



# Study of the evolutionary dynamics of plant beneficial bacteria

Beatriz Manriquez

## ► To cite this version:

Beatriz Manriquez. Study of the evolutionary dynamics of plant beneficial bacteria. Microbiology and Parasitology. Université de Lyon, 2021. English. NNT : 2021LYSE1007 . tel-03626849

**HAL Id: tel-03626849**

**<https://theses.hal.science/tel-03626849>**

Submitted on 31 Mar 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

N°d'ordre NNT :  
2021LYSE1007



## **THESE de DOCTORAT DE L'UNIVERSITE DE LYON**

opérée au sein de  
**l'Université Claude Bernard Lyon 1**

**Ecole Doctorale N° 341**  
**Evolution Ecosystème Microbiologie Modélisation**

**Spécialité de doctorat** : Evolution expérimentale, interactions, microorganismes  
**Discipline** : Microbiologie

Soutenue publiquement le 14/01/2021, par :

**Beatriz MANRIQUEZ**

---

# **Analyse de la dynamique évolutive d'une communauté synthétique de souches phytobénéfiques de *Pseudomonas* lors de son interaction avec le maïs**

---

Devant le jury composé de :

Guidot Alice, Directrice de recherche, INRAE Toulouse  
Rivilla Rafael, Professeur, Université Autonome de Madrid (Espagne)  
Sarniguet, Alain, Directeur de Recherche, INRAE Angers  
Fraissinet-Tachet Laurence, Professeure, université Lyon 1  
Muller Daniel, Maître de conférences, Université Lyon 1  
Prigent-Combaret Claire, Directrice de recherche, CNRS, Lyon

Rapporteur  
Rapporteur  
Rapporteur  
Examinatrice  
Co-encadrant  
Directrice



## **Université Claude Bernard – LYON 1**

Président de l'Université	M. Frédéric FLEURY
Président du Conseil Académique	M. Hamda BEN HADID
Vice-Président du Conseil d'Administration	M. Didier REVEL
Vice-Président du Conseil des Etudes et de la Vie Universitaire	M. Philippe CHEVALLIER
Vice-Président de la Commission de Recherche	M. Jean-François MORNEX
Directeur Général des Services	M. Damien VERHAEGHE

### **COMPOSANTES SANTE**

Faculté de Médecine Lyon-Est – Claude Bernard	Doyen : M. Gilles RODE
Faculté de Médecine et Maïeutique Lyon Sud Charles. Mérieux Doyenne :	Mme Carole BURILLON
UFR d'Odontologie	Doyenne : Mme Dominique SEUX
Institut des Sciences Pharmaceutiques et Biologiques Directrice :	Mme Christine VINCIGUERRA
Institut des Sciences et Techniques de la Réadaptation Directeur :	M. Xavier PERROT
Département de Formation et Centre de Recherche en Biologie Humaine	Directrice : Mme Anne-Marie SCHOTT

### **COMPOSANTES & DEPARTEMENTS DE SCIENCES & TECHNOLOGIE**

UFR Biosciences	Directrice : Mme Kathrin GIESELER
Département Génie Electrique et des Procédés	Directrice : Mme Rosaria FERRIGNO
Département Informatique	Directeur : M. Behzad SHARIAT
Département Mécanique	Directeur M. Marc BUFFAT
UFR - Faculté des Sciences	Ad. prov. : M. Bruno ANDRIOLETTI
UFR (STAPS)	Directeur : M. Yannick VANPOULLE
Observatoire de Lyon	Directrice : Mme Isabelle DANIEL
Ecole Polytechnique Universitaire Lyon 1	Directeur : Emmanuel PERRIN
Ecole Supérieure de Chimie, Physique, Electronique	Directeur : Gérard PIGNAULT
Institut Universitaire de Technologie de Lyon 1	Directeur : M. Christophe VITON
Institut de Science Financière et d'Assurances	Directeur : M. Nicolas LEBOISNE
ESPE Administrateur Provisoire :	M. Pierre CHAREYRON

---

## Analyse de la dynamique évolutive d'une communauté synthétique de souches phyto-bénéfiques de *Pseudomonas* lors de son interaction avec le maïs

---

Au niveau de leurs racines, les plantes interagissent avec un microbiote qui impacte fortement leur développement et leur adaptation à leur environnement. Au sein du microbiote rhizosphérique, certaines bactéries appartenant à des espèces différentes sont qualifiées de PGPR (*Plant Growth-Promoting Rhizobacteria*), pour leur aptitude à stimuler la croissance de la plante. En fonction de la souche PGPR, différentes fonctions phyto-bénéfiques peuvent être exprimées allant de l'amélioration de la nutrition minérale de la plante, la modulation des balances hormonales végétales, à l'inhibition de pathogènes (production d'antimicrobiens et/ou induction des réponses de défense de la plante). Bien que certaines PGPR possèdent une large gamme de ces fonctions, leur inoculation ne conduit pas toujours à des effets bénéfiques significatifs sur la plante. En effet, les mécanismes favorisant l'établissement des PGPR dans la rhizosphère et l'expression de leurs propriétés restent largement méconnus. Plusieurs travaux ont suggéré que les plantes sélectionneraient des PGPR avec un nombre élevé de fonctions phyto-bénéfiques, du fait d'effets phyto-stimulateurs potentiellement additifs. Pourtant, précédant mes travaux de thèse, une étude a démontré que dans la rhizosphère du maïs, ce sont plutôt des PGPR avec un faible nombre de fonctions phyto-bénéfiques qui sont sélectionnées. Pour mieux comprendre le fonctionnement des coopérations plantes-PGPR et confirmer quel type de PGPR sont préférentiellement sélectionnées dans la rhizosphère du maïs, nous avons mis en place une approche d'évolution expérimentale avec une communauté synthétique constituée de 10 souches de PGPR du genre *Pseudomonas*, inoculées sur le maïs, la moitié des souches possédant un nombre élevé de fonctions phyto-bénéfiques et l'autre un nombre modéré. Dans un premier temps, nous avons suivi les dynamiques de populations de la communauté synthétique lors de 27 cycles d'inoculations successives sur des plants de maïs, en conditions gnotobiotiques. En parallèle, pour confirmer que les dynamiques de population observées sont dues à la plante, nous avons réalisé une évolution expérimentale dans un milieu de culture riche, en absence de plante. Nous avons observé des dynamiques de populations rapides mais différentes pour chacune de ces évolutions expérimentales. Dans un deuxième temps, l'impact de l'évolution expérimentale *in planta* sur les génomes des souches a été évalué. Pour cela nous avons analysé les mutations des génomes des populations récupérées après le 1<sup>er</sup> et le dernier cycle de l'évolution expérimentale *in planta*. Nous avons également analysé les génomes de variants des souches majoritaires, isolés au cycle 27. Nous avons observé des évolutions convergentes dans les génomes des variants isolés des communautés synthétiques différentes. Notre étude a permis de mieux comprendre les mécanismes adaptatifs mis en place par les *Pseudomonas* lors de leur interaction conjointe avec le maïs.

---

## **Analysis of the evolutionary dynamics of a synthetic community of phytobeneficial strains of *Pseudomonas* during its interaction with maize.**

---

At the root level, plants interact with their microbiota and this has a strong impact on their development and their adaptation to their environment. Within rhizospheric microbiota, certain bacteria belonging to different species are classified as PGPR (Plant Growth-Promoting Rhizobacteria) as a result of their ability to stimulate plant growth. Depending on the PGPR strain, different plant beneficial functions can be expressed, ranging from the improvement of the plant mineral nutrition, to the modulation of plant hormonal balances, to the inhibition of pathogens (production of antimicrobials and/or induction of the plant defense responses). Although some PGPR possess a wide range of these functions, their inoculation does not always lead to significant beneficial effects to the plant. Indeed, the mechanisms favoring the establishment of PGPR in the rhizosphere and the expression of their properties remain largely unknown. Several studies have suggested that plants would select PGPR with a high number of plant beneficial functions, possibly due to additive plant-stimulating effects. However, prior to the present research project, a previous study showed that within the maize rhizosphere, it is in fact PGPR with a low number of plant beneficial functions that are selected. To better understand the functioning of plant-PGPR cooperations and to thus confirm which type of PGPR are preferentially selected in the maize rhizosphere, an experimental evolutionary approach was set-up: a synthetic community made up of 10 strains of PGPR of the *Pseudomonas* genus inoculated on maize, half of the strains having a high number of plant beneficial functions and the other half a moderate number. The population dynamics of the synthetic community were first followed during 27 successive inoculation cycles on maize plants under gnotobiotic conditions. At the same time, to confirm that the population dynamics observed are due to the plant, an experimental evolution was carried in a rich culture medium, in the absence of the plant. Rapid but differing population dynamics were observed for each of these experimental evolutions. In a second step, the impact of the experimental evolution *in planta* on the genomes of the strains was evaluated. To do this, the mutations in the genomes of the populations which were recovered following the first and last cycles of the experimental evolution *in planta* were analyzed. The genomes of variants of the dominant strains, isolated at the end of cycle 27 were analyzed as well. Convergent evolution was observed in the genomes of variants isolated from different synthetic communities. This study has led to a better understanding of the adaptive mechanisms put in place by *Pseudomonas* when they interact with maize.

---

## REMERCIEMENTS - ACKNOWLEDGEMENTS

---

Après toutes ces années à dire que j'arrêtais la thèse, nous y voilà !

First, I would like to thank Alice Guidot, Alain Sarniguet, Rafael Rivilla who agreed to evaluate this thesis work as well as Laurence Fraissinet-Tachet who will join them for the thesis defense jury.

Merci également au directeur de l'UMR CNRS 5557, Yvan Moëgne-Loccoz, pour m'avoir permis d'intégrer le laboratoire. Je voulais également remercier les membres de mon comité de pilotage, Denis Faure, Franck Bertolla et Nathalie Poussereau pour leurs conseils et leurs idées. Jordan et FloFlo, je ne vous ai pas oublié, vous arrivez plus tard.

Un grand merci à mes chefs pour ces 4 ans partagés, d'abord en M2 et puis en thèse. Merci d'avoir cru en moi et de m'avoir préparé au concours. Le chemin n'a pas toujours été simple mais je ne regrette en aucun cas ces 3 ans et demi. Merci pour votre soutien, vos connaissances, et surtout pour votre patience face à mes questions. Merci pour vos compétences respectives et vos conseils, j'ai beaucoup appris à vos côtés. Mention spéciale pour Daniel et ses blagues qui m'ont toujours fait rire (même si j'ai mis un moment à les comprendre !).

Je continue avec Ibio, un grand merci à Danis et Audrey sans vous, eh bien, il n'y aurait tout simplement pas de résultats dans cette thèse ! Danis un grand merci pour ta disponibilité ces derniers mois (et pour les analyses qui restent à faire d'ici à la soutenance). Audrey, t'es partie et t'a laissé un vide dans le bureau et dans le labo en général, merci d'avoir joué les psychologues et de m'avoir toujours poussé vers l'avant.

Merci également aux personnes du CESN, en particulier à Marjo, pour ta bienveillance et toutes ces racines lyophilisées. Merci aussi à la compta, je ne vous ai pas beaucoup vu pendant ces années mais vous avez toujours été là quand je comprenais rien aux formulaires d'envoi des échantillons. Un grand merci à Elise pour m'avoir expliqué 100 fois comment marchait un phytotron et pour tes conseils avisés.

J'aimerais aussi remercier Ludo et Quentin pour m'avoir pris en stage en M1 (et désolée si je n'ai pas été la stagiaire la plus facile !), pour votre patience et votre bienveillance. Si vous ne m'aviez pas donné cette opportunité, je pense que je ne serai pas là. Merci aussi à Ofo pour ses conseils et ses mots croisés.

Merci à Corinne pour tous les milieux autoclavés et pour tous les tubes décontaminés. Sans toi il

n'y aurait pas eu de manipes. Un énorme merci à FloFlo, t'a été d'une aide énorme au début de ma thèse, plus que de l'aide technique t'as été une aide humaine. T'as toujours su me motiver, me guider, m'accompagner et surtout me redonner le sourire. Et tu m'as aussi appris à écrire les 8 comme une grande fille (compétence que je pourrai rajouter dans mon CV). Toi aussi quand t'es partie t'a laissé un vide et je suis un peu deg de ne pas pouvoir profiter de ta compagnie maintenant que t'es revenue. Merci aussi à Jeanne qui est venue me prêter main forte pour la deuxième partie de cette thèse, merci pour ce nombre de clones incalculables qu'on a trié. Même en étant partagée entre 2 équipes t'as toujours su trouver du temps pour m'aider. J'aimerais aussi remercier Charline, qui m'a aussi aidé au cours de ces travaux de thèse.

J'aimerais aussi remercier Pauline, ma seule et unique stagiaire, vu que le covid m'a empêché de rencontrer Justine. Merci pour tes questions et pour ton envie d'apprendre, et merci pour cette manipe réalisée à la serre, on ne croyait pas s'en sortir avec tous ces pots.

Merci aux membres de l'équipe 3, je pense notamment à Flo WD qui partage ma passion pour la couleur violette (et qui hérite de ma chaise de labo). Merci pour ta rigueur et pour tes cours en Master, tu nous faisais un peu peur mais t'as su nous transmettre ta passion pour la science. Merci aussi pour tes conseils, tes remarques et tes questions lors de mes présentations. Juliana, te quería agradecer las conversaciones que compartimos. Gracias a tus consejos (científicos y personales) y por mostrarme que hay otras opciones cuando una termina la tesis.

Merci également aux étudiants, à ceux que j'ai vu grandir (oui, oui, je suis un vieux crouton maintenant) : Cocotte, Mélanie, Aline, Cécile, Zélia, Flavien, Marine, Théo, et à ceux qui sont arrivés quand je n'étais plus trop au labo, je pense notamment à Adrian, Amandine et Irena. Mention spéciale pour Cécile, avec qui j'ai dû partager Daniel et le bureau. Je te souhaite plein de réussite et j'espère qu'un jour tu trouveras tes ARN. Un conseil pour vous tous : serrez-vous les coudes, communiquez, il ne faut pas rester seul en thèse. Je vous transmets les sages paroles que Guillaume Schwob m'a dit avant de partir à l'autre bout du monde : « c'est un marathon, pas un sprint » donc pensez à prendre du temps pour vous.

Je remercie aussi les vieux croutons qui sont partis avant moi, vous avez rendu mon M2 et mes premières années de thèse fantastiques. Je pense notamment à la fameuse soirée d'intégration des M2 et à tous ces jeudis soir passés au Toï Toï, plus les vendredis matin difficiles. J'aimerais aussi remercier mon groupe « Pause-Café » pour toutes ces conversations partagées, ma consommation en caféine a bien diminué depuis que vous êtes parties alors que ma productivité a augmenté ! Tant que je reste parmi les vieux, j'aimerais remercier Flo V pour toutes ses tapenades et pour ses encouragements. Merci aussi à Elo pour vieillir avec moi parmi tous ces jeunes, merci aussi pour tes blagues et tes conseils. Je rajoute aussi Laura R, je n'oublierai jamais



ce road trip à la cote Saint André avec ton permis fraîchement obtenu. Un grand merci à Soso pour les rires, les soirées et les tatouages partagés. Je ne comprends pas pourquoi on se calculait pas quand j'étais en M2, c'est du temps gaspillé ! T'as été ma copine de labo par excellence (et maintenant ma demoiselle d'honneur :)) merci de m'avoir donné un aperçu de la 3eme année...c'était bien pire haha. J'aurais tellement aimé faire un Lisbonne 2.0 pour ce rendu de manuscrit, mais ce n'est que partie remise. Et merci d'avoir corrigé ces remerciements, du coup si vous voyez des fautes...et bah ce n'est pas ma faute hehe.

Merci à mes camarades de M2, notamment à Colin qui remplit mon feed insta de belles photos depuis son départ. Un grand merci à ma keupiiiiine pour avoir voulu m'expliquer R un million de fois (et avoir échoué à chaque fois), finalement cette thèse n'a pas de stats et c'est parfait ! Je pense aussi à Lucas et Céline qui restent encore en thèse, accrochez-vous, on se voit de de l'autre côté. Je continue avec Clem, que je connais depuis la licence. Puis en master on a rencontré Laura C et depuis on est inséparables. Quelle fierté de vous voir grandir et devenir docteur les meufs, et aussi quelle fierté de vous avoir comme demoiselles d'honneur (si le covid le veut bien). Merci pour ces années partagées, et vous avez intérêt à ne pas m'oublier quand vous partirez en post doc.

Avant de sortir du labo j'aimerais remercier les membres de l'AssoDocE2M2, dommage qu'on n'ait pas organiser nos évènements cette année. Je remercie également tous les gens que j'ai oublié (oops).

En dehors du labo, j'aimerais remercier du fond du cœur Jordan Vacheron, qui a cru en moi quand je n'étais qu'une M1 (ou qui m'a fait croire qu'il croyait en moi, je ne saurai jamais). Merci pour ton accompagnement, tes conseils, ton encouragement, si je finis cette thèse c'est en grande partie grâce à toi. Ça n'a pas toujours été facile entre nous, mais malgré tout, t'as toujours été là quand j'en avais le plus besoin. Je pense que je ne te remercierai jamais assez.

Merci à mon groupe de potes qui me suis depuis le lycée : Antho, Biquet, Guillaume et Thomas et aux filles, Alex, Marine, Maud, qui se sont rajoutées au fil des années (je vous ai mis par ordre alphabétique, je ne veux pas de jaloux !). Merci à vous tous qui n'avez jamais compris sur quoi je bosse mais qui voulez depuis toujours assister à ma soutenance. Merci aussi à Etienne qui s'est rajouté plus tard, merci pour tes mouchoirs qui ont bien servi lors de mes mental breakdowns et pour tes conseils en insomnie. Toi aussi tu vas t'en sortir. Merci pour toutes les soirées (à l'Ayers boat et ailleurs), journées, BBQ ensoleillées et pluvieuses, sorties au lac pour découvrir, merci pour ce faux mariage merveilleux.

Muchas gracias a Yina, por esta amistad que ya tiene más de 12 años, gracias por todo. Es que no me alcanzan las palabras para decirte todo lo que siento. Gracias por motivarme, por levantarme cada vez que me caí, por siempre estar a mi lado, por ser uno de mis cables a tierra. Gracias a mis papas y a mi hermana por su apoyo durante estos estudios interminables. Gracias por su paciencia, y por preguntarme por mis experimentos, aunque creo que nunca los entendieron. Gracias a mi hermana por el “survival kit” que me salvo estas últimas semanas. Quería agradecer especialmente a mi papa, por sus esfuerzos por darnos un futuro mejor, sin ti, yo no estaría aquí escribiendo esto. Un grand merci aussi à mes beaux-parents, pour leur soutien au cours de cette thèse. Eric, toi qui a fait une thèse, t'aurais pu me dire ce qui m'attendait !

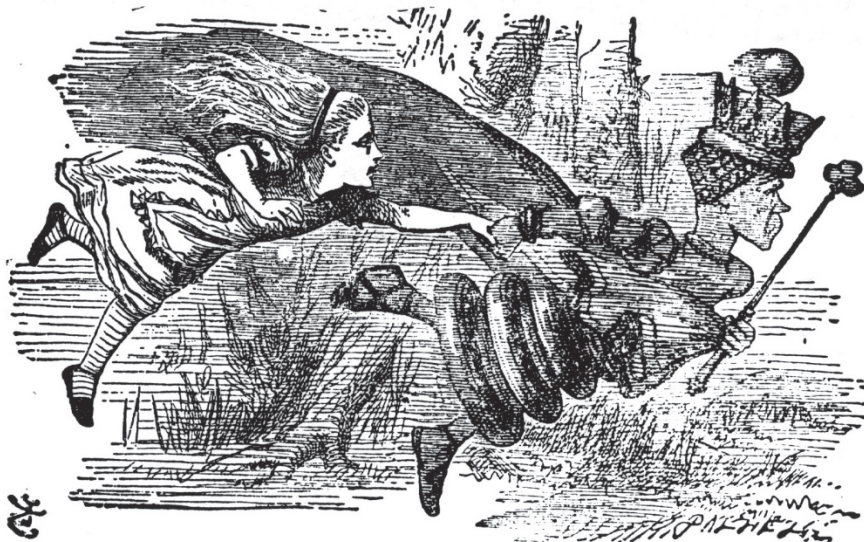
Et le meilleur pour la fin, un grand merci à Nico qui partage ma vie depuis 10 ans déjà. Merci pour ton soutien et tes encouragements malgré le nombre de fois où j'ai voulu baisser les bras. Merci de me dire que t'étais fier de moi quand moi j'avais juste envie de tout balancer. Merci pour ton aide, ta patience et ton amour inconditionnel pendant ce confinement rédactionnel. Et merci pour les références (très belle preuve d'amour d'ailleurs !). I can't wait to share the rest of my life with you.

Adieu les moches !



“Well, in our country,” said Alice, still panting a little, “you’d generally get to somewhere else — if you run very fast for a long time, as we’ve been doing.”

“A slow sort of country!” said the Queen. “Now, here, you see, it takes all the running you can do, to keep in the same place. If you want to get somewhere else, you must run at least twice as fast as that!”



Alice's adventures in Wonderland, Lewis Carroll, 1865

This passage was the inspiration for “The Red Queen hypothesis” (Van Valen, 1977), an evolutionary theory stating that species must constantly adapt in order to survive the ever-evolving opposing species.

---

## LIST OF FIGURES

---

### INTRODUCTION

Figure 1: Rhizomicrobiota recruitment by plants.

Figure 2: Lenski's long-term experimental evolution

### CHAPTER 1: Experimental evolution: a tool for deciphering the functioning and evolution of microbial communities

Figure 1: Experimental Evolution set up.

Figure 2: Plant selection of beneficial bacteria.

Figure 3: Black Queen Hypothesis in the rhizosphere.

#### Designing synthetic communities

Figure 1: Interaction network of a cellulose degrading synthetic community.

Figure 2: Surface assembly of a functional minicellulosome by intracellular complementation.

#### Synthetic communities for plants

Figure 1: Designing synthetic communities for plants.

Figure 2: Functional redundancy within microbial communities.

### CHAPTER 2: Deciphering maize selection of plant beneficial bacteria via an experimental evolution approach

Figure 1: Experimental evolution set up.

Figure 2: Phylogenetic tree of the ten *Pseudomonas* strains used in this study.

Figure 3: Populations proportions obtained for the initial SynCom and for the *in vitro* EE.

Figure 4: Average PBFC genes/clusters distribution for the initial SynCom (yellow) and for the *in vitro* EE SynCom recovered at the cycles 1, 5, 15 and 20.

Figure 5: Evolution of *Pseudomonas* populations during cycles of the *in planta* experimental evolution.

Figure 6: *Pseudomonas* population proportions obtained for the initial SynCom and for the *in planta* EE (A) for SynCom 1, 3 and 6 from cycles 1 to 27 and (B) for SynCom 2, 4, 5 and 7 from cycles 1 to 27.

Figure 7: Average PBFC genes/clusters distribution for the initial SynCom (yellow) and for the *in planta* SynCom 1, 3 and 6 and SynCom 2, 4, 5 and 7 at the end of the 27th cycle of the EE.

Figure S1: Growth curves of each *Pseudomonas* strain on a maize-based medium using bioscreen

(Lab systems Helsinki, Finland).

Figure S2: Average PBFC genes/clusters distribution for the initial SynCom and for the SynCom 1, 3 and 6 at the end of cycles 1 to 6 and 27 of the *in planta* EE.

Figure S3: Average PBFC genes/clusters distribution for the initial SynCom and for the SynCom 2, 4, 5 and 7 at the end of cycles 1 to 6 and 27 of the *in planta* EE.

### **CHAPTER 3: Experimental evolution of a synthetic community of *Pseudomonas*: different trajectories to reach the same functional network**

Figure 1: *Pseudomonas strains* relative abundance in each EE for A) the initial SynCom and for the *in planta* EE for SynCom 1, 2, 6 and 7 at the end of the 1<sup>st</sup> cycle and B) for the *in planta* EE for SynCom 1, 2, 6 and 7 at the end of the 27<sup>th</sup> cycle.

Figure 2: Sector chart representing the proportions of mutated genes (23) at the end of the 27<sup>th</sup> cycle belonging to the 3 functional EGGNOG categories.

Figure 3: Sector chart representing the proportions of mutated genes (42) detected in all of the evolved clones of *P. protegens* JV245A at the end of the 27<sup>th</sup> cycle belonging to the 3 functional EGGNOG categories.

Figure S4: *Pseudomonas* population proportions obtained for the initial SynCom and for the *in planta* EE (A) for MG-SynCom 1,3 and 6 for cycles 1 and 27 and (B) for MG-SynCom 2 and 7 for cycles 1 and 27.

Figure S5: Average PBFC genes/clusters distribution for the initial SynCom (yellow) and for the *in planta* MG-SynCom 1,3 and 6 and MG-SynCom 2 and 7 at the end of the 27<sup>th</sup> cycle.

### **CHAPTER 4: GENERAL DISCUSSION AND PERSPECTIVES**

Figure 1: Summary figure of the *in planta* experimental evolution study and work perspectives.

Figure 2: Experimental summary figure for the *in vitro* experimental evolution and work perspectives.

---

## LIST OF TABLES

---

### CHAPTER 1: Experimental evolution: a tool for deciphering the functioning and evolution of microbial communities

Table 1: Survey of microbial experimental evolutions.

#### Designing synthetic communities

Table 1: Five metabolism based bacterial interactions.

### CHAPTER 2: Deciphering maize selection of plant beneficial bacteria via an experimental evolution approach

Table 1: Different positions of nucleotide variations between the 10 studied strains of *Pseudomonas* for the *rpoD* gene.

Table 2: Analyzed samples for both experimental evolutions.

Table 3: Plant Beneficial Contributing Genes of the 10 studied strains of *Pseudomonas*.

Table 4: General genome features of the 10 studied strains of *Pseudomonas*.

Table S1: Secondary metabolites and corresponding gene clusters of *Pseudomonas* strains.

Table S2: Pairwise comparison of the 10 *Pseudomonas* genomes vs. type strain genomes.

Table S3: Type based species and subspecies clustering.

Table S4: Biosynthesis gene clusters for secondary metabolites predicted with AntiSMASH.

Table S5: Number of *rpoD* reads obtained for the initial SynCom.

Table S6: Relative abundances (expressed in percent) of each of the 10 *Pseudomonas* strains in the initial SynCom.

Table S7: Number of *rpoD* reads assigned to each *Pseudomonas* population, obtained for each SynCom 1 to 5 for the *in vitro* EE, at (A) cycle 1, (B) cycle 5, (C) cycle 15 and (D) cycle 20.

Table S8: Relative abundances (expressed in percent) of each of the 10 *Pseudomonas* strains obtained for the *in vitro* EE at the end of the cycles 1, 5, 15 and 20.

Table S9: Number of *rpoD* reads assigned to each *Pseudomonas* population obtained for (A) SynCom 1, 3 and 6, and (B) SynCom 2, 4, 5 and 7 at different cycles of the *in planta* EE.

Table S10: Relative abundances obtained for (A) SynCom 1, 3 and 6 and (B) SynCom 2, 4, 5 and 7 at different cycles of the *in planta* EE.

Table S11: Number of total reads and number of specific strain reads assigned to *Pseudomonas* populations obtained for the initial SynCom.

Table S12: Relative abundances (expressed in percent) for the initial SynCom.

Table S13: Number of total reads and number of specific strain reads assigned to *Pseudomonas* populations, obtained for SynCom 1, 2, 3, 6 and 7 at the end of the 1<sup>st</sup> cycle of the *in planta* EE (= MG-SynCom X-1 with X= 1, 2, 3, 6 or 7).

Table S14: Relative abundances (expressed in percent) for SynCom 1, 2, 3, 6 and 7 at the end of the 1<sup>st</sup> cycle of the *in planta* EE (= MG-SynCom X-1 with X= 1, 2, 3, 6 or 7).

Table S15: Number of total reads and number of specific strain read assigned to *Pseudomonas* population, obtained for SynCom 1, 2, 3, 6 and 7 at the end of the 27<sup>th</sup> cycle of the *in planta* EE (= MG-SynCom X-1 with X= 1, 2, 3, 6 or 7).

Table S16: Relative abundances (expressed in percent) for SynCom 1, 2, 3, 6 and 7 at the end of the 27<sup>th</sup> cycle of the *in planta* EE (= MG-SynCom X-1 with X= 1, 2, 3, 6 or 7).

Table S17: Number of specific genes for each strain.

### **CHAPTER 3: Experimental evolution of a synthetic community of *Pseudomonas*: different trajectories to reach the same functional network**

Table 1: Relative abundances (expressed in percent) for the initial SynCom, for SynComs from the 1<sup>st</sup> cycle (A) and from the 27<sup>th</sup> cycle (B).

Table 2: Number of mutated genes predicted by Breseq software for strains *P. fluorescens* JV391D10, JV391D17, and *P. protegens* JV245A for SynComs retrieved at the end of the 1<sup>st</sup> and 27<sup>th</sup> cycle of EE.

Table 3: Survey of mutated genes for strains *P. fluorescens* JV391D10, JV391D17 and *P. protegens* JV245A found in the SynComs retrieved at the end of the 1<sup>st</sup> cycle (A and B) and at the end of the 27<sup>th</sup> cycle (C, D and E).

Table 4: Survey of the mutations found for the evolved clones of *P. protegens* JV245A.

Table 5: Intergenic mutations found in the evolved clones of *P. protegens* JV245A.

Table 6: Non-silent mutations found for the *gacA*.

Table S1: Non-silent mutations found for *P. fluorescens* JV391D10 at the end of the 27<sup>th</sup> cycle.

Table S2: Non-silent mutations found for *P. fluorescens* JV391D17 at the end of the 27<sup>th</sup> cycle.

Table S3: Non-silent mutations found for *P. protegens* JV245A at the end of the 27<sup>th</sup> cycle.



---

## GLOSSARY

---

12	List of tables and figures
20	<b>GENERAL INTRODUCTION</b>
25	Objectives of the work and structure of the manuscript
28	<b>CHAPTER 1</b>
	<b>Experimental evolution: a tool for deciphering the functioning and evolution of microbial communities</b>
30	<b>INTRODUCTION</b>
34	<b>1. Plant-associated microorganisms in experimental evolution</b>
36	<u>1.1 Plant pathogens</u>
37	<u>1.2 Non-pathogenic plant-root-associated microorganisms</u>
39	<b>2. Experimental evolution to decipher the adaptive evolution of natural communities</b>
39	<u>2.1 Adaptive responses and parallel or convergent evolution</u>
40	<u>2.2 Co-evolution between populations in a community</u>
41	<b>3. Towards a full consideration of microbial communities in experimental evolution studies</b>
41	<u>3.1 From single microorganisms to complex communities</u>
43	<u>3.2 From simple to somewhat more complex environments</u>
45	<b>CONCLUDING REMARKS</b>
52	<b>Designing synthetic communities</b>
52	<b>INTRODUCTION</b>
53	<b>1. Community function as the foundation for SynCom design</b>
53	<b>2. SynComs constructed based on interaction networks</b>
54	<u>2.1 Metabolism and cross-feeding based cooperation</u>
56	<u>2.2 Environmental factors driven cooperation</u>
57	<b>3. Application of SynComs</b>
57	<u>3.1 SynComs for fundamental research: adaptation to new environments</u>
59	<u>3.2 SynComs for applied research: bioproduction of molecules of interest</u>
62	<b>CONCLUDING REMARKS</b>

63	<b>Synthetic Communities for plants</b>
63	<b>INTRODUCTION</b>
64	1. The importance of plant beneficial properties
66	2. Functional redundancy
67	3. Rhizosphere establishment of the plant-beneficial bacteria
69	4. Diversity and species richness within the SynCom
70	5. Plant recognition of inoculated beneficial microorganisms
71	<b>CONCLUDING REMARKS</b>
78	<b>CHAPTER 2</b>
	<b>Deciphering maize selection of plant beneficial bacteria via an experimental evolution approach</b>
80	<b>INTRODUCTION</b>
82	<b>MATERIAL AND METHODS</b>
91	<b>RESULTS</b>
91	General genomes features
91	Plant beneficial properties distribution within the initial SynCom
92	Figuring out the initial SynCom composition
93	<i>In vitro</i> experimental evolution
95	<u><i>Pseudomonas</i> population abundances after the 1<sup>st</sup> cycle of the <i>in vitro</i> EE</u>
95	<u><i>Pseudomonas</i> population dynamics after 5, 15 and 20 cycles of the <i>in vitro</i> EE</u>
95	<i>In planta</i> experimental evolution
97	<u>Population dynamics for SynCom 1, 3 and 6</u>
98	<u>Population dynamics for SynCom 2, 4, 5 and 7</u>
99	<u>PBCF distribution in <i>in planta</i>-evolved SynComs</u>
101	<b>DISCUSSION</b>
111	<b>SUPPLEMENTARY DATA</b>
133	<b>SUPPLEMENTARY RESULTS</b>
133	<b>MATERIAL AND METHODS</b>
134	<b>RESULTS</b>
134	Initial SynCom composition via the analysis of the total read counts
136	Analyses of the <i>in planta</i> evolution of the SynCom on total read counts
139	Analyses of the <i>in planta</i> evolution of the SynCom on strain-specific read counts
139	<u>Population dynamics for SynComs 1, 3 and 6</u>

140	<u>Population dynamics for SynComs 2 and 7</u>
140	<b>PBCF distribution in <i>in planta</i>-evolved SynComs</b>
142	<b>DISCUSSION</b>
147	<b>CHAPTER 3</b>
	<b>Experimental evolution of a synthetic community of <i>Pseudomonas</i>: different trajectories to reach the same functional network</b>
149	<b>INTRODUCTION</b>
150	<b>MATERIAL AND METHODS</b>
154	<b>RESULTS AND DISCUSSION</b>
154	<b>Genetic variations within Synthetic Community metagenomes data</b>
159	<b>Evolution of genetic variations for each strain (comparative genomics)</b>
160	<u><i>P. fluorescens</i> JV391D10</u>
162	<u><i>P. fluorescens</i> JV391D17</u>
163	<u><i>P. protegens</i> JV245A</u>
165	<b>Genome variations of variants isolated after the 27<sup>th</sup> cycle</b>
165	<b>Genomic variation for the evolved clones from JV245A</b>
168	<u>Intergenic mutations</u>
168	<u>Poorly characterized category</u>
170	<u>Metabolism category</u>
170	<u>Information storage and processing category</u>
170	<u>Cellular process and signalling category</u>
172	<b>CONCLUDING REMARKS</b>
178	<b>SUPPLEMENTARY DATA</b>
184	<b>CHAPTER 4</b>
	<b>GENERAL DISCUSSION AND PERSPECTIVES</b>
187	<b>Population dynamics of the synthetic community</b>
191	<b>Impact of the <i>in planta</i> experimental evolution on the strains genomes</b>

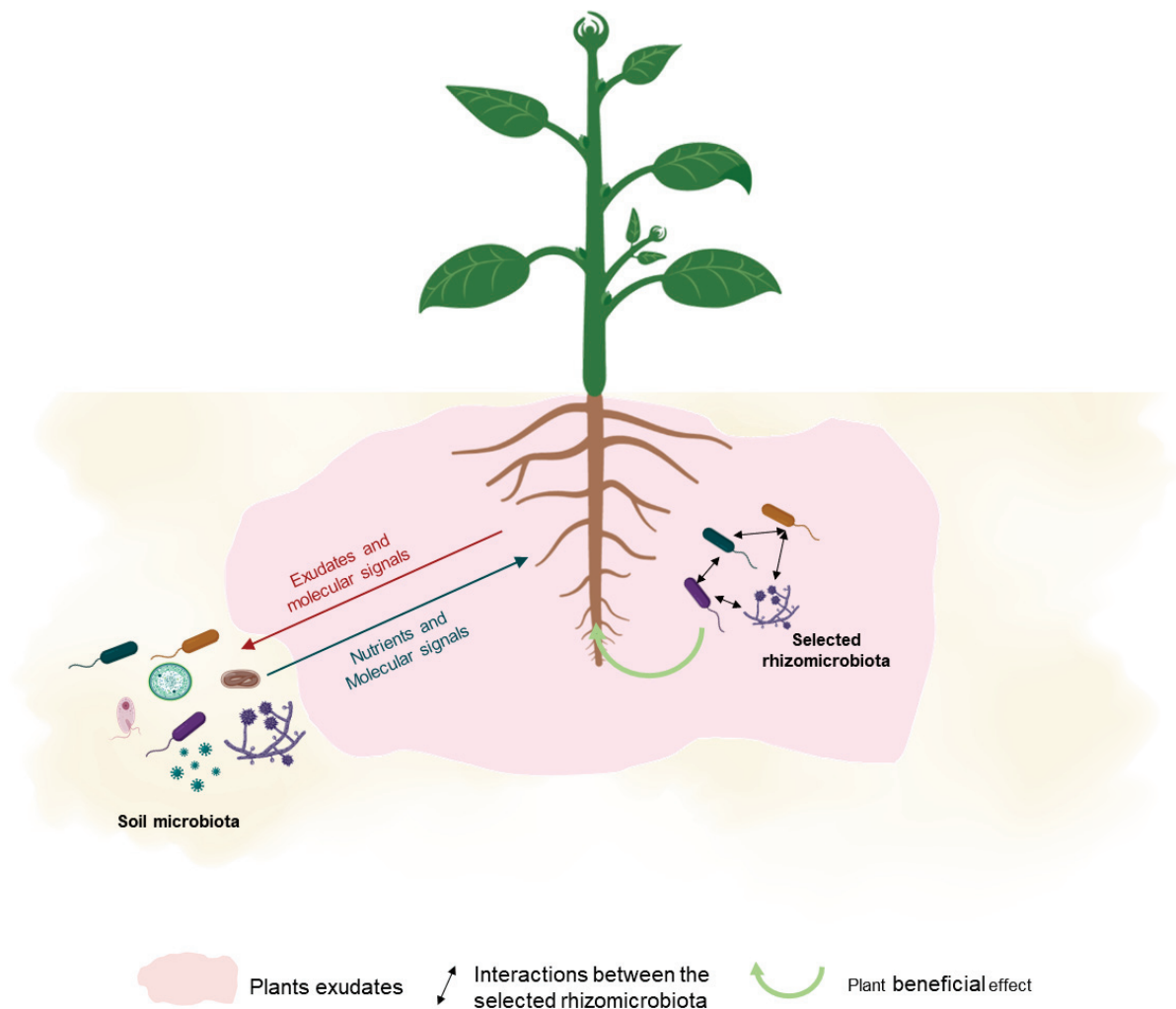


---

## GENERAL INTRODUCTION

---

Crop yield is one of the main concerns in agriculture as it has become essential to keep elevated levels of food production to feed Earth's never ending growing population (Parnell *et al.*, 2016). Technical advances such as plant breeding, crop protection, fertilization and agronomic management have helped us to overcome different agricultural crises. Among these technological advances, chemical pesticides and fertilizers have been widely used since the green revolution (1950-1960) as they offer the best result to maintain crop productivity. However, these chemical compounds are expensive and can have great negative impacts on the environment. The perturbation of molecular interactions between plants and soil beneficial microorganisms, such as nitrogen-fixing and phosphorus solubilizing microorganisms, has for instance been observed in soils contaminated by pesticides. Chemical pesticides are also believed to be responsible for decline of functions such as the mineralization of organic matter (Hussain *et al.*, 2009). In the early 1990, the European Union legislation (Jess *et al.*, 2014; Cray *et al.*, 2016) became more severe, and began to forbid the use of several pesticides and fertilizers that accumulate in the environment and can damage ecological systems, in addition to ban those that are highly mutagenic and/or cancer-causing. Furthermore, recent studies have reported that the excessive use of these chemical compounds can lead to soil leaching after their application. Among the most controversial molecules, dichlorodiphenyl-trichloroethane, commonly known as DDT, was widely used after World War II with great success as an insecticide and to control malaria and typhus. DDT rapidly became infamous when it was discovered that it was an endocrine disruptor and possible carcinogen, in addition to have caused egg shell thinning among many bird species. DDT utilisation mainly stopped in 1970 and the European legislation forbade its use in 1986 (Miyagawa *et al.*, 2016). However, traces of this persistent compound can still be found in soils. It is thus imperative to develop more sustainable cropping systems to preserve the environment and our health. The most important constraint limiting crop yield is low soil fertility. For an optimal plant growth, nutrients must also be available in sufficient and balanced conditions. Plant microbiota have drawn attention as these microbial inhabitants have deep impact on plant nutrition and thus, on its health and development.



**Figure 1: Rhizomicrobiota recruitment by plants.** Plants recruit their rhizomicrobiota, from the available pool of soil microorganisms (fungi, algae, bacteria, viruses, *etc.*) through the exudation of compounds (exudates) acting as nutrients, signals or antimicrobials (brown arrow). In response to these molecules, microorganisms will trigger a response in the plant via molecular signals (blue arrow). This communication will lead to plant colonization by certain microorganisms and thus, to the formation of the rhizomicrobiota. Selected microorganisms will share network of interactions, and the observed plant-beneficial effects (on growth and protection against pathogens) will be the results of those interactions.

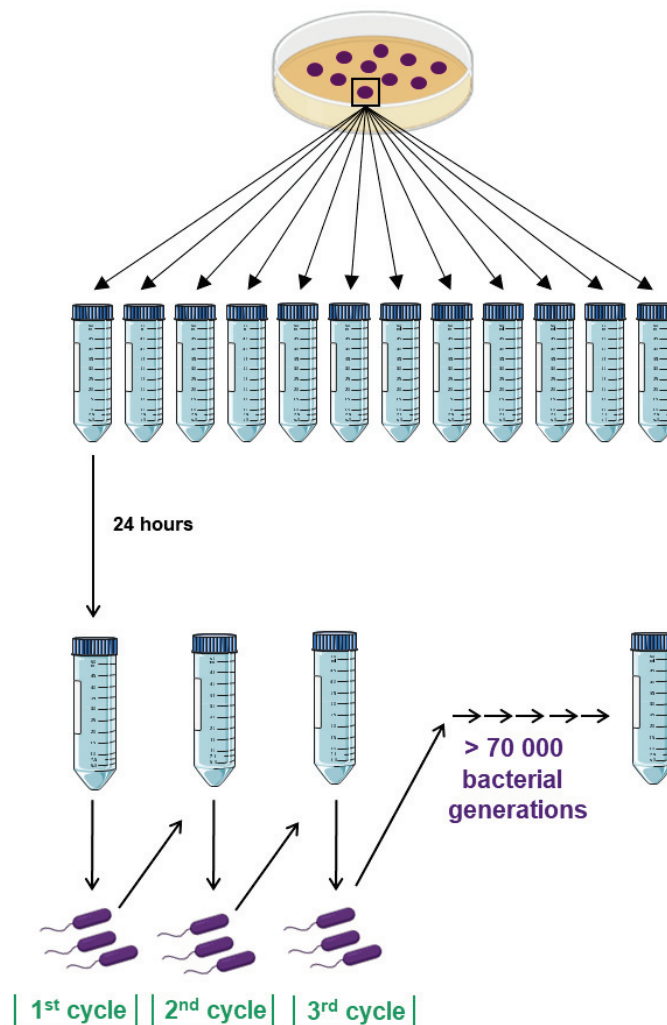
Plant roots are among the most studied natural ecosystems as they interact with many soil microorganisms offering a rich and complex network of interactions. Plants most likely recruit microorganisms via an array of exudates but microorganisms can also activate a response in the plant, leading to a specific plant-microorganisms communication involving molecular signals that may modulate their colonisation of the plant tissues. Within the selected microbiota, some microorganisms can contribute to optimal plant development (Figure 1). These beneficial interactions can be carried by mutualistic organisms or by free living bacteria able to colonize the roots (Meena, 2017; Etesami and Maheshwari, 2018).

This last category includes Plant Growth-Promoting Rhizobacteria (PGPR) which improve plant growth and health even under adverse conditions via a large range of plant beneficial properties (hormone production, plant nutrition and protection against plant pathogens, (Vacheron *et al.*, 2016). As non-strict mutualistic bacteria, PGPR are able to interact with a large panel of host plant species and have drawn attention in the last decades as they can act as biocontrol agents and fertilizers, protecting and increasing crop productivity. Although these bacteria have a wide range of beneficial functions, their inoculation does not always lead to significant beneficial effects on the plant. This may be because, under natural conditions, plants provide multiple ecological niches, so that any population tends to evolve and interact with the other sympatric microorganisms. Different interactions networks will develop (positive and negative interactions) and thus, the observed plant growth-promoting effect will be the result of these interactions. Therefore, it is crucial to unravel how inoculated bacteria colonize and persist on the plant roots despite the settled natural microbial communities. The inherent complexity of network interactions within the microbial communities makes it difficult to understand how plant microbiota assemble and function. As a first step, the inoculated microorganisms must adapt to a new abiotic and biotic environment, shaped by the roots and the microbial community of the roots.

In order to adapt to a new environment, organism must be capable of using the present resources in a more efficient way than their neighbours. To accomplish this, organisms can acquire new functions or modify the pre-existent ones. The organisms with the most adapted traits will be the more fitted in a given environment and therefore will reproduce faster (increased fitness). This process is known as natural selection. When an adaptive trait is acquired, it will require a certain number of generations to be fixed and passed onto the descendants. Fixation time and therefore adaptation will depend on the studied organism and on its generation time. This process being rather slow for macro-organism, evolutionary biology has long relied on comparative studies and deduction, rather than on evolutive experiment. Back in his time, Darwin used a comparative approach by necessity. Indeed, he studied collateral descendants from the same original parent but always pointed out the importance of observing the actual lineal ancestor when studying the process of adaptation by natural selection. Due to the limited resources and available techniques, Darwin focused on rather big animals with long generations times, leaving aside microorganisms and specially bacteria. However, their short replication time makes them ideal candidates to study evolution in real time.

Moreover, bacteria are very easy to handle in laboratory: they can develop in liquid or solid media and their fast replication rate allows to have thousands of generations in little time as long as fresh media is continuously incorporated into the system. In 1983, Richard Lenski had the idea of studying evolution backward. Instead of studying fossils to unravel passed evolution, he decided to create these fossils via an experimental evolution approach. Experimental evolution studies are

setups in which an organism is let evolved, and in order to let natural selection acts, only the environment can be imposed and controlled. Lenski's experiment is thus based on a very simple protocol: one single ancestor of *E. coli* was inoculated in several tubes containing a glucose-minimal medium; each tube was incubated and after 24 hours, a small fraction of the culture broth is retrieved, diluted and inoculated into a new tube containing fresh media (Figure 2; Lenski *et al.*, 1991). At first, Lenski thought the experiment would not last more than 8 years as *E. coli* would eventually adapt completely to its environment. Nowadays, the experimental evolution is still running and has passed 70 000 bacterial generations. In human time, it is like we could observe, in the lab, the human evolution since the time of Neanderthal men.



**Figure 2: Lenski's long-term experimental evolution.** One single clone of *E. coli* was inoculated 12 times in tubes containing a glucose-minimal medium. Each tube was incubated 24h and afterwards, a fraction of each bacterial suspension was retrieved, diluted and re-inoculated onto fresh media. This process corresponds to one cycle of evolution and is repeated every day, since 1989 (Lenski 1991). Each serial passages corresponds to 6.7 bacterial generations (Lenski 2017) and has now exceed 70 000 generations, which equals more than 10 447 evolution cycles.



Coupled with the advances in genome sequencing, experimental evolution studies have helped unravel the genetic bases of adaptation, by tracking mutations and by correlating genetic and phenotypic changes. Since Lenski's pioneer experimental evolution, this approach has been used to study an array of ecological questions such as the evolutionary adaptation during biotic interactions (as co-evolution, see (Brockhurst and Koskella, 2013) or adaptation of pathogens to new host plants (Guidot *et al.*, 2014). Experimental evolution has also been used to understand the underlying mechanisms of virus-bacteria co-evolution (Paterson *et al.*, 2010; Scanlan *et al.*, 2011) and adaptation of bacterial pathogens to humans (Yang *et al.*, 2011). Overall, this approach has brought a deeper understanding of ecological interactions and evolutionary processes such as co-evolution between species.

However, most of experimental evolution studies focus on genetic diversity within species rather than species diversity within communities. Recent studies have highlighted the impact communities can have on pairwise interactions and on the adaptation rate of species (Cairns *et al.*, 2018; Smith and Schuster, 2019). Ecological and evolutionary process can be modified in complex microbial communities, such as the ones found in plants, due to more complex competitive interactions and multiple selection pressures. Moreover, multispecies communities house some key ecological features (diversity, stability, succession and high order species interactions) that are not observable in single organism studies. Therefore, there is an increasing need to expand the biotic complexity of experimental setups in order to unravel the mechanics of natural ecosystems whether it is for the purpose of understanding natural systems or for more applied studies looking to engineer a community that can better perform a function when re introduced in its natural habitat. In an effort to mimic natural environmental conditions, synthetic communities have risen up in the field of plant research these last few years. Synthetic communities are defined as artificially created communities of 2 or more microorganisms and offer a great opportunity: that of unravelling the network of interactions within microbial communities and the possibility to engineer an inoculum with a desired function. Furthermore, as synthetic communities offer a simplified representation of natural communities they can also help to understand the impact abiotic (*e.g.* salinity, temperature, *etc.*) and biotic (*e.g.* cooperation, antagonism, *etc.*) factors can have on natural microbial communities' complexity (abundance, diversity of species and functioning).

## Objectives of the work and structure of the manuscript

Here we propose an experimental evolution carried on maize plants for 27 evolution cycles, with a synthetic community composed of 10 *Pseudomonas* PGPR. *Pseudomonas* possess a wide-range of plant beneficial functions (Sarniguet *et al.*, 1995; Sánchez-Contreras *et al.*, 2002; Loper *et al.*, 2012; Biessy *et al.*, 2019), however inoculation does not always lead to significant effects on the plant. Indeed, the mechanisms favoring the establishment of PGPR in the rhizosphere and the expression of their properties remain largely unknown. A study prior to this work revealed that within the maize rhizosphere, *Pseudomonas* PGPR with a low number of plant-beneficial functions and belonging to the FMJK group (*i.e.* *P. fluorescens*, *P. mandelii*, *P. jessenii*, *P. koreensis* subgroups) are more present in the rhizosphere than those harboring a high number of plant-beneficial functions, suggesting that there might be a trade-off between rhizosphere prevalence and the ability to maintain a large number of plant-beneficial properties in the rhizosphere (Vacheron *et al.*, 2016). The main aim of this study was thus to provide insights into plant-*Pseudomonas* cooperation by taking into account this issue. To do this, an initial synthetic community was designed with 10 strains from 7 different species, with 3 strains belonging to *P. fluorescens* species (*i.e.* FMJK group) and 2 to *P. protegens* (*i.e.* CPC group for *P. corrugata*, *P. protegens* and *P. chlororaphis* subgroups) species in order to give insight into intra-specific interactions as well as inter-specific interactions. Furthermore, strains were chosen so half of the *Pseudomonas* harbour a low number of plant-beneficial function (10 or less) in their genomes and the other half, a high number (11 or more). The hypothesis of this work was that plants select *Pseudomonas* PGPR populations adapted to cooperate with each other under rhizosphere conditions, maximizing the taxonomic diversity and functional redundancy of *Pseudomonas* populations. The objectives were thus to evaluate the population dynamics of the different *Pseudomonas* strains evolving within the rhizosphere of their host plant, over the course of an experimental evolution approach. Afterwards, evolved clones were selected from the last cycle of the experimental evolution and screened for genetic changes.

This manuscript is organized in 4 chapters. Firstly, a literature survey is presented. This bibliographic introduction is organized into 3 parts. The first part is a review of the literature, providing a non-exhaustive overview of experimental evolution studies carried out between plants and plant-associated microorganisms, as well as highlighting the importance of taking into account microbial communities and complex environments in this type of studies. This review is untitled 'Experimental evolution: a tool for deciphering the functioning and evolution of microbial communities' and has been submitted to Frontiers in Microbiology in October 2020.

The second part of this section focuses on the design and study of synthetic communities. The main principles on which synthetic communities are often built on are discussed as well as the application of these synthetic communities. In a third part, a focus is made on the design of synthetic communities aiming to mimic natural plant communities.

The second chapter aims to give insights into plant-PGPR cooperation by studying the population dynamics of a synthetic community of 10 *Pseudomonas* on maize roots and to unravel the role of the plant within this cooperation. To do so, the initial synthetic community was evolved for 27 cycles and population dynamics were tracked. In parallel, another experimental evolution was carried with the same synthetic community in rich liquid media to confirm the role of the plant on the *Pseudomonas* population dynamics.

Then, the purpose of the third chapter was to study the eventual genetic adaptation of *Pseudomonas* when they evolved to the contact of maize roots. In order to do this, genetic changes in the *Pseudomonas* genomes were studied at two levels, first at the level of metagenomes of the evolved communities and second, at the level of individual evolved clones isolated from the 27<sup>th</sup> cycle of evolution.

Finally, in a fourth chapter, results from the second and third chapter were discussed together, in effort to present a more complete answer to the initial hypothesis. Finally, several perspectives are proposed as this study is a groundwork offering new research avenues regarding the root-habitat adaptation of *Pseudomonas* species.

---

# CHAPTER 1

---

## Experimental evolution studies

This first part of the bibliographic chapter corresponds to a literature survey of studies done using experimental evolution approaches that has been submitted to *Frontiers in Microbiology* in October 2020.

### Experimental evolution: a tool for deciphering the functioning and evolution of microbial communities

**Beatriz Manriquez<sup>1</sup>, Daniel Muller<sup>1</sup>, and Claire Prigent-Combaret<sup>1\*</sup>**

<sup>1</sup> Univ Lyon, Université Claude Bernard Lyon 1, CNRS, INRAE, VetAgro Sup, UMR5557 Ecologie Microbienne, 43 bd du 11 novembre 1918, F-69622 Villeurbanne, France

**Running title:** Experimental evolution for deciphering adaptive evolution of natural communities

**\*Corresponding author:** UMR CNRS 5557 Ecologie Microbienne, Université Lyon 1, 43 bd du 11 Novembre 1918, 69622 Villeurbanne cedex, France. Phone +33 4 72 43 13 49. E-mail: [claire.prigent-combaret@univ-lyon1.fr](mailto:claire.prigent-combaret@univ-lyon1.fr)

## **ABSTRACT**

In natural environments, microbial communities must constantly adapt to fluctuating environmental conditions. The adaptive responses of microbial communities to environmental changes have long been difficult to decipher until development of experimental evolution studies. The latter aim to analyse in real-time the evolution of microbial populations in response to imposed environmental factors or during the interaction with a host, by screening for phenotypic and genotypic changes over a multitude of identical experimental cycles. Experimental evolution (EE) coupled with comparative genomics has facilitated the monitoring of bacterial genetic evolution and the understanding of adaptive evolution processes. Basically EE studies had long been done on single strains, allowing to reveal the dynamics and genetic targets of natural selection and to uncover the correlation between genetic and phenotypic adaptive changes. However, species are always evolving in relation with other species and have to adapt not only to the environment itself but also to the biotic environment dynamically shaped by the other species. Nowadays, there is a growing interest to apply EE on microbial communities evolving under natural environments. In this review, after having given a non-extensive overview of EE studies realized between plants and plant-associated microorganisms, and analysing the mechanisms controlling the functioning of microbial species and their adaptive responses to environment changes, we examine the importance of considering bacterial communities and complex environments in EE studies.

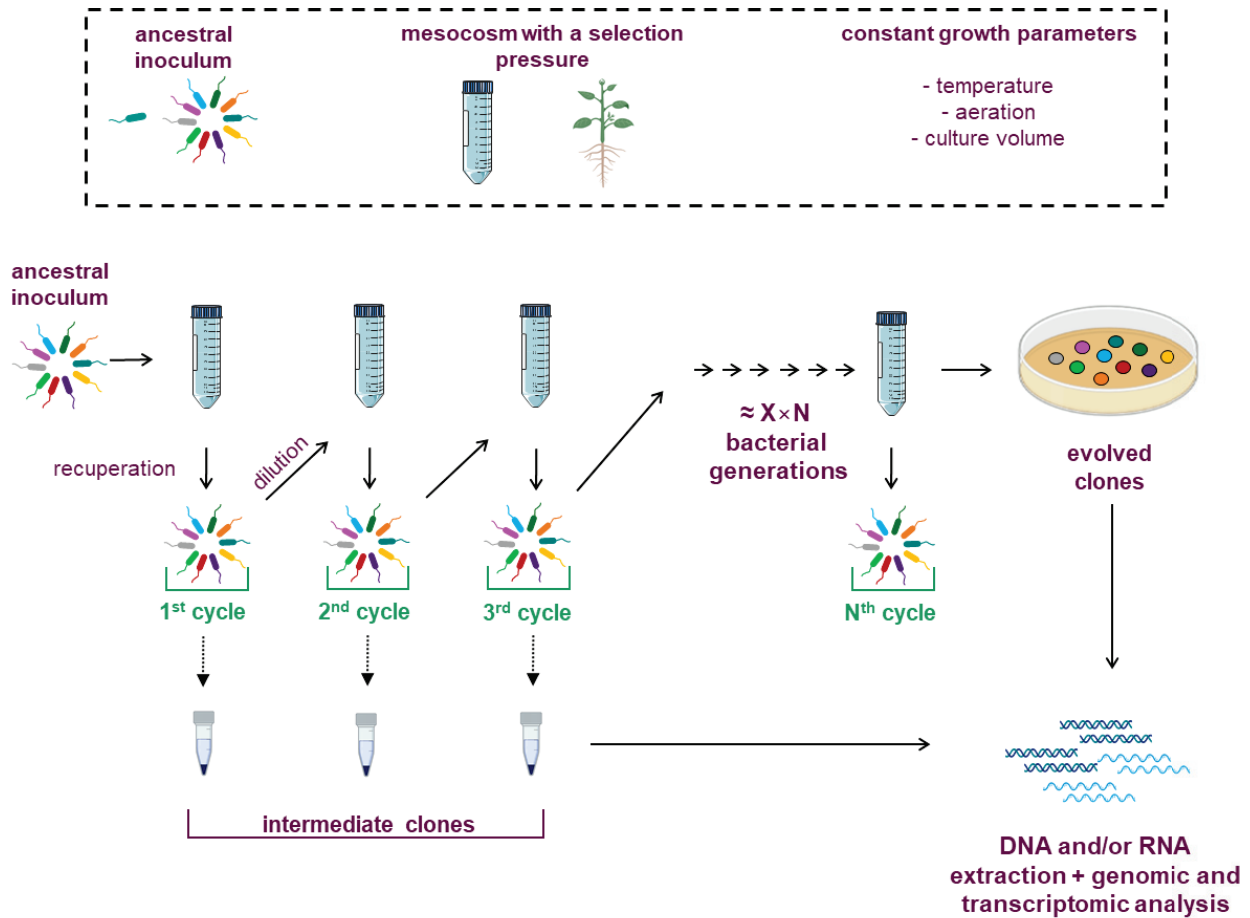
**Keywords:** synthetic community, interaction network, evolutive adaptation, microbiota, holobiont

## INTRODUCTION

In order to survive in a changing environment, species need to interact with each other. These interactions have given rise to major changes in the evolution of prokaryotes and eukaryotes such as during the transition from aquatic to terrestrial ecosystems (Battistuzzi and Hedges, 2009; Wisniewski-Dyé *et al.*, 2011). The evolutionary process that allows an organism to fit its environment is known as adaptation. Adaptation is the result of mutations in genes that either improve an existing function (point mutations) or enable the organism to acquire new functions (*i.e.* duplication followed by gene mutation, or more often acquisition by horizontal gene transfer). When individual genomes acquire adaptive mutations (genetic variations) that allow a gain in fitness, and a better competitiveness in the niche, these mutations will be fixed over generations. This process is known as natural selection. Nevertheless, these genetic variations are rare and often non-adaptive. As an example, the rate for point mutations in *Escherichia coli* is below  $10^{-11}$  per base-pair per generation (Drake, 1991; Wielgoss *et al.*, 2011). However, their rapid generation time increases the chances of pointing out a beneficial mutation. This process can be complex to identify in natural populations of species with long generation times.

One representative example of monitoring rapid bacterial adaptation was assessed with *E. coli* on “The Microbial Evolution and Growth Arena” plate. This plate is composed by several regions starting with an antibiotic-free section in the outermost part and with progressively increasing concentrations towards the middle of the plate (3, 30, 300, 300x wild-type Minimal Inhibitory Concentration, MIC). *E. coli* was initially inoculated on the antibiotic-free section and within 10 days, evolved mutants with a surprisingly 10,000-fold increase in MIC were sampled from the highest concentration region (300x wild-type MIC) (Baym *et al.*, 2016). Adaptation is a driving force of evolution and can be studied at a genetic level but, we often see the functional consequences of the adaptation and not the process leading to this adaptation.

Since seventy years, the advent of experimental evolution (EE) studies (Kawecki *et al.*, 2012) - that can be briefly defined as the analyses of the evolution of individuals in response to imposed environmental conditions over a multitude of identical experimental cycles (Figure 1)-made it possible to monitor the real-time adaptation of a population to its environment by detecting phenotypic changes between individuals in the population or for the invasion of mutated alleles within the population. EE approaches have come a long way since the first experiments with continuous-culture employed in the early 1950s focused on describing dominant phenotypes (reviewed in Adams and Rosenzweig, 2014). In the decade that followed, scientists attempted to correlate the observed phenotypic changes with gene mutations or duplications, revealing that genetic adaptation is the basis of evolutionary adaptation. During the 1970s and 1980s, the development of genetic analysis techniques led to a better resolution of mutations.



**Figure 1: Experimental Evolution set up.** One or several microorganisms are inoculated into a mesocosm (liquid or solid media or eukaryotic hosts). We refer to this initial inoculum as “ancestral” in order to differentiate from the evolved microorganisms that will be recovered at the end of each experimental cycle. All the other parameters of the experimental set up (*i.e.* temperature, aeration, culture volume, *etc.*) are kept constant, thus adaptation is confined to only one selection pressure (nutrient availability in this example). An experimental cycle is composed of a number of generations ( $X$ ), which is linked to the time needed for the microorganisms to utilize all the nutrients present in the mesocosm. After each experimental cycle, microorganisms are retrieved and diluted before being re-inoculated into a fresh mesocosm and let evolve for another experimental cycle. At each experimental cycle, samples can be stored. DNA and/ or RNA can be extracted and analysed to compare ancestral, intermediate and/or evolved microorganisms’ genomes and transcriptomes. The total number of bacterial generations ( $N$ ) depends on each experiment, however, literature seem to agree on several hundreds to thousands of generations in order to observe genetic modifications.

These advances showed that the evolutionary adaptation of a monoculture in a unique and limited resource environment can happen within the first hundred cell generations (Lenski and Travisano, 1994; Cooper and Lenski, 2000; Wiser *et al.*, 2013; Kram *et al.*, 2017; Lenski, 2017). Finally, with the advent of mass sequencing, the mechanisms of evolutionary adaptation are understood at the level of complete genomes. EE coupled with comparative genomics has facilitated the monitoring of bacterial genetic evolution and the understanding of adaptive evolution processes (Box1; Table 1; Kawecki *et al.*, 2012; Bailey and Bataillon, 2016). These experiments are being used to track



mutation rates, to reveal the dynamics and genetic targets of natural selection and to uncover the correlation between genetic and phenotypic changes. Nowadays, experimental evolution is a tool used to answer more complex ecological questions, such as evolutionary adaptation during biotic interactions (as co-evolution see Brockhurst and Koskella, 2013) or adaptation of pathogens to new host plants (Guidot *et al.*, 2014) (Table 1). It is also a powerful tool for understanding the underlying mechanisms of virus-bacteria co-evolution (Paterson *et al.*, 2010; Scanlan *et al.*, 2011) and adaptation of pathogens to humans (Yang *et al.*, 2011) but the latter topics are not discussed in the present review. Here, first we focus on EE studies realized between prokaryotes and eukaryote hosts that can achieve their complete life cycle under controlled conditions in a short time, like between plants and plant-associated microorganisms. Second, we analysed the mechanisms controlling the functioning and evolution of microbial species and their adaptive responses to environment changes. Finally, we examine the importance of considering bacterial communities and complex environments in EE studies.

### **Box 1: The genetic mechanics underlying adaptation in bacteria**

Experimental evolution has mostly been carried out on very simple conditions with simplified systems (Barrick *et al.*, 2009). EE shows real time adaptation of a given organism to its environment thanks to mutations and the action of natural selection. In order to allow natural selection to work, the experimenter can only impose the environmental conditions in an EE assay (rather than selecting organisms on a particular trait). The experimenter controls the environmental conditions but does not impose the selection; selection is here the outcome of the struggle for life (natural selection). Adaptation happens via mutations that can improve the performance (and reproduction) of an organism in its environment and therefore improve its fitness (Lang *et al.*, 2011). Different conditions impose different selection pressures and determine whether or not the mutation is beneficial or detrimental in a given niche at a given time. If a beneficial mutation happens on a single clone it will need a certain amount of time to rise to the population level (fixation). The time required is inversely proportional to the advantages it gives (Fisher, 1930; Tenaillon, 2014). Moreover, random drift also plays an important role in fixation of mutation. This evolutive force represents the role of chance in deciding which organisms survive and reproduce (despite the difference in their fitness). Genetic variation is therefore reduced randomly and many mutations disappear before they can be fixed (Elena and Lenski 2003). Overall, the fate of a beneficial mutation greatly depends on its own fitness: the more advantageous the mutation is the less it will be subjected to random drift as it will be rapidly fixed (Lenski *et al.*, 1991). In this context, bacteria are powerful candidates to pinpoint mutations as they offer short generation time and large population size, so multiple mutations can be present simultaneously.

**Table 1: Survey of microbial experimental evolutions.** The duration of experiments is given in number of generations when known or by the number of serial transfer.

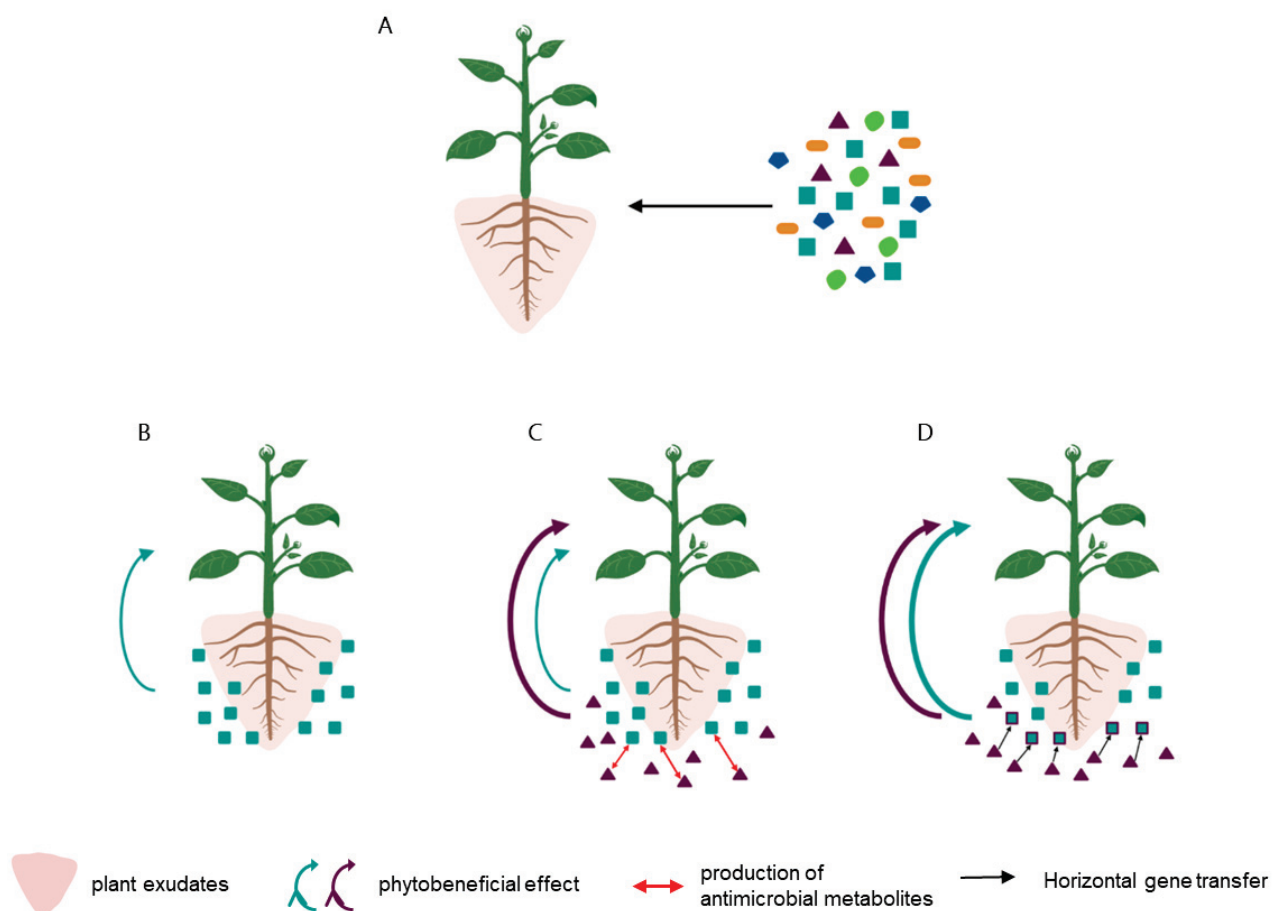
System	Organism	Description	Reference
Simple media	<i>E. coli</i> single clone (for a more thorough review please refer to Lenski 2017)	Ongoing EE carried with <i>E. coli</i> on a glucose-limited minimal salts medium.	Lenski <i>et al.</i> , 1991; Lenski and Travisano, 1994; de Visser and Lenski, 2002; Khan <i>et al.</i> , 2011
		EE carried with <i>E. coli</i> on a glucose-limited environment for > 20 000 generations revealed parallel changes in terms of beneficial mutations.	Pelosi <i>et al.</i> , 2006
		EE carried with <i>E. coli</i> on minimal medium for 660 generations. Sequencing of 5 evolved clones highlighted the presence of parallel mutations.	Herring <i>et al.</i> , 2006
		EE carried for 1000 generations on maltose with <i>E. coli</i> lineages previously evolved on glucose medium showed convergent adaptation.	Travisano <i>et al.</i> , 1995
		EE carried with <i>E. coli</i> on an acidic-alkaline media with pH variation between pH 5.3 and 7.8 for 2000 generations revealed adaptation to novel environment.	Hughes <i>et al.</i> , 2007
		EE carried with <i>E. coli</i> on amino-acid containing environments for 2000 generations showed revealed rapid evolution of auxotrophic genotypes.	D'Souza and Kost, 2016
	Single clone	EE carried with <i>P. salmonis</i> on a cell-free culture medium for 200 serial transfer revealed genomic and transcriptomic rearrangements.	Valenzuela-Miranda <i>et al.</i> , 2020
		EE carried for 60 serial transfer with <i>S. cerevisiae</i> on rich media showed a transition to multicellularity life form.	Ratcliff <i>et al.</i> , 2012
		EE carried with <i>S. enterica</i> on rich media for up to 270 serial transfer revealed genome reduction.	Nilsson <i>et al.</i> , 2005
	Mutant strain	EE carried for 32 serial transfer with an <i>M. xanthus</i> strain unable to make a pili (necessary for socially dependent swarming) on rich media showed a new mechanistic basis for cooperative swarming.	Velicer and Yu, 2003
Complex media	Single clone	EE carried with <i>E. coli</i> with serial transfers every 4 days in order to go through all the cell growth phases and for 300 generations showed adaptation to the changing environment.	Kram <i>et al.</i> , 2017
		EE carried with <i>P. fluorescens</i> in 15 environments of different complexity (1 to 8 different carbon sources) for 900 generations showed adaptation to multiple carbon sources.	Barrett <i>et al.</i> , 2005
		EE carried with <i>E. coli</i> for at least 10 generations on an antibiotic supplemented media with increasing concentrations showed that different mutations are selected depending on the size of the inoculum.	Garoff <i>et al.</i> , 2020
		EE carried with a five species communities isolated from beech roots and made evolving on a beech leaf-extract media for 70 generations showed adaptation to novel environment.	Lawrence <i>et al.</i> , 2012
	Synthetic community	EE carried with 33 bacterial strains on a complex liquid media with serial transfer every 96h and for 12 serial transfers showed the importance of biotic interaction network. EE carried with a 62 bacterial strain on antibiotic supplemented media for 10 serial transfers revealed the importance of spatial structure for gene transfer.	Cairns <i>et al.</i> , 2018a Cairns <i>et al.</i> , 2018b
Eukaryotic host	Single clone	EE carried with a single clone of <i>R. solanaceum</i> on three original and two distant hosts for 300 generations.	Guidot <i>et al.</i> , 2014
	Engineered strain	EE realized with a clone of <i>R. solanaceum</i> carrying a symbiotic plasmid from <i>C. taiwanensis</i> for 16 serial transfers showing the adaptation of a non-symbiotic bacterium into a legume symbiont.	Doin de Moura <i>et al.</i> , 2020,
	Strain + bacteriophages	EE carried with <i>P. syringae</i> on its natural host and on a distant one in presence or absence of bacteriophages for 4 serial transfers revealed pathogen adaptation.	Meaden and Koskella, 2017
Mesocosm	Natural microbial community	EE carried for 16 serial transfers leading to the selection of soil microbiomes inducing a high or low plant biomass on <i>A. thaliana</i> .	Swenson <i>et al.</i> , 2000
		EE carried for 10 serial transfers leading to the selection of soil microbiomes inducing an early or late flowering time on <i>A. thaliana</i> and <i>Brassica rapa</i> .	Panke-Buisse <i>et al.</i> , 2015

Also, genomic and molecular data are available for many bacterial species, as well as techniques for their precise genetic analysis and manipulation. Moreover, bacteria are easy to track (*i.e.* isolation, numeration) and store. A regular collection and conservation of samples during the experiment (for example at each cycle) and further genetic and/or phenotypic analyses allow to decipher the evolution process (Marchetti *et al.*, 2017). This storage capacity is a compelling advantage as in EE assays, two properties are widely studied: size and mean fitness. Cell size is a morphological trait that reflects many of the functional properties of an organism and fitness is the most important property of any organism according to the theory of evolution (Lenski and Travisano, 1994). Fitness of the evolved organism is expressed relative to its ancestor. The two forms are allowed to compete and the outcome of this competition is accounted for in Petri dishes and the relative fitness of evolved clones, defined as the ratio of the growth rates of the evolved and ancestral forms, is calculated. Despite their numerous advantages, bacteria reproduce asexually, so the different genotypes that rise in the population do not recombine. Take for example a clonal population (founded from a single clone) in which several mutations arise; among the different mutated clones, only one will reach fixation within the population by outcompeting its fellow mutated clones and their common ancestral clone. This phenomenon is known as clonal interference and influences the dynamics of evolution as clones with beneficial mutations will interfere with each other's spread in the population (Gerrish and Lenski, 1998). Experimental evolution using bacteria can help to unlock the mysteries of the adaptive process as this set up shows the kinetics of evolution, allowing to follow directly the evolutionary mechanisms by observing evolution in real time for ten, hundreds or even thousands of generations. Thus, unlike carrying out genomic analysis on existing natural organisms and then interpreting their evolution, EE makes it possible to transform evolutionary genetics into a prospective undertaking, and to decipher the genetic bases of adaptation. For a more thorough review on the genetic mechanisms underlying adaptation please refer to Van den Bergh *et al.* (2018).

## **1. Plant-associated microorganisms in experimental evolution**

Many soil microorganisms are capable of colonizing the plant rhizosphere -the volume of soil under the influence of root activities (Hiltner, 1904)- and plant exerts a selection by secreting a wide range of compounds known as plant exudates (Figure 2A, 2B) (Haichar *et al.*, 2014). Root exudates represent up to 21% of photosynthetically fixed carbon (Derrien *et al.*, 2004). Thus, the rhizosphere is a medium rich in nutrients, stimulating the development of many microbial populations, making it a hotspot for horizontal gene transfer (van Elsas *et al.*, 2003). Whole genome sequences have revealed that horizontal gene transfer has had a great impact on bacterial evolution (Wisniewski-Dyé *et al.*, 2011; Lassalle *et al.*, 2017). Horizontal gene transfers, mutations and gene arrangements will result in population genotype and phenotype changes, in the modification of the structure and function of the community and hence, will have profound

effects on the long-term evolution of communities. Selected fixed mutations will induce greater fitness in the given niche, favouring biotic interactions between community members in the rhizosphere such as competition (Figure 2C). As previously stated, evolution is a stochastic process driven by random mutations and fixation of the ones with a beneficial outcome in a given environmental context (see Box1). New genes will only be fixed if its benefit is greater than its cost. Rhizosphere populations are in constant competition for nutrients, facing energy limitations, which makes the expression of any additional trait costly. For plant-associated microorganisms, the newly acquired trait has to be advantageous against other microorganisms competing for the same resources or has to drive adaptation to new ecological niches (Figure 2D) (Wisniewski-Dyé *et al.*, 2011; Lassalle *et al.*, 2017).



**Figure 2: Plant selection of beneficial bacteria.** (A) Plants mainly select their microbiome via their exudates by stimulating the growth of many bacterial populations. (B) Some of the selected bacteria, such as PGPR (Plant-Growth Promoting Rhizobacteria), have a beneficial effect on the plant (protection against plant pathogens, plant hormones production, *etc.*). (C) Being a highly competitive environment, an established community is always at risk of being invaded by more performant microorganisms. Hence, bacteria have to constantly adapt in order to remain in the current niche (production of antimicrobial metabolites, red arrows) while still being advantageous for the plant to keep. (D) This rich environment is also known for being a HGT (horizontal gene transfer) hotspot; DNA exchanges (purple outlines blue bacteria) have already been proved to have an impact in bacterial evolution and in the structure of communities. Still, new DNA will only be fixed if the benefit is greater than the cost.

Plant exudates mostly select plant beneficial microorganisms but because of the wide variety of compounds, plant pathogens can also be attracted. Plant pathogens present many strategies to attack plants and have largely driven attention as understanding their interaction with plants may help restrain them (Griffin *et al.*, 2004). Here we review how EE assays have helped to better understand plant-microorganism's interactions, first for plant pathogens and secondly for plant-beneficial bacteria.

## 1.1 Plant pathogens

The EE approach has been widely used to study the ecology of plant pathogens. Indeed, the adaptive history and evolutionary mechanisms involved during their relationships with host must be unravelled in order to better understand their modes of action on the plant and, in an applied point of view, to potentially discover new ways to inhibit their pathogenicity (Griffin *et al.*, 2004). The genetic bases of pathogen adaptation are critical to understand disease emergence and the acquisition of novel traits by the pathogen when colonizing its plant host or different hosts (Toft and Andersson, 2010; Barroso-Batista *et al.*, 2014). Indeed, the dynamics of adaptation are influenced by the pathogen-host interaction (Gandon *et al.*, 2013), especially when the pathogen is interacting with a new host. For example, pathogens must constantly face plant defences; this translates into a real arms race as the pathogen evolves. The production and regulation of virulence factors are associated with high costs (energy or fitness). EE has highlighted in these situations the appearance of cheaters in pathogen populations. These cheaters are pathogens adapted to take advantage of the production of virulence factors by the rest of the population without producing them themselves (this topic will be discussed later on, Tannières *et al.*, 2017). Here, a focus is made on *in planta* EE experiments done with bacterial plant pathogens (Table 1), but others studies have also focused on fungal (Gilbert and Parker, 2010) or virus plant pathogens (Bedhomme *et al.*, 2012). When colonizing a distant host plant, less adapted microorganisms might see their fitness increase rapidly at first and then their fitness tends to decelerate until reaching an optimum (Elena and Lenski 2003). Fisher's geometrical model of adaptation (Fisher, 1930; Tenaillon, 2014) proposed an explanation to this phenomenon: (1) the pool of available beneficial mutations in a novel host is bigger than in an original host and (2) mutations offering greater advantages will be fixed first, followed by a decrease of advantageous mutations as the population becomes more adapted to the new environment over time.

In an attempt to experimentally study the genetic basis of adaptation to new host, Guidot *et al.*, (2014) carried an EE assays with *Ralstonia solanacearum*, a plant pathogen with a continuously broadening host spectrum. A single clone of the model strain GMI1000 was inoculated on three native host plants (tomato, eggplant and pelargonium) where the pathogen causes disease and two distant plants (cabbage and bean) where it grows asymptotically. The pathogen was transferred serially to the same plant line in order to maintain the pathogen on the same host for

300 bacterial generations (26 serial passages). Although evolved strains showed an increase in fitness (pathogenicity) in both host and non-host plants, this increase was greater in the non-host plant (cabbage and beans), to which the pathogen was originally not adapted. Whole-genome analysis and comparison of the ancestral GMI1000 clone with 9 evolved strains (3 from the tomato host and 6 from the bean) highlighted that only few genes contribute to plant adaptation. In particular, the transcriptional regulator encoding gene *efpR* was identified as important for the adaptation of *R. solanacearum* to bean as 3 bean evolved clones presented different allelic version of this gene. EfpR was thereafter identified as a global catabolic repressor and regulator of virulence traits (Perrier *et al.*, 2016). Further analysis revealed that the mutations in *efpR* can be fixed after 26 serial passages. In a second experiment, strains harbouring the evolved *efpR* alleles were co-inoculated with the wild-type clone on bean. Results showed that after only 3 serial passages, evolve clones had a greater fitness compared to the wild type clone during co-infection of bean plants (Guidot *et al.*, 2014). It was found that the *efpR* mutations allow an enhanced metabolic versatility and adaptation to host vascular tissues of new hosts (Perrier *et al.*, 2016). Overall this results are consistent with Fisher's geometril model, as fitness increased in original and distant host but achieved greater fitness in distant hosts. This study also highlighted the importance of combining EE with whole-genome sequencing in order to unravel the genetic basis of pathogen adaptation.

Another EE was setup to study how the adaptation of a pathogen to a specific host affects the growth and virulence of this pathogen evolving on another host. To this end, *Pseudomonas syringae*, a well-known plant pathogen, has been used to infect a distant host (*Arabidopsis*) or a native host (tomato) over several infection cycles in the presence or absence of phages. Bacteriophages impose evolutionary trade-offs on the bacteria they infect, e.g. via the modification of bacterial surface receptors that subsequently may impact *P. syringae* interactions with distinct hosts. Selection on *Arabidopsis* through serial passages led to greater pathogen growth in both hosts than selection on tomato plants do. Phages had nevertheless little direct impact on the bacterial populations during this experiment. These results point out that a given association between a host plant and its pathogen may affect the growth rate of the pathogen on the plants it will subsequently infect (Meaden and Koskella, 2017).

## 1.2 Non-pathogenic plant-root-associated microorganisms

EE studies are not only of interest for understanding the adaptations of pathogenic microorganisms to new hosts. For example, EE have been carried out to decipher the adaptation of a non-symbiotic bacterium into a legume symbiont. To do this, *R. solanacearum*, hosting the *Cupriavidus taiwanensis* symbiotic plasmid of *Mimosa pudica*, was repeatedly inoculated on the plant host (see for a review Doin de Moura *et al.*, 2020) (Table 1). After a series of *Mimosa pudica* infection cycles, an evolved symbiote well adapted to the host (without having acquired the ability to fix nitrogen) was obtained. Sequencing of intermediate and final forms revealed that the



symbiosis-adaptive mutations happened in global regulatory proteins, leading to a reworking of the regulatory systems in *R. solanaceum*. These adaptive mutations included the inactivation of the type III secretion systems (the main virulence factor of *R. solanaceum* (Genin and Denny, 2012) and modifications on the expression of *efpR* gene. Genomic analyses revealed that is not only the *efpR* gene that is mutated but also its upstream region. Altogether these mutations led to metabolic and transcriptomic changes (as seen by Perrier *et al.*, 2016) as well as nodulation and nodule cell infection by these new symbionts (Capela *et al.*, 2017; Doin de Moura *et al.*, 2020). These findings highlight how the modification of regulatory systems controlling virulence can determine the ecological function of microorganism. In this pseudo-symbiosis, plant immunity and nutrition of both the plant and the bacterial partner have a great impact on the outcome of the symbiosis (Guan *et al.*, 2013). However, the evolutionary forces that drive selection and intracellular accommodation are still poorly known and EE can actually help identify the evolutionary pathways that drive the evolution of symbiotic functions in rhizobia.

Other plant-beneficial bacteria are implied in less specific interactions with plant. These bacteria known as Plant Growth-Promoting Rhizobacteria (PGPR) are colonizing the plant rhizosphere. While EE assays carried out *in planta* with plant-inoculated PGPR are uncommon, numerous studies have taken an interest in plant-associated microbiota, illustrating the ability of the microorganisms to positively influence plant health and developmental traits such as disease resistance, abiotic stress tolerance and growth (Mendes *et al.*, 2011). Plants and their microbiota should be viewed and analysed as a holobiont, namely as a meta-organism unit where all partners are influencing each other, evolve or even co-evolve together (Vandenkoornhuyse *et al.*, 2015; Theis *et al.*, 2016; Sánchez-Cañizares *et al.*, 2017). Over time, new forms of interactions can be observed between the partners, and these interactions can have an impact on plant development (Theis *et al.*, 2016). For instance, an artificial mesocosm selection carried on of *A. thaliana* showed that plant biomass levels can be modified by plant-associated microorganisms. Several *A. thaliana* were inoculated with non-sterile soil batch and let grow for 35 days, which corresponds to one mesocosm cycle. After each cycle, the plants presenting the highest and lowest biomasses were selected. The soil from these plants was retrieved and used to inoculate the next batch of plants. The experiment was carried for 16 mesocosm cycles and both artificial selections (high or low biomass) were analysed. *Arabidopsis* inoculated with the soil community from “high biomass” plants presented indeed higher biomasses than those inoculated with the soil community from “low biomass” plants. After 13 cycles, a soil analysis revealed different soil characteristics for high and low biomasses, reflecting differences in the biotic component (Swenson *et al.*, 2000).

Plant flowering is another trait that can be modulated by plant-associated microorganisms and the contributing microbiota's populations can be experimentally selected and enriched from one mesocosm cycle to the next. In a similar way to Swenson and collaborators' experiment, several microcosms of *A. thaliana* Col0 were created with seeds placed on sterile soil. In this experiment,

mesocosms were selected for either an early or late flowering phenotype. Plants were harvested and the soil retrieved as soon as all the plants of the mesocosm flowered, therefore, the duration of the mesocosm cycle depended on the flowering time. After ten cycles on *A. thaliana* Col0, microorganisms were retrieved and inoculated on other *A. thaliana* genotypes and on a related crucifer, *Brassica rapa*. Plant-associated microorganisms induced on these plants the same phenotype (early flowering) as the one induced on the genotype (Col0) used for selecting the evolved community (Panke-Buisse *et al.*, 2015). Rather than imposing no selective pressure on the plant host, in these studies, researchers selected a particular plant trait (biomass or flowering) to select plant-associated microorganisms contributing to this plant trait (Table 1). To our knowledge, none of the EE studies have yet been done without taking into consideration improved plant health, growth or development phenotype as a selective pressure outcome of the evolutive process. One can suggest that letting the plant holobiont evolve naturally, without any additional selective pressure, might allow to get a larger diversity of evolutionary routes in the plant-associated microbiota and to better understand the holobiont's mechanistic functioning (Vandenkoornhuyse *et al.*, 2015; Hassani *et al.*, 2018).

## **2. Experimental evolution to decipher the adaptive evolution of natural communities**

### **2.1 Adaptive responses and parallel or convergent evolution**

EE studies provide essential knowledge for adaptive changes, but an important question arises from these studies: are they reproducible? Do same genotypic and phenotypic changes arise in the evolved microbial populations? In order to better understand evolutionary adaptation, genotype-to-phenotype correlations need to be more fully deciphered. This can be done by focusing on parallel and convergent changes. Parallel evolution is defined as the independent evolution of similar phenotypic traits in lineages closely related to each other and involving changes in orthologous genes, whereas convergent evolution concerns non-related phylogenetic lineages and changes in non-homologous genes (Bai *et al.*, 2015; Pickersgill, 2018). Identical or very similar changes reaching high frequency and even fixation in independent lineages evolving under identical conditions are hallmarks of natural selection and adaptation. Hence, convergent genetic and/or functional traits emerging in independent populations are thus good indicators of adaptive evolution (Pelosi *et al.*, 2006).

In accordance with that, many studies have shown that a single clonal ancestor can give rise to independently evolved populations sharing similar traits under the same environmental conditions (Schluter *et al.*, 2004; Colosimo, 2005; Zhang *et al.*, 2005). Lenski's experimental evolution study is composed of 12 replicate populations of *E. coli* B, which are currently still evolving on a glucose-limited minimal salts medium (Lenski *et al.*, 1991) (Table 1). DNA microarray analyses of 2 evolved



populations showed similar changes in the transcription of 59 genes after 20,000 generations; the genetic bases of these changes were investigated and relevant parallel mutations were found on many of the independently evolved populations (Cooper *et al.*, 2003). Whole-genome comparisons revealed that the vast majority of mutations that reach high frequencies have an adaptive effect (Wichman, 1999). Evolution acts on biological functions; in simple organisms such as viruses, these functions are encoded by single genes, whereas, in higher forms of life, more complex regulatory networks are involved. In bacteria and yeast, complex functions are often related to modules of genes (Hartwell *et al.*, 1999). Reflecting this, parallel evolution will not always mean changes in the same genes but rather similar changes in related gene modules. Herring *et al.* (2006) explored the parallel changes in metabolic and regulatory networks that arose in five *E. coli* populations that evolved separately. They proved, via a mutant approach and whole-genome resequencing of five clones of *E. coli*, that 13 different spontaneous mutations were responsible for improved fitness during adaptation to a glycerol synthetic growth medium (Herring *et al.*, 2006). This study provides a clear example of how changes might propagate through complex networks and how different changes can have similar phenotypic effects (Table 1). Dissimilar genetic changes can lead to parallel phenotypes, meaning that functional connections within and between genetic modules can be established by pairing experimental evolution to whole genome sequencing (Segrè *et al.*, 2006; Yeh *et al.*, 2006).

However, clonal populations, evolving in an identical environment, may reach different fitness rates and may diverge (Lenski *et al.*, 1991). This divergence reveals that there may be more than one adaptive strategies under the same environmental conditions (Elena and Lenski, 2003). On the other hand, we could expect that experiments ran with genetically different organisms led to different fitness. Travisiano *et al.*, (1995) evaluated fitness on maltose for several *E. coli* lines that had evolved in glucose for 2000 generations. Lines that started with the lower fitness in maltose improved faster, but all the lines tend eventually to converge to a similar fitness on maltose. Of course, the adaptive responses, fitness and evolutionary routes of bacterial populations will depend on the different partners that evolve together under the tested environmental conditions.

## 2.2 Co-evolution between populations in a community

Microorganisms are indeed in constant interaction with other organisms in their environment, so that any population in nature evolves with the other sympatric microbial populations. Co-evolution can be defined as the process of reciprocal adaptation by interacting species (Buckling and Rainey, 2002). Evolving together creates adaptive coevolutionary dependencies in a changing environment: “it takes all the running you can do to stay in the same place”, meaning that species have to constantly adapt to the other evolving species in order to survive / maintain their fitness (*i.e.* “Red Queen hypothesis”; Van Valen, 1977; Strotz *et al.*, 2018). There is abundant evidence

that coevolution drives fast evolution between species with strong ecological interaction (Brockhurst *et al.*, 2003; Guan *et al.*, 2013).

Initially, coevolution experiments mainly focused on antagonistic microorganisms, in part due to the pressure of the reciprocal selection that leads to rapid evolution and diversification of the partners. Nowadays coevolution experiments focus on mutualistic interactions, because these interactions are widespread in nature (Zbinden *et al.*, 2008). Hillesland and Stahl (2010) co-cultured two microorganisms on lactate, forcing them to collaborate in order to produce an energy-yielding reaction. Populations largely fluctuated for the first 300 generations before stabilizing. The coevolved populations had greater fitness on lactate compared to the ancestral populations (Hillesland and Stahl, 2010). The existence of co-evolution between plants and soil micro- and macro-organisms has been suggested being mediated by the production and sensing of signal molecules emitted either by the soil organisms and the plant (Blouin, 2018). This complex dialog has not been fully deciphered because these signal molecules operate at very low concentrations in a complex matrix that is the rhizosphere. In connection with this rhizosphere molecular signalling network, one can suggest that free-living soil organisms do co-evolve with plants in a more “diffuse” way than plant physically-associated organisms (Janzen, 1980). Coevolution experiments are often carried out with only two interactions species, whereas, in nature, interaction networks are more complex. Networks of species interactions shape the evolution and stability of the community as a whole and the observed effects on the plant could be the outcome of complex interactions between free-living organisms, plants and their plant-associated microorganisms in the rhizosphere (Janzen, 1980). Deciphering the principles that underpin ecological and evolutionary properties of microbial communities can allow to build predictive models of ecological dynamics of microbial communities.

### **3. Towards a full consideration of microbial communities in experimental evolution studies**

#### **3.1 From single microorganisms to complex communities**

When it comes to microbial adaptation, a major part of studies has considered one species, or at most two-interacting species. These studies have taken into account only genetic diversity within species rather than species diversity within communities. Nearly all biogeochemical cycles are the outcome of a joint effort by microorganisms with different functional roles (De Roy *et al.*, 2014). There are two important aspects of microbial communities: their taxonomic structure and how they function (Little *et al.*, 2008). Structure is defined as the diversity and abundance of each individual population within the community and function as the community’s behaviour and activities, and how the community interacts with its environment. The mechanisms that dictate interactions among microorganisms are largely responsible for the properties of the community as a whole

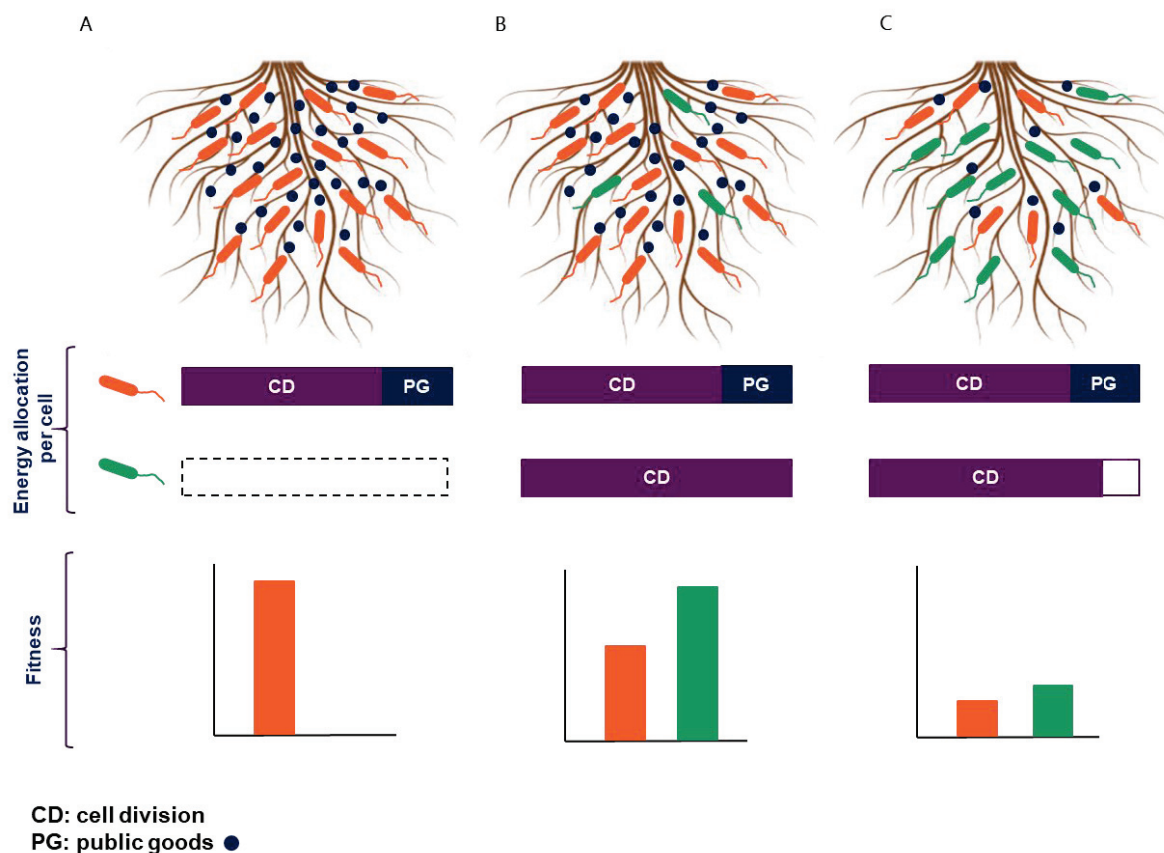
and for its evolution. Species interactions can influence how species evolve and adapt to environmental changes. For example, within a microbial community, some species can use the waste products generated by others (Lawrence *et al.*, 2012). These interactions can become so relevant that a given microorganism can present a lower growth rate or even may not grow when cultivated alone, compared to when cultivated within a community (Smith and Schuster, 2019).

In EE studies done on single strains, one species has to adapt to a new environment whereas in microbial communities' EE experiments, species have to adapt not only to the environment itself but also to the biotic environment dynamically shaped by the other species. Within a microbial community many interactions take place, these interactions can have positive or negative outcomes. Sharing the same ecological niche implies sharing some resources and this inevitably promotes competition within the group. Synthetic community approaches aim to mimic natural microbiota (Vorholt *et al.*, 2017) and can help to better understand the functioning of biotic interaction network within natural communities. Recently, Cairns and co-workers (2018a) serially transferred a synthetic community of 33 bacterial strains on a complex liquid media. Over half of the strains from different species were lost in 16 days, after which the community was relatively undisturbed until the end of the experiment (48 days). Within the evolved community, 14 strains co-existed with the predominance of three strains. The synthetic community shared high diversity at different levels (*i.e.* taxonomic, metabolic and functional levels) (Table 1). A version of this community was also used to track the mobility of antibiotic resistance genes; the authors evidenced the importance of spatial structuring of environment in gene transfer (Cairns *et al.*, 2018b).

Competition between species results in a modification of the abundance of the species within the community and subsequently engenders the evolutionary adaptation of the community. In the evolved community, out-competing species have smaller numbers and slower uptake of beneficial mutations (Johansson, 2008; Mazancourt *et al.*, 2008). Moreover, some species are becoming more fitted to a given niche than others and they will have an advantage allowing them to develop at the expense of the less pre-adapted species. Adaption to other competitive species may imply the production of antimicrobial metabolites to be protected against the others, but a trade-off can arise due to the substantial energy-cost of their production (Yoshida *et al.*, 2003).

Thus, when adapting to a new ecological niche, metabolite production by microbial populations can be modified and some compounds can be costly to produce for one individual but beneficial for all the members of the community. Microbes have thus developed multicellular cooperative behaviours, like biofilm formation and quorum sensing, along with nutrition acquisition, and the outcome of these interactions are referred as public goods (Smith and Schuster, 2019). Public goods take many forms from large proteins to small metabolites and can be actively or passively secreted, but one of their main features is that their benefit increases with population density

(Figure 3A). Unfortunately, cooperative behaviour is not an evolutionarily stable strategy because cheaters may appear and invade the community. Cheaters are non-cooperative individuals that benefit from the public goods without producing them. By not paying the cost of their production, cheaters could have more energy allocated to their growth, and therefore their relative fitness increases (Figure 3B). The underlying mechanisms of cheaters' loss of function often involved a selective gene loss to optimize their adaptation to the environment (*i.e.* the “Black Queen Hypothesis”, Mas *et al.*, 2016). This evolutionary strategy has great impact on long term interactions because the fitness of cheaters depends on the public goods provided by the cooperating microbes (Figure 3C).



**Figure 3: Black Queen Hypothesis in the rhizosphere.** (A) When colonizing a new niche, a group of microorganisms can present multicellular cooperative behaviors leading to the production of public goods, accessible for all the members of the community. (B) However, cooperative behavior is not a stable strategy and cheaters (green bacteria) may arise, these bacteria can allocate more energy to their growth while they still benefit from the public goods, therefore, their fitness is increased. (C) If cheaters invade the community and the production of public goods cannot be sustained by the few remaining cooperative bacteria, fitness of cheaters drops as it depends on the public goods provided by the cooperative bacteria.

### 3.2 From simple to somewhat more complex environments

The stability of the community can only be maintained if the proportion of cheaters remains low within the community. Ultimately, cheaters will reduce the effective population size of the cooperating microbes, reducing the rate of public goods production, but also the rate at which beneficial mutations would arise in the community and the species would be able to adapt to a novel environment. The evolution of cooperation is a tricky issue since all microorganisms will tend, according to the natural selection theory, to maximize their own fitness. However cooperative behaviour is ubiquitous even though selfish interests have always been a source of conflict (Sachs and Wilcox, 2006; Burt *et al.*, 2009). A deeper insight into biotic interactions within microbial communities is thus crucial to understand adaptation to novel environments. It can help in deciphering the dynamics of living systems and to predict responses to anthropogenic changes in the natural environment (Winder and Schindler, 2004; Davis *et al.*, 2005; Berg *et al.*, 2010).

In addition to the importance of species interactions within microbial communities, the environment can also have an impact on species interactions. In the past, EE studies were mostly realized in simple environments, composed of only one essential resource (such as carbon or nitrogen), available to all individuals, and whose concentration controls the bacterial growth rate (Hillesland and Stahl, 2010; Yu *et al.*, 2017). This resource would be consumed progressively, changing the environment, and boosting individuals able to grow on lower concentrations. At last, only the most frugal individuals would remain until the resource ran out. Individuals can evolve either as generalists (able to consume a large range of the available resources) or as specialists (able to grow in a short range of the available resources). In complex environments that contain several resources, cell populations may become wide specialists able to consume various carbon substrates. The importance of environmental complexity and evolution of niche width have been studied by Barrett *et al.*, (2005), who compared the evolution of more than one hundred replicate lines of the bacterium *Pseudomonas fluorescens*, over ~900 generations in 15 environments of different complexity (Table 1). To do this, authors used a chemically defined growth medium but containing between one and eight distinct carbon substrates and compared the genetic evolution of *P. fluorescens* lineages in simple and complex environmental growth media. They showed that genotypes from populations selected in complex environments evolved with greater fitness on all carbon sources than those from populations selected in simple environments. In addition, there was a higher genetic fitness variance within populations selected in complex environments. Indeed, in simple environment, lineages became adapted to a single carbon substrate while in complex media lineages were able to adapt to several substrates simultaneously without any appreciable loss of functions. These results suggest that evolution in complex environments will lead to the emergence of imperfect generalist overlapping lines adapted to a certain range of substrates but not to all (Barrett *et al.*, 2005).

However, in nature, most communities grow under conditions where many substrates are simultaneously available. In environments containing a great variety of substrates, microbial growth can be supported by any of these substrates. Greater fitness in this environment will depend on the temporal and spatial distribution of the available substrates. For example, the rhizosphere is a rich and complex environment where many terminal electron acceptors can be found. Some microorganisms are capable of both aerobic and anaerobic respiration (e.g. *deltaproteobacteria*). The latter will first use the most energetic substrates but when the corresponding substrates run out, they will be able to use less energetic substrates to sustain growth (Lecomte *et al.*, 2018).

In the laboratory, because of the complexity of evolutionary processes, most of the EE assays have been carried out in a constant environment, using chemostats or continuous culture systems, so the growth conditions would remain the same for the whole experiment (Treves *et al.*, 1998; Maharjan *et al.*, 2006; Gresham and Hong, 2015). Recently, Westphal *et al.* (2018) used a batch culture system to study the impact of four changing environments on the adaptation of *E. coli*. In these systems there is no addition of nutrients, therefore the cultures experience fluctuation in nutrient availability. Therefore, as nutrients are consumed, waste products are released and less energetically favourable metabolisms become important for survival. Their data revealed that different environments may select different mutations; this emphasizes the importance of performing experimental evolution in complex and ever-changing environments. To take these findings one step further, Kram (2017) and his team carried out an EE assay with *E. coli* in a complex and changing environment (*i.e.* in a rich medium LB and with serial passages every four days). This scheme allows cells to go through all phases of growth and to adapt to different stresses (nutrient limitation, oxidative stress and pH variation). This experiment showed that after only 30 generations, evolved populations presented changes in growth rates but also adaptive mutations allowing the cells to cope with the varying stresses arising during the culture (Table 1). They also evidenced parallel changes in evolved populations (Kram *et al.*, 2017).

## CONCLUDING REMARKS

Under environmental conditions, individual microbial species evolve in the context of communities constituting of many species co-evolving together. Understanding the selection pressures driving species interactions will bring an evolutionary understanding of microbial communities. Although EE assays have historically mostly focused on fitness outcomes, sometimes combined with the analyses of some genetic or functional markers, whole-genome sequencing has been largely used during the last decades (Barrick *et al.*, 2010; Dettman *et al.*, 2012), even in the case of eukaryotes, so that it is now possible to reveal the genetic basis separating an evolved clone from its original ancestor strain (Braslavsky *et al.*, 2003; Shendure, 2005; Segrè *et al.*, 2006). The application of whole-genome sequencing to the study of microbial evolution allows to reveal the complex



functional networks of mutations that underlie adaptation. Nevertheless, one technological barrier still remains: whole genome sequencing of many clones from one population together is still technologically impossible (Schlötterer *et al.*, 2015). Overriding this problem will shed better light on potential clonal interference phenomenon (Gerrish and Lenski, 1998). The next step in microbial evolution studies would be taking them outside of the laboratory conditions and adding the selection pressures imposed by the environment. Future studies may be able to use genome-sequencing technologies to answer some crucial questions at the interface between cellular processes and the observed evolutionary responses. A larger concept also arises: seeing the community as a whole, like the entire unit upon which selection acts. Natural selection and random drift operate on populations leading to the fixation of beneficial mutations, altering the genetic composition of populations and indirectly affecting species interactions that dictate community ecology. The intersection of ecology and evolution is key to understand microbial communities.

## ACKNOWLEDGEMENT

Beatriz Manriquez was supported by a Ph.D. fellowship from the French Ministère de l'Education Nationale, de l'Enseignement Supérieur et de la Recherche. The authors acknowledge funding from the French National Programme EC2CO 'Structuring initiative Continental and Coastal Ecosphere' - Microbien 2018-2019.

## CONFLICT OF INTEREST STATEMENT

The authors declare they have no conflict of interest.

## REFERENCES

- Adams, J., and Rosenzweig, F. (2014). Experimental microbial evolution: History and conceptual underpinnings. *Genomics* 104, 393–398. doi: 10.1016/j.ygeno.2014.10.004.
- Bai, Y., Müller, D. B., Srinivas, G., Garrido-Oter, R., Potthoff, E., Rott, M., *et al.*, (2015). Functional overlap of the *Arabidopsis* leaf and root microbiota. *Nature* 528, 364–369. doi:10.1038/nature16192.
- Bailey, S. F., and Bataillon, T. (2016). Can the experimental evolution programme help us elucidate the genetic basis of adaptation in nature? *Molecular Ecology* 25, 203–218. doi:10.1111/mec.13378.
- Barrett, R. D. H., MacLean, R. C., and Bell, G. (2005). Experimental evolution of *Pseudomonas fluorescens* in simple and complex environments. *The American Naturalist* 166, 470–480. doi:10.1086/444440.
- Barrick, J. E., Kauth, M. R., Strelisoff, C. C., and Lenski, R. E. (2010). *Escherichia coli rpoB* mutants have increased evolvability in proportion to their fitness defects. *Mol Biol Evol* 27, 1338–1347. doi:10.1093/molbev/msq024.
- Barroso-Batista, J., Sousa, A., Lourenço, M., Bergman, M.-L., Sobral, D., Demengeot, J., *et al.*, (2014). The first steps of adaptation of *Escherichia coli* to the gut are dominated by soft sweeps. *PLoS Genetics* 10, e1004182. doi: 10.1371/journal.pgen.1004182.
- Battistuzzi, F. U., and Hedges, S. B. (2009). A major clade of prokaryotes with ancient adaptations to life on land. *Molecular Biology and Evolution* 26, 335–343. doi:10.1093/molbev/msn247.
- Baym, M., Lieberman, T. D., Kelsic, E. D., Chait, R., Gross, R., Yelin, I., *et al.*, (2016). Spatiotemporal microbial evolution on antibiotic landscapes. *Science* 353, 1147–1151. doi:10.1126/science.aag0822.
- Bedhomme, S., Lafforgue, G., and Elena, S. F. (2012). Multihost experimental evolution of a

plant RNA virus reveals local adaptation and host-specific mutations. *Mol Biol Evol* 29, 1481–1492. doi:10.1093/molbev/msr314.

Berg, M. P., Kiers, E. T., Driessen, G., van der Heijden, M., Kooi, B. W., Kuenen, F., *et al.*, (2010). Adapt or disperse: understanding species persistence in a changing world. *Global Change Biology* 16, 587–598. doi:10.1111/j.1365-2486.2009.02014.x.

Blouin, M. (2018). Chemical communication: an evidence for co-evolution between plants and soil organisms. *Applied Soil Ecology* 123, 409–415. doi:10.1016/j.apsoil.2017.10.028.

Braslavsky, I., Hebert, B., Kartalov, E., and Quake, S. R. (2003). Sequence information can be obtained from single DNA molecules. *Proc. Natl. Acad. Sci.* 100, 3960–3964. doi:10.1073/pnas.0230489100.

Brockhurst, M. A., and Koskella, B. (2013). Experimental coevolution of species interactions. *Trends in Ecology & Evolution* 28, 367–375. doi:10.1016/j.tree.2013.02.009.

Brockhurst, M. A., Morgan, A. D., Rainey, P. B., and Buckling, A. (2003). Population mixing accelerates coevolution: population mixing accelerates coevolution. *Ecology Letters* 6, 975–979. doi:10.1046/j.1461-0248.2003.00531.x.

Buckling, A., and Rainey, P. B. (2002). Antagonistic coevolution between a bacterium and a bacteriophage. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 269, 931–936. doi:10.1098/rspb.2001.1945.

Burt, A., and Trivers, R. (2009). Genes in conflict: The biology of selfish genetic elements. *Harvard University Press*. doi: 10.4159/9780674029118.

Cairns, J., Jokela, R., Hultman, J., Tamminen, M., Virta, M., and Hiltunen, T. (2018a). Construction and characterization of synthetic bacterial community for experimental ecology and evolution. *Front. Genet.* 9. doi:10.3389/fgene.2018.00312.

Cairns, J., Ruokolainen, L., Hultman, J., Tamminen, M., Virta, M., and Hiltunen, T. (2018b). Ecology determines how low antibiotic concentration impacts community composition and horizontal transfer of resistance genes. *Communications Biology* 1. doi:10.1038/s42003-018-0041-7.

Capela, D., Marchetti, M., Clérissi, C., Perrier, A., Guetta, D., Gris, C., *et al.*, (2017). Recruitment of a lineage-specific virulence regulatory pathway promotes intracellular infection by a plant pathogen experimentally evolved into a legume

symbiont. *Mol Biol Evol* 34, 2503–2521. doi:10.1093/molbev/msx165.

Colosimo, P. F. (2005). Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. *Science* 307, 1928–1933. doi:10.1126/science.1107239.

Cooper, T. F., Rozen, D. E., and Lenski, R. E. (2003). Parallel changes in gene expression after 20,000 generations of evolution in *Escherichia coli*. *Proc. Natl. Acad. Sci.* 100, 1072–1077. doi:10.1073/pnas.0334340100.

Cooper, V. S., and Lenski, R. E. (2000). The population genetics of ecological specialization in evolving *Escherichia coli* populations. *Nature* 407, 736–739. doi:10.1038/35037572.

Davis, M. B., Shaw, R. G., and Etterson, J. R. (2005). Evolutionary responses to changing climate. *Ecology* 86, 1704–1714. doi:10.1890/03-0788.

De Roy, K., Marzorati, M., Van den Abbeele, P., Van de Wiele, T., and Boon, N. (2014). Synthetic microbial ecosystems: an exciting tool to understand and apply microbial communities: synthetic microbial ecosystems. *Environ Microbiol* 16, 1472–1481. doi:10.1111/1462-2920.12343.

de Visser, J. A. G., and Lenski, R. E. (2002). Long-term experimental evolution in *Escherichia coli*. Rejection of non-transitive interactions as cause of declining rate of adaptation. *BMC Evol. Biol.* 2, 19. doi:10.1186/1471-2148-2-19.

Derrien, D., Marol, C., and Balesdent, J. (2004). The dynamics of neutral sugars in the rhizosphere of wheat. An approach by <sup>13</sup>C pulse-labelling and GC/C/IRMS. *Plant and Soil* 267, 243–253. doi:10.1007/s11104-005-5348-8.

Dettman, J. R., Rodrigue, N., Melnyk, A. H., Wong, A., Bailey, S. F., and Kassen, R. (2012). Evolutionary insight from whole-genome sequencing of experimentally evolved microbes: experimental evolution and genome sequencing. *Molecular Ecology* 21, 2058–2077. doi:10.1111/j.1365-294X.2012.05484.x.

Doin de Moura, G. G., Remigi, P., Masson-Boivin, C., and Capela, D. (2020). Experimental evolution of legume symbionts: What have we learnt? *Genes* 11, 339. doi:10.3390/genes11030339.

Drake, J. W. (1991). A constant rate of spontaneous mutation in DNA-based microbes. *Proc. Natl. Acad. Sci.* 88, 7160–7164. doi:10.1073/pnas.88.16.7160.

D'Souza, G., and Kost, C. (2016). Experimental evolution of metabolic dependency in bacteria.



PLoS Genetics 12, e1006364.  
doi:10.1371/journal.pgen.1006364.

Elena, S. F., and Lenski, R. E. (2003). Microbial genetics: evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. *Nature Reviews Genetics* 4, 457–469. doi:10.1038/nrg1088.

Fisher, R. A. (1930). The genetical theory of natural selection. Oxford, UK: Oxford Univ. Press.

Gandon, S., Hochberg, M. E., Holt, R. D., and Day, T. (2013). What limits the evolutionary emergence of pathogens? *Philosophical Transactions of the Royal Society B: Biological Sciences* 368, 20120086. doi:10.1098/rstb.2012.0086.

Garoff, L., Pietsch, F., Huseby, D. L., Lilja, T., Brandis, G., and Hughes, D. (2020). Population bottlenecks strongly influence the evolutionary trajectory to fluoroquinolone resistance in *Escherichia coli*. *Mol Biol Evol* 37, 1637–1646. doi:10.1093/molbev/msaa032.

Genin, S., and Denny, T. P. (2012). Pathogenomics of the *Ralstonia solanacearum* species complex. *Annual Review of Phytopathology* 50, 67–89. doi:10.1146/annurev-phyto-081211-173000.

Gerrish, P. J., and Lenski, R. E. (1998). The fate of competing beneficial mutations in an asexual population. *Mutation and Evolution* 127–144. doi:10.1007/978-94-011-5210-5\_12.

Gilbert, G. S., and Parker, I. M. (2010). Rapid evolution in a plant-pathogen interaction and the consequences for introduced host species: pathogen evolution and novel hosts. *Evolutionary Applications* 3, 144–156. doi:10.1111/j.1752-4571.2009.00107.x.

Gresham, D., and Hong, J. (2015). The functional basis of adaptive evolution in chemostats. *FEMS Microbiol Rev* 39, 2–16. doi:10.1111/1574-6976.12082.

Griffin, A. S., West, S. A., and Buckling, A. (2004). Cooperation and competition in pathogenic bacteria. *Nature* 430, 1024–1027. doi:10.1038/nature02744.

Guan, S. H., Gris, C., Cruveiller, S., Pouzet, C., Tasse, L., Leru, A., et al., (2013). Experimental evolution of nodule intracellular infection in legume symbionts. *The ISME Journal* 7, 1367–1377. doi:10.1038/ismej.2013.24.

Guidot, A., Jiang, W., Ferdy, J.-B., Thébaud, C., Barberis, P., Gouzy, J., et al., (2014). Multihost experimental evolution of the pathogen *Ralstonia solanacearum* unveils genes involved in

adaptation to plants. *Molecular Biology and Evolution* 31, 2913–2928. doi:10.1093/molbev/msu229.

Haichar, F. el Z., Santaella, C., Heulin, T., and Achouak, W. (2014). Root exudates mediated interactions belowground. *Soil Biology and Biochemistry* 77, 69–80. doi:10.1016/j.soilbio.2014.06.017.

Hartwell, L. H., Hopfield, J. J., Leibler, S., and Murray, A. W. (1999). From molecular to modular cell biology. *Nature* 402, C47–C52. doi:10.1038/35011540.

Hassani, M. A., Durán, P., and Hacquard, S. (2018). Microbial interactions within the plant holobiont. *Microbiome* 6, 58. doi:10.1186/s40168-018-0445-0.

Herring, C. D., Raghunathan, A., Honisch, C., Patel, T., Applebee, M. K., Joyce, A. R., et al., (2006). Comparative genome sequencing of *Escherichia coli* allows observation of bacterial evolution on a laboratory timescale. *Nature Genetics* 38, 1406–1412. doi:10.1038/ng1906.

Hillesland, K. L., and Stahl, D. A. (2010). Rapid evolution of stability and productivity at the origin of a microbial mutualism. *Proc. Natl. Acad. Sci.* 107, 2124–2129. doi:10.1073/pnas.0908456107.

Hughes, B. S., Cullum, A. J., and Bennett, A. F. (2007). Evolutionary adaptation to environmental pH in experimental lineages of *Escherichia coli*. *Evolution* 61, 1725–1734. doi:10.1111/j.1558-5646.2007.00139.x.

Janzen, D. (1980) When is it co-evolution? *Evolution*, 34 (3) 611-612. doi: 10.1111/j.1558-5646.1980.tb04849.x.

Johansson, J. (2008). Evolutionary responses to environmental changes: how does competition affect adaptation? *Evolution* 62, 421–435. doi:10.1111/j.1558-5646.2007.00301.x.

Kawecki, T. J., Lenski, R. E., Ebert, D., Hollis, B., Olivieri, I., and Whitlock, M. C. (2012). Experimental evolution. *Trends in Ecology & Evolution* 27, 547–560. doi:10.1016/j.tree.2012.06.001.

Khan, A. I., Dinh, D. M., Schneider, D., Lenski, R. E., and Cooper, T. F. (2011). Negative epistasis between beneficial mutations in an evolving bacterial population. *Science* 332, 1193–1196. doi:10.1126/science.1203801.

Kram, K. E., Geiger, C., Ismail, W. M., Lee, H., Tang, H., Foster, P. L., et al., (2017). Adaptation of *Escherichia coli* to long-term serial passage in complex medium: evidence of parallel evolution. *Systems* 2. doi:10.1128/mSystems.00192-16.

- Lang, G. I., Botstein, D., and Desai, M. M. (2011). Genetic variation and the fate of beneficial mutations in asexual populations. *Genetics* 188, 647–661. doi:10.1534/genetics.111.128942.
- Lassalle, F., Planel, R., Penel, S., Chapulliot, D., Barbe, V., Dubost, A., *et al.*, (2017). Ancestral genome estimation reveals the history of ecological diversification in *Agrobacterium*. *Genome Biol Evol* 9, 3413–3431. doi:10.1093/gbe/evx255.
- Lawrence, D., Fiegna, F., Behrends, V., Bundy, J. G., Phillimore, A. B., Bell, T., *et al.*, (2012). Species interactions alter evolutionary responses to a novel environment. *PLoS Biology* 10, e1001330. doi:10.1371/journal.pbio.1001330.
- Lecomte, S. M., Achouak, W., Abrouk, D., Heulin, T., Nesme, X., and Haichar, F. el Z. (2018). Diversifying anaerobic respiration strategies to compete in the rhizosphere. *Frontiers in Environmental Science* 6. doi:10.3389/fenvs.2018.00139.
- Lenski, R. E. (2017). Experimental evolution and the dynamics of adaptation and genome evolution in microbial populations. *The ISME Journal* 11, 2181–2194. doi:10.1038/ismej.2017.69.
- Lenski, R. E., Rose, M. R., Simpson, S. C., and Tadler, S. C. (1991). Long-term experimental evolution in *Escherichia coli*, adaptation and divergence during 2,000 generations. *The American Naturalist* 138, 1315–1341. doi:10.1086/285289.
- Lenski, R. E., and Travisano, M. (1994). Dynamics of adaptation and diversification: a 10,000-generation experiment with bacterial populations. *Proc. Natl. Acad. Sci.* 91, 6808–6814. doi:10.1073/pnas.91.15.6808.
- Little, A. E. F., Robinson, C. J., Peterson, S. B., Raffa, K. F., and Handelsman, J. (2008). Rules of engagement: interspecies interactions that regulate microbial communities. *Annual Review of Microbiology* 62, 375–401. doi:10.1146/annurev.micro.030608.101423.
- Maharjan, R., Seeto, S., Notley-McRobb, L., and Ferenci, T. (2006). Clonal adaptive radiation in a constant environment. *Science* 313, 514–517. doi:10.1126/science.1129865.
- Marchetti, M., Clerissi, C., Yousfi, Y., Gris, C., Bouchez, O., Rocha, E., *et al.*, (2017). Experimental evolution of rhizobia may lead to either extra- or intracellular symbiotic adaptation depending on the selection regime. *Molecular Ecology* 26, 1818–1831. doi:10.1111/mec.13895.
- Mas, A., Jamshidi, S., Lagadeuc, Y., Eveillard, D., and Vandenkoornhuyse, P. (2016). Beyond the Black Queen hypothesis. *The ISME Journal* 10, 2085–2091. doi:10.1038/ismej.2016.22.
- Mazancourt, C. D., Johnson, E., and Barraclough, T. G. (2008). Biodiversity inhibits species' evolutionary responses to changing environments. *Ecology Letters* 11, 380–388. doi:10.1111/j.1461-0248.2008.01152.x.
- Meaden, S., and Koskella, B. (2017). Adaptation of the pathogen, *Pseudomonas syringae*, during experimental evolution on a native vs. alternative host plant. *Molecular Ecology* 26, 1790–1801. doi:10.1111/mec.14060.
- Mendes, R., Kruijt, M., de Bruijn, I., Dekkers, E., van der Voort, M., Schneider, J. H. M., *et al.*, (2011). Deciphering the rhizosphere microbiome for disease-suppressive bacteria. *Science* 332, 1097–1100. doi:10.1126/science.1203980.
- Nilsson, A. I., Koskineniemi, S., Eriksson, S., Kugelberg, E., Hinton, J. C. D., and Andersson, D. I. (2005). Bacterial genome size reduction by experimental evolution. *Proc. Natl. Acad. Sci.* 102, 12112–12116. doi:10.1073/pnas.0503654102
- Panke-Buisse, K., Poole, A. C., Goodrich, J. K., Ley, R. E., and Kao-Kniffin, J. (2015). Selection on soil microbiomes reveals reproducible impacts on plant function. *The ISME Journal* 9, 980–989. doi:10.1038/ismej.2014.196.
- Paterson, S., Vogwill, T., Buckling, A., Benmayor, R., Spiers, A. J., Thomson, N. R., *et al.*, (2010). Antagonistic coevolution accelerates molecular evolution. *Nature* 464, 275–278. doi:10.1038/nature08798.
- Pelosi, L., Kühn, L., Guetta, D., Garin, J., Geiselmann, J., Lenski, R. E., *et al.*, (2006). Parallel changes in global protein profiles during long-term experimental evolution in *Escherichia coli*. *Genetics* 173, 1851–1869. doi:10.1534/genetics.105.049619.
- Perrier, A., Peyraud, R., Rengel, D., Barlet, X., Lucasson, E., Gouzy, J., *et al.*, (2016). Enhanced *in planta* fitness through adaptive mutations in EfpR, a dual regulator of virulence and metabolic functions in the plant pathogen *Ralstonia solanacearum*. *PLoS Pathogens* 12, e1006044. doi:10.1371/journal.ppat.1006044.
- Pickersgill, B. (2018). Parallel vs. convergent evolution in domestication and diversification of crops in the americas. *Front. Ecol. Evol.* 6. doi:10.3389/fevo.2018.00056.
- Ratcliff, W. C., Denison, R. F., Borrello, M., and Travisano, M. (2012). Experimental evolution of

- multicellularity. *Proc. Natl. Acad. Sci.* 109, 1595–1600. doi:10.1073/pnas.1115323109.
- Sachs, J. L., and Wilcox, T. P. (2006). A shift to parasitism in the jellyfish symbiont *Symbiodinium microadriaticum*. *Proceedings of the Royal Society B: Biological Sciences* 273, 425–429. doi:10.1098/rspb.2005.3346.
- Sánchez-Cañizares, C., Jorrín, B., Poole, P. S., and Tkacz, A. (2017). Understanding the holobiont: The interdependence of plants and their microbiome. *Current Opinion in Microbiology* 38, 188–196. doi:10.1016/j.mib.2017.07.001.
- Scanlan, P. D., Hall, A. R., Lopez-Pascua, L. D. C., and Buckling, A. (2011). Genetic basis of infectivity evolution in a bacteriophage. *Molecular Ecology* 20, 981–989. doi:10.1111/j.1365-294X.2010.04903.x.
- Schlötterer, C., Kofler, R., Versace, E., Tobler, R., and Franssen, S. U. (2015). Combining experimental evolution with next-generation sequencing: a powerful tool to study adaptation from standing genetic variation. *Heredity* 114, 431–440. doi:10.1038/hdy.2014.86.
- Schluter, D., Clifford, E., Nemethy, M., and Mckinnon, J. (2004). Parallel evolution and inheritance of quantitative traits. *The American naturalist* 163, 809–22. doi:10.1086/383621.
- Segrè, A. V., Murray, A. W., and Leu, J.-Y. (2006). High-resolution mutation mapping reveals parallel experimental evolution in yeast. *PLoS Biology* 4, e256. doi:10.1371/journal.pbio.0040256.
- Shendure, J. (2005). Accurate multiplex polony sequencing of an evolved bacterial genome. *Science* 309, 1728–1732. doi:10.1126/science.1117389.
- Smith, P., and Schuster, M. (2019). Public goods and cheating in microbes. *Current Biology* 29, R442–R447. doi:10.1016/j.cub.2019.03.001.
- Strotz, L. C., Simões, M., Girard, M. G., Breitzkreuz, L., Kimmig, J., and Lieberman, B. S. (2018). Getting somewhere with the Red Queen: chasing a biologically modern definition of the hypothesis. *Biology Letters* 14, 20170734. doi:10.1098/rsbl.2017.0734.
- Swenson, W., Wilson, D. S., and Elias, R. (2000). Artificial ecosystem selection. *Proc. Natl. Acad. Sci.* 97, 9110–9114. doi:10.1073/pnas.150237597.
- Tannières, M., Lang, J., Barnier, C., Shykoff, J. A., and Faure, D. (2017). Quorum-quenching limits quorum-sensing exploitation by signal-negative invaders. *Scientific Reports* 7, 40126. doi:10.1038/srep40126.
- Tenaillon, O. (2014). The utility of Fisher's geometric model in evolutionary genetics. *Annual Review of Ecology, Evolution, and Systematics* 45, 179–201. doi:10.1146/annurev-ecolsys-120213-091846.
- Theis, K. R., Dheilly, N. M., Klassen, J. L., Brucker, R. M., Baines, J. F., Bosch, T. C. G., et al., (2016). Getting the hologenome concept right: an eco-evolutionary framework for hosts and their microbiomes. *mSystems* 1, e00028-16. doi:10.1128/mSystems.00028-16.
- Toft, C., and Andersson, S. G. E. (2010). Evolutionary microbial genomics: insights into bacterial host adaptation. *Nature Reviews Genetics* 11, 465–475. doi:10.1038/nrg2798.
- Travisano, M., Mongold, J., Bennett, A., and Lenski, R. (1995). Experimental tests of the roles of adaptation, chance, and history in evolution. *Science* 267, 87–90. doi:10.1126/science.7809610.
- Treves, D. S., Manning, S., and Adams, J. (1998). Repeated evolution of an acetate-crossfeeding polymorphism in long-term populations of *Escherichia coli*. *Mol Biol Evol* 15, 789–797. doi:10.1093/oxfordjournals.molbev.a025984.
- Valenzuela-Miranda, D., Valenzuela-Muñoz, V., Nuñez-Acuña, G., and Gallardo-Escárate, C. (2020). Long-term serial culture of *Piscirickettsia salmonis* leads to a genomic and transcriptomic reorganization affecting bacterial virulence. *Aquaculture* 529, 735634. doi:10.1016/j.aquaculture.2020.735634.
- Van den Bergh, B., Swings, T., Fauvart, M., and Jan, M. (2018). Experimental design, population dynamics, and diversity in microbial experimental evolution. *Microbiology and Molecular Biology Reviews*. <https://mmbr.asm.org/content/82/3/e00008-18.short>.
- Van Elsas, J. D., Turner, S., and Bailey, M. J. (2003). Horizontal gene transfers in the phytosphere. *New Phytologist* 157, 525–537. doi:10.1046/j.1469-8137.2003.00697.x.
- Van Valen, L. (1977). The Red Queen. *The American Naturalist* 111, 809–810. doi:10.1086/283213.
- Vandenkoornhuyse, P., Quaiser, A., Duhamel, M., Van, A. L., and Dufresne, A. (2015). The importance of the microbiome of the plant holobiont. *New Phytologist* 206, 1196–1206. doi:10.1111/nph.13312.

- Velicer, G. J., and Yu, Y. N. (2003). Evolution of novel cooperative swarming in the bacterium *Myxococcus xanthus*. *Nature* 425, 75–78. doi:10.1038/nature01908.
- Vorholt, J. A., Vogel, C., Carlström, C. I., and Müller, D. B. (2017). Establishing causality: opportunities of synthetic communities for plant microbiome research. *Cell Host & Microbe* 22, 142–155. doi:10.1016/j.chom.2017.07.004.
- Westphal, L. L., Lau, J., Negro, Z., Moreno, I. J., Ismail Mohammed, W., Lee, H., *et al.*, (2018). Adaptation of *Escherichia coli* to long-term batch culture in various rich media. *Research in Microbiology* 169, 145–156. doi:10.1016/j.resmic.2018.01.003.
- Wichman, H. A. (1999). Different trajectories of parallel evolution during viral adaptation. *Science* 285, 422–424. doi:10.1126/science.285.5426.422.
- Wielgoss, S., Barrick, J. E., Tenaillon, O., Cruveiller, S., Chane-Woon-Ming, B., Médigue, C., *et al.*, (2011). Mutation rate inferred from synonymous substitutions in a long-term evolution experiment with *Escherichia coli*. *Genes|Genomes|Genetics* 1, 183–186. doi:10.1534/g3.111.000406.
- Winder, M., and Schindler, D. E. (2004). Climate change uncouples trophic interactions in an aquatic ecosystem. *Ecology* 85, 2100–2106. doi:10.1890/04-0151.
- Wiser, M. J., Ribeck, N., and Lenski, R. E. (2013). Long-term dynamics of adaptation in asexual populations. *Science* 342, 1364–1367. doi:10.1126/science.1243357.
- Wisniewski-Dyé, F., Borziak, K., Khalsa-Moyers, G., Alexandre, G., Sukharnikov, L. O., Wuichet, K., *et al.*, (2011). *Azospirillum* genomes reveal transition of bacteria from aquatic to terrestrial environments. *PLoS Genetics* 7, e1002430. doi:10.1371/journal.pgen.1002430.
- Yang, L., Jelsbak, L., Marvig, R. L., Damkiaer, S., Workman, C. T., Rau, M. H., *et al.*, (2011). Evolutionary dynamics of bacteria in a human host environment. *Proc. Natl. Acad. Sci.* 108, 7481–7486. doi:10.1073/pnas.1018249108.
- Yeh, P., Tschumi, A. I., and Kishony, R. (2006). Functional classification of drugs by properties of their pairwise interactions. *Nat Genet* 38, 489–494. doi:10.1038/ng1755.
- Yoshida, T., Jones, L. E., Ellner, S. P., Fussmann, G. F., and Hairston, N. G. (2003). Rapid evolution drives ecological dynamics in a predator–prey system. *Nature* 424, 303–306. doi:10.1038/nature01767.
- Yu, Z., Beck, D. A. C., and Chistoserdova, L. (2017). Natural selection in synthetic communities highlights the roles of *Methylococcaceae* and *Methylophilaceae* and suggests differential roles for alternative methanol dehydrogenases in methane consumption. *Front. Microbiol.* 8. doi:10.3389/fmicb.2017.02392.
- Zbinden, M., Haag, C. R., and Ebert, D. (2008). Experimental evolution of field populations of *Daphnia magna* in response to parasite treatment. *Journal of Evolutionary Biology* 21, 1068–1078. doi:10.1111/j.1420-9101.2008.01541.x.
- Zhang, Z., Gosset, G., Barabote, R., Gonzalez, C. S., Cuevas, W. A., and Saier, M. H. (2005). Functional interactions between the carbon and iron utilization regulators, Crp and Fur, in *Escherichia coli*. *Journal of Bacteriology* 187, 980–990. doi:10.1128/JB.187.3.980-990.2005.





# Designing synthetic communities

## INTRODUCTION

Laboratory systems with single strains offer high controllability, however they were culture-based; yet, only a small fraction of microorganisms can be cultured. Also, microorganisms present different behaviours depending on how they are grown: pure cultures or within a microbial community (De Roy *et al.*, 2014). In nature, microorganisms are under constant interaction with other organisms, causing that any population tends to evolve with the other sympatric microorganisms (Brockhurst *et al.*, 2003; Guan *et al.*, 2013). Furthermore, ecosystems functioning is the result of a joint effort of different microorganisms with different functional roles (De Roy *et al.*, 2014). These insights caused a shift from single microorganism to whole community studies. Yet, researchers have taken interest in plant – associated microbial communities since 1940 (Timonin, 2011) and microbial assemblies have largely been used in food industry for sausage and dairy products fermentation (Bassi *et al.*, 2015; Leroy *et al.*, 2006). However, it is only recently, with the development of high-throughput sequencing of amplicons and molecular fingerprinting techniques as well as metagenomics, metatranscriptomics and metaproteomics approaches that communities can be targeted as a whole and be characterized. Nowadays, SynCom are being assembled to unravel ecological questions. Synthetic communities (SynCom) approaches aim to mimic natural microbial communities and are defined as artificial communities created by co-culturing two or more strains under a well-defined medium. The chosen microorganisms can be of the same species as well as genetically engineered or not. These SynCom offer a lower complexity than natural community and are generally grown in highly controlled environments, achieving high reproducibility (Großkopf and Soyer, 2014). SynCom have become popular these last decades when it comes to gaining an insight into natural microbial communities' functioning as they offer a simplified representation conserving microbial community's key features. Moreover, natural microbial communities also face numerous abiotic *e.g.* salinity, temperature, *etc.*) and biotic factors (cooperation, antagonism, *etc.*) that can modify their complexity (abundance, diversity of species and functioning). SynCom can help to unravel the impact abiotic and biotic factors can have on communities' complexity and functioning by closely monitoring them. Understanding these factors will help to engineer SynCom with a desired function.

Here we focus on SynCom with bacteria, yeast and fungi mostly. We review the two main approaches that stand out when working with SynCom. Both approaches begin with a question of interest (*i.e.* looking to assemble a cellulose degrading community). The first approach aims to answer this question by defining the function of the community (*i.e.* cellulose degradation) and choosing the strains that can perform that function. In the second approach, the focus is on

species interaction; researchers look for an interaction network (*i.e.* a cooperative network) that will answer the question. The interaction network can be obtained by controlling the strain's metabolism or environment. Later on, we will focus on plant microbial communities and what to look for when building a plant-intended SynCom.

## **1. Community function as the foundation for SynCom design**

A synthetic community can be set up to perform a specific biological function or process (*e.g.* lactate production, pesticide degradation, *etc.*). Often, these functions can be monitored and therefore different community assemblies performing the same function can be compared (Großkopf and Soyer, 2014). For wastewater treatments, the reduction of organic matter is a process that is easily measurable. Microbes use oxygen to break down organic compounds. Thus, by measuring the biochemical oxygen demand, the performance of microbial consortia to complete the process can be monitored (Briones and Raskin, 2003). This approach enables to unravel the relations between species diversity and functionality, both key properties that conditioning community functioning. The strength of this approaches to fulfil a biological process is the ability to fine-tune community diversity. Thus, replacing one microbial strain with another as well as increasing or decreasing the number of populations, enables to measure the impact of these modifications on community functioning. For example, increasing the diversity of a community has been shown to increase the removal of mercury from polluted wastewater (von Canstein *et al.*, 2002). Possible bottlenecks within the synthetic community can be foreseen with this approach as the effect of each member can be assessed independently. This functional approach offers insight into community structure as species interactions condition whether the function is carried efficiently or not. This mechanistic understanding gives the opportunity to detect key species for the performance of a given function.

## **2. SynComs constructed based on interaction networks**






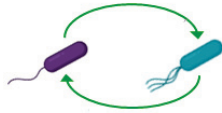


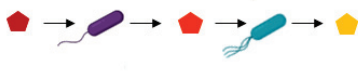



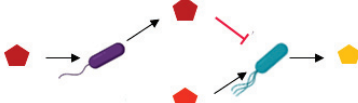
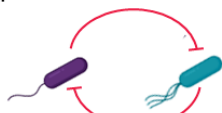
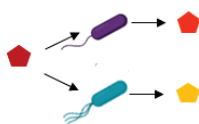
Interacting microorganisms can have positive, negative or neutral effects on each other. Hence, interactions between microorganisms determine the community functioning. Unravelling the interactions that take place within a community is essential to understand how a community carries a function and how it responds to perturbations. In nature, there are 5 basic bacterial interactions: cooperation, commensalism, antagonism, amensalism, competition (table1). Among these interactions, cooperation, defined as the action or process of working together to the same end, is the most controversial. Indeed, natural selection pushes onto maximizing the own fitness of microbial populations. However, cooperation is wide spread across microbial systems (Sachs, 2006). It is possible to exploit this biotic interaction and design synthetic cooperative communities with increased function performance. Interactions can be either metabolism-based, as many metabolites are shared within a community, and therefore be directly impacted by the environment.

## 2.1 Metabolism and cross-feeding based cooperation

Metabolism is a major driver of microbial interactions as it takes part in all 5 basic bacterial interactions cooperation, commensalism, antagonism, amensalism, competition between any microorganisms (table 1). Cooperative behaviour can easily be achieved by dividing metabolic roles within the community. In cooperative SynCom, strains can be modified to become auxotrophic for one essential metabolite. Auxotrophic strains are unable to grow by themselves as they cannot synthesize one essential metabolite. Within a bacterial community, auxotrophic strains for different metabolites can complement one another's growth by cross-feeding the missing metabolite, sometimes achieving greater growth rates than the wild strain. A collection of 46 auxotrophic *Escherichia coli* (Wintermute and Silver, 2010), each auxotrophic for one different essential metabolite (synthesis of amino acids, nucleotides, cofactors, *etc.*), were grown in co-culture (2 strains at a time). The expected results were that each auxotrophic strain would supply their partner's missing metabolite. Dissimilar co-culture combinations (n=1035) were tested and greater growth rates were found for 176 co-cultures (up to 50-fold increase when compared to the corresponding monoculture). Auxotrophic cooperative strain present higher growth rates because by helping their partner, strains increase the production of their own essential metabolite.

Another investigation with 14 *E. coli* strains (Mee *et al.*, 2014), each auxotrophic for one different essential amino acid, also showed that co-cultures of 2 or 3 auxotrophic strains presented synergistic growth (up to 90 fold than the control condition). The syntrophic network was reconstructed based on the synergistic growth rates and showed that amino acids that are costly to produce (methionine, leucine) seem to promote more and stronger cooperative interactions than biosynthetically cheaper amino acids. In parallel, the 14 strains were experimentally evolved as a SynCom on minimal medium for 50 days (around 400 generations) and population abundances were tracked by qPCR. After only 3 days, there is a shift in population abundances with only 4 dominant members (A, B, C, D). Population dynamics then stabilised for the rest of the experiment. To get a deeper insight into these interactions, four 13-members-SynCom, were designed leaving out one of the dominant strains (A, B, C, D). These 13-members SynCom lead to different population dynamics, with 2 to 3 different dominant strains depending on the 13-member SynCom. These results show the importance of community composition and populations dynamics on cross-feeding interactions. This study also highlights the importance of studying communities as a whole when trying to understand microbial ecosystems.



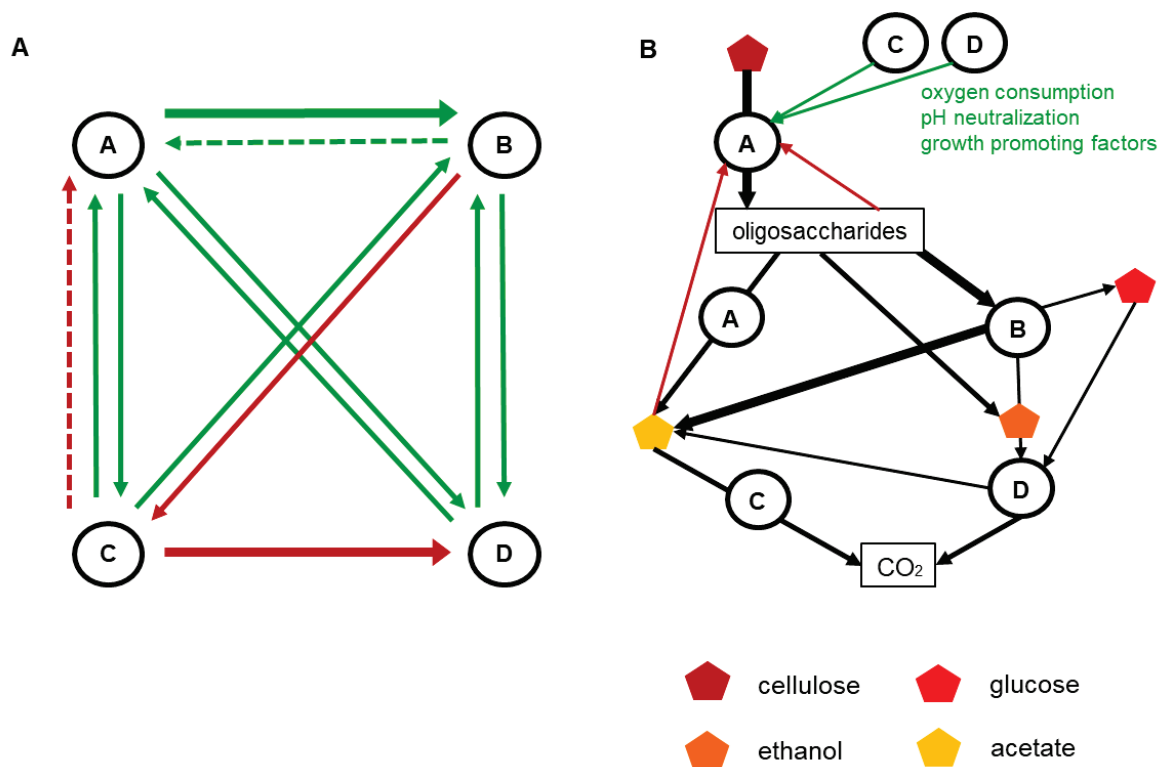
<div style="display: flex; justify-content: space-around; align-items: center;"> <div style="text-align: center;">Species A </div> <div style="text-align: center;">Species B </div> <div style="text-align: center;">Metabolites   </div> </div>			
Effect	Ecological Interaction	Metabolic Interaction	Description
<b>+/+</b>	Cooperation 		The violet microorganism is inhibited by its own waste product, the light blue microorganisms is capable of utilizing this waste product and therefore lifting the inhibition.
<b>0/+</b>	Commensalism 		The light blue microorganism is capable of using the product synthesized by the violet microorganism without having any impact on this microorganism.
<b>-/+</b>	Antagonism 		The product synthesized by the light blue microorganism can inhibit the violet one.
<b>0/-</b>	Amensalism 		The product of the violet microorganism inhibits the light blue one although it is not the same food chain.
<b>-/-</b>	Competition 		Both microorganisms feed on the same metabolite.

**Table 1: Five metabolism based bacterial interactions.** Violet and light blue circles represent two different interacting species and the pentagons represent metabolites. Green and red arrows represent positive or negative interactions respectively.

## 2.2 Environmental factors driven cooperation

Environmental changes can influence species interactions and some environmental conditions are more favourable to cooperation. Some environmental parameters generate selective pressures such that they force microorganisms to cooperate. Fixing these parameters makes it possible to obtain consortia of cooperating strains. The aim is to create artificial microbial ecosystems without engineering the microorganisms but rather by predicting their growth rate in a given environment. *Desulfovibrio vulgaris* and *Methanococcus maripaludis* share a positive interaction: *D. vulgaris* consumes lactate and produces formate and hydrogen, which are consumed by *M. maripaludis* to produce methane. *D. vulgaris* benefits from the hydrogen reduction as it allows it to continue its metabolic processes (Klitgord and Segrè, 2010; Taffs *et al.*, 2009). Flux balance analysis, a mathematical modelling approach used to quantitatively simulate microbial metabolism, enabled to design 6 simple media sustaining the growth of both microorganisms together but not separately. All of these media contained lactate plus a source of nitrogen and sulphur. At the time of the analysis, only one of the predicted media had been experimentally tested to study these microorganisms. (Klitgord and Segrè, 2010). The remaining media were predicted to favour commensal or mutualistic interactions, offering 5 other environmental conditions that could be used to unravel other aspects of these interactions. This computer model based on fundamental physiochemical laws, takes into account the stoichiometry of the environment and should be used when designing SynCom as it can foresee different outcomes depending on the tested environmental conditions.

Other studies have combined approaches based on both species' function and on biotic interactions to create SynCom. This combined approach was for instance used to study the network of relationships of a four strain synthetic-consortia degrading cellulose (Kato *et al.*, 2008). The synthetic community was composed of a cellulose-degrading anaerobe (strain A), a saccharide-utilizing anaerobe (strain B), a peptide- and acetate-utilizing aerobe (strain C) and a peptide-, glucose-, and ethanol-utilizing aerobe (strain D). They unravelled the interspecies interaction by analysing the effects of culture filtrates of one strain on the growth of the other strains and found out a complex network of positive and negative, as well as direct or indirect interactions (Figure 1A and B). This study pointed out that all the interactions, even the negatives ones were well balanced to maintain community stability. However, the less numerous species are in SynCom, the more resilient they need to be, in order for the community to resist to environmental perturbations (this topic is discussed later on). Combining the function-based and the interaction-based approaches to build SynCom gives insight into the relations between function, structure and dynamics of the communities.



**Figure 1: Interaction network of a cellulose degrading synthetic community.** (A) Substrate utilization was deduced for each strain by constructing “knockout communities” (communities missing one strain). The flow of the substrate (cellulose) is indicated by the solid arrows. The thickness of the arrows represents the relative contribution of a given strain for cellulose degradation. Finally, the green and red arrows represent promoting or inhibiting effects of intermediate metabolites or strains regarding cellulose degradation. Figure modified from Kato *et al.*, 2005. (B) Substrate utilization was deduced for each strain by constructing “knockout communities” (communities missing one strain). The flow of the substrate (cellulose) is indicated by the solid arrows. The thickness of the arrows represents the relative contribution of a given strain for cellulose degradation. Finally, the green and red arrows represent promoting or inhibiting effects of intermediate metabolites or strains regarding cellulose degradation. Figure modified from Kato *et al.*, 2005.

### 3. Application of SynComs

#### 3.1 SynComs for fundamental research: adaptation to new environments

Because of their reduced complexity and high controllability, SynCom are preferred in laboratory research to test ecological theories and to get a deeper insight into the complex networks of interactions within natural communities. SynCom have been used to study adaptation to new environments, as species interactions can greatly influence the adaptive responses of microbial communities (Lawrence *et al.*, 2012). For instance, five naturally co-occurring bacteria, isolated from the roots of beech trees and belonging to the families of *Sphingobacteriaceae*, *Flavobacteriaceae*, *Enterobacteriaceae*, *Pseudomonadaceae* and *Sphingomonadaceae* were evolved for 70 generations in monocultures and polycultures on beech leaf extract (Lawrence *et al.*, 2012). Species in communities evolved to use the waste products generated by others, such

that growth rates in the polycultures went down when one species was removed. These results highlight the importance of species interactions-when adapting to a new environment (from beech trees roots to beech leaf extract medium).

Along the same lines, bacterial adaptation to antibiotics has extensively been assessed (Baquero *et al.*, 2008; Levy and Marshall, 2004), and there is evidence that sub-minimal inhibiting concentration (sub-MIC) can increase variance in fitness among the members of a community, consequently reducing taxa diversity (Abeles *et al.*, 2016; Hall and Corno, 2014; Trindade *et al.*, 2012). Bacteria mostly become resistant via horizontal gene transfers (HGT), in which they acquire mobile genetic elements coding for antibiotic resistance, such as plasmids (Martínez *et al.*, 2015). Several factors can modulate HGT: presence of antibiotic, species interactions (such as predation) and spatial heterogeneity. Cairns *et al.* (2018a) studied the effects these parameters can have on community composition and plasmid transfer by evolving a 62-member SynCom on an antibiotic supplemented media, structured via glass beads and in presence of a ciliated protist predator. Each of these parameters was tested separately and among the 62 members, only one carried broad-host-range conjugative plasmid. All the parameters modify diversity and the conjugative plasmid was detected in 35% of the total taxa of the community. Antibiotic reduced community composition and increased the number of plasmid recipient taxa, however, the plasmid recipient taxa depended on the tested parameter (antibiotic, glass beads, predator). These results bring up the importance of abiotic (antibiotic, glass beads) and biotic (predator) factors can have when adapting to a new environment.

In this sub-section we discussed the importance of species interactions when adapting to a new environment. In the above examples, the environment is defined as the abiotic support on which the SynCom evolves (*i.e.* beech leaf extract medium or antibiotic supplemented media). However, SynCom also adapt to their biotic environment, that is, the other members of the community. Microorganisms live as communities and multispecies communities' adaptive evolution is difficult to predict even with stable laboratory-imposed environmental conditions. To overcome this problem, an experiment was carried with a synthetic community of six soil bacteria species (*Enterobacter aerogenes*, *Serratia marcescens*, *Pseudomonas fluorescens*, *Pseudomonas aurantiaca*, *Pseudomonas veronii* and *Pseudomonas putida*) and with 96 identical replicates of the ancestral community (Celiker and Gore, 2014). All of the 96 identical ancestral communities underwent a short-term and a long-term evolution. For the short-term evolution, the evolved communities were plated after roughly 40 generations. Co-existence of 5 out of 6 of the initial species was found for all 96 replicates, only *P. putida* was represented at less than 1% abundance. The long-term evolution (roughly 400 generations) showed high discrepancy among the replicates. There was still a strong co-existence of the 5 species (*E. aerogenes*, *S. marcescens*, *P. fluorescens*, *P. aurantiaca* and *P. veronii*). However, final species abundances could be clustered into 4 possible outcomes, with different dominant strains among the 5 co-

existing species. Further analyses showed that these possible outcomes are determined by a single species gaining an advantage at a given moment and increasing in frequency. Surprisingly, *P. putida* became the dominant strain in 9 out of the 96 replicates after 400 generations despite its very low initial abundance (less than 1% after 40 generations). Additional analyses showed a rare advantageous mutational event in this *P. putida* strain before it went extinct allowing it to increase its frequency within the SynCom. These results suggest that adaptive evolution in multispecies communities is not random, but it rather follows one of many possible outcomes depending on the individual adaptation of their species and on the resulting interactions network. Understanding how species interaction and abundance of community members evolve is crucial to predict how complex communities will respond to an environmental change.

### 3.2 SynComs for applied research: bioproduction of molecules of interest

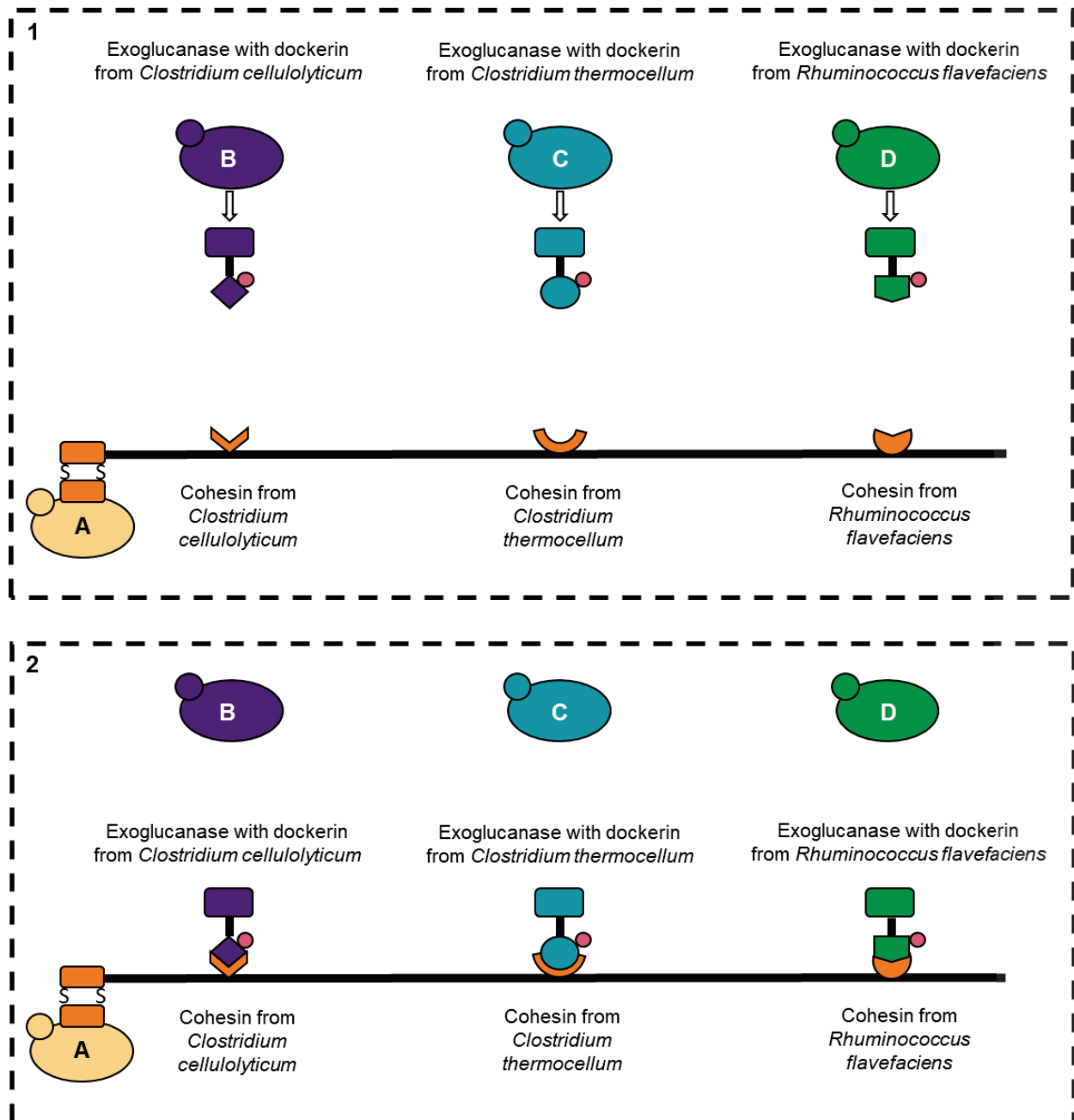
When it comes to increasing the production of a molecule of interest, biotechnology development has long relied in a so-called “consolidated bioprocessing” approach. The idea of this approach is to construct a “superbug” which would harbour all the required functionalities to assure the biological processes. Despite being a cost-effective approach, creating a “superbug” turned out to be challenging and yields are often non-commercially viable (Minty *et al.*, 2013). “Superbugs” are not often found in nature because microorganisms live within rich communities with synergistic interactions that allow the community to carry a function. Communities offer attractive properties such as stability, functional robustness, and the ability to perform complex tasks (Mee and Wang, 2012; Zuroff and Curtis, 2012). These properties can be exploited when using synthetic microbial communities: communities can be assembled so that interactions will increase the yield of a certain reaction or by creating overlapping metabolic food chains (*i.e.* using the waste products from another microorganism). In the biotechnology industry, SynCom are generally used to compartmentalize metabolic pathways between several microorganisms in order to achieve individual optimisation and a higher final yield.

SynCom have drawn attention in the research field of biofuels. This is especially the case for the degradation of lignocellulosic biomass, or dry plant matter, the number one candidate for biofuel production as it is the most abundant raw material on Earth. Ethanol can be produced from lignocellulosic biomass and may lessen the world wide dependency on fossil fuels (Lynd *et al.*, 2005). The exploitation of this natural resource is mainly slowed down by the high cost of enzymes required for hydrolysis of cellulose and fermentation of the released sugars into ethanol. Initially, *S. cerevisiae* was considered to be a good candidate to create a “superbug” capable of cellulose hydrolysis and ethanol fermentation, however none of the attempts to engineer this microorganism were satisfactory (Den Haan *et al.*, 2007; Jeon *et al.*, 2009; Yanase *et al.*, 2010). In an effort to surpass the limitations of the consolidated bioprocess approach, a synthetic community of 4 *S. cerevisiae* displaying a mini-cellulosome, capable of cellulose hydrolysis, was

assembled. Cellulosomes are naturally occurring enzymes complex capable of cellulose hydrolysis. In the case of this synthetic community, each yeast was genetically engineered so that the consortia would display an exoglucanase, an endoglucanase and a  $\beta$ -glucosidase, all required enzymes for cellulose hydrolysis and ethanol production (Figure 2). One of the engineered yeasts displayed a trifunctional scaffoldin, carrying three divergent cohesive domains from *Clostridium cellulolyticum*, *Clostridium thermocellum*, and *Ruminococcus flavefaciens* (yeast A). The other three were engineered to secrete one of the three corresponding dockerin-tagged cellulases (exoglucanase, endoglucanase and  $\beta$ -glucosidase, yeast B, C, D). The secreted cellulases are docked onto a trifunctional scaffoldin in an organized manner leading to a functional mini-cellulosome on the surface of the yeast displaying the scaffoldin. Each yeast provides a unique building block necessary for the mini-cellulosome structure in such a way that the cellulosome assembly, cellulose hydrolysis and ethanol production can be fine-tuned by modifying the ratio of each population (Goyal *et al.*, 2011; Tsai *et al.*, 2009, 2010). The synthetic community displayed a 3-fold increase of ethanol production compared to the 3 single yeasts. The use of synthetic complex communities may improve ethanol production, taking bio fuel production one step further.

Plants also play a role in the biotechnology industry as they can synthesize compounds with therapeutic properties (e.g. taxol, a potent anticancer drug produced by *Taxus brevifolia*, a pacific yew tree). Microbial metabolisms have been modified to achieve a heterologous production of this type of molecules. It has been challenging in bacteria, eukaryotic enzymes being often required for the synthesis of complex compounds, moreover, protein folding is far more complex and difficult to achieve in prokaryote models (Chang *et al.*, 2007; Gething, 1997). On the other hand, for some biosynthetic products such as isoprenoid, higher yield is often achieved in bacteria than in yeasts. Division of labour among eukaryotic and prokaryotic microorganisms has been found in nature to accomplish metabolic and energy demanding tasks (Agapakis *et al.*, 2012). In order to produce complex compounds with high yield, it could be interesting to reconstruct a heterologous metabolic pathway within a SynCom in which one microorganism synthesizes a metabolic intermediate that is then translocated to another microorganism or several other microorganisms to realize the other steps of the reaction. It would achieve a spatial pathway module segregation, as each microorganism is specialised in one fraction of the metabolic pathway. The production of the anticancer drug taxol was increased by splitting its synthetic pathway between engineered *E. coli* and *Saccharomyces cerevisiae*. *E. coli* was engineered to overproduce an intermediate inactive molecule, taxadiene, which is then taken up by *S. cerevisiae*, the preferable host for expressing the oxygenation reactions needed to functionalize taxadiene. Dividing the metabolic synthetic pathways between these 2 microorganisms combines rapid production of taxadiene by *E. coli* with efficient oxygenation by *S. cerevisiae* (Zhou *et al.*, 2015). Engineered consortia with portioned and well-defined function may be more efficient in achieving specialized tasks than “superbugs”.





**Figure 2: Surface assembly of a functional minicellulosome by intracellular complementation.** (1) Four yeast strains were engineered to be able to either display a trifunctional scaffoldin carrying three divergent cohesin domains from *C. cellulolyticum*, *C. thermocellum*, and *R. flavefaciens* (A) or to secrete one of the three corresponding dockerin-tagged cellulases: exoglucanase (B), endoglucanase (C) and b-glucosidase (D). (2) The secreted cellulases are docked onto the trifunctional scaffoldin in a highly organized manner leading to a functional mini-cellulosome on the surface of the yeast (A). Figure modified from Tsai *et al.*, 2010.

## CONCLUDING REMARKS

SynCom offers thus simplified and controlled models to better understand the functioning of biotic interaction network within natural communities. Initially, studies using SynCom have focused on population dynamics and outcomes of the interactions between populations. With increasing knowledge of community's structure and functioning, SynCom studies evolved onto more applied purposes, opening the door to biotechnological applications. However, one of the remaining barriers is predicting what will happen once the SynCom is introduced in a natural complex environment as interactions within communities in nature are driven by a plethora of environmental factors. Reciprocally, microbiota have an impact on their ecological niche or host, to such extent that host-microorganism's interactions have played key roles in adaptation. For instances, arbuscular mycorrhizal interactions are believed to be the origin of plants terrestrialization (Heckman *et al.*, 2001; Selosse and Le Tacon, 1998). Since this discovery, it has become clear that microorganisms have a profound impact on their host and both partners should be viewed as a holobiont. A holobiont is defined as the plant, its microbiota and the interactions existing between them Zilber-Rosenberg and Rosenberg, 2008).



# Synthetic communities for plants

## INTRODUCTION

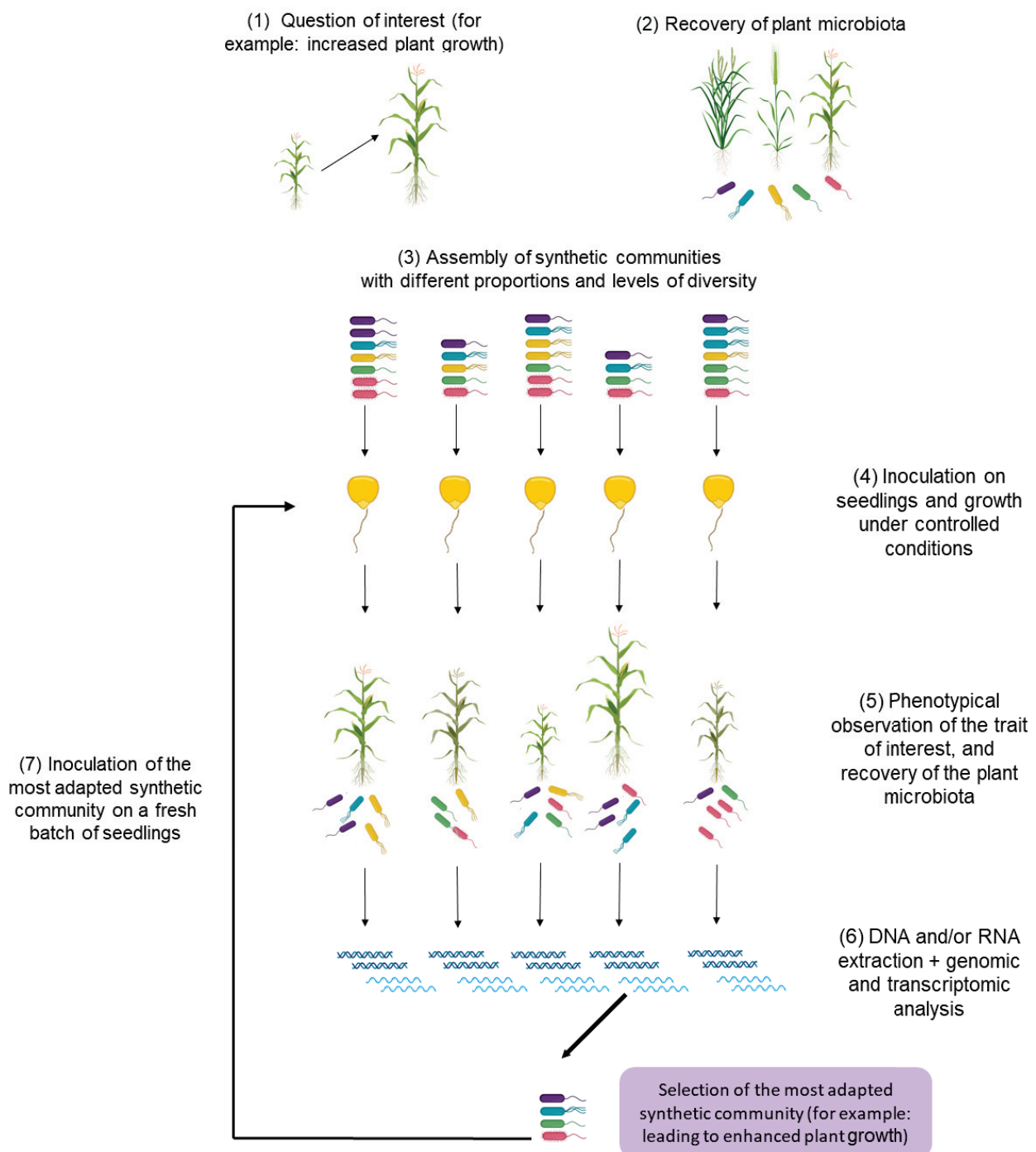
Plants present different types of environment for microorganisms. For environments under the influence of plant exudations and secretions, we distinguish aerial (phyllosphere), root (rhizosphere), floral and even an intra-plant environment which is called endosphere. Thus plant is hosting different microbiota (phyllospheric, rhizospheric, endophytic and flower microbiota; Bulgarelli *et al.*, 2013), which can be passed from one generation to the next or recruited from the surrounding microbial reservoirs (soil, air, other plants...). Among the recruited microorganisms, plant growth-promoting rhizobacteria (PGPR) colonize the rhizosphere of many plants and improve plant growth even under adverse conditions via numerous plant-beneficial properties (Vacheron, 2015). Since microbiota composition conditions the interactions within the community (positive or negative), the observed plant growth-promoting effect will be the result of all the interactions. Consequently, understanding the mechanisms controlling microbial colonisation and persistence into already settled natural microbial communities is crucial. Interactions between the plant and its microbiota, and within the microbiota itself are known to be complex which makes understanding its assembling and functioning challenging. SynCom offer the opportunity to unravel the network of interactions within microbial communities and the possibility to engineer an inoculum with a desired function. Here we briefly discuss how plant SynCom are assembled and the main points that should be taken into account when assembling SynCom destined to increase growth or to protect plants.

When building SynCom, microorganisms are chosen in order to answer the ecological question that is being asked (e.g. increased plant growth; Figure 1-1). Different microorganisms are recovered from plants and screening strategies are used to select candidates harbouring plant-beneficial properties that measure up to the question (Figure 1-2). The selected microorganisms are assembled in different proportions and level of diversity (Figure 1-3), with up to now, mostly effort made on taxonomic diversity rather than functional diversity. SynComs can be inoculated on the plant of interest and grown under controlled conditions (Figure 1-4; growth chambers with pre-defined day and night cycles) or in the field (Vorholt *et al.*, 2017). The extent of the experiments depends on the plant, but it has to be long enough for roots to develop and for the microorganisms to colonize them. This time lapse is referred to as a selection cycle. At the end of each selection cycle, the SynComs are retrieved and phenotypical traits are observed if possible (Figure 1-5; *i.e.* enhanced plant-growth). DNA and/or RNA can be extracted from the SynComs and sequenced (Figure 1-6). Genomic reads will reveal the evolution dynamics of the SynCom: Which strains are still present? Are there any mutations fixed within the genomes? The transcriptomes will reveal which genes were expressed at the time of root retrieval. Finally, the results of these two data sets (phenotypical observations and genomic analysis) helps choosing

the SynCom that best answers the question of interest (*i.e.* enhanced growth). Afterwards, the selected SynCom can be re-inoculated on a fresh batch of seedlings (Figure 1-7). As discussed before (section 2.1), community composition is crucial as species abundance and diversity can have an impact on species interactions, and thus on the capacity of the SynCom to adapt to changes. Therefore, several factors such as the number of plant-beneficial properties shared by microorganisms or their ability to colonize plant roots, need to be accounted for when choosing the microorganisms. Here we expose the principles that should guide the choice of strains to construct an efficient plant-beneficial synthetic community.

## **1. The importance of plant beneficial properties**

The rhizosphere harbours a rich and diversified microbial community (Gans, 2005; Gomes *et al.*, 2010; Roesch *et al.*, 2008). As the plant liberates exudates, it actively selects its own microbiota. Among the selected microorganisms some will be beneficial for plant health and growth through direct or indirect mechanisms (Couillerot *et al.*, 2009). The interaction between the plant and these microorganisms are termed as symbiotic and the cost and benefits are shared between the partners (Bulgarelli *et al.*, 2013). There are mainly two types of symbiotic interactions: mutualistic obligate interactions between some microorganisms and a narrow range of compatible host plants (*e.g.* rhizobia and *Fabaceae* interaction, (Parniske, 2008), and secondly, cooperation. Cooperation is characterized by less specific and obligated interactions (Drogue *et al.*, 2012). SynCom designed for plants are focused on cooperative microorganisms that are able to colonize the roots and display beneficial effects on the plant.



**Figure 1: Designing synthetic communities for plants.** Firstly, the question of interest is defined (1). Then, microorganisms recovered from plants are selected depending on the question of interest (2) and different synthetic communities are assembled with different proportions and levels of diversity (3). Synthetic communities are inoculated on sterilized seedlings and grown under controlled conditions. (4). At the end of the selection cycle, phenotypal observations are made if possible and plant microbiota are recovered (5), DNA and/or RNA are extracted and analysed (6). Combining these approaches (5 and 6) helps selecting the most suited synthetic community to answer the question of interest. The selected synthetic community can then be inoculated on a new batch of seedlings (7).

Beneficial microorganisms are defined as such because of the plant-beneficial properties they harbour. Plant beneficial properties can be classified in different categories. First, there are properties involved in enhancing the plant nutrition, such as nitrogen fixation, potassium acquisition, phosphate solubilisation, *etc.* (Glick, 2010; Lin *et al.*, 2009; Richardson, 2001; Vance, 2001). These minerals are widely present in the soil but are not readily accessible for plants (Rai, 2006). Secondly, there are properties able to enhance root development and plant growth, and involve the secretion of hormones like auxins, cytokinins, *etc.* Lastly, some plant associated microorganisms can protect the plant from abiotic stresses, for instance some PGPR are able to depollute heavy metal contaminated soils by rendering them inactive (Tak 2013). PGPR can also protects plants from biotic stresses like plant pathogens, by the production of antimicrobial compounds (2,4-diacetylphloroglucinol (DAPG), phenazines (Couillerot *et al.*, 2009). Plant protection can also be assured by PGPR harbouring type III secretion systems, a molecular syringe capable of host manipulation by injecting virulence factors directly into eukaryotic cells (Rezzonico et al 2005).

Even though the right combination of plant-beneficial properties may create a functional SynCom, it does not consider the interactions the selected microorganisms may have between them. In order to build an efficient synthetic community, many other factors will influence the interactions and need to be taken into account, such as the ability of community members to colonize and persist on the host plant or the presence of type VI secretion systems, a molecular syringe capable of interbacterial competition (Bernal et al 2018).

## **2. Functional redundancy**

Microbial communities are complex and several parameters must be taken into account to understand their functioning and persistence in the environment. These include species richness and functional diversity. The latter is defined as the capacity of communities to perform a wide range of functions. Disturbance of species richness and/or functional diversity can have a major impact on ecosystem functioning (Rosenfeld, 2002). SynCom are being assembled to be used, in the long term, in natural environments where stochastic changes can be radical to the point that some species will not be able to execute their role. In order to resist to stochastic changes, it would be interesting thus to choose microorganisms with functional redundancy when assembling SynCom.

Functional redundancy can be defined as the ability to perform a given function by multiple means (Rosenfeld, 2002). However, functional redundancy can also be found at the level of the community, where several species are capable of performing the same function (Figure 2). In the rhizosphere, plant-beneficial bacteria colonize the roots of plants and may protect them from

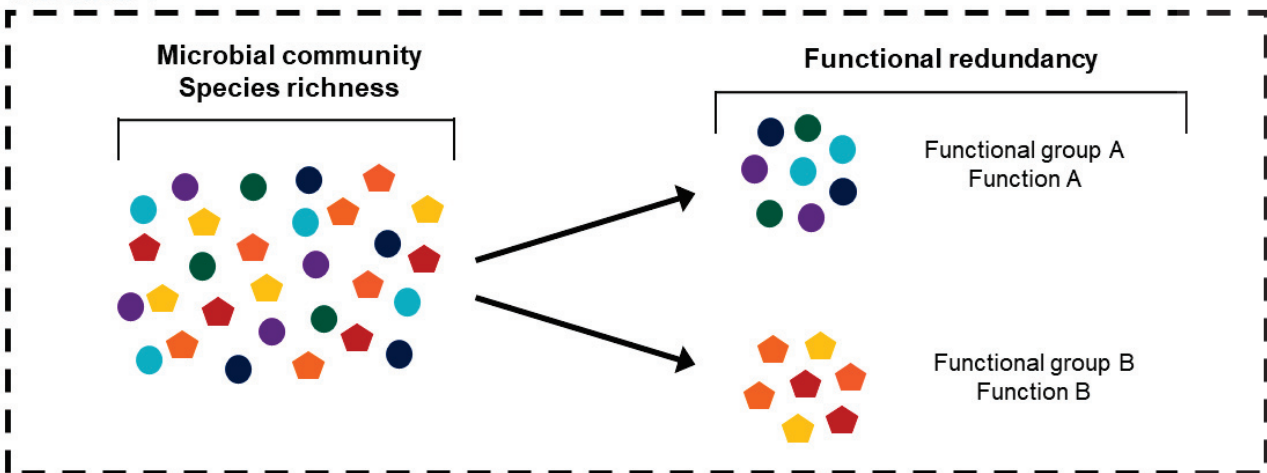
different stresses. Their effectiveness can be influenced by their ability to sense the presence of other organisms, whether plant-beneficial or pathogenic ones. This can be done via signal molecules, such as *N*-acylhomoserine lactones (AHL) in the case of Gram negative bacteria, which allow communication between cells so that the population or community acts as a multicellular entity (Fuqua *et al.*, 1994). This is known as *quorum sensing* communication systems. These systems are widespread among bacteria. Often exogenous AHLs, produced by bacteria belonging to other populations, are able to activate/regulate the same functions than the endogenous AHLs. In *Pseudomonas aureofaciens*, the expression of the phenazine genes, coding for a metabolite involved in the protection of plants against certain pathogens, can be activated by AHLs from other bacteria belonging to the same genus (Pierson *et al.*, 1998). Plant-beneficial bacteria are never alone in the rhizosphere and the fact that different exogenous AHLs can be perceived by bacterial populations makes it possible to regulate the production of metabolites more effectively. AHLs have a production cost. Possessing a receptor allowing the use of exogenous AHLs but not the AHLs synthases allows the bacteria to modify their behaviour more rapidly without having to pay the price of their biosynthesis. Thus, this type of cross-regulations plays an important role in determining community structure and rhizosphere functioning.

It is important to consider functional redundant species within SynCom so the function will still be assured despite perturbations. The disappearance of a redundant species will affect community composition, but will not lead to the loss of community functions because the other taxa in the functional group compensate for this extinction. However, functional redundancy for a given function does not mean that the loss of certain taxa is without consequence for the functioning of the functional group, since different species performing the same function do not respond in the same way to fluctuating environmental conditions (Walker, 1995) Indeed, they might be identical in terms of their role but not in terms of their acclimatization to environmental conditions. Thus, depending on the disturbance, functional redundancy may allow or not the SynCom to perform as well as before the disturbance despite the loss of certain species (Allison and Martiny, 2008).

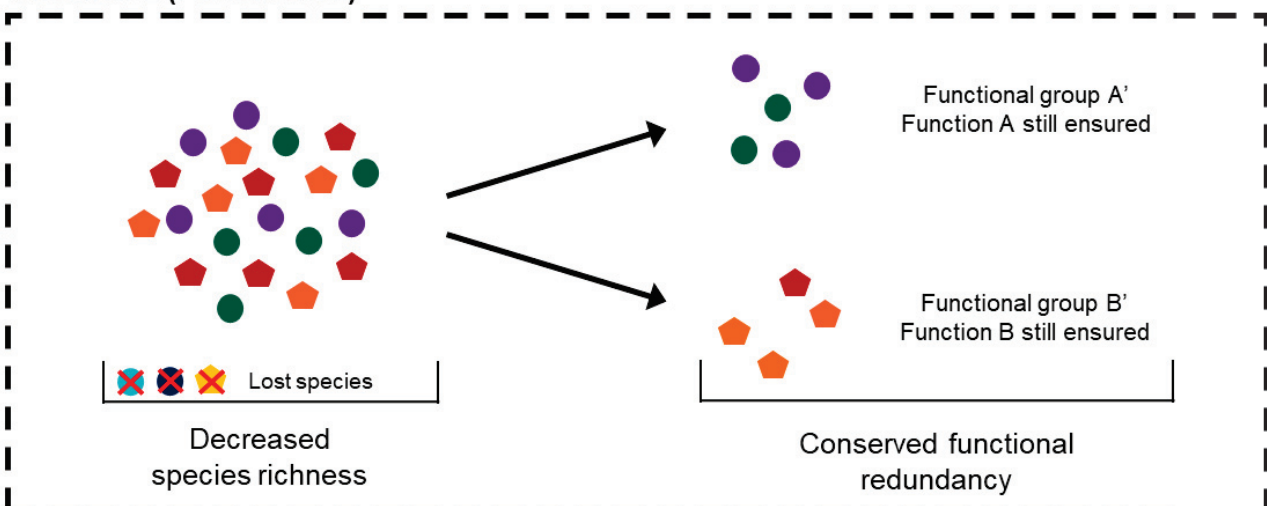
### **3. Rhizosphere establishment of the plant-beneficial bacteria**

Before the synthetic community can manifest any plant growth promoting effects on the inoculated plant, the exogenous microorganisms have to invade, colonize and persist on the plant. In order to persist, the inoculated synthetic community must adapt to the indigenous microbial population. Natural plant microbiota is believed to be assembled in two steps: a general recruitment from the vicinity of the roots and a second selection that involves genetic factors (Lemanceau *et al.*, 2017). However, the underlying mechanics of microbial community assemblies in plants is still at the initial stages as there are several microbiota in one plant: phyllospheric, rhizospheric, endophytic and flower microbiota (Bulgarelli *et al.*, 2013).

### Situation 1



### Situation 2 (Perturbation)



**Figure 2: Functional redundancy within microbial communities.** Situation 1 represents a microbial community harbouring high species richness (each species is represented by a different colour). Within the different species, several are capable of performing the same function: function A (circles) or function B (pentagons). The species of a same functional group are functionally redundant. Situation 2 a perturbation occurred leading to the loss of 3 species. However, there are still species capable of performing function A and B within each functional group. Despite decreased species richness, the functional redundancy is conserved limiting the impact of perturbation on ecosystem functioning.

Often, the phyllospheric, endospheric and rhizospheric microbiota is dependent of the host genotype. For instance, the effects of host genotype on the phyllospheric community of *Arabidopsis thaliana* have been assessed using SynCom (Bodenhausen *et al.*, 2014). To do this, 55 *Arabidopsis* mutants with different alterations (e.g. in cuticle structure, defense signalling, pathogen recognition; among others) were inoculated with a SynCom composed of seven strains representing the most abundant phyla in the *Arabidopsis*' phyllosphere and belonging to *Alphaproteobacteria* and *Betaproteobacteria*. Mutants for cuticle formation showed the most



perturbed microbial composition and a higher bacterial abundance relative to the wild type line. Mutants for cuticle formation plants are presenting an increased cell wall permeability, which would lead to an increased nutrient availability, favouring the growth of the *Betaproteobacterium Variovorax*, capable of feeding on many substrates. In parallel, *Rhodococcus* which feeds on cuticle components (which may be less available on these mutants), presented a decreased population level. Overall these results showed that different bacteria can benefit or be handicapped by genetic modifications of the host and that host genotype is another factor that needs to be accounted for when studying the composition of host-associated communities. However other studies that have tried to explain the factors structuring microbiota have presented contradictory results: 27 modern maize inbreds were grown in 5 different locations and the impact of heritability, sample type (bulk soil or maize rhizosphere) and geographical location was assessed on microbial diversity. Although there is a proportion of heritable variation between the replicates of the inbred lines most of the variation between the microbial populations can be imputed to geographical locations (Peiffer *et al.*, 2013).

Moreover, root exudate composition can also have an impact on microbial diversity as it changes along the root system but also with the development of the plant. These modifications greatly impact the composition of the communities selected by the plants (Bouffaud *et al.*, 2012; Bulgarelli *et al.*, 2013; Chaparro *et al.*, 2013; Redford and Fierer, 2009). This suggest that plant microbiota assemblies are influenced by many different factors (genotype, geographic variation, season variation, etc) and fluctuate during the different stages of plant development. Although these studies have shed some light on the assembling of plant communities they also show that there will not be “one size fits all” consortia as there is great variation in the selection of plant microbiota.

#### **4. Diversity and species richness within the SynCom**

Microbial diversity can widen the availability of potential plant beneficial properties. However, expression of the genes coding for this functions can be modified by the biotic interactions with the endogenous community (Garbeva *et al.*, 2011; Jousset *et al.*, 2014). For example, co-existing species may have positive interactions leading to a more performant community (Bruno *et al.*, 2003).

Literature has pointed out the positive relation between biodiversity and ecosystem functioning: highly diverse communities are expected to be more performant as they are more efficient in capturing limiting resources (Cadotte *et al.*, 2008; Cardinale *et al.*, 2007; Loreau, 2010). Highly diverse communities also imply a higher capacity to colonize diverse plant niches and therefore to use different resources proposed by the plant (Cadotte *et al.*, 2008; Loreau and Hector, 2001). However, limited space and nutrients can also favour antagonistic interactions, reducing

community functioning at high diversity. Plant protection can be affected by the diversity of plant-associated communities: diverse communities are believed to be highly resistant to invasion of an exogenous pathogen microorganism due to the high number of species interactions (Case, 1990) and limited space (high competition for the available niches, (van Elsas *et al.*, 2012). Yet, studies have also reported negative relationships between biodiversity and community productivity. For instance, Becker *et al* manipulated the genotypic diversity of *P. fluorescens* communities and showed that antagonistic interactions increased drastically with species richness leading to a complete loss of plant protection (Becker *et al.*, 2012). In an effort to understand these interactions and the impact of the rhizospheric microbiota on plant health and growth, maize axenic seedlings were inoculated with a synthetic community (Niu *et al.*, 2017) composed of *Enterobacter cloacae*, *Stenotrophomonas maltophilia*, *Ochrobactrum pituitosum*, *Herbaspirillum frisingense*, *Pseudomonas putida*, *Curtobacterium pusillum*, and *Chryseobacterium indologenes*. The role each strain plays in community assembly was evaluated by removing one strain at a time coupled to a culture-dependent approach. Only the removal of *E. cloacae* produces great changes in community composition, suggesting that this strain may be necessary for the persistence of the others. When removing the other species, population dynamics are similar to those found with the seven original species synthetic community. Furthermore, this community successfully protected maize against *Fusarium verticillioides* (causing agent of seedling blight disease). Each separated strain was able to protect maize to varying degrees and *E. cloacae* displayed the highest efficiency against the pathogen. The ability to track the impact of each separated strain within the synthetic community may give insight into the mechanisms underlying community assemblies and beneficial effects of plant microbiota. It could also help to identify species that play key roles within the community, such as *E. cloacae*.

Higher diversity equals a more natural representation of the studied microbiota. However, no matter how complex the synthetic community is; it will always be less complex than the natural population it is trying to mimic. Yet, increased diversity will prevent from missing out on key members and therefore, on the interactions they may trigger. Controlled diversity offers reproducibility to experiments carried with SynCom. Therefore, the lower the diversity, the more reproducible the system will be. Indeed, stochastic events are minimized for low diversity communities (Vorholt *et al.*, 2017).

## **5. Plant recognition of inoculated beneficial microorganisms**

Besides adapting to the indigenous microbial communities, the inoculated microorganisms must not trigger the plant defence responses. Plant are in constant interaction with many microorganisms and rely on two types of defences, physical barriers (cell wall and cuticle, preventing microbial colonization) and their immune system. Plants immune system recognize



conserved structures of microbes, often essential for their development (Newman *et al.*, 2013). These molecules are referred to as microbe associated molecular patterns (MAMPs) are recognized by the plant immune systems pattern recognition receptors (PRRs). Plants can discriminate pathogens from non-pathogens and this recognition triggers an immune response: plants will either support, resist or ignore microbial growth on or within plant tissue. *Arabidopsis* immune system was tested by inoculating various non-pathogenic bacteria from its own phyllosphere (Vogel *et al.*, 2016). Different plant transcriptional responses were found depending on the inoculated bacteria. Colonization by *Methylobacterium extorquens* lightly modified the expression of less than 10 genes, whereas *Sphingomonas melonis* modified the expression of over 400 genes. Surprisingly, among the 400 genes, some overlapped those triggered in presence of the plant pathogen *P. syringae* DC3000. It is argued that these differential induced expressions of genes may be due to different MAMPs expressed by the inoculated bacteria, or a different accessibility from the PRRs. Nonetheless *S. melonis* may act somewhat like a vaccine, inducing a pattern-triggered immunity which can lead to long-lasting induction of defence genes, contributing to plant health upon pathogen encounter.

Phytohormones also play a crucial role in the plant immune system. Salicylic acid, jasmonic acid and gaseous ethylene mediate localized and systemic plant immune responses (Belkhadir *et al.*, 2014; Pieterse *et al.*, 2014). *Arabidopsis* that share similar root associated microbiota can display different defense phytohormones profiles (Bulgarelli *et al.*, 2012; Lundberg *et al.*, 2012). To understand the impact of the plant defence phytohormone, salicylic acid, on microbiota composition, wild type *A. thaliana* root microbiota was compared to *A. thaliana* mutants lacking biosynthesis of, and/or signal dependent salicylic acid. Despite being a foliar defence phytohormone, salicylic acid and/ or salicylic acid-mediated events drive microbial selection from the available microbial pool and are needed to sculpt the root microbiota. These data suggest that plant immune responses play an active role in selecting plant microbiota from the available pool microbiota and should be taking into account when designing SynCom.

## CONCLUDING REMARKS

The mechanisms underlying communities functioning and structure are shaped by many abiotic and biotic factors. Interactions within the community can vary from positive to negative depending on the spatial scale and resource heterogeneity or with species richness, diversity and functional redundancy. In plants, these factors are poorly understood due to the inherent complexity of the model. SynCom have become a promising approach to understand the underlying mechanisms of plant-microbiota interactions as they aim to mimic natural microbiota, although they have a significantly lower complexity than natural populations and they might be missing significant community members (Vorholt *et al.*, 2017). Despite their reduced complexity, synthetic consortia

offer a laboratory sized model mimicking natural microbiota where the consequences of abiotic (*i.e.* changes in the environment) and biotic perturbations (*i.e.* modifications in population proportions) can be closely monitored. SynCom are an innovative tool meant to complement field studies by unravelling the molecular basis of environmental observation and offering a laboratory-sized system to test ecological theories. This approach can help unravel interspecies interactions under laboratory conditions and give a deeper insight into how these interactions affect microbiota assembly dynamics. In the long run, understanding the principles underlying the functioning and evolution of microbial communities in natural environment can eventually, guide next-generations applications in agriculture using SynCom.

## REFERENCES:

- Abeles, S.R., Jones, M.B., Santiago-Rodriguez, T.M., Ly, M., Klitgord, N., Yooseph, S., *et al.*, (2016). Microbial diversity in individuals and their household contacts following typical antibiotic courses. *Microbiome* 4, 39. 10.1186/s40168-016-0187-9.
- Agapakis, C.M., Boyle, P.M., and Silver, P.A. (2012). Natural strategies for the spatial optimization of metabolism in synthetic biology. *Nat. Chem. Biol.* 8, 527–535. 10.1038/nchembio.975.
- Allison, S.D., and Martiny, J.B.H. (2008). Resistance, resilience, and redundancy in microbial communities. *Proc. Natl. Acad. Sci.* 105, 11512–11519. 10.1073/pnas.0801925105.
- Baquero, F., Martínez, J.-L., and Cantón, R. (2008). Antibiotics and antibiotic resistance in water environments. *Curr. Opin. Biotechnol.* 19, 260–265. 10.1016/j.copbio.2008.05.006.
- Bassi, D., Puglisi, E., and Cocconcelli, P.S. (2015). Comparing natural and selected starter cultures in meat and cheese fermentations. *Curr. Opin. Food Sci.* 2, 118–122. 10.1016/j.cofs.2015.03.002.
- Becker, J., Eisenhauer, N., Scheu, S., and Jousset, A. (2012). Increasing antagonistic interactions cause bacterial communities to collapse at high diversity. *Ecol. Lett.* 15, 468–474. 10.1111/j.1461-0248.2012.01759.x.
- Belkhadir, Y., Yang, L., Hetzel, J., Dangl, J.L., and Chory, J. (2014). The growth-defense pivot: crisis management in plants mediated by LRR-RK surface receptors. *Trends Biochem. Sci.* 39, 447–456. 10.1016/j.tibs.2014.06.006.
- Bodenhausen, N., Bortfeld-Miller, M., Ackermann, M., and Vorholt, J.A. (2014). A synthetic community approach reveals plant genotypes affecting the phyllosphere microbiota. *PLoS Genet.* 10, e1004283. 10.1371/journal.pgen.1004283.
- Bouffaud, M., Kyselková, M., Gouesnard, B., Grundmann, G., Muller, D., and Moëgne-Loccoz, Y. (2012). Is diversification history of maize influencing selection of soil bacteria by roots? *Mol. Ecol.* 21, 195–206. 10.1111/j.1365-294X.2011.05359.x.
- Briones, A., and Raskin, L. (2003). Diversity and dynamics of microbial communities in engineered environments and their implications for process stability. *Curr. Opin. Biotechnol.* 14, 270–276. 10.1016/S0958-1669(03)00065-X.
- Brockhurst, M.A., Morgan, A.D., Rainey, P.B., and Buckling, A. (2003). Population mixing accelerates coevolution. *Ecol. Lett.* 6, 975–979. 10.1046/j.1461-0248.2003.00531.x.
- Bruno, J.F., Stachowicz, J.J., and Bertness, M.D. (2003). Inclusion of facilitation into ecological theory. *Trends Ecol. Evol.* 18, 119–125. 10.1016/S0169-5347(02)00045-9.
- Bulgarelli, D., Rott, M., Schlaeppi, K., Ver Loren van Themaat, E., Ahmadinejad, N., Assenza, F., *et al.*, (2012). Revealing structure and assembly cues for *Arabidopsis* root-inhabiting bacterial microbiota. *Nature* 488, 91–95. 10.1038/nature11336.
- Bulgarelli, D., Schlaeppi, K., Spaepen, S., van Themaat, E.V.L., and Schulze-Lefert, P. (2013). Structure and functions of the bacterial microbiota of plants. *Annu. Rev. Plant Biol.* 64, 807–838. 10.1146/annurev-arplant-050312-120106.
- Cadotte, M.W., Cardinale, B.J., and Oakley, T.H. (2008). Evolutionary history and the effect of biodiversity on plant productivity. *Proc. Natl.*

*Acad. Sci.* 105, 17012–17017.  
10.1073/pnas.0805962105.

Von Canstein, H., Kelly, S., Li, Y., and Wagner-Döbler, I. (2002). Species diversity improves the efficiency of mercury-reducing biofilms under changing environmental conditions. *Appl. Environ. Microbiol.* 68, 2829–2837. 10.1128/AEM.68.6.2829-2837.2002.

Cardinale, B.J., Wright, J.P., Cadotte, M.W., Carroll, I.T., Hector, A., Srivastava, D.S., *et al.*, (2007). Impacts of plant diversity on biomass production increase through time because of species complementarity. *Proc. Natl. Acad. Sci.* 104, 18123–18128. 10.1073/pnas.0709069104.

Case, T.J. (1990). Invasion resistance arises in strongly interacting species-rich model competition communities. *Proc. Natl. Acad. Sci.* 87, 9610–9614. 10.1073/pnas.87.24.9610.

Celiker, H., and Gore, J. (2014). Clustering in community structure across replicate ecosystems following a long-term bacterial evolution experiment. *Nat. Commun.* 5, 4643. 10.1038/ncomms5643.

Chang, M.C.Y., Eachus, R.A., Trieu, W., Ro, D.-K., and Keasling, J.D. (2007). Engineering *Escherichia coli* for production of functionalized terpenoids using plant P450s. *Nat. Chem. Biol.* 3, 274–277. 10.1038/nchembio875.

Chaparro, J.M., Badri, D.V., Bakker, M.G., Sugiyama, A., Manter, D.K., and Vivanco, J.M. (2013). Root exudation of phytochemicals in *Arabidopsis* follows specific patterns that are developmentally programmed and correlate with soil microbial functions. *PLoS ONE* 8, e55731. 10.1371/journal.pone.0055731.

Couillerot, O., Prigent-Combaret, C., Caballero-Mellado, J., and Moënne-Loccoz, Y. (2009). *Pseudomonas fluorescens* and closely-related fluorescent Pseudomonads as biocontrol agents of soil-borne phytopathogens. *Lett. Appl. Microbiol.* 48, 505–512. 10.1111/j.1472-765X.2009.02566.x.

De Roy, K., Marzorati, M., Van den Abbeele, P., Van de Wiele, T., and Boon, N. (2014). Synthetic microbial ecosystems: An exciting tool to understand and apply microbial communities. *Environ. Microbiol.* 16, 1472–1481. 10.1111/1462-2920.12343.

Den Haan, R., Rose, S.H., Lynd, L.R., and van Zyl, W.H. (2007). Hydrolysis and fermentation of amorphous cellulose by recombinant *Saccharomyces cerevisiae*. *Metab. Eng.* 9, 87–94. 10.1016/j.ymben.2006.08.005.

Drogue, B., Doré, H., Borland, S., Wisniewski-Dyé, F., and Prigent-Combaret, C. (2012). Which

specificity in cooperation between phytostimulating rhizobacteria and plants? *Res. Microbiol.* 163, 500–510. 10.1016/j.resmic.2012.08.006.

Van Elsas, J.D., Chiurazzi, M., Mallon, C.A., Elhottova, D., Kristufek, V., and Salles, J.F. (2012). Microbial diversity determines the invasion of soil by a bacterial pathogen. *Proc. Natl. Acad. Sci.* 109, 1159–1164. 10.1073/pnas.1109326109.

Fuqua, W.C., Winans, S.C., and Greenberg, E.P. (1994). *Quorum* sensing in bacteria: the LuxR-LuxI family of cell density-responsive transcriptional regulators. *J. Bacteriol.* 176, 269–275. 10.1128/jb.176.2.269-275.1994.

Gans, J. (2005). Computational improvements reveal great bacterial diversity and high metal toxicity in soil. *Science* 309, 1387–1390. 10.1126/science.1112665

Garbeva, P., Silby, M.W., Raaijmakers, J.M., Levy, S.B., and Boer, W. de (2011). Transcriptional and antagonistic responses of *Pseudomonas fluorescens* Pf0-1 to phylogenetically different bacterial competitors. *ISME J.* 5, 973–985. 10.1038/ismej.2010.196

Gething, M.-J. (1997). The difference with prokaryotes. *Nature* 388, 329–331. 10.1038/40979.

Glick, B.R. (2010). Using soil bacteria to facilitate phytoremediation. *Biotechnol. Adv.* 28, 367–374. 10.1016/j.biotechadv.2010.02.001.

Gomes, N.C.M., Cleary, D.F.R., Pinto, F.N., Egas, C., Almeida, A., Cunha, A., *et al.*, (2010). Taking root: enduring effect of rhizosphere bacterial colonization in mangroves. *PLoS ONE* 5, e14065. 10.1371/journal.pone.0014065.

Goyal, G., Tsai, S.-L., Madan, B., DaSilva, N.A., and Chen, W. (2011). Simultaneous cell growth and ethanol production from cellulose by an engineered yeast consortium displaying a functional mini-cellulosome. *Microb. Cell Factories* 10, 89. 10.1186/1475-2859-10-89.

Großkopf, T., and Soyer, O.S. (2014). Synthetic microbial communities. *Curr. Opin. Microbiol.* 18, 72–77. 10.1016/j.mib.2014.02.002.

Guan, S.H., Gris, C., Cruveiller, S., Pouzet, C., Tasse, L., Leru, A., *et al.*, (2013). Experimental evolution of nodule intracellular infection in legume symbionts. *ISME J.* 7, 1367–1377. 10.1038/ismej.2013.24.

Hall, A.R., and Corno, G. (2014). Tetracycline modifies competitive interactions in experimental microcosms containing bacteria isolated from

freshwater. *FEMS Microbiol. Ecol.* 90, 168–174. 10.1111/1574-6941.12388.

Heckman, D.S., Geiser, D.M., Eidell, B.R., Stauffer, R.L., Kardos, N.L., and Hedges, S.B. (2001). Molecular evidence for the early colonization of land by fungi and plants. *Science* 293, 1129–1133. 10.1126/science.1061457.

Hug, L.A., Baker, B.J., Anantharaman, K., Brown, C.T., Probst, A.J., Castelle, C.J., *et al.*, (2016). A new view of the tree of life. *Nat. Microbiol.* 1, 1–6. 10.1038/nmicrobiol.2016.48.

Jeon, E., Hyeon, J.-E., Suh, D.J., Suh, Y.-W., Kim, S.W., Song, K.H., *et al.*, (2009). Production of cellulosic ethanol in *Saccharomyces cerevisiae* heterologous expressing *Clostridium thermocellum* endoglucanase and *Saccharomycopsis fibuligera* beta-glucosidase genes. *Mol. Cells* 28, 369–373. 10.1007/s10059-009-0131-y.

Jessup, C.M., Forde, S.E., and Bohannan, B.J.M. (2005). Microbial experimental systems in ecology. *Advances in ecological research, Elsevier*, pp. 273–307. DOI: 10.1016/S0065-2504(04)37009-1.

Johansson, J. (2008). Evolutionary responses to environmental changes: how does competition affect adaptation? *Evolution* 62, 421–435. 10.1111/j.1558-5646.2007.00301.x.

Jousset, A., Becker, J., Chatterjee, S., Karlovsky, P., Scheu, S., and Eisenhauer, N. (2014). Biodiversity and species identity shape the antifungal activity of bacterial communities. *Ecology* 95, 1184–1190. 10.1890/13-1215.1.

Kato, S., Haruta, S., Cui, Z.J., Ishii, M and Igarashi, Y. (2005). Stable coexistence of five bacterial strains as a cellulose-degrading community. *Appl. Environ. Microbiol.* Vol. 71, No. 11, 7099–7106. doi:10.1128/AEM.71.11.7099–7106.2005

Kato, S., Haruta, S., Cui, Z.J., Ishii, M., and Igarashi, Y. (2008). Network relationships of bacteria in a stable mixed culture. *Microb. Ecol.* 56, 403–411. 10.1007/s00248-007-9357-4

Klitgord, N., and Segrè, D. (2010). Environments that induce synthetic microbial ecosystems. *PLoS Comput. Biol.* 6, e1001002. 10.1371/journal.pcbi.1001002.

Lawrence, D., Fiegna, F., Behrends, V., Bundy, J.G., Phillimore, A.B., Bell, T., *et al.*, (2012). Species interactions alter evolutionary responses to a novel environment. *PLoS Biol.* 10. 10.1371/journal.pbio.1001330

Lemanceau, P., Blouin, M., Muller, D., and Moënné-Loccoz, Y. (2017). Let the core

microbiota be functional. *Trends Plant Sci.* 22, 583–595. 10.1016/j.tplants.2017.04.008.

Leroy, F., Verluyten, J., and De Vuyst, L. (2006). Functional meat starter cultures for improved sausage fermentation. *Int. J. Food Microbiol.* 106, 270–285. 10.1016/j.ijfoodmicro.2005.06.027.

Levy, S.B., and Marshall, B. (2004). Antibacterial resistance worldwide: causes, challenges and responses. *Nat. Med.* 10, S122–S129. 10.1038/nm1145.

Lin, L., Li, Z., Hu, C., Zhang, X., Chang, S., Yang, L., *et al.*, (2009). Plant growth-promoting nitrogen-fixing enterobacteria are in association with sugarcane plants growing in Guangxi, China. *Microbes Environ. advpub*, 1204160379–1204160379. 10.1264/jisme2.ME11275.

Loreau, M. (2010). Linking biodiversity and ecosystems: towards a unifying ecological theory. *Philos. Trans. R. Soc. B Biol. Sci.* 365, 49–60. 10.1098/rstb.2009.0155.

Loreau, M., and Hector, A. (2001). Partitioning selection and complementarity in biodiversity experiments. *Nature* 412, 72–76. 10.1038/35083573.

Lundberg, D.S., Lebeis, S.L., Paredes, S.H., Yourstone, S., Gehring, J., Malfatti, S., *et al.*, (2012). Defining the core *Arabidopsis thaliana* root microbiome. *Nature* 488, 86–90. 10.1038/nature11237.

Lynd, L.R., van Zyl, W.H., McBride, J.E., and Laser, M. (2005). Consolidated bioprocessing of cellulosic biomass: an update. *Curr. Opin. Biotechnol.* 16, 577–583. 10.1016/j.copbio.2005.08.009.

Martínez, J.L., Coque, T.M., and Baquero, F. (2015). What is a resistance gene? Ranking risk in resistomes. *Nat. Rev. Microbiol.* 13, 116–123. 10.1038/nrmicro3399.

de Mazancourt, C., Johnson, E., and Barraclough, T.G. (2008). Biodiversity inhibits species' evolutionary responses to changing environments. *Ecol. Lett.* 11, 380–388. 10.1111/j.1461-0248.2008.01152.x.

Mee, M.T., and Wang, H.H. (2012). Engineering ecosystems and synthetic ecologies. *Mol. Biosyst.* 8, 2470–2483. 10.1039/c2mb25133g.

Mee, M.T., Collins, J.J., Church, G.M., and Wang, H.H. (2014). Syntrophic exchange in synthetic microbial communities. *Proc. Natl. Acad. Sci.* 111, E2149–E2156. 10.1073/pnas.1405641111.



- Minty, J.J., Singer, M.E., Scholz, S.A., Bae, C.-H., Ahn, J.-H., Foster, C.E., *et al.*, (2013). Design and characterization of synthetic fungal-bacterial consortia for direct production of isobutanol from cellulosic biomass. *Proc. Natl. Acad. Sci.* 110, 14592–14597. 10.1073/pnas.1218447110.
- Newman, M.-A., Sundelin, T., Nielsen, J.T., and Erbs, G. (2013). MAMP (microbe-associated molecular pattern) triggered immunity in plants. *Front. Plant Sci.* 4. 10.3389/fpls.2013.00139.
- Niu, B., Paulson, J.N., Zheng, X., and Kolter, R. (2017). Simplified and representative bacterial community of maize roots. *Proc. Natl. Acad. Sci.* 114, E2450–E2459. 10.1073/pnas.1616148114.
- Parniske, M. (2008). Arbuscular mycorrhiza: the mother of plant root endosymbioses. *Nat. Rev. Microbiol.* 6, 763–775. 10.1038/nrmicro1987.
- Peiffer, J.A., Spor, A., Koren, O., Jin, Z., Tringe, S.G., Dangi, J.L., *et al.*, (2013). Diversity and heritability of the maize rhizosphere microbiome under field conditions. *Proc. Natl. Acad. Sci.* 110, 6548–6553. 10.1073/pnas.1302837110.
- Pierson, E.A., Wood, D.W., Cannon, J.A., Blachere, F.M., and Pierson, L.S. (1998). Interpopulation signaling via *n*-acyl-homoserine lactones among bacteria in the wheat rhizosphere. *Mol. Plant-Microbe Interactions* 11, 1078–1084. 10.1094/MPMI.1998.11.11.1078
- Pieterse, C.M.J., Zamioudis, C., Berendsen, R.L., Weller, D.M., Van Wees, S.C.M., and Bakker, P.A.H.M. (2014). Induced systemic resistance by beneficial microbes. *Annu. Rev. Phytopathol.* 52, 347–375.
- Rai, M. (2006). *Handbook of Microbial Biofertilizers* (CRC Press).
- Redford, A.J., and Fierer, N. (2009). Bacterial succession on the leaf surface: a novel system for studying successional dynamics. *Microb. Ecol.* 58, 189–198. 10.1007/s00248-009-9495-y.
- Richardson, A.E. (2001). Prospects for using soil microorganisms to improve the acquisition of phosphorus by plants. *Funct. Plant Biol.* 28, 897–906. 10.1071/pp01093.
- Roesch, L.F.W., Camargo, F.A.O., Bento, F.M., and Triplett, E.W. (2008). Biodiversity of diazotrophic bacteria within the soil, root and stem of field-grown maize. *Plant Soil* 302, 91–104. 10.1007/s11104-007-9458-3
- Rosenfeld, J.S. (2002). Functional redundancy in ecology and conservation. *Oikos* 98, 156–162. 10.1034/j.1600-0706.2002.980116.x
- Sachs, J.L. (2006). Cooperation within and among species. *J. Evol. Biol.* 19, 1415–1418. 10.1111/j.1420-9101.2006.01152.x.
- Schmidt, T.M. (2006). The maturing of microbial ecology. *Int Microbiol.* 9(3):217-23. <https://core.ac.uk/reader/159084644>.
- Selosse, M.-A., and Le Tacon, F. (1998). The land flora: a phototroph-fungus partnership? *Trends Ecol. Evol.* 13, 15–20. 10.1016/S0169-5347(97)01230-5.
- Taffs, R., Aston, J.E., Briley, K., Jay, Z., Klatt, C.G., McGlynn, S., *et al.*, (2009). In silico approaches to study mass and energy flows in microbial consortia: a syntrophic case study. *BMC Syst. Biol.* 3, 114. 10.1186/1752-0509-3-114.
- Timonin, M.I. (2011). The interaction of higher plants and soil micro-organisms: Microbial population of rhizosphere of seedlings of certain cultivated plants. *Can. J. Res.* 10.1139/cjr40c-031.
- Trindade, S., Sousa, A., and Gordo, I. (2012). Antibiotic resistance and stress in the light of Fisher's model. *Evolution* 66, 3815–3824. 10.1111/j.1558-5646.2012.01722.x.
- Tsai, S.-L., Oh, J., Singh, S., Chen, R., and Chen, W. (2009). Functional assembly of minicellulosomes on the *Saccharomyces cerevisiae* cell surface for cellulose hydrolysis and ethanol production. *Appl. Environ. Microbiol.* 75, 6087–6093. 10.1128/AEM.01538-09.
- Tsai, S.-L., Goyal, G., and Chen, W. (2010). Surface display of a functional minicellulosome by intracellular complementation using a synthetic yeast consortium and its application to cellulose hydrolysis and ethanol production. *Appl. Environ. Microbiol.* 76, 7514–7520. 10.1128/AEM.01777-10.
- Vacheron, J. (2015). Sélection des rhizobactéries phytostimulatrices par la plante : impact sur la distribution des propriétés phytobénéfiques chez les *Pseudomonas* fluorescents. *phD Thesis. Université Claude Bernard - Lyon 1*.
- Vance, C.P. (2001). Symbiotic nitrogen fixation and phosphorus acquisition. plant nutrition in a world of declining renewable resources. *Plant Physiol.* 127, 390–397. 10.1104/pp.010331.
- Vogel, C., Bodenhausen, N., Griessem, W., and Vorholt, J.A. (2016). The *Arabidopsis* leaf transcriptome reveals distinct but also overlapping responses to colonization by phyllosphere commensals and pathogen infection with impact on plant health. *New Phytol.* 212, 192–207. 10.1111/nph.14036.

Vorholt, J.A., Vogel, C., Carlström, C.I., and Müller, D.B. (2017). Establishing causality: opportunities of synthetic communities for plant microbiome research. *Cell Host Microbe* 22, 142–155. 10.1016/j.chom.2017.07.004.

Walker, B. (1995). Conserving biological diversity through ecosystem resilience. *Conserv. Biol.* 9, 747–752. 10.1046/j.1523-1739.1995.09040747.x.

Wintermute, E.H., and Silver, P.A. (2010). Emergent cooperation in microbial metabolism. *Mol. Syst. Biol.* 6, 407. 10.1038/msb.2010.66.

Yanase, S., Yamada, R., Kaneko, S., Noda, H., Hasunuma, T., Tanaka, T., *et al.*, (2010). Ethanol production from cellulosic materials using cellulase-expressing yeast. *Biotechnol. J.* 5, 449–455. 10.1002/biot.200900291.

Zhou, K., Qiao, K., Edgar, S., and Stephanopoulos, G. (2015). Distributing a metabolic pathway among a microbial consortium enhances production of natural products. *Nat. Biotechnol.* 33, 377–383. 10.1038/nbt.3095.

Zilber-Rosenberg, I., and Rosenberg, E. (2008). Role of microorganisms in the evolution of animals and plants: the hologenome theory of evolution. *FEMS Microbiol. Rev.* 32, 723–735. 10.1111/j.1574-6976.2008.00123.x.

Zuroff, T.R., and Curtis, W.R. (2012). Developing symbiotic consortia for lignocellulosic biofuel production. *Appl. Microbiol. Biotechnol.* 93, 1423–1435. 10.1007/s00253-011-3762-9.

---

## CHAPTER 2

---



## **Deciphering maize selection of plant beneficial bacteria via an experimental evolution approach**

**Beatriz Manriquez<sup>1</sup>, Danis Abrouk<sup>1</sup>, Audrey Dubost, Florence Gerin<sup>1</sup>, Jordan Vacheron<sup>2</sup>, Daniel Muller<sup>1</sup>, Claire Prigent Combaret<sup>1</sup> \***

<sup>1</sup> Univ Lyon, Université Claude Bernard Lyon 1, CNRS, INRAE, VetAgro Sup, UMR5557 Ecologie Microbienne, 43 bd du 11 novembre 1918, F-69622 Villeurbanne, France

<sup>2</sup>Department of Fundamental Microbiology, University of Lausanne, Lausanne, Switzerland

**Running title:** Deciphering maize selection of plant beneficial bacteria via an experimental evolution approach

**\*Corresponding author:** UMR CNRS 5557 Ecologie Microbienne, Université Lyon 1, 43 bd du 11 Novembre 1918, 69622 Villeurbanne cedex, France. Phone +33 4 72 43 13 49. E-mail: [claire.prigent-combaret@univ-lyon1.fr](mailto:claire.prigent-combaret@univ-lyon1.fr)

## ABSTRACT

Plants are in association with microbial communities that strongly influence their development and their adaptation to the environment in which they evolve. Within the plant microbiota, certain bacteria, known as PGPR (for Plant Growth-Promoting Rhizobacteria), are capable of stimulating plant growth and protecting the plant against diseases via numerous plant beneficial functions. Although these bacteria produce a wide variety of beneficial functions, the mechanisms favouring the establishment of PGPRs in the rhizosphere and the expression of their properties are still largely unknown. The aim of this study was to offer insight into the evolutionary processes involved in root colonisation since in the rhizosphere, bacterial behaviour depends on the interaction networks PGPR share with other microbial populations colonizing plant roots. To get insight into the functioning of PGPR-plant cooperations, we set up an experimental evolution approach, *in vitro* and *in planta*, with a synthetic community. Our synthetic community was composed of 10 *Pseudomonas* strains of PGPR possessing a high or moderate number of plant beneficial functions. Our results show a rapid evolution, with half of the strains being lost after only 3 serial passages, and with a selection of different species between the experimental evolution performed under *in vitro* conditions and the ones performed in presence of maize as a model plant.

**Keywords:** Experimental Evolution, synthetic community, PGPR, plant microbiota, plant beneficial function contributing genes.

**Abbreviations:** experimental evolution (EE); plant growth-promoting rhizobacteria (PGPR); plant beneficial function contributing (PBFC); Plant beneficial function (PBF); hours (h); minutes (min); seconds (s); 1-aminocyclopropane-1-carboxylate deaminase (ACC deaminase)

## INTRODUCTION

Microbial communities are in perpetual evolution with their environment. At each generation, some microorganisms may acquire new functions that may increase their ability to colonise the environment, for example, by using the available resources in a more efficient way than their neighbours. The most adapted microorganisms will have an increased fitness and will thus reproduce faster than the others in the environment. These changes will disturb the balance of the microbial community, which in turn will lead to changes in other microorganisms and even, as a consequence, in the environment. The evolutionary process of adaptation of microbial communities to a new biotic or abiotic environment can only be seen through observation of variations of ecological traits of microorganisms over time (Elena and Lenski, 2003; Buckling *et al.*, 2009).

Among microorganisms, bacteria are ideal candidates to study adaptation in real time as they are very easy to handle in laboratory conditions. They can develop in liquid or solid media and their fast replication rate allows them to produce thousands of generations in little time. Bacteria have been used in experimental evolution (EE) studies since 1970, with Lenski's EE of *Escherichia coli* in a glucose-minimal medium (Lenski *et al.*, 1991) being the most striking example. In EE assays, an initial population is grown in a given environment (*i.e.* media or a eukaryotic host) during one evolutive cycle, usually defined by the time required for the initial population to consume the environmental resources. At the end of the evolutive cycle, a proportion of the population is recovered and transferred to a fresh batch of the same environment. EE allows researchers to observe real-time evolution of individuals in response to imposed environmental conditions over a multitude of identical experimental cycles (Kawecki *et al.*, 2012). This approach has been used to study an array of promising microorganisms and ecological principles (Travisano *et al.*, 1995; Adams and Rosenzweig, 2014; D'Souza and Kost, 2016; Lenski, 2017; Cairns *et al.*, 2018). However, most EE studies have focused on the evolutionary outcomes of a microbial population descending from one single organism and mainly on genetic diversity within a population. Yet, in microbial communities, ecological and evolutionary processes can be altered due to more complex competitive interactions occurring between populations and multiple selection pressures than within a single population.

Recent studies have highlighted the impact communities can have on pairwise interactions and on the adaptation rate of species co-existing within communities (Smith and Schuster, 2019; Manriquez *et al.*, submitted). Moreover, multispecies communities house some key ecological features such as diversity, stability, succession and high order species interactions that are not observable in single organism studies. Therefore, there is an increasing need to expand the biotic complexity of EE setups in order to unravel the mechanics of natural microbial communities in ecosystems (De Roy *et al.*, 2014). In a more applied point of view, it can also allow to engineer less complex synthetic microbial communities that can successfully perform a function when re-

introduced in their natural habitat. Among the most studied natural systems, plant roots have drawn attention as they interact with many soil microorganisms offering a rich and complex network of interactions. Within the plant-selected microbiota, some microorganisms inhabiting the rhizosphere can benefit to the plant and contribute to its optimal development. These beneficial interactions can be carried by mutualistic organisms or by free living plant-beneficial bacteria able to colonize the roots. This last category includes plant growth-promoting rhizobacteria (PGPR, Kloepper *et al.*, 1989) which promote plant growth and health via a large range of plant beneficial function contributing genes (PBFC genes/clusters; Bruto *et al.*, 2014). PBFC genes/clusters can be divided into 3 main categories: (1) plant nutrition, (2) plant hormonal balance and (3) antimicrobial compounds and siderophores. As non-strict mutualistic bacteria, PGPR are able to interact with a large panel of host plant species and have drawn attention in the last decades. Despite the progress that has been made, understanding the impact of PGPR on roots and on the plant still remains challenging; particularly when it comes to the question of relationships between plant beneficial properties' number and the effect on plants. A higher number was thought to provide better effects on the plant (Bashan and de-Bashan, 2010) as the effects of different modes of action may add up, or as some PBFC genes/clusters would be expressed in particular environmental conditions, which could imply a larger prevalence of PGPR with a higher number of PBFC genes/clusters in the rhizosphere. However, a previous work (Vacheron *et al.*, 2016) proved that maize preferentially selected *Pseudomonas* with few PBFC genes/clusters (up to 5) and results also suggested a trade-off between rhizosphere prevalence and the ability to maintain a large number of PBFC genes/clusters in such a competitive environment, leading to a great majority of rhizosphere isolates displaying a low number of co-occurring PBFC genes/clusters.

Here we wanted to test if maize selects PGPR at a taxonomical or functional level, and if the number of PBFC genes/clusters from PGPR populations matters in the selection of particular assembly of PGPR. To monitor the selection dynamics, we carried out two EE assays with a synthetic community composed of 10 *Pseudomonas* PGPR isolated from maize rhizosphere: firstly, on a liquid rich medium in absence of the host plant (*in vitro*) and secondly, on axenic maize seedlings (*in planta*). Half of the isolates possessed a low number of plant beneficial properties (10 or less) and the other half a high number (more than 11), giving a functional redundancy within the initial (or ancestral) synthetic community. The EE was carried out *in vitro* for 370 and *in planta* for 400 bacterial generations and the evolutionary dynamics of the synthetic community were assessed by a metabarcoding approach on the *rpoD* gene. We compare the simultaneous population dynamics of the 10 PGPR from the synthetic community evolving over time in a commonly used rich medium and on alive maize roots, as well as the impact of these evolutionary dynamics on PBFC genes/clusters distribution.

## **MATERIAL AND METHODS**

### **Bacterial strains and growth conditions**

The synthetic community was composed of 10 *Pseudomonas* strains (*i.e.* *P. jessenii* JV251A, *P. inefficax* JV551A1, *P. kribbensis* JV359A8, *P. fluorescens* JV391D10, JV449, JV391D17, *P. wadenswilerensis* JV244B, *P. brassicacerum* JV551A7, *P. protegens* JV245A and JV359A1). All the strains were isolated from maize rhizosphere in a previous study (Vacheron *et al.*, 2016) and screened for plasmid presence using the Eckhardt gel technique (Hynes *et al.*, 1985). All the strains were initially screened for the ability to grow individually at 28°C in two liquid media: Tryptic Soy Broth (TSB; Conda, Pronadisa, Spain) and in a maize extract-based medium. This medium was prepared by incubating surface-disinfected maize seeds (n=10) in 20 mL of sterilized water for 72 h. The water was retrieved by freeze-drying during 72 h (Christ Alpha 1-4, Fisher Bioblock, Illkirch, France), and the remaining extracts were suspended in 5 mL of sterile water. For the EE assays, all the strains were grown overnight in liquid TSB medium at 28°C under agitation at 300rpm and 1 mL of a suspension at 10<sup>5</sup> CFU/mL (estimated *via* optical density measures and calibration, data not shown) of each strain was used to create an initial synthetic community composed of the 10 strains in equivalent proportions. At the end of each cycle of both experimental *in vitro* and *in planta* evolutions, strains were enumerated on King's B agar medium (Conda, Pronadisa, Spain).

### **Plant material and growth conditions**

*Zea mays* seeds of the PR37Y15 genotype (Pioneer Semences SAS, France) were surface-sterilised in a 0.34% chlorine active solution for 45 min followed by 5 washes in sterilised water and a 2 h incubation in the water of the last wash. To check the surface-sterility of the seeds, 250 µL of the last wash were spread on King's B agar medium and the plates were incubated at 28°C. The disinfected seeds were then placed on plant agar plates (4 g/L), embryo face up, and incubated in the dark for 72 h at 28°C with 75% humidity. Germinated seeds with 1 to 2 cm long roots were selected to be further used. Each selected germinated seed was placed in tubes containing 30 mL of plant agar (2 g/L) and incubated in a growth chamber for 24 h at 75% humidity, 16 h of light and 8 h of darkness, and 25°C. With this experimental setup roots of maize seedlings grew in a plant agar based gnotobiotic system with no added nutritive solution.

### ***In vitro* experimental evolution**

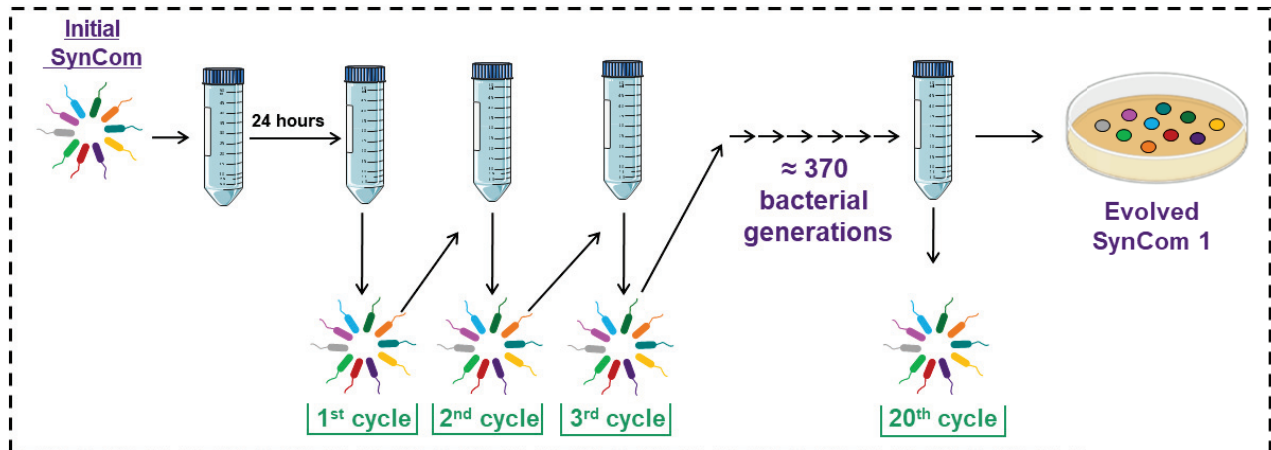
The initial synthetic community (initial SynCom) was inoculated ( $v = 50 \mu\text{L}$ ) in 5 tubes (*i.e.* 5 independent series in parallel) containing 9.950 mL of TSB media. The inoculated tubes were incubated at 28°C for 24 h on a static rack in order to avoid agitation induced bias and to mimic rhizosphere lifestyle. The next day, equivalent to one cycle of the *in vitro* EE, each tube was centrifuged at 5000 rpm for 4 min at room temperature. Bacterial pellets were suspended into 1

mL which were subsequently divided into different fractions. A 250  $\mu$ L bacterial suspension was stored into glycerol (50% v/v) at  $-80^{\circ}\text{C}$ ; a 100  $\mu$ L suspension was serially diluted to approximately  $10^5$  CFU/mL and 50  $\mu$ L of the dilution were re-inoculated into 9.950 mL of fresh TSB media. Serial dilutions were plated onto King's B and incubated at  $28^{\circ}\text{C}$  in order to determine the output concentration and to calculate the number of generations obtained after 24 h of incubation. The number of generations obtained at each cycle was calculated by comparing the number of cells inoculated at the beginning of the cycle and the number of cells recovered after 24 h. At each cycle, DNA samples were extracted and stored. This EE was conducted five times simultaneously with five series of independent conical tubes (*in vitro* SynCom 1 to 5, figure 1). Each serial passage included a non-inoculated tube that received 50  $\mu$ L of sterilised water instead of the bacterial suspension. The population retrieved at the 20<sup>th</sup> cycle is referred to as the evolved *in vitro* SynCom.

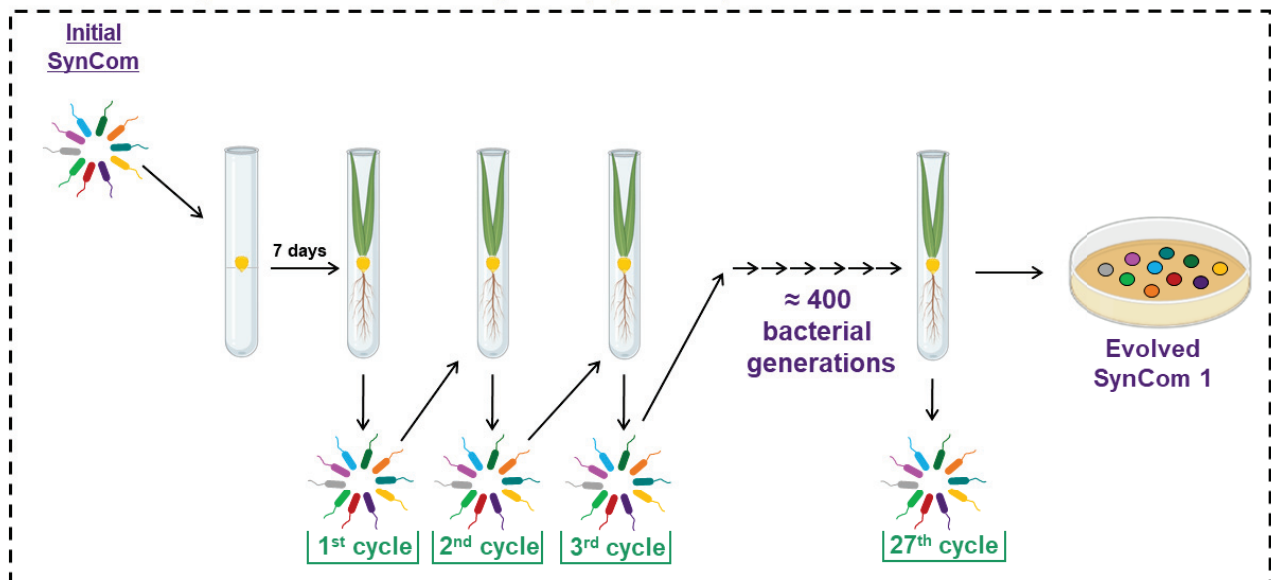
### ***In planta* experimental evolution**

The initial SynCom was inoculated on 7 germinated seeds (*i.e.* 7 independent series in parallel), placed 24 h earlier in tubes containing 30 mL of plant agar. The inoculated plants were incubated in a growth chamber for 7 days under the previously described conditions. After seven days (equivalent to one cycle of the *in planta* EE), roots of maize seedlings from each plant were vortexed in 5 mL of sterile water plus 2 metal beads, for 5 min, in order to detach bacteria from the roots. The recovered bacterial suspension was centrifuged and condensed into 1 mL, and 100  $\mu$ L of the suspension were diluted to approximately  $10^5$  CFU/mL and 50  $\mu$ L of the dilution were re-inoculated onto 7 newly germinated seeds. Serial dilutions of the suspension were plated onto King's B medium and incubated at  $28^{\circ}\text{C}$  in order to determine the output concentration and to calculate the number of generations obtained after 7 days of incubation. The number of generations obtained at each cycle was estimated by comparing the effective number of cells injected at the beginning of the cycle and the effective number of cells recovered after seven days. Nonetheless, the number of generations, and therefore the population size, might be underestimated as the bacterial cells were not recovered from the whole plant but from roots, and the recovering protocol may not be exhaustive. At each cycle, 250  $\mu$ L of the recovered bacterial suspension were stored into glycerol (50% v/v) at  $-80^{\circ}\text{C}$ . DNA samples were also stored at each cycle. This EE was conducted seven times simultaneously, with seven series of independent maize plants (*in planta* SynCom 1 to 7, figure 1). Each serial passage included 3 non-inoculated plants that received only sterilised water. The root population retrieved at the 27<sup>th</sup> cycle is referred to as the evolved *in planta* synthetic community.

### A series of experimental evolution *in vitro* (SynCom 1)



### A series of experimental evolution *in planta* (SynCom 1)



**Figure 1: Experimental evolution set up.** 5 independent series of experimental evolution were carried for the *in vitro* experience and 7 series for the *in planta* experience. For each series of independent *in vitro* and *in planta* experimental evolutions, the ancestral mix was inoculated in one conical tube or on one pre-germinated maize seed and incubated for 24 hours or 7 days, respectively (=one cycle). Promptly, the bacterial suspensions or the roots populations were recovered, condensed and diluted before being inoculated onto a new conical tube or a new pre-germinated seeds. Overall, there are 5 SynCom for the *in vitro* EE and 7 SynCom for the *in planta* EE.



## Microbiota diversity analysed through *rpoD* diversity analysis

The *rpoD* gene coding for RNA polymerase sigma 70 (sigma D) factor was chosen to distinguish all the strains one from another. All the *rpoD* genes were aligned and primers (*rpoD*\_F: TCGCCAAGAAGTACACCAAC and *rpoD*\_R: CCATGGAGATCGGCTCTT) were designed to amplify a 353bp fragment of *rpoD* (position 1179 to 1532). In the 353bp fragments, from 1201 to 1513, 32 positions displayed different nucleotides in at least in two *Pseudomonas* (Table 1). Microbiota total DNA was extracted from the samples (*in vitro* SynCom 1 to 5 and *in planta* SynCom 1 to 7 at different cycles of evolution) using a NucleoSpin Tissue extraction kit (Macherey-Nagel, Germany). Extracted DNA concentration was quantified using a NanoPhotometer apparatus (Implen, Germany) and subsequently diluted or concentrated to at least 15 ng/μL.

**Table 1: Different positions of nucleotide variations between the 10 studied strains of *Pseudomonas* for the *rpoD* gene.** Primers (*rpoD*\_F: TCGCCAAGAAGTACACCAAC and *rpoD*\_R: CCATGGAGATCGGCTCTT) were designed to amplify a fragment from positions 1179 to 1532 allowing to distinguish the 10 strains on a minimum of 1 position between strains.

	1201	1210	1219	1240	1243	1244	1252	1255	1258	1279	1282	1288	1297	1300	1315	1318
<i>P. jessenii</i> JV251A	T	A	T	C	C	T	A	G	A	C	C	A	T	T	C	T
<i>P. inefficax</i> JV551A1	T	G	C	C	C	T	A	G	G	C	C	G	C	C	C	C
<i>P. kribbensis</i> JV359A8	T	A	T	C	T	C	A	G	G	T	T	G	T	T	C	T
<i>P. fluorescens</i> JV391D10	T	A	T	C	C	T	G	T	G	C	C	G	T	T	C	T
<i>P. fluorescens</i> JV449	T	A	T	C	C	T	G	T	G	C	C	G	T	T	C	T
<i>P. fluorescens</i> JV391D17	T	A	T	C	C	T	G	T	G	C	C	G	T	T	C	T
<i>P. wadenswilerensis</i> JV244B	C	A	C	C	C	C	G	G	G	T	T	G	C	T	T	C
<i>P. brassicacerum</i> JV551A7	C	A	C	T	C	C	G	G	G	C	C	G	C	T	C	T
<i>P. protegens</i> JV245A	T	G	C	C	C	T	A	C	G	C	C	G	T	T	C	T
<i>P. protegens</i> JV359A1	T	G	C	C	C	T	A	C	G	C	C	G	T	T	C	T

	1330	1351	1360	1363	1366	1384	1405	1414	1421	1447	1462	1486	1489	1492	1502	1513
<i>P. jessenii</i> JV251A	T	C	C	T	T	G	C	C	T	C	C	C	G	C	T	T
<i>P. inefficax</i> JV551A1	C	A	C	T	T	G	T	C	C	C	C	C	A	C	C	C
<i>P. kribbensis</i> JV359A8	T	C	C	T	T	G	C	C	C	C	T	C	G	C	T	C
<i>P. fluorescens</i> JV391D10	T	C	C	T	T	A	T	G	T	T	C	T	A	C	T	T
<i>P. fluorescens</i> JV449	T	C	C	T	T	G	T	G	T	C	C	T	A	C	T	T
<i>P. fluorescens</i> JV391D17	T	C	C	T	T	G	T	G	T	T	C	T	A	C	T	T
<i>P. wadenswilerensis</i> JV244B	C	T	C	T	T	G	C	G	T	C	T	C	G	C	T	C
<i>P. brassicacerum</i> JV551A7	T	C	C	C	C	G	T	G	C	G	C	C	G	T	T	T
<i>P. protegens</i> JV245A	T	C	C	T	T	G	T	G	C	T	C	C	G	C	T	C
<i>P. protegens</i> JV359A1	T	C	T	T	T	G	T	G	C	T	C	C	G	C	T	C

All of the analysed samples underwent one DNA extraction. The same initial SynCom was used for both EE and its composition was assessed 4 times in total. For each EE, the initial SynCom underwent 2 parallel DNA extractions, and independent *rpoD* sequencing analyses were carried for each extraction. Samples are referred to by the series of EE (*i.e.* SynCom 1) and by the time of recovery (*i.e.* cycle 3), giving the following code: SynCom 1-3 for the population recovered from the first series of EE at the end of the 3<sup>rd</sup> cycle of evolution. The total number of analysed samples per cycle and per independent series of EE can be found in Table 2.

**Table 2: Analyzed samples for both experimental evolutions.** Samples are referred to by the series of EE (*i.e.* SynCom 1) and by the time of recovery (*i.e.* cycle 3). Giving the following code, SynCom 1-3 designates the population recovered from the first series of EE at the end of the 3rd cycle of evolution.

		EXPERIMENTAL EVOLUTION <i>in vitro</i>			
Series of EE		CYCLE 1	CYCLE 5	CYCLE 15	CYCLE 20
SynCom 1					
SynCom 2					
SynCom 3					
SynCom 4					
SynCom 5					

		EXPERIMENTAL EVOLUTION <i>in planta</i>						
Series of EE		CYCLE 1	CYCLE 2	CYCLE 3	CYCLE 4	CYCLE 5	CYCLE 6	CYCLE 27
SynCom 1								
SynCom 2								
SynCom 3								
SynCom 4								
SynCom 5								
SynCom 6								
SynCom 7								

DNA extracts were sent to MR DNA laboratory ([www.mrdnalab.com](http://www.mrdnalab.com); Shallowater, TX) for sequencing. For *rpoD* sequencing, PCR primers *rpoD\_F*/*rpoD\_R* were used, the forward primer carrying a barcode. Primers were used in a 30-cycle PCR (5 cycles implemented on PCR products), using the HotStarTaq Plus Master Mix Kit (Qiagen, Valencia, CA) under the following conditions: 94°C for 3 min, followed by 28 cycles of 94°C for 30 s, 53°C for 40 s and 72°C for 1 min, with a final elongation step at 72°C for 5 min. PCR products were checked in 2% agarose gel to determine amplification success and relative band intensity. Multiple samples were pooled together in equal proportions based on their molecular weight and DNA concentrations. Pooled samples were purified using calibrated Ampure XP beads and used to prepare a DNA library following Illumina TruSeq DNA library preparation protocol. Sequencing was performed on a MiSeq following the manufacturer's guidelines.

Sequence data were processed using the analysis pipeline of MR DNA. Briefly, sequences were depleted of barcodes; sequences < 150 bp or with ambiguous base calls were removed. The remaining sequences were denoised, operational taxonomic units (OTUs; defined at 0% divergence threshold for the *rpoD* gene) generated, and chimeras removed. Final OTUs were classified using an in-house *rpoD* database, containing the 10 *rpoD* sequences from the *Pseudomonas* strains used in this study. For all analyzed samples, the ratios between the number of *rpoD* sequences assigned to one strain and the total number of *rpoD* sequences per samples were calculated and expressed in percent.

## ***Pseudomonas* genome sequencing and PBFC gene annotation**

Genomic DNA extraction was done from an overnight culture in TSB medium at 37°C using a Nucleospin tissue kit (Macherey-Nagel, France). Genomic DNA was sequenced at MR DNA (Shallowater, TX, USA) using Illumina MiSeq technology, generating a 2X300-bp paired-end library. NGen version 14 software (DNASTar, Inc.) was used for trimming sequences (default settings) and for *de novo* assembly. Genome annotation was done with the MicroScope platform version 3.14.1 (Vallenet *et al.*, 2020), coverage depth varied from 86X (*P. fluorescens* JV449) to 440X (*P. fluorescens* JV391D17), and number of contigs from 9 (*P. protegens* JV359A1) to 206 (*P. wadenswilerensis* JV244B).

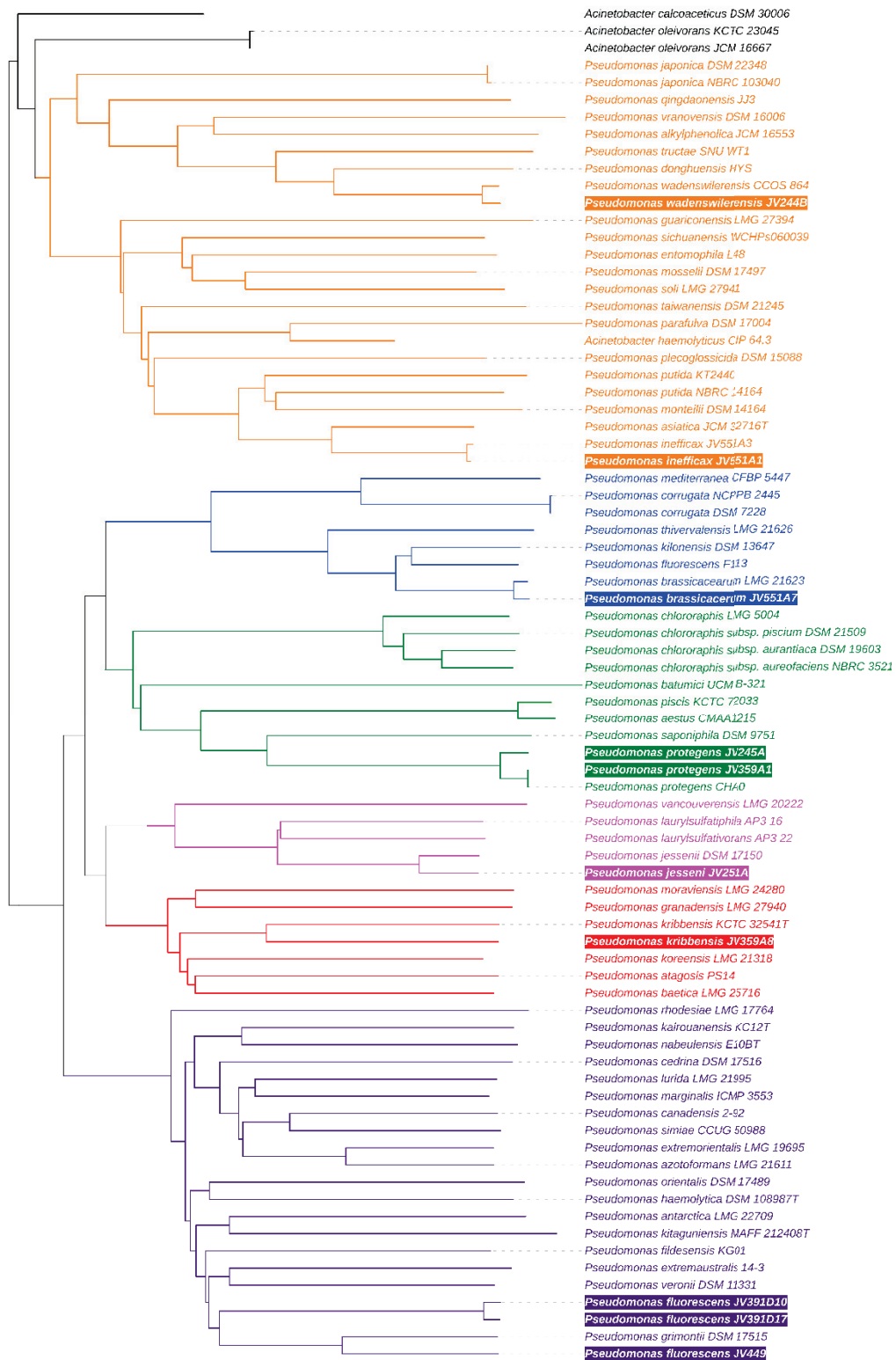
Gene cluster coding for secondary metabolites was identified using antiSMASH version 5.0.0 and MIBiG version 1.4 (Medema *et al.*, 2011; Blin *et al.*, 2019; Kautsar *et al.*, 2020). PBFC genes/clusters distribution and annotation were assessed for all of the strains used in this study by comparing (*i.e. via* blastP) protein sequences (coverage >80%; identity >50% in most cases) with protein sequences of closest reference strains on the MicroScope platform version 3.14.1 (Vallenet *et al.*, 2020). For *P. fluorescens* JV391D10, JV391D17 and JV449 and *P. brassicacerum* JV551A7 the reference strain was *P. kilonensis* F113. For strains *P. protegens* JV245A and JV359A1 and *P. jessenii* JV251A the reference strain used was *P. protegens* CHA0. For strain *P. inefficax* JV551A1 the reference strain used was *P. putida* KT2440. For *P. wadenswilerensis* JV244B it was *P. putida* KT2440, *P. kilonensis* F113, *P. fluorescens* Pf-5 and *P. aeruginosa* PAO1. Lastly, for strain *P. kribbensis* JV359A8, protein sequences from *P. kilonensis* F113, *P. fluorescens* Pf-5 and *P. fluorescens* SS101 were used based on the phylogenetic affiliation of the strain and as no *P. kribbensis* genome has been accurately annotated and listed on the MicroScope platform. All in all, the strains from this study harbour 4 (1 isolate) to 17 co-occurring PBFC genes/clusters (*i.e.* genes harboured by a same isolate, Table 3 and corresponding PBFC coding loci in Table S1). Half of the isolates possessed a low number of plant beneficial properties (10 or less) and the other half a high number (11 or more), giving a functional redundancy within the initial SynCom.



## Phylogenetic tree reconstruction

A phylogenetic tree (Figure 2) was reconstructed using the Type Strain Genome Server. The genome sequence data were uploaded to the Type (Strain) Genome Server (TYGS; <https://tygs.dsmz.de>), for a whole genome-based taxonomic analysis (Meier-Kolthoff and Göker, 2019). In brief, the TYGS analysis was subdivided into the following steps. First, determination of closest type strain genomes was done in two complementary ways. All our 10 *Pseudomonas* strain genomes were compared against all type strain genomes available in the TYGS database via the MASH algorithm, a fast approximation of intergenomic relatedness (Ondov *et al.*, 2016), and, the ten type strains with the smallest MASH distances were chosen per our *Pseudomonas* strain genomes. Then, an additional set of ten closely related type strains was determined via the 16S rDNA gene sequences. Thus, 16S rDNA gene sequences were extracted from the studied *Pseudomonas* genomes using RNAmmer (Lagesen *et al.*, 2007) and each sequence was subsequently BLASTed (Camacho *et al.*, 2009) against the 16S rDNA gene sequence of each of the type strains available in the TYGS database. This was used as a proxy to find the best 50 matching type strains (according to the bitscore) for each studied *Pseudomonas* genome and to subsequently calculate precise distances using the Genome BLAST Distance Phylogeny approach (GBDP) under the algorithm 'coverage' and distance formula d5 (Meier-Kolthoff *et al.*, 2013). These distances were finally used to determine the 10 closest type strain genomes for each of our 10 *Pseudomonas* strain genomes. Second, all pairwise comparisons among the set of genomes were conducted using GBDP and accurate intergenomic distances inferred under the algorithm 'trimming' and distance formula d5 (Meier-Kolthoff *et al.*, 2013). One hundred distance replicates were calculated each. Digital DDH values and confidence intervals were calculated using the recommended settings of the GGDC 2.1 (Meier-Kolthoff *et al.*, 2013). Third, the resulting intergenomic distances were used to infer a balanced minimum evolution tree with branch support via FASTME 2.1.4 including SPR post processing (Lefort *et al.*, 2015). Branch support was inferred from 100 pseudo-bootstrap replicates each. The trees were rooted at the midpoint (Farris, 1972) and visualized with PhyD3 (Kreft *et al.*, 2017). Fourth, the type-based species clustering using a 70% dDDH radius around each of the 97 type strains was done as previously described (Meier-Kolthoff and Göker, 2019). The resulting groups are shown in Table S2 and S3. Subspecies clustering was done using a 79% dDDH threshold as previously introduced (Meier-Kolthoff *et al.*, 2014).

Tree scale: 0.01



**Figure 2: Phylogenetic tree of the ten *Pseudomonas* strains used in this study.** Tree inferred with FastME 2.1.6.1 from GBDP distances calculated from genome sequences. The branch lengths are scaled in terms of GBDP distance formula d5 (100 iterations). The tree was rooted at the midpoint. Each subgroup of *Pseudomonas* is presented in different colors. The highlighted names correspond to the 10 strains of this study.



## RESULTS

### General genomes features

Genome size fluctuate from 5889653 bp (*P. inefficax* JV551A1) to 7442190 bp (*P. wadenswilerensis* JV244B). The average G+C content is 61.57%. No plasmids were detected for any of the strains using the Eckhardt gel technique (Hynes *et al.*, 1985). The studied genomes harbour between 5782 (*P. inefficax* JV551A1) to 7132 (*P. wadenswilerensis* JV244B) coding sequences, among which 76% to 79% are classified in at least one cluster of orthologous group (COG). Strains present a variable number of rRNA (5s\_rRNA, 16s\_rRNA and 23s\_rRNA), whereas 55 tRNA (*P. brassicacerum* JV551A7 and *P. fluorescens* JV449) to 64 (*P. protegens* JV245A and *P. inefficax* JV551A1) were detected. Furthermore, between 5 (*P. inefficax* JV551A1) to 23 (*P. wadenswilerensis* JV244B) biosynthesis gene clusters for secondary metabolites were predicted with AntiSMASH (Table S4). The general genome data for each strain is presented in Table 4.

### Plant beneficial properties distribution within the initial SynCom

Distribution of PBFC genes/clusters was searched in genomes of all the strains (Table 4 and corresponding PBFC coding loci in Table S1). Clusters were classified into 3 categories of PBF: plant nutrition; plant hormonal balance; antimicrobial compounds and siderophores. In addition, strains genomes were also screened for the presence of additional genes involved in biotic interactions, like genes encoding entomotoxins, secretion systems and other non-ribosomal peptides synthetase (NRPS) that were not classified in literature as plant beneficial functions (PBF) but that are contributing to PGPR establishment in the rhizosphere. All of them were throughout designed as PBFC genes/clusters. The strains were classified so half of them possess a low number of PBFC genes/clusters (10 or less; strains JV251A, JV551A1, JV359A8, JV391D10, and JV449) and the other half harbour a high number of PBFC genes (11 or more; strains JV391D17, JV244B, JV551A7, JV245A, JV359A1). The minimal number of PBF harboured per strain is 4 (for JV251A) and the maximal number is 17 (for JV359A1). In the case of auxin production, the strains selected for this study harbour either the indole-3-acetamide pathway (*iaaM*) and/or the indole-3-pyruvate pathway (*ipdC*). Of a total of 20 PBFC genes/clusters, only 4 (phosphate solubilisation (*gcd*), auxin (*iaaM* and/or *ipdC*) and pyoverdine production (*pvd*) are shared by all of the strains.



**Table 4: General genome features of the 10 studied strains of *Pseudomonas*.** These features were recovered using the Microscope platform. CDS: Coding DNA Sequences; COG: Clusters of Orthologous Group. For AntiSMASH, the total number of clusters were reported.

	total genome length	total number of bases	coverage depth	total number of reads	GC%	contig number	genomic objects	CDS	protein coding density (%)
<i>P. jessenii</i> JV251A	6509003	2363623234	363	7826567	59.66	60	6353	6197	88.22
<i>P. inefficax</i> JV551A1	5889653	665698718	113	2458024	62.75	55	5949	5782	89.25
<i>P. kribbensis</i> JV359A8	6337157	2745463573	433	6006245	60.66	24	6292	6128	90.4
<i>P. fluorescens</i> JV391D10	7165607	2730939640	381	6043047	60.54	24	7094	6928	89.67
<i>P. fluorescens</i> JV449	6832692	590112499	86	1067395	60.92	35	6733	6558	90.72
<i>P. fluorescens</i> JV391D17	6921048	3045199502	440	6408128	60.68	25	6797	6640	89.92
<i>P. wadenswilerensis</i> JV244B	7442190	2285210746	307	7566923	62.76	206	7280	7132	90.92
<i>P. brassicacerum</i> JV551A7	6735578	773790040	115	1600226	60.9	32	6655	6470	89.51
<i>P. protegens</i> JV245A	6383769	1452335719	228	4902011	63.44	83	7659	7008	83.5
<i>P. protegens</i> JV359A1	6850953	2613742377	382	5718299	63.42	9	6594	6448	89.7

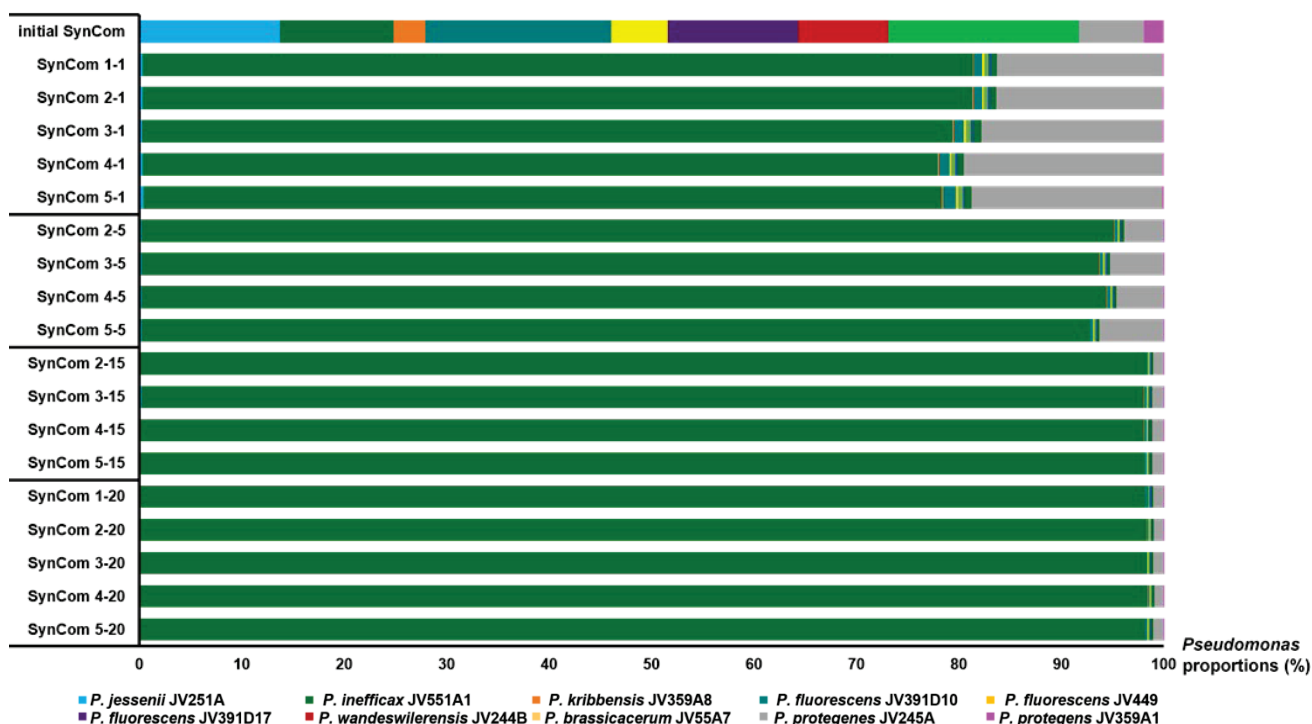
  

COG classification								
	COG classification percentage (%)	Cellular processes and signaling	Information storage and processing	Metabolism	Poorly characterized	16s_rRNA	tRNA	AntiSMASH
<i>P. jessenii</i> JV251A	77.58	19.35	15.57	40.08	20.30	1	58	9
<i>P. inefficax</i> JV551A1	76.32	20.20	15.42	38.58	19.92	4	64	5
<i>P. kribbensis</i> JV359A8	76.26	21.41	14.72	36.75	20.85	1	65	11
<i>P. fluorescens</i> JV391D10	77.39	21.30	15.83	40.12	19.43	3	56	14
<i>P. fluorescens</i> JV449	78.48	22.51	15.03	40.54	19.73	2	55	10
<i>P. fluorescens</i> JV391D17	77.47	21.72	15.49	40.52	19.05	2	56	16
<i>P. wadenswilerensis</i> JV244B	79.44	22.09	15.91	39.38	20.84	2	54	23
<i>P. brassicacerum</i> JV551A7	76.11	21.56	15.01	38.33	20.12	1	55	10
<i>P. protegens</i> JV245A	76.97	21.47	15.04	39.98	19.63	4	64	16
<i>P. protegens</i> JV359A1	79.2	21.08	15.82	40.29	21.68	1	60	15

## Figuring out the initial SynCom composition

First, the individual growth of the 10 strains on a maize extract based medium was found identical showing that all the *Pseudomonas* strains could grow in presence of compounds released by maize seeds (Figure S1). To establish the initial SynCom, attempted equal content of cells of each strain (estimated based on OD<sub>600</sub> and calibration curves established for all strains between OD<sub>600</sub> and CFU) were mixed. *rpoD* based diversity of the initial SynCom was analysed 4 times in total at 2 independent times (*i.e.* when doing each EE) and similar results were obtained for both EE: 34924 and 34887 reads (*in vitro* EE) and 49864 and 37911 reads (*in planta* EE, Table S5). The mean was calculated for the 4 analysed samples (Table S6). The obtained strain proportions were as it follows (Figure 3): *P. jessenii* JV251A (13.77 ± 1.36 %), *P. inefficax* JV551A1 (11.09 ± 1.18 %), *P. kribbensis* JV359A8 (3.05 ± 0.49 %), *P. fluorescens* JV391D10 (18.14 ± 0.55 %), JV449 (5.56 ± 0.20 %), JV391D17 (12.78 ± 0.19 %), *P. wadenswilerensis* JV244B (8.75 ± 0.92 %), *P. brassicacerum* JV551A7 (18.63 ± 0.16 %), *P. protegens* JV245A (6.32 ± 1.28 %) and JV359A1 (1.91 ± 0.12 %).

As the presence/absence of PBFC genes varies depending on the strains (e.g. the indole-3-pyruvate pathway found in only 2 strains), the composition of the initial SynCom will determine how represented each PBFC gene/cluster is (Figure 4). The representation of a PBFC gene/cluster is the addition of the proportions of strains carrying this PBFC gene/cluster.

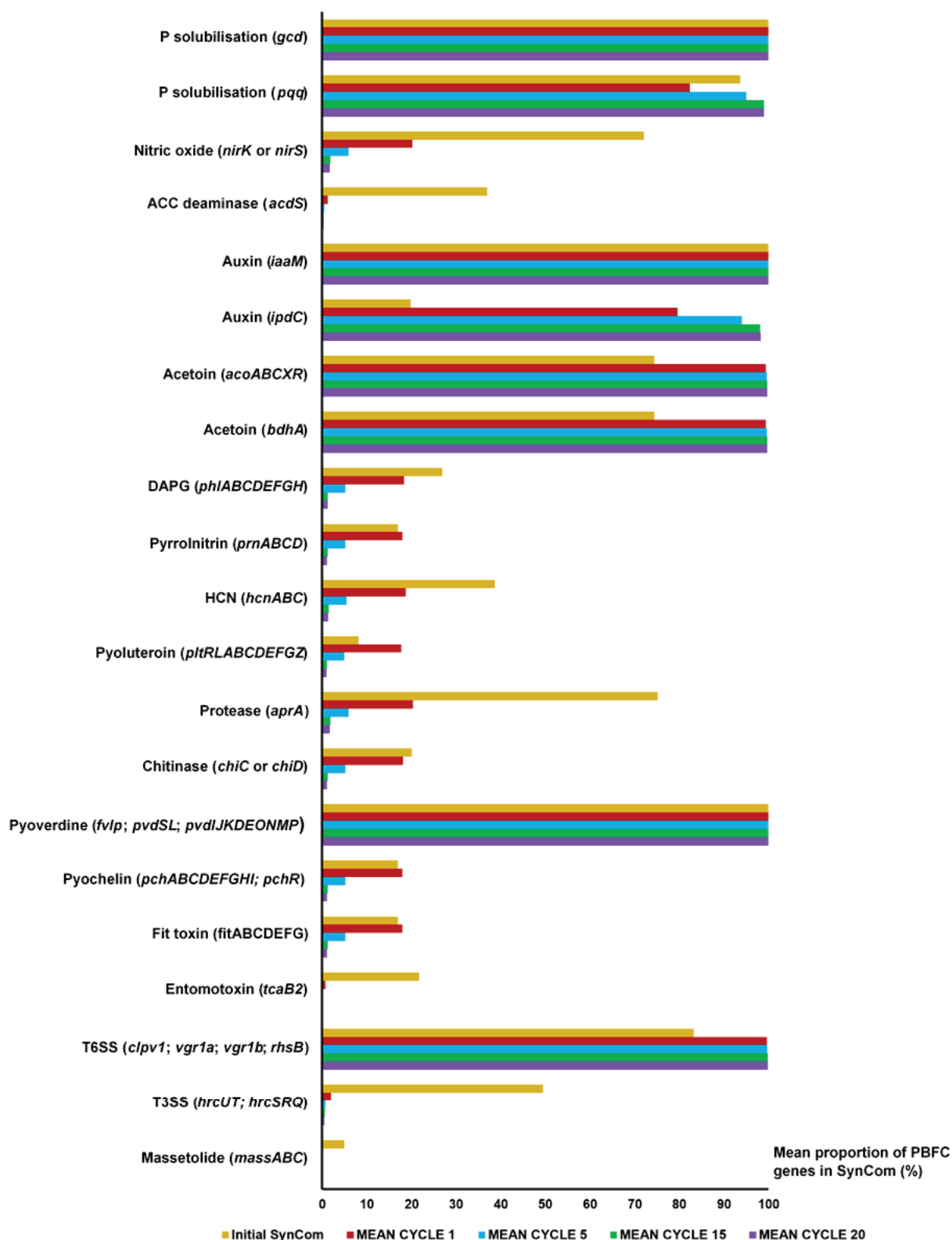


**Figure 3: *Pseudomonas* population proportions obtained for the initial SynCom and for the *in vitro* EE for SynCom 1 to 5 and for cycles 1, 5, 15 and 20.** The proportions are expressed in percent and represent the number of *rpoD* sequences assigned to one strain divided by the total number of *rpoD* sequences in each sample.

Indeed, the *phl* operon (*i.e.* 2,4-diacetylphloroglucinol (DAPG) production) is represented at 26.87% in the initial SynCom as this PBFC cluster is carried by 3 strains: *P. brassicacerum* JV551A7 (18.63%), *P. protegens* JV245A (6.32%) and JV359A1 (1.91%). Overall, 4 PBFC genes/clusters: phosphate solubilisation (*gcd* and *pqq*), auxin (*iaaM*) and pyoverdine (*pvd*) production are over represented (>90%) as they are carried by the majority of the strains.

### ***In vitro* experimental evolution**

The population dynamics of the *Pseudomonas* SynCom grown serially in TBS medium (5 independent series of EE) were analysed via *rpoD* metabarcoding of samples from cycles 1, 5, 15 and 20. The obtained number of reads and proportions are presented in Table S7 and S8. Globally, the 5 series of EE converge to the same population distribution with *P. inefficax* JV551A1 becoming the predominant strain in all of the 5 EE (Figure 3).



**Figure 4: Average PBFC genes/clusters distribution for the initial SynCom (yellow) and for the *in vitro* EE SynCom recovered at the cycles 1, 5, 15 and 20.** Results are expressed in percent and represent the average percent of all strains possessing the PBFC genes/clusters in all the SynCom (1 to 5) at each studied cycles. For each PBC, the genes or cluster of genes are indicated into brackets. P: Phosphate.

### *Pseudomonas* population abundances after the 1<sup>st</sup> cycle of the *in vitro* EE

Population composition evolved drastically (Figure 3) after only 24 h (= end of cycle 1). *P. inefficax* JV551A1 became the preponderant strain for all of the analysed SynCom ( $79.33 \pm 1.61$  %). The second more abundant strain was *P. protegens* JV245A ( $17.65 \pm 1.44$  %). All the other strains (*P. jessenii* JV251A, *P. kribbensis* JV359A8, *P. fluorescens* JV391D10, JV391D17 and JV449, *P. wadenswilerensis* JV244B, *P. brassicacerum* JV551A7, and *P. protegens* JV359A1) were poorly detected after 24 h ( $0.08 \pm 0.02$  % to  $0.98 \pm 0.19$  %) whereas these strains together represented 82.58% of the initial SynCom. These changes in population abundances impact the distribution of PBFC genes/clusters (figure 4) as the dominant strain *P. inefficax* JV551A1 carries 8 PBFC genes/clusters. This selection leads to an evolved SynCom with the similar PBFC genes/clusters distribution as the initial SynCom (>80%) with presence of genes for phosphate solubilisation (*gcd* and *pqq*), auxin (*iaaM*) and pyoverdine production, in addition to acetoin metabolism (*aco* and *bdhA*), auxin (*ipdC*) production and type 6 secretion system (T6SS), three properties carried by *P. inefficax* JV551A1. The second strain, *P. protegens* 245A (17.65%), harbours one of the highest number of PBFC genes/cluster and allows for other PBF to be represented in the evolving SynCom but with a lower abundancy. PBFC genes/clusters that are not carried by either of these 2 strains saw their proportions decrease drastically: 1-aminocyclopropane-1-carboxylic acid (ACC) deaminase (*acdS*), entomotoxin (*tcaB*), type 3 secretion system (T3SS) and massetolide coding genes.

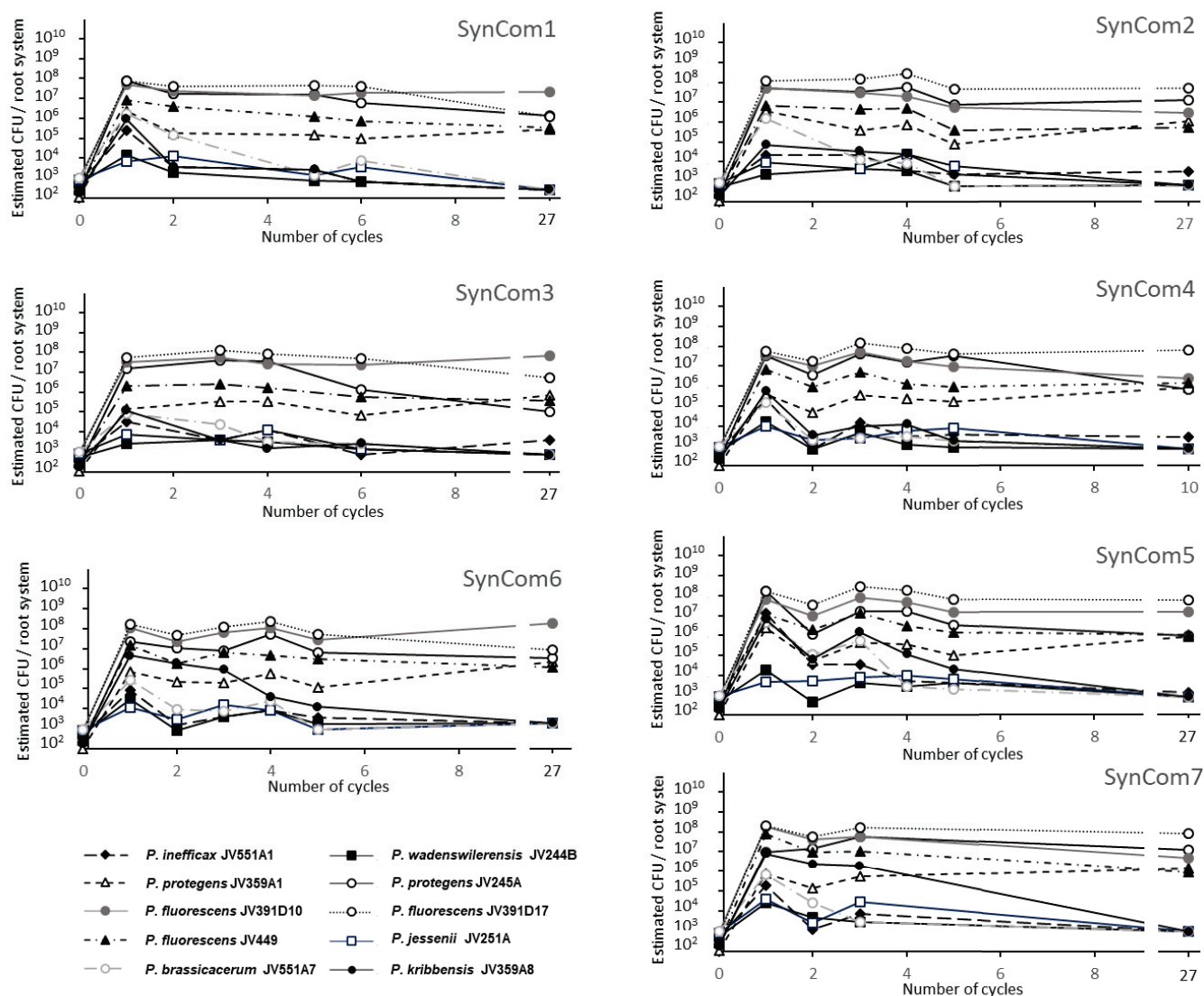
### *Pseudomonas* population dynamics after 5, 15 and 20 cycles of the *in vitro* EE

As to get a more comprehensive view of the population dynamics, SynCom were also analysed at 3 other cycles (5, 15 and 20) and similar results were obtained: *P. inefficax* JV551A1 remained the predominant strain for all of the SynCom and for all of the cycles (Figure 3). This strain's proportions increased from  $93.82 \pm 0.52$  % (5<sup>th</sup> cycle) to  $98.13 \pm 0.05$  % (20<sup>th</sup> cycle). For *P. protegens* JV245A, proportions rapidly decreased across the cycles: from  $4.98 \pm 0.51$  % (5<sup>th</sup> cycle) to  $0.98 \pm 0.03$  % (20<sup>th</sup> cycle). However, this strain still remains the second most abundant as all of the others strains were barely detectable (0.00% to 0.30%) in the 5<sup>th</sup>, 15<sup>th</sup> or 20<sup>th</sup> cycle (Figure 3). PBFC genes/clusters distribution remains globally the same, as *P. inefficax* JV551A1 was still the predominant strain for all of the SynCom (Figure 4). Moreover, as *P. protegens* JV245A proportions decreased across the cycles, so did the representation of PBFC genes/clusters carried by this strain.

### ***In planta* experimental evolution**

To assess population dynamics of the initial SynCom inoculated on maize seedlings, samples from cycles 1 to 6, and from the last cycle (27<sup>th</sup>) were analysed for all of the plants. The obtained number of reads and *Pseudomonas* population proportions are presented in Table S9 and S10.

Overall, out of the 7 independent series of EE, two differential evolutive dynamics stand out (Figure 5) at the end of the *in planta* EE. For SynCom 1, 3 and 6, *P. fluorescens* JV391D10 was the predominant strain at the 27<sup>th</sup> cycle, whereas for SynCom 2, 4, 5 and 7 it was *P. fluorescens* JV391D17. However, despite presenting convergent outcomes, each SynCom went through different population dynamics as presented below.



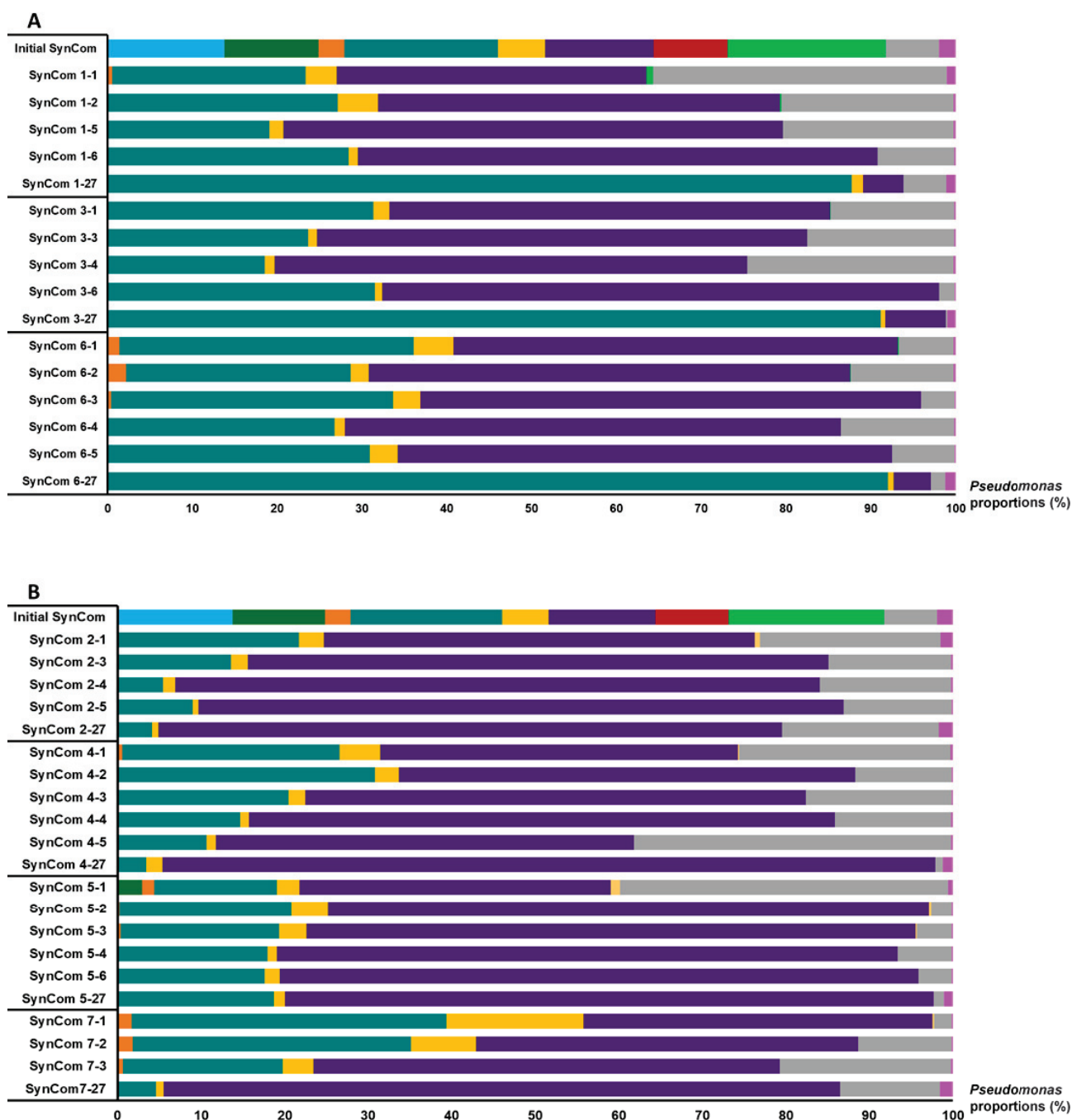
**Figure 5: Evolution of *Pseudomonas* populations during cycles of the *in planta* experimental evolution.** Each graph represents the data obtained for each of the 7 SynComs carried out independently over serial cycles. The level of each population was determined by multiplying the total number of cells recovered from roots by the percent of *rpoD* sequences assigned to one strain in each sample.

### Population dynamics for SynCom 1, 3 and 6

After only seven days, the SynCom composition highly changed (Figure 6A). Three strains became dominant, *P. fluorescens* JV391D17 (i.e. 36.53% for SynCom 1-1, 51.91% for SynCom 3-1, 52.42% for SynCom 6-1), *P. fluorescens* JV391D10 (22.83% for SynCom 1-1, 31.17% SynCom 3-1, 34.69% for SynCom 6-1), or *P. protegens* JV245A (34.63% for SynCom 1-1, 14.59% for SynCom 3-1, 6.49% for SynCom 6-1). Then, strains *P. fluorescens* JV449 (3.68% for SynCom 1-1, 1.93% for SynCom 3-1, 4.69% for SynCom 6-1), *P. protegens* JV359A1 (1.04% for SynCom 1-1, 0.15% for SynCom 3-1, 0.21% for SynCom 6-1) or *P. kribbensis* JV359A8 (0.41% for SynCom 1-1, 0.12% for SynCom 3-1, 1.37% for SynCom 6-1) were present in lower proportions. All the other species (*P. jessenii* JV251A, *P. inefficax* JV551A1, *P. wadenswilerensis* JV244P and *P. brassicacerum* JV551A7) were hardly detected after seven days (0% to 0.75%). These selection dynamics remain globally the same for cycles 2 to 6. *P. fluorescens* JV391D17 (61.31% for SynCom 1-6, 65.66% for SynCom 3-6, 58.30% for SynCom 6-6) and JV391D10 (28.40% for SynCom 1-6, 31.56% for SynCom 3-6, 30.91% for SynCom 6-6) remained the dominant strains. On the other hand, *P. protegens* JV245A proportions increased up to 24.32% and 13.38% at the end of the 4<sup>th</sup> cycle for SynCom 3-4 and SynCom 6-4, before dropping to 1.85% for SynCom 3-6, 9.05% for SynCom 1-6 and 7.37% for SynCom 6-5. *P. fluorescens* JV449 (1.10% for SynCom 1-6, 0.82% for SynCom 3-6, 3.28% for SynCom 6-6) was the fourth strain in lead whereas the *P. protegens* JV359A1 proportions decreased (between 0.1% and 0.14%) at the end of the 6<sup>th</sup> cycle. All the other strains (*P. jessenii* JV251A, *P. inefficax* JV551A1, *P. kribbensis* JV359A8, *P. brassicacerum* JV551A7, *P. wadenswilerensis* JV244B) were difficult to detect (0% to 0.13%).

At the end of the 27<sup>th</sup> cycle, a shift in population proportions was observed. *P. fluorescens* JV391D10 instead of JV391D17 became the dominant strain (87.75% for SynCom 1-27, 91.15% for SynCom 3-27, 92.05% for SynCom 6-27), followed by JV391D17 (4.79% for SynCom 1-27, 7.17% for SynCom 3-27, 4.41% for SynCom 6-27). Together these 2 strains represented between 92.54% and 98.32% of the recovered *Pseudomonas* populations. *P. protegens* JV245A population proportions decreased between 0.15% and 5.04% at the end of the 27<sup>th</sup> cycle. Proportions slightly increased for *P. protegens* JV359A1 reaching between 0.96% and 1.18% of *Pseudomonas* populations and remained stable for *P. fluorescens* JV449 (between 0.56% and 1.34%). The remaining strains (*P. jessenii* JV251A, *P. inefficax* JV551A1, *P. kribbensis* JV359A8, *P. brassicacerum* JV551A7 and *P. wadenswilerensis* JV244B) stayed mainly not detected (less than 0.01%).





**Figure 6: *Pseudomonas* population proportions obtained for the initial SynCom and for the *in planta* EE (A) for SynCom 1, 3 and 6 from cycles 1 to 27 and (B) for SynCom 2, 4, 5 and 7 from cycles 1 to 27. The proportions are expressed in percent and represent the number of *rpoD* sequences assigned to one strain divided by the total number of *rpoD* sequences in each sample.**

### Population dynamics for SynCom 2, 4, 5 and 7

Population dynamics also changed rapidly for these SynComs (Figure 6B). One strain was the dominant strain in 3 out of the 4 SynComs, *P. fluorescens* JV391D17 (with proportions reaching between 37.26% and 51.56% of SynComs) at the end of the 1<sup>st</sup> cycle. This strain was closely followed by *P. fluorescens* JV391D10 (with proportions between 14.71% and 37.72%) or *P. protegens* JV245A (with proportions between 21.59% and 39.31% except in SynCom 7-1 where



it was only representing 2.09% of this SynCom). *P. fluorescens* JV449 was the third strain in lead (with proportions varying between 2.70% and 16.35%). *P. protegens* JV359A1 and *P. kribbensis* JV359A8 were still present but in lower proportions (between 0.16 and 1.49% for JV359A1 and between 0.41% and 1.63% for JV359A8 except for SynCom 2-1 where it was at only 0.03%) and all the other strains were barely detectable (proportions lower than 0.66%). This selection pattern is also found in cycles 2 to 6.

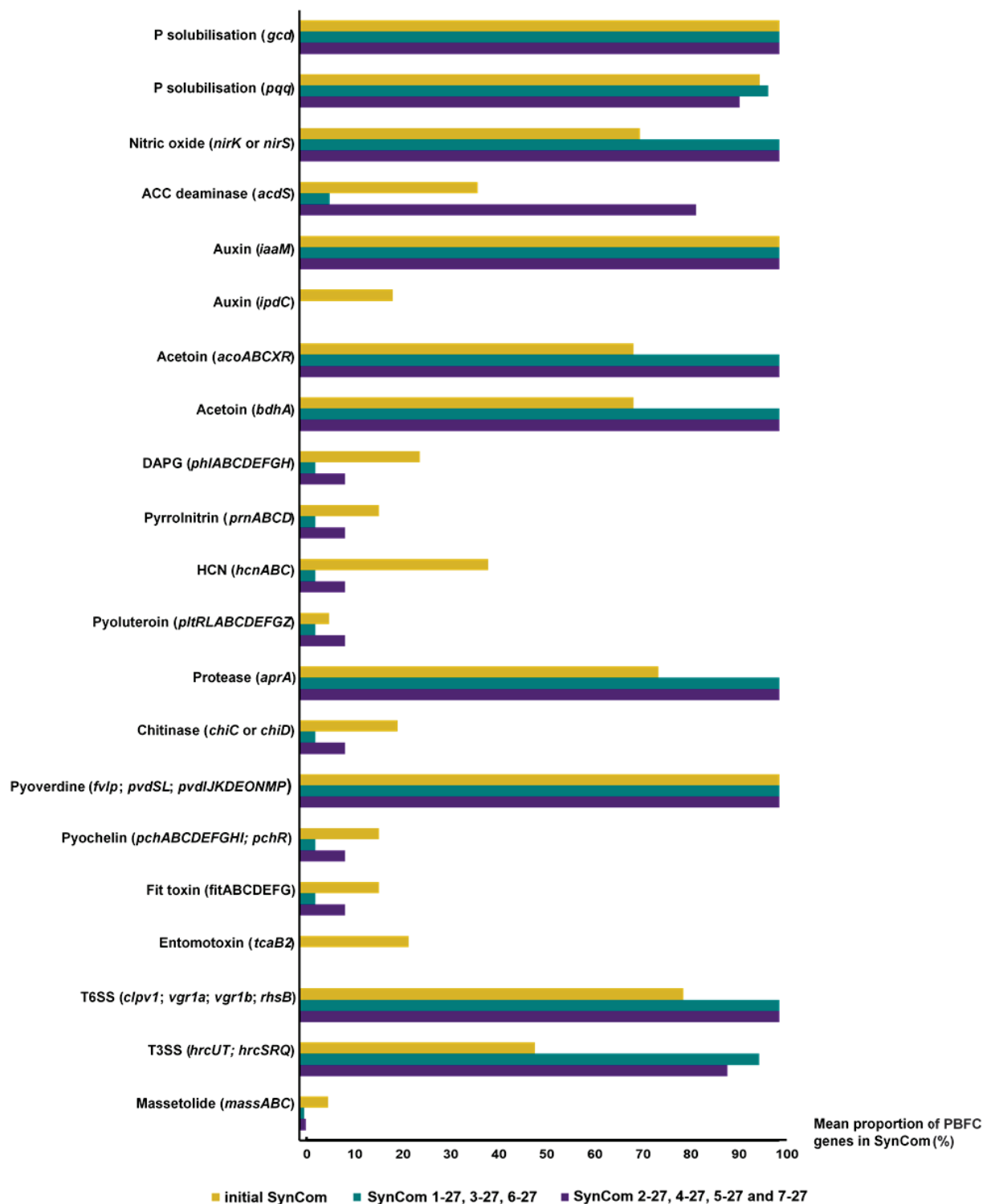
At the end of the 5<sup>th</sup> (SynComs 2-5 and 4-5), 6<sup>th</sup> (SynCom 5-6) or 3<sup>rd</sup> (SynCom 7-3) cycles, *P. fluorescens* JV391D17 became the dominant strain in all SynComs 2, 4, 5 and 7 with proportions fluctuating between 50.05% and 77.25%. Proportions decreased for *P. fluorescens* JV391D10 in 3 out of the 4 SynComs (reaching between 8.99% and 19.08% of the SynCom's populations) and for JV449 in all cases (reaching between 0.66% and 3.70%).

For *P. protegens* JV245A, two cases were observed: in one case its proportions decreased to 12.95% for SynCom 2-5 and 3.98% for SynCom 5-6 whereas it increased to 37.96% for SynCom 4-5 and 20.54% for SynCom 7-3. For JV359A1 its proportions remained low but detectable (at between 0.12% and 0.21%) while the remaining strains are mostly not detected (inferior to 0.01%).

At the end of the 27<sup>th</sup> cycle, *P. fluorescens* JV391D17 remained the dominant strains (representing between 74.64% to 92.57% of the SynComs). The second strain in lead was either *P. protegens* JV245A (18.79% for SynCom 2-27 and 11.95% for SynCom 7-27) or *P. fluorescens* JV391D10 (3.44% for SynCom 4-27 and 18.71% for SynCom 5-27). The 3 strains represented more than 96.95% of the recovered populations in all SynComs 2, 4, 5 and 7. Proportions of *P. protegens* JV359A1 and *P. fluorescens* JV449 were fluctuating between 1.07% and 1.66% and between 0.80% and 1.90%, respectively. The remaining strains (*P. jessenii* JV251A, *P. inefficax* JV551A1, *P. kribbensis* JV359A8, *P. wadenswilerensis* JV244B and *P. brassicacerum* JV551A7) were not detected (0%).

### PBCF distribution in *in planta*-evolved SynComs

At the end of the 27<sup>th</sup> cycle, two selection patterns are distinguished among the SynCom: (1) *P. fluorescens* JV391D10 is the dominant strain for *in planta* SynComs 1, 3 and 6 and (2) *P. fluorescens* JV391D17 is the dominant strain for *in planta* SynComs 2, 4, 5 and 7. However, as these 2 strains are sharing a quite similar number of PBFC genes/clusters, the SynComs were all in all harbouring a quite similar profile (Figure 7). The detail of the PBFC genes/clusters distribution for each cycle is available in Figures S2 and S3.



**Figure 7: Average PBFC genes/clusters distribution for the initial SynCom (yellow) and for the *in planta* SynCom 1, 3 and 6 and SynCom 2, 4, 5 and 7 at the end of the 27<sup>th</sup> cycle of the EE.** Results are expressed in percent and represent the average percent of all strains possessing the PBFC genes/clusters in the referred SynCom at the 27<sup>th</sup> cycle. For each PBC, the genes or gene clusters are indicated into brackets. P: Phosphate.

In SynComs 1, 3 and 6, 8 main PBFC genes/clusters were overrepresented (Figure 7, proportions superior to 90%): phosphate solubilisation (*gcd* and *pqq*), and productions of nitric oxide (*nirK* or *nirS*), auxin (*iaaM*), acetoin metabolism (*aco* and *bdhA*), production of protease (*aprA*) and pyoverdine (*pvd*). In addition, 2 secretion system types, T3SS and T6SS were also in proportions superior to 90%. In SynComs 2, 4, 5 and 7, *P. fluorescens* JV391D17 being the dominant strain, an additional PBFC gene was overrepresented: *acdS*, coding for the ACC deaminase activity (82.7% vs 6.3% in SynComs 1, 3 and 6). T3SS was reaching a proportion of 89.2% (rather than 95.8% in SynComs 1, 3 and 6). Finally, except 2 PBFC genes (*i.e.* *tcaB* and *ipdC*) that were lost, all the remaining PBFC genes/clusters were at 9.6% (vs 3.4% in SynComs 1, 3 and 6).

## DISCUSSION

Plants are associated with microbial communities, which strongly influence their development and their adaptation to the environment. Plant microbiota composition results from an active selection by the plant of microorganisms from the pool available in the soil (Lebeis *et al.*, 2015; de Souza *et al.*, 2016; Wagner *et al.*, 2016). A part of the plant microbiota is contributing to the stimulation of plant growth and protection of crops from diseases (Vacheron *et al.*, 2013; Bouffaud *et al.*, 2016, 2018; Lemanceau *et al.*, 2017). In nature, micro-organisms are in constant interaction with other organisms, so that any population tends to evolve with the other sympatric microorganisms. There is abundant evidence (Brockhurst *et al.*, 2003; Guan *et al.*, 2013) that this proximity drives fast evolution between interacting species. Commonly, most coevolution experiments are carried out with only two interacting species (Hillesland and Stahl, 2010; Blouin, 2018), whereas in nature, microbial interacting networks are far more complex. To get insight into the functioning of PGPRs-plant cooperation, an EE approach has been carried out with a synthetic community of 10 *Pseudomonas* PGPR isolated from the maize rhizosphere that we have let evolve either in liquid growth medium or into contact with a living plant, maize. This original strategy is a first step to better understand the processes involved in rhizosphere competence since, in the rhizosphere, bacterial behaviour depends on the interaction networks that PGPR will share together when colonizing plant roots (Mas *et al.*, 2016; Karimi *et al.*, 2017). In an effort to decipher this intricate interaction network, we present this work, where we provide a robust and representative study of population dynamic of a ten-strains synthetic community on the roots of sterilized maize seedlings and check whether the plant has any kind of control on these dynamics by comparing the results to the observed dynamics in absence of the plant.

Rapid but different evolutionary population dynamics were observed for each EE. The populations stabilise rapidly and even faster under *in vitro* conditions than *in planta*. For the *in vitro* EE, *P. inefficax* JV551A1 became the dominant strain from the 1st cycle (up to 81.02%) and remained in the lead till the last cycle (20<sup>th</sup>), representing up to 98.29% of the recovered populations, whereas it only represented 11.09% of the initial SynCom. In presence of the plant,

selection dynamics shift to the selection of a duo of dominant species: *P. fluorescens* and *P. protegens* whereas *P. inefficax* was absent after 27 cycles. These two species together represent on average  $98.32 \pm 0.62$  % of the recovered populations at the end of the 1<sup>st</sup> cycle and  $100.00 \pm 0.00$ % at the end of the 27<sup>th</sup> cycle, whereas they composed only  $44.71 \pm 1.74$  % of the initial SynCom. These findings are consistent with a previous study (Vacheron, 2015; Vacheron *et al.*, 2016) in which up to 60% of the retrieved populations from maize rhizosphere belonged to the FMJK cluster (composed of *P. fluorescens*, *P. mandelii*, *P. jessenii*. and *P. koreensis* species). Furthermore, *P. fluorescens* was the most represented species within this cluster.

The differential dynamic selection between both EE might be explained by differences of available catabolizable nutrient sources. However, preliminary bacterial growth tests on TSB and on a maize custom media were carried out to ensure that there are similar growth rates for the initial members of the SynCom (Figure S1). Similar growth was observed. In contrast to plant roots, both of these media were liquid, offering a homogenous distribution and accessibility of nutrients. But nutrient heterogeneity and accessibility is known to impact microbial network interactions (Kram *et al.*, 2017; Westphal *et al.*, 2018). Yet, the *in vitro* EE was carried in a liquid complex media (TSB), where all the strains have access to the nutrient in the same way. On the other hand, the SynCom from the *in planta* EE had only access to axenic maize seedlings as nutrient supply. Plants secrete a wide range of compounds that can easily sustain microbial life (Haichar *et al.*, 2014). However, secretion varies through the age of the individuals and alongside the roots, offering a heterogeneous nutrient distribution that may explain the different selection dynamics between both experiments (Jones *et al.*, 2004).

To better understand this differential selection between the two EE, we looked into the metabolism of the strains. Among the metabolic profiles of the 10 strains generated using the MicroScope platform, none major metabolic specificity was observed for *P. inefficax* JV551A1 that rapidly became dominant *in vitro* (Data not shown). The only particular point that we observed is its potential ability to biosynthesize zeaxanthin from beta-carotene (*crtz* gene), a product with antioxidant properties belonging to the family of carotenoids (Fidan and Zhan, 2019). However, the other carotenoid biosynthesis *crt* genes (Fukaya *et al.*, 2018) have not been found inside its genomes. Zeaxanthin is a valuable molecule in the food industry (Beuttler *et al.*, 2011), however to our knowledge, it does not present any antimicrobial properties nor has a role in degradation of nutrients.

Regarding a more functional point of view, there were 4 main PBFC genes/clusters (proportion higher than 90%) in the initial SynCom involved in phosphate solubilisation (*gcd* and *pqq*), production of pyoverdine (*pvd*) and auxin (*iaaM*). Phosphate solubilisation genes are common among PGPR *Pseudomonas* (Bruto *et al.*, 2014) since phosphorus is essential for plant growth and hard to reach from the soil (An and Moe, 2016). Genes encoding pyoverdine production are classified as a biocontrol PBFC cluster in this study but they can also be classified as a plant nutrition PBFC cluster as this siderophore is used for iron uptake which is essential for microbial

life development, iron availability in the rhizosphere being limited (Deori *et al.*, 2018). Lastly, auxin (*iaaM*) production is another well-spread PBF function among fluorescent *Pseudomonas* in the maize rhizosphere and a long recognized phytohormone that stimulates root development (Kepinski and Leyser, 2005; Bruto *et al.*, 2014; Vacheron *et al.*, 2016; Ortiz-Castro *et al.*, 2020). For the *in vitro* EE, PBFC genes/clusters distribution slightly changed compared to the initial SynCom, as *P. inefficax* JV55A1 is the dominant strain across all cycles. This strain harbours the same 4 afore mentioned PBFC genes/clusters plus genes involved in acetoin metabolism (*aco* and *bdhA*) and T6SS (*i.e.* a key function involved in bacterial interactions as described later). Operon *aco* is involved in the catabolism of acetoin. Acetoin is a volatile organic compound known for stimulating growth and induced systemic resistance on the plant as it can interfere with salicylic acid and ethylene pathways (Frag *et al.*, 2013), but in some bacteria can be used as a carbon source (Xiao *et al.*, 2009). In addition, *bdhA*, encoding a butanediol dehydrogenase, is implicated in the production of 2,3-butanediol by degradation of acetoin. 2,3-butanediol is a volatile compound also involved in plant growth promotion and induced resistance (Ryu *et al.*, 2003) and therefore classified as a plant hormonal balance PBF. The maintenance of the genes involved in acetoin metabolism may offer a supplementary carbon source for *P. inefficax* JV55A1.

In both the initial and evolved *in vitro* SynCom (recovered after the 20<sup>th</sup> cycle), auxin production is over-represented. However, it is only the *iaaM* gene that is overrepresented in the initial SynCom whereas both *iaaM* and *ipdC* genes are overrepresented in the evolved SynComs. Indole is a well-known communication signal in *E. coli* (Lee and Lee, 2010) and Bianco and co-workers (2006) showed that indole induces expression of genes helpful to survive under stress conditions. These findings may suggest that, as *P. inefficax* JV55A1 possess both the *iaaM* and *ipdC* genes, the production of indolic compounds could induce the expression of other genes, like type VI secretion system genes, offering a better adaptation to its environment (Spaepen and Vanderleyden, 2011; Van Puyvelde *et al.*, 2011). Furthermore, PBFC genes/cluster that are not directly related to the plant are quickly lost in the *in vitro* EE. For example, proportions for the gene *acdS* quickly drop. This gene codes for ACC deaminase, an enzyme capable of regulating the plant hormonal balance by degrading the ethylene precursor, ACC. Yet, there is no ACC present in the liquid media.

On the other hand, *in planta* EE SynComs present a more variable distribution of PBFC genes/clusters. After only one week of EE, 3 additional PBFC genes/clusters become over-represented in all of the SynComs: nitric oxide (*nirK* or *nirS*), acetoin metabolism (*aco* and *bdhA*) and protease (*aprA*) production as well as the T6SS coding genes. At the end of the *in planta* EE, T3SS coding genes also becomes over-represented. Nitric oxide is known to play an important role in plant growth development, signal transduction, adaptive response to biotic and abiotic stresses (Besson-Bard *et al.*, 2008; Baudouin and Hancock, 2014). Nitric oxide is also believed to serve as a long-distance signal for the activation of systemic acquired resistance,

making its producers, more effective biocontrol agents. Protease (*aprA*) is another PBF classified as a biocontrol type function in this study as it has an antimicrobial effect but it can also catalyse reactions necessary for microorganisms life and growth by degrading root exudates, a source of amino acid and small peptides (Lugtenberg and Kamilova, 2009; Egamberdieva *et al.*, 2010). Lastly, secretion systems also become over-represented through both EE: T6SS becomes predominant in both EE and type 3 secretion system (T3SS) prevails only in the *in planta* EE. T3SS is famous for allowing induction of disease for plant pathogens (Marchi *et al.*, 2013; Almario *et al.*, 2014) but this molecular syringe can also be used in symbiotic interactions with plants (Rezzonico *et al.*, 2005) as the introduced factors can hijack the host cell in a way that is beneficial to the invading bacteria. T6SS is also an injectosome-like secretion system, that has previously been described in some PGPR (Boyer *et al.*, 2009; Barret, 2011). Its main function is promoting interbacterial antagonisms (Vacheron *et al.*, 2019) by injecting toxic effectors into nearby cells, a T6SS active organism prevents its self-intoxication by producing immunity proteins. T6SS can also be involved in host manipulation, in order to fight the plant endophytic bacteria, making it therefore important for root colonisation (Russell *et al.*, 2013; Bernal *et al.*, 2018).

Finally, T6SS is common among our strains (8 out of 10). This is consistent with the phylogenetic analyses carried by Bernal *et al.*, (2018) which points out that T6SS is wide-spread among *Pseudomonas*, underlying that this system probably plays an important role in *Pseudomonas* physiology and fitness. Taking this into account T6SS can give a significant fitness advantage in a bacterial SynCom, whether in terms of root colonisation or interbacterial interactions in both EE. According to the literature, effectors vary among species, which could explain the overall and rapid selection of *P. inefficax* JV551A1 in the *in vitro* EE. The distinctive selection between two *P. fluorescens* strains JV391D10 and JV391D17 leads to believe that effectors and immunity proteins may differ between strains from a same species.

Furthermore, for the *in planta* EE, *P. fluorescens* strains seem to have an impact on *P. protegens*. The proportion of this species decreased when there was an overall selection of the *P. fluorescens* JV391D10 strain and increased when *P. fluorescens* JV391D17 was mainly selected. However, this is only true for SynComs 2-27 and 7-27 as SynCom 4-27 and 5-27 also present a majority of *P. fluorescens* JV391D17, but *P. protegens* JV245A proportions are as low as the proportions found in plants 1-27, 3-27 and 6-27, where *P. fluorescens* JV391D10 is the dominant strain. In any case, *P. protegens* JV245A plays an important role in the functional redundancy of the evolved SynComs as it possesses one of the highest number of PBFC genes/clusters (15). Therefore, SynComs 2-27 and 7-27 present the most functionally redundant populations as *P. fluorescens* JV391D17 and *P. protegens* JV245A have complementary PBFC genes/ clusters (*i.e.* JV391D17 has ACC deaminase, lacking in *P. protegens*). Together these strains harbour 17 out of 20 PBFC genes/clusters studied in this work. The only lacking PBFC genes are the NRPS massetolide (a cyclic lipopeptide surfactant) genes (*mass*), the entomotoxin (*tcaB2*) production genes (Rangel *et al.*, 2016). In the tested conditions neither of these PBFC genes seem to be



advantageous for the SynComs to keep. Finally, auxin production coded by *ipdC* is also missing in the *in planta* EE (Spaepen *et al.*, 2007; McClerklin *et al.*, 2018).

Regardless of a different species selection between both EE (*P. inefficax* for the *in vitro* EE and *P. fluorescens* et *P. protegens* for the *in planta* EE), all the PBFC genes/clusters over represented *in vitro* at the 20<sup>th</sup> cycle are also over represented *in planta* at the 27<sup>th</sup> cycle, except for the indole-3-pyruvate (*ipdC*) pathway for auxin production, which is only over-represented in the *in vitro* EE. All these results show that, despite the fact that SynCom evolutions in the presence of the plant can select different sets of species of *Pseudomonas*, the PBFC genes/clusters distribution remains globally the same. These findings suggest that strain selection by plant is driven in order to maintain the main types of PBFC genes/clusters within the plant-associated community. The next step would be to analyse the difference of PBFC genes/clusters selection between both EE, especially in terms of gene expression in order to see which PBFC gene/cluster is really expressed and useful in presence of the plant. Moreover, T6SS is present in the dominant strains of both EE and should be studied further, especially in confrontations between each strain to see if the missing strains are eliminated by the effectors of the dominant strains.

Even though these *Pseudomonas* bacteria possess a great diversity of plant beneficial functions, their inoculation does not always present significant beneficial effects on their hosts (Prigent-Combaret and Muller, unpublished results). However, mono-inoculations have long been proven to be effective on maize growth stimulation (Gholami *et al.*, 2009), yet microorganisms live in communities and their potential beneficial effects are the result of joint efforts. Consortium bacteria inoculation has risen up in the last years and recent studies have formulated consortiums of PGPR that successfully stimulate growth of maize and other plants (Chauhan and Bagyaraj, 2015; Saia *et al.*, 2015; Molina-Romero *et al.*, 2017; Amogou *et al.*, 2019). But studies using synthetic communities of PGPR are rather new and need to be strengthened. Contrary to what was expected, plant selection leans toward PGPR with a high number of PBFC genes/clusters as they outcompeted the other *Pseudomonas* of this study. Our work aims to advance the understanding of rhizosphere colonisation by the soil microbial pool and unravel the interaction network between the plant and its microbiota, and within the microbiota itself. Deciphering the principles that underpin ecological and evolutionary properties of microbial communities can allow us to build predictive models of ecological dynamics of microbial communities.

## ACKNOWLEDGEMENTS

Beatriz Manriquez was supported by a Ph.D. fellowship from the French Ministère de l'Education Nationale, de l'Enseignement Supérieur et de la Recherche. The authors acknowledge funding from the French National Programme EC2CO 'Structuring initiative Continental and Coastal Ecosphere' - Microbien 2018-2019. We are grateful to Jeanne Doré for technical help, as well as Denis Faure (I2BC, France) for helpful discussion. We also acknowledge the LABGeM and the



National Infrastructure “France Genomique” for support within the MicroScope annotation platform. The authors declare no conflict of interest.

## DATA ACCESSIBILITY

Illumina MiSeq paired-end reads have been deposited in the European Bioinformatics Institute (EBI) database and waiting for accession numbers for *rpoD*. Genome accessibility: PRJNA572594; PRJNA572595; PRJNA566351; PRJNA566305; PRJNA572600; PRJNA566269; PRJNA572570; PRJNA566300 and PRJNA572600.

## AUTHOR CONTRIBUTIONS

BM, DM and CPC designed the project, BM and FG carried out field EE set-up, DA and AD implemented bioinformatic analyses, BM, DM, DA and CPC analyzed data, BM, DM and CPC prepared the first draft of the manuscript, which was finalized by all authors.

## REFERENCES

- Adams, J., and Rosenzweig, F. (2014). Experimental microbial evolution: history and conceptual underpinnings. *Genomics* 104, 393–398. doi:10.1016/j.ygeno.2014.10.004.
- Almario, J., Gobbin, D., Défago, G., Moënnelocoz, Y., and Rezzonico, F. (2014). Prevalence of type III secretion system in effective biocontrol *Pseudomonads*. *Res. Microbiol.* 165, 300–304. doi:10.1016/j.resmic.2014.03.008.
- Amogou, O., Agbodjato, N. A., Dagbénonbakin, G., Noumavo, P. A., Sina, H., Sylvestre, A. A., *et al.*, (2019). Improved maize growth in condition controlled by PGPR inoculation on ferruginous soil in central Benin. *Food Nutr. Sci.* 10, 1433–1451. doi:10.4236/fns.2019.1012102.
- An, R., and Moe, L. A. (2016). Regulation of pyrroloquinoline quinone-dependent glucose dehydrogenase activity in the model rhizosphere-dwelling bacterium *Pseudomonas putida* KT2440. *Appl. Environ. Microbiol.* 82, 4955–4964. doi:10.1128/AEM.00813-16.
- Barret (2011). Genomic analysis of the type VI secretion systems in *Pseudomonas spp.*: novel clusters and putative effectors uncovered. *Microbiology Society*.
- Bashan, Y., and de-Bashan, L. E. (2010). How the plant growth-promoting bacterium *Azospirillum* promotes plant growth—a critical assessment. *Advances in Agronomy* (Elsevier), 77–136. doi:10.1016/S0065-2113(10)08002-8.
- Baudouin, E., and Hancock, J. (2014). Nitric oxide signaling in plants. *Front. Plant Sci.* 4. doi:10.3389/fpls.2013.00553.
- Bernal, P., Llamas, M. A., and Filloux, A. (2018). Type VI secretion systems in plant-associated bacteria. *Environ. Microbiol.* 20, 1–15. doi:10.1111/1462-2920.13956.
- Besson-Bard, A., Pugin, A., and Wendehenne, D. (2008). New insights into nitric oxide signaling in plants. *Annu. Rev. Plant Biol.* 59, 21–39. doi:10.1146/annurev.arplant.59.032607.092830.
- Beuttler, H., Hoffmann, J., Jeske, M., Hauer, B., Schmid, R. D., Altenbuchner, J., *et al.*, (2011). Biosynthesis of zeaxanthin in recombinant *Pseudomonas putida*. *Appl. Microbiol. Biotechnol.* 89, 1137–1147. doi:10.1007/s00253-010-2961-0.
- Bianco, C., Imperlini, E., Calogero, R., Senatore, B., Amoresano, A., Carpentieri, A., *et al.*, (2006). Indole-3-acetic acid improves *Escherichia coli*'s defences to stress. *Arch. Microbiol.* 185, 373–382. doi:10.1007/s00203-006-0103-y.
- Blin, K., Shaw, S., Steinke, K., Villebro, R., Ziemert, N., Lee, S. Y., *et al.*, (2019). AntiSMASH 5.0: Updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res.* 47, W81–W87. doi:10.1093/nar/gkz310.
- Blouin, M. (2018). Chemical communication: An evidence for co-evolution between plants and soil organisms. *Appl. Soil Ecol.* 123, 409–415. doi:10.1016/j.apsoil.2017.10.028.
- Bouffaud, M.-L., Renoud, S., Dubost, A., Moënnelocoz, Y., and Muller, D. (2018). 1-

aminocyclopropane-1-carboxylate deaminase producers associated to maize and other *Poaceae* species. *Microbiome* 6, 114. doi:10.1186/s40168-018-0503-7.

Bouffaud, M.-L., Renoud, S., Moënné-Loccoz, Y., and Muller, D. (2016). Is plant evolutionary history impacting recruitment of diazotrophs and *nifH* expression in the rhizosphere? *Sci. Rep.* 6, 21690. doi:10.1038/srep21690.

Boyer, F., Fichant, G., Berthod, J., Vandenbrouck, Y., and Attree, I. (2009). Dissecting the bacterial type VI secretion system by a genome wide *in silico* analysis: what can be learned from available microbial genomic resources? *BMC Genomics* 10, 104. doi:10.1186/1471-2164-10-104.

Brockhurst, M. A., Morgan, A. D., Rainey, P. B., and Buckling, A. (2003). Population mixing accelerates coevolution: population mixing accelerates coevolution. *Ecol. Lett.* 6, 975–979. doi:10.1046/j.1461-0248.2003.00531.x.

Bruto, M., Prigent-Combaret, C., Muller, D., and Moënné-Loccoz, Y. (2014). Analysis of genes contributing to plant beneficial functions in plant growth-promoting rhizobacteria and related Proteobacteria. *Sci. Rep.* 4, 6261. doi:10.1038/srep06261.

Buckling, A., Craig Maclean, R., Brockhurst, M. A., and Colegrave, N. (2009). The *Beagle* in a bottle. *Nature* 457, 824–829. doi:10.1038/nature07892.

Cairns, J., Ruokolainen, L., Hultman, J., Tamminen, M., Virta, M., and Hiltunen, T. (2018). Ecology determines how low antibiotic concentration impacts community composition and horizontal transfer of resistance genes. *Commun. Biol.* 1. doi:10.1038/s42003-018-0041-7.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., *et al.*, (2009). BLAST+: architecture and applications. *BMC Bioinformatics* 10, 421. doi:10.1186/1471-2105-10-421.

Chauhan, H., and Bagyaraj, D. J. (2015). Inoculation with selected microbial consortia not only enhances growth and yield of french bean but also reduces fertilizer application under field condition. *Sci. Hortic.* 197, 441–446. doi:10.1016/j.scienta.2015.10.001.

De Roy, K., Marzorati, M., Van den Abbeele, P., Van de Wiele, T., and Boon, N. (2014). Synthetic microbial ecosystems: an exciting tool to understand and apply microbial communities. *Environ. Microbiol.* 16, 1472–1481. doi:10.1111/1462-2920.12343.

de Souza, R. S. C., Okura, V. K., Armanhi, J. S. L., Jorrín, B., Lozano, N., da Silva, M. J., *et al.*, (2016). Unlocking the bacterial and fungal communities assemblages of sugarcane microbiome. *Sci. Rep.* 6, 28774. doi:10.1038/srep28774.

Deori, M., Jayamohan, N. S., and Kumudini, B. S. (2018). Pyoverdine production, characterization and iron binding affinity of hydroxamate siderophores from rhizosphere associated fluorescent *Pseudomonas*. *J. Plant Prot. Res.* 58. doi:10.24425/119116.

D'Souza, G., and Kost, C. (2016). Experimental evolution of metabolic dependency in bacteria. *PLoS Genet.* 12, e1006364. doi:10.1371/journal.pgen.1006364.

Egamberdieva, D., Renella, G., Wirth, S., and Islam, R. (2010). Protease enzyme activities in the rhizosphere of plants. *Soil Enzymology* Soil Biology., eds. G. Shukla and A. Varma (Berlin, Heidelberg: Springer Berlin Heidelberg), 149–166. doi:10.1007/978-3-642-14225-3\_8.

Elena, S. F., and Lenski, R. E. (2003). Microbial genetics: Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. *Nat. Rev. Genet.* 4, 457–469. doi:10.1038/nrg1088.

Farag, M. A., Zhang, H., and Ryu, C.-M. (2013). Dynamic chemical communication between plants and bacteria through airborne signals: Induced resistance by bacterial volatiles. *J. Chem. Ecol.* 39, 1007–1018. doi:10.1007/s10886-013-0317-9.

Farris, J. S. (1972). Estimating phylogenetic trees from distance matrices. *Am. Nat.* 106, 645–668. doi:10.1086/282802.

Fidan, O., and Zhan, J. (2019). Discovery and engineering of an endophytic *Pseudomonas* strain from *Taxus chinensis* for efficient production of zeaxanthin diglucoside. *J. Biol. Eng.* 13, 66. doi:10.1186/s13036-019-0196-x.

Fukaya, Y., Takemura, M., Koyanagi, T., Maoka, T., Shindo, K., and Misawa, N. (2018). Structural and functional analysis of the carotenoid biosynthesis genes of a *Pseudomonas* strain isolated from the excrement of *Autumn darter*. *Biosci. Biotechnol. Biochem.* 82, 1043–1052. doi:10.1080/09168451.2017.1398069.

Gholami, A., Shahsavani, S., and Nezarat, S. (2009). The effect of plant growth promoting rhizobacteria (PGPR) on germination, seedling growth and yield of Maize. 3, 6. *International Journal of Agricultural and Biosystems Engineering*, Vol.3, No.1, 2009

- Guan, S. H., Gris, C., Cruveiller, S., Pouzet, C., Tasse, L., Leru, A., *et al.*, (2013). Experimental evolution of nodule intracellular infection in legume symbionts. *ISME J.* 7, 1367–1377. doi:10.1038/ismej.2013.24.
- Haichar, F. el Z., Santaella, C., Heulin, T., and Achouak, W. (2014). Root exudates mediated interactions belowground. *Soil Biol. Biochem.* 77, 69–80. doi:10.1016/j.soilbio.2014.06.017.
- Hillesland, K. L., and Stahl, D. A. (2010). Rapid evolution of stability and productivity at the origin of a microbial mutualism. *Proc. Natl. Acad. Sci.* 107, 2124–2129. doi:10.1073/pnas.0908456107.
- Hynes, M. F., Simon, R., and Pühler, A. (1985). The development of plasmid-free strains of *Agrobacterium tumefaciens* by using incompatibility with a *Rhizobium meliloti* plasmid to eliminate *pAtc58*. *Plasmid* 13, 99–105. doi:10.1016/0147-619X(85)90062-9.
- Jones, D. L., Hodge, A., and Kuzyakov, Y. (2004). Plant and mycorrhizal regulation of rhizodeposition. *New Phytol.* 163, 459–480. doi:10.1111/j.1469-8137.2004.01130.x.
- Karimi, B., Maron, P. A., Chemidlin-Prevost Boure, N., Bernard, N., Gilbert, D., *et al.*, (2017). Microbial diversity and ecological networks as indicators of environmental quality. *Environ. Chem. Lett.* 15, 265–281. doi:10.1007/s10311-017-0614-6.
- Kautsar, S. A., Blin, K., Shaw, S., Navarro-Muñoz, J. C., Terlouw, B. R., van der Hooft, J. J. J., *et al.*, (2020). MIBiG 2.0: a repository for biosynthetic gene clusters of known function. *Nucleic Acids Res.* 48, D454–D458. doi:10.1093/nar/gkz882.
- Kawecki, T. J., Lenski, R. E., Ebert, D., Hollis, B., Olivieri, I., and Whitlock, M. C. (2012). Experimental evolution. *Trends Ecol. Evol.* 27, 547–560. doi:10.1016/j.tree.2012.06.001.
- Kepinski, S., and Leyser, O. (2005). The *Arabidopsis* F-box protein TIR1 is an auxin receptor. *Nature* 435, 446–451. doi:10.1038/nature03542.
- Kloepper, J. W., Lifshitz, R., and Zablotowicz, R. M. (1989). Free-living bacterial inocula for enhancing crop productivity. *Trends Biotechnol.* 7, 39–44. doi:10.1016/0167-7799(89)90057-7.
- Kram, K. E., Geiger, C., Ismail, W. M., Lee, H., Tang, H., Foster, P. L., *et al.*, (2017). Adaptation of *Escherichia coli* to long-term serial passage in complex medium: evidence of parallel evolution. *mSystems* 2. doi:10.1128/mSystems.00192-16.
- Kreft, Ł., Botzki, A., Coppens, F., Vandepoele, K., and Van Bel, M. (2017). PhyD3: a phylogenetic tree viewer with extended phyloXML support for functional genomics data visualization. *Bioinformatics* 33, 2946–2947. doi:10.1093/bioinformatics/btx324.
- Lagesen, K., Hallin, P., Rødland, E. A., Stærfeldt, H.-H., Rognes, T., and Ussery, D. W. (2007). RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.* 35, 3100–3108. doi:10.1093/nar/gkm160.
- Lebeis, S. L., Paredes, S. H., Lundberg, D. S., Breakfield, N., Gehring, J., McDonald, M., *et al.*, (2015). Salicylic acid modulates colonization of the root microbiome by specific bacterial taxa. *Science* 349, 860–864. doi:10.1126/science.aaa8764.
- Lee, J.-H., and Lee, J. (2010). Indole as an intercellular signal in microbial communities. *FEMS Microbiol. Rev.* 34, 426–444. doi:10.1111/j.1574-6976.2009.00204.x.
- Lefort, V., Desper, R., and Gascuel, O. (2015). FastME 2.0: a comprehensive, accurate, and fast distance-based phylogeny inference program. *Mol. Biol. Evol.* 32, 2798–2800. doi:10.1093/molbev/msv150.
- Lemanceau, P., Blouin, M., Muller, D., and Moëgne-Loccoz, Y. (2017). Let the core microbiota be functional. *Trends Plant Sci.* 22, 583–595. doi:10.1016/j.tplants.2017.04.008.
- Lenski, R. E. (2017). Experimental evolution and the dynamics of adaptation and genome evolution in microbial populations. *ISME J.* 11, 2181–2194. doi:10.1038/ismej.2017.69.
- Lenski, R. E., Rose, M. R., Simpson, S. C., and Tadler, S. C. (1991). Long-term experimental evolution in *Escherichia coli*. I. adaptation and divergence during 2,000 generations. *Am. Nat.* 138, 1315–1341. doi:10.1086/285289.
- Lugtenberg, B., and Kamilova, F. (2009). Plant-growth promoting Rhizobacteria. *Annu. Rev. Microbiol.* 63, 541–556. doi:10.1146/annurev.micro.62.081307.162918.
- Marchi, M., Boutin, M., Gazengel, K., Rispe, C., Gauthier, J.-P., Guillermin-Erckelboudt, A.-Y., *et al.*, (2013). Genomic analysis of the biocontrol strain *Pseudomonas fluorescens* Pf29Arp with evidence of T3SS and T6SS gene expression on plant roots. *Environ. Microbiol. Rep.* 5, 393–403. doi:10.1111/1758-2229.12048.
- Mas, A., Jamshidi, S., Lagadeuc, Y., Eveillard, D., and Vandenkoornhuyse, P. (2016). Beyond the Black Queen Hypothesis. *ISME J.* 10, 2085–2091. doi:10.1038/ismej.2016.22.
- McClerklin, S. A., Lee, S. G., Harper, C. P., Nwumeh, R., Jez, J. M., and Kunkel, B. N.

(2018). Indole-3-acetaldehyde dehydrogenase-dependent auxin synthesis contributes to virulence of *Pseudomonas syringae* strain DC3000. *PLoS Pathog.* 14, e1006811. doi:10.1371/journal.ppat.1006811.

Medema, M. H., Blin, K., Cimermancic, P., de Jager, V., Zakrzewski, P., Fischbach, M. A., *et al.*, (2011). AntiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Res.* 39, W339–W346. doi:10.1093/nar/gkr466.

Meier-Kolthoff, J. P., Auch, A. F., Klenk, H.-P., and Göker, M. (2013). Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics* 14, 60. doi:10.1186/1471-2105-14-60.

Meier-Kolthoff, J. P., and Göker, M. (2019). TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nat. Commun.* 10, 2182. doi:10.1038/s41467-019-10210-3.

Meier-Kolthoff, J. P., Hahnke, R. L., Petersen, J., Scheuner, C., Michael, V., Fiebig, A., *et al.*, (2014). Complete genome sequence of DSM 30083T, the type strain (U5/41T) of *Escherichia coli*, and a proposal for delineating subspecies in microbial taxonomy. *Stand. Genomic Sci.* 9, 2. doi:10.1186/1944-3277-9-2.

Molina-Romero, D., Baez, A., Quintero-Hernández, V., Castañeda-Lucio, M., Fuentes-Ramírez, L. E., Bustillos-Cristales, M. del R., *et al.*, (2017). Compatible bacterial mixture, tolerant to desiccation, improves maize plant growth. *PLoS ONE* 12, e0187913. doi:10.1371/journal.pone.0187913.

Ondov, B. D., Treangen, T. J., Melsted, P., Mallonee, A. B., Bergman, N. H., Koren, S., *et al.*, (2016). Mash: fast genome and metagenome distance estimation using MinHash. *Genome Biol.* 17, 132. doi:10.1186/s13059-016-0997-x.

Ortiz-Castro, R., Campos-García, J., and López-Bucio, J. (2020). *Pseudomonas putida* and *Pseudomonas fluorescens* influence *Arabidopsis* root system architecture through an auxin response mediated by bioactive cyclodipeptides. *J. Plant Growth Regul.* 39, 254–265. doi:10.1007/s00344-019-09979-w.

Rangel, L. I., Henkels, M. D., Shaffer, B. T., Walker, F. L., Davis, E. W., Stockwell, V. O., *et al.* (2016). Characterization of toxin complex gene clusters and insect toxicity of bacteria representing four subgroups of *Pseudomonas fluorescens*. *PLoS ONE* 11. doi:10.1371/journal.pone.0161120.

Rezzonico, F., Binder, C., Défago, G., and Moëgne-Loccoz, Y. (2005). The Type III secretion system of biocontrol *Pseudomonas fluorescens* KD targets the phytopathogenic chromista *Pythium ultimum* and promotes cucumber protection. *Mol. Plant. Microbe Interact.* 18, 991–1001. doi:10.1094/MPMI-18-0991.

Russell, A. B., LeRoux, M., Hathazi, K., Agnello, D. M., Ishikawa, T., Wiggins, P. A., *et al.*, (2013). Diverse type VI secretion phospholipases are functionally plastic antibacterial effectors. *Nature* 496, 508–512. doi:10.1038/nature12074.

Ryu, C. M., Farag, M. A., Hu, C. H., Reddy, M. S., Paré, P. W., and Kloepper, J. W. (2003). Volatiles produced by PGPR elicit plant growth promotion and induced resistance in *Arabidopsis*. 6<sup>th</sup> International PGPR Workshop.8.

Saia, S., Rappa, V., Ruisi, P., Abenavoli, M. R., Sunseri, F., Giambalvo, D., *et al.*, (2015). Soil inoculation with symbiotic microorganisms promotes plant growth and nutrient transporter genes expression in durum wheat. *Front. Plant Sci.* 6. doi:10.3389/fpls.2015.00815.

Smith, P., and Schuster, M. (2019). Public goods and cheating in microbes. *Curr. Biol.* 29, R442–R447. doi:10.1016/j.cub.2019.03.001.

Spaepen, S., and Vanderleyden, J. (2011). Auxin and plant-microbe interactions. *Cold Spring Harb. Perspect. Biol.* 3, a001438–a001438. doi:10.1101/cshperspect.a001438.

Spaepen, S., Vanderleyden, J., and Remans, R. (2007). IPDC Indole-3-acetic acid in microbial and microorganism-plant signaling. *FEMS Microbiol. Rev.* 31, 425–448. doi:10.1111/j.1574-6976.2007.00072.x.

Travisano, M., Mongold, J., Bennett, A., and Lenski, R. (1995). Experimental tests of the roles of adaptation, chance, and history in evolution. *Science* 267, 87–90. doi:10.1126/science.7809610.

Vacheron, J. (2015). Sélection des rhizobactéries phytostimulatrices par la plante : impact sur la distribution des propriétés phytobénéfiques chez les *Pseudomonas* fluorescents. *phD Thesis. Université Claude Bernard - Lyon 1*.

Vacheron, J., Desbrosses, G., Bouffaud, M.-L., Touraine, B., Moëgne-Loccoz, Y., Muller, D., *et al.*, (2013). Plant growth-promoting rhizobacteria and root system functioning. *Front. Plant Sci.* 4. doi:10.3389/fpls.2013.00356.

Vacheron, J., Moëgne-Loccoz, Y., Dubost, A., Gonçalves-Martins, M., Muller, D., and Prigent-Combaret, C. (2016). Fluorescent *Pseudomonas*



strains with only few plant-beneficial properties are favored in the maize rhizosphere. *Front. Plant Sci.* 7. doi:10.3389/fpls.2016.01212.

Vacheron, J., Péchy-Tarr, M., Brochet, S., Heiman, C. M., Stojiljkovic, M., Maurhofer, M., *et al.*, (2019). T6SS contributes to gut microbiome invasion and killing of an herbivorous pest insect by plant-beneficial *Pseudomonas protegens*. *ISME J.* 13, 1318–1329. doi:10.1038/s41396-019-0353-8.

Vallenet, D., Calteau, A., Dubois, M., Amours, P., Bazin, A., Beuvin, M., *et al.*, (2020). MicroScope: an integrated platform for the annotation and exploration of microbial gene functions through genomic, pangenomic and metabolic comparative analysis. *Nucleic Acids Res.* 48, D579–D589. doi:10.1093/nar/gkz926.

Van Puyvelde, S., Cloots, L., Engelen, K., Das, F., Marchal, K., Vanderleyden, J., *et al.*, (2011). Transcriptome analysis of the rhizosphere bacterium *Azospirillum brasilense* reveals an

extensive auxin response. *Microb. Ecol.* 61, 723–728. doi:10.1007/s00248-011-9819-6.

Wagner, M. R., Lundberg, D. S., del Rio, T. G., Tringe, S. G., Dangl, J. L., and Mitchell-Olds, T. (2016). Host genotype and age shape the leaf and root microbiomes of a wild perennial plant. *Nat. Commun.* 7, 12151. doi:10.1038/ncomms12151.

Westphal, L. L., Lau, J., Negro, Z., Moreno, I. J., Ismail Mohammed, W., Lee, H., *et al.*; (2018). Adaptation of *Escherichia coli* to long-term batch culture in various rich media. *Res. Microbiol.* 169, 145–156. doi:10.1016/j.resmic.2018.01.003.

Xiao, Z., Ma, C., Xu, P., and Lu, J. R. (2009). Acetoin catabolism and acetylbutanediol formation by *Bacillus pumilus* in a chemically defined medium. *PLoS ONE* 4, e5627. doi:10.1371/journal.pone.0005627.

**Table S1: Secondary metabolites and corresponding gene clusters of *Pseudomonas* strains.**

Secondary metabolite	Genes involved in biosynthesis	Reference Stain †	JV391D10	JV391D17	JV449	JV359A8	JV551A7	JV245A	JV359A1	JV251A	JV551A1	JV244B
Phosphate solubilisation	<i>gcd</i>	<i>P. kilonensis</i> F113 PSF113_4702	JV391D10_v1_50044	JV391D17_v1_140048	JV449_v1_50249	JV359A8_v1_210038	JV551A7_v1_170009					
		<i>P. protegens</i> CHA0 PFLCHA0_c48970						JV245A_v1_170010	JV359A1_v1_20719	JV251A_v1_100062		
		<i>P. putida</i> KT2440 PP_1444							JV551A1_v1_470112			JV244B_v1_110123
		<i>P. kilonensis</i> F113 PSF113_1629	JV391D10_v1_10919	JV391D17_v1_10736	JV449_v1_230050	JV359A8_v1_50318	JV551A7_v1_100108					
		<i>P. protegens</i> CHA0 PFLCHA0_c4393_0							JV359A1_v1_20202	JV251A_v1_10642		
Nitric Oxide Production	<i>nirK</i> or <i>nirS</i>	<i>P. putida</i> KT2440 PP_1762								JV244B_v1_120062	JV551A1_v1_190072	JV244B_v1_150136
		<i>P. kilonensis</i> F113 PSF113_3754	JV391D10_v1_40462	JV391D17_v1_100026	JV449_v1_190023		JV551A7_v1_170009					
		<i>P. protegens</i> CHA0 PFLCHA0_c54550						JV245A_v1_540286	JV359A1_v1_60086			
1-aminocyclo propane-1-carboxylate deaminase (ACC)	<i>acdS</i>	<i>P. kilonensis</i> F113 PSF113_3500					JV551A7_v1_130269					
		<i>P. kilonensis</i> F113 PSF113_5381	JV391D10_v1_20058	JV391D17_v1_120057	JV449_v1_340057	JV359A8_v1_100115	JV551A7_v1_120061					
		<i>P. protegens</i> CHA0 PFLCHA0_c56220						JV245A_v1_180081	JV359A1_v1_30068			
Auxin Production	<i>iaaM</i>	<i>P. putida</i> KT2440 PP_0383									JV551A1_v1_40028	JV244B_v1_140107
		<i>P. putida</i> KT2440 PP_2552							JV551A1_v1_230065			JV244B_v1_260079

Secondary metabolite	Genes involved in biosynthesis	Reference Stain <sup>†</sup>	JV391D10	JV391D17	JV449	JV359A8	JV551A7	JV245A	JV359A1	JV251A	JV551A1	JV244B
2,4-diacetylphloroglucinol (DAPG)	<i>phl</i> <i>ABCDEFGH</i>	<i>P. kilonensis</i> F113					JV551A7_v1_140107-140114					
		PSF113 2458 – 2468										
		<i>P. protegens</i> CHA0						JV245A_v1_110027-110036	JV359A1_v1			
		PFLCHA0 c59050 – c59120							30369-30376			
Pyrrolnitrin	<i>prnABCD</i>	<i>P. protegens</i> CHA0							JV359A1_v1			JV244B_v1_540011-40014
		PFLCHA0 c36450 - c36480						JV245A_v1_830017-30021	13013-13016			
Acetoin	<i>acoABCXR</i>	<i>P. kilonensis</i> F113	JV391D10_v1	JV391D17_v1_110162-10166	JV449_v1_130015-130018		JV551A7_v1_170051-70054					
		PSF113 3712 - 3716	30418-30420									
		<i>P. protegens</i> CHA0						JV245A_v1_210111-10116	JV359A1_v1			
		PFLCHA0 c22140– c22190							11509 –1513			
		<i>P. putida</i> KT2440									JV551A1_v1_50057-0059	
	<i>bdhA</i>	<i>P. kilonensis</i> F113	JV391D10_v1_11073	JV391D17_v1_170017	JV449_v1_30026		JV551A7_v1_170050					
		PSF113 3717										
		<i>P. protegens</i> CHA0						JV245A_v1_210110	JV359A1_v1			
		PFLCHA0 c22200– c33950							11514-12747			
		<i>P. putida</i> KT2440									JV551A1_v1_50054	
Hydrogen Cyanide Production (HCN)	<i>hcnABC</i>	<i>P. kilonensis</i> F113				JV359A8_v1	JV551A7_v1_60072-60074					JV244B_v1_20134-20136
		PSF113 2367 - 2369				20729-0731						
		<i>P. protegens</i> CHA0						JV245A_v1_260089-60093	JV359A1_v1			
		PFLCHA0 c26420- c26440							11959-1961			



Secondary metabolite	Genes involved in biosynthesis	Reference Stain <sup>†</sup>	JV391D10	JV391D17	JV449	JV359A8	JV551A7	JV245A	JV359A1	JV251A	JV551A1	JV244B
Pyoluteroin	<i>pttRLABCDE FGZ</i>	<i>P. protegens</i> CHA0 PFLCHA0_c28500						JV245A_v1_310242-10252	JV359A1_v1_12176-12185			JV244B_v1_440005 JV244B_v1_420038 JV244B_v1_120142
Protease	<i>aprA</i>	<i>P. kilonensis</i> F113 PSF113_2949	JV391D10_v1_10510	JV391D17_v1_20902	JV449_v1_170162	JV359A8_v1_40499	JV551A7_v1_150068					
Protease	<i>aprA</i>	<i>P. protegens</i> CHA0 PFLCHA0_c3240_0						JV245A_v1_530061 JV245A_v1_260197	JV359A1_v1_11859			
Chitinase	<i>chiC</i> or <i>chiD</i>	<i>P. fluorescens</i> Pf-5 PFL2091				JV359A8_v1_20780						
		<i>P. protegens</i> CHA0 PFLCHA0_c21380						JV245A_v1_150024	JV359A1_v1_11431			JV244B_v1_50204
		<i>P. fluorescens</i> Pf-5 PFL2980 – 2987										JV244B_v1_310046-310053
Fit Toxin	<i>fit</i> <i>ABCDEFGH</i>	<i>P. protegens</i> CHA0 PFLCHA0_c30230 – 30300						JV245A_v1_320059	JV359A1_v1_12359-12366			
TcaB2	<i>tcaB2</i>	<i>P. fluorescens</i> Q8r-96 AHPOv1_51192				JV359A8_v1_90103	JV551A7_v1_280059					
Pyoverdine	<i>fpvI</i> ; <i>pvdA</i> ; <i>pvdIJKDEON MP</i> ; <i>pvdSL</i> <i>fpvI</i> ; <i>pvdA</i> ; <i>pvdIJKDEON MP</i> ; <i>pvdSL</i>	<i>P. kilonensis</i> F113 PSF113_1856-1860; PSF113_1836-1847; PSF113_1749-1750	JV391D10_v1_11443-11448; JV391D10_v1_30406-0411; JV91D10_v1_10836-10383	JV391D17_v1_10188 – 10183; JV391D17_v1_110173-110178; JV391D17_v1_20648-20650	JV449_v1_190443-190448; JV449_v1_160036; JV449_v1_130004-130009; JV449_v1_25009-250011	JV359A8_v1_160033-160037; JV359A8_v1_160042; JV359A8_v1_50462-50463						

Secondary metabolite	Genes involved in biosynthesis	Reference Stain <sup>†</sup>	JV391D10	JV391D17	JV449	JV359A8	JV551A7	JV245A	JV359A1	JV251A	JV551A1	JV244B
Pyoverdine	<i>fpvI</i> ; <i>pvdA</i> ; <i>pvdIJKDEON</i>	<i>P. kilonensis</i> <b>F113</b>					JV551A7_v1_220127-220137;					
	<i>MP</i> ; <i>pvdSL</i>	PSF113					JV551A7_v1_250002-50005;					
	<i>fpvI</i> ; <i>pvdA</i> ;	PSF113					JV551A7_v1_220127					
	<i>pvdIJKDEON</i> <i>MP</i> ; <i>pvdSL</i>	1836- 1847; PSF113 1749-1750					JV551A7_v1_120100-120101					
Pyoverdine	<i>fpvI</i> ; <i>pvdA</i> ; <i>pvdIJKDEON</i>	<i>P. protegens</i> <b>CHA0</b>							JV359A1_v1_50011-50015;	JV251A_v1_40070-40074;		
	<i>MP</i> ; <i>pvdSL</i>	PFLCHA0					JV245A_v1_10197-10203;		JV359A1_v1_70005;	JV251A_v1_40050;		
	<i>fpvI</i> ; <i>pvdA</i> ;	c41390-41420;					JV245A_v1_10230;		JV359A1_v1_50003-50008;	JV251A_v1_40065-40069;		
	<i>pvdIJKDEON</i> <i>MP</i> ; <i>pvdSL</i>	PFLCHA0 c41450 – 41500; PFLCHA0 c42540-42530					JV245A_v1_270166-270167		JV359A1_v1_20048-20049	JV251A_v1_10782-10783		
Pyoverdine	<i>fpvI</i> ; <i>pvdA</i> ; <i>pvdIJKDEON</i>	<i>P. putida</i> <b>KT4420</b>									JV551A1_v1_390026-90029;	
	<i>MP</i> ; <i>pvdSL</i>	PP_3796; PP_4208-4211; PP_3807; PP_4212 - 4216; PP_4243 - 4244									JV551A1_v1_340007;JV551A_v1_340018;	
											JV551A1_v1_450006;	
											JV551A1_v1_420082;	
											JV551A1_v1_40001-40003;	
											JV551A1_v1_400017-400018	

Secondary metabolite	Genes involved in biosynthesis	Reference Stain <sup>†</sup>	JV391D10	JV391D17	JV449	JV359A8	JV551A7	JV245A	JV359A1	JV251A	JV551A1	JV244B
Pyoverdine	<i>fpvI</i> ; <i>pvdA</i> ; <i>pvdJKDEON</i> <i>MP</i> ; <i>pvdSL</i>	<i>P. putida</i> <b>KT4420</b> PP_3796; PP_4208-4211; PP_3807; PP_4212-4216; PP_4243 - 4244										JV244B_v1_390026-390029 JV244B_v1_800003-800006; JV244B_v1_530011-30015; JV244B_v1_340007; JV244B_v1_340018; JV244B_v1_270005; JV244B_v1_109003; JV244B_v1_60197; JV244B_v1_450006; JV244B_v1_420082; JV244B_v1_270004-270008 JV244B_v1_530003-30008; JV244B_v1_400001-400003; JV244B_v1_800002; JV244B_v1_40237-40239; JV244B_v1_400017-400018

[illegible]

Secondary metabolite	Genes involved in biosynthesis	Reference Stain †	JV391D10	JV391D17	JV449	JV359A8	JV551A7	JV245A	JV359A1	JV251A	JV551A1	JV244B
Rhs proteins	<i>rhsB</i>	<i>P. kilonensis</i> F113 PSF113_5809; PSF113_5811	JV391D10_ v1_20515; JV391D10_ v1_20517	JV391D17_ v1_30095; JV391D17_ v1_30097	JV449_v1_10069	JV359A8	JV551A7_ v1_320221; JV551A7_ v1_320223	JV245A_ v1_370227; JV245A_ v1_370228; JV245A_ v1_370231	JV359A1_ v1_30526	JV251A	JV551A1	JV244B
		<i>P. protegens</i> CHA0 PFLCHA0_c6055_0										
Rhs proteins												
Rhs proteins	<i>rhsB</i>	<i>P. aeruginosa</i> PAO1 PA2684									JV551A1_ v1_460141	
Type III secretion system core apparatus	<i>hrcU, hrcT, hrcS, hrcR, hrcQb, hrcQa, hrpP, hrpO, hrcN, hrpQ, hrcV, hrpJ, hrpL,</i>	<i>P. kilonensis</i> F113 PSF113_5599-5610	JV391D10_ v1_130097-130108α	JV391D17_ v1_11016-11027			JV551A7_ v1_320009-320018 α					
Massetolide A	<i>massA, massB, massC</i>	<i>P. fluorescens</i> SS101 AHPNV1_5051; AHPNV1_3771-3772							JV359A1_v1_11481-114883			

† Annotation of secondary metabolite biosynthesis clusters of all the strains was made by comparison (coverage > 80%; identity > 60%) with protein sequences of the reference strains. For *P. fluorescens* JV391D10, JV391D17 and JV449 the reference strain was *P. kilonensis* F113. For strains *P. protegens* JV245A and *P. jessenii* JV251A the reference strain used was *P. protegens* CHA0. For *P. brassicacerum* JV551A7 the reference strains were *P. kilonensis* F113 and *P. fluorescens* Q8r-96. For strain *P. kribbensis* JV359A8 the reference strains were *P. kilonensis* F113, *P. fluorescens* Pf-5 and *P. fluorescens* Q8r-96. For *P. protegens* JV359A1 the reference strain were *P. protegens* CHA0 and *P. fluorescens* SS101. For *P. inefficax* JV551A1 the reference strain used was *P. putida* KT2440 and *P. aeruginosa* PAO1. Lastly, for *P. wadenswilerensis* JV244B the reference strains were *P. putida* KT2440, *P. kilonensis* F113, *P. protegens* CHA0, *P. fluorescens* Pf-5 and *P. aeruginosa* PAO1.  
α annotation number presenting this symbol have an identity lower than 60% with the reference strains. (45%)

**Table S2: Pairwise comparison of the 10 *Pseudomonas* genomes vs. type strain genomes.**

Query strain	Subject strain	dDDH (d0, in %)	C.I. (d0, in %)	dDDH (d4, in %)	C.I. (d4, in %)	dDDH (d6, in %)	C.I. (d6, in %)	G+C content difference (in %)
'Pseudomonas JV359A1'	Pseudomonas protegens CHA0	100	[100.0 - 100.0]	100	[100.0 - 100.0]	100	[100.0 - 100.0]	0.03
'Pseudomonas JV359A1'	'Pseudomonas protegens CHA0'	100	[100.0 - 100.0]	100	[100.0 - 100.0]	100	[100.0 - 100.0]	0.03
'Pseudomonas protegens CHA0'	Pseudomonas protegens CHA0	100	[100.0 - 100.0]	100	[100.0 - 100.0]	100	[100.0 - 100.0]	0
'Pseudomonas fluorescens Q8r1-96'	Pseudomonas brassicacearum LMG 21623	98.1	[96.8 - 98.9]	99.6	[99.4 - 99.8]	99.1	[98.4 - 99.4]	0.15
'Pseudomonas sp. JV551A7'	Pseudomonas inefficax Pseudomonas sp. JV55'	97.3	[95.8 - 98.3]	97.8	[96.8 - 98.4]	98.5	[97.6 - 99.0]	0.08
'Pseudomonas fluorescens Q8r1-96'	Pseudomonas sp. JV551A7	93.8	[91.2 - 95.6]	96.1	[94.8 - 97.1]	96	[94.3 - 97.2]	0.06
'Pseudomonas sp. JV551A7'	Pseudomonas brassicacearum LMG 21623	93.9	[91.3 - 95.7]	95.4	[93.9 - 96.6]	96	[94.3 - 97.2]	0.09
'Pseudomonas JV391D10'	'Pseudomonas JV391D17'	93.6	[91.0 - 95.5]	94.5	[92.8 - 95.8]	95.7	[94.0 - 97.0]	0.13
'Pseudomonas protegens CHA0'	'Pseudomonas protegens H78'	96.9	[95.2 - 98.0]	91.7	[89.6 - 93.4]	97.7	[96.5 - 98.5]	0.01
'Pseudomonas JV359A1'	'Pseudomonas protegens H78'	96.8	[95.0 - 97.9]	91.7	[89.5 - 93.4]	97.6	[96.3 - 98.4]	0.02
'Pseudomonas protegens H78'	Pseudomonas protegens CHA0	96.9	[95.2 - 98.0]	91.7	[89.6 - 93.4]	97.7	[96.5 - 98.5]	0.01
'Pseudomonas protegens H78'	'Pseudomonas protegens JV245A'	90.1	[86.9 - 92.6]	90.9	[88.7 - 92.7]	92.8	[90.4 - 94.6]	0.04
'Pseudomonas fluorescens Pf-5'	'Pseudomonas protegens H78'	93	[90.3 - 95.0]	90.1	[87.8 - 92.0]	94.8	[92.8 - 96.2]	0.1
'Pseudomonas protegens CHA0'	'Pseudomonas protegens JV245A'	89.7	[86.4 - 92.3]	89.8	[87.5 - 91.8]	92.3	[89.9 - 94.2]	0.05
'Pseudomonas protegens JV245A'	Pseudomonas protegens CHA0	89.7	[86.4 - 92.3]	89.8	[87.5 - 91.8]	92.3	[89.9 - 94.2]	0.05
'Pseudomonas JV359A1'	'Pseudomonas protegens JV245A'	89.5	[86.2 - 92.1]	89.8	[87.4 - 91.7]	92.1	[89.7 - 94.0]	0.02
'Pseudomonas fluorescens Pf-5'	Pseudomonas protegens CHA0	94.5	[92.1 - 96.2]	89.6	[87.3 - 91.6]	95.8	[94.0 - 97.0]	0.09
'Pseudomonas fluorescens Pf-5'	'Pseudomonas JV359A1'	94.3	[91.8 - 96.1]	89.6	[87.2 - 91.5]	95.6	[93.9 - 96.9]	0.12
'Pseudomonas fluorescens Pf-5'	Pseudomonas protegens CHA0	94.5	[92.1 - 96.2]	89.6	[87.3 - 91.6]	95.8	[94.0 - 97.0]	0.09
'Pseudomonas fluorescens Pf-5'	'Pseudomonas protegens JV245A'	86.2	[82.5 - 89.2]	88.9	[86.5 - 90.9]	89.4	[86.6 - 91.7]	0.14
'Pseudomonas JV244B'	Pseudomonas wadsworthii DSM 13647	79.4	[75.4 - 82.8]	88.1	[85.6 - 90.2]	83.8	[80.5 - 86.6]	0.37
'Pseudomonas protegens Cab57'	Pseudomonas protegens CHA0	97.3	[95.8 - 98.3]	87.4	[84.8 - 89.6]	97.6	[96.3 - 98.4]	0.05
'Pseudomonas protegens Cab57'	'Pseudomonas protegens CHA0'	97.3	[95.8 - 98.3]	87.4	[84.8 - 89.6]	97.6	[96.3 - 98.4]	0.05
'Pseudomonas JV359A1'	'Pseudomonas protegens Cab57'	97.2	[95.5 - 98.2]	87.3	[84.8 - 89.5]	97.4	[96.2 - 98.3]	0.09
'Pseudomonas protegens Cab57'	'Pseudomonas protegens H78'	96	[93.9 - 97.4]	86.5	[83.8 - 88.7]	96.5	[94.9 - 97.6]	0.07
'Pseudomonas protegens Cab57'	'Pseudomonas protegens JV245A'	88.3	[84.9 - 91.1]	85.4	[82.7 - 87.7]	90.6	[87.9 - 92.7]	0.11
'Pseudomonas fluorescens Pf-5'	'Pseudomonas protegens Cab57'	93.7	[91.1 - 95.6]	85	[82.3 - 87.4]	94.7	[92.6 - 96.1]	0.03
'Pseudomonas JV251A'	Pseudomonas jessenii DSM 17150	79	[75.1 - 82.5]	78.3	[75.4 - 81.0]	81.8	[78.5 - 84.7]	0.04
'Pseudomonas putida F1'	'Pseudomonas putida KT2440'	75.6	[71.6 - 79.2]	75	[72.0 - 77.8]	78.3	[74.8 - 81.4]	0.34
'Pseudomonas fluorescens Pf0-1'	Pseudomonas glyciniae MS586	85	[81.3 - 88.1]	65.1	[62.2 - 67.9]	84.1	[80.9 - 86.9]	0.04
'Pseudomonas fluorescens F113'	Pseudomonas kilnensis DSM 13647	75.1	[71.1 - 78.7]	64.5	[61.6 - 67.3]	75.6	[72.1 - 78.8]	0.08
'Pseudomonas fluorescens F113'	'Pseudomonas fluorescens Q8r1-96'	78.3	[74.3 - 81.8]	60.9	[58.1 - 63.7]	77.4	[74.0 - 80.6]	0.17
'Pseudomonas fluorescens F113'	Pseudomonas sp. JV551A7	78.3	[74.3 - 81.8]	60.7	[57.9 - 63.5]	77.4	[73.9 - 80.5]	0.11
'Pseudomonas fluorescens F113'	Pseudomonas brassicacearum LMG 21623	78.6	[74.6 - 82.1]	60.5	[57.7 - 63.3]	77.6	[74.1 - 80.7]	0.02
'Pseudomonas sp. JV551A7'	Pseudomonas kilnensis DSM 13647	75.9	[72.0 - 79.5]	59.1	[56.3 - 61.8]	75	[71.5 - 78.2]	0.03
'Pseudomonas fluorescens Q8r1-96'	Pseudomonas kilnensis DSM 13647	76	[72.0 - 79.6]	59	[56.2 - 61.8]	75	[71.5 - 78.2]	0.09
'Pseudomonas sp. JV551A1'	Pseudomonas asiatica JCM 32716T	70.8	[66.9 - 74.5]	57.4	[54.6 - 60.1]	70.2	[66.8 - 73.4]	0.21
'Pseudomonas JV359A8'	Pseudomonas glyciniae MS586	81.2	[77.3 - 84.6]	57	[54.3 - 59.8]	78.8	[75.4 - 81.9]	0.18
'Pseudomonas fluorescens Pf0-1'	'Pseudomonas JV359A8'	80.6	[76.7 - 84.0]	56.7	[53.9 - 59.4]	78.2	[74.8 - 81.3]	0.14
'Pseudomonas sp. JV449'	Pseudomonas grimmii DSM 17515	67	[63.2 - 70.7]	54.6	[51.9 - 57.3]	66.2	[62.8 - 69.4]	0.85
'Pseudomonas fluorescens F113'	Pseudomonas thivervalensis LMG 21626	71.2	[67.3 - 74.9]	46.7	[44.1 - 49.3]	67.2	[63.8 - 70.5]	0.41
'Pseudomonas sp. JV551A7'	Pseudomonas thivervalensis LMG 21626	75.5	[71.5 - 79.1]	46.2	[43.6 - 48.7]	70.5	[67.1 - 73.7]	0.3
'Pseudomonas fluorescens Q8r1-96'	Pseudomonas thivervalensis LMG 21626	75.2	[71.2 - 78.8]	46	[43.5 - 48.6]	70.2	[66.8 - 73.5]	0.24
'Pseudomonas JV251A'	Pseudomonas laurylsulfatiphila AP3 16	66.4	[62.5 - 70.0]	45.9	[43.3 - 48.4]	63	[59.7 - 66.2]	0.49
'Pseudomonas JV251A'	Pseudomonas laurylsulfativorans AP3 22	65.8	[62.0 - 69.4]	45.8	[43.2 - 48.4]	62.5	[59.2 - 65.7]	0.04
'Pseudomonas fluorescens Pf0-1'	Pseudomonas kribbensis KCTC 32541T	75	[71.0 - 78.6]	42.6	[40.1 - 45.2]	68.6	[65.2 - 71.9]	0.02
'Pseudomonas putida GB-1'	Pseudomonas putida NBRC 14164	70.1	[66.2 - 73.8]	42.4	[39.9 - 44.9]	64.7	[61.4 - 67.9]	0.39
'Pseudomonas JV359A8'	Pseudomonas kribbensis KCTC 32541T	73	[69.0 - 76.6]	42.4	[39.9 - 45.0]	67	[63.6 - 70.2]	0.12
'Pseudomonas fluorescens Q8r1-96'	Pseudomonas viciae 11K1	65.2	[61.4 - 68.8]	39.4	[37.0 - 42.0]	59.6	[56.4 - 62.8]	0.62
'Pseudomonas sp. JV551A7'	Pseudomonas viciae 11K1	65.6	[61.8 - 69.2]	39.2	[36.7 - 41.7]	59.8	[56.6 - 63.0]	0.56
'Pseudomonas JV359A8'	Pseudomonas koreensis LMG 21318	66	[62.2 - 69.6]	34.2	[31.8 - 36.7]	57.8	[54.6 - 60.9]	0.13
'Pseudomonas JV359A8'	Pseudomonas koreensis JCM 14769	66	[62.2 - 69.6]	34.2	[31.7 - 36.7]	57.8	[54.6 - 60.9]	0.08
'Pseudomonas fluorescens Pf0-1'	Pseudomonas atagosis PS14	67.5	[63.6 - 71.1]	34	[31.5 - 36.5]	58.8	[55.5 - 61.9]	0.9
'Pseudomonas JV359A8'	Pseudomonas atagosis PS14	67.2	[63.3 - 70.8]	33.8	[31.4 - 36.3]	58.4	[55.2 - 61.6]	1.04
'Pseudomonas JV359A1'	Pseudomonas piscis KCTC 72033	65.5	[61.7 - 69.1]	31.9	[29.5 - 34.4]	56.2	[53.0 - 59.3]	0.15
'Pseudomonas protegens CHA0'	Pseudomonas piscis KCTC 72033	65.6	[61.8 - 69.3]	31.9	[29.5 - 34.4]	56.3	[53.1 - 59.4]	0.18
'Pseudomonas protegens Cab57'	Pseudomonas piscis KCTC 72033	65.8	[61.9 - 69.4]	31.8	[29.4 - 34.3]	56.3	[53.1 - 59.4]	0.23
'Pseudomonas JV359A1'	Pseudomonas aestus CMAA1215	66	[62.2 - 69.6]	31.7	[29.2 - 34.2]	56.4	[53.2 - 59.5]	0.38
'Pseudomonas protegens CHA0'	Pseudomonas aestus CMAA1215	65.9	[62.1 - 69.5]	31.7	[29.2 - 34.2]	56.3	[53.1 - 59.4]	0.41
'Pseudomonas protegens Cab57'	Pseudomonas aestus CMAA1215	66.2	[62.4 - 69.9]	31.5	[29.1 - 34.0]	56.5	[53.3 - 59.6]	0.47

Table S3: Type based species and subspecies clustering. Briefly, the clustering yielded 99 species clusters and the provided query strains were assigned to 12 of these. Moreover, the strains from this study were located in 13 of 103 subspecies cluster.

TYGS ID	Kind	Species cluster	Subspecies cluster	Preferred name	Deposit	Authority	Other deposits	Synonymous taxon names	Base pairs	Percent G+C	No. proteins	Goldstamp	Bioproject accession	Biosample accession	Assembly accession	IMG OID	BacDive
1033	type strain	1	0	<i>Pseudomonas putida</i>	NBRC 14164	(Trevisan 1889) Migula 1895	CFBP 2086; LMG 2257; CCUG 12690; DSM 291; DSM 7314; JCM 13063; JCM 20120; ATCC 12633; IFO 14164; NCTC 10936; CIP 52191; HAMBI 7; IAM 1236; NCIM B.01634; NCCB 68020; NCCB 72066	<i>Bacillus putidus</i> ; <i>Pseudomonas putida</i>	6,156,701	62.33	5449	Gp0014652	PRJDB191	SAMD00061028	GCA_000412675	2554235365	strain page
10406	type strain	2	1	<i>Pseudomonas alagosis</i>	PS14	Morimoto et al. 2020	LMG 31496; CECT 9940	<i>Pseudomonas alagosis</i>	6,115,729	59.62	5423		PRJNA565185	SAMN12736367	GCA_011369465	species page	
10758	type strain	3	2	<i>Pseudomonas kairouanensis</i>	KC12T	Oueslati et al. 2020	CFBP 8682; CECT 9766	<i>Pseudomonas kairouanensis</i>	6,694,223	60.38	6077		PRJNA485038	SAMN0781217	GCA_004682055	species page	
10800	type strain	4	3	<i>Pseudomonas asiatica</i>	JCM 32716T	Tohya et al. 2019 emend. Tohya et al., 2020	DSM 107182; RYU5	<i>Pseudomonas asiatica</i>	5,969,036	62.55	5447		PRJDB7090	SAMD00127555	GCA_009932335	species page	
11045	type strain	5	4	<i>Pseudomonas baetica</i>	LMG 25716	López et al. 2012	a390; CECT 7720	<i>Pseudomonas baetica</i>	6,902,809	58.76	5980	Gp0112912	PRJNA303516	SAMN04488362	GCA_002813455	strain page	
11118	type strain	6	5	<i>Pseudomonas caspiana</i>	FBF102	Burquets et al. 2017	CCUG 69273; CECT 9164	<i>Pseudomonas caspiana</i>	6,053,735	58.95	5539	Gp0323546	PRJNA224116	SAMN04317010	GCF_002158995	strain page	
11462	type strain	7	6	<i>Pseudomonas grimontii</i>	DSM 17515	Balda et al. 2002	ATCC BAA-140; CFML 97-514; CIP 106645	<i>Pseudomonas grimontii</i>	7,107,743	60.07	6457	Gp0116961	PRJNA224116	SAMN04490186	GCF_900101085	strain page	
11595	type strain	8	7	<i>Pseudomonas reidhizensis</i>	CCOS 865	Frasson et al. 2017	LMG 28328	<i>Pseudomonas reidhizensis</i>	6,161,977	64.09	5441	Gp0365760	PRJEB28254	SAMEA4836241	GCA_900536025	strain page	
12479	type strain	9	102	<i>Pseudomonas brassicaearum</i>	LMG 21623	Achouak et al. 2000 emend. Ivanova et al., 2009	CFBP 11706; DSM 13227; JCM 11938; CIP 107059; strain DBK11	<i>Pseudomonas brassicaearum</i> ; <i>Pseudomonas brassicaearum</i> subsp. <i>brassicaearum</i>	6,776,559	60.8	5951	Gp0116954	PRJNA224116	SAMN04490180	GCF_900103245	strain page	
U70608	user strain	9	102	<i>Pseudomonas fluorescens</i> C6r1-96					6,602,511	60.96	5783						
U70624	user strain	9	102	<i>Pseudomonas</i> sp. JVS17A7					6,735,578	60.9	5964						
12637	type strain	10	8	<i>Pseudomonas chlororaphis</i> subsp. <i>aurantiaca</i>	DSM 19603	(Nakhimovskaya 1946) Peix et al. 2007	ATCC 33663; VKM B-876; CIP 106718; NCIB 10068; NCIMB 10068	<i>Pseudomonas chlororaphis</i> subsp. <i>aurantiaca</i> ; <i>Pseudomonas chlororaphis</i> subsp. <i>aurantiaca</i>	7,109,779	62.86	6297	Gp0442278	PRJNA224116	SAMN08359190	GCF_003851835	strain page	
4463	type strain	10	57	<i>Pseudomonas aureofaciens</i>	NBRC 3521	(Kluyver 1956) Peix et al. 2007	CFBP 2133; LMG 1245; DSM 6698; ATCC 13885; NCIMB 9030	<i>Pseudomonas aureofaciens</i> ; <i>Pseudomonas aureofaciens</i> subsp. <i>aureofaciens</i>	6,973,338	62.75	6245	Gp0023849	PRJDB203	SAMD00018760	GCA_000813225	strain page	
12777	type strain	11	9	<i>Pseudomonas laurylultrapilla</i>	AP3_16	Furnaczkyk et al. 2019	DSM 105097; PCM 2903	<i>Pseudomonas laurylultrapilla</i>	6,684,644	60.15	5879	Gp0359215	PRJNA390162	SAMN07222753	GCA_002934665	species page	
13725	type strain	12	10	<i>Pseudomonas avellanae</i>	BPIC 631	Janae et al. 1987	NCPBP 3487; DSM 11809; JCM 11937; BPIC 631; CIP 105176; F11	<i>Pseudomonas avellanae</i>	5,847,420	58.71	4713	Gp0020557	PRJNA94293	SAMN02471966	GCA_000302915	2531839710	strain page
13736	type strain	13	11	<i>Pseudomonas extremorientalis</i>	LMG 19695	Ivanova et al. 2002	DSM 15824; KMM 3447	<i>Pseudomonas extremorientalis</i>	6,345,601	60.93	5686		PRJNA224116	SAMN05561041	GCF_001870465	strain page	
13746	type strain	14	12	<i>Acinetobacter lactuca</i>	NRRL B-41902	Rooney et al. 2016	CCUG 68785	<i>Acinetobacter lactuca</i>	3,916,154	38.64	3541	Gp0145455	PRJNA309605	SAMN04037545	GCA_001605885	strain page	
17614	type strain	14	26	<i>Acinetobacter dikshorniae</i>	JVAP01	Osogaya et al. 2016	LMG 29805; CECT 9134	<i>Acinetobacter dikshorniae</i>	3,905,717	38.83	3467	Gp0145294	PRJNA224116	SAMN04038486	GCF_001595745	strain page	
13937	type strain	15	13	<i>Pseudomonas nabeulensis</i>	E10BT	Oueslati et al. 2020	CFBP 8661; CECT 9765	<i>Pseudomonas nabeulensis</i>	6,758,655	60.29	6049		PRJNA485041	SAMN09781218	GCA_004682045	species page	
1406	type strain	16	14	<i>Pseudomonas vanovensis</i>	DSM 16006	Tyrzová et al. 2006	282; CCM 7279	<i>Pseudomonas vanovensis</i>	5,701,836	61.52	5242	Gp0021956	PRJNA188914	SAMN02441181	GCA_000425805	2523533545	strain page
1408	type strain	17	15	<i>Pseudomonas parvula</i>	DSM 17004	Uchino et al. 2002	AJ 2128; NBRC 501; JCM 11244; IFO 16336; NBRC 16336; CB-1; CIP 107617	<i>Pseudomonas parvula</i>	4,952,945	62.49	4459	Gp0021954	PRJNA188912	SAMN02441530	GCA_000425765	2523533547	strain page
14146	type strain	18	16	<i>Pseudomonas alkylphenolica</i>	JCM 16553	Mulet et al. 2015	KCTC 22206; KL28	<i>Pseudomonas alkylphenolica</i>	5,764,622	60.63	5281	Gp0101030	PRJNA224116	SAMN02929205	GCF_000746525	species page	
14196	type strain	19	17	<i>Pseudomonas reinkei</i>	MT1	Cámara et al. 2007	CCUG 53116; DSM 18361	<i>Pseudomonas reinkei</i>	6,249,573	59.17	5603	Gp0359321	PRJNA359931	SAMN06196581	GCA_001945365	strain page	
1422	type strain	20	18	<i>Pseudomonas taiwanensis</i>	DSM 21245	Wang et al. 2010	BCRC 17751; CMS	<i>Pseudomonas taiwanensis</i>	5,414,466	61.86	4974	Gp0021955	PRJNA188913	SAMN02440964	GCA_000425785	2523533561	strain page
14344	type strain	21	55	<i>Pseudomonas japonica</i>	DSM 22348	Furugami et al. 2008	DSM 22348; JCM 21532; NBRC 103040; IAM 15071; TISTR 1526; WL 103040; IAM 15071; TISTR 1526; WL 103040	<i>Pseudomonas japonica</i>	6,718,127	64.14	5758	Gp0148719	PRJNA331484	SAMN05443352	GCA_900188455	2700988710	strain page
4028	type strain	21	55	<i>Pseudomonas japonica</i>	NBRC 103040	Furugami et al. 2008	DSM 22348; JCM 21532; NBRC 103040; IAM 15071; TISTR 1526; WL 103040	<i>Pseudomonas japonica</i>	6,663,130	64.16	5740	Gp0024160	PRJDB212	SAMD00018431	GCA_000730585	strain page	
14418	type strain	22	74	<i>Pseudomonas amygdali</i>	CFBP 3205	Psallidas and Panagopoulos 1975	NCPBP 2807; LMG 13184; LMG 2123; ICMP 3918; CCUG 32770; DSM 7298; ATCC 33614; CIP 106734	<i>Pseudomonas amygdali</i>	5,719,124	58.29	5288	Gp0120625	PRJNA274666	SAMN03328995	GCA_000935645	strain page	
5875	type strain	22	74	<i>Pseudomonas meliae</i>	CFBP 3225	Ogimi 1981	NCPBP 3033; LMG 2220; ICMP 6289; DSM 6759; MAFF 301463; ATCC 33050; No. 2; Ogimi 2	<i>Pseudomonas meliae</i>	5,143,623	58.4	4803	Gp0120629	PRJNA224116	SAMN03328996	GCF_000935675	species page	
5891	type strain	22	74	<i>Pseudomonas savastanoi</i>	ICMP 4352	(Janse 1982) Gardan et al. 1992	CFBP 1670; NCPBP 639; LMG 2209; DSM 19341; DSM 50286; ATCC 13522; CIP 103721	<i>Pseudomonas savastanoi</i> ; <i>Pseudomonas syringae</i> subsp. <i>savastanoi</i>	6,022,960	57.99	5578	Gp0146403	PRJNA224116	SAMN03976268	GCF_001401285	strain page	
14671	type strain	23	19	<i>Pseudomonas kribbensis</i>	KCTC 32541T	Chang et al. 2016	46-2; DSM 100278	<i>Pseudomonas kribbensis</i>	6,324,282	60.55	5661	Gp0365917	PRJNA224116	SAMN09244908	GCF_003352185	species page	
14783	type strain	24	33	<i>Pseudomonas piscis</i>	KCTC 72033	Liu et al. 2020	MCCC 1103575; MCCC 1103575; MCCC 1103575	<i>Pseudomonas piscis</i>	6,921,954	63.57	6294		PRJNA578493	SAMN13064563	GCA_000930155	species page	
18793	type strain	24	33	<i>Pseudomonas aestus</i>	CMAA1215	Vasconcelos et al. 2017	NRRL B-653100; CMAA-11962	<i>Pseudomonas aestus</i>	6,658,223	63.8	6111		PRJNA215329	SAMN02318190	GCA_000474765	species page	



Table S3 (continued)

TYGS ID	Kind	Species cluster	Subspecies cluster	Preferred name	Deposit	Authority	Other deposits	Synonymous taxon names	Base pairs	Percent G+C	No. proteins	Goldstamp	Bioproject accession	Biosample accession	Assembly accession	IMG OID	BackDive
15027	type strain	25	20	<i>Pseudomonas grigoloniensis</i>	J13	Wang et al. 2019	KCTC 42384 <sup>†</sup> ; JCM 32579; CGMCC 1.16493	<i>Pseudomonas grigoloniensis</i>	5,734,615	64.22	5201		PRJNA419834	SAMN08096959	GCA_002806685		species page
15057	type strain	26	101	<i>Pseudomonas</i> <i>as</i> sp. <i>inflexax</i> <i>JV551A3</i>		Keshavarz-Tohid et al. 2019	CFBP9493; DSM 108619; JV551A3	<i>Pseudomonas inflexax</i>	6,212,057	62.83	6116		PRJEB24815	SAMEA104569200	GCA_900277125		species page
U70623	user strain	26	101	<i>Pseudomonas</i> sp. <i>JV551A1</i>					5,889,653	62.75	5288						
15008	type strain	27	21	<i>Pseudomonas moorei</i>	DSM 12847	Cámara et al. 2007	CCUG 53114; RW10	<i>Pseudomonas moorei</i>	6,546,438	59.66	5989	Gp0116970	PRJNA224116	SAMN04490195	GCF_900102045		strain page
16162	type strain	28	22	<i>Pseudomonas antiaetica</i>	LMG 22709	Reddy et al. 2004	CCUG 49625; DSM 15318; OMS 35; MTC 4992	<i>Pseudomonas antiaetica</i>	6,377,074	59.52	5785	Gp0116953	PRJNA224116	SAMN04490179	GCF_900103795		strain page
16543	type strain	29	23	<i>Pseudomonas truceae</i>	SNU WT1	Oh et al. 2019	KCTC 72285; JCM 33436	<i>Pseudomonas truceae</i>	5,685,196	61.83	5033		PRJNA522009	SAMN10922936	GCA_004214895		species page
1757	type strain	30	24	<i>Pseudomonas montellii</i>	DSM 14164	Elomar et al. 1997	CCUG 38736; JCM 21585; ATCC 700476; NBRC 103158; CFML 90-60; CIP 104883	<i>Pseudomonas montellii</i>	6,310,507	61.49	5953	Gp0021953	PRJNA221052	SAMN02743994	GCA_000621245	2556921620	strain page
17600	type strain	31	25	<i>Pseudomonas cedrina</i>	DSM 17516	Dabbousi et al. 2002	CFML 96-198; CIP 105541	<i>Pseudomonas cedrina</i>	6,513,567	60.71	5835	Gp0376649	PRJNA224116	SAMN05945970	GCF_001993175		strain page
1769	type strain	32	27	<i>Pseudomonas plecoglossoides</i>	DSM 15088	Nishimori et al. 2000	CCUG 49623; JCM 21584; ATCC 700383; NBRC 103162; CIP 106493; FPC 951	<i>Pseudomonas plecoglossoides</i>	5,349,463	62.97	4847	Gp0040000	PRJNA221049	SAMN02743900	GCA_000688275	2556921648	strain page
1770	type strain	33	28	<i>Pseudomonas fulva</i>	DSM 17717	Iizuka and Komagata 1963	LMG 11722; NRIC 180; JCM 11242; ATCC 31418; IFO 16637; NBRC 16637; CIP 106785; CIP 106819; IAM 1529	<i>Pseudomonas fulva</i>	4,766,154	61.77	4278	Gp0039998	PRJNA221053	SAMN02743980	GCA_000621265	2556921649	strain page
1774	type strain	34	29	<i>Pseudomonas mosseli</i>	DSM 17497	Dabbousi et al. 2002	ATCC BAA-99; CFML 90-83; CIP 105259	<i>Pseudomonas mosseli</i>	6,257,170	64.01	5763	Gp0039999	PRJNA221051	SAMN02743979	GCA_000621225	2556921653	strain page
17880	type strain	35	30	<i>Pseudomonas laryngis/oleovorans</i>	AP3_22	Furmanczyk et al. 2019	DSM 105086; PCM 2804	<i>Pseudomonas laryngis/oleovorans</i>	6,680,726	59.62	5877	Gp0224266	PRJNA369606	SAMN06289919	GCA_002906155		species page
17980	type strain	36	31	<i>Acinetobacter oleovorans</i>	KCTC 23045	Kang et al. 2011	KCTC 23045; JCM 16687; DR1	<i>Acinetobacter oleovorans</i>	4,102,937	38.62	3841	Gp0130481	PRJDB1754	SAMN00019889	GCA_000836035		species page
17981	type strain	36	31	<i>Acinetobacter oleovorans</i>	JCM 16667	Kang et al. 2011	KCTC 23045; JCM 16667; DR1	<i>Acinetobacter oleovorans</i>	4,152,543	38.73	3852	Gp0006987	PRJNA224116	SAMN02803211	GCF_000196795		species page
18212	type strain	37	32	<i>Pseudomonas fildensis</i>	KG01	Pavlov et al. 2020	CECT 9084; DSM 102036	<i>Pseudomonas fildensis</i>	6,310,821	60.14	5594		PRJNA287706	SAMN03785346	GCA_001050345		species page
19005	type strain	38	34	<i>Pseudomonas moravensis</i>	LMG 24280	Tizová et al. 2006	1B4; DSM 16007; JCM 14770; CCM 7280	<i>Pseudomonas moravensis</i>	6,092,441	60.17	5351	Gp0116971	PRJNA224116	SAMN04490196	GCF_900105805		strain page
19344	type strain	39	92	<i>Pseudomonas waderswieriensis</i>	CCOS 864	Frasson et al. 2017	LMG 29327	<i>Pseudomonas waderswieriensis</i>	5,963,749	62.39	5436	Gp0357349	PRJEB27830	SAMEA4796555	GCA_900497695		strain page
U70609	user strain	39	92	<i>Pseudomonas klugii</i>	MAFF 212408T	Sawada et al. 2020	ICMP 23530	<i>Pseudomonas klugii</i>	7,442,190	62.76	6767		PRJNA224116	SAMN12688074	GCF_009296165		species page
19903	type strain	40	35	<i>Pseudomonas sielensis</i>	A3	Kaminski et al. 2018	DSM 103370; PCM 2856	<i>Pseudomonas sielensis</i>	5,861,721	59.48	5538	Gp0137126	PRJNA315635	SAMN04566726	GCA_001061075		strain page
20396	type strain	41	36	<i>Pseudomonas frederiksbergensis</i>	LMG 19851	Andersen et al. 2000	DSM 13022; ATCC BAA-257; CIP 108887; JAJ28	<i>Pseudomonas frederiksbergensis</i>	6,823,538	59.58	5875	Gp0116960	PRJNA224116	SAMN04490185	GCF_900106495		strain page
20462	type strain	42	37	<i>Pseudomonas schuaniensis</i>	WCHP5060	Qin et al. 2019	ONCTC 7662; GDACC 11424	<i>Pseudomonas schuaniensis</i>	6,094,336	58.93	5783	Gp0367339	PRJNA415331	SAMN039302769	GCA_003231305		species page
20536	type strain	43	38	<i>Pseudomonas canadensis</i>	2-92	Tambong et al. 2017	LMG 28496; DOAB 798	<i>Pseudomonas canadensis</i>	6,018,116	63.87	5465	Gp0070925	PRJNA224116	SAMN02469303	GCF_000503215		strain page
22735	type strain	44	39	<i>Pseudomonas unsorgensis</i>	DSM 18611	Kwon et al. 2003	LMG 21317; KACC 10847; Ps 3-10	<i>Pseudomonas unsorgensis</i>	4,937,085	60.38	4442	Gp0365413	PRJNA390488	SAMN07236649	GCA_002236105		strain page
23191	type strain	45	40	<i>Pseudomonas jessenii</i>	DSM 17150	Verhille et al. 1999	CCUG 42059; ATCC 700870; CFML 95-307; CIP 105274	<i>Pseudomonas jessenii</i>	6,701,403	59.73	5937	Gp0365413	PRJNA390488	SAMN07236649	GCA_002236105		strain page
23196	type strain	46	41	<i>Pseudomonas jessenii</i>	DSM 17150	Behrendt et al. 2003	CCUG 42059; ATCC 700870; CFML 95-307; CIP 105274	<i>Pseudomonas jessenii</i>	6,537,206	59.7	5826	Gp0365412	PRJNA390484	SAMN07236640	GCA_002236115		strain page
U70610	user strain	46	93	<i>Pseudomonas lutea</i>	LMG 21995	Behrendt et al. 2007	DSM 15835; P 51318	<i>Pseudomonas lutea</i>	6,509,003	59.66	5855						
23686	type strain	47	42	<i>Pseudomonas rhodesiae</i>	LMG 17764	Corleier et al. 1997	CCUG 38732; DSM 14020; JCM 11940; CIP 104664	<i>Pseudomonas rhodesiae</i>	6,311,648	60.61	5821	Gp0112929	PRJNA303504	SAMN004488484	GCA_002563895		strain page
23732	type strain	48	43	<i>Acinetobacter calcoaceticus</i>	DSM 30006	(Beljunc 1911)	LMG 1046; CCUG 12804; CALM 17; JCM 6842; ATCC 23055; NCTC 12983; CIP 81 08; NCCB 22016	<i>Acinetobacter calcoaceticus</i>	6,417,699	60.17	5908	Gp0116988	PRJNA224116	SAMN004490209	GCF_900105575		strain page
3039	type strain	49	44	<i>Acinetobacter baumannii</i>	DSM 30006	Baumann et al. 1986	LMG 996; CCUG 688; DSM 6862; ATCC 17906; NCTC 12155; B40; Mannheim 2446160; NCCB 85026	<i>Acinetobacter baumannii</i>	3,926,967	38.65	3654	Gp0034167	PRJNA183274	SAMN01628159	GCA_000368965	2534682017	species page
3043	type strain	50	45	<i>Acinetobacter pili</i>	CIP 64.3	(ex-Sterzel and Mannheim 1963)	LMG 1035; DSM 25618; ATCC 19004; NIPH 519; RUH 2206	<i>Acinetobacter pili</i>	3,485,216	38.65	3322	Gp0033080	PRJNA183265	SAMN01628153	GCA_000369065	2537561554	strain page
3047	type strain	51	46	<i>Acinetobacter pili</i>	CIP 70.29	Nemec et al. 2011	LMG 1035; DSM 25618; ATCC 19004; NIPH 519; RUH 2206	<i>Acinetobacter pili</i>	3,796,866	38.8	3502	Gp0032907	PRJNA183266	SAMN01628154	GCA_000369045	2537561555	strain page
3125	type strain	52	47	<i>Acinetobacter brisquii</i>	CIP 110357	Avandham et al. 2011	BYN5-8; DSM 18516; KACC 11602	<i>Acinetobacter brisquii</i>	3,145,929	41.72	2938	Gp0047563	PRJNA217359	SAMN02335354	GCA_000488275	2578778708	strain page
3346	type strain	53	48	<i>Pseudomonas extremaustralis</i>	14-3	López et al. 2010	DSM 17835; CIP 109839; strain CT14; CFBP 1392; NCPPB 281; LMG 1247; NRRL B-1631; ICMP 3023; CCUG 14279; DSM 10604; DSM 6893; ATCC 19310; CIP 106898; NCAM B.01398	<i>Pseudomonas extremaustralis</i>	6,586,243	60.66	5870	Gp0017727	PRJNA77729	SAMN02469910	GCA_000242115	2548876812	strain page
3385	type strain	54	49	<i>Pseudomonas syringae</i>	KCTC 12500	van Hall 1902	DSM 17835; CIP 109839; strain CT14; CFBP 1392; NCPPB 281; LMG 1247; NRRL B-1631; ICMP 3023; CCUG 14279; DSM 10604; DSM 6893; ATCC 19310; CIP 106898; NCAM B.01398	<i>Pseudomonas syringae</i>	6,150,048	58.93	5082	Gp0070928	PRJNA227265	SAMN02406416	GCA_000507185	2571042114	species page
3423	type strain	55	50	<i>Pseudomonas marginalis</i>	ICMP 3553	(Brown 1916) Stevens 1925	CFBP 1387; CFBP 2037; NCPPB 667; LMG 2210; LMG 2215; DSM 13124; ATCC 10844; CIP 106712	<i>Pseudomonas marginalis</i>	6,613,668	60.41	5807	Gp0150108	PRJNA286744	SAMN04102629	GCA_001645105		strain page

Table S3 (continued)

TVGS ID	Kind	Species cluster	Subspecies cluster	Preferred name	Deposit	Authority	Other deposits	Synonymous taxon names	Base pairs	Percent G+C	No. proteins	Goldstamp	Bioproject accession	Biosample accession	Assembly accession	IMG OID	BacDive
3514	type strain	56	51	<i>Pseudomonas thivervalensis</i>	LMG 21626	Achouak et al. 2000	CFBP 11261; DSM 13194; JCM 11941; SBK 26	<i>Pseudomonas thivervalensis</i>	6,551,379	61.2	5609	Gp0149496	PRJNA298287	SAMN04017594	GCA_001637285		strain page
3579	type strain	57	52	<i>Pseudomonas mediterranea</i>	CFBP 5447	Calara et al. 2002	ICMP 14184; DSM 16733	<i>Pseudomonas mediterranea</i>	6,306,612	61.23	5355	Gp0102154	PRJNA210952	SAMN03112705	GCA_000774145		strain page
37260	type strain	58	53	<i>Pseudomonas glyciniae</i>	MS586	Jia et al. 2020	LMG 30275; NRRL B-65441	<i>Pseudomonas glyciniae</i>	6,396,728	60.48	5674		PRJNA309549	SAMN04455621	GCA_001594225		species page
3779	type strain	59	96	<i>Pseudomonas protegens</i>	CHAO	Ranette et al. 2012	CFBP 6505; DSM 19095	<i>Pseudomonas protegens</i>	6,867,980	63.39	6115	Gp0014459	PRJNA78307	SAMN02603271	GCA_000397205	2554235341	strain page
U70607	user strain	59	96	<i>Pseudomonas fluorescens Pf-5</i>					7,074,893	63.3	6276						
U70611	user strain	59	96	<i>Pseudomonas JY359A1</i>					6,850,953	63.42	6136						
U70615	user strain	59	96	<i>Pseudomonas protegens Cab57</i>					6,827,892	63.34	6137						
U70616	user strain	59	96	<i>Pseudomonas protegens CHAO</i>					6,867,980	63.39	6141						
U70617	user strain	59	96	<i>Pseudomonas protegens H78</i>					7,032,394	63.4	6300						
U70618	user strain	59	96	<i>Pseudomonas protegens JY245A</i>					6,383,769	63.44	6614						
3878	type strain	60	54	<i>Acinetobacter seifertii</i>	NIPH 973	Nemec et al. 2015	CCUG 34785; CCM 8535; CIP 110471; Geimer-Smidt 10090; LUH 1472	<i>Acinetobacter seifertii</i>	4,212,819	38.6	4051	Gp0039588	PRJNA183323	SAMN01828183	GCA_000388065		strain page
44545	type strain	61	56	<i>Pseudomonas hutereaui</i>	XWS2T	Xiang et al. 2019	OCTCCAB 2018189; KACC 19898	<i>Pseudomonas hutereaui</i>	6,167,580	61.08	5523		PRJNA174423	SAMN09303047	GCA_004025535		species page
4611	type strain	62	58	<i>Pseudomonas donghaiensis</i>	HYS	Gao et al. 2015	NRRL B-59108; CCTCC AB 2012141	<i>Pseudomonas donghaiensis</i>	5,639,475	62.42	5138	Gp0021513	PRJNA89717	SAMN02472130	GCA_000259195		strain page
19103	type strain	63	59	<i>Pseudomonas koreensis</i>	LMG 21318	Kwon et al. 2003	LMG 21318; DSM 16510; JCM 14769; KACC 10846; Ps 9-14	<i>Pseudomonas koreensis</i>	6,123,813	60.53	5467	Gp0116964	PRJNA224116	SAMN04490189	GCF_900101415		strain page
48496	type strain	63	59	<i>Pseudomonas koreensis</i>	JCM 14769	Kwon et al. 2003	LMG 21318; DSM 16510; JCM 14769; KACC 10846; Ps 9-14	<i>Pseudomonas koreensis</i>	6,064,848	60.58	5520		PRJDB10510	SAMD00245249	GCA_014646855		strain page
50913	type strain	64	60	<i>Pseudomonas viciae</i>	11K1	Zhao et al. 2020	GDMMC 1,1743; KACC 21650	<i>Pseudomonas viciae</i>	6,704,877	60.34	5778		PRJNA514417	SAMN10724295	GCA_004786035		species page
5120	type strain	65	61	<i>Pseudomonas poae</i>	LMG 21465	Behrendt et al. 2003	DSM 14936; P 5271/3	<i>Pseudomonas poae</i>	5,996,363	60.41	5301	Gp0375626	PRJNA338673	SAMN05560560	GCA_001730605		strain page
5121	type strain	66	62	<i>Pseudomonas trivialis</i>	LMG 21464	Behrendt et al. 2003	DSM 14937; P 513/19	<i>Pseudomonas trivialis</i>	5,771,272	60.93	5053	Gp0375592	PRJNA338684	SAMN05560640	GCA_001730655		strain page
5123	type strain	67	63	<i>Pseudomonas simiae</i>	COUG 50988	Vela et al. 2006	CECT 7078; DSM 18881; strain OLI	<i>Pseudomonas simiae</i>	6,289,498	60.3	5710	Gp0375583	PRJNA338691	SAMN05560726	GCA_001730615		strain page
5483	type strain	68	64	<i>Pseudomonas granadensis</i>	LMG 27940	Pascual et al. 2015	DSM 28040; F-278; F-278, 770	<i>Pseudomonas granadensis</i>	5,943,070	60.16	5191	Gp0127197	PRJEB16372	SAMN05216579	GCA_900105485		strain page
5568	type strain	69	65	<i>Pseudomonas mandelii</i>	LMG 21607	Verhille et al. 1999	COUG 42058; DSM 17967; JCM 21619; ATCC 700871; NBRC 103147; CFML 95-303; CIP 105273	<i>Pseudomonas mandelii</i>	7,041,658	59.2	6279	Gp0120260	PRJEB16502	SAMN04489801	GCA_900106065		strain page
5597	type strain	70	66	<i>Pseudomonas glaucocens</i>	LMG 27394	Toro et al. 2013	CECT 9262; DSM 28410; PCAUV11	<i>Pseudomonas glaucocens</i>	5,076,874	62.6	4618	Gp0127199	PRJEB15687	SAMN05216185	GCA_900102675		strain page
5671	type strain	71	67	<i>Pseudomonas veronii</i>	DSM 11331	Elomari et al. 1996	LMG 17761; COUG 43519; JCM 11942; ATCC 700272; ATCC 700474; CFML 92-134; CIP 104663	<i>Pseudomonas veronii</i> <i>Pseudomonas veronii</i> subsp. <i>veronii</i>	6,997,867	60.72	6228	Gp0133281	PRJNA224116	SAMN03328800	GCF_001439695		strain page
5672	type strain	72	68	<i>Pseudomonas orientalis</i>	DSM 17489	Dabboussi et al. 2002	CFBP 4863; CFML 96-170; CIP 105540	<i>Pseudomonas orientalis</i>	6,186,637	60.51	5471	Gp0133518	PRJNA224116	SAMN03328802	GCF_001439815		strain page
5682	type strain	73	69	<i>Pseudomonas kilonensis</i>	DSM 13647	Sikorski et al. 2001	CCUG 12536; DSM 18862; JCM 2777; IFO 12693; NBRC 12693; CIP 106744; IAM 1603	<i>Pseudomonas kilonensis</i>	6,385,809	60.86	5584	Gp0145008	PRJNA280438	SAMN03891674	GCA_001269885		strain page
5761	type strain	74	70	<i>Pseudomonas azobiformans</i>	LMG 21611	Iizuka and Komagata 1963	CFBP 2341; NCPPB 2445; LMG 2172; ICMP 5819; CCUG 23367; CCUG 32778; DSM 7228; ATCC 29736; CIP 105514	<i>Pseudomonas azobiformans</i>	6,727,286	60.98	5943	Gp0120257	PRJEB16206	SAMN04489799	GCA_900103345		strain page
5772	type strain	75	71	<i>Pseudomonas cannabina</i>	ICMP 2823	(ex Sutilis and Dowson 1959) Gardan et al. 1999 emend. Bull et al. 2010	CFBP 2341; NCPPB 2445; LMG 2172; ICMP 5819; CCUG 23367; CCUG 32778; DSM 7228; ATCC 29736; CIP 105514	<i>Pseudomonas cannabina</i>	6,386,314	58.46	5635	Gp0127163	PRJEB15982	SAMN05216597	GCA_900100365		strain page
5888	type strain	76	73	<i>Pseudomonas corrugata</i>	NCPPB 2445	Roberts and Scarlett 1981 emend. Saita et al. 1997	CFBP 2431; NCPPB 2445; LMG 2172; ICMP 5819; CCUG 23367; CCUG 32778; DSM 7228; ATCC 29736; CIP 105514	<i>Pseudomonas corrugata</i>	6,080,326	60.63	5314		PRJNA224116	SAMN03879927	GCF_001411965		strain page
5890	type strain	76	73	<i>Pseudomonas corrugata</i>	DSM 7228	Roberts and Scarlett 1981 emend. Saita et al. 1997	CFBP 2431; NCPPB 2445; LMG 2172; ICMP 5819; CCUG 23367; CCUG 32778; DSM 7228; ATCC 29736; CIP 105514	<i>Pseudomonas corrugata</i>	6,126,718	60.62	5310	Gp0145018	PRJNA280438	SAMN03891677	GCA_001269905		strain page
5889	type strain	77	72	<i>Pseudomonas chlororaphis</i>	LMG 5004	(Guignard and Sauvageau 1894) Bergqvist et al. 1930 emend. Pek et al. 2007	CFBP 2132; CCUG 552; CCUG 552 B; DSM 50083; JCM 2778; ATCC 9446; IFO 3904; NBRC 3904; NCTC 13002; VKM B-1246; CIP 63.22; HAMBI 2011; NCCB 76041; NCIMB 9392	<i>Bacillus chlororaphis</i> ; <i>Pseudomonas chlororaphis</i> ; <i>Pseudomonas chlororaphis</i> subsp. <i>chlororaphis</i>	6,792,468	63.01	6059	Gp0145021	PRJNA280438	SAMN03891669	GCA_001269625		species page

Table S3 (continued)

TVGS ID	Kind	Species cluster	Subspecies cluster	Preferred name	Deposit	Authority	Other deposits	Synonymous taxon names	Base pairs	Percent G+C	No. proteins	Goldstamp	Bioproject accession	Biosample accession	Assembly accession	IMG OID	BacDive
5904	Type strain	78	75	<i>Pseudomonas protekii</i> LMG 26867	LMG 26867	Kosina et al. 2014	CCM 7990; strain AN/28/1 CFBP 5592; LMG 2336; NRRL B-14; NRRL B-731; CCUG 560; DSM 19125; JCM 20238; ATCC 4683; NCTC 10697; VKM B-904; CIP 103298; IAM 1653	<i>Pseudomonas protekii</i>	6,091,955	59.56	5360	Gp0127210	PRJEB16427	SAMN05216222	GCA_900105155		species page
5941	Type strain	79	76	<i>Pseudomonas taetrolens</i> DSM 21104	DSM 21104	Haynes 1957		<i>Pseudomonas taetrolens</i>	4,921,904	58.26	4472	Gp0133291	PRJNA224116	SAMN03328798	GCF_001042915		strain page
5942	Type strain	80	77	<i>Pseudomonas lini</i> DSM 16768	DSM 16768	Delorme et al. 2002	CFBP 5737; ICMP 14138; CCUG 51522; DLE11J	<i>Pseudomonas lini</i>	6,495,504	58.82	5812	Gp0132912	PRJNA224116	SAMN03328801	GCF_001042905		strain page
5946	Type strain	81	78	<i>Pseudomonas chlororaphis subsp. piscium</i> DSM 21509	DSM 21509	Burr et al. 2010	JF3835; NCIMB 14478	<i>Pseudomonas chlororaphis subsp. piscium</i>	7,044,604	62.68	6252	Gp0386345	PRJNA290438	SAMN03891664	GCA_001269555		subspecies page
5995	Type strain	82	79	<i>Pseudomonas caricapapayae</i> ICMP 2855	ICMP 2855	Robbs 1956	CFBP 3204; NCPPB 1873; LMG 2152; CCUG 32775; DSM 21109; ATCC 33615; CIP 106736	<i>Pseudomonas caricapapayae</i>	6,257,512	58.32	5542	Gp0146395	PRJNA292453	SAMN03976590	GCA_001400735		strain page
6057	Type strain	83	80	<i>Pseudomonas congelans</i> DSM 14939	DSM 14939	Behrendt et al. 2003	LMG 21466; P 538/23	<i>Pseudomonas congelans</i>	5,731,984	59.33	4980	Gp0127158	PRJEB16279	SAMN05216596	GCA_900103225		strain page
6632	Type strain	84	81	<i>Pseudomonas vancouverensis</i> LMG 20222	LMG 20222	Mohn et al. 1999	CCUG 49675; DSM 17555; ATCC 700686; CIP 106707; DHA-51	<i>Pseudomonas vancouverensis</i>	6,424,805	60.11	5822	Gp0116982	PRJNA224116	SAMN05216558	GCF_900105825		strain page
6949	Type strain	85	82	<i>Acinetobacter halodolerans</i> JCM 31009	JCM 31009	Dahal et al. 2017	KACC 18453; KEMB 9005-333; R160	<i>Acinetobacter halodolerans</i>	3,284,402	39.97	3038	Gp0039157	PRJNA387062	SAMN10908556	GCA_001484935		species page
7210	Type strain	86	83	<i>Pseudomonas izuensis</i> IZPS43_300_3	IZPS43_300_3	Lu et al. 2020	LMG 31527; CECT 9963	<i>Pseudomonas izuensis</i>	6,857,708	59.62	6149		PRJNA594796	SAMN13530344	GCA_009861505		species page
739	Type strain	87	84	<i>Pseudomonas entomophila</i> L48	L48	Mulet et al. 2012	CCUG 61470; CECT 7985; DSM 28517	<i>Pseudomonas entomophila</i>	5,888,780	64.16	5134	Gp0000348	PRJNA16800	SAMEA3138225	GCA_000026105		strain page
7903	Type strain	88	85	<i>Pseudomonas saponiphila</i> DSM 9751	DSM 9751	Lang et al. 2012	CIP 109856	<i>Pseudomonas saponiphila</i>	7,375,852	62.79	6831	Gp0127188	PRJEB16433	SAMN05216778	GCA_900105185		strain page
8269	Type strain	89	86	<i>Pseudomonas mohinii</i> DSM 18327	DSM 18327	Cámara et al. 2007	CCUG 53115; lpa-2	<i>Pseudomonas mohinii</i>	6,592,588	59.62	5947	Gp0127185	PRJEB16418	SAMN05216205	GCA_900105115		strain page
8573	Type strain	90	87	<i>Pseudomonas soli</i> LMG 27941	LMG 27941	Pascual et al. 2015	DSM 28043; F-279; F-279.208	<i>Pseudomonas soli</i>	5,640,705	64	5104	Gp0127213	PRJEB16975	SAMN05216230	GCA_900110655		strain page
8943	Type strain	91	88	<i>Pseudomonas fluorescens batumici</i> DSM 10987T	DSM 10987T	Hofmann et al. 2020	LMG 31232; WS 5063	<i>Pseudomonas fluorescens batumici</i>	6,007,255	60.03	5325		PRJNA224116	SAMN12399692	GCF_009659625		species page
9391	Type strain	92	89	<i>Pseudomonas batumici</i> UCM B-321	UCM B-321	Kiprianova et al. 2011		<i>Pseudomonas batumici</i>	6,561,844	61.8	5831	Gp0118174	PRJNA270768	SAMN03273282	GCA_000820515		species page
U70605	user strain	93	90	<i>Pseudomonas fluorescens F113</i>				<i>Pseudomonas fluorescens F113</i>	6,845,832	60.78	6045						
U70606	user strain	94	91	<i>Pseudomonas fluorescens Pf0-1</i>				<i>Pseudomonas fluorescens Pf0-1</i>	6,438,405	60.52	5739						
U70612	user strain	95	94	<i>Pseudomonas JY35948</i>				<i>Pseudomonas JY35948</i>	6,337,157	60.66	5717						
U70613	user strain	96	95	<i>Pseudomonas JY391D10</i>				<i>Pseudomonas JY391D10</i>	7,165,607	60.54	6462						
U70614	user strain	96	95	<i>Pseudomonas JY391D17</i>				<i>Pseudomonas JY391D17</i>	6,821,048	60.68	6192						
U70620	user strain	97	98	<i>Pseudomonas putida GB-1</i>				<i>Pseudomonas putida GB-1</i>	6,078,430	61.94	5445						
U70619	user strain	98	97	<i>Pseudomonas putida F1</i>				<i>Pseudomonas putida F1</i>	5,955,964	61.86	5294						
U70621	user strain	98	99	<i>Pseudomonas putida KT2440</i>				<i>Pseudomonas putida KT2440</i>	6,181,873	61.52	5584						
U70622	user strain	99	100	<i>Pseudomonas sp. JY449</i>				<i>Pseudomonas sp. JY449</i>	6,832,692	60.92	6134						

**Table S4: Biosynthesis gene clusters for secondary metabolites predicted by AntiSMASH for each of the 10 *Pseudomonas* strains.**

Replicon name	Replicon type	Begin	End	Length	Region type	MIBiG	Completion	Product	Type
JV251A.1	WGS	638591	653430	14840	NAGGN	-	-	-	-
JV251A.1	WGS	784549	837547	52999	NRPS	BGC0000424	0.231	Serobactins	NRPS
JV251A.1	WGS	2154486	2219744	65259	NRPS	BGC0001312	0.312	Viscosin	NRPS
JV251A.1	WGS	2445983	2469262	23280	betalactone	BGC0001103	0.2	Mycosubtilin	NRPS-transatpks
JV251A.1	WGS	4842567	4853409	10843	bacteriocin	BGC0001098	0.222	Iturin	NRPS-transatpks
JV251A.1	WGS	5183510	5231082	47573	NRPS, T1PKS	-	-	-	-
JV251A.1	WGS	5715625	5759230	43606	arylpolyyene	BGC0000837	0.4	APE Vf	other
JV251A.1	WGS	5982085	6024922	42838	NRPS-like	-	-	-	-
JV251A.1	WGS	6436933	6455003	18071	NRPS-like	BGC0000387	0.714	Mangotoxin	NRPS
Replicon name	Replicon type	Begin	End	Length	Region type	MIBiG	Completion	Product	Type
JV551A1.1	WGS	1496495	1511345	14851	NAGGN	-	-	-	-
JV551A1.1	WGS	2148499	2159347	10849	bacteriocin	-	-	-	-
JV551A1.1	WGS	3146541	3166957	20417	phenazine	-	-	-	-
JV551A1.1	WGS	4148893	4201116	52224	NRPS	BGC0000424	0.231	Serobactins	NRPS
JV551A1.1	WGS	4476928	4551027	74100	NRPS	BGC0001347	1*	Poaemide B	NRPS
Replicon name	Replicon type	Begin	End	Length	Region type	MIBiG	Completion	Product	Type
JV359A8.1	WGS	547337	558233	10897	bacteriocin	-	-	-	-
JV359A8.1	WGS	1797471	1829995	32525	NRPS-like	BGC0000387	0.714	Mangotoxin	NRPS
JV359A8.1	WGS	2109639	2153253	43615	arylpolyyene	BGC0000837	0.4	APE Vf	other
JV359A8.1	WGS	2472624	2534836	62213	NRPS	BGC0001132	1	Xenotetrapeptide	NRPS
JV359A8.1	WGS	2566749	2578602	11854	siderophore	-	-	-	-
JV359A8.1	WGS	3326820	3341593	14774	NAGGN	-	-	-	-
JV359A8.1	WGS	3480211	3533212	53002	NRPS	BGC0000457	0.222	Vicibactin	NRPS
JV359A8.1	WGS	4015412	4078895	63484	NRPS	BGC0000438	1*	Syringopeptin	NRPS
JV359A8.1	WGS	4194238	4205077	10840	bacteriocin	-	-	-	-
JV359A8.1	WGS	5612727	5636020	23294	betalactone	BGC0001103	0.2	Mycosubtilin	NRPS-transatpks
JV359A8.1	WGS	5873533	5934757	61225	NRPS	BGC0001312	0.312	Viscosin	NRPS
Replicon name	Replicon type	Begin	End	Length	Region type	MIBiG	Completion	Product	Type
JV391D10.1	WGS	103316	134828	31513	thiopeptide	BGC0000774	0.057	Lipopolysaccharide	saccharide
JV391D10.1	WGS	824973	895326	70354	NRPS	BGC0000984	0.286	Delftibactin	NRPS-t1pks
JV391D10.1	WGS	914111	929286	15176	NAGGN	-	-	-	-
JV391D10.1	WGS	1142981	1186502	43522	NRPS-like	-	-	-	-
JV391D10.1	WGS	2147603	2158448	10846	bacteriocin	-	-	-	-
JV391D10.1	WGS	2591403	2614511	23109	betalactone	BGC0001103	0.2	Mycosubtilin	NRPS-transatpks
JV391D10.1	WGS	2802518	2868471	65954	NRPS	BGC0001312	0.562	Viscosin	NRPS
JV391D10.1	WGS	4222185	4265760	43576	arylpolyyene	BGC0000837	0.45	APE Vf	other
JV391D10.1	WGS	4453714	4494925	41212	ladderane, hserlactone	BGC0001047	0.286	Syringolin A	NRPS-t1pks
JV391D10.1	WGS	5277008	5288933	11926	siderophore	-	-	-	-
JV391D10.1	WGS	5561216	5572094	10879	bacteriocin	-	-	-	-
JV391D10.1	WGS	5815223	5842297	27075	NRPS-like	BGC0000287	1*	2-amino-4-methoxy-trans-3-butenic acid	NRPS
JV391D10.1	WGS	6502266	6522841	20576	hserlactone	-	-	-	-
JV391D10.1	WGS	7159979	7165607	5629	NRPS	BGC0001758	1	Rhizomide A / rhizomide B / rhizomide C	NRPS
Replicon name	Replicon type	Begin	End	Length	Region type	MIBiG	Completion	Product	Type
JV449.1	WGS	152733	162037	9305	bacteriocin	-	-	-	-
JV449.1	WGS	257824	298559	40736	NRPS-like	BGC0000387	0.714	Mangotoxin	NRPS
JV449.1	WGS	613676	657251	43576	arylpolyyene	BGC0000837	0.45	APE Vf	other
JV449.1	WGS	1718244	1729119	10876	bacteriocin	-	-	-	-
JV449.1	WGS	2602275	2677290	75016	NRPS	BGC0001758	1	omide A / rhizomide B / rhizomide C	NRPS
JV449.1	WGS	3411671	3423521	11851	siderophore	-	-	-	-
JV449.1	WGS	4167237	4190316	23080	betalactone	BGC0001103	0.2	Mycosubtilin	NRPS-transatpks
JV449.1	WGS	4762823	4777630	14808	NAGGN	-	-	-	-
JV449.1	WGS	4811835	4864731	52897	NRPS	BGC0001189	0.167	Taiwachelin	NRPS
Replicon name	Replicon type	Begin	End	Length	Region type	MIBiG	Completion	Product	Type
JV391D17.1	WGS	492348	535869	43522	NRPS-like	-	-	-	-
JV391D17.1	WGS	743098	758072	14975	NAGGN	-	-	-	-
JV391D17.1	WGS	1243617	1254495	10879	bacteriocin	-	-	-	-
JV391D17.1	WGS	1610766	1622691	11926	siderophore	-	-	-	-
JV391D17.1	WGS	1710708	1763592	52885	NRPS	BGC0000984	0.286	Delftibactin	NRPS-t1pks
JV391D17.1	WGS	2160387	2171232	10846	bacteriocin	-	-	-	-
JV391D17.1	WGS	3502436	3546011	43576	arylpolyyene	BGC0000837	0.45	APE Vf	other
JV391D17.1	WGS	4009380	4050591	41212	ladderane, hserlactone	BGC0001047	0.286	Syringolin A	NRPS-t1pks
JV391D17.1	WGS	4080400	4111574	31175	NRPS	BGC0000413	0.046	Pyoverdine	NRPS
JV391D17.1	WGS	4289952	4313055	23104	betalactone	BGC0001103	0.2	Mycosubtilin	NRPS-transatpks
JV391D17.1	WGS	5458289	5486745	28457	NRPS	BGC0001758	1	Rhizomide A / rhizomide B / rhizomide C	NRPS
JV391D17.1	WGS	5696434	5717009	20576	hserlactone	-	-	-	-
JV391D17.1	WGS	6787667	6814975	27309	NRPS-like	BGC0000387	0.571	Mangotoxin	NRPS
JV391D17.1	WGS	6845691	6874281	28591	thiopeptide	BGC0000774	0.057	Lipopolysaccharide	saccharide
JV391D17.1	WGS	6906823	6915721	8899	NRPS	BGC0000413	0.064	Pyoverdine	NRPS
JV391D17.1	WGS	6915822	6921048	5227	NRPS	-	-	-	-

Table S4 (continued)

Replicon name	Replicon type	Begin	End	Length	Region type	MIBiG	Completion	Product	Type
JV244B.1	WGS	964803	1006002	41200	T3PKS	-	-	-	-
JV244B.1	WGS	1054611	1095779	41169	NRPS	BGC0001189	0.111	Taiwachelin	NRPS
JV244B.1	WGS	1268321	1297554	29234	betalactone	-	-	-	-
JV244B.1	WGS	1306956	1349458	42503	NRPS-like	BGC0000095	0.4	Micromonolactam	t1pks
JV244B.1	WGS	1538726	1585032	46307	NRPS	BGC0001754	0.4	Ralsolamycin	NRPS-t1pks
JV244B.1	WGS	2727351	2742086	14736	NAGGN	-	-	-	-
JV244B.1	WGS	3248624	3301285	52662	NRPS	BGC0000412	1	Pyochelin	NRPS
JV244B.1	WGS	3428153	3471766	43614	arylpolyene	BGC0000837	0.4	APE Vf	other
JV244B.1	WGS	4417786	4435562	17777	PpyS-KS	-	-	-	-
JV244B.1	WGS	4609604	4640158	30555	LAP	BGC0000413	0.028	Pyoverdine	NRPS
JV244B.1	WGS	4752641	4776508	23868	NRPS	BGC0001312	0.188	Viscosin	NRPS
JV244B.1	WGS	6096628	6117194	20567	hserlactone	BGC0001437	0.143	Obafuorin	NRPS
JV244B.1	WGS	6233943	6266829	32887	other	BGC0000924	1	Pyrrolnitrin	other
JV244B.1	WGS	6639180	6660335	21156	NRPS-like	BGC0000387	0.714	Mangotoxin	NRPS
JV244B.1	WGS	6915641	6931108	15468	NRPS-like	BGC0000413	0.11	Pyoverdine	NRPS
JV244B.1	WGS	7114947	7124671	9725	NRPS	BGC0001347	1*	Poaamide B	NRPS
JV244B.1	WGS	7196661	7204628	7968	NRPS	BGC0000413	0.046	Pyoverdine	NRPS
JV244B.1	WGS	7295468	7300692	5225	NRPS	BGC0001608	0.375	Gldopeptin	NRPS
JV244B.1	WGS	7346529	7349557	3029	NRPS	-	-	-	-
JV244B.1	WGS	7382834	7385100	2267	NRPS	-	-	-	-
JV244B.1	WGS	7385201	7387331	2131	NRPS	-	-	-	-
JV244B.1	WGS	7387432	7389556	2125	NRPS	-	-	-	-
JV244B.1	WGS	7403197	7404746	1550	NRPS	-	-	-	-

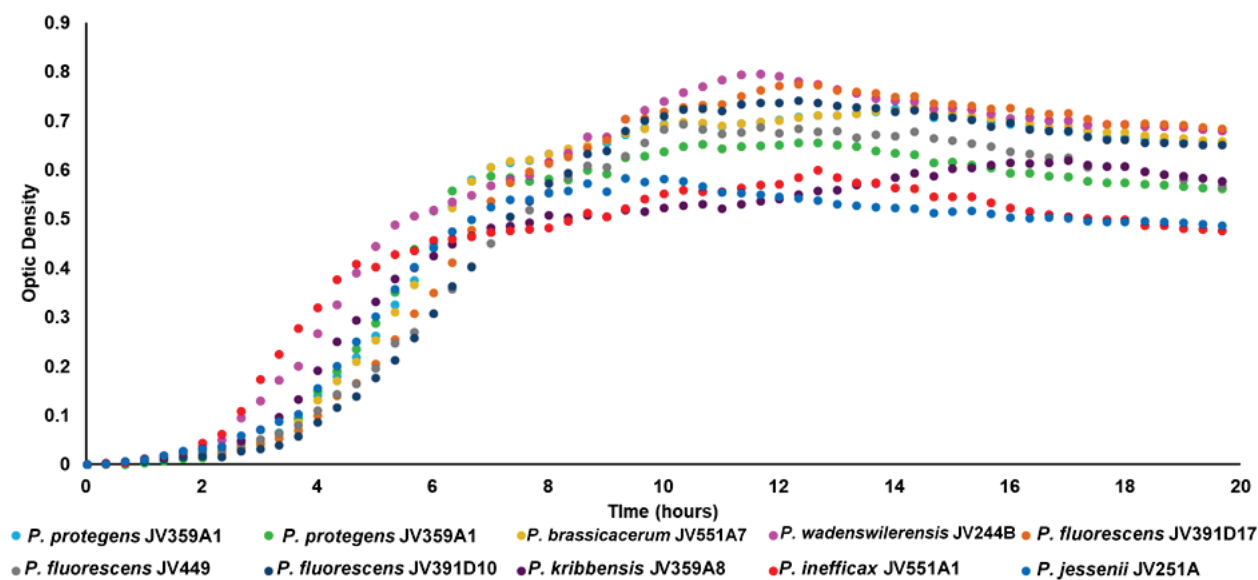
Replicon name	Replicon type	Begin	End	Length	Region type	MIBiG	Completion	Product	Type
JV551A7.1	WGS	30642	73213	42572	NRPS-like	BGC0000387	0.714	Mangotoxin	NRPS
JV551A7.1	WGS	345716	389327	43612	arylpolyene	BGC0000837	0.4	APE Vf	other
JV551A7.1	WGS	1273113	1283982	10870	bacteriocin	-	-	-	-
JV551A7.1	WGS	1686984	1701897	14914	NAGGN	-	-	-	-
JV551A7.1	WGS	1932327	1984308	51982	NRPS	BGC0000424	0.154	Serobactins	NRPS
JV551A7.1	WGS	2178819	2202544	23726	betalactone	BGC0001103	0.2	Mycosubtilin	NRPS-transatpks
JV551A7.1	WGS	2722845	2763894	41050	T3PKS	BGC0000281	1	2,4-Diacetylphloroglucinol	t3pks
JV551A7.1	WGS	2810543	2875593	65051	NRPS	BGC0000330	0.176	Cupriachelin	NRPS
JV551A7.1	WGS	2927451	2940879	13429	butyrolactone	-	-	-	-
JV551A7.1	WGS	4663877	4741480	77604	NRPS	BGC0000984	0.571	Delftibactin	NRPS-t1pks

Replicon name	Replicon type	Begin	End	Length	Region type	MIBiG	Completion	Product	Type
JV245A.1	WGS	159636	230608	70973	NRPS	BGC0000438	1*	Syringopeptin	NRPS
JV245A.1	WGS	346854	356034	9181	bacteriocin	-	-	-	-
JV245A.1	WGS	536840	557589	20750	CDPS	-	-	-	-
JV245A.1	WGS	631770	646602	14833	NAGGN	-	-	-	-
JV245A.1	WGS	862436	903485	41050	T3PKS	BGC0000281	1*	2,4-Diacetylphloroglucinol	t3pks
JV245A.1	WGS	1143378	1174464	31087	NRPS	BGC0001379	0.071	BD-12	NRPS
JV245A.1	WGS	2131076	2174257	43182	NRPS	BGC0001128	1	Luminmide	NRPS
JV245A.1	WGS	2458897	2511646	52750	T1PKS	BGC0000127	1*	Pyoluteorin	t1pks
JV245A.1	WGS	2974707	2985552	10846	bacteriocin	-	-	-	-
JV245A.1	WGS	3474324	3514623	40300	arylpolyene	-	-	-	-
JV245A.1	WGS	3761976	3774809	12834	hserlactone	BGC0001437	0.143	Obafuorin	NRPS
JV245A.1	WGS	4138227	4183853	45627	NRPS	BGC0000324	0.364	Coelibactin	NRPS
JV245A.1	WGS	4477299	4527462	50164	NRPS	BGC0001347	1*	Poaamide B	NRPS
JV245A.1	WGS	5440206	5463420	23215	betalactone	BGC0001103	0.3	Mycosubtilin	NRPS-transatpks
JV245A.1	WGS	5575343	5618960	43618	arylpolyene	BGC0000837	0.4	APE Vf	other
JV245A.1	WGS	6364337	6383769	19433	other	BGC0000924	1*	Pyrrolnitrin	other

Replicon name	Replicon type	Begin	End	Length	Region type	MIBiG	Completion	Product	Type
JV359A1.1	WGS	615682	626572	10891	bacteriocin	-	-	-	-
JV359A1.1	WGS	716608	737357	20750	CDPS	-	-	-	-
JV359A1.1	WGS	1518651	1612159	93509	NRPS	BGC0001347	1*	Poaamide B	NRPS
JV359A1.1	WGS	2253213	2305962	52750	T1PKS	BGC0000127	1	Pyoluteorin	t1pks
JV359A1.1	WGS	3042087	3116308	74222	NRPS	BGC0001801	0.267	Thiazostatin / watasemycin A / watasemycin B / 2-hydroxyphenylthiazoline enantiopyochelin / isopyochelin	NRPS
JV359A1.1	WGS	3206175	3247260	41086	other	BGC0000924	1	Pyrrolnitrin	other
JV359A1.1	WGS	3472780	3523775	50996	NRPS	BGC0000413	0.183	Pyoverdine	NRPS
JV359A1.1	WGS	3670932	3685764	14833	NAGGN	-	-	-	-
JV359A1.1	WGS	3981590	4032619	51030	NRPS	BGC0001379	0.071	BD-12	NRPS
JV359A1.1	WGS	4876297	4917346	41050	T3PKS	BGC0000281	1	2,4-Diacetylphloroglucinol	t3pks
JV359A1.1	WGS	5147877	5158722	10846	bacteriocin	-	-	-	-
JV359A1.1	WGS	5623979	5667596	43618	arylpolyene	BGC0000837	0.4	APE Vf	other
JV359A1.1	WGS	6055271	6086020	30750	NRPS	BGC0001312	0.312	Viscosin	NRPS
JV359A1.1	WGS	6231024	6254238	23215	betalactone	BGC0001103	0.2	Mycosubtilin	NRPS-transatpks
JV359A1.1	WGS	6755332	6795204	39873	NRPS	BGC0001189	1*	Taiwachelin	NRPS



**Figure S1: Growth curves of each *Pseudomonas* strain on a maize-based medium using bioscreen (Lab systems Helsinki, Finland).**



**Table S5: Number of *rpoD* reads obtained for the initial SynCom.** The same initial SynCom was used for both *in vitro* and *in planta* EE and its composition was assessed 4 times in total. For each EE, the initial SynCom underwent 2 parallel DNA extractions, and independent *rpoD* sequencing analyses were carried for each extraction.

STRAINS	number of <i>rpoD</i> reads			
	EE <i>in vitro</i> (A)	EE <i>in vitro</i> (B)	EE <i>in planta</i> (A)	EE <i>in planta</i> (B)
<i>P. jessenii</i> JV251A	3995	3986	7901	6209
<i>P. inefficax</i> JV551A1	4693	4471	4608	3366
<i>P. kribbensis</i> JV359A8	755	775	1973	1457
<i>P. fluorescens</i> JV391D10	6640	6681	8515	6562
<i>P. fluorescens</i> JV449	2115	1976	2565	2038
<i>P. fluorescens</i> JV391D17	4335	4390	6392	5039
<i>P. wadenswilerensis</i> JV244B	2517	2498	5364	3740
<i>P. brassicacerum</i> JV551A7	6386	6429	9447	7150
<i>P. protegens</i> JV245A	2930	3030	2058	1553
<i>P. protegens</i> JV359A1	558	651	1041	797
<b>TOTAL READ COUNTS</b>	<b>34924</b>	<b>34887</b>	<b>49864</b>	<b>37911</b>

**Table S6: Relative abundances (expressed in percent) of each of the 10 *Pseudomonas* strains in the initial SynCom.** The same initial SynCom was used for both *in vitro* and *in planta* EE and its composition was assessed 4 times in total. For each EE, the initial SynCom underwent 2 parallel DNA extractions, and independent *rpoD* sequencing analyses were carried for each extraction. The relative abundance represents the number of *rpoD* sequences assigned to one strain divided by the total number of *rpoD* sequences in each sample.

STRAINS	relative abundance (%) of <i>Pseudomonas</i> strains					
	EE <i>in vitro</i> (A)	EE <i>in vitro</i> (B)	EE <i>in planta</i> (A)	EE <i>in planta</i> (B)	MEAN	STANDARD ERROR
<i>P. jessenii</i> JV251A	11.44	11.43	15.85	16.38	13.77	1.36
<i>P. inefficax</i> JV551A1	13.44	12.82	9.24	8.88	11.09	1.18
<i>P. kribbensis</i> JV359A8	2.16	2.22	3.96	3.84	3.05	0.49
<i>P. fluorescens</i> JV391D10	19.01	19.15	17.08	17.31	18.14	0.55
<i>P. fluorescens</i> JV449	6.06	5.66	5.14	5.38	5.56	0.20
<i>P. fluorescens</i> JV391D17	12.41	12.58	12.82	13.29	12.78	0.19
<i>P. wadenswilerensis</i> JV244B	7.21	7.16	10.76	9.87	8.75	0.92
<i>P. brassicacerum</i> JV551A7	18.29	18.43	18.95	18.86	18.63	0.16
<i>P. protegens</i> JV245A	8.39	8.69	4.13	4.10	6.32	1.28
<i>P. protegens</i> JV359A1	1.60	1.87	2.09	2.10	1.91	0.12
<b>TOTAL</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>-</b>

**Table S7: Number of *rpoD* reads assigned to each *Pseudomonas* population, obtained for each SynCom 1 to 5 for the *in vitro* EE, at (A) cycle 1, (B) cycle 5, (C) cycle 15 and (D) cycle 20.**

STRAINS	CYCLE 1 - number of <i>rpoD</i> reads			
	SynCom 1-1	SynCom 2-1	SynCom 3-1	SynCom 4-1
<i>P. jessenii</i> JV251A	122	124	110	116
<i>P. inefficax</i> JV551A1	28347	28332	27656	27162
<i>P. kribbensis</i> JV359A8	38	37	41	29
<i>P. fluorescens</i> JV391D10	290	281	339	361
<i>P. fluorescens</i> JV449	73	60	87	70
<i>P. fluorescens</i> JV391D17	135	153	153	138
<i>P. wadenswilerensis</i> JV244B	89	80	109	84
<i>P. brassicacerum</i> JV551A7	207	208	265	192
<i>P. protegens</i> JV245A	5667	5687	6192	6806
<i>P. protegens</i> JV359A1	19	22	28	29
TOTAL READ COUNTS	34987	34984	34980	34987

STRAINS	CYCLE 5 - number of <i>rpoD</i> reads				
	SynCom 5-1	SynCom 2-5	SynCom 3-5	SynCom 4-5	SynCom 5-5
<i>P. jessenii</i> JV251A	141	66	79	76	70
<i>P. inefficax</i> JV551A1	27275	33235	32724	32964	32373
<i>P. kribbensis</i> JV359A8	36	23	13	21	14
<i>P. fluorescens</i> JV391D10	448	89	107	110	118
<i>P. fluorescens</i> JV449	76	23	30	24	30
<i>P. fluorescens</i> JV391D17	152	58	65	64	53
<i>P. wadenswilerensis</i> JV244B	73	52	61	48	52
<i>P. brassicacerum</i> JV551A7	218	80	71	63	77
<i>P. protegens</i> JV245A	6526	1350	1833	1604	2187
<i>P. protegens</i> JV359A1	40	10	9	11	6
TOTAL READ COUNTS	34985	34986	34992	34985	34980

STRAINS	CYCLE 15 - number of <i>rpoD</i> reads			
	SynCom 2-15	SynCom 3-15	SynCom 4-15	SynCom 5-15
<i>P. jessenii</i> JV251A	50	71	57	40
<i>P. inefficax</i> JV551A1	34332	34237	34265	34279
<i>P. kribbensis</i> JV359A8	9	16	12	12
<i>P. fluorescens</i> JV391D10	75	82	79	87
<i>P. fluorescens</i> JV449	15	26	21	24
<i>P. fluorescens</i> JV391D17	44	53	39	58
<i>P. wadenswilerensis</i> JV244B	42	54	52	42
<i>P. brassicacerum</i> JV551A7	64	62	73	66
<i>P. protegens</i> JV245A	350	374	389	378
<i>P. protegens</i> JV359A1	6	8	4	8
TOTAL READ COUNTS	34987	34983	34991	34994

STRAINS	CYCLE 20 - number of <i>rpoD</i> reads				
	SynCom 1-20	SynCom 2-20	SynCom 3-20	SynCom 4-20	SynCom 5-20
<i>P. jessenii</i> JV251A	56	63	50	49	52
<i>P. inefficax</i> JV551A1	34294	34353	34332	34386	34303
<i>P. kribbensis</i> JV359A8	12	9	6	10	14
<i>P. fluorescens</i> JV391D10	83	65	58	61	69
<i>P. fluorescens</i> JV449	19	25	20	19	25
<i>P. fluorescens</i> JV391D17	50	48	45	42	40
<i>P. wadenswilerensis</i> JV244B	55	39	37	39	44
<i>P. brassicacerum</i> JV551A7	64	58	71	61	72
<i>P. protegens</i> JV245A	343	324	366	314	366
<i>P. protegens</i> JV359A1	8	3	4	5	6
TOTAL READ COUNTS	34984	34987	34989	34986	34991

**Table S8: Relative abundances (expressed in percent) of each of the 10 *Pseudomonas* strains obtained for the *in vitro* EE at the end of the cycles 1, 5, 15 and 20.** The relative abundance represent the number of *rpoD* sequences assigned to one strain divided by the total number of *rpoD* sequences in each sample.

STRAINS	CYCLE 1		CYCLE 5		CYCLE 15		CYCLE 20	
	MEAN	STANDARD ERROR	MEAN	STANDARD ERROR	MEAN	STANDARD ERROR	MEAN	STANDARD ERROR
<i>P. jessenii</i> JV251A	0.35	0.01	0.21	0.01	0.16	0.02	0.15	0.01
<i>P. inefficax</i> JV551A1	79.33	1.61	93.82	0.52	97.97	0.06	98.13	0.05
<i>P. kribbensis</i> JV359A8	0.10	0.01	0.05	0.01	0.04	0.00	0.03	0.00
<i>P. fluorescens</i> JV391D10	0.98	0.19	0.30	0.02	0.23	0.01	0.19	0.01
<i>P. fluorescens</i> JV449	0.21	0.03	0.08	0.01	0.06	0.01	0.06	0.00
<i>P. fluorescens</i> JV391D17	0.42	0.03	0.17	0.01	0.14	0.01	0.13	0.01
<i>P. wadsworthii</i> JV244B	0.25	0.04	0.15	0.01	0.14	0.01	0.12	0.01
<i>P. brassicacearum</i> JV551A7	0.62	0.08	0.21	0.01	0.19	0.01	0.19	0.01
<i>P. protegens</i> JV245A	17.65	1.44	4.98	0.51	1.07	0.02	0.98	0.03
<i>P. protegens</i> JV359A1	0.08	0.02	0.03	0.00	0.02	0.00	0.01	0.00
TOTAL	100.00	-	100	-	100	-	100	-

**Table S9: Number of *rpoD* reads assigned to each *Pseudomonas* population obtained for (A) SynCom 1, 3 and 6, and (B) SynCom 2, 4, 5 and 7 at different cycles of the *in planta* EE.**

		number of <i>rpoD</i> reads										A									
		EE in <i>in planta</i> SynCom 1					EE in <i>in planta</i> SynCom 3					EE in <i>in planta</i> SynCom 6									
STRAINS		SynCom 1-1	SynCom 1-2	SynCom 1-3	SynCom 1-4	SynCom 1-5	SynCom 1-6	SynCom 1-7	SynCom 3-1	SynCom 3-2	SynCom 3-3	SynCom 3-4	SynCom 3-5	SynCom 3-6	SynCom 3-7	SynCom 6-1	SynCom 6-2	SynCom 6-3	SynCom 6-4	SynCom 6-5	SynCom 6-7
<i>P. jessenii</i> JV251A	1	7	1	2	0	0	0	0	3	1	4	4	1	0	0	2	2	4	1	0	0
<i>P. inefficax</i> JV551A1	36	2	2	0	0	0	0	0	12	1	4	0	0	3	15	0	1	1	2	0	0
<i>P. kribbensis</i> JV359A8	129	2	2	0	0	0	0	0	47	0	0	0	0	0	782	1273	1273	226	5	7	0
<i>P. fluorescens</i> JV391D10	7187	12394	10413	10391	42444	12142	13766	8584	17265	17265	8584	17265	17265	48745	19768	15288	17492	13215	15499	49355	0
<i>P. fluorescens</i> JV449	1159	2167	916	402	647	751	629	522	259	2671	1219	1695	597	1643	332	29874	32807	31050	28632	29233	2362
<i>P. fluorescens</i> JV391D17	11499	21632	32147	22433	2317	20222	33656	25789	35917	3835	29874	32807	31050	28632	29233	29874	32807	31050	28632	29233	2362
<i>P. wadsworthensis</i> JV244B	2	1	0	0	0	0	0	0	1	1	1	1	1	0	5	0	1	1	1	0	0
<i>P. brassicacerum</i> JV551A7	236	77	1	4	0	0	0	32	6	6	1	2	0	0	46	6	2	3	0	0	0
<i>P. protegens</i> JV245A	10903	9272	11018	3310	2439	5685	10080	11250	1013	81	3701	7014	2064	6598	3695	834	63	74	63	633	934
<i>P. protegens</i> JV359A1	328	96	110	50	524	58	84	106	52	142	49	516	122	516	122	49	516	122	49	516	122
TOTAL READ COUNTS	31480	45650	54610	36592	48371	38953	58225	46261	54704	57752	52804	49327	50143	53616	50730	56986	57752	52804	49327	50143	53616

		number of <i>rpoD</i> reads										B									
		EE in <i>in planta</i> SynCom 2					EE in <i>in planta</i> SynCom 4					EE in <i>in planta</i> SynCom 4									
STRAINS		SynCom 2-1	SynCom 2-2	SynCom 2-3	SynCom 2-4	SynCom 2-5	SynCom 2-6	SynCom 2-7	SynCom 4-1	SynCom 4-2	SynCom 4-3	SynCom 4-4	SynCom 4-5	SynCom 4-6	SynCom 4-7	SynCom 4-8	SynCom 4-9	SynCom 4-10	SynCom 4-11	SynCom 4-12	SynCom 4-13
<i>P. jessenii</i> JV251A	2	1	3	0	5	2	0	0	3	3	0	2	4	0	0	0	0	0	0	0	0
<i>P. inefficax</i> JV551A1	5	5	0	2	2	0	0	0	75	1	5	2	1	2	2	0	0	0	0	0	0
<i>P. kribbensis</i> JV359A8	16	8	3	2	2	0	0	0	174	5	3	4	1	0	0	0	0	0	0	0	0
<i>P. fluorescens</i> JV391D10	11068	6948	2260	4479	1745	11153	15135	10944	5786	5786	10944	5786	4872	1745	1745	1745	1745	1745	1745	1745	1745
<i>P. fluorescens</i> JV449	1525	1036	593	328	342	2093	1384	1115	430	511	963	511	963	511	963	511	963	511	963	511	963
<i>P. fluorescens</i> JV391D17	26327	35562	32093	38509	31789	18360	26861	32105	27706	22912	46959	22912	46959	22912	46959	22912	46959	22912	46959	22912	46959
<i>P. wadsworthensis</i> JV244B	0	1	0	0	0	0	0	0	5	1	0	0	0	0	0	0	0	0	0	0	0
<i>P. brassicacerum</i> JV551A7	336	3	1	0	0	0	0	48	3	0	1	1	1	0	0	0	0	0	0	0	0
<i>P. protegens</i> JV245A	11026	7536	6532	6455	8003	10857	5672	9335	5489	17379	478	17379	478	17379	478	17379	478	17379	478	17379	478
<i>P. protegens</i> JV359A1	759	92	89	67	707	128	70	80	74	95	583	74	95	583	74	95	583	74	95	583	74
TOTAL READ COUNTS	51064	51192	41574	49847	42588	42896	49135	53585	39483	45777	50730	45777	50730	45777	50730	45777	50730	45777	50730	45777	50730

		number of <i>rpoD</i> reads										C									
		EE in <i>in planta</i> SynCom 5					EE in <i>in planta</i> SynCom 7					EE in <i>in planta</i> SynCom 7									
STRAINS		SynCom 5-1	SynCom 5-2	SynCom 5-3	SynCom 5-4	SynCom 5-5	SynCom 5-6	SynCom 5-7	SynCom 7-1	SynCom 7-2	SynCom 7-3	SynCom 7-4	SynCom 7-5	SynCom 7-6	SynCom 7-7	SynCom 7-8	SynCom 7-9	SynCom 7-10	SynCom 7-11	SynCom 7-12	SynCom 7-13
<i>P. jessenii</i> JV251A	0	6	1	2	3	0	0	0	3	1	4	0	0	0	0	0	0	0	0	0	0
<i>P. inefficax</i> JV551A1	1368	41	4	1	2	2	1	16	0	0	1	0	0	0	0	0	0	0	0	0	0
<i>P. kribbensis</i> JV359A8	699	77	177	23	10	0	0	614	914	263	0	0	0	0	0	0	0	0	0	0	0
<i>P. fluorescens</i> JV391D10	6913	11865	9259	9489	7219	11553	14219	16405	7730	2085	16405	7730	2085	16405	7730	2085	16405	7730	2085	16405	7730
<i>P. fluorescens</i> JV449	1270	2519	1821	593	732	827	6165	3805	1499	415	1499	415	1499	415	1499	415	1499	415	1499	415	1499
<i>P. fluorescens</i> JV391D17	17510	41524	35594	39470	31419	47944	15773	22538	22625	36548	22625	36548	22625	36548	22625	36548	22625	36548	22625	36548	22625
<i>P. wadsworthensis</i> JV244B	2	0	0	0	2	0	2	0	2	0	0	0	0	0	0	0	0	0	0	0	0
<i>P. brassicacerum</i> JV551A7	504	139	64	0	1	0	56	10	0	0	0	0	0	0	0	0	0	0	0	0	0
<i>P. protegens</i> JV245A	18475	1426	2065	3427	1633	758	788	5522	8321	5395	8321	5395	8321	5395	8321	5395	8321	5395	8321	5395	8321
<i>P. protegens</i> JV359A1	259	63	52	75	659	62	55	78	689	689	689	689	689	689	689	689	689	689	689	689	689
TOTAL READ COUNTS	47000	57660	48837	53080	41072	61742	37698	48252	40521	45132	45132	45132	45132	45132	45132	45132	45132	45132	45132	45132	45132

**Table S10: Relative abundances (expressed in percent) obtained for (A) SynCom 1,3 and 6 and (B) SynCom 2, 4, 5 and 7 at different cycles of the *in planta* EE.** The relative abundance represents the number of *rpoD* sequences assigned to one strain divided by the total number of *rpoD* sequences in each sample

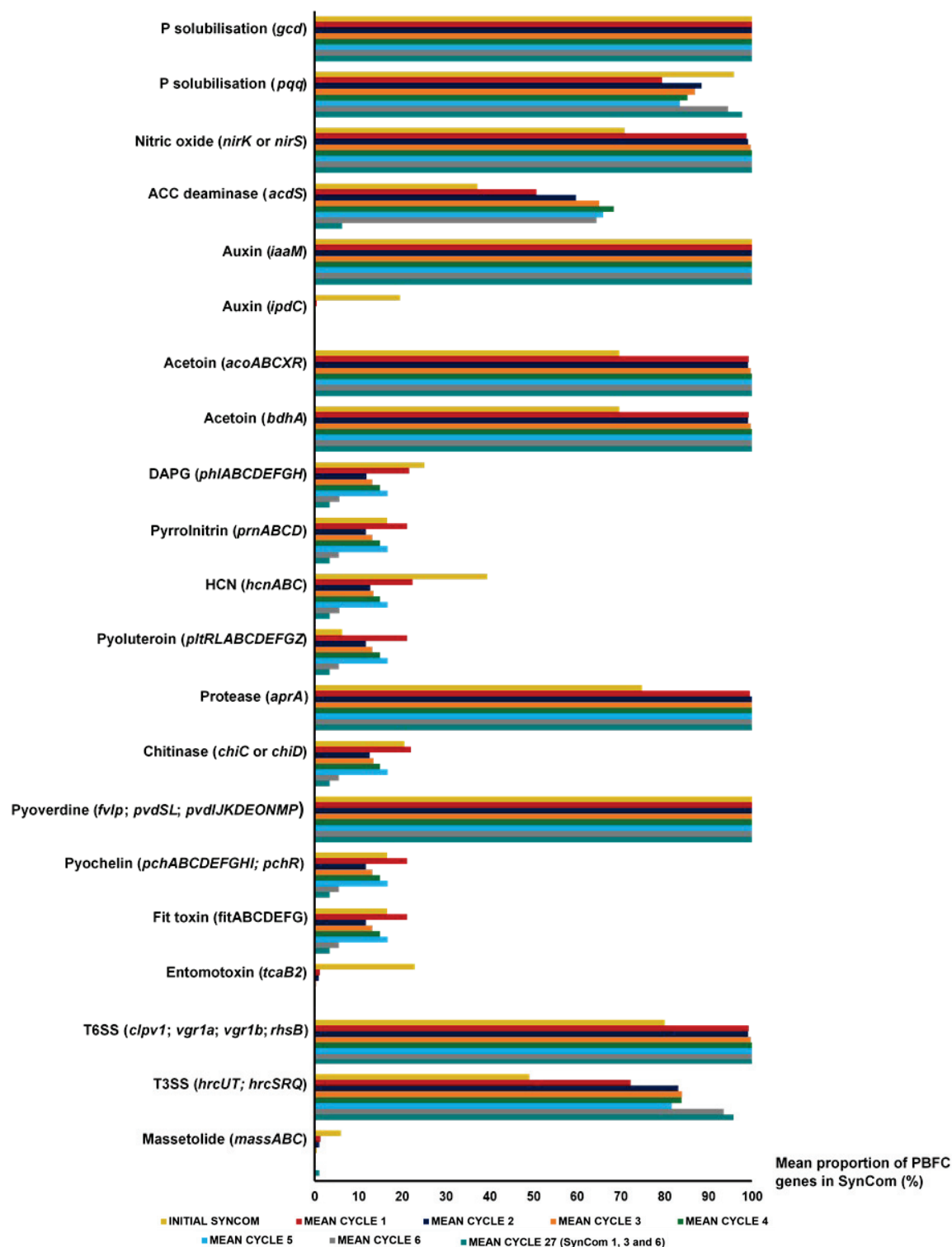
STRAINS	relative abundance (%) of <i>Pseudomonas</i> strain															
	EE <i>in planta</i> SynCom 1								EE <i>in planta</i> SynCom 6							
	SynCom 1-1	SynCom 1-2	SynCom 1-5	SynCom 1-6	SynCom 1-27	SynCom 3-1	SynCom 3-3	SynCom 3-4	SynCom 3-6	SynCom 3-27	SynCom 6-1	SynCom 6-2	SynCom 6-3	SynCom 6-4	SynCom 6-5	SynCom 6-27
<i>P. jessenii</i> JV251A	0.00	0.02	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
<i>P. inflexa</i> JV551A1	0.11	0.00	0.00	0.00	0.00	0.03	0.00	0.01	0.00	0.01	0.03	0.00	0.00	0.00	0.00	0.00
<i>P. kribbensis</i> JV359A8	0.41	0.00	0.00	0.00	0.00	0.12	0.00	0.00	0.00	0.00	1.37	2.20	0.43	0.01	0.01	0.00
<i>P. fluorescens</i> JV391D10	22.83	27.15	19.07	28.40	87.75	31.17	23.64	18.56	31.56	91.15	34.69	26.47	33.25	26.79	30.91	92.05
<i>P. fluorescens</i> JV449	3.68	4.75	1.68	1.10	1.34	1.93	1.08	1.13	0.82	0.56	4.69	2.11	3.22	1.21	3.28	0.62
<i>P. fluorescens</i> JV391D17	36.53	47.39	58.87	61.31	4.79	51.91	57.80	55.75	65.66	7.17	52.42	56.81	59.03	58.45	58.30	4.41
<i>P. waderswilerensis</i> JV244B	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00
<i>P. brassicacelerum</i> JV551A7	0.75	0.17	0.00	0.01	0.00	0.08	0.01	0.00	0.00	0.00	0.08	0.01	0.00	0.01	0.00	0.00
<i>P. protegens</i> JV245A	34.63	20.31	20.18	9.05	5.04	14.59	17.31	24.32	1.85	0.15	6.49	12.15	3.96	13.36	7.37	1.74
<i>P. protegens</i> JV359A1	1.04	0.21	0.20	0.14	1.08	0.15	0.14	0.23	0.10	0.96	0.21	0.25	0.09	0.15	0.13	1.18
TOTAL	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

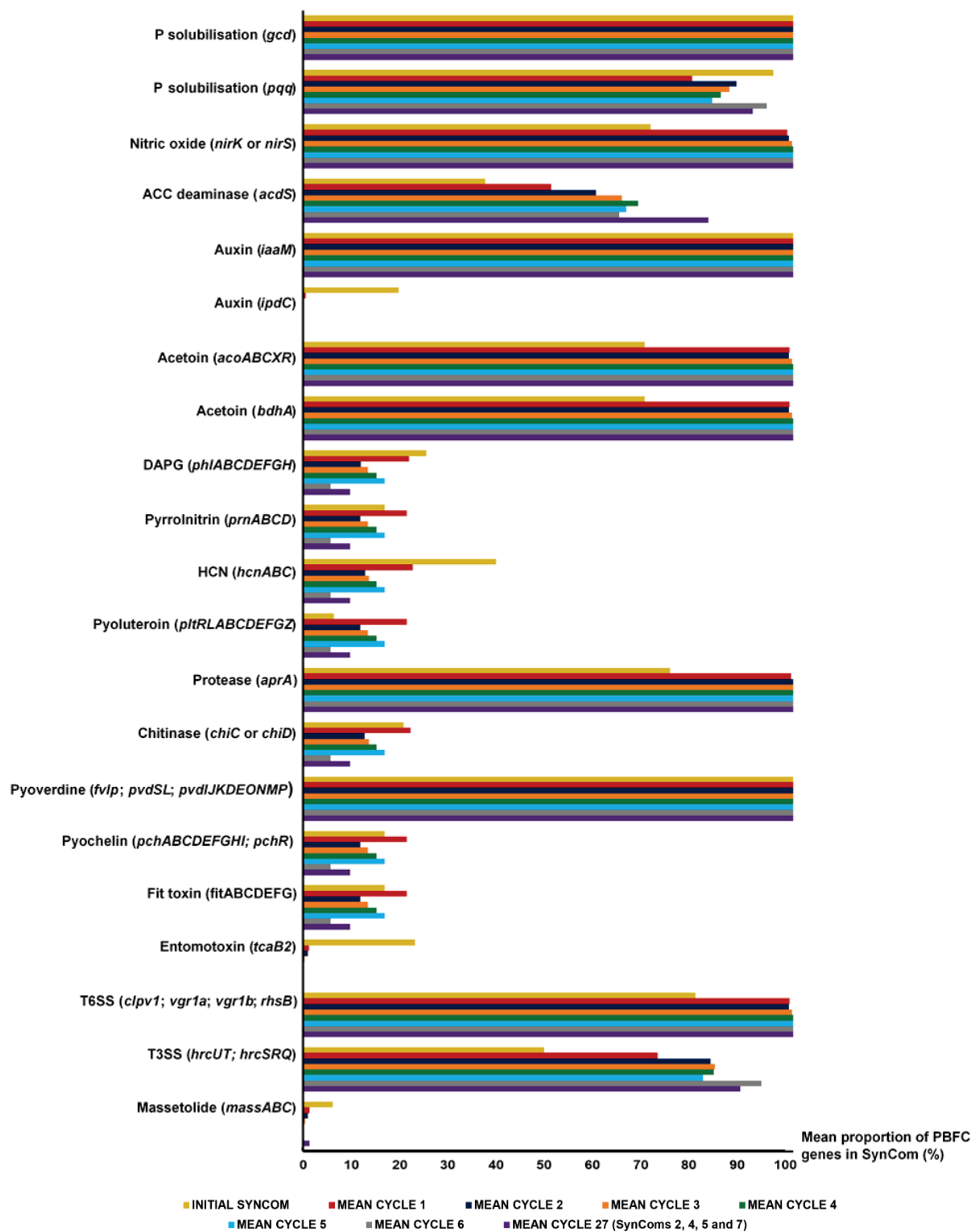
STRAINS	relative abundance (%) of <i>Pseudomonas</i> strain															
	EE <i>in planta</i> SynCom 2								EE <i>in planta</i> SynCom 4							
	SynCom 2-1	SynCom 2-3	SynCom 2-4	SynCom 2-5	SynCom 2-27	SynCom 4-1	SynCom 4-2	SynCom 4-3	SynCom 4-4	SynCom 4-5	SynCom 4-27					
<i>P. jessenii</i> JV251A	0.00	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.01	0.00					
<i>P. inflexa</i> JV551A1	0.01	0.01	0.00	0.00	0.00	0.17	0.00	0.01	0.00	0.00	0.00					
<i>P. kribbensis</i> JV359A8	0.03	0.02	0.01	0.00	0.00	0.41	0.01	0.00	0.01	0.00	0.00					
<i>P. fluorescens</i> JV391D10	21.67	13.57	5.44	8.99	4.10	26.00	30.80	20.42	14.65	10.64	3.44					
<i>P. fluorescens</i> JV449	2.99	2.02	1.43	0.66	0.80	4.88	2.82	2.08	1.09	1.12	1.90					
<i>P. fluorescens</i> JV391D17	51.56	69.47	77.19	77.25	74.64	42.80	54.67	59.91	70.15	50.05	92.57					
<i>P. waderswilerensis</i> JV244B	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00					
<i>P. brassicacelerum</i> JV551A7	0.66	0.01	0.00	0.00	0.00	0.11	0.01	0.00	0.00	0.00	0.00					
<i>P. protegens</i> JV245A	21.59	14.72	15.71	12.95	18.79	25.31	11.54	17.42	13.90	37.96	0.94					
<i>P. protegens</i> JV359A1	1.49	0.18	0.21	0.13	1.66	0.30	0.14	0.15	0.19	0.21	1.15					
TOTAL	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00					

STRAINS	relative abundance (%) of <i>Pseudomonas</i> strain															
	EE <i>in planta</i> SynCom 5								EE <i>in planta</i> SynCom 7							
	SynCom 5-1	SynCom 5-2	SynCom 5-3	SynCom 5-4	SynCom 5-6	SynCom 5-27	SynCom 7-1	SynCom 7-2	SynCom 7-3	SynCom 7-5	SynCom 7-27					
<i>P. jessenii</i> JV251A	0.00	0.01	0.00	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.00					
<i>P. inflexa</i> JV551A1	2.91	0.07	0.01	0.00	0.00	0.00	0.04	0.00	0.00	0.00	0.00					
<i>P. kribbensis</i> JV359A8	1.49	0.13	0.36	0.04	0.02	0.00	1.63	1.86	0.65	0.00	0.00					
<i>P. fluorescens</i> JV391D10	14.71	20.58	18.96	17.88	17.58	18.71	37.72	33.31	19.08	4.62	0.00					
<i>P. fluorescens</i> JV449	2.70	4.37	3.32	1.12	1.78	1.34	16.35	7.73	3.70	0.92	0.00					
<i>P. fluorescens</i> JV391D17	37.26	72.02	72.88	74.36	76.50	77.65	41.84	45.76	55.84	80.98	0.00					
<i>P. waderswilerensis</i> JV244B	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00					
<i>P. brassicacelerum</i> JV551A7	1.07	0.24	0.13	0.00	0.00	0.00	0.15	0.02	0.00	0.00	0.00					
<i>P. protegens</i> JV245A	39.31	2.47	4.23	6.46	3.98	1.23	2.09	11.21	20.54	11.95	0.00					
<i>P. protegens</i> JV359A1	0.55	0.11	0.11	0.14	0.12	1.07	0.16	0.11	0.19	1.53	0.00					
TOTAL	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00					



**Figure S2: Average PBFC genes/clusters distribution for the initial SynCom and for the SynCom 1, 3 and 6 at the end of cycles 1 to 6 and 27 of the *in planta* EE.** Results are expressed in percent and represent the average percent of all strains possessing the PBFC genes/clusters in the referred SynCom at the 27th cycle. For each PBC, the genes or gene clusters are indicated into brackets. P: Phosphate.



**Figure S3: Average PBFC genes/clusters distribution for the initial SynCom and for the SynCom 2, 4, 5 and 7 at the end of cycles 1 to 6 and 27 of the *in planta* EE.** Results are expressed in percent and represent the average percent of all strains possessing the PBFC genes/clusters in the referred SynCom at the 27th cycle. For each PBC, the genes or gene clusters are indicated into brackets. P: Phosphate.



## SUPPLEMENTARY RESULTS

In order to validate the previously presented metabarcoding results, the initial SynCom as well as SynComs 1, 2, 3, 6 and 7 from the 1<sup>st</sup> and 27<sup>th</sup> cycle were analysed via a metagenomic approach. This approach was based on the specific reads that could be attributed to each *Pseudomonas* within the SynCom. The advantage of this method is that a larger number of specific reads were analysed for each *Pseudomonas* as the whole metagenome is analysed instead of only one conserved gene (*i.e.* *rpoD*).

## MATERIAL AND METHODS

### Microbiota diversity analysed through metagenomic analysis

Microbial total DNA samples from the *in planta* SynCom 1, 2, 3, 6 and 7 from the 1<sup>st</sup> and the 27<sup>th</sup> cycle underwent 2 independent metagenomic analyses per DNA extracted samples. The composition of the initial SynCom was assessed 4 times in total. Samples are referred according to the following code, MG-SynCom 1-3, for the metagenomes recovered from the first series of EE at the end of the 3<sup>rd</sup> cycle of evolution.

### Shotgun sequencing from metagenomic DNA

DNA extracts were sent to MR DNA laboratory ([www.mrdnalab.com](http://www.mrdnalab.com); Shallowater, TX) for sequencing. Metagenome sequencing steps include isolation and purification of genomic DNA, fragmentation, ligation to sequencing adapters and purification. Following the amplification and denaturation steps, libraries can be pooled and sequenced. We used 50 ng of DNA from each sample to prepare the libraries using Nextera DNA Sample Preparation Kit (Illumina). Library insert size was determined by Experion Automated Electrophoresis Station (Bio-Rad). The insert size of the libraries ranged from 300 to 850 bp (average 500bp). Pooled library (12pM) was loaded to a 600 Cycles v3 Reagent cartridge (Illumina) and the sequencing was performed on MiSeq (Illumina). Specific reads were attributed to each strain by mapping against the reference genomes available on the MicroScope platform. The reads that could not be separated belong either to the core genome shared at least by two strains, or to unknown sequences (out of the ten genomes).

### *Pseudomonas* genome sequencing and PBFC gene annotation

Gene cluster coding for secondary metabolites was identified using antiSMASH version 5.0.0 and MIBiG version 1.4 (Medema *et al.*, 2011; Blin *et al.*, 2019; Kautsar *et al.*, 2020) as previously described. The presence/absence of PBFC genes varies depending on the strains (e.g. the indole-3-pyruvate pathway found in only 2 strains), the composition of the SynCom will determine how represented each PBFC gene/cluster is (Figure 4). The representation of a

PBFC gene/cluster is the addition of the proportions of strains carrying this PBFC gene/cluster.

## RESULTS

### Initial SynCom composition via the analysis of the total read counts

The obtained results were similar for the 4 samples (table S11 and table S12), the total read counts per sample ranged from 16 398 640 to 18 120 035 reads, among which 12 406 780 (71.9%  $\pm$  0.23 %) corresponded to multiple genomes: reads that could not be matched specifically to only one strain. Around 259 343 reads were classified as 'no hits' (1,5  $\pm$  0.12 %). Further analyses showed these 'no hits' reads matched mostly with *Zea mays* genome (maize cells are probably detached from the roots during the recovery protocol). Sequence reads that matched multiple genomes (sequences common to 2 or more *Pseudomonas* genomes) and 'no hits' reads (no matches against the 10 *Pseudomonas* genomes, essentially maize sequences) were discarded, in order to focus on strain-specific reads.

**Table S11: Number of total reads and number of specific strain reads assigned to *Pseudomonas* populations obtained for the initial SynCom.** The same initial SynCom was used for both *in vitro* and *in planta* EE and its composition was assessed 4 times in total. For each EE, the initial SynCom underwent 2 parallel DNA extractions, and independent metagenomic sequencing analyses were carried for each extraction.

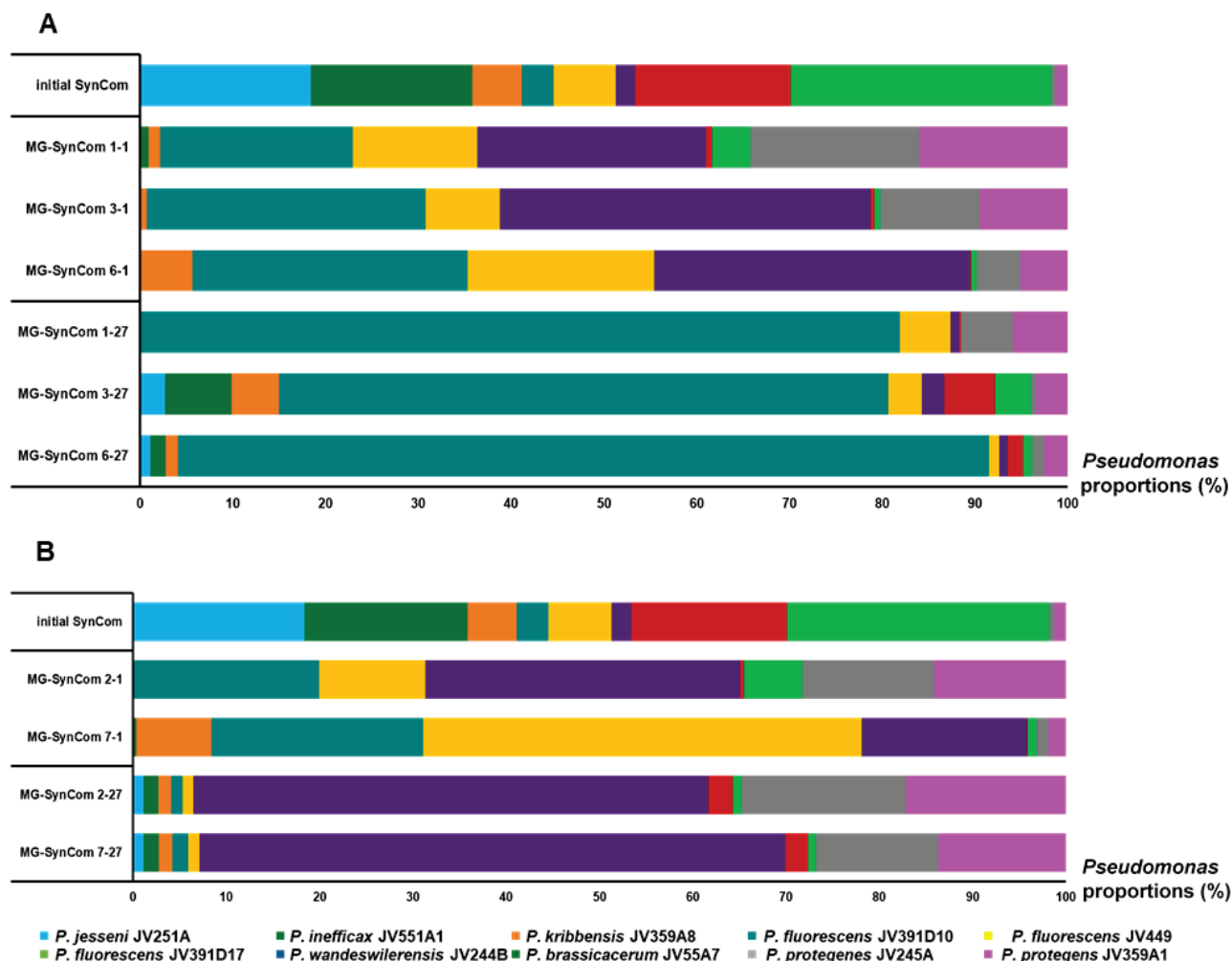
	number of reads				MEAN
	EE <i>in vitro</i> (A)	EE <i>in vitro</i> (B)	EE <i>in planta</i> (A)	EE <i>in planta</i> (B)	
<b>TOTAL READ COUNTS</b>	16398676	16398640	18120035	18120035	17259347
<b>MULTIPLE GENOMES</b>	11893629	11768415	13016654	12948423	12406780
<b>NO HITS</b>	206623	285336	240996	304416	259343
<b>STRAIN SPECIFIC READS</b>	4298424	4344889	4862385	4867196	4593224
<i>P. jessenii</i> JV251A	788106	793569	898451	896055	844045
<i>P. inefficax</i> JV551A1	762470	767560	839543	837823	801849
<i>P. kribbensis</i> JV359A8	225806	228550	258538	258817	242928
<i>P. fluorescens</i> JV391D10	146472	146707	166873	166683	156684
<i>P. fluorescens</i> JV449	269070	272574	349530	349604	310195
<i>P. fluorescens</i> JV391D17	90717	91015	103820	103432	97246
<i>P. wadsworthii</i> JV244B	729845	736858	803881	803680	768566
<i>P. brassicacearum</i> 551A7	1216047	1226082	1361181	1357749	1290265
<i>P. protegens</i> JV245A	22868	23154	26230	26204	24614
<i>P. protegens</i> JV359A1	47023	58820	54338	67149	56833

**Table S12: Relative abundances (expressed in percent) for the initial SynCom.** The relative abundance represents the number of specific sequences assigned to one strain divided by the total number of strain specific sequences in each sample. The relative abundance represents the number of sequences that were assigned to one category (multiple genomes, no hits or strain specific reads) divided by the total number of sequences in each sample.

	relative abundance (%)				MEAN	STANDARD ERROR
	EE <i>in vitro</i> (A)	EE <i>in vitro</i> (B)	EE <i>in planta</i> (A)	EE <i>in planta</i> (B)		
TOTAL READ COUNTS	100	100	100	100	100	-
MULTIPLE GENOMES	72.53	71.76	71.84	71.46	71.90	0.23
NO HITS	1.26	1.74	1.33	1.68	1.50	0.12
STRAIN SPECIFIC READS	26.21	26.50	26.83	26.86	26.60	0.15
<i>P. jessenii</i> JV251A	18.33	18.26	18.48	18.41	18.37	0.05
<i>P. inefficax</i> JV551A1	17.74	17.67	17.27	17.21	17.47	0.13
<i>P. kribbensis</i> JV359A8	5.25	5.26	5.32	5.32	5.29	0.02
<i>P. fluorescens</i> JV391D10	3.41	3.38	3.43	3.42	3.41	0.01
<i>P. fluorescens</i> JV449	6.26	6.27	7.19	7.18	6.73	0.27
<i>P. fluorescens</i> JV391D17	2.11	2.09	2.14	2.13	2.12	0.01
<i>P. wadenswilerensis</i> JV244B	16.98	16.96	16.53	16.51	16.75	0.13
<i>P. brassicacerum</i> 551A7	28.29	28.22	27.99	27.90	28.10	0.09
<i>P. protegens</i> JV245A	0.53	0.53	0.54	0.54	0.54	0.00
<i>P. protegens</i> JV359A1	1.09	1.35	1.12	1.38	1.24	0.08

Concerning the specific read counts, the initial SynCom was analysed 4 times in total and similar results were obtained: 4298424 and 4344889 reads (*in vitro* EE) and 4862385 and 4867196 reads (*in planta* EE, Table S11). The mean was calculated for the 4 analysed samples (Table S12) and obtained strain proportions were as follows (Figure S4): *P. jessenii* JV251A ( $18.37 \pm 0.05$  %), *P. inefficax* JV551A1 ( $17.47 \pm 0.13$  %), *P. kribbensis* JV359A8 ( $5.29 \pm 0.02$  %), *P. fluorescens* JV391D10 ( $3.41 \pm 0.01$  %), JV449 ( $6.73 \pm 0.27$  %), JV391D17 ( $2.12 \pm 0.01$  %), *P. wadenswilerensis* JV244B ( $16.75 \pm 0.13$  %), *P. brassicacerum* JV551A7 ( $28.1 \pm 0.09$  %), *P. protegens* JV245A ( $0.54 \pm 0$  %) and JV359A1 ( $1.24 \pm 0.08$  %).

Overall, 4 PBF, *i.e.* phosphate solubilisation (*gcd* and *pqq*), auxin (*iaaM*) and pyoverdine (*pvd*) production, are over represented (>90%) as they are carried by the majority of the strains.



**Figure S4: *Pseudomonas* population proportions obtained for the initial SynCom and for the in planta EE (A) for MG-SynCom 1,3 and 6 for cycles 1 and 27 and (B) for MG-SynCom 2 and 7 for cycles 1 and 27. The proportions are expressed in percent and represent the number of strain specific sequences assigned to one strain divided by the total number of strain specific sequences in each sample**

### **Analyses of the *in planta* evolution of the SynCom on total read counts**

To confirm the evolutionary dynamics of the initial SynCom observed with the *rpoD* gene, total DNA samples from the first and last cycle (27<sup>th</sup>) were sequenced twice per sample (table S13, and S14). For the first cycle, read counts varied from 12 644 381 to 20 665 912 reads (Table S13). For each sample, between  $89.59 \pm 0.29$  % (MG-SynCom 7-1) to  $91.64 \pm 0.22$  % (MG-SynCom 3-1) of the reads correspond to multiple genomes (Table S14). The percentage of 'no hits' reads varied from  $2.05 \pm 0.22$  % (MG-SynCom 7-1) to  $4.05 \pm 0.14$  % (MG-SynCom 2-1). Removing both of these categories leaves only the reads that were matched specifically to the genome of one strain:  $5.76 \pm 0.06$  % for MG-SynCom 1-1;  $5.42 \pm 0.03$  % for MG-SynCom 2-1;  $4.79 \pm 0.04$  % for MG-SynCom 3-1;  $5.93 \pm 0.04$  % for MG-SynCom 6-1 and  $8.36 \pm 0.06$  % for MG-SynCom 7-1.

**Table S13: Number of total reads and number of specific strain reads assigned to *Pseudomonas* populations, obtained for SynCom 1, 2, 3, 6 and 7 at the end of the 1<sup>st</sup> cycle of the *in planta* EE (= MG-SynCom X-1 with X= 1, 2, 3, 6 or 7). Each SynCom underwent 1 DNA extraction, and independent metagenomic sequencing analyses were carried.**

	number of reads									
	MG-SynCom 1-1 (A)	MG-SynCom 1-1 (B)	MG-SynCom 2-1 (A)	MG-SynCom 2-1 (B)	MG-SynCom 3-1 (A)	MG-SynCom 3-1 (B)	MG-SynCom 6-1 (A)	MG-SynCom 6-1 (B)	MG-SynCom 7-1 (A)	MG-SynCom 7-1 (B)
<b>TOTAL READ COUNTS</b>	14720875	14720875	20665912	20665912	12644381	12644381	16961743	16961743	14598091	14598091
<b>MULTIPLE GENOMES</b>	13534771	13430796	18745553	18673526	11615346	11560219	15532450	15450792	13119871	13036411
<b>NO HITS</b>	347412	434265	808037	865901	428644	474164	430828	498675	267145	332836
<b>STRAIN SPECIFIC READS</b>	838692	855814	1112322	1126485	600391	609998	998465	1012276	1211075	1228844
<i>P. jessenii</i> JV251A	122	157	170	184	75	95	203	224	400	389
<i>P. inefficax</i> JV551A1	7190	7280	38	41	356	371	880	887	3555	3616
<i>P. kribbensis</i> JV359A8	10787	10952	1221	1215	3774	3751	55144	55370	98767	98704
<i>P. fluorescens</i> JV391D10	175668	176158	221824	221475	181816	182073	298911	298604	276022	276372
<i>P. fluorescens</i> JV449	112723	114889	127447	127653	48303	48444	201371	202479	570306	576898
<i>P. fluorescens</i> JV391D17	208907	209203	378021	377575	242463	242000	342725	342186	216607	216612
<i>P. wadenswilerensis</i> JV244B	5299	5325	5053	5082	2119	2168	1585	1592	554	576
<i>P. brassicacerum</i> 551A7	34775	35156	69896	69567	4049	4039	5899	5918	12464	12483
<i>P. protegens</i> JV245A	154145	155159	158162	158152	64770	65100	47573	47822	14542	14533
<i>P. protegens</i> JV359A1	129076	141535	150490	165541	52666	61957	44174	57194	17858	28661

**Table S14: Relative abundances (expressed in percent) for SynCom 1, 2, 3, 6 and 7 at the end of the 1<sup>st</sup> cycle of the *in planta* EE (= MG-SynCom X-1 with X= 1, 2, 3, 6 or 7). The relative abundance represents the number of specific sequences assigned to one strain divided by the total number of strain specific sequences in each sample. The relative abundance represents the number of sequences that were assigned to one category (multiple genomes, no hits or strain specific reads) divided by the total number of sequences in each sample.**

	relative abundance (%)									
	MG-SynCom 1-1		MG-SynCom 2-1		MG-SynCom 3-1		MG-SynCom 6-1		MG-SynCom 7-1	
	MEAN	STANDARD ERROR	MEAN	STANDARD ERROR	MEAN	STANDARD ERROR	MEAN	STANDARD ERROR	MEAN	STANDARD ERROR
<b>TOTAL READ COUNTS</b>	100.00	-	100.00	-	100.00	-	100.00	-	100.00	-
<b>MULTIPLE GENOMES</b>	91.59	0.35	90.53	0.17	91.64	0.22	91.33	0.24	89.59	0.29
<b>NO HITS</b>	2.65	0.29	4.05	0.14	3.57	0.18	2.74	0.20	2.05	0.22
<b>STRAIN SPECIFIC READS</b>	5.76	0.06	5.42	0.03	4.79	0.04	5.93	0.04	8.36	0.06
<i>P. jessenii</i> JV251A	0.02	0.00	0.02	0.00	0.01	0.00	0.00	0.00	0.01	0.00
<i>P. inefficax</i> JV551A1	0.85	0.00	0.00	0.00	0.06	0.00	0.01	0.00	0.10	0.00
<i>P. kribbensis</i> JV359A8	1.28	0.00	0.11	0.00	0.62	0.01	2.45	0.03	1.34	0.06
<i>P. fluorescens</i> JV391D10	20.76	0.18	19.80	0.14	30.07	0.22	0.04	0.22	3.35	0.15
<i>P. fluorescens</i> JV449	13.43	0.01	11.39	0.06	7.99	0.05	6.11	0.08	13.54	0.07
<i>P. fluorescens</i> JV391D17	24.68	0.23	33.75	0.23	40.03	0.36	2.67	0.26	7.96	0.13
<i>P. wadenswilerensis</i> JV244B	0.63	0.00	0.45	0.00	0.35	0.00	0.10	0.00	0.06	0.00
<i>P. brassicacerum</i> 551A7	4.13	0.02	6.23	0.05	0.67	0.01	0.04	0.00	0.22	0.01
<i>P. protegens</i> JV245A	18.25	0.12	14.13	0.09	10.73	0.06	2.95	0.02	1.76	0.01
<i>P. protegens</i> JV359A1	15.96	0.57	14.11	0.58	9.46	0.69	2.87	0.61	2.09	0.43

For the 27<sup>th</sup> cycle, read counts varied from 13 408 279 to 23 618 242 (Table S15). Reads corresponding to multiple genomes represented (Table S16)  $90.84 \pm 0.17\%$  for MG-SynCom 1-27, whereas it was significantly lower for MG-SynCom 2-27 ( $55.64 \pm 0.11\%$ ), MG-SynCom 3-27 ( $40.85 \pm 0.20\%$ ), MG-SynCom 6-27 ( $55.19 \pm 0.19\%$ ) and for MG-SynCom 7-27 ( $57.17 \pm 0.12\%$ ). On the other hand, MG-SynCom 1-27 ( $4.24 \pm 0.14\%$ ) presents lower percentages for the 'no hits' reads than MG-SynCom 2-27 ( $41.45 \pm 0.08\%$ ), MG-SynCom 3-27 ( $56.07 \pm 0.17\%$ ), MG-SynCom 6-27 ( $41.34 \pm 0.16\%$ ) and MG-SynCom 7-27 ( $40.07 \pm 0.09\%$ ). Strain specific read correspond to  $4.92 \pm 0.03$  for MG-SynCom 1-27;  $2.90 \pm 0.03\%$  for MG-SynCom 2-27;  $3.08 \pm 0.02$  for MG-SynCom 3-27;  $3.47 \pm 0.03$  for MG-SynCom 6-27 and  $2.76 \pm 0.03\%$  for MG-SynCom 7-27.

**Table S15: Number of total reads and number of specific strain read assigned to *Pseudomonas* population, obtained for SynCom 1, 2, 3, 6 and 7 at the end of the 27<sup>th</sup> cycle of the *in planta* EE (= MG-SynCom X-1 with X= 1, 2, 3, 6 or 7). Each SynCom underwent 1 DNA extraction, and independent metagenomic sequencing analyses were carried.**

	number of reads									
	MG-SynCom 1-27 (A)	MG-SynCom 1-27 (B)	MG-SynCom 2-27 (A)	MG-SynCom 2-27 (B)	MG-SynCom 3-27 (A)	MG-SynCom 3-27 (B)	MG-SynCom 6-27 (A)	MG-SynCom 6-27 (B)	MG-SynCom 7-27 (A)	MG-SynCom 7-27 (B)
<b>TOTAL READ COUNTS</b>	13408279	13408279	18175918	18175918	18534324	18534324	14854800	14854800	23618242	23618242
<b>MULTIPLE GENOMES</b>	12203900	12157540	10133438	10094349	7607321	7533700	8227409	8170222	13530836	13472094
<b>NO HITS</b>	549739	587282	7521194	7548458	10360687	10425557	6117206	6164742	9442573	9485085
<b>STRAIN SPECIFIC READS</b>	654640	663457	521286	533111	566316	575067	510185	519836	644833	661063
<i>P. jessenii</i> JV251A	101	95	5936	5724	15356	14953	5919	5691	7286	7257
<i>P. inefficax</i> JV551A1	384	399	8634	8558	41334	40514	8342	8047	11089	10731
<i>P. kribbensis</i> JV359A8	35	42	7251	7146	29352	29041	6869	6645	9493	9334
<i>P. fluorescens</i> JV391D10	539676	538516	6255	6175	374962	375337	450329	450608	11080	11086
<i>P. fluorescens</i> JV449	36318	36268	6148	6077	20693	20074	5792	5661	7807	7816
<i>P. fluorescens</i> JV391D17	5942	5974	291533	290800	13991	13930	4884	4861	410039	410115
<i>P. wadsworthii</i> JV244B	1313	1281	13747	13938	31703	31003	8632	8476	15802	15963
<i>P. brassicacearum</i> 551A7	188	203	4915	4880	22312	22013	5191	4779	5608	5443
<i>P. protegens</i> JV245A	36761	36798	93020	92781	3020	2920	6587	6631	86082	85981
<i>P. protegens</i> JV359A1	33922	43881	83847	97032	13593	25282	7640	18437	80547	97337

**Table S16: Relative abundances (expressed in percent) for SynCom 1, 2, 3, 6 and 7 at the end of the 27<sup>th</sup> cycle of the *in planta* EE (= MG-SynCom X-1 with X= 1, 2, 3, 6 or 7). The relative abundance represents the number of specific sequences assigned to one strain divided by the total number of strain specific sequences in each sample. The relative abundance represents the number of sequences that were assigned to one category (multiple genomes, no hits or strain specific reads) divided by the total number of sequences in each sample.**

	relative abundance (%)									
	MG-SynCom 1-27		MG-SynCom 2-27		MG-SynCom 3-27		MG-SynCom 6-27		MG-SynCom 7-27	
	MEAN	STANDARD ERROR	MEAN	STANDARD ERROR	MEAN	STANDARD ERROR	MEAN	STANDARD ERROR	MEAN	STANDARD ERROR
<b>TOTAL READ COUNTS</b>	100.00	-	100.00	-	100.00	-	100.00	-	100.00	-
<b>MULTIPLE GENOMES</b>	90.84	0.17	55.64	0.11	40.85	0.20	55.19	0.19	57.17	0.12
<b>NO HITS</b>	4.24	0.14	41.45	0.08	56.07	0.17	41.34	0.16	40.07	0.09
<b>STRAIN SPECIFIC READS</b>	4.92	0.03	2.90	0.03	3.08	0.02	3.47	0.03	2.76	0.03
<i>P. jessenii</i> JV251A	0.01	0.00	1.11	0.03	2.66	0.06	1.13	0.03	1.11	0.02
<i>P. inefficax</i> JV551A1	0.06	0.00	1.63	0.03	7.17	0.13	1.59	0.04	1.67	0.05
<i>P. kribbensis</i> JV359A8	0.01	0.00	1.37	0.03	5.12	0.07	1.31	0.03	1.44	0.03
<i>P. fluorescens</i> JV391D10	81.80	0.64	1.18	0.02	65.74	0.47	87.48	0.79	1.70	0.02
<i>P. fluorescens</i> JV449	5.51	0.04	1.16	0.02	3.57	0.08	1.11	0.02	1.20	0.01
<i>P. fluorescens</i> JV391D17	0.90	0.00	55.24	0.69	2.45	0.02	0.95	0.01	62.81	0.77
<i>P. wadsworthii</i> JV244B	0.20	0.00	2.63	0.01	5.49	0.10	1.66	0.03	2.43	0.02
<i>P. brassicacearum</i> 551A7	0.03	0.00	0.93	0.01	3.88	0.06	0.97	0.05	0.85	0.02
<i>P. protegens</i> JV245A	5.58	0.03	17.62	0.22	0.52	0.01	1.28	0.01	13.18	0.17
<i>P. protegens</i> JV359A1	5.90	0.72	17.14	1.06	3.40	1.00	2.52	1.02	13.61	1.12

## **Analyses of the *in planta* evolution of the SynCom on strain-specific read counts**

### Population dynamics for SynComs 1, 3 and 6

Population composition changed drastically after only seven days (=end of cycle 1; Figure S4): *P. fluorescens* JV391D17 became the dominant strain in all the SynCom (*i.e.*  $24.68 \pm 0.23\%$  for MG-SynCom 1-1,  $40.03 \pm 0.36\%$  for MG-SynCom 3-1,  $34.06\% \pm 0.26\%$  for MG-SynCom 6-1). The second strain in lead was *P. fluorescens* JV391D10 ( $20.76 \pm 0.18\%$  for MG-SynCom 1-1,  $30.07 \pm 0.22\%$  MG-SynCom 3-1,  $29.72 \pm 0.22\%$  for MG-SynCom 6-1). The third and fourth most represented strains were *P. protegens* JV245A for MG-SynCom 1-1 ( $18.25 \pm 0.12\%$ ) and MG-SynCom 3-1 ( $10.73 \pm 0.06\%$ ) and *P. fluorescens* JV449 for MG-SynCom 6-1 ( $20.09 \pm 0.08\%$ ). *P. fluorescens* JV359A1 was also represented up to  $15.96 \pm 0.57\%$  for MG-SynCom 1-1 while it was in lower proportions for MG-SynCom 3-1 ( $9.46 \pm 0.69\%$ ) and MG-SynCom 6-1 ( $5.04 \pm 0.61\%$ ). Lastly, *P. brassicacerum* JV551A7 and *P. kribbensis* JV359A8 represented  $4.13 \pm 0.02\%$  and  $1.28 \pm 0.00\%$  for MG-SynCom 1-1;  $0.67 \pm 0.01\%$  and  $0.62 \pm 0.01\%$  for MG-SynCom 3-1; and  $0.59 \pm 0\%$  and  $5.50 \pm 0.03\%$  for MG-SynCom 6-1. The remaining strains (*P. jessenii* JV251A, *P. inefficax* JV551A1 and *P. wadenswilerensis* JV244B) were poorly represented (0.02% to 0.85%).

At the end of the 27<sup>th</sup> cycle, a shift in population proportions was observed and *P. fluorescens* JV391D10 became the dominant strain ( $81.80 \pm 0.64\%$  for MG-SynCom 1-27,  $65.74 \pm 0.47\%$  for MG-SynCom 3-27,  $87.48 \pm 0.79\%$  for MG-SynCom 6-27) whereas *P. fluorescens* JV391D17 percentages dropped down to  $0.90 \pm 0\%$  (MG-SynCom 1-27),  $2.45 \pm 0.02\%$  (MG-SynCom 3-27) and  $0.95 \pm 0.01\%$  (MG-SynCom 6-27). The second most represented strain was *P. protegens* JV359A1 ( $5.90 \pm 0.72\%$  for MG-SynCom 1-27,  $3.40 \pm 1\%$  for MG-SynCom 3-27 and  $2.52 \pm 1.02\%$  for MG-SynCom 6-27). JV245A represented  $5.58 \pm 0.03\%$  in MG-SynCom 1-27, whereas it was less represented in MG-SynCom 3-27 ( $0.52 \pm 0.01\%$ ) and in MG-SynCom 6-27 ( $1.28 \pm 0.01\%$ ). The fourth strain in lead was *P. fluorescens* JV449 ( $5.51 \pm 0.04\%$  for MG-SynCom 1-27,  $3.57 \pm 0.08\%$  for MG-SynCom 3-27,  $1.11 \pm 0.02\%$  for MG-SynCom 6-27). MG-SynCom 3-27 presented higher percentages of *P. jessenii* JV251A ( $2.66 \pm 0.06\%$ ), *P. inefficax* JV551A1 ( $7.17 \pm 0.13\%$ ), *P. kribbensis* JV359A8 ( $5.12 \pm 0.07\%$ ), *P. wadenswilerensis* JV244B ( $5.49 \pm 0.1\%$ ) and *P. brassicacerum* JV55A7 ( $3.88 \pm 0.06\%$ ) than MG-SynCom 1-27 and 6-27. The remaining strains were hardly detected (0.01% to 1.66%).



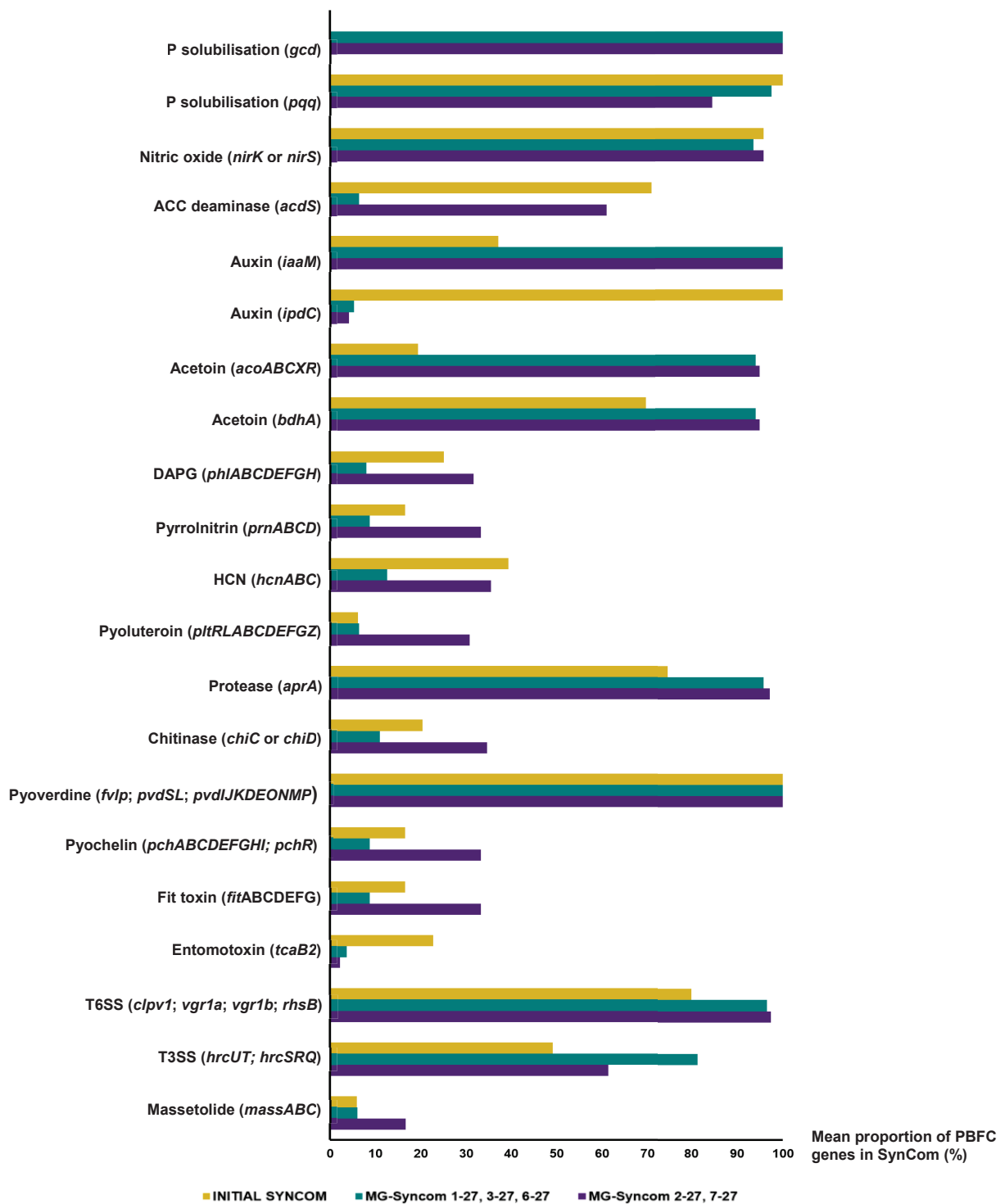
### Population dynamics for SynComs 2 and 7

Population dynamics also changed rapidly for both these SynCom (Figure S5). *P. fluorescens* JV391D17 was the dominant strain ( $33.75 \pm 0.23\%$ ) for MG-SynCom 2-1, whereas it was JV449 ( $47.02 \pm 0.07\%$ ) for MG-SynCom 7-1. In each case, these strains were closely followed by the other *P. fluorescens* strains:  $19.80 \pm 0.14\%$  (JV391D10) and  $11.39 \pm 0.06\%$  (JV449) for MG-SynCom 2-1 and  $22.64 \pm 0.15\%$  (JV391D10) and  $17.76 \pm 0.13\%$  (JV391D17) for MG-SynCom 7-1. *P. protegens* JV245A and JV359A1 represented  $14.13 \pm 0.09\%$  and  $14.11 \pm 0.58\%$  for MG-SynCom 2-1 and  $1.19 \pm 0.01\%$  and  $1.90 \pm 0.43\%$  for MG-SynCom 7-1. Lastly, percentages for *P. kribbensis* JV359A8 were higher for MG-SynCom 7-1 ( $8.09 \pm 0.06\%$ ) than for MG-SynCom 2-1 ( $0.11 \pm 0\%$ ) whereas *P. brassicacerum* JV551A7 percentages were higher for SynCom 2-1 ( $6.23 \pm 0.05\%$ ) than for MG-SynCom 7-1 ( $1.02 \pm 0.01\%$ ). The other strains (*P. jessenii* JV251A, *P. inefficax* JV551A1, and *P. wadenswilerensis* JV244B) were poorly represented ( $0.02\%$  to  $0.45\%$ ).

At the end of the 27<sup>th</sup> cycle, *P. fluorescens* JV391D17 remained the dominant strain ( $55.24 \pm 0.69\%$  for MG-SynCom 2-27 and  $62.81 \pm 0.77\%$  for MG-SynCom 7-27), whereas percentages for JV391D10 and JV449 dropped down to  $1.18 \pm 0.02\%$  and  $1.16 \pm 0.02\%$  (MG-SynCom 2-27) and  $1.70 \pm 0.02\%$  and  $1.20 \pm 0.01\%$  (MG-SynCom 7-27). *P. protegens* JV245A and JV359A1 are the second strains in lead with  $17.62 \pm 0.22\%$  and  $17.14 \pm 1.06\%$  (MG-SynCom 2-27) and  $13.18 \pm 0.17\%$  and  $13.61 \pm 1.12\%$  (MG-SynCom 7-27), respectively. All the other strains (*P. jessenii* JV251A, *P. inefficax* JV551A1, *P. kribbensis* JV359A8, *P. brassicacerum* JV551A7 and *P. wadenswilerensis* JV244B) were poorly represented ( $0.85\%$  to  $2.63\%$ ).

### **PBCF distribution in *in planta*-evolved SynComs**

In SynCom 1, 3 and 6, 7 main PBFC genes/clusters were overrepresented (Figure S5, proportions superior to 90%). They are involved in phosphate solubilisation (*gcd* and *pqq* genes), productions of nitric oxide (*nirK* or *nirS*), auxin (*iaaM*), protease (*aprA*) and pyoverdine (*pvd*), metabolism of acetoin (*aco* and *bdhA*), plus the type 6 secretion system. In SynCom 2 and 7, the same PBFC genes/clusters distribution was observed except for the *pqq* gene (represented in proportion less than 90%). An additional PBFC gene was also most presented in SynCom 2 and 7 compared to SynCom 1, 3 and 6: the ACC deaminase encoding gene *acdS*, ( $61.09\%$  in SynCom 2 and 7 vs  $6.46\%$  in SynCom 1, 3 and 6). Type 3 secretion system genes were represented at 81% for SynCom 1, 3 and 6 and at 61% for SynCom 2 and 7. Finally, no PBFC gene/ cluster was lost.



**Figure S5: Average PBFC genes/clusters distribution for the initial SynCom (yellow) and for the *in planta* MG-SynCom 1,3 and 6 and MG-SynCom 2 and 7 at the end of the 27<sup>th</sup> cycle.** Results are expressed in percent and represent the average percent of all strains possessing the PBFC genes/clusters in the referred SynCom at the 27<sup>th</sup> cycle. For each PBC, the genes or gene clusters are indicated into brackets. P: Phosphate.

## DISCUSSION

Overall, the number of reads obtained with the metagenomic approach is 100 times higher (roughly 4 million specific reads per SynCom) than the number of reads obtained with the barcoding approach (roughly 40 000 reads per SynCom). This is because the metagenomic approach targets the total genome of each strain, whereas the barcoding approach is only based on the portion of the *rpoD* gene. The number of specific reads is higher for the initial SynCom ( $26.60 \pm 0.15\%$  of total reads) than in the evolved SynComs as this sample was constructed in sterile media and all 10 strains are present, giving an overall higher percentage of specific reads (e.g. 4593224 reads in average). After only 7 days of evolution, the number of specific reads dropped down to between  $4.79 \pm 0.04\%$  (MG-SynCom 3-1) and  $8.36 \pm 0.06\%$  (MG-SynCom 7-1) as the initial SynCom quickly evolved toward a majority of *P. fluorescens* strains (JV391D10, JV391D17 and 449) which share together a high number of genes, increasing the proportion of reads corresponding to multiple genomes from  $71.9 \pm 0.23\%$  (initial SynCom) to between  $89.59 \pm 0.29\%$  (MG-SynCom 7-1) and  $91.64 \pm 0.22\%$  (MG-SynCom 3-1). In parallel, the number of 'no hits' reads increased between the initial SynCom and the SynCom from the 1<sup>st</sup> cycle. Most of these reads matched with *Zea mays* genome (BalstN analysis against database) and are probably due to the recovery protocol which may provoke detachment of maize cells from the roots increasing the number of 'no hits'.

For the SynComs from the 27<sup>th</sup> cycle, only MG-SynCom 1-27 presents a similar percentage of multiple genomes reads to those observed in the 1<sup>st</sup> cycle, whereas the number of multiple genomes and 'no hits' reads vary across the MG-SynCom 2-27, 3-27, 6-27 and 7-27. Further analysis revealed that most of the 'no hits' reads matched with a maize endophyte: *Pantoea ananatis* (in addition to maize hits). *P. ananatis* reads are classified as 'no hits' as they cannot be mapped to any specific strain of the SynCom. The presence of this endophyte impacts the number of multiple genomes reads and specific reads as it represents up to  $56.07 \pm 0.17\%$  of the total reads.

Regarding the initial SynCom composition, proportions vary between the 2 approaches presenting similar percentages for only *P. fluorescens* JV449 and *P. protegens* JV359A1 to those obtained with the metabarcoding approach on the *rpoD* gene. This is most likely due to the fact that the metagenomic approach is done only on the specific reads of each strain and the number of specific genes to each strain is variable. Indeed, *P. fluorescens* JV391D10 and JV391D17 are the less represented strains in initial SynCom through the metagenomic approach as they carry the smallest number of specific genes (341 and 296 genes respectively, table S17).

**Table S17: Number of specific genes for each strain.** The number of genes was calculated by comparing the genomes of each strain against all other (coverage > 80%; identity <30 %).

	number of specific genes
<i>P. jessenii</i> JV251A	789
<i>P. inefficax</i> JV551A1	848
<i>P. kribbensis</i> JV359A8	641
<i>P. fluorescens</i> JV391D10	341
<i>P. fluorescens</i> JV449	580
<i>P. fluorescens</i> JV391D17	296
<i>P. wadenswilerensis</i> JV244B	654
<i>P. brassicacerum</i> 551A7	831
<i>P. protegens</i> JV245A	900
<i>P. protegens</i> JV359A1	265

Despite obtaining different population proportions for the initial SynCom, both approaches show the same trend in population dynamics: an overall selection of *P. fluorescens* and *P. protegens*. These two species together represent on average  $93.74\% \pm 0.22\%$  of the recovered populations at the end of the 1<sup>st</sup> cycle. These results are consistent with the values obtained via the metabarcoding approach, where both these species represented  $98,32 \pm 0.62\%$  of the recovered populations at the end of the 1<sup>st</sup> cycle. At end of the EE, the percentage for these duo of species remains on average the same ( $94.47 \pm 0.37\%$ ) for all the SynCom except MG-SynCom 3-27 where these 2 species only represent  $75.68\% \pm 0.32\%$  of the recovered populations, in favour of other strains (*P. inefficax* JV551A1, *P. kribbensis* JV359A8, *P. wadenswilerensis* JV244B and *P. brassicacerum* JV551A7).

The differential selection of *P. fluorescens* JV391D10 or JV391D17 is also observable with the metagenomic approach: (1) *P. fluorescens* JV391D10 is the dominant strain for MG-SynCom 1-27, 3-27 and 6-27 and (2) *P. fluorescens* JV391D17 is the dominant strain for MG-SynCom 2-27 and 7-27. However, using the metagenomic approach, both *P. protegens* strains JV245A and JV359A1 were present with the JV391D10 or the JV391D17 strains, whereas with the metabarcoding approach only *P. protegens* JV245A was observed and JV359A1 remained low for all of the experiment.

One relevant difference between the two approaches is that with the metabarcoding, some strains quickly disappear (0%). This may be because in this approach, only specific reads corresponding to the *rpoD* gene are targeted. Moreover, there is a bias as the samples undergo a PCR amplification before sequencing, which tends to amplify the most represented *rpoD* sequences. This bias makes this method less sensitive to minor changes in population dynamics leading to incomplete conclusions. In our case, *P. jessenii* JV251A, *P. inefficax* JV551A1, *P. kribbensis* JV359A8, *P. wadenswilerensis* JV244B and *P. brassicacerum* JV551A7 are no longer detected in the last cycle (0%) whereas these strains are still present (0.01% to 5.12%) in the metagenomic sequencing analyses.

The increase sensibility of the whole-genome approach was also reflected when focusing on

PBFC genes/clusters distribution. Type 3 secretion system genes were no longer over-represented, using this approach (represented at less than 90%). On the other hand, no PBFC genes/clusters are lost as all the strains are detected at the 27<sup>th</sup> cycle. Finally, SynCom 2 and 7 harbour more diverse populations as they present higher proportions of *P. protegens* JV245A and JV359A1. These strains carry the higher number of PBFC genes/clusters, which increased the proportions of PBFC genes/clusters (*i.e.* *fit* and *pch*) that were less represented with the metabarcoding approach.

Altogether these results confirm the populations dynamics observed with the metabarcoding approach and reveal the limits of it (less sensitive to minor changes). Both approaches are complementary and offer independent validation of the evolved community compositions.

---

## CHAPTER 3

---

To better understand the functioning of plant-PGPR cooperations, and especially PGPR root colonization, we setup an EE approach with a SynCom of 10 strains of *Pseudomonas*. In the second chapter of this thesis we studied the population dynamics underwent by the SynComs by using an experimental approach *in planta*. We carried 7 independent series of EE and found rapid and different population dynamics: SynComs 1-27, 3-27 and 6-27 harboured *P. fluorescens* JV391D10 and *P. protegens* JV245 as the dominant strains whereas, it was *P. fluorescens* JV391D17 and *P. protegens* JV245A for SynComs 2-27, 4-27, 5-27 and 7-27. In parallel, to confirm that the observed population dynamics were due to the plant, we performed an identical EE but in a rich liquid medium in absence of the plant (chapter 2) which led to *P. inefficax* JV551A1 as the dominant strain. These different population dynamics are the result of population interactions and adaptation within the SynComs. Historically, biotic interactions are known to have given rise to major evolutive changes (*i.e.* for instance the transition from aquatic to terrestrial ecosystems; Battistuzzi and Hedges, 2009; Wisniewski-Dyé *et al.*, 2011). In this study, despite a differential selection of set of *Pseudomonas* species, the PBFC genes/clusters distribution remains similar for all the *in planta*-evolved SynComs. This results suggest that the plant shapes a microbial selection in order to maintain a certain type of PBFC genes/clusters within the plant associated community, rather than selecting a particular strain. Through the EE, the initial SynCom evolved in order to adapt to the gnotobiotic system. Evolution happens through adaptation to new abiotic (e.g. salinity, temperature, *etc.*) and biotic (the other microorganisms) environments. Adaptation is the result of mutations in genes that either improve an existing function or enable the organism to acquire new functions. To get further insight into the populations evolution underwent by the initial SynCom during the *in planta* EE, we sought for changes in the genomes of the strains, particularly in genes involved in plant-bacterial interaction. To do this, different sets of genomic data were analysed with Breseq (Deatherage *et al.*, 2015), a software for predicting mutations (SNPs, deletions, insertions and substitutions). Firstly, the specific reads of each strain from the metagenomics data of SynComs 1-1, 2-1, 6-1, 7-1, and 1-27, 2-27, 6-27, 7-27 were analysed. In this study, only the analysis obtained for the dominant strains of these SynComs are presented: *P. fluorescens* JV391D10 or JV391D17 and *P. protegens* JV245A. Secondly, evolved clones were isolated from SynComs 1-27, 2-27, 6-27, 7-27. The isolated clones corresponded to the dominant strains (*P. fluorescens* JV391D10 or JV391D17 or *P. protegens* JV245A). The genome of this variants were screened for point mutations with Breseq.

The aim of this study was to highlight the impact the EE had on the strain's genomes and to reveal if there were parallel changes between the strains or between the evolved clones from a same species



## Experimental evolution of a synthetic community of *Pseudomonas*: different trajectories to reach the same functional network

**Auteurs** Beatriz Manriquez, Aurélie Peticca, Jeanne Doré, Danis Abrouk, Claire Prigent-Combaret\*, Daniel Muller

**Affiliations** Univ Lyon, Université Claude Bernard Lyon 1, CNRS, INRAE, VetAgro Sup, UMR5557 Ecologie Microbienne, 43 bd du 11 novembre 1918, F-69622 Villeurbanne, France

**\*Corresponding author:** UMR CNRS 5557 Ecologie Microbienne, Université Lyon 1, 43 bd du 11 Novembre 1918, 69622 Villeurbanne cedex, France. Phone +33 4 72 43 13 49. E-mail: [claire.prigent-combaret@univ-lyon1.fr](mailto:claire.prigent-combaret@univ-lyon1.fr)

## ABSTRACT

Plants offer a multitude of ecological niches, providing microorganisms with an ever-changing environment shaped by the other sympatric evolving microorganisms. The evolutionary process of adaptation can only be seen through observation of variations of ecological traits. In order to remain in the same habitat, microorganisms need to constantly adapt by using the available resources in a more efficient way than their neighbors. Adaptation happens through mutations and if a mutation is advantageous, it will be fixed and will rise to the population level after several generations. In order to investigate the underlying mechanics of plant microbiota adaptation to maize roots, an experimental evolution with a synthetic community of 10 strains of *Pseudomonas* was previously carried. The obtained results showed two different evolved synthetic communities after 27 cycles of evolution (roughly 400 bacterial generations). In order to get deeper insights into the evolutionary adaptation of microorganisms during biotic interactions we sought for genetic changes at 2 levels within the evolved synthetic communities: first, by an analysis of metagenome sequence data of the evolved synthetic communities and second, by the analysis of the genomes of evolved clones isolated at the end of the 27th cycle. Fast adaptation was observed with the metagenomic data: most of the mutations were fixed and involved in metabolism, indicating that 400 bacterial generations are long enough to adapt to a new environment (maize roots). Furthermore, various transcriptional regulators were found mutated for both analyses, with parallel mutations for *gacA* in the evolved clones. Overall, these results evidence rapid adaptive genetic changes in *Pseudomonas* populations colonizing maize roots for a sufficient period of time.

**Keywords:** Experimental Evolution, synthetic community, PGPR, plant microbiota, plant-beneficial function contributing genes.

**Abbreviations:** Experimental Evolution (EE); Synthetic Community (SynCom); Plant Growth-Promoting Rhizobacteria (PGPR); Single Nucleotide Polymorphism (SNP).

## INTRODUCTION

EE shows real time adaptation of a given organism to its environment. Adaptation happens when advantageous mutations are fixed via the action of natural selection. In order to let natural selection act, the initial SynCom co-evolved in a stable environment (a gnotobiotic system composed only of pre-germinated maize seedling with no additional nutrient source). Co-evolution can be defined as the process of reciprocal adaptation by interacting species (Buckling and Rainey, 2002). Evolving together creates adaptive coevolutionary dependencies in a changing environment, meaning that species have to constantly adapt to the other evolving species in order to survive (*i.e.* “Red Queen hypothesis”; Van Valen, 1977; Strotz *et al.*, 2018). Co-evolution drives fast evolution between species with strong ecological interaction (Brockhurst *et al.*, 2003; Guan *et al.*, 2013). Eventually, the most adapted strains should be the most abundant in the evolved synthetic communities. Adaptation results from mutations in genes that either improve an existing function or enable the organism to acquire new functions. These mutations happen randomly in some individuals of the population and if they are advantageous, they will require a certain time to rise to the population level and to be fixed. The time required is inversely proportional to the advantages it gives (Fisher, 1930; Tenaillon, 2014). Moreover, another evolutive force, random drift, also plays an important role in fixation of mutations: it represents the role of chance in deciding which organisms survive and reproduce (despite the difference in their fitness). Genetic variation is therefore reduced randomly and many mutations disappear before they can be fixed (Elena and Lenski, 2003). In summary, the fate of a mutation depends on the advantages it gives: the more advantageous the mutation is, the less it will be subjected to random drift as it will be rapidly fixed (Lenski *et al.*, 1991). Bacteria are powerful candidates to pinpoint mutations as they offer short generation time and large population size, so multiple mutations can be present simultaneously.

The rhizosphere of plants is the home to many microorganisms. Bacteria, fungi and other microorganisms co-exist together in plant roots assuring a variety of functions such as plant nutrition and protection against plant pathogens (Vacheron *et al.*, 2016). Plant-growth promoting rhizobacteria are among the most remarkable plant beneficial microorganisms as they can harbor a panel of plant beneficial contributing genes within their genomes (Bruto *et al.*, 2014). However, their inoculation does not always lead to significant effect on the plant. One of the reasons for this is that PGPRs are often inoculated into already well-established microbiota and are likely to be “lost” among all the microorganisms already present. To unravel PGPR establishment in the rhizosphere of maize we carried out an *in planta* experimental evolution (EE) with a synthetic community (SynCom) composed of 10 strains of *Pseudomonas*. In our previous work (see chapter 2), the synthetic community was evolved *in planta* for 27 cycles of evolution ( $\approx 400$  bacterial generations). At the end of each cycle of evolution, the evolved synthetic community was retrieved from the roots and re-inoculated on fresh

seedlings. Our results show a rapid evolution, with half of the strains being lost after only 3 serial passages. There were two different selection dynamics: *P. fluorescens* JV391D10 and *P. protegens* JV245A were the dominant strain at the end of the 27<sup>th</sup> for SynComs 1-27 and 6-27, whereas it was *P. fluorescens* JV391D17 and *P. protegens* JV245A for SynComs 2-27 and 7-27).

To get further insight into the populations dynamics underwent by the initial SynCom during the *in planta* EE, we sought for changes in the genomes of the strains. Evolution within the genomes of the different *Pseudomonas* found in cycle 27 was analyzed by 2 approaches: i) an analysis of metagenome sequence data, then refined ii) by the analysis of the genomes of evolved clones isolated in cycle 27. First, metagenome data from evolved SynCom 1-1 (SynCom 1 retrieved from maize roots at the end of the 1<sup>st</sup> cycle), 2-1, 6-1, 7-1, and 1-27, 2-27, 6-27, 7-27 were screened for the presence of mutations of the Single Nucleotide Polymorphism (SNP) type, in the strains *P. fluorescens* JV391D10, 391D17 and *P. Protegens* JV245A. Secondly, individual clones were isolated from evolved SynCom 1-27, 2-27, 6-27 and 7-27. Although the analysis of metagenome data made it possible to characterise the presence of 5 different strains co-existing at the end of the 27<sup>th</sup> cycle of the *in planta* EE (*P. fluorescens* JV391D10, JV391D17, JV449 and *P. protegens* JV245A, JV359A1), only clones identified as *P. fluorescens* JV391D10, JV391D17 and *P. protegens* JV245A could be isolated from the communities resulting from the 27<sup>th</sup> cycle. Variants of the latter were specifically studied because this strain possess the higher number of plant beneficial function contributing genes amongst the three strains. In addition, a better sequencing effort was devoted to the *P. protegens* JV245A ancestral strain leading to a higher genome quality than that of *P. fluorescens* JV391D10 and JV391D17. This point is primordial when comparing genome sequences from evolved clones and the ancestral strain in order to identify the evolutive changes that occurred during the EE.

## **MATERIALS AND METHODS**

### **Bacterial Strains and Growth Conditions**

The strains from this study are ancestral and evolved clones of strains *P. fluorescens* JV391D10, *P. fluorescens* JV391D17, and *P. protegens* JV245A. The ancestral clones were used to form an initial SynCom composed of 10 strains (*i.e.* *P. jessenii* JV251A, *P. inefficax* JV551A1, *P. kribbensis* JV359A8, *P. fluorescens* JV391D10, JV449, JV391D17, *P. wadenswilerensis* JV244B, *P. brassicacerum* JV551A7, *P. protegens* JV245A and JV359A1).

This initial SynCom was evolved on maize seedling during 27 cycles of evolution (see chapter 2). The evolved clones were retrieved from plant roots at the end of the 27<sup>th</sup> cycle of evolution. All strains were grown onto King's B agar medium (Conda, Pronadisa, Spain) and incubated at 28°C for 48h, or grown in liquid TSB medium (Conda, Pronadisa, Spain), overnight, at 28°C under agitation at 300rpm.

### **Isolation and identification of evolved clones from the 27<sup>th</sup> cycle of the *in planta* EE**

One hundred microliters of the bacterial suspensions, corresponding to SynComs 1, 2, 6 and 7 from the 27<sup>th</sup> cycle (previously stored in glycerol (50% v/v) at -80°C, see chapter 2) were serially diluted and plated onto King's B medium. For each SynCom, 96 clones were isolated, and 30 clones were selected, representing selected morphotypes (*i.e.* yellow smooth colonies that colour the medium for *P. fluorescens* and smooth beige colonies for *P. protegens*). The selected clones were identified by PCR with previously designed primers (*rpoD*\_F: TCGCCAAGAAGTACACCAAC and *rpoD*\_R: CCATGGAGATCGGCTCTT) that amplify a 353bp fragment of *rpoD* (position 1179 to 1532; see chapter 2). Briefly, bacterial DNA was extracted from the selected clones using a NucleoSpin Tissue extraction kit (Macherey-Nagel, Germany). Extracted DNA concentration was quantified using a nanophotometer (Implen, Germany) and subsequently diluted or concentrated to at least 15 ng/μL. The *rpoD* primers were used to amplify a 353bp fragment of *rpoD* (position 1179 to 1532) using the HotStarTaq Plus Master Mix Kit (Qiagen, Valencia, CA) under the following conditions: 94°C for 3 min, followed by 28 cycles of 94°C for 30 s, 53°C for 40 s and 72°C for 1 min, with a final elongation step at 72°C for 5 min. PCR products were sent to Biofidal ([www.biofidal.com](http://www.biofidal.com); Vaulx-en-Velin, France) for sequencing. Overall, 8 evolved clones were selected for *P. fluorescens* JV391D10 (*i.e.* corresponding to genome sequencing samples A to H, later described), 9 for JV391D17 (*i.e.* corresponding to samples I to Q) and 17 for *P. protegens* JV245A (*i.e.* corresponding to samples A to Q).

### **DNA extraction of *Pseudomonas* communities' samples**

All of the analysed *Pseudomonas* SynCom samples underwent one DNA extraction and 2 independent metagenomic analysis, except for the initial SynCom which was analysed twice (two parallel DNA extractions and two independent metagenomics analysis for each extraction were performed). Samples are referred to by the plant series of EE (*i.e.* SynCom 1) and by the time of recovery (*i.e.* cycle 27), giving the following code: SynCom 1-27 for the population recovered from the first plant series of EE at the end of the 27<sup>th</sup> cycle of evolution. Microbiota total DNA was extracted from the samples (v= 50 μL, initial SynCom, evolved SynCom 1-1, 2-

1, 6-1, 7-1, and 1-27, 2-27, 6-27, 7-27) using a NucleoSpin Tissue extraction kit (Macherey-Nagel, Germany). Extracted DNA concentration was quantified and subsequently diluted or concentrated to at least 15 ng/μL.

### **Shotgun sequencing from SynCom DNA**

DNA extracts were sent to MR DNA laboratory ([www.mrdnalab.com](http://www.mrdnalab.com); Shallowater, TX) for sequencing. Metagenome or genome sequencing steps include isolation and purification of genomic DNA, fragmentation, ligation to sequencing adapters and purification. Following the amplification and denaturation steps, libraries can be pooled and sequenced. We used 50 ng of DNA from each sample to prepare the libraries using Nextera DNA Sample Preparation Kit (Illumina). Library insert size was determined by Experion Automated Electrophoresis Station (Bio-Rad). The insert size of the libraries ranged from 300 to 850 bp (average 500 bp). Pooled library (12 pM) was loaded to a 600 Cycles v3 Reagent cartridge (Illumina) and the sequencing was performed on MiSeq (Illumina).

Reads from SynComs metagenomes were sorted by using the genomes available on the MicroScope platform (Vallenet *et al.*, 2020) as references. During this separation analysis, the reads corresponding to the core genome could not be separated and attributed to one particular strain. There was also a 'no hits' reads category that did not match to any of the 10 *Pseudomonas* genome. 'No hits' reads were analyzed by blastn (nr), revealing that these reads mostly correspond to maize sequences. Therefore, the reads were classified into 3 categories: (1) multiple genome reads; these are the reads that could not be attributed to one *Pseudomonas* specifically, (2) 'no hits' reads; the reads that did not map to any of the *Pseudomonas* genomes and (3) specific reads; in this last category we can find all the reads that were successfully attributed to one particular *Pseudomonas*. Genetic variations were studied on the specific reads only.

### **Single Nucleotide Polymorphism (SNP) analysis on SynCom metagenomes**

Sequence variation was assessed on the specific reads by Breseq (version v0.35.4; <https://barricklab.org/twiki/bin/view/Lab/ToolsBacterialGenomeResequencing>; Deatherage *et al.*, 2015), a software for predicting mutations (SNPs, deletions, insertions and substitutions). Breseq uses as reference the genomes available on the MicroScope platform version 3.14.1 (Vallenet *et al.*, 2020).

### **Whole genome sequencing of the ancestral clones**

The whole-genome sequencing of ancestral *P. protegens* JV245 was carried out using Illumina NovaSeq and Pacific Biosciences platforms. Hybrid de novo assembly of the 10.3 million reads

from NovaSeq sequencing and 25 493 reads from PacBio sequencing was performed using Unicycler (Wick *et al.*, 2017). An unique circular contig that was indicated as complete sequence was obtained from the assembly and was annotated using prokka (Seemann, 2014). On the other hand, the whole genome sequencing of the strains JV391D10 and JV391D17 was carried out using Illumina NovaSeq. The sequencing run produce 9.3 and 9.1 million reads of 150 bp paired ends and de novo assembly was performed using Unicycler. The hybrid assembly for *P. protegens* JV245A was used as the reference genome for the Breseq analysis.

## **Shotgun sequencing of the isolated clones from the 27<sup>th</sup> cycle of the *in planta* EE**

DNA extracts were mixed to form pairs with clones that had been identified as being of different species. There were 8 mixtures of *P. fluorescens* JV391D10 and *P. protegens* JV245A (mixtures A to H) plus 9 mixtures of *P. fluorescens* JV391D17 and *P. protegens* JV245A (mixtures I to Q). As an example, the mixture A corresponded to one clone of *P. fluorescens* JV391D10 and one of *P. protegens* JV245A, both clones are referred to by the letter A. DNA mixtures were sent to Novogene laboratory (<https://en.novogene.com>; London, UK) for sequencing. Metagenomes or genome sequencing steps include isolation and purification of genomic DNA, fragmentation, ligation to sequencing adapters and purification. Following the amplification and denaturation steps, libraries can be pooled and sequenced. We used 50 ng of DNA from each sample to prepare the libraries using Nextera DNA Sample Preparation Kit (Illumina). Library insert size was determined by Experion Automated Electrophoresis Station (Bio-Rad). The insert size of the libraries ranged from 300 to 850 bp (average 500bp). Pooled library (12pM) was loaded to a 600 Cycles v3 Reagent cartridge (Illumina) and the sequencing was performed on MiSeq (Illumina). Afterwards, the reads from each mixture were separated using the genomes available on MicroScope platform as reference (for strains *P. fluorescens* JV391D10 and JV391D17). During this separation analysis, the reads corresponding to the core genome cannot be separated and attributed to a particular strain. Therefore, the rest of the analysis is done only on the specific reads that were successfully classified.

## **Single Nucleotide Polymorphism (SNP) analysis on SynCom metagenomes**

Sequence variation was assessed in the specific reads of each variant by Breseq, using the genomes available on the MicroScope platform for strains *P. fluorescens* JV391D10 and JV391D17 and the hybrid assembly genome of *P. protegens* JV245, as references. These reference genomes correspond to the genomes of the strains that composed the initial SynCom. Only strain-specific genome regions were studied, as the core genome sequences of the strains could not be specifically attributed to one given strain.



## RESULTS AND DISCUSSION

The EE with the initial SynCom of 10 strains of *Pseudomonas* was carried for 27 cycles of evolution (roughly 400 bacterial generations). The *in planta* EE led to rapid changes in population dynamics with half of the strains being lost after 3 serial passages. Furthermore, at the end of the 27<sup>th</sup> cycle, two main types of communities stood out: SynCom 1 and 6 were dominated by *P. fluorescens* JV391D10 and *P. protegens* JV245A, whereas it was *P. fluorescens* JV39D17 and *P. protegens* that prevailed in SynComs 2 and 7 (see chapter 2). Genetic variation was assessed for each strain separately and in every SynCom. The aim was to see if, for a given *Pseudomonas* line, there were genetic variation common to SynComs that shared the same evolutive outcome.

### Genetic variations within Synthetic Community metagenomes data

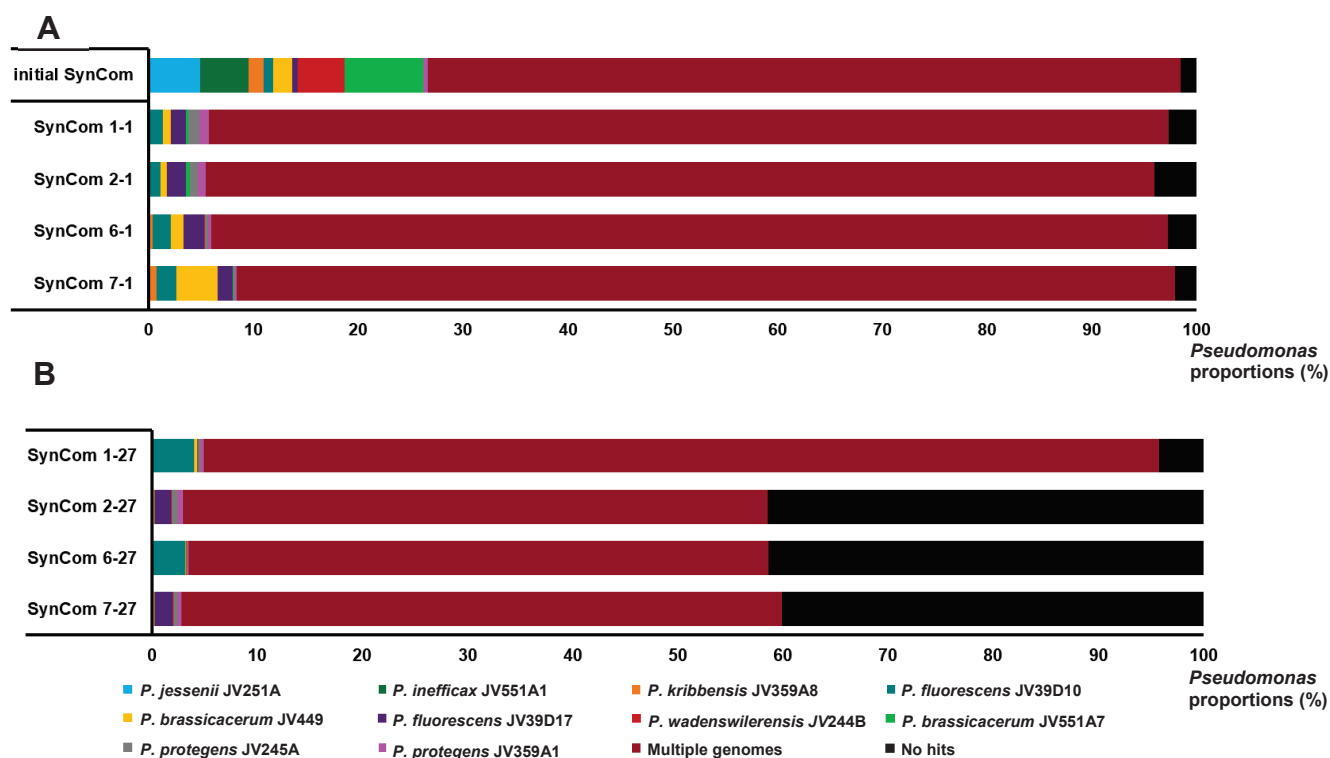
For the initial SynCom, specific read counts to a given *Pseudomonas* strains (cf. the supplementary data of chapter 2) varied from 4 298 424 to 4 867 196. These reads represented around  $26.6 \pm 0.15\%$  of the total reads (Table 1A and Figure 1A). For the SynComs from the 1<sup>st</sup> cycle, the total number of specific reads varied from 838 692 to 1 228 844, representing the following percentages compared to the total number of reads:  $5.76 \pm 0.06\%$  for SynCom 1-1;  $5.42 \pm 0.03\%$  for SynCom 2-1;  $5.93 \pm 0.04\%$  for SynCom 6-1 and  $8.36 \pm 0.06\%$  for SynCom 7-1 (Table 1A and Figure 1A). For all the EE analysed, the percentage of specific readings drops rapidly during the first EE cycle (from  $26.6 \pm 0.15\%$  to less than 10%), this is because the EE led to rapid population dynamics where 2 *P. fluorescens* and 2 *P. protegens* quickly became dominants (chapter 2), increasing the proportion of reads that belong to their core genome (multiple genome reads category) and in turn dropping the number of specific reads. At the end of the EE (27<sup>th</sup> cycle), specific read counts varied from 510 185 to 663 457, representing the following percentages compared to the total number of reads:  $4.92 \pm 0.03\%$  for SynCom 1-27;  $2.90 \pm 0.03\%$  for SynCom 2-27;  $3.47 \pm 0.03\%$  for SynCom 6-27 and  $2.76 \pm 0.03\%$  for SynCom 7-27 (Table 1B and Figure 1B) whereas the 'no hits' reads categories arise up to half of the total reads for SynComs 2-27, 6-27 and 7-27.

**Table 1: Relative abundances (expressed in percent) for the initial SynCom, for SynComs from the 1<sup>st</sup> cycle (A) and from the 27<sup>th</sup> cycle (B).** Mean obtained for the initial SynCom which was analysed 4 times in total: 2 parallel DNA extractions, and independent metagenomic sequencing analyses were carried for each extraction. Mean obtained for each SynCom which were analyzed twice: 1 DNA extraction and 2 independent metagenomic sequencing. The relative abundance represents the number of sequences that were assigned to one category (multiple genomes, no hits or strain specific reads) divided by the total number of sequences in each sample.

A									
relative abundance (%)									
initial SynCom	STANDARD ERROR	SynCom 1-1	STANDARD ERROR	SynCom 2-1	STANDARD ERROR	SynCom 6-1	STANDARD ERROR	SynCom 7-1	STANDARD ERROR
MULTIPLE GENOMES	71.90	0.23	91.59	0.35	90.53	0.17	91.33	0.24	89.59
NO HITS	1.50	0.12	2.65	0.29	4.05	0.14	2.74	0.20	2.05
STRAIN SPECIFIC READS	26.60	0.15	5.76	0.06	5.42	0.03	5.93	0.04	8.36
<i>P. jessenii</i> JV251A	4.89	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00
<i>P. inefficax</i> JV551A1	4.65	0.01	0.05	0.00	0.00	0.00	0.01	0.00	0.02
<i>P. kribbensis</i> JV359A8	1.41	0.01	0.07	0.00	0.01	0.00	0.33	0.00	0.68
<i>P. fluorescens</i> s JV391D10	0.91	0.01	1.19	0.00	1.07	0.00	1.76	0.00	1.89
<i>P. fluorescens</i> JV449	1.79	0.08	0.77	0.01	0.62	0.00	1.19	0.00	3.93
<i>P. fluorescens</i> JV391D17	0.56	0.01	1.42	0.00	1.83	0.00	2.02	0.00	1.48
<i>P. wadsworthii</i> JV244B	4.45	0.01	0.04	0.00	0.02	0.00	0.01	0.00	0.00
<i>P. brassicacervum</i> 551A7	7.47	0.02	0.24	0.00	0.34	0.00	0.03	0.00	0.09
<i>P. protegens</i> JV245A	0.14	0.00	1.05	0.00	0.77	0.00	0.28	0.00	0.10
<i>P. protegens</i> JV359A1	0.33	0.02	0.92	0.04	0.76	0.04	0.30	0.04	0.16

B									
relative abundance (%)									
SynCom 1-27	STANDARD ERROR	SynCom 2-27	STANDARD ERROR	SynCom 6-27	STANDARD ERROR	SynCom 7-27	STANDARD ERROR	STANDARD ERROR	STANDARD ERROR
MULTIPLE GENOMES	90.84	0.17	55.64	0.11	55.19	0.19	57.17	0.12	0.12
NO HITS	4.24	0.14	41.45	0.08	41.34	0.16	40.07	0.09	0.09
STRAIN SPECIFIC READS	4.92	0.03	2.90	0.03	3.47	0.03	2.76	0.03	0.03
<i>P. jessenii</i> JV251A	0.00	0.00	0.03	0.00	0.04	0.00	0.03	0.00	0.00
<i>P. inefficax</i> JV551A1	0.00	0.00	0.05	0.00	0.06	0.00	0.05	0.00	0.00
<i>P. kribbensis</i> JV359A8	0.00	0.00	0.04	0.00	0.05	0.00	0.04	0.00	0.00
<i>P. fluorescens</i> s JV391D10	4.02	0.00	0.03	0.00	3.03	0.00	0.05	0.00	0.00
<i>P. fluorescens</i> JV449	0.27	0.00	0.03	0.00	0.04	0.00	0.03	0.00	0.00
<i>P. fluorescens</i> JV391D17	0.04	0.00	1.60	0.00	0.03	0.00	1.74	0.00	0.00
<i>P. wadsworthii</i> JV244B	0.01	0.00	0.08	0.00	0.06	0.00	0.07	0.00	0.00
<i>P. brassicacervum</i> 551A7	0.00	0.00	0.03	0.00	0.03	0.00	0.02	0.00	0.00
<i>P. protegens</i> JV245A	0.27	0.00	0.51	0.00	0.04	0.00	0.36	0.00	0.00
<i>P. protegens</i> JV359A1	0.29	0.04	0.50	0.04	0.09	0.04	0.38	0.04	0.04



**Figure 1: *Pseudomonas* strains relative abundance in each EE for A) the initial SynCom and for the *in planta* EE for SynCom 1, 2, 6 and 7 at the end of the 1<sup>st</sup> cycle and B) for the *in planta* EE for SynCom 1, 2, 6 and 7 at the end of the 27<sup>th</sup> cycle.** The proportions are expressed in percent and represent the number of reads that were assigned to specific reads of a given *Pseudomonas* strain, to multiple genomes (*i.e.* reads shared by a least two different *Pseudomonas*), ‘no hits’ (reads that did not match to any *Pseudomonas*) divided by the total number of sequences in each sample.

Further analysis revealed these reads correspond to maize genome but also to the genome of an endophyte of maize: *Pantoea ananatis* (Sheibani-Tezerji *et al.*, 2015). Presence of *P. ananatis* was validated during the isolation of evolved *Pseudomonas* variants at the end of the 27<sup>th</sup> cycle. As *rpoD* primers did not work on some colonies, these colonies were tested by *rrs* PCR and sequencing showed that they belonged to the *P. ananatis* species. The presence of this endophyte does not seem to have a major impact on the dynamics of the communities that underwent the *in planta* EE. Indeed, for the SynComs hosting *Pantoea* in cycle 27, SynComs 6-27 was dominated by *P. fluorescens* JV391D10 and SynCom 2-27 and 7-27 by *P. fluorescens* JV391D17, while no colony (nor metagenomic sequence read) of *P. ananatis* was found for SynCom 1-27 also dominated by *P. fluorescens* JV391D10.

The number of silent and non-silent (SNP or insertion of 1 or 2 nucleotides) mutations that could be detected among the specific sequence reads matching to genome of the 3 dominant strains was compared between SynComs from cycles 1 and 27 within metagenomes data (table S1-S3). The number of mutated genes and mutations (*i.e.* among specific reads) for

each SynCom are variable (Table 2). Overall, when combining all the mutation data observed at one cycle (fusion of all the EE) a higher number of new mutated genes are observed at the end of the 27<sup>th</sup> (23 mutated genes) vs. the end of the 1st cycle (2 mutated genes). At a global level (combining all the EE data) for SNP apparition during cycle 1 of the EE, no mutated genes were found for *P. fluorescens* JV391D10, whereas, 1 mutated gene was found for *P. fluorescens* JV391D17 and for *P. protegens* JV245A (table 3A and 3B). Regarding the SynComs from the 27<sup>th</sup> cycle (table 3C to 3E), 8 new mutated genes appear for *P. fluorescens* JV391D10, 7 for JV391D17, 8 for *P. protegens* JV245A.

**Table 2: Number of mutated genes predicted by Breseq software for strains *P. fluorescens* JV391D10, JV391D17, and *P. protegens* JV245A for SynComs retrieved at the end of the 1st and 27th cycle of EE.** Numbers represent the appearance of a new mutated gene

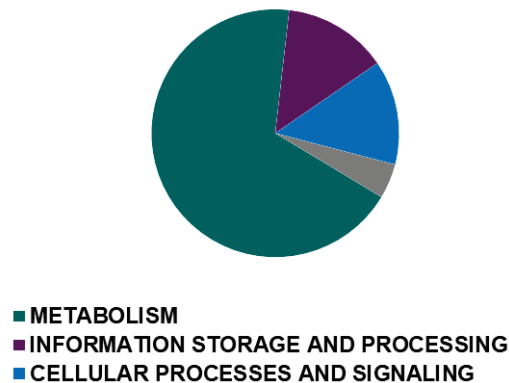
	number of mutated genes		
	JV391D10	JV391D17	JV245A
<b>SynCom 1-1</b>	-	-	-
<b>SynCom 2-1</b>	-	-	-
<b>SynCom 6-1</b>	-	-	2
<b>SynCom 7-1</b>	-	-	-
<b>SynCom 1-27</b>	-	1	-
<b>SynCom 2-27</b>	8	4	4
<b>SynCom 6-27</b>	-	2	1
<b>SynCom 7-27</b>	-	-	2

Several of these mutated genes are shared between different SynComs (Tables 3A to 3E) and silent and non-silent mutations are detected for each gene. A silent mutation in a coding sequence, although has no impact on the amino acid sequence, may impact translation if the bacteria do not code the tRNA corresponding to the new codon. A non-silent mutation changes the amino acid which has a direct impact on the protein and potentially on its activity, as well as on its structure. As a first step, we verified that all of the mutations were truly fixed; in other words, that the same non-silent mutation was predicted for all the available reads that mapped on that particular gene position (table S1, S2 and S3). This was not always the case, and these particular cases are discussed below.

**Table 3: Survey of mutated genes for strains *P. fluorescens* JV391D10, JV391D17 and *P. protegens* JV245A found in the SynComs retrieved at the end of the 1<sup>st</sup> cycle (A and B) and at the end of the 27<sup>th</sup> cycle (C, D and E). Numbers indicate the presence of a silent or non-silent mutation within a SynCom. S: silent mutation; NS: non-silent mutation**

<i>P. fluorescens</i> JV391D17 - 1st cycle										A			
GENE	LOCUS	PRODUCT	CLASS DESCRIPTION (COG)		EGGNOG	SYNCOM 1-1		SYNCOM 2-1		SYNCOM 6-1		SYNCOM 7-1	
						S	NS	S	NS	S	NS	S	NS
-	<a href="#">JV391D17_v1_10348</a>	DNA methyltransferase	Replication, recombination	L	INFORMATION STORAGE AND PROCESSING	1	-	-	-	-	-	1	-
<i>P. protegens</i> JV245A - 1st cycle										B			
GENE	LOCUS	PRODUCT	CLASS DESCRIPTION (COG)		EGGNOG	SYNCOM 1-1		SYNCOM 2-1		SYNCOM 6-1		SYNCOM 7-1	
						S	NS	S	NS	S	NS	S	NS
-	<a href="#">JV245A_v1_200008</a>	conserved hypothetical protein	Amino acid transport and metabolism	E	METABOLISM	-	-	-	-	20	2	15	10
<i>P. fluorescens</i> JV391D10 - 27th cycle										C			
GENE	LOCUS	PRODUCT	CLASS DESCRIPTION (COG)		EGGNOG	SYNCOM 1-27		SYNCOM 2-27		SYNCOM 6-27		SYNCOM 7-27	
						S	NS	S	NS	S	NS	S	NS
<i>ydhB</i>	<a href="#">JV391D10_v1_210016</a>	putative DNA-binding transcriptional regulator; LysR-type	Transcription	K	INFORMATION STORAGE AND PROCESSING	-	-	4	4	8	4	5	4
<i>accA</i>	<a href="#">JV391D10_v1_10828</a>	acetyl CoA carboxylase, carboxytransferase, alpha subunit	Lipid transport and metabolism	I	METABOLISM	-	-	25	16	-	-	25	16
<i>pepA</i>	<a href="#">JV391D10_v1_50070</a>	aminopeptidase A, a cyteinyglycinase	Amino acid transport and metabolism	E	METABOLISM	-	-	17	22	-	-	17	22
<i>proP</i>	<a href="#">JV391D10_v1_10040</a>	proline/glycine betaine transporter	Carbohydrate transport and metabolism	G	METABOLISM	-	-	1	1	-	-	-	-
<i>dmlR</i>	<a href="#">JV391D10_v1_10781</a>	D-malate degradation protein R	Transcription	K	INFORMATION STORAGE AND PROCESSING	-	-	23	2	-	-	23	2
<i>hrpA</i>	<a href="#">JV391D10_v1_90183</a>	ATP-dependent RNA helicase HrpA	Replication, recombination and repair	L	INFORMATION STORAGE AND PROCESSING	-	-	15	5	-	-	15	5
-	<a href="#">JV391D10_v1_40309</a>	Major/phage tail sheath protein	Function unknown	S	POORLY CHARACTERIZED	-	-	10	6	-	-	11	6
-	<a href="#">JV391D10_v1_10500</a>	conserved protein of unknown function	Function unknown	S	POORLY CHARACTERIZED	-	-	-	-	-	-	-	3
<i>P. fluorescens</i> JV391D17 - 27th cycle										D			
GENE	LOCUS	PRODUCT	CLASS DESCRIPTION (COG)		EGGNOG	SYNCOM 1-27		SYNCOM 2-27		SYNCOM 6-27		SYNCOM 7-27	
						S	NS	S	NS	S	NS	S	NS
<i>lon</i>	<a href="#">JV391D17_v1_70250</a>	DNA-binding ATP-dependent protease	Post translational modification, protein turnover, chaperones	O	CELLULAR PROCESSES AND SIGNALING	-	-	7	0	23	9	7	-
-	<a href="#">JV391D17_v1_20264</a>	Urea carboxylase	Lipid transport and metabolism	I	METABOLISM	-	-	14	5	14	5	14	7
-	<a href="#">JV391D17_v1_10344</a>	conserved protein of unknown function	Cell cycle control, cell division, chromosome partitioning	D	CELLULAR PROCESSES AND SIGNALING	2	-	-	-	-	-	-	-
<i>frmA</i>	<a href="#">JV391D17_v1_20124</a>	Alcohol dehydrogenase class 3 (Alcohol dehydrogenase class III) (S-(hydroxymethyl)glutathione dehydrogenase) (Glutathione-dependent formaldehyde dehydrogenase) (FDH) (FALDH)	Energy production and conversion	C	METABOLISM	-	-	4	-	-	-	-	-
<i>glpD</i>	<a href="#">JV391D17_v1_80020</a>	sn-glycerol-3-phosphate dehydrogenase, aerobic, FAD/NAD(P)-binding	Energy production and conversion	C	METABOLISM	-	-	1	1	-	-	-	-
-	<a href="#">JV391D17_v1_40123</a>	CAIB/BAIF family CoA transferase	Energy production and conversion	C	METABOLISM	-	-	-	-	14	7	-	-
-	<a href="#">JV391D17_v1_150113</a>	Urea ABC transporter permease	Amino acid transport and metabolism	E	METABOLISM	-	-	-	-	14	7	-	-
<i>P. protegens</i> JV245A - 27th cycle										E			
GENE	LOCUS	PRODUCT	CLASS DESCRIPTION (COG)		EGGNOG	SYNCOM 1-27		SYNCOM 2-27		SYNCOM 6-27		SYNCOM 7-27	
						S	NS	S	NS	S	NS	S	NS
<i>acnB</i>	<a href="#">JV245A_v1_260023</a>	aconitate hydratase 2	Energy production and conversion	C	METABOLISM	-	-	13	5	13	5	13	5
<i>aroP</i>	<a href="#">JV245A_v1_630056</a>	aromatic amino acid transporter	Amino acid transport and metabolism	E	METABOLISM	-	-	10	11	10	11	10	11
<i>betB</i>	<a href="#">JV245A_v1_180204</a>	NAD+-dependent betaine aldehyde dehydrogenase	Energy production and conversion	C	METABOLISM	-	-	13	13	13	13	13	16
<i>phaD</i>	<a href="#">JV245A_v1_260053</a>	putative K(+)H(+) antiporter subunit D	Inorganic ion transport and metabolism	P	METABOLISM	-	-	15	5	13	5	16	6
<i>pgdA</i>	<a href="#">JV245A_v1_70066</a>	Peptidoglycan deacetylase	Hydrate transport and metabolism	G	METABOLISM	-	-	13	8	13	8	14	8
<i>hisC</i>	<a href="#">JV245A_v1_270012</a>	Histidinol-phosphate aminotransferase 2	Amino acid transport and metabolism	E	METABOLISM	-	-	-	-	14	12	14	12
-	<a href="#">JV245A_v1_510013</a>	2,3-diketo-5-methylthio-1-phosphopentane phosphatase 5-	Amino acid transport and metabolism	E	METABOLISM	-	-	12	10	12	10	-	-
<i>metE</i>	<a href="#">JV245A_v1_560012</a>	methyltetrahydropteroyltryglutamate-	Amino acid transport and metabolism	E	METABOLISM	-	-	-	-	-	-	20	4

Mutated genes can be classified into 3 functional categories (EGGNOG): (1) Metabolism, (2) Information storage and processing, (3) Cellular processes and signalling. At the end of the 1<sup>st</sup> cycle, *P. protegens* JV245A presents 1 mutated gene and it seems to be involved in metabolism (JV245A\_v1\_200008). *P. fluorescens* JV391D17 only presents 1 mutated gene and it seems to be involved in information and storage. At the end of the 27<sup>th</sup> cycle, the 23 new mutated genes can be classified into all 3 categories (figure 2). Most of these genes (15 out of 23) are classified as involved in metabolism.



**Figure 2: Sector chart representing the proportions of mutated genes (23) at the end of the 27<sup>th</sup> cycle belonging to the 3 functional EGGNOG categories.**

### Evolution of genetic variations for each strain (comparative genomics)

For the 1<sup>st</sup> cycle, few genetic variations are observed. For *P. fluorescens* JV391D17, the locus JV391D17\_v1\_10348, coding for a DNA methyltransferase presents mutations in the SynCom 1-1 and 7-1, however, all of the mutations are silent (table 3C). For *P. protegens* JV245A (table 3D), SynCom 6-1 and 7-1 harbour several silent and non-silent mutations for the locus JV245A\_v1\_200008 which encodes a conserved protein of unknown function potentially involved in amino acid transport and metabolism (with FIG029529, extracellular solute-binding protein, family 3).

Few to none genetic variation was expected for the SynComs from the 1<sup>st</sup> cycle as these samples were retrieved after roughly 15 bacterial cell generations (=7 days of interaction with the plant). As a reference, in Lenski's long-term EE, mutation rate was calculated to be  $8.9 \times 10^{-11}$  per base-pair per generation in clones that had already evolved for 40 000 generations (Wielgoss *et al.*, 2011). After 27 cycles (> 400 generations) of *in planta* evolution, most of the mutations were distributed across SynComs 1-27, 2-27 and 7-27, while only 1 silent mutation was found in SynCom 1-27 (JV391D17\_v1\_10344). The main difference between SynCom 1-27 community composition and the other SynComs from the 27<sup>th</sup> cycle is the absence of the endophyte *P. ananatis*. These results may suggest that despite the fact that *P. ananatis* does not seem to affect the evolutive outcomes of the EE, it may have applied an additional selection pressure for SynCom 2-27, 6-27 and 7-27.

### *P. fluorescens* JV391D10

When comparing the *P. fluorescens* JV391D10 specific reads from the SynComs of the 27<sup>th</sup> cycle to the ancestral genome of *P. fluorescens* JV391D10 (inoculated into the initial SynCom), a total of 8 different genes were found mutated (table 3C). Only one gene (*ydhB*) is common to SynCom 2-27, 6-27 and 7-27. This gene encodes a transcriptional regulator of the LysR family, suggesting potential changes of gene expression compared to the ancestral strain in the evolved clones. Six other mutant genes are only shared between SynCom 2-27 and 7-27 which share the same evolutive outcome at the end of the *in planta* EE (*P. fluorescens* JV391D17 and *P. protegens* JV245A as the dominant strains). These genes are respectively coding for lipid, amino acid, or carbohydrate transporters (*accA*, *pepA* and *proP*), involved in transcription, replication, recombination and repair (*dmlR* and *hrpA*), and the locus JV391D10\_v1\_40309 (*gpf1*) encodes a major/phage tail sheath protein of *Myoviridae* phages. This protein can self-assemble into tubular structures. The gene is found in a region containing several other phage genes that may be implicated in the production of phage particles. SynCom 7-27 is the only sample to presents mutations on the loci JV391D10\_v1\_10500, encoding a conserved protein of unknown function.

To get deeper insights into the genetic changes, non-silent mutations were analysed for the genes that were shared between 2 or more SynCom with the same evolutive outcome or not. First, the gene *ydhB*, shared between SynCom 2-27, 6-27 and 7-27 (*P. ananatis* was detected in these 3 SynCom), presented 5 non-silent mutations in each SynCom (table S1A). Only the fifth mutation (at position 15.581) differed, being a substitution in SynCom 2-27 and an insertion of 2 bp for SynCom 6-27 and 7-27, which shifts the reading frame. The *ydhB* gene codes for a LysR-type transcriptional regulator, which is the most abundant family of regulators in the prokaryotic kingdom that are known for regulating different genes/operons involved in metabolisms, virulence, *quorum* sensing and motility (Maddocks and Oyston, 2008). In *ydhB* case, these non-silent mutations may condition the expression of genes necessary to adapt or to persist on the plant roots during the EE. Another study, demonstrated that in *E. coli*, *ydhB* binds to an upstream region of the *ydhC* gene, an adenosine transporter, in presence of adenine. Moreover, orthologs of these genes were identified in *Pseudomonas* spp. and it was hypothesised that they would have the same role. In plant cells, adenine nucleotides are present in high levels as they are used as co-factors in metabolism as well as in the synthesis of nucleic acids (Rita *et al.*, 2006). Adenine may perhaps be hijacked by *Pseudomonas* for their own metabolisms as adenine can be used as a source of ammonium (Rodionova *et al.*, 2020).

In JV391D10, genes *accA*, *pepA*, *dmlR*, and *hrpA* (Tables S1B, S1C, S1D and S1E) shared the same mutations between SynCom 2-27 and 7-27. In *Pseudomonas aeruginosa*, the type-II fatty acid synthesis (FAS) system is carried by the enzyme Acetyl-coA carboxylase (ACC)



which is encoded by the operon *accABCD* (Lee *et al.*, 2012). Fatty acid synthesis is a conserved process among all living organisms and its products are mainly used for membrane biosynthesis. However, the products of fatty acid synthesis can also play a role in bacterial *quorum* sensing and post-translational modification (Chan and Vogel, 2010). ACC catalyses the first step of fatty acid biosynthesis (Chen *et al.*, 2019) and mutations in *accA* could lead to a non-functional biosynthesis pathway of fatty acids. Regarding the gene *pepA*, it encodes an aminopeptidase in *E. coli* and even if its functions is not well characterized it is believed to be involved in the metabolism of peptides produced by protein degradation within the cell or externally supplied (Charlier *et al.*, 2000). Concerning *dmlR* and *hrpA*, both possess regulatory properties: *dmlR* is a LysR type transcriptional regulator required for *dmlA* expression in *E. coli*. The gene *dmlA* encodes a D-malate dehydrogenase which can convert D-malate to pyruvate (Lukas *et al.*, 2010), allowing growth under anaerobic conditions. In this study, strains developed on plants roots, where microaerophilic conditions could be encountered. The *hrpA* gene encodes an ATP-dependent RNA helicase, involved in mRNA processing in *E. coli* (Koo *et al.*, 2004). In general, RNA helicases are essential for most processes of RNA metabolism, and can regulate gene expression at a post-transcriptional level by i) associating with the RNA degradosome or by ii) un-winding secondary structure in the mRNAs of target genes (Butland *et al.*, 2005; Kaberdin and Bläsi, 2013; Vakulskas *et al.*, 2014). Furthermore, in *Pseudomonas aeruginosa*, another ATP-dependent RNA helicase (*deaD*) was shown to enhance at a post-transcriptional level the *exsA* expression, involved in the synthesis of the central regulator of T3SS (Intile *et al.*, 2015). *P. fluorescens* JV391D10 possess a T3SS, and there is evidence that it can be used in symbiotic interaction (Rezzonico *et al.*, 2005) to hijack host cells. A mutation on *hrpA* may lead to a less efficient activation of this secretion system.

Finally, SynCom 2-27 and 7-27, also share common non-silent mutations on the locus JV391D10\_v1\_40309 (encoding for a major/phage tail sheath protein). Two of these mutations differ, being insertions for SynCom 2-27 and substitutions for SynCom 7-27 (Table S1F). Furthermore, SynCom 7-27 presented one additional substitution inside this locus. Phages can impose significant mortality on their bacterial hosts but bacteria can evolve resistance to phage attack (Koskella and Brockhurst, 2014). The observed mutations (in JV391D10\_v1\_40309) could be a result of co-evolution of the bacterium with its phage.

For *P. fluorescens* JV391D10, a bigger number of mutated genes were found to be shared between SynCom 2-27 and 7-27, both these SynCom showed similar populations evolution with *P. fluorescens* JV391D17 and *P. protegens* JV245A as dominating strains (chapter 2). This community structuration with another *P. fluorescens* as the dominant strain may have imposed a selection pressure *P. fluorescens* JV391D10.

### *P. fluorescens* JV391D17

At the end of the 27<sup>th</sup> cycle, 7 mutated genes were found for *P. fluorescens* JV391D17 (table 3D) among the specific reads of the analysed SynComs when compared to the initial genome (inoculated into the original SynCom). Only 2 out of the 7 detected genes were common to SynCom 2-27, 6-27 and 7-27 (*P. ananatis* was detected in these 3 SynCom): *lon* and the locus JV391D17\_v1\_20264. The first is involved in post translational modification and the second is implicated in the metabolism (lipid transfer). All the other genes are only mutated in one of the analysed SynCom. For SynCom 1-27, the locus JV391D17\_v1\_10344, is involved in cell cycle control, cell division and chromosome portioning. Regarding the remaining genes, they all play roles in metabolism: *frmA* and *glpD* (energy production and conversion, SynCom 6-27) plus loci JV391D17\_v1\_40123 and JV391D17\_v1\_150113 (energy production and conversion and amino acid transport, SynCom 7-27). The detail of the non-silent mutations for these genes/locus are presented in tables S2.

In order to get deeper insights, we analysed the non-silent mutations of the genes shared between SynCom 2-27, 6-27 and 7-27. The *lon* gene only presents non-silent mutations (9 different) in SynCom 6-27 (table S2A). Among these 9 mutations, 5 of them are insertions and the other 4 are substitutions. Among the 5 insertions, all add 2 or 4 bp, shifting the reading-frame, except for the last one (position 258.845) which adds 3 bp. Only the substitution in position 258.800 is not fixed (98.9%). Lon protein is an ATP-dependant serine protease, widespread among bacteria. Best characterized in *E. coli*, this protein degrades abnormal, or non-functional intracellular proteins, as well as certain short-lived regulatory proteins (Gottesman and Maurizi, 1992; Gottesman, 1996; Lee and Suzuki, 2008). This highly conserved protease is important for cellular homeostasis and for survival from DNA damages and developmental changes induced by stress. Lon protein has also been reviewed in *P. syringe*, as being involved in regulation of the type III secretion system (T3SS), necessary for pathogenicity in plants. *P. syringe* Lon mutants rapidly elicited plant response and hyper secrete effectors proteins (Bretz *et al.*, 2002). Some of the strains used for the *in planta* EE, and particularly *P. fluorescens* JV391D17, harbour T3SS coding genes in its genome. Despite being well-known for plant pathogenicity, this molecular syringe can also be used in symbiotic interactions (Rezzonico *et al.*, 2005) to take over the host cells in a way that is beneficial for the bacteria. A mutation in this gene could lead to an over secretion of some effectors and allow for a more efficient root colonization of *P. fluorescens* JV391D17. Moreover, pyoluteorin production and stress response in *P. protegens* Pf-5 are controlled by global regulators (*gacA*, *gacS* and *rpoS*). Whistler *et al.* (2000) found that a homolog of *lon* can also control these functions in Pf-5. *P. fluorescens* JV391D17 does not harbour this antibiotic in its genome. Nonetheless, the association of antimicrobial compound production and stress responses are crucial for overcoming environmental stresses *in planta*.

Regarding the locus JV391D17\_v1\_20264, which encodes a urea carboxylase, involved in lipid transport and urea metabolism, there are 5 common non-silent mutations (substitution and insertions; table S2B) between SynComs 2-27, 6-27 and 7-27, plus, an additional substitution and insertion of 2 bp found in SynCom 7-27. Many microorganisms possess this enzyme which allows bacteria to consume urea as nitrogen source (Kanamori *et al.*, 2004). Urea can be found in plants as derivative from arginine metabolism (Witte, 2011). A modification in this gene could imply a different consumption of nitrogen sources.

Contrary to what was observed with *P. fluorescens* JV39D10, mutated genes in *P. fluorescens* JV39D17 do not seem to be shared between SynCom with the same evolutive outcomes (SynComs 1-27 and 6-27 plus SynComs 2-27 and 7-27). However, 2 mutated genes are common to SynCom 2-27, 6-27 and 7-27, where *P. ananatis* was recovered, suggesting that the endophyte might add a selection pressure. The other mutations seem to have appeared rather randomly and were fixed because of the advantages they may offer.

### *P. protegens* JV245A

When compared to the initial genomes, 8 mutated genes were found for *P. protegens* JV245A at the end of the 27<sup>th</sup> cycle (table 3E) among the specific reads from the metagenomics data of the SynComs. In all of the analysed SynComs, *P. fluorescens* JV245A was the second most abundant strain, far behind a dominating *P. fluorescens* strain that may have exerted a selection pressure. Therefore, we expected to find mutations common to all the SynCom. However, none of the 8 mutated genes resulting from the EE, was common to the 4 independent SynComs. But all 8 genes are involved in metabolism and 5 are common between SynCom 2-27, 6-27 and 7-27 (*P. ananatis* was detected in these SynComs): *acnB* and *betB* (energy production and conversion), *aroP*, *phaD* and *pgdA* (amino acid, ions and carbohydrate transport). Regarding the remaining genes: *hisC* (SynCom 6-27 and 7-27), the locus JV245A\_v1\_510013 (SynCom 2-27 and 6-27) and *metE* (SynCom 7-27) are all involved in amino acid transport.

There is no segregation of the mutated gene types among the SynComs with the same evolutive outcome (SynComs 1 and 6 in one hand and SynComs 2 and 7 in the other hand) as 5 out of 9 of the mutated genes are shared between SynCom 2-27, 6-27 and 7-27 (these SynComs host *P. ananatis*). Moreover, two genes (*hisC* and the locus JV245A\_v1\_510013) are shared between SynCom that presented different evolution *in planta*: *hisC* is shared between SynComs 6-27 (in which *P. fluorescens* JV391D10 was the dominant strain) and 7-27 (where *P. fluorescens* JV391D17 was the dominant strain). This combination of mutated genes in *P. protegens* JV245A may suggest that the evolutive outcome of the *in planta* EE with *P. fluorescens* JV391D10 or JV391D17 emerging as the dominant strain does not have great impact on the mutations observed for this strain.

To get deeper insights into the genomic changes, non-silent mutations were analysed for the genes that were shared between 2 or more SynCom. Among the genes shared by SynCom 2-27, 6-27 and 7-27, identical mutations were found for genes *acnB* and *aroP* (Tables S3A, and S3B). First, *acnB* is involved in energy production and conversion, and it encodes a cytosolic aconitase, an enzyme that plays a role in iron metabolism and virulence in *Bacillus* (Alen and Sonenshein, 1999). In *Pseudomonas aeruginosa*, a low level of iron inactivates aconitase activity, increasing the synthesis of its exotoxin A (its most powerful virulence factor). As *P. protegens* JV245A does not possess such toxin, a mutation on this gene could imply a modification on the iron metabolism, which is essential for growth in the rhizosphere. Regarding *aroP*, it encodes an aromatic amino acid transporter and the same 11 mutations are found in all 3 SynComs (table S3B). However, 8 out of 11 mutations seem not to be fixed in SynCom7-27. The genes *betB* and *phaD*, which are involved in energy production and in ion transport are also shared by SynCom 2-27, 6-27 and 7-27. However, both genes harbour additional non-silent mutations in SynCom 7-27: 3 for *betB* (table S3C) and 1 for *phaD* (table S3D). The gene *pgdA* (table S3E), is also mutated in SynCom 2-27, 6-27 and 7-27 and encodes for a peptidoglycan deacetylase, or more precisely, an *N*-acetylglucosamine deacetylase and presented the same mutations in all SynComs except for SynCom 7-27 which harbours one additional substitution (Table S3E). Peptidoglycan is a polymer formed of sugar and amino acids that forms an outside layer, known as the cell-wall (Rogers *et al.*, 1980). The sugars are composed of *N*-acetylglucosamine (NAG) and *N*-acetylmuramic acid (NAM). *N*-acetylglucosamine deacetylation has been reported to contribute to lysozyme and bacteriocins resistance (Benachour *et al.*, 2012; Grishin *et al.*, 2020). However, to our knowledge, this has only been proved in some gram-positive bacteria and *Pseudomonas* are gram-negative. In gram-negative bacteria, lysozyme resistance is more often mediated via specific inhibitory proteins that bind to the active site of peptidoglycan-degrading enzymes (Grishin *et al.*, 2020). Mutations in this gene could mean that these clones are less resistant to lysozymes, however, few information regarding this gene are found for gram-negative bacteria.

Next, SynCom 6-27 and 7-27 share the same mutations on the gene *hisC* (table S3F), which is part of the *his* operon and encodes for a histidine phosphate aminotransferase, responsible for the eighth step in histidine biosynthesis (Winkler and Ramos-Montañez, 2009). The histidine biosynthesis is an ancient metabolic pathway prior to the diversification of Bacteria, Archaea, and Eucarya. Histidine is a crucial metabolite for the cell functioning of many organisms as it is connected to the synthesis of purine nucleotides and the nitrogen metabolism (Sivaraman *et al.*, 2001). Finally, SynComs 2-27 and 6-27 share the same mutations for the locus JV245A\_v1\_510013 (table S3G), only the mutation in position 12.770 is not fixed. This locus encodes a 2,3-diketo-5-methylthio-1-phosphopentane phosphatase. These transmembranes proteins are involved in the methionine salvage pathway, used by

some bacteria to recycle the end-products of spermine metabolism and ethylene biosynthesis. In *Kelbsiella pneumonia*, these products are metabolized back to methionine (Heilbronn *et al.*, 1999; Dai *et al.*, 2001; Sekowska and Danchin, 2002). A functional methionine salvage pathway was experimentally demonstrated in *P. aeruginosa* (Sekowska *et al.*, 2004).

All the afore-mentioned genes are involved in different metabolic pathways. These mutations may indicate that *P. fluorescens* JV245A underwent a metabolic adaptation, perhaps, to best utilize the resources offered by the other strains and the plant. Overall, the evolutive outcome of the SynComs (*P. fluorescens* JV391D10 or JV391D17 as the dominant strain) do not seem to condition the appearance of mutations. On the other hand, *P. ananatis* seems to add a selection pressure as 5 out of the 7 mutated genes are shared between the SynComs (2-27, 6-27 and 7-27) hosting the endophyte.

### **Genome variations of variants isolated after the 27<sup>th</sup> cycle**

All of the evolved clones for *P. fluorescens* JV391D10 were isolated from SynCom 1-27 and 6-27, where this strain was dominant at the end of the 27<sup>th</sup> cycle. The same was done with *P. fluorescens* JV391D17. All evolved clones were isolated from SynCom 2-27 and 7-27, where this strain was dominant at the end of the EE. On the contrary, the evolved clones for *P. protegens* JV245A were isolated from all the SynComs.

Unfortunately, the results obtained for the genomic variations of the evolved of *P. fluorescens* JV391D10 and JV391D17 were not satisfactory. The predicted mutations for most of the genes presented 'N' instead of nucleotide substitution, insertion or deletion, implying that Breseq could not find the corresponding reference sequence in the ancestral genomes. This indicated that the genomes available on the MicroScope platform need to be re-sequenced. Indeed, the *Pseudomonas* used in the EE belong to a library of 698 *Pseudomonas* isolated from maize rhizosphere (Vacheron *et al.*, 2016). Some of the strains of this study (in particular, *P. fluorescens* JV391D10 and JV391D17) were sequenced prior to this thesis work. As these strains were isolated together there is a possibility that they were not properly isolated from one another before sequencing. In consequence, the results concerning the evolved variants of these strains are not presented.

### **Genomic variation for the evolved clones from JV245A**

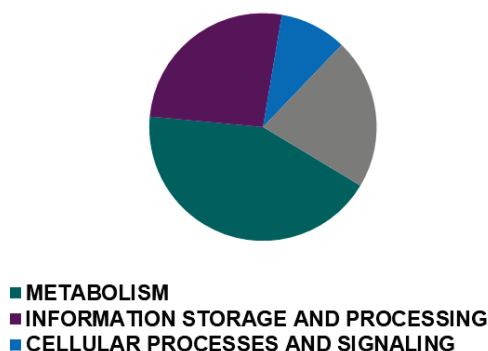
In total, 17 evolved clones of *P. protegens* JV245A were isolated and are named from A to Q. As *P. protegens* JV245A was the second dominant strain in all the SynCom, half of the clones (9) were isolated from SynCom 1-27 (A to C) and 5 from SynCom 6-27 (D to H), where both *P. fluorescens* JV391D10 and *P. protegens* JV245A were dominant. The remaining clones (8) were isolated from SynCom 2-27 (I to M) and SynCom 7-27 (N to Q), where *P. fluorescens* JV391D17 and *P. protegens* JV245A were the dominant strains. Each of these clones were



sequenced in a mixture composed of one clone of *P. protegens* JV245A and one of *P. fluorescens* JV391D10 (A to H) or one of *P. fluorescens* JV391D17 (I to Q). After sequencing, the reads from each mixture were separated using the hybrid assembly for *P. protegens* JV245A as reference. During this separation analysis, the reads corresponding to the core genome of the two strains could not be separated. Therefore, Breseq analysis was done only on the specific reads that were successfully classified for *P. protegens* JV245A.

The total number of mutated genes and mutations (silent and non-silent) varied among all analysed evolved clones (from 1 to 11, table 4). There were no genes in common with those found in the SynCom metagenomics data analysis. This may be due to the fact that for the Breseq carried on the metagenomics reads of the SynComs, an old genome version of *P. protegens* JV245 (available on MicroScope platform) was used. Whereas, the Breseq carried on the evolved clones used the hybrid assembly of *P. protegens* as a reference, which is of much better quality. In order to check this hypothesis, the Breseq carried on the metagenomics read of the SynComs is being re-run with the hybrid assembly of *P. protegens* JV245A as a reference (ongoing study).

In this genomics approach, a gene in a given genome is actually an allele of that gene in the evolved population. Also for the same SynComs, all sequenced clones must share exactly the same mutation to say that this mutation has been fixed. But here we could find that several mutations are not fixed in the strain population of the SynCom. As no mutation was common to all the evolved clones, we classified these clones based on the SynCom from which they were isolated (table 4). There were 42 mutated genes in total and also 4 intergenic mutations were detected (table 4 and 5). All the mutated genes and corresponding mutations could be classified into 3 functional categories (EGGNOG): (1) Metabolism, (2) Information storage and processing, (3) Cellular processes and signalling (Figure 3). Many of these genes (18 out of 42) are classified in the metabolism category. In the framework of this study we focused on the non-silent mutations detected for each clone in each of the afore mentioned COG categories.



**Figure 3: Sector chart representing the proportions of mutated genes (42) detected in all of the evolved clones of *P. protegens* JV245A at the end of the 27<sup>th</sup> cycle belonging to the 3 functional EGGNOG categories.**

**Table 4: Survey of the mutations found for the evolved clones of *P. protegens* JV245A. Numbers indicate the presence of a silent or non-silent mutation within a SynCom. SC: SynCom; S: silent; NS: non-silent.**

GENE	LOCUS	PRODUCT	CLASS DESCRIPTION (COG)	EGNOG	JV245A as the dominant strain at the end of the 27 cycle										JV245A as the dominant strain at the end of the 27 cycle									
					SC 1-27										SC 2-27									
					A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
					S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	T
<i>yefK_1</i>	<a href="#">JV245A.v1_670228</a>	Nucleoid-associated protein PFL_1090	Function unknown	S	POORLY CHARACTERIZED	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-	<a href="#">JV245A.v1_10142</a>	protein of unknown function	Function unknown	S	POORLY CHARACTERIZED	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-	<a href="#">JV245A.v1_50024</a>	protein of unknown function	Function unknown	S	POORLY CHARACTERIZED	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-	<a href="#">JV245A.v1_270149</a>	protein of unknown function	Function unknown	S	POORLY CHARACTERIZED	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-	<a href="#">JV245A.v1_310280</a>	protein of unknown function	Function unknown	S	POORLY CHARACTERIZED	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-	<a href="#">JV245A.v1_730065</a>	ABC-type ribitol:fructose:carbamate transport system substrate-binding protein	Function unknown	S	POORLY CHARACTERIZED	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>gda</i>	<a href="#">JV245A.v1_390071</a>	Cobalt transporter subunit CbaA	Function unknown	S	POORLY CHARACTERIZED	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-	<a href="#">JV245A.v1_150046</a>	protein of unknown function	Function unknown	S	POORLY CHARACTERIZED	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>hcp1</i>	<a href="#">JV245A.v1_370241</a>	Protein hcp1 (modular protein)	Function unknown	S	POORLY CHARACTERIZED	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>sdhD</i>	<a href="#">JV245A.v1_780020</a>	succinylglutamic semialdehyde dehydrogenase	Energy production and conversion	C	METABOLISM	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>ehfD</i>	<a href="#">JV245A.v1_200001</a>	electron transfer flavoprotein ubiquinone oxidoreductase	Energy production and conversion	C	METABOLISM	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>acdD</i>	<a href="#">JV245A.v1_210018</a>	Acetylaldehyde dehydrogenase 2	Energy production and conversion	C	METABOLISM	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>acd_2</i>	<a href="#">JV245A.v1_180171</a>	putative N-methylproline demethylase	Energy production and conversion	C	METABOLISM	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>gcbB</i>	<a href="#">JV245A.v1_180118</a>	malate synthase G	Energy production and conversion	C	METABOLISM	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>conA1_1</i>	<a href="#">JV245A.v1_810017</a>	Cob3 type cytochrome c oxidase subunit Con1	Energy production and conversion	C	METABOLISM	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>gah</i>	<a href="#">JV245A.v1_670124</a>	Leucine dehydrogenase	Amino acid transport and metabolism	E	METABOLISM	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-	<a href="#">JV245A.v1_170056</a>	fragment of putative methionine aminotransferase, PLP-dependent	Amino acid transport and metabolism	E	METABOLISM	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>pyrF</i>	<a href="#">JV245A.v1_470056</a>	oxindole 5-phosphate decarboxylase	Nucleotide transport and metabolism	F	METABOLISM	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>adp</i>	<a href="#">JV245A.v1_130008</a>	adenosine deaminase	Nucleotide transport and metabolism	F	METABOLISM	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>agaA</i>	<a href="#">JV245A.v1_670255</a>	N-acetylglucosamine 6-phosphate deacetylase	Carbohydrate transport and metabolism	G	METABOLISM	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>gpxX</i>	<a href="#">JV245A.v1_310116</a>	Glycogen operon protein GpxX	Carbohydrate transport and metabolism	G	METABOLISM	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>phoC</i>	<a href="#">JV245A.v1_670153</a>	Major phosphate-irrepressible acid phosphatase	Lipid transport and metabolism	I	METABOLISM	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>znhB_2</i>	<a href="#">JV245A.v1_780116</a>	zinc transport protein ZnhB	Inorganic ion transport and metabolism	P	METABOLISM	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>ygdD</i>	<a href="#">JV245A.v1_780007</a>	putative oxidoreductase YgdD	Inorganic ion transport and metabolism	P	METABOLISM	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>lgrD</i>	<a href="#">JV245A.v1_10217</a>	TP-dependent D-leucine adenylyase	Secondary metabolites biosynthesis, transport and catabolism	Q	METABOLISM	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>copA_2</i>	<a href="#">JV245A.v1_130021</a>	Copper exporting P-type ATPase A	Secondary metabolites biosynthesis, transport and catabolism	Q	METABOLISM	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-	<a href="#">JV245A.v1_120038</a>	4-hydroxyphenylacetate isomerase	Secondary metabolites biosynthesis, transport and catabolism	Q	METABOLISM	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>hrcC</i>	<a href="#">JV245A.v1_720021</a>	protein chain initiation factor IF1	Translation, ribosomal structure and biogenesis	J	INFORMATION STORAGE AND PROCESSING	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>cdhR_7</i>	<a href="#">JV245A.v1_180193</a>	HTH type transcriptional regulator CdhR	Transcription	K	INFORMATION STORAGE AND PROCESSING	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-	<a href="#">JV245A.v1_450048</a>	Transcriptional regulator, MerR family	Transcription	K	INFORMATION STORAGE AND PROCESSING	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-	<a href="#">JV245A.v1_380056</a>	Transcriptional regulator, LysR family	Transcription	K	INFORMATION STORAGE AND PROCESSING	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>lepA</i>	<a href="#">JV245A.v1_670240</a>	GTP binding protein	Cell wall/membrane/envelope biogenesis	M	CELLULAR PROCESSES AND SIGNALING	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-	<a href="#">JV245A.v1_670174</a>	putative alginate C-acylase AlgI	Cell wall/membrane/envelope biogenesis	M	CELLULAR PROCESSES AND SIGNALING	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>udg</i>	<a href="#">JV245A.v1_320161</a>	putative UDP-glucose 6-dehydrogenase	Cell wall/membrane/envelope biogenesis	M	CELLULAR PROCESSES AND SIGNALING	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>dnaJ</i>	<a href="#">JV245A.v1_130139</a>	curved DNA-binding protein, DnaJ homologue that functions as a co-chaperone of DnaK	Post-translational modification, protein turnover, chaperones	O	CELLULAR PROCESSES AND SIGNALING	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>gacA</i>	<a href="#">JV245A.v1_180045</a>	Two-component transcriptional regulator, LysR family	Signal transduction mechanisms	T	CELLULAR PROCESSES AND SIGNALING	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>acg_4</i>	<a href="#">JV245A.v1_100054</a>	Methyl-accepting chemotaxis protein	Signal transduction mechanisms	T	CELLULAR PROCESSES AND SIGNALING	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>norC_1</i>	<a href="#">JV245A.v1_630028</a>	Nitrogen assimilation regulatory protein NtrC	Signal transduction mechanisms	T	CELLULAR PROCESSES AND SIGNALING	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>rocC_5</i>	<a href="#">JV245A.v1_310009</a>	Vibrance sensor protein Bvg	Signal transduction mechanisms	T	CELLULAR PROCESSES AND SIGNALING	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>fecR</i>	<a href="#">JV245A.v1_570099</a>	Signal factor regulatory protein, FecR-PupR family	Signal transduction mechanisms	T	CELLULAR PROCESSES AND SIGNALING	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-	<a href="#">JV245A.v1_50018</a>	Signaling protein	Signal transduction mechanisms	T	CELLULAR PROCESSES AND SIGNALING	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>lsoD_2</i>	<a href="#">JV245A.v1_210014</a>	Uncharacterized ABC transporter ATP-binding protein	Defense mechanisms	V	CELLULAR PROCESSES AND SIGNALING	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-



### Intergenic mutations

A total of 4 intergenic mutations were detected for the 17 evolved clones of *P. protegens* JV245A. (table 5). The mutated intergenic regions concern genes involved in metabolism, information storage and processing or genes that were poorly characterized. Only 3 of them are shared between three or more clones and there seems to be a segregation for two of them (JV245A\_v1\_310226/JV245A\_v1\_310225 and JV245A\_v1\_540102 / JV245A\_v1\_540100) among the clones isolated from SynComs sharing the same evolutive outcome. The first intergenic mutation is common to all the evolved clones isolated from SynCom 1-27 (*i.e.* a fixed mutation) and is found between the loci JV245A\_v1\_310226 and JV245A\_v1\_310225. In all 3 clones it is a substitution in the same position (C→A). The locus JV245A\_v1\_310226 codes for an N-ethylmaleimide reductase, involved in energy production and conversion. The locus JV245A\_v1\_310225 corresponds to the gene *sdsB*, encoding for a transcriptional regulator belonging to the LysR family (Jovcic *et al.*, 2010), known for regulating the expression of several genes. This intergenic mutation could imply a differential expression of *sdsB*. Next, there is an identical mutation (insertion of CTC) concerning the evolved clones F (isolated from SynCom 6-27) and clones N and P (SynCom 7-27) in the intergenic region between the loci JV245A\_v1\_830015 and JV245A\_v1\_830013. The first correspond to the gene *ntaB* which encodes a nitrilotriacetate monooxygenase, one of the components required for degradation of nitriloacetate, which can be used as a source of carbon by some *Pseudomonas* (Tiedje *et al.*, 1973; Knobel *et al.*, 1996). Lastly, clones L (isolated from SynCom 2-27) and clones N, O, P, Q (isolated from SynCom 7-27) share a mutation (C → T) between the locus JV245A\_v1\_540102 and JV245A\_v1\_540100. The locus JV245A\_v1\_540102 encodes for an endoribonuclease HigB. Endoribonucleases participate in RNA maturation and degradation (Baek *et al.*, 2019). The other gene is not annotated on the MicroScope platform.

Intergenic regions have drawn attention in prokaryotes and structural non-coding RNAs (tRNA, rRNA and sRNA; Wassarman *et al.*, 2001; Sridhar *et al.*, 2011), as multiple roles in the survival of the cell have been identified, along with transposons (Siguier *et al.*, 2006), pseudogenes, inverted repeats (Sharples and Lloyd, 1990), integrase and transcription factors binding sites (Doublet *et al.*, 2008). Further analysis of this intergenic mutations are required in order to see if there is a modification of the expression of genes that are on each side.

### Poorly characterized category

There are 9 genes that were classified as poorly characterized (Table 4), none of them was fixed in for any SynCom (*i.e.* found in all the clones isolated from one SynCom).

Table 5: Intergenic mutations found in the evolved clones of *P. protegens* JV245A. Numbers indicate the presence of a silent or non-silent mutation within a SynCom. SC: SynCom; S: silent; NS: non-silent

GENE	LOCUS	PRODUCT	CLASS DESCRIPTION (COG)	EGNOCG	JV391D10 as the dominant strain at the end of the 27 cycle										JV391D17 as the dominant strain at the end of the 27 cycle																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																							
					SC 1-27					SC 2-27					SC 6-27					SC 7-27					SC 7-27																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																													
					A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	
					S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS</

Table 6: Non-silent mutations found for the *gacA*. Clones A, B and C were isolated from SynCom 1-27 whereas clones JLM were isolated from SynCom 2-27.

GENE	LOCUS	TYPE OF MUTATION	VARIANT 245A - A				VARIANT 245A - B				VARIANT 245A - C			
			POSITION	MUTATION	MUTATED READS (%)	ANNOTATION	POSITION	MUTATION	MUTATED READS (%)	ANNOTATION	POSITION	MUTATION	MUTATED READS (%)	ANNOTATION
<i>gacA</i>	JV245A v1. 180045	NS	3,261,368	G→T	100%	G81C (GGC→TGC)	3,261,368	G→T	100%	G81C (GGC→TGC)	3,261,368	G→T	100%	G81C (GGC→TGC)
GENE	LOCUS	TYPE OF MUTATION	VARIANT 245A - J				VARIANT 245A - L				VARIANT 245A - M			
			POSITION	MUTATION	MUTATED READS (%)	ANNOTATION	POSITION	MUTATION	MUTATED READS (%)	ANNOTATION	POSITION	MUTATION	MUTATED READS (%)	ANNOTATION
<i>gacA</i>	JV245A v1. 180045	NS	3,261,267	G→T	100%	L114F (TTG→TTT)	3,261,267	G→T	100%	L114F (TTG→TTT)	3,261,267	G→T	100%	L114F (TTG→TTT)

### Metabolism category

Among the 18 genes that were classified as involved in metabolism, 6 participate in energy production and conversion, 2 in amino acid transport, 2 in nucleotide transport, 1 in lipid transport, 2 in inorganic transport and 3 in secondary metabolites biosynthesis. None of these genes were mutated in more than one clone per SynCom. (=un-fixed mutation in the strain population).

### Information storage and processing category

For this category, only 4 genes were predicted as mutated (Table 4). Among them only 1 is involved in translation, ribosomal structure and biogenesis, whereas the other 3 are involved in transcription. All these mutated genes are only present in one evolved clone. Nevertheless, 2 of these genes are transcriptional regulators belonging to well-known families: LysR and MerR. LysR-type regulators influence the expression of different genes/operons involved in metabolisms, virulence, *quorum* sensing and motility (Maddocks and Oyston, 2008). MerR regulators are known to respond to environmental signals such as oxidative stress, heavy metals or antibiotics. In the context of our EE, the plant offers a changing environment to which these regulators can react. Mutations in these genes probably impact the expression of other genes that are not mutated, and this hypothesis requires further investigation. With the metagenomics analysis, another LysR-type regulator (encoded by *ydhB*) was found mutated in *P. fluorescens* JV39D10 in SynComs 2-27, 6-27 and 7-27. These results suggest that these regulators may play important roles in the evolved clones as, despite the strains or the population dynamics underwent by the SynComs, they are mutated. However, these mutations are not fixed in the isolated evolved clones of *P. protegens* JV245A.

### Cellular process and signalling category

There were 11 genes in this category; 3 were involved in cell wall/ membrane/ envelope biogenesis, 1 in post translational modification, 6 in signal transduction mechanisms and 1 in defence mechanisms. Among the genes involved in signal transduction there is the gene *gacA* which is mutated in 6 different clones. 3 of these clones (A, B and C) were isolated from SynCom 1-27 and the other 3 (J, L and M) from SynCom 2-27, suggesting that this mutation are fixed in the strain populations of these EE (SynCom 1 and 2). Each clone presents a non-silent mutation (table 6). In both cases it is a substitution, however it is a different mutation depending whether the clone comes from SynCom 1-27 or 2-27. Contrary to what was observed with the SynCom metagenome data, the evolutive outcome of the EE seems to have an impact on the observed mutations for *P. protegens* JV245. As a reminder, *P. fluorescens* JV391D10 is the dominant strain for SynCom 1-27 at the end of the 27<sup>th</sup> cycle, whereas it is JV391D17 for SynCom 2-27.

In *Pseudomonas*, *gacA* works with *gacS* to form a two-component system, where *gacS* encodes

a membrane-bound sensor kinase which upon activations, phosphorylates the response regulator GacA. The regulatory system controls the biosynthesis of various secondary metabolites (Cheng *et al.*, 2013). The GacS/GacA two-component system (Gac system) can also regulate the expression of several genes involved in virulence, biofilm formation, motility, *quorum* sensing, stress responses and survival (Kinscherf and Willis, 1999; Haas and Keel, 2003; Raaijmakers *et al.*, 2010; Yamazaki *et al.*, 2012). In *P. fluorescens*, activated GacA is known to enhance the transcription of 4 small RNA molecules: RsmX, RsmY, RsmZ and RgsA. RsmX, RsmY and RsmZ bind to transcriptional repressor proteins of the RsmA/CsrA family, RsmA and RsmE. This binding relieves translation repression by inducing the dissociation of these proteins from the Shine-Dalgarno sequences of their target genes (Heeb *et al.*, 2002; Valverde *et al.*, 2003; Kay *et al.*, 2005). Furthermore, the Gac system has been reported as capable of influencing iron metabolism in *P. protegens* Pf-5 (formerly *P. fluorescens* Pf-5) by enhancing transcript levels of siderophores-coding genes. Whereas, *P. protegens* Pf-5 mutants for *gacA* showed low level of transcripts of genes involved in production of hydrogen cyanide, of the antibiotic pyoluteorin and of the extracellular protease AprA (Hassan *et al.*, 2010). In *P. protegens* and *P. chlororaphis*, the Gac system also controls the production of 2,4- diacetylphloroglucinol (*phl* operon; Zhang *et al.*, 2020). All of the afore-mentioned plant-beneficial contributing genes are present in *P. protegens* JV245A genome and their expression is potentially modified by the predicted mutations.

Besides regulating antimicrobial production, Gac system has also been described as repressing motility in *P. kilonensis* F113 (formerly *P. fluorescens* F113) in such a way that F113 hypermotile mutants have been reported as more competitive in colonization (Barahona *et al.*, 2011). Furthermore, overexpression of 2 recombinases that control phenotypic variation in *P. kilonensis* F113 led to a large number of variants after selection by long-term cultivation or by rhizosphere passage. All the generated variants presented mutations either on *gacA* or/and *gacS* and were more motile than the wild-type strain, showing that mutations on the Gac system are selected in both conditions. Further tests revealed that these variants were more competitive than the wild-type strain in root colonization (Martínez-Granero *et al.*, 2006). This system has also been proved to act as a negative regulator of indole-3-acetic acid (IAA) production in *P. chlororaphis* (Kang *et al.*, 2006). Perhaps a mutation of the Gac system and the consequently potential increase in IAA production can also modulate gene expression in *Pseudomonas*.

Overall, Gac system has been reported as having a central role in the complex regulatory networks of *Pseudomonads*. Genome-wide transcriptomic analyses revealed this system can modify 10% to 15% of transcripts level (including numerous transcriptional genes). GacA seems to have the largest impact on transcript levels and it is likely to be mediated via other intermediate transcriptional regulators including *quorum* sensing circuits (Wang *et al.*, 2013; Wei *et al.*, 2013). In conclusion, the two-component systems regulate a wide variety of genes and the implication

of substitution in the evolved clones of *P. protegens* JV245A are likely to be more important and require further study.

## CONCLUDING REMARKS

Adaptation is the driving force of evolution and it happens via the fixation of mutations. Mutations happen via selection pressures and need a certain amount of generations to be fixed. In the present study, few mutations were observed at the end of the 1<sup>st</sup> cycle of evolution (≈15 bacterial generations), whereas several new mutations were detected at the end of the 27<sup>th</sup> cycle (≈400 bacterial generations). These results imply that 27 successive cycles of evolution are long enough for mutations to rise and settle within the population. Most of the genetic changes found (for the metagenomics data or the evolved clones) concern mostly genes involved in the metabolism. This suggests that the selection pressures undergone during the EE caused the strains to adapt to a new environment: the roots of maize. Although all the strains from this study were initially isolated from maize rhizosphere, our EE setup is quite different from natural rhizosphere: the only nutrient supply comes from maize, whereas in natural environments, soil offers a richer environment. An additional selection pressure seems to be imposed by *P. ananatis*, as most of the mutations found within the SynComs are distributed across the SynComs that harbour this endophyte (SynCom 2-27, 6-27 and 7-27).

On the other hand, the evolutive outcome of the SynComs at the end of the 27<sup>th</sup> cycle (*P. fluorescens* JV391D10 or JV391D17 as the dominant strain) does not seem to impose a selective pressure. Only *P. fluorescens* JV391D10 presents a segregation of its mutations in SynComs sharing the same population dynamics (SynCom 2-27 and 7-27). However, when analysing the evolved clones of *P. protegens* JV245A this segregation was clearly found for the gene *gacA* (Gac System). Besides the *gacA*, other genes involved in transcription and translation are also mutated in both approaches particularly, LysR-type regulators. The next step would be to study the expression of the regulated genes to see the actual effect of these mutations by assessing the fitness of the evolved clones *in planta*. Studies similar to ours have showed that in EE not many genetic variations are detected. However, when a regulator is mutated it often implies a modification on the transcription of other genes belonging to their regulon (Guidot *et al.*, 2014; Perrier *et al.*, 2016).

## ACKNOWLEDGEMENTS

Beatriz Manriquez was supported by a Ph.D. fellowship from the French Ministère de l'Education Nationale, de l'Enseignement Supérieur et de la Recherche. The authors acknowledge funding from the French National Programme EC2CO 'Structuring initiative Continental and Coastal Ecosphere' - Microbien 2018-2019. We are grateful to Jeanne Doré for technical help, as well as Denis Faure (I2BC, France) for helpful discussion. We also acknowledge the LABGeM and the National Infrastructure "France Génomique" for support within the MicroScope annotation platform. The authors declare no conflict of interest.

## DATA ACCESSIBILITY

The ancestral genomes have been deposited in the European Bioinformatics Institute (EBI) database: PRJNA572594; PRJNA572595; PRJNA566351; PRJNA566305; PRJNA572600; PRJNA566269; PRJNA572570; PRJNA566300 and PRJNA572600. The metagenomes and evolved genomes have not been deposited yet.

## AUTHOR CONTRIBUTIONS

BM, DM and CPC designed the project, BM and JD realized the molecular extractions, AP and DA implemented bioinformatic analyses. BM, DM, DA and CPC analyzed data; BM, DM and CPC prepared the first draft of the manuscript, which was finalized by all authors.

## REFERENCES

- Alen, C., and Sonenshein, A. L. (1999). *Bacillus subtilis* aconitase is an RNA-binding protein. *Proc. Natl. Acad. Sci.* 96, 10412–10417. doi:10.1073/pnas.96.18.10412.
- Baek, Y. M., Jang, K.-J., Lee, H., Yoon, S., Baek, A., Lee, K., *et al.*, (2019). The bacterial endoribonuclease RNase E can cleave RNA in the absence of the RNA chaperone Hfq. *Journal of Biological Chemistry* 294, 16465–16478. doi:10.1074/jbc.RA119.010105.
- Barahona, E., Navazo, A., Martínez-Granero, F., Zea-Bonilla, T., Pérez-Jiménez, R. M., Martín, M., *et al.*, (2011). *Pseudomonas fluorescens* F113 mutant with enhanced competitive colonization ability and improved biocontrol activity against fungal root pathogens. *American Society for Microbiology*. doi:10.1128/AEM.00320-11
- Battistuzzi, F. U., and Hedges, S. B. (2009). A major clade of prokaryotes with ancient adaptations to life on land. *Molecular Biology and Evolution* 26, 335–343. doi:10.1093/molbev/msn247.
- Benachour, A., Ladjouzi, R., Jeune, A. L., Hébert, L., Thorpe, S., Courtin, P., *et al.*, (2012). The lysozyme-induced peptidoglycan N-acetylglucosamine deacetylase PgdA (EF1843) is required for *Enterococcus faecalis* virulence. *Journal of Bacteriology* 194, 6066–6073. doi:10.1128/JB.00981-12.
- Bianco, C., Imperlini, E., Calogero, R., Senatore, B., Amoresano, A., Carpentieri, A., *et al.*, (2006). Indole-3-acetic acid improves *Escherichia coli*'s defences to stress. *Arch Microbiol* 185, 373–382. doi:10.1007/s00203-006-0103-y.
- Bretz, J., Losada, L., Lisboa, K., and Hutcheson, S. W. (2002). Lon protease functions as a negative regulator of type III protein secretion in *Pseudomonas syringae*. *Mol Microbiol* 45, 397–409. doi:10.1046/j.1365-2958.2002.03008.x.
- Brockhurst, M. A., Morgan, A. D., Rainey, P. B., and Buckling, A. (2003). Population mixing accelerates coevolution: Population mixing accelerates coevolution. *Ecology Letters* 6, 975–979. doi:10.1046/j.1461-0248.2003.00531.x.



- Bruto, M., Prigent-Combaret, C., Muller, D., and Moëne-Loccoz, Y. (2014). Analysis of genes contributing to plant-beneficial functions in plant growth-promoting rhizobacteria and related Proteobacteria. *Scientific Reports* 4, 6261. doi:10.1038/srep06261.
- Buckling, A., and Rainey, P. B. (2002). Antagonistic coevolution between a bacterium and a bacteriophage. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 269, 931–936. doi:10.1098/rspb.2001.1945.
- Butland, G., Peregrín-Alvarez, J. M., Li, J., Yang, W., Yang, X., Canadien, V., et al. (2005). Interaction network containing conserved and essential protein complexes in *Escherichia coli*. *Nature* 433, 531–537. doi:10.1038/nature03239.
- Chan, D. I., and Vogel, H. J. (2010). Current understanding of fatty acid biosynthesis and the acyl carrier protein. *Biochemical Journal* 430, 1–19. doi:10.1042/BJ20100462.
- Charlier, D., Kholti, A., Huysveld, N., Gigot, D., Maes, D., Thia-Toong, T.-L., et al. (2000). Mutational analysis of *Escherichia coli* PepA, a multifunctional DNA-binding aminopeptidase. *Journal of Molecular Biology* 302, 409–424. doi:10.1006/jmbi.2000.4067.
- Chen, D., Yuan, X., Liang, L., Liu, K., Ye, H., Liu, Z., et al. (2019). Overexpression of acetyl-CoA carboxylase increases fatty acid production in the green alga *Chlamydomonas reinhardtii*. *Biotechnol Lett* 41, 1133–1145. doi:10.1007/s10529-019-02715-0.
- Cheng, X., Bruijn, I. de, Voort, M. van der, Loper, J. E., and Raaijmakers, J. M. (2013). The Gac regulon of *Pseudomonas fluorescens* SBW25. *Environmental Microbiology Reports* 5, 608–619. doi:https://doi.org/10.1111/1758-2229.12061.
- Dai, Y., Pochapsky, T. C., and Abeles, R. H. (2001). Mechanistic studies of two dioxygenases in the methionine salvage pathway of *Klebsiella pneumoniae*. *Biochemistry* 40, 6379–6387. doi:10.1021/bi010110y.
- Deatherage, D. E., Traverse, C. C., Wolf, L. N., and Barrick, J. E. (2015). Detecting rare structural variation in evolving microbial populations from new sequence junctions using Breseq. *Front. Genet.* 5. doi:10.3389/fgene.2014.00468.
- Doublet, B., Golding, G. R., Mulvey, M. R., and Cloeckert, A. (2008). Secondary chromosomal attachment site and tandem integration of the mobilizable salmonella genomic island 1. *PLoS ONE* 3, e2060. doi:10.1371/journal.pone.0002060.
- Elena, S. F., and Lenski, R. E. (2003). Microbial genetics: evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. *Nature Reviews Genetics* 4, 457–469. doi:10.1038/nrg1088.
- Fisher, R. A. (1930). The genetical theory of natural selection. Oxford, UK: Oxford Univ. Press.
- Gottesman, S. (1996). Proteases and their targets in *Escherichia coli*. *Annual Review of Genetics* 30, 465–506. doi:10.1146/annurev.genet.30.1.465.
- Gottesman, S., and Maurizi, M. R. (1992). Regulation by proteolysis: energy-dependent proteases and their targets. *Microbiology and Molecular Biology Reviews* 56, 592–621.
- Grishin, A. V., Karyagina, A. S., Vasina, D. V., Vasina, I. V., Gushchin, V. A., and Lunin, V. G. (2020). Resistance to peptidoglycan-degrading enzymes. *Critical Reviews in Microbiology* 46, 703–726. doi:10.1080/1040841X.2020.1825333.
- Guan, S. H., Gris, C., Cruveiller, S., Pouzet, C., Tasse, L., Leru, A., et al. (2013). Experimental evolution of nodule intracellular infection in legume symbionts. *The ISME Journal* 7, 1367–1377. doi:10.1038/ismej.2013.24.
- Guidot, A., Jiang, W., Ferdy, J.-B., Thébaud, C., Barberis, P., Gouzy, J., et al. (2014). Multihost experimental evolution of the pathogen *Ralstonia solanacearum* unveils genes involved in adaptation to plants. *Molecular Biology and Evolution* 31, 2913–2928. doi:10.1093/molbev/msu229.
- Haas, D., and Keel, C. (2003). Regulation of antibiotic production in root-colonizing *Pseudomonas* spp. and relevance for biological control of plant disease. *Annu. Rev. Phytopathol.* 41, 117–153. doi:10.1146/annurev.phyto.41.052002.095656.
- Hassan, K. A., Johnson, A., Shaffer, B. T., Ren, Q., Kidarsa, T. A., Elbourne, L. D. H., et al. (2010). Inactivation of the GacA response regulator in *Pseudomonas fluorescens* Pf-5 has far-reaching transcriptomic consequences. *Environmental Microbiology* 12, 899–915. doi:https://doi.org/10.1111/j.1462-2920.2009.02134.x.
- Heeb, S., Blumer, C., and Haas, D. (2002). Regulatory RNA as mediator in GacA/RsmA-dependent global control of exoproduct formation in *Pseudomonas fluorescens* CHA0. *Journal of Bacteriology* 184, 1046–1056. doi:10.1128/jb.184.4.1046-1056.2002.



- Heilbronn, J., Wilson, J., and Berger, B. J. (1999). Tyrosine aminotransferase catalyzes the final step of methionine recycling in *Klebsiella pneumoniae*. *Journal of Bacteriology* 181, 1739–1747. doi:10.1128/JB.181.6.1739-1747.1999.
- Intile, P. J., Balzer, G. J., Wolfgang, M. C., and Yahr, T. L. (2015). The RNA helicase DeaD stimulates ExsA translation to promote expression of the *Pseudomonas aeruginosa* type III secretion system. *Journal of Bacteriology* 197, 2664–2674. doi:10.1128/JB.00231-15.
- Jovcic, B., Venturi, V., Davison, J., Topisirovic, L., and Kojic, M. (2010). Regulation of the sdsA alkyl sulfatase of *Pseudomonas* sp. ATCC19151 and its involvement in degradation of anionic surfactants. *Journal of Applied Microbiology* 109, 1076–1083. doi:https://doi.org/10.1111/j.1365-2672.2010.04738.x.
- Kaberlin, V. R., and Bläsi, U. (2013). Bacterial helicases in post-transcriptional control. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* 1829, 878–883. doi:10.1016/j.bbapap.2012.12.005.
- Kanamori, T., Kanou, N., Atomi, H., and Imanaka, T. (2004). Enzymatic characterization of a prokaryotic urea carboxylase. *Journal of Bacteriology* 186, 2532–2539. doi:10.1128/JB.186.9.2532-2539.2004.
- Kang, B. R., Yang, K. Y., Cho, B. H., Han, T. H., Kim, I. S., Lee, M. C., et al. (2006). Production of indole-3-acetic acid in the plant-beneficial strain *pseudomonas chlororaphis* O6 is negatively regulated by the global sensor kinase GacS. *Curr Microbiol* 52, 473–476. doi:10.1007/s00284-005-0427-x.
- Kay, E., Dubuis, C., and Haas, D. (2005). Three small RNAs jointly ensure secondary metabolism and biocontrol in *Pseudomonas fluorescens* CHA0. *Proc. Natl. Acad. Sci.* 102, 17136–17141. doi:10.1073/pnas.0505673102.
- Kinscherf, T. G., and Willis, D. K. (1999). Swarming by *Pseudomonas syringae* B728a requires gacS (IemA) and gacA but not the acyl-homoserine lactone biosynthetic GeneAII. *Journal of Bacteriology* 181, 4133–4136. doi:10.1128/JB.181.13.4133-4136.1999.
- Knobel, H. R., Egli, T., and van der Meer, J. R. (1996). Cloning and characterization of the genes encoding nitrilotriacetate monooxygenase of *Chelatobacter heintzii* ATCC 29600. *Journal of bacteriology* 178, 6123–6132. doi:10.1128/JB.178.21.6123-6132.1996.
- Koo, J. T., Choe, J., and Moseley, S. L. (2004). HrpA, a DEAH-box RNA helicase, is involved in mRNA processing of a fimbrial operon in *Escherichia coli*. *Molecular Microbiology* 52, 1813–1826. doi:https://doi.org/10.1111/j.1365-2958.2004.04099.x.
- Koskella, B., and Brockhurst, M. A. (2014). Bacteria–phage coevolution as a driver of ecological and evolutionary processes in microbial communities. *FEMS Microbiol Rev* 38, 916–931. doi:10.1111/1574-6976.12072.
- Lee, I., and Suzuki, C. K. (2008). Functional mechanics of the ATP-dependent Lon protease-lessons from endogenous protein and synthetic peptide substrates. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* 1784, 727–735. doi:10.1016/j.bbapap.2008.02.010.
- Lee, J.-H., and Lee, J. (2010). Indole as an intercellular signal in microbial communities. *FEMS Microbiol Rev* 34, 426–444. doi:10.1111/j.1574-6976.2009.00204.x.
- Lee, S., Jeon, E., Jung, Y., and Lee, J. (2012). Heterologous Co-expression of *accA*, *fabD*, and thioesterase genes for Improving long-chain fatty acid production in *Pseudomonas aeruginosa* and *Escherichia coli*. *Appl Biochem Biotechnol* 167, 24–38. doi:10.1007/s12010-012-9644-5.
- Lenski, R. E., Rose, M. R., Simpson, S. C., and Tadler, S. C. (1991). Long-term experimental evolution in *Escherichia coli*. I. Adaptation and divergence during 2,000 generations. *The American Naturalist* 138, 1315–1341. doi:10.1086/285289.
- Lukas, H., Reimann, J., Kim, O. B., Grimpö, J., and Uden, G. (2010). Regulation of aerobic and anaerobic D-malate metabolism of *Escherichia coli* by the LysR-Type regulator DmlR (YeaT). *Journal of Bacteriology* 192, 2503–2511. doi:10.1128/JB.01665-09.
- Maddocks, S. E., and Oyston, P. C. F. (2008). Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins. *Microbiology* 154, 3609–3623. doi:10.1099/mic.0.2008/022772-0.
- Martínez-Granero, F., Rivilla, R., and Martín, M. (2006). Rhizosphere selection of highly motile phenotypic variants of *Pseudomonas fluorescens* with enhanced competitive colonization ability. *Appl. Environ. Microbiol.* 72, 3429–3434. doi:10.1128/AEM.72.5.3429-3434.2006.
- Perrier, A., Peyraud, R., Rengel, D., Barlet, X., Lucasson, E., Gouzy, J., et al. (2016). Enhanced *in planta* fitness through adaptive mutations in EfpR, a dual regulator of virulence and metabolic functions in the plant pathogen *Ralstonia*

*solanacearum*. *PLoS Pathogens* 12, e1006044. doi:10.1371/journal.ppat.1006044

Raaijmakers, J. M., De Bruijn, I., Nybroe, O., and Ongena, M. (2010). Natural functions of lipopeptides from *Bacillus* and *Pseudomonas*: more than surfactants and antibiotics. *FEMS Microbiol Rev* 34, 1037–1062. doi:10.1111/j.1574-6976.2010.00221.x.

Rezzonico, F., Binder, C., Défago, G., and Moënné-Loccoz, Y. (2005). The type III secretion system of biocontrol *Pseudomonas fluorescens* KD targets the phytopathogenic chromista *Pythium ultimum* and promotes cucumber protection. *Molecular Plant-Microbe Interactions* 18, 991–1001. doi:10.1094/MPMI-18-0991.

Rita, Z., Mark, S., Uwe, S., and Ralf, B. (2006). Pyrimidine and purine biosynthesis and degradation in plants. doi:10.1146/annurev.arplant.57.032905.105421.

Rodionova, I. A., Gao, Y., Sastry, A., Yoo, R., Rodionov, D. A., Saier, M. H., *et al.* (2020). Synthesis of the novel transporter YdhC, is regulated by the YdhB transcription factor controlling adenosine and adenine uptake. *bioRxiv*, 2020.05.03.074617. doi:10.1101/2020.05.03.074617.

Rogers, H. J., Perkins, H. R., and Ward, J. B. (1980). "Biosynthesis of peptidoglycan" in *Microbial Cell Walls and Membranes*, eds. H. J. Rogers, H. R. Perkins, and J. B. Ward (Dordrecht: Springer Netherlands), 239–297. doi:10.1007/978-94-011-6014-8\_8.

Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30, 2068–2069. doi:10.1093/bioinformatics/btu153.

Sekowska, A., and Danchin, A. (2002). The methionine salvage pathway in *Bacillus subtilis*. *BMC Microbiol* 2, 8. doi:10.1186/1471-2180-2-8.

Sekowska, A., Dénervaud, V., Ashida, H., Michoud, K., Haas, D., Yokota, A., *et al.* (2004). Bacterial variations on the methionine salvage pathway. *BMC Microbiol* 4, 9. doi:10.1186/1471-2180-4-9.

Sharples, G. J., and Lloyd, R. G. (1990). A novel repeated DNA sequence located in the intergenic regions of bacterial chromosomes. *Nucleic Acids Res* 18, 6503–6508. doi:10.1093/nar/18.22.6503.

Sheibani-Tezerji, R., Naveed, M., Jehl, M.-A., Sessitsch, A., Rattei, T., and Mitter, B. (2015). The genomes of closely related *Pantoea ananatis* maize seed endophytes having different effects on the host plant differ in secretion system

genes and mobile genetic elements. *Front. Microbiol.* 6. doi:10.3389/fmicb.2015.00440.

Siguier, P., Filée, J., and Chandler, M. (2006). Insertion sequences in prokaryotic genomes. *Current Opinion in Microbiology* 9, 526–531. doi:10.1016/j.mib.2006.08.005.

Sivaraman, J., Li, Y., Larocque, R., Schrag, J. D., Cygler, M., and Matte, A. (2001). Crystal structure of histidinol phosphate aminotransferase (HisC) from *Escherichia coli*, and its covalent complex with pyridoxal-5'-phosphate and l-histidinol phosphate. *Journal of Molecular Biology* 311, 761–776. doi:10.1006/jmbi.2001.4882.

Sridhar, J., Sabarinathan, R., Balan, S. S., Rafi, Z. A., Gunasekaran, P., and Sekar, K. (2011). Junker: an intergenic explorer for bacterial genomes. *Genomics Proteomics Bioinformatics* 9, 179–182. doi:10.1016/S1672-0229(11)60021-1.

Strotz, L. C., Simões, M., Girard, M. G., Breitreuz, L., Kimmig, J., and Lieberman, B. S. (2018). Getting somewhere with the Red Queen: chasing a biologically modern definition of the hypothesis. *Biology Letters* 14, 20170734. doi:10.1098/rsbl.2017.0734.

Tenaillon, O. (2014). The Utility of Fisher's geometric model in evolutionary genetics. *Annual Review of Ecology, Evolution, and Systematics* 45, 179–201. doi:10.1146/annurev-ecolsys-120213-091846.

Tiedje, J. M., Mason, B. B., Warren, C. B., and Malec, E. J. (1973). Metabolism of nitrilotriacetate by cells of *Pseudomonas* species. *Applied Microbiology* 25, 811–818. doi:10.1128/AEM.25.5.811-818.1973.

Vacheron, J., Moënné-Loccoz, Y., Dubost, A., Gonçalves-Martins, M., Muller, D., and Prigent-Combaret, C. (2016). Fluorescent *Pseudomonas* strains with only few plant-beneficial properties are favored in the maize rhizosphere. *Front. Plant Sci.* 7. doi:10.3389/fpls.2016.01212.

Vakulskas, C. A., Pannuri, A., Cortés-Selva, D., Zere, T. R., Ahmer, B. M., Babitzke, P., *et al.* (2014). Global effects of the DEAD-box RNA helicase DeaD (CsdA) on gene expression over a broad range of temperatures. *Molecular Microbiology* 92, 945–958. doi:https://doi.org/10.1111/mmi.12606.

Vallenet, D., Calteau, A., Dubois, M., Amours, P., Bazin, A., Beuvin, M., *et al.* (2020). MicroScope: an integrated platform for the annotation and exploration of microbial gene functions through genomic, pangenomic and metabolic

comparative analysis. *Nucleic Acids Res* 48, D579–D589. doi:10.1093/nar/gkz926.

Valverde, C., Heeb, S., Keel, C., and Haas, D. (2003). RsmY, a small regulatory RNA, is required in concert with RsmZ for GacA-dependent expression of biocontrol traits in *Pseudomonas fluorescens* CHA0. *Molecular Microbiology* 50, 1361–1379. doi:https://doi.org/10.1046/j.1365-2958.2003.03774.x.

Van Valen, L. (1977). The Red Queen. *The American Naturalist* 111, 809–810. doi:10.1086/283213.

Wang, D., Lee, S.-H., Seeve, C., Yu, J. M., Pierson, L. S., and Pierson, E. A. (2013). Roles of the Gac-Rsm pathway in the regulation of phenazine biosynthesis in *Pseudomonas chlororaphis* 30-84. *MicrobiologyOpen* 2, 505–524. doi:https://doi.org/10.1002/mbo3.90.

Wassarman, K. M., Repoila, F., Rosenow, C., Storz, G., and Gottesman, S. (2001). Identification of novel small RNAs using comparative genomics and microarrays. *Genes Dev.* 15, 1637–1651. doi:10.1101/gad.901001.

Wei, X., Huang, X., Tang, L., Wu, D., and Xu, Y. (2013). Global control of GacA in secondary metabolism, primary metabolism, secretion systems, and motility in the rhizobacterium *Pseudomonas aeruginosa* M18. *Journal of Bacteriology* 195, 3387–3400. doi:10.1128/JB.00214-13.

Whistler, C. A., Stockwell, V. O., and Loper, J. E. (2000). Lon protease influences antibiotic production and UV tolerance of *Pseudomonas fluorescens* Pf-5. *Appl Environ Microbiol* 66, 2718–2725.

Wick, R. R., Judd, L. M., Gorrie, C. L., and Holt, K. E. (2017). Unicycler: Resolving bacterial

genome assemblies from short and long sequencing reads. *PLoS Computational Biology* 13, e1005595. doi:10.1371/journal.pcbi.1005595.

Wielgoss, S., Barrick, J. E., Tenaillon, O., Cruveiller, S., Chane-Woon-Ming, B., Médigue, C., et al. (2011). Mutation rate inferred from synonymous substitutions in a long-term evolution experiment with *Escherichia coli*. *G3&#58; Genes|Genomes|Genetics* 1, 183–186. doi:10.1534/g3.111.000406.

Winkler, M. E., and Ramos-Montañez, S. (2009). Biosynthesis of histidine. *EcoSal Plus* 3. doi:10.1128/ecosalplus.3.6.1.9.

Wisniewski-Dyé, F., Borziak, K., Khalsa-Moyers, G., Alexandre, G., Sukharnikov, L. O., Wuichet, K., et al. (2011). *Azospirillum* genomes reveal transition of bacteria from aquatic to terrestrial environments. *PLoS Genetics* 7, e1002430. doi:10.1371/journal.pgen.1002430.

Witte, C.-P. (2011). Urea metabolism in plants. *Plant Science* 180, 431–438. doi:10.1016/j.plantsci.2010.11.010.

Yamazaki, A., Li, J., Zeng, Q., Khokhani, D., Hutchins, W. C., Yost, A. C., et al. (2012). Derivatives of plant phenolic compound affect the type III secretion system of *Pseudomonas aeruginosa* via a GacS-GacA two-component signal transduction system. *Antimicrobial Agents and Chemotherapy* 56, 36–43. doi:10.1128/AAC.00732-11.

Zhang, Y., Zhang, B., Wu, X., and Zhang, L.-Q. (2020). Characterization the role of GacA-dependent small RNAs and RsmA family proteins on 2,4-diacetylphloroglucinol production in *Pseudomonas fluorescens* 2P24. *Microbiological Research* 233, 126391. doi:10.1016/j.micres.2019.126391.

SUPPLEMENTARY DATA

Table S1: Non-silent mutations found for *P. fluorescens* JV391D10 at the end of the 27<sup>th</sup> cycle. Only the genes that were shared between 2 or more SynCom are presented.

I	SynCom 2-27						SynCom 6-27						SynCom 7-27						A
	GENE	LOCUS	TYPE OF MUTATION	POSITION	MUTATION	ANNOTATION	MUTATED READS (%)	POSITION	MUTATION	ANNOTATION	MUTATED READS (%)	POSITION	MUTATION	ANNOTATION	MUTATED READS (%)	POSITION	MUTATION	ANNOTATION	
	<i>adhB</i>	JV391D10_v1_210016	NS	15.521	2 bp→AC	coding (75 76/915 nt)	100%	15.521	2 bp→AC	coding (75 76/915 nt)	100%	15.521	2 bp→AC	coding (75 76/915 nt)	100%	15.521	2 bp→AC	coding (75 76/915 nt)	100%
	<i>adhB</i>	JV391D10_v1_210016	NS	15.55	2 bp→CC	coding (46 47/915 nt)	100%	15.55	2 bp→CC	coding (46 47/915 nt)	100%	15.55	2 bp→CC	coding (46 47/915 nt)	100%	15.55	2 bp→CC	coding (46 47/915 nt)	100%
	<i>adhB</i>	JV391D10_v1_210016	NS	15.573	A→T	D8E (GAI→GAA)	100%	15.573	A→T	D8E (GAI→GAA)	100%	15.573	A→T	D8E (GAI→GAA)	100%	15.573	A→T	D8E (GAI→GAA)	100%
	<i>adhB</i>	JV391D10_v1_210016	NS	15.579	G→C	S6A (TCC→GCC)	100%	15.579	G→C	S6A (TCC→GCC)	100%	15.579	G→C	S6A (TCC→GCC)	100%	15.579	G→C	S6A (TCC→GCC)	100%
	<i>adhB</i>	JV391D10_v1_210016	NS	15.581	A→C	S6A (TCC→GCG)	100%	15.581	2 bp→CA	coding (15 16/915 nt)	100%	15.581	2 bp→CA	coding (15 16/915 nt)	100%	15.581	2 bp→CA	coding (15 16/915 nt)	100%

B	SynCom 2-27						SynCom 7-27						SynCom 7-27					
	GENE	LOCUS	TYPE OF MUTATION	POSITION	MUTATION	ANNOTATION	MUTATED READS (%)	POSITION	MUTATION	ANNOTATION	MUTATED READS (%)	POSITION	MUTATION	ANNOTATION	MUTATED READS (%)	POSITION	MUTATION	ANNOTATION
	<i>accA</i>	JV391D10_v1_10828	NS	33.703	6 bp→6 bp	coding (361 366/948 nt)	100%	33.703	6 bp→6 bp	coding (361 366/948 nt)	100%	33.703	6 bp→6 bp	coding (361 366/948 nt)	100%	33.703	6 bp→6 bp	coding (361 366/948 nt)
	<i>accA</i>	JV391D10_v1_10828	NS	33.714	2 bp→GA	coding (372 373/948 nt)	100%	33.714	2 bp→GA	coding (372 373/948 nt)	100%	33.714	2 bp→GA	coding (372 373/948 nt)	100%	33.714	2 bp→GA	coding (372 373/948 nt)
	<i>accA</i>	JV391D10_v1_10828	NS	33.717	G→T	V125I (GTG→ATL)	100%	33.717	G→T	V125I (GTG→ATL)	100%	33.717	G→T	V125I (GTG→ATL)	100%	33.717	G→T	V125I (GTG→ATL)
	<i>accA</i>	JV391D10_v1_10828	NS	33.738	4 bp→AGCA	coding (396 399/948 nt)	100%	33.738	4 bp→AGCA	coding (396 399/948 nt)	100%	33.738	4 bp→AGCA	coding (396 399/948 nt)	100%	33.738	4 bp→AGCA	coding (396 399/948 nt)
	<i>accA</i>	JV391D10_v1_10828	NS	33.761	+G	coding (419/948 nt)	100%	33.761	+G	coding (419/948 nt)	100%	33.761	+G	coding (419/948 nt)	100%	33.761	+G	coding (419/948 nt)
	<i>accA</i>	JV391D10_v1_10828	NS	33.766	Δ1 bp	coding (424/948 nt)	100%	33.766	Δ1 bp	coding (424/948 nt)	100%	33.766	Δ1 bp	coding (424/948 nt)	100%	33.766	Δ1 bp	coding (424/948 nt)
	<i>accA</i>	JV391D10_v1_10828	NS	33.84	2 bp→CG	coding (498 499/948 nt)	100%	33.84	2 bp→CG	coding (498 499/948 nt)	100%	33.84	2 bp→CG	coding (498 499/948 nt)	100%	33.84	2 bp→CG	coding (498 499/948 nt)
	<i>accA</i>	JV391D10_v1_10828	NS	33.845	2 bp→GT	coding (503 504/948 nt)	100%	33.845	2 bp→GT	coding (503 504/948 nt)	100%	33.845	2 bp→GT	coding (503 504/948 nt)	100%	33.845	2 bp→GT	coding (503 504/948 nt)
	<i>accA</i>	JV391D10_v1_10828	NS	33.858	4 bp→TGGT	coding (516 519/948 nt)	100%	33.858	4 bp→TGGT	coding (516 519/948 nt)	100%	33.858	4 bp→TGGT	coding (516 519/948 nt)	100%	33.858	4 bp→TGGT	coding (516 519/948 nt)
	<i>accA</i>	JV391D10_v1_10828	NS	33.865	2 bp→TC	coding (523 524/948 nt)	100%	33.865	2 bp→TC	coding (523 524/948 nt)	100%	33.865	2 bp→TC	coding (523 524/948 nt)	100%	33.865	2 bp→TC	coding (523 524/948 nt)
	<i>accA</i>	JV391D10_v1_10828	NS	33.88	T→C	W180R (TGG→CGT)	100%	33.88	T→C	W180R (TGG→CGT)	100%	33.88	T→C	W180R (TGG→CGT)	100%	33.88	T→C	W180R (TGG→CGT)
	<i>accA</i>	JV391D10_v1_10828	NS	33.882	G→T	W180R (TGG→CGI)	100%	33.882	G→T	W180R (TGG→CGI)	100%	33.882	G→T	W180R (TGG→CGI)	100%	33.882	G→T	W180R (TGG→CGI)
	<i>accA</i>	JV391D10_v1_10828	NS	33.893	2 bp→AA	coding (551 552/948 nt)	100%	33.893	2 bp→AA	coding (551 552/948 nt)	100%	33.893	2 bp→AA	coding (551 552/948 nt)	100%	33.893	2 bp→AA	coding (551 552/948 nt)
	<i>accA</i>	JV391D10_v1_10828	NS	33.898	G→T	A186S (GCC→TCC)	100%	33.898	G→T	A186S (GCC→TCC)	100%	33.898	G→T	A186S (GCC→TCC)	100%	33.898	G→T	A186S (GCC→TCC)
	<i>accA</i>	JV391D10_v1_10828	NS	33.901	C→G	R187G (CGC→GGT)	100%	33.901	C→G	R187G (CGC→GGT)	100%	33.901	C→G	R187G (CGC→GGT)	100%	33.901	C→G	R187G (CGC→GGT)
	<i>accA</i>	JV391D10_v1_10828	NS	33.903	C→T	R187G (CGC→GGI)	100%	33.903	C→T	R187G (CGC→GGI)	100%	33.903	C→T	R187G (CGC→GGI)	100%	33.903	C→T	R187G (CGC→GGI)

Table S1 (continued)

SynCom 2-27												SynCom 7-27				C
GENE	LOCUS	TYPE OF MUTATION	POSITION	MUTATION	ANNOTATION	MUTATED READS (%)	POSITION	MUTATION	ANNOTATION	MUTATED READS (%)	MUTATED READS (%)					
pepA	JV391D10.v1.50070	NS	90.174	A→G	I315V (ATC→GTG)	100%	90.174	A→G	I315V (ATC→GTG)	100%	100%					
pepA	JV391D10.v1.50070	NS	90.176	C→G	I315V (ATC→GTG)	100%	90.176	C→G	I315V (ATC→GTG)	100%	100%					
pepA	JV391D10.v1.50070	NS	90.181	C→T	A317V (GCC→GTC)	100%	90.181	C→T	A317V (GCC→GTC)	95.50%	95.50%					
pepA	JV391D10.v1.50070	NS	90.183	T→G	C318G (TGC→GGC)	100%	90.183	T→G	C318G (TGC→GGC)	100%	100%					
pepA	JV391D10.v1.50070	NS	90.186	2 bp→TG	coding (955 956/1491 nt)	100%	90.186	2 bp→TG	coding (955 956/1491 nt)	100%	100%					
pepA	JV391D10.v1.50070	NS	90.201	2 bp→GA	coding (970 971/1491 nt)	100%	90.201	2 bp→GA	coding (970 971/1491 nt)	100%	100%					
pepA	JV391D10.v1.50070	NS	90.207	4 bp→AAAT	coding (976 979/1491 nt)	100%	90.207	4 bp→AAAT	coding (976 979/1491 nt)	100%	100%					
pepA	JV391D10.v1.50070	NS	90.214	2 bp→TG	coding (983 984/1491 nt)	100%	90.214	2 bp→TG	coding (983 984/1491 nt)	100%	100%					
pepA	JV391D10.v1.50070	NS	90.227	2 bp→CG	coding (986 987/1491 nt)	100%	90.227	2 bp→CG	coding (986 987/1491 nt)	100%	100%					
pepA	JV391D10.v1.50070	NS	90.23	2 bp→GC	coding (999 1000/1491 nt)	100%	90.23	2 bp→GC	coding (999 1000/1491 nt)	100%	100%					
pepA	JV391D10.v1.50070	NS	90.233	C→G	V334M (GTC→CTG)	100%	90.233	C→G	V334M (GTC→CTG)	100%	100%					
pepA	JV391D10.v1.50070	NS	90.24	C→A	L337M (CTC→ATG)	100%	90.24	C→A	L337M (CTC→ATG)	100%	100%					
pepA	JV391D10.v1.50070	NS	90.242	3 bp→GTC	coding (1011 1013/1491 nt)	100%	90.242	3 bp→GTC	coding (1011 1013/1491 nt)	100%	100%					
pepA	JV391D10.v1.50070	NS	90.261	A→G	I344V (ATC→GTG)	100%	90.261	A→G	I344V (ATC→GTG)	100%	100%					
pepA	JV391D10.v1.50070	NS	90.263	2 bp→GT	coding (1032 1033/1491 nt)	100%	90.263	2 bp→GT	coding (1032 1033/1491 nt)	100%	100%					
pepA	JV391D10.v1.50070	NS	90.316	2 bp→TG	coding (1085 1086/1491 nt)	100%	90.316	2 bp→TG	coding (1085 1086/1491 nt)	100%	100%					
pepA	JV391D10.v1.50070	NS	90.327	A→G	K366D (AAG→GAT)	100%	90.327	A→G	K366D (AAG→GAT)	100%	100%					
pepA	JV391D10.v1.50070	NS	90.329	G→T	K366D (AAG→GAT)	100%	90.329	G→T	K366D (AAG→GAT)	100%	100%					
pepA	JV391D10.v1.50070	NS	90.332	2 bp→GG	coding (1101 1102/1491 nt)	100%	90.332	2 bp→GG	coding (1101 1102/1491 nt)	100%	100%					
pepA	JV391D10.v1.50070	NS	90.335	G→A	Q368E (CAG→GAA)	100%	90.335	G→A	Q368E (CAG→GAA)	100%	100%					
pepA	JV391D10.v1.50070	NS	90.337	C→T	A369V (GCG→GTG)	100%	90.337	C→T	A369V (GCG→GTG)	100%	100%					
pepA	JV391D10.v1.50070	NS	90.348	A→G	I373V (ATC→GTT)	100%	90.348	A→G	I373V (ATC→GTT)	100%	100%					
pepA	JV391D10.v1.50070	NS	90.35	C→T	I373V (ATC→GTT)	100%	90.35	C→T	I373V (ATC→GTT)	100%	100%					
pepA	JV391D10.v1.50070	NS	90.372	A→G	I381V (ATC→GTT)	100%	90.372	A→G	I381V (ATC→GTT)	100%	100%					
pepA	JV391D10.v1.50070	NS	90.374	2 bp→TA	coding (1143 1144/1491 nt)	100%	90.374	2 bp→TA	coding (1143 1144/1491 nt)	100%	100%					
SynCom 2-27												SynCom 7-27				D
GENE	LOCUS	TYPE OF MUTATION	POSITION	MUTATION	ANNOTATION	MUTATED READS (%)	POSITION	MUTATION	ANNOTATION	MUTATED READS (%)	MUTATED READS (%)					
dmfR	JV391D10.v1.10781	NS	787.382	C→A	Q112K (CAA→AAA)	100%	787.382	C→A	Q112K (CAA→AAA)	100%	100%					
dmfR	JV391D10.v1.10781	NS	787.431	A→G	N128S (AAC→AGC)	100%	787.431	A→G	N128S (AAC→AGC)	100%	100%					
SynCom 2-27												SynCom 7-27				E
GENE	LOCUS	TYPE OF MUTATION	POSITION	MUTATION	ANNOTATION	MUTATED READS (%)	POSITION	MUTATION	ANNOTATION	MUTATED READS (%)	MUTATED READS (%)					
hrpA	JV391D10.v1.90183	NS	186.713	2 bp→GA	coding (1195 1196/3912 nt)	100%	186.713	2 bp→GA	coding (1195 1196/3912 nt)	100%	100%					
hrpA	JV391D10.v1.90183	NS	186.729	C→G	A394P (GCG→CCG)	100%	186.729	C→G	A394P (GCG→CCG)	100%	100%					
hrpA	JV391D10.v1.90183	NS	186.754	G→C	A385T (GCC→ACG)	100%	186.754	G→C	A385T (GCC→ACG)	100%	100%					
hrpA	JV391D10.v1.90183	NS	186.756	C→T	A385T (GCC→ACG)	100%	186.756	C→T	A385T (GCC→ACG)	100%	100%					
hrpA	JV391D10.v1.90183	NS	186.791	3 bp→GGA	coding (1116 1118/3912 nt)	100%	186.791	3 bp→GGA	coding (1116 1118/3912 nt)	100%	100%					
hrpA	JV391D10.v1.90183	NS	186.805	3 bp→TTT	coding (1102 1104/3912 nt)	100%	186.805	3 bp→TTT	coding (1102 1104/3912 nt)	100%	100%					
SynCom 2-27												SynCom 7-27				F
GENE	LOCUS	TYPE OF MUTATION	POSITION	MUTATION	ANNOTATION	MUTATED READS (%)	POSITION	MUTATION	ANNOTATION	MUTATED READS (%)	MUTATED READS (%)					
-	JV391D10.v1.40309	NS	337.987	4 bp→TAAA	coding (1069 1072/1176 nt)	92.30%	337.987	4 bp→TAAA	coding (1069 1072/1176 nt)	100%	100%					
-	JV391D10.v1.40309	NS	338.01	G→C	T350S (ACC→AGC)	92.30%	338.01	G→C	T350S (ACC→AGC)	100%	100%					
-	JV391D10.v1.40309	NS	338.012	G→C	D349E (GAC→GAG)	100%	338.012	G→C	D349E (GAC→GAG)	100%	100%					
-	JV391D10.v1.40309	NS	338.025	A→G	V345A (GTC→GCC)	100%	338.025	A→G	V345A (GTC→GCC)	100%	100%					
-	JV391D10.v1.40309	NS	338.028	2 bp→CT	coding (1030 1031/1176 nt)	100%	338.028	T→C	D344G (GAC→GGC)	100%	100%					
-	JV391D10.v1.40309	NS	338.033	3 bp→GTC	coding (1024 1026/1176 nt)	100%	338.033	C→G	P342A (CCG→GCC)	100%	100%					
-	JV391D10.v1.40309	NS	-	-	-	-	338.035	G→C	P342A (CCG→GCC)	100%	100%					









Table S3 (continued)

---

## CHAPTER 4

---

---

## GENERAL DISCUSSION AND PERSPECTIVES

---

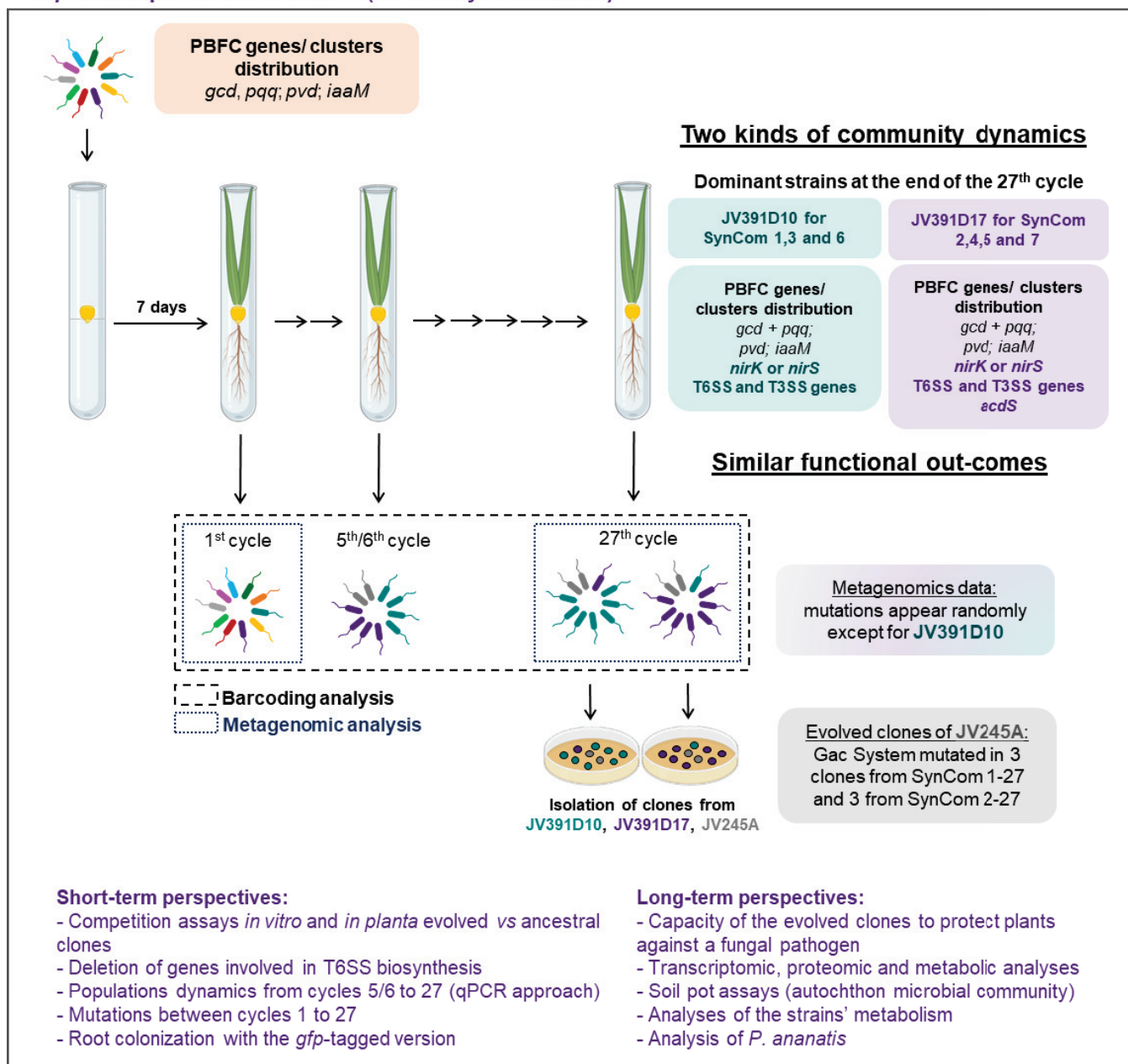
Plants roots recruit, via their exudates, a plethora of microorganisms that influence their development and adaptation to the environment. Rhizomicrobiota composition varies according to different biotic and abiotic factors. The surrounding soil is the main determinant as it offers many diverse microbial populations that may interact with the plant (Alabouvette and Steinberg, 2006; Bulgarelli *et al.*, 2012; Lundberg *et al.*, 2012; Schlaeppi *et al.*, 2014). Plants also regulate rhizomicrobiota composition via their exudates (Haichar *et al.*, 2008). Exudation profiles vary between two plants of different families, genera, species or varieties (Lesuffleur *et al.*, 2007; Preece and Peñuelas, 2020) and change over time, according to the stage of development of the plant (Jones *et al.*, 2004). Other studies have pointed out that plant genotype and geographic location can also condition rhizomicrobiota composition (Peiffer *et al.*, 2013; Bodenhausen *et al.*, 2014). Furthermore, rhizomicrobiota are also shaped by soil properties (availability of carbon, nutrient, pH value, salinity; Hansel *et al.*, 2008; Chu *et al.*, 2010; Eilers *et al.*, 2012), and anthropological intervention in agricultural systems can modify these properties (Li *et al.*, 2014). Finally, the evolutionary history of plants, including their domestication, is a factor that has also influenced the composition of the rhizomicrobiota (Bouffaud *et al.*, 2012, 2014; Pérez-Jaramillo *et al.*, 2016). Thus, the interaction between the plant and its environment is one of the most important factor influencing the assembly of the rhizomicrobiota (Peiffer *et al.*, 2013). Moreover, in microbial communities, microorganisms are always evolving in relation with other microorganisms and have to adapt not only to the environment itself but also to the biotic environment dynamically shaped by the other microorganisms. Therefore, bacterial behaviours depend on the interactions networks shared between microorganisms (Smith and Schuster, 2019). Consequently, plants and their microbiota should no longer be seen as separated organisms but rather as one interacting entity: a holobiont (Vandenkoornhuyse *et al.*, 2015).

Among the recruited microorganisms, PGPR are capable of stimulating plant growth and/or protecting plants against diseases via numerous plant beneficial properties (Bashan and de-Bashan, 2010; Vacheron, 2015; Lemanceau *et al.*, 2017). Despite harbouring a high or low number of plant beneficial functions, PGPR inoculation does not always lead to significant effects on the plant, probably because the observed plant growth-promoting effects may depend on the interactions of PGPR with the already settled microbial populations. Consequently, in order to get a deeper insight into plant-PGPR cooperations, understanding the mechanisms controlling microbial colonisation and persistence into plant-associated microbial communities is crucial. Interactions between the plant and its microbiota, and within

the microbiota itself are known to be complex. Synthetic communities (SynCom) offer a simplified version of natural communities. They can be defined in order to conserve their key features and to unravel the network of interactions within microbial communities. The hypothesis of this work was that plants, via their root system, would shape a community of cooperative populations. To test this, we first tracked the population dynamics of a PGPR synthetic community on maize roots via an experimental evolution approach. Secondly, we studied the consequences the EE had on the strains' genomes.

The initial synthetic community was composed of 10 strains belonging to 7 different species of the *Pseudomonas fluorescens* lineage (2 species of the *P. putida* group: *P. wadenswilerensis*, *P. inefficax* and 5 species of the *P. fluorescens* group: *P. brassicacerum*, *P. protegens*, *P. jessenii*, *P. kribbensis* and *P. fluorescens*). *Pseudomonas* are ubiquitous soil bacteria and among this genus, *P. fluorescens* lineage is found in the rhizosphere of many plants (Couillerot *et al.*, 2009; Haney *et al.*, 2015). The strains of this lineage harbor a wide range of plant beneficial functions with different modes of action on the plant (Ahmad *et al.*, 2008; Naik *et al.*, 2008; Loper *et al.*, 2012), which grants them a crucial ecological role in the rhizosphere. Due to the diverse modes of action of bacteria belonging to the *Pseudomonas fluorescens* lineage, they have been considered as first-order candidates for biological control since the 1970s (Weller, 2007). *Pseudomonas* can be used as biostimulants: products of preventive use which facilitates nutrient uptake, plant growth and resistance to abiotic stresses such as salinity and drought (Kumar and Alope, 2020). Indeed, *Pseudomonas* are present in some commercialized PGPR-based products (Kumar and Alope, 2020). Furthermore, *Pseudomonas* are dominant in maize rhizoplane, where they represent 67% of the operational taxonomic units (Ofek *et al.*, 2014) and in maize rhizosphere (Bouffaud *et al.*, 2012; García-Salamanca *et al.*, 2013). Moreover, previous works (Vacheron, 2015; Vacheron *et al.*, 2016) showed that up to 60% of the retrieved populations from maize rhizosphere belonged to the *P. fluorescens* group FMJK cluster, from which strains *P. fluorescens* JV391D10, JV391D17 and JV449, *P. jessenii* JV251A and *P. kribbensis* JV359A8 belong to. All the strains that composed the initial SynCom come from a library of 698 cultivated *Pseudomonas* that were initially isolated from maize (Vacheron *et al.*, 2016). Maize is a worldwide crop of great importance, and its associated microbial communities have intrigued many researchers. The studies that have analysed maize associated communities have shown that its rhizospheric community is quite different (in terms of relative abundance and diversity) from the adjacent bulk soil ones (Sanguin *et al.*, 2006; Haichar *et al.*, 2008; Bouffaud *et al.*, 2012; García-Salamanca *et al.*, 2013; Peiffer *et al.*, 2013).

## In planta experimental evolution (for one SynCom series)



**Figure 1: Summary figure of the *in planta* experimental evolution study and work perspectives.** One SynCom series is presented as an example for the *in planta* experimental evolution. The metabarcoding analysis was done for the initial SynCom as well as for SynComs from cycles 1 to 5/6 and 27 (black dotted lines). Metagenomics analysis was carried on the initial SynCom and on SynComs from cycle 1 and 27 (blue lines). Results for both these approaches were presented in chapter 2. The metagenomics data used to validate the metabarcoding approach in chapter 2 was used in chapter 3 to assess the genetic changes found within the SynComs in strains *P. fluorescens* JV391D10, JV391D17 and *P. protegens* JV245A. Afterwards, evolved clones of these strains were isolated from SynComs from the 27 cycles, however genetic changes were only analysed for the evolved clones of *P. protegens* JV245A. The obtained results are summarized in the coloured boxes and the discussed perspectives are presented in the framed boxes.

PBFC genes: plant-beneficial function contributing genes; T6SS: type 6 secretion system; T3SS: type 3 secretion system.

## Population dynamics of the synthetic community

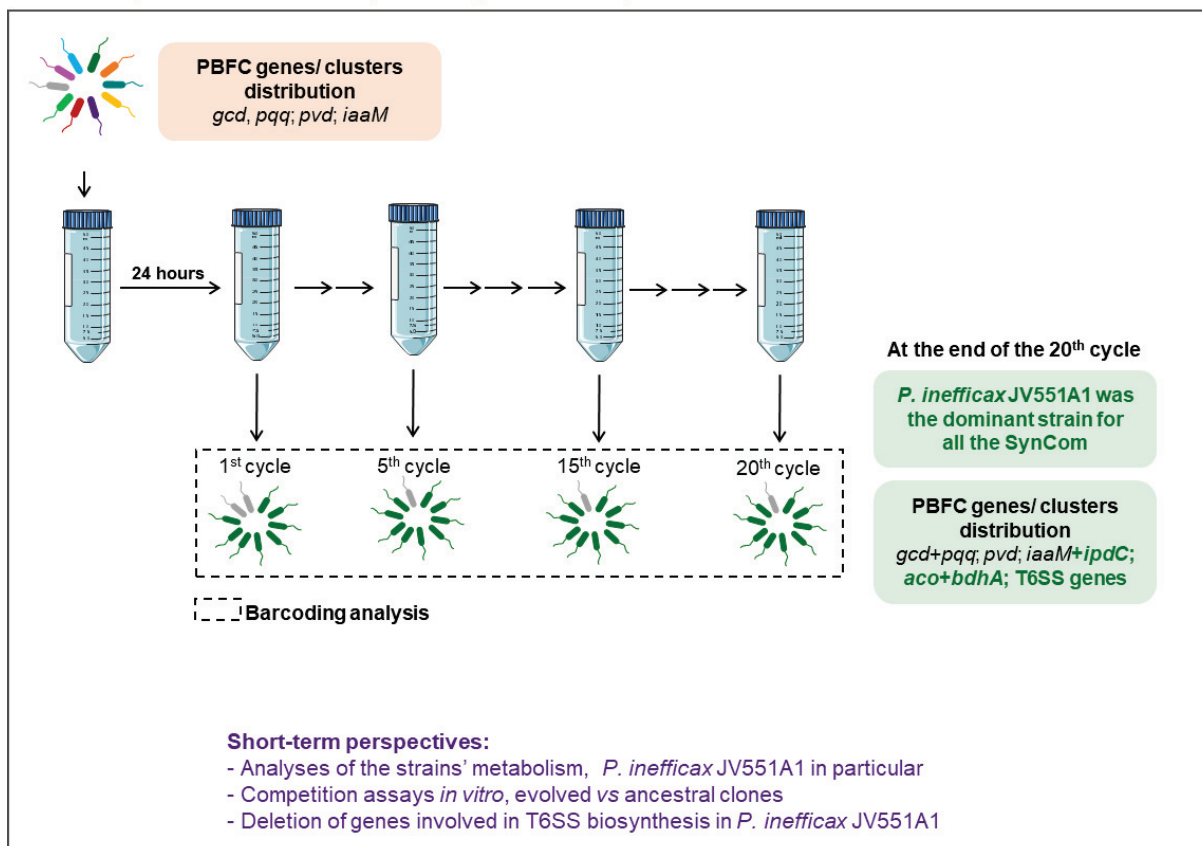
The *in planta* EE was carried for 27 cycles of evolution, at the end of each cycle, plant root populations were recovered, diluted and re-inoculated on fresh maize seedlings (Figure 1). Rapid dynamics populations were observed as half of the strains disappeared after only 3 cycles of evolution (based on the metabarcoding analysis) leading to a duo of dominant species: *P. fluorescens* and *P. protegens*. In total, 7 series of EE (SynComs) were carried independently and different population dynamics were observed among the SynComs at the end of the 27<sup>th</sup> cycle: (1) SynComs 1-27, 3-27 and 6-27 harboured *P. fluorescens* JV391D10 and *P. protegens* JV245A as the dominant strains whereas, it was (2) *P. fluorescens* JV391D17 *P. protegens* JV245A for SynComs 2-27, 4-27, 5-27 and 7-27. Population dynamics during the *in planta* EE were tracked using a metabarcoding approach on the gene *rpoD*, which encodes the sigma 70 factor subunit of the RNA polymerase. When it comes to *Pseudomonas* identification, *rpoD* offers a better discrimination as this gene evolves faster than 16S rRNA gene (Yamamoto *et al.*, 2000). *Pseudomonas* are ubiquitous bacteria in nature and primers on *rpoD* have successfully identified *Pseudomonas* strains using polymerase chain reaction-based approach (Mulet *et al.*, 2009). To confirm the observed population dynamics, a metagenomic analysis was carried on the initial SynCom and on SynComs retrieved at the end of the 1<sup>st</sup> and 27<sup>th</sup> cycle. The metagenomics approach showed a rather different composition of the initial SynCom. Contrary to what was observed in the initial SynCom with the metabarcoding approach, metagenomics revealed a SynCom where *P. fluorescens* and *P. protegens* were the least represented strains. For this approach only, the reads that could specifically be attributed to only one strain were analysed. The number of specific genes (and therefore the number of reads that can be mapped) varies among the strains, in such a way that strains with a higher number of specific genes tend to be over-represented. Despite the differences in the initial SynCom, the metagenomic approach presented the same population dynamics as observed with the meta-barcoding approach: SynComs 1-27, 3-27 and 6-27 harboured *P. fluorescens* JV391D10 and *P. protegens* JV245A (plus JV359A1) as the dominant strains whereas, it was *P. fluorescens* JV391D17 and *P. protegens* JV245A (plus JV359A1) for SynComs 2-27 and 7-27. Furthermore, the metagenomics approach brought insight into minor changes that may have happened during the EE, revealing that no strain had disappear but were all still present in very low proportions. With the metabarcoding approach, the number of analysed reads was 100 smaller (roughly 4 million specific reads per SynCom for the metagenomic approach and roughly 40 000 reads for the metabarcoding approach) making the method less sensitive to minor changes in the population (half of the strains quickly became undetectable). Additionally, the metagenomics approach revealed the presence of an endophyte (*P. ananatis*) in SynCom 2-27, 3-27, 6-27 and 7-27 which could not have been seen with the metabarcoding approach as the designed primers only targeted the *rpoD* gene of



*Pseudomonas*. Overall, both approaches are complementary and allow independent validations to describe the composition of the synthetic communities.

In order to unravel the role of the plant on these population dynamics, an *in vitro* EE was carried in parallel, on a liquid rich media with the same initial SynCom (Figure 2).

#### *In vitro* experimental evolution (for one SynCom series)



**Figure 2: Experimental summary figure for the *in vitro* experimental evolution and work perspectives.** One SynCom series is presented as an example for the *in vitro* experimental evolution. The metabarcoding analysis was done for the initial SynCom as well as for SynComs from cycles 1, 5, 10 and 15 (black dotted lines). Results for this approach were presented in chapter 2. The obtained results are summarized in the coloured boxes and the discussed perspectives are presented in the framed boxes.

PBFC genes: plant-beneficial function contributing genes; T6SS: type 6 secretion system.

Population dynamics were tracked with the metabarcoding approach on the *rpoD* gene and rapid population dynamics were detected: *P. inefficax* JV551A1 became the only dominant strain at the end of the 1<sup>st</sup> cycle of evolution and for all the SynCom (n=5) done in parallel and stayed dominant till the end of the *in vitro* EE. This result show that the population dynamics observed *in planta* (*P. fluorescens* and *P. protegens* as the dominant species) are most likely due to the selection maize roots exert on the synthetic community rather than on the inherent bacterial interactions shared by the SynCom populations.



We first thought these different selection dynamics were due to the nutrient sources: *in planta*, they were furnished by the roots (by the rhizodeposits) whereas *in vitro* it was a rich liquid media (TSB). Preliminary tests showed similar growth rates for each strain on a maize-based liquid medium and on a rich liquid medium. Plant root exudate secretion varies through the age of individuals and alongside the roots (Jones *et al.*, 2004), giving a heterogeneous distribution of nutrients that can impact microbial networks interactions (Kram *et al.*, 2017). To get a deeper insight into the differential population dynamics between both EE, the metabolism of the strains was screened *in silico*, but no main significant differences were found between the 10 *Pseudomonas* strains. Moreover, to go further, all the ancestral strains have been *gfp*-tagged and could be used to track their root colonisation patterns (alone or in combination), in particular for the strains that are dominant at the end of the *in planta* EE, to see if they colonize different zones of the root which could explain their compatibility.

Regarding plant beneficial function contributing (PBFC) genes, there were 4 main plant beneficial functions that were represented in the initial SynCom: phosphate solubilisation (*pqq* and *gcd*), production of pyoverdine (*pvd*) and auxin (*iaaM*). *In vitro*, *P. inefficax* JV55A1 rapidly became dominant, this strains harbours 8 PBFC genes/clusters: the same 4 afore mentioned plus genes involved in acetoin metabolism (*aco* and *bdhA*), in auxin production (*iaaM* and *ipdC*) and type VI secretion system (T6SS) coding genes. Acetoin is a volatile organic compound that may be used as an additional carbon source (Xiao *et al.*, 2009). At the end of the *in vitro* EE, PBFC genes/clusters encoding for antimicrobial compounds are poorly represented, maybe because no fungal pathogens were used in the experimental set-up. PBFC genes/clusters encoding entomotoxins (*tcaB*) are quickly lost as there are no insects in this set-up too. In general, PBFC genes/clusters that are directly related to the plant are lost, such as *acdS* (encoding ACC deaminase) or type III secretion system (T3SS). In the same vein, auxin production is a plant beneficial function that should be lost as there is no need for auxin in the *in vitro* EE. However, the presence of *iaaM* and *ipdC* genes may lead to potentially higher levels of indole, a well-known communication signal in *E. coli* (Lee and Lee, 2010), which could induce the expression of other genes, like the T6SS genes, which are known to play a role in *Pseudomonas* fitness as it can be used for interbacterial competition (Bernal *et al.*, 2018). Furthermore, indole and auxin can modify gene expression in several microorganisms (Spaepen *et al.*, 2007) and more precisely induce the expression of DAPG coding genes in *P. protegens* (de Werra *et al.*, 2011), a potent antimicrobial compound.

*In planta*, despite different evolutive dynamics (*P. fluorescens* JV391D10 or JV391D17 as the dominant strain), the PBFC genes/clusters distribution remains similar. Indeed, *in planta*, *P. fluorescens* JV391D10 or JV391D17 dominance led to an over-representation of the same PBFC genes found for the *in vitro* EE, plus 2 to 3 additional PBFC genes: nitric oxide (*nirK* or *nirS*) and protease (*aprA*) production and ACC deaminase (*acdS*) for SynCom 2, 4, 5 and 7

(where *P. fluorescens* JV391D17 was the dominant strain). All of these additional PBFC genes/cluster are related to the plant as they can modify plant hormonal balance (*nirK* or *nirS* and *acdS*) or degrade root exudates (*aprA* and also *acdS*). *In planta*, community selection seems to be driven in order to maintain the main types of PBFC genes/cluster distribution. Furthermore, both *P. fluorescens* shared the SynComs with *P. protegens*, which harbours 15 PBFC genes/cluster in its genomes giving a higher representation to other PBFC genes/clusters that were lost in the *in vitro* EE. The different population selection between the *in vitro* and *in planta* EE remains unresolved and might mostly be explained by the nutrient sources availability in both EE. Further in-depth analysis on the metabolism of the strains and particularly that of *P. inefficax* JV551A1 or of its functions could shed some light into the different selection dynamics.

Another line of inquiry would be the presence of T6SS genes which are found in 8 out of 10 strains. This molecular syringe, wide-spread among *Pseudomonas* (Marchi *et al.*, 2013; Bernal *et al.*, 2018), has been described as promoting interbacterial competition (Vacheron *et al.*, 2019). It would be interesting to carry competition tests between the ancestral strains (the ones used for the initial SynCom) as it could reveal if certain strains already had an advantage before the EE. It could be possible that *in vitro* the *P. inefficax* JV551A1 might be most competitive in dual competition assays than the other strains; while *in planta*, it is likely that *P. fluorescens* will be more competitive in dual competition assays. Additionally, competition tests between the evolved strains and their ancestral version could reveal if the evolved clones are indeed more competitive *in vitro* and/or *in planta*. Another short-term perspective would be to engineer mutant strains for the T6SS of *P. inefficax* JV551A1, *P. fluorescens* JV391D10 and JV391D17. Such mutants would shed some light over the implications of this molecular syringe in the observed population dynamics.

This study has focused on the presence or absence of a given PBFC gene/cluster within the genome of each strain. The next step would be to analyse gene expression to see which PBFC genes/clusters are really expressed and useful in presence of the plant. In order to do so, RNA samples were stored every 2 cycles for all the *in planta*-evolved SynComs. Moreover, the *in planta* EE was built to mimic natural conditions by working with a synthetic community, however it was a gnotobiotic simplified system (test tube with plant agar as a support for the pre-germinated seedling). In nature, bulk soil and plant rhizosphere harbour hundreds of thousands of microorganisms. Even though these EE have given insight into the population dynamics underwent by a group of *Pseudomonas* it is crucial to translocate this setup into soil pot assays. Soil pot assays would give insight into the *Pseudomonas* SynCom colonization and its settlement among the native microbial community associated to maize which is mainly composed of *Enterobacter*, *Stenotrophomonas*, *Ochrobactrum*, *Herbaspirillum*, *Curtobacterium* and *Chryseobacterium* (Niu *et al.*, 2017). This imply testing the specificity of

the *rpoD* primers designed for this work in order to see if they are specific enough when tested against soil microbial populations. A further perspective would be to add a plant fungal pathogen to this kind of setup as many of the *Pseudomonas* of the study harbour PBFC genes that encode for antifungal metabolites.

### **Impact of the *in planta* experimental evolution on the strains genomes**

In the second part of these work, genetic changes in the genomes of strains were analysed in order to get a deeper insight into the population evolution. These changes were analysed with Breseq software on 2 different sets of data. First, on the specific reads of each strain from the metagenomic data of the initial SynCom, plus SynComs 1, 2, 6 and 7 from the 1<sup>st</sup> and 27<sup>th</sup> cycle. This is the same data that was used to validate the metabarcoding approach. In this Breseq analysis, only the specific reads of *P. fluorescens* JV391D10, JV391D17 and *P. protegens* JV245A were analysed. The choice of the analysed strains was based on the evolutive outcomes observed at the end of the *in planta* EE. Secondly, genetic variation was assessed on 17 evolved clones of *P. protegens* JV245A isolated from SynCom 1-27, 2-27, 6-27 and 7-27.

Regarding the metagenomic data, the evolutive outcomes of the SynComs do not seem to have an impact on mutations and they appear rather randomly from one SynCom to another. Only *P. fluorescens* JV391D10 presents a segregation of its mutations in SynComs sharing the same population dynamics (*P. fluorescens* JV391D17; SynCom 2-27 and 7-27). The presence of the endophyte *P. ananatis* seems however to impose a selective pressure as 7 to 9 mutations were found in the specific reads from SynCom 2-27, 6-27 and 7-27 whereas, only 2 mutations are detected in SynCom 1-27, where the endophyte was absent.

Evolved clones were isolated for these 3 strains. For both *P. fluorescens*, evolved clones were isolated from SynComs where they were dominant (SynComs 1-27 and 6-27 for *P. fluorescens* JV391D10 and SynComs 2-27 and 7-27 for JV391D17). Only 96 evolved clones were isolated from each SynCom. A larger sample size would perhaps have allowed to isolate evolved clones of *P. fluorescens* JV391D10 and JV391D17 from SynComs where these strains were in the minority and may reveal new mutations. Breseq analysis revealed some problems for these two strains and their reference genomes need to be re-sequenced.

On the other hand, *P. protegens* JV245A genome was re-sequenced during this thesis work and a hybrid assembly was used as the reference genome for the Breseq analysis on the evolved clones of this strain. For these evolved clones, mutations also seemed to appear rather randomly except for the ones found on the gene *gacA*, where 2 distinct mutations were found in 3 clones isolated from SynCom 1-27 (where *P. fluorescens* JV391D10 was the dominant strain) and in 3 clones isolated from SynCom 2-27 (where *P. fluorescens* JV391D17 was the dominant strain). The Gac systems regulates the expression of many genes, particularly those

involved in antimicrobial production (Hassan *et al.*, 2010; Zhang *et al.*, 2020). Therefore, it would be important to test the impact these mutations may have on the evolved clone's fitness via *in planta* competition assays, confronting the ancestral clones of *P. protegens* JV245A against the evolved clones. In the long term, it would also be interesting to test the effects of these mutations on the clone's capacity to protect plants against fungal pathogens as well as to analyse the expression of genes encoding antimicrobial biosynthesis pathways.

Evolution happens through adaptation and adaptation happens through the fixation of mutations. In order to be fixed, a mutation needs time to rise to the population level via certain number of generations. The duration of the EE ( $\approx 400$  bacterial generations) was based on previous works (Lenski and Travisano, 1994; Cooper and Lenski, 2000; Wiser *et al.*, 2013; Kram *et al.*, 2017; Lenski, 2017) which have shown that adaptation happens in the first hundreds of generations. Indeed, few mutations were found at the end of the 1<sup>st</sup> cycle ( $\approx 15$  generations), whereas more mutations were found at the end of the EE ( $\approx 400$  bacterial generations). Regarding this particular point, new mutations keep rising in Lenski's ongoing EE (which has now passed the 60 000 generations) (Good *et al.*, 2017), implying that despite reaching a plateau, adaptation keeps happening, at a much slower rate, even in a constant environment. Following that line of reasoning, our EE could have been kept going for more than 400 generations in order to see if new adaptation would rise of, if population dynamics would change as the metagenomics approach revealed that all the strains are still present at very low proportions even at the end of the 27<sup>th</sup> cycle. However, before considering a long-term EE, it would be wiser to analyse if any new mutations appear between the 1<sup>st</sup> and the 6<sup>th</sup> cycle, as the metabarcoding on *rpoD* showed that most of the population dynamic happened between these cycles. Furthermore, in this study only cycles 1 to 6 and 27 were analysed, and even if population dynamics seemed to have stabilized at the end of cycles 5 and 6, a shift between dominant populations was observed in SynCom 1, 3 and 6 when analysed at the end of the 27<sup>th</sup> cycle. All the SynComs retrieved from the other cycles of evolution (7 to 26) have been stored: glycerol samples for the bacterial populations that were retrieved and corresponding DNA samples. The latter can be used to get a more in-depth look at the EE, as perhaps the population dynamics changes between cycles 7 to 26 before giving the proportions that were observed at the end of 27<sup>th</sup> cycle. During this work, a quantitative PCR approach was considered based on specific primers designed for each strain that would allow to quantify the proportions of each strain within the SynComs. Unfortunately, preliminary test showed that the designed primers were not specific as thought. These primers were designed based on the genomes available in the MicroScope platform, or, Breseq analysis pointed out that the quality of some of these genomes was not as expected.

The aim of this study was to offer insight into plant-PGPR cooperation. Altogether these results offer a more comprehensive view of population dynamics that takes place within a synthetic

community in interaction with the plant. In this study, maize seems to select PGPR at both taxonomic and functional level. From a taxonomic point of view, *P. protegens* and *P. fluorescens* JV391D10 or JV391D7 were the dominant strains. At the beginning of this work, *P. fluorescens* JV391D10 was classified in the low number of plant beneficial properties category as it harbours 10 PBFC genes/clusters in its genome. Whereas, *P. fluorescens* JV391D17 and *P. protegens* JV245A were classified as having a high number of plant beneficial properties as each harbour 11 and 15 PBFC genes/clusters respectively. This taxonomic selection directly impacted PBFC genes/clusters distribution. At the end of the *in planta* EE, all PBFC genes/clusters (except for *ipdC* and *tcaB2*) were still present. This is due mainly to the selection of *P. protegens* which carries most of the PBFC (15) genes/clusters analysed in this work.

A work prior to this one revealed that in maize rhizosphere, most of the selected *Pseudomonas* harbour a low number of plant beneficial properties (5 or less), as there might exist a metabolic cost to maintaining PBFC genes/clusters. Yet, our results suggest that in the tested conditions (axenic maize seedlings) there is no particular cost to maintain a high number of PBFC genes/clusters. However, the system presented in this work is very different from natural maize rhizosphere. In order to get deeper understanding of plant-PGPR interactions, transcriptomic analysis, as well as soil pot assays are required. As long-term perspectives, proteomics and metabolomics analyses can be considered in order to get insight into the underlying functional networks. Furthermore, it would be interesting to carry the EE in the presence of a plant pathogen or with additional selection pressures to see if the population dynamics and genetic adaptation of the evolved microbiota differ from the presented results. Studies using SynComs of PGPR are rather new and need to be strengthened, for example by the addition of a plant pathogen into the system. Nevertheless, approaches such as experimental evolution and synthetic communities are powerful tools to unravel the intricate network of interactions that conditions plant colonization and settlement of the inoculated bacteria within the autochthon plant microbiota.



## REFERENCES (INTRODUCTION)

- Biessy, A., Novinscak, A., Blom, J., Léger, G., Thomashow, L. S., Cazorla, F. M., *et al.* (2019). Diversity of phytobeneficial traits revealed by whole-genome analysis of worldwide-isolated phenazine-producing *Pseudomonas* spp. *Environmental Microbiology* 21, 437–455. doi:https://doi.org/10.1111/1462-2920.14476.
- Brockhurst, M. A., and Koskella, B. (2013). Experimental coevolution of species interactions. *Trends in Ecology & Evolution* 28, 367–375. doi:10.1016/j.tree.2013.02.009.
- Cairns, J., Ruokolainen, L., Hultman, J., Tamminen, M., Virta, M., and Hiltunen, T. (2018). Ecology determines how low antibiotic concentration impacts community composition and horizontal transfer of resistance genes. *Communications Biology* 1. doi:10.1038/s42003-018-0041-7.
- Cray, J. A., Connor, M. C., Stevenson, A., Houghton, J. D. R., Rangel, D. E. N., Cooke, L. R., *et al.* (2016). Biocontrol agents promote growth of potato pathogens, depending on environmental conditions. *Microbial Biotechnology* 9, 330–354. doi:10.1111/1751-7915.12349.
- Etesami, H., and Maheshwari, D. K. (2018). Use of plant growth promoting rhizobacteria (PGPRs) with multiple plant growth promoting traits in stress agriculture: Action mechanisms and future prospects. *Ecotoxicology and Environmental Safety* 156, 225–246. doi:10.1016/j.ecoenv.2018.03.013.
- Guidot, A., Jiang, W., Ferdy, J.-B., Thébaud, C., Barberis, P., Gouzy, J., *et al.* (2014). Multihost experimental evolution of the pathogen *Ralstonia solanacearum* unveils genes involved in adaptation to plants. *Molecular Biology and Evolution* 31, 2913–2928. doi:10.1093/molbev/msu229.
- Hussain, S., Siddique, T., Saleem, M., Arshad, M., and Khalid, A. (2009). Impact of pesticides on soil microbial diversity, enzymes, and biochemical reactions. Chapter 5. *Advances in Agronomy* (Academic Press), 159–200. doi:10.1016/S0065-2113(09)01005-0.
- Jess, S., Kildea, S., Moody, A., Rennick, G., Murchie, A. K., and Cooke, L. R. (2014). European union policy on pesticides: implications for agriculture in ireland: European union policy on pesticides. *Pest Management Science* 70, 1646–1654. doi:10.1002/ps.3801.
- Lenski, R. E., Rose, M. R., Simpson, S. C., and Tadler, S. C. (1991). Long-term experimental evolution in *Escherichia coli*, adaptation and divergence during 2,000 generations. *The American Naturalist* 138, 1315–1341. doi:10.1086/285289.
- Loper, J. E., Hassan, K. A., Mavrodi, D. V., Li, E. W. D., Lim, C. K., Shaffer, B. T., *et al.* (2012). Comparative genomics of plant-associated *Pseudomonas* spp.: Insights into diversity and inheritance of traits involved in multitrophic interactions. *PLoS Genetics* 8, e1002784. doi:10.1371/journal.pgen.1002784.
- Meena (2017). Abiotic stress responses and microbe-mediated mitigation in plants: The omics Strategies. *Frontiers. Plant Science*. https://doi.org/10.3389/fpls.2017.00172
- Miyagawa, S., Sato, T., and Iguchi, T. (2016). “Dichlorodiphenyltrichloroethane,” subchapter D101 in *Handbook of Hormones*, eds. San Diego: Academic Press, 579–580. doi:10.1016/B978-0-12-801028-0.00242-7.
- Parnell, J. J., Berka, R., Young, H. A., Sturino, J. M., Kang, Y., Barnhart, D. M., *et al.* (2016). From the lab to the farm: An industrial perspective of plant beneficial microorganisms. *Front. Plant Sci.* 7. doi:10.3389/fpls.2016.01110.
- Paterson, S., Vogwill, T., Buckling, A., Benmayor, R., Spiers, A. J., Thomson, N. R., *et al.* (2010). Antagonistic coevolution accelerates molecular evolution. *Nature* 464, 275–278. doi:10.1038/nature08798.
- Sánchez-Contreras, M., Martín, M., Villaceros, M., O’Gara, F., Bonilla, I., and Rivilla, R. (2002). Phenotypic selection and phase variation occur during alfalfa root colonization by *Pseudomonas fluorescens* F113. *Journal of Bacteriology* 184, 1587–1596. doi:10.1128/JB.184.6.1587-1596.2002.
- Sarniguet, A., Kraus, J., Henkels, M. D., Muehlchen, A. M., and Loper, J. E. (1995). The sigma factor sigma s affects antibiotic production and biological control activity of *Pseudomonas fluorescens* Pf-5. *PNAS* 92, 12255–12259. doi:10.1073/pnas.92.26.12255.
- Scanlan, P. D., Hall, A. R., Lopez-Pascua, L. D. C., and Buckling, A. (2011). Genetic basis of infectivity evolution in a bacteriophage. *Molecular Ecology* 20, 981–989. doi:10.1111/j.1365-294X.2010.04903.x.
- Smith, P., and Schuster, M. (2019). Public goods and cheating in microbes. *Current Biology* 29, R442–R447. doi:10.1016/j.cub.2019.03.001.

Vacheron, J., Moëgne-Loccoz, Y., Dubost, A., Gonçalves-Martins, M., Muller, D., and Prigent-Combaret, C. (2016). Fluorescent *Pseudomonas* strains with only few plant-beneficial properties are favored in the maize rhizosphere. *Front. Plant Sci.* 7. doi:10.3389/fpls.2016.01212.

Yang, L., Jelsbak, L., Marvig, R. L., Damkiaer, S., Workman, C. T., Rau, M. H., *et al.* (2011). Evolutionary dynamics of bacteria in a human host environment. *Proc. Natl. Acad. Sci.* 108, 7481–7486. doi:10.1073/pnas.1018249108.

## REFERENCES (GENERAL DISCUSSION AND PERSPECTIVES)

Ahmad, F., Ahmad, I., and Khan, M. S. (2008). Screening of free-living rhizospheric bacteria for their multiple plant growth promoting activities. *Microbiological Research* 163, 173–181. doi:10.1016/j.micres.2006.04.001.

Alabouvette, C., and Steinberg, C. (2006). The soil as a reservoir for antagonists to plant diseases. *An Ecological and Societal Approach to Biological Control* Progress in Biological Control., eds. (Dordrecht: Springer Netherlands), 123–144. doi:10.1007/978-1-4020-4401-4\_8.

Bashan, Y., and de-Bashan, L. E. (2010). How the plant growth-promoting bacterium *Azospirillum* promotes plant growth—a critical assessment. *Advances in Agronomy* (Elsevier), 77–136. doi:10.1016/S0065-2113(10)08002-8.

Bernal, P., Llamas, M. A., and Filloux, A. (2018). Type VI secretion systems in plant-associated bacteria. *Environmental Microbiology* 20, 1–15. doi:10.1111/1462-2920.13956.

Bodenhausen, N., Bortfeld-Miller, M., Ackermann, M., and Vorholt, J. A. (2014). A synthetic community approach reveals plant genotypes affecting the phyllosphere microbiota. *PLoS Genetics* 10, e1004283. doi:10.1371/journal.pgen.1004283.

Bouffaud, M., Kyselková, M., Gouesnard, B., Grundmann, G., Muller, D., and Moëgne-Loccoz, Y. (2012). Is diversification history of maize influencing selection of soil bacteria by roots? *Molecular Ecology* 21, 195–206. doi:10.1111/j.1365-294X.2011.05359.x.

Bouffaud, M.-L., Poirier, M.-A., Muller, D., and Moëgne-Loccoz, Y. (2014). Root microbiome relates to plant host evolution in maize and other *Poaceae*. *Environmental Microbiology* 16, 2804–2814. doi:https://doi.org/10.1111/1462-2920.12442.

Bulgarelli, D., Rott, M., Schlaeppi, K., Ver Loren van Themaat, E., Ahmadinejad, N., Assenza, F.,

*et al.*, (2012). Revealing structure and assembly cues for *Arabidopsis* root-inhabiting bacterial microbiota. *Nature* 488, 91–95. doi:10.1038/nature11336.

Chu, H., Fierer, N., Lauber, C. L., Caporaso, J. G., Knight, R., and Grogan, P. (2010). Soil bacterial diversity in the Arctic is not fundamentally different from that found in other biomes. *Environmental Microbiology* 12, 2998–3006. doi:https://doi.org/10.1111/j.1462-2920.2010.02277.x.

Cooper, V. S., and Lenski, R. E. (2000). The population genetics of ecological specialization in evolving *Escherichia coli* populations. *Nature* 407, 736–739. doi:10.1038/35037572.

Couillerot, O., Prigent-Combaret, C., Caballero-Mellado, J., and Moëgne-Loccoz, Y. (2009). *Pseudomonas fluorescens* and closely-related fluorescent Pseudomonads as biocontrol agents of soil-borne phytopathogens. *Letters in Applied Microbiology* 48, 505–512. doi:10.1111/j.1472-765X.2009.02566.x.

de Werra, P., Huser, A., Tabacchi, R., Keel, C., and Maurhofer, M. (2011). Plant- and microbe-derived compounds affect the expression of genes encoding antifungal compounds in a *Pseudomonad* with biocontrol activity. *Applied and Environmental Microbiology* 77, 2807–2812. doi:10.1128/AEM.01760-10.

Eilers, K. G., Debenport, S., Anderson, S., and Fierer, N. (2012). Digging deeper to find unique microbial communities: The strong effect of depth on the structure of bacterial and archaeal communities in soil. *Soil Biology and Biochemistry* 50, 58–65. doi:10.1016/j.soilbio.2012.03.011.

García-Salamanca, A., Molina-Henares, M. A., Dillewijn, P. van, Solano, J., Pizarro-Tobías, P., Roca, A., *et al.* (2013). Bacterial diversity in the rhizosphere of maize and the surrounding carbonate-rich bulk soil. *Microbial Biotechnology* 6, 36–44. doi:https://doi.org/10.1111/j.1751-7915.2012.00358.x.

Good, B. H., McDonald, M. J., Barrick, J. E., Lenski, R. E., and Desai, M. M. (2017). The dynamics of molecular evolution over 60,000 generations. *Nature* 551, 45–50. doi:10.1038/nature24287.

Haichar, F. el Z., Marol, C., Berge, O., Rangel-Castro, J. I., Prosser, J. I., Balesdent, J., *et al.* (2008). Plant host habitat and root exudates shape soil bacterial community structure. *The ISME Journal* 2, 1221–1230. doi:10.1038/ismej.2008.80.



- Haney, C. H., Samuel, B. S., Bush, J., and Ausubel, F. M. (2015). Associations with rhizosphere bacteria can confer an adaptive advantage to plants. *Nat Plants* 1. doi:10.1038/nplants.2015.51.
- Hansel, C. M., Fendorf, S., Jardine, P. M., and Francis, C. A. (2008). Changes in bacterial and archaeal community structure and functional diversity along a geochemically variable soil profile. *Appl. Environ. Microbiol.* 74, 1620–1633. doi:10.1128/AEM.01787-07.
- Hassan, K. A., Johnson, A., Shaffer, B. T., Ren, Q., Kidarsa, T. A., Elbourne, L. D. H., et al. (2010). Inactivation of the GacA response regulator in *Pseudomonas fluorescens* Pf-5 has far-reaching transcriptomic consequences. *Environmental Microbiology* 12, 899–915. doi:https://doi.org/10.1111/j.1462-2920.2009.02134.x.
- Jones, D. L., Hodge, A., and Kuzyakov, Y. (2004). Plant and mycorrhizal regulation of rhizodeposition. *New Phytologist* 163, 459–480. doi:10.1111/j.1469-8137.2004.01130.x.
- Kram, K. E., Geiger, C., Ismail, W. M., Lee, H., Tang, H., Foster, P. L., et al. (2017). Adaptation of *Escherichia coli* to long-term serial passage in complex medium: Evidence of parallel evolution. *mSystems* 2. doi:10.1128/mSystems.00192-16.
- Kumar, H. D., and Alope, P. (2020). Role of biostimulant formulations in crop production: An Overview. *International Journal of Agricultural Sciences and Veterinary Medicine*. Volume 8
- Lee, J.-H., and Lee, J. (2010). Indole as an intercellular signal in microbial communities. *FEMS Microbiol Rev* 34, 426–444. doi:10.1111/j.1574-6976.2009.00204.x.
- Lemanceau, P., Blouin, M., Muller, D., and Moëne-Loccoz, Y. (2017). Let the core microbiota be functional. *Trends in Plant Science* 22, 583–595. doi:10.1016/j.tplants.2017.04.008.
- Lenski, R. E. (2017). Experimental evolution and the dynamics of adaptation and genome evolution in microbial populations. *The ISME Journal* 11, 2181–2194. doi:10.1038/ismej.2017.69.
- Lenski, R. E., and Travisano, M. (1994). Dynamics of adaptation and diversification: a 10,000-generation experiment with bacterial populations. *PNAS* 91, 6808–6814. doi:10.1073/pnas.91.15.6808.
- Lesuffleur, F., Paynel, F., Bataillé, M.-P., Le Deunff, E., and Cliquet, J.-B. (2007). Root amino acid exudation: measurement of high efflux rates of glycine and serine from six different plant species. *Plant Soil* 294, 235–246. doi:10.1007/s11104-007-9249-x.
- Li, C., Yan, K., Tang, L., Jia, Z., and Li, Y. (2014). Change in deep soil microbial communities due to long-term fertilization. *Soil Biology and Biochemistry* 75, 264–272. doi:10.1016/j.soilbio.2014.04.023.
- Loper, J. E., Hassan, K. A., Mavrodi, D. V., Li, E. W. D., Lim, C. K., Shaffer, B. T., et al. (2012). Comparative genomics of plant-associated *Pseudomonas* spp.: Insights into diversity and inheritance of traits involved in multitrophic interactions. *PLoS Genetics* 8, e1002784. doi:10.1371/journal.pgen.1002784.
- Lundberg, D. S., Lebeis, S. L., Paredes, S. H., Yourstone, S., Gehring, J., Malfatti, S., et al. (2012). Defining the core *Arabidopsis thaliana* root microbiome. *Nature* 488, 86–90. doi:10.1038/nature11237.
- Mulet, M., Bennasar, A., Lalucat, J., and García-Valdés, E. (2009). An *rpoD*-based PCR procedure for the identification of *Pseudomonas* species and for their detection in environmental samples. *Molecular and Cellular Probes* 23, 140–147. doi:10.1016/j.mcp.2009.02.001.
- Naik, P. R., Raman, G., Narayanan, K. B., and Sakthivel, N. (2008). Assessment of genetic and functional diversity of phosphate solubilizing fluorescent *Pseudomonads* isolated from rhizospheric soil. *BMC Microbiol* 8, 230. doi:10.1186/1471-2180-8-230.
- Niu, B., Paulson, J. N., Zheng, X., and Kolter, R. (2017). Simplified and representative bacterial community of maize roots. *Proc. Natl. Acad. Sci.* 114, E2450–E2459. doi:10.1073/pnas.1616148114.
- Ofek, M., Voronov-Goldman, M., Hadar, Y., and Minz, D. (2014). Host signature effect on plant root-associated microbiomes revealed through analyses of resident vs. active communities. *Environmental Microbiology* 16, 2157–2167. doi:https://doi.org/10.1111/1462-2920.12228.
- Peiffer, J. A., Spor, A., Koren, O., Jin, Z., Tringe, S. G., Dangl, J. L., et al. (2013). Diversity and heritability of the maize rhizosphere microbiome under field conditions. *Proc. Natl. Acad. Sci.* 110, 6548–6553. doi:10.1073/pnas.1302837110.
- Pérez-Jaramillo, J. E., Mendes, R., and Raaijmakers, J. M. (2016). Impact of plant domestication on rhizosphere microbiome assembly and functions. *Plant Molecular Biology* 90, 635–644. doi:10.1007/s11103-015-0337-7.

Preece, C., and Peñuelas, J. (2020). A return to the wild : root exudates and food security. *Trends in plant science* 25, 14–21. doi:10.1016/j.tplants.2019.09.010.

Sanguin, H., Remenant, B., Dechesne, A., Thioulouse, J., Vogel, T. M., Nesme, X., *et al.* (2006). Potential of a 16S rRNA-based taxonomic microarray for analyzing the rhizosphere effects of maize on *Agrobacterium* spp. and bacterial communities. *Applied and Environmental Microbiology* 72, 4302–4312. doi:10.1128/AEM.02686-05.

Schlaeppli, K., Dombrowski, N., Oter, R. G., Ver Loren van Themaat, E., and Schulze-Lefert, P. (2014). Quantitative divergence of the bacterial root microbiota in *Arabidopsis thaliana* relatives. *Proc. Natl. Acad. Sci.* 111, 585–592. doi:10.1073/pnas.1321597111.

Smith, P., and Schuster, M. (2019). Public goods and cheating in microbes. *Current Biology* 29, R442–R447. doi:10.1016/j.cub.2019.03.001.

Spaepen, S., Vanderleyden, J., and Remans, R. (2007). Indole-3-acetic acid in microbial and microorganism-plant signaling. *FEMS Microbiology Reviews* 31, 425–448. doi:10.1111/j.1574-6976.2007.00072.x.

Vacheron, J. (2015). Sélection des rhizobactéries phytostimulatrices par la plante : impact sur la distribution des propriétés phytobénéfiques chez les *Pseudomonas* fluorescents. <https://tel.archives-ouvertes.fr/tel-01433752/document>.

Vacheron, J., Moënné-Loccoz, Y., Dubost, A., Gonçalves-Martins, M., Muller, D., and Prigent-Combaret, C. (2016). Fluorescent *Pseudomonas* strains with only few plant-beneficial properties are favored in the maize rhizosphere. *Front. Plant Sci.* 7. doi:10.3389/fpls.2016.01212.

Vacheron, J., Péchy-Tarr, M., Brochet, S., Heiman, C. M., Stojiljkovic, M., Maurhofer, M., *et al.* (2019). T6SS contributes to gut microbiome invasion and killing of an herbivorous pest insect by plant-beneficial *Pseudomonas* protegens. *ISME J* 13, 1318–1329. doi:10.1038/s41396-019-0353-8.

Vandenkoornhuyse, P., Quaiser, A., Duhamel, M., Van, A. L., and Dufresne, A. (2015). The importance of the microbiome of the plant holobiont. *New Phytologist* 206, 1196–1206. doi:10.1111/nph.13312.

Weller, D. M. (2007). *Pseudomonas* biocontrol agents of soilborne pathogens: Looking back over 30 years. *Phytopathology* 97, 250–256. doi:10.1094/PHYTO-97-2-0250.

Wiser, M. J., Ribeck, N., and Lenski, R. E. (2013). Long-term dynamics of adaptation in asexual populations. *Science* 342, 1364–1367. doi:10.1126/science.1243357.

Xiao, Z., Ma, C., Xu, P., and Lu, J. R. (2009). Acetoin catabolism and acetylbutanediol formation by *Bacillus pumilus* in a chemically defined medium. *PLoS One* 4, e5627. doi:10.1371/journal.pone.0005627.

Yamamoto, S., Kasai, H., Arnold, D. L., Jackson, R. W., Vivian, A., and Harayama, S. (2000). Phylogeny of the genus *Pseudomonas*: intragenomic structure reconstructed from the nucleotide sequences of *gyrB* and *rpoD* genes. *Microbiology*, 146, 2385–2394. doi:10.1099/00221287-146-10-2385.

Zhang, Y., Zhang, B., Wu, X., and Zhang, L.-Q. (2020). Characterization the role of GacA-dependent small RNAs and RsmA family proteins on 2,4-diacetylphloroglucinol production in *Pseudomonas fluorescens* 2P24. *Microbiological Research* 233, 126391. doi:10.1016/j.micres.2019.126391.