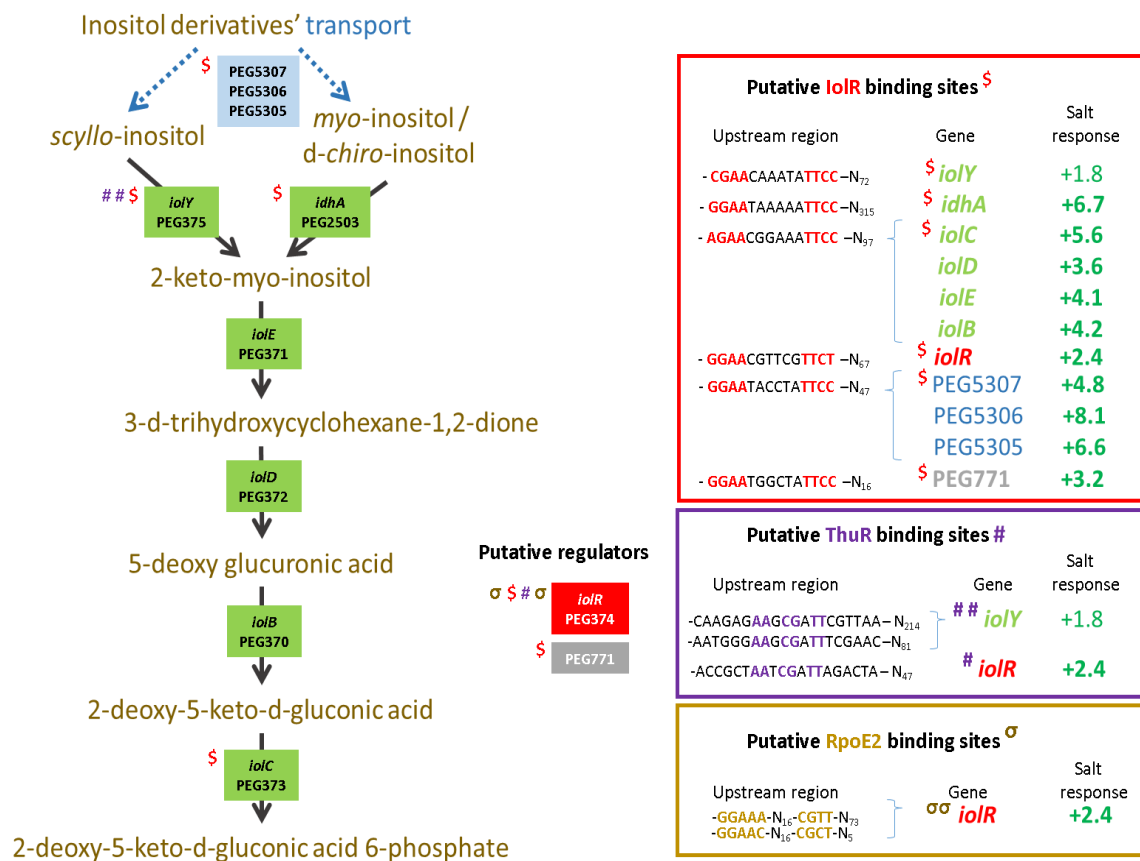


Figure 8. Catabolism of Inositol. *Ensifer aridi* predicted encoding genes (PEGs) involved in the transport and catabolism of inositol are shown in the metabolic pathway. The presence of putative *IolR*, *ThuR* and *RpoE2* binding sites upstream of genes (shown on the right boxes) are indicated respectively by “\$”, “#” and “σ” on the left of genes when identified and the position of conserved residues are bolded and colored accordingly in boxes. The salt response refers to fold change previously found by RNAseq [15]. Positive fold changes indicating upregulation (green) in salt containing medium are bolded when ≥ 2 .

Among putative *ThuR* binding sites identified, we also found the presence of two motifs upstream of the PEP carboxykinase encoding gene *pcKA* (PEG3563) (data not shown) that was also found strongly repressed by salt [15]. Interestingly Di Cenzo and colleagues [45] showed that in *S. meliloti*, *pcKA* was regulated by PckR, another LacI-type transcriptional regulator that binds to a motif containing the following conserved residues (5'- TNNAANCGNTT -3') that overlaps the predicted *ThuR* binding conserved residues. Previous experimental data suggested a complex regulatory mechanism that relied in the presence of simple or double binding of the PckR regulator in the target promoter sequences which would result respectively in an induction or a repression of

regulated genes in the presence of PEP, and inversely in the absence of the effector [45]. PckR was shown to regulate several genes of the glycolytic and gluconeogenic pathways among which *pckA* and *fbaB* that possessed 2 PckR binding domains, and *zwf*, *mgsA* and *eda2* that possessed a single PckR binding domain that were inversely regulated as compared to *pckA* and *fbaB*. Similarly, in *E. aridi*, both the presence and the number of PckR (PEG3052 in LMR001) binding domains could be identified in the promoter regions of these genes (data not shown) and similar expression profiles were found in *E. aridi*. Indeed, we found previously that a salt stress resulted in a strong repression of both *pckA* (PEG3563) and *fbaB* (PEG2954) and an induction of *zwf* (PEG714), *mgsA* (PEG118) and *eda2* (PEG3069) [15] which corroborates Di Cenzo and colleagues' work on *S. meliloti* and would suggest, according to their regulatory model, that under salt stress, *E. aridi* cells contain high levels of PEP. It should be noted that one of the 2 predicted ThuR binding sites upstream of *thuR* and *aglE* contained the PckR conserved motif suggesting a possible involvement of PckR in ThuR mediated regulation. di Cenzo and colleagues proposed several scenarios including one that was based on a displacement of another negative regulator thus enabling RNA polymerase to access the promoter. Given the fact that the predicted PckR binding motif includes conserved nucleotides of the ThuR binding motif, the ThuR regulator could be involved in modulating PckR regulated genes, however, much work remains to be done to further characterize this complex regulatory mechanism and a possible interplay.

The transcriptional regulation of *rsiA1* and *rsiB1* were also studied in the various LMR001 derivatives (Figure 7G and 7H respectively). Regarding the *rsiA1* relative expression, we found that in comparison to the LMR001 strain, three mutants ($\Delta nesR\Delta sahR$, $\Delta thuR$ and $\Delta otsB$) presented a lower expression while in contrast $\Delta sahR$ and more importantly $\Delta iolR$ presented a higher relative expression. Regarding the *rsiB1* expression, three mutants ($\Delta nesR\Delta sahR$, $\Delta iolR$ and $\Delta otsB$) presented an increased transcription level as compared with LMR001. It is interesting to note that both *rsiA1* and *rsiB1* relative expressions were increased in the $\Delta iolR$ strain as compared to the WT. Even if additional experimental data are required, the presence of a putative IolR binding site

located in their promoter regions may be involved in modulating their transcriptional regulation and supports a role of IolR in GSR tuning. Surprisingly, both $\Delta nesR\Delta sahR$ and $\Delta otsB$ strains presented a reduced expression of *rsiA1* while we observed stronger expressions of *rsiB1* as compared to the WT strain. Because the anti-anti-sigma factor RsiB1 promotes RpoE2 recruitment to its operators through competitive binding to the anti-sigma factor RsiA1, our results suggest that GSR regulation may be altered in these mutants. Nevertheless, *rpoE2* is co-transcribed with *rsiA1* and the GSR is finely tuned and requires RsiB1 phosphorylation through a phosphorylation cascade that is yet not fully characterized, and studies are needed to functionally demonstrate the effects that these mutations may have on GSR. Interestingly, both strains presented a strong induction of the *otsBA* operon (Figure 7I) shown to be RpoE2 mediated suggesting that GSR was activated in these mutants. Surprisingly, with the exception of the *iolC* gene that was highly induced upon *otsB* deletion, the relative expression profiles obtained for $\Delta nesR\Delta sahR$ and $\Delta otsB$ strains were highly similar which contrast with their responses to stress or their symbiotic performances. The fact that the *otsB* deletion resulted in an induction of the *otsBA* operon suggests that in this mutant the cells accumulate the phosphor-sugar trehalose-6P as the phosphatase OtsB is nonfunctional. Whether tolerance to salt and symbiosis efficiency improvements are due to higher level of trehalose-6P in the $\Delta otsB$ strain remains to be determined but further support that trehalose and its phosphorylated derivative are key metabolites for stress adaptation and plant growth promotion notably under abiotic stress [46]. Interestingly, Sengupta and colleagues previously showed in yeast that trehalose phosphate synthase methylation by a Cystein methyl transferase resulted in increased trehalose-6 phosphate biosynthesis [47] further supporting possible interplay between SAM generating methionine cycling and trehalose metabolism. Much work is however necessary to test whether such type of post-translational regulation also occurs in bacteria.

4. Conclusions

In order to decipher the complex response of *E. aridi* to an osmotic upshift, a previous study based on RNA-seq enabled the identification of several metabolisms that were altered by salt. Despite activation of regulatory genes involved in the GSR, a deletion of the extra cytoplasmic function *rpoE2* sigma factor encoding gene did not result in phenotypic alterations which suggested that alternative regulatory mechanisms could be recruited [15]. Here, a series of mutants in genes involved in the metabolisms of trehalose (*thuR*, *otsB*), inositol (*iolR*) and SAM generating methionine (*nesR* and *sahR*) that appeared regulated upon salt stress were produced and phenotyped. To further describe the possible regulatory circuits between these metabolisms that were differentially regulated upon salt stress in the wild type LMR001 strain, the transcriptional state of selected genes and notably regulators involved in these metabolisms were assessed in these mutant backgrounds using a promoter-gfp fusion based reporter system. The results show that mutations in genes involved in the methionine cycling regulation altered the symbiosis development and functioning with two compatible hosts suggesting that a fine tuned regulation of the methionine and possibly SAM is required. These mutants also showed a reduced growth in a medium containing 0.4 M of NaCl and their motility were more strongly inhibited by salt than that of the wild type LMR001 strain which may reduce plant colonization and in planta rhizobial fitness. Surprisingly, when the two genes *nesR* and *sahR* were deleted, the strain appeared more tolerant to detergent, oxidative or acid stresses suggesting an alteration of the cellular membranes. In contrast, mutants in the *thuR* or *otsB* genes resulted in improved symbiosis showing that a modulation of the trehalose metabolism represents an interesting way to improve symbiosis and thus plant growth which was also correlated with an improved growth in a medium containing 0.4 M of NaCl. The transcription of genes encoding repressors *IolR*, *ThuR* and *SahR* were induced when they were deleted suggesting autoregulation which was shown in other bacteria for *IolR* and *SahR* and predicted for *ThuR*. We also found that *IolR* was required for proper *iolC* induction, that was, in contrast, induced in $\Delta nesR$, $\Delta sahR$; $\Delta thuR$ and $\Delta otsB$ suggesting that trehalose and methionine metabolisms regulation can alter inositol catabolism that necessitate however additional experimental evidences.

Nevertheless, if *nesR* promoter activity remained low in all mutant backgrounds, we also found that the transcriptional state of the putative SahR repressor was also induced upon *thuR* or *iolR* mutations which further demonstrate interplay between studied metabolisms' regulations. We also found that deletion of the *iolR* regulatory gene altered *thuR* and *thuE* transcriptions. When the transcriptional regulation of the *rsiA1* and *rsiB1* genes was assessed in the studied mutants, different profiles were found depending on the mutant. Indeed, the *thuR* deletion resulted in a lower transcription of *rsiA1* that was also found in the $\DeltaotsB$ and $\Delta nesR\Delta sahR$ strains. In contrast, the two later strains presented a strong induction of the *rsiB1* gene suggesting that despite no or little effect single mutants $\Delta nesR$ or $\Delta sahR$ had on these genes' transcriptions, a deregulation of the methionine and trehalose metabolisms alter the GSR. Finally, GSR may also be influenced by Inositol as both *rsiA1* and *rsiB1* genes appeared induced in the $\Delta iolR$ strain. Whether these regulations are under the direct control of studied transcriptional regulators, that possessed in some cases several putative binding motifs in their promoter regions, or indirect through the involvement of yet unknown factors may now be further explored. Nevertheless, the present study clearly shows that the stress responsive genes are regulated by complex regulatory circuits and suggest interplay of stress responsive genes involved in methionine, inositol and trehalose regulatory pathways.

Supplementary Materials: The following are available online at www.mdpi.com/xxx/s1, Figure S1: Genetic map of the GSR regulatory genes *rsiA1*, *rsiB1* and *rpoE2*.

Author Contributions: Conceptualization, AFM, ALQ; methodology, MB, ALQ; formal analysis, MB, AFM, ALQ.; writing—original MB, ALQ; writing—review and editing, MB, AFM, ALQ; supervision AFM, ALQ; project administration, AFM; funding acquisition, AFM and ALQ. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the French Institute of Research for the Development (IRD) through the LMI LBMV and an ARTS PhD grant for M.B.

Acknowledgments: The authors would like to acknowledge Samir El Qaidi and Hassan Ait-Benhassou for methodological support for the construction of the mutants and Abderrahim Ferradous from the Centre Régional de Recherches Forestière in Marrakech for *Vachellia* seeds.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Hakim, S.; Imran, A.; Mirza, M. S. Phylogenetic Diversity Analysis Reveals *Bradyrhizobium Yuanmingense* and *Ensifer aridi* as Major Symbionts of Mung Bean (*Vigna radiata* L.) in Pakistan. *Braz. J. Microbiol.* **2021**, *52* (1), 311–324. <https://doi.org/10.1007/s42770-020-00397-9>.
2. Lamin, H.; Alami, S.; Bouhnik, O.; ElFaik, S.; Abdelmoumen, H.; Bedmar, E. J.; Missbah-El Idrissi, M. Nodulation of *Retama monosperma* by *Ensifer aridi* in an Abandoned Lead Mine Soils in Eastern Morocco. *Front. Microbiol.* **2019**, *10*, 1456. <https://doi.org/10.3389/fmicb.2019.01456>.
3. Rocha, G.; Le Queré, A.; Medina, A.; Cuéllar, A.; Contreras, J.-L.; Carreño, R.; Bustillos, R.; Muñoz-Rojas, J.; Villegas, M. del C.; Chaintreuil, C.; Dreyfus, B.; Munive, J.-A. Diversity and Phenotypic Analyses of Salt- and Heat-Tolerant Wild Bean *Phaseolus filiformis* Rhizobia Native of a Sand Beach in Baja California and Description of *Ensifer aridi* Sp. Nov. *Arch. Microbiol.* **2020**, *202* (2), 309–322. <https://doi.org/10.1007/s00203-019-01744-7>.
4. Sakrouhi, I.; Belfquih, M.; Sbabou, L.; Moulin, P.; Bena, G.; Filali-Maltouf, A.; Le Queré, A. Recovery of Symbiotic Nitrogen Fixing Acacia Rhizobia from Merzouga Desert Sand Dunes in South East Morocco – Identification of a Probable New Species of *Ensifer* Adapted to Stressed Environments. *Syst. Appl. Microbiol.* **2016**, *39* (2), 122–131. <https://doi.org/10.1016/j.syapm.2016.01.001>.

5. Tak, N.; Awasthi, E.; Bissa, G.; Meghwal, R. R.; James, E. K.; Sprent, J. S.; Gehlot, H. S. Multi Locus Sequence Analysis and Symbiotic Characterization of Novel *Ensifer* Strains Nodulating *Tephrosia* Spp. in the Indian Thar Desert. *Syst. Appl. Microbiol.* **2016**, *39* (8), 534–545. <https://doi.org/10.1016/j.syapm.2016.08.002>.
6. Le Quéré, A.; Tak, N.; Gehlot, H. S.; Lavire, C.; Meyer, T.; Chapulliot, D.; Rathi, S.; Sakrouhi, I.; Rocha, G.; Rohmer, M.; Severac, D.; Filali-Maltouf, A.; Munive, J.-A. Genomic Characterization of *Ensifer aridi*, a Proposed New Species of Nitrogen-Fixing Rhizobium Recovered from Asian, African and American Deserts. *BMC Genomics* **2017**, *18* (1), 85. <https://doi.org/10.1186/s12864-016-3447-y>.
7. Cytryn, E. J.; Sangurdekar, D. P.; Streeter, J. G.; Franck, W. L.; Chang, W.; Stacey, G.; Emerich, D. W.; Joshi, T.; Xu, D.; Sadowsky, M. J. Transcriptional and Physiological Responses of *Bradyrhizobium Japonicum* to Desiccation-Induced Stress. *J. Bacteriol.* **2007**, *189* (19), 6751–6762. <https://doi.org/10.1128/JB.00533-07>.
8. López-Leal, G.; Tabche, M.; Castillo-Ramírez, S.; Mendoza-Vargas, A.; Ramírez-Romero, M. A.; Dávila, G. RNA-Seq Analysis of the Multipartite Genome of *Rhizobium etli* CE3 Shows Different Replicon Contributions under Heat and Saline Shock. *BMC Genomics* **2014**, *15* (1), 770. <https://doi.org/10.1186/1471-2164-15-770>.
9. Pérez-Montaña, F.; del Cerro, P.; Jiménez-Guerrero, I.; López-Baena, F. J.; Cubo, M. T.; Hungria, M.; Megías, M.; Ollero, F. J. RNA-Seq Analysis of the *Rhizobium tropici* CIAT 899 Transcriptome Shows Similarities in the Activation Patterns of Symbiotic Genes in the Presence of Apigenin and Salt. *BMC Genomics* **2016**, *17* (1), 198. <https://doi.org/10.1186/s12864-016-2543-3>.
10. Alexandre, A.; Laranjo, M.; Oliveira, S. Global Transcriptional Response to Heat Shock of the Legume Symbiont *Mesorhizobium loti* MAFF303099 Comprises Extensive Gene Downregulation. *DNA Res.* **2014**, *21* (2), 195–206. <https://doi.org/10.1093/dnares/dst050>.
11. Laranjo, M.; Alexandre, A.; Oliveira, S. Genes Commonly Involved in Acid Tolerance Are Not Overexpressed in the Plant Microsymbiont *Mesorhizobium loti*

- MAFF303099 upon Acidic Shock. *Appl. Microbiol. Biotechnol.* **2014**, 98 (16), 7137–7147. <https://doi.org/10.1007/s00253-014-5875-4>.
12. Laranjo, M.; Alexandre, A.; Oliveira, S. Global Transcriptional Response to Salt Shock of the Plant Microsymbiont *Mesorhizobium loti* MAFF303099. *Res. Microbiol.* **2017**, 168 (1), 55–63. <https://doi.org/10.1016/j.resmic.2016.07.006>.
 13. Liu, X.; Luo, Y.; Mohamed, O. A.; Liu, D.; Wei, G. Global Transcriptome Analysis of *Mesorhizobium alhagi* CCNWXJ12-2 under Salt Stress. *BMC Microbiol.* **2014**, 14 (1), 319. <https://doi.org/10.1186/s12866-014-0319-y>.
 14. Maynaud, G.; Brunel, B.; Mornico, D.; Durot, M.; Severac, D.; Dubois, E.; Navarro, E.; Cleyet-Marel, J.-C.; Le Quéré, A. Genome-Wide Transcriptional Responses of Two Metal-Tolerant Symbiotic *Mesorhizobium* Isolates to Zinc and Cadmium Exposure. *BMC Genomics* **2013**, 14 (1), 292. <https://doi.org/10.1186/1471-2164-14-292>.
 15. Belfquih, M.; Sakrouhi, I.; Ait-Benhassou, H.; Dubois, E.; Severac, D.; Filali-Maltouf, A.; Le Quere, A. *Ensifer aridi* LMR001T Symbiosis and Tolerance to Stress Do Not Require the Alternative Sigma Factor RpoE2. *Agronomy* **2021**, 11 (9), 1787. <https://doi.org/10.3390/agronomy11091787>.
 16. de Lucena, D. K.; Pühler, A.; Weidner, S. The Role of Sigma Factor RpoH1 in the PH Stress Response of *Sinorhizobium meliloti*. *BMC Microbiol.* **2010**, 10 (1), 265. <https://doi.org/10.1186/1471-2180-10-265>.
 17. Domínguez-Ferreras, A.; Pérez-Arnedo, R.; Becker, A.; Olivares, J.; Soto, M. J.; Sanjuán, J. Transcriptome Profiling Reveals the Importance of Plasmid PSymbB for Osmoadaptation of *Sinorhizobium meliloti*. *J. Bacteriol.* **2006**, 188 (21), 7617–7625. <https://doi.org/10.1128/JB.00719-06>.
 18. Hellweg, C.; Pühler, A.; Weidner, S. The Time Course of the Transcriptomic Response of *Sinorhizobium meliloti* 1021 Following a Shift to Acidic PH. *BMC Microbiol.* **2009**, 9 (1), 37. <https://doi.org/10.1186/1471-2180-9-37>.
 19. Lu, M.; Jiao, S.; Gao, E.; Song, X.; Li, Z.; Hao, X.; Rensing, C.; Wei, G. Transcriptome Response to Heavy Metals in *Sinorhizobium meliloti* CCNWSX0020 Reveals New Metal Resistance Determinants That Also Promote Bioremediation by *Medicago*

- lupulina* in Metal-Contaminated Soil. *Appl. Environ. Microbiol.* **2017**, 83 (20).
<https://doi.org/10.1128/AEM.01244-17>.
20. Sauviac, L.; Philippe, H.; Phok, K.; Bruand, C. An Extracytoplasmic Function Sigma Factor Acts as a General Stress Response Regulator in *Sinorhizobium meliloti*. *J. Bacteriol.* **2007**, 189 (11), 4204–4216. <https://doi.org/10.1128/JB.00175-07>.
 21. Vriezen, J. A. C.; de Bruijn, F. J.; Nüsslein, K. Responses of Rhizobia to Desiccation in Relation to Osmotic Stress, Oxygen, and Temperature. *Appl. Environ. Microbiol.* **2007**, 73 (11), 3451–3459. <https://doi.org/10.1128/AEM.02991-06>.
 22. Beringer, J. E. R Factor Transfer in *Rhizobium leguminosarum*. *Microbiology* **1974**, 84 (1), 188–198. <https://doi.org/10.1099/00221287-84-1-188>.
 23. Sambrook, J.; Fritsch, E.; Maniatis, T. *Molecular Cloning: A Laboratory Manual: Vol. 2*, 2. ed.; Cold Spring Harbor: S.I., 1989.
 24. Hanahan, D. Studies on Transformation of *Escherichia coli* with Plasmids. *J. Mol. Biol.* **1983**, 166 (4), 557–580. [https://doi.org/10.1016/S0022-2836\(83\)80284-8](https://doi.org/10.1016/S0022-2836(83)80284-8).
 25. Simon, R.; Priefer, U.; Pühler, A. A Broad Host Range Mobilization System for In Vivo Genetic Engineering: Transposon Mutagenesis in Gram Negative Bacteria. *Bio/Technology* **1983**, 1 (9), 784–791. <https://doi.org/10.1038/nbt1183-784>.
 26. Quandt, J.; Hynes, M. F. Versatile Suicide Vectors Which Allow Direct Selection for Gene Replacement in Gram-Negative Bacteria. *Gene* **1993**, 127 (1), 15–21.
[https://doi.org/10.1016/0378-1119\(93\)90611-6](https://doi.org/10.1016/0378-1119(93)90611-6).
 27. Figurski, D. H.; Helinski, D. R. Replication of an Origin-Containing Derivative of Plasmid RK2 Dependent on a Plasmid Function Provided in Trans. *Proc. Natl. Acad. Sci. U. S. A.* **1979**, 76 (4), 1648–1652. <https://doi.org/10.1073/pnas.76.4.1648>.
 28. Miller, W. G.; Leveau, J. H. J.; Lindow, S. E. Improved *Gfp* and *InaZ* Broad-Host-Range Promoter-Probe Vectors. *Mol. Plant-Microbe Interact.* **2000**, 13 (11), 1243–1250.
<https://doi.org/10.1094/MPMI.2000.13.11.1243>.
 29. Chen, W.; Kuo, T. A Simple and Rapid Method for the Preparation of Gram-Negative Bacterial Genomic DNA. *Nucleic Acids Res.* **1993**, 21 (9), 2260–2260.
<https://doi.org/10.1093/nar/21.9.2260>.

30. Broughton, W. J.; Dilworth, M. J. Control of Leghaemoglobin Synthesis in Snake Beans. *Biochem. J.* **1971**, 125 (4), 1075–1080. <https://doi.org/10.1042/bj1251075>.
31. Patankar, A. V.; González, J. E. An Orphan LuxR Homolog of *Sinorhizobium meliloti* Affects Stress Adaptation and Competition for Nodulation. *Appl. Environ. Microbiol.* **2009**, 75 (4), 946–955. <https://doi.org/10.1128/AEM.01692-08>.
32. Novichkov, P. S.; Li, X.; Kuehl, J. V.; Deutschbauer, A. M.; Arkin, A. P.; Price, M. N.; Rodionov, D. A. Control of MéthionineMetabolism by the SahR Transcriptional Regulator in Proteobacteria: Control of MéthionineMetabolism in Proteobacteria. *Environ. Microbiol.* **2014**, 16 (1), 1–8. <https://doi.org/10.1111/1462-2920.12273>.
33. Robledo, M.; García-Tomsig, N. I.; Matia-González, A. M.; García-Rodríguez, F. M.; Jiménez-Zurdo, J. I. Synthetase of the Methyl Donor S-Adenosylméthioninefrom Nitrogen-Fixing α -Rhizobia Can Bind Functionally Diverse RNA Species. *RNA Biol.* **2021**, 18 (8), 1111–1123. <https://doi.org/10.1080/15476286.2020.1829365>.
34. Ravcheev, D. A.; Khoroshkin, M. S.; Laikova, O. N.; Tsoy, O. V.; Sernova, N. V.; Petrova, S. A.; Rakhmaninova, A. B.; Novichkov, P. S.; Gelfand, M. S.; Rodionov, D. A. Comparative Genomics and Evolution of Regulons of the LacI-Family Transcription Factors. *Front. Microbiol.* **2014**, 5. <https://doi.org/10.3389/fmicb.2014.00294>.
35. Suárez, R.; Wong, A.; Ramírez, M.; Barraza, A.; Orozco, M. del C.; Cevallos, M. A.; Lara, M.; Hernández, G.; Iturriaga, G. Improvement of Drought Tolerance and Grain Yield in Common Bean by Overexpressing Trehalose-6-Phosphate Synthase in Rhizobia. *Mol. Plant-Microbe Interact.* **2008**, 21 (7), 958–966. <https://doi.org/10.1094/MPMI-21-7-0958>.
36. Moussaid, S.; Domínguez-Ferreras, A.; Muñoz, S.; Aurag, J.; Berraho, E. B.; Sanjuán, J. Increased Trehalose Biosynthesis Improves *Mesorhizobium ciceri* Growth and Symbiosis Establishment in Saline Conditions. *Symbiosis* **2015**, 67 (1–3), 103–111. <https://doi.org/10.1007/s13199-015-0338-y>.

37. Paul, M. J.; Watson, A.; Griffiths, C. A. Trehalose 6-Phosphate Signalling and Impact on Crop Yield. *Biochem. Soc. Trans.* **2020**, *48* (5), 2127–2137.
<https://doi.org/10.1042/BST20200286>.
38. Baena-González, E.; Lunn, J. E. SnRK1 and Trehalose 6-Phosphate – Two Ancient Pathways Converge to Regulate Plant Metabolism and Growth. *Curr. Opin. Plant Biol.* **2020**, *55*, 52–59. <https://doi.org/10.1016/j.pbi.2020.01.010>.
39. da-Silva, J. R.; Alexandre, A.; Brígido, C.; Oliveira, S. Can stress response genes be used to improve the symbiotic performance of rhizobia? *AIMS Microbiol.* **2017**, *3* (3), 365–382. <https://doi.org/10.3934/microbiol.2017.3.365>.
40. Kohler, P. R. A.; Zheng, J. Y.; Schoffers, E.; Rossbach, S. Inositol Catabolism, a Key Pathway in *Sinorhizobium meliloti* for Competitive Host Nodulation. *Appl. Environ. Microbiol.* **2010**, *76* (24), 7972–7980. <https://doi.org/10.1128/AEM.01972-10>.
41. Kohler, P. R. A.; Choong, E.-L.; Rossbach, S. The RpiR-Like Repressor IolR Regulates Inositol Catabolism in *Sinorhizobium meliloti*. *J. Bacteriol.* **2011**, *193* (19), 5155–5163. <https://doi.org/10.1128/JB.05371-11>.
42. Dong, Y.; Li, S.; Zhao, D.; Liu, J.; Ma, S.; Geng, J.; Lu, C.; Liu, Y. IolR, a Negative Regulator of the Myo-Inositol Metabolic Pathway, Inhibits Cell Autoaggregation and Biofilm Formation by Downregulating RpmA in *Aeromonas hydrophila*. *Npj Biofilms Microbiomes* **2020**, *6* (1), 22. <https://doi.org/10.1038/s41522-020-0132-3>.
43. Lang, C.; Barnett, M. J.; Fisher, R. F.; Smith, L. S.; Diodati, M. E.; Long, S. R. Most *Sinorhizobium meliloti* Extracytoplasmic Function Sigma Factors Control Accessory Functions. *mSphere* **2018**, *3* (5). <https://doi.org/10.1128/mSphereDirect.00454-18>.
44. Jensen, J. B.; Peters, N. K.; Bhuvaneshwari, T. V. Redundancy in Periplasmic Binding Protein-Dependent Transport Systems for Trehalose, Sucrose, and Maltose in *Sinorhizobium meliloti*. *J. Bacteriol.* **2002**, *184* (11), 2978–2986.
<https://doi.org/10.1128/JB.184.11.2978-2986.2002>.
45. diCenzo, G. C.; Muhammed, Z.; Østerås, M.; O'Brien, S. A. P.; Finan, T. M. A Key Regulator of the Glycolytic and Gluconeogenic Central Metabolic Pathways in

Sinorhizobium meliloti. *Genetics* **2017**, 207 (3), 961–974.

<https://doi.org/10.1534/genetics.117.300212>.

46. Onwe, R. O.; Onwosi, C. O.; Ezugworie, F. N.; Ekwealor, C. C.; Okonkwo, C. C. Microbial Trehalose Boosts the Ecological Fitness of Biocontrol Agents, the Viability of Probiotics during Long-Term Storage and Plants Tolerance to Environmental-Driven Abiotic Stress. *Sci. Total Environ.* **2022**, 806, 150432.
<https://doi.org/10.1016/j.scitotenv.2021.150432>.
47. Sengupta, S.; Banerjee, S.; Lahiri, S.; Dutta, T.; Dhar, T. K.; Ghosh, A. K. Purification, Characterization, Sequencing and Molecular Cloning of a Novel Cysteine Methyltransferase That Regulates Trehalose-6-Phosphate Synthase from *Saccharomyces cerevisiae*. *Biochim. Biophys. Acta* **2014**, 1840 (6), 1861–1871.
<https://doi.org/10.1016/j.bbagen.2014.01.005>.

1. Introduction

La bactérie *Ensifer aridi* LMR001 isolée du désert marocain est dotée de capacités multiples intéressantes pour la survie dans des conditions environnementales difficiles. L'étude effectuée sur cette espèce a révélé des éléments importants sur les mécanismes moléculaires utilisés dans l'adaptation au stress générée par ces milieux de vies. Une étude préliminaire réalisée sur les données des séquençages des ARNseq a révélé que la Réponse Générale au Stress (par le RpoE2 sigma facteur) est induite par le stress salin. Cette première étude a également montré que l'activité du gène *rpoE2* est non essentielle pour la bactérie pour persister en réponse au stress et pour développer une symbiose. D'autres fonctions aussi importantes et induites par le stress salin, comme le métabolisme et transport du tréhalose (gène *ostB* et *thuR*), l'activité du cycle de la méthionine par l'étude de ses deux principaux régulateurs (gène *nesR* et *sahR*), le catabolisme et la régulation du catabolisme d'inositol (gène *iolR* et *iolC*) et d'autres fonctions apparaissent réprimées comme la mobilité. Nous avons muté plusieurs gènes révélés induits par le stress puis évalué leur implication dans l'adaptation au stress et le développement de la symbiose.

2. Résultats et discussion

Les voies métaboliques des acides aminés contenant du soufre sont souvent actives en l'absence de dioxygène comme la voie de la biosynthèse de la méthionine et de l'Adomet qui sont impliquées dans la machinerie de la biosynthèse des macromolécules et dans le métabolisme intermédiaire (Marsh et al., 2010; Sekowska et al., 2019). En effet, la méthionine est un précurseur métabolique de la S-adénosylméthionine (AdoMet ou SAM), donneur universel de groupements méthyles et qui sert de cofacteur essentiel à diverses méthyltransférases. L'activité des méthyltransférases via la SAM agit sur le contrôle du métabolisme de l'ADN et de l'ARN ainsi que dans l'expression des gènes, sur la formation de nombreux métabolites de petite taille et sur des modifications post-traductionnelles des protéines (Novichkov et al., 2014). La S-adénosyl-L-méthionine est un substrat produit par une série d'enzymes impliquées dans le cycle de la biosynthèse de la méthionine. Cette molécule transfère le méthyle et a de nombreuses fonctions biologiques. Aussi, elle joue le rôle de précurseur des composés chélateurs des ions métalliques nicotianamine et phytosidérophores, des hormones végétales gazeuse (éthylène) et la biosynthèse des polyamines spermidine et spermine. AdoMet est aussi la source des radicaux catalytiques, 5'-désoxyadénosyle, qui sont des produits intermédiaires de la réaction de la superfamille des enzymes radicalaires (Roje 2006; Tehlivets et al., 2013). L'activité de la molécule S-adénosyl-L-méthionine (AdoMet) produit la méthylation qui régule de multiples processus biologiques qui

peuvent être cruciaux pour la survie de la cellule. Plus de 50 méthyltransférases dépendantes de l'AdoMet ont été découvertes et méthyle un très grand spectre de composés cellulaires y compris : des acides nucléiques, des protéines et des lipides. La réaction de méthylation par le transfert du groupe méthyle aboutit à la production d'un sous-produit, par la conversion de l'Adomet en AdoHcy (la S-adenosyl-L-homocystéine). Ce composé intermédiaire est défini comme un inhibiteur qui doit être recyclée en Adomet et est soumis à une régulation fine. La S-adenosyl-L-homocystéine hydrolase est la seule enzyme eucaryote et procaryote avec un potentiel d'hydrolyse AdoHcy en adénosine et homocystéine. Elle produit une réaction réversible et donne par conséquence une réduction d'inhibition de l'activité de l'AdoHcy (Tehlivets et al., 2013 ; Danchin 2017; Sekowska et al., 2019). Le processus de recyclage de l'homocystéine en méthionine, par les processus de piégeage du soufre et la trans-sulfuration inverse, est associé à divers procédés de récupération du soufre réduit (Cho et al., 2015 ; Sun et al., 2016). Il nécessite des donneurs de méthyle et la présence de divers composés, qui varient selon les organismes. Une méthyl-tétrahydrofolate, méthyl-cobalamine et une variété de petites molécules méthylées [p. Ex. choline, glycine bêtaïne (Zou et al., 2016), S-méthylméthionine (Li et al., 2016)] ce qui montre de multiples associations fonctionnelles liées étroitement à l'activité de la cellule (Sekowska et al., 2019).

Deux gènes régulateurs de réponse *nesR (luxR)* et *sahR* du cycle de la méthionine ont été identifiés chez la souche *Ensifer aridi LMR001* associé respectivement dans l'induction et la répression des gènes impliqués dans le cycle méthionine. Chez cette nouvelle espèce, comme d'autres bactéries types, le cycle de la méthionine est régulé par plusieurs systèmes. Les régulateurs sont dépendants des régulateurs de l'ADN (par exemple, des facteurs de transcription) et des régulateurs de l'ARN (Novichkov et al., 2014). La régulation de la transcription des gènes impliqués dans la voie métabolique de la biosynthèse de la méthionine/SAM est complexe et reste mal comprise. Afin d'accroître nos connaissances sur la régulation de ce métabolisme clé, la mutation des deux principaux régulateurs *nesR (luxR)* et *sahR* est essentiel pour décrire et comprendre la réponse transcriptionnelle liée à l'activité du cycle et son effet sur la physiologie et la réponse phénotypique au stress de la souche *Ensifer aridi LMR001*.

Nous avons pu montrer un effet négatif de la mutation des deux gènes régulateurs du cycle de la méthionine, *nesR* et *sahR*, sur le développement symbiotique et la mobilité de la bactérie ainsi que sur la croissance de la souche en réponse au stress salin. Par contre la réponse des simples mutants au stress oxydatif, au stress de détergent et à l'acide n'a pas été affectée en comparaison à la réponse de la souche sauvage aux mêmes stress. En revanche, la double mutation des deux gènes résulte en une tolérance accrue aux stress induits par des détergents, des oxydants et l'acide, et a généré en même temps une diminution drastique de la tolérance au stress hyperosmotique salin.

Cet impact de la double mutation des deux régulateurs du cycle de la méthionine sur l'adaptation des mutants est peut-être dû à des changements sur les composants de la paroi de la cellule bactérienne. La double mutation a généré aussi une diminution sévère de la capacité des doubles mutants à développer une symbiose et à contrôler sa mobilité en fonction des changements des concentrations de la salinité dans le milieu de culture. La délétion de gènes impliqués dans la régulation du cycle de la méthionine (*nesR* et *sahR*) a considérablement réduit le poids des pousses des plantes pour *Vachellia* par rapport à la souche sauvage. De plus, la nodulation était également significativement inférieure à celle de la souche *LMR001* sauvage chez *V. tortilis* ou chez *V. gummifera*, lorsque les deux gènes régulateurs ont été mutés. Ces données suggèrent que la régulation du cycle de la méthionine générant le donneur universel de méthyle SAM est importante pour le bon développement de la symbiose et du fonctionnement des nodules. Nous avons également pu montrer que le gène *metK* codant pour la SAM synthétase, MetK, impliquée dans le cycle méthionine était régulé par les deux gènes régulateurs, *nesR* et *sahR*. Il est intéressant de noter que d'après les données transcriptomiques obtenues (Belfquih et al., 2021), le gène *metK* a été montré comme le gène le plus induit par le stress hyperosmotique. Il a été montré que cette enzyme pouvait interagir avec l'ARN chaperon Hfq et ainsi que d'autres ARNm. Ceci suggère que la régulation post-transcriptionnelle épigénétique implique MetK et démontre la complexité de la régulation du fonctionnement cellulaire notamment en réponse au stress.

Pour étudier le mécanisme de l'accumulation du tréhalose par la souche *E.aridi LMR001* via son transport ou sa biosynthèse. Le régulateur du transport et de l'utilisation du tréhalose, appelé *thuR*, et Le gène *ostB* qui code pour l'enzyme Tréhalose-6-phosphatase ont été mutés. Le tréhalose est un disaccharide constitué de deux unités de glucose liées par une liaison α, α -1, 1 (ou) '1, 1- α -glycosurique. À ce jour, cinq voies de synthèse biosynthétique du tréhalose sont connues (Avonce et al., 2006). La voie la plus étudiée de la synthèse du tréhalose comprend une enzyme tréhalose-6-phosphate-synthase (TPS), responsable de la production de la molécule intermédiaire tréhalose-6-phosphate (Tre6P), codée par l'expression du gène *ostA* et une enzyme Tre6p-phosphatase (TPP) codée par l'expression du gène *ostB* qui permet la production de tréhalose. Cette voie est particulièrement répandue chez les champignons et les micro-organismes vivants dans les milieux arides. L'enzyme tréhalose-6-phosphate synthase est une glycosyltransférase capable d'utiliser un donneur UDP-glucose pour catalyser la synthèse de l'alpha, alpha-1, 1-tréhalose-6-phosphate avec du glucose-6-phosphate. Cette enzyme est considérée comme l'enzyme clé dans la voie de synthèse du tréhalose. L'activité de la tréhalose-6-phosphatase est fortement liée à plusieurs conditions de stress comme la dessiccation, la déshydratation, la chaleur, le froid et l'oxydation. Toutes les études

d'exploration évolutive basées sur les gènes de biosynthèse du tréhalose mettent en évidence la formation de complexes protéiques ainsi que la bi-fonctionnalité de ces protéines dans certains groupes de bactéries. Les conséquences physiologiques de la fonction et de la fusion des protéines sont encore mal connues (Avonce et al., 2010). À côté de cette voie de synthèse de tréhalose, chez *Ensifer aridi* LMR001, une deuxième voie de synthèse TreY/TreZ a été identifiée. L'expression des deux gènes *treY* et *treZ* qui convertissent des maltodextrines en tréhalose est induite sous stress chez *E. aridi* LMR001. Les analyses de transcription suggèrent que le système était induit par la présence du maltose/maltodextrine et réprimé par la disponibilité du glucose dans le milieu de croissance de la bactérie. L'utilisation du maltose-maltodextrine est impliquée dans la formation du biofilm d'entérocoque et la colonisation du tractus gastro-intestinal lors d'une infection naturelle par *L. monocytogene* (Gopal et al., 2010). Les enzymes de transport et/ou d'utilisation du tréhalose sont codées par l'opéron (*thuEFGKAB*). L'opéron est divisé en gènes *thuEFGK* pour le transport du tréhalose et les deux gènes *thuAB* d'utilisation. Il est situé sur le chromide *pSymB* chez les bactéries *E. meliloti*, *M. loti*, et *Agrobacterium tumefaciens*. Des analyses génétiques et biochimiques ont montré que ces bactéries assimilent les disaccharides et produisent des dérivés 3-cétodisaccharides. Elles ont aussi montré que les gènes *thuAB* codent pour les enzymes de conversion des disaccharides en cétodisaccharides (Ampomah et al., 2013). Les gènes *thuA* et *thuB* sont induits in vitro par le tréhalose mais pas par le saccharose et l'étendue de son induction dépend de la concentration en tréhalose disponible dans le milieu. Généralement, l'accumulation des glucides augmente pendant la phase G1 du cycle de la croissance, lorsque les taux de croissance des bactéries sont faibles et la disponibilité en sources de carbone ou en azote est minime. L'accumulation est aussi corrélée avec l'activation transcriptionnelle des gènes liés au glycogène et au tréhalose (Hazelwood et al., 2009). Chez la souche *Ensifer aridi*, la biosynthèse et l'accumulation du tréhalose, comme pour d'autres espèces bactériennes, occupent des rôles importants (Makihara et al., 2005). Le tréhalose est un élément présent dans la rhizosphère. Il aide les bactéries (comme l'*E. meliloti*) à améliorer leur niveau d'adaptation aux stress et à produire une meilleure interaction avec la plante (Murphy et al., 2005). Les gènes *thuA* et *thuB* codent respectivement pour la glutamine amido transférase et la NAD (H) déshydrogénase dépendante. L'induction de la synthèse et de l'accumulation de tréhalose dépend principalement de la concentration de tréhalose disponible dans le milieu de vie de la bactérie. Ainsi, l'induction in planta du gène *thuB* est visible lors de l'infection et ne se produit pas dans les bactéroïdes (Ampomah et al., 2013). Une étude comparative entre les bactéries mutées au niveau du gène *thuB* et les bactéries sauvages de la même espèce a révélé une différence de réponse lors de l'infection. Les bactéries *thuB* mutées ont montré une absence de compétitivité avec les bactéries sauvages,

avec une répression des gènes cataboliques et de synthèse du tréhalose chez les mutés et une induction de ces gènes chez les bactéries sauvages.

Les deux mutants *E.aridi* Δ *thuR* et *E.aridi* Δ *otsB* ont généré une augmentation de la capacité de la souche *Ensifer aridi* LMR001 à développer la symbiose comparée à la souche sauvage, avec une augmentation très importante de la croissance des plantes et de la nodulation. Aussi, les deux mutants ont révélé une forte capacité à tolérer les fortes concentrations de sel avec une croissance bactérienne rapide dans les concentrations qui dépassent les 0,4M de NaCl. Ces résultats relèvent l'importance du contrôle de l'accumulation du tréhalose via la régulation de son transport et la régulation de ces voies métaboliques dans l'amélioration des capacités des bactéries en réponse à des environnements difficiles et fluctuants. Le phénotype du mutant *ostB* révélé par les résultats d'inoculation et les résultats de la tolérance au stress salin relève un intérêt biotechnologique de la molécule phospho-sucre, Tréhalose-6-phosphate, dans l'amélioration des capacités adaptatives dans les environnements arides. Récemment, il a été démontré que Tréhalose-6-phosphate agissait non seulement comme soluté compatible mais aussi comme molécule de signalisation dans la plante. En modulant l'état de phosphorylation de la SnRK1, une protéine kinase centrale impliquée dans la coordination du développement de la plante. Le Tréhalose-6-phosphate conduit à une amélioration de la croissance de la plante en situation de stress. D'où la nécessité de faire des études plus approfondies sur le mutant *ostB* pour voir si les effets sont liés à l'amélioration de la condition physique et/ou de la signalisation.

Les gènes impliqués dans le catabolisme de l'inositol (*iolA* et *iolRCDEB*) ont été identifiés comme fortement exprimés sous stress par l'étude transcriptomique. Ce métabolisme est identifié aussi dans les génomes de *Ensifer meliloti* et de *Bacillus subtilis*. Les bactéries *Ensifer meliloti* et *Ensifer aridi* ont le potentiel d'utiliser la molécule myo inositol comme la seule source de carbone. Une étude fonctionnelle a été réalisée chez *E. meliloti* et a montré que les gènes *iolA*, *smc01163* et *iolRCDE* nécessaire pour la croissance de la bactérie avec le myo-isomère, et aussi pour une croissance avec les deux molécules scyllo- et D-chiro-inositol comme source de carbone. La myo - et surtout la scyllo-inositol sont décrites comme des sources de carbone excellentes pour la croissance d'*E.meliloti* comparablement à d'autres sources de C. Tels que le glucose, le glycérol ou le succinate. A l'inverse, la croissance de la bactérie *B. subtilis* ne s'avère pas efficace avec le myo/ scyllo-inositol comparé avec le glucose. Les études ont montré aussi que le gène *iolR* régulateur est situé en amont de la séquence des gènes (*iolCDEB*). Il joue un rôle de régulation négative de l'expression de ces gènes. L'expression des deux enzymes myo - et les

déshydrogénases scyllo-inositol était 4 à 7 fois plus élevée chez le mutant *iolR* que chez le sauvage (Kohler et al., 2010, Kohler et al., 2013).

L'étude transcriptionnelle réalisée par l'utilisation de fusions des régions promotrices de gènes d'intérêt au gène *-gfp* dans les mutants décrits dans cet article a montré qu'il existait des interactions entre les voies métaboliques de l'inositol, la méthionine et du tréhalose. Nous avons constaté que la mutation du gène *iolR*, répresseur des gènes du catabolisme d'inositol entraîne une induction de sa propre transcription, et une répression de la transcription du gène *iolC*. Ceci suggère que la régulation du catabolisme de l'inositol implique d'autres facteurs. Une séquence proche du site de fixation d'IolR a été trouvée dans la région intergénique des gènes *rsiA1-rsiB1*, avec un chevauchement avec le site de fixation de la protéine régulatrice RpoE2 prédit située en amont du *rsiB1*. Ceci suggère également que le répresseur IolR est potentiellement impliqué dans la modulation de la réponse générale au stress GSR. De plus, nous avons également trouvé deux motifs proches de la séquence consensus du site de liaison du facteur *ecfG* sigma RpoE2 situés en amont du gène *iolR*, avec un chevauchement avec le site de liaison du répresseur IolR. Ces données suggèrent la présence d'une régulation croisée entre les deux régulateurs IolR et RpoE2. Nous avons également pu identifier un site de fixation de IolR en amont d'un système de transport d'inositol de type ABC prédit (*PEG5307-5306-5305*) qui est également induit par le stress salin. Nous avons aussi trouvé un site de liaison de IolR en amont du gène *nesR* suggérant la présence d'une régulation potentielle du cycle de la méthionine dépendante de l'activité de la protéine IolR. Également, la transcription du gène *iolC* s'est révélée plus élevée dans les souches *ΔnesR*, *ΔsahR*, *ΔthuR* et *ΔotsB*, tandis que sa transcription s'est révélée très basse chez le double mutant *ΔnesR ΔsahR*. Ces données indiquent la présence d'une régulation complexe et interdépendante de ces métabolismes, mais des travaux de recherche supplémentaires sont encore nécessaires.

Dans la caractérisation de l'expression des gènes du transport et du métabolisme du tréhalose. Nous avons constaté que l'expression relative du gène *thuR* est induite dans le mutant *thuR*, avec une forte induction du gène *thuE*. En revanche, l'induction du gène *thuR* dans le mutant *ΔiolR* a généré une répression du gène *thuE* par rapport à *LMR001*. Ces résultats suggèrent que le gène *thuR* est un répresseur de sa propre transcription et de la transcription des gènes *thuEFGKAB*. Nous avons aussi constaté une forte répression de *thuR* dans le mutant *ΔotsB* et dans le double mutant *ΔnesR ΔsahR*. La transcription du gène *thuR* est dépendante de RpoE2 chez *E.aridi* *LMR001*. La séquence du motif GGAAA-N17-AGAT de fixation de la RpoE2 trouvée a été identifiée entre les deux motifs de reconnaissance de ThuR en amont du gène *thuR*. Des sites de liaison de la ThuR ont été aussi trouvés en amont des deux gènes *fold* et *metC* impliqué dans le

métabolisme de la méthionine. L'expression du gène *sahR* s'est montrée induite dans le mutant *thuR*, soutenant l'implication de la ThuR dans la régulation du cycle de la méthionine. De plus, nous avons également trouvé le motif de liaison de la ThuR putatif une fois en amont de *iolR* et deux fois avant *iolY*. Ces études de l'activité transcriptionnelle par fusion de promoteurs à la *gfp* suggèrent une implication de ThuR dans la régulation du catabolisme de l'inositol. Car la transcription du gène *iolR* est induite dans le mutant $\Delta thuR$. L'expression relative de *nesR* est faible pour les trois mutants des régulateurs du cycle de la méthionine ce qui rend difficile l'interprétation. En revanche, le régulateur *sahR* a montré des niveaux d'expressions significativement différentes dans les différents mutants testés. Avec une expression très élevée dans le $\Delta sahR$, une induction relativement élevée chez le $\Delta nesR$ et une induction très basse pour le double mutant. L'induction du gène *sahR* chez le mutant $\Delta sahR$ une autorégulation de l'activité transcriptionnelle de ce gène. Nous avons relevé la présence du site de fixation de la protéine SahR en amont du gène *sahR*, et en amont des gènes cibles du cycle de la méthionine *ahcY*, *metH*, *metK* et *bhmT*. L'activité transcriptionnelle du gène *sahR* s'est montrée aussi élevée dans les mutants $\Delta iolR$ et $\Delta thuR$, et dans le mutant étudié dans l'article 1 $\Delta rpoE2$ ce qui montre la complexité des réseaux de régulation impliqués chez la bactérie dans le contrôle de ces métabolismes.

La régulation transcriptionnelle du gène *rsiA1* dans la souche LMR001 par rapport aux trois mutants $\Delta nesR$ $\Delta sahR$, $\Delta thuR$ et $\Delta otsB$ présente une expression plus faible. Au contraire, pour la souche $\Delta iolR$, elle s'est révélée induite. Le gène *rsiB1* apparaît moins exprimé dans la souche sauvage par rapport aux trois mutants ($\Delta nesR$ $\Delta sahR$, $\Delta iolR$ et $\Delta otsB$). Les deux gènes *rsiA1* et *rsiB1* étant plus exprimés dans la souche $\Delta iolR$ que dans la souche sauvage et ayant identifié un motif de liaison de IolR situé au niveau de leur région promotrice, il est donc tentant de spéculer sur une possible implication de IolR dans la modulation de la GSR. D'autre part, chez les deux mutants $\Delta nesR$ $\Delta sahR$ et $\Delta otsB$, l'induction de l'expression du facteur anti-anti-sigma RsiB1 suggère une stimulation de l'activité régulatrice de RpoE2 par une liaison compétitive au facteur anti-sigma RsiA1. Cependant, la cascade de phosphorylation n'est pas encore entièrement caractérisée et de nombreuses études sont nécessaires afin de mieux cerner les bases de la régulation de la GSR, et identifier les déterminants impliqués.

3. Conclusion

L'étude transcriptomique basée sur le séquençage des ARNs a été réalisée pour explorer les mécanismes/fonctions potentiellement impliqués dans la réponse adaptative de la nouvelle espèce *E. aridi* au stress osmotique. Elle a permis d'identifier et de sélectionner des métabolismes affectés par l'ajout de sel dans le milieu de culture. La mutagenèse dirigée par délétion des gènes ciblés

différentiellement exprimés en réponse au stress salin a été réalisée. Ainsi, des gènes impliqués dans le cycle de la méthionine générateur de la SAM (gènes régulateurs *nesR* et *sahR*), la régulation du transport et de l'utilisation du tréhalose (gène régulateur *thuR*) ou impliqué dans sa biosynthèse, (*otsB*) et le catabolisme de l'inositol (gène régulateur *iolR*) ont été mutés et étudiés. Des tests phénotypiques ont été réalisés sur les mutants des gènes cibles et comparés à la souche sauvage. Les mutations des gènes régulateurs du cycle de la méthionine ont modifié la capacité de la souche à développer des symbioses fonctionnelles avec leurs plantes hôtes compatibles. Ils ont également montré une réduction de la tolérance et la motilité en réponse au sel et une induction de la tolérance aux stress détergents, oxydatifs ou acides. Ceci suggère que la régulation de la méthionine, et le contrôle de l'activité et de la synthèse de la SAM, sont essentiels pour la colonisation et l'infection des plantes. Par contre, la mutation des deux gènes impliqués dans la synthèse et la régulation du transport et de l'utilisation du tréhalose a généré une amélioration de la symbiose et de la tolérance aux stress hyperosmotique suggérant une relation étroite entre l'accumulation du tréhalose et l'amélioration des potentiels adaptatifs de la bactérie en réponse aux conditions de stress extrêmes, à l'état libre et en symbiose. L'étude de la transcription des gènes régulateurs *iolR*, *thuR* et *sahR* a révélé une induction de ces gènes après leur mutation suggérant des mécanismes d'autorégulation. Nous avons également constaté qu'il existait probablement une régulation croisée entre les voies métaboliques du tréhalose, méthionine et l'inositol. En effet, la mutation des gènes *thuR* et *iolR* résulte dans l'induction du gène répresseur du cycle méthionine *sahR*, et la mutation du gène *iolR* a généré une modification de l'expression des gènes *thuR*, *thuE*. La mutation des gènes *thuR*, *otsB* et la double mutation des gènes *nesR* et *sahR* a réduit l'expression du gène *rsiA1* et induit l'expression du gène *rsiB1*. Bien qu'il reste de nombreuses études à réaliser afin de déterminer les mécanismes impliqués dans la régulation de ces métabolismes, nos données suggèrent une interdépendance entre la régulation du métabolisme du tréhalose et du cycle de la méthionine avec la régulation de la réponse générale au stress GSR. Plus globalement, la présente étude montre clairement qu'il existe une interaction dans les schémas de régulation des gènes impliqués dans les métabolismes de la méthionine, de l'inositol et du tréhalose et il reste à déterminer si cette régulation est directe ou indirecte et quels sont les facteurs additionnels impliqués.

CHAPITRE V

Discussion, conclusion générale et perspectives

1. Discussion

L'étude transcriptomique effectuée sur *Ensifer aridi* LMR001, a montré que sous l'effet d'un stress modéré salin (milieu TY ajusté avec 250 mM de NaCl), ou hydrique (milieu TY ajusté avec du PEG8000 à 12%), de nombreux gènes étaient régulés. On a pu noter, en particulier une répression sévère des gènes impliqués dans la motilité et la chimiotaxie. De plus, ces stress ont affecté l'activité des métabolismes de la méthionine, de l'inositol, le transport et l'accumulation du tréhalose, l'activité du système de régulation lié à la membrane cellulaire de la bactérie *E. aridi* (Belfquih et al., 2021).

Le stress salin, en particulier active des gènes impliqués dans la voie du cycle de la méthionine, dont deux gènes de régulation nommés *nesR* et *sahR*. Un troisième régulateur potentiel du cycle de la méthionine, appelé *metR* a été identifié mais s'est montré légèrement réprimé chez *E. aridi* LMR001. Le gène *bhmt* qui code pour la bétaine-homocystéine-S méthyltransférase est également surexprimé en condition de stress salin. Le cycle de la méthionine est un processus responsable de la méthylation (le transfert du groupe méthyle) à plusieurs molécules clés, comme les acides nucléiques, les dérivés de la molécule d'inositol et peut être impliqué dans la biosynthèse des polyamines ou la production de la glycine bétaine (Sánchez-Aguayo et al., 2004). Ces réactions métaboliques peuvent être importantes pour l'adaptation au stress, par exemple, l'utilisation de la glycine bétaine comme osmoprotecteur ou des dérivés d'inositol comme piègeurs des radicaux libres (Bohnert et al., 1995; Bohnert et al., 1996). Les changements de la transcription des gènes liés au cycle de la méthionine sont corrélés avec des réponses adaptatives de la cellule au stress. Par exemple les plantes de tomate sous stress ont révélé que la production des molécules SAM (S- adénosyl-méthionine) augmentait parallèlement avec l'augmentation de la production des tissus lignifiant de la plante pour protéger les tissus végétatifs (Sánchez-Aguayo et al., 2004). Une étude a montré que les feuilles d'orge soumises à des stress salins, secs et froids montraient une induction du gène *hvM*, homologue au gène *metH*. Le gène *metH* code pour la S-méthionine synthase (Narita et al., 2004). Chez les rhizobiums également, le gène *metH* est lié à la tolérance au stress. En effet, chez *E. fredii*, la mutation du gène *metH* a généré une sensibilité du mutant en réponse au stress salin. De même, chez *E. meliloti*, le mutant 102F34 $\Delta metH$ s'est avéré très sensible au stress salin. Le potentiel de croissance des mutants 102F34 $\Delta metH$ en présence du sel a été restauré avec l'ajout de la glycine bétaine (Barra et al., 2006 ; Jiang et al., 2004). Au niveau de la relation symbiotique plante-microbe, il a été montré que pour *E. meliloti*, l'altération de la synthèse de la méthionine génère des défauts de compétitivité pour la nodulation. Le mutant *E. meliloti* 102F34 $\Delta metH$ a

montré une compétitivité déficiente pour la nodulation en co-inoculation avec la souche sauvage (Barra et al., 2006). De plus, la mutation du gène *metA*, impliqué dans la biosynthèse de la méthionine chez *E. meliloti* Rm2021 a montré une nodulation retardée par rapport à la souche sauvage (Barra et al., 2006, Pobigaylo et al., 2008). Ces travaux montrent que l'activité du cycle de la méthionine a un effet sur la capacité d'infecter les racines et à former des nodules fonctionnels (Barra et al., 2006 ; Pobigaylo et al., 2008). La plante a également la capacité de produire de la méthionine qui peuvent être utilisées par les bactéries pour augmenter les chances de succès de l'interaction symbiotique (Pobigaylo et al., 2008 ; Barra et al., 2006 ; Abbas et al., 2002 ; Kerppola et al., 1988 ; Kim et al., 1988 ; Taté et al., 1999). Les trois mutants *E. aridi* *LMR001 ΔnesR*, *ΔsahR* et *ΔsahRΔnesR* ont donné des réponses physiologiques et phénotypiques. Une faible croissance cellulaire sous effet du stress salin et une modification de la sensibilité en réponse aux stress induits par un détergent, de l'acide et au stress oxydatif ont pu être observées. De plus, nous avons constaté une formation abondante des nodules de taille très petite et blanche, après un temps de croissance de 6 mois des plantes inoculées confirmant que la dérégulation du cycle de la méthionine impliqué dans la production du SAM et donc de la capacité de méthylation de rhizobium impacte leur capacité à fixer l'azote. La mobilité des mutants s'est montrée aussi très affectée sous l'effet du stress salin. La sensibilité de la souche double mutée *ΔsahRΔnesR* en réponse aux stress acide et oxydatif s'est révélée différente des souches simples mutants. Cette variabilité a été aussi visible au niveau de sa mobilité en réponse au stress salin. Lorsque le double mutant montre une faible sensibilité au stress d'acide et au stress oxydatif, la mobilité a augmenté dans le milieu sans sel alors qu'elle est réduite sous stress salin. Ces résultats montrent également que le dérèglement de la régulation du cycle de la méthionine via la mutation des deux régulateurs a peut-être provoqué une altération au niveau de la membrane cellulaire. Cette altération a agi sur la régulation de la mobilité, le contrôle des systèmes d'adaptation au stress et l'efficacité de la bactérie à développer une symbiose fonctionnelle avec les plantes.

La régulation du cycle de la méthionine est très complexe. Plusieurs facteurs environnementaux et régulateurs agissent sur son fonctionnement et sur le contrôle de ces activités. Dans le cas de la bactérie *Ensifer aridi*, on a pu identifier deux principaux régulateurs *nesR* et *sahR* ainsi qu'un homologue de *metR*. Chez *E. meliloti*, le gène *nesR* est un régulateur de type « orphelins ». Il code pour la protéine NesR, paralogue de ExpR, qui agit sans l'activité du gène *sinR*. Les gènes *sinR/I* et *expR* sont directement impliqués dans le contrôle de la production d'exopolysaccharides EPS II, un élément essentiel pour le processus symbiotique (Marketon et al., 2003 ; Pellock et al., 2002). Ils sont également impliqués dans la répression de la motilité, la chimiotaxie et la production de

succinoglycans de faible poids moléculaire (Hoang, Gurich, et González 2008; Hoang, Becker, et González 2004; Glenn et al., 2007)). La souche *E. aridi* LMR001 a montré une induction partielle des gènes impliqués dans les exopolysaccharides en réponse au stress salin et hydrique. Le locus *sinR/I* est impliqué dans la production et la régulation d'une gamme d'AHL (C12-HSL, C14-HSL, 3O-C14-HSL, C16:1-HSL, 3O-C16:1-HSL et C18-HSL) (Marketon et al., 2002 ; González et al., 2003 ; Teplitski et al., 2003). L'expression d'*expR* est impliqué dans la régulation de la motilité via deux protéines de type LuxR appelées VisN et VisR. L'activité des gènes *visN/visR* induit l'activité des gènes impliqués dans la formation des flagelles, la régulation de la chimiotaxie (Bahlawane et al., 2008 ; Hoang, Gurich, et González 2008 ; Sourjik et al., 2000). Ces deux gènes se sont révélés chez *Ensifer aridi* LMR001 réprimés par l'effet du stress. Le processus d'inhibition de la mobilité chez *Ensifer aridi* peut être expliqué par la répression des deux gènes *visN/visR*. L'activité de la protéine ExpR en cas de la forte densité cellulaire peut aussi agir sur la régulation de la mobilité. VisN va agir ensuite par inhibition de l'expression des gènes responsable de l'assemblage et de l'activité des flagelles via la protéine Rem qui génère par conséquence une diminution de la motilité et de la chimiotaxie (Bahlawane et al., 2008). Chez *Ensifer aridi* LMR001, deux systèmes de régulation activés ont été identifiés. Avec une activation du système LuxR/LuxI en situation de stress hydrique ; une activation du gène *sinR* qui est un régulateur orphelin de type LuxR en situation de stress salin. Les protéines de la famille LuxR sont composées de deux domaines fonctionnels : le domaine C-terminal avec un motif de liaison à l'ADN hélice-tour-hélice (HTH) et le domaine N-terminal de liaison AHL. Fait intéressant, il existe également un autre sous-groupe de protéines de la famille des *luxRs* orphelines qui ont la même structure de domaine de liaison avec l'AHL classique et un C-terminal et le domaine de liaison à l'ADN HTH (González et al., 2013; Patankar et González 2009; Subramoni et Venturi 2009; Fuqua 2006). Les protéines LuxR solo permettent aux bactéries de détecter des signaux endogènes et exogènes et contrôlent la régulation de la transcription de plusieurs gènes cibles (Sabag-Daigle et al., 2015). Chez plusieurs bactéries, la régulation du système QS est importante pour le développement de la symbiose, la virulence, la mobilité, la biosynthèse des métabolites secondaires et la formation de biofilm (Miller et al., 2001; Whitehead et al., 2001). Chez *E. meliloti*, en absence de l'expression du gène de type solo LuxR *nesR*, l'expression de nombreux gènes tels que *metH*, *bmt*, *metK*, *metF* et *ahcY* se révèle non induite même sous stress (Qian et al., 2014). Chez *Ensifer aridi*, ces gènes ont été induits par le sel et l'activité du gène *nesR* s'est également révélée aussi surexprimée. Ces résultats suggèrent que le gène *nesR* est un gène régulateur activateur des gènes impliqués dans le cycle de la méthionine. L'expression des deux gènes *metK* et *metF* est réduite de six à cinq fois pour le mutant *nesR*. Également l'expression du gène *bhmt* apparaît réduite de quatre fois pour le

mutant *nesR* par rapport à la souche sauvage chez *E. meliloti* (Qian et al., 2014). De plus, l'expression du gène *metH* a également été réduite par la mutation du gène *nesR*. La mutation de *nesR* n'a pas affecté le potentiel de la croissance des bactéries sans stress. Cependant, en présence d'une forte concentration de sel, la croissance du mutant *nesR* était très faible par rapport à la souche sauvage (Qian et al., 2014). Des réponses similaires ont été montrées dans des expériences appliquées sur la croissance du mutant *Ensifer aridi* $\Delta nesR$ dans les conditions de croissance avec stress salin. En plus de *nesR*, nous avons identifié un autre gène régulateur de transcription des gènes impliqués dans le cycle de la méthionine nommé *sahR*. Ce gène appartient à la famille des régulateurs de type ArsR. L'étude de Novichkov a permis de montrer que la liaison de la protéine SahR aux promoteurs des gènes cibles produit une inhibition de leurs expressions. Une étude de génomique comparative de 62 protéobactéries suggère que le régulon SahR est conservé. Il comprend les enzymes requises pour le cycle du SAM : SAH hydrolase AhaY, les enzymes de biosynthèse de la méthionine MetE/MetH/MetF et la SAM synthétase MetK. Le gène *sahR* représente un nouveau type de régulateur transcriptionnel qui contrôle le métabolisme des acides aminés soufrés. Il a été démontré chez *Delta-proteobacterium-Desulfo-vibrioalaskensis* que SahR est un régulateur inhibiteur d'expression qui répond à l'accumulation de la S-adénosylhomocystéine (SAH) dans la cellule. Le niveau élevé de la SAH (S-adénosylhomocystéine) empêche la liaison de la protéine SahR avec ses opérateurs d'ADN et induit par conséquence l'expression de tous les gènes du cycle S-Adénosyl-Méthionine régulés par le gène *sahR*. La SAH se produit de la SAM après le transfert du groupement méthyle (Novichkov et al., 2014). Dès lors le rapport SAM/SAH dans la cellule est finement régulé. La fonction physiologique de la SahR chez *E. aridi* LMR001 reste à démontrer mais ces travaux suggèrent une fonction similaire et permet à la cellule de réguler l'activité du cycle de la méthionine.

Les gènes impliqués dans le catabolisme de l'inositol (*iola* et *iolRCDEB*) ont également été induits sous stress pour *Ensifer aridi* LMR001. Ces gènes sont également présents dans le génome d'*Ensifer meliloti* et dans le génome de la bactérie *Bacillus subtilis*. L'étude fonctionnelle par mutagenèse des gènes *iola*, *smc01163* et *iolRCDE* de *Ensifer meliloti* a montré une relation entre l'expression génique de ces gènes et la croissance de la bactérie. L'expression des gènes impliqués dans le catabolisme d'inositol est essentielle pour la croissance de la bactérie pour l'utilisation du myo-isomère, ainsi qu'à la croissance avec deux molécules scyllo- et D-chiro-inositol comme sources de carbone. La myo-inositol déshydrogénase est responsable de l'oxydation du D-chiro-inositol. Le gène *smc01163* est également essentiel à la croissance avec le myo-inositol et code pour la famille *IdhA/MocA/GFO*. Le myo-inositol et surtout le scyllo-inositol sont révélés comme de bonnes sources de carbone pour *E. meliloti* comparés à d'autres sources de C, comme le glucose,

le glycérol, ou le succinate. Une étude faite sur un locus génétique requis pour le catabolisme de l'inositol comme seule source de carbone, qui a été cloné à partir d'*E. meliloti* (Kohler et al., 2010), a confirmé que le gène de régulation *iolR* situé en amont des gènes (*iolCDEB*) joue le rôle de répresseur. En effet, l'expression des enzymes du catabolisme d'inositol était 4 à 7 fois plus élevée dans la souche *iolR* mutante en comparaison de la souche sauvage. Les auteurs ont également trouvé que la mutation du gène *iolR* réduit la capacité de compétition avec la souche sauvage pour leur propre plante hôte (Kohler et al., 2010). Une grande variété de bactéries associée à la vie dans le sol et dans les plantes, a été décrite avec la capacité de dégrader le myo-inositol, comme *Klebsiella pneumoniae*, *R. leguminosarum* bv. *viciae*, *B. subtilis* et *E. meliloti* (Moreira et al., 2008).

Chez *K. pneumoniae* et *R. leguminosarum* bv. *viciae*, la voie de dégradation de l'inositol implique les enzymes myo-inositol déshydrogénase et 2-céto-myo-inositol déshydratase. Comme aussi identifié pour la souche *Ensifer aridi* LMR001, où le gène du myo-inositol déshydrogénase (*idhA*) qui a été montré comme essentiel pour l'utilisation du myo-inositol. De plus, les souches mutantes de *E. meliloti* avec une insertion dans le gène de la myo-inositol déshydrogénase se sont avérées incapables d'utiliser la rhizopine. Les souches qui abritaient une mutation dans le gène du myo-inositol déshydrogénase et ne présentaient pas d'activité myo-inositol déshydrogénase étaient incapables d'utiliser le myo-inositol comme seule source d'énergie carbonée et incapables de cataboliser la rhizopine. La rhizopine est un dérivé de l'inositol qui joue un rôle dans les associations symbiotiques fixatrices d'azote entre certains membres des *Rhizobiaceae* et les légumineuses dans les nodules (Avonce et al., 2006). Il est intéressant de noter que la présence des gènes de synthèse et de catabolisme de la rhizopine est rare parmi les différentes souches de *Rhizobiaceae*. Seules trois souches d'*E. meliloti* de 50 isolats testés se sont avérées capables de synthétiser et de cataboliser la rhizopine, et de posséder les gènes de synthèse et de catabolisme de la rhizopine (*rnoc*). Dans une autre étude appliquée sur plus de 300 souches, seulement 12 souches de *E. meliloti* et neuf de *Rhizobium leguminosarum* bv. *viciae* se sont avérées capables de cataboliser et de synthétiser la rhizopine. Les souches rhizobiennes capables de synthétiser et de cataboliser la rhizopine présentent un avantage compétitif dans la rhizosphère par rapport aux autres microorganismes. Dans notre étude, la mutation au niveau du gène régulateur de cette voie catabolique *iolR* chez *E. aridi* génère une croissance bactérienne similaire à la souche sauvage avec une légère retardation de la phase stationnaire en réponse au stress salin à une concentration de 0,4M de NaCl. Au niveau de la mobilité et la sensibilité aux stress acide/oxydatif/ détergent, ou du développement symbiotique, aucun effet significatif n'a été révélé par la mutation mis à part une meilleure croissance racinaire chez le mutant de *iolR* par rapport à la souche LMR001.

L'absence de séquence du gène *iolR* a perturbé l'expression de plusieurs gènes, *rpoE2*, *rsiA1*, *rsib1*, *iolC* et les gènes de la biosynthèse et du transport du tréhalose ce qui montre la complexité des réseaux de régulation chez les rhizobiums qui apparaissent interdépendants mais nécessite encore des études afin d'être clarifiés.

La voie de la biosynthèse et le transport du tréhalose sont aussi induits sous effet du stress chez *E. aridi* LMR001. La mutation au niveau du gène *otsB* qui code pour une enzyme glycosyles phospho-transférase (la tréhalose-6-phosphatase) chez la bactérie *E. aridi* a révélé une amélioration de la croissance dans les conditions de stress à 0,4M. L'efficacité symbiotique estimée par la croissance végétale des plantes hôtes ainsi que la nodulation sont également améliorées dans le mutant *otsB*. L'accumulation des glucides s'accroît généralement pendant la phase G1 du cycle de la croissance bactérienne. Lorsque les taux de croissance des bactéries sont faibles et les sources en carbone sont très fortes (Paalman et al., 2003). Ceci est corrélé avec la disponibilité et l'accumulation du glycogène et du tréhalose (Hazelwood et al., 2009 ; Brauer et al., 2008). Il est intéressant de savoir que pendant la phase stationnaire de croissance bactérienne, l'accumulation de tréhalose ne s'accompagne pas avec une activité élevée des enzymes tréhalose 6-P synthase. Il a été démontré que Tre6P est impliqué dans la régulation du métabolisme des phospholipides, la formation des vésicules, le métabolisme du carbone et de l'énergie et l'homéostasie Pi/ATP. Ainsi, on peut suggérer que le contrôle du métabolisme de la biosynthèse et l'accumulation de glycogène et du tréhalose est fortement lié aux changements de conditions de vie de la cellule bactérienne et les phases de croissance cellulaire.

La molécule de tréhalose a une propriété unique. Elle peut protéger les cellules membranaires de la dessiccation en stabilisant les protéines (Singer et Lindquist 1998). La propriété de ce disaccharide a aussi le potentiel d'activer le régulateur transcriptionnel pour la réponse au choc thermique ou pour la réponse au stress salin chez les eucaryotes. Elle est aussi définie essentielle dans la production de la protéine Hsf1 (Conlin et Nelson 2007; Bulman et Nelson 2005; Amorós et Estruch 2001). Le tréhalose peut protéger les cellules contre les radicaux et contre la toxicité de l'éthanol (Voorst et al., 2006 ; Benaroudj et al., 2001). L'accumulation de tréhalose a été aussi liée à la virulence chez les levures pathogènes (Al-Bader et al., 2010). Chez *P. aeruginosa*, il a été montré que le tréhalose favorise l'utilisation de l'azote lors de l'infection des plantes (Djonović et al., 2013). Chez *Aspergillus fumigatus*, la protéine tréhalose-6-phosphate synthase est codée par quatre gènes *tpsA*, *tpsB*, *tpsC* et *tpsD*. L'expression des deux gènes *tpsA* et *tpsB* s'accompagne d'une accumulation de tréhalose au cours du cycle de vie du champignon pathogène. Sous les effets du choc thermique, *A. fumigatus* accumule une grande quantité de tréhalose qui favorise la viabilité

et la croissance des spores. De plus, une mutation des gènes *tpsAB* diminue la viabilité du champignon et provoque un défaut de croissance et un défaut de la sporulation du champignon. La mutation affecte donc la virulence, la morphologie et la paroi cellulaire des champignons (Al-Bader et al., 2010). Également, la quantité de tréhalose accumulée par les rhizobiums influence l'abondance des nodules dans la plante et la survie des bactéroïdes dans les nodules. La teneur en tréhalose dans les bactéroïdes de *B. japonicum* augmente avec l'âge de la plante hôte (Streeter 1985a). Ici, les données transcriptomiques suggèrent l'accumulation du tréhalose par le biais de la synthèse *de novo* et la réduction des gènes d'utilisation (*thuAB*) pour la bactérie *E. aridi* LMR001 ce qui pourrait contribuer à améliorer le niveau de la tolérance au stress et la symbiose. Dans le cadre de nos travaux, nous avons pu montrer que les mutants des gènes *otsB* générant probablement une accumulation de tre6P et de la régulation du transport et de l'utilisation du tréhalose *thuR* génèrent une amélioration de la croissance des plantes inoculées suggérant que non seulement le trehalose mais également le tre6P représentent des molécules d'intérêt en biotechnologie pour l'amélioration de la symbiose fixatrice d'azote.

Nous avons également montré une réduction de la mobilité chez *E. aridi* sous l'effet du stress. La motilité bactérienne est liée à l'activité des gènes de la biosynthèse et du fonctionnement des flagelles. Dans le sol, les bactéries sont soit attirées par les nutriments comme les sucres, les acides aminés, l'oxygène ou repoussées par des substances répulsives (Kentner et al., 2006). La motilité améliore aussi le niveau de la pathogénèse des microbes. Elle permet aux bactéries de se disperser et explorer de multiples zones et facilite ainsi l'identification de sites d'infection. Le fait que la bactérie réduise sa mobilité à la suite d'un stress hyperosmotique peut en partie expliquer l'impact négatif que ce type de stress sur le développement symbiotique et en particulier les phases précoces de l'infection. Cela montre aussi qu'il est préférable d'inoculer en masse des microorganismes symbiotiques afin d'assurer un contact entre les racines et les rhizobiums si l'on souhaite améliorer la symbiose dans un sol où l'on retrouve ce type de stress.

La présente étude a aussi mis en évidence l'implication au niveau transcriptionnel de la GSR avec l'activation des gènes codant pour PhyP-PhyR-NepR- σ E chez la bactérie symbiotique *E. aridi* LMR001. La cascade de signalisation PhyR-NepR- σ E est connue comme facteur très important dans le processus de la réception des informations requises des changements externes vers l'espace intracellulaire. Il a été montré chez *E. meliloti* que l'activité de RpoE2 est régulée négativement par deux facteurs anti-sigma paralogues des RsiA1 et RsiA2, produits respectivement de SMc01505 et SMc04884, et inversement induite par deux régulateurs de réponse paralogues RsiB1 et RsiB2 produits respectivement de SMc01504 et SMc00794 (Bastiat et al., 2010). Comme

indiqué plus haut, les régulateurs de la réponse RsiB1 et RsiB2 n'agissent pas au niveau de la transcription du gène *rpoE2* mais agissent en anti-anti-sigma pour libérer RpoE2 des anti-sigma RsiA1 et RsiA2. Il est intéressant de noter que les transcrits d'ARN putatifs du gène *rsiA2* ont été identifiés dans une étude chez *E. meliloti* pendant la phase exponentielle de la croissance bactérienne (Mao et al., 2008). Parmi les gènes activés lors du passage de la souche *E. meliloti* en phase de croissance stationnaire les auteurs ont ainsi pu identifier le gène *rpoE2* (smc01506) induit également par diverses conditions de stress. Ils ont mis en évidence que le gène *rpoE2* contrôle sa propre transcription liée à la phase stationnaire de la croissance. Cette phase est caractérisée par une privation en carbone ou en azote (stress dû à la famine) et par une limitation en oxygène (Sauviac et al., 2007). Des résultats indiquent que la réduction de l'activité phosphatase de PhyP (Sensor kinase) conduit à l'activation de la cascade PhyR-NepR- σ E et permet par la suite de déclencher le mécanisme via l'activation du système de signalisation. PhyP contient un domaine périplasmique putatif et un domaine HAMP de liaison typique des histidines kinases associées à la membrane. Cependant, il existe d'autres modulateurs de la phosphorylation de PhyR, tels que des histidine kinases à capteur apparentés ou des petits donneurs phospho-endogènes (Kaczmarczyk et al., 2011). Durant le test de la croissance du mutant *E. aridi* LMR001 Δ *rpoE2*, la mutation du gène *rpoE2* n'a pas montré de réponse phénotypique différente à la souche de type sauvage. Cette absence de différence dans la réponse phénotypique est peut-être due à la présence d'autres gènes qui codent également pour un facteur sigma ECF et qui assurent la même fonction que celle-ci. Un article décrit que le génome de *E. meliloti* contient plusieurs gènes qui codent pour les facteurs sigma ECF. Il serait intéressant de mener des études complémentaires pour tester si certains d'autres gènes codant pour ce type de facteurs sigma sont présents chez *E. aridi* et peuvent compléter l'absence du facteur sigma RpoE2 comme observé également dans d'autres systèmes de signalisation des facteurs sigma ECF chez *B. subtilis* (R. K. Poole 2002; Sauviac et al., 2007). Il a été suggéré que certains gènes pourraient être régulés par RpoE et RpoS chez *Salmonella* sous condition de stress environnemental, mais les preuves expérimentales font défaut. Les facteurs sigma RpoE (σ E) et RpoS (σ S) gouvernent l'initiation de la transcription de deux ensembles de gènes chez *Escherichia* et *Salmonella* (Du et al., 2011). La régulation de l'expression génique par les facteurs sigma RpoE et RpoS chez *Salmonella enterica* ont été identifiés au cours de la réponse au stress hyperosmotique. Une étude des profils de transcription globaux réalisée à l'aide d'une micropuce à ADN génomique sur des mutants de *Salmonella* Δ *rpoE*, Δ *rpoS*, Δ *rpoE* / Δ *rpoS* a révélé qu'après 30 min de stress hyperosmotique, les niveaux d'expression de 38 gènes étaient nettement réduits pendant le stress hyperosmotique chez la souche double mutante Δ *rpoE* / Δ *rpoS*, et les niveaux d'expression n'étaient pas significativement affectés pour les simples

mutations *ΔrpoE* ou *ΔrpoS*. Ceci a été confirmé pour plusieurs gènes clés par RT-PCR. Cette étude a donc indiqué une diaphonie entre les facteurs sigma RpoE et RpoS chez *S. Typhi* dans des conditions hyperosmotiques et a fourni de nouvelles informations sur les réseaux de régulation (Du et al., 2011). Dans *Salmonella*, la régulation de l'expression et de l'activité de RpoS est nécessaire à la survie dans des conditions extrêmes et à la virulence de *S. enterica sérovar typhimurium* (Nickerson et Curtiss 1997). Le facteur RpoS est absent chez *E. meliloti*, ce qui suggère que la régulation pourrait être effectuée uniquement par des facteurs RpoE. Chez *E. meliloti*, 11 gènes *rpoE* sont prédits (Galibert et al., 2001) et 44 gènes dépendants de RpoE2 ont été identifiés. De plus, le gène *rpoE2* s'est révélé induit par un stress osmotique faible et également par le stress hyperosmotique (Sauviac et al., 2007). Ces résultats suggèrent que RpoE2 est impliquée dans la régulation d'autres gènes chez *E. aridi* LMR001 qu'une étude transcriptomique sur le mutant *rpoE2* pourrait permettre d'identifier. Cependant, nous avons testé l'effet de l'absence de la protéine RpoE2 sur la transcription de gènes cibles par fusion de leurs promoteurs avec le gène *gfp*-rapporteurs de transcription et évaluation par spectrofluormétrie. Les promoteurs de l'opéron *rsiA1-rpoE2* et du gène *rsiB1*, portés par le plasmide *pPROBE NT* et d'autres gènes liés à d'autres voies métaboliques régulées par le stress ont ainsi pu être étudiés et comparés entre le mutant *ΔrpoE2* et la souche sauvage. Nous avons ainsi pu confirmer que l'activation du promoteur des gènes *rsiA1-rpoE2* est dépendante de l'activité de la protéine RpoE2. De plus nous avons pu constater que l'activité du facteur sigma RpoE2 agit sur la biosynthèse et l'accumulation du principal osmoprotecteur utilisée par *E. aridi* (le tréhalose). En effet, la transcription du promoteur des gènes *otsBA* est réprimée dans les deux conditions de croissance dans la souche la LMR001 Δ *rpoE2* par rapport à la souche sauvage. L'activité des gènes *otsBA* génère deux réactions enzymatiques consécutives impliquées dans la synthèse du tréhalose. L'enzyme tréhalose-6-phosphate-synthase (TPS) codée par le gène *ostA* et une deuxième enzyme Tre6p-phosphatase (TPP) codée par le gène *otsB*, pour donner au final le tréhalose. Cette voie métabolique est identifiée up-réglée dans notre souche *E. aridi* LMR001 sous stress salin et hydrique par le test transcriptomique. Les deux enzymes tréhalose 6-P synthase et phosphatase sont généralement activées pour l'accumulation et la mobilisation du tréhalose dans la cellule bactérienne, et pour le maintien du taux de cette molécule sous condition de stress. L'adaptation au stress osmotique par l'accumulation de solutés compatibles facilite le maintien de la turgescence et la stabilisation des macromolécules. La régulation génétique de l'accumulation de solutés est mal comprise et n'a été décrite au niveau moléculaire que chez les entérobactéries. Dans *E. meliloti*, ils ont montré l'importance du facteur sigma alternatif RpoE2 dans l'osmoadaptation en réponse au stress hyperosmotique via l'induction du gène *otsA*. La construction et la caractérisation de *E. meliloti*

mutée $\Delta rpoE2$ a révélé une croissance compromise dans un milieu hyperosmotique. Ce défaut a été dû à l'absence de la molécule de tréhalose au cours de la phase initiale et la phase stationnaire de la croissance bactérienne des mutants. Ils ont démontré que les trois voies de synthèse du tréhalose dépendent de l'activité du RpoE2. De plus, ils ont confirmé que l'absence d'induction d'*otsA* dépendant de RpoE2 induction est la cause directe du phénotype osmotique du mutant *rpoE2*. En conclusion, l'article a montré que malgré la faible concentration en tréhalose dans *E. meliloti*, cette molécule reste le soluté le plus important utilisé par le mutant en réponse aux conditions hyperosmotiques (Flechar et al., 2010). Trois voies de synthèse du tréhalose ont été aussi décrites chez *E. meliloti* dans l'étude effectuée par (Streeter 1985b). Ils n'ont pas caractérisé tous les gènes correspondant à ces voies. Mais ils ont analysé l'expression d'un gène spécifique dans chaque voie (*treS*, *treY* et *otsA*). Ces trois gènes se sont révélés tous régulés par le gène *rpoE2*. Cette étude a montré que l'inactivation du gène *treS* n'a pas affecté la synthèse de tréhalose en milieu GAS, comme décrit dans *Propionibacterium freudenreichii*. D'autres chercheurs ont émis l'hypothèse que TreS pourrait être impliqué dans la dégradation du tréhalose (Cardoso et al., 2007). Seule la mutation *otsA* a aboli complètement la synthèse de tréhalose et a affecté directement la croissance du mutant en milieu hyperosmotique (Streeter 1985b). La synthèse de tréhalose osmorégulée est aussi médiée par l'activité de TreYZ chez *Corynebacterium glutamicum* (Wolf, Krämer, et Morbach 2003). Néanmoins, dans la plupart des bactéries possédant les trois voies du tréhalose, la voie *OtsAB* occupe toujours la principale et la majeure voie de synthèse du tréhalose. Comme observé chez *Mycobacterium tuberculosis* (Murphy et al., 2005), *Rhodobacter sphaeroides* (Makihara et al., 2005) et *P. freudenreichii* (Cardoso et al., 2007). Dans les bactéries qui possèdent le gène *rpoS* sigma factor, les gènes *otsA/B* se révèlent dépendants du RpoS et régulés positivement au cours d'un stress (Kandror et al., 2002; King et al., 2014). Cependant, La régulation de la synthèse du tréhalose a été étudiée au niveau moléculaire uniquement chez *E. coli* et *Salmonella typhimurium* qui possédaient la protéine RpoS. Dans ces deux bactéries apparentées, le tréhalose est exclusivement produit par l'activité de la voie *OtsBA*. L'opéron *otsBA* est identifié régulé par l'activité de RpoS, qui a été identifié comme un facteur de stress (Hengge et al., 1991). Ainsi l'activité de la RpoS permet la transcription de l'opéron *otsBA* en phase stationnaire où les cellules bactériennes sont soumises à un stress osmotique plus fort et au cours d'une croissance exponentielle où la multiplication des cellules bactériennes est plus active. Il a été identifié une accumulation très basse du tréhalose dans *E.coli* au cours de la phase stationnaire dans des milieux à faible osmolarité. Le tréhalose est produit à un niveau bas pendant la phase stationnaire dans un milieu non stressant et ce niveau de synthèse augmente suite à l'exposition à un stress osmotique élevé. Le gène *rpoS* n'est pas présent dans le génome d' *E. meliloti* (Galibert et al., 2001), mais la

synthèse du tréhalose chez *E. meliloti* présente aussi des caractéristiques similaires à celle observée chez *E. coli*. La synthèse du tréhalose est contrôlée par un facteur sigma alternatif RpoS pour les *entérobactéries* et RpoE2 pour *E. meliloti*. Ces facteurs sigma ont un rôle plus général dans la régulation du stress. Car ils contrôlent les gènes impliqués dans l'adaptation à d'autres stress (Sauviac et al., 2007). À cet égard, RpoE2 joue un rôle similaire chez *E. meliloti* et chez *E. aridi* à celui du RpoS chez *E. coli*. Le rôle protecteur du tréhalose peut également être important dans la réponse à un stress thermique (Hengge et al., 1991 ; Kandrór et al., 2002; Phadtare 2004). Une étude récente (Dominguez-Ferreras et al., 2009), réalisée dans des milieux de croissance différents, a montré que OtsA était le facteur principal de la synthèse du tréhalose dans des conditions de stress salin, et a montré également la participation de TreS et de TreY dans la synthèse de tréhalose sous stress en fonction des conditions et l'influence de la composition du milieu de culture. L'influence de la composition du milieu sur la synthèse du tréhalose a été décrite chez *E. meliloti* (Streeter 1985b) et d'autres micro-organismes (Beppu et al., 2004). TreS utilise le maltose comme substrat (Wolf et al., , 2003); TreY dépend de la disponibilité en glycogène (Carpinelli, Kramer, et Agosin 2006) ; d'autres études ont montré que les gènes de la biosynthèse du glycogène chez *E. meliloti* sont induits par la famine et sont dépendant de l'activité de RpoE2 (Sauviac et al., 2007). Les protéines régulatrices de la transduction du signal RpoE2 restent cependant à identifier. Les facteurs sigma extracytoplasmiques sont régulent la transcription de nombreux gènes lors du stress. Chez *E. meliloti*, RpoE2 est montré impliqué dans la tolérance acquise en phase stationnaire en réponse à un choc thermique et à l'effet de l'agent détergent l'H₂O₂. La résistance acquise à l'H₂O₂ est due à l'induction de la catalase KatC sous la dépendance de RpoE2 (Flechard et al., 2010). De ce fait, ce facteur sigma agit aussi sur le contrôle de la réponse au stress de l'enveloppe chez les bactéries. Dans le cas d'*E. coli* K-12, la partie du régulon conservée dans les génomes responsables de la synthèse, l'assemblage et l'homéostasie du lipopolysaccharide et des porines de la membrane externe, des constituants clés de la membrane externe des bactéries à Gram négatif est contrôlé par le facteur sigma E. Bien que le mécanisme moléculaire de l'activation de RpoE2 ne soit pas encore complètement décrit. Des articles confirment que son activation est directement déclenchée après la perturbation de l'enveloppe cellulaire en réponse à des stress, physiques et chimiques générant deux transductions de signal par différents facteurs (Sauviac et al., 2007). Chez *Salmonella typhimurium*, la mutagenèse par transposon et le profilage génétique ont révélé que les systèmes de réponse au stress de l'enveloppe cellulaire sigma 24 (RpoE) et Rcs phosphorelay-dépendants détectent des modifications des lipopolysaccharides (Imke Spöring et al., 2018). Il est possible que la perturbation de l'enveloppe de la bactérie *E. aridi* en condition de stress déclenche le même type de réponses via l'activation du sigma E.

Pour assurer leur viabilité et faire face aux problèmes de déshydratation. L'accumulation de tréhalose est impliquée dans la résistance à la dessiccation de *Rhizobium leguminosarum* et améliore les conditions de vie des cellules pendant la phase de croissance stationnaire. Les mutants qui n'accumulent pas de tréhalose présentent généralement des sévères formes de sensibilités en réponse à la dessiccation (McIntyre et al., 2007 ; Vriezen et al., 2007). Cela correspond à la corrélation étroite entre l'induction maximale du gène *rpoE2* et l'accumulation de tréhalose par *E. meliloti*. La mutation du gène *rpoE2* est liée à la réduction de la survie des microbes en réponse à la dessiccation (Humann et al., 2009). Nos résultats suggèrent que RpoE2 participe à la régulation de la synthèse du tréhalose. Il se peut que cette régulation soit aussi liée chez *E. aridi* à la tolérance à la dessiccation. Cela suggère que la participation du tréhalose à l'osmoadaptation ne se limite pas à la récupération de la turgescence, mais implique également des propriétés cyto-protectrices contre la dénaturation des macromolécules et la perte des propriétés énergétiques tout en offrant une disponibilité de carbone en cas de manque et une rétention de l'eau en cas de sécheresse (Yancey et al., 2005; Humann et al., 2009).

2. Conclusion générale et perspectives

La première étape de cette étude basée sur le séquençage des RNAs a déterminé les métabolismes affectés par les stress salin et hydrique dans *E. aridi* LMR001. Les fonctions affectées par le stress sont la motilité, la biosynthèse et le transport du tréhalose, le cycle de la méthionine et du SAM, le catabolisme de l'inositol. D'autres fonctions apparaissent altérées mais plus légèrement comme la production des exopolysaccharides. Une étude de génétique fonctionnelle a été entreprise durant ma thèse. Premièrement, on a réalisé des mutations par délétion des gènes sélectionnés dans certains métabolismes révélés affectés par le stress. Ensuite, on a effectué des tests phénotypiques comparatifs entre les mutants et la sauvage et une analyse de transcription des gènes cibles par l'utilisation des fusions de leurs promoteurs avec le gène *gfp*. Le premier article a également montré que la mutation par délétion du gène *rpoE2* codant pour le sigma facteur impliqué dans la Réponse générale au stress avec les gènes *rsiB1* et *rsiA1* n'impacte pas la tolérance de la bactérie aux stress salin, oxydatifs, acide ou générés par un détergent. Les mutants *rpoE2* ne présentent pas une altération de leur capacité à développer la symbiose. De plus, Les données transcriptionnelles relevées de l'expérience faite sur une comparaison entre la sauvage et les $\Delta rpoE$ a suggéré que l'activité du σ facteur RpoE2 induit le gène *thuR* et *otsA*. La mutation du gène *rpoE2* induit également l'expression des gènes impliqués dans le catabolisme de l'inositol et des deux gènes régulateurs du cycle de la méthionine ce qui suggère une interaction dans les réseaux de régulation du sigma facteur RpoE2 et de l'accumulation du tréhalose, du catabolisme de l'inositol et de la

régulation du cycle de la méthionine. Par la suite, une série de mutation a été réalisé sur *E. aridi* LMR001, par une délétion de deux gènes impliqués dans le métabolisme et de la régulation du transport et de l'utilisation du tréhalose (appelés respectivement *otsB* et *thuR*), du gène régulateur du le catabolisme d'inositol (*iolR*) et des deux régulateurs du cycle de la méthionine (*nesR* et *sahR*). Les résultats présentés dans l'article 2 montrent que les mutations des gènes régulateurs du cycle de la méthionine ont altéré sévèrement le développement et le fonctionnement de la symbiose avec deux plantes hôtes. Ces mutants ont également montré une très faible croissance dans un milieu hyperosmotique avec une dérégulation de leur motilité. Cette faible tolérance due à la perturbation de la régulation du cycle de la méthionine et donc de la synthèse du donneur universel SAM est potentiellement responsable de la réduction des capacités de la souche sauvage LMR001 à s'adapter aux stress et à produire une meilleure colonisation des plantes. En revanche, la mutation de ces deux gènes *nesR* et *sahR* accentue la tolérance aux stress détergents, oxydatifs ou acides, suggérant une altération de la membrane cellulaire.

La mutation des gènes *thuR* et *otsB* responsables de l'accumulation du tréhalose-6-P a accentué les formes de tolérance des mutants et aussi a augmenté les capacités symbiotiques de la bactérie et sa mobilité. Ces données confirment l'importance du tréhalose et probablement du tre6P dans l'adaptation au stress salin qui pourrait représenter un facteur déterminant de l'adaptation de de *E. aridi* aux conditions retrouvées dans les déserts. En particulier, il serait intéressant de poursuivre des études spécifiques sur l'accumulation de la molécule tréhalose-6-phosphate afin de déterminer son rôle dans l'interaction symbiotique notamment dans des environnements sous contraintes.

RÉFÉRENCES BIBLIOGRAPHIQUES

- Abbas, Basil A., K. E. Vineetha, C. Krishna Prasad, Neeraj Vij, Raad Hassani, et Gursharn S. Randhawa. 2002. « Symbiotic Characteristics of Cysteine and Méthionine Auxotrophs of *Sinorhizobium Meliloti* ». *Indian Journal of Experimental Biology* 40 (10): 1121-30.
- Abdullahi et al., 2020. « Symbiotic N₂ Fixation As An Alternative Source Of Nitrogen - A Review ».
- Al-Bader, Nadia, Ghyslaine Vanier, Hong Liu, Fabrice N. Gravelat, Mirjam Urb, Christopher M.-Q. Hoareau, Paolo Campoli, Joséé Chabot, Scott G. Filler, et Donald C. Sheppard. 2010. « Role of Trehalose Biosynthesis in *Aspergillus Fumigatus* Development, Stress Response, and Virulence ». *Infection and Immunity* 78 (7): 3007-18. <https://doi.org/10.1128/IAI.00813-09>.
- Amorós, M., et F. Estruch. 2001. « Hsf1p and Msn2/4p Cooperate in the Expression of *Saccharomyces Cerevisiae* Genes HSP26 and HSP104 in a Gene- and Stress Type-Dependent Manner ». *Molecular Microbiology* 39 (6): 1523-32.
- Ampomah, Osei Yaw, Anna Avetisyan, Espen Hansen, Johan Svenson, Thomas Huser, John Beck Jensen, et T. V. Bhuvaneswari. 2013. « The *ThuEFGKAB* Operon of *Rhizobia* and *Agrobacterium Tumefaciens* Codes for Transport of Trehalose, Maltitol, and Isomers of Sucrose and Their Assimilation through the Formation of Their 3-Keto Derivatives ». *Journal of Bacteriology* 195 (17): 3797-3807. <https://doi.org/10.1128/JB.00478-13>.
- Avonce, Nelson, Alfredo Mendoza-Vargas, Enrique Morett, et Gabriel Iturriaga. 2006. « Insights on the evolution of trehalose biosynthesis ». *BMC Evolutionary Biology* 6: 109. <https://doi.org/10.1186/1471-2148-6-109>.
- Avonce, Nelson, Jan Wuyts, Katrien Verschooten, Lies Vandesteene, et Patrick Van Dijck. 2010. « The *Cytophaga Hutchinsonii* ChTPSP: First Characterized Bifunctional TPS-TPP Protein as Putative Ancestor of All Eukaryotic Trehalose Biosynthesis Proteins ». *Molecular Biology and Evolution* 27 (2): 359-69. <https://doi.org/10.1093/molbev/msp241>.
- Ayangbenro, Ayansina Segun, et Olubukola Oluranti Babalola. 2021. « Reclamation of Arid and Semi-Arid Soils: The Role of Plant Growth-Promoting Archaea and Bacteria ». *Current Plant Biology* 25 (janvier): 100173. <https://doi.org/10.1016/j.cpb.2020.100173>.
- Ba, Salif, Anne Willems, Philippe de Lajudie, Philippe Roche, Habib Jeder, Paola Quatrini, Marc Neyra, et al., 2002. « Symbiotic and Taxonomic Diversity of *Rhizobia* Isolated from *Acacia Tortilis* Subsp. *Raddiana* in Africa ». *Systematic and Applied Microbiology* 25 (1): 130-45. <https://doi.org/10.1078/0723-2020-00091>.
- Bahlawane, Christelle, Matthew McIntosh, Elizaveta Krol, et Anke Becker. 2008. « *Sinorhizobium Meliloti* Regulator MucR Couples Exopolysaccharide Synthesis and Motility ». *Molecular Plant-Microbe Interactions: MPMI* 21 (11): 1498-1509. <https://doi.org/10.1094/MPMI-21-11-1498>.
- Balser, T. C., K. D. McMahon, D. Bart, D. Bronson, D. R. Coyle, N. Craig, M. L. Flores-Mangual, et al., 2006. « Bridging the Gap between Micro - and Macro-Scale Perspectives on the Role of Microbial Communities in Global Change Ecology ». *Plant and Soil* 289 (1-2): 59-70. <https://doi.org/10.1007/s11104-006-9104-5>.
- Barra, Lise, Catherine Fontenelle, Gwennola Ermel, Annie Trautwetter, Graham C. Walker, et Carlos Blanco. 2006. « Interrelations between Glycine Betaine Catabolism and Méthionine Biosynthesis in *Sinorhizobium Meliloti* Strain 102F34 ». *Journal of Bacteriology* 188 (20): 7195-7204. <https://doi.org/10.1128/JB.00208-06>.

- Bastiat, B., L. Sauviac, et C. Bruand. 2010. « Dual Control of Sinorhizobium Meliloti RpoE2 Sigma Factor Activity by Two PhyR-Type Two-Component Response Regulators ». *Journal of Bacteriology* 192 (8): 2255-65. <https://doi.org/10.1128/JB.01666-09>.
- Bauchrowitz, M. 1995. Lectin genes from *Medicago truncatula*: cloning, characterization and analysis of expression patterns during Rhizobium-legume symbiosis. *Thèse de Doctorat (Université de Freiburg, Allemagne)*
- Belfquih, Meryem, Abdelkarim Filali-Maltouf, et Antoine Le Quéré. 2022. « Analysis of Ensifer Aridi Mutants Affecting Regulation of Methionine, Trehalose, and Inositol Metabolisms Suggests a Role in Stress Adaptation and Symbiosis Development ». *Microorganisms* 10 (2): 298. <https://doi.org/10.3390/microorganisms10020298>.
- Belfquih, Meryem, Ilham Sakrouhi, Hassan Ait-Benhassou, Emeric Dubois, Dany Severac, Abdelkarim Filali-Maltouf, et Antoine Le Quere. 2021. « Ensifer Aridi LMR001T Symbiosis and Tolerance to Stress Do Not Require the Alternative Sigma Factor RpoE2 ». *Agronomy* 11 (9): 1787. <https://doi.org/10.3390/agronomy11091787>.
- Benaroudj, N., D. H. Lee, et A. L. Goldberg. 2001. « Trehalose Accumulation during Cellular Stress Protects Cells and Cellular Proteins from Damage by Oxygen Radicals ». *The Journal of Biological Chemistry* 276 (26): 24261-67. <https://doi.org/10.1074/jbc.M101487200>.
- Ben Rebah et al., 2007. Agro-industrial waste 588 materials and wastewater sludge for rhizobial inoculant production:” *A review. Bioresource Technology* 18, 3535-3546.
- Beppu, T., K. Toyomasu, H.-O. Ogawa, H. Kakuta, K. Hosono, K. Ueda, A. Seto, et H. Yoshijima. 2004. « Effective Extracellular Trehalose Production by *Cellulosimicrobium Cellulans* ». *Applied Microbiology and Biotechnology* 64 (6): 794-99. <https://doi.org/10.1007/s00253-003-1507-0>.
- Blanco, Julien, Mohamed Alifriqui, Kamil Hassan, Didier Genin, et Geneviève Michon. 2016. *Acacias et agropasteurs sahariens: façonnage d’un terroir et enjeux de conservation. IRD.*
- Bohnert, H. J., D. E. Nelson, et R. G. Jensen. 1995. « Adaptations to Environmental Stresses ». *The Plant Cell* 7 (7): 1099-1111. <https://doi.org/10.1105/tpc.7.7.1099>.
- Bohnert, Hans J., et Richard G. Jensen. 1996. « Strategies for engineering water-stress tolerance in plants ». *Trends in Biotechnology* 14 (3): 89-97. [https://doi.org/10.1016/0167-7799\(96\)80929-2](https://doi.org/10.1016/0167-7799(96)80929-2).
- Booth, I. R. 2002. Stress and the single cell : intrapopulation diversity is a mechanism to ensure survival upon exposure to stress. *Int J Food Microbiology*. 78(1-2):19-30.
- Bordeleau L. M. and D. Prevost. 1994. Nodulation and nitrogen fixation in extreme environments. *Plant Soil* 161:115-124
- Boscari A., Mandon K., Dupont L., Poggi M. C., and Le Rudulier D. 2002. BetS is a major glycine betaine/proline betaine transporter required for early osmotic adjustment in *Sinorhizobium meliloti*. *J. Bacteriol.* 184:2654–2663.
- Botumile, Amantle, Demel Teketay, Witness Mojeremane, et Thembinkosi Mathowa. 2020. « OVERCOMING SEED DORMANCY OF *SENEGALIA GALPINII* AND *VACHELLIA ROBUSTA* THROUGH SCARIFICATION PRE-SOWING TREATMENTS ». *The Journal « Agriculture and Forestry »* 66 (1). <https://doi.org/10.17707/AgricultForest.66.1.15>.
- Brauer, Matthew J., Curtis Huttenhower, Edoardo M. Airolidi, Rachel Rosenstein, John C. Matese, David Gresham, Viktor M. Boer, Olga G. Troyanskaya, et David Botstein. 2008. « Coordination of Growth Rate, Cell Cycle, Stress Response, and Metabolic Activity in Yeast ». *Molecular Biology of the Cell* 19 (1): 352-67. <https://doi.org/10.1091/mbc.E07-08-0779>.

- Bremer, E., D., and R. Krämer. 2000. Coping with osmotic challenges: osmoregulation through accumulation and release of compatible solutes in bacteria. In Stroz G., Hengge Aronis R (eds). *Bacterial Stress Responses*. ASM Press Washington DC: 79:97.
- Broughton, W. J., et M. J. Dilworth. 1971. « Control of Leghaemoglobin Synthesis in Snake Beans ». *Biochemical Journal* 125 (4): 1075-80. <https://doi.org/10.1042/bj1251075>.
- Bulman, Amanda L., et Hillary C. M. Nelson. 2005. « Role of Trehalose and Heat in the Structure of the C-Terminal Activation Domain of the Heat Shock Transcription Factor ». *Proteins* 58 (4): 826-35. <https://doi.org/10.1002/prot.20371>.
- Burghardt, Liana T. 2020. « Evolving Together, Evolving Apart: Measuring the Fitness of Rhizobial Bacteria in and out of Symbiosis with Leguminous Plants ». *New Phytologist* 228 (1): 28-34. <https://doi.org/10.1111/nph.16045>.
- Calatrava-Morales, Nieves, Matthew McIntosh, et María Soto. 2018. « Regulation Mediated by N-Acyl Homoserine Lactone Quorum Sensing Signals in the Rhizobium-Legume Symbiosis ». *Genes* 9 (5): 263. <https://doi.org/10.3390/genes9050263>.
- Cardoso, F. S., R. F. Castro, N. Borges, et H. Santos. 2007. « Biochemical and Genetic Characterization of the Pathways for Trehalose Metabolism in *Propionibacterium Freudenreichii*, and Their Role in Stress Response ». *Microbiology* 153 (1): 270-80. <https://doi.org/10.1099/mic.0.29262-0>.
- Carpinelli, J., R. Kramer, et E. Agosin. 2006. « Metabolic Engineering of *Corynebacterium Glutamicum* for Trehalose Overproduction: Role of the TreYZ Trehalose Biosynthetic Pathway ». *Applied and Environmental Microbiology* 72 (3): 1949-55. <https://doi.org/10.1128/AEM.72.3.1949-1955.2006>.
- Chang, W.-S., M. van de Mortel, L. Nielsen, G. Nino de Guzman, X. Li, et L. J. Halverson. 2007. « Alginate Production by *Pseudomonas Putida* Creates a Hydrated Microenvironment and Contributes to Biofilm Architecture and Stress Tolerance under Water-Limiting Conditions ». *Journal of Bacteriology* 189 (22): 8290-99. <https://doi.org/10.1128/JB.00727-07>.
- Cho, Kyuil, Bradley S. Evans, B. McKay Wood, Ritesh Kumar, Tobias J. Erb, Benjamin P. Warlick, John A. Gerlt, et Jonathan V. Sweedler. 2015. « Integration of Untargeted Metabolomics with Transcriptomics Reveals Active Metabolic Pathways ». *Metabolomics* 11 (3): 503-17. <https://doi.org/10.1007/s11306-014-0713-3>.
- Cloutier J. Prevost D., Nadeau P., Antoun H. 1992. Heat and shock protein synthesis in arctic and temperature strains of rhizobia. *Appl. Environ. Microbiol.* 58, 2846-2853.
- Comben, D. F., G. A. McCulloch, G. K. Brown, et G. H. Walter. 2020. « Phylogenetic Placement and the Timing of Diversification in Australia's Endemic *Vachellia* (Caesalpinioideae, Mimosoid Clade, Fabaceae) Species ». *Australian Systematic Botany* 33 (1): 103. <https://doi.org/10.1071/SB19013>.
- Conlin, Laura K., et Hillary C. M. Nelson. 2007. « The Natural Osmolyte Trehalose Is a Positive Regulator of the Heat-Induced Activity of Yeast Heat Shock Transcription Factor ». *Molecular and Cellular Biology* 27 (4): 1505-15. <https://doi.org/10.1128/MCB.01158-06>.
- Danchin, Antoine. 2017. « Coping with Inevitable Accidents in Metabolism ». *Microbial Biotechnology* 10 (1): 57-72. <https://doi.org/10.1111/1751-7915.12461>.
- Davies, Simon. 2005. « Evaluation of *Dorycnium* Spp », 188.
- Dicenzo, George C., Maryam Zamani, Alice Checcucci, Marco Fondi, Joel S. Griffiths, Turlough M. Finan, et Alessio Mengoni. 2019. « Multidisciplinary Approaches for Studying Rhizobium-Legume Symbioses ». *Canadian Journal of Microbiology* 65 (1): 1-33. <https://doi.org/10.1139/cjm-2018-0377>.

- Djonović, Slavica, Jonathan M. Urbach, Eliana Drenkard, Jenifer Bush, Rhonda Feinbaum, Jonathan L. Ausubel, David Traficante, et al., 2013. « Trehalose Biosynthesis Promotes *Pseudomonas aeruginosa* Pathogenicity in Plants ». *PLOS Pathog* 9 (3): e1003217. <https://doi.org/10.1371/journal.ppat.1003217>.
- Dominguez-Ferreras, A., M. J. Soto, R. Perez-Arnedo, J. Olivares, et J. Sanjuan. 2009. « Importance of Trehalose Biosynthesis for *Sinorhizobium Meliloti* Osmotolerance and Nodulation of Alfalfa Roots ». *Journal of Bacteriology* 191 (24): 7490-99. <https://doi.org/10.1128/JB.00725-09>.
- Du, Hong, Min Wang, Zhe Luo, Bin Ni, Fei Wang, Yanchen Meng, Shungao Xu, et Xinxiang Huang. 2011. « Coregulation of Gene Expression by Sigma Factors RpoE and RpoS in *Salmonella Enterica* Serovar Typhi During Hyperosmotic Stress ». *Current Microbiology* 62 (5): 1483-89. <https://doi.org/10.1007/s00284-011-9890-8>.
- Elboutahiri, Nadia, Imane Thami-Alami, et Sripada M Udupa. 2010. « Phenotypic and Genetic Diversity in *Sinorhizobium Meliloti* and *S. Medicae* from Drought and Salt Affected Regions of Morocco ». *BMC Microbiology* 10 (1): 15. <https://doi.org/10.1186/1471-2180-10-15>.
- Errington, J., 1993. *Bacillus subtilis* sporulation: Regulation of gene expression and control of morphogenesis. *Microbiological Reviews* 31: 637-656.
- Fagorzi, Camilla, Alexandru Ilie, Francesca Decorosi, Lisa Cangilioli, Carlo Viti, Alessio Mengoni, et George C diCenzo. 2020. « Symbiotic and Nonsymbiotic Members of the Genus *Ensifer* (Syn. *Sinorhizobium*) Are Separated into Two Clades Based on Comparative Genomics and High-Throughput Phenotyping ». Édité par Esperanza Martinez-Romero. *Genome Biology and Evolution* 12 (12): 2521-34. <https://doi.org/10.1093/gbe/evaa221>.
- Finn, Sarah, Orla Condell, Peter McClure, Alejandro Amézquita, et Séamus Fanning. 2013. « Mechanisms of Survival, Responses and Sources of *Salmonella* in Low-Moisture Environments ». *Frontiers in Microbiology* 4. <https://doi.org/10.3389/fmicb.2013.00331>.
- Flechard, M., C. Fontenelle, C. Blanco, R. Goude, G. Ermel, et A. Trautwetter. 2010. « RpoE2 of *Sinorhizobium Meliloti* Is Necessary for Trehalose Synthesis and Growth in Hyperosmotic Media ». *Microbiology* 156 (6): 1708-18. <https://doi.org/10.1099/mic.0.034850-0>.
- Francez-Charlot, Anne, Julia Frunzke, Christian Reichen, Judith Zingg Ebnetter, Benjamin Gourion, et Julia A. Vorholt. 2009. « Sigma factor mimicry involved in regulation of general stress response ». *Proceedings of the National Academy of Sciences* 106 (9): 3467-72.
- Fridovich, I., 1998. Oxygen toxicity: A radical explanation. *Journal of experimental Biology*. 210:1203-1209.
- Fuqua, Clay. 2006. « The QscR Quorum-Sensing Regulon of *Pseudomonas Aeruginosa*: An Orphan Claims Its Identity ». *Journal of Bacteriology* 188 (9): 3169-71. <https://doi.org/10.1128/JB.188.9.3169-3171.2006>.
- Gage, D. J. 2004. « Infection and Invasion of Roots by Symbiotic, Nitrogen-Fixing Rhizobia during Nodulation of Temperate Legumes ». *Microbiology and Molecular Biology Reviews* 68 (2): 280-300. <https://doi.org/10.1128/MMBR.68.2.280-300.2004>.
- Galardini, Marco, Alessio Mengoni, Matteo Brilli, Francesco Pini, Antonella Fioravanti, Susan Lucas, Alla Lapidus, et al., 2011. « Exploring the symbiotic pangenome of the nitrogen-fixing bacterium *Sinorhizobium meliloti* ». *BMC Genomics* 12: 235. <https://doi.org/10.1186/1471-2164-12-235>.

- Galibert, F. 2001. « The Composite Genome of the Legume Symbiont *Sinorhizobium Meliloti* ». *Science* 293 (5530): 668-72. <https://doi.org/10.1126/science.1060966>.
- Glenn, Sarah A., Nataliya Gurich, Morgan A. Feeney, et Juan E. González. 2007. « The ExpR/Sin Quorum-Sensing System Controls Succinoglycan Production in *Sinorhizobium Meliloti* ». *Journal of Bacteriology* 189 (19): 7077-88. <https://doi.org/10.1128/JB.00906-07>.
- González, Juan E., et Melanie M. Marketon. 2003. « Quorum Sensing in Nitrogen-Fixing Rhizobia ». *Microbiology and Molecular Biology Reviews: MMBR* 67 (4): 574-92.
- González, Juan F., et Vittorio Venturi. 2013. « A Novel Widespread Interkingdom Signaling Circuit ». *Trends in Plant Science* 18 (3): 167-74. <https://doi.org/10.1016/j.tplants.2012.09.007>.
- Gopal, Shubha, Daniela Berg, Nicole Hagen, Eva-Maria Schriefer, Regina Stoll, Werner Goebel, et Jürgen Kreft. 2010. « Maltose and Maltodextrin Utilization by *Listeria monocytogenes* Depend on an Inducible ABC Transporter which Is Repressed by Glucose ». *PLOS ONE* 5 (4): e10349. <https://doi.org/10.1371/journal.pone.0010349>.
- Gourion, B., A. Francez-Charlot, et J. A. Vorholt. 2008. « PhyR Is Involved in the General Stress Response of *Methylobacterium Extorquens* AM1 ». *Journal of Bacteriology* 190 (3): 1027-35. <https://doi.org/10.1128/JB.01483-07>.
- Grouzis, Michel, et Edouard Le floch. 2003. « Un arbre au désert : *Acacia raddiana* », 319.
- Gupta, Radhey S., et Amy Mok. 2007. « Phylogenomics and Signature Proteins for the Alpha Proteobacteria and Its Main Groups ». *BMC Microbiology* 7 (1): 106. <https://doi.org/10.1186/1471-2180-7-106>.
- Halverson, L. J., et M. K. Firestone. 2000. « Differential Effects of Permeating and Nonpermeating Solutes on the Fatty Acid Composition of *Pseudomonas Putida* ». *Applied and Environmental Microbiology* 66 (6): 2414-21. <https://doi.org/10.1128/AEM.66.6.2414-2421.2000>.
- Hazelwood, Lucie A., Michael C. Walsh, Marijke A. H. Luttik, Pascale Daran-Lapujade, Jack T. Pronk, et Jean-Marc Daran. 2009. « Identity of the Growth-Limiting Nutrient Strongly Affects Storage Carbohydrate Accumulation in Anaerobic Chemostat Cultures of *Saccharomyces Cerevisiae* ». *Applied and Environmental Microbiology* 75 (21): 6876-85. <https://doi.org/10.1128/AEM.01464-09>.
- Hengge-Aronis, R., W. Klein, R. Lange, Ms Rimmele, et W. Boos. 1991. « Trehalose synthesis genes are controlled by the putative sigma factor encoded by *rpoS* and are involved in stationary-phase thermotolerance in *Escherichia coli*. » *Journal of Bacteriology* 173 (24): 7918-24.
- Hoang, Hanh H., Anke Becker, et Juan E. González. 2004. « The LuxR Homolog ExpR, in Combination with the Sin Quorum Sensing System, Plays a Central Role in *Sinorhizobium Meliloti* Gene Expression ». *Journal of Bacteriology* 186 (16): 5460-72. <https://doi.org/10.1128/JB.186.16.5460-5472.2004>.
- Hoang, Hanh H., Nataliya Gurich, et Juan E. González. 2008. « Regulation of Motility by the ExpR/Sin Quorum-Sensing System in *Sinorhizobium Meliloti* ». *Journal of Bacteriology* 190 (3): 861-71. <https://doi.org/10.1128/JB.01310-07>.
- Hopkins W.G. 2003. *Physiologie végétale*. Université des Sciences et Technologie de Lille. *Edition de boeck*, 514 p.
- Humann, J. L., H. T. Ziemkiewicz, S. N. Yurgel, et M. L. Kahn. 2009. « Regulatory and DNA Repair Genes Contribute to the Desiccation Resistance of *Sinorhizobium Meliloti* Rm1021 ». *Applied and Environmental Microbiology* 75 (2): 446-53. <https://doi.org/10.1128/AEM.02207-08>.

- Imke Spöring, Sebastian Felgner, Matthias Preuße, Denitsa Eckweiler, Manfred Rohde, Susanne Häussler, Siegfried Weiss, et Marc Erhardt. 2018. « Regulation of Flagellum Biosynthesis in Response to Cell Envelope Stress in *Salmonella enterica* Serovar Typhimurium ».
- Janczarek, Monika. 2011. « Environmental Signals and Regulatory Pathways That Influence Exopolysaccharide Production in *Rhizobia* ». *International Journal of Molecular Sciences* 12 (11): 7898-7933. <https://doi.org/10.3390/ijms12117898>.
- Jiang, Ju Quan, Wei Wei, Bing Hai Du, Xiao Hong Li, Lei Wang, et Su Sheng Yang. 2004. « Salt-Tolerance Genes Involved in Cation Efflux and Osmoregulation of *Sinorhizobium Fredii* RT19 Detected by Isolation and Characterization of Tn5 Mutants ». *FEMS Microbiology Letters* 239 (1): 139-46. <https://doi.org/10.1016/j.femsle.2004.08.029>.
- Juergen Prell. 2006. Metabolic changes of rhizobia in legume nodules. *TRENDS in Microbiology*. Vol.14 No.4
- Kaczmarczyk, A., S. Campagne, F. Danza, L. C. Metzger, J. A. Vorholt, et A. Francez-Charlot. 2011. « Role of *Sphingomonas* Sp. Strain Fr1 PhyR-NepR- EcfG Cascade in General Stress Response and Identification of a Negative Regulator of PhyR ». *Journal of Bacteriology* 193 (23): 6629-38. <https://doi.org/10.1128/JB.06006-11>.
- Kandror, Olga, Ann DeLeon, et Alfred L. Goldberg. 2002. « Trehalose synthesis is induced upon exposure of *Escherichia coli* to cold and is essential for viability at low temperatures ». *Proceedings of the national academy of sciences* 99 (15): 9727-32.
- Kentner, David, Sebastian Thiem, Markus Hildenbeutel, et Victor Sourjik. 2006. « Determinants of Chemoreceptor Cluster Formation in *Escherichia Coli* ». *Molecular Microbiology* 61 (2): 407-17. <https://doi.org/10.1111/j.1365-2958.2006.05250.x>.
- Kerppola, T. K., et M. L. Kahn. 1988. « Symbiotic Phenotypes of Auxotrophic Mutants of *Rhizobium Meliloti* 104A14 ». *Journal of General Microbiology* 134 (4): 913-19. <https://doi.org/10.1099/00221287-134-4-913>.
- Kiliç NK, Stensballe A, Otzen DE, Dönmez G, 2009. Proteomic changes in response to chromium (VI) toxicity in *Pseudomonas aeruginosa*. *Bioresource technology* 101: 2134-2140.
- Kim, Choong-Hyun, L. David Kuykendall, Ketan S. Shah, et Donald L. Keister. 1988. « Induction of Symbiotically Defective Auxotrophic Mutants of *Rhizobium Fredii* HH303 by Transposon Mutagenesis ». *Applied and Environmental Microbiology* 54 (2): 423-27.
- King, Thea, Chawalit Kocharunchitt, Kari Gobius, John P. Bowman, et Tom Ross. 2014. « Global Genome Response of *Escherichia Coli* O157:H7 Sakai during Dynamic Changes in Growth Kinetics Induced by an Abrupt Temperature Downshift ». Édité par Dipankar Chatterji. *PLoS ONE* 9 (6): e99627. <https://doi.org/10.1371/journal.pone.0099627>.
- Kinkema M. 2006. Legume nodulation: successful symbiosis through short and long distance signaling. *Func. Plant Biol.* 33:707-721
- Kohler, Petra R. A., Jasmine Y. Zheng, Elke Schoffers, et Silvia Rossbach. 2010. « Inositol Catabolism, a Key Pathway in *Sinorhizobium Meliloti* for Competitive Host Nodulation ». *Applied and Environmental Microbiology* 76 (24): 7972-80. <https://doi.org/10.1128/AEM.01972-10>.
- Kohler, Petra, et Silvia Rossbach. 2013. « Bacterial Inositol Catabolism—A Sweet Ride into the Host ».
- Kuykendall, L. David, Fawzy M. Hashem, et En Tao Wang. 2005. « *Sinorhizobium* Chen, Yan and Li 1988b, 396VP Emend. de Lajudie, Willems, Pot, Dewettinck, Maestrojuan,

Neyra, Collins, Dreyfus, Kersters and Gillis 1994, 732 ». In *Bergey's Manual® of Systematic Bacteriology*, édité par Don J. Brenner, Noel R. Krieg, George M. Garrity, James T. Staley, David R. Boone, Paul Vos, Michael Goodfellow, Fred A. Rainey, et Karl-Heinz Schleifer, 358-61. New York: Springer-Verlag. https://doi.org/10.1007/0-387-29298-5_88.

- Landini, P. 2009. Cross-talk mechanisms in biofilm formation and responses to environmental and physiological stress in *Escherichia coli*. *Research in Microbiology* 160:259-266.
- Le Quéré, Antoine, Lavire Celine, Meyer Thibault, Chapulliot David, Rathie Sonam, Sakrouhi Ilham, Rocha Guadalupe, et al., 2017. « Genomic characterization of *Ensifer aridi*, a proposed new species of nitrogen-fixing rhizobium recovered from Asian, African and American deserts ». *figshare*, janvier. <https://doi.org/10.6084/m9.figshare.c.3666331>.
- Lebre, Pedro H., Pieter De Maayer, et Don A. Cowan. 2017. « Xerotolerant Bacteria: Surviving through a Dry Spell ». *Nature Reviews Microbiology* 15 (5): 285-96. <https://doi.org/10.1038/nrmicro.2017.16>.
- Li, K., G. Li, L. M. T. Bradbury, A. D. Hanson, et S. D. Bruner. 2016. « Crystal Structure of the Homocysteine Methyltransferase MmuM from *Escherichia coli* ». *Biochemical Journal* 473 (3): 277-84. <https://doi.org/10.1042/BJ20150980>.
- Liu, Y., W. Gao, Y. Wang, L. Wu, X. Liu, T. Yan, E. Alm, et al., 2005. « Transcriptome Analysis of *Shewanella oneidensis* MR-1 in Response to Elevated Salt Conditions ». *Journal of Bacteriology* 187 (7): 2501-7. <https://doi.org/10.1128/JB.187.7.2501-2507.2005>.
- Makihara, Fumihito, Minoru Tsuzuki, Kiichi Sato, Shinji Masuda, Kenji V. P. Nagashima, Mitsuru Abo, et Akira Okubo. 2005. « Role of Trehalose Synthesis Pathways in Salt Tolerance Mechanism of *Rhodobacter sphaeroides* f. sp. *denitrificans* IL106 ». *Archives of Microbiology* 184 (1): 56-65. <https://doi.org/10.1007/s00203-005-0012-5>.
- Mao, Chunhong, Clive Evans, Roderick V Jensen, et Bruno WS Sobral. 2008. « Identification of New Genes in *Sinorhizobium meliloti* Using the Genome Sequencer FLX System ». *BMC Microbiology* 8 (1): 72. <https://doi.org/10.1186/1471-2180-8-72>.
- Marketon, Melanie M., Sarah A. Glenn, Anatol Eberhard, et Juan E. González. 2003. « Quorum Sensing Controls Exopolysaccharide Production in *Sinorhizobium meliloti* ». *Journal of Bacteriology* 185 (1): 325-31. <https://doi.org/10.1128/JB.185.1.325-331.2003>.
- Marketon, Melanie M., Matthew R. Gronquist, Anatol Eberhard, et Juan E. González. 2002. « Characterization of the *Sinorhizobium meliloti* SinR/SinI Locus and the Production of Novel N-Acyl Homoserine Lactones ». *Journal of Bacteriology* 184 (20): 5686-95. <https://doi.org/10.1128/JB.184.20.5686-5695.2002>.
- Marsh, E. Neil G., Dustin P. Patterson, et Lei Li. 2010. « Adenosyl Radical: Reagent and Catalyst in Enzyme Reactions ». *ChemBioChem* 11 (5): 604-21. <https://doi.org/10.1002/cbic.200900777>.
- McIntyre, H. J., H. Davies, T. A. Hore, S. H. Miller, J.-P. Dufour, et C. W. Ronson. 2007. « Trehalose Biosynthesis in *Rhizobium leguminosarum* Bv. *trifolii* and Its Role in Desiccation Tolerance ». *Applied and Environmental Microbiology* 73 (12): 3984-92. <https://doi.org/10.1128/AEM.00412-07>.
- Miller, M. B., et B. L. Bassler. 2001. « Quorum Sensing in Bacteria ». *Annual Review of Microbiology* 55: 165-99. <https://doi.org/10.1146/annurev.micro.55.1.165>.
- Moreira, Fátima, Anderson Lange, Osmar Klauber-Filho, José O. Siqueira, Rafaela SA Nóbrega, et Adriana S. Lima. 2008. « Associative diazotrophic bacteria in grass roots and soils from heavy metal contaminated sites ». *Anais da Academia Brasileira de Ciências* 80 (4): 749-61.

- Muller S. H. and Pereira P. A. A. 1995. Nitrogen fixation of common bean (*Phaseolus vulgaris* L.) as affected by mineral nitrogen supply at different growth stages. *Plant and Soil* .177, 55-61.
- Murphy, Helen N., Graham R. Stewart, Vladimir V. Mischenko, Alexander S. Apt, Richard Harris, Mark S. B. McAlister, Paul C. Driscoll, Douglas B. Young, et Brian D. Robertson. 2005. « The OtsAB Pathway Is Essential for Trehalose Biosynthesis in Mycobacterium Tuberculosis ». *Journal of Biological Chemistry* 280 (15): 14524-29. <https://doi.org/10.1074/jbc.M414232200>.
- Narita, Yukio, Hiromu Taguchi, Toshihide Nakamura, Akihiro Ueda, Weiming Shi, et Tetsuko Takabe. 2004. « Characterization of the salt-inducible méthioninesynthase from barley leaves ». *Plant Science* 167 (5): 1009-16. <https://doi.org/10.1016/j.plantsci.2004.05.039>.
- Nickerson, Cheryl A., et Roy Curtiss. 1997. « Role of sigma factor RpoS in initial stages of Salmonella typhimurium infection. » *Infection and immunity* 65 (5): 1814-23.
- Novichkov, Pavel S., Xiaoqing Li, Jennifer V. Kuehl, Adam M. Deutschbauer, Adam P. Arkin, Morgan N. Price, et Dmitry A. Rodionov. 2014. « Control of MéthionineMetabolism by the SahR Transcriptional Regulator in Proteobacteria ». *Environmental Microbiology* 16 (1): 1-8. <https://doi.org/10.1111/1462-2920.12273>.
- Nowak, Agnieszka, Izabela Greń, et Agnieszka Mrozik. 2016. « Changes in Fatty Acid Composition of Stenotrophomonas Maltophilia KB2 during Co-Metabolic Degradation of Monochlorophenols ». *World Journal of Microbiology and Biotechnology* 32 (12). <https://doi.org/10.1007/s11274-016-2160-y>.
- OByrne, C.P., and I.R Booth. 2002. Osmoregulation and its importance to food-borne microorganisms. *Int J Food Microbiology*. 74(3):203-16.
- Oldroyd, Giles E. D., et J. Allan Downie. 2004. « Calcium, Kinases and Nodulation Signalling in Legumes ». *Nature Reviews Molecular Cell Biology* 5 (7): 566-76. <https://doi.org/10.1038/nrm1424>.
- Oldroyd G. E., Murray J.D. 2013. Rhizobial infection is associated with the development of peripheral vasculature in nodules of Medicago truncatula. *Plant Physiology* .162 (1)107-15
- Paalman, Johannes W. G., René Verwaal, Sjoukje H. Slofstra, Arie J. Verkleij, Johannes Boonstra, et C. Theo Verrips. 2003. « Trehalose and Glycogen Accumulation Is Related to the Duration of the G1 Phase of Saccharomyces Cerevisiae ». *FEMS Yeast Research* 3 (3): 261-68.
- Patankar, Arati V., et Juan E. González. 2009. « Orphan LuxR Regulators of Quorum Sensing ». *FEMS Microbiology Reviews* 33 (4): 739-56. <https://doi.org/10.1111/j.1574-6976.2009.00163.x>.
- Patriarca E.J., Tate R., Ferraioli S. and Iaccarino M. 2004. Organogenesis of legume root nodules. *Int. Rev. Cytol.* 234: 201-62
- Payakapong W., Tittabutr P., Teaumroong N., Boonkerd N., Singleton P.W. and Borthakur D. 2006. Identification of Two Clusters of Genes Involved in Salt Tolerance in Sinorhizobium sp. Strain BL3. *Symbiosis* 41: 47-51
- Pellock, Brett J., Max Teplitski, Ryan P. Boinay, W. Dietz Bauer, et Graham C. Walker. 2002. « A LuxR Homolog Controls Production of Symbiotically Active Extracellular Polysaccharide II by Sinorhizobium Meliloti ». *Journal of Bacteriology* 184 (18): 5067-76.
- Perret X., Staehelin C. and Broughton W.J. 2000. Molecular basis of symbiotic promiscuity. *Microbiol. Mol. Biol. Rev.* 64:180-201.

- Phadtare, Sangita. 2004. « Recent developments in bacterial cold-shock response ». *Current issues in molecular biology* 6 (2): 125-36.
- Pobigaylo, Nataliya, Silke Szymczak, Tim W. Nattkemper, et Anke Becker. 2008. « Identification of Genes Relevant to Symbiosis and Competitiveness in *Sinorhizobium Meliloti* Using Signature-Tagged Mutants ». *Molecular Plant-Microbe Interactions: MPMI* 21 (2): 219-31. <https://doi.org/10.1094/MPMI-21-2-0219>.
- Poole, Philip, Vinoy Ramachandran, et Jason Terpolilli. 2018. « Rhizobia: From Saprophytes to Endosymbionts ». *Nature Reviews Microbiology* 16 (5): 291-303. <https://doi.org/10.1038/nrmicro.2017.171>.
- Poole, R.K. 2002. *Advances in Microbial Physiology*. Vol. 46. San Diego: Academic Press.
- Qian, Guoliang, Feifei Xu, Vittorio Venturi, Liangcheng Du, et Fengquan Liu. 2014. « Roles of a Solo LuxR in the Biological Control Agent *Lysobacter Enzymogenes* Strain OH11 ». *Phytopathology* 104 (3): 224-31. <https://doi.org/10.1094/PHYTO-07-13-0188-R>.
- Rao DLN, Giller K.E., Yeo A.R. and Flowers T.J. 2002. The effect of salinity and sodicity upon nodulation and nitrogen fixation in chickpea (*Cicer arietinum*). *Ann. Bot.* 89: 563- 570.
- Rasanen L. 2002. Biotic and abiotic factors influencing the development of N₂-fixing symbioses between rhizobia and the woody legumes *Acacia*." Thèse de doctorat de l'université de Helsinki. *Finlan Oldroyd*. 220p.
- Reid, A. N., R.Palyada, L.Whitworth, E. Doukhanine & A. Stintzi. 2008. Identification of *Campylobacter jejuni* genes contributing to acid adaptation by transcriptional profiling and genome-wide mutagenesis. *Applied and Environmental Microbiology*. 74:1598-1612
- Rejili, Mokhtar, Mohamed Ali BenAbderrahim, et Mohamed Mars. 2020. « Phylogenomic Review of Root Nitrogen-Fixing Symbiont Population Nodulating Northwestern African Wild Legumes ». In *Nitrogen Fixation*, édité par Everlon Cid Rigobelo et Ademar Pereira Serra. IntechOpen. <https://doi.org/10.5772/intechopen.87082>.
- Rocha, Guadalupe, Antoine Le Queré, Arturo Medina, Alma Cuéllar, José-Luis Contreras, Ricardo Carreño, Rocío Bustillos, et al., 2020. « Diversity and Phenotypic Analyses of Salt- and Heat-Tolerant Wild Bean *Phaseolus Filiformis* Rhizobia Native of a Sand Beach in Baja California and Description of *Ensifer Aridi* Sp. Nov. » *Archives of Microbiology* 202 (2): 309-22. <https://doi.org/10.1007/s00203-019-01744-7>.
- Rodriguez-Llorente, Ignacio, Miguel A Caviedes, Mohammed Dary, Antonio J Palomares, Francisco M Cánovas, et José M Peregrín-Alvarez. 2009. « The Symbiosis Interactome: A Computational Approach Reveals Novel Components, Functional Interactions and Modules in *Sinorhizobium Meliloti* ». *BMC Systems Biology* 3 (1): 63. <https://doi.org/10.1186/1752-0509-3-63>.
- Roje, Sanja. 2006. « S-Adenosyl-l-Methionine: Beyond the Universal Methyl Group Donor ». *Phytochemistry* 67 (15): 1686-98. <https://doi.org/10.1016/j.phytochem.2006.04.019>.
- Ruiz, Bryan, Alexandre Le Scornet, Laurent Sauviac, Antoine Rémy, Claude Bruand, et Eliane Meilhoc. 2019. « The Nitrate Assimilatory Pathway in *Sinorhizobium Meliloti*: Contribution to NO Production ». *Frontiers in Microbiology* 10 (juillet): 1526. <https://doi.org/10.3389/fmicb.2019.01526>.
- Sabag-Daigle, Anice, Jessica L. Dyszel, Juan F. Gonzalez, Mohamed M. Ali, et Brian M. M. Ahmer. 2015. « Identification of *sdiA*-regulated genes in a mouse commensal strain of *Enterobacter cloacae* ». *Frontiers in Cellular and Infection Microbiology* 5 (mai). <https://doi.org/10.3389/fcimb.2015.00047>.
- Sadowsky MJ. 2005. Soil stress factors influencing symbiotic nitrogen fixation. In: Werner D, Newton WE (eds) *Nitrogen fixation in agriculture, forestry, ecology, and the environment*. Springer. *The Netherlands*. pp 89-112

- Sakrouhi, Ilham, Meryem Belfquih, Laïla Sbabou, Patricia Moulin, Gilles Bena, Abdelkarim Filali-Maltouf, et Antoine Le Quéré. 2016. « Recovery of Symbiotic Nitrogen Fixing Acacia Rhizobia from Merzouga Desert Sand Dunes in South East Morocco – Identification of a Probable New Species of Ensifer Adapted to Stressed Environments ». *Systematic and Applied Microbiology* 39 (2): 122-31. <https://doi.org/10.1016/j.syapm.2016.01.001>.
- Sánchez-Aguayo, Inmaculada, José Manuel Rodríguez-Galán, Remedios García, José Torreblanca, et José Manuel Pardo. 2004. « Salt Stress Enhances Xylem Development and Expression of S-Adenosyl-L-Méthionine Synthase in Lignifying Tissues of Tomato Plants ». *Planta* 220 (2): 278-85. <https://doi.org/10.1007/s00425-004-1350-2>.
- Sankhla, Indu Singh, Nisha Tak, Raju Ram Meghwal, Sunil Choudhary, Alkesh Tak, Sonam Rathi, Janet I. Sprent, Euan K. James, et Hukam Singh Gehlot. 2017. « Molecular Characterization of Nitrogen Fixing Microsymbionts from Root Nodules of Vachellia (Acacia) Jacquemontii, a Native Legume from the Thar Desert of India ». *Plant and Soil* 410 (1-2): 21-40. <https://doi.org/10.1007/s11104-016-2838-9>.
- Santos, Helena, et Milton S. da Costa. 2002. « Compatible Solutes of Organisms That Live in Hot Saline Environments ». *Environmental Microbiology* 4 (9): 501-9. <https://doi.org/10.1046/j.1462-2920.2002.00335.x>.
- Sauviac, L., H. Philippe, K. Phok, et C. Bruand. 2007. « An Extracytoplasmic Function Sigma Factor Acts as a General Stress Response Regulator in Sinorhizobium Meliloti ». *Journal of Bacteriology* 189 (11): 4204-16. <https://doi.org/10.1128/JB.00175-07>.
- Sekowska, Agnieszka, Hiroki Ashida, et Antoine Danchin. 2019. « Revisiting the Méthionine Salvage Pathway and Its Paralogues ». *Microbial Biotechnology* 12 (1): 77-97. <https://doi.org/10.1111/1751-7915.13324>.
- Singer, M. A., et S. Lindquist. 1998. « Multiple Effects of Trehalose on Protein Folding in Vitro and in Vivo ». *Molecular Cell* 1 (5): 639-48.
- Sleator, R. D., and C. Hill. 2002. Bacterial osmoadaptation: the role of osmolytes in bacterial stress and virulence. *FEMS Microbiology Rev.* 26(1):49-71
- Sourjik, Victor, Paul Muschler, Birgit Scharf, et Rüdiger Schmitt. 2000. « VisN and VisR Are Global Regulators of Chemotaxis, Flagellar, and Motility Genes in Sinorhizobium (Rhizobium) Meliloti ». *Journal of Bacteriology* 182 (3): 782-88. <https://doi.org/10.1128/JB.182.3.782-788.2000>.
- Staron, Anna, Heidi J. Sofia, Sascha Dietrich, Luke E. Ulrich, Heiko Liesegang, et Thorsten Mascher. 2009. « The Third Pillar of Bacterial Signal Transduction: Classification of the Extracytoplasmic Function (ECF) σ Factor Protein Family: ECF σ Factor Classification ». *Molecular Microbiology* 74 (3): 557-81. <https://doi.org/10.1111/j.1365-2958.2009.06870.x>.
- Streeter, J. G. 1985a. « Accumulation of alpha,alpha-trehalose by Rhizobium bacteria and bacteroids. » *Journal of Bacteriology* 164 (1): 78-84.
- . 1985b. « Accumulation of alpha,alpha-trehalose by Rhizobium bacteria and bacteroids. » *Journal of Bacteriology* 164 (1): 78-84.
- Subramoni, Sujatha, et Vittorio Venturi. 2009. « LuxR-Family “Solos”: Bachelor Sensors/Regulators of Signalling Molecules ». *Microbiology (Reading, England)* 155 (Pt 5): 1377-85. <https://doi.org/10.1099/mic.0.026849-0>.
- Sun, Jing, Jonathan D. Todd, J. Cameron Thrash, Yanping Qian, Michael C. Qian, Ben Temperton, Jiazhen Guo, et al., 2016. « The Abundant Marine Bacterium Pelagibacter Simultaneously Catabolizes Dimethylsulfoniopropionate to the Gases Dimethyl Sulfide and Methanethiol ». *Nature Microbiology* 1 (8). <https://doi.org/10.1038/nmicrobiol.2016.65>.

- Taté, R., A. Riccio, E. Caputo, M. Iaccarino, et E. J. Patriarca. 1999. « The Rhizobium Etli MetZ Gene Is Essential for MéthionineBiosynthesis and Nodulation of Phaseolus Vulgaris ». *Molecular Plant-Microbe Interactions: MPMI* 12 (1): 24-34. <https://doi.org/10.1094/MPMI.1999.12.1.24>.
- Tehlivets, Oksana, Nermina Malanovic, Myriam Visram, Tea Pavkov-Keller, et Walter Keller. 2013. « S-Adenosyl-L-Homocysteine Hydrolase and Methylation Disorders: Yeast as a Model System ». *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease* 1832 (1): 204-15. <https://doi.org/10.1016/j.bbadis.2012.09.007>.
- Teplitski, Max, Anatol Eberhard, Matthew R. Gronquist, Mengsheng Gao, Jayne B. Robinson, et Wolfgang D. Bauer. 2003. « Chemical Identification of N-Acyl Homoserine Lactone Quorum-Sensing Signals Produced by Sinorhizobium Meliloti Strains in Defined Medium ». *Archives of Microbiology* 180 (6): 494-97. <https://doi.org/10.1007/s00203-003-0612-x>.
- Timmers A. C.J., Soupene E., Auriac M.C., De Billy F., Vasse J., Boistard P. and Truchet G. 2000. Saprophytic intracellular rhizobia in alfalfa nodules. *Mol. Plant-Microbe Interact.* 13: 1204-1213.
- Verstraeten, Natalie, Kristien Braeken, Bachaspatimayum Debkumari, Maarten Fauvart, Jan Fransaer, Jan Vermant, et Jan Michiels. 2008. « Living on a Surface: Swarming and Biofilm Formation ». *Trends in Microbiology* 16 (10): 496-506. <https://doi.org/10.1016/j.tim.2008.07.004>.
- Voorst, Frank van, Jens Houghton-Larsen, Lars Jønson, Morten C. Kielland-Brandt, et Anders Brandt. 2006. « Genome-Wide Identification of Genes Required for Growth of Saccharomyces Cerevisiae under Ethanol Stress ». *Yeast (Chichester, England)* 23 (5): 351-59. <https://doi.org/10.1002/yea.1359>.
- Vriezen, J. A. C., F. J. de Bruijn, et K. Nusslein. 2007. « Responses of Rhizobia to Desiccation in Relation to Osmotic Stress, Oxygen, and Temperature ». *Applied and Environmental Microbiology* 73 (11): 3451-59. <https://doi.org/10.1128/AEM.02991-06>.
- Whitehead, N. A., A. M. Barnard, H. Slater, N. J. Simpson, et G. P. Salmond. 2001. « Quorum-Sensing in Gram-Negative Bacteria ». *FEMS Microbiology Reviews* 25 (4): 365-404.
- Wolf, Andreas, Reinhard Krämer, et Susanne Morbach. 2003. « Three Pathways for Trehalose Metabolism in Corynebacterium Glutamicum ATCC13032 and Their Significance in Response to Osmotic Stress: Trehalose Synthesis in Corynebacterium Glutamicum ». *Molecular Microbiology* 49 (4): 1119-34. <https://doi.org/10.1046/j.1365-2958.2003.03625.x>.
- Yancey, P. H. 2005. « Organic Osmolytes as Compatible, Metabolic and Counteracting Cytoprotectants in High Osmolarity and Other Stresses ». *Journal of Experimental Biology* 208 (15): 2819-30. <https://doi.org/10.1242/jeb.01730>.
- Zahran H. 2003. Polyphasic taxonomy of symbiotic rhizobia from wild leguminous plants growing in Egypt. *Folia Microbiol.* 48:510-520.
- Zahran H. H. 1999. Rhizobium-legume symbiosis and nitrogen fixation under severe conditions and in an arid climate. *Microbiol. Mol. Biol.* 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., T. C. Charles, B. Pan, and D. L. Smith. 1996. Inhibition of the expression of *Bradyrhizobium japonicum nod* genes at low temperatures. *Soil Biol. Biochem.* 28, 1579-1583.
- Zerhari, K., J. Aurag, B. Khbaya, D. Kharchaf, et A. Filali-Maltouf. 2000. « Phenotypic Characteristics of Rhizobia Isolates Nodulating Acacia Species in the Arid and Saharan Regions of Morocco ». *Letters in Applied Microbiology* 30 (5): 351-57. <https://doi.org/10.1046/j.1472-765x.2000.00730.x>.

Zou, Huibin, Ningning Chen, Mengxun Shi, Mo Xian, Yimin Song, et Junhong Liu. 2016. « The Metabolism and Biotechnological Application of Betaine in Microorganism ». *Applied Microbiology and Biotechnology* 100 (9): 3865-76. <https://doi.org/10.1007/s00253-016-7462-3>.

1. Culture bactérienne et test phénotypique : Tolérance aux stress abiotiques

1.1 Stress salin en milieu solide/ en milieu liquide

Les souches de la bactérie *E.aridi* LMR001 WT et les mutants ont été cultivées dans des cultures TY (5g/l tryptone, 3g/l extrait de levure avec 8 g/l CaCl₂) liquides jusqu'à à une DO (600 nm) de 0,7, puis 1 ml de la culture bactérienne a été utilisé pour inoculer 50 ml de milieu TY liquide avec ou sans NaCl. Les concentrations finales de sel dans les milieux de culture varient de 0,25 M, 0,3 M et 0,4 M). Ensuite les cultures ont été incubées dans un roto-agitateur fixé à 180 RPM dans une température de 28 °C. Les croissances étaient mesurées à une DO (600 nm) et enregistrée après 4, 8, 24, 48, 120, et 144h d'incubation. Les moyennes des DO (600 nm) des trois cultures indépendantes ont été calculées pour toutes les souches et à chaque instant avec leurs écarts-types respectifs.

Pour le test sur milieu solide, spotter 5µl de chaque souche diluée à 10⁸ bactéries/ml sur du TY solide contenant les différentes concentrations de sel et laisser incuber à 28°C. Les résultats sont notés après 24, 48 et 72 heures de croissance.

1.2 Test du stress acide (HCL), détergent (SDS) et oxydatif (H2O2)

Le test de sensibilité en réponse aux stress oxydant, acide et détergent était induit respectivement par H₂O₂, HCl et SDS pour les souches LMR001 sauvage et les mutants. 4 ml de cultures bactériennes à une OD 600 nm de 0,7 en milieu liquide TY est ajoutées à 16 ml de gélose molle TY à 0,7% d'agar. Une fois le milieu de culture solidifié, un papier disque de 6 mm de diamètre a été placé sur la plaque et 5 µL de H₂O₂ (2 M), HCl (5,5 M), ou du SDS (10 % p / v) a été déposé dessus. Les plaques des trois répétitions ont été incubées à 28 °C pendant trois jours. Les diamètres des zones d'inhibition étaient mesurés.

1.3 Test de motilité

Les souches de la bactérie *E.aridi* LMR001 WT et les mutants ont été cultivés à 28 °C dans le milieu de culture TY sous agitation à 180 tr/min jusqu'à une DO (600 nm) de 0,7. Après incubation, les cellules ont été concentrées par centrifugation, lavées deux fois avec de l'eau stérile, et additionnées avec 10 volumes d'eau stérile. Le test de la motilité bactérienne a été lancé sur du

milieu TY à faible pourcentage d'agar de (0,2%) et ajusté à des concentrations différentes de NaCl (0M, 0.1M, 0.25M).

Pour réaliser ce test :

- « Spotter » 5 µl de chaque souche à la dilution (N°1) au centre de la boîte et veiller à faire entrer la suspension bactérienne à l'intérieur du milieu de culture utilisé.
- Laisser incubé la culture dans l'étuve à 28°C.
- Les résultats sont pris par la mesure du diamètre du halo bactérien (qui correspond à la distance de propagation) après 2, 3 et 5 jours de croissance. Le diamètre moyen et les écarts types de trois biologiques ont été calculés.

1.4 Test de nodulation

Les graines des plantes endémiques du Maroc, *Vachellia tortilis subsp. raddiana* et de la *Vachellia gummifera* (Blanco et al., 2016) ont été scarifiées et stérilisées avec l'acide sulfurique concentré (95%) sous agitation (1 h pour *V. tortilis* et 30 min pour *V. gummifera*) puis rincé abondamment à l'eau stérile. Les graines ont été mises après à germer sur des plaques de l'eau gélosée pour être incubé pendant 5 jours à 25 °C dans l'obscurité. Après la germination, les semis ont été plantés dans des pots contenant du sable préalablement stérilisé par autoclavage pendant 1 h à 120 °C. Les souches d'*Ensifer aridi LMR001 WT* et les mutants étaient pré-cultivée en milieu TY jusqu'à une densité optique (DO) 600 nm de 0,7 puis lavé et remis en suspension dans de l'eau stérilisée. L'inoculation est réalisée par 1 ml de cellules des cultures bactériennes lavées après la plantation dans le sable stérile. Six plantes individuelles ont été inoculées avec chaque souche de rhizobium ou avec l'eau stérile pour les contrôles négatifs et positifs. Les plantes ont été incubées dans une chambre de croissance à 26 °C sous une photopériode de 16 h jour et 8 h nuit et arrosées avec de la solution nutritive (Broughton et Dilworth 1971) avec ou sans azote KNO₃ (3 mM) pour les témoins positifs et les témoins négatifs. Après 6 mois de croissance, les plantes ont été récoltées, et le nombre de nodules, les pousses de plantes et les poids frais des racines ont été enregistrés.

2. Étude transcriptomique

L'étude transcriptomique, sous forme de la technique des microarrays et le séquençage des ARN séquences (RNA- seq), ont été utilisé pour étudier les réponses transcriptionnelles à une large gamme de conditions génétiques et environnementales. Pour décrire les expressions génétiques de la bactérie *Ensifer aridi LMR001T* dans les situations de stress et étudier les circuits des régulations transcriptionnels clés. Les séquences transcriptionnelles ont décrit une gamme de phénotypes, y

compris la croissance, la formation de biofilm, la répression de la motilité et les changements associés à la dépolarisation membranaire. L'étude transcriptomique est également utilisée pour déchiffrer le rôle des gènes surexprimés et réprimés dans la régulation complexes de différentes chaînes métaboliques, tel que : le processus de la symbiose et le processus de la tolérance aux stress abiotiques à l'état libre dans l'environnement. Afin d'examiner l'influence de ces variations génétiques sur le comportement et les capacités de la bactérie en réponse aux variations environnementales. Ces études sont essentiels pour mettre en évidence les effets importants de la variation des paramètres abiotiques et biotiques dans un temps donné sur l'activité transcriptionnelle de la bactérie (Dicenzo et al., 2019 ; Belfquih et al., 2021 ; Belfquih, Filali-Maltouf, et Le Quéré 2022).

La technique de l'analyse génomique utilisée dans cette étude est basée sur une étude transcriptomique. Elle a été faite pour caractériser le transcriptome de la souche *Ensifer aridi* LMR001 et pour interpréter les éléments fonctionnels du génome dans un état de croissance bactérienne définis. C'est une approche de séquençage des ARNs (ARNm messagers, ARNt de transfert et ARNr ribosomique). La réponse transcriptomique globale de la souche a été faite dans deux types d'environnement stressants (stress hydrique et stress salin). L'extraction des ARNs est réalisée à partir de trois différentes conditions de croissance dans des cultures avec 6 répliquas biologiques. En utilisant des kits d'extraction d'ARN bactériens commerciaux (Ribopure, Ambion). Chaque répliquât sert d'inoculum pour 3 cultures : (milieu TY, milieu TY+ 250mM de NaCl pour le stress salin, milieu TY+12% PEG8000 pour le stress hydrique). Les cultures bactériennes ont été préparées à partir des cultures fraîches d'une nuit de croissance d'*Ensifer aridi* LMR001T et arrivées à la phase exponentielle et ensuite incubées pendant un temps de génération estimé de 2H30min. Les résultats du séquençage des ARNs ont révélé une présence de majorité des ARN ribosomiaux avec un pourcentage de 95% des ARNs. Ensuite, Une séparation entre les ARNs et l'ARNr a été réalisée par le kit (Microb Express, Ambion). Après est venue l'étape finale du séquençage des ARNs traité et enrichis avec les ARN messagers. Cette étape est réalisée en France par le service (Genomix, Unité de Service CNRS, Montpellier, France), en utilisant la technique Illumina, HiSeq2000 (100pb/ lecture). Les données du séquençage ont été traitées et normalisées par la technique médiane du logarithme base 2 (log2) pour sélectionner les séquences géniques différentiellement exprimées entre les deux différentes conditions du stress et la condition de non-stress. Les reads séquence ont été importés dans CLC genomics plateforme pour l'annotation et après sur le programme RAST- Rapid Annotations using Subsystems Technology pour la cartographie et l'annotation des gènes. Les gènes putatifs codant pour les protéines (PEGs) ont été gardés dans l'étude et les ARNs qui présentaient une similarité de 95% avec plusieurs PEG

ont été rejetés. Au final, la table de lecture finale obtenue lors du mappage a été normalisée par une transformation LOESS. Les données sélectionnées pour chaque gène, la couverture de séquençage, le ratio du nombre de read (lecture) sont ensuite calculées et analysées par le programme Excel, suivant des paramètres bien définis. Les gènes qui montrent une expression 4x différente à celle révélée par le témoin de non-stress étaient sélectionnés pour identification. Le calcul d'Excel correspond à $(\log_2(\text{ratio}(\text{nb reads condition sous stress} / \text{nb de reads condition contrôle})) > 2 \text{ ou } < -2)$. Cette étape va nous permettre de sélectionner et de cribler les gènes et les fonctions les plus affectées par le stress. Elle va permettre d'explorer les changements majeurs associés aux transcriptomes. Le protéome prédit de la souche LMR001 a été annoté par les classifications COG et KEGG. Les gènes différentiellement régulés ont ensuite été utilisés pour estimer les principales classes et fonctions associées aux stress hyperosmotiques induits par le sel et le PEG8000 (Sakrouhi et al., 2016, Le Quéré et al., 2017, Belfquih et al., 2021, Belfquih et al., 2022). Ainsi, la caractérisation fonctionnelle du génome de la souche *E. aridi* LMR001 a généré des explications sur les différences des phénotypes et les différences de l'expression génétiques, via la mutation par la suppression globale des gènes cibles, en réponse aux stress. L'utilisation des données de cette étude a servi également à l'analyse de l'activité transcriptionnelle des gènes à l'échelle du génome et a aidé à produire les réseaux de régulation inter/intra métaboliques dans différents stades de la vie du *rhizobium*, et dans différents environnements de stress

3. Mutagenèse dirigée et le clonage des promoteurs de gènes ciblés

A. PRINCIPE : mutagenèse dirigée par délétion et construction des promoteurs

La mutagenèse dirigée est une technique utilisée en génie génétique pour cibler directement un gène au niveau du génome de l'organisme étudié et générer sa mutation. Cette technique permet d'éliminer un gène par délétion propre afin d'étudier le fonctionnement du génome et voir quels rôles occupent le gène dans le processus métabolique, d'observer le phénotype révélé sur l'organisme en réponse aux changements externes. Le principe de la technique est basé sur deux étapes : 1- l'amplification des séquences bordantes de la séquence génique ciblée avec des amorces modifiées, c'est-à-dire la séquence 5' et la séquence 3' du gène cible à muter. La réalisation de cette amplification nécessite trois PCR successives. La première et la deuxième PCR vont donner les deux séquences 5' et 3' connues et d'une taille bien définie. Qui par la suite vont être appelées séquence AB (pour la 5') et séquence CD (pour la 3') « la nomenclature était basée sur les noms des amorces dessinées par l'utilisation du programme Bio-Edit ». La troisième et la dernière PCR va amplifier la séquence finale (non-fonctionnelle). En utilisant comme matrice les

deux séquences produites par la PCR N°1 et la PCR N°2. L'amplification des deux séquences AB et CD va générer la séquence finale nommée séquence AD. Le nom indique que les séquences ont été rassemblées dans une seule séquence. 2- la deuxième étape se base sur deux recombinaisons génétiques entre les deux séquences AB (5') et CD (3') de la séquence finale AD produite par PCR et les deux séquences bordantes de la séquence du gène ciblé. Deux recombinaisons sont nécessaires, la première intègre la séquence AB (non fonctionnelle) portée par le plasmide dans le génome de la cellule et la deuxième recombinaison par la séquence CD élimine complètement la séquence du gène pour la remplacer au final par la séquence AD. Pour la création des promoteurs on a choisi d'utiliser le produit d'amplification de la partie 5' du gène d'intérêt qui contient le promoteur du gène pour analyser l'expression du gène étudié. Cette séquence 5' va être insérée dans un vecteur avec un gène rapporteur de fluorescence (*gfp gène*) puis cloner dans la bactérie *LMR001* sauvage et la bactérie mutante de la même espèce.

B. METHODOLOGIE : Mutagenèse dirigée par PCR et construction des promoteurs

B.1. Dessin d'amorce et architecture moléculaire pour l'amplification et l'obtention des parties 5' et 3' du gène d'intérêt à muter par PCR successives

Les 4 amorces spécifiques nommées A, B, C et D utilisées pour l'amplification des deux régions flanquantes du gène d'intérêt, les deux régions vont être appelées par H1 pour la séquence en amont du gène (5') amplifiée avec les deux amorces A et B et les H2 pour la séquence en aval (3'), la taille des séquences H1 et H2 est proche de ~500 pb. Donc la taille de la séquence finale AB (H1+H2) va être ~1000 pb (1Kb). Ces amorces présentent des modifications telles que décrites ci-dessous : Afin de faciliter le clonage dans le vecteur qui sera utilisé par la suite, les amorces externes A et D contiennent en 5' un site de restriction reconnaissant 6 bases et qui doit être présent dans le site de clonage du vecteur qui sera ultérieurement utilisé pour le clonage. Le site de restriction utilisé ne doit pas être présent dans la séquence finale qui sera sous-clonée dans un second vecteur suicide ; en effet, si présente la séquence désirée ne pourra être clonée dans le vecteur. Ces sites sont rajoutés en 5' de séquence des amorces A et D. La séquence de l'amorce B doit contenir en 5' une partie complémentaire de 15 bases de la partie 5' de la séquence de l'amorce C. Ainsi, ces modifications permettront de lier les deux régions 5' et 3' bordantes du gène à muter et de créer une seule séquence. Donc, à partir de la séquence génomique du gène d'intérêt et des régions 5' et 3' adjacentes (400bp de chaque côté) que l'on récupère sur le site du génoscope où est stocké le génome de la souche *Ensifer aridi LMR001*, on recherche des amorces A, B, C et D utilisées pour le clonage. On doit aussi générer la séquence contenant les trois régions (ex : Contig

metK-flank). Ensuite, Les amorces sont recherchées avec l'aide du programme « BioEdit », en copiant les 4 séquences (*iolC*-5', *iolc*, *iolc*-3' et le contig) en manipulant les séquences (reverse complément pour les amorces B et D). La séquence 5' amplifiée va être aussi utilisée pour générer un promoteur afin d'étudier l'expression des gènes d'intérêt identifiés durant l'étude.

B.2. Création d'un antibiogramme d'*Ensifer aridi*

La résistance de la souche étudiée *Ensifer aridi* est testé vis-à-vis de divers antibiotiques pour déterminer la concentration minimale inhibitrice (CMI) de la croissance bactérienne à la présence de chaque antibiotique. Afin de choisir les deux antibiotiques à utiliser dans la sélection des mutants positifs. Les deux antibiotiques sélectionnés doivent accomplir un rôle complémentaire, ne doivent pas affecter la croissance de la souche *Ensifer aridi* mutée. Dont un doit être porté par le vecteur suicide utilisé pour produire les deux recombinaisons. Au début les cellules bactériennes sauvages doivent être résistantes à l'un des deux antibiotiques et sensible au deuxième porté par le vecteur suicide. La souche e-coli porteuse du vecteur et qui va transférer le vecteur suicide vers la souche à muter doit être sensible à l'antibiotique auquel *l'Ensifer aridi* est résistante. La bactérie e-coli utilisée pour le transfert du vecteur est la S17-1. Pour cela, un antibiogramme a été construit dans un premier temps en utilisant des disques d'antibiotiques, pour tester une large gamme d'antibiotique. Cette méthode consiste à construire un milieu TY Top-Agar (le fond du milieu est solide et la surface est formé par un mélange de milieu semi-solide à une concentration d'agar de 0,7% et la suspension bactérienne de la souche à une DO $\approx$ 0,7). Puis après séchage de l'ensemble, on dépose sur le milieu des disques d'antibiotique. Les antibiotiques utilisés pour ce test sont : Ampicilline 10 μ g, Rifampicine 30UI, Kanamycine 30UI, Spectinomycine 10UI, Cefotaxime 30 μ g, Tétracycline 30 UI, Nalidixic acid 30 μ g, chloramphenicol 30 μ g, Gentamycine 20 μ g et polymexine 300UI). Ensuite, un autre antibiogramme est réalisé sur milieu solide TY additionné avec une concentration connue d'un antibiotique donné. Le deuxième antibiogramme est effectué avec des dilutions par strie des colonies pures sur un milieu solide additionné à des concentrations croissantes d'antibiotique choisit. Les antibiotiques choisis sont kanamycine, Gentamycine et Tétracycline. Cette étape est importante pour tester la viabilité de la souche en réponse à des différentes concentrations d'antibiotique, afin de choisir la concentration exacte de sélection.

B.3. Extraction et purification des acides nucléiques

B.3.1. Extraction de l'ADN chromosomique

L'ADN bactérien chromosomique est extrait par un tampon Tris-EDTA (TE) (40mM de Tris-Acetate pH 7,8, 20mM Sodium acetate, EDTA 1 mM et 1%SDS) additionné à 20 μ g/ml d'ARNase.

Lancer une culture d'une colonie pure dans 5 à 10 ml de milieu TY liquide et incuber jusqu'à obtention de la phase stationnaire. La suspension est centrifugée à une température ambiante à la vitesse de 12000 à 13000 tr/min pendant 3 minutes. Éliminer le surnageant et ajouter sur le culot 300µl de tampon de lyse. Cette étape permettra la lyse de la cellule et libérera les acides nucléiques. Incuber la solution au bain marie à 37°C pendant une heure et agiter manuellement de temps à autre. Puis ajouter 100µl de NaCl à 5M et agiter manuellement, puis centrifuger pour 15 min à une vitesse de 12000 tr/min sous une température de 4°C. Transférer le surnageant dans un nouveau tube eppendorf et ajouter le même volume en chloroforme : isoamyl alcool (24 : 1). Le chloroforme est une solution organique qui va permettre de séparer les composés hydrophiles (dont l'ADN) des composés hydrophobes (protéines composés membranaires dont les acides gras). Mélanger fortement le tube (50 fois) par inversion ou vortexer jusqu'à l'obtention d'une solution à couleur blanche. Puis centrifuger pendant 15 minutes à haute vitesse 12000 tr/min et transférer en pipetant la phase aqueuse (supérieure) doucement dans un nouveau tube Eppendorf (éviter de pipetter l'interphase et la phase inférieure organique qui contient les protéines). Ajouter 2 volumes (soit 1200µl) d'éthanol absolu et inverser plusieurs fois et centrifuger pendant 10 minutes à vitesse maximale afin de faire culoter l'ADN. Éliminer le surnageant et rincer deux fois le culot avec 200µl d'éthanol à 70%. Centrifuger le mélange de la même manière, éliminer le surnageant et sécher à 37°C pour éliminer les traces d'éthanol. Finalement ajouter au culot d'ADN bien sec 50µl d'eau distillée stérile. Une fois l'ADN est extrait, lavé et dissout un dosage est demandé pour estimer la concentration en acides nucléiques et pour déterminer le niveau de la pureté de l'ADN. Afin de définir la qualité et la quantité de l'ADN, la mesure est effectuée avec le spectrophotomètre UV (le nanodrope), en déposant un volume de 1 à 2 µl du produit de l'extraction dans une cellule de l'appareil. La lecture de la DO se fait à 260 nm pour l'ADN et à 220 et 320 pour estimer la pureté. Puis le rapport 260/280 indique la contamination par les protéines (doit être compris entre 1,7 et 1,9) et le rapport 260/230 indique la contamination par les autres constituants (doit être ≤ 1).

B.3.2. Préparation d'ADN plasmidique

L'extraction d'ADN plasmidique à partir d'une culture liquide *d'e-coli* faite avec le kit (The Wizard® Plus SV Minipreps DNA Purification System) est une méthode simple et rapide. Consiste à préparer dans un premier temps, une culture de la souche de *E. coli* contenant le plasmide vide est préparée dans 10ml de milieu LB (10g/l tryptone, 5g /l extrait de levure et 10g/l NaCl) avec un antibiotique de sélection (la gentamycine à 10 µg/ml de milieu) et incubé toute la nuit jusqu'à la phase stationnaire. Puis la culture bactérienne est centrifugée pendant 5 min à 10000 tr/min et le culot est bien séché et récupéré pour ajouter 250µl de CLA (solution de lyse alcaline).

Par la suite la solution est bien mixée avec la suspension bactérienne juste en mélangeant 4 fois à la main ou avec la pipette. Sur l'étape qui suit ajouter 10µl d'alcaline protéase puis mixer à la main et incuber à température ambiante pendant 5min (ne pas dépasser le temps indiqué car autrement de l'ADN génomique sera aussi présent dans la solution aqueuse. Ajouter 350µl de la solution neutralisante NSB et mixer le mélange à la main et centrifuger le tout pendant 10min à une vitesse de 14000 tr/min sous une température ambiante. Transférer ensuite le surnageant (environ 850µl) dans une colonne. La colonne va permettre de lier les acides nucléiques et ainsi les purifier en les séparant des autres constituants puis centrifuger le tout à 14000 tr/min pendant 1 min. Retirer la colonne et jeter le liquide qui se trouve dans le tube, après il faut remettre la colonne au-dessus du tube et rajouter 750µl de solution de lavage contenant de l'éthanol (CWA column wash solution) et centrifuger à une température ambiante (14000 tr/min à 1min). Répéter les deux dernières étapes mais en ajoutant cette fois 250µl de CAW puis centrifuger pendant 2min à 14000 tr/min. A l'étape finale transférer la colonne dans un nouveau tube de 1,5 ml stérile et ajouter 50µl de la nucléase free water déjà chauffé à 55°C.

Laisser reposer la solution pendant une minute à température ambiante puis centrifuger à 14000 tr/min pendant une minute et mesurer la densité optique au spectrophotomètre (NanoDrop 2000) puis conserver le tube contenant l'extrait de plasmide à -20°C. Une fois le plasmide pur obtenu, nous devons procéder à l'étape de clonage.

B.4. Caractérisation moléculaire

L'identification moléculaire est nécessaire pour vérifier le matériel biologique utilisé, elle a été réalisée au Laboratoire de Biologie Moléculaire et Microbiologie (LMBM), à l'Université Mohammed V de Rabat, faculté des sciences. Elle se base sur une détermination génétique de l'espèce par l'amplification et le séquençage d'une séquence d'ADN qui code pour l'ARN ribosomique 16S, (l'ARNr 16S) est la séquence la plus conservée au sein de l'ADN de toutes les espèces (à un pourcentage supérieur de 90%). Ce pourcentage d'apparenté donne le degré de similitude entre deux organismes étudiés et indique directement le lien de parenté relative existant. Ainsi, cette méthode consiste premièrement à amplifier le gène avec des amorces spécifiques, et par la suite à vérifier la qualité et la taille de la séquence amplifiée sur un gel de migration. Après la réalisation du séquençage, la séquence est annotée et corrigée avec le programme BioEdit. Puis comparée avec d'autre séquence connue existantes dans une base de données numériques (des banques internationales telle : NCBI, National Center Biology Information -USA).

B.4.1. Amplification par PCR

La PCR va nous permettre d'augmenter la concentration de la séquence génique d'ARNr 16S de l'ADN cible par une amplification spécifique. Le principe de la PCR consiste en une répétition de trois cycles (la dénaturation des deux brins d'ADN matrice, hybridation des deux amorces en amont et en aval de la séquence ciblée et le troisième cycle consiste sur l'élongation du brin complémentaire par l'activité d'ADN polymérase en présence de quatre dNTP). Ainsi, la PCR est réalisée dans un thermocycleur, après addition des réactifs pour un volume de 25 µl.

Préparation du mélange :

- Tampon concentré 5 X (additionné avec MgCl₂ et dNTP (2,5 mM)) : 5 µl,
- Taq-polymérase (5U) : 0,125 µl,
- ADN de l'échantillon 1 µl,
- Amorces spécifiques du gène à amplifier (10 µM) : 2,5 µl (de chaque amorce),
- Compléter avec H₂O (eau ultra pure stérile) jusqu'au volume final de 25 µl par échantillons.

Les amorces varient selon le gène d'intérêt à amplifier. Dans ce cas les deux amorces spécifiques utilisées durant cette PCR sont 41F (5'-GCTCAGATTGAACGCTGGCG-3') et 1488R (5'-CGGTTACCTTGTTACGACTTCACC-3') (InvitroGen).

Le programme de la PCR est le suivant : le premier cycle de la dénaturation (5min à 95°C), puis 35 cycles qui vont succéder et comprennent chacun trois étapes de variation de température (dénaturation de 30 s à 94°C, hybridation de 1mn à 64°C et élongation de 1mn 30 à 72°C). Avant la fin, une étape d'élongation en un seul cycle de 10 min à 72°C est nécessaire pour achever la synthèse des séquences en cours et terminer le programme de la PCR en réglant le thermocycleur à une température de conservation de 16°C.

Pour vérifier la PCR, 5 µl de chaque produit d'amplification est déposé sur gel d'agarose à 1% (p/v) et 20 µl restante est laissée pour l'envoi au séquençage, si la bande visualisée après migration est de la bonne taille et en bonne concentration. Le Gel d'agarose utilisé contient quelques gouttes de la solution de bromure d'éthidium (à 0,5 mg/l). La migration se fait pendant 45 mn à 100V dans du TAE (Tris acétate EDTA : 2M Tris, 1M acide acétique glacial, 0,5M à PH 8 d'EDTA). Le gel est visualisé sous lumière UV et photographié avec un système de capture digital.

B.4.2. Séquençage et Analyse des séquences

Le volume restant (20µl) du produit de l'amplification du gène de l'ARNr 16S est envoyé à la société Genoscreen pour séquençage. A la réception du résultat des séquençage, les séquences partielles de l'ARNr 16S des souches obtenues sont corrigé avec le programme BioEdit et comparées avec les séquences des ARNr 16S disponibles sur la banque génomique excitante sur (<http://blast.ncbi.nlm.nih.gov>) par Blast. Le programme blast sur le site NCBI compare et ressorte sur les premières lignes les souches qui portent la séquence génique d'ARNr la plus proche de la séquence traitée dans la souche à vérifier dans l'étude. Par cette technique la souche est identifiée.

B.5. Création de l'amplicon de la séquence H1-H2 obtenus par PCR

L'ADN chromosomique utilisé durant les amplifications par PCR est dilué à l'ordre de 1/100 dans l'eau pure stérile. Pour la création des amplicons des fragments on a produit 3 PCR organisées en deux étapes : durant la première étape deux mixtes PCR sont préparées séparément pour obtenir la région 5'(H1) et la région 3' (H2) en utilisant les amorces A-B et D-C et la deuxième étape est basée sur une seule PCR réalisée avec les amorces A-C afin d'amplifier le fragment H1+H2 destiné pour la génération de la mutation par délétion.

Ces PCR sont réalisées avec un mix de Taq particulier, le Go Taq long PCR master mix, ce mix à pour avantage de ne pas ajouter d'adénine en fin de polymérisation. L'absence d'adénine en fin de polymérisation est due à la présence d'une ADN polymérase qui possède une activité exonucléase 3' 5' dans le Go Taq long PCR mix. Cette caractéristique est importante car les deux fragments générés par les deux premières PCR seront par la suite hybridés ensembles durant la troisième PCR.

Pour un volume réactionnel de 25µl par PCR, on prépare :

- ✓ 5X buffer...5µl. (contient 7,5 mM de MgCl₂)
- ✓ dNTP (2,5 mM).....2,5µl.
- ✓ GoTaq (5U/µl)...0,125µl.
- ✓ DNA.....2µl.
- ✓ Primer forward (10µM)....2, 5µl.
- ✓ Primer reverse (10µM).....2,5µl.
- ✓ H₂O diluée.....12,375µl.

Lancer l'amplification PCR dans le thermocycleur selon le programme suivant :

95°C -2min

5x (95°C-30sec/56°C-30sec/ 72°C-1min).

25x (95°C- 30 sec/ 56°C - 30sec / 72°C sec).

72°C- 5min / puis 4°C.

Lancer une migration sur gel à 2% d'agar et déposer tout le produit PCR, afin de vérifier que les fragments amplifiés H1 (A-B) et H2 (D-C) ont bien les tailles attendues pour les extraire et les purifier du gel et utiliser le mélange des deux produits finaux comme matrice pour la dernière PCR. La troisième amplification va être faite par le couple d'amorces A et D. Cette étape devrait permettre cette fois et grâce aux séquences complémentaires présentes dans les extrémités 3' de AB et de 5' de CD de générer un seul amplifiât correspondant au fragment désiré de AB lié à CD soit le fragment AB : CD. On va utiliser un volume de 25µl avec le mélange réactionnel qui comprend :

- ✓ 5X buffer 5µl. (contient 7,5 mM de MgCl₂)
- ✓ dNTP (2,5 mM) 2,5µl.
- ✓ Matrice 2µl. (A-B).
..... 2µl. (D-C).
- ✓ Amorce A (10µM).....2,5µl.
- ✓ Amorce D (10µM) 2,5µl.
- ✓ Go Taq (5U).....0,125µl.
- ✓ H₂O.....7,125µl.

On utilise le programme de PCR suivant :

95°C -5min

5x [95°C- 30sec/ 50°C- 30 sec/72°C- 50sec]

25x [95°C- 30sec/ 50°C- 30 sec/72°C- 50sec]

72°C- 5min / puis 4°C.

A la fin de la PCR, préparer un gel à 2% d'agarose dans le tampon TAE_x1. Déposer 10µl du produit PCR et vérifier que la taille de la bande AD à l'aide du marqueur de taille 1Kb correspond bien à 1000pb. Si la taille est bonne, on charge sur un nouveau gel d'agarose tout le reste de produit PCR à 75Volts. A la fin de l'électrophorèse on procède à l'extraction de la bande correspondant au produit AB : CD à partir du gel.

B.5.1. Extraction et purification des fragments AB : CD à partir du gel d'agarose

L'extraction de l'ADN amplifié après PCR se fait avec un kit d'extraction de QIAGEN (**QIAquick® Gel Extraction kit**) et on procède comme suit : découper à l'aide d'un scalpel propre la bande sous UV et transférer la bande dans un tube stérile qui a été taré puis peser le morceau

d'agarose contenant l'ADN et ajouter 3 volumes en μl de tampon (solution QG) pour chaque mg de gel ($1\text{mg} \approx 1\mu\text{l}$). Incuber le tout à 50°C pendant 10 min bien vortexer de temps à autre jusqu'à que le gel soit bien dissous, après cette étape, ajouter dans le tube un volume d'isopropanol puis transférer le mélange sur une colonne contenant une membrane permettant le piégeage des acides nucléiques. Centrifuger à 14000 tr/min pour une minute et jeter le liquide. Ajouter par la suite sur la colonne 750 μl de solution de lavage (Tampon PE) à laquelle de l'éthanol absolu a été rajouté et centrifuger (14000 tr/min, 1 min) puis jeter le liquide. Jeter le liquide et centrifuger une nouvelle fois pour éliminer toute trace de solution de lavage (contenant de l'éthanol) pendant 1min à la même vitesse puis transférer la colonne sur un nouveau tube eppendorf stérile de 1,5ml. Finalement, éluer l'ADN avec 50 μl de Tampon EB (10mM Tris.Cl, pH 8,5) ou avec l'eau (nucléase free water) et centrifuger (14000tr/min, 1min) après avoir incubé pendant 1min jusqu'au 4 minutes pour augmenter la concentration des éluas d'ADN. Noter que si les bandes sont trop faibles après visualisation sur gel, une dernière amplification peut être effectuée afin d'augmenter la concentration du produit AB::CD final. Dans ce cas, refaire le même mix PCR que précédemment mais utiliser comme matrice l'éluât

B.6. Manipulation des fragments amplifiés

B.6.1. Digestion et la ligation de la séquence H1+H2 (AB ::CD) dans pGEMT-Easy et le plasmide pJQ200KS+

Les séquences H1+H2 (AB :: CD) amplifiées vont être clonées dans le plasmide **pGEMT-Easy** puis dans le plasmide suicide *pJQ200KS+*. Dans un premier temps on réalise une extraction du plasmide *pGEMT-Easy* et du plasmide *pJQ200KS+* et on vérifie leurs tailles par une migration lente sur un gel d'agarose de 0,8%, puis on fait une digestion des deux plasmides et de la séquence AB:: CD. La digestion se fait en utilisant les composants suivants : Préparer un tube eppendorf de 1,5ml en mettant dans l'ordre indiqué :

- 1 μg de plasmide
- 2 μl tampon x10.
- 1 μl d'enzyme *XbaI* (10U).
- 0,2 μl BSA (10 $\mu\text{g}/\mu\text{l}$).

Compléter avec de l'eau distillée stérile jusqu'à un volume final de 20 μl .

Mettre le tube à 37°C (dans le bain-marie préchauffer à cette température) pendant 3 heures.

Suite à la digestion, on réalise ensuite une étape de déphosphorylation du vecteur. Afin de réaliser cette étape de déphosphorylation, on ajoute au mélange réactionnel du plasmide digéré 2µl d'alkaline phosphatase. On vortexe le mélange et on met le tout en incubation à 37°C pour une heure. Après l'achèvement de la durée d'incubation le mélange réactionnel est placé dans le bain sec à une température de 56°C pendant 15min. Puis on ajoute 1µl d'alkaline protéase, on incube le mélange à 37°C pendant 30 minutes et à 75°C pour 10min. Ensuite, on fait migrer le vecteur digéré et déphosphorylé pendant 45 min dans un gel d'agarose (0.8%) afin d'isoler la bande qui correspond à la bonne taille. La purification de la bande d'intérêt est faite à partir du gel comme décrit précédemment.

Une fois le vecteur et l'insert purifiés, on procède à leur quantification puis on prépare la réaction de ligation comme suit sachant que nous voulons utiliser au maximum 200 ng d'ADN et que nous voulons un ratio molaire Vecteur / Insert de 1/3.

La ligase qui est l'enzyme utilisée pour la ligation coûte très chère, nous utilisons un petit volume 10 µl de réaction finale avec un maximum de ligase à 0.7 µl.

Par exemple si nous avons une concentration en plasmide de 50 ng/ul et que notre insert à une concentration de 10 ng/ul, et que notre vecteur à une taille de 5.2 kb et que notre insert à une taille de 400 bp ; il faudra préparer la réaction suivante

Calculs pour obtenir le bon ratio :

$(\text{ng vecteur} \times \text{taille insert (kb)} / \text{taille du vecteur (kb)} \text{ vecteur}) \times \text{ratio of } 3/1 = \text{ng insert}$

Si on prend 100 ng de vecteur, la quantité en (ng) d'insert sera de :

$(100\text{ng} \times 400 \text{ pb} / 5200\text{pb}) \times \text{ratio } (3 / 1) = X \text{ ng d'insert avec } X = 23.07$

- Si on utilise 100ng de vecteur il faudra 23 ng d'insert pour un ratio molaire (Vecteur/ insert) de 1/3.

Ce qui correspond à un volume de 2µl de plasmide et 2.3µl d'insert.

Dans ces conditions, nous devons mettre dans la réaction de ligation les composés suivants :

- Ligase buffer (10X): 1 µl
- Ligase: 0.7 µl
- Vecteur : 2 µl
- Insert : 2.3 µl
- BSA : 0,3µl

Compléter avec de l'eau distillée stérile jusqu'à un volume final de 10µl.

La réaction de ligation est finalement incubée à 16°C sur la nuit dans un bain marie réfrigéré.
Le lendemain, nous procédons à la transformation.

B.6.2. Digestion et la ligation de la séquence H1 (AB promoteur) dans le plasmide TOPO vecteur et le plasmide pPROBE NT

Les séquences H1 (AB) amplifiées vont être clonées dans le plasmide TOPO® TA Cloning® Kit for Sequencing vector, puis dans le plasmide *pPROBE-NT-gfp*. On réalise une extraction du plasmide *pPROBE-NT-gfp* puis on fait une digestion du plasmide extrait et de la séquence AB. On vérifie la taille des séquences par une migration lente sur un gel d'agarose de 0,8%.

La digestion se fait en utilisant les composants suivants : Préparer un tube eppendorf de 1,5ml en mettant dans l'ordre indiqué :

- ✓ 1µg d'insert.
- ✓ 2µl de Tampon X10
- ✓ 1µl d'enzyme *EcoRI* (10U)
- ✓ 0,2 µl BSA (10µg/µl)

Compléter avec de l'eau distillée stérile jusqu'à un volume final de 20µl.

Mettre le tube à 37°C (dans le bain-marie préchauffer à cette température) pendant 3 heures et vérifier la concentration de la séquence.

Dans un premier temps, ajouter les éléments suivants dans le tube de 1,5ml dans l'ordre suivant : 0.5 à 2 µl du produit de PCR frais, ajouter 0.5 µl du tampon NaCl 1.2M + 0.06M MgCl₂ (kit: Salt solution) puis 1.5 µl d'eau à partir de kit avec 0.5 µl *TOPO vector*. Homogénéiser bien et doucement les solutions par pipetage et incubé à 22-23 °C pour 5 – 30 minutes (noter bien que 15 minutes ça suffit pour augmenter le rendement). Placer les tubes dans la glace et puis lancer la transformation des bactéries E-coli DH5α par le produit TOPO:: insert-EcoRI. (Les *DH5alpha* sont chimio-compétentes).

Ensuite, les séquences *TOPO:: insert-EcoRI* sont extraites des clones *DH5 alpha* et digérées. Pour isoler la séquence insert-EcoRI et produire une ligation avec le plasmide *pPROBE-NT-gfp*. Déphosphoryler le plasmide *pPROBE-NT* digéré et réaliser une ligation comme indiqué précédemment.

B.7. Transformation bactérienne

B.7.1. Escherichia coli

Le clonage moléculaire est une des bases du génie génétique. Il consiste à insérer un fragment d'ADN (dénommé insert) dans un vecteur approprié comme un plasmide. Le nouveau plasmide ainsi créé sera ensuite introduit dans une cellule hôte, en générale la bactérie *Escherichia coli* chimio-compétente. Celle-ci sera alors sélectionnée et multipliée afin d'obtenir une grande quantité du plasmide d'intérêt.

Durant nos expériences nous avons utilisé trois différentes souches d'*Escherichia coli* (*DH5 α* et *XL2- Blue* Ultracompetent cells chimio-compétentes et la S17-1 électro-compétente)

Il faut avoir au préalable préparé des cellules d'*E.coli DH5 α* compétentes et la S17 électro-compétentes. La *XL2- Blue* Ultracompetent cells est utilisée directement à partir du kit selon un protocole prescrit.

La souche d'*Escherichia coli DH5 α* a été développée pour une utilisation de clonage au laboratoire. C'est une souche classiquement utilisée pour la transformation car elle comporte plusieurs mutations. Ces mutations sont : La *lacZ Delta M15* : Permet de dépistage bleu-blanc pour les cellules recombinantes. La *recA1* mutation : réduit la recombinaison homologue pour une insertion plus stable. L'*endA1* mutation : permet la dégradation de l'endonucléase qui assure des taux de transfert plus élevés de plasmide.

J'ai pu tester différentes méthodes de transformation bactérienne, la transformation par électrochoc, la transformation par choc thermique des cellules chimio compétentes et la transformation par conjugaison bi-parentale.

Voici le protocole de préparation des cellules chimio-compétentes d'*E.coli DH5 α* qui ont été utilisées pour les clonages des constructions *plasmide-insert*.

Tout d'abord, il faut repiquer à partir d'un tube de *DH5 α* stocké à -80°C sur LB solide pour l'incuber à 37°C. A l'apparition des colonies lancer une préculture bactérienne sur milieu LB liquide. Après une croissance marquée par un trouble bactérien mettre 2 ml de la préculture dans 50ml de LB liquide et lancer par la suite une incubation à 37°C sous agitation forte.

Après 2 heures d'incubation mesurer la densité optique qui doit être entre 0,2 et 0,3. Ensuite, incuber dans la glace pendant 15 minutes la culture, une solution de CaCl₂ (0,1M) stérile et aussi le mélange CaCl₂ (0,1M) complémenté avec du glycérol à 15%. A la fin d'incubation glaciale, centrifuger la culture à une vitesse de 8000 rpm, dans une température de 4°C, pendant 10 minutes. Par la suite éliminer le surnageant et ajouter au culot 25 ml de CaCl₂ (0,1M) glacé et incuber le tube pendant 30 minutes dans la glace, centrifugé à nouveau dans les mêmes conditions que la fois

précédente. Jeter le surnageant et ajouter 6ml du mélange CaCl_2 (0,1M) avec glycérol 15%. Finalement, on aliquote 0,5 ml de solution de cellules compétentes dans des tubes eppendorf de 1,5 ml stériles puis les conservés à -80°C .

Remarque :

- ✓ Les ions Ca^{++} apportés par la solution de transformation CaCl_2 neutralisent les charges négatives répulsives de la membrane bactérienne et de la molécule d'ADN du plasmide, favorisant ainsi son adhésion.
- ✓ L'incubation initiale dans la glace stabilise les mouvements moléculaires dans la membrane bactérienne.
- ✓ Le choc thermique (passage de 0°C à 42°C) accroît la perméabilité de la membrane bactérienne ce qui permet le passage des molécules d'ADN du milieu extracellulaire au milieu intracellulaire.

Ces étapes favorisent la pénétration du plasmide dans la cellule.

Le protocole de préparation des cellules électro-compétente d'*E. coli* S17-1 utilisées pour clonage des constructions *pPROBE-NT-gfp*-insert AB (H1) et aussi des constructions *pJQ200KS*+AB : CD (H1-H2) se résume en plusieurs étapes. Premièrement, il faut commencer par un repiquage de la souche conservées à -80°C sur du LB solide puis préparer à partir d'un clone pur une pré-culture dans 10 ml de LB sans antibiotiques. Le lendemain ajouter 2,5 ml de la pré-culture sur 250 ml de LB et laisser incuber jusqu'à une DO de 0,5 ou 0,6. Ensuite, incuber la culture bactérienne, l'eau stérile et le glycérol 10% dans la glace pour 15 minutes (à partir de cette étape les échantillons doivent être manipulés au froid). Après refroidissement, centrifuger la culture à (4000 tr/min pour 15 min/ 4°C) puis éliminer le surnageant et laver les cellules en ajoutant de l'eau stérile. Ajouter sur le culot le volume équivalent au milieu éliminé de l'eau stérile puis agiter avec vortex et centrifuger (même condition de centrifugation précédente). Jeter le surnageant et ajouter 50 ml de glycérol 10% (50 ml pour les 250 ml préparés au début) et Centrifuger/15min/ 4°C . Ensuite, ajouter 500 μl ou 300 μl de glycérol de 10% sur le culot (le volume ajouté dépend de la quantité du culot) et distribuer 50 μl de suspension bactérienne par tube pour les stocker à -80°C . Lors de la transformation S17-1, il faut utiliser les cuvettes à diamètre de 0.2 mm refroidis et mettre dans la cuvette 50 μl de la souche S17 puis ajouter 2 ou 3 à 4 μl de la construction à cloner (plasmide+insert). L'électroporation est effectuée selon les paramètres suivants : 15 KVolts, 700ms (haut niveau) 200ms (bas niveau). Resuspendre les bactéries rapidement dans 500 μl de milieu LB liquide et mélanger avec pipetage. Par la suite la suspension bactérienne transformée est transférée sur un tube pour une incubation à 37°C pendant 1 heure sous agitation. A la fin de

l'incubation 10 et 100µl de la suspension bactérienne sont étalées sur deux boîtes LB-agar contenant l'antibiotique adéquat au plasmide utilisé. Incuber à 37°C jusqu'à apparition de colonie.

B.7.2. *Ensifer aridi* LMR001

Nous avons réalisé une transformation par une conjugaison bactérienne biparentale entre la *S17-1* et l'*Ensifer aridi* LMR001 pour avoir le transfert de l'information génétique sous forme de plasmide modifié d'une bactérie donneuse vers la bactérie receveuse. En commençant, Etaler la bactérie *Ensifer aridi* LMR001 résistante à la Rifampicine sur milieu TY solide sélectif par l'addition d'antibiotique approprié et incuber la culture à une température de 28°C jusqu'à l'apparition des colonies (à ne pas dépasser les 3jours). Le deuxième jour, Etaler la bactérie *S17-1* qui contient le plasmide que l'on veut transférer sur milieu LB Solide avec l'antibiotique de sélection, puis on incube la culture à 37°C pendant une nuit. Lancer une préculture à partir des colonies bien pure dans du TY liquide pour la LMR001 et LB liquide pour la *S17-1* additionnées avec les antibiotiques adéquats pendant une nuit. Le lendemain, centrifuger les précultures à 4000 tr/min pour 15 minutes puis accéder à un lavage cellulaire avec de l'eau distillées stériles (deux lavages sont nécessaires pour éliminer les traces des antibiotiques ajoutés et après chaque lavage centrifuger à 4000 tr/min pour 15 minutes). Eliminer les traces de l'eau et ajouter du TY-liquide sans antibiotique puis mélanger 5 volumes de bactéries receveuse avec un volume de bactérie dans un tube de 2 ml. Centrifuger la suspension et éliminer les surnageant pour laisser un volume de 50µl qui va être déposé sous forme d'un goutte au centre de la surface du milieu TY-solide. Incuber la boîte à 32°C pour 3 jours. Le quatrième jour, on prend toute la croissance bactérienne dans 2ml de TY puis on étale 50 µl de trois dilutions D0, D3 et D 5 sur TY additionné avec les antibiotiques choisis. Les clones apparus sur les boites sont sélectionnés par PCR spécifiques en utilisant les amorces spécifiques A et D utilisés précédemment pour la création des fragments H1 et H2 afin de sélectionner les souches mutées, aussi l'amorce GFP et l'amorce spécifique B pour la sélection des souches qui ont reçues le plasmide avec le promoteur.

B.8. Mutagénèse par double recombinaison homologue

B.8.1. Principe de la sélection

Cette approche se base sur un mécanisme moléculaire nommé la double recombinaison qui permet d'intégrer des fragments d'ADN provenant d'un plasmide dans le chromosome d'une bactérie par recombinaison au niveau de séquences homologues entre le chromosome et le plasmide. La procédure de sélection se fait en deux étapes. Dans une première étape, l'ensemble du plasmide est intégré dans le chromosome. Lors de cette étape, une première recombinaison homologue

intervient entre deux séquences homologues portées sur le chromosome d'une part et sur le plasmide d'autre part. Dans le cas de la mutagenèse dirigée et comme indiqué plus haut, nous avons cloné dans un vecteur suicide, deux fragments d'ADN correspondants aux régions 5' et 3' du gène à muter (régions H1 et H2). Afin de transférer le plasmide *pJQ200SK+*, nous avons utilisé la méthode de la conjugaison bactérienne biparentale utilisant la souche *LMR001* résistante à la rifampicine. À la suite de la première recombinaison qui a lieu (événement rare) entre les deux premières régions homologues, il y a insertion de la totalité du plasmide dans le chromosome. Comme le plasmide ne peut se répliquer dans la souche de rhizobium, il est facile de sélectionner les cellules ayant intégré le plasmide dans le chromosome car elles auront acquis la résistance à l'antibiotique qui est porté par ce plasmide. Dans le cas du plasmide *pJQ200SK+*, il s'agit de la gentamycine. Nous avons donc sélectionné des clones de *LMR001* résistants à la rifampicine et à la gentamycine. Pour la deuxième étape, Une fois une bactérie *LMR001* tolérante à la gentamycine obtenue, il va falloir sélectionner les bactéries qui auront subi la deuxième recombinaison au niveau de la deuxième région clonée afin d'obtenir des mutants. La sélection des mutants ayant perdu le plasmide se fait en utilisant le saccharose (sucre) à une concentration de 5-7 % dans le milieu de culture. Pour cette étape, le gène *sacB* de *Bacillus subtilis* et qui est présent dans le plasmide suicide *pJQ200KS+* est utilisé comme marqueur de contre sélection.

Pour cela, on cultive dans un premier temps sur boîte contenant du TY (20µg/ml Gentamycine, 100µg/ml Rifampicine) la souche de *LMR001* contenant le plasmide intégré. Il faut effectuer ensuite une série de dilutions (50µl de la suspension bactérienne dans 450µl de milieu TY liquide pour la première dilution...). Puis étaler chaque dilution sur la surface du milieu TY solide sélectif avec sucrose 5% et 100µg/ml Rifampicine. Incuber la culture dans l'étuve à une température de 28°C jusqu'à apparition des colonies. A noter que cette deuxième recombinaison peut avoir lieu au niveau de la deuxième région mais également au niveau de la première région (H1). Dans le cas d'une deuxième recombinaison au niveau de la première région, le génotype sauvage est alors obtenu. Afin d'obtenir des mutants, il faudra donc identifier des clones ayant subits la deuxième recombinaison au niveau de la deuxième région en plus de la première recombinaison.

B.8.2. Sélection des clones des *LMR001* mutants

Après croissance de colonies, il faut dans un premier temps isoler chaque colonie et les repiquer à la fois sur TY solide non sélectif et sur TY solide avec Gentamycine de sorte à avoir une copie de la souche de chaque colonie isolée sur boîte de TY solide et de vérifier que chaque clone a bien perdu le plasmide qui était intégré dans le génome. Comme le gène *sacB* a un taux de mutation

assez élevé, cette étape nous permettra de s'assurer que la deuxième recombinaison a bien eu lieu. En effet, si la bactérie pousse sur milieu avec saccharose, en théorie, cela veut dire que *sacB* n'est plus présent dans le génome de la souche. Hors si le gène *sacB* a muté (ce qui l'a rendu non fonctionnel), on pourra le savoir car la bactérie possèdera toujours la résistance à la gentamycine qui est aussi présente dans le plasmide *pJQ200SK+* inséré. Finalement, il faudra donc sélectionner les clones qui poussent sur saccharose mais qui ne poussent pas sur gentamycine pour sélectionner que les clones mutés.

On procède ensuite à un criblage par PCR des clones ayant subi la double recombinaison aboutissant au mutant désiré et non au génotype sauvage. Pour cela on utilise des amorces externes aux fragments clonés : les amorces A et D qui doivent dans le cas des mutants amplifier des fragments de taille inférieure à l'amplification des clones sauvages.

Pour préparer la PCR des clones recombinés, on procède comme suit.

Dans un premier temps, mettre en suspension la colonie à l'aide d'un matériel fin et stérile (cure dent) dans 50µl d'eau stérile. Puis transférer le tube dans un bain marie sec préchauffer à 99°C et incubé pendant 10min puis dans la glace. Le mélange va nous servir comme matrice pour la PCR. Il suffit d'ajouter 3µl de la matrice à 27µl de mixte.

Master mixte, PCR : pour un volume final de 30µl

- ✓ 5X buffer 5µl. (contient 7,5 mM de MgCl₂)
- ✓ dNTP (2,5 mM).....3µl.
- ✓ Matrice 3µl. (bactéries lysées)
- ✓ Amorce A (10µM) ...3µl.
- ✓ Amorce D (10µM) ...3µl.
- ✓ Go-Taq (5U) 0,15µl.
- ✓ H₂O.....10,35µl.

Programme PCR :

95°C- 5min.

35x [95°C- 30sec/ 56°C- 30 sec/72°C- 50 sec]

72°C- 5min/ puis 4°C.

Finalement, comparer la taille des bandes obtenues et vérifier qu'elles correspondent aux tailles attendues.

Sélectionner un mutant, et le mettre en collection à -80°C.

Extraire l'ADN génomique comme indiqué plus haut et refaire les PCR de la région mutée et vérifier la mutation par séquençage de la région mutée.

4. Étude de l'expression de gènes cibles en utilisant le gène rapporteur *gfp*

Ce test est basé sur la mesure d'expression différentielle révélée par l'intensité de la fluorescence du gène rapporteur *gfp*. Après la réalisation du transfert de la construction *pPROBE NT-promoteur-gfp* dans *Ensifer aridi LMR001* sauvage et dans *Ensifer aridi LMR001* mutée. Nous avons comparé la réponse transcriptionnelle du gène rapporteur *gfp* entre les deux souches dans les conditions stress et les conditions de culture de non-stress. Le stress salin est produit avec l'ajout du NaCl dans le milieu TY-liquide à une concentration de 0,25 M ; la même concentration de sel utilisée durant l'analyse transcriptomique. La mesure de l'intensité de la fluorescence du gène *gfp* est réalisée par l'automate TECAN microplaque spectro-fluomètre, sous une excitation de 485 nm et une émission fixée à 535 nm. La valeur de l'expression relative des promoteurs est mesurée par une soustraction de la valeur tirée de la fluorescence qui correspond au test contrôle non-inoculé et normalisée avec la densité cellulaire de la croissance bactérienne mesurée par la TECAN machine. Pour la vérification des données d'expression des gènes ciblés dans cette étude *gfp*-rapporteur. Nous avons fait une PCR quantitative de certains gènes ciblés par la méthode *gfp* fluorescence en comparant leurs niveaux d'expression avec l'expression des gènes de ménage (voir Article 2 et 3 pour la procédure en détail).