



HAL
open science

Lifestyle of *Agrobacterium tumefaciens* in the tumor

Almudena González Mula

► **To cite this version:**

Almudena González Mula. Lifestyle of *Agrobacterium tumefaciens* in the tumor. *Phytopathology and phytopharmacy*. Université Paris Saclay (COmUE), 2017. English. NNT: 2017SACLS130. tel-03505903

HAL Id: tel-03505903

<https://theses.hal.science/tel-03505903v1>

Submitted on 1 Jan 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.

NNT : 2017SACLS130

THESE DE DOCTORAT
DE L'UNIVERSITE PARIS-SACLAY,
préparée à l'Université Paris-Sud

ÉCOLE DOCTORALE N° 567
Sciences du Végétal : du Gène à l'Ecosystème

Spécialité de doctorat : Biologie

Par

Mme. Almudena GONZÁLEZ MULA

Mode de vie d'*Agrobacterium tumefaciens* dans la tumeur

Thèse présentée et soutenue à Gif -sur-Yvette, le Jeudi 08 Juin 2017 :

Composition du Jury :

M, RATET, Pascal	Directeur de recherche CNRS, Gif-sur-Yvette,	Président du jury
Mme, HOMMAIS, Florence	Maître de Conférences, Univ de Lyon,	Rapporteur
M, NORMAND, Philippe	Directeur de recherche CNRS, Lyon,	Rapporteur
Mme, MASSON, Catherine	Directrice de recherche INRA, Toulouse,	Examinatrice
M, FAURE, Denis	Directeur de recherche CNRS, Gif-sur-Yvette,	Directeur de thèse



*Todo pasa y todo queda,
pero lo nuestro es pasar,
pasar haciendo caminos,
caminos sobre la mar.*

Antonio Machado

Acknowledgements

First, I would like to thank Florence Hommais and Philippe Normand for agreeing to evaluate my work. I would like also to thank the other members of the jury, Catherine Masson and Pascal Ratet. And finally, the members of my thesis committee, Marie-Anne Barny and Xavier Nesme for their discussions and advice.

I especially would like to thank Denis Faure for giving me the opportunity to do my PhD work in his team. Thank you very much for your dedication, since without it all this would not have been possible. And finally, thank you for your faith in me even when I did not have it.

Thanks to Yves Dessaux for your advice and your time, always with a big smile and a good mood. I would also like to thank Peter Mergaert, Benoît Alunni, Catherine Grandclement, Tania Timtchenko, and Nicolas Mirouze for your help and support.

Thanks to Pauline simply for being you. Thank you for sharing this experience with me. For all the laughs, the tears, moments shared, for your friendship, your love and your immense affection.

Thanks to all the team for your support, advice and care. Thank you Jérémy, Florian, Quentin, Kévin and Joy, for the laughs, the beers, your support and concern especially at the end of this work. Thanks also to all the members who have gone through the team: Julien for your advice and your special humor, Anthony for your generosity, Yannick for all the shared laughs, Fabienne, Nadia, Claudie and Elisa for your friendship and affection, and all other team members: Marta, Slimane, Said, Quentin, Leo, Samuel, Noura, Teik Min, Kar Wai, Audren and all the other people who passed along these nearly 4 years.

To all members of the old ISV and the new I2BC: administrative services, informatics services, technical service, greenhouse, and laundry service that do a great job which makes it easy to work with.

Many thanks to Yves for all your love. Thank you for not letting me surrender when I did not believe it possible. Thanks for being by my side. Thanks also to your family for its great affection.

Y sobre todo y muy especialmente a mis padres y a mi hermana, muchísimas gracias por todo el apoyo, la confianza y el cariño incluso en la distancia. Gracias por estar SIEMPRE ahí.

ABBREVIATIONS

3OC8HSL: N-3-oxo-octanoyl homoserine lactone
ABm: *Agrobacterium* Broth minimal medium
AND-T: Transfer DNA
At: *Agrobacterium tumefaciens*
CDS: Coding sequence
DNA: Deoxyribonucleic acid
GABA: γ -Aminobutyric acid
GHB: γ -Hydroxybutyrate
Gm: Gentamicin
IAA: Indole-3-acetic acid
LBm: Luria Bertani modified
Mb: Megabase
NAHL: N-acyl-homoserine lactone
OD₆₀₀: Optical density (measured at a wavelength of 600 nm)
PAMP: Pathogen-associated molecular patterns
pAt: At plasmid
PBP: Periplasmic binding protein
PCR: Polymerase chain reaction
pTi: Ti plasmid
QS: Quorum-sensing
Rif: Rifampicin
RNA: Ribonucleic acid
ROS: Reactive oxygen species
SSA: Succinic semialdehyde
TCA: Tricarboxylic Acid
Ti: Tumor-inducing
Tn-Seq: Transposon sequencing
CFU: Colony-forming unit

INDEX

Avant propos (Synthèse en français).....	1
Chapter I: Introduction.....	3
1. <i>Agrobacterium tumefaciens</i> species.....	5
2. <i>The Agrobacterium tumefaciens C58 genome</i>	5
3. <i>A. tumefaciens and host plant interactions</i>	8
4. <i>Bibliography</i>	17
Chapter II	23
1. INTRODUCTION	27
2. EXPERIMENTAL PROCEDURES	29
3. RESULTS AND DISCUSSION.....	33
4. CONCLUSIONS	51
5. BIBLIOGRAPHY	52
Chapter III	57
1. INTRODUCTION	61
2. MATERIAL AND METHODS	62
3. RESULTS	67
4. DISCUSSION	77
5. BIBLIOGRAPHY	79
Chapter IV	81
1. INTRODUCTION	85
2. MATERIAL AND METHODS	85
3. RESULTS	91
4. DISCUSSION.....	105
5. CONCLUSIONS	107
6. BIBLIOGRAPHY	108
Chapter V: Discussion	115
1. Transcriptome <i>in planta</i>	117
2. Transcriptome <i>in vitro</i> and Tn-Seq analysis of GHB and GABA cultures	121
3. Experimental evolution in <i>A. tumefaciens</i> C58	123
PERSPECTIVES	125
BIBLIOGRAPHY	126
ANNEXES.....	129

AVANT PROPOS

(Synthèse en français)

Agrobacterium tumefaciens, l'agent de la galle du collet (tumeur végétale) est bien connu pour sa capacité à transformer certaines cellules eucaryotes avec une partie de son ADN génomique, l'ADN-T. Cette découverte a été la source d'une avancée biotechnologique sans précédent, permettant pour la première fois d'obtenir des organismes végétaux génétiquement modifiés. Depuis lors, les premières étapes de l'interaction entre la bactérie et sa plante hôte ont fait l'objet de nombreuses études visant à identifier les principaux acteurs impliqués dans ce processus de transfert de gènes. Malgré ces nombreuses œuvres et découvertes, *A. tumefaciens* continue d'être un organisme fascinant à étudier pour les scientifiques.

Dans ce manuscrit de thèse, nous présentons de nouveaux aspects des interactions entre *A. tumefaciens* et les plantes hôtes. La première partie consistera en une analyse transcriptomique globale des cellules d'*A. tumefaciens* provenant d'une tumeur végétale. Ce travail a permis de démontrer l'état cellulaire d'*A. tumefaciens* lorsqu'il colonise la tumeur de la plante.

La deuxième partie analysera le rôle du GHB et du GABA lorsqu'il est utilisé comme source d'énergie pour les bactéries. À l'aide de deux méthodes différentes, transcriptomiques et Tn-Seq, nous avons pu proposer une voie de dégradation et d'assimilation du GHB.

La troisième partie traitera d'une expérience d'évolution par des passages en série de trois souches différentes du pathogène sur la plante hôte *Solanum lycopersicum*. Cette expérience a démontré la capacité de cette bactérie à évoluer dans une plante hôte et à augmenter sa valeur sélective.

Enfin, dans la dernière partie, nous reprendrons et discuterons des principaux résultats de cette thèse.

Chapter I

Chapter I: General Introduction

1. *Agrobacterium tumefaciens* species

Agrobacterium tumefaciens is a soil-borne plant pathogen. It is responsible for the crown gall disease, characterized by tissue overgrowths, commonly called tumor. *A. tumefaciens* shows a broad host plants spectrum, among more than 90 families including most dicotyledonous (Escobar and Dandekar, 2003).

1.1. Classification

Bacteria of the genus *Agrobacterium* are α -*Proteobacteria* belonging to the *Rhizobiaceae* family. The first taxonomy studies classified them based on pathogenicity tests (Tzfira and Citovsky, 2008). All pathogenic bacteria inducing crown gall were named *A. tumefaciens*, *A. rubi* induced the raspberry gall, *A. rhizogenes* induced hairy roots, *A. vitis* the gall of the grapevine, *A. larrymoorei* induced the gall of the fig tree and finally all non-pathogenic strains were named *A. radiobacter*. Most recent studies, based on the analysis of polymorphism and phylogenetic relationships, have ordered *Agrobacterium* strains in different genomic groups (Costechareyre *et al.*, 2010; Mougel *et al.*, 2002). According to this classification, *A. tumefaciens* is a complex composed by ten distinct genomic species (G1 to G9, G13). The model strain *A. tumefaciens* C58 belongs to G8 genomic species. The focus of recent research has been on identifying specific genes for each of the defined genomic groups and, in some cases, the characterization of possible ecological adaptations (Lassalle *et al.*, 2011). For these reasons, it has been proposed, but still not approved, that the genomic group G8 forms a fully-fledged species, renamed *Agrobacterium fabrum*.

2. The *Agrobacterium tumefaciens* C58 genome

A. tumefaciens C58 genome is composed by four replicons: a circular chromosome of about 2.8 Mb (*atu0001* to *atu2835* genes), a linear chromosome of approximately 2 Mb (*atu3000* to *atu4898* genes) and two dispensable plasmids: the At (*Agrobacterium tumefaciens*) plasmid of around 0.5 Mb (*atu5000* to *atu5549* genes) and the Ti (tumor-inducing) plasmid of about 0.2 Mb (*atu6000* to *atu6197* genes). Although the Ti plasmid is not essential for bacterial survival, it is necessary for the pathogenicity of the microorganism. The Ti plasmid contains the T-DNA that is transferred to the plant cell during infection and the *vir* genes encoding the T-DNA transfer machinery. In addition, there are other important genes, such as the operons linked to the catabolism of opines

and the *trb* regulon, controlled by *tral* and *traR* genes of the quorum-sensing (QS) system, which strongly regulates the Ti plasmid copy number and its conjugation to other *Agrobacteria* (Piper *et al.*, 1993).

In comparison, At plasmid has been less thoroughly studied. This plasmid, able to replicate thanks to the *repABC* system, also contains all the genes necessary for its own conjugation (Chen *et al.*, 2002). The complete genome of *A. tumefaciens* C58 (Goodner *et al.*, 2001; Wood *et al.*, 2001) describes that the At plasmid codes for many transport systems and is enriched with transposable elements. Not only the At plasmid is important for the ecology of the bacteria, since it contains genes involved in the degradation of molecules present in plants and rhizosphere, such as γ -butyrolactone (GBL) and γ -hydroxybutyrate (GHB) (Chai *et al.*, 2007) and deoxy-fructosyl glutamine (DFG) (Baek *et al.*, 2005); but also seems to have a positive effect on the virulence of the Ti plasmid (Matthysse *et al.*, 2008; Nair *et al.*, 2003).

2.1. Species-specific genes

As mentioned above, *A. tumefaciens* C58 belongs to the G8 genomic group. As shown recently by (Lassalle *et al.*, 2011), there are 196 genes specific to G8 genomovar which seem to play a role in the adaptation of the bacteria to different environments. They are located in 7 different genomic islands. Almost all of them (72% of total genes) were found on the linear chromosome, and only one of these regions was found on the circular chromosome (28% of total genes), suggesting a widely plasticity of the linear chromosome. These clusters respond to various functions, as the production of secondary metabolites, the detoxification or catabolism of different compounds.

Three of these clusters are implicated in catabolism of carbohydrates: SpG8-4, SpG8-1a and SpG8-5. This last group of genes contains an alanine racemase that could be involved in the catabolism of opines or related compounds or Amadori compounds, derivatives of aminodeoxysugars present in decaying plant materials.

SpG8-3 cluster is the most extensive of the 7 and appears to be involved in the biosynthesis, release and uptake of siderophores (Rondon *et al.*, 2004). The results suggest that this cluster is induced in iron limiting condition.

SpG8-6 group of genes has also been investigated. Its main function seems to be the detoxification of various compounds, since it contains 3 putative multidrug transport systems. Among them the *tetR-tetA* system, which seems to confer resistance to tetracycline, has been well characterized. *tetA* codes for an efflux pump able to confer

resistance to tetracycline, and *tetR* is the repressor that inhibits the expression of *tetA* (Luo and Farrand, 1999).

SpG8-2a genomic group principal function is the biosynthesis of an exopolysaccharide called curdlan (Lassalle *et al.*, 2011). Curdlan is a water insoluble β -D-1, 3-glucan produced by the *Agrobacterium* genus. In *A. tumefaciens* ATCC31749, curdlan could be crucial for biofilm formation under oxidative stress/low pH and/or limited nitrogen with abundant sugar, indicating that the biosynthesis of this exopolysaccharide could have an evolutionary origin (Yu *et al.*, 2015).

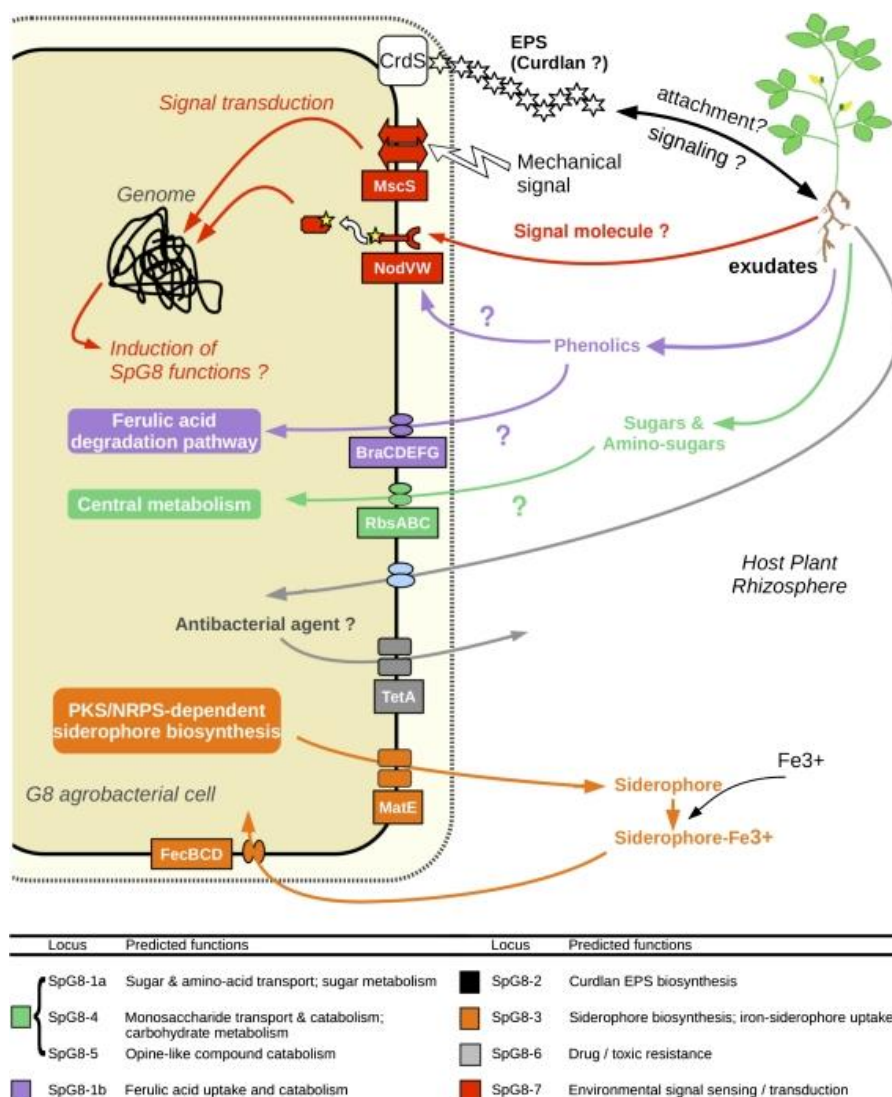


Figure 1. Hypothetical integrated functioning of SpG8 genes. From Lassalle *et al.*, 2011.

3. *A. tumefaciens* and host plant interactions

3.1. Plant transformation

A. tumefaciens and plant cell interactions consist on several steps: recognition, attachment to the host cell, virulence (*vir*) gene expression, preparation and transfer of T-DNA and finally, the integration of T-DNA into the plant cell genome (Sheng and Citovsky, 1996).

When a plant is wounded, it synthesizes and releases into the rhizosphere phenolic compounds such acetosyringone that attract *Agrobacteria* by chemotaxis to this area. *A. tumefaciens* does not invade plant cells but it attaches to the cell wall. This attachment to the host plant requires different molecules. Three chromosomal genes have been identified as important for this process. This is *chvA*, *chvB* and *pscA* involved in the synthesis and/or localization of beta-1,2 glucans (Pitzschke and Hirt, 2010). On the other hand, several *A. thaliana* mutants resistant to the infection of *A. tumefaciens* has been described (Zhu *et al.*, 2003). These mutants are named *rat* (recalcitrant to *Agrobacterium* transformation). Additional analysis of these mutants should help find out how host plants can recognize *Agrobacterium* cells.

virA and *virG* genes encode a two-component signal system that can be activated by various phenolic compounds, acidic pH, low phosphate concentration or certain sugars, but phenolic compounds seem to be the only signal completely required. There are many monosaccharides diffuse by the wound site and are recognize by the *chvE* gene (located on the circular chromosome). This gene encodes a periplasmic sugar binding protein. As soon as sugars are attached to ChvE, ChvE activates the VirA-VirG system (Hu *et al.*, 2013).

Once activated, the VirA-VirG system induces the expression of *vir* genes which generally consist of 5 different complexes: *virB*, *virC*, *virD*, *virE* and *virF* (Gelvin, 2012; Pitzschke and Hirt, 2010). The VirB complex, a type IV secretion system (T4SS), is necessary for virulence and it is constituted by at least 12 different proteins. The T-DNA, before being transferred, is prepared as a single-stranded molecule. The VirD1 and VirD2 nucleases are responsible for this process. VirD2 will then remain attached to the 5' end of the T-DNA, which facilitates its transport. This complex is called a T-strand. Once the T-strand is translocated into the plant cell by the type 4 secretion system, this DNA fragment is to be directed to the plant nucleus for further integration into the host genome (Pitzschke and Hirt, 2010). At the nucleus, T-DNA is randomly integrated into the plant genome, but with a preference for transcriptionally active chromatin or gene-

rich regions (Kim *et al.*, 2007).

3.2. Disease process: tumorigenesis

Once the T-DNA is integrated into the genome of the host plant, genes located on it are expressed. Two classes of genes are present in the T-DNA: oncogenes and opine-related genes. Oncogenes are involved in biosynthesis of the phytohormones auxin and cytokinin. *iaaH* and *iaaM* codes for a indole-3-acetamide hydrolase and tryptophan monooxygenase, respectively, and they contribute to the production of 3-indole acetic acid (IAA). Besides, the *ipt* gene codes for a isopentenyl transferase that catalyzes the synthesis of cytokinin (Tzfira and Citovsky, 2008). The accumulation of auxin and cytokinin in plant tissue creates a hormonal imbalance leading to cell dedifferentiation and the tumor development (Veselov *et al.*, 2003). The second class of T-DNA genes are involved in the synthesis of low molecular weight amino acids and sugar phosphate derivatives called opines (Dessaux *et al.*, 1998). It exists more that 20 different opines depending on the Ti plasmid carried by *Agrobacterium* strains (Dessaux *et al.*, 1993). *A. tumefaciens* C58 carries a Ti plasmid of nopaline T-DNA type. The synthesis of nopaline and nopalique acid are provided by the *nos* gene (nopaline synthase), whereas the production of agrocinopines (agrocinopine A and B) is due to gene *acs* (agrocinopines synthase). Hence, the production of opines in transformed plant cells supposes a distinct ecological niche for each infecting strain of *Agrobacterium*. Therefore, although opines are not directly involved in tumor formation, they provide a source of nutrients for the bacteria as well as help the transfer of conjugal Ti plasmid and chemotaxis (Kim *et al.*, 2001).

In addition to the production of hormones and opines, an interesting feature of the tumor formation is the development of a new vascular system within the tumor. This fact had already been extensively studied in *Ricinus communis* L plant (Aloni *et al.*, 1995) and recently been observed in *Arabidopsis thaliana* tumors (Lang, Gonzalez-Mula, *et al.*, 2016) induced by *A. tumefaciens*. Neovascularization is a critical step in tumor genesis and more evidence that the ethylene molecule whose production is stimulated by auxin and cytokinin plays a crucial role here (Wächter *et al.*, 2003). Another hormone that seems to be important in tumor development is abscisic acid. Tumors show marked evaporation. It has now been established that abscisic acid that accumulates in tumors promotes the synthesis of osmoprotectants (including proline) and suberin that limit the effects of water stress (Efetova *et al.*, 2007).

3.2.1. Quorum sensing

The quorum-sensing (QS) is a mechanism connecting cell population to gene expression via N-acylhomoserine lactones (NAHLs) (Fuqua *et al.*, 1994). LuxI/LuxR system of *Vibrio fischeri* was the first QS system described (Nealson *et al.*, 1970). It has described many homologues of LuxI and LuxR proteins in other bacterial species like *Pseudomonas aeruginosa* and *A. tumefaciens*. The quorum-sensing system of *A. tumefaciens* C58 is one of the most studied, using this bacterium as model of study.

As described before, after transfer and integration of the T-DNA, plant cells synthesize opines. Some opines are required for the synthesis of NAHLs (Piper *et al.*, 1999). The TraI/TraR quorum sensing system is homologous to LuxI/LuxR, and is located on the Ti plasmid. *traI* gene (*atu6042*) codes for an acyl-homoserine-lactone synthase, which produces the N-3-oxo-octanoyl homoserine (3OC8HSL) signal molecule (Hwang *et al.*, 1994). This gene is located on the *trb* operon and is induced by agrocinopines A and B (Piper *et al.*, 1999), more precisely their by-product arabinose-2-P (El Sahili *et al.*, 2015). TraR (*atu6134*), the transcriptional activator protein can sense 3OC8HSL signals produced by *A. tumefaciens* C58 and, and is located on the *arc* operon. At high cell density, the QS signals reach a threshold concentration and are then perceived by the TraR receptors. The TraR/OC8HSL complex activates the transcription of the *trb* operon, containing the *traI* gene, which means TraR/OC8HSL complex induces the synthesis of the QS signals. TraR acts also a transcriptional regulator and activates the expression of QS target genes (Qin *et al.*, 2000).

QS signaling pathway in *A. tumefaciens* C58 is controlled by two transcriptional regulators, TraR (*atu6134*) and AccR (*atu6138*). TraR is activated by de-repression or by activation. Regarding AccR, it can contribute to the repression of the QS regulation system. AccR repressor suppresses the expression of the *acc* operon, involved in the transport and catabolism of opines, and the *arc* operon, including the *traR* gene. In the presence of agrocinopines, i.e. in tumors, repression by AccR is detached and the transcription of both operons *acc* and *arc*, and therefore *traR* is induced. Agrocinopines bind to AccR and reduce drastically the affinity of AccR for the promoter sequence of the *acc* operon. The expression of the transcriptional regulator TraR is therefore inhibited by the AccR repressor in the absence of opines.

An *accR* mutant overproduces OC8HSL because the expression of TraR is no longer suppressed and the expression of the operon *tra* is induced. This mutant can also transfer its Ti plasmid constitutively, because it not necessary any longer the presence

of opines to remove the repression of TraM (*atu6131*). Hence, this regulator plays a crucial role in controlling signaling and optimizing the expression of QS-regulated bacterial functions during interaction with the host plant.

In *A. tumefaciens*, QS controls different important functions. The first one is the control of the copy number of the Ti plasmid (Li and Farrand, 2000), by increasing the expression of *repA* gene. The *repABC* operon, adjacent to the *trb* operon, is responsible for the replication of the Ti plasmid, although the transcription of the *rep* genes does not depend entirely on the QS regulation. The TraR/3OC8HSL complex then stimulates the expression of *rep* operon which leads to an increase in the number of copies of the pTi, (Pappas, 2008). A mutation in AccR, the *traR* repressor, also results in an increase in the copy number of the pTi (Li and Farrand, 2000). Another function of QS in *A. tumefaciens* is an increase of its aggressiveness due to increment in the Ti plasmid copy number (Pappas, 2008). Finally, the last function controlled by QS is the Ti plasmid transfer. In *Rhizobiaceae*, the conjugation system is generally encoded by the *tra* and *trb* genes, whose expression is activated by TraR. The cluster *tra* contains the transfer origin (*oriT*) between the two divergent operons *traAFB* and *traCDG*. The *tra* operon is involved in the synthesis of the type IV secretion system and the *trb* operon in the transfer of the Ti plasmid between two bacteria (Li *et al.*, 1999). The TraR/OC8HSL complex activates the *tra* and *trb* operons encoding the proteins necessary for the conjugation mechanism.

A. tumefaciens C58 harbors two lactonases, BlcC (*atu5139*) and AiiB (*atu6071*), which have the capacity to degrade QS signals by opening the lactone ring of the OC8HSL molecule. BlcC lactonase is part of the *blcABC* operon located on the At plasmid. Its expression is controlled by the transcriptional regulator *blcR* (*atu5136*, also called AttJ) located upstream of the operon. This operon is involved in the assimilation of gamma-butyrolactone (GBL) and gamma-hydroxybutyrate (GHB)(Carlier *et al.*, 2004).

GABA is produced in high different biotic or abiotic stress conditions. It has been demonstrated the induction of *blcC* operon by this compound (Chevrot *et al.*, 2006). GABA, in bacterial cells, is converted to a semialdehyde succinic (SSA) which reduces BlcC repression and thus activates the transcription of the operon and consequently induces QS signal degradation. *blcC* gene is found in many strains of *Agrobacterium* (Haudecoeur *et al.*, 2009; Khan and Farrand, 2009). The high frequency of *blcC* in agrobacterial populations could be explained, among other things, by the involvement of the *blcABC* operon in the detoxification of molecules produced by plants such as SSA or

GHB. GABA produced in plant tumors is transported into the bacteria by two specific and non-specific transporters: ABC transporter Bra, is involved in the uptake of GABA and Proline (Planamente *et al.*, 2010) and ABC transporter Gts, which does not compete with any other compound (Planamente *et al.*, 2012). A recent study revealed that altering the ratio GABA:Proline in *A. thaliana* tumors, we can modify the expression of the BlcC lactonase and reduce the dissemination of the Ti plasmid (Lang, Gonzalez-Mula, *et al.*, 2016).

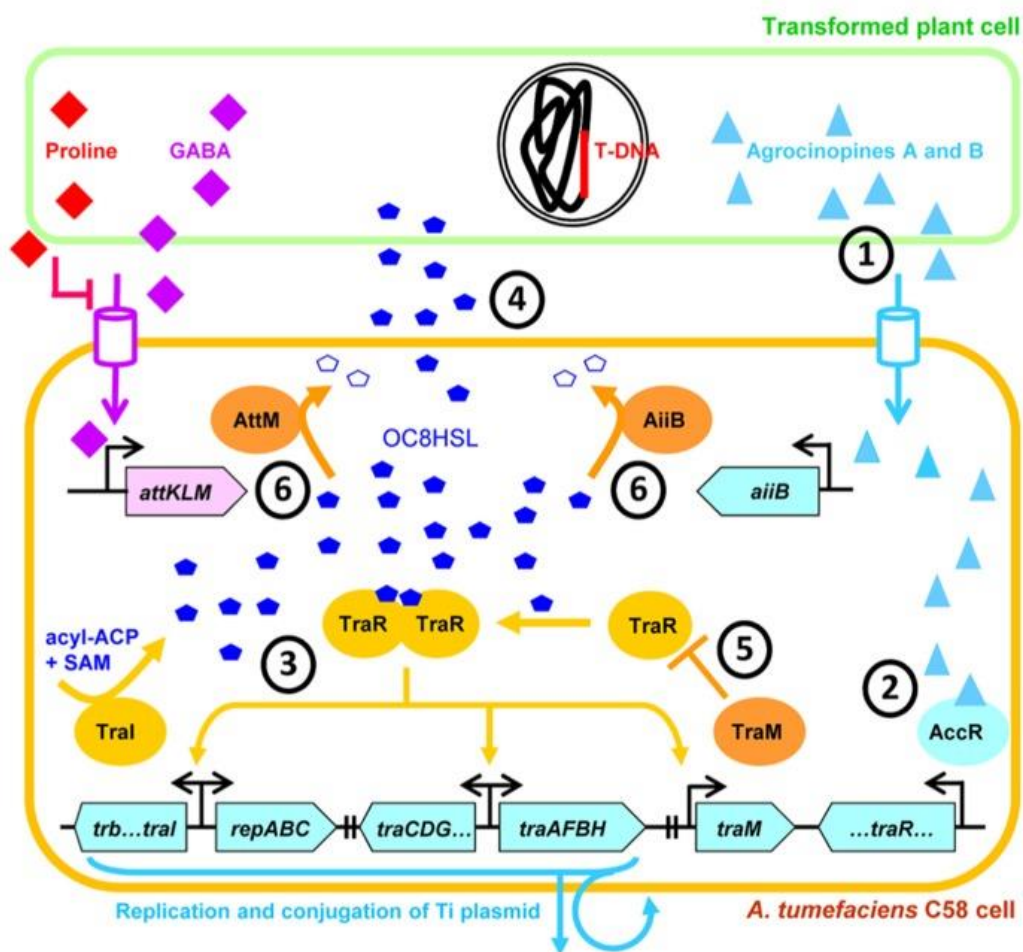


Figure 2. Representation of the QS regulation in *A. tumefaciens* C58 strain. From Lang and Faure, 2004.

3.2.2. *Opines catabolism*

Tumors possess a particular metabolism, since in the transformed tissues are synthesized specific compounds of low molecular weight called opines. These molecules are products of the condensation of simple sugars, organic acids and amino acids. Its production in tumors is triggered by *A. tumefaciens* and they are used by the bacterium as an important source of nitrogen and energy (Dessaux *et al.*, 1998).

Studies of several tumors induced by different strains of *A. tumefaciens* have shown that the nature of opines produced in tumors can be very varied. To date more than 20 different opines have been identified. 7 different groups have been established depending on which opines that can synthesize their T-DNA: octopine, nopaline, agropine, succinamopine, lippie, chrysopine/succinamopine and chrysopine/nopaline.

Opines are synthesized in tumors by plant cells transformed by genes located in the T-DNA transferred to the plant from the bacteria. On the other hand, genes required for their catabolism are in a region of the Ti plasmid that is not transferred to the plant. In general, there is a close relationship between the plasmid that causes the transformation and the ability of this plasmid to metabolize only the same types of opines as those it induces (Dessaux *et al.*, 1998). Thus, in the octopine-type plasmids, the Occ operon allows the assimilation of octopine, octopinic acid and lysopine. In nopaline-type plasmids, the Noc region is responsible for the importation and degradation of nopaline, and the transport and catabolism of agrocinopines are ensured by the products of the Acc operon. The expressions of the Occ, Noc and Acc regions are controlled by the transcriptional regulators OccR, NocR and AccR respectively. The promoters of these three regions are also inducible by their corresponding opines (Dessaux *et al.*, 1998).

3.2.3. *The opine concept*

This concept assigns an ecological role to opines and establishes that the transformation of plant tissues, which leads to the production of opines, creates a favorable niche for the strains that possess the genes of their catabolism, favoring their clonal multiplication. In addition, by inducing the conjugative transfer of the Ti plasmid, they favor the dissemination of the same to initially non-pathogenic strains (Petit *et al.*, 1978).

Subsequently it has been found that the opines are also chemoattractants of *Agrobacterium*, which allow the strains to explore the surroundings of the tumors (Kim and Farrand, 1998).

Although some bacterial species different from *Agrobacteria*, such as *Pseudomonas sp.*, are also capable of assimilating opines (Nautiyal *et al.*, 1991), this concept would explain the theoretical advantages for the survival of pathogenic strains within tumors and possibly competing with other bacteria lacking such a system. A recent research has demonstrated that opines confer that selective advantage in tumor environments. This study shows that mutants affected in the binding (*nocT*) and assimilations (*oca*) of nopaline in *A. tumefaciens* C58 are less competitive we co-inoculate them with the wild-type strain in tomato tumors (Lang *et al.*, 2014). In addition, wild-type Ti plasmids were preferentially transfer than nopaline-mutants Ti plasmids, even if the Ti-plasmid transfer is not regulated by nopaline. These results reveal an advantageous transfer of this Ti-plasmid amongst the *Agrobacteria* that colonize plant tumors.

Although the concept of opine has been validated, latest research has shown that having a comprehensive opine assimilation mechanism does not always lead to a selective advantage of these populations. In this work, it was shown that although the NocT transport protein can bind both nopaline and octopine, it does not confer an advantage on other octopine-type Ti plasmid strains when they are co-infected (Lang, Vigouroux, *et al.*, 2016).

The predominance of non-pathogenic strains in soil would imply either that the Ti plasmid must carry genes that would make the bacterium more susceptible to physical or biological stresses outside of the tumors or that some genes directly involved in tumor induction would produce that decrease in the adaptability of the strains in the soil. Consequently, loss of the plasmid or significant mutations in certain genes would increase its competitiveness in this habitat, causing its predominance.

3.3. Symptoms of the disease

According to the tumor characteristics, plants infected with *A. tumefaciens* may exhibit various symptoms associated with slow growth. By inhibiting the physiological functions of the plant such as the transport of water and nutrients, tumors are responsible for a significant decrease in the quantity and quality of plant production. In addition, tumors are often easy entry points for other pathogens of the plant.

To define the host range of *A. tumefaciens*, we should distinguish the host that *A. tumefaciens* can transform with its T-DNA from the host in which it causes the crown gall disease. In fact, not only plant species are natural hosts for the transfer of T-DNA, the transformation of animal cells and fungi have also been carried out under laboratory conditions (Lacroix *et al.*, 2006).

It has been identified more than 600 plant species, belonging to 331 genera and 93 different families, susceptible to infection by *A. tumefaciens* (De Cleene and De Ley, 1976). The clear majority of these species are dicotyledons, although some monocotyledons may also be hosts of the pathogenic bacterium. This definition of the host range of *A. tumefaciens* must nevertheless be qualified by the fact that there is a relatively large diversity in the pathogenic species and that consequently certain strains of *A. tumefaciens* could specifically induce the disease in several specific hosts. More generally, the molecular factors involved in infectious and tumor development processes are major determinants of the ability of an *A. tumefaciens* strain to infect a given plant.

Economically, crown gall frequently cause significant losses in nursery production of fruit trees, rose and vines, in several countries of the world. It is the main bacterial disease of stone fruit trees in the Mediterranean countries (Pionnat *et al.*, 1999). To combat this damage, different detection strategies (by PCR for example) and biocontrol (infection with *A. tumefaciens* antagonist strains) have been and are still being developed (Dandurishvili *et al.*, 2011; Puławska and Sobiczewski, 2005).

3.4. Plant responses to *A. tumefaciens* infection

Once *A. tumefaciens* has infected plant cells, tumors evolve from aerobic and autotrophic system into anaerobic and heterotrophic system. Genes involved in photosynthetic electron transport are downregulated and energy production switches to fermentation (Deeken *et al.*, 2006). They accumulate many compounds including anions, sugars and amino acids. This increase also correlates with the expression profiles of different genes encoding enzymes and transporters of these key metabolic pathways (Ditt *et al.*, 2005). Early studies of plant defense responses to *A. tumefaciens* suggest that these responses were completely inhibited in the early stages of infection. Nevertheless, recent studies confirm it. It has been demonstrated that expression of defense genes could be both induced and repressed when analyzing the transcriptome of *A. thaliana* cell cultures 48 hours after *A. tumefaciens* infection (Ditt *et al.*, 2006). Another study focused on hormone pathways in plants demonstrated that after the integration of the T-DNA into the *A. thaliana* genome, the levels of ethylene and salicylic acid in the transformed cells increased continuously during tumor development (Lee *et al.*, 2009). However, this increase was not correlated with the activation of the usual signaling pathways of these two hormones. Thus, even if the accumulation of salicylic acid in the transformed cells could cause a limitation of *A. tumefaciens* disease development, this defense process remained independent of the induction of the defense genes usually regulated by salicylic acid and did not involve the implementation of Systemic Acquired Resistance (SAR). Another important result of this study was the lack of accumulation of jasmonate in *A. thaliana* tumors. Finally, the possibility that *A. tumefaciens* induced an immune response of the PTI (PAMPs-Triggered Immunity) type in the host plant was explored. PTI-type immunity is based on the ability of plants to recognize certain molecular motifs associated with pathogens. It involves PRR (Pattern Recognition Receptor), receptors composed of an extracellular LRR domain (Leucine-Rich Repeat) and an intracellular kinase domain. Bacterial flagellins are typically potent PTI elicitors in plants. Surprisingly, flagellins of *A. tumefaciens* are not recognized by the PTI receptors of *A. thaliana*. On the other hand, plants harboring the EFR receptor (EF-Tu receptor) can perceive the elongation factor EF-Tu of *A. tumefaciens* and trigger a defense cascade resulting in a restriction of the transformation efficiency by T-DNA (Zipfel *et al.*, 2006).

Altogether, data mentioned above try to describe an original phytopathogenic process in which interactions between *A. tumefaciens* and the host plant reveals some plant defense responses limitation, thus allowing a cohabitation of both organisms in tumors.

4. Bibliography

- Aloni, R., Pradel, K. and Ullrich, C.** (1995) The three-dimensional structure of vascular tissues in *Agrobacterium tumefaciens*-induced crown galls and in the host stems of *Ricinus communis* L. *Planta* **196**, 597–605.
- Baek, C.H., Farrand, S.K., Park, D.K., Lee, K.E., Hwang, W. and Kim, K.S.** (2005) Genes for utilization of deoxyfructosyl glutamine (DFG), an amadori compound, are widely dispersed in the family *Rhizobiaceae*. *FEMS Microbiol. Ecol.* **53**, 221–233.
- Carlier, A., Chevrot, R., Dessaux, Y. and Faure, D.** (2004) The assimilation of gamma-butyrolactone in *Agrobacterium tumefaciens* C58 interferes with the accumulation of the N-acyl-homoserine lactone signal. *Mol. Plant. Microbe. Interact.* **17**, 951–7.
- Chai, Y., Ching, S.T., Cho, H. and Winans, S.C.** (2007) Reconstitution of the biochemical activities of the AttJ repressor and the AttK, AttL, and AttM catabolic enzymes of *Agrobacterium tumefaciens*. *J. Bacteriol.* **189**, 3674–3679.
- Chen, L., Chen, Y., Wood, D.W. and Nester, E.W.** (2002) A new type IV secretion system promotes conjugal transfer in *Agrobacterium tumefaciens*. *J. Bacteriol.* **184**, 4838–4845.
- Chevrot, R., Rosen, R., Haudecoeur, E., Cirou, A., Shelp, B.J., Ron, E. and Faure, D.** (2006) GABA controls the level of quorum-sensing signal in *Agrobacterium tumefaciens*. *Proc. Natl. Acad. Sci.* **103**, 7460–7464.
- Cleene, M. De and Ley, J. De** (1976) The host range of crown gall. *Bot. Rev.* **42**, 389–464.
- Costechareyre, D., Rhouma, A., Lavire, C., Portier, P., Chapulliot, D., Bertolla, F., Boubaker, A., Dessaux, Y. and Nesme, X.** (2010) Rapid and Efficient Identification of *Agrobacterium* Species by recA Allele Analysis. *Microb. Ecol.* **60**, 862–872.
- Dandurishvili, N., Toklikishvili, N., Ovadis, M., et al.** (2011) Broad-range antagonistic rhizobacteria *Pseudomonas fluorescens* and *Serratia plymuthica* suppress *Agrobacterium* crown gall tumours on tomato plants. *J. Appl. Microbiol.* **110**, 341–352.
- Deeken, R., Engelmann, J.C., Efetova, M., et al.** (2006) An Integrated View of Gene Expression and Solute Profiles of *Arabidopsis* Tumors: A Genome-Wide Approach. *Plant Cell Online* **18**, 3617–3634.
- Dessaux, Y., Petit, A., Farrand, S.K. and Murphy, P.J.** (1998) Opines and Opine-Like Molecules Involved in Plant-*Rhizobiaceae* Interactions. In *The Rhizobiaceae.*, pp. 173–197. Dordrecht: Springer Netherlands.
- Dessaux, Y., Petit, A. and Tempe, J.** (1993) Chemistry and biochemistry of opines, chemical mediators of parasitism. *Phytochemistry* **34**, 31–38.

- Ditt, R.F., Kerr, K.F., Figueiredo, P. de, Delrow, J., Comai, L. and Nester, E.W.** (2006) The *Arabidopsis thaliana* transcriptome in response to *Agrobacterium tumefaciens*. *Mol. Plant. Microbe. Interact.* **19**, 665–81.
- Ditt, R.F., Nester, E. and Comai, L.** (2005) The plant cell defense and *Agrobacterium tumefaciens*. *FEMS Microbiol. Lett.* **247**, 207–213.
- Efetova, M., Zeier, J., Riederer, M., Lee, C.-W., Stingl, N., Mueller, M., Hartung, W., Hedrich, R. and Deeken, R.** (2007) A central role of abscisic acid in drought stress protection of *Agrobacterium*-induced tumors on *Arabidopsis*. *Plant Physiol.* **145**, 853–62.
- Escobar, M.A. and Dandekar, A.M.** (2003) *Agrobacterium tumefaciens* as an agent of disease. *Trends Plant Sci.* **8**, 380–386.
- Fuqua, W.C., Winans, S.C. and Greenberg, E.P.** (1994) MINIREVIEW Quorum Sensing in Bacteria: the LuxR-LuxI Family of Cell Density-Responsive Transcriptional Regulators. *J. Bacteriol.* **176**, 269–275.
- Gelvin, S.B.** (2012) Traversing the cell: *Agrobacterium* T-DNA's journey to the host genome. *Plant-Microbe Interact.* **3**, 52.
- Goodner, B., Hinkle, G., Gattung, S., et al.** (2001) Genome sequence of the plant pathogen and biotechnology agent *Agrobacterium tumefaciens* C58. *Science* **294**, 2323–8.
- Haudecoeur, E., Tannieres, M., Cirou, A., Raffoux, A., Dessaux, Y. and Faure, D.** (2009) Different regulation and roles of lactonases AiiB and AttM in *Agrobacterium tumefaciens* C58. *Mol Plant Microbe Interact* **22**, 529–537.
- Hu, X., Zhao, J., DeGrado, W.F. and Binns, A.N.** (2013) *Agrobacterium tumefaciens* recognizes its host environment using ChvE to bind diverse plant sugars as virulence signals. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 678–83.
- Hwang, I., Li, P.L., Zhang, L., Piper, K.R., Cook, D.M., Tate, M.E. and Farrand, S.K.** (1994) Tral, a LuxI homologue, is responsible for production of conjugation factor, the Ti plasmid N-acylhomoserine lactone autoinducer. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 4639–4643.
- Khan, S.R. and Farrand, S.K.** (2009) The BlcC (AttM) lactonase of *Agrobacterium tumefaciens* does not quench the quorum-sensing system that regulates Ti plasmid conjugative transfer. *J. Bacteriol.* **191**, 1320–1329.
- Kim, H. and Farrand, S.K.** (1998) Opine catabolic loci from *Agrobacterium* plasmids confer chemotaxis to their cognate substrates. *Mol. Plant. Microbe. Interact.* **11**, 131–143.

- Kim, K.-S., Baek, C.-H., Lee, J.K., Yang, J.M. and Farrand, S.K.** (2001) Intracellular Accumulation of Mannopine, an Opine Produced by Crown Gall Tumors, Transiently Inhibits Growth of *Agrobacterium tumefaciens*. *Mol. Plant-Microbe Interact.* **14**, 793–803.
- Kim, S.I., Veena and Gelvin, S.B.** (2007) Genome-wide analysis of *Agrobacterium* T-DNA integration sites in the *Arabidopsis* genome generated under non-selective conditions. *Plant J.* **51**, 779–791.
- Lacroix, B., Tzfira, T., Vainstein, A. and Citovsky, V.** (2006) A case of promiscuity: *Agrobacterium*'s endless hunt for new partners. *Trends Genet.* **22**, 29–37.
- Lang, J., Gonzalez-Mula, A., Taconnat, L., Clement, G. and Faure, D.** (2016) The plant GABA signaling downregulates horizontal transfer of the *Agrobacterium tumefaciens* virulence plasmid. *New Phytol.* **210**, 974–983.
- Lang, J., Vigouroux, A., Planamente, S., Sahili, A. El, Blin, P., Aumont-Nicaise, M., Dessaux, Y., Moréra, S. and Faure, D.** (2014) *Agrobacterium* Uses a Unique Ligand-Binding Mode for Trapping Opines and Acquiring A Competitive Advantage in the Niche Construction on Plant Host. *PLoS Pathog.* **10**.
- Lang, J., Vigouroux, A., Sahili, A. El, Kwasiborski, A., Aumont-Nicaise, M., Dessaux, Y., Shykoff, J.A., Moréra, S. and Faure, D.** (2016) Fitness costs restrict niche expansion by generalist niche-constructing pathogens. *ISME J.*, 1–12.
- Lassalle, F., Campillo, T., Vial, L., et al.** (2011) Genomic species are ecological species as revealed by comparative genomics in *Agrobacterium tumefaciens*. *Genome Biol. Evol.* **3**, 762–781.
- Lee, C.-W., Efetova, M., Engelmann, J.C., Kramell, R., Wasternack, C., Ludwig-Müller, J., Hedrich, R. and Deeken, R.** (2009) *Agrobacterium tumefaciens* promotes tumor induction by modulating pathogen defense in *Arabidopsis thaliana*. *Plant Cell* **21**, 2948–62.
- Li, P.L. and Farrand, S.K.** (2000) The replicator of the nopaline-type Ti plasmid pTiC58 is a member of the repABC family and is influenced by the TraR-dependent quorum-sensing regulatory system. *J. Bacteriol.* **182**, 179–188.
- Li, P.L., Hwang, I., Miyagi, H., True, H. and Farrand, S.K.** (1999) Essential components of the Ti plasmid trb system, a type IV macromolecular transporter. *J. Bacteriol.* **181**, 5033–41.
- Luo, Z.Q. and Farrand, S.K.** (1999) Cloning and characterization of a tetracycline resistance determinant present in *Agrobacterium tumefaciens* C58. *J. Bacteriol.* **181**, 618–626.

- Matthysse, A.G., Jaeckel, P. and Jeter, C.** (2008) *attG* and *attC* mutations of *Agrobacterium tumefaciens* are dominant negative mutations that block attachment and virulence. *Can. J. Microbiol.* **54**, 241–247.
- Mougel, C., Thioulouse, J., Perrière, G. and Nesme, X.** (2002) A mathematical method for determining genome divergence and species delineation using AFLP. *Int. J. Syst. Evol. Microbiol.* **52**, 573–586.
- Nair, G.R., Liu, Z. and Binns, A.N.** (2003) Reexamining the role of the accessory plasmid pAtC58 in the virulence of *Agrobacterium tumefaciens* strain C58. *Plant Physiol.* **133**, 989–99.
- Nautiyal, C.S., Dion, P. and Chilton, W.S.** (1991) Mannopine and mannopinic acid as substrates for *Arthrobacter* sp. strain MBA209 and *Pseudomonas putida* NA513. *J. Bacteriol.* **173**, 2833–2841.
- Nealson, K.H., Platt, T. and Hastings, J.W.** (1970) Cellular Control of the Synthesis and Activity of the Bacterial Luminescent System1. *J. Bacteriol.* **104**, 313–322.
- Pappas, K.M.** (2008) Cell-cell signaling and the *Agrobacterium tumefaciens* Ti plasmid copy number fluctuations. *Plasmid* **60**, 89–107.
- Petit, A., Tempe, J., Kerr, A., Holsters, M., Montagu, M. Van and Schell, J.** (1978) Substrate induction of conjugative activity of *Agrobacterium tumefaciens* Ti plasmids. *Nature* **271**, 570–572.
- Pionnat, S., Keller, H., Hélicher, D., Bettachini, A., Dessaux, Y., Nesme, X. and Poncet, C.** (1999) Ti plasmids from *Agrobacterium* characterize rootstock clones that initiated a spread of crown gall disease in mediterranean countries. *Appl. Environ. Microbiol.* **65**, 4197–4206.
- Piper, K.R., Beck von Bodman, S. and Farrand, S.K.** (1993) Conjugation factor of *Agrobacterium tumefaciens* regulates Ti plasmid transfer by autoinduction. *Nature* **362**, 448–50.
- Piper, K.R., Beck Von Bodman, S., Hwang, I. and Farrand, S.K.** (1999) Hierarchical gene regulatory systems arising from fortuitous gene associations: Controlling quorum sensing by the opine regulon in *Agrobacterium*. *Mol. Microbiol.* **32**, 1077–1089.
- Pitzschke, A. and Hirt, H.** (2010) New insights into an old story: *Agrobacterium*-induced tumour formation in plants by plant transformation. *EMBO J.* **29**, 1021–32.
- Planamente, S., Mondy, S., Hommais, F., Vigouroux, A., Moréra, S. and Faure, D.** (2012) Structural basis for selective GABA binding in bacterial pathogens. *Mol. Microbiol.* **86**, 1085–1099.

- Planamente, S., Vigouroux, A., Mondy, S., Nicaise, M., Faure, D. and Moréra, S.** (2010) A conserved mechanism of GABA binding and antagonism is revealed by structure-function analysis of the periplasmic binding protein Atu2422 in *Agrobacterium tumefaciens*. *J. Biol. Chem.* **285**, 30294–30303.
- Puławska, J. and Sobiczewski, P.** (2005) Development of a semi-nested PCR based method for sensitive detection of tumorigenic *Agrobacterium* in soil. *J. Appl. Microbiol.* **98**, 710–721.
- Qin, Y., Luo, Z.Q., Smyth, a J., Gao, P., Beck von Bodman, S. and Farrand, S.K.** (2000) Quorum-sensing signal binding results in dimerization of TraR and its release from membranes into the cytoplasm. *EMBO J.* **19**, 5212–21.
- Rondon, M.R., Ballering, K.S. and Thomas, M.G.** (2004) Identification and analysis of a siderophore biosynthetic gene cluster from *Agrobacterium tumefaciens* C58. *Microbiology* **150**, 3857–3866.
- Sahili, A. El, Li, S.Z., Lang, J., et al.** (2015) A Pyranose-2-Phosphate Motif Is Responsible for Both Antibiotic Import and Quorum-Sensing Regulation in *Agrobacterium tumefaciens*. *PLoS Pathog.* **11**, 1–24.
- Sheng, J. and Citovsky, V.** (1996) *Agrobacterium*-plant cell DNA transport: have virulence proteins, will travel. *Plant Cell* **8**, 1699–710.
- Tzfira, T. and Citovsky, V.** (2008) *Agrobacterium* From Biology to Biotechnology,.
- Veselov, D., Langhans, M., Hartung, W., et al.** (2003) Development of *Agrobacterium tumefaciens* C58-induced plant tumors and impact on host shoots are controlled by a cascade of jasmonic acid, auxin, cytokinin, ethylene and abscisic acid. *Planta* **216**, 512–22.
- Wächter, R., Langhans, M., Aloni, R., et al.** (2003) Vascularization, high-volume solution flow, and localized roles for enzymes of sucrose metabolism during tumorigenesis by *Agrobacterium tumefaciens*. *Plant Physiol.* **133**, 1024–37.
- Wood, D.W., Setubal, J.C., Kaul, R., et al.** (2001) The genome of the natural genetic engineer *Agrobacterium tumefaciens* C58. *Science* **294**, 2317–23.
- Yu, X., Zhang, C., Yang, L., Zhao, L., Lin, C., Liu, Z. and Mao, Z.** (2015) CrdR function in a curdlan-producing *Agrobacterium* sp. ATCC31749 strain. *BMC Microbiol.* **15**, 25.
- Zhu, Y., Nam, J., Humara, J.M., et al.** (2003) Identification of *Arabidopsis* rat mutants. *Plant Physiol.* **132**, 494–505.
- Zipfel, C., Kunze, G., Chinchilla, D., Caniard, A., Jones, J.D.G., Boller, T. and Felix, G.** (2006) Perception of the Bacterial PAMP EF-Tu by the Receptor EFR Restricts *Agrobacterium*-Mediated Transformation. *Cell* **125**, 749–760.

Chapter II

Chapter II: The *Agrobacterium tumefaciens* C58 lifestyle in plant tumor

L'interaction entre la plante hôte et l'agent pathogène est un processus complexe dans lequel des aspects très différents des deux partenaires entrent en jeu. La réponse de la plante à l'infection par *Agrobacterium* a été analysée dans plusieurs études (Ditt *et al.*, 2006; Gohlke and Deeken, 2014). Contre l'attaque de ces pathogènes, la plante utilisera divers mécanismes pour se défendre. En réaction, l'agent pathogène doit être en mesure de surmonter ou de supprimer les défenses de la plante pour pouvoir la coloniser et se développer de manière satisfaisante. *Agrobacterium tumefaciens* a été étudié pendant des années pour sa capacité à transférer une partie de son ADN à la plante hôte. Outre cette caractéristique, très importante en biotechnologie, d'autres aspects de la relation entre la bactérie et la plante ont également été étudiés. Parmi eux se trouve le mécanisme de quorum-sensing ou la synthèse et la dégradation de certains métabolites appelés opines, dont le rôle est également important dans le processus d'infection.

L'objectif de cette étude était d'explorer de nouveaux aspects du mode de vie de *A. tumefaciens*, et en particulier l'interaction entre la bactérie et son hôte végétal. Pour cela, une analyse transcriptomique d'une tumeur induite par de *A. tumefaciens* C58 a été réalisée dans la plante hôte *Arabidopsis thaliana*.

Julian Lang a effectué les expériences liées à l'obtention des données du transcriptome que j'ai analysé dans ce chapitre. Certains des mutants utilisés dans l'étude étaient disponibles en laboratoire. Par ailleurs, j'ai réalisé la construction de 10 mutants chez *A. tumefaciens* C58 et mesuré la compétitive de ces mutants dans les tumeurs induites chez *A. thaliana*. Le texte a bénéficié de la relecture critique de Denis Faure.

1. INTRODUCTION

The interaction between the host plant and the pathogen is a complex process in which very different aspects of both agents come into play. Against the attack of these pathogens, the plant will use various mechanisms to defend itself. In contrast, the pathogenic agent must be able to overcome or suppress the plant's defenses to be able to colonize and develop in it satisfactorily. *Agrobacterium tumefaciens* has been studied for years for its ability to transfer part of its DNA to the host plant. Besides this characteristic, very important in biotechnology, other aspects of the relationship between the bacterium and plant have also been widely studied. Among them are the mechanism of quorum-sensing or the synthesis and degradation of some metabolites called opines, whose role is also important in the process of infection. On the other hand, the plant's response to *Agrobacterium* infection has been analyzed in several studies (Ditt *et al.*, 2006; Gohlke and Deeken, 2014). However, few studies have analyzed the later stages of the infection cycle and above all, there is no study that reflects the adaptation of the bacteria to the tumor lifestyle.

A. tumefaciens is the causal agent of the plant disease called crown gall, and it's able to infect more than 90 families of dicotyledonous plants. It is an α -*Proteobacterium* and belongs to the *Rhizobiaceae* family. *A. tumefaciens* is a complex of different species grouped in 10 genomovars (G1 to G8, and G13). *A. tumefaciens* C58 (also named as *A. fabrum* C58) belongs to the G8 group. Its genome consists in 4 replicons: 1 chromosome circular, 1 chromosome linear and 2 dispensable plasmids: pAt (for *A. tumefaciens*) and pTi (for Tumor inducing), which is required for virulence. A recent article published in *Genome Biol. Evol.* (Lassalle *et al.*, 2011) has shown that there are 196 genes specific to the G8 group. These genes are called species-specific genes. Most of them are clustered into 7 genomic islands: one on the circular chromosome and the other six in the linear chromosome. These clusters contain genes that codes for ecologically relevant roles, and we can classify them depending on these global functional units.

The role of both plasmids seems to be fundamental for *A. tumefaciens* C58. As mentioned above, most of the virulence genes are located in the Ti plasmid and have been the subject of numerous studies. Although harboring this plasmid supposes an important energetic cost, the advantage that confers in the infection process, makes that compensate both aspects. On the contrary, the plasmid At is not so well characterized. It seems that its importance lies in the ability of *Agrobacterium* to adapt to the rhizosphere. The cost of maintaining this plasmid is even higher than that of the Ti plasmid, but on the contrary it is easier to find strains of *Agrobacterium* that possess only

this plasmid or strains in which both coexist (Platt *et al.*, 2014).

The purpose of this study is to give a broader understanding of the adaptation process to the tumor lifestyle and to elucidate some possible functions of the At plasmid in this process.

2. EXPERIMENTAL PROCEDURES

2.1. Bacterial strains and culture conditions

A. tumefaciens C58 and its derivatives were cultivated at 28°C in Luria-Bertani modified medium (LBm; NaCl 5 g/L), TY medium (Bacto tryptone, 5 g/L ; yeast extract, 3 g/L) or *Agrobacterium* broth (AB) minimal medium (K₂HPO₄, 3 g/L; NaH₂PO₄, 1 g/L; MgSO₄·7H₂O, 0.3 g/L; KCl, 0.15 g/L; CaCl₂, 0.01 g/L; FeSO₄·7H₂O, 2.5 mg/L; pH7) (Chilton *et al.*, 1974) supplemented with ammonium chloride (1 g/L) and mannitol (2 g/L) as nitrogen and carbon sources. *Escherichia coli* DH5α was cultivated at 37°C in LBm and was the routine host for DNA cloning. Antibiotics were used at the following concentrations: gentamycin, 25 µg/mL; ampicillin, 50 µg/mL; rifampicin, 100 µg/mL.

2.2. RNA extraction

RNAs were extracted from two growth conditions of *A. tumefaciens* C58: exponential growth and plant tumor. Exponential culture was performed in AB medium with NH₄Cl as a nitrogen source and mannitol as a carbon source. Bacterial cells were collected when growth culture reached an optical cell density (OD₆₀₀) of 0.4. At 18 days post-infection, *A. thaliana* Columbia ecotype (Col-0) tumors were crushed in 20% Qiagen RNA protect Bacteria Reagent and 80% of phosphate buffer 50 mM. Plant tissues and bacteria were separated by a gel gradient using Gentodenz (Centaur, Paris, France) according the procedure described by Chapelle *et al.* (2015). All RNA extractions were carried out using the Qiagen RNeasy kit per the supplier's instructions.

2.3. Transcriptomics

Three hundred ng of total RNAs were used for the synthesis of the cDNA conforming to the protocol of the WTA1 TransPlex Whole Transcriptome Amplification Kit (Sigma). Resulting cDNA samples were purified using GenElute™ Clean-Up kit (Sigma) and labelled with Cy3 monoreactive dye. Purified cDNAs were then hybridized to NimbleGen 4-plex microarrays (Roche) containing probe sets for 5355 genes that represent the entire *A. tumefaciens* C58 genome. After hybridation, microarrays were washed three times and then scanned (Roche-NimbleGen MS200 scanner) to measure the fluorescence of cyanine 3 at the PartnerChip platform (GénopoleEvry, France). Images were processed using Nimblegen's NimbleScan software and further processed and normalized with GeneSpring GX11 analysis tool. The qualitative controls and the statistical analysis of the data were carried out using the same software. Normalized data of different samples were pairwise compared and P-values corresponding to statistical t-test were attributed for each gene. When the exponential growth and tumor life style were compared, only genes with an absolute fold change equal to or higher than 5 (i.e.

$\log_2 = 2.3$) and with a p-value lower than 0.05 were considered for the global analysis of transcriptomes. However, in the description of some pathways, absolute fold change value between 2 and 5 and with a p-value lower than 0.05 were indicated.

2.4. Calculation of factorial changes

Factorial change is the natural logarithm of the ratio of the percentage of differentially expressed genes in a functional category to the percentage of all genes of that category. To calculate it we used the mathematical formula described by (Deeken *et al.*, 2006): factorial change = $\ln((DE_a/DE_t)/(T_a/T_t))$ where DE_a is the number of differentially expressed genes annotated in the category a, DE_t the total number of differentially expressed genes in our experiment, T_a the number of genes annotated in the category a and T_t the total number of annotated genes in *A. tumefaciens* C58.

2.5. RNA quantification by RT-qPCR

cDNA was prepared from 1 μ g of bacterial or RNA using RevertAid H Minus First Strand cDNA Synthesis Kit (Fermentas, Saint-Remy-les-Chevres, France) following the manufacturer's instructions. RT-qPCRs were performed with Lightcycler® 96 (Roche) apparatus. The data were processed using the $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen, 2001). The internal control used was the gene *lepA* (*atu0241*) shown to be unaffected by our experimental conditions.

2.6. Construction of *A. tumefaciens* C58 mutants

The *A. tumefaciens* C58 derivative C107 with a gentamycin cassette cloned in a non-coding region of the Ti plasmid was available already (Haudecoeur, Tannieres, *et al.*, 2009). The gentamicin-resistance cassette from p34S-Gm (Dennis and Zylstra, 1998) was introduced into the *atu4206*, *atu4215*, *atu5061*, *atu5245*, *atu5344*, *atu5414*, *atu5502*, *atu5503* genes which were cloned into the *pGEM-T* Easy vector (Promega). The resulting plasmids were electroporated in *A. tumefaciens* C58 cells. Marker exchange was selected using Gm resistance, and double crossing-overs were verified by PCR. The plasmid p6000 (Maurhofer *et al.*, 1998) was used to drive expression of genes *atu1789*, *atu4215* and *atu4732* in *A. tumefaciens* C58.

2.7. Virulence assays in *A. thaliana* plants

A. thaliana plants (Col-0 ecotype) were grown in a climatic room at 22°C with a short day (8/16 h day/night) and 65% hygrometry. Three weeks after sowing, the emerging *A. thaliana* inflorescences were wounded and infected with ca. 10^6 *A. tumefaciens* cells as previously described (Lang *et al.*, 2016). Twenty-one days post-infection, *A. thaliana*

tumors, were collected, measured and crushed into NaCl 0.8%. Suspension-dilutions were plated on TY medium with appropriate antibiotics for *A. tumefaciens* CFU numbering.

2.8. Competition assays in *A. thaliana* tumor

A mixture of the *A. tumefaciens* C58 derivative C107 with each of the tested *A. tumefaciens* C58 KO-mutants (*atu4206*, *atu4215*, *atu5061*, *atu5245*, *atu5344*, *atu5414*, *atu5502*, *atu5503*) was used for inoculating the 21 days old *A. thaliana* Col-0 plants. Similar experiments were conducted by infecting plants with a mixture of the *A. tumefaciens* C58 harboring the empty vector p6000 with each the *A. tumefaciens* C58 derivatives harboring the p6000 vector expressing the *atu1789*, *atu4215* or *atu4732* gene. Twenty-eight days post-infection, the tumors were collected, measured and crushed into NaCl 0.8% and suspensions-dilutions were plated on the TY medium with appropriate antibiotics. Proportion of the inoculated genotypes were determined in the inoculum and bacteria collected from tumors by testing around 50 colonies by PCR using gene-specific primers (Table S1). Experiments were performed in duplicate.

2.9. Growth competition assays

Fresh TY liquid medium was inoculated at an optical cell density equal to 0.05 (OD₆₀₀) by a mixture (ratio 1:1) of the *A. tumefaciens* C58 derivative C107 with each of the tested *A. tumefaciens* C58 KO-mutants. After growing overnight at 28°C, suspension dilutions were spotted onto selective agar media to enumerate the bacterial populations. Proportion of both genotypes were determined before and after the TY culture using PCR primers on bacterial colonies. For statistical analyses, CI values (Macho *et al.*, 2010) were first transform to a square root so the values follow a normal distribution (Shapiro-Wilk normality test, P value > 0.05). Each CI (square root transformed) was analyzed using Student's t test and the null hypothesis: mean CI was not significantly different from 1 (P value = 0.05 was used).

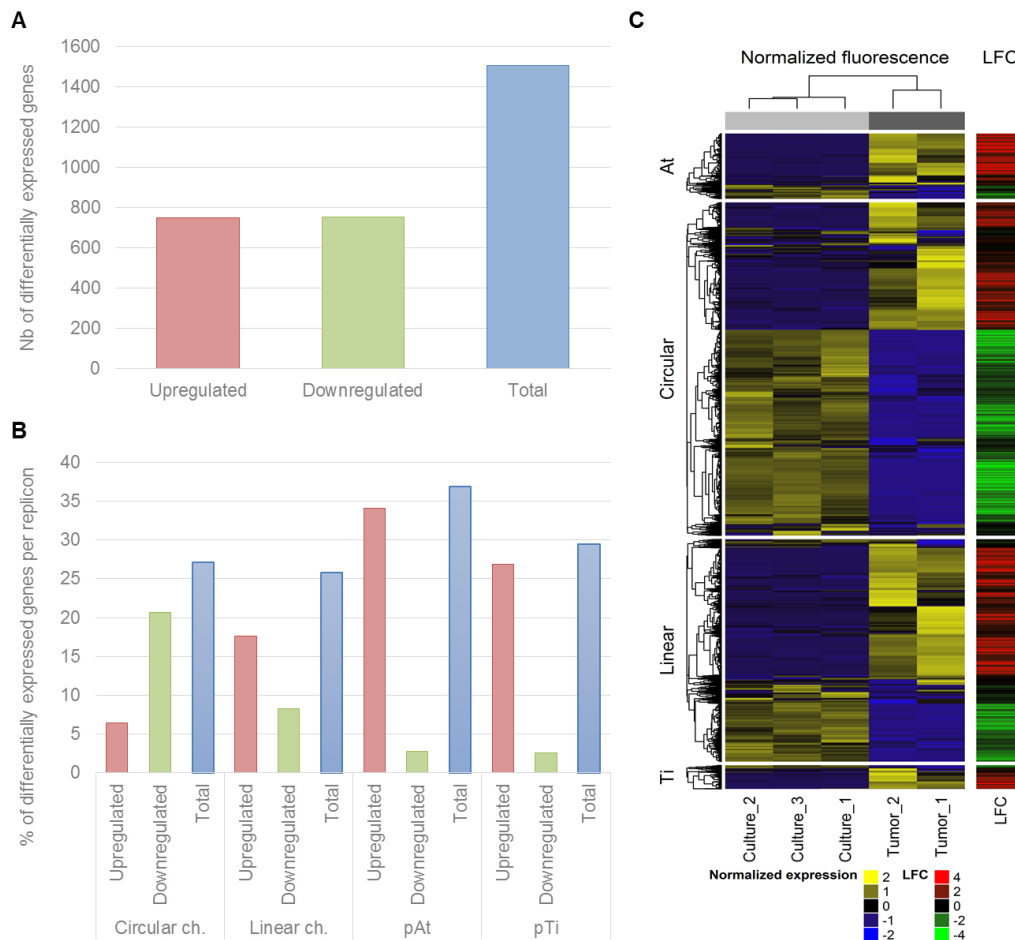


Figure 1. Comparative transcriptomics of *A. tumefaciens* C58 tumor cells versus C58 planktonic cells.
 A. Total number of differentially expressed genes.
 B. Percentage of differentially expressed genes by replicon.
 C. Heatmap of transcriptomic profile data. Heatmap was produced using the ComplexHeatMap package (v 1.12, Gu *et al.*, 2016).

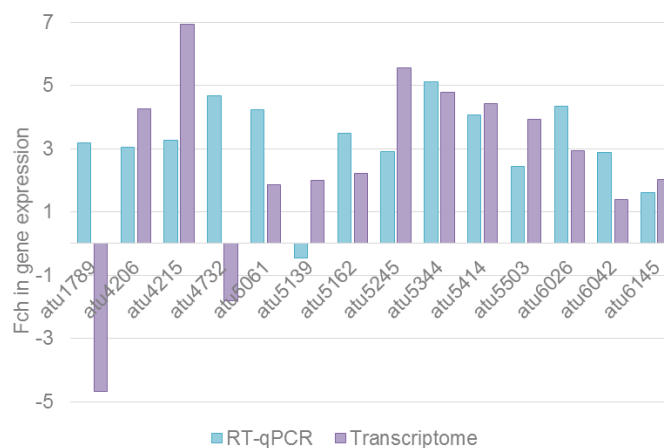


Figure 2. Comparison of the expression fold change (Fch) in *A. tumefaciens* C58 cells in plant tumor versus C58 planktonic cell for 14 genes using quantitative RT-PCR (RT-qPCR) or microarray technique.

3. RESULTS AND DISCUSSION

3.1. A massive reshaping of *A. tumefaciens* transcriptome in the plant tumor

A total of 1 504 genes were differentially expressed (absolute fold change ≥ 5) when the *A. tumefaciens* C58 cells in exponential growth condition in a synthetic liquid medium were compared with those cells colonizing the *A. thaliana* tumor (Figure 1a). The number of upregulated (749) and downregulated genes (755) in plant tumor condition were similar. All the differentially expressed genes represented more than a quarter (28%) of *A. tumefaciens* C58 genes, indicating a massive reshaping of transcriptome. All four replicons were affected (Figure 1b). In circular (762 of 2815 total genes) and linear (484 of 1875 total genes) chromosomes, 27% and 26% of genes are differentially expressed, respectively. In both plasmids, the percentage is slightly higher, being 29% for the Ti plasmid (58 of 197 total genes) and 37% for the At plasmid (200 of 542 total genes). Remarkably, most of the differentially expressed genes in both plasmids are upregulated, revealing a relative higher abundance of plasmid gene transcripts during plant tumor colonization. Gene expression of some genes was confirmed by qRT-PCR. Transcriptome and RT-qPCR data showed a similar gene expression profile (Figure 2).

Using the COG database, *A. tumefaciens* C58 genes were assigned to 21 functional categories. Each gene was assigned to a single functional category (Figure 3). To identify the functional classes which were the most affected in plant tumor condition, factorial changes were calculated for each gene category. This method compares the distribution of differentially expressed genes in functional categories with the distribution of all genes classified in *A. tumefaciens* C58. Thus, positive factorial changes indicate that a large part of the genes of that category were differentially expressed, and vice versa. Among upregulated genes, of the 22 categories shown in Figure 3, 6 categories contained a higher number of differentially expressed genes: amino acid transport and metabolism; defense mechanisms; inorganic ion transport; metabolism and intracellular trafficking, secretion and vesicular transport; secondary metabolites biosynthesis, transport and catabolism. Aside unclassified genes, the category associated to transport and metabolism of amino acids and of carbohydrates exhibited the highest number of upregulated genes. Most of them belong to linear chromosome and plasmids. On the other hand, downregulated genes were enriched in cellular motility; translation, ribosomal structure and biogenesis; energy production and conversion; cell cycle control and cell division; and posttranscriptional modification. These data insinuate that bacteria are in a state of “slow growth” within the plant tumor, with a reduced energy activity but

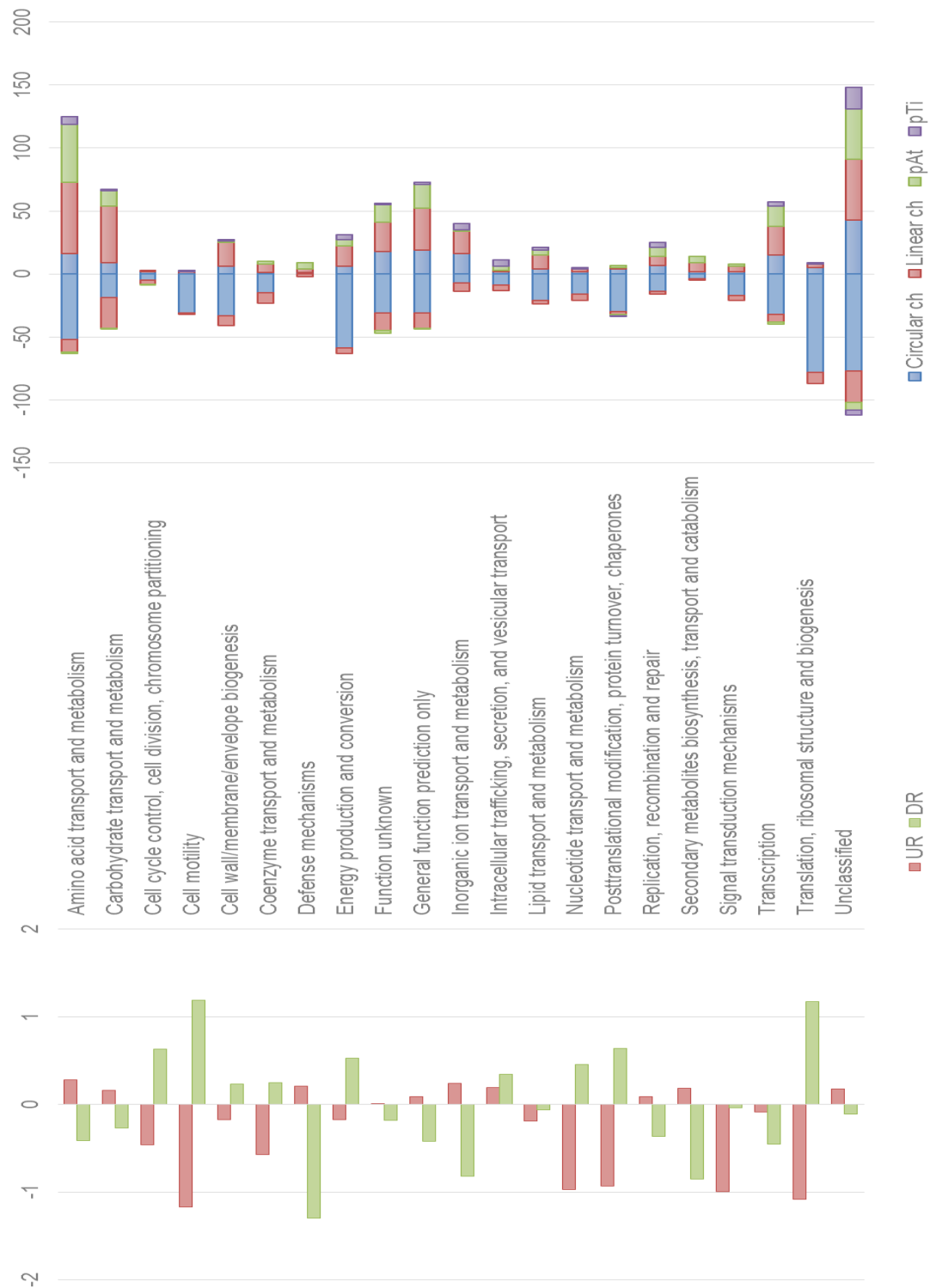


Figure 3. Transcriptome analysis of *A. tumefaciens* C58 tumor cells versus C58 planktonic cells.

- Factorial changes for the 21 functional categories of upregulated and downregulated genes in tumor condition. Functional categories with a positive factorial change have a higher fraction of differentially expressed genes than would be expected from the total number of genes assigned to that category. A negative value indicates a lower number of differentially regulated genes in the respective category than expected.
- Number of differentially expressed genes per replicon for each functional category. In positive axis are represented the upregulated genes and in the negative axis the downregulated genes for each category.

prepared to overcome the possible attack of the plant.

3.2. Reduction of cellular activity

Examination of the downregulated genes suggested a strong decrease of central metabolic and cellular processes, hence a reduced growth of *A. tumefaciens* cells in plant tumor.

Expression of the main aerobic respiratory chains for generating a proton motive force that drives and ATP synthesis was decreased. This is highlighted by downregulation of the operons *nuoABCDEFGHIJKLMN* (*atu1268-atu1283*) coding for NADH dehydrogenase, *sdhCDAB* (*atu2643-atu2645*) for succinate dehydrogenase, *atpDGAH* (*atu2622-atu2625*) for ATP synthesis, and *cydAB* (*atu4091-atu4092*), *coxAB-ctaBG-coxC* (*atu0767-atu0772*) and *fbcCBF* (*atu2237-atu2239*) for cytochromes and cytochrome c, bc and o oxidases. Most of the TCA genes (*acnB* = *atu4734*, *sucABCD* = *atu2635-atu2638*, *sdhCDAB* = *atu2643-atu2645*, *mdh* = *atu2639*) as well as those required to produce acetyl-CoA and oxaloacetate from pyruvate and phosphoenolpyruvate (*pykA* = *atu3762*, *pdhAB* = *atu1429-atu1430*, *lpdA* = *atu1434*, *pycA* = *atu2726*) were downregulated. Thirty-two genes of the lipid transport and metabolism were also downregulated, including those involved in malonyl-CoA synthesis (*accD* = *atu0020*, *accC* = *atu1330*, *accB* = *atu1331*) and fatty acid chain elongation (*fabF* = *atu1097*, *fabG* = *atu1095*, *fabZ* = *atu1383*, *fabI* = *atu0757*, *fabI* = *atu0149*, *acpP* = *atu1096*, *fabH* = *atu1179*, *plsX* = *atu1178*). Only a few of lipid synthesis were upregulated *accC* (*atu4273*), *fabF* (*atu4216*) and *fabG2* (*atu5250*).

The observed downregulation of many ribosomal and other translational genes is a solid indicator of a growth slowdown. Expression of genes coding for 53 ribosomal proteins was downregulated, 22 of them were present among the 50 highest downregulated genes. In addition, expression of ribosome binding factor A (*rbfA* = *atu0086*), ribonuclease E (*rne* = *atu1339*), ATP-dependent helicase (*rhIE* = *atu1833*), translation initiation factors (*infA* = *atu0534*, *infB* = *atu0087*, *infC* = *atu0254*), translation elongation factors (*tsf* = *atu1375*, *tufB* = *atu1948*, *fusA* = *atu1949*, *efp* = *atu2553*), the ribosome recycling factor (*rrf*=*atu1377*) was repressed. TufB (EF-Tu) is the best known PAMP that activates plant defense in response to *A. tumefaciens* infection (Zipfel *et al.*, 2006). Genes involved in protein folding (*groEL* = *atu0682*, *groES* = *atu0683*, *tig* = *atu1664*) and recycling (*clp2* = *atu1258*, *clpX* = *atu1259*, *clpP3* = *atu2270*, *clpA* = *atu1364*, *ftsH* = *atu3710*) were also repressed. A decreased expression of the genes involved in proteolysis could indicate that little recycling is required in slow-growing *A. tumefaciens*

cells.

Transcriptional key genes were repressed in *Agrobacteria* collected from plant tumor, especially those coding for RNA polymerase (*rpoA* = *atu1923*, *rpoB* = *atu1956*, *rpoC* = *atu1955*, *rpoD* = *atu2167*), the transcriptional terminaison factor Rho (*atu2833*), the transcription antiterminaison factor NusG (*atu1961*). In contrast, the ribonuclease H gene *rnhA* (= *atu3177*) was upregulated. This contributes to a reduction in global transcriptional activity and RNA content. A reduced cell metabolism is supported by a downregulation of 16 genes involved in nucleotide transport and metabolism, including *adk* (*atu1926*), *nrdI* (*atu0069*) and *pyrH* (*atu1376*), *nrdJ* (*atu1733*). In contrast, protelomerase Tera (*Atu2523*) which resolves a concatenated dimer of the linear chromosome as the last step of chromosome replication was upregulated and should contribute to increase efficiency of linear chromosome processing before cell division. Each replicon (circular and linear chromosomes and At and Ti plasmids) carries a copy of *dnaE* (*dnaE*=*atu1292*; *dnaE1*=*atu3228*. *dnaE2*=*atu5100*; *dnaE3*=*atu6093*) coding for alpha subunit of replicative DNA polymerase III. Only the At plasmid *dnaE2* was upregulated. This could be linked to high expression of replication and conjugation of At plasmid genes (see below).

3.3. Downregulation of cell division and developmental genes

Cell division in *A. tumefaciens* occurs via a polar budding process rather than binary fission that is described in many other rod-shaped bacteria (Grangeon *et al.*, 2015; Zupan *et al.*, 2013). In pant-tumor colonizing cells, growth slow-down was highlighted by the repression of genes involved in cell cycle control and cell division, as *parA* (*atu2136*) coding a chromosome partitioning protein, likewise *ftsZ2* (*atu2086*), *ftsA* (*atu2087*), *ftsQ* (*atu2088*) and *ftsZ1* (*atu4673*) coding for cell division proteins. Some master regulators of the coordination of division and developmental pathway were also affected. Gene expression of the master regulator CtrA (*Atu2434*) exhibited the third highest decrease with an absolute fold change of 140 in plant tumor colonizing cells. In the related α -proteobacterium *Caulobacter crescentus*, asymmetrical division results in high CtrA activity in the swarmer cell and decreased CtrA activity in the staked cell. In *A. tumefaciens*, the relative abundance and activity of CtrA could be a characteristic distinguishing sessile and plant adherent cells. Among the other coordination of division and developmental genes, the two response regulators DivK (*Atu1296*) and PleD (*Atu1297*) were downregulated, while those coding for the sensor histidine kinases were not affected. DivK and PleD regulate the phosphorelay components and levels of the internal second messenger bis-(3'-5')-cyclic dimeric guanosine monophosphate (c-di-

GMP). In an *A. tumefaciens* *divK* KO-mutant, biofilm formation was strongly reduced, and flagellar placement and swimming were slightly altered (Kim *et al.*, 2013).

3.4. Downregulation of motility and chemotaxis

A. tumefaciens exhibits a sparse tuft of four to six flagellar filaments (Chesnokova *et al.*, 1997). *A. tumefaciens* moves toward plant compounds and plant wounding site for inducing virulence gene expression (Cangelosi *et al.*, 1990; Shaw, 1991) but an *A. tumefaciens* mutant devoid of flagella was moderately reduced in virulence (Chesnokova *et al.*, 1997). Chemotaxis and flagellar-based motility contributes to attachment and biofilm formation in *A. tumefaciens*. In the plant-tumor colonizing cells, three major motility and chemotaxis clusters were downregulated : (i) cluster *atu0514-atu0526* (including the methyl-accepting chemotaxis gene *atu0514*, *cheA* = *atu0517*, *cheR* = *atu0518*, *fliF* = *atu0523*, *visN* = *atu0524*, *mclA* = *atu0526*), (ii) cluster *atu0542-atu0558* (including *fla* = *atu0542*, *flaB* = *atu0543*, *fliL* = *atu0547*, *flgG* = *atu0552*, *fliE* = *atu0553*, *flgC* = *atu0554*), and (iii) cluster *atu0560-atu0582* (including *motA*=*atu0560*, *fliM* = *atu0561*, *fliN* = *atu0562*, *fliG* = *atu0563*, *flaD* = *atu0567*, *motD* = *atu0571*, *flgE* = *atu0574*, *flgK* = *atu0575*, *flgL* = *atu0576*, *flaF* = *atu0577*, *flbT* = *atu0578*, *flgD* = *atu0579*, *fliQ* = *atu0580*). Other downregulated motility and chemotaxis genes were the methyl-accepting chemotaxis gene *atu0373*, *mcpA* (*atu0387*), *mcpG* (*atu0738*), *mcpC* (*atu0872*), *mcpV* (*atu1027*), *mclA* (*atu1912*), *cheW* (*atu2075*), *mcpA* (*atu2173*), *mcpA2* (*atu2223*) and methylation accepting chemotaxis gene *atu3363*. Additional genes were downregulated with an absolute fold change value between 2 and 5: *cheB* (*atu0519*), *cheD* (*atu0521*), *flgI* (*atu0550*), *flgA* (*atu0551*), *flgF* (*atu0558*), *fliI* (*atu0557*), *flhB* (*atu0564*), *motB* (*atu0569*), *fliR* (*atu0582*). Apart from the gene *atu3363*, all of motility and chemotaxis downregulated genes belong to the circular chromosome.

In *A. tumefaciens*, motility and chemotaxis are regulated at a transcriptional level by the VisN and VisR regulators (Xu *et al.*, 2013), as well as at a post-transcriptional level by the RNA-chaperone Hfq (Atu1450) (Möller *et al.*, 2014). In a Hfq-defective mutant, production of several motility proteins is decreased and motility is altered (Möller *et al.*, 2014; Wilms *et al.*, 2012). The expression of the *hfq* gene strongly decreased in the plant tumor colonizing *A. tumefaciens* cells (absolute fold change of 108, the fourth highest downregulated gene). In addition to VisNR transcriptional control, Hfq depletion could contribute to a reduced synthesis of motility and chemotaxis proteins when the pathogen lives in a plant tumor.

Most of the downregulated genes coding for the above functions (central metabolism and processes, cell division and development, and motility and chemotaxis) are coded by the circular chromosome (Figure 3). Among the first hundred of the highest downregulated genes, 87 belong to these functions. This explains why circular chromosome exhibits a high percentage of downregulated genes when the *A. tumefaciens* cells living under planktonic and plant tumor conditions were compared (Figure 1b).

3.5. Shift of carbon metabolism

Mannitol was the unique carbon and energy source in cell culture condition. In plant tumor, nutrient resources were expected to be more diverse, but more limited as bacteria only acceded to the extracellular compartment of plant tumor (surface and apoplast). In host-associated *A. tumefaciens*, mannitol assimilation was downregulated: this pathway encompasses the genes *atu4447-atu4450* coding for mannitol ABC-transporter and gene *atu4451* for mannitol 2-dehydrogenase involved in mannitol conversion into keto-D-fructose. Some genes involved in fructose assimilation via glycolysis (*atu1620*, *atu3737*, *atu3739*, *atu3449*, *atu0156*) and pentose phosphate pathway (*atu0599*, *atu600*, *atu1526*, *atu1613*, *atu3736*, *atu4464*) were also downregulated. Nutrient diversification became visible through the increased expression of genes coding for several glycoside hydrolases: *melA* (*atu4665*) for alpha-galactosidase, *cga* (*atu4833*) for glucoamylase, *picA* (*atu3128*) for rhamnogalacturonyl hydrolase, and *cscA* (*atu0944*) for sucrose hydrolase. Using random mutagenesis with a Mu transposon carrying promoterless *lacZYA* genes, the *picA* locus was previously identified as inducible by certain acidic polysaccharides found in a plant extract (Rong *et al.*, 1990; Rong *et al.*, 1991). Other activated genes were involved in the assimilation of galactose (*atu5455-5456*), rhamnose (*atu3491-3492*), myoinositol (*atu5434-5436*), tagatose (*atu4748*) and uncharacterized hexa and pentaglycosides (*atu2719*, *atu5407*). The upregulated gene *accG* (*atu6196*) coded for arabinose phosphate phosphatase that achieves the ultimate step of the agrocinnopine degradation (Kim and Farrand, 1997).

While most of the TCA genes decreased in their expression as described above, the glyoxylate pathway was activated, as *aceA* (*atu0607*) coding for the conversion of isocitrate into glyoxylate and succinate was induced. The *aceA* gene was reported to respond to acidic conditions (Yuan *et al.*, 2008), a physical constraint that is encountered in plant apoplast. Activation of glyoxylate pathway suggested a reorientation of carbon fluxes that could be mobilized for gluconeogenesis and fatty acid synthesis. This hypothesis was supported by the increased expression of adjacent genes coding for the

aldolase Atu3875 and glycerol-3-phosphate dehydrogenase GlpD (Atu3876) directing the entry of glycerol-3-phosphate into lipid synthesis.

Differential expression of carbohydrate transport systems also supported uptake of diverse nutrients. While 18 genes of this category were downregulated, 48 other genes were upregulated. Downregulated genes included those coding for mannitol (*atu4447-4450*), fructose (*atu0063-0066*) and ribose (*atu3818-3820*) ABC transporters, suggesting that the planktonic cells could mobilize different transporters for assimilating a unique sugar. Among the highly-upregulated genes were *atu3096-3099* on linear chromosome and *atu5059-5062* on At plasmid coding for two glycerol-3-phosphate ABC transporters. The adjacent gene *atu5061* code for a putative glycerophosphoryl diester phosphodiesterase (with a positive fold change of 4). The other upregulated transport genes were *atu4744-4747* coding for a sugar ABC transporter adjacent to the upregulated gene *atu4748* (tagatose-3-epimerase), *atu0392-0394* and *atu4370-4372* for two uncharacterized sugar ABC transporters. Activation of two ABC transporters for glycerol-3-phosphate uptake highlighted this compound as a resource for carbohydrate and lipid synthesis, as well as a phosphate resource for energy and nucleic pathways. A third potential advantage of glycerol-3-phosphate uptake and degradation by the *A. tumefaciens* pathogen would be linked to its role in systemic activation of plant defense (Chanda *et al.*, 2011).

In addition to sugars, some upregulated genes were involved in the utilization of organic acids. These included the three gene clusters: *atu5334-5335* coding for converting hydroxypyruvate into 2-phospho-glycerate (directed to glycolysis/gluconeogenesis), *atu4770-4772* coding for two hypothetical proteins and a periplasmic oxalate decarboxylase Atu4771 releasing formate and CO₂ (Shen *et al.*, 2008), and *atu5326-5328* cluster coding for hypothetical protein, transcriptional regulator and D-isomer 2-hydroxyacid dehydrogenase that catalyzes the oxidation of lactate into pyruvate in the presence of an unknown acceptor. Other phosphorylated compounds such as phosphonates would be assimilated via the ABC transporter coded by the genes *atu0171-0174* that were upregulated as the gene *atu0175* coding for an acetyltransferase. The phosphate transport regulatory protein PhoU (*atu0424*) was also induced.

3.6. High diversification of the nitrogen sources

Ammonium was the unique nitrogen source (none amino acids were added) in cell culture condition. Several *glnA* genes (*atu1770*, *atu0602*, *atu2416*) coding for condensation of ammonium and glutamate into glutamine were downregulated in plant host-associated cells. Some amino acid synthesis genes were also downregulated: *dapA* genes (*atu1024*, *atu0371*) involved in the first step of the lysine synthesis from pyruvate, *hisE* (*atu0038*) and *hisI* (*atu1750*) in the histidine synthesis from ribose-5-phosphate, and *aroQ1* (*atu1332*) and *atu1610* (*dhs*) in the chorismate synthesis from erythrose-4-phosphate. Expression of *tyrA* (*atu2698*) coding for the conversion of chorismate to prephenate, a precursor of phenylalanine and tyrosine, was reduced. Precursors of these amino synthesis pathways (namely, pyruvate, ribose-5-phosphate and erythrose-4-phosphate) derive from the glycolysis and pentose phosphate pathway. The observed downregulation of the lysine, histidine and chorismate synthesis pathways could be interpreted as a consequence of the global decrease of sugar metabolism when planktonic and host-associated cells were compared. Another non-exclusive hypothesis is that the hosted bacteria could import all the proteogenic amino acids they need, or most of them. This hypothesis was supported by the high number of amino acid transport genes that were upregulated in host-associated *A. tumefaciens* cells.

Tens of amino acid transporter genes were differentially expressed (n=119), most of them were upregulated (n=101). Among the downregulated genes, only the 4 genes *atu1577-1580* coding for a putative amino acid ABC transporter were downregulated together. Most of the other repressed ABC genes code for PBPs only (*atu0609*, *atu2276*, *atu2281*, *atu2365*, *atu2422*, *atu4113*, *atu4421*, *atu4431*), suggesting a different regulation of the PBP and channel/ATPase encoding genes of ABC transporters. Putative ligands of these PBPs were proteogenic and non-proteogenic amino acids, amines, oligopeptides. Out of the down-regulated PBPs, the deepest characterized is *Atu2422* which is involved in the competitive binding of GABA and proline (Haudecoeur, Planamente, *et al.*, 2009; Planamente *et al.*, 2010).

The upregulated amino acid transporter genes were either ABC transporter genes (n=87) or permeases of major facilitator superfamily (n=14) that are uniporters, symporters, and antiporters. Permease expression suggested complex in and out exchanges between bacterial cells and surrounding plant environment. A majority of these upregulated genes (ABC transporters and permeases) belonged to the linear chromosome and *At* plasmid, each harboring around forty upregulated genes. This highlights the still poorly investigated contribution of these two replicons in the *A. tumefaciens* lifestyle. Specific

or putative functions could be assigned to these upregulated genes, such as uptake of proteogenic amino acids, oligopeptides and amino acids derivatives such as nopaline. These functions were exemplified by a selection of upregulated genes.

Two upregulated ABC transporters harbored by the plasmid At would be involved in the importation of proteogenic amino acids: *atu5218-5220* for proline, glycine, betaine and *atu5521-5524* for branched amino acids such as valine, leucine, isoleucine. The upregulation of these transporters would potentially compensate a decrease of the amino acids (proline, valine, glycine, and betaine) importation as expression the broad ligand range PBP *Atu2422* decreased. The genes coding for *atu4221-4224* coding for the sarcosine oxidase that converts N-methylglycine (sarcosine) into glycine were also upregulated. Several ABC transporters would be involved in the importation of oligopeptides and nitrogen-compounds according to annotation of the PBP ligand and adjacent genes. The ABC transporter genes *atu5413-5411* were upregulated as the adjacent genes *atu5414* and *atu5415* coding for peptidase and hypothetical protein. Noticeably, At plasmid harbored another upregulated peptidase-encoding gene *atu5193* co-induced with *atu5194* coding for a transcriptional regulator. An upregulated cluster *atu5341-atu5348* associates ABC transporter genes and two genes coding for putative esterase (*Atu5348*) and 3-hydroxyacyl-CoA dehydrogenase/thioesterase (*Atu5344*). The transcriptional regulator and ABC transporter genes *atu5419-atu5423* would be involved in the uptake of amines (spermidine/putrescine) while ABC transporter genes *atu5246-atu5450* that were co-expressed with the putative nicotinamidase *Atu5245* would be involved in the uptake and assimilation of nicotinamide/chorismate related compounds. The ABC transporter genes *atu5531-5535* and *atu5536* coding for a putative nitrilase and ABC transporter genes *atu3444-3442* adjacent to *atu3445* and *atu3441* coding for hippurate hydrolase and nitriloacetate monooxygenase were upregulated. A large cluster *atu3036-3050* coding for two ABC transporters and a putative glycine/D-amino acid oxidase (*Atu3046*) were activated, as well as *atu5234-5238* coding for an ABC transporter and succinylglutamate desuccinylase (*Atu5238*).

Some upregulated genes are well characterized as they are involved in the assimilation of deoxy-fructosyl-glutamine (Amadori compound) and nopaline of which synthesis is encoded by T-DNA (Baek *et al.*, 2003; Marty *et al.*, 2016; Schardl and Kado, 1983). Those are genes *atu5003-5007* required for transport and degradation of deoxy-fructosyl-glutamine (Amadori compound) into fructose and glutamine. The others are *nocMPQTR* (*atu6025-6029*) coding for nopaline ABC transporter and transcriptional regulator of the nopaline regulon. The *noxAB* genes (*atu6019-6021*) coding for cleavage

of nopaline into alpha-ketoglutarate and arginine were upregulated with a fold change between 3 and 4. The degradation of arginine into urea and ornithine is operated by two arginases Arc (Atu6018) and ArcA (Atu4007) of which the expression increased in host-colonizing *A. tumefaciens*. Nopaline assimilation confers a selective advantage to *A. tumefaciens* when it colonizes the plant tumor (Lang *et al.*, 2014). Another remarkable upregulated gene was *kamA* (*atu2555*) coding for the conversion of lysine into beta-lysine. This gene is closed to *poxA* (*atu2554*) and *efp* (*atu2553*) coding for a lysyl-tRNA-synthase homolog and elongation factor P, respectively. In *Salmonella enterica*, elongation factor P is modified with (R)- β -lysine by the lysyl-tRNA synthetase paralog PoxA (Roy *et al.*, 2011). In this enterobacterium, *poxA* and *yjeK* mutants exhibit extensive phenotypic pleiotropy, including attenuated virulence (Navarre *et al.*, 2010).

3.7. Remodeling of cell wall and surface components

A. tumefaciens produces at least six polysaccharide types, several of which play roles in attachment and biofilm formation: outer membrane polysaccharide (LPS), succinoglycan, unipolar polysaccharide (UPP), cellulose, succinoglycal, cyclic β -1,2-glucans and β -1,3-glucan (curdlan) (Heindl *et al.*, 2014; Wu *et al.*, 2016). In LPS pathway, genes involved in synthesis of the glucosamine precursor were downregulated (*lpxC* = *atu2085*, *lpxD* = *atu1382*, *lpxB* = *atu1386*), but expression of those for core saccharide precursor (*kdsB* = *atu0100*, *kdsA* = *atu1024*, *kpsF* = *atu3774*) and lipid A (*lpxK* = *atu0697*, *kdtA* = *atu0695*, *msbB* = *atu1594*, *acpXL* = *atu1600*) remained unchanged. In contrast, several genes associated to surface and membrane polysaccharides were upregulated, including *gmd* (*atu4789*), genes *rfe*, *bme2*, *bme3*, *bme4* (*atu4811-4814*), genes *amsJ* and *pssF* (*atu4808-4806*), response regulator gene *atu4803*, genes *bme23*, *rfbC*, *bme10* (*atu4800-4798*), *bme12* (*atu4794*) and a hypothetical protein (*atu4792*). Hence, they could modify the surface polysaccharides that contribute to host interactions (adhesion and recognition).

In the succinoglycan pathway (Wu *et al.*, 2016), genes involved in the biosynthesis of nucleotide sugar precursors (*exoC* = *atu4074*, *exoB* = *atu0530*, *exoN* = *atu4050*) and condensation of galactose and lipid (*exoYF*) were downregulated in the host-associated *A. tumefaciens*. This repression could be in relation to a general slowdown of metabolic activity. In contrast, the expression of the succinoglycan elongation (*exoALMOUW*), decoration (*exoZHV*), polymerization and secretion (*exoKPQT*) genes remained globally unchanged. A noticeable variation is an enhanced expression of the poorly characterized *exsDC* genes (*atu3344-3345*) which would be involved in some particular modifications of succinoglycan. Gene expression of unipolar polysaccharide (*uppABCDEFGF* = *atu1235-*

1240), cellulose (*celABC* = *atu3307-3309*, *celG* = *atu8186*) and cyclic β -1,2-glucan (*chvB* = *atu2730*) were weakly or unaffected. By contrast, the curdlan genes *crdASC* (*atu3055-3057*) were upregulated in host-associated *A. tumefaciens*, while those involved in synthesis of nucleotide sugar precursor were down regulated (*exoC* = *atu4074*, *galU* = *atu3778*).

Surface components such as type IV pili (*ctpA* = *atu0224*, *ctpB* = *atu0223*) and the major outer membrane protein Omp1 (BamA = *Atu1381*) were down regulated in host-associated bacteria, as well as the general secretion systems Sec (*secD* = *atu1562*, *secE* = *1962*, *secG* = *atu1619*, *secY* = *atu1927*, *yajG* = *atu1563*) and Tat (*tatB* = *atu1705*, *tatC* = *atu1704*). The *A. tumefaciens* twin-arginine-dependent translocation is important for virulence, flagellation, and chemotaxis but not type IV secretion (Ding and Christie, 2003). In contrast, *atu4838* coding for an outer membrane protein and *atu0472* coding for a putative fasciclin (adhesion protein) were upregulated, as well as *atu5364*, *atu5502* and *atu5503* genes coding for exported hypothetical proteins that could be involved in the interaction with the host plant.

3.8. Response to and exploitation of iron and gas

Two transcriptional repressors Fur (*Atu0354*) and RirA (*Atu0201*) control the homeostasis genes of manganese and iron, respectively (Hibbing and Fuqua, 2011; Kitphati *et al.*, 2007). In a less extend, Fur regulon is also involved in iron homeostasis (Kitphati *et al.*, 2007). Transcription of *rirA* is also negatively controlled by the master regulator Irr (*Atu0153*). In host-associated transcriptome, *fur* and *sitABCD* genes (*atu4468-4471*) coding for the manganese ABC transporter were downregulated indicating a sufficient manganese resource in plant tumor. In the same context, *rirA* was downregulated, while genes coding for iron ABC transporters *atu2473-2476* and Dps protein (*atu2477*) that protects DNA and cells from oxidative damage (Ceci *et al.*, 2003) were upregulated. The *Agrobacterium* siderophore-uptake *fhuA-fep* genes (*atu5311-5316*) were also activated, while the iron ABC transporter *fbp* genes (*atu0408-0406*) and siderophore synthesis operons *atu3670-3673* and *atu3675-atu3685* (Rondon *et al.*, 2004) were partially induced. Gene expression of the bacterioferritin Bfr used for iron storage (*atu2771*) was decreased, as well as that of *hemA* (*atu2613*) coding the first step of heme biosynthesis. Transcriptome data highlighted a reduced abundance of iron in plant tumor, a feature that is supported by a decrease of iron utilization and storage and an activation of uptake systems for trapping plant iron. Fur and RirA regulons are also involved in sensitivity to oxidative stress, as their overexpression increase intracellular iron, hence production of cell damaging hydroxyl radicals via the Fenton reaction

(Kitphati *et al.*, 2007). Regarding at the response to oxidative stress, the OhrR-type regulator genes *atu5064* and *atu5211* were upregulated, as well as the adjacent gene *cpo2* (*atu5065*) coding for non-heme peroxidase, but expression of other genes coding for catalases, peroxidases and OhrR-type regulators remained unchanged. *A. tumefaciens* cells colonizing tumor did not face an increased oxidative stress as compared to their planktonic lifestyle.

A. tumefaciens C58 is a partial denitrifier that lacks nitrous oxide reductase. In vitro, expression of the nitrite and nitric oxide reductases (Nir and Nor) requires both low levels of oxygen and a nitrogen oxide. When infiltrated in *A. thaliana* leaves, expression of the *nor-gfp* fusion was observed in response to plant nitric oxide, but not to nitrogen oxide produced from nitrite by *A. tumefaciens* cells themselves (Baek and Shapleigh, 2005). Remarkably, in *A. thaliana* tumor-associated bacterial cells, transcriptomic comparison revealed an increased expression of nitric oxide reductase (*norBC* = *atu4388-4389*), but none nitrate reductase (*NapAB* = *Atu4408-4409*), nor nitrite reductase (*NirK* = *Atu4382*), suggesting that only nitric oxide was sensed by *A. tumefaciens*. Because nitric oxide synthesis is activated during plant infection by pathogens and involved in the activation of some plant defense (Delledonne *et al.*, 1998; Durner and Klessig, 1999), the Nor expression in *A. tumefaciens* could be a response to plant nitric oxide or a tentative to manipulate nitric oxide-regulated plant defense.

3.9. Virulence and defense functions

Virulence genes involved in the T-DNA transfer were upregulated including *virD1* and *virD2* coding endonuclease activity promoting T-DNA excision. In *A. tumefaciens* strain overexpressing *virD1* and *virD2*, increased amounts of T-strand could be detected and they showed the ability to transform plants with higher efficiency (Wang *et al.*, 1990). Other induced T-DNA T4SS genes were *virB1-B2-B3* (*atu6167-6169*), *virB8* (*atu6174*), *virD5* (*atu6185*), as well as, with an absolute fold change value > 2, *virB4* (*atu6170*), *virB6* (*atu6172*), *virC1-C2* (*atu6179-6180*), *virD3* (*atu6183*), suggesting that the T-DNA transfer was still active in plant tumor. Expression of the gene *tzs* (*atu6164*) was also upregulated, this codes for cytokinin biosynthesis, hence contributing to tumor development in concert with cytokinin and auxins produced by T-DNA-encoded proteins (Hwang *et al.*, 2013). Other upregulated genes were the *att* genes on the At plasmid, more precisely the ABC transporter genes *atu5126-5132* (*attA1A2BCEF*) with the exception of *attD* (*atu5130*). A *attC* mutant is affected for attachment to tomato roots and virulence by a still unknown mechanism (Matthysse *et al.*, 2008).

Different upregulated genes would contribute to the resistance against toxic compounds produced by host plant or microbiote. These encompass the *tetA* (*atu4206*) gene coding for a major facilitator family transporter, *atu1347* for TetM/TetO subfamily tetracycline resistance protein, *atu2319* for quaternary ammonium compound-resistance efflux protein, the adjacent gene *atu2320* coding for TetR transcriptional regulator, *atu3465* for beta-lactamase hydrolase-like protein, and *atu5085* for metallo-beta-lactamase. Gene *virH2* (*atu6151*) coding for P450-monooxygenase converting ferulic acid into caffeic acid was also induced (with an absolute fold change of 4.7). Ferulic acid is far more toxic than caffeic acid to the *A. tumefaciens* (Kalogeraki *et al.*, 1999).

3.10. Genome plasticity: At and Ti plasmid transfer, insertion sequences

Most of *avh* (*atu5162-5172*) and *tra* (*atu5109-5111*) genes coding for the At plasmid transfer were upregulated: those were *avhB1B2B3B5B6B7B8B9B11* and *traD2*, while *avhB4B5B8B9* and *traC2A2* exhibited an absolute fold change between 2 and 5. Similarly, the Ti plasmid conjugation genes were activated, but at a lower fold change value: *trbI* (*atu6031*), *traA* and *traC* (*atu6126-6127*) at a fold change value higher than 5, and *trbHGFL* (*atu6032-6035*), *trbBC* (*atu6040-6041*), *traG* (*atu6124*) and *traB* (*atu6129*) at a fold change value between 2 and 5. The expression of these two conjugative systems correlated the observed transfer and co-transfer of the At and Ti plasmid in *A. thaliana* plant tumors. In *A. thaliana* tumor, transfer of At plasmid is tenfold more efficient than that of Ti plasmid (Lang *et al.*, 2013).

The At plasmid conjugation is not strictly dependent on quorum-sensing, but is slightly enhanced when quorum-sensing is strongly expressed (Lang *et al.*, 2013). In contrast, Ti plasmid conjugation requires quorum sensing which is activated in the presence of the opine agrocinopine, especially its arabinose-2-phosphate by product (El Sahili *et al.*, 2015). The *tral* gene (*atu6042*) coding for synthesis of the quorum sensing signal 3-oxo-octanoylhomoserine lactone was slightly activated (fold change was 2.3). In contrast, genes being expected to be induced by agrocinopine, such as *traR* (*atu6134*) and *accABCDEF* (*atu6139-6144*) genes were not activated in the transcriptomic data. This observation suggested that, at the sampling time, the agrocinopine resource was exhausted or/and that an only fraction of the *Agrobacterium* population responded to quorum sensing (hence diluting the relative abundance of transcripts). Quorum-quenching lactonases BlcC (Atu5139) and AiiB (Atu6071) were also expressed, also contributing to attenuate quorum-sensing and Ti plasmid conjugation in plant tumor, as described in *A. thaliana* (Lang *et al.*, 2016).

None insertion sequence genes were down regulated. Nine insertion sequence genes which belongs to the IS3, IS21, IS30, IS66, IS136, IS426, IS911 families, were upregulated (*atu1488*, *atu2107*, *atu4601-4602*, *atu4494*, *atu4861*, *atu5028*, *atu5353*, *atu5366-5367*). Gene disruption by IS136 transposition was observed in *A. tumefaciens* C58 (Rawat *et al.*, 2009; Schneider *et al.*, 2000).

3.11. Expression of genomic species-specific genes

The genome of *A. tumefaciens* C58 contains genes specific to the genomic group (G8 proposed as *A. fabrum* novel species) to which it belongs. Of the 196 specific genes of G8 group, 40 genes showed a differential expression between plant tumors comparing with planktonic condition. Practically, all of these genes are located on the linear chromosome. Among them, 34 were overexpressed in plant tumor. These genes are all located in the linear plasmid. Category SpG8-6 seemed to be important, since almost half of its genes were overexpressed. This category appears to be involved in resistance to drugs or toxic compounds. One of them is *tetA* gene (*atu4206*) that codes for a tetracycline resistance protein (Luo and Farrand, 1999). The partial *ftsZ2* overexpressed gene (*atu4215*) also located on this genomic island is putatively involved in cell division (Zupan *et al.*, 2013). This contradicts the fact that most of the genes involved in cell division were downregulated. Considering that in the genome of *A. tumefaciens* C58 there are 3 *ftsZ* genes on both chromosomes, and two of them were repressed, we could propose that this gene have another function in the adaptation of the bacteria to plant tumor condition. Other important species-specific gene highly upregulated is *crdS* (*atu3056*) located on SpG8-2a group which main function is the biosynthesis of the exopolysaccharide called curdlan. This extracellular polysaccharide is produced by *Agrobacterium* during low pH condition and nitrogen starvation (Wu *et al.*, 2016). Along genes from SpG8-6, *metE* on the SpG8-4 involved in ribose transport; monosaccharide catabolism and carbohydrate metabolism, *menG* of the SpG8-7a related to environmental signal sensing/transduction and finally *ephA* on the SpG8-3 involved in siderophore biosynthesis were those that present a higher value of fold change. The other 6 differentially expressed genes were downregulated. 4 of them belong to the SpG8-4 group and the others are from the SpG8-1b and SpG8-6a categories. We can imagine that the bacteria need these genes to defend itself against the attack of the plant and that evolutionarily has been acquiring them for a better adaptation to new ecological niches.

3.12. *The aggressiveness and fitness of the constructed mutants*

Involvement of several upregulated (*atu4206*, *atu4215*, *atu5061*, *atu5245*, *atu5344*, *atu5414*, *atu5502*, *atu5503*) and downregulated (*atu1789*, *atu4215* or *atu4732*) genes in aggressiveness and fitness was studied using constructed KO-mutants and constitutive expression on plasmids, respectively. Among the upregulated genes, we selected two species-specific genes *atu4206* (= *tetA*) coding for a major facilitator family transporter that could be involved in efflux of toxic plant compounds and *atu4215* (= *ftsZ2*) coding a putative cell division protein. All the other upregulated genes belong to At plasmid that exhibit a highest percentage of upregulated genes when the four replicons were compared. These genes were *atu5061* coding for a putative glycerophosphoryl diester phosphodiesterase, *atu5245*, a putative nicotinamidase, *atu5344*, a putative esterase, *atu5414*, a putative peptidase, each of these them being associated to an upregulated ABC transporter. The last selected upregulated genes were *atu5502* and *atu5503* coding for putative exported proteins that could interact with the host plant. Among the downregulated genes, we overexpressed a lipoprotein, *atu1789*, a fimbrial chaperone, *atu4732*. We also constructed an overexpression mutant of putative cell division protein gene, *atu4215* (= *ftsZ2*).

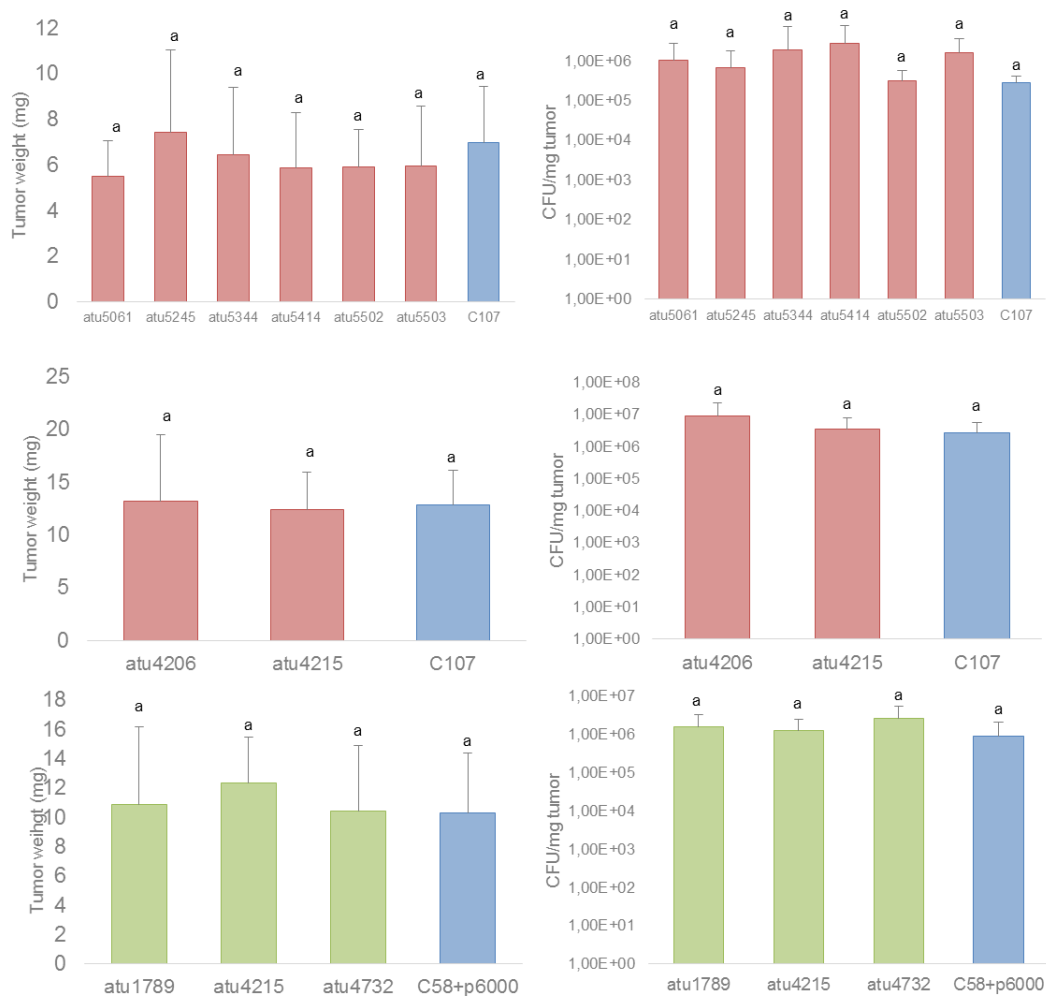


Figure 4. Virulence assay on *Arabidopsis thaliana* plants.

A. Mass of 21 days post-infection tumors induced with KO-mutants (red color) and overexpression mutants (green color) and their respective control strains.

B. Colonization efficiency (bacterial numeration) in 21 days post-infection tumors.

Mean values and standard deviations (SD) of two independent experiments are presented. Nonparametric Kruskal-Wallis and post-hoc Dunnet tests ($14 < n < 22$; $P < 0.05$) were used and different letters indicate statistical significance.

We compared *A. tumefaciens* aggressiveness by performing infection on *A. thaliana* plants (Figure 4). Tumor size and bacterial charge did not differ from the aggressiveness of the control strain. In addition, to determine if the genes were involved in tumor colonization fitness, a competition assay was performed where the control and mutant strains were co-inoculated. The tumors were collected after 21 days and the proportion of each strain in the tumor was analyzed by PCR. All KO-mutants of the At plasmid, presented a higher fitness than the control strain as shown in Figure 5. To verify if this effect depended on the response to the plant or on the contrary was intrinsic to the KO-mutants, we performed an *in vitro* competition test. In the case, the results showed that there was no *in vitro* competition of mutants and control strain (Figure 6). This suggests that the plant would play an important role in the regulation of these genes.

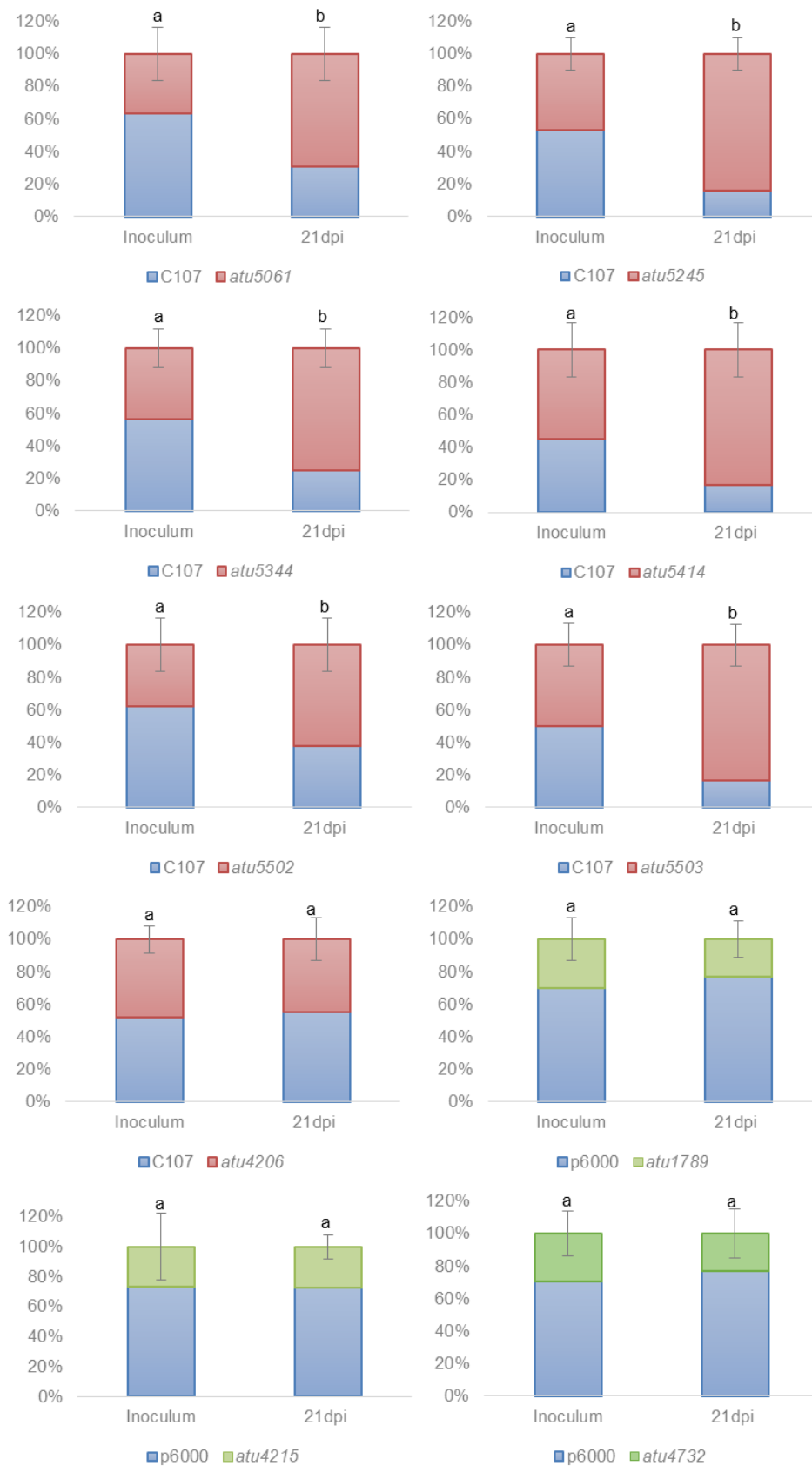


Figure 5. Competition assays in *A. thaliana* plant of KO-mutants (red color) and overexpression mutants (green color). All experiments were done on 28 days post-infection *A. thaliana* plant tumors. The experiment was done in two independent replicates ($5 < n < 11$). Statistical differences using the Fisher's Exact test ($p < 0.05$) are noted by different letters. Mean and SD of inoculum was calculated from 3 independent samples.

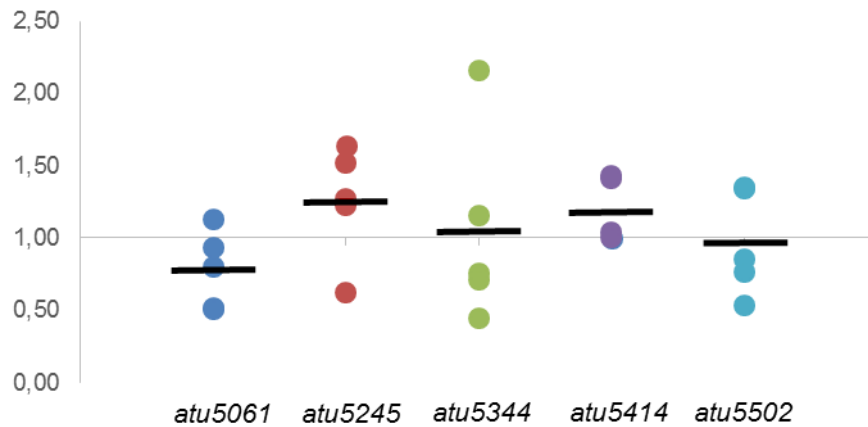


Figure 6. Fitness of KO-mutant versus *A. tumefaciens* (control) strain in AB mannitol medium. CI is defined as the mutant-to-wt output ratio divided by the mutant-to-wt input ratio.

4. CONCLUSIONS

The transcriptome analysis established the cell status of *A. tumefaciens* when it colonized the plant tumor. While some remarkable functions were downregulated, such as motility, chemotaxis, and cell division and developmental pathway, some other such as the major virulence function that is T-DNA transfer and horizontal transfer of At and Ti plasmids were upregulated. Surface components especially O-antigen were remodeled and curdlan synthesis was activated. In addition, *A. tumefaciens* responded to specific constraints, such as iron limitation, nitric oxide, acidic pH, and several plant compounds (sucrose, ferulic acid, rhamnolactoside, opines ...) that are assimilated as nutrients or detoxified. This work highlighted the unique lifestyle of *A. tumefaciens* in plant tumor.

5. BIBLIOGRAPHY

- Baek, C.H., Farrand, S.K., Lee, K.E., Park, D.K., Lee, J.K. and Kim, K.S.** (2003) Convergent evolution of Amadori opine catabolic systems in plasmids of *Agrobacterium tumefaciens*. *J. Bacteriol.* **185**, 513–524.
- Baek, S.H. and Shapleigh, J.P.** (2005) Expression of nitrite and nitric oxide reductases in free-living and plant-associated *Agrobacterium tumefaciens* C58 Cells. *Appl. Environ. Microbiol.* **71**, 4427–4436.
- Cangelosi, G.A., Ankenbauer, R.G. and Nester, E.W.** (1990) Sugars induce the *Agrobacterium* virulence genes through a periplasmic binding protein and a transmembrane signal protein. *Proc. Natl. Acad. Sci. U. S. A.* **87**, 6708–12.
- Ceci, P., Ilari, A., Falvo, E. and Chiancone, E.** (2003) The Dps protein of *Agrobacterium tumefaciens* does not bind to DNA but protects it toward oxidative cleavage. X-ray crystal structure, iron binding, and hydroxyl-radical scavenging properties. *J. Biol. Chem.* **278**, 20319–20326.
- Chanda, B., Xia, Y., Mandal, M.K., et al.** (2011) Glycerol-3-phosphate is a critical mobile inducer of systemic immunity in plants. *Nat. Genet.* **43**, 421–7.
- Chapelle, E., Alunni, B., Malfatti, P., Solier, L., Pédrón, J., Kraepiel, Y. and Gijsegem, F. Van** (2015) A straightforward and reliable method for bacterial in planta transcriptomics: Application to the *Dickeya dadantii/Arabidopsis thaliana* pathosystem. *Plant J.* **82**, 352–362.
- Chesnokova, O., Coutinho, J.B., Khan, I.H., Mikhail, M.S. and Kado, C.I.** (1997) Characterization of flagella genes of *Agrobacterium tumefaciens*, and the effect of a bald strain on virulence. *Mol. Microbiol.* **23**, 579–590.
- Chilton, M.D., Currier, T.C., Farrand, S.K., Bendich, A.J., Gordon, M.P. and Nester, E.W.** (1974) *Agrobacterium tumefaciens* DNA and PS8 bacteriophage DNA not detected in crown gall tumors. *Proc. Natl. Acad. Sci. U. S. A.* **71**, 3672–6.
- Deeken, R., Engelmann, J.C., Efetova, M., et al.** (2006) An Integrated View of Gene Expression and Solute Profiles of *Arabidopsis* Tumors: A Genome-Wide Approach. *Plant Cell Online* **18**, 3617–3634.
- Delledonne, M., Xia, Y., Dixon, R. a and Lamb, C.** (1998) Nitric oxide functions as a signal in plant disease resistance. *Nature* **394**, 585–588.
- Dennis, J.J. and Zylstra, G.J.** (1998) Plasmids: modular self-cloning minitransposon derivatives for rapid genetic analysis of gram-negative bacterial genomes. *Appl. Environ. Microbiol.* **64**, 2710–5.
- Ding, Z. and Christie, P.J.** (2003) *Agrobacterium tumefaciens* twin-arginine-dependent translocation is important for virulence, flagellation, and chemotaxis but not type IV secretion. *J. Bacteriol.* **185**, 760–771.

- Ditt, R.F., Kerr, K.F., Figueiredo, P. de, Delrow, J., Comai, L. and Nester, E.W.** (2006) The *Arabidopsis thaliana* transcriptome in response to *Agrobacterium tumefaciens*. *Mol. Plant. Microbe. Interact.* **19**, 665–81.
- Durner, J. and Klessig, D.F.** (1999) Nitric oxide as a signal in plants [published erratum appears in *Curr Opin Plant Biol* 1999 Dec;2(6):525]. *Curr. Opin. Plant Biol.* **2**, 369–374.
- Gohlke, J. and Deeken, R.** (2014) Plant responses to *Agrobacterium tumefaciens* and crown gall development. *Front. Plant Sci.* **5**, 155.
- Grangeon, R., Zupan, J.R., Anderson-Furgeson, J. and Zambryski, P.C.** (2015) PopZ identifies the new pole, and PodJ identifies the old pole during polar growth in *Agrobacterium tumefaciens*. *Proc. Natl. Acad. Sci. United States* **112**, 11666–11671.
- Haudecoeur, E., Planamente, S., Cirou, A., Tannières, M., Shelp, B.J., Moréra, S. and Faure, D.** (2009) Proline antagonizes GABA-induced quenching of quorum-sensing in *Agrobacterium tumefaciens*. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 14587–92.
- Haudecoeur, E., Tannieres, M., Cirou, A., Raffoux, A., Dessaux, Y. and Faure, D.** (2009) Different regulation and roles of lactonases AiiB and AttM in *Agrobacterium tumefaciens* C58. *Mol Plant Microbe Interact* **22**, 529–537.
- Heindl, J.E., Wang, Y., Heckel, B.C., Mohari, B., Feirer, N. and Fuqua, C.** (2014) Mechanisms and regulation of surface interactions and biofilm formation in *Agrobacterium*. *Front. Plant Sci.* **5**, 176.
- Hibbing, M.E. and Fuqua, C.** (2011) Antiparallel and interlinked control of cellular iron levels by the Irr and RirA regulators of *Agrobacterium tumefaciens*. *J. Bacteriol.* **193**, 3461–3472.
- Hwang, H.-H., Yang, F.-J., Cheng, T.-F., Chen, Y.-C., Lee, Y.-L., Tsai, Y.-L. and Lai, E.-M.** (2013) The Tzs Protein and Exogenous Cytokinin Affect Virulence Gene Expression and Bacterial Growth of *Agrobacterium tumefaciens*. *Phytopathology* **103**, 888–899.
- Kalogeraki, V.S., Zhu, J., Eberhard, A., Madsen, E.L. and Winans, S.C.** (1999) The phenolic vir gene inducer ferulic acid is O-demethylated by the VirH2 protein of an *Agrobacterium tumefaciens* Ti plasmid. *Mol. Microbiol.* **34**, 512–522.
- Kim, H. and Farrand, S.K.** (1997) Characterization of the *acc* operon from the nopaline-type Ti plasmid pTiC58, which encodes utilization of agrocinopines A and B and susceptibility to agrocin 84. *J. Bacteriol.* **179**, 7559–7572.

- Kim, J., Heindl, J.E. and Fuqua, C.** (2013) Coordination of Division and Development Influences Complex Multicellular Behavior in *Agrobacterium tumefaciens*. *PLoS One* **8**.
- Kitphati, W., Ngok-ngam, P., Suwanmaneerat, S., Sukchawalit, R. and Mongkolsuk, S.** (2007) *Agrobacterium tumefaciens* fur has important physiological roles in iron and manganese homeostasis, the oxidative stress response, and full virulence. *Appl. Environ. Microbiol.* **73**, 4760–4768.
- Lang, J., Gonzalez-Mula, A., Tacconat, L., Clement, G. and Faure, D.** (2016) The plant GABA signaling downregulates horizontal transfer of the *Agrobacterium tumefaciens* virulence plasmid. *New Phytol.* **210**, 974–983.
- Lang, J., Planamente, S., Mondy, S., Dessaux, Y., Moréra, S. and Faure, D.** (2013) Concerted transfer of the virulence Ti plasmid and companion At plasmid in the *Agrobacterium tumefaciens*-induced plant tumour. *Mol. Microbiol.* **90**, 1178–1189.
- Lang, J., Vigouroux, A., Planamente, S., Sahili, A. El, Blin, P., Aumont-Nicaise, M., Dessaux, Y., Moréra, S. and Faure, D.** (2014) *Agrobacterium* Uses a Unique Ligand-Binding Mode for Trapping Opines and Acquiring A Competitive Advantage in the Niche Construction on Plant Host. *PLoS Pathog.* **10**.
- Lassalle, F., Campillo, T., Vial, L., et al.** (2011) Genomic species are ecological species as revealed by comparative genomics in *Agrobacterium tumefaciens*. *Genome Biol. Evol.* **3**, 762–781.
- Livak, K.J. and Schmittgen, T.D.** (2001) Analysis of relative gene expression data using real-time quantitative PCR and. *Methods* **25**, 402–408.
- Luo, Z.Q. and Farrand, S.K.** (1999) Cloning and characterization of a tetracycline resistance determinant present in *Agrobacterium tumefaciens* C58. *J. Bacteriol.* **181**, 618–626.
- Macho, A.P., Guidot, A., Barberis, P., Beuzón, C.R. and Genin, S.** (2010) A competitive index assay identifies several *Ralstonia solanacearum* type III effector mutant strains with reduced fitness in host plants. *Mol. Plant. Microbe. Interact.* **23**, 1197–1205.
- Marty, L., Vigouroux, A., Aumont-Nicaise, M., Dessaux, Y., Faure, D. and Moréra, S.** (2016) Structural Basis for High Specificity of Amadori Compound and Mannopine Opine Binding in Bacterial Pathogens. *J. Biol. Chem.* **291**, 22638–22649.
- Matthysse, A.G., Jaeckel, P. and Jeter, C.** (2008) *attG* and *attC* mutations of *Agrobacterium tumefaciens* are dominant negative mutations that block attachment and virulence. *Can. J. Microbiol.* **54**, 241–247.

- Maurhofer, M., Reimann, C., Schmidli-Sacherer, P., Heeb, S., Haas, D. and Défago, G.** (1998) Salicylic Acid Biosynthetic Genes Expressed in *Pseudomonas fluorescens* Strain P3 Improve the Induction of Systemic Resistance in Tobacco Against Tobacco Necrosis Virus. *Phytopathology* **88**, 678–84.
- Möller, P., Overlöper, A., Förstner, K.U., Wen, T.N., Sharma, C.M., Lai, E.M. and Narberhaus, F.** (2014) Profound impact of Hfq on nutrient acquisition, metabolism and motility in the plant pathogen *Agrobacterium tumefaciens*. *PLoS One* **9**.
- Navarre, W.W., Zou, S.B., Roy, H., et al.** (2010) PoxA, yjeK, and elongation factor P coordinately modulate virulence and drug resistance in *Salmonella enterica*. *Mol. Cell* **39**, 209–21.
- Planamente, S., Vigouroux, A., Mondy, S., Nicaise, M., Faure, D. and Moréra, S.** (2010) A conserved mechanism of GABA binding and antagonism is revealed by structure-function analysis of the periplasmic binding protein Atu2422 in *Agrobacterium tumefaciens*. *J. Biol. Chem.* **285**, 30294–30303.
- Platt, T.G., Morton, E.R., Barton, I.S., Bever, J.D. and Fuqua, C.** (2014) Ecological dynamics and complex interactions of *Agrobacterium* megaplasmids. *Front. Plant Sci.* **5**, 635.
- Rawat, P., Kumar, S., Pental, D. and Burma, P.K.** (2009) Inactivation of a transgene due to transposition of insertion sequence (IS136) of *Agrobacterium tumefaciens*. *J. Biosci.* **34**, 199–202.
- Rondon, M.R., Ballering, K.S. and Thomas, M.G.** (2004) Identification and analysis of a siderophore biosynthetic gene cluster from *Agrobacterium tumefaciens* C58. *Microbiology* **150**, 3857–3866.
- Rong, L., Karcher, S.J., O’Neal, K., Hawes, M.C., Yerkes, C.D., Jayaswal, R.K., Hallberg, C.A. and Gelvin, S.B.** (1990) *picA*, a novel plant-inducible locus on the *Agrobacterium tumefaciens* chromosome. *J. Bacteriol.* **172**, 5828–36.
- Rong, L.J., Karcher, S.J. and Gelvin, S.B.** (1991) Genetic and molecular analyses of *picA*, a plant-inducible locus on the *Agrobacterium tumefaciens* chromosome. *J. Bacteriol.* **173**, 5110–20.
- Roy, H., Zou, S.B., Bullwinkle, T.J., Wolfe, B.S., Gilreath, M.S., Forsyth, C.J., Navarre, W.W. and Ibba, M.** (2011) The tRNA synthetase paralog PoxA modifies elongation factor-P with (R)- β -lysine. *Nat. Chem. Biol.* **7**, 667–9.
- Sahili, A. El, Li, S.Z., Lang, J., et al.** (2015) A Pyranose-2-Phosphate Motif Is Responsible for Both Antibiotic Import and Quorum-Sensing Regulation in *Agrobacterium tumefaciens*. *PLoS Pathog.* **11**, 1–24.

- Schardl, C.L. and Kado, C.I.** (1983) A functional map of the nopaline catabolism genes on the Ti plasmid of *Agrobacterium tumefaciens* C58. *MGG Mol. Gen. Genet.* **191**, 10–16.
- Schneider, D., Faure, D., Noirclerc-Savoye, M., Barrière, A.-C., Coursange, E. and Blot, M.** (2000) A Broad-Host-Range Plasmid for Isolating Mobile Genetic Elements in Gram-Negative Bacteria. *Plasmid* **44**, 201–207.
- Shaw, C.H.** (1991) Swimming against the tide: chemotaxis in *Agrobacterium*. *Bioessays* **13**, 25–9.
- Shen, Y.-H., Liu, R.-J. and Wang, H.-Q.** (2008) Oxalate decarboxylase from *Agrobacterium tumefaciens* C58 is translocated by a twin arginine translocation system. *J. Microbiol. Biotechnol.* **18**, 1245–51.
- Wang, K., Herrera-Estrella, A. and Montagu, M. Van** (1990) Overexpression of *virD1* and *virD2* genes in *Agrobacterium tumefaciens* enhances T-complex formation and plant transformation. *J. Bacteriol.* **172**, 4432–4440.
- Wilms, I., Möller, P., Stock, A.M., Gurski, R., Lai, E.M. and Narberhaus, F.** (2012) Hfq influences multiple transport systems and virulence in the plant pathogen *Agrobacterium tumefaciens*. *J. Bacteriol.* **194**, 5209–5217.
- Wu, D., Li, A., Ma, F., Yang, J. and Xie, Y.** (2016) Genetic control and regulatory mechanisms of succinoglycan and curdlan biosynthesis in genus *Agrobacterium*. *Appl. Microbiol. Biotechnol.* **100**, 6183–6192.
- Xu, J., Kim, J., Koestler, B.J., Choi, J.-H., Waters, C.M. and Fuqua, C.** (2013) Genetic analysis of *Agrobacterium tumefaciens* unipolar polysaccharide production reveals complex integrated control of the motile-to-sessile switch. *Mol. Microbiol.* **89**, 929–48.
- Yuan, Z.C., Liu, P., Saenkham, P., Kerr, K. and Nester, E.W.** (2008) Transcriptome profiling and functional analysis of *Agrobacterium tumefaciens* reveals a general conserved response to acidic conditions (pH 5.5) and a complex acid-mediated signaling involved in *Agrobacterium*-plant interactions. *J. Bacteriol.* **190**, 494–507.
- Zipfel, C., Kunze, G., Chinchilla, D., Caniard, A., Jones, J.D.G., Boller, T. and Felix, G.** (2006) Perception of the Bacterial PAMP EF-Tu by the Receptor EFR Restricts *Agrobacterium*-Mediated Transformation. *Cell* **125**, 749–760.
- Zupan, J.R., Cameron, T. a., Anderson-Furgeson, J. and Zambryski, P.C.** (2013) Dynamic FtsA and FtsZ localization and outer membrane alterations during polar growth and cell division in *Agrobacterium tumefaciens*. *Proc. Natl. Acad. Sci.* **110**, 9060–9065.

Chapter III

Chapter III: Identification of essential genes for GHB metabolism in *A. tumefaciens* C58 combining transcriptomics and Tn-Seq analysis.

Le GABA est un acide aminé non-protéique présent dans les eucaryotes ainsi que les procaryotes. La dégradation de ce composé aboutit à une production de semialdéhyde succinique (SSA) et d'acide gamma-hydroxybutyrique (GHB) (Chevrot *et al.*, 2006). Le GABA et ses sous-produits s'accumulent dans les tumeurs de *A. tumefaciens* (Deeken *et al.*, 2006; Lang *et al.*, 2016) et il a été démontré que l'expression de la lactonase BlcC, codée par le plasmide At, était induite par le GABA et ses sous-produits GHB et SSA (Carlier *et al.*, 2004; Chevrot *et al.*, 2006). En raison de l'importance du GABA et du GHB, l'objectif de ce travail était de fournir plus d'informations quant à la réponse bactérienne à ces deux composés lorsqu'ils sont utilisés comme source d'énergie.

L'originalité de ce travail réside dans la combinaison de deux techniques différentes mais complémentaires pour l'étude génétique des bactéries : (1) l'analyse transcriptomique donne un aperçu de l'expression des gènes d'*A. tumefaciens* lorsque ces deux composés sont utilisés dans le milieu de culture ; (2) l'exploitation d'une banque Tn-Seq offre la possibilité d'identifier les gènes essentiels à la survie et la compétitivité d'*A. tumefaciens* en condition d'assimilation de ces deux composés.

Anthony Kwasiborski a effectué des cultures cellulaires et extrait l'ARN dans le milieu GABA avant l'analyse transcriptomique. Fabienne Pierre a réalisé la construction d'un des mutants GABA. Florian Lamouche a réalisé les analyses d'expression statistique et différentielle des données transcriptomiques. Leo Mathias a effectué la banque de mutants Tn-Seq. J'ai réalisé les analyses des données transcriptomiques GABA dans ce chapitre, la construction des deux autres mutants GABA et leurs analyses phénotypiques. J'ai aussi effectué les cultures cellulaires, les extractions d'ARN et les analyses des données dans le milieu GHB. Finalement, j'ai réalisé les cultures des mutants Tn-Seq dans les milieux GABA et GHB et leurs analyses correspondantes.

1. INTRODUCTION

In the last years, numerous studies have appeared analyzing data coming from a RNA-Seq transcriptomic approach. This technique is able to give us accurate data on the level of transcription of genes under certain conditions. And although it is a very interesting method, it also has its disadvantages, since not all overexpressed genes are equally important. The technique of Transposon sequencing (Tn-Seq) is being widely used in the last few years. In this case, this method allows us to know if a gene is essential or not in a certain condition. When we compare it with the technique of RNA-Seq, we realize that both give us complementary results, i.e. some genes overexpressed in transcriptomics are the genes of which the mutants are going to disappear in the Tn-Seq method and therefore the genes essentials.

The Ti plasmid of *Agrobacterium tumefaciens* carries almost all genes required for virulence. In it are found genes necessary for the transfer of the T-DNA, genes involved in the catabolism of opines produced by the tumor (and that will serve as source of food for the bacteria) and genes involved in the quorum sensing (QS) communication system that the bacteria will use to regulate, among other things, the diffusion of this plasmid. Besides the Ti plasmid, *A. tumefaciens* usually harbor a companion plasmid called At and it can be co-transfer with the virulent plasmid (Lang *et al.*, 2013). The *blcC* gene, located in this pAt plasmid, codes for a lactonase, which confers the ability cleave the QS signal. The expression of *blcC* is induced by gamma-aminobutyrate (GABA) and its by-products as gamma-hydroxybutyric acid (GHB) and succinic semialdehyde (SSA) (Chevrot *et al.*, 2006). These compounds can be accumulated in the plant tissue. GABA and GABA-related metabolites accumulate in *A. tumefaciens* tumors (Deeken *et al.*, 2006; Lang *et al.*, 2016) and it has been observed that the expression of the lactonase BlcC was induced by GABA and gamma-butyrolactone and their byproducts, GHB and SSA (Carlier *et al.*, 2004; Chevrot *et al.*, 2006).

Here, we present the combination of both methods in the analysis of *Agrobacterium tumefaciens* C58 when it is grown in the presence of two metabolites accumulated in plant tumors, GABA and GHB.

2. MATERIAL AND METHODS

2.1. Bacterial strains and culture conditions

A. tumefaciens C58 was the strain used in this work. *A. tumefaciens* NTLR4 (Cha *et al.*, 1998) was used as acyl-HSL biosensors for oxo-C8HSL. *A. tumefaciens* cells were cultivated at 28°C in TY medium (Bacto tryptone, 5 g/liter; yeast extract, 3 g/liter; agar, 15 g/liter) or *Agrobacterium* broth (AB) minimal medium (K₂HPO₄, 3g/L; NaH₂PO₄, 1g/L; MgSO₄.7H₂O, 0.3 g/L; KCl, 0.15 g/L; CaCl₂, 0.01 g/L; FeSO₄.7H₂O, 2.5 mg/L; pH7) (Chilton *et al.*, 1974) supplemented with sucrose or γ -hydroxybutyric acid (GHB) at 10 mM as carbon source; and ammonium chloride or γ -Aminobutyric acid (GABA) (20mM) as nitrogen source. *Escherichia coli* MFDpir harboring pEGL55 plasmid, auxotroph for diaminopimelic acid (DAP), was used as transposon donor for mutagenesis. *E. coli* DH5 α was the routine host for cloning. *E. coli* cells were cultivated at 37°C in Luria-Bertani modified medium (LBm; NaCl 5 g/L). Antibiotic concentrations were used as the follow concentrations: gentamycin, 25 μ g/mL; kanamycin, 50 μ g/mL; ampicillin, 50 μ g/mL, rifampicin, 100 μ g/mL; DAP (300 μ g/mL).

The growth curves of *A. tumefaciens* were computed by monitoring the optical density at 600 nm.

2.2. Bacterial isolation and RNA extraction and transcriptomic analysis

An overnight culture of *A. tumefaciens* C58 was grown in AB medium with sucrose and ammonium chloride as sources of carbon and nitrogen respectively. This culture was washed twice with NaCl 0.8% and served to inoculate 3 replicates of these different media: AB with sucrose and ammonium chloride; AB with GHB and ammonium chloride; and AB with sucrose and GABA. When these cultures achieved their exponential phase, they were centrifuged and their RNA extracted with the MasterPure™ Complete DNA and RNA Purification Kit according to the supplier's instructions. RNAseq analysis were performed using NextSeq 500 (Illumina, CA, USA) at the I2BC platform (Gif-sur-Yvette, France) using the 75-cycles NextSeq 500 High Output Kit.

2.3. Statistical and differential expression analyses

Count tables have been filtered to retain only genes with a gene count over 1 cpm (count per million) in half of the samples of the dataset. Normalization and differential analyses were performed using generalized linear models as described in DESeq2 package (version 1.12.4, (Love *et al.*, 2014)). The cutoff chosen for differentially expressed genes (DEG) are a False Discovery Rate (FDR) < 0.01 and a log₂ fold change > 2.

2.4. Quantitative RT-qPCR

cDNA was prepared from 1 µg of bacterial or RNA using RevertAid™ H Minus First Strand cDNA Synthesis Kit (Fermentas, Saint-Remy-les-Chevres, France) following the manufacturer's instructions. RT-qPCRs were performed with Lightcycler® 96 (Roche) apparatus. The data were processed using the $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen, 2001).

2.5. Construction of *A. tumefaciens* C58 high density transposon library

The mutagenesis was performed by conjugation in which *Agrobacterium tumefaciens* C58, resistant to rifampicin, was taken as the recipient strain and as the donor strain *E. coli* MFDpir carrying the pEGL55 for transposon *Himar1* mutagenesis. *A. tumefaciens* cells were cultured overnight in TY liquid medium, and *E. coli* MFDpir pEGL55 was cultured for 4 hours in the presence of DAP (300 µg/mL). Both cultures were centrifuged and adjusted to 1 OD₆₀₀. Equivalent amounts of both cultures were mixed, centrifuged and resuspended in TY supplemented with DAP (300 µg / mL). 200 µL of culture medium was incubated on Millipore HAWP01300, type HA, 0.45 µm diameter nitrocellulose filters on a TY plate. The plates were incubated overnight at 28°C. Subsequently, the mass of cells deposited in the filter was resuspended in NaCl 0.8% solution and plated in TY medium supplied only with rifampicin and gentamycin. Cascade dilutions were performed to determine the number of mutants. After 72 hours of incubation, mutants were collected, homogenized and then aliquoted and stored at -80°C in 50 % (v/v) glycerol.

2.6. Inoculation of different mediums with the of *A. tumefaciens* C58 Tn-seq library

Four different aliquots of the Tn-Seq library were selected from the collection at -80 ° C and cultured on TY medium to revive them for 4 hours. Once this process was performed, each culture was washed twice with 0.8% NaCl and DO₆₀₀ was measured. 4 replicates of 10 mL of 3 different mediums were inoculated at an initial OD of 0.05 with each aliquot of the Tn-Seq library: AB medium with GHB and NH₄Cl; AB sucrose and GABA; and AB sucrose and NH₄Cl. They were grown at 28 ° C for 24 hours. Finally, they were centrifuged and stored for further DNA extractions.

2.7. Tn-Seq library generation and sequencing

Genomic DNA was extracted using the DNeasy Blood & Tissue Kit from QIAGEN. 2 µg of gDNA were digested with the MmeI enzyme (BioLabs) during 1 hour at 37°C. After that, digested DNA was incubated for 1 hour at 37°C with FastAP Thermosensitive Alkaline phosphatase (ThermoScientific) and finally the enzyme was deactivated by heating the sample at 75°C during 5 minutes. gDNA samples were then purified using QIAquick PCR purification kit from QIAGEN. At the same time, the p-adapters containing the barcodes needed for sequencing were prepared by an annealing step (Table 1). Once the p-adapters were ready, they were ligated overnight at 16 ° C to the purified gDNA. Finally, the ligation product was used to perform a PCR with primers P7 (CAAGCAGAAGACGGCATAACGATAGACCGGGACTTATCATCCAACCTGT) and P5 (AATGATACGGCGACCACCGAGATCTACACTCTTCCCTACACGACGCTCTCCGATCT). The PCR product, of about 130 base pairs, was separated on agarose gel and purified with the QIAquick gel extraction kit from QIAGEN. The final samples were mixed in equimolar amounts and sequenced by Illumina NextSeq 500 machine 2 x 75 paired-end.

Primer	Sequence
PM13	TTCCCTACACGACGCTCTTCCGATCTATGCTNN
PM13-phosphorylated	p-AGCATAGATCGGAAGAGCGTCGTGTAGGGAAAGAGT-p
PM14	TTCCCTACACGACGCTCTTCCGATCTGTAGTNN
PM14-phosphorylated	p-ACTACAGATCGGAAGAGCGTCGTGTAGGGAAAGAGT-p
PM15	TTCCCTACACGACGCTCTTCCGATCTCTACANN
PM15-phosphorylated	p-TGTAGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGT-p
PM16	TTCCCTACACGACGCTCTTCCGATCTCAGANN
PM16-phosphorylated	p-TCTGAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGT-p
PM17	TTCCCTACACGACGCTCTTCCGATCTGACTANN
PM17-phosphorylated	p-TAGTCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGT-p
PM18	TTCCCTACACGACGCTCTTCCGATCTTCGATNN
PM18-phosphorylated	p-ATCGAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGT-p
PM19	TTCCCTACACGACGCTCTTCCGATCTCGTACNN
PM19-phosphorylated	p-GTACGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGT-p
PM20	TTCCCTACACGACGCTCTTCCGATCTTAGACNN
PM20-phosphorylated	p-GTCTAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGT-p
PM21	TTCCCTACACGACGCTCTTCCGATCTCATGCNN
PM21-phosphorylated	p-GCATGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGT-p
PM22	TTCCCTACACGACGCTCTTCCGATCTTCACGNN
PM22-phosphorylated	p-CGTGAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGT-p
PM23	TTCCCTACACGACGCTCTTCCGATCTGATCGNN
PM23-phosphorylated	p-CGATCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGT-p
PM24	TTCCCTACACGACGCTCTTCCGATCTCCGATGNN
PM24-phosphorylated	p-CATCGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGT-p

Table 1. Primers used for preparing p-adapters.

2.8. Tn-Seq analysis

We used the ARTIST pipeline (Pritchard *et al.*, 2014) in Matlab software (The MathWorks, Natick, MA) to perform the Tn-Seq analysis of the raw data provided by the Illumina sequencing. Two different analysis were carried out, one called EL-ARTIST (for Essential Loci analysis), which defines all loci that are required for growth (i.e., regions with few or no associated transposon insertions) in our *A. tumefaciens* C58 Tn-Seq library samples. The second one, called Con-ARTIST (for Conditionally essential loci analysis), was used to compared our Tn-Seq library samples with the samples growth in the AB mediums, to define which loci that were only required for survival under these conditions.

2.9. Construction of mutants

The gentamicin-resistance cassette from p34S-Gm and kanamycin-resistance cassette from p34S-Km (Dennis and Zylstra, 1998) were introduced into the *atu3407* and the *atu4761* genes respectively harbored by the pGEM-T Easy vector (Promega). The plasmids were eletroporated in *A. tumefaciens* C58 cells. Marker exchange was selected using Gm and Km resistance, and double crossing-overs were verified by specific PCR primers.

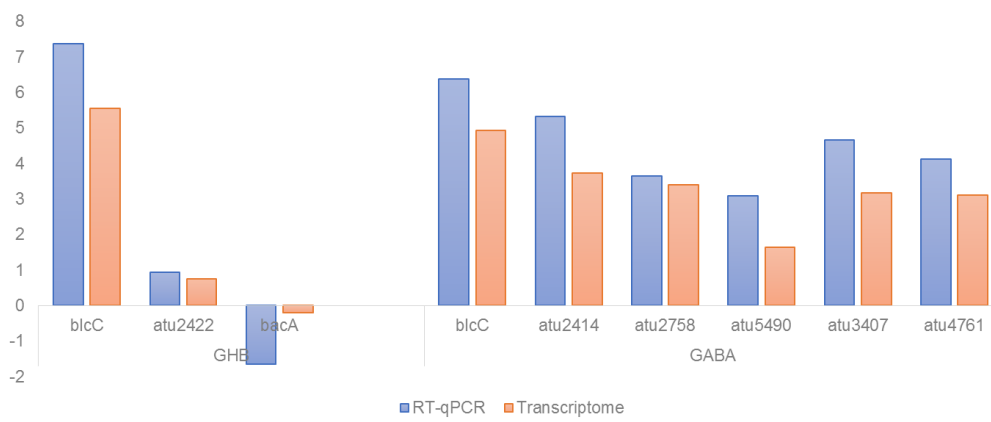


Figure 1. Comparison of the expression fold change (Log₂ FCh) in *A. tumefaciens* C58 cells in GHB and GABA conditions using quantitative RT-PCR (RT-qPCR) or RNAseq techniques.

3. RESULTS

3.1. Transcriptomic analysis of *A. tumefaciens* cell cultures

3.1.1. *A. tumefaciens* C58 in GHB medium

To understand the basic aspects of the metabolism of GHB in *Agrobacterium tumefaciens*, a transcriptomic comparison of two cultures of *A. tumefaciens* C58 in a minimal medium with sucrose or GHB as carbon source was carried out during this work. Total RNA isolated from both cell cultures in exponential phase was used for RNA-Seq analysis, as described above. Gene expression of some genes was confirmed by qRT-PCR. Transcriptome and RT-qPCR data showed a similar gene expression profile (Figure 1).

180 genes were differentially expressed in GHB condition compared to sucrose culture, among them 98 genes were upregulated (data in Table, Annexes) and 82 downregulated (data in Table S2, Annexes). Almost all the differentially expressed genes were located on both chromosomes, 76 in the circular chromosome and 98 in the linear chromosome. Only 6 of them, all upregulated, were located in the At plasmid. Each gene was assigned to a single functional category, resulting in clustering into 18 distinct categories.

Concerning the upregulated genes, the most represented functional categories were involved metabolism processes. These include 14 genes of nucleotide transport and metabolism category, including *serA* (*atu3706*) involved in serine biosynthesis, *argH2* (*atu4651*) involved in arginine biosynthesis, and several ABC transporters as *dppA* (*atu2518*) and *socA* (*atu5006*), involved in the transport of deoxy-fructosyl-glutamine (Amadori compound) (Marty *et al.*, 2016). We found also 11 genes of carbohydrate transport and metabolism, including a fructose biphosphate aldolase (*atu3740*), a pyrophosphate-fructose-6-phosphate 1-phosphotransferase (*pfp* = *atu2115*) involved in the transport of fructose, and some ABC transports as *rbsA* (*atu3818*) and *rsbB* (*atu3821*). 21 genes were classified in the energy production and conversion class, which means ¼ of all upregulated genes. Most of the genes belonging to this category were genes involved in the oxidative phosphorylation pathway, as *cydABC* (*atu4090-atu4092*), *fixP* (*atu1534*), *fixO* (*atu1536*), *sdhCD* (*atu2654-atu2655*). These three last genes have shown to be also upregulated in the presence of salicylic acid (SA), a plant defense signal synthesized by plants in response to the attack of pathogens that induces the expression of the operon *blcABC* (Yuan, Haudecoeur, *et al.*, 2008). Likewise, we found *napABC* (*atu4408-atu4410*), involved in the nitrate reduction. *PckA* (*atu0035*), a phosphoenolpyruvate carboxykinase, and *dctA* (*atu3298*), a C4-dicarboxylate transport protein were found to be highly upregulated. Two remarkable genes in this category were

blcB (*atu5138*), a gamma hydroxybutyrate dehydrogenase, and *blcA* (*atu5137*), a NAD-dependent succinyl-semialdehyde dehydrogenase. These two genes, together with the lactonase *BlcC* (*atu5139*), the most upregulated gene and classified in the general function category, form the *blcABC* operon. Its overexpression confirms the validity and robustness of the transcriptome, since this operon was observed to be induced by GABA and its byproducts (Chevrot *et al.*, 2006).

From the set of downregulated genes, inorganic ion transport and metabolism category, with 21 genes, was the most represented. These genes are involved in tolerance and transport of copper, as *copC* (*atu3990*), *actP* (*atu1195*) and *actP2* (*atu0937*); and iron-siderophore transport as *fecBCD* (*atu3688-atu3690*) and *hmuVUT* (*atu2458-atu2460*). 14 genes were classified in carbohydrate transport and metabolism category, and we can highlight that most of them were involved in transport as the operon *aglEFGAK* (*atu0591-atu0595*).

3.1.2. *A. tumefaciens* C58 in GABA culture

Tumors induced by *Agrobacterium* infection accumulate a wide variety of compound including opines, sugars and amino-acids as gamma-aminobutyric acid (GABA) (Deeken *et al.*, 2006). Within them, GABA promotes degradation of quorum-sensing (QS) signals by the induction of *blc* operon (Chevrot *et al.*, 2006), and probably following its conversion to SSA by a still uncharacterized GABA-transaminase (Khan and Farrand, 2009). To investigate the gene(s) that could be involved in this process we performed a transcriptomic analysis of *A. tumefaciens* C58 culture using GABA as sole nitrogen source. Only genes with a fold change greater than 2 and a p-value less than 0.05 were considered in the study. The results showed that there were only 112 differentially expressed genes, of which the majority were overexpressed, 90 in total (data Table S3 and S4, Annexes). 21 of these genes belonged to the amino acid transport and metabolism category from COG classification. Most of them are involved in transport as the operon *atu1388-atu1391*. Genes implicated in nitrate/nitrite reduction as *nasA* (*atu3900*) and *nasB* (*atu3902*) were also upregulated. These genes belonged to the energy production category. In this category, we can find also *blcA* (*atu5137*) and *blcB* (*atu5138*), which together with *blcC* (*atu5139*) were the highest upregulated genes, and as previously stated, GABA activates this operon. Nearly all the rest of upregulated genes are hypothetical proteins, 29 in total.

Of all the upregulated genes, two appealed our attention because they were the only ones with a certain transaminase activity and they have a high percentage of identity with GABA transaminases of *Rhizobium* species. These were the *atu3407* and *atu4761*

genes, one putative succinyldiaminopimelate aminotransferase and a hypothetical protein, respectively. Therefore, we decided to use them as candidate genes for our functional study. In contrast to some reports in the literature, the GABA transporters *bra* (*atu2422* and *atu2424*) were very slightly upregulated.

Among all the 34 downregulated genes, the category most represented was energy production, with genes involved in nitrate reduction highly represented as *napABCDE* (*atu4405-atu4410*). These results are in contrast with those found in the GHB transcriptome, since in this last one, these genes are overexpressed. This may be because in both transcriptomes nitrogen source is different. Carbohydrate transport and metabolism genes as *rbs* operon (*atu4369-atu4372*) were also downregulated, as they were in GHB transcriptome.

Gene expression of some genes was also confirmed by qRT-PCR. Transcriptome and RT-qPCR data showed a similar gene expression profile (Figure 1).

3.2. Tn-seq analysis of *A. tumefaciens* C58

3.2.1. Bioinformatic analysis of *A. tumefaciens* C58 genome

Before realizing the transposon insertion, it was necessary to check the insertion probability of the transposon. The insertion of the *Himar1* transposon requires TA bases at their integration site, without no other position bias (Barquist *et al.*, 2013). The circular chromosome (2 841 580 pb), linear chromosome (2 075 577bp), the plasmid At and plasmid Ti contain 55 348, 41 503, 13 084 and 5 590 TA sites, respectively. Each replicon harbors 2760, 1848, 543 and 97 CDS respectively (Figure 2). In addition, the counting of TA sites along the genome was necessary to estimate the required number of mutants for adequate coverage. Saturation in the case of *A. tumefaciens* C58 was about 10 times the number of TA sites, i.e. 1 155 250 independent mutants. In conclusion, the Tn-seq mutagenesis applied to *A. tumefaciens* C58 seems to be a powerful tool when genes are large enough, but there could be a significant risk that small genetic elements with or without a function would be forgotten.

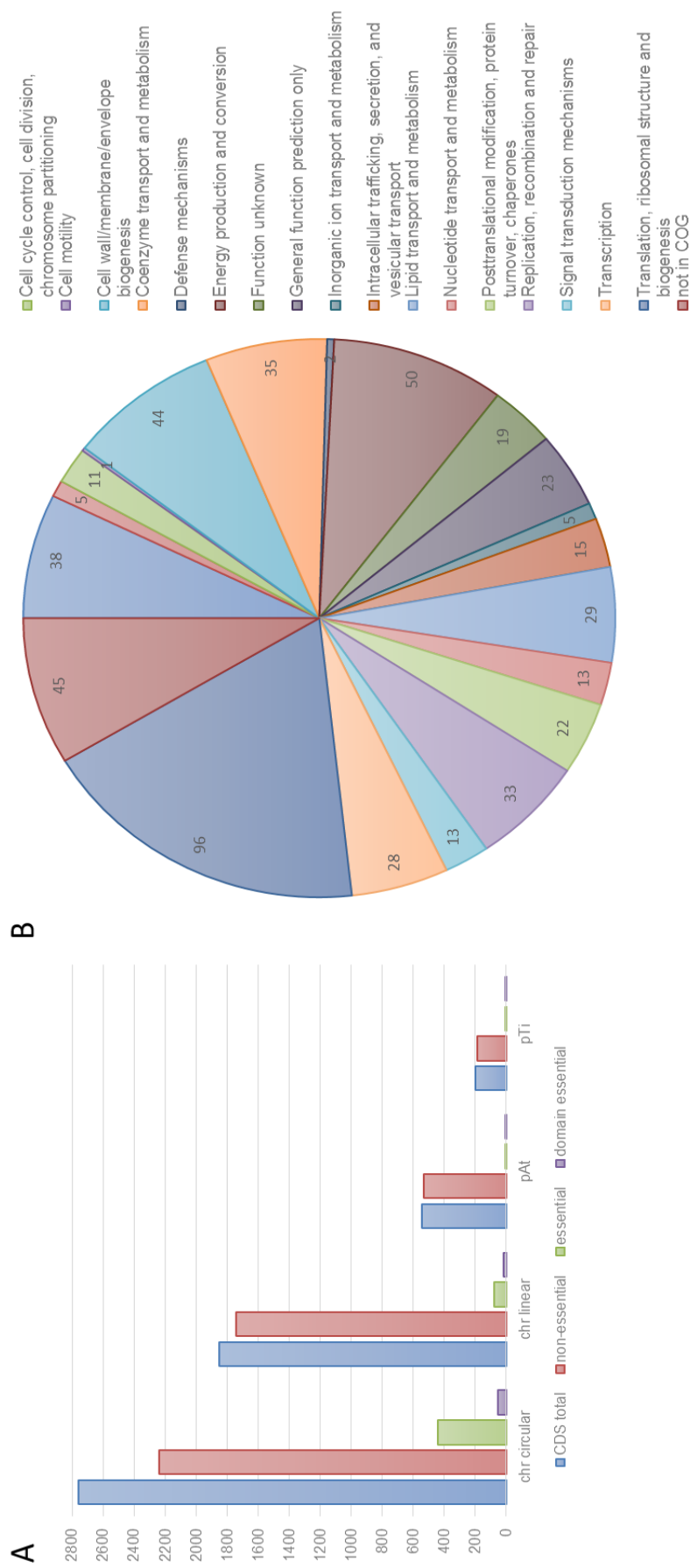


Figure 2. Determination of gene essentiality. A. The distribution of essential genes across the replicons of *A. tumefaciens* C58 using EL-ARTIST analysis. The number of ambiguous genes are not represented in the graphic.

B.) Classification of essential genes by COG (Clusters of Orthologous Groups of proteins) categories. Each number correspond to the total of essential genes of each category in *A. tumefaciens* C58 genome

3.2.2. Tn-seq analysis of *A. tumefaciens* C58 library

Essential loci are typically identified as regions (e.g., genes) in a high density transposon library that lack or have a significantly lower insertion frequency than the average even when the library is grown in optimal laboratory conditions (Chao *et al.*, 2016).

A. tumefaciens C58 *in vitro* library was analyzed by EL-ARTIST and essential and non-essential loci were defined. A total of 527 genes were found to be essential for the survival of *A. tumefaciens* C58 under the Tn-seq library. This represents 10% of the total genes of *A. tumefaciens* C58. These essential genes were mostly located on the circular chromosome (Figure 2a). They were assigned to 19 different metabolic categories (Figure 2b). Among them, the category most represented was the translation, ribosomal structure and biogenesis with genes codifying for most of the ribosomal proteins as *atu1928-atu1951*. Also, genes involved in energy production and conversion were found abundant in the analysis with EL-ARTIST, as for example the *nuoABCDEFGHIJKLMN* operon (*atu1268-atu1283*). We found also some genes essential for the maintenance of both plasmids in *A. tumefaciens* C58. These genes are involved in the replication of both plasmids, the operon *repABC* (*atu5000-atu5002* for pAt and *atu6043-atu6045* for pTi), as well as some hypothetical protein such as *atu5013* (that forms part of a toxin/antitoxin system and located in pAt), and *atu8054* (pTi); *atu5353* (IS3 family transposase in pAt) and finally in Ti plasmid *atu6009*, an indole-3-lactate synthase and *atu6082*, a AAA ATPase family protein.

3.2.3. Tn-seq analysis of *A. tumefaciens* C58 GHB cell cultures

The objective of microbial genetics is to understand the relationship between genotype and phenotype. But gene screens can be an extremely long and tedious job. The recent development of random transposon insertion mutagenesis followed by deep sequencing has accelerated this process. This method allows the analysis of essential genes in a specific life condition using saturated transposon mutant libraries. Once we obtained the transposon library as described in Material and Methods, we used 4 different aliquots to inoculate a poor medium with GHB as a carbon source. After 24 hours of growth, we performed the Tn-seq library for subsequent sequencing of the samples. The raw data was then analyzed by Con-ARTIST pipeline. The ARTIST pipeline predicts the essential genes for the bacteria in a given condition. Our results for *Agrobacterium* GHB samples predict that there were 34 essential genes of *A. tumefaciens* C58. Among the essential genes, 20 of them were classified in the category of transport and metabolism of amino acids, since we compared samples coming from a prototrophic condition (Tn-seq library samples) with samples from auxotrophic condition, evidently the entire metabolism of

amino acids is essential in the new condition. Within them, we found genes involved in the biosynthesis of tryptophan, histidine, phenylalanine, cysteine, valine/isoleucine, serine, aspartate and leucine (Table 2). The results also showed essential genes involved in carbohydrate transport and metabolism as *pckA* (*atu0035*), *eno* (*atu1426*), *tktA* (*atu3736*), *gapA* (*atu3737*) and a fructose bisphosphate aldolase (*atu3740*). BlcA (*atu5137*) and BlcB (*atu5138*) proteins, belonging to the *blcABC* operon, seemed to be also essential in GHB culture. We observed also that no transporters were found to be essential for GHB growth condition. The results were consistent with the transcriptome analysis and gives us an idea of the complexity of the transcriptome, since an overexpressed gene does not directly mean that it is an essential gene.

An analysis of the Tn-seq raw data (number of reads) was also carried out, but with the same technique used for the analysis of the transcriptome data. The results obtained are compatible with the transcriptome and those of the analysis by the ARTIST program. So, it seems that it could be possible to analyze Tn-seq sequencing data as if they were transcriptomic data.

3.2.4. Tn-seq analysis of GABA cell cultures

GABA was used as a sole nitrogen source in cell cultures. Nearly all the 41 genes found to be essential for the survival of *A. tumefaciens* C58 in this condition were involved in amino acid and nucleotide transport and metabolism (Table 3). Most of the genes of the tryptophan (*atu0017*, *atu0018*, *atu1687* and *atu2289*), leucine (*atu2264*, *atu2790* and *atu2791*) biosynthesis were essential. Some others amino acid synthesis genes seemed also to be essential, like genes involved in histidine, phenylalanine and serine biosynthesis. The two-component system *ntrBC* (*atu1445* and *atu1446*) was found also essential in GABA cultures. This system responds to a low nitrogen availability. Neither genes from *blcABC* operon nor the GABA transport genes *bra* (*atu2422* and *atu2424*) appeared to be essential. We found no genes with putative transaminase activity and only two of them were also among the upregulated genes of the transcriptome: *ilvL* (*atu2036*), a acetolactate synthase 3 catalytic subunit and *leuA* (*atu2264*) coding for a 2-isopropylmalate synthase. These data revealed us the limits of the technic. As well as for the GHB data, we analyzed the raw data in the same way of the transcriptomic data with same results.

As well as for the GHB cultures, a transcriptome analysis of the raw data was performed and resemble to be compatible with the Con-ARTIST results.

Gene	Name	Product	COG classification
<i>atu0017</i>	<i>trpF</i>	N-(5'-phosphoribosyl)anthranilate isomerase	Amino acid transport and metabolism
<i>atu0018</i>	<i>trpB</i>	Tryptophan synthase beta chain	Amino acid transport and metabolism
<i>atu0035</i>	<i>pckA</i>	Phosphoenolpyruvate carboxykinase [ATP]	Energy production and conversion
<i>atu0040</i>	<i>hisA</i>	Imidazole-4-carboxamide isomerase	Amino acid transport and metabolism
<i>atu0041</i>	<i>hisH</i>	Imidazole glycerol phosphate synthase subunit HisH	Amino acid transport and metabolism
<i>atu8115</i>		Hypothetical protein	Unclassified
<i>atu0043</i>	<i>hisB</i>	Imidazoleglycerol-phosphate dehydratase	Amino acid transport and metabolism
<i>atu0099</i>	<i>pheA</i>	Prephenate dehydratase	Amino acid transport and metabolism
<i>atu0432</i>		O-succinylhomoserine sulfhydrylase	Amino acid transport and metabolism
<i>atu0434</i>	<i>dcd</i>	2'-deoxycytidine 5'-triphosphate deaminase	Nucleotide transport and metabolism
<i>atu0678</i>	<i>hisZ</i>	Putative histidyl-tRNA synthetase	Amino acid transport and metabolism
<i>atu0679</i>	<i>hisG</i>	ATP phosphoribosyltransferase	Amino acid transport and metabolism
<i>atu0775</i>	<i>thrB</i>	Homoserine kinase	General function prediction only
<i>atu1426</i>	<i>eno</i>	Enolase	Carbohydrate transport and metabolism
<i>atu1427</i>		Hypothetical protein	Cell cycle control, cell division, chromosome partitioning
<i>atu1571</i>	<i>cysE</i>	Serine acetyltransferase	Amino acid transport and metabolism
<i>atu1589</i>		Aminotransferase	Amino acid transport and metabolism
<i>atu1918</i>	<i>ilvD</i>	Dihydroxy-acid dehydratase	Amino acid transport and metabolism
<i>atu2019</i>	<i>ilvC</i>	Ketol-acid reductoisomerase	Amino acid transport and metabolism
<i>atu2020</i>		TetR family transcriptional regulator	Transcription
<i>atu2036</i>	<i>ilvI</i>	Acetolactate synthase 3 catalytic subunit	Amino acid transport and metabolism
<i>atu2040</i>	<i>serB</i>	Ahosphoserine phosphatase	Amino acid transport and metabolism
<i>atu2196</i>	<i>aatA</i>	Aspartate aminotransferase	Amino acid transport and metabolism
<i>atu2264</i>	<i>leuA</i>	2-isopropylmalate synthase	Amino acid transport and metabolism
<i>atu2289</i>	<i>trpE(G)</i>	Anthranilate synthase	Amino acid transport and metabolism
<i>atu2790</i>	<i>leuD</i>	3-isopropylmalate dehydratase small subunit	Amino acid transport and metabolism
<i>atu2823</i>	<i>purH</i>	Bifunctional purine biosynthesis protein	Nucleotide transport and metabolism
<i>atu3707</i>	<i>serC</i>	Phosphoserine aminotransferase	Amino acid transport and metabolism
<i>atu3736</i>	<i>tktA</i>	Transketolase	Carbohydrate transport and metabolism
<i>atu3737</i>	<i>gapA</i>	Glyceraldehyde 3-phosphate dehydrogenase	Carbohydrate transport and metabolism
<i>atu3740</i>		Fructose bisphosphate aldolase	Carbohydrate transport and metabolism
<i>atu3754</i>	<i>purE</i>	Phosphoribosylaminoimidazole carboxylase catalytic subunit	Nucleotide transport and metabolism
<i>atu5137</i>	<i>blcA</i>	NAD-dependent succinyl-semialdehyde dehydrogenase	Energy production and conversion
<i>atu5138</i>	<i>blcB</i>	Gamma hydroxybutyrate dehydrogenase	Energy production and conversion

Table 1. *A. tumefaciens* essential genes for growth and survival in GHB condition. All genes in *A. tumefaciens* were categorized by Con-ARTIST for their contributions towards growth in AB medium supplemented with GHB as source of carbon. Color blue represents common genes in both transcriptomic and Tn-seq data.

Gene	Name	Product	COG classification
<i>atu0017</i>	<i>trpF</i>	N-(5'-phosphoribosyl)anthranilate isomerase	Amino acid transport and metabolism
<i>atu0018</i>	<i>trpB</i>	Tryptophan synthase beta chain	Amino acid transport and metabolism
<i>atu0040</i>	<i>hisA</i>	Imidazole-4-carboxamide isomerase	Amino acid transport and metabolism
<i>atu0041</i>	<i>hisH</i>	Imidazole glycerol phosphate synthase subunit HisH	Amino acid transport and metabolism
<i>atu8115</i>		Hypothetical protein	Unclassified
<i>atu0043</i>	<i>hisB</i>	Imidazoleglycerol-phosphate dehydratase	Amino acid transport and metabolism
<i>atu0099</i>	<i>pheA</i>	Prephenate dehydratase	Amino acid transport and metabolism
<i>atu0399</i>	<i>pyrC</i>	Dihydroorotase	Nucleotide transport and metabolism
<i>atu0400</i>	<i>pyrE</i>	Orotate phosphoribosyltransferase	Nucleotide transport and metabolism
<i>atu0432</i>		O-succinylhomoserine sulfhydrylase	Amino acid transport and metabolism
<i>atu0434</i>	<i>dcd</i>	2'-deoxycytidine 5'-triphosphate deaminase	Nucleotide transport and metabolism
<i>atu0678</i>	<i>hisZ</i>	Putative histidyl-tRNA synthetase	Amino acid transport and metabolism
<i>atu0679</i>	<i>hisG</i>	ATP phosphoribosyltransferase	Amino acid transport and metabolism
<i>atu0775</i>	<i>thrB</i>	Homoserine kinase	General function prediction only
<i>atu1307</i>	<i>pyrC</i>	Dihydroorotase	Nucleotide transport and metabolism
<i>atu1308</i>	<i>pyrB</i>	Aspartate carbamoyltransferase	Nucleotide transport and metabolism
<i>atu1445</i>	<i>ntrB</i>	Two component sensor kinase	Signal transduction mechanisms
<i>atu1446</i>	<i>ntrC</i>	Two component response regulator	Signal transduction mechanisms
<i>atu1588</i>	<i>hom</i>	Homoserine dehydrogenase	Amino acid transport and metabolism
<i>atu1589</i>		Aminotransferase	Amino acid transport and metabolism
<i>atu1687</i>	<i>trpD</i>	Anthranilate phosphoribosyltransferase	Amino acid transport and metabolism
<i>atu1843</i>	<i>purC1</i>	Phosphoribosylaminoimidazole-succinocarboxamide synthase 1	Nucleotide transport and metabolism
<i>atu1844</i>	<i>purS</i>	Nucleotide transport and metabolism	Unclassified
<i>atu1845</i>	<i>purQ</i>	Phosphoribosylformylglycinamide synthase I	Nucleotide transport and metabolism
<i>atu2020</i>		TetR family transcriptional regulator	Transcription
<i>atu2036</i>	<i>ilvI</i>	Acetolactate synthase 3 catalytic subunit	Amino acid transport and metabolism
<i>atu2196</i>	<i>aatA</i>	Aspartate aminotransferase	Amino acid transport and metabolism
<i>atu2264</i>	<i>leuA</i>	2-isopropylmalate synthase	Amino acid transport and metabolism
<i>atu2289</i>	<i>trpE(G)</i>	Anthranilate synthase	Amino acid transport and metabolism
<i>atu8028</i>		Hypothetical protein	Unclassified
<i>atu2790</i>	<i>leuD</i>	3-isopropylmalate dehydratase small subunit	Amino acid transport and metabolism
<i>atu2791</i>	<i>leuB</i>	3-isopropylmalate dehydrogenase	Energy production and conversion
<i>atu2823</i>	<i>purH</i>	Bifunctional purine biosynthesis protein PurH	Nucleotide transport and metabolism
<i>atu3289</i>	<i>thiE</i>	Thiamine-phosphate synthase	Coenzyme transport and metabolism
<i>atu3290</i>	<i>thiD</i>	Phosphomethylpyrimidine kinase	Coenzyme transport and metabolism
<i>atu3291</i>		Lpr family transcriptional regulator	Transcription
<i>atu3292</i>	<i>dadB</i>	Alanine racemase, catabolic	Cell wall/membrane/envelope biogenesis
<i>atu3707</i>	<i>serC</i>	Phosphoserine aminotransferase	Amino acid transport and metabolism
<i>atu3753</i>		Hypothetical protein	Unclassified
<i>atu3754</i>	<i>purE</i>	Phosphoribosylaminoimidazole carboxylase catalytic subunit	Nucleotide transport and metabolism
<i>atu3755</i>	<i>purK</i>	Phosphoribosylaminoimidazole carboxylase ATPase subunit	Nucleotide transport and metabolism

Table 2. *A. tumefaciens* essential genes for growth and survival in GABA condition. All genes in *A. tumefaciens* were categorized by Con-ARTIST for their contributions towards growth in AB medium supplemented with GABA as source of nitrogen. Color blue represents genes in common with transcriptomic data.

3.3. *atu3407* and *atu4761* genes are not involved in the degradation of GABA to SSA

To evaluate the influence of the two candidate GABA transaminases, we constructed single mutants of both genes and a double mutant as described in Material and methods. When we grew the mutants in minimal medium with GABA as a sole source of nitrogen, their growth was similar to that of the wild type strain (Figure 3), suggesting that they were not affected in GABA degradation. In addition, a degradation test of QS signals was also performed (data not shown). This test had as purpose to check if the degradation of the signals of QS was degradation when we compared with the wild type strain. All together seem to point out that none of both genes are involved in the conversion of GABA to SSA by *Agrobacterium tumefaciens* C58.

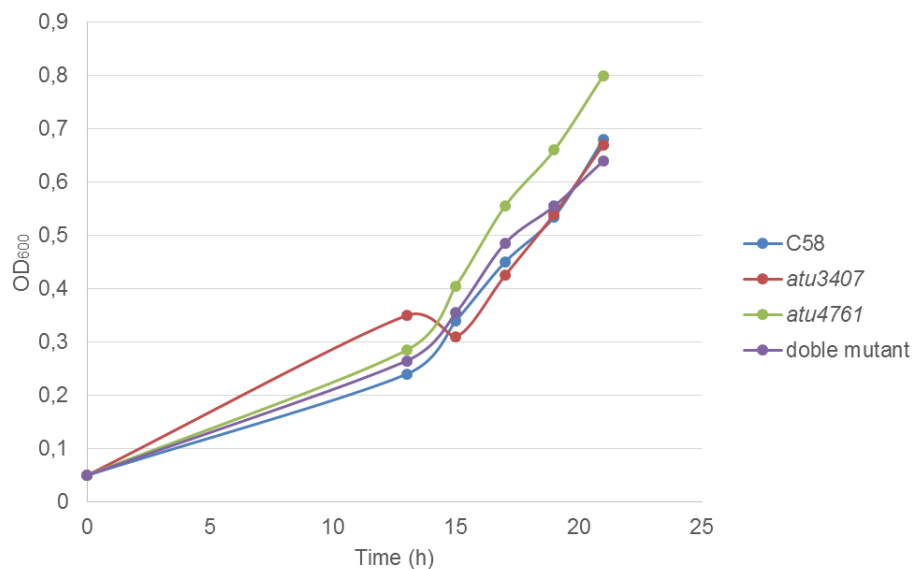


Figure 3. Growth of *A. tumefaciens* C58 wild type and putative transaminases *atu3407*, *atu4761* and double mutant in the presence of GABA as sole nitrogen source.

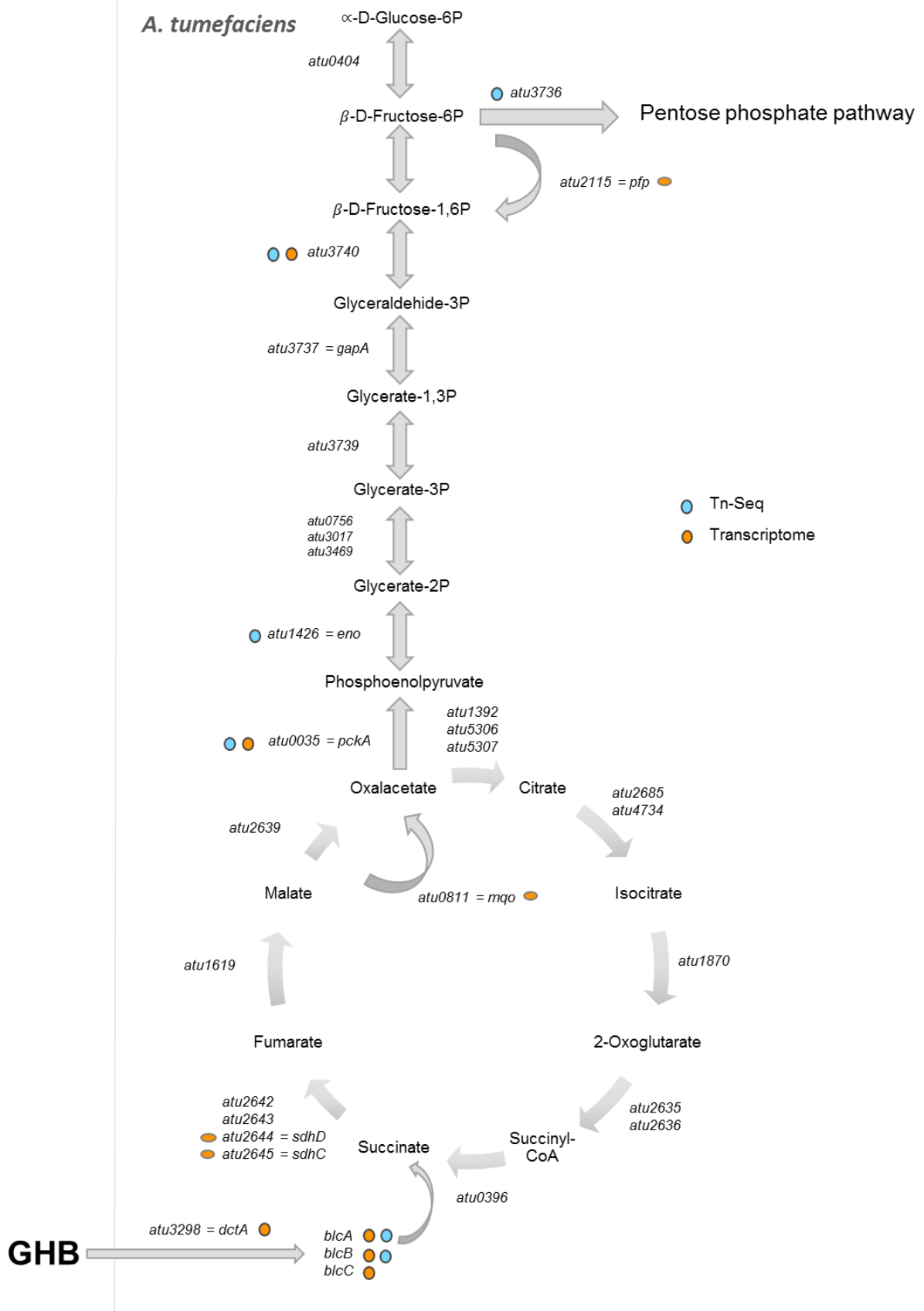


Figure 4. Schema of the gluconeogenesis pathway. Upregulated genes in transcriptome data are colored in orange and essential genes found in Tn-Seq analysis are colored in blue.

4. DISCUSSION

Transcriptomics offers a comprehensive way of analyzing changes in gene expression of an organism under different conditions (Schulze *et al.*, 2016). The widespread adoption of the RNA-seq technique has become a standard part of the toolkit used in life sciences research (Conesa *et al.*, 2016). Moreover, the use of Tn-seq technic has now been applied to several different microbial species. By comparing mutant frequencies of initial and final population, obtained by high throughput sequencing, it is possible to determine the contribution of a gene in fitness (van Opijnen *et al.*, 2014). In the last decade, the role of GABA and its related metabolites have been extensively studied due to their impact in the virulence of *Agrobacterium tumefaciens* (Haudecoeur and Faure, 2010). In this study, we examine the GHB and GABA degradation pathways combining two different approaches: transcriptomics and transposon-insertion sequencing.

The first objective of this study was to gain a broader understanding of the GHB (γ -hydroxybutyric acid) uptake and degradation pathways in *A. tumefaciens* C58 cells. PckA, a phosphoenolpyruvate carboxykinase, seems to be responsible of the reversible decarboxylation and phosphorylation of oxaloacetate to form phosphoenolpyruvate (PEP), which will be used in the citrate cycle (TCA cycle, Krebs cycle). GHB is, as oxaloacetate, a four-carbon carboxylic acid, so we could propose that could be primarily converted in PEP by this gene. Another transcriptomic study revealed that *pckA* was also induced in acidic conditions (Yuan, Liu, *et al.*, 2008). Moreover, it has been shown that *pckA* defected mutants were highly attenuated in virulence and also in *vir* gene expression (Liu *et al.*, 2005). The fact that this gene was highly upregulated in the transcriptome and also found to be essential in the Tn-seq analysis, suggest that this enzyme plays a key role for the metabolism of GHB. Similarly, was the case of *atu3740* gene, a fructose bisphosphate aldolase. This gene, involved in the transformation of glyceraldehyde-3P to β -D-fructose 1-phosphate, is both upregulated and essential. Another highly upregulated gene was *dctA*. This gene appears to be involved in the transport of C4 metabolites. In *Sinorhizobium meliloti*-alfalfa symbiosis, nodules formed by *dctA* mutants cannot fix nitrogen because the TCA cycle dicarboxylic acid intermediates, that serve the bacteria as source of energy, are supplied by the host plant (Yarosh *et al.*, 1989). Also, a recent study of an *A. tumefaciens* A6 microarray in the presence of succinic semialdehyde (SSA) showed that this gene is also highly induced by this compound (Wang *et al.*, 2006). But the evidence that, in the Tn-seq analysis, there are not essential genes involved in metabolites transport led us think that several

transport systems could be involved in the transport of GHB towards the bacteria. As shown in Figure 4, also genes *sdhCD* (*atu2654-atu2655*) coding for succinate dehydrogenase of the TCA cycle were found upregulated in the transcriptome. The expression of part of this aerobic respiratory chain demonstrates that this process is important in GHB metabolism but not essential, as none of them appeared in the Tn-seq analysis. GHB is transformed to SSA by BlcB, but this compound is extremely toxic for the bacteria and can generate reactive oxygen species (ROS). That is why to protect bacteria against ROS-induced damage, these genes were found to be upregulated. To conclude, data obtained from the transcriptomic analysis shows that the *blcABC* operon was highly upregulated and both *blcA* and *blcB* are essential in the metabolism of GHB. The results obtained from the transcriptomic analysis suggest that GHB could be degraded and incorporated via the gluconeogenesis pathway. This hypothesis is supported by the results of the Tn-Seq analysis which has shown that there are also some genes of this pathway that are essential for *A. tumefaciens* cells in the presence of GHB (Figure 4).

In contrast to the GHB results, the combination of transcriptomic and Tn-seq methods seems not to be sufficient to establish possible GABA assimilation pathways. The fact that we found only two genes with putative transaminase activity in the transcriptome and none in the Tn-Seq analysis demonstrates us that both techniques have certain limits. The main limitation of the Tn-seq experimental results is that we compared a prototrophic to an auxotrophic condition. Clearly, further study of both conditions compared to a minimal medium with sucrose and NH_4 as energy sources will be required to prove which the most essential genes in each condition are. Also, the fact that only completely essential genes, i.e. genes without a possible complementation by others or redundant genes; could be a disadvantage of this technique, since important pathways or reactions could be unnoticed.

5. BIBLIOGRAPHY

- Barquist, L., Boinett, C.J. and Cain, A.K.** (2013) Approaches to querying bacterial genomes with transposon-insertion sequencing. *RNA Biol.* **10**, 1161–9.
- Carlier, A., Chevrot, R., Dessaux, Y. and Faure, D.** (2004) The assimilation of gamma-butyrolactone in *Agrobacterium tumefaciens* C58 interferes with the accumulation of the N-acyl-homoserine lactone signal. *Mol. Plant. Microbe. Interact.* **17**, 951–7.
- Cha, C., Gao, P., Chen, Y.C., Shaw, P.D. and Farrand, S.K.** (1998) Production of acyl-homoserine lactone quorum-sensing signals by gram-negative plant-associated bacteria. *Mol. Plant. Microbe. Interact.* **11**, 1119–29.
- Chao, M.C., Abel, S., Davis, B.M. and Waldor, M.K.** (2016) The design and analysis of transposon insertion sequencing experiments. *Nat. Rev. Microbiol.* **14**, 119–128.
- Chevrot, R., Rosen, R., Haudecoeur, E., Cirou, A., Shelp, B.J., Ron, E. and Faure, D.** (2006) GABA controls the level of quorum-sensing signal in *Agrobacterium tumefaciens*. *Proc. Natl. Acad. Sci.* **103**, 7460–7464.
- Chilton, M.D., Currier, T.C., Farrand, S.K., Bendich, A.J., Gordon, M.P. and Nester, E.W.** (1974) *Agrobacterium tumefaciens* DNA and PS8 bacteriophage DNA not detected in crown gall tumors. *Proc. Natl. Acad. Sci. U. S. A.* **71**, 3672–6.
- Conesa, A., Madrigal, P., Tarazona, S., et al.** (2016) A survey of best practices for RNA-seq data analysis. *Genome Biol* **17**, 13.
- Deeken, R., Engelmann, J.C., Efetova, M., et al.** (2006) An Integrated View of Gene Expression and Solute Profiles of *Arabidopsis* Tumors: A Genome-Wide Approach. *Plant Cell Online* **18**, 3617–3634.
- Dennis, J.J. and Zylstra, G.J.** (1998) Plasposons: modular self-cloning minitransposon derivatives for rapid genetic analysis of gram-negative bacterial genomes. *Appl. Environ. Microbiol.* **64**, 2710–5.
- Haudecoeur, E. and Faure, D.** (2010) A fine control of quorum-sensing communication in *Agrobacterium tumefaciens*. *Commun. Integr. Biol.* **3**, 84–88.
- Khan, S.R. and Farrand, S.K.** (2009) The B1cC (AttM) lactonase of *Agrobacterium tumefaciens* does not quench the quorum-sensing system that regulates Ti plasmid conjugative transfer. *J. Bacteriol.* **191**, 1320–1329.
- Lang, J., Gonzalez-Mula, A., Tacconnat, L., Clement, G. and Faure, D.** (2016) The plant GABA signaling downregulates horizontal transfer of the *Agrobacterium tumefaciens* virulence plasmid. *New Phytol.* **210**, 974–983.
- Lang, J., Planamente, S., Mondy, S., Dessaux, Y., Moréra, S. and Faure, D.** (2013) Concerted transfer of the virulence Ti plasmid and companion At plasmid in the *Agrobacterium tumefaciens*-induced plant tumour. *Mol. Microbiol.* **90**, 1178–1189.

- Liu, P., Wood, D. and Nester, E.W.** (2005) Phosphoenolpyruvate carboxykinase is an acid-induced, chromosomally encoded virulence factor in *Agrobacterium tumefaciens*. *J. Bacteriol.* **187**, 6039–45.
- Livak, K.J. and Schmittgen, T.D.** (2001) Analysis of relative gene expression data using real-time quantitative PCR and. *Methods* **25**, 402–408.
- Love, M.I., Huber, W. and Anders, S.** (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550.
- Marty, L., Vigouroux, A., Aumont-Nicaise, M., Dessaux, Y., Faure, D. and Moréra, S.** (2016) Structural Basis for High Specificity of Amadori Compound and Mannopine Opine Binding in Bacterial Pathogens. *J. Biol. Chem.* **291**, 22638–22649.
- Opijnen, T. van, Lazinski, D.W. and Camilli, A.** (2014) Genome-Wide Fitness and Genetic Interactions Determined by Tn-seq, a High-Throughput Massively Parallel Sequencing Method for Microorganisms. *Curr. Protoc. Mol. Biol.* **106**, 7.16.1-24.
- Pritchard, J.R., Chao, M.C., Abel, S., Davis, B.M., Baranowski, C., Zhang, Y.J., Rubin, E.J. and Waldor, M.K.** (2014) ARTIST: High-Resolution Genome-Wide Assessment of Fitness Using Transposon-Insertion Sequencing. *PLoS Genet.* **10**.
- Schulze, S., Schleicher, J., Guthke, R. and Linde, J.** (2016) How to predict molecular interactions between species? *Front. Microbiol.* **7**, 1–13.
- Wang, C., Zhang, H.B., Wang, L.H. and Zhang, L.H.** (2006) Succinic semialdehyde couples stress response to quorum-sensing signal decay in *Agrobacterium tumefaciens*. *Mol. Microbiol.* **62**, 45–56.
- Yarosh, O.K., Charles, T.C. and Finan, T.M.** (1989) Analysis of C₄-dicarboxylate transport genes in *Rhizobium meliloti*. *Mol. Microbiol.* **3**, 813–823.
- Yuan, Z.C., Haudecoeur, E., Faure, D., Kerr, K.F. and Nester, E.W.** (2008) Comparative transcriptome analysis of *Agrobacterium tumefaciens* in response to plant signal salicylic acid, indole-3-acetic acid and γ -amino butyric acid reveals signalling cross-talk and *Agrobacterium*-plant co-evolution. *Cell. Microbiol.* **10**, 2339–2354.
- Yuan, Z.C., Liu, P., Saenkham, P., Kerr, K. and Nester, E.W.** (2008) Transcriptome profiling and functional analysis of *Agrobacterium tumefaciens* reveals a general conserved response to acidic conditions (pH 5.5) and a complex acid-mediated signaling involved in *Agrobacterium*-plant interactions. *J. Bacteriol.* **190**, 494–507.

Chapter IV

Chapter III: Experimental evolution in *Agrobacterium tumefaciens* C58

L'évolution expérimentale est utilisée depuis plusieurs années pour analyser les processus d'évolution qui sont derrière l'adaptation d'un organisme à une condition déterminée imposée par l'expérimentateur. Cette approche peut être utilisée pour tester différentes hypothèses (Kawecki *et al.*, 2012). Récemment, sa combinaison avec les techniques de séquençage à haut débit a permis d'identifier facilement les variations génétiques pour lesquelles il a été possible d'établir des liens entre les phénotypes et les génotypes sélectionnés (Guidot *et al.*, 2014).

Dans ce travail, nous avons entrepris d'analyser les changements génétiques qui émergent lors de cultures successives d'*A. tumefaciens* C58 dans des tumeurs induites sur plantes de tomate pendant 2 années. Pour cela, la dynamique évolutive de 3 génotypes d'*A. tumefaciens* C58 ont été comparés: une population de type sauvage, une autre population incapable de produire des signaux QS, donc incapable de transférer son plasmide Ti, et une dernière population capable de transférer de manière constitutive son plasmide Ti. En les comparants, nous aussi avons cherché à déterminer si le QS avait un impact sur la nature des variantes sélectionnées.

Les analyses de séquençage ont été réalisées en collaboration avec les plateformes technologiques de l'Université de Malaisie et les plateformes techniques Imagif. Certains des mutants utilisés dans l'étude étaient disponibles en laboratoire. J'ai réalisé toute l'expérience d'évolution et aussi les analyses des données de séquençage.

1. INTRODUCTION

Agrobacterium tumefaciens genome consists of a circular chromosome, a linear chromosome and two dispensable plasmids, the Ti (tumor-inducing) plasmid and plasmid At (*Agrobacterium tumefaciens*). *A. tumefaciens* is well known for its ability to transform plant cells by transferring a piece of DNA, the T-DNA, which is part of the Ti plasmid. *Agrobacterium* cells are also able to transfer its plasmids to other bacteria via conjugation. Previous research have shown that the quorum-sensing activates the dissemination of the Ti plasmid and also that it exits a co-transfer of both plasmids. To evaluate dynamics of the genome and plasmidome in the course of infection and the role of the quorum-sensing in the same, we performed evolution experiments by serial passages of three different strains in tumor induced on wounded stem of the host plant *Solanum lycopersicum*. A wild type strain and two mutant strains affected in the quorum-sensing has been used: a Tral defective strain for producing quorum-sensing signals and a strain in which the master transcriptional repressor AccR of the quorum-sensing is inactivated.

2. MATERIAL AND METHODS

2.1. Bacterial strains and growth conditions

A. tumefaciens C58 was the strain used in this work. *A. tumefaciens* cells were cultivated at 28°C in TY medium (bacto tryptone, 5 g/L; yeast extract, 3 g/L; agar, 15 g/L) and in AB minimal medium (Chilton *et al.*, 1974) supplemented with NH₄Cl (1g/L) and sucrose (2 g/L). Antibiotic concentrations were used as the follow concentrations: gentamycin, 25 µg/mL; kanamycin, 100 µg/mL; rifampicin, 100 µg/mL. Cycloheximide (100 µg/mL) was used as an antifungal agent.

For growth curve experiments of *A. tumefaciens*, we grew strains in 5 mL of TY or AB medium. Growth was followed by the determination of OD₆₀₀ values, and viable counts (colony-forming units, CFU) were obtained by plating onto agar medium suitable dilutions made in 0.8% NaCl.

2.2. Construction of ancestor strains

A. tumefaciens C58 was used to construct the pTi-*tral*::Gm mutant as described in (Haudecoeur, Tannières, *et al.*, 2009). A spontaneous rifampicin-resistant mutant, *A. tumefaciens* C58RpTi-*tral*::Gm, was isolated from *A. tumefaciens* C58pTi-*tral*::Gm. This rifampicin-resistant mutant was used as the Tral-defective ancestor for the evolution experiment. However, the Ti plasmid was cured from this ancestor and then the resulting strain *A. tumefaciens* C58R was used as a recipient strain to receive either the plasmid pTi-*atu6148*::Gm or pTi-*accR*::Gm (Haudecoeur, Tannières, *et al.*, 2009). Ti plasmid

conjugation was performed in tomato plants *Solanum lycopersicum* plants (F1 hybrid Dona, Vilmorin) by co-infecting the recipient *A. tumefaciens* C58R with each of the donors *A. tumefaciens* C58pTi-*atu6148*::Gm and C58pTi-*accR*::Gm (Haudecoeur, Tannières, *et al.*, 2009). Co-infected tumors were crushed and plated on TY Rif Gm agar medium to select the transconjugants. Their genotypes were verified by PCR with specific primers. Finally, all the three ancestors *A. tumefaciens* C58RpTi-*tral*::Gm, C58RpTi-*atu6148*::Gm and C58RpTi-*accR*::Gm were stored at -80°C for subsequent use in evolution experiment.

2.3. Experimental evolution

The experimental evolution study was carried out on *S. lycopersium* plants (F1 hybrid Dona, Vilmorin). Tomato plants were grown in greenhouse under long day conditions and controlled temperature (24–26°C). TY overnight cultures of each of the ancestors *A. tumefaciens* C58RpTi-*tral*::Gm, C58RpTi-*accR*::Gm and C58RpTi-*atu6148*::Gm were adjusted to 1 OD₆₀₀, plated on TY agar and incubated after 24 hours at 28°C. Each genotype was used to inoculate six different 4 weeks old tomato plants which were wounded with a scalpel between the first and the second stem nodes as described previously (Planamente *et al.*, 2010). Hence, 18 independent lines were started to propagate.

Between 32 to 45 days after infection, the 18 tumors were collected separately and their length and weight were measured. Then, tumors were crushed in a mortar with 3 mL of 0.8% NaCl. 3 samples of 100 µL of the bacterial suspension were taken and two serial dilutions were made in 900 µL of 0.8% NaCl. 200 µL of each dilution were plated on TY supplemented with gentamycin, rifampicin and cycloheximide. 48 hours later, CFU number was counted to determine size population in plant tumor. Following we collected all the bacteria population in NaCl 0.8% and divided the sample in three different parts: one that was used for the subsequent inoculation on a new 4 weeks old tomato plant, another one for DNA extraction and the last one to store in glycerol 50% at -80°C. Finally, at the passages 7 and 18, 30 clones of each line were randomly picked from plates, separately inoculated for growth in 150 µL of TY Gm in microwell plate and stored in glycerol 50% at -80°C for cell conservation and pool sequencing.

2.4. Plasmidome

The plasmidome study was performed also on tomato plants to observe whether the dissemination of the Ti plasmid depended on possible variations found in the genome. Samples recovered from passage number 7 (L1 to 18 lines) were used as donor strains. As recipient strain for the Ti plasmid, we employed a strain deprived of the Ti plasmid and carrying a Km cassette on the circular chromosome (*atu1411=braE*) (Haudecoeur, Planamente, *et al.*, 2009). The strains were mixed at a ratio of 1: 1.5 for the donor strain and the recipient strain, respectively. The same infection procedure that for the evolution experiment was performed. 32 days after inoculation, the tumors were collected independently and measured their length and weight. To recover bacterial populations from tumors, they were crushed in a mortar with 3 mL of 0.8% NaCl. For C58pTi-*traI*::Gm and C58RpTi-*atu6148*::Gm lines, we plated all the bacterial suspension in order to recover as many transconjugants as possible. For C58RpTi-*accR*::Gm (hyperconjugative strain), 3 samples of 100 μ L of the bacterial suspension were taken and two serial dilutions were made in 900 μ L of 0.8% NaCl. 200 μ L of each dilution were plated. All samples were plated on selective medium TY supplemented with gentamycin, kanamycin and cycloheximide. 72 hours after, we counted the CFU number to determine the number of transconjugants. Finally, 32 clones were randomly chosen, cultivated in 150 μ L of TYGm of microwell plates, and stored in glycerol 50% at -80°C for cell conservation and pool sequencing.

2.5. DNA extraction and sequencing

DNA was extracted using the MasterPure™ Complete DNA Purification Kit according to the supplier's instructions. All genomic sequences were produced using Illumina technology. Paired-end libraries were prepared following the protocol recommended by Illumina Inc. (<http://www.illumina.com>). Sequence mapping and coverage calculations were performed using CLC Genomic Workbench 8 software (CLC bio, Aarhus, Denmark) with the following parameters for the mapping: 80% of identity on 50% of read length. I polymorphisms were checked by PCR amplification and Sanger sequencing.

2.6. High-resolution melting analysis

The HRM analysis was used for distinguishing single nucleotide variation between an ancestor and its evolved variant. HMR DNA-amplification was performed on individual clones using in a final volume of 15 μ l. The HRM reaction mix contained 5 ng of sample DNA, 0.4 μ M of each primer, 2.5 mM MgCl₂ and the LightCycler® 480 High Resolution Melting Dye from Roche. PCR conditions were as follows: 10 min activation at 95°C, 45 cycles 95°C for 10 s, 60°C for 15 s and 72°C for 15 s and a final step at 95°C for 1 min and cooling to 40°C for 1 min. HRM was performed from 65°C to 97°C, rising at 2.2°C/s with 15 acquisitions per degree. The analyses of the data were performed using the LightCycler® 96 System. Primer list found in Table S6, Annexes.

2.7. Fitness assay on tomato plants

Independent TY overnight cultures of both ancestor and evolved variants were adjusted to an optical cell density of 1 and mixed at ratio 1:1. Serial dilutions of the bacterial suspension were plated onto TYRifGm medium and incubated at 28°C to determine the effective cell number. A total of 32 colonies were isolated from each inoculum for HRM identification. Eight plants were wounded with and scalpel between the first and the second stem nodes and then inoculated by the mixture of the ancestor and evolved variants. Tomato plants were grown in greenhouse under long day conditions and controlled temperature (24–26 °C). 32 days post-infection, tumors were crushed into NaCl 0.8% and plated in TYGmRif (supplemented with cycloheximide) plates to enumerate and recover 20 clones of each tumor. Genotypes were identified using HRM analysis. Experiments were performed in duplicate, hence a total of 16 plants for each competition.

2.8. Statistical analysis of fitness assay

For each evolved variants, two independent competitive assays were carried out, each one with 8 plants. Competitive index (CI) values were calculated as previously described Macho *et al.*, (2010) and then transformed to a square root so the values follow a normal distribution (Shapiro-Wilk normality test, P value > 0.05). A Wilkason test was used to analyze the CI values (square root transformed) from both experiments (the null hypothesis postulates that the samples come from populations with the same probability distribution). If no difference was detected, the values were pooled and the CI shown are the average of the all replicates from both independent experiment. Each CI (square root transformed) was analyzed using Student's t test and the null hypothesis: mean CI was not significantly different from 1 (P value < 0.05 was used).

2.9. Swimming

Swimming phenotypes were tested on AB medium agar plates (100-mm) containing 0.3% agar and supplemented with mannitol, glucose, fructose and sucrose at 10mM as carbon source, and NH₄Cl at 20 mM as nitrogen source.

2.10. Cell size measurement

Overnight cultures of ancestor C58RpTi-*atu6148*::Gm and its evolved variants L18 clones were grown in TY medium and incubated for 10 min in live/dead staining solution (5 mM SYTO9 and 30 mM PI in 50 mM Tris pH 7.0 buffer; Live/Dead BacLight, Invitrogen). Images were acquired at 1024x1024 pixels resolution with a Leica confocal laser scanning microscope TCS SP2. Automated measurements of microbial cells were performed by using OUFTI software (Paintdakhi *et al.*, 2016).

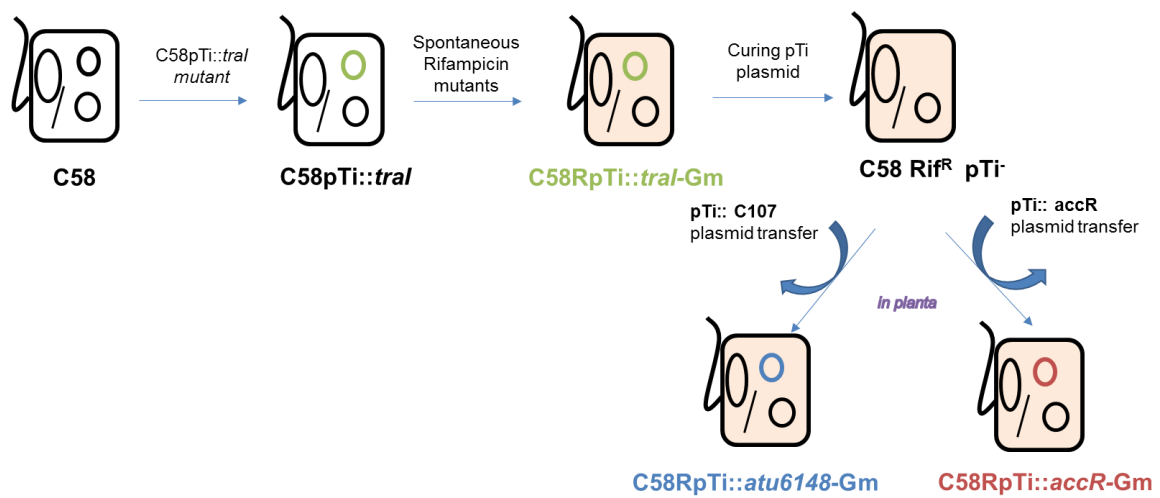


Figure 1. Scheme of the construction of the strains used for this study. In color are the three strains subject to the experimental evolution: C58RpTi::tral-Gm (in green), C58RpTi::accR-Gm (in red) and C58RpTi::atu6148-Gm (in blue).

3. RESULTS

3.1. Experimental evolution

Three ancestors *A. tumefaciens* C58RpTi-*atu6148*::Gm, C58RpTi-*tral*::Gm, and C58RpTi-*accR*::Gm were used for starting the evolution experiment in tomato plant tumor. Except the Ti plasmid, they shared the same genetic background (Figure 1). *A. tumefaciens* C58RpTi-*atu6148*::Gm (pTi-Gm) mutant carries a gentamicin cassette in a non-coding region of Ti plasmid and it has the same phenotypic characteristics of the wild type strain *A. tumefaciens* C58. *A. tumefaciens* pTi-*tral*::Gm mutant is defective for the synthesis of quorum-sensing signals and is not capable to transfer its Ti plasmid (Fuqua and Winans, 1994; Hwang *et al.*, 1994; Piper *et al.*, 1993). *A. tumefaciens* C58RpTi-*accR*::Gm (pTi-*accR*) mutant had enhanced the production of the quorum-sensing signals and the Ti plasmid transfer (Beck von Bodman *et al.*, 1992; Piper and Farrand, 2000).

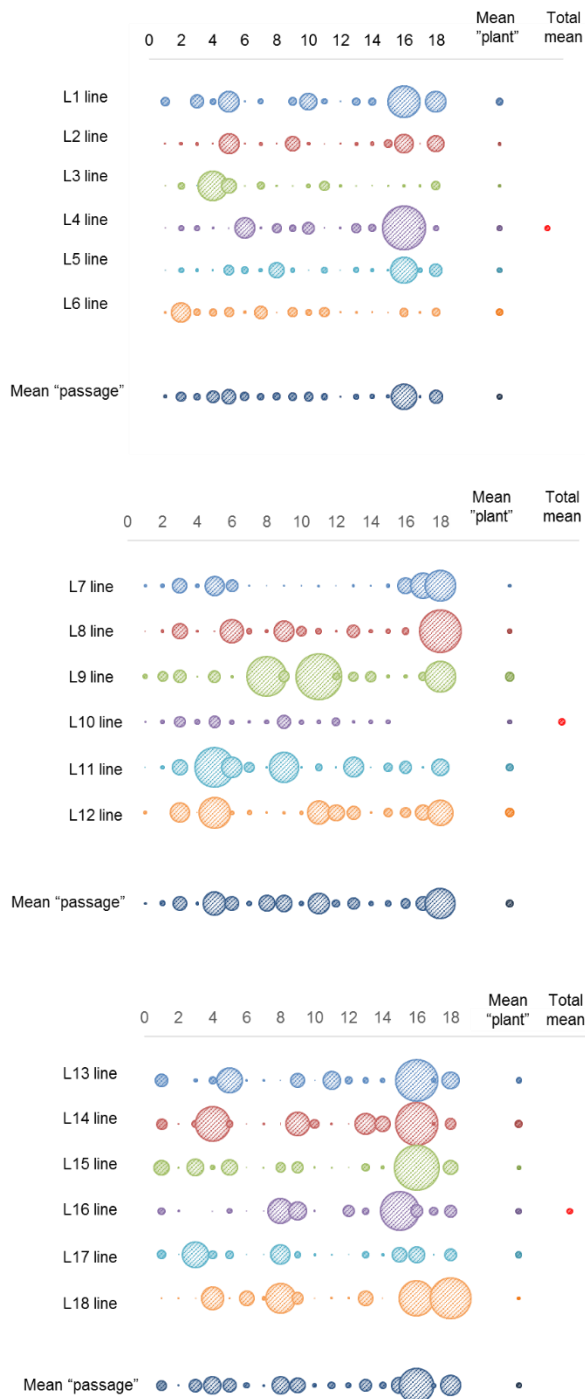
Each ancestor generated six lineages of evolved populations: L1 to L6 from the pTi-*tral* ancestor, L7 to L12 from pTi-*accR* ancestor and L13 to L18 from pTi-Gm ancestor. A total of 18 consecutive inoculations or passages were performed during the 2-year study (723 days exactly).

In the course of the experiment, no difference in tumor weight was observed. The mean tumor weight for all lines was 300 mg for pTi-*accR*, 305 mg for pTi-*tral* and 316 mg for pTi-Gm (Figure 2b).

Concerning the bacterial charge, differences became more remarkable. The average of the bacterial charge of all lines throughout the experiment reflected that for both pTi-*tral* ($1,68E+05$ CFU/tumor) and pTi-Gm ($1,46E+05$ CFU/tumor) is almost twice as high as pTi-*accR* ($8,50E+04$ CFU/tumor). These data indicate that the number of bacteria that colonize the tumor is very variable (Figure 2a) during time and lines, and that there is no direct relationship between the number of bacteria found inside the tumor and the symptoms they cause.

In the course of the experiment, the extinction of two lines happened. The first case, in passage 2, L1 pTi-*tral* could not be recovered due to the amount of contamination found in TYGmRif plates (different bacteria and fungi). Therefore, L2 was used as inoculum for two lines, L2 itself and a new L1, during passage 3. For the second case, no colonies of

A



B

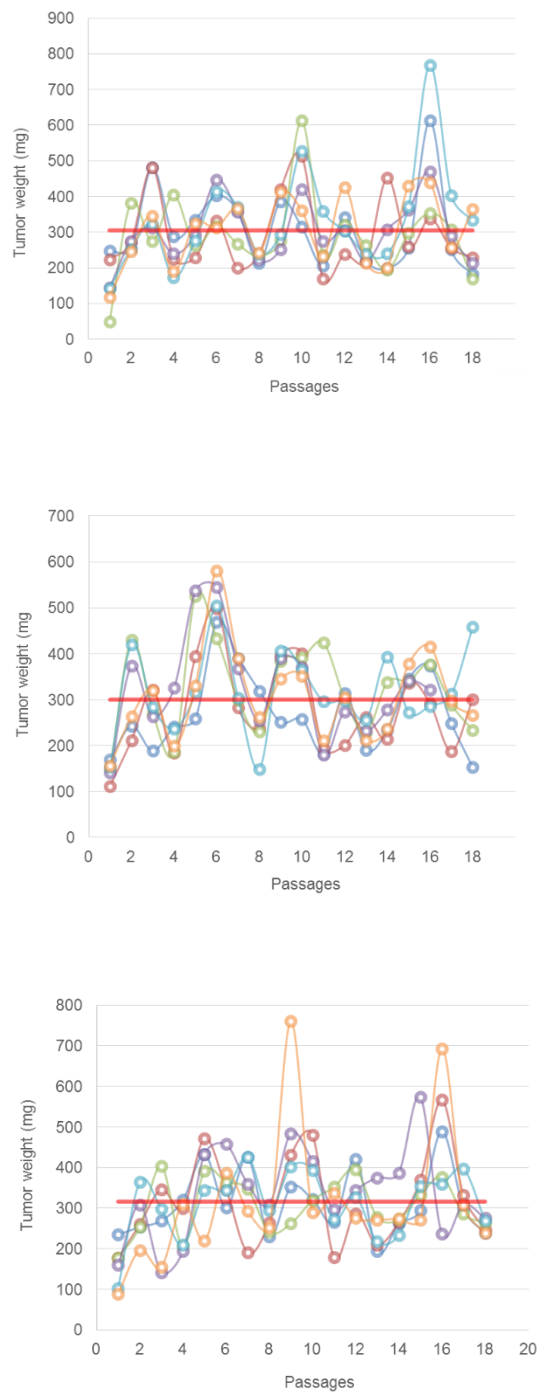


Figure 2. Experimental evolution assay.

- A. Bacterial charge (CFU/tumor) of each different line in each different passage. Also the average of each passage is represented ("mean passage"), as well as total average of each line ("mean" plant). Finally, the average of the whole experiment for each strain was represented (total mean).
- B. Tumor weight (mg/tumor) of each line for each different passage. The red line shows the total average of the experiment for each strain.

L10 pTi-*accR* were recovered in passage 16, hence we decided to finish L10 experiment whereas we continued with the other 5 lines pTi-*accR* until passage 18.

In the meanwhile, we observed a morphotype variation of the bacteria in only case, pTi-Gm line 18. All L18 colonies presented, in passage 5, a smaller size and different color/texture compared with the ones of the other lines of the same ancestor (Figure 3). After confirming by PCR that clones belonged to pTi-Gm line, and they were not a contamination of other bacteria, this characteristic was presented in all the following passages of the L18 lineage.

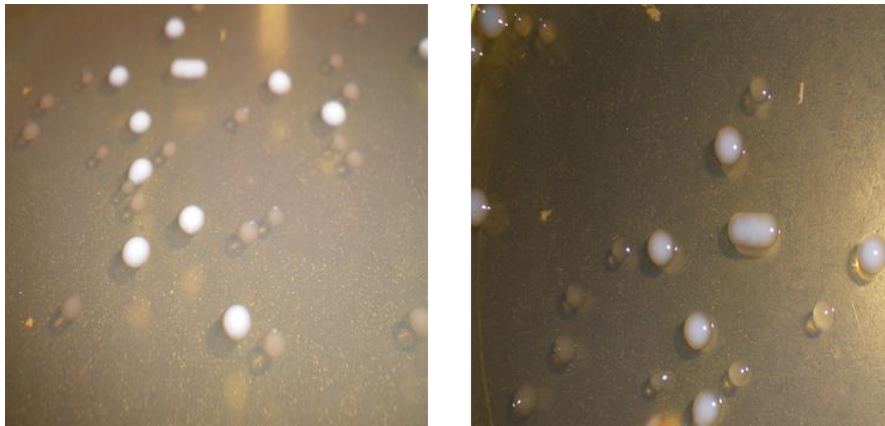


Figure 3. Phenotypic differences between C58RpTi::*atu6148*-Gm: large white colonies belong to the ancestor and the smaller and translucent colonies belong to the line L18.

3.2. Analysis of an early stage of the experimental evolution assay

3.2.1. Pool sequencing of early stage passage

To determine the genetic variation that occurred during the experiment, we sequenced the three ancestor clones and a pool of 30 random clones of each independent line. Sequenced data were mapped to the *A. tumefaciens* C58 NCBI genome and analyzed as described in Material and methods. Two different sequencing points were performed, one in an early passage (the 7th) and another one in the last passage of the experiment (the 18th).

The first sequencing analysis was carried out in passage 7, nearly one year after the first inoculation. Most of the lines did not present any variation respecting to their ancestor (Table 1). A total of 55 different variations were present in 7 different lines: L1, L2, L8, L9, L14, L16 and L18. Of all of them, 38 corresponded to non-conservative variations, it means, a change to an amino acid with similar physiochemical properties or a vastly different amino acid. Generally, each line contains between one to three different variations, except for the L18 line, which contains 45 different SNP/InDels in 38 different genes. Furthermore, no variations were found in both plasmids. Genes appeared to experience genetic changes belonged to different classes, but most of them were involved in amino acid transport and metabolism, transcription or proteins of unknown function.

Concerning the pTi-*traI* lines, only L1 and L2 lines presented variations. L1 exhibited only one SNP in *atu3932* gene, a hypothetical protein. We could observe that this variation entails a stop codon after the glutamic acid in position 56 of 266 amino acids that define this protein. This variation gives rise to a truncated protein and it was presented in 45% of the 30 isolated clones that conformed the pool, i.e. in at least 14 of those clones. In the case of L2, two different variations were founded if we compared the evolved clones to the pTi-*traI* ancestor. These two non-conservatives SNPs are in *mcpC* gene (*atu0872*), a methyl-accepting chemotaxis protein and in a RNA polymerase sigma factor (*atu4160*).

For pTi-*accR* lines, only clones coming from L8, L9 and L10 lines presented variations in their genome. L8 line showed a single SNP in *atu0499* gene, a hypothetical protein. As reported by (Tomlinson and Fuqua, 2009), this gene codes for a polar development protein J, PodJ. This protein is a membrane protein and is a crucial factor for polar growth. It has been demonstrated that a deletion of this gene give rises to ectopic polar growth (Anderson-Furgeson *et al.*, 2016). In L9 line, we found 2 different SNPs and an

Table 1. Genetic variants appeared in passage t7 for each different line (table continues on next page).

Ancestor	Line	Pool size	Replicon	Position	Type	Freq (%)	Allele	Amino acid change	atu code	Gene	Product	COG Classification	
pTi-tral	L1	30	linear	1035985	SNV	45	C → A	Glu56*	atu3932		Hypothetical Protein	Amino acid transport and metabolism	
	L2	30	circular	868634	SNV	36	T → C	Phe156Ser	atu0872	mcpC	Methyl-Accepting Chemotaxis Protein	Cell motility	
linear			1279040	SNV	36	G → A	Ala156Val	atu4160			RNA Polymerase Sigma Factor	Transcription	
pTi-accR	L8	30	circular	486970	SNV	64	G → T	Thr884Asn	atu0499		Hypothetical Protein	Cell motility	
	L9	2	circular	1166756	SNV	100	G → T	Leu312Phe	atu1173		MFS Permease	Amino acid transport and metabolism	
			circular	1275277	Insertion	80	- → C	Arg47fs	atu1287			Hypothetical Protein	Unclassified
			linear	1279040	SNV	35	G → A	Ala156Val	atu4160			RNA Polymerase Sigma Factor	Transcription
L10	30	circular	1375173	SNV	39	A → G		atu1381	omp1	Putative Outer Membrane Protein Assembly Factor	Cell wall/membrane/envelope biogenesis		
		linear	1279040	SNV	36	G → A	Ala156Val	atu4160			RNA Polymerase Sigma Factor	Transcription	
pTi-Gm	L14	30	linear	523952	SNV	41	A → G				Between Atu3475 And Atu3476	Unclassified	
	L16	30	circular	1915883	SNV	100	A → G				Between Atu1947 And Atu1948	Unclassified	
			circular	1934926	SNV	100	T → C	Gln12Arg	atu1962	secE	Preprotein Translocase Subunit	Intracellular trafficking, secretion, and vesicular transport	
	L18	30	circular	502200	SNV	95	T → A	Ile315Phe	atu8131	phaA	Monovalent Cation/H+ Antiporter Subunit A	Energy production and conversion	
			circular	826605	SNV	95	G → A				Between Atu0827 And Atu0828	Unclassified	
			circular	1021837	SNV	96	C → T				Between Atu1028 / Atu1029	Unclassified	
			circular	1023751	SNV	94	G → A	Arg364His	atu1030		GTP Pyrophosphohydrolase/Synthetase	Transcription	
			circular	1368399	SNV	97	A → G	Lys132Arg	atu1374	rpsB	30S Ribosomal Protein S2	Translation, ribosomal structure and biogenesis	
			circular	1381967	SNV	95	G → T	Arg161Leu	atu1387		Gntr Family Transcriptional Regulator	Transcription	
			circular	1562277	SNV	96	G → A				Between Atu1575 / Atu1577	Unclassified	
			circular	1562966	SNV	94	A → G	Thr167Ala	atu1577		ABC Transporter, Substrate Binding Protein (Amino Acid)	Amino acid transport and metabolism	
			circular	1563380	Deletion	88	CTG → -	Leu305del	atu1577		ABC Transporter, Substrate Binding Protein (Amino Acid)	Amino acid transport and metabolism	
			circular	1633942	SNV	96	T → C				Between Atu1649 / Atu1650	Unclassified	
			circular	1873736	SNV	93	G → A		atu1897		Short Chain Dehydrogenase	Lipid transport and metabolism	
			circular	1907071	SNV	95	A → T	Val35Asp	atu1929	rpsE	30S Ribosomal Protein S5	Translation, ribosomal structure and biogenesis	
			circular	1920197	SNV	95	T → C	Lys43Arg	atu1951	rpsL	30S Ribosomal Protein S12	Translation, ribosomal structure and biogenesis	
			circular	1929719	SNV	95	C → T	Arg539His	atu1956	rpoB	DNA-Directed RNA Polymerase Subunit Beta	Transcription	
			circular	2174067	SNV	95	T → C	Val49Ala	atu2200	cspA	Cold Shock Protein	Transcription	
circular	2425790	SNV	93	A → T	Ser185Thr	atu2454		iron-chelator utilization protein	Inorganic ion transport and metabolism				
circular	2615103	SNV	95	C → A	Ala117Ser	atu2631	cybB	cytochrome b561	Energy production and conversion				
linear	4480	Deletion	92	A → -	Phe111fs	atu3006		conserved hypothetical protein, membrane protein	Function unknown				
linear	46862	SNV	36	G → A	Thr100Ile	atu3049		putative ABC transporter, substrate binding protein (oligopeptide), TaT signal	Amino acid transport and metabolism				
linear	153903	Deletion	97	T → -	Ile73fs	atu3146		Putative Xaa-Pro Aminopeptidase	Amino acid transport and metabolism				
linear	155876	SNV	96	G → A	Ser161Leu	atu3147		Oxidoreductase	General function prediction only				
linear	530170	SNV	95	T → C		atu3480	hmgL	Hydroxymethylglutaryl-Coa Lyase	Amino acid transport and metabolism				

Table 1. Genetic variants appeared in passage t7 for each different line (continuation).

Ancestor	Line	Pool size	Replicon	Position	Type	Freq (%)	Allele	Amino acid change	atu code	Gene	Product	COG Classification		
pTi-Gm	L18	30	linear	552125	SNV	96	T → C	Asp237Gly	<i>atu3501</i>	<i>cysA2</i>	Sulfate/Thiosulfate Import ATP-Binding Protein Cysa 2	Inorganic ion transport and metabolism		
			linear	676317	SNV	97	A → C	Val40Gly	<i>atu3621</i>	<i>cobT</i>	Cobyrinic Acid Synthase	Coenzyme transport and metabolism		
			linear	685673	SNV	95	T → C	Ile83Val	<i>atu3633</i>			Hypothetical Protein	General function prediction only	
			linear	721932	SNV	95	G → A	Ala207Val	<i>atu3671</i>			Putative Luciferase-Like Monooxygenase (PP-Binding)	Energy production and conversion	
			linear	735472	SNV	96	T → C		<i>atu3677</i>			Putative Glutamate-1-Semialdehyde 2,1-Aminomutase	Coenzyme transport and metabolism	
			linear	750532	SNV	96	T → C	Val272Ala	<i>atu3683</i>			Putative Non-Ribosomal Peptide Synthetase (NRPS)	Secondary metabolites biosynthesis, transport and catabolism	
			linear	1090505	SNV	78	C → T		<i>atu3977</i>	<i>ina</i>			Ice Nucleation-Like Protein	Intracellular trafficking, secretion, and vesicular transport
			linear	1090511	SNV	81	A → G		<i>atu3977</i>	<i>ina</i>			Ice Nucleation-Like Protein	Intracellular trafficking, secretion, and vesicular transport
			linear	1090520	SNV	83	A → G		<i>atu3977</i>	<i>ina</i>			Ice Nucleation-Like Protein	Intracellular trafficking, secretion, and vesicular transport
			linear	1090523	SNV	84	G → A		<i>atu3977</i>	<i>ina</i>			Ice Nucleation-Like Protein	Intracellular trafficking, secretion, and vesicular transport
			linear	1090526	SNV	84	C → G		<i>atu3977</i>	<i>ina</i>			Ice Nucleation-Like Protein	Intracellular trafficking, secretion, and vesicular transport
			linear	1090529	MNV	84	CGC → GAT	Ala337Ile	<i>atu3977</i>	<i>ina</i>			Ice Nucleation-Like Protein	Intracellular trafficking, secretion, and vesicular transport
			linear	1090536	SNV	89	G → T	Thr335Asn	<i>atu3977</i>	<i>ina</i>			Ice Nucleation-Like Protein	Intracellular trafficking, secretion, and vesicular transport
			linear	1149844	SNV	96	G → A	Ala223Thr	<i>atu4034</i>				Two Component Sensor Kinase	Signal transduction mechanisms
			linear	1199120	SNV	97	T → C	Asn400Ser	<i>atu4078</i>	<i>glgP</i>			Glycogen Phosphorylase	Carbohydrate transport and metabolism
			linear	1231723	SNV	95	C → A		<i>atu4112</i>				Proline Dipeptidase	Amino acid transport and metabolism
			linear	1326408	SNV	94	A → G	Val101Ala	<i>atu4204</i>				Putative DNA-Binding Transcriptional Regulator; Prophage	Transcription
			linear	1358427	SNV	96	T → C	Ile6Thr	<i>atu4233</i>				ABC Transporter, Substrate Binding Protein (Amino Acid)	Amino acid transport and metabolism
			linear	1387019	SNV	96	T → C	Thr129Ala	<i>atu4260</i>				ABC Transporter, Membrane Spanning Protein	Inorganic ion transport and metabolism
			linear	1661950	SNV	96	A → G		<i>atu4514</i>				Rpir Family Transcriptional Regulator	Transcription
			linear	1829950	Insertion	58	- → G	Leu342fs	<i>atu4668</i>				Putative ABC Transporter, Membrane Spanning Protein	Defense mechanisms
			linear	1953053	SNV	96	T → C	Ile18Val	<i>atu4781</i>				Putative Two Component Sensor Kinase	Signal transduction mechanisms
linear	1953541	SNV	95	C → T	Asp103Asn	<i>atu4782</i>				Putative Two Component Response Regulator	Transcription			

insertion. The gene *atu4061* was again affected in L9 by the same variation as in L2 line, and it was again presented in L10 line. This variation didn't appear to be completely fixed in the samples since it appeared at a frequency of approximately 35%. The other SNP of L9 line was found in *atu1173* gene, an MFS permease. In addition, the insertion that appears in the *atu1287* gene, a hypothetical protein, resulted in a mutation with a displacement of the reading frame (frame shift). Finally, the aforementioned variation of L10 line was presented together with the SNP in the putative outer membrane protein assembly factor Omp1 (*atu1381*), which was not affected by the variation for being this one conservative.

It was in pTi-Gm lines where we found most of the total variations. In L14 line, we found a conservative variation in an intergenic region. In L16 line, we could observe 2 different variation, a SNP in another intergenic area and also in *atu1962* gene, *secE*. SecE encodes a preprotein translocase subunit of the general secretion systems Sec. This variation seemed to be fixed in all clones from the sample. The most particular case is L18 line. As it was described above L18 colonies were smaller and had a different color/texture than its ancestor or the other lines. It contained 45 different SNPs/InDels in 38 different genes. Seven different variations were found in *ina* (*atu3977*) gene, but only two of them were non-conservative. Also in *atu1577* gene, an ABC transporter, we observed 1 SNP and a deletion, both conferring an amino acid modification. This gene appears to be overexpressed under acidic pH conditions (Yuan *et al.*, 2008), as is the case of the plant apoplast. In the remaining affected genes, only 1 single variation was found. These genes were involved in different metabolic pathways as the respiratory chain as *rpsB* (*atu1374*), *rpsE* (*atu1929*), *rpsL* (*atu1951*) and *rpoB* (*atu1956*), iron uptake as *atu2454* and *atu3671* and both upregulated under iron limitation (Heindl *et al.*, 2016), ABC transporters, transcriptional factors, etc.

3.2.2. Pool plasmidome sequencing of early stage

To determine if possible changes in the genome of *A. tumefaciens* could confer an improvement of the transfer of the Ti plasmid, an experiment of “capture” of this plasmid was carried out. For this purpose, an *A. tumefaciens* C58 strain lacking the Ti plasmid and with a different selection cassette on the circular chromosome was used as recipient strain of the Ti plasmid and samples from each line of Evolution experiment as donor strains.

As expected, once the samples were plated, no transconjugants were found from the pTi-*traI* plants. This strain is not capable of producing QS signals that are necessary for the transfer of the Ti plasmid. In the case of pTi-Gm plants, we found transconjugants in only 4 plants and we also could recover transconjugants from 5 plants.

The Illumina results are analogous to the data found in step 7 of the Evolution assay. Since no changes were found in the Ti plasmids of any of the lines, no difference was found in the transconjugants plasmidome.

3.2.3. Single clone sequencing

To validate the occurrence of some of these variations, some isolated clones of each line were checked by PCR amplification and followed by Sanger sequencing (primer list in Table S5, Annexes). We could detect some clones of each line that contained the variations, except for the gene of *atu3939* of L9 line (pTi-*accR*) that could not be verified by Sanger sequencing. This could be possible due to the existence of 4 identical copies of this gene in *A. tumefaciens* genome.

In addition, one isolated clones of L18 line (pTi-Gm) was re-sequenced in its entirety by the Illumina method to find out how many SNPs found in the pool sample appeared in it. Results showed that most of the variations of the pool sample were also fixed in the isolated clone (38 SNPs of 45 SNPs of line18). We could observe also some that were not present in it and some new that we could not detected in the pool sample. Perhaps the low incidence of these new variations in the pool sample was the reason why it could not be detected in the first sequencing.

Finally, one validated clone of each line was used for fitness assays.

3.2.4. Fitness on tomato plants

To evaluate changes in fitness ability, a competition experiment was carried out on tomato plants. Evolved clones of the early passage were co-inoculated with their respective ancestor clone. The evolved clones chosen for the experiment were those with a non-conservative amino acid change. Hence, the clones chosen were L1, L2, L8, L9, L14 and L18 lines (Table 2). Each pair of ancestor/evolved clones was co-inoculated to 4 weeks old tomato plants at a ratio of 1:1 and the fitness of each evolved clone was measured by counting the number of evolved and ancestral clones before and after the infection using qPCR High-resolution melting analysis. This technique allows to identify variations in nucleic acid sequences. The method is based on detecting small differences in PCR melting (dissociation) curves.

Line	Clone	Validation method	SNPs validated
L1	L1.2	Sanger	<i>atu3932</i>
L2	L2.12	Sanger	<i>atu0872</i>
L8	L8.2	Sanger	<i>atu0499</i>
L9	L9.2	Sanger	<i>atu1173; atu1287</i>
L14	L14.2	Sanger	<i>atu3476 and Between atu3475</i>
L16	L16.1	Sanger	<i>atu1962 and Between atu1947 and atu1948</i>
L18	L18.1	Illumina	39 SNPs (38 found in L18 line)

Table 2. List of clones used for fitness assays on tomato plants. The data show the lines where they come from, the method used for validating the SNPs and the SNPs validated.

As showed in Figure 4, three of the evolved clones tested were found to improve in fitness respecting their ancestral strain. CI average values of L1.2, L2.12 and L18.1 clones differed significantly from 1 (Student's t test, P value < 0.05). For L8 and L14 clones, CI was not significantly difference from 1. This indicated that their fitness in tomato plants did not differ from their ancestor strains. In the case of L9.2 clone, the result from Wilkoston test showed that both independent experiments had a different probability distribution. In this case, no pool was done for the CI values and we represented the mean values of both assays separately.

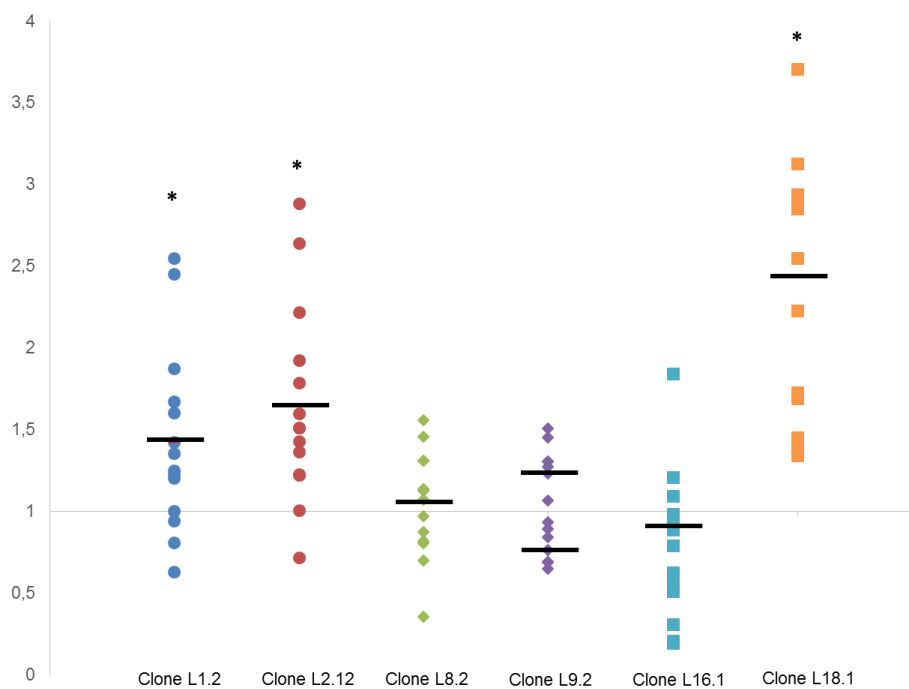


Figure 4. Competition assays of evolved clones from t7 passage and their respective ancestor strains. Asterisks indicate that indexes are significantly different, as established using Student's t-test (P < 0.05).

3.2.5. *mcpC* mutation could improve bacterial chemotaxis in plant tumors

As previously reported, L2 evolved clone is able to better colonize the tumor as the ancestor strain. To test whether this new trait was due to a chemotaxis capacity different from that of the ancestor, we performed a swimming assay on AB medium agar plates supplemented with different carbon sources. As we can observe in Figure 5, evolved clone formed significantly bigger swim rings on 0.3% swim agar when we used mannitol, glucose or fructose as carbon source. When sucrose was present on the medium, the ancestor had a better ability to swim compared with the evolved clone.

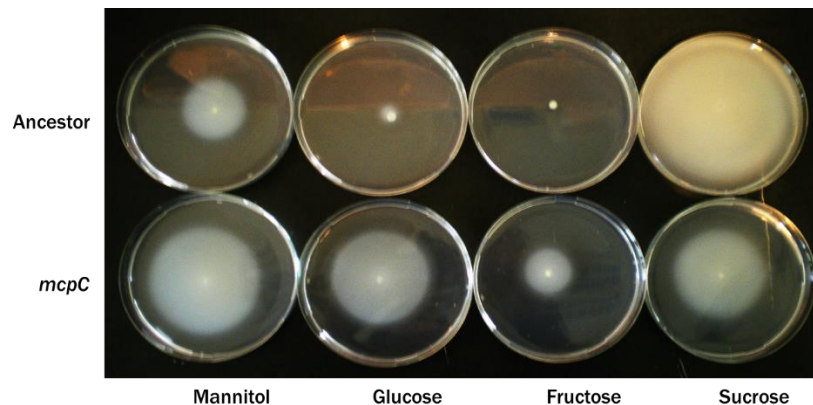


Figure 5. Swimming assays of L2.12 evolved clone and its ancestor in AB 0.3% agar plates supplemented with different sugars.

3.2.6. Growth rate is not affected in L18 evolved clone

Colonies from L18 line had a different morphotype respecting the ancestor strain, as mentioned above. To analyze if colony size was due to a slower rate growth than the ancestor, we grew both clones on two different mediums. No growth differences were found during the course of the experiment between the ancestor strain and the L18 clone (Figure 6), although, as expected, we did find differences on growth when the culture media were varied. Therefore, differences in colonies size did not seem to be related to a lower growth rate of the evolved clone.

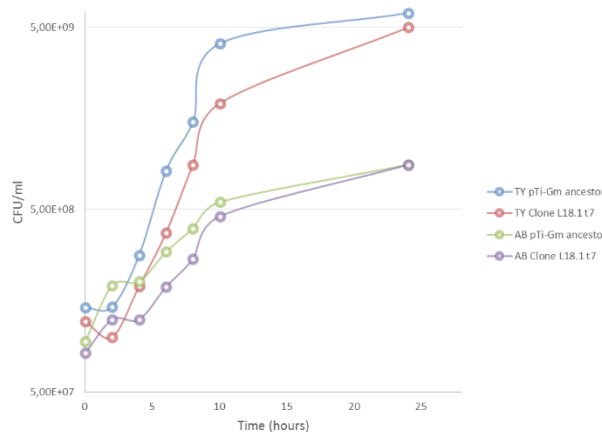


Figure 6. Growth curves of clone L18.1 and its ancestor on TY and AB sucrose medium.

3.2.7. Bacteria cell size of L18 clone differed from pTi-Gm ancestor strain

As difference in colony size of L18 clone was not due to a decrease in the growth rate, we wondered if L18 clone exhibited a smaller size of the bacterial cells. A microscopy experiment was performed and results are shown in Figure 7. As we could verify the size of both cells is different. L18 line cells are slightly smaller than those of the ancestral clone and this variation is statistically significant.

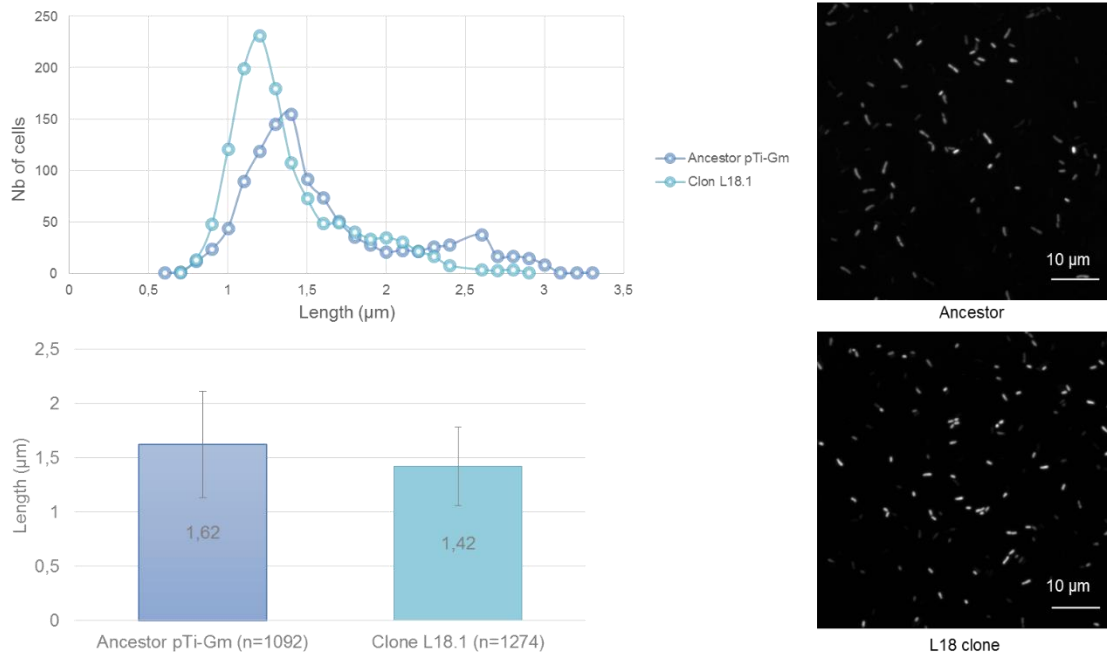


Figure 7. Microscopy assay of clone L18.1 and its ancestor for size measurements.

3.3. Analysis of the last stage of the experimental evolution assay

Table 3 summarizes all variants founded on the last sequencing point of the experiment, the one corresponding with the last passage.

This last sequencing point was performed in passage 18, two years after the beginning of the experiment. All lines presented changes in their genome respecting to their ancestor, excepting L10 line (pTi-*accR*) that was already extinguished. 74 different variations were found within all lines, 49 corresponding to non-conservative variations. As we observed in passage 7, L18 line contained the highest amount on variation in its genome. This time, we could find changes also in the At plasmid, but none in Ti plasmid. Most of the affected genes belong to the amino acid transport and metabolism and transcription categories.

Respecting pTi-*traI* lines, L1 conserved the variation on *atu3932* gene fixed in the whole population. Besides, this change was present on the L3 line. L2 line only maintained the change on a RNA polymerase sigma factor (*atu4160*), and was also found on the rest of pTi-*traI* lines, indicating the appearance of parallel evolution. The variation on *chvE* (*atu2348*), a multiple sugar-binding periplasmic receptor resulted in a stop codon after a proline in position 235 of 354 amino acids that contains this protein. Other non-conservative changes appeared in *bexD* (*atu4329*) a polysaccharide export protein, *rpoB* (*atu1956*) a DNA-directed RNA polymerase subunit beta and *dgcA* (*atu1257*) a diguanylate cyclase, control of the motile-to-sessile switch and involved in biofilm formation (Heindl *et al.*, 2014).

For pTi-*accR* lines, the variations appeared on passage 7 were conserved except the insertion that appeared in the *atu1287* gene. All changes were non-conservatives, and we also found parallel variations on a RNA polymerase sigma factor (*atu4160*), the same variation that in pTi-*traI* lines.

As we have seen on passage 7, on pTi-Gm lines we still found most of the total variations. The only change conserved from passage 7 was on *secE* (*atu1962*) of L14 line, which was already well fixed. The RNA polymerase sigma factor (*atu4160*) found on the rest of lines was also in all pTi-Gm lines except of L18. This last one, L18 line, conserved most of the variations appeared in passage 7. We observed 31 different SNPs/InDels in 30 different genes. It conserved 27 different variations, and we could found 3 new changes. The first one was on a LysR family transcriptional regulator (*atu2559*). The Atu3205 conserved protein presented also an SNP. This protein was found on a

proteome study of *A. tumefaciens* exposed to plant roots (Rosen *et al.*, 2003). Finally, the last new variation was on *atu4163* gene, a two component sensor kinase, also present on gaba⁺ spontaneous mutants as shown by (Planamente *et al.*, 2012).

Table 2. Genetic variants appeared in passage t18 for each different line. Blue color indicates that the variation was present in passage 7 (table continues on next page).

Ancestor	Line	Pool size	Replicon	Position	Type	Freq (%)	Allele	Amino acid change	atu code	Gene	Product	COG classification
pTi- <i>tral</i>	L1	30	circular	399872	SNV	65	C → A		<i>atu0404</i>	<i>pgi</i>	Glucose-6-phosphate isomerase	Carbohydrate transport and metabolism
			circular	2326018	Deletion	80	C → -	Pro235fs	<i>atu2348</i>	<i>chvE</i>	Multiple sugar-binding periplasmic receptor	Carbohydrate transport and metabolism
			circular	2489145	SNV	68	G → A		<i>atu2519</i>		Dipeptidase	Amino acid transport and metabolism
			linear	1035985	SNV	100	C → A	Glu56*	<i>atu3932</i>		Hypothetical protein	Amino acid transport and metabolism
			linear	1279040	SNV	90	G → A	Ala156Val	<i>atu4160</i>		RNA polymerase sigma factor	Transcription
	L2	30	linear	1279040	SNV	93	G → A	Ala156Val	<i>atu4160</i>		RNA polymerase sigma factor	Transcription
	L3	30	circular	860664	SNV	40	C → G		<i>atu0863</i>		Hypothetical protein	Transcription
			circular	2326018	Deletion	99	C → -	Pro235fs	<i>atu2348</i>	<i>chvE</i>	Multiple sugar-binding periplasmic receptor	Carbohydrate transport and metabolism
			linear	1035985	SNV	100	C → A	Glu56*	<i>atu3932</i>		Hypothetical protein	Amino acid transport and metabolism
			linear	1279040	SNV	89	G → A	Ala156Val	<i>atu4160</i>		RNA polymerase sigma factor	Transcription
			pAt	384553	SNV	99	A → T				Between <i>atu5390</i> and <i>atu5391</i>	Unclassified
	L4	30	circular	1872057	SNV	41	A → G				Between <i>atu1893</i> and <i>atu1894</i>	Unclassified
			circular	1928217	SNV	99	A → G	Tyr1040His	<i>atu1956</i>	<i>rpoB</i>	DNA-directed RNA polymerase subunit beta	Transcription
			linear	1279040	SNV	92	G → A	Ala156Val	<i>atu4160</i>		RNA polymerase sigma factor	Transcription
			linear	1457896	SNV	99	G → T	Asp148Tyr	<i>atu4329</i>	<i>bexD</i>	Polysaccharide export protein	Cell wall/membrane/envelope biogenesis
			pAt	244470	SNV	40	T → G		<i>atu5248</i>	<i>potA</i>	ABC transporter, nucleotide binding/ATPase protein (spermidine/putrescine)	Amino acid transport and metabolism
	L5	30	circular	860664	SNV	36	C → G		<i>atu0863</i>		Hypothetical protein	Transcription
			circular	1245541	SNV	88	C → T	Gly92Ser	<i>atu1257</i>	<i>dgcA</i>	Diguanylate cyclase, control of the motile-to-sessile switch	Signal transduction mechanisms
			linear	1279040	SNV	91	G → A	Ala156Val	<i>atu4160</i>		RNA polymerase sigma factor	Transcription
	L6	30	circular	1181850	SNV	87	C → T		<i>atu1189</i>		Hypothetical protein	Nucleotide transport and metabolism
circular			1872057	SNV	86	A → G				Between <i>atu1893</i> and <i>atu1894</i>	Unclassified	
circular			1928217	SNV	87	A → G	Tyr1040His	<i>atu1956</i>	<i>rpoB</i>	DNA-directed RNA polymerase subunit beta	Transcription	
linear			1279040	SNV	91	G → A	Ala156Val	<i>atu4160</i>		RNA polymerase sigma factor	Transcription	
linear			1457896	SNV	88	G → T	Asp148Tyr	<i>atu4329</i>	<i>bexD</i>	Polysaccharide export protein	Cell wall/membrane/envelope biogenesis	
L7	30	circular	860664	SNV	38	C → G		<i>atu0863</i>		Hypothetical protein	Transcription	
		circular	2559499	SNV	44	T → C	Val48Ala	<i>atu2584</i>		Hypothetical protein	Unclassified	
		linear	1279040	SNV	93	G → A		<i>atu4160</i>		RNA polymerase sigma factor	Transcription	
		pAt	519871	SNV	69	G → A				Between <i>atu5527</i> and <i>atu5528</i>	Unclassified	
L8	30	circular	340266	SNV	68	C → T		<i>atu0346</i>	<i>glnD</i>	[Protein-PilI] uridylyltransferase	Posttranslational modification, protein turnover, chaperones	
		circular	486970	SNV	100	G → T	Thr884Asn	<i>atu0499</i>		Hypothetical protein	Cell motility	
		circular	860664	SNV	38	C → G		<i>atu0863</i>		Hypothetical protein	Transcription	
L9	30	linear	1279040	SNV	1	G → A	Ala156Val	<i>atu4160</i>		RNA polymerase sigma factor	Transcription	
		circular	860664	SNV	38	C → G		<i>atu0863</i>		Hypothetical protein	Transcription	
L11	30	circular	860664	SNV	38	C → G		<i>atu0863</i>		Hypothetical protein	Transcription	
		circular	1166756	SNV	100	G → T	Leu312Phe	<i>atu1173</i>		MFS permease	Amino acid transport and metabolism	
		linear	1279040	SNV	92	G → A	Ala156Val	<i>atu4160</i>		RNA polymerase sigma factor	Transcription	
		circular	860664	SNV	35	C → G		<i>atu0863</i>		Hypothetical protein	Transcription	
		circular	1881485	SNV	37	T → G		<i>atu1903</i>	<i>glpK</i>	Glycerol kinase 2	Energy production and conversion	
L11	30	linear	1279040	SNV	91	G → A	Ala156Val	<i>atu4160</i>		RNA polymerase sigma factor	Transcription	
		linear	1698017	SNV	36	T → G				Between <i>atu4551</i> and <i>atu4552</i>	Unclassified	
		linear	1702982	SNV	39	A → T	Ser347Cys	<i>atu4556</i>	<i>purR</i>	LacI family transcription regulator	Transcription	

Table 2. Genetic variants appeared in passage t18 for each different line. Blue color indicates that the variation was present in passage 7 (table continues on next page).

Ancestor	Line	Pool size	Replicon	Position	Type	Freq (%)	Allele	Amino acid change	atu code	Gene	Product	COG classification
pTi- <i>accR</i>	L12	30	circular	860664	SNV	39	C → G		<i>atu0863</i>		Hypothetical protein	Transcription
			circular	1321214	SNV	45	G → C	Asp333His	<i>atu1334</i>	<i>aspB</i>	Aspartate transaminase	Amino acid transport and metabolism
			linear	1279040	SNV	93	G → A	Ala156Val	<i>atu4160</i>		RNA polymerase sigma factor	Transcription
			pAt	188646	SNV	66	G → A	Arg73His	<i>atu5193</i>		Dipeptidyl aminopeptidase/acylaminoacyl peptidase	Amino acid transport and metabolism
pTi-Gm	L13	30	circular	860664	SNV	38	C → G		<i>atu0863</i>		Hypothetical protein	Transcription
			linear	1279040	SNV	93	G → A	Ala156Val	<i>atu4160</i>		RNA polymerase sigma factor	Transcription
			linear	1448669	SNV	61	G → T		<i>atu4320</i>	<i>rbsB</i>	ABC transporter, substrate binding protein (ribose)	Carbohydrate transport and metabolism
			linear	2029361	SNV	66	T → C				Between <i>atu4848</i> and <i>atu4849</i>	Unclassified
	L14	30	circular	1682130	SNV	99	A → G	Met252Val	<i>atu1696</i>		Hypothetical protein	Amino acid transport and metabolism
			circular	1682873	SNV	99	G → A	Gly176Asp	<i>atu1697</i>		MFS permease	Amino acid transport and metabolism
			linear	1279040	SNV	92	G → A	Ala156Val	<i>atu4160</i>		RNA polymerase sigma factor	Transcription
	L15	30	circular	1034419	SNV	53	G → A		<i>atu1042</i>		Hypothetical protein	General function prediction only
			circular	1911781	SNV	100	T → C	Ile111Val	<i>atu1940</i>	<i>rpsC</i>	30S ribosomal protein S3	Translation, ribosomal structure and biogenesis
			circular	1929013	MNV	98	GA → TC	Val774Gly	<i>atu1956</i>	<i>rpoB</i>	DNA-directed RNA polymerase subunit beta	Transcription
			linear	484966	SNV	78	G → A	Val102Ile	<i>atu3438</i>		Putative oxidoreductase	Energy production and conversion
	L16	30	linear	1279040	SNV	93	G → A	Ala156Val	<i>atu4160</i>		RNA polymerase sigma factor	Transcription
circular			860664	SNV	36	C → G		<i>atu0863</i>		Hypothetical protein	Transcription	
circular			1915883	SNV	100	A → G				Between <i>atu1947</i> and <i>atu1948</i>	Unclassified	
circular			1934926	SNV	100	T → C	Gln12Arg	<i>atu1962</i>	<i>secE</i>	Preprotein translocase subunit	Intracellular trafficking, secretion, and vesicular transport	
linear			1279040	SNV	91	G → A	Ala156Val	<i>atu4160</i>		RNA polymerase sigma factor	Transcription	
L17	30	linear	1279040	SNV	82	G → A	Ala156Val	<i>atu4160</i>		RNA polymerase sigma factor	Transcription	
L18	30	circular	502200	SNV	99	T → A	Ile315Phe	<i>atu8131</i>	<i>phaA</i>	monovalent cation/H⁺ antiporter subunit A	Energy production and conversion	
		circular	826605	SNV	100	G → A				Between <i>atu0827</i> and <i>atu0828</i>	Unclassified	
		circular	860664	SNV	37	C → G		<i>atu0863</i>		Hypothetical protein	Transcription	
		circular	1021837	SNV	100	C → T				Between <i>Atu1028</i> / <i>Atu1029</i>	Unclassified	
		circular	1023751	SNV	100	G → A	Arg364His	<i>atu1030</i>		GTP pyrophosphohydrolase/synthetase	Transcription	
		circular	1368399	SNV	100	A → G	Lys132Arg	<i>atu1374</i>	<i>rpsB</i>	30S ribosomal protein S2	Translation, ribosomal structure and biogenesis	
		circular	1381967	SNV	100	G → T	Arg161Leu	<i>atu1387</i>		GntR family transcriptional regulator	Transcription	
		circular	1562277	SNV	100	G → A				Between <i>Atu1575</i> / <i>Atu1577</i>	Unclassified	
		circular	1562966	SNV	100	A → G	Thr167Ala	<i>atu1577</i>		ABC transporter, substrate binding protein (amino acid)	Amino acid transport and metabolism	
		circular	1563380	Deletion	88	CTG → -	Leu305del	<i>atu1577</i>		ABC transporter, substrate binding protein (amino acid)	Amino acid transport and metabolism	
		circular	1633942	SNV	99	T → C				Between <i>Atu1649</i> / <i>Atu1650</i>	Unclassified	
		circular	1873736	SNV	100	G → A		<i>atu1897</i>		short chain dehydrogenase	Lipid transport and metabolism	
		circular	1907071	SNV	100	A → T	Val35Asp	<i>atu1929</i>	<i>rpsE</i>	30S ribosomal protein S5	Translation, ribosomal structure and biogenesis	

Table 2. Genetic variants appeared in passage t18 for each different line. Blue color indicates that the variation was present in passage 7.

Ancestor	Line	Pool size	Replicon	Position	Type	Freq (%)	Allele	Amino acid change	atu code	Gene	Product	COG classification
pTi-Gm	L18	30	circular	1920197	SNV	100	T → C	Lys43Arg	<i>atu1951</i>	<i>rpsL</i>	30S ribosomal protein S12	Translation, ribosomal structure and biogenesis
			circular	1929719	SNV	99	C → T	Arg539His	<i>atu1956</i>	<i>rpoB</i>	DNA-directed RNA polymerase subunit beta	Transcription
			circular	2174067	SNV	99	T → C	Val49Ala	<i>atu2200</i>	<i>cspA</i>	cold shock protein	Transcription
			circular	2425790	SNV	100	A → T	Ser185Thr	<i>atu2454</i>		iron-chelator utilization protein	Inorganic ion transport and metabolism
			circular	2531634	SNV	99	T → G	Thr198Pro	<i>atu2559</i>		LysR family transcriptional regulator	Transcription
			circular	2615103	SNV	100	C → A	Ala117Ser	<i>atu2631</i>	<i>cybB</i>	cytochrome b561	Energy production and conversion
			linear	4480	Deletion	96	A → -	Phe111fs	<i>atu3006</i>		conserved hypothetical protein, membrane protein	Function unknown
			linear	153903	Deletion	99	T → -	Ile73fs	<i>atu3146</i>		putative Xaa-Pro aminopeptidase	Amino acid transport and metabolism
			linear	155876	SNV	99	G → A	Ser161Leu	<i>atu3147</i>		oxidoreductase	General function prediction only
			linear	217854	SNV	100	G → T	Leu252Met	<i>atu3205</i>		Conserved protein of unknown function	General function prediction only
			linear	303731	SNV	99	G → T		<i>atu3278</i>		aryl-alcohol dehydrogenase	General function prediction only
			linear	530170	SNV	100	T → C		<i>atu3480</i>	<i>hmgL</i>	hydroxymethylglutaryl-CoA lyase	Amino acid transport and metabolism
			linear	552125	SNV	100	T → C	Asp237Gly	<i>atu3501</i>	<i>cysA2</i>	Sulfate/thiosulfate import ATP-binding protein CysA 2	Inorganic ion transport and metabolism
			linear	676317	SNV	100	A → C	Val40Gly	<i>atu3621</i>	<i>cobT</i>	cobyrinic acid synthase	Coenzyme transport and metabolism
			linear	685673	SNV	99	T → C	Ile83Val	<i>atu3633</i>		hypothetical protein	General function prediction only
			linear	721932	SNV	97	G → A	Ala207Val	<i>atu3671</i>		Putative luciferase-like monooxygenase (PP-binding)	Energy production and conversion
			linear	735472	SNV	100	T → C		<i>atu3677</i>		Putative glutamate-1-semialdehyde 2,1-aminomutase, porphyrin biosynthesis protein	Coenzyme transport and metabolism
			linear	750532	SNV	100	T → C	Val272Ala	<i>atu3683</i>		Putative non-ribosomal peptide synthetase (NRPS)	Secondary metabolites biosynthesis, transport and catabolism
			linear	1090536	SNV	96	G → T	Thr335Asn	<i>atu3977</i>	<i>ina</i>	ice nucleation-like protein	Intracellular trafficking, secretion, and vesicular transport
			linear	1149844	SNV	100	G → A	Ala223Thr	<i>atu4034</i>		two component sensor kinase	Signal transduction mechanisms
			linear	1199120	SNV	100	T → C	Asn400Ser	<i>atu4078</i>	<i>glgP</i>	glycogen phosphorylase	Carbohydrate transport and metabolism
			linear	1231723	SNV	100	C → A		<i>atu4112</i>		proline dipeptidase	Amino acid transport and metabolism
			linear	1281458	SNV	100	T → C	Val240Ala	<i>atu4163</i>		two component sensor kinase	Signal transduction mechanisms
			linear	1326408	SNV	99	A → G	Val101Ala	<i>atu4204</i>		putative DNA-binding transcriptional regulator; prophage	Transcription
			linear	1358427	SNV	100	T → C	Ile6Thr	<i>atu4233</i>		ABC transporter, substrate binding protein (amino acid)	Amino acid transport and metabolism
			linear	1387019	SNV	100	T → C	Thr129Ala	<i>atu4260</i>		ABC transporter, membrane spanning protein	Inorganic ion transport and metabolism
			linear	1661950	SNV	99	A → G		<i>atu4514</i>		RpiR family transcriptional regulator	Transcription
			linear	1953053	SNV	99	T → C	Ile18Val	<i>atu4781</i>		putative two component sensor kinase	Signal transduction mechanisms
			linear	1953541	SNV	99	C → T	Asp103Asn	<i>atu4782</i>		putative two component response regulator	Transcription

4. DISCUSSION

In this work, we have used an experimental evolution assay to analyze the mechanisms used by *A. tumefaciens* for a long-term adaptation to tumor lifestyle.

4.1. Experimental evolution assay shows an increase in genome variations during the course of the experiment

Combining genome sequencing methods with our experimental evolution study we could observe that variations in the genome were accumulated during the passages. Many of these variations, which at first moment were not totally fixed in the population, became installed until they were totally established in the final population. Most of the mutations were non-conservative. This possibly produces beneficial variations, as they are not neutral and influence the protein sequence that could lead to improve fitness (Barrick *et al.*, 2009).

Parallel mutations were detected within the replicates of the same strain and between the replicates of the different strains. Such variations confirm the action of natural selection on these genes. These mutations on this small scale, the same nucleotide, are quite rare (Dettman *et al.*, 2012). Experimentally, although a gene is under a strong selection, mutations generally occur in different residues within the same gene. A recent *in vitro* study on *A. tumefaciens* C58 revealed that under certain conditions, bacteria are capable of presenting mutations within the same gene but in different positions (Tannières *et al.*, 2017). Another study in *Escherichia coli* long term experiment reflects this kind of parallel mutations (Woods *et al.*, 2006).

4.2. QS has no impact on the genome variation of *A. tumefaciens* C58

pTi-*accR* strain is able to transfer constitutively its Ti plasmid. This characteristic could make us think that the number of variations in the genome would be higher due to this gene transfer. At the same time, the contrary could be possible for the pTi-*traI* strain. Sequencing data reflect that there is no relationship between the number of variations and the QS system.

4.3. Variation in *A. tumefaciens* C58 genome lead to beneficial mutations that improve fitness on tomato tumors.

In both sequencing points, we observed that in most cases, bacterial populations carried only a few variations in their genome. It has been demonstrated that most evolved strains harbor just only some SNPs and this fact is correlated with better fitness (Dettman *et al.*, 2012). *Ralstonia solanacearum* evolved clones from host plants presented no more than 3 different changes, reinforcing this hypothesis (Guidot *et al.*, 2014).

Fitness assays of some evolved clones demonstrated that mutations on *atu3932*, *atu0872*, *atu4160* and *atu0499* genes result in an increase of fitness of evolved clones respecting their ancestors.

atu0872 gene coding for a methyl-accepting chemotaxis protein had a higher fitness than pTi-*traI*. This result together with those that showed a different chemotaxis capacity (with mannitol, glucose and fructose as a carbon source) than its ancestor can explain this fitness improvement. Especially important is this chemotaxis ability because it has already been proven that at least in *A. thaliana* tumors glucose and fructose are the most abundant carbon compounds accumulated (Deeken *et al.*, 2006).

atu0499 gene encodes a polar development protein J. A clone carrying this variation also colonize better tomato tumors than pTi-*accR* ancestor. The implication of this protein in polar growth (Grangeon *et al.*, 2015) could suggest also a better attachment to plant surfaces.

As has been shown, the L18 clone has a smaller size than its ancestor, pTi-Gm. This result together with the improved fitness may suggest that its ability to colonize better the tumor may be due to this smaller size. The fact that more than 40 SNPs were found within its genome makes it impossible to reject the effect of one or more of these variations.

5. CONCLUSIONS

This work presents the first study of the evolutionary dynamics of the genome of *A. tumefaciens* C58 inside the tumor. The experiment in tomato plants has demonstrated the ability of this bacteria to evolve in a host plant and increase the fitness in it. Although certain aspects could not be strongly studied, as the generation number, this method can be used to identify new factors involved in the adaptation of the bacteria to the tumor lifestyle. A more in-depth study of the last stage of the experiment, as well as the study of the different variations presented in line 18, can further clarify the determinants of niche adaptation.

6. BIBLIOGRAPHY

- Anderson-Furgeson, J.C., Zupan, J.R., Grangeon, R. and Zambryski, P.C.** (2016) Loss of PodJ in *Agrobacterium tumefaciens* leads to ectopic polar growth, branching, and reduced cell division. *J. Bacteriol.* **198**, 1883–1891.
- Barrick, J., Yu, D., Yoon, S., Jeong, H., Oh, T., Schneider, D., Lenski, R. and Kim, J.** (2009) Genome evolution and adaptation in a long-term experiment with *Escherichia coli*. *Nature* **461**, 1243–1247.
- Beck von Bodman, S., Hayman, G.T. and Farrand, S.K.** (1992) Opine catabolism and conjugal transfer of the nopaline Ti plasmid pTiC58 are coordinately regulated by a single repressor. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 643–647.
- Chilton, M.D., Currier, T.C., Farrand, S.K., Bendich, A.J., Gordon, M.P. and Nester, E.W.** (1974) *Agrobacterium tumefaciens* DNA and PS8 bacteriophage DNA not detected in crown gall tumors. *Proc. Natl. Acad. Sci. U. S. A.* **71**, 3672–6.
- Deeken, R., Engelmann, J.C., Efetova, M., et al.** (2006) An Integrated View of Gene Expression and Solute Profiles of *Arabidopsis* Tumors: A Genome-Wide Approach. *Plant Cell Online* **18**, 3617–3634.
- 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. *Mol. Ecol.* **21**, 2058–2077.
- Fuqua, W.C. and Winans, S.C.** (1994) A LuxR-LuxI type regulatory system activates *Agrobacterium* Ti plasmid conjugal transfer in the presence of a plant tumor metabolite. *J. Bacteriol.* **176**, 2796–2806.
- Grangeon, R., Zupan, J.R., Anderson-Furgeson, J. and Zambryski, P.C.** (2015) PopZ identifies the new pole, and PodJ identifies the old pole during polar growth in *Agrobacterium tumefaciens*. *Proc. Natl. Acad. Sci. United States* **112**, 11666–11671.
- Guidot, A., Jiang, W., Ferdy, J.B., Thébaud, C., Barberis, P., Gouzy, J. and Genin, S.** (2014) Multihost experimental evolution of the pathogen *Ralstonia solanacearum* unveils genes involved in adaptation to plants. *Mol. Biol. Evol.* **31**, 2913–2928.
- Haudecoeur, E., Planamente, S., Cirou, A., Tannières, M., Shelp, B.J., Moréra, S. and Faure, D.** (2009) Proline antagonizes GABA-induced quenching of quorum-sensing in *Agrobacterium tumefaciens*. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 14587–92.
- Haudecoeur, E., Tannieres, M., Cirou, A., Raffoux, A., Dessaux, Y. and Faure, D.** (2009) Different regulation and roles of lactonases AiiB and AttM in *Agrobacterium tumefaciens* C58. *Mol Plant Microbe Interact* **22**, 529–537.

- Heindl, J.E., Hibbing, M.E., Xu, J., Natarajan, R., Buechlein, A.M. and Fuqua, C.** (2016) Discrete responses to limitation for iron and manganese in *Agrobacterium tumefaciens*: Influence on attachment and biofilm formation. *J. Bacteriol.* **198**, 816–829.
- Heindl, J.E., Wang, Y., Heckel, B.C., Mohari, B., Feirer, N. and Fuqua, C.** (2014) Mechanisms and regulation of surface interactions and biofilm formation in *Agrobacterium*. *Front. Plant Sci.* **5**, 176.
- Hwang, I., Li, P.L., Zhang, L., Piper, K.R., Cook, D.M., Tate, M.E. and Farrand, S.K.** (1994) Tral, a LuxI homologue, is responsible for production of conjugation factor, the Ti plasmid N-acylhomoserine lactone autoinducer. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 4639–4643.
- 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.
- Macho, A.P., Guidot, A., Barberis, P., Beuzón, C.R. and Genin, S.** (2010) A competitive index assay identifies several *Ralstonia solanacearum* type III effector mutant strains with reduced fitness in host plants. *Mol. Plant. Microbe. Interact.* **23**, 1197–1205.
- Paintdakhi, A., Parry, B., Campos, M., Irnov, I., Elf, J., Surovtsev, I. and Jacobs-Wagner, C.** (2016) Oufiti: an integrated software package for high-accuracy, high-throughput quantitative microscopy analysis. *Mol. Microbiol.* **99**, 767–77.
- Piper, K.R., Beck von Bodman, S. and Farrand, S.K.** (1993) Conjugation factor of *Agrobacterium tumefaciens* regulates Ti plasmid transfer by autoinduction. *Nature* **362**, 448–50.
- Piper, K.R. and Farrand, S.K.** (2000) Quorum sensing but not autoinduction of Ti plasmid conjugal transfer requires control by the opine regulon and the antiactivator TraM. *J. Bacteriol.* **182**, 1080–1088.
- Planamente, S., Mondy, S., Hommais, F., Vigouroux, A., Moréra, S. and Faure, D.** (2012) Structural basis for selective GABA binding in bacterial pathogens. *Mol. Microbiol.* **86**, 1085–1099.
- Planamente, S., Vigouroux, A., Mondy, S., Nicaise, M., Faure, D. and Moréra, S.** (2010) A conserved mechanism of GABA binding and antagonism is revealed by structure-function analysis of the periplasmic binding protein Atu2422 in *Agrobacterium tumefaciens*. *J. Biol. Chem.* **285**, 30294–30303.
- Rosen, R., Matthysse, A.G., Becher, D., Biran, D., Yura, T., Hecker, M. and Ron, E.Z.** (2003) Proteome analysis of plant-induced proteins of *Agrobacterium tumefaciens*. *FEMS Microbiol. Ecol.* **44**, 355–360.

- Tannières, M., Lang, J., Barnier, C., et al.** (2017) Quorum-quenching limits quorum-sensing exploitation by signal-negative invaders. *Sci. Rep.* **7**, 40126.
- Tomlinson, A.D. and Fuqua, C.** (2009) Mechanisms and regulation of polar surface attachment in *Agrobacterium tumefaciens*. *Curr. Opin. Microbiol.* **12**, 708–14.
- Woods, R., Schneider, D., Winkworth, C.L., Riley, M. a and Lenski, R.E.** (2006) Tests of parallel molecular evolution in a long-term experiment with *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 9107–9112.
- Yuan, Z.C., Liu, P., Saenkham, P., Kerr, K. and Nester, E.W.** (2008) Transcriptome profiling and functional analysis of *Agrobacterium tumefaciens* reveals a general conserved response to acidic conditions (pH 5.5) and a complex acid-mediated signaling involved in *Agrobacterium*-plant interactions. *J. Bacteriol.* **190**, 494–507.

Chapter V

Chapter V: Discussion

1. Transcriptome *in planta*

During the last years, many transcriptomic works have been carried out with the objective of analyzing the interactions of species, the adaptation and the results of the evolutionary processes (Kammenga *et al.*, 2007).

1.1. *A. tumefaciens* cells are in a slow growth state in *A. thaliana* tumors.

Transcriptome analysis results allowed us to clarify the state of the bacterium inside an *A. thaliana* mature tumor. *A. tumefaciens* cells are in a general state of slow growth within the tumor, due to the large number of downregulated genes associated with growth processes and cellular activity.

As we have seen, many of the important functions of the bacteria are suppressed. This indicates that the bacteria are inside the tumor in slow growth or “dormancy”. *Rhizobium etli* in bacteroids presents also a state of non-growth, with its most important metabolic functions repressed (Vercruyssen *et al.*, 2011). *R. etli* belongs also to the *Rhizobiaceae* family, as *A. tumefaciens*. These results highlight the similarities between the two bacteria, which had a common ancestor and finally they evolved giving rise to a symbiotic bacterium and a pathogenic bacterium.

The repression of the main respiratory chain genes not only signals a reduction in the cellular activity by the decrease of the ATP production, but could also be a consequence of the decrease of the oxygen concentration inside the tumor as described by (Deeken *et al.*, 2006). The increase in the expression of nitric oxide reductases may also be due to not only as a response to the nitric oxide produced by the plant as a defense product, but could also be used as electron acceptors in this new microaerobic environment. The low content of nitrate in plant tumors (Deeken *et al.*, 2006) explains the fact that none nitrate reductase nor nitrite reductase genes were differentially expressed on tumor condition.

Tumor tissue of *A. thaliana* is a heterotrophic environment. The total amino acids concentration in the tumor tissues of is nearly 9 times higher compared with a control tissue as shown in (Deeken *et al.*, 2006). The transcriptome data reveal that genes involved in the amino acid synthesis were downregulated. It is possible that *A. tumefaciens* can import them inside the cell due to the amount of ABC transporters are

overexpressed. In tumor cells of *A. thaliana*, glucose appeared to be the main carbon source. *A. tumefaciens* carbon metabolism was downregulated evidencing that the plant can supply all these compounds to the bacteria that colonize the tumor.

The two best characterized bacterial PAMPs are peptides derived from flagellin (flg22) and EF-Tu elongation factor (Felix *et al.*, 1999; Kunze *et al.*, 2004). It has been demonstrated that the EF-Tu elongation factor can induce 35 genes in *A. thaliana* (Lee *et al.*, 2009). In mature tumors, 28 of these genes were overexpressed. This contrasts with the evidence that this gene is downregulated in the transcriptome of *A. tumefaciens*. One hypothesis is that although it is able to modulate the expression of this gene, the plant is able to recognize it, but this recognition is also quite weak. This suggests a possible co-evolution of both participants.

Agrobacterium appears to be able to suppress early accumulation of H₂O₂, but not when during the more advanced stages of the tumor, where it is accumulated (Lee *et al.*, 2009). In our study, *A. tumefaciens* genes involved in oxidative stress, such as catalases or peroxidases, are not differentially expressed in tumor condition. It seems that the accumulation of H₂O₂ in the tumor tissue is due to the strengthening of the cell walls than the induction of an HR (Gohlke and Deeken, 2014).

Motility and chemotaxis were also strongly affected in the tumor condition. Once the bacteria has entered inside the host plant, plant cells modify their organization and start to grow cell walls (Deeken *et al.*, 2006). It is also known as the bacteria are located in the apoplast of plant tissue. *A. tumefaciens* is not able to completely colonize the tumor, but is predominantly established on the external part of the tumor (Lang *et al.*, 2016). Also genes involved in attachment of the bacteria to surfaces as unipolar polysaccharide (UPP) were found downregulated. UPP plays an important role in soil and plant colonization (Matthysse, 2014). It seems that *A. tumefaciens* use some chemoreceptors to detect different substances in the soil (Guo *et al.*, 2017). The repression of these functions that seem to be important in the rhizosphere colonization by *A. tumefaciens* points out the differences between both environments. Furthermore, it indicates that the bacterium is able to adapt and modify its gene expression depending on the environment.

1.2. Transcriptome analysis of *A. tumefaciens* tumor cells reveals the importance of the plasmidome in tumor lifestyle.

The genome of *A. tumefaciens* consists of more than 5600 putative genes. The results obtained from the comparison between *A. tumefaciens* cells living under planktonic and plant tumor conditions show that almost a third of the genome of the bacterium is differentially expressed with respect to the free-living condition. The genome of *A. tumefaciens* consists of more than 5600 putative genes. Almost a third of the genome of the bacterium was differentially expressed in plant tumor. Similarly, the distribution of genes follows a “pattern”: most genes of both chromosomes are generally repressed and the genes of the At and Ti plasmids are mainly upregulated, suggesting a that both plasmid are important for tumor colonization.

Both At and Ti plasmids of *A. tumefaciens* C58 have similar *repABC* replication systems (Wood *et al.*, 2001). It is supposed that two plasmids that share the same replication system are incompatible and cannot remain in the same bacterial cell (Bouet *et al.*, 2007). In our results, both replication systems were upregulated, reinforcing the research of Lang *et al.*, (2013), in which it was demonstrated the co-transfer of the two plasmids. In addition, results obtained in the Tn-Seq analysis have remarked that these genes are essential for the maintenance of both plasmids in the bacterium, that is the reason why these *rep* mutants disappeared in the Tn-Seq EL-ARTIST analysis. The cost associated with the maintenance of both plasmids has already been analyzed in different studies (Morton *et al.*, 2013; Platt *et al.*, 2012),. So, the advantage of maintaining them is due to the benefit they provide in various environments. For the Ti plasmid is essential for virulence and the At plasmid confers an advantage on the rhizosphere (Morton *et al.*, 2014).

Ti plasmid is essential for *A. tumefaciens* virulence. In the transcriptome analysis, genes involved in virulence and T-DNA transfer were upregulated. This indicates that even in mature tumors, bacteria still have the ability to transfer their T-DNA and transform new plant cells and thus continue the infection. This result coincides with the studies conducted by Rezmer *et al.*, (1999) and Deeken *et al.*, (2006) in which it was observed that more than 95% of the tumor cells contained the T-DNA. Genes involve in opine degradation and production of QS signals were also upregulated, showing that most important functions of the Ti plasmid are upregulated in the tumor and its relevance in the infection process.

As already mentioned, At plasmid contains important functions for the ecology of the *A. tumefaciens*, but it entails also a high cost for the bacteria, although a recent study has shown that the presence of the plasmid Ti mitigates the cost of the pAt (Platt *et al.*, 2014). It has been shown that a deletion of a broad region of the At plasmid can take place and has been associated with fitness gain under cell culture conditions (Morton *et al.*, 2013). It is known that plasmid At contains numerous ABC transporters (Wood *et al.*, 2001) and seems that these systems allow bacteria to import a broad variety of energy sources, some of which may be specific to plant tumors. The transcriptome analysis shows that a large part of them are upregulated in plant tumors. At plasmid also harbors genes involved in the assimilation of Amadori compounds, that are present in rhizosphere and plant tumors (Baek *et al.*, 2005). BlcC lactonase encoded by the At plasmid and AiiB on the Ti plasmid, were also upregulated. They contribute to the attenuation of the transfer of the Ti plasmid (Lang *et al.*, 2016) and allow a fitness gain within the tumor (Haudecoeur, Tannières, *et al.*, 2009). In addition, most of the genes from the *att* operon were also upregulated. It has been demonstrated that *attC* is involved in virulence due to its role in bacterial attachment (Matthysse *et al.*, 2008; Matthysse and McMahan, 1998). The particularity that the conservation of the At plasmid supposes a high cost to the bacterium can explain the fact that the constructed KO-mutants presented an increase of their fitness respecting the wild type strain.

All these data, demonstrate that the role of the plasmidome is crucial for the adaptation of *A. tumefaciens* to the plant tumor, and in particular the companion At plasmid.

2. Transcriptome *in vitro* and Tn-Seq analysis of GHB and GABA cultures

QS system of *A. tumefaciens* has been extensively studied. The conjugative role of opines (Piper *et al.*, 1999) as well as the existence of two lactonases (AiiB and BlcC) (Haudecoeur, Tannières, *et al.*, 2009) has also been already documented

γ -Aminobutyric acid (GABA) is a non-protein amino acid presented in bacteria, plants and animals. It has several possible roles in plants as: (i) contributing to carbon/nitrogen balance; regulation of pH; defense against ROS and insect attack; osmoregulation and signal molecule (Bouché and Fromm, 2004).

As described in the previous section, large amounts of amino acids accumulate within the tumor tissue and the non-protein nitrogen compound most abundant is GABA (Deeken *et al.*, 2006). In addition, we can also highlight the accumulation of GABA and GABA-related compounds (GHB and SSA) as shown by (Lang *et al.*, 2016). GHB, SSA and GABA induce the expression of the BlcC lactonase and consequently its impact on the bacterial QS (Chevrot *et al.*, 2006).

In chapter 2, we have demonstrated that the metabolism of GHB when it was used as a carbon source relies primarily on the gluconeogenesis pathway. The comparison of two different but complementary techniques has helped us to highlight the key steps in the degradation of this molecule. Recent research using the Tn-Seq technique established that *Pseudomonas aeruginosa* is highly vulnerable to changes that interrupt the central carbon-energy metabolism (Lee *et al.*, 2015).

Transcriptomic results emphasize the fact that GABA and GHB induce the operon *blcABC*. This agrees with the fact that GABA is a derepressor and not the substrate of any of these genes. On the contrary, GHB is substrate of BlcB. The product of this reaction is SSA, toxic to the bacterium, and converted to succinate by BlcA. Analyzing these results with Tn-Seq data, we could observe that in GHB condition *blcA* and *blcB* genes are essential for the bacterial survival. On the contrary, none of those genes were essential in GABA condition.

Uptake and degradation pathway of GABA has been well characterized. In *A. tumefaciens* cells, GABA is degraded into SSA by a still unknown transaminase (Haudecoeur and Faure, 2010). Transcriptome analysis of GABA cultures pointed to two putative transaminases involved in its degradation. KO-mutants phenotypes did not differ from the wild type strain. Examining Tn-Seq results, we could not find any gene with a

putative transaminase function. No GABA transport system was found differentially expressed or found essential. In *A. tumefaciens* tumors, *atu2422* gene was downregulated. GABA compete with proline, another abundant amino acid in plant tumor, using the same periplasmic binding protein (PBP) Atu2422 (Haudecoeur, Planamente, *et al.*, 2009). Transcriptome analysis of GHB revealed several different ABC transports upregulated, but no essential genes were found neither in Tn-Seq analysis. The uptake of GHB by *A. tumefaciens* is still also unknown.

The analysis of essential genes from the initial library resulted in the establishment of 527 essential genes in the initial library, i.e. 527 genes in which the transposon was not inserted or it was on a very small scale. Among all of them, most were present on the circular chromosome. When we cross these data with the *in planta* transcriptome data, we observed that from the 440 essential genes, 334 were differentially expressed, being the majority downregulated genes. This result confirms the different status of the bacteria depending on the environment.

Results obtained by the transcriptomes coincide only partially with data obtained by the Tn-Seq technique. This reflects the limits of both methods. Understanding the transcriptome results is essential for interpreting the functional elements of the genome and for understanding bacteria state. Limits due to the knowledge about the genome sequence, genome coverage and bioinformatics analysis (Wang *et al.*, 2009). In the case of the Tn-Seq method, the effect depending on the number of mutants the of the initial library may be limiting the fitness for some genes (Barquist *et al.*, 2013). Besides, redundant functions are not represented in this analysis, so important genes could not be detected (Meeske *et al.*, 2016). Finally, bottlenecks could also be a handicap of this method because it would be difficult to know whether the mutant has disappear because the reduction of fitness or by chance (Stephens *et al.*, 2015).

3. Experimental evolution in *A. tumefaciens* C58

Experimental evolution is the use of laboratory or controlled field manipulations to explore evolutionary dynamics. Although it has been used for decades, last years this method has been increasingly used to study different hypotheses. Among them we can highlight: mutation and adaptation, genetic drift and inbreeding, environmental variability, sexual selection, cooperation, behavior, speciation and host-parasite evolution (Kawecki *et al.*, 2012), in which our study is located.

During two years, we performed an experimental evolution study of three different populations of *A. tumefaciens* C58 in the host plant *S. lycopersicum*. The reason to perform this experiment in a host plant is because *A. tumefaciens* has a wide spectrum of guests and especially the interest of knowing if an evolutionary process is possible in which the bacteria increase their fitness. Different scientific studies have been carried out on plant pathogens associated to different hosts. In many cases, the aim of these experiments is the adaptation processes of bacteria to different environments. An example is the work done by (Marchetti *et al.*, 2010) in which the *Ralstonia solanacearum* became a symbiotic bacterium. Another work using the same pathogen had as objective to find out possible mechanisms of the adaptation of the bacteria to both host and non - host plants (Guidot *et al.*, 2014).

Three different populations were used for this study. In addition to the wild strain, two other strains with different properties with respect to QS were used. The ability to transfer the Ti plasmid and with them the mobilization and transfer of genes was the reason why both were used.

The confirmation that only a few variations appeared in the genome leads to the belief that they may be related to a change in the phenotype and in the fitness of the same (Dettman *et al.*, 2012). With the passage of time it was also possible to verify that many of those first changes were fixed in the whole population and that some other disappeared (at least of the sample). In addition, we confirmed, by performing fitness tests, that some of these variations showed higher fitness than their ancestor strains.

Another evolution experiment was carried out in our laboratory using both the wild strain and the *accR* strain. This in vitro experimental evolution allowed to discover different adaptation processes among those that are centered: complete loss of Ti plasmid, mutations in the regulon *accR* and a mutation in the *atu1360* gene (Tannières *et al.*, 2017). It was shown that bacteria that constitutively produce QS signals can eliminate

this high cost using QS signals from neighboring bacteria, thus becoming QS-hijacking.

As confirmed in the first chapter, in the host plant the maintenance of both plasmids is essential for the adaptation of the bacteria to the tumor environment. Therefore, most of the genes subject to variations were found in both chromosomes. Moreover, many of these changes were placed on downregulated genes of the plant transcriptome data. This allows to suggest that for a better adaptation to the tumor lifestyle, bacteria will try to modify these downregulated genes to either completely silence them or modify the protein for an improvement of fitness.

The slow growth *in planta* and the population setting in the apoplastic environment, could lead to bottlenecking during the experiment. This problem occurred during the course of the experiment. Moreover, due to this slow growth, it was impossible to determine the generation number.

Finally, the presence of parallel mutations in different lines of different strains highlights the importance of this evolution experiment.

PERSPECTIVES

The combination of the transcriptomic analysis and the Tn-Seq have allowed to find genes involved and in turn essential in the process of adaptation of the bacteria to a certain condition. Therefore, the Tn-Seq technique could be used to associate the genes that are essential in the tumor to the results of the in-plant transcriptome. Independently, the Tn-seq bank could also be used to monitor the infection of the bacteria in the rhizosphere.

The necessity to compare the results of the Tn-seq analysis to a sample in which the bacterium has been grown in minimal medium would avoid finding in the results many genes that are essential for that passage from heterotroph to auxotroph environment.

Finally, a study on the fitness of the new variations of the experimental evolution would allow to consolidate the already acquired knowledge and to increase the information pertaining to the adaptation of the bacterium to the tumor.

BIBLIOGRAPHY

- Baek, C.H., Farrand, S.K., Park, D.K., Lee, K.E., Hwang, W. and Kim, K.S.** (2005) Genes for utilization of deoxyfructosyl glutamine (DFG), an amadori compound, are widely dispersed in the family *Rhizobiaceae*. *FEMS Microbiol. Ecol.* **53**, 221–233.
- Barquist, L., Boinett, C.J. and Cain, A.K.** (2013) Approaches to querying bacterial genomes with transposon-insertion sequencing. *RNA Biol.* **10**, 1161–9.
- Bouché, N. and Fromm, H.** (2004) GABA in plants: Just a metabolite? *Trends Plant Sci.* **9**, 110–115.
- Bouet, J.Y., Nordström, K. and Lane, D.** (2007) Plasmid partition and incompatibility - The focus shifts. *Mol. Microbiol.* **65**, 1405–1414.
- Chevrot, R., Rosen, R., Haudecoeur, E., Cirou, A., Shelp, B.J., Ron, E. and Faure, D.** (2006) GABA controls the level of quorum-sensing signal in *Agrobacterium tumefaciens*. *Proc. Natl. Acad. Sci.* **103**, 7460–7464.
- Deeken, R., Engelmann, J.C., Efetova, M., et al.** (2006) An Integrated View of Gene Expression and Solute Profiles of *Arabidopsis* Tumors: A Genome-Wide Approach. *Plant Cell Online* **18**, 3617–3634.
- 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. *Mol. Ecol.* **21**, 2058–2077.
- Felix, G., Duran, J.D., Volko, S. and Boller, T.** (1999) Plants have a sensitive perception system for the most conserved domain of bacterial flagellin. *Plant J.* **18**, 265–276.
- Gohlke, J. and Deeken, R.** (2014) Plant responses to *Agrobacterium tumefaciens* and crown gall development. *Front. Plant Sci.* **5**, 155.
- Guidot, A., Jiang, W., Ferdy, J.B., Thébaud, C., Barberis, P., Gouzy, J. and Genin, S.** (2014) Multihost experimental evolution of the pathogen *Ralstonia solanacearum* unveils genes involved in adaptation to plants. *Mol. Biol. Evol.* **31**, 2913–2928.
- Guo, M., Huang, Z. and Yang, J.** (2017) Is there any crosstalk between the chemotaxis and virulence induction signaling in *Agrobacterium tumefaciens*? *Biotechnol. Adv.*
- Haudecoeur, E. and Faure, D.** (2010) A fine control of quorum-sensing communication in *Agrobacterium tumefaciens*. *Commun. Integr. Biol.* **3**, 84–88.
- Haudecoeur, E., Planamente, S., Cirou, A., Tannières, M., Shelp, B.J., Moréra, S. and Faure, D.** (2009) Proline antagonizes GABA-induced quenching of quorum-sensing in *Agrobacterium tumefaciens*. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 14587–92.

- Haudecoeur, E., Tannieres, M., Cirou, A., Raffoux, A., Dessaux, Y. and Faure, D.** (2009) Different regulation and roles of lactonases AiiB and AttM in *Agrobacterium tumefaciens* C58. *Mol Plant Microbe Interact* **22**, 529–537.
- Kammenga, J.E., Herman, M.A., Ouborg, N.J., Johnson, L. and Breitling, R.** (2007) Microarray challenges in ecology. *Trends Ecol. Evol.* **22**, 273–279.
- 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.
- Kunze, G., Zipfel, C., Robatzek, S., Niehaus, K., Boller, T. and Felix, G.** (2004) The N terminus of bacterial elongation factor Tu elicits innate immunity in *Arabidopsis* plants. *Plant Cell* **16**, 3496–507.
- Lang, J., Gonzalez-Mula, A., Tacconat, L., Clement, G. and Faure, D.** (2016) The plant GABA signaling downregulates horizontal transfer of the *Agrobacterium tumefaciens* virulence plasmid. *New Phytol.* **210**, 974–983.
- Lang, J., Planamente, S., Mondy, S., Dessaux, Y., Moréra, S. and Faure, D.** (2013) Concerted transfer of the virulence Ti plasmid and companion At plasmid in the *Agrobacterium tumefaciens*-induced plant tumour. *Mol. Microbiol.* **90**, 1178–1189.
- Lee, C.-W., Efetova, M., Engelmann, J.C., Kramell, R., Wasternack, C., Ludwig-Müller, J., Hedrich, R. and Deeken, R.** (2009) *Agrobacterium tumefaciens* promotes tumor induction by modulating pathogen defense in *Arabidopsis thaliana*. *Plant Cell* **21**, 2948–62.
- Lee, S.A., Gallagher, L.A., Thongdee, M., Staudinger, B.J., Lippman, S., Singh, P.K. and Manoil, C.** (2015) General and condition-specific essential functions of *Pseudomonas aeruginosa*. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 5189–94.
- Marchetti, M., Capela, D., Glew, M., et al.** (2010) Experimental evolution of a plant pathogen into a legume symbiont. *PLoS Biol.* **8**.
- Matthysse, A.G.** (2014) Attachment of *Agrobacterium* to plant surfaces. *Front. Plant Sci.* **5**, 252.
- Matthysse, A.G., Jaeckel, P. and Jeter, C.** (2008) *attG* and *attC* mutations of *Agrobacterium tumefaciens* are dominant negative mutations that block attachment and virulence. *Can. J. Microbiol.* **54**, 241–247.
- Matthysse, A.G. and McMahan, S.** (1998) Root colonization by *Agrobacterium tumefaciens* is reduced in *cel*, *attB*, *attD*, and *attR* mutants. *Appl. Environ. Microbiol.* **64**, 2341–2345.
- Meeske, A.J., Rodrigues, C.D.A., Brady, J., Lim, H.C., Bernhardt, T.G. and Rudner, D.Z.** (2016) High-Throughput Genetic Screens Identify a Large and Diverse Collection of New Sporulation Genes in *Bacillus subtilis*. *PLoS Biol.* **14**, 1–

- Morton, E.R., Merritt, P.M., Bever, J.D. and Fuqua, C.** (2013) Large deletions in the pAtC58 megaplasmid of *Agrobacterium tumefaciens* can confer reduced carriage cost and increased expression of virulence genes. *Genome Biol. Evol.* **5**, 1453–1464.
- Morton, E.R., Platt, T.G., Fuqua, C. and Bever, J.D.** (2014) Non-additive costs and interactions alter the competitive dynamics of co-occurring ecologically distinct plasmids. *Proc. Biol. Sci.* **281**, 20132173.
- Piper, K.R., Beck Von Bodman, S., Hwang, I. and Farrand, S.K.** (1999) Hierarchical gene regulatory systems arising from fortuitous gene associations: Controlling quorum sensing by the opine regulon in *Agrobacterium*. *Mol. Microbiol.* **32**, 1077–1089.
- Platt, T.G., Bever, J.D. and Fuqua, C.** (2012) A cooperative virulence plasmid imposes a high fitness cost under conditions that induce pathogenesis. *Proc. Biol. Sci.* **279**, 1691–9.
- Platt, T.G., Morton, E.R., Barton, I.S., Bever, J.D. and Fuqua, C.** (2014) Ecological dynamics and complex interactions of *Agrobacterium* megaplasmids. *Front. Plant Sci.* **5**, 635.
- Rezmer, C., Schlichting, R., Wächter, R. and Ullrich, C.I.** (1999) Identification and localization of transformed cells in *Agrobacterium tumefaciens* - induced plant tumors. *Planta* **209**, 399–405.
- Stephens, W.Z., Wiles, T.J., Martinez, E.S., Jemielita, M., Burns, A.R., Parthasarathy, R., Bohannon, B.J.M. and Guillemin, K.** (2015) Identification of Population Bottlenecks and Colonization Factors during Assembly of Bacterial Communities within the Zebrafish Intestine. *MBio* **6**, e01163-15.
- Tannières, M., Lang, J., Barnier, C., et al.** (2017) Quorum-quenching limits quorum-sensing exploitation by signal-negative invaders. *Sci. Rep.* **7**, 40126.
- Vercruyse, M., Fauvart, M., Beullens, S., Braeken, K., Cloots, L., Engelen, K., Marchal, K. and Michiels, J.** (2011) Bacteroids: Specific Gene Expression During Symbiotic Nongrowth. *Mol. Plant-Microbe Interact.* **24**, 1553–1561.
- Wang, Z., Gerstein, M. and Snyder, M.** (2009) RNA-Seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet.* **10**, 57–63.
- Wood, D.W., Setubal, J.C., Kaul, R., et al.** (2001) The genome of the natural genetic engineer *Agrobacterium tumefaciens* C58. *Science* **294**, 2317–23.

Annexes

Table S1. Transcriptomic data of Chapter III. Upregulated genes in GHB medium.

atu code	log2FC	padj	Gene	Product	COG class
atu5139	5,54	NA		Zn-dependent gamma butyryl lactone lactonase	General function prediction only
atu0035	5,44	2,33E-26	<i>pckA</i>	Phosphoenolpyruvate carboxykinase [ATP]	Energy production and conversion
atu5138	4,77	8,42E-99	<i>bcbB</i>	Gamma hydroxybutyrate dehydrogenase	Energy production and conversion
atu4092	4,63	6,61E-186	<i>cydB</i>	Cytochrome d oxidase subunit II	Energy production and conversion
atu1527	4,57	1,38E-46	<i>fixS</i>	Nitrogen fixation protein fixS	Inorganic ion transport and metabolism
atu1528	4,53	4,04E-93	<i>fixI</i>	Nitrogen fixation protein fixI	Inorganic ion transport and metabolism
atu1534	4,48	2,40E-67	<i>fixP</i>	Cytochrome-c oxidase, fixp chain	Energy production and conversion
atu1536	4,40	3,02E-67	<i>fixO</i>	Cytochrome C oxidase, fixo chain	Energy production and conversion
atu5137	4,09	8,49E-102	<i>bicA</i>	NAD-dependent succinyl-semialdehyde dehydrogenase	Energy production and conversion
atu4408	3,80	1,46E-71	<i>napA</i>	Periplasmic nitrate reductase	Energy production and conversion
atu1529	3,79	1,57E-61	<i>fixH</i>	Nitrogen fixation protein fixH	Unclassified
atu3191	3,72	1,79E-45		Outer membrane protein	Cell wall/membrane/envelope biogenesis
atu4410	3,61	1,94E-45	<i>napC</i>	Periplasmic nitrate reductase, cytochrome c-type protein	Energy production and conversion
atu4091	3,56	NA	<i>cydA</i>	Cytochrome d oxidase	Energy production and conversion
atu1535	3,53	1,08E-07	<i>fixQ</i>	Cytochrome c oxidase, fixq chain	Unclassified
atu1537	3,45	1,75E-13	<i>fixN</i>	Cbb3-type cytochrome c oxidase subunit I	Posttranslational modification, protein turnover, chaperones
atu0036	3,32	1,45E-26		Peptidyl-trna hydrolase domain protein	Translation, ribosomal structure and biogenesis
atu8036	3,30	9,02E-34		Protein YBGT-related protein	Unclassified
atu4524	3,29	5,06E-24		ABC transporter, membrane spanning protein (oligopeptide)	Amino acid transport and metabolism
atu4653	3,22	2,49E-44		ABC transporter, membrane spanning protein (sugar)	Carbohydrate transport and metabolism
atu2283	3,21	4,95E-54		Pseudouridine	Energy production and conversion
atu0626	3,16	4,67E-52	<i>adhP</i>	Alcohol dehydrogenase	General function prediction only
atu2333	3,15	2,68E-49		ATP-dependent RNA helicase	Translation, ribosomal structure and biogenesis
atu4708	3,03	2,84E-48	<i>fdhF</i>	Formate dehydrogenase alpha subunit	General function prediction only
atu4090	3,02	5,52E-32	<i>cydC</i>	ABC transporter, nucleotide binding/atpase protein	Energy production and conversion
atu0251	3,02	7,63E-26		Conserved hypothetical protein	Function unknown
atu3190	3,01	1,16E-40		Hypothetical protein	Unclassified
atu4409	3,01	1,65E-11	<i>napB</i>	Periplasmic nitrate reductase small subunit	Energy production and conversion
atu4652	2,89	1,41E-30		ABC transporter, membrane spanning protein (sugar)	Carbohydrate transport and metabolism
atu3740	2,89	1,28E-74		Fructose biphosphate aldolase	Carbohydrate transport and metabolism
atu4709	2,79	2,16E-17	<i>fdsD</i>	NAD-dependent formate dehydrogenase delta subunit	Unclassified
atu2020	2,78	6,19E-19		Tetr family transcriptional regulator	Transcription
atu1601	2,75	6,82E-23	<i>hemN</i>	Coproporphyrinogen III oxidase	Coenzyme transport and metabolism
atu3473	2,72	2,22E-36	<i>bkdA1</i>	2-oxoisovalerate dehydrogenase alpha subunit	Energy production and conversion
atu3822	2,72	8,09E-19	<i>ncl</i>	Unclassified	Unclassified
atu3298	2,71	5,06E-46	<i>dctA</i>	C4-dicarboxylate transport protein	Energy production and conversion
atu1530	2,71	7,90E-24	<i>fixG</i>	Nitrogen fixation protein fixG	Energy production and conversion
atu3123	2,68	6,31E-39	<i>rpsU2</i>	30S ribosomal protein S21 2	Translation, ribosomal structure and biogenesis
atu1585	2,66	1,37E-08		Trna-Glu	Unclassified
atu3124	2,65	1,16E-47		Hypothetical protein	Unclassified
atu4171	2,64	NA	<i>deaD</i>	Cold-shock dead-box protein A	Translation, ribosomal structure and biogenesis
atu4710	2,61	2,42E-22		MFS permease	Amino acid transport and metabolism
atu3912	2,61	4,65E-42		Hypothetical protein	Amino acid transport and metabolism
atu0983	2,58	4,03E-17		Hypothetical protein	Function unknown
atu4525	2,54	3,67E-53		ABC transporter, substrate binding protein (oligopeptide)	Amino acid transport and metabolism
atu2115	2,53	8,93E-70	<i>ppf</i>	Pyrophosphate-fructose-6-phosphate 1-phosphotransferase	Carbohydrate transport and metabolism
atu3706	2,50	8,48E-27	<i>serA</i>	D-3-phosphoglycerate dehydrogenase	Amino acid transport and metabolism
atu1899	2,48	7,64E-23	<i>rbsB</i>	ABC transporter, substrate binding protein (ribose)	Carbohydrate transport and metabolism
atu5498	2,48	7,81E-12		Putative Abscisic-aldehyde oxidase	Energy production and conversion
atu0623	2,48	2,09E-08		Hypothetical protein	Unclassified
atu4651	2,43	3,72E-20	<i>argH2</i>	Argininosuccinate lyase 2	Amino acid transport and metabolism
atu2644	2,40	9,72E-25	<i>sdhD</i>	Succinate dehydrogenase hydrophobic membrane anchor	Energy production and conversion
atu3821	2,39	4,00E-19	<i>rbsB</i>	ABC transporter, substrate binding protein (ribose)	Carbohydrate transport and metabolism
atu4424	2,39	1,75E-37		ABC transporter, nucleotide binding/atpase protein (spermidine/putrescine)	Amino acid transport and metabolism
atu2394	2,39	2,02E-17	<i>sinR</i>	Regulator of biofilm formation, Fnr family	Signal transduction mechanisms
atu1600	2,37	2,15E-21	<i>acpXL</i>	Acyl carrier protein acpXL	Lipid transport and metabolism
atu1997	2,37	NA	<i>cysJ</i>	Sulfite reductase	Function unknown
atu1468	2,36	6,94E-05		Hypothetical protein	Function unknown
atu1583	2,35	4,06E-05		Trna-Glu	Unclassified
atu3472	2,34	1,30E-09	<i>bkdA2</i>	2-oxoisovalerate dehydrogenase beta subunit	Energy production and conversion
atu4008	2,28	5,65E-20	<i>arcB*</i>		Amino acid transport and metabolism
atu1951	2,25	8,09E-17	<i>rpsL</i>	30S ribosomal protein S12	Translation, ribosomal structure and biogenesis
atu3818	2,25	7,20E-18	<i>rbsA</i>	Ribose/galactose/methyl galactoside import ATP-binding protein 3	Carbohydrate transport and metabolism
atu0396	2,25	1,48E-21		Coenzyme A transferase	Energy production and conversion
atu4654	2,24	1,22E-21		ABC transporter, substrate binding protein (sugar)	Carbohydrate transport and metabolism
atu4193	2,24	2,63E-04		ABC transporter, membrane spanning protein (oligopeptide)	Amino acid transport and metabolism
atu2492	2,23	3,86E-33	<i>mtbA</i>	MFS permease	Amino acid transport and metabolism
atu3913	2,22	NA		Biotin carboxylase	Lipid transport and metabolism
atu4425	2,21	1,72E-30	<i>iunH</i>	Inosine-uridine preferring nucleoside hydrolase	Nucleotide transport and metabolism
atu3485	2,21	6,15E-21		Short chain dehydrogenase	Function unknown
atu0118	2,16	3,03E-16		Hypothetical protein	Function unknown
atu2645	2,16	1,14E-33	<i>sdhC</i>	Succinate dehydrogenase cytochrome B-556 subunit	Energy production and conversion
atu2260	2,16	6,08E-25		Acetyltransferase	General function prediction only
atu2518	2,15	1,22E-09	<i>dppA</i>	ABC transporter, substrate binding protein (dipeptide)	Amino acid transport and metabolism
atu0323	2,15	4,56E-18	<i>rpsT</i>	30S ribosomal protein S20	Translation, ribosomal structure and biogenesis
atu0624	2,13	3,32E-19	<i>guaB</i>	Inosine 5'-monophosphate dehydrogenase	Nucleotide transport and metabolism
atu3617	2,11	9,11E-20	<i>rpmB</i>	50S ribosomal protein L28	Translation, ribosomal structure and biogenesis
atu4399	2,11	NA		Hypothetical protein	Coenzyme transport and metabolism
atu4423	2,10	4,15E-13		ABC transporter, membrane spanning protein	Amino acid transport and metabolism
atu0284	2,10	3,14E-13	<i>tspO</i>	Sensory protein	Signal transduction mechanisms
atu5049	2,09	1,19E-27		Putative DNA helicase ATP dependent	Function unknown
atu1685	2,07	5,12E-06		Hypothetical protein	Unclassified
atu2704	2,07	3,36E-05		Hypothetical protein	Unclassified
atu3487	2,07	2,57E-18		ABC transporter, substrate binding protein (sugar)	Carbohydrate transport and metabolism
atu4382	2,07	1,70E-06	<i>nirK</i>	Nitrite reductase, copper-containing	Secondary metabolites biosynthesis, transport and catabolism
atu3911	2,06	8,69E-14		UPF0317 protein Atu3911	Function unknown
atu1950	2,06	NA	<i>rpsG</i>	30S ribosomal protein S7	Translation, ribosomal structure and biogenesis
atu4422	2,05	7,07E-31		ABC transporter, membrane spanning protein (spermidine/putrescine)	Amino acid transport and metabolism
atu4407	2,05	2,36E-14	<i>napD</i>	Periplasmic nitrate reductase, napD protein	Inorganic ion transport and metabolism
atu0811	2,04	7,95E-20	<i>mgo</i>	Probable malate:quinone oxidoreductase	General function prediction only
atu1467	2,04	2,81E-04		Hypothetical protein	Unclassified
atu3198	2,04	5,44E-15		ABC transporter, substrate binding protein (ribose)	Carbohydrate transport and metabolism
atu3525	2,03	1,56E-12		Conserved hypothetical glycine-rich protein	Unclassified
atu4161	2,02	7,42E-25		Conserved hypothetical protein	Unclassified
atu0116	2,02	3,92E-16		Putative periplasmic substrate-binding protein	Function unknown
atu0667	2,01	2,84E-26	<i>glcF</i>	Glycolate oxidase iron-sulfur subunit	Energy production and conversion
atu3489	2,01	6,25E-04		ABC transporter, membrane spanning protein	Carbohydrate transport and metabolism
atu5006	2,00	5,68E-10	<i>socA</i>	Deoxyfructosyl-amino acid transporter periplasmic binding protein	Amino acid transport and metabolism

Table S2. Transcriptomic data of Chapter III. Downregulated genes in GHB medium.

atu code	log2FC	padj	Gene	Product	COG class
atu4097	-2.00	1,73E-18	nadB	L-aspartate oxidase	Coenzyme transport and metabolism
atu2348	-2.00	NA	chvE	Multiple sugar-binding periplasmic receptor chve	Carbohydrate transport and metabolism
atu3326	-2.02	1,19E-21	exoF	Exopolysaccharide production protein	Cell wall/membrane/envelope biogenesis
atu0287	-2.03	2,96E-18		Hypothetical protein	Energy production and conversion
atu0407	-2.06	1,92E-20	fbpA	ABC transporter, substrate binding protein (iron)	Inorganic ion transport and metabolism
atu0288	-2.12	1,78E-17		Hypothetical protein	Inorganic ion transport and metabolism
atu4375	-2.15	4,16E-19		Hypothetical protein	General function prediction only
atu4098	-2.15	7,05E-21	nadA	Quinolinate synthase A	Coenzyme transport and metabolism
atu2463	-2.17	1,74E-13		Putative heat shock protein	Posttranslational modification, protein turnover, chaperones
atu2462	-2.18	1,54E-05		Conserved hypothetical protein	Inorganic ion transport and metabolism
atu3166	-2.19	NA	scrK	Fructokinase	Carbohydrate transport and metabolism
atu3325	-2.23	3,84E-21	exoQ	Exopolysaccharide production protein	Cell wall/membrane/envelope biogenesis
atu4373	-2.26	2,15E-13	cytR	Laci family transcriptional regulator	Transcription
atu4537	-2.26	7,72E-10		ABC transporter, membrane spanning protein (amino acid)	Amino acid transport and metabolism
atu4374	-2.40	1,12E-11		Hypothetical protein	Carbohydrate transport and metabolism
atu0163	-2.41	3,79E-05	tonB	Tonb protein	Cell wall/membrane/envelope biogenesis
atu0938	-2.44	1,35E-40	cueR	Merr family transcriptional regulator	Transcription
atu0410	-2.47	6,20E-06		Arac family transcriptional regulator	Transcription
atu4443	-2.50	3,34E-33		Conserved hypothetical protein	General function prediction only
atu2453	-2.52	3,14E-06	mhqO	Putative ring-cleaving dioxygenase mhqo	Amino acid transport and metabolism
atu0145	-2.55	7,65E-32	glb1	Putative glutamate synthase large subunit	Amino acid transport and metabolism
atu3390	-2.57	3,97E-23		ABC transporter, substrate binding protein (iron)	Inorganic ion transport and metabolism
atu3687	-2.61	2,52E-35	fecA	Putative tonb-dependent receptor (iron transport)	Inorganic ion transport and metabolism
atu8200	-2.62	6,27E-10		Hypothetical protein	Unclassified
atu4376	-2.62	7,37E-40		Hypothetical protein	Carbohydrate transport and metabolism
atu2471	-2.62	1,50E-31		Hypothetical protein	Function unknown
atu0146	-2.63	3,97E-40	cysJ	Sulfite reductase (NADPH) flavoprotein alpha-component	Inorganic ion transport and metabolism
atu4055	-2.64	2,09E-24	exoK	Endo-1,3-1,4-beta-glycanase	Carbohydrate transport and metabolism
atu2460	-2.68	3,46E-05	hmuT	ABC transporter, substrate binding protein (hemin)	Inorganic ion transport and metabolism
atu3388	-2.72	7,63E-31	fecE	ABC transporter, nucleotide binding/ATPase protein (iron)	Coenzyme transport and metabolism
atu0162	-2.80	5,09E-26	exbD	Biopolymer transport protein exbd	Intracellular trafficking, secretion, and vesicular transport
atu3691	-2.81	1,34E-37	fecE	Iron (III)-siderophore ABC transporter, nucleotide binding/ATPase protein	Coenzyme transport and metabolism
atu3688	-2.93	3,18E-48	fecB	Putative iron(III)-siderophore ABC transporter, periplasmic substrate-binding protein	Inorganic ion transport and metabolism
atu3689	-2.95	3,35E-34	fecC	Iron(III)-siderophore ABC transporter, membrane spanning protein	Inorganic ion transport and metabolism
atu0229	-2.98	1,27E-24		Conserved hypothetical inner membran protein	Function unknown
atu2454	-3.04	4,01E-09		Iron-chelator utilization protein	Inorganic ion transport and metabolism
atu3389	-3.14	2,98E-19		ABC transporter, membrane spanning protein (iron)	Inorganic ion transport and metabolism
atu4372	-3.19	1,15E-22	rhsC	ABC transporter, membrane spanning protein (ribose)	Carbohydrate transport and metabolism
atu0590	-3.19	1,77E-43	aglR	Transcriptional regulator repressor	Transcription
atu4494	-3.24	1,68E-38	kdgA	Keto-hydroxyglutarate-aldolase/keto-deoxy- phosphogluconate aldolase	Carbohydrate transport and metabolism
atu2459	-3.25	4,38E-07	hmuU	ABC transporter, membrane spanning protein (hemin)	Inorganic ion transport and metabolism
atu3385	-3.25	4,00E-24	hasR	Heme receptor	Inorganic ion transport and metabolism
atu1196	-3.26	1,83E-61		Putative naphthalene 1,2-dioxygenase	Energy production and conversion
atu4371	-3.30	6,15E-10		ABC transporter, membrane spanning protein (sugar)	Carbohydrate transport and metabolism
atu0593	-3.31	1,64E-55	aglG	ABC transporter, membrane spanning protein, Alpha-glucoside transport system	Carbohydrate transport and metabolism
atu3690	-3.33	1,56E-30	fecD	Putative iron (III)-siderophore ABC transporter, membrane spanning protein	Inorganic ion transport and metabolism
atu2561	-3.39	5,29E-21	modA	ABC transporter, substrate binding protein (molybdate)	Inorganic ion transport and metabolism
atu0595	-3.43	4,09E-101	aglK	ABC transporter, nucleotide binding/ATPase protein	Carbohydrate transport and metabolism
atu0594	-3.44	8,92E-100	aglA	Alpha-glucosidase	Carbohydrate transport and metabolism
atu2383	-3.47	NA		Conserved hypothetical protein	Unclassified
atu3686	-3.48	7,34E-53		Conserved hypothetical protein; putative siderophore biosynthesis protein	Unclassified
atu2382	-3.58	5,08E-09		Conserved protein of unknown function	Unclassified
atu3384	-3.62	8,00E-23		Hypothetical protein	Secondary metabolites biosynthesis, transport and catabolism
atu0592	-3.71	1,37E-60	aglF	ABC transporter, membrane spanning protein	Carbohydrate transport and metabolism
atu3674	-3.77	1,90E-10	nclssifite	Unclassified	Unclassified
atu3992	-3.94	1,11E-74	copB	Copper tolerance protein	Cell wall/membrane/envelope biogenesis
atu3667	-4.06	1,89E-35	panD	Aspartate 1-decarboxylase	Coenzyme transport and metabolism
atu3685	-4.32	1,64E-68		Putative thioesterase type II, putative siderophore biosynthesis protein	Secondary metabolites biosynthesis, transport and catabolism
atu0591	-4.32	6,72E-87	aglE	ABC transporter, substrate binding protein (alpha-glucoside)	Carbohydrate transport and metabolism
atu1203	-4.52	1,11E-115		Putative Heavy metal transport/detoxification protein	Inorganic ion transport and metabolism
atu3683	-4.54	9,56E-76		Putative non-ribosomal peptide synthetase (NRPS) ; Siderophore biosynthesis protein	Secondary metabolites biosynthesis, transport and catabolism
atu3684	-4.56	2,65E-76		Putative non-ribosomal peptide synthetase (NRPS); siderophore biosynthesis protein	Secondary metabolites biosynthesis, transport and catabolism
atu3991	-4.84	5,38E-70	copA	Multicopper oxidase	Secondary metabolites biosynthesis, transport and catabolism
atu2287	-4.87	1,07E-20		Outer membrane heme receptor	Inorganic ion transport and metabolism
atu3669	-4.88	1,66E-67		Putative Na ⁺ -dependent cation transporter, putative siderophore biosynthesis protein	Defense mechanisms
atu0937	-4.91	9,71E-164	actP2	Copper transporting ATPase	Inorganic ion transport and metabolism
atu2458	-4.98	1,69E-66	hmuV	Hemin import ATP-binding protein hmuV	Inorganic ion transport and metabolism
atu4370	-5.01	3,40E-32		ABC transporter, nucleotide binding/ATPase protein (sugar)	Carbohydrate transport and metabolism
atu3676	-5.10	NA		Putative SAM-dependent methyltransferase; Siderophore biosynthesis protein	Secondary metabolites biosynthesis, transport and catabolism
atu3682	-5.17	2,73E-29		Putative non-ribosomal peptide synthetase (NRPS) ; Siderophore biosynthesis protein	Secondary metabolites biosynthesis, transport and catabolism
atu3675	-5.23	1,02E-21		Putative non-ribosomal peptide synthetase (NRPS) ; siderophore biosynthesis protein	Secondary metabolites biosynthesis, transport and catabolism
atu4369	-5.33	1,45E-119	rhsB	ABC transporter, substrate binding protein (sugar)	Carbohydrate transport and metabolism
atu3673	-5.42	5,65E-31		Non-ribosomal peptide synthase (NRPS); Siderophore biosynthesis protein	Lipid transport and metabolism
atu3678	-5.53	6,15E-61		Conserved hypothetical protein, siderophore biosynthesis protein	Function unknown
atu1195	-5.54	3,15E-181	actP	Copper-transporting P-type ATPase	Inorganic ion transport and metabolism
atu3670	-5.56	NA		Putative non-ribosomal peptide synthase (NRPS); siderophore biosynthesis protein	Secondary metabolites biosynthesis, transport and catabolism
atu3677	-5.86	6,75E-96		Putative glutamate-1-semialdehyde 2,1-aminomutase	Coenzyme transport and metabolism
atu3672	-5.95	1,58E-102		Putative nosb-like polyketide synthase (PKS); siderophore biosynthesis protein	Secondary metabolites biosynthesis, transport and catabolism
atu3679	-5.96	2,04E-50		Conserved hypothetical protein, putative siderophore biosynthesis protein	Unclassified
atu3990	-5.97	2,45E-159	copC	Copper tolerance protein	Inorganic ion transport and metabolism
atu3681	-6.06	1,75E-94		Putative polyketide synthetase (PKS) ; Siderophore biosynthesis protein	Secondary metabolites biosynthesis, transport and catabolism
atu4442	-6.14	1,41E-157		Putative protease	General function prediction only
atu3671	-6.20	8,35E-115		Putative luciferase-like monooxygenase ; putative siderophore biosynthesis protein	Energy production and conversion
atu3680	-6.31	4,21E-91		Conserved hypothetical protein, putative siderophore biosynthesis protein	Unclassified

Table S3. Transcriptomic data of Chapter III. Upregulated genes in GABA medium.

atu code	log2FC	padj	Gene	Product	COG class
atu5139	5.54	NA	<i>b1cC</i>	Zn-dependent gamma butyryl lactone lactonase	General function prediction only
atu0035	5.45	2,33E-26	<i>pckA</i>	Phosphoenolpyruvate carboxykinase [ATP]	Energy production and conversion
atu5138	4.77	8,42E-99	<i>b1cB</i>	Gamma hydroxybutyrate dehydrogenase	Energy production and conversion
atu4092	4.64	6,61E-186	<i>cydB</i>	Cytochrome d oxidase subunit II	Energy production and conversion
atu1527	4.57	1,38E-46	<i>fixS</i>	Nitrogen fixation protein fixS	Inorganic ion transport and metabolism
atu1528	4.53	4,04E-93	<i>fixI</i>	Nitrogen fixation protein fixI	Inorganic ion transport and metabolism
atu1534	4.48	2,40E-67	<i>fixP</i>	Cytochrome-c oxidase, fixP chain	Energy production and conversion
atu1536	4.40	3,02E-67	<i>fixO</i>	Cytochrome C oxidase, fixO chain	Energy production and conversion
atu5137	4.09	8,49E-102	<i>b1cA</i>	NAD-dependent succinyl-semialdehyde dehydrogenase	Energy production and conversion
atu4408	3.80	1,46E-71	<i>napA</i>	Periplasmic nitrate reductase	Energy production and conversion
atu1529	3.79	1,57E-61	<i>fixH</i>	Nitrogen fixation protein fixH	Unclassified
atu3191	3.72	1,79E-45		Outer membrane protein	Cell wall/membrane/envelope biogenesis
atu4410	3.61	1,94E-45	<i>napC</i>	Periplasmic nitrate reductase, cytochrome c-type protein	Energy production and conversion
atu4091	3.56	NA	<i>cydA</i>	Cytochrome d oxidase	Energy production and conversion
atu1535	3.53	1,08E-07	<i>fixQ</i>	Cytochrome c oxidase, fixQ chain	Unclassified
atu1537	3.45	1,75E-13	<i>fixN</i>	Cbb3-type cytochrome c oxidase subunit I	Posttranslational modification, protein turnover, chaperones
atu0036	3.32	1,45E-26		Peptidyl-trna hydrolase domain protein	Translation, ribosomal structure and biogenesis
atu8036	3.30	9,02E-34		Protein YBGT-related protein	Unclassified
atu4524	3.29	5,06E-24		ABC transporter, membrane spanning protein (oligopeptide)	Amino acid transport and metabolism
atu4653	3.22	2,49E-44		ABC transporter, membrane spanning protein (sugar)	Carbohydrate transport and metabolism
atu2283	3.21	4,95E-54		Pseudoazurin	Energy production and conversion
atu0626	3.16	4,67E-52	<i>adhP</i>	Alcohol dehydrogenase	General function prediction only
atu2333	3.15	2,68E-49		ATP-dependent RNA helicase	Translation, ribosomal structure and biogenesis
atu4708	3.03	2,84E-48	<i>fdhF</i>	Formate dehydrogenase alpha subunit	General function prediction only
atu4090	3.02	5,52E-32	<i>cydC</i>	ABC transporter, nucleotide binding/ATPase protein	Energy production and conversion
atu0251	3.02	7,63E-26		Conserved hypothetical protein, rubrerythrin-related with DUF 125 domain	Function unknown
atu3190	3.01	1,16E-40		Hypothetical protein	Unclassified
atu4409	3.01	1,65E-11	<i>napB</i>	Periplasmic nitrate reductase small subunit	Energy production and conversion
atu4652	2.89	1,41E-30		ABC transporter, membrane spanning protein (sugar)	Carbohydrate transport and metabolism
atu3740	2.89	1,28E-74		Fructose bisphosphate aldolase	Carbohydrate transport and metabolism
atu4709	2.79	2,16E-17	<i>fdsD</i>	NAD-dependent formate dehydrogenase delta subunit	Unclassified
atu2020	2.78	6,19E-19		Tetr family transcriptional regulator	Transcription
atu1601	2.75	6,82E-23	<i>hemN</i>	Coproporphyrinogen III oxidase	Coenzyme transport and metabolism
atu3473	2.72	2,22E-36	<i>bkdA1</i>	2-oxoisovalerate dehydrogenase alpha subunit	Energy production and conversion
atu3822	2.72	8,09E-19			Unclassified
atu3298	2.71	5,06E-46	<i>dctA</i>	C4-dicarboxylate transport protein	Energy production and conversion
atu1530	2.71	7,90E-24	<i>fixG</i>	Nitrogen fixation protein fixG	Energy production and conversion
atu3123	2.68	6,31E-39	<i>rpsU2</i>	30S ribosomal protein S21 2	Translation, ribosomal structure and biogenesis
atu1585	2.66	1,37E-08		Trna-Glu	Unclassified
atu3124	2.65	1,16E-47		Hypothetical protein	Unclassified
atu4171	2.64	NA	<i>deaD</i>	Cold-shock dead-box protein A	Translation, ribosomal structure and biogenesis
atu4710	2.61	2,42E-22		MFS permease	Amino acid transport and metabolism
atu3912	2.61	4,65E-42		Hypothetical protein	Amino acid transport and metabolism
atu0963	2.58	4,03E-17		Hypothetical protein	Function unknown
atu4525	2.54	3,67E-53		ABC transporter, substrate binding protein (oligopeptide)	Amino acid transport and metabolism
atu2115	2.53	8,93E-70	<i>ppf</i>	Pyrophosphate--fructose-6-phosphate 1-phosphotransferase	Carbohydrate transport and metabolism
atu3706	2.50	8,48E-27	<i>serA</i>	D-3-phosphoglycerate dehydrogenase	Amino acid transport and metabolism
atu1899	2.48	7,64E-23	<i>rbsB</i>	ABC transporter, substrate binding protein (ribose)	Carbohydrate transport and metabolism
atu5498	2.48	7,81E-12		Putative Abscisic-aldehyde oxidase	Energy production and conversion
atu0623	2.48	2,09E-08		Hypothetical protein	Unclassified
atu4651	2.43	3,72E-20	<i>argH2</i>	Argininosuccinate lyase 2	Amino acid transport and metabolism
atu2644	2.40	9,72E-25	<i>sdhD</i>	Succinate dehydrogenase hydrophobic membrane anchor	Energy production and conversion
atu3821	2.39	4,00E-19	<i>rbsB</i>	ABC transporter, substrate binding protein (ribose)	Carbohydrate transport and metabolism
atu4424	2.39	1,75E-37		ABC transporter, nucleotide binding/ATPase protein (spermidine/putrescine)	Amino acid transport and metabolism
atu2394	2.39	2,02E-17	<i>sinR</i>	Regulator of biofilm formation, Fnr family	Signal transduction mechanisms
atu1600	2.37	2,15E-21	<i>acpXL</i>	Acyl carrier protein acpXL	Lipid transport and metabolism
atu1997	2.37	NA	<i>cysJ</i>	Sulfite reductase	Function unknown
atu1468	2.36	6,94E-05		Hypothetical protein	Function unknown
atu1583	2.35	4,06E-05		Trna-Glu	Unclassified
atu3472	2.34	1,30E-09	<i>bkdA2</i>	2-oxoisovalerate dehydrogenase beta subunit	Energy production and conversion
atu4008	2.28	5,65E-20	<i>arcB*</i>		Amino acid transport and metabolism
atu1951	2.25	8,09E-17	<i>rpsL</i>	30S ribosomal protein S12	Translation, ribosomal structure and biogenesis
atu3818	2.25	7,20E-18	<i>rbsA</i>	Ribose/galactose/methyl galactoside import ATP-binding protein 3	Carbohydrate transport and metabolism
atu0396	2.25	1,48E-21		Coenzyme A transferase	Energy production and conversion
atu4654	2.24	1,22E-21		ABC transporter, substrate binding protein (sugar)	Carbohydrate transport and metabolism
atu4193	2.24	2,63E-04		ABC transporter, membrane spanning protein (oligopeptide)	Amino acid transport and metabolism
atu2492	2.23	3,86E-33	<i>mtbA</i>	MFS permease	Amino acid transport and metabolism
atu3913	2.22	NA		Biotin carboxylase	Lipid transport and metabolism
atu4425	2.21	1,72E-30	<i>iunH</i>	Inosine-uridine preferring nucleoside hydrolase	Nucleotide transport and metabolism
atu3485	2.21	6,15E-21		Short chain dehydrogenase	Function unknown
atu0118	2.16	3,03E-16		Hypothetical protein	Function unknown
atu2645	2.16	1,14E-33	<i>sdhC</i>	Succinate dehydrogenase cytochrome B-556 subunit	Energy production and conversion
atu2260	2.16	6,08E-25		Acetyltransferase	General function prediction only
atu2518	2.15	1,22E-09	<i>dppA</i>	ABC transporter, substrate binding protein (dipeptide)	Amino acid transport and metabolism
atu0323	2.15	4,56E-18	<i>rpsT</i>	30S ribosomal protein S20	Translation, ribosomal structure and biogenesis
atu0624	2.13	3,32E-19	<i>guaB</i>	Inosine 5'-monophosphate dehydrogenase	Nucleotide transport and metabolism
atu3617	2.11	9,11E-20	<i>rpmB</i>	50S ribosomal protein L28	Translation, ribosomal structure and biogenesis
atu4399	2.11	NA		Hypothetical protein	Coenzyme transport and metabolism
atu4423	2.10	4,15E-13		ABC transporter, membrane spanning protein	Amino acid transport and metabolism
atu0284	2.10	3,14E-13	<i>tspO</i>	Sensory protein	Signal transduction mechanisms
atu5049	2.09	1,19E-27		Putative DNA helicase ATP dependent	Function unknown
atu1685	2.07	5,12E-06		Hypothetical protein	Unclassified
atu2704	2.07	3,36E-05		Hypothetical protein	Unclassified
atu3487	2.07	2,57E-18		ABC transporter, substrate binding protein (sugar)	Carbohydrate transport and metabolism
atu4382	2.07	1,70E-06	<i>nirK</i>	Nitrite reductase, copper-containing	Secondary metabolites biosynthesis, transport and catabolism
atu3911	2.06	8,69E-14		UPF0317 protein Atu3911	Function unknown
atu1950	2.06	NA	<i>rpsG</i>	30S ribosomal protein S7	Translation, ribosomal structure and biogenesis
atu4422	2.05	7,07E-31		ABC transporter, membrane spanning protein (spermidine/putrescine)	Amino acid transport and metabolism
atu4407	2.05	2,36E-14	<i>napD</i>	Periplasmic nitrate reductase, napD protein	Inorganic ion transport and metabolism
atu0811	2.04	7,95E-20	<i>mgo</i>	Probable malate:quinone oxidoreductase	General function prediction only
atu1467	2.04	2,81E-04		Hypothetical protein	Unclassified
atu3198	2.04	5,44E-15		ABC transporter, substrate binding protein (ribose)	Carbohydrate transport and metabolism
atu3525	2.03	1,56E-12		Conserved hypothetical glycine-rich protein	Unclassified
atu4161	2.02	7,42E-25		Conserved hypothetical protein	Unclassified
atu0116	2.02	3,92E-16		Putative periplasmic substrate-binding protein	Function unknown
atu0667	2.01	2,84E-26	<i>glcF</i>	Glycolate oxidase iron-sulfur subunit	Energy production and conversion
atu3489	2.01	6,25E-04		ABC transporter, membrane spanning protein	Carbohydrate transport and metabolism
atu5006	2.00	5,68E-10	<i>socA</i>	Deoxyfructosyl-amino acid transporter periplasmic binding protein	Amino acid transport and metabolism

Table S4. Transcriptomic data of Chapter III. Downregulated genes in GABA medium.

atu code	log2FC	padj	Gene	Product	COG class
atu4097	-2.00	1,73E-18	<i>nadB</i>	L-aspartate oxidase	Coenzyme transport and metabolism
atu2348	-2.00	NA	<i>chvE</i>	Multiple sugar-binding periplasmic receptor chve	Carbohydrate transport and metabolism
atu3326	-2.02	1,19E-21	<i>exoF</i>	Exopolysaccharide production protein	Cell wall/membrane/envelope biogenesis
atu0287	-2.03	2,96E-18		Hypothetical protein	Energy production and conversion
atu0407	-2.06	1,92E-20	<i>fbpA</i>	ABC transporter, substrate binding protein (iron)	Inorganic ion transport and metabolism
atu0288	-2.12	1,78E-17		Hypothetical protein	Inorganic ion transport and metabolism
atu4375	-2.15	4,16E-19		Hypothetical protein	General function prediction only
atu4098	-2.15	7,05E-21	<i>nadA</i>	Quinolinate synthase A	Coenzyme transport and metabolism
atu2463	-2.17	1,74E-13		Putative heat shock protein	Posttranslational modification, protein turnover, chaperones
atu2462	-2.18	1,54E-05		Conserved hypothetical protein	Inorganic ion transport and metabolism
atu3166	-2.19	NA	<i>scrK</i>	Fructokinase	Carbohydrate transport and metabolism
atu3325	-2.23	3,84E-21	<i>exoQ</i>	Exopolysaccharide production protein	Cell wall/membrane/envelope biogenesis
atu4373	-2.26	2,15E-13	<i>cytR</i>	Laci family transcriptional regulator	Transcription
atu4537	-2.26	7,72E-10		ABC transporter, membrane spanning protein (amino acid)	Amino acid transport and metabolism
atu4374	-2.40	1,12E-11		Hypothetical protein	Carbohydrate transport and metabolism
atu0163	-2.41	3,79E-05	<i>tonB</i>	Tonb protein	Cell wall/membrane/envelope biogenesis
atu0938	-2.44	1,35E-40	<i>cueR</i>	Merr family transcriptional regulator	Transcription
atu0410	-2.47	6,20E-06		Arac family transcriptional regulator	Transcription
atu4443	-2.50	3,34E-33		Conserved hypothetical protein	General function prediction only
atu2453	-2.52	3,14E-06	<i>mhqO</i>	Putative ring-cleaving dioxygenase mhqo	Amino acid transport and metabolism
atu0145	-2.55	7,65E-32	<i>gltB1</i>	Putative glutamate synthase large subunit	Amino acid transport and metabolism
atu3390	-2.57	3,97E-23		ABC transporter, substrate binding protein (iron)	Inorganic ion transport and metabolism
atu3687	-2.61	2,52E-35	<i>fecA</i>	Putative tonb-dependent receptor (iron transport)	Inorganic ion transport and metabolism
atu8200	-2.62	6,27E-10		Hypothetical protein	Unclassified
atu4376	-2.62	7,37E-40		Hypothetical protein	Carbohydrate transport and metabolism
atu2471	-2.62	1,50E-31		Hypothetical protein	Function unknown
atu0146	-2.63	3,97E-40	<i>cysJ</i>	Sulfite reductase (NADPH) flavoprotein alpha-component	Inorganic ion transport and metabolism
atu4055	-2.64	2,09E-24	<i>exoK</i>	Endo-1,3-1,4-beta-glycanase	Carbohydrate transport and metabolism
atu2460	-2.68	3,46E-05	<i>hmuT</i>	ABC transporter, substrate binding protein (hemin)	Inorganic ion transport and metabolism
atu3388	-2.72	7,63E-31	<i>fecE</i>	ABC transporter, nucleotide binding/atpase protein (iron)	Coenzyme transport and metabolism
atu0162	-2.80	5,09E-26	<i>exbD</i>	Biopolymer transport protein exbd	Intracellular trafficking, secretion, and vesicular transport
atu3691	-2.81	1,34E-37	<i>fecE</i>	Iron (III)-siderophore ABC transporter, nucleotide binding/atpase protein	Coenzyme transport and metabolism
atu3688	-2.93	3,18E-48	<i>fecB</i>	Putative iron(III)-siderophore ABC transporter, periplasmic substrate-binding protein	Inorganic ion transport and metabolism
atu3689	-2.95	3,35E-34	<i>fecC</i>	Iron(III)-siderophore ABC transporter, membrane spanning protein	Inorganic ion transport and metabolism
atu0229	-2.98	1,27E-24		Conserved hypothetical inner membran protein	Function unknown
atu2454	-3.04	4,01E-09		Iron-chelator utilization protein	Inorganic ion transport and metabolism
atu3389	-3.14	2,98E-19		ABC transporter, membrane spanning protein (iron)	Inorganic ion transport and metabolism
atu4372	-3.19	1,15E-22	<i>rhsC</i>	ABC transporter, membrane spanning protein (ribose)	Carbohydrate transport and metabolism
atu0590	-3.19	1,77E-43	<i>aglR</i>	Transcriptional regulator repressor	Transcription
atu4494	-3.24	1,68E-38	<i>kdgA</i>	Keto-hydroxyglutarate-aldolase/keto-deoxy- phosphogluconate aldolase	Carbohydrate transport and metabolism
atu2459	-3.25	4,38E-07	<i>hmuU</i>	ABC transporter, membrane spanning protein (hemin)	Inorganic ion transport and metabolism
atu3385	-3.25	4,00E-24	<i>hasR</i>	Heme receptor	Inorganic ion transport and metabolism
atu1196	-3.26	1,83E-61		Putative naphthalene 1,2-dioxygenase	Energy production and conversion
atu4371	-3.30	6,15E-10		ABC transporter, membrane spanning protein (sugar)	Carbohydrate transport and metabolism
atu0593	-3.31	1,64E-55	<i>aglG</i>	ABC transporter, membrane spanning protein, Alpha-glucoside transport system	Carbohydrate transport and metabolism
atu3690	-3.33	1,56E-30	<i>fecD</i>	Putative iron (III)-siderophore ABC transporter, membrane spanning protein	Inorganic ion transport and metabolism
atu2561	-3.39	5,29E-21	<i>mocA</i>	ABC transporter, substrate binding protein (molybdate)	Inorganic ion transport and metabolism
atu0595	-3.43	4,09E-101	<i>aglK</i>	ABC transporter, nucleotide binding/atpase protein	Carbohydrate transport and metabolism
atu0594	-3.44	8,92E-100	<i>aglA</i>	Alpha-glucosidase	Carbohydrate transport and metabolism
atu2383	-3.47	NA		Conserved hypothetical protein	Unclassified
atu3686	-3.48	7,34E-53		Conserved hypothetical protein; putative siderophore biosynthesis protein	Unclassified
atu2382	-3.58	5,08E-09		Conserved protein of unknown function	Unclassified
atu3384	-3.62	8,00E-23		Hypothetical protein	Secondary metabolites biosynthesis, transport and catabolism
atu0592	-3.71	1,37E-60	<i>aglF</i>	ABC transporter, membrane spanning protein	Carbohydrate transport and metabolism
atu3674	-3.77	1,90E-10			Unclassified
atu3992	-3.94	1,11E-74	<i>copB</i>	Copper tolerance protein	Cell wall/membrane/envelope biogenesis
atu3667	-4.06	1,89E-35	<i>panD</i>	Aspartate 1-decarboxylase	Coenzyme transport and metabolism
atu3685	-4.32	1,64E-68		Putative thioesterase type IIs; putative siderophore biosynthesis protein	Secondary metabolites biosynthesis, transport and catabolism
atu0591	-4.32	6,72E-87	<i>aglE</i>	ABC transporter, substrate binding protein (alpha-glucoside)	Carbohydrate transport and metabolism
atu1203	-4.52	1,11E-115		Putative Heavy metal transport/detoxification protein	Inorganic ion transport and metabolism
atu3683	-4.54	9,56E-76		Putative non-ribosomal peptide synthetase (NRPS); Siderophore biosynthesis protein	Secondary metabolites biosynthesis, transport and catabolism
atu3684	-4.56	2,65E-76		Putative non-ribosomal peptide synthetase (NRPS) ; siderophore biosynthesis protein	Secondary metabolites biosynthesis, transport and catabolism
atu3991	-4.84	5,38E-70	<i>copA</i>	Multicopper oxidase	Secondary metabolites biosynthesis, transport and catabolism
atu2287	-4.87	1,07E-20		Outer membrane heme receptor	Inorganic ion transport and metabolism
atu3669	-4.88	1,66E-67		Putative Na ⁺ -dependent cation transporter ; siderophore biosynthesis protein	Defense mechanisms
atu0937	-4.91	9,71E-164	<i>actP2</i>	Copper transporting ATPase	Inorganic ion transport and metabolism
atu2458	-4.98	1,69E-66	<i>hmuV</i>	Hemin import ATP-binding protein HmuV	Inorganic ion transport and metabolism
atu4370	-5.01	3,40E-32		ABC transporter, nucleotide binding/ATPase protein (sugar)	Carbohydrate transport and metabolism
atu3676	-5.10	NA		Putative SAM-dependent methyltransferase; Siderophore biosynthesis protein	Secondary metabolites biosynthesis, transport and catabolism
atu3682	-5.17	2,73E-29		Putative non-ribosomal peptide synthetase (NRPS); Siderophore biosynthesis protein	Secondary metabolites biosynthesis, transport and catabolism
atu3675	-5.23	1,02E-21		Putative non-ribosomal peptide synthetase (NRPS); siderophore biosynthesis protein	Secondary metabolites biosynthesis, transport and catabolism
atu4369	-5.33	1,45E-119	<i>rhsB</i>	ABC transporter, substrate binding protein (sugar)	Carbohydrate transport and metabolism
atu3673	-5.42	5,65E-31		Non-ribosomal peptide synthase (NRPS) ; Siderophore biosynthesis protein	Lipid transport and metabolism
atu3678	-5.53	6,15E-61		Conserved hypothetical protein siderophore biosynthesis protein	Function unknown
atu1195	-5.54	3,15E-181	<i>actP</i>	Copper-transporting P-type ATPase	Inorganic ion transport and metabolism
atu3670	-5.56	NA		Putative non-ribosomal peptide synthase (NRPS) ; siderophore biosynthesis protein	Secondary metabolites biosynthesis, transport and catabolism
atu3677	-5.86	6,75E-96		Putative glutamate-1-semialdehyde 2,1-aminomutase	Coenzyme transport and metabolism
atu3672	-5.95	1,58E-102		Putative NosB-like polyketide synthase (PKS) ; siderophore biosynthesis protein	Secondary metabolites biosynthesis, transport and catabolism
atu3679	-5.96	2,04E-50		Conserved hypothetical protein, putative siderophore biosynthesis protein	Unclassified
atu3990	-5.97	2,45E-159	<i>copC</i>	Copper tolerance protein	Inorganic ion transport and metabolism
atu3681	-6.06	1,75E-94		Putative polyketide synthetase (PKS); Siderophore biosynthesis protein	Secondary metabolites biosynthesis, transport and catabolism
atu4442	-6.14	1,41E-157		Putative protease	General function prediction only
atu3671	-6.20	8,35E-115		Putative luciferase-like monooxygenase; siderophore biosynthesis protein	Energy production and conversion
atu3680	-6.31	4,21E-91		Conserved hypothetical protein, putative siderophore biosynthesis protein	Unclassified

Table S5. Primers list used for Single clone sequencing in Chapter IV

Line	SNP	Gene	Primer 3' → 5'
L1	<i>atu3932</i>	HP	for CCGTGCCGATGTTGAAATCC rev GCTGTGCTGCGTGTGTAG
L2	<i>atu0872</i>	<i>mcpC</i>	for GCCGCCAGATGTCTTTCAG rev GTCGCCTTCCACGCAAATG
L8	<i>atu0499</i>	HP	for CCGATGGTGCATTGTCAATG rev CGGAAGGACGGACGGATTTC
L9	<i>atu1173</i>	MFS	for AAATTGCCCGACCCGAAGAC rev AAGGACCAGGCGATGAAACG
	<i>atu1287</i>	HP	for GCGCCAACCACATTCAAAGG rev TGATTGGGTGCGACATAATGC
L14	<i>atu3475/atu3476</i>		for TCAATGTCAGAAACCGGAATAGC rev TGGCGTGGATCAGGTAATCG
L16	<i>atu1947/atu1948</i>		for CGTGAGGAGAGCGGTTGAC rev ACTACCGTCCGCAGTTCTAC
	<i>atu1962</i>	<i>secE</i>	for CCTTCTACCGGGCCTTTGAC rev CCCTGCCACTTTCCATTTCG

Table S6. Primers list used for qPCR High-resolution melting analysis of fitness assays on tomato plants in Chapter IV

Line	SNP	Gene	Primer 3' → 5'
L1	<i>atu3932</i>	HP	for CACGTAGCGGTGGAATGTTG rev GACAACGGGAGGATTTTCGC
L2	<i>atu0872</i>	<i>mcpC</i>	for CAGCATCGCCAGTGTGCG rev GGTCTGGCCCTGATCGAGAG
L8	<i>atu0499</i>	HP	for TGGTGCATTGTCAATGATGTTTC rev TCCCGAAACCGACGATAAGG
L9	<i>atu1173</i>	MFS	for GCATGTTGATTTCCCGATTTCG rev ACCGAAATTCCAGCCGAGAC
L16	<i>atu1962</i>	<i>secE</i>	for CGCAGCAAAGAAGAAGACAGAGC rev CGTGGGGCTGATTAGTTTCGG

Titre : Mode de vie d'*Agrobacterium tumefaciens* dans la tumeur

Mots clés : *Agrobacterium tumefaciens*, Tumeur, Transcriptomique, Evolution expérimental, Tn-Seq

Résumé Le phytopathogène *Agrobacterium tumefaciens* est l'agent causal de la maladie appelée galle du collet, et est capable d'infecter plus de 90 familles de plantes dicotylédones. Cette α -proteobactérie appartient à la famille des *Rhizobiaceae*. *A. tumefaciens* est décrit comme un complexe de différentes espèces regroupées en 10 génomovars (G1 à G8 et G13). *A. tumefaciens* C58 appartient au groupe du G8. Son génome est constitué de 4 réplicons : 1 chromosome circulaire, 1 chromosome linéaire et les plasmides: pAt (pour *A. tumefaciens*) et pTi (pour tumor inducing, qui est requis pour la virulence).

Pour explorer de nouveaux aspects du mode de vie d'*A. tumefaciens*, et en particulier l'interaction entre la bactérie et sa plante hôte,

deux approches différentes ont été utilisées pour identifier, caractériser et analyser les gènes qui pourraient jouer un rôle dans l'adaptation des bactéries à la tumeur. Une expérience de l'évolution par des passages en série de trois souches différentes de l'agent pathogène sur la plante hôte *Solanum lycopersicum* a été effectuée afin de clarifier la dynamique évolutive du génome au cours de l'infection. Parallèlement, une étude de différents transcriptomes (*in planta* et *in vitro*) a été réalisée et étudiée pour élucider des gènes bactériens candidats impliqués dans l'interaction de la bactérie avec la plante et divers composés produits dans la tumeur. Ce travail tente de donner une vue plus générale du processus d'adaptation de la bactérie à la niche écologique qui est la tumeur.

Title : Lifestyle of *Agrobacterium tumefaciens* in the tumor

Keywords : *Agrobacterium tumefaciens*, Tumor, Transcriptomics, Experimental evolution, Tn-Seq

Abstract: *Agrobacterium tumefaciens* is the causal agent of the plant disease called crown gall, and it's able to infect more than 90 families of dicotyledonous plants. This α -*Proteobacterium* belongs to the *Rhizobiaceae* family. *A. tumefaciens* is a complex of different species grouped in 10 genomovars (G1 to G8, and G13). *A. tumefaciens* C58 belongs to the G8 group. Its genome consists in 4 replicons: 1 chromosome circular, 1 chromosome linear and 2 dispensable plasmids: pAt (for *A. tumefaciens*) and pTi (for Tumor inducing), which is required for virulence. To explore new aspects of the *A. tumefaciens* lifestyle, and in particular the interaction between the bacteria and its plant host, two different approaches have been used to

identify, characterize and analyze genes that could play a role in the adaptation of the bacteria to tumor lifestyle. An evolution experiment by serial passages of three different strains of the pathogen on the host plant *Solanum lycopersicum* has been carried out to clarify the evolutionary dynamics of the genome during the course of infection. In parallel, a study of different transcriptomes (*in planta* and *in vitro*) was performed and studied to elucidate bacterial candidate genes involved in the interaction of the bacteria with the plant and various compounds produced in the tumor. This work attempts to give a more general view of the process of adaptation of the bacteria to the ecological niche that is the tumor.